Latent variable interactions

Bengt Muthén & Tihomir Asparouhov

Mplus

www.statmodel.com

November 2, 2015

1 Latent variable interactions

Structural equation modeling with latent variable interactions has been discussed with respect to maximum-likelihood estimation in Klein and Moosbrugger (2000). Multivariate normality is assumed for the latent variables. The ML computations are heavier than for models without latent variable interactions because numerical integration is needed. For an overview of the ML approach and various estimators suggested in earlier work, see Marsh et al. (2004). Arminger and Muthén (1998), Klein and Muthén (2007), Cudeck et al. (2009), and Mooijaart and Bentler (2010) discuss alternative estimators and algorithms. This section discusses interpretation, model testing, explained variance, standardization, and plotting of effects for models with latent variable interactions.

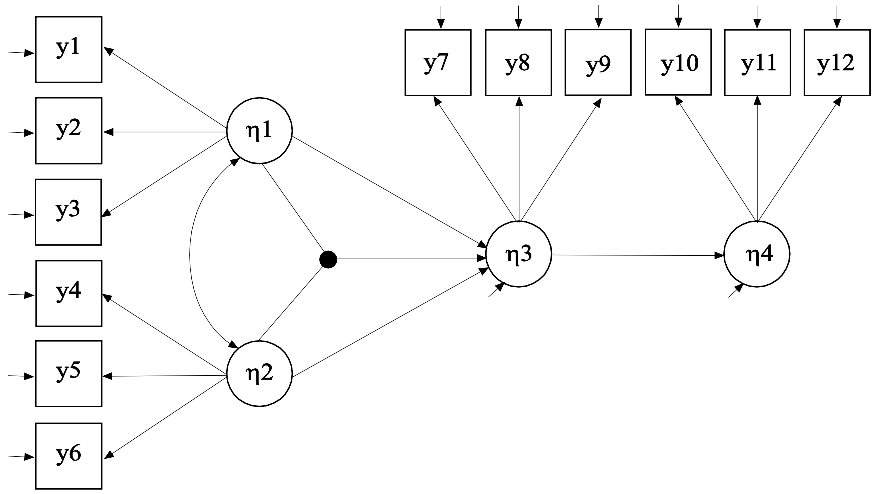
1.1 Model interpretation

As an example, consider the latent variable interaction model of Figure 1. The figure specifies that the factor η_3 is regressed on η_1 and η_2 as well as the interaction between η_1 and η_2, as shown by the structural equation

$$\eta_3 = \beta_1\, \eta_1 + \beta_2\, \eta_2 + \beta_3\, \eta_1 \times \eta_2 + \zeta_3. \tag{1}$$

The interaction variable $\eta_1 \times \eta_2$ involves only one parameter, the slope β_3. The interaction variable does not have a mean or a variance parameter. It does not have parameters for covariances with other variables. It can also not be a dependent variable. As is seen in Figure 1, the model also contains a second structural equation where η_4 is linearly regressed on η_3, so that there is no direct effect on

Figure 1: Structural equation model with interaction between latent variables



η_4 from η_1 and η_2, or their interaction.

For ease of interpretation the (1) regression can be re-written in the equivalent form

$$\eta_3 = (\beta_1 + \beta_3\,\eta_2)\,\eta_1 + \beta_2\,\eta_2 + \zeta_3, \tag{2}$$

where $(\beta_1 + \beta_3\,\eta_2)$ is a moderator function (Klein & Moosbrugger, 2000) so that the β_1 strength of influence of η_1 on η_3 is moderated by $\beta_3\,\eta_2$. The choice of moderator when translating (1) to (2) is arbitrary from an algebraic point of view, and is purely a choice based on ease of substantive interpretation. As an example, Cudeck et al. (2009) considers school achievement (η_3) influenced by general reasoning (η_1), quantitative ability (η_2), and their interaction. In line with (2) the interaction is expressed as quantitative ability moderating the influence of general reasoning on school achievement. Plotting of interactions further aids the interpretation as discussed in Section 1.5.

1.2 Model testing

As pointed out in Mooijaart and Satorra (2009), the likelihood-ratio χ^2 obtained by ML for models without latent variable interactions is not sensitive to incorrectly leaving out latent variable interactions. For example, the model of Figure 1 without the interaction term $\beta_3\ \eta_1 \times \eta_2$ fits data generated as in (1) perfectly. This is due to general maximum-likelihood results on robustness to non-normality (Satorra, 1992, 2002). Misfit can be detected only by considering higher-order moments than the second-order variances and covariances of the outcomes. Without involving higher-order moments, a reasonable modeling strategy is to first fit a model without interactions and obtain a good fit in terms of the ML likelihood-ratio χ^2. An interaction term can then be added and the β_3 significance of the interaction significance tested by either a z-test or a likelihood-ratio χ^2 difference test (Klein & Moosbrugger, 2000). Likelihood-ratio or Wald tests can be used to test the joint significance of several interaction terms.

1.3 Mean, variance, and R^2

To compute a dependent variable mean, variance, and R^2 for models with latent variable interactions, the following results are needed. As discussed in Chapter **??**, the covariance between two variables x_j and x_k is defined as

$$Cov(x_j, x_k) = E(x_j\ x_k) - E(x_j)\ E(x_k), \tag{3}$$

so that the variance is obtained as

$$Cov(x_j, x_j) = V(x_j) = E(x_j^2) - [E(x_j)]^2. \tag{4}$$

With $E(x_j) = 0$ or $E(x_k) = 0$, (3) gives the mean of a product

$$E(x_j\ x_k) = Cov(x_j, x_k). \tag{5}$$

Assuming multivariate normality for four random variables x_i, x_j, x_k, x_l any third-order moment about the mean (μ) is zero (see, e.g., Anderson, 1984),

$$E(x_i - \mu_i)(x_j - \mu_j)(x_k - \mu_k) = 0, \tag{6}$$

while the fourth-order moment about the mean is a function of covariances,

$$E(x_i - \mu_i)(x_j - \mu_j)(x_k - \mu_k)(x_l - \mu_l) = \sigma_{ij}\ \sigma_{kl} + \sigma_{ik}\ \sigma_{jl} + \sigma_{il}\ \sigma_{jk}, \tag{7}$$

where for example $\sigma_{jk} = Cov(x_j, x_k)$ and $\sigma_{kk} = Var(x_k)$. This gives

$$E(x_j\ x_k\ x_j\ x_k) = V(x_j)\ V(x_k) + 2\ [Cov(x_j, x_k)]^2, \tag{8}$$

so that the variance of a product is obtained as

$$\begin{aligned} V(x_j\ x_k) &= E(x_j\ x_k\ x_j\ x_k) - [E(x_j\ x_k)]^2 && (9) \\ &= V(x_j)\ V(x_k) + 2\ [Cov(x_j, x_k)]^2 - [Cov(x_j, x_k]^2 && (10) \\ &= V(x_j)\ V(x_k) + [Cov(x_j, x_k)]^2, && (11) \end{aligned}$$

Consider the application of these results to the mean and variance of the factor η_3 in (1) of Figure 1. Because of zero factor means, using (5) the mean of η_3 in

(1) is obtained as

$$E(\eta_3) = \beta_1\,0 + \beta_2\,0 + \beta_3\,E(\eta_1\,\eta_2) + 0 \tag{12}$$

$$= \beta_3\,Cov(\eta_1, \eta_2). \tag{13}$$

Using (4), the variance of η_3 is

$$V(\eta_3) = E(\eta_3\,\eta_3) - [E(\eta_3)]^2, \tag{14}$$

where the second term has already been determined. As for the first term, multiplying the right-hand side of (1) by itself results in products of two, three, and four factors. Expectations for three- and four-factor terms are simplified by the following two results, assuming bivariate normality and zero means for η_1 and η_2. All third-order moments $E(\eta_i\,\eta_j\,\eta_k)$ are zero by (6). The formula (8) is used to obtain the result

$$E(\eta_1\,\eta_2\,\eta_1\,\eta_2) = V(\eta_1)\,V(\eta_2) + 2\,[Cov(\eta_1, \eta_2)]^2. \tag{15}$$

Collecting terms, it follows that the variance of η_3 is obtained as

$$V(\eta_3) = \beta_1^2\,V(\eta_1) + \beta_2^2\,V(\eta_2) + 2\,\beta_1\,\beta_2\,Cov(\eta_1, \eta_2) + \beta_3^2\,V(\eta_1\,\eta_2) + V(\zeta_3), \tag{16}$$

where by (11)

$$V(\eta_1\,\eta_2) = V(\eta_1)\,V(\eta_2) + [Cov(\eta_1, \eta_2)]^2, \tag{17}$$

R-square for η_3 can be expressed as usual as

$$[V(\eta_3) - V(\zeta_3)]/V(\eta_3). \tag{18}$$

Using (16), the proportion of $V(\eta_3)$ contributed by the interaction term can be quantified as (cf. Mooijaart & Satorra, 2009; p. 445)

$$\beta_3^2\ [V(\eta_1)\ V(\eta_2) + [Cov(\eta_1, \eta_2)]^2]/V(\eta_3). \tag{19}$$

Consider as a hypothetical example the latent variable interaction model of Figure 2. Here, the latent variable interaction is between an exogenous and an endogenous latent variable. This example is useful to study the details of how to portray the model. The structural equations are

$$\eta_1 = \beta\ \eta_2 + \zeta_1, \tag{20}$$

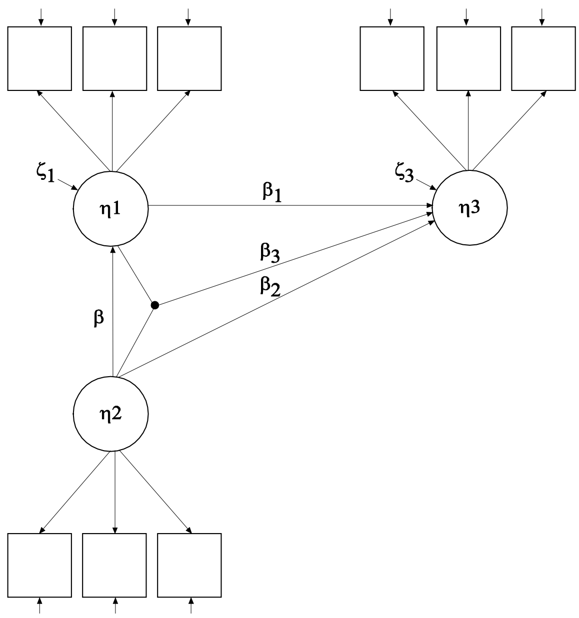
$$\eta_3 = \beta_1\ \eta_1 + \beta_2\ \eta_2 + \beta_3\ \eta_1 \times \eta_2 + \zeta_3. \tag{21}$$

Let $\beta = 1$, $\beta_1 = 0.5$, $\beta_2 = 0.7$, $\beta_3 = 0.4$, $V(\eta_2) = 1$, $V(\zeta_1) = 1$, and $V(\zeta_3) = 1$. This implies that $V(\eta_1) = \beta^2\ V(\eta_2) + V(\zeta_1) = 1^2 \times 1 + 1 = 2$ and $Cov(\eta_1, \eta_2) = \beta\ V(\eta_2) = 1 \times 1 = 1$. Using (16), $V(\eta_3) = 3.17$. The η_3 R-square is 0.68 and the variance percentage due to the interaction is 15%.

1.4 Standardization

Because latent variables have arbitrary metrics, it is useful to also present interaction effects in terms of standardized latent variables. Noting that (21)

Figure 2: Structural equation model with interaction between an exogenous and an endogenous latent variable



is identical to (1), the model interpretation is aided by considering the moderator function $(\beta_1 + \beta_3\ \eta_2)\ \eta_1$ of (2), so that η_2 moderates the η_1 influence on η_3.

As usual, standardization is obtained by dividing by the standard deviation of the dependent variable and multiplying by the standard deviation of the independent variable. The standardized β_1 and β_3 coefficients in the term $(\beta_1 + \beta_3\ \eta_2)\ \eta_1$ are obtained by dividing both by $\sqrt{V(\eta_3)} = \sqrt{3.17}$, multiplying β_1 by $\sqrt{V(\eta_1)} = \sqrt{2}$, and multiplying β_3 by $\sqrt{V(\eta_1)}\ \sqrt{V(\eta_2)} = \sqrt{2}$. This gives a standardized $\beta_1 = 0.397$ and a standardized $\beta_3 = 0.318$. The standardization of β_3 is in line with Wen, Marsh, and Hau (2010; equation 10). These authors discuss why standardization of β_3 using $\sqrt{V(\eta_1)}\ \sqrt{V(\eta_2)}$ is preferred over using $\sqrt{V(\eta_1 \times \eta_2)}$.

The standard deviation change in η_3 as a function of a one standard deviation change in η_1 can now be evaluated at different values of η_2 using the moderator function. At the zero mean of η_2, a standard deviation increase in η_1 leads to a 0.397 standard deviation increase in η_3. At one standard deviation above the mean of η_2, a standard deviation increase in η_1 leads to a $0.397 + 0.318 \times 1 = 0.715$ standard deviation increase in η_3. At one standard deviation below the mean of η_2, a standard deviation increase in η_1 leads to a $0.397 - 0.318 \times 1 = 0.079$ standard deviation increase in η_3. In other words, the biggest effect of η_1 on η_3 occurs for subjects with high values on η_2.

1.5 Plotting of interactions

The interaction can be plotted as in Figure 3. Using asterisks to denote standardization, consider the rearranged (21),

$$\eta_3^* = (\beta_1^* + \beta_3^* \, \eta_2^*) \, \eta_1^* + \beta_2^* \, \eta_2^* + \zeta_3^*. \tag{22}$$

Using (22), the three lines in the figure are expressed as follows in terms of the conditional expectation function for η_3^* at the three levels of η_2^*,

$$E(\eta_3^* | \eta_1^*, \eta_2^* = 0) = \beta_1^* \, \eta_1^*, \tag{23}$$

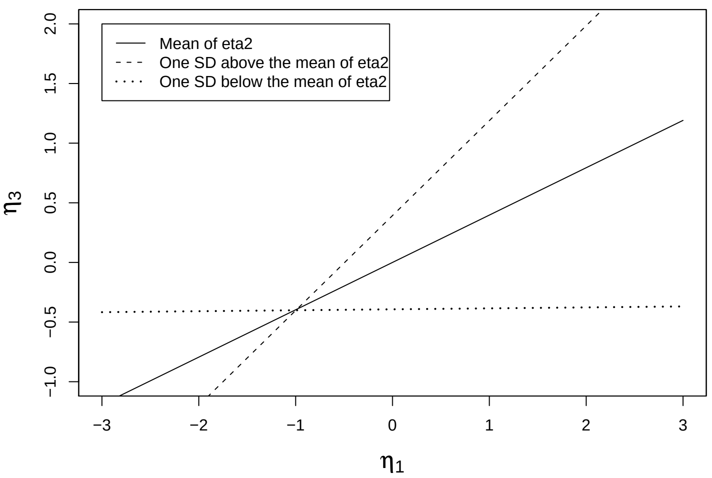
$$E(\eta_3^* | \eta_1^*, \eta_2^* = 1) = (\beta_1^* + \beta_3^*) \, \eta_1^* + \beta_2^*, \tag{24}$$

$$E(\eta_3^* | \eta_1^*, \eta_2^* = -1) = (\beta_1^* - \beta_3^*) \, \eta_1^* - \beta_2^*. \tag{25}$$

$$\tag{26}$$

Here, the standardized value $\beta_2^* = \beta_2 \times \sqrt{V(\eta_2)} / \sqrt{V(\eta_3)} = 0.7 \times 1 / \sqrt{3.17} = 0.393$.

Figure 3: Interaction plot for structural equation model with interaction between an exogenous and an endogenous latent variable



2 Standardization in matrix terms

This section describes standardization of the general model in Mplus when latent variable interactions are present. Suppose that Y is a vector of dependent variables, X is a vector of covariates, η is a vector of latent variables. All residuals are assumed normal.

Suppose that the variables (Y,η,X) are split into two disjoint set of variables V_1 and V_2. V_1 set of dependent variables that are not part of interaction terms and V_2 is a set of variables that are a part of interaction terms. Suppose that V_1 is a vector of size p_1 and V_2 is a vector of size p_2. The SEM equation is given by these two equations

$$V_1 = \alpha_1 + B_1 V_1 + C_1 V_2 + \sum_{i=1}^{k} D_i (V_{2,f(i)} V_{2,g(i)}) + \varepsilon_1$$

$$V_2 = \alpha_2 + B_2 V_2 + \varepsilon_2$$

where $\alpha_1, B_1, C_1, D_i, \alpha_2, B_2$ are model parameters. The vectors α_1, D_i are of length p_1 while the vector α_2 is of length p_2. The matrices B_1, C_1 and B_2 are of size $p_1 \times p_1$, $p_1 \times p_2$ and $p_2 \times p_2$ respectively.

The residual variable ε_1 has zero mean and variance covariance Θ and ε_2 has zero mean and variance covariance Ψ. The residuals ε_1 are not considered independent of the residuals ε_2. Let's call the covariance $F = Cov(\varepsilon_1, \varepsilon_2)$. The functions $f(i)$ and $g(i)$ simply define the interaction terms, i.e., $f(i)$ and $g(i)$ are integers between 1 and p_2 and k is the number of interaction terms in the model.

We can assume that all covariates X are in the V_2 vector and the V_1 vector consists only of η and Y variables that are regressed on interaction terms, while the remaining η and Y variables are in vector V_2. We can compute the model implied mean and variance for these variables as follows. For the variables V_2 we get

$$E(V_2) = \mu_2 = (1 - B_2)^{-1}\alpha_2$$

$$Var(V_2) = \Sigma_2 = (1 - B_2)^{-1}\Psi((1 - B_2)^{-1})^T$$

For V_1 we get

$$E(V_1) = (1-B_1)^{-1}\alpha_1 + (1-B_1)^{-1}C_1\mu_2 + (1-B_1)^{-1}\sum_{i=1}^{k} D_i(\mu_{2,f(i)}\mu_{2,g(i)} + \Sigma_{2,f(i),g(i)}).$$

Denote by

$$V_{20} = V_2 - \mu_2$$

$$\mu_{10} = (1 - B_1)^{-1}\alpha_1 + (1 - B_1)^{-1}C_1\mu_2 + (1 - B_1)^{-1}\sum_{i=1}^{k} D_i(\mu_{2,f(i)}\mu_{2,g(i)})$$

$$V_{10} = (1-B_1)^{-1}C_1V_{20} + (1-B_1)^{-1}\varepsilon_1 + (1-B_1)^{-1}\sum_{i=1}^{k} D_i(\mu_{2,f(i)}V_{20,g(i)}) +$$

$$(1-B_1)^{-1}\sum_{i=1}^{k} D_i(V_{20,f(i)}\mu_{2,g(i)}).$$

Then

$$V_1 = \mu_{10} + V_{10} + (1-B_1)^{-1}\sum_{i=1}^{k} D_i(V_{20,f(i)}V_{20,g(i)}).$$

Another representation for V_{10} is

$$V_{10} = QV_{20} + (1-B_1)^{-1}\varepsilon_1$$

where the matrix Q combines all the coefficients from the terms involving V_{20}. The above equation is essentially the definition of Q. Note now that

$$Cov(\varepsilon_1, V_{20}) = F((1-B_2)^{-1})^T$$

and thus

$$Var(V_{10}) = Q\Sigma_2Q^T + (1-B_1)^{-1}\Theta((1-B_1)^{-1})^T + Q(1-B_2)^{-1}F^T((1-B_1)^{-1})^T +$$

$$(1-B_1)^{-1}F((1-B_2)^{-1})^TQ^T.$$

Using the fact that the covariance between $V_{20,f(i)}V_{20,g(i)}$ and V_{20} and the covariance between $V_{20,f(i)}V_{20,g(i)}$ and ε_1 are zero we get that

$$Var(V_1) = Var(V_{10}) + \sum_{i,j} D_iCov(V_{20,f(i)}V_{20,g(i)}, V_{20,f(j)}V_{20,g(j)})D_j^T =$$

$$Var(V_{10}) + \sum_{i,j} D_i D_j^T (\Sigma_{2,f(i),f(j)} \Sigma_{2,g(i),g(j)} + \Sigma_{2,f(i),g(j)} \Sigma_{2,g(i),f(j)}).$$

Note also that

$$Cov(V_1, V_2) = Cov(V_{10}, V_{20}) = Q\Sigma_2 + (1 - B_1)^{-1} F((1 - B_2)^{-1})^T.$$