DSEM in a Nutshell

Tihomir Asparouhov, Bengt Muthén and Ellen Hamaker

July 10, 2017

Overview

- The general DSEM model and the new RDSEM model
- The MEAR model
- Covariates in DSEM and RDSEM

DSEM model

- Includes three separate models: single level, twolevel , cross-classified
- Main decomposition equation

$$Y_{it} = Y_{1,it} + Y_{2,i} + Y_{3,t}$$

- $Y_{2,i}$, $Y_{3,t}$ are the "individual" and "time" specific contribution. These are latent variables. $Y_{1,it}$ is the residual.
- Includes three separate models:
  - single level DSEM: type=general, N=1, $Y_{2,i}$, $Y_{3,t}$ are removed
  - two-level DSEM: type=twolevel, $Y_{3,t}$ is removed
  - cross-classified DSEM: type=cross, full version

DSEM model

- The within level model includes latent variables and observed variables from the previous L (lag) periods

$$Y_{1,it} = \nu_1 + \sum_{l=0}^{L} \Lambda_{1,lit} \eta_{1,i,t-l} + \sum_{l=0}^{L} R_{lit} Y_{1,i,t-l} + \sum_{l=0}^{L} K_{1,lit} X_{1,i,t-l} + \varepsilon_{1,it}$$

$$\eta_{1,it} = \alpha_{1,it} + \sum_{l=0}^{L} B_{1,lit} \eta_{1,i,t-l} + \sum_{l=0}^{L} Q_{lit} Y_{1,i,t-l} + \sum_{l=0}^{L} \Gamma_{1,lit} X_{1,i,t-l} + \xi_{1,it}$$

- Note that all predictors are centered i.e. $Y_{1,i,t-l}$ is not $Y_{i,t-l}$ (covariates X are optional)
- The model is a combination of the state space model with the dynamic factor model, merged with full SEM functionality, e.g., covariates, path analysis, regression among latent variables, CFA, MIMIC, correlated uniqueness, lagged loadings, in addition to the core extensions of two-level and cross-classified modeling as well as categorical variables.

DSEM model

- The usual structural equations at level 2 and 3.

$$Y_{2,i} = \nu_2 + \Lambda_2 \eta_{2,i} + \varepsilon_{2,i}$$

$$\eta_{2,i} = \alpha_2 + B_2 \eta_{2,i} + \Gamma_2 x_{2,i} + \xi_{2,i}$$

$$Y_{3,t} = \nu_3 + \Lambda_3 \eta_{3,t} + \varepsilon_{3,t}$$

$$\eta_{3,t} = \alpha_3 + B_3 \eta_{3,t} + \Gamma_3 x_t + \xi_{3,t}$$

- These include not just between parts of Y_{it} but also observed between level variables

Residual DSEM (RDSEM), available in future Mplus release

No change in the between level model. The within level model further splits the autoregressive and the structural part

$$Y_{1,it} = Y_{0,it} + \hat{Y}_{1,it}$$

$$\eta_{1,it} = \eta_{0,it} + \hat{\eta}_{1,it}$$

- The variables $Y_{0,it}$ and $\eta_{0,it}$ represent the linear predictor part (no random element)
- The variables $\hat{Y}_{1,it}$ and $\hat{\eta}_{1,it}$ represent the auto-regressive part and can be thought of as being the residuals

Residual DSEM model

The linear predictor model for $Y_{0,it}$ and $\eta_{0,it}$

$$Y_{0,it} = \nu_1 + \sum_{l=0}^{L} K_{1,lit} X_{1,i,t-l}$$

$$\eta_{0,it} = \alpha_{1,it} + \sum_{l=0}^{L} \Gamma_{1,lit} X_{1,i,t-l}$$

The auto-regressive model for $\hat{Y}_{1,it}$ and $\hat{\eta}_{1,it}$

$$\hat{Y}_{1,it} = \sum_{l=0}^{L} \Lambda_{1,lit} \hat{\eta}_{1,i,t-l} + \sum_{l=0}^{L} R_{lit} \hat{Y}_{1,i,t-l} + \varepsilon_{1,it}$$

$$\hat{\eta}_{1,it} = \sum_{l=0}^{L} B_{1,lit} \hat{\eta}_{1,i,t-l} + \sum_{l=0}^{L} Q_{lit} \hat{Y}_{1,i,t-l} + \xi_{1,it}$$

ARMA(1,1) and the measurement error AR(1) models

- The ARMA(1,1) model

$$Y_t = \mu + \beta Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

$$\sigma = Var(\varepsilon_t)$$

- The measurement error AR(1) = MEAR(1)

$$Y_t = \mu + f_t + \varepsilon_t$$

$$f_t = \beta f_{t-1} + \xi_t$$

$$\sigma_1 = Var(\varepsilon_t), \sigma_2 = Var(\xi_t)$$

- The two models are equivalent

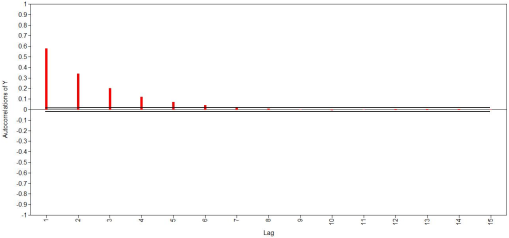
$$\sigma_1 = -\frac{\theta\sigma}{\beta}$$

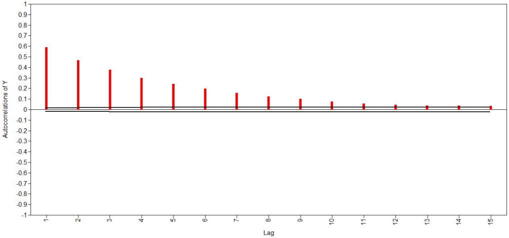
$$\sigma_2 = (1+\theta^2)\sigma + \frac{(1+\beta^2)\theta\sigma}{\beta}$$

ARMA(1,1) and the measurement error AR(1) models

- ... as long as $\sigma_1 > 0$ and $\sigma_2 > 0$
- The MEAR(1) shows how traditional SEM logic doesn't hold for DSEM: one indicator factor model is perfectly identified
- Two reasons to prefer MEAR(1) v.s. ARMA(1,1)
  - More efficient Mplus estimation
  - Easier to interpret - SEM like flavor
- Two reasons to prefer MEAR(1)/ARMA(1,1) v.s. AR(1)
  - AR(1) exponential decay of autocorrelation is not realistic
  - ARMA(1,1) is a two-parameter fit for the autocorrelation function , v.s., AR(1) which is one parameter
- Easy to test MEAR(1)/ARMA(1,1) v.s. AR(1) using significance of parameter.

AR(1) v.s. ARMA(1,1) autocorrelation decay





ARMA(1,1) simulation study

Table: Bias(coverage) MEAR(1) / ARMA(1,1), N=1

parameter	True value	$T = 100$	$T = 200$	$T = 300$	$T = 500$
μ	0	-.09(.82)	-.01(.89)	-.04(.85)	-.02(.87)
β	.8	-.07(.96)	-.04(.92)	-.03(.87)	-.01(.95)
σ_1	1	-.10(.97)	-.09(.94)	-.08(.88)	-.04(.90)
σ_2	1	.25(.95)	.17(.92)	.14(.91)	.08(.90)

- $T \geq 200$ recommended for small bias and acceptable coverage levels
- For two-level models smaller T are acceptable as long as not all four parameters are subject specific - typically σ_1 and σ_2 will not be subject specific

How to add a covariate in ARMA(1,1) and AR(1) models

- The same applies for AR(1) and ARMA(1,1)/MEAR(1). Three ways to do it
- Direct model

$$Y_t = \mu + f_t + \beta_1 X_t + \xi_t$$

$$f_t = \phi f_{t-1} + \varepsilon_t.$$

- Indirect model

$$Y_t = \mu + f_t + \xi_t$$

$$f_t = \phi f_{t-1} + \beta_2 X_t + \varepsilon_t.$$

- Full model

$$Y_t = \mu + f_t + \beta_1 X_t + \xi_t$$

$$f_t = \phi f_{t-1} + \beta_2 X_t + \varepsilon_t.$$

How to add a covariate in ARMA(1,1) and AR(1) models

- Consider the fundamental difference between the models by what it implies for $E(Y_t|X)$
- Direct model - no effect beyond the last value of X

$$E(Y_t|X) = \mu + \beta_1 X_t.$$

- Indirect model - accumulation effect of X with diminishing effects

$$E(Y_t|X) = \mu + \beta_2(X_t + \phi X_{t-1} + \phi^2 X_{t-2} + \phi^3 X_{t-3} + ...).$$

- Full model - direct and accumulated effect

$$E(Y_t|X) = \mu + \beta_1 X_t + \beta_2(X_t + \phi X_{t-1} + \phi^2 X_{t-2} + \phi^3 X_{t-3} + ...).$$

How to add a covariate in ARMA(1,1) and AR(1) models. The special case of $X_t = t$, linear growth model

- Hamaker, E.L. (2005) Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. Sociological Methods and Research, 33, 3, 404 - 418.
- It is shown in this paper that the direct and the indirect models are equivalent, i.e., the full model is not identified

How to add a covariate in ARMA(1,1) and AR(1) models. The special case of $X_t = t$, linear growth model

- The direct linear growth AR(1) model is

$$Y_t = \gamma_0 + \gamma_1 t + \xi_t$$
$$\xi_t = \phi \xi_{t-1} + \varepsilon_t$$

- The indirect linear growth AR(1) model is

$$Y_t = \beta_0 + \beta_1 t + \phi Y_{t-1} + \varepsilon_t$$

- Here t affects Y_t through it's effect on Y_{t-1} in addition to the direct effect β_1

$$Y_{t-1} = \beta_0 + \beta_1 (t-1) + \phi Y_{t-2} + \varepsilon_{t-1}$$

- The two models are algebraically equivalent and the full model is unidentified

$$\gamma_0 = \frac{\beta_0}{1-\phi} - \frac{\phi \beta_1}{(1-\phi)^2}, \quad \gamma_1 = \frac{\beta_1}{1-\phi}$$

AR(1) quadratic growth model, $X_t = (t, t^2)$

- The direct quadratic growth AR(1) model is

$$Y_t = \gamma_0 + \gamma_1 t + \gamma_2 t^2 + \xi_t$$
$$\xi_t = \phi \xi_{t-1} + \varepsilon_t$$

- The indirect quadratic growth AR(1) model is

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \phi Y_{t-1} + \varepsilon_t$$

- The two models are algebraically equivalent and the full model is unidentified

$$\gamma_0 = \frac{\beta_0}{1-\phi} - \frac{\phi \beta_1}{(1-\phi)^2} + \frac{\beta_2 \phi (1+\phi)}{(1-\phi)^3}$$
$$\gamma_1 = \frac{\beta_1}{1-\phi} - \frac{2\phi \beta_2}{(1-\phi)^2}$$
$$\gamma_2 = \frac{\beta_2}{1-\phi}$$

Two-level ARMA(1,1) full model simulation study

$$Y_{it} = \mu_i + f_t + \beta_1 X_{it} + \xi_{it}$$

$$f_{it} = \phi f_{i,t-1} + \beta_2 X_{it} + \varepsilon_{it}$$

$$\mu_i \sim N(\mu, \sigma_b)$$

- $\beta_1 = 0.3$, $\beta_2 = 0.4$, $\phi = 0.5$, $\mu = 0$, $\sigma_b = 0.7$, $Var(\xi_{it}) = Var(\varepsilon_{it}) = 1$
- The covariate is generated using AR(1) process with $Var(X_{it}) = 1$ and autocorrelation $r_x = 0, 0.5, 0.8$
- We analyze the data using the full, direct and indirect models

Two-level ARMA(1,1) simulation study - full model results

Table: Two-level full ARMA(1,1) with covariate, N=200, T=100

parameter	r_x	True value	Estimate(Coverage)
β_1	0	.30	.30(.87)
β_1	0.5	.30	.30(.96)
β_1	0.8	.30	.31(.89)
β_2	0	.40	.40(.87)
β_2	0.5	.40	.40(.93)
β_2	0.8	.40	.40(.90)
ϕ	0	.50	.50(.88)
ϕ	0.5	.50	.50(.93)
ϕ	0.8	.50	.50(.93)

No bias. Good coverage. Model is well identified.

Two-level ARMA(1,1) simulation study - direct model results

Table: Two-level full ARMA(1,1) with covariate analyzed as direct, N=200, T=100

parameter	r_x	True value	Estimate(Coverage)
β_1	0	.70	.65(.00)
β_1	0.5	.70	.74(.07)
β_1	0.8	.70	.88(.00)
ϕ	0	.50	.50(.92)
ϕ	0.5	.50	.51(.85)
ϕ	0.8	.50	.52(.83)

Both parameters are biased. Coverage is low. Bias depends on r_x.

Two-level ARMA(1,1) simulation study - indirect model results

Table: Two-level full ARMA(1,1) with covariate analyzed as indirect, N=200, T=100

parameter	r_x	True value	Estimate(Coverage)
β_2	0	.70	.69(.92)
β_2	0.5	.70	.67(.21)
β_2	0.8	.70	.65(.07)
ϕ	0	.50	.36(.00)
ϕ	0.5	.50	.38(.00)
ϕ	0.8	.50	.41(.00)

Both parameters are biased. Coverage is low. Bias depends on r_x.

How to add a covariate in ARMA(1,1) and AR(1) models

- The direct and the indirect models are not a reparameterization of each other.
- The dilemma for direct and the indirect models should not be resolved based on what is simpler to interpret, software availability, or tradition - rather the data should decide that.
- For the AR(1) model the issue is more complicated as the direct model requires fixing $Var(\xi_i)$ to 0 (small value near zero) which leads to slow convergence, i.e., we have to specify the model as MEAR(1)
- For the AR(1) model the Residual DSEM is the appropriate approach - fast convergence, no fixing residuals to zero, no MEAR(1) usage. RDSEM absorbs the complexity - more complicated to estimate than DSEM.

Simulation study: direct two-level AR(1) model

$$T = 50, N = 50, rep = 50, X_{it} \sim AR(1), r_x = 0.5$$
$$Y_{it} = \mu_i + \beta X_{it} + f_{it}$$
$$f_{it} = \phi f_{it-1} + \varepsilon_{it}$$

- Three methods: DSEM-MEAR(1) with fixed residual variance to a small value, DSEM with free residual variance, RDSEM (new input)

```
MODEL:

%WITHIN%
y^ on y^1*0.7;
y on x*1; y*1;

%BETWEEN%
y*1; [y*3];
```

Simulation study: direct two-level AR(1) model

Table: Two-level AR(1) with covariate bias(coverage)

method	DSEM-fixed	DSEM-free	RDSEM
β	.00(.60)	.00(.55)	.00(.90)
ϕ	.03(.62)	.01(.92)	.00(.94)
time per rep in sec	3	33	0.5

Clearly RDSEM outperforms DSEM for this model due to quality of mixing - only 200 MCMC iterations until convergence.

Linear growth AR(1) simulation

In this example we illustrate the following 6 concepts

- Equivalence of direct and indirect AR(1) models
- The dependence of Yule-Walker output: residual/tech4/stand on the stationarity of the autoregressive portion of the model
- How to setup the model so that the trend is not included in the autoregressive portion of the model using the MEAR concept for the purpose of obtaining correct Yule-Walker output
- We illustrate the advantages of RDSEM
- Challenges with fixing variance to zero/small values
- We illustrate 4 ways to run the model in Mplus and obtain the exact same model estimates

We generate data 500 points using the indirect linear growth model

$$Y_t = \beta_0 + \beta_1 t + \phi Y_{t-1} + \varepsilon_t$$

$\beta_0 = 1$, $\beta_1 = 0.3$, $\phi = 0.5$, $\theta = Var(\varepsilon_t) = 1$, $t = 0.1, 0.2, 0.3, ..., 50.$

Linear growth AR(1) simulation

Estimate the following 4 equivalent models

- Indirect linear growth, followed by model parameter transformation, using DSEM

$$Y_t = \beta_0 + \beta_1 t + \phi Y_{t-1} + \varepsilon_t$$

- Direct linear growth using RDSEM

$$Y_t = \gamma_0 + \gamma_1 t + \varepsilon_t$$

$$\varepsilon_t = \phi \varepsilon_{t-1} + \xi_t$$

- MEAR(1) model with $Var(\zeta_t)$ fixed to a small value 0.05, DSEM

$$Y_t = \gamma_0 + \gamma_1 t + f_t + \varepsilon_t$$

$$f_t = \phi f_{t-1} + \xi_t$$

- MEAR(1) model with $Var(\zeta_t)$ estimated as a free parameter, DSEM

Indirect linear growth AR(1)

```
variable:   names=y t;
            lagged=y(1);
data:       file=1.dat;

analysis:   estimator=bayes;
            proc=2; fbiter=50000;

model:
 y on y&1 (phi)
      t   (b1);
 [y] (b0);

model constraint: new(g0 g1);
 g0=b0/(1-phi)-0.1*b1*phi/(1-phi)**2;
 g1=b1/(1-phi);

output: residual;
```

Direct linear growth AR(1)

```
variable: names = y t;
          lagged=y(1);

data: file=1.dat;

analysis: estimator=bayes;
          proc=2; fbiter=50000;

model:
 y on t*0.6;
 y^ on y^1;
 [y*0];

output: residual;
```

MEAR(1) linear growth (free)

```
variable:  NAMES ARE y t;

data: file=1.dat;

analysis:  estimator=bayes;
           proc=2; fbiter=50000; thin=50;

model:
 y on t;
 f on f&1*0.5 ;
 f by y (&1);
 y*0.05;
 [y*0];

output: residual;
```

Linear growth AR(1) simulation: results

	True value	Indirect DSEM Full AR	Direct RDSEM Resid AR	MEAR(1) free Full AR	MEAR(1) fixed Full AR
convergence		fast	fast	slow	slow
γ_0	1.94	1.94(.18)	1.94(.19)	1.93(.19)	1.93(.19)
γ_1	.6	.6(.006)	.6(.006)	.6(.006)	.6(.006)
ϕ	.5	.51(.04)	.51(.04)	.53(.04)	.53(.04)
θ	1	1(.06)	1(.06)	.93 (.09)	.94(.06)
$E(Y)$	16.94	23.33	16.94	16.94	16.94
$Var(Y)$	76.23	67.06	76.05	76.12	76.17
DIC		1421	1420	N/A	N/A

Linear growth AR(1) simulation: conclusions

- All four methods produce the same model estimates, MEAR models produce approximation
- Indirect AR(1) DSEM model yields incorrect residual / tech4 / stand results due to violating the assumptions of Yule-Walker
- MEAR(1) free is preferred over MEAR(1) fixed because you don't need to decide on the small value, also safeguard against the small variance not being zero, i.e., safeguard against the model not being AR(1) but rather ARMA(1,1)
- MEAR(1) formulation shows slow convergence, however, in real data sets most likely the residual variance will not be exactly zero / much better convergence
- MEAR(1) methods do not produce usable DIC due to conditioning on the within level factor
- Fixing variances to zero is not harmless in Bayes estimation as it is in ML - most likely causing poor mixing and convergence
- RDSEM is the best method - not available in Mplus 8, available in Mplus 8.1

Comparing Bayesian estimation of DSEM and RDSEM

The RDSEM estimation is more complex than the DSEM estimation. Consider this lag 2 model

$$Y_t = \alpha + BX_t + R_1 Y_{t-1} + R_2 Y_{t-2} + \varepsilon_t$$

the posterior distribution for B is as in a standard regression coefficient given all other quantities. Consider now the RDSEM model

$$Y_t = \alpha + BX_t + \hat{Y}_t$$

$$\hat{Y}_t = R_1 \hat{Y}_{t-1} + R_2 \hat{Y}_{t-2} + \varepsilon_t$$

$$Y_t = \alpha + BX_t + R_1(Y_{t-1} - \alpha - BX_{t-1}) + R_2(Y_{t-2} - \alpha - BX_{t-2}) + \varepsilon_t$$

$$Y_t = (I - R_1 - R_2)\alpha + BX_t + R_1 Y_{t-1} - R_1 BX_{t-1} + R_2 Y_{t-2} - R_2 BX_{t-2} + \varepsilon_t$$

New sufficient statistics are needed. Matrix multiplication causes a single parameter to end up in many of the regression coefficients combinations. Solution: bayesian estimation of regression coefficients under linear constraints. Similar complications arise for latent variables and missing data updating.

New models coming in future Mplus release

- Residual DSEM
- Bayesian Multilevel Mixture Models
- Bayesian Multilevel Latent Transition Models with cluster specific transition probabilities
- Multilevel Mixtures of Dynamic Structural Equation Models
- Dynamic Latent Class Analysis
- Multilevel Hidden Markov Model
- Multilevel Markov Switching Autoregressive Models
- Multilevel Markov Switching DSEM Models
- These are at various stages of development.
- Asparouhov, T., Hamaker, E.L. & Muthen, B. (2017). Dynamic Latent Class Analysis, Structural Equation Modeling: A Multidisciplinary Journal, 24:2, 257-269