Modeling misspecification as a parameter in Bayesian structural equation models

James Ohisei Uanhoro
james.uanhoro@unt.edu

May 4, 2023

Department of Educational Psychology
University of North Texas

Outline

Introduction

Model

Simulation studies

Closing

Introduction

Study context

Misspecification is active research area for Bayesian SEMs

- Levy (2011): Posterior predictive checking using the likelihood ratio test; Bayesian SRMR
- Hoofs, van de Schoot, Jansen, and Kant (2018): Bayesian RMSEA
- Garnier-Villarreal and Jorgensen (2020): Several translations of frequentist fit indices: RMSEA, CFI, TLI, ...
- Cain and Zhang (2019): Deviance information criterion

Current approach

We want to model misspecification

- Wu and Browne (2015) have done this in frequentist context:
    - Parameter estimates have greater uncertainty about them reflecting degree of model misspecification
    - Estimates the RMSEA as a model parameter
- We could simply recreate Wu and Browne (2015) in a Bayesian context, but:
    - The RMSEA is not so easy to interpret (e.g. Chen, Curran, Bollen, Kirby, & Paxton, 2008; Savalei, 2012)
    - The SRMR has clearer interpretations.
    - Model misspecification so parameter estimates reflect degree of model incorrectness, while estimating an SRMR-type index.

Current approach

We want to model misspecification

- Wu and Browne (2015) have done this in frequentist context:
  - Parameter estimates have greater uncertainty about them reflecting degree of model misspecification
  - Estimates the RMSEA as a model parameter
- We could simply recreate Wu and Browne (2015) in a Bayesian context, but:
  - The RMSEA is not so easy to interpret (e.g. Chen et al., 2008; Savalei, 2012)
  - The SRMR has clearer interpretations.
  - Model misspecification so parameter estimates reflect degree of model incorrectness, while estimating an SRMR-type index.

Current approach

We want to model misspecification

- Wu and Browne (2015) have done this in frequentist context:
  - Parameter estimates have greater uncertainty about them reflecting degree of model misspecification
  - Estimates the RMSEA as a model parameter
- We could simply recreate Wu and Browne (2015) in a Bayesian context, but:
  - The RMSEA is not so easy to interpret (e.g. Chen et al., 2008; Savalei, 2012)
  - The SRMR has clearer interpretations.
  - Model misspecification so parameter estimates reflect degree of model incorrectness, while estimating an SRMR-type index.

Current approach

We want to model misspecification

- Wu and Browne (2015) have done this in frequentist context:
  - Parameter estimates have greater uncertainty about them reflecting degree of model misspecification
  - Estimates the RMSEA as a model parameter
- We could simply recreate Wu and Browne (2015) in a Bayesian context, but:
  - The RMSEA is not so easy to interpret (e.g. Chen et al., 2008; Savalei, 2012)
  - The SRMR has clearer interpretations.
  - Model misspecification so parameter estimates reflect degree of model incorrectness, while estimating an SRMR-type index.

Current approach

We want to model misspecification

- Wu and Browne (2015) have done this in frequentist context:
  - Parameter estimates have greater uncertainty about them reflecting degree of model misspecification
  - Estimates the RMSEA as a model parameter
- We could simply recreate Wu and Browne (2015) in a Bayesian context, but:
  - The RMSEA is not so easy to interpret (e.g. Chen et al., 2008; Savalei, 2012)
  - The SRMR has clearer interpretations.
  - Model misspecification so parameter estimates reflect degree of model incorrectness, while estimating an SRMR-type index.

Current approach

We want to model misspecification

- Wu and Browne (2015) have done this in frequentist context:
  - Parameter estimates have greater uncertainty about them reflecting degree of model misspecification
  - Estimates the RMSEA as a model parameter
- We could simply recreate Wu and Browne (2015) in a Bayesian context, but:
  - The RMSEA is not so easy to interpret (e.g. Chen et al., 2008; Savalei, 2012)
  - The SRMR has clearer interpretations.
  - Model misspecification so parameter estimates reflect degree of model incorrectness, while estimating an SRMR-type index.

Current approach

We want to model misspecification

- Wu and Browne (2015) have done this in frequentist context:
  - Parameter estimates have greater uncertainty about them reflecting degree of model misspecification
  - Estimates the RMSEA as a model parameter
- We could simply recreate Wu and Browne (2015) in a Bayesian context, but:
  - The RMSEA is not so easy to interpret (e.g. Chen et al., 2008; Savalei, 2012)
  - The SRMR has clearer interpretations.
  - Model misspecification so parameter estimates reflect degree of model incorrectness, while estimating an SRMR-type index.

Model

CFA as an example

Begin with Muthén and Asparouhov (2012):

$$\boldsymbol{\Sigma} = \underbrace{\boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^{\mathsf{T}} + \boldsymbol{\Delta}}_{\text{Standard}} + \boldsymbol{\Psi} \tag{1}$$

$\boldsymbol{\Lambda}$: Loading matrix, $\boldsymbol{\Phi}$: Interfactor correlation matrix, $\boldsymbol{\Delta}$: Standard residual covariance matrix (often diagonal)

$\boldsymbol{\Psi}$: Residual covariance matrix, with all off-diagonal elements estimated. Assumed to reflect the influence of minor factors (MacCallum & Tucker, 1991)

$\boldsymbol{\Psi}$ is not identified: Muthén and Asparouhov (2012) apply an inverse-Wishart prior with known parameters to attempt to shrink elements in $\boldsymbol{\Psi}$ to zero.

CFA as an example

Begin with Muthén and Asparouhov (2012):

$$\mathbf{\Sigma} = \underbrace{\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathsf{T}} + \mathbf{\Delta}}_{\text{Standard}} + \mathbf{\Psi} \tag{1}$$

$\mathbf{\Lambda}$: Loading matrix, $\mathbf{\Phi}$: Interfactor correlation matrix, $\mathbf{\Delta}$: Standard residual covariance matrix (often diagonal)

$\mathbf{\Psi}$: Residual covariance matrix, with all off-diagonal elements estimated. Assumed to reflect the influence of minor factors (MacCallum & Tucker, 1991)

$\mathbf{\Psi}$ is not identified: Muthén and Asparouhov (2012) apply an inverse-Wishart prior with known parameters to attempt to shrink elements in $\mathbf{\Psi}$ to zero.

CFA as an example

Begin with Muthén and Asparouhov (2012):

$$\mathbf{\Sigma} = \underbrace{\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathsf{T}} + \mathbf{\Delta}}_{\text{Standard}} + \mathbf{\Psi} \tag{1}$$

$\mathbf{\Lambda}$: Loading matrix, $\mathbf{\Phi}$: Interfactor correlation matrix, $\mathbf{\Delta}$: Standard residual covariance matrix (often diagonal)

$\mathbf{\Psi}$: Residual covariance matrix, with all off-diagonal elements estimated. Assumed to reflect the influence of minor factors (MacCallum & Tucker, 1991)

$\mathbf{\Psi}$ is not identified: Muthén and Asparouhov (2012) apply an inverse-Wishart prior with known parameters to attempt to shrink elements in $\mathbf{\Psi}$ to zero.

Goal is to model misspecification

$$\boldsymbol{\Sigma} = \underbrace{\boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^{\mathsf{T}} + \boldsymbol{\Delta}}_{\text{Standard}} + \boldsymbol{\Psi}$$

Let ψ_{ij} $(i \neq j)$ be off diagonal elements in $\boldsymbol{\Psi}$, reflecting misspecification / minor factor influences.

To model minor factor influences:

$$\underbrace{\frac{\psi_{ij}}{\sqrt{\sigma_{jj}\sigma_{ii}}}}_{\text{SRCs}} \sim \mathcal{N}(0, \tau), \ \tau \sim \mathcal{N}^{+}(0, 1)$$

$\sigma_{ii/jj}$: indicator variances i.e. SRCs: standardized residual covariances.

Goal is to model misspecification

$$\mathbf{\Sigma} = \underbrace{\mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}^{\mathsf{T}} + \mathbf{\Delta}}_{\text{Standard}} + \mathbf{\Psi}$$

Let ψ_{ij} ($i \neq j$) be off diagonal elements in $\mathbf{\Psi}$, reflecting misspecification / minor factor influences.

To model minor factor influences:

$$\underbrace{\frac{\psi_{ij}}{\sqrt{\sigma_{jj}\sigma_{ii}}}}_{\text{SRCs}} \sim \mathcal{N}(0, \tau),\ \tau \sim \mathcal{N}^{+}(0, 1)$$

$\sigma_{ii/jj}$: indicator variances i.e. SRCs: standardized residual covariances.

A hierarchical model for SRCs

$$\underbrace{\frac{\psi_{ij}}{\sqrt{\sigma_{jj}\sigma_{ii}}}}_{\text{SRCs}} \sim \mathcal{N}(0, \tau), \ \tau \sim \mathcal{N}^{+}(0, 1)$$

Implications

- SRCs are assumed to be zero on average
- SRC scale parameter (τ) is the root mean square error of SRCs with location constrained to be zero
- τ is akin to the correlation root mean square residual (CRMR)
- Smaller values of τ reflect a model with smaller SRCs, lower misspecification OR smaller influence of minor factors
- $\tau < 0.05 \implies$ most SRCs < 0.10

A hierarchical model for SRCs

$$\underbrace{\frac{\psi_{ij}}{\sqrt{\sigma_{jj}\sigma_{ii}}}}_{\text{SRCs}} \sim \mathcal{N}(0, \tau),\ \tau \sim \mathcal{N}^{+}(0, 1)$$

Implications

- SRCs are assumed to be zero on average
- SRC scale parameter (τ) is the root mean square error of SRCs with location constrained to be zero
- τ is akin to the correlation root mean square residual (CRMR)
- Smaller values of τ reflect a model with smaller SRCs, lower misspecification OR smaller influence of minor factors
- $\tau < 0.05 \implies$ most SRCs < 0.10

A hierarchical model for SRCs

$$\underbrace{\frac{\psi_{ij}}{\sqrt{\sigma_{jj}\sigma_{ii}}}}_{\text{SRCs}} \sim \mathcal{N}(0, \tau),\ \ \tau \sim \mathcal{N}^{+}(0, 1)$$

Implications

- SRCs are assumed to be zero on average
- SRC scale parameter (τ) is the root mean square error of SRCs with location constrained to be zero
- τ is akin to the correlation root mean square residual (CRMR)
- Smaller values of τ reflect a model with smaller SRCs, lower misspecification OR smaller influence of minor factors
- $\tau < 0.05 \implies$ most SRCs < 0.10

Simulation studies

Three simulation studies

List of simulation studies

1. Can the proposed approach recover model parameters?
   - Yes, upcoming
2. Can differences in τ between competing models be used for model selection?
   - Lacks power, more promising to use approximate leave-one-out cross-validation to select between competing models.
3. Can the Ψ matrix be used to detect specific residual covariances that are too large?
   - Yes, especially for larger samples

Three simulation studies

List of simulation studies

1. Can the proposed approach recover model parameters?
    - Yes, upcoming
2. Can differences in τ between competing models be used for model selection?
    - Lacks power, more promising to use approximate leave-one-out cross validation to select between competing models.
3. Can the Ψ matrix be used to detect specific residual covariances that are too large?
    - Yes, especially for larger samples.

List of simulation studies

1. Can the proposed approach recover model parameters?
   - Yes, upcoming
2. Can differences in τ between competing models be used for model selection?
   - Lacks power, more promising to use approximate leave-one-out cross validation to select between competing models.
3. Can the Ψ matrix be used to detect specific residual covariances that are too large?
   - Yes, especially for larger samples.

Three simulation studies

List of simulation studies

1. Can the proposed approach recover model parameters?
   - Yes, upcoming
2. Can differences in τ between competing models be used for model selection?
   - Lacks power, more promising to use approximate leave-one-out cross validation to select between competing models.
3. Can the Ψ matrix be used to detect specific residual covariances that are too large?
   - Yes, especially for larger samples.

Three simulation studies

List of simulation studies

1. Can the proposed approach recover model parameters?
   - Yes, upcoming
2. Can differences in τ between competing models be used for model selection?
   - Lacks power, more promising to use approximate leave-one-out cross validation to select between competing models.
3. Can the Ψ matrix be used to detect specific residual covariances that are too large?
   - Yes, especially for larger samples.

Three simulation studies

List of simulation studies

1. Can the proposed approach recover model parameters?
    - Yes, upcoming
2. Can differences in τ between competing models be used for model selection?
    - Lacks power, more promising to use approximate leave-one-out cross validation to select between competing models.
3. Can the $\mathbf{\Psi}$ matrix be used to detect specific residual covariances that are too large?
    - Yes, especially for larger samples.

Three simulation studies

List of simulation studies

1. Can the proposed approach recover model parameters?
    - Yes, upcoming
2. Can differences in τ between competing models be used for model selection?
    - Lacks power, more promising to use approximate leave-one-out cross validation to select between competing models.
3. Can the $\boldsymbol{\Psi}$ matrix be used to detect specific residual covariances that are too large?
    - Yes, especially for larger samples.

Study 1: Recovery of structural parameters

Simulate data with minor factor influences ($\mathbf{\Psi}$)

5 modeling approaches

- Two **ignore** minor factor influences:
  1. *Freq*: Standard frequentist
  2. *Baseline*: Standard Bayesian
- One **fixes** the size of minor factor influences:
  3. *AZ*: Muthén and Asparouhov (2012)
- Two **estimate** the size of minor factor influences:
  4. *LKJ*: another proposed method (in paper)
  5. *NRM*: hierarchical normal method (focus here)

Results: Bias adequate; Coverage poor for methods that ignore or fix the size of the influence of minor factors

Study 1: Recovery of structural parameters

Simulate data with minor factor influences ($\boldsymbol{\Psi}$)

5 modeling approaches

- Two **ignore** minor factor influences:
  1. *Freq*: Standard frequentist
  2. *Baseline*: Standard Bayesian
- One **fixes** the size of minor factor influences:
  3. *AZ*: Muthén and Asparouhov (2012)
- Two **estimate** the size of minor factor influences:
  4. *LKJ*: another proposed method (in paper)
  5. *NRM*: hierarchical normal method (focus here)

Results: Bias adequate; Coverage poor for methods that ignore or fix the size of the influence of minor factors

Study 1: Recovery of structural parameters

Simulate data with minor factor influences ($\mathbf{\Psi}$)

5 modeling approaches

- Two **ignore** minor factor influences:
  1. *Freq*: Standard frequentist
  2. *Baseline*: Standard Bayesian
- One **fixes** the size of minor factor influences:
  3. *AZ*: Muthén and Asparouhov (2012)
- Two **estimate** the size of minor factor influences:
  4. *LKJ*: another proposed method (in paper)
  5. *NRM*: hierarchical normal method (focus here)

Results: Bias adequate; Coverage poor for methods that ignore or fix the size of the influence of minor factors

Study 1: Recovery of structural parameters

Simulate data with minor factor influences ($\mathbf{\Psi}$)

5 modeling approaches

- Two **ignore** minor factor influences:
  1. *Freq*: Standard frequentist
  2. *Baseline*: Standard Bayesian
- One **fixes** the size of minor factor influences:
  3. *AZ*: Muthén and Asparouhov (2012)
- Two **estimate** the size of minor factor influences:
  4. *LKJ*: another proposed method (in paper)
  5. *NRM*: hierarchical normal method (focus here)

Results: Bias adequate; Coverage poor for methods that ignore or fix the size of the influence of minor factors

Study 1: Recovery of structural parameters

Simulate data with minor factor influences ($\boldsymbol{\Psi}$)

5 modeling approaches

- Two **ignore** minor factor influences:
  1. *Freq*: Standard frequentist
  2. *Baseline*: Standard Bayesian
- One **fixes** the size of minor factor influences:
  3. *AZ*: Muthén and Asparouhov (2012)
- Two **estimate** the size of minor factor influences:
  4. *LKJ*: another proposed method (in paper)
  5. *NRM*: hierarchical normal method (focus here)

Results: Bias adequate; Coverage poor for methods that ignore or fix the size of the influence of minor factors

Study 1: Recovery of structural parameters

Simulate data with minor factor influences ($\mathbf{\Psi}$)

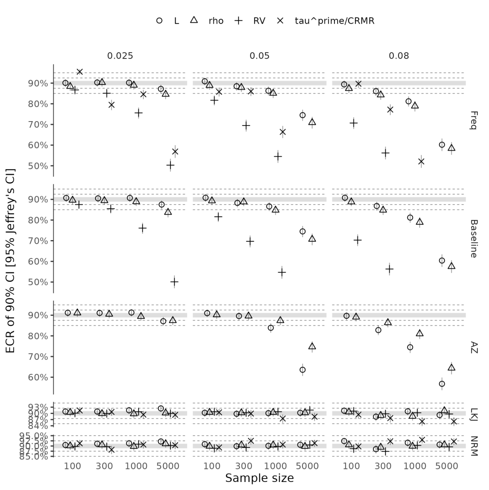
5 modeling approaches

- Two **ignore** minor factor influences:
  1. *Freq*: Standard frequentist
  2. *Baseline*: Standard Bayesian
- One **fixes** the size of minor factor influences:
  3. *AZ*: Muthén and Asparouhov (2012)
- Two **estimate** the size of minor factor influences:
  4. *LKJ*: another proposed method (in paper)
  5. *NRM*: hierarchical normal method (focus here)

Results: Bias adequate; Coverage poor for methods that ignore or fix the size of the influence of minor factors

Study 1: Empirical coverage rate of 90% CI.

Coverage of structural parameters and τ'



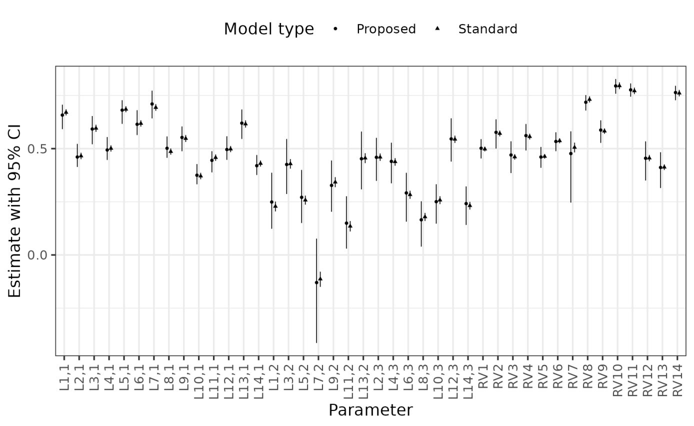
L: loading; RV: residual variance; rho: interfactor correlation

Findings:

- *Freq* and *Baseline* models fail when sample size or τ is large
- Performance of *AZ* model deteriorates at large τ
- Methods that estimate the size of minor factor influences perform as expected

An empirical example: Hospital Anxiety and Depression scale

Bifactor model parameters, $n = 21820, \tau = 0.028$. Note narrow intervals for the *Standard* approach.



LX.Y are loadings of factor Y on item X; RV: residual variances. Standard estimates are from a model without Ψ.

Closing

Discussion

Discussion points / future work

- There is an R package for this: minorbsem
- Package also includes:
  - Bayesian implementation of Wu and Browne (2015) method.
  - Other prior assumptions for standardized residual covariances e.g. lasso, generalized double Pareto (global-local)
  - Global-local prior on all cross-loadings, relaxing simple structure
  - These features are yet to be tested via simulation
- Extensions to other data types are also worth considering
  - Additional issues/extensions are on GitHub: `https://github.com/jamesuanhoro/minorbsem/issues/3`

References

Cain, M. K., & Zhang, Z. (2019). Fit for a Bayesian: An evaluation of PPP and DIC for structural equation modeling. *Structural Equation Modeling*, *26*(1), 39–50. doi: 10.1080/10705511.2018.1490648

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008, May). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, *36*(4), 462–494. doi: 10.1177/0049124108314720

Garnier-Villarreal, M., & Jorgensen, T. D. (2020, February). Adapting fit indices for Bayesian structural equation modeling: Comparison to maximum likelihood. *Psychological Methods*, *25*(1), 46–70. doi: 10.1037/met0000224

Hoofs, H., van de Schoot, R., Jansen, N. W. H., & Kant, I. (2018, August). Evaluating model fit in Bayesian confirmatory factor analysis with large samples: Simulation study introducing the BRMSEA. *Educational and Psychological Measurement*, *78*(4), 537–568. doi: 10.1177/0013164417709314

Levy, R. (2011, October). Bayesian data-model fit assessment for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(4), 663–685. doi: 10.1080/10705511.2011.607723

MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in the common-factor model: Implications for theory and practice. *Psychological Bulletin*, *109*, 502–511. doi: 10.1037/0033-2909.109.3.502

Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*(3), 313–335.

doi: 10.1037/a0026802

Savalei, V. (2012, December). The relationship between root mean square error of approximation and model misspecification in confirmatory factor analysis models. *Educational and Psychological Measurement*, *72*(6), 910–932. doi: 10.1177/0013164412452564

Wu, H., & Browne, M. W. (2015, September). Quantifying adventitious error in a covariance structure as a random effect. *Psychometrika*, *80*(3), 571–600. doi: 10.1007/s11336-015-9451-3