Recent progress on Bayesian structural equation models

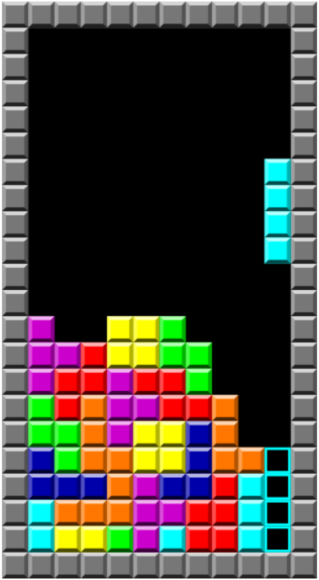
Ed Merkle

Michigan State, October 5, 2021

Introduction

- blavaan: An R package for Bayesian SEM, started in 2015.
- Initial goal: Automatically generate JAGS code from a lavaan object (focus on Bollen Political Democracy model).
- Subsequent goals are based on tricky problems encountered during development.

Subsequent Goals



Introduction

- Subsequent goals:
  - MCMC efficiency and handling of latent variables during model estimation
  - Use of different likelihoods for model comparison indices
  - Helping other researchers build on blavaan

Bayesian

- Why Bayesian SEM?
  - Include prior information/expectations in analyses
  - Address model convergence/sample size issues (prior information can help here)
  - Handle uncertainty: Ease of describing uncertainty in key results (latent variables, functions of parameters)
  - Flexibility/extensibility: As models increase in complexity, Bayesian methods can be easier to extend to new situations

Method comparison

- Traditional (frequentist):
  - Specify model
  - Fit model to data
  - Obtain *point* estimate and standard error for each model parameter.

Method comparison

- Bayesian:
  - Specify model
  - Specify prior beliefs (distributions) about model parameters
  - Fit model to data (Markov chain Monte Carlo)
  - Obtain posterior *distributions* for each model parameter, reflecting how prior beliefs are updated by data

Initial example

blavaan

(from Richard McElreath)

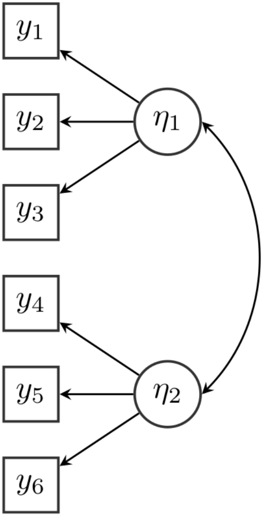
Example

- *blavaan* is intended to work like *lavaan*. We will demonstrate via a factor analysis model.

- `HolzingerSwineford1939` data from *lavaan*: Test scores of 301 students; 3 tests of visual perception and 3 tests of verbal comprehension.

- One factor underlying the visual tests, a second factor underlying the verbal tests.

Path diagram



lavaan

- Model specification and estimation in *lavaan*:

```
library("lavaan")

HS.model <- ' visual  =~ x1 + x2 + x3
              verbal  =~ x4 + x5 + x6 '

fit <- cfa(HS.model, data = HolzingerSwineford1939)
```

blavaan

- If you use all the defaults, *blavaan* is almost exactly the same:

```
library("blavaan")

HS.model <- ' visual  =~ x1 + x2 + x3
              verbal  =~ x4 + x5 + x6 '

bfit <- bcfa(HS.model, data = HolzingerSwineford1939)
```

blavaan

- But you shouldn't rely on defaults! blavaan provides functionality for things like
  - Choosing number of burnin (warmup) and sampling iterations.
  - Choosing your own prior distributions, and sampling from the priors.
  - Sampling the factors (latent variables), along with other parameters.
  - Assessing chain convergence.

Detailed example

Examples from Laird et al, 2003

Child Development, May/June 2003, Volume 74, Number 3, Pages 752–768

Parents' Monitoring-Relevant Knowledge and Adolescents' Delinquent Behavior: Evidence of Correlated Developmental Changes and Reciprocal Influences

Robert D. Laird, Gregory S. Pettit, John E. Bates, and Kenneth A. Dodge

Links between parental knowledge and adolescent delinquent behavior were tested for correlated rates of developmental change and reciprocal associations. For 4 years beginning at age 14, adolescents ($N = 396$) reported on their delinquent behavior and on their parents' knowledge of their whereabouts and activities. Parents completed measures of their adolescents' delinquent behavior. Knowledge was negatively correlated with delinquent behaviors at baseline, and increases over time in knowledge were negatively correlated with increases in parent-reported delinquent behavior. Reciprocal associations indicate that low levels of parental knowledge predict increases in delinquent behavior and that high levels of delinquent behavior predict decreases in knowledge. Discussion considers both youth-driven and parent-driven processes that may account for the correlated developmental changes and reciprocal associations.

Data

- At each of grades 9–12, adolescents and parents report on the adolescents' delinquent behavior problems. They also respond to questions that ask how much the parents know about the adolescents.
- We focus here on change in adolescent ratings of delinquent behavior over time.
- Adolescent scale scores range from 0–22; parent scale scores range from 0–26. Assumed continuous here, but item analysis is also possible!

Data

```
## read in data (R's working directory must be set to
##                the location of the data file)
dat <- read.csv("PDF2020_Data_Laird_CDP.csv")
```

Data I

```r
str(dat)
```

```
## 'data.frame':    585 obs. of  27 variables:
##  $ p12mon  : int  NA 13 11 NA NA 5 9 NA NA NA ...
##  $ a10know : int  NA 14 12 NA NA 7 NA NA NA NA ...
##  $ a11know : int  NA 14 9 NA NA NA 13 NA NA NA ...
##  $ y9del   : int  NA 10 7 NA NA 3 5 NA NA NA ...
##  $ y10del  : int  NA 4 11 NA NA 2 NA NA NA NA ...
##  $ y11del  : int  NA 4 11 NA NA NA 3 NA NA NA ...
##  $ y12del  : int  NA 6 9 NA NA 1 8 NA NA NA ...
##  $ m9del   : int  NA 3 6 NA NA 4 2 NA NA NA ...
##  $ m10del  : int  NA 4 17 NA NA 2 NA NA NA NA ...
##  $ m11del  : int  NA 4 22 NA NA NA 4 NA NA NA ...
##  $ m12del  : int  NA 3 4 NA NA 7 4 NA NA NA ...
##  $ a9knowr : int  NA 12 15 NA NA 9 13 NA NA NA ...
##  $ a12knowr: int  NA 12 6 NA NA 10 10 NA NA NA ...
##  $ y8del_l : num  NA 2.4 2.08 NA NA ...
##  $ y10del_l: num  NA 1.61 2.48 NA NA ...
##  $ y11del_l: num  NA 1.61 2.48 NA NA ...
##  $ y12del_l: num  NA 1.95 2.3 NA NA ...
##  $ m9del_l : num  NA 1.39 1.95 NA NA ...
##  $ m10del_l: num  NA 1.61 2.89 NA NA ...
##  $ m11del_l: num  NA 1.61 3.14 NA NA ...
##  $ m12del_l: num  NA 1.39 1.61 NA NA ...
##  $ d1tcsex : chr  NA "Male" "Male" NA ...
##  $ p9mon   : int  NA 13 13 NA NA 8 13 NA NA NA ...
##  $ d1med   : chr  "12 yrs." "12 yrs." "12 yrs." "12 yrs." ...
##  $ d1fed   : chr  "16-17 yrs." "16-17 yrs." NA "12 yrs." ...
##  $ d1tcrace: chr  "White" "White" "Black" "White" ...
##  $ d1ses   : num  63 50 17 32 34.5 14 32 31.5 36 37 ...
```

Strategy

- Modeling strategy: Latent growth curves from Grades 9 to 12.
  - Similar to mixed/hierarchical/multilevel models for repeated measures.
  - Describe how delinquent behavior changes over time, for adolescent reports.
  - Extensions (not presented here) could describe how parent reports correspond to adolescent reports, or how adolescent delinquent behavior corresponds to how much parents know about them.

Visualization

- Aside: Data visualization is always important, but SEM can make this difficult.
  - SEM typically wants data in wide format; graphing software wants data in long format.
  - So we need to do a conversion, which seldom works on the first try.
  - A side benefit of R: modeling and graphing all in one place.

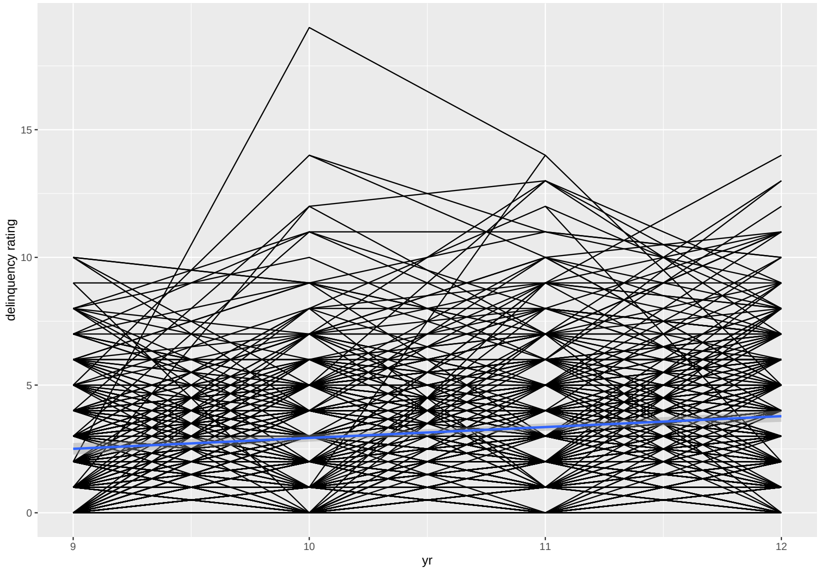
Visualization

- We need to do a wide-to-long conversion. There are many ways to do this; below will require some study but is very concise. (Converts both adolescent and parent delinquency ratings to long format.)

```
library("tidyr")

dat$ID <- 1:nrow(dat)
longdat <- pivot_longer(dat[,c(4:11,28)],
                        cols = c('y9del', 'y10del', 'y11del', 'y12del',
                                 'm9del', 'm10del', 'm11del', 'm12del'),
                        names_to = c("respondent", "yr", "qtype"),
                        names_pattern = "(.)([[:digit:]]{1,2})(.*)",
                        values_to = "value")
longdat$yr <- as.numeric(longdat$yr)
```

```r
library("ggplot2")
ggplot(longdat[longdat$respondent=="y",], aes(x = yr, y = value)) +
  geom_line(aes(group = ID)) + geom_smooth(method='lm') + ylab("delinquency rating")
```



Adolescents' growth model

```r
## a model that uses too many defaults
library("blavaan")

model <- ' i =~ 1*y9del + 1*y10del + 1*y11del + 1*y12del
           s =~ 0*y9del + 1*y10del + 2*y11del + 3*y12del '

## this spits out lots of intermediate output (shows sampling progress):
fit1 <- bgrowth(model, data = dat)
```

Adolescents' growth model

- We first check model convergence. Roughly, R-hat statistics below 1.05 and effective sample sizes $> 100\times$(number of chains)

```
round(blavInspect(fit1, 'rhat'), 3)
```

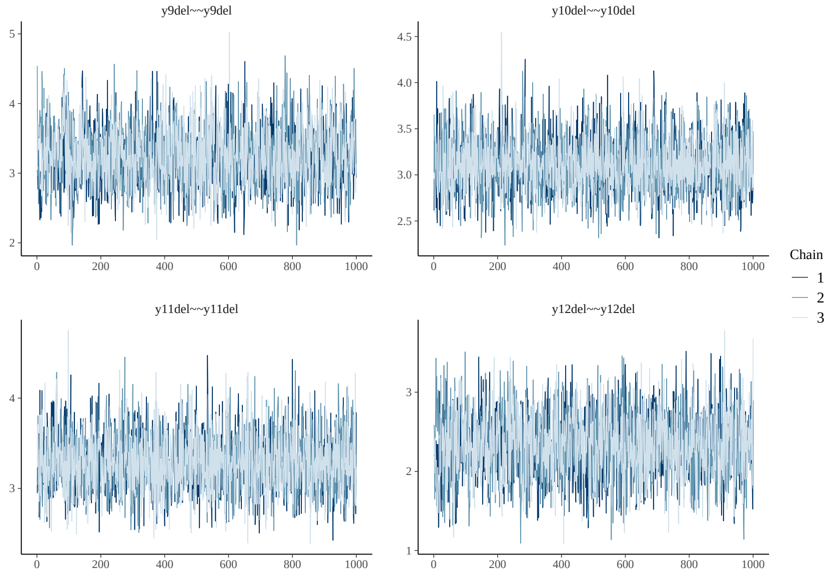
```
##   y9del~~y9del y10del~~y10del y11del~~y11del y12del~~y12del           i~~i
##          1.001          0.999          0.999          1.000          1.000
##           s~~s           i~~s            i~1            s~1
##          1.001          1.000          1.000          1.001
```

```
round(blavInspect(fit1, 'neff'), 3)
```

```
##   y9del~~y9del y10del~~y10del y11del~~y11del y12del~~y12del           i~~i
##       1680.926       2892.099       2891.099       1864.512       1958.665
##           s~~s           i~~s            i~1            s~1
##       1551.771       1588.392       2625.513       2602.230
```

Adolescents' growth model

```r
## posterior samples of the first four model parameters; ordering found by coef()
plot(fit1, 1:4)
```



Adolescents' growth model I

```r
summary(fit1)
```

```
## blavaan (0.3-17) results of 1000 samples after 500 adapt/burnin iterations
##
##   Number of observations                           396
##
##   Number of missing patterns                         5
##
##   Statistic                                 MargLogLik         PPP
##   Value                                      -3354.697       0.308
##
## Latent Variables:
##                    Estimate  Post.SD pi.lower pi.upper     Rhat    Prior
##   i =~
##     y9del             1.000                                  NA
##     y10del            1.000                                  NA
##     y11del            1.000                                  NA
##     y12del            1.000                                  NA
##   s =~
##     y9del             0.000                                  NA
##     y10del            1.000                                  NA
##     y11del            2.000                                  NA
##     y12del            3.000                                  NA
##
## Covariances:
##                    Estimate  Post.SD pi.lower pi.upper     Rhat    Prior
##   i ~~
##     s                 0.037    0.155    -0.28    0.316    1.000     lkj_corr(1)
##
## Intercepts:
##                    Estimate  Post.SD pi.lower pi.upper     Rhat    Prior
##    .y9del             0.000                                  NA
##    .y10del            0.000                                  NA
```

Adolescents' growth model II

```
##    .y11del           0.000                                      NA
##    .y12del           0.000                                      NA
##     i                2.527    0.113    2.298    2.744    1.000    normal(0,10)
##     s                0.413    0.048    0.318    0.507    1.001    normal(0,10)
##
## Variances:
##                   Estimate  Post.SD pi.lower pi.upper     Rhat    Prior
##    .y9del            3.224    0.432    2.424     4.12    1.001 gamma(1,.5)[sd]
##    .y10del           3.117    0.307    2.553    3.758    0.999 gamma(1,.5)[sd]
##    .y11del           3.295    0.325    2.693    3.964    0.999 gamma(1,.5)[sd]
##    .y12del           2.324    0.416    1.534     3.14    1.000 gamma(1,.5)[sd]
##     i                2.789    0.415    2.023    3.672    1.000 gamma(1,.5)[sd]
##     s                0.301    0.096    0.122    0.488    1.001 gamma(1,.5)[sd]
```

Adolescents' growth model

- The mean change in delinquency ratings over time is estimated to be $2.5 + .41 \times$ year.
- This model is like a multilevel model (intercept and slope vary by person), while also allowing for residual heterogeneity across time points.

Adolescents' growth model

- One way to judge model fit is by the posterior predictive p-value (not the best way, but an easy way). Values closer to .5 indicate good fit; values closer to 0 indicate poor fit.
- Also see `blavFitIndices()`, which computes Bayesian versions of some SEM fit metrics (contributed by Garnier-Villareal & Jorgensen).

```r
fitMeasures(fit1, 'ppp')
```

```
##   ppp
## 0.308
```

Adolescents' growth model

- At this point, we would want to do further model checking. Perhaps we should consider autocorrelation between time points, perhaps linear trends do not make sense (treat time points as discrete?), etc.

- Below, we will stick with our simple model to illustrate a few more blavaan features. But many extensions and alternatives are possible.

Further modeling

Further features

- We have seen that blavaan with default options looks similar to lavaan. But some things are different from lavaan:
  - Prior distributions
  - Model comparison metrics
  - Model extension/modification

Prior distributions

- For the growth model that we just examined, how could we add information/expectations to our prior distributions?
  - Intercepts (grade 9 scores): Should be positive and relatively low, say `normal(4, sd=1.5)`
  - Slopes: Would be surprising if negative, and values of 4 or 5 would push us from the bottom of the scale to the top from grade 9 to 12. `normal(2.5, sd=1)`

Prior distributions

- Across-person variability in intercepts: At grade 9, we would expect adolescents to exhibit variability in delinquency. Say, somewhere around an SD of 3 on the 0–22 scale. `gamma(1.5, .5)`
- Across-person variability in slopes: We also expect variability in adolescents' change over time. We just said that slopes of 4 or 5 would be surprising, and many negative values would be surprising. So let's stick with SDs around 1 here. `gamma(1, 1)`

Prior distributions

- We can encode these priors in the model specification, via `prior()`.

```
modelp <- ' i =~ 1*y9del + 1*y10del + 1*y11del + 1*y12del
            s =~ 0*y9del + 1*y10del + 2*y11del + 3*y12del

            i ~ prior("normal(4, 1.5)")*1
            s ~ prior("normal(2.5, 1)")*1
            i ~~ prior("gamma(1.5,.5)[sd]")*i
            s ~~ prior("gamma(1,1)[sd]")*s '

fit1p <- bgrowth(modelp, data = dat)
```

Model comparison

- blavaan can easily make use of new metrics for Bayesian model comparison, including WAIC and LOOIC (from the loo package, developed by Vehtari, Gelman, and collaborators).
- Strategy: Fit multiple models and compute a metric for each model. The model with the lowest value is preferred, though we should also consider uncertainty in the values.

Model comparison

- Below, we fit a new model with no change over time. We can then compare this model to the previous model via WAIC, LOOIC, or the Bayes factor.

```
model2 <- ' i =~ 1*y9del + 1*y10del + 1*y11del + 1*y12del '

fit2 <- bgrowth(model2, data = dat)
```

Model comparison

- Metrics imply that the model with a nonzero slope is better:

```r
blavCompare(fit1, fit2)
```

```
##
## WAIC estimates:
##  object1:  6684.679
##  object2:  6789.432
##
## WAIC difference & SE:
##   -52.376   13.763
##
## LOO estimates:
##  object1:  6684.444
##  object2:  6789.282
##
## LOO difference & SE:
##   -52.419   13.746
##
## Laplace approximation to the log-Bayes factor
## (experimental; positive values favor object1):   48.807
```

Model extensions

- You might be unhappy because blavaan does not provide some feature that you desire.
- Then you can specify a related model and use `mcmcfile = TRUE` to output the Stan (or JAGS) model syntax and necessary data. This provides you a starting point for doing what you want.
- Disclaimer: The Stan model is especially complicated, because it is written to handle any SEM that the user requests. I advise to start at the model block and work backwards.

Conclusions

Conclusions

- So far, blavaan has led to some improvements and tightening in Bayesian SEM estimation and model comparison. It has also provided researchers with tools for applying new MCMC methods to their own data, and for developing/implementing new procedures.
- Future development is aided by a new grant from the Institute of Education Sciences, U.S. Department of Education.

Future

- The near future: Ordinal SEM (coming soon); multilevel SEM (next year)
- Other possibilities
  - Parallelization in Stan
  - Latent variable interactions and quadratic effects
  - Modeling framework closer to GLLAMM
  - Your contribution!

Conclusions

- Software development can be an interesting way to do research.
  - Good research topics often arise during implementation!
  - Fit in the traditional academic incentive structure is not always clear. Impact vs impact factor.
  - Users become collaborators and contributors.

References

- Merkle, E. C., Fitzsimmons, E., Uanhoro, J., & Goodrich, B. (in press). Efficient Bayesian structural equation modeling in Stan. *Journal of Statistical Software*.

- Merkle, E. C., Furr, D., & Rabe-Hesketh, S. (2019). Bayesian comparison of latent variable models: Conditional vs marginal likelihoods. *Psychometrika, 84*, 802–829.

- Merkle, E. C. & Rosseel, Y. (2018). blavaan: Bayesian structural equation models via parameter expansion. *Journal of Statistical Software, 85(4)*, 1–30.

Thank you!

Try it yourself:

```r
install.packages("blavaan")
```

Further information:

https://ecmerkle.github.io/blavaan/

Extra Slides

Bayesian methods

- "But the prior"
  - Often, people say they don't have any prior beliefs. So they use uninformative prior distributions, which are often the software defaults.
  - But even if you don't have strong prior beliefs about what the model parameters should be, you often have beliefs about what the model parameters *should not* be. That is a good thing to encode in a prior ("mildly informative priors").
  - Sensitivity analysis (sensitivity of results to different priors) is recommended, as well as prior predictive analysis.

Likelihoods

- Marginal (traditional) SEM likelihood (marginalize over latent variables):

$$\boldsymbol{y}_i \sim \mathrm{N}(\boldsymbol{\nu} + \boldsymbol{\Lambda}\boldsymbol{\alpha}, \boldsymbol{\Lambda}(\boldsymbol{I} - \boldsymbol{B})^{-1}\boldsymbol{\Psi}(\boldsymbol{I} - \boldsymbol{B}^\top)^{-1}\boldsymbol{\Lambda}^\top + \boldsymbol{\Theta})$$

- Conditional SEM likelihood (sample latent variables):

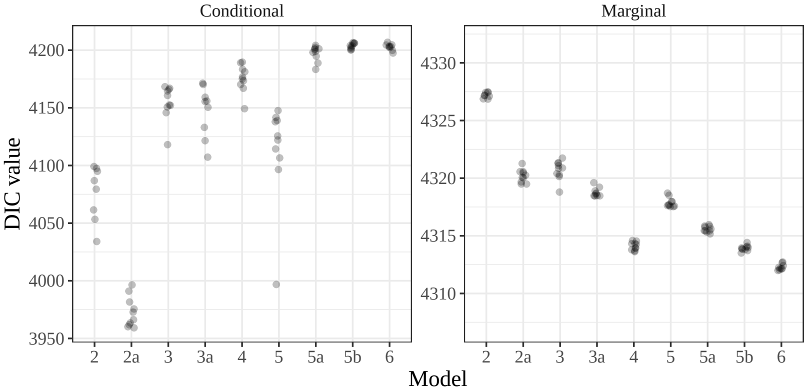
$$\begin{aligned}
\boldsymbol{y}_i &\sim \mathrm{N}(\boldsymbol{\nu} + \boldsymbol{\Lambda}\eta_i, \boldsymbol{\Theta}) \\
\eta_i &\sim \mathrm{N}((\boldsymbol{I} - \boldsymbol{B})^{-1}\alpha, (\boldsymbol{I} - \boldsymbol{B})^{-1}\boldsymbol{\Psi}(\boldsymbol{I} - \boldsymbol{B}^\top)^{-1})
\end{aligned}$$

Information criteria

- Those conditional and marginal likelihoods play a role in computing Bayesian information criteria like DIC and WAIC.
- Basic idea: For a single model fit to a single dataset, you can compute one DIC using the conditional likelihood and another DIC using the marginal likelihood. Most Bayesian SEM researchers use whatever their software produces.

CFA and DIC

- DIC computations for nine CFA models from Wicherts et al. (10 replications each)



Major results

- DIC/WAIC/LOO-CV computations and conclusions depend on use of marginal vs conditional likelihood, and DIC also differs between BUGS and JAGS.
- Marginal is recommended as default, because it is most related to the model's ability to generalize to new people. And marginal criteria exhibit much less Monte Carlo error.
- Merkle, E. C., Furr, D., & Rabe-Hesketh, S. (2019). Bayesian comparison of latent variable models: Conditional vs marginal likelihoods. *Psychometrika, 84*, 802–829.

CFA and DIC

- Effective number of parameters for nine CFA models