Bayesian Structural Equation Modeling Workshop - Exercises

Sara van Erp

09-03-2022

Introduction

These exercises are part of the workshop "Bayesian Structural Equation Modeling". In these exercises, you will conduct a Bayesian SEM analysis on an example data set, with a special focus on conducting a prior sensitivity analysis. If you have your own data you wish to analyze, feel free to adapt the exercises to apply them to your own data instead of the example data set. Make sure you have the `blavaan` package loaded.

```
library(blavaan)
```

Data

In these exercises, we will conduct a confirmatory factor analysis (CFA) with the classic Holzinger and Swineford data. You can find more information about the data as follows:

```
?HolzingerSwineford1939
```

Model

We will fit the following three-factor CFA model to these data:

```
HS.model <- ' visual  =~ x1 + x2 + x3
              textual =~ x4 + x5 + x6
              speed   =~ x7 + x8 + x9 '
```

Priors

Before we can fit our model, we need to consider the priors. Think about all the parameters in the model. Use the `dpriors()` function to check the default priors in `blavaan` and write down which default prior each parameter would get. It can be helpful to do so by creating a table with in one column the original parameter (e.g. `visual =~ x1`), in one column the `blavaan` notation (e.g. `lambda`), and then a column with the `blavaan` default prior for that parameter.

Most of the prior distributions should look familiar. Note that the `dpriors()` function reports a Wishart prior for `ibpsi`, the inverse covariance matrix of blocks of latent variables. However, this prior is not used in the default `stan` implementation of `blavaan`. Instead, the `lkj_corr(1)` prior is used as default. This can be seen by sampling from the priors or posterior and then calling the summary on the resulting fitobject. The Lewandowski-Kurowicka-Joe (LKJ) distribution is often used for correlation matrices. Its hyperparameter $\eta = 1$ by default which implies a uniform (beta) prior on the correlations.

Visualizing the priors

Sample from the default prior distributions and plot the prior for the latent variable covariances.

```
priors.def <- bcfa(HS.model,
                   data = HolzingerSwineford1939,
```

```
                   prisamp = TRUE,
                   sample = 1000)

plot(priors.def,
     pars = c(19:21),
     plot.type = "areas")
```

Consider the resulting priors. Do they result in values for the latent variable covariances that are reasonable? You might find it difficult to determine whether the resulting priors for the latent variable covariances are reasonable or not. In this situation, it might be more insightful to consider the priors on the latent variable correlations since correlations have known bounds (i.e., a correlation must lie between -1 and 1). Add the argument `std.lv = TRUE` to the code to sample from the prior. This will ensure that the model is identified by fixing the mean and variance of the latent variables. Plot the priors for the latent variable correlations. Are the resulting priors reasonable?

Adapting the priors

Now, let's adapt the priors on the latent variable correlations. The `lkj_corr(1)` prior specified by default implies a `beta(1, 1)` prior which implies a uniform prior with support from -1 to 1. We can do so by adapting the model specification as follows:

```
HS.model2 <- ' visual  =~ x1 + x2 + x3
              textual =~ x4 + x5 + x6
              speed   =~ x7 + x8 + x9
              visual ~~ prior("beta(1,1)")*textual
              visual ~~ prior("beta(1,1)")*speed
              textual ~~ prior("beta(1,1)")*speed'
```

By changing the hyperparameters of the beta distributions in the model above, we can change the priors for the latent variable correlations. To do so, it can be helpful to first plot the beta prior using different hyperparameter settings. Note that, normally the beta distribution has support from 0 to 1, but the implementation for correlations in `blavaan` is such that the support runs from -1 to 1. Therefore, we cannot use the regular `dbeta` function to visualize the prior. Instead, the following code can be used:

```
s1 <- 1 # hyperparameter 1
s2 <- 1 # hyperparameter 2
x <- seq(.01, .99, .01)
plot(-1 + 2*x,
     dbeta(x, shape1 = s1, shape2 = s2),
     type="l",
     yaxt = "n",
     xlab = "",
     ylab = "",
     xlim = c(-1, 1))
```

Try out different values for the hyperparameters. Make sure to include values smaller than 1 as well as values larger than 1. Also make sure to use two equal and two different hyperparameters. Make sure that you understand the impact of changing these hyperparameters. We will use these results later on in the prior sensitivity analysis.

Fit the model

Next, fit the model with the priors you find reasonable. Check the convergence. If the model has converged, check the estimates and the fit indices.

Prior sensitivity analysis

Next, we will conduct a prior sensitivity analysis in which we will focus on the priors for the latent variable correlations. Run the model with the following hyperparameter settings for the beta prior and compare the estimates for the latent variable correlations. Do they vary greatly?

Hyperparameter 1	Hyperparameter 2
1	1
10	10
0.1	0.1
1	10
10	1