Bayesian Estimation of Multidimensional Latent Variable Models

Lecture 8

## Today’s Lecture Objectives

1. Show how to estimate multidimensional latent variable models using MCMC (Stan)
2. Describe why Stan needs to use the Cholesky decomposition for covariance matrices of latent variables

## Example Data: Simulated Two Dimensions

Today we will be using the simulated data we started class with and used in Lecture 7

* 10 items
* 1000 examinees
* 2 dimensions
* Both continuous(for CFA) and categorical data (for IRT)
* Q-matrix:

Qmatrix

theta1 theta2  
item1 1 0  
item2 1 0  
item3 1 0  
item4 1 0  
item5 1 0  
item6 0 1  
item7 0 1  
item8 0 1  
item9 0 1  
item10 0 1

## Our Example: Multivariate Normal Distribution

For our example, we will assume the set of traits follows a multivariate normal distribution

Where:

* is the number of latent variables (dimensions)
* is the covariance matrix of the latent variables
  + is the inverse of the covariance matrix
* is the mean vector of the latent variables
* is the matrix determinant function
* is the matrix transpose operator

Alternatively, we would specify ; but, we cannot always estimate and

## Observed Item Model

For the IRT analyses, we will use a two parameter logistic item response model where:

For the CFA analyses, we will use a CFA model where:

Where:

* is the response of person to item
* is the score of person on dimension
* is the intercept of item
* is the factor loading of item on dimension
* is the Q-matrix value for item on dimension
* is the residual error of person on item
* is the number of dimensions
* is the number of items

## Stan versus JAGS

Bayesian analyses can be implemented in many different software packages

* I focus today on Stan (but I will leave JAGS code in the folder)
* Stan:
  + Hamiltonian Monte Carlo (a more efficient type of Metropolis algorithm)
  + Can directly model continuous parameters only (no discrete parameters–DCMs/mixture models are harder)
  + Has two R interfaces
    - rstan – older and not developed as much, if at all, any longer
    - cmdstanr
  + Compiles the model to C++ code (so it is fast)
  + Has built-in LKJ prior for correlation matrices
* JAGS
  + Uses a number of different sampling methods (Gibbs and Metropolis-Hastings based)
  + Slower than Stan (standalone executable)
  + Can model continuous and discrete parameters
  + Difficult to estimate models with standardized factors and correlations

## LKJ Priors for Correlation Matrices

From [Stan’s Functions Reference](https://mc-stan.org/docs/functions-reference/lkj-correlation.html), for a correlation matrix

Correlation Matrix Properties:

* Positive definite (determinant greater than zero;
* Symmetric
* Diagonal values are all one

LKJ Prior, with hyperparameter , is proportional to the determinant of the correlation matrix

Where:

* Density is uniform over correlation matrices with
* With , identity matrix is modal (moves correlations toward zero)
* With , density has a trough at the identity (moves correlations away from zero)

For this example, we set , noting a uniform prior over all correlation matrices

## Cholesky Decomposition

The functions we are using do not use the correlation matrix directly

#### The problem:

* The issue: We need the inverse and determinant to calculate the density of a multivariate normal distribution
  + Inverses and determinants are numerically unstable (computationally) in general form

#### The solution:

* Convert to a triangular form using the Cholesky decomposition

## Cholesky Decomposition Math

* (note: product becomes sum for log-likelihood)
* For the inverse, we work with the term in the exponent of the MVN pdf:

Then, we solve by back substitution:

* We can then multiply this result by itself to form the term in the exponent (times )
* Back substitution involves far fewer steps, minimizing the amount of rounding error in the process of forming the inverse

Most algorithms using MVN distributions use some variant of this process (perhaps with a different factorization method such as QR)

## Estimation Notes and Links

I teach a Bayesian Psychometric Models course on occasion:

* JAGS Version: <https://jonathantemplin.com/bayesian-psychometric-modeling-spring-2019/>
* Stan Version: <https://jonathantemplin.com/bayesian-psychometric-modeling-fall-2022/>

Much more detail in both courses about how Bayesian estimation works