Lecture 3: Bayes + LVM/FA

Jeff Rouder

June, 2024

Why Bayes For FA and SEM?

- lavaan is quick, easy, seamless, well documented
- you certainly don't need me here to show you how to use it
- And Bayes FA/SEM
    - not quick
    - not easy
    - not seamless
    - not well documented

Why Bayes For FA and SEM

1. Fixes Heywood cases.

- Variances are negative
- Correlations are bigger than 1.0 or negative

Why Bayes For FA and SEM

2. Most of our data are not cute little matrices $\boldsymbol{Y}_{[I\times J]}$
3. Preprocessing the data to make $\boldsymbol{Y}$

- could be simple, like aggregation
- could be complicated, like deriving a drift rate in a diffusion model
- resulting $\boldsymbol{Y}$ in real data is often too noisy to support FA/SEM

4. Needed: a fully integrated approach where FA affects preprocessing and preprocessing affects FA.
5. Enter Bayes

Bayes For Cute Little Score Matrices

Here is our data set:

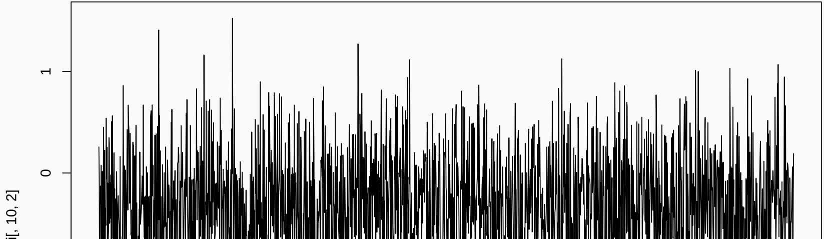
```r
set.seed(123)
I=200
J=8
D=2
lambda=matrix(nrow=J,ncol=D)
lambda[,1]=seq(1,0,length=J)
lambda[,2]=seq(0,1,length=J)
Sigma=crossprod(t(lambda))+diag(rep(1^2,J))
y=rmvnorm(I,rep(0,J),Sigma)
```

Your turn

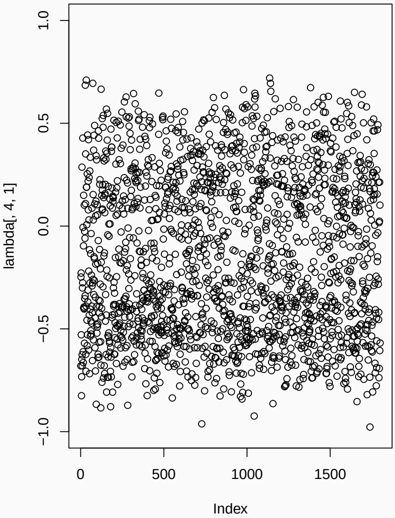
- Program up a Bayes sampler in JAGS or stan (I will do JAGS) for recovering lambda
- Use the conditional formulation
  - write out the model
  - implement it in JAGS or stan
  - run it and see if you can document issues
- When I program up all Gibbs steps, I get a lot of autocorrelation.

My turn

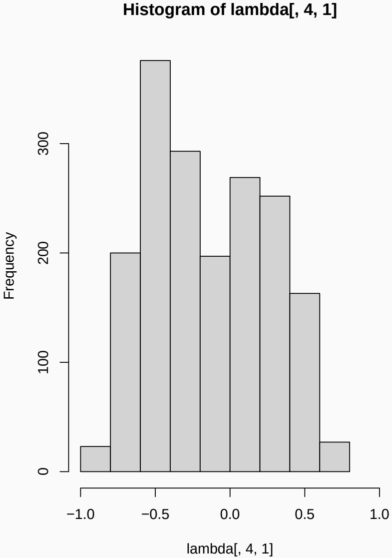
```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 1600
##    Unobserved stochastic nodes: 432
##    Total graph size: 7056
##
## Initializing model
```



Lambda [4,1]



lambda[, 4, 1]
1.0
0.5
0.0
−0.5
−1.0
0
500
1000
1500
Index



Histogram of lambda[, 4, 1]
Frequency
0
100
200
300
−1.0
−0.5
0.0
0.5
1.0
lambda[, 4, 1]

Rotations!

- each iteration is corresponding to a different rotation.
- what to do?
- old way, fix loading to lower triangle

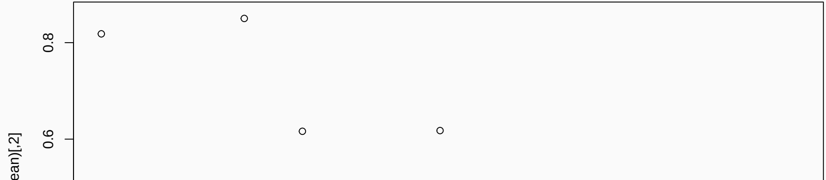
Lower Triangle for 3 Factors

$+$	0	0
*	$+$	0
*	*	$+$
*	*	*
*	*	*
*	*	*
*	*	*
*	*	*

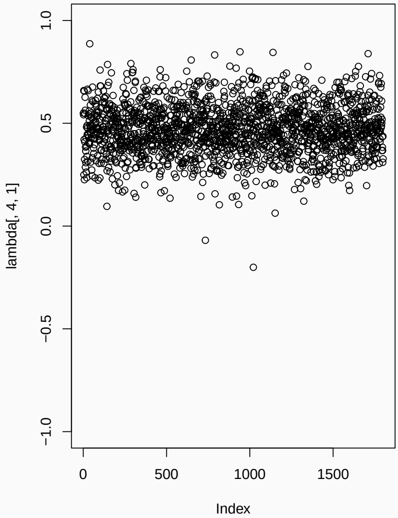
Your Turn

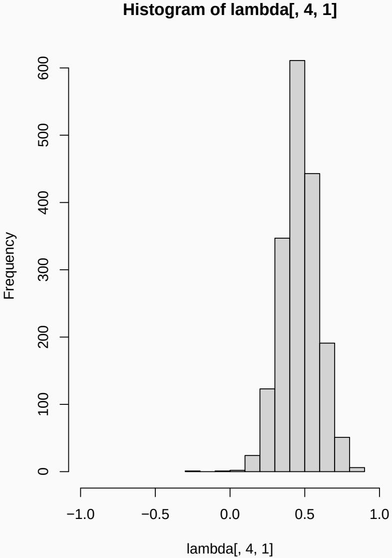
Adapt your code for lower triangle.

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 1600
##    Unobserved stochastic nodes: 432
##    Total graph size: 7103
##
## Initializing model
```



Lower Triangle





Post-Sampling Rotations

- Very new approach
- Align each iteration to a common rotation after the fact.
- Papastamoulis, P., & Ntzoufras, I. (2022). On the identifiability of Bayesian factor analytic models. Statistics and Computing, 32(2), 23. doi:10.1007/s11222-022-10084-4
- Poworoznek, E., Ferrari, F., & Dunson, D. (2021, July 29). Efficiently resolving rotational ambiguity in Bayesian matrix sampling with matching. Retrieved November 21, 2023, from http://arxiv.org/abs/2107.13783