Bayesian Lasso Confirmatory Factor Analysis

Lijin Zhang
Sun Yat-sen University

Joint work with Junhao Pan
Sun Yat-sen University
Edward Haksing Ip
Wake Forest University School of Medicine

23 May 2022

Overview

Introduction

Confirmatory Factor Analysis

Suppose $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n$ are independent random observations, and each $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{ip})^T$ satisfies the following factor analysis model:

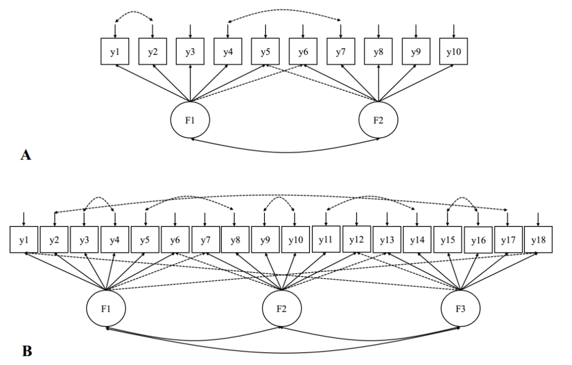
$$\mathbf{y}_i = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\omega}_i + \boldsymbol{\epsilon}_i, i = 1, 2, ..., n, \tag{1}$$

- $\boldsymbol{\mu} : p \times 1$ vector of intercepts.
- $\boldsymbol{\Lambda} : p \times q$ factor loading matrix, reflects the relation of observed variables in y_i with the $q \times 1$ latent factors in ω_i.
- $\boldsymbol{\omega}_i \sim N[0, \Phi]$.
- $\boldsymbol{\epsilon}_i : p \times 1$ random vector of measurement errors, $\sim N[0, \Psi]$, independent of ω_i.

Model Violations

While theory-based CFA is more compelling in many ways, sometimes the theory being tested does not fit the data well.

- violation of local independence (residual correlations)
- missing cross-loadings



Model Violations

- Revert to exploratory factor analysis (EFA), compare the results derived from the two different approaches, and make changes for a separate round of exercise in CFA.
- Use modification indexes (MIs) for identifying components in the model that could be tweaked for the purpose of improving overall goodness-of-fit, known as post hoc model modification (PMM, Kaplan, 1990; Sörbom, 1989).

Post-hoc Model Modification

There are several advantages of PMM from a practical point of view (Bentler & Bonett, 1980; MacCallum, 1995; Sörbom, 1989).

Several problems of using the PMM methodology:

- the use of modification indexes can be easily influenced by the researchers' subjective choices.
- over-fitting problem.
- parameters must be modified sequentially, causes difficulties in finding the global optimal model (Chou & Bentler, 1990).
- there is no guarantee that the modified covariance matrix is positive definite.

Bayesian CFA (Muthén & Asparouhov, 2012)

Relax the strict constraints in traditional CFA using small variance priors

- Cross-loadings: zero mean, small variance prior (e.g., $N[0, 0.01]$).
- Residual covariances: inverse-Wishart prior ($IW(I, df)$ with $df = p + 6, p =$ number of items, gives a prior standard deviation of 0.1)

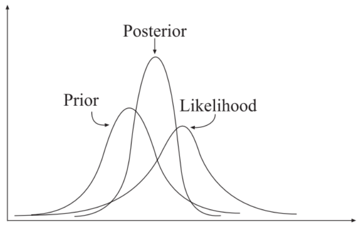


Figure 1. Prior, likelihood, and posterior for a parameter.

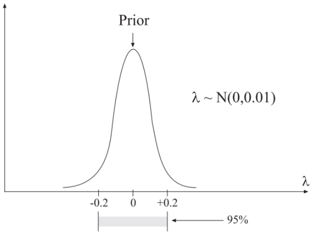


Figure 2. Informative prior for a factor loading parameter.

Lasso

In the frequentist framework, the Lasso approach implements regularization by adding a penalty term to the usual likelihood so that the model would move toward a solution that contains fewer parameters.

$$PL(\boldsymbol{\theta}) = \log(p(\mathbf{y} \mid \boldsymbol{\theta}, M)) + \lambda \sum_{j=1}^{p} |\theta_j| = LL(\boldsymbol{\theta}) + \lambda \sum_{j=1}^{p} |\theta_j| \quad (2)$$

Bayesian Lasso

In the Bayesian framework, the key quantity is the posterior distribution

$$p(\boldsymbol{\theta} \mid \mathbf{y}, M) \propto p(\mathbf{y} \mid \boldsymbol{\theta}, M) \times p(\boldsymbol{\theta} \mid M) \tag{3}$$

The log posterior in a Bayesian approach takes the general form

$$log(p(\mathbf{y} \mid \boldsymbol{\theta}, M)) + log(p(\boldsymbol{\theta} \mid M)) = LL(\boldsymbol{\theta}) + LPrior(\boldsymbol{\theta}) \tag{4}$$

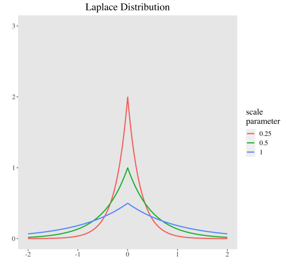
If the appropriate form of the prior distribution is chosen, the log prior distribution in Bayesian analysis tends to play the role of the penalty function in Lasso.

Bayesian Covariance Lasso

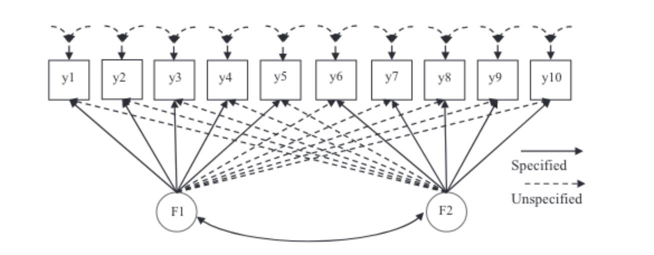
$$\sigma_{ii} \sim \frac{\lambda}{2} exp(-\frac{\lambda}{2}\sigma_{ii}), \sigma_{ij} \sim \frac{\lambda}{2} exp(-\lambda \mid \sigma_{ij} \mid), i < j \qquad (5)$$

where $\Sigma = \Psi^{-1} = (\sigma_{ij})_{p \times p}$

$$\begin{pmatrix}
\psi_{11} & 0.0 & 0.0 & 0.0 & 0.0 & \psi_{16} & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & \psi_{22} & 0.0 & 0.0 & 0.0 & 0.0 & \psi_{27} & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & \psi_{33} & 0.0 & \psi_{35} & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & \psi_{44} & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & \psi_{53} & 0.0 & \psi_{55} & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
\psi_{61} & 0.0 & 0.0 & 0.0 & 0.0 & \psi_{66} & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & \psi_{72} & 0.0 & 0.0 & 0.0 & 0.0 & \psi_{77} & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & \psi_{88} & 0.0 & \psi_{8,10} \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & \psi_{99} & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & \psi_{10,8} & 0.0 & \psi_{10,10}
\end{pmatrix}.$$



Bayesian Lasso CFA

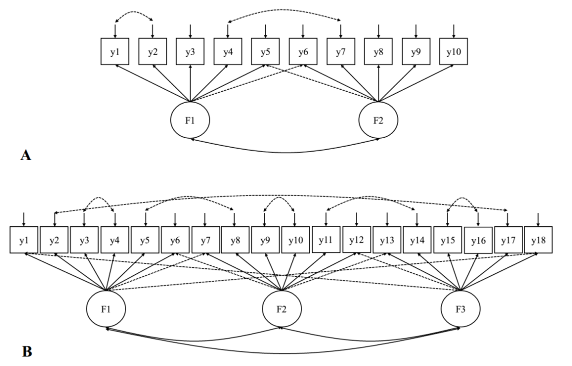


- Detects all the significant residual covariances and cross-loadings in one estimation, thus, circumvents the problem of having to handle model violations sequentially.
- Achieves model parsimony as well as an identifiable model.
- The detection of residual covariances and cross-loadings can reduce the bias in structural estimates.

Simulation Studies

Simulation Study

Purpose: Test the performance of blcfa in parameter recovery

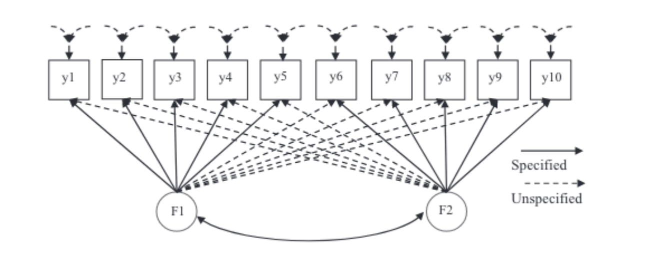


Non-zero cross-loadings were set at 0.5

Non-zero residual covariances were set at 0.3

$N = 250, 500$

Model Estimation



Simple Model

Par	True	N = 250 BIAS	RMSE	*SE*	SIG%	N = 500 BIAS	RMSE	*SE*	SIG%
λ_{11}	0.7	0.081	0.092	0.137	1.000	0.078	0.087	0.123	1.000
λ_{21}	0.7	0.084	0.096	0.137	1.000	0.078	0.086	0.123	1.000
λ_{31}	0.7	−0.067	0.081	0.131	1.000	−0.058	0.068	0.115	1.000
λ_{41}	0.7	−0.087	0.096	0.130	0.995	−0.082	0.089	0.116	1.000
λ_{51}	0.5	−0.039	0.064	0.133	0.995	−0.043	0.059	0.116	0.995
λ_{61}	0.5	−0.128	0.139	0.139	0.780	−0.105	0.114	0.125	0.875
λ_{52}	0.5	−0.082	0.102	0.151	0.775	−0.061	0.075	0.131	0.910
λ_{62}	0.5	0.013	0.050	0.134	0.985	0.001	0.036	0.117	0.995
λ_{72}	0.7	−0.054	0.070	0.137	0.990	−0.046	0.056	0.117	1.000
λ_{82}	0.7	−0.029	0.053	0.134	0.990	−0.015	0.035	0.112	1.000
λ_{92}	0.7	−0.025	0.052	0.134	0.990	−0.016	0.036	0.112	1.000
$\lambda_{10,2}$	0.7	−0.027	0.054	0.135	1.000	−0.017	0.036	0.112	1.000
λ_0	0	−0.008	0.043	0.111	0.000	−0.008	0.037	0.104	0.000
ϕ_{12}	0.3	0.069	0.100	0.191	0.560	0.086	0.101	0.172	0.850
ψ_{21}	0.3	−0.109	0.118	0.146	0.005	−0.102	0.110	0.138	0.040
ψ_{74}	0.3	−0.077	0.086	0.098	0.755	−0.070	0.076	0.089	0.915
ψ_{55}	0.35	0.106	0.116	0.114	1.000	0.083	0.091	0.095	1.000
ψ_{66}	0.35	0.097	0.109	0.111	1.000	0.082	0.090	0.093	1.000
ψ_{jj}	0.51	0.025	0.091	0.133	1.000	0.014	0.075	0.118	1.000
ψ_0	0	0.023	0.043	0.082	0.000	0.019	0.036	0.071	0.000
δ_l	—	7.584	0.488	2.653	1.000	7.543	0.355	2.528	1.000
δ_s	—	2.188	0.125	0.354	1.000	2.276	0.100	0.360	1.000

Note. ψ_{jj} averaged across elements from $j = 1$ to 4 and 7 to 10; λ_0 averaged across all zero loading estimates; for δ_l and δ_s, BIAS = mean; RMSE = *SD*.

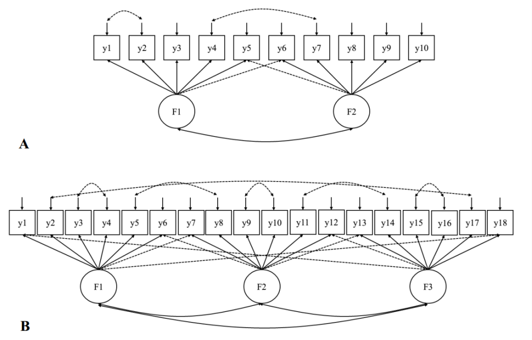
Complex Model

Par	λ_{11}	λ_{21}	λ_{31}	λ_{41}	λ_{51}	λ_{61}	λ_{71}	$\lambda_{18,1}$	λ_{62}	λ_{72}	λ_{82}	λ_{92}	$\lambda_{10,2}$	$\lambda_{11,2}$	$\lambda_{12,2}$	
TRUE	0.5	0.7	0.7	0.7	0.7	0.5	0.5	0.5	0.5	0.5	0.7	0.7	0.7	0.7	0.5	
BIAS	-0.01	-0.035	0.082	0.084	-0.034	-0.007	-0.056	-0.052	-0.061	-0.017	-0.042	0.08	0.087	-0.038	-0.012	
RMSE	0.051	0.053	0.093	0.093	0.054	0.049	0.081	0.076	0.084	0.055	0.062	0.09	0.096	0.055	0.049	
SE	0.096	0.104	0.106	0.106	0.104	0.096	0.099	0.099	0.1	0.097	0.107	0.108	0.108	0.105	0.098	
SIG%	1	1	1	1	1	1	1	0.995	1	1	1	1	1	1	1	
Par	$\lambda_{13,2}$	λ_{13}	$\lambda_{12,3}$	$\lambda_{13,3}$	$\lambda_{14,3}$	$\lambda_{15,3}$	$\lambda_{16,3}$	$\lambda_{17,3}$	$\lambda_{18,3}$	λ_0	$f_{kk'}$	ψ_{jj}	ψ_{ii}	ψ_w	ψ_b	ψ_0
TRUE	0.5	0.5	0.5	0.5	0.7	0.7	0.7	0.7	0.5	0	0.3	0.35	0.51	0.3	0.3	0
BIAS	-0.057	-0.054	-0.057	-0.012	-0.042	0.081	0.082	-0.041	-0.014	-0.007	0.031	0.071	-0.009	-0.09	-0.062	0.005
RMSE	0.08	0.078	0.078	0.054	0.061	0.091	0.092	0.06	0.053	0.038	0.074	0.084	0.081	0.101	0.071	0.028
SE	0.101	0.101	0.1	0.097	0.106	0.108	0.108	0.105	0.096	0.086	0.135	0.073	0.105	0.106	0.07	0.053
SIG%	0.99	1	0.995	1	1	1	1	1	1	0	0.805	1	1	0.702	1	0

Simulation Study

To investigate the performance of blcfa in parameter identification, we manipulated the following factors in the simulation study.

- Sample size: 200, 500, 1000
- Model Size: 2 factors and 10 items, 3 factors and 18 items
- Effect Sizs: 0, 0.1, 0.2, 0.3 for cross-loadings; 0, 0.1, 0.3, 0.7 for residual correlations



Model Estimation

To avoid the possible confounding effect, the conditions of non-zero cross-loadings and non-zero residual correlations were separately generated and analyzed.

- M_1: model with some non-zero cross-loadings and diagonal residual covariance matrix
- M_2: model with some non-zero, off-diagonal residual covariance entries but no cross-loading.

Parameter Identification

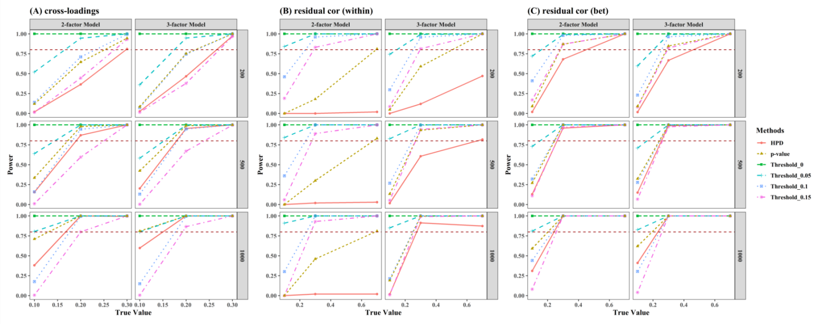
- Thresholds of magnitude 0, 0.05, 0.1, and 0.15 with the decision rule to include if the absolute value of the standard estimate is larger than the cutoff.
- A p-value with α=0.05, with the decision rule to include if p ¡ 0.05. The p-value can be different from the frequentist p-value, it is one-tailed and is based on MCMC samples rather than the z-test.
- A 95% HPD interval, with the decision rule to include if the point 0.0 is outside the 95% HPD interval.

Evaluation Criteria

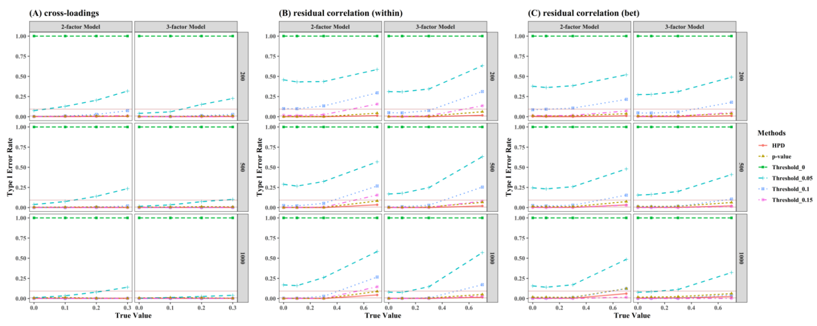
- Power: the probability of correctly identifying the cross-loadings/residual correlations when the parameters are non-zero (Muthén & Asparouhov, 2012).
- Type-I error rates: the probability of erroneously identifying the cross-loadings/residual correlations when the parameters are zero.

Results: Power

- The HPD interval and *p*-value rule have similar power problems in detecting within-factor residual correlations.
- The thresholding rule is more robust to sample sizes.



The general pattern was almost a mirror image of that of power.



Detect Cross-loadings & Residual Covariances Simultaneously

- The co-existence of cross-loadings and residual correlations may be expected in practice.
- The phenomenon of low power of p-value and HPD interval methods was still present for within-factor residual correlations.
- Compared to performance when only one kind of parameter was present, we found the HPD interval and p-value methods were more sensitive to model size and provided lower power.

Recommendation

Sample Size	Parameters	Threshold 0.1		Threshold 0.15		HPD Interval		*p*-value	
		Power	Type I	Power	Type I	Power	Type I	Power	Type I
200	Cross-loadings[1]	×	√	×	√	×	√	×	√
	Residual correlations	√	×	√	√	×	√	×	√
500	Cross-loadings	√	√	×	√	√	√	√	√
	Residual correlations	√	√	√	√	×	√	×	√
1000	Cross-loadings	√	√	√	√	√	√	√	√
	Residual correlations	√	√	√	√	×	√	×	√

Note: Type I: Type I Error Rate; √: acceptable in most conditions, ×: unacceptable in many conditions, shaded: the best criterion in the corresponding condition.

[1]None of the criteria can provide sufficient power under this condition.

Usage of The Thresholding Rule

- The 0.1 cutoff value, which we recommend for detecting cross-loading, can be a candidate for application to parameter identification for path coefficients in SEM.
- We conjecture that using a cutoff of 0.15 for other correlation parameters would be appropriate if the Bayesian lasso model is adopted.
- For exploratory analysis in SEM in which the purpose is to extract as many potentially important relationships as possible, the cutoff value can be lowered.

R Package blcfa

R Package blcfa

To make use of the advantages of Bayesian Lasso CFA in detecting residual covariances and cross-loadings, we propose a two-steps method for model modifications:

- (1) detect significant cross-loadings and/or residual covariances different from zero by Bayesian Lasso CFA;
- (2.1) free the identified significant parameters;
- (2.2) automatically feed the output from (2.1) into Mplus to obtain an appropriately modified CFA model using Maximum likelihood (ML) estimator or Bayesian estimation.

We built an R package named 'blcfa' to facilitate the application of this method.

Example

Detailed Illustration: https://github.com/zhanglj37/blcfa

Installation

install.packages("devtools")
library(devtools)
install_github("zhanglj37/blcfa")

Detect Cross-loadings and Residual Covariances

Social Support Scale, 5-points Likert scale, 17 items, three factors

```
library(blcfa)

filename = "ss.txt"
varnames = c("gender",paste("y", 1:17, sep = ""))  # variables in dataset
usevar = c(paste("y", 1:17, sep = ""))  # variables used in the analysis
NZ = 3  # number of factors
IDY = matrix(c(
  9,-1,-1,
  1,-1,-1,
  1,-1,-1,
  1,-1,-1,
  1,-1,-1,
  -1,9,-1,
  -1,1,-1,
  -1,1,-1,
  -1,1,-1,
  -1,1,-1,
  -1,1,-1,
  -1,-1,9,
  -1,-1,1,
  -1,-1,1,
  -1,-1,1,
  -1,-1,1,
  -1,-1,1
),ncol=NZ,byr=T)
# NZ: number of factors
# 9: fixed at one for identifing the factor
# 1: estimate this parameter without shrinkage
# -1: estimate this parameter using lasso shrinkage
# 0: fixed at zero.
```

Example

Function:

blcfa(filename, varnames, usevar, IDY, estimation = 'Bayes', ms = -9, interval = TRUE)

estimation (= 'ML' / 'Bayes', the default value is 'Bayes')

ms represents missing value

interval: Detect significant residual correlations and cross-loadings based on HPD interval or threshold

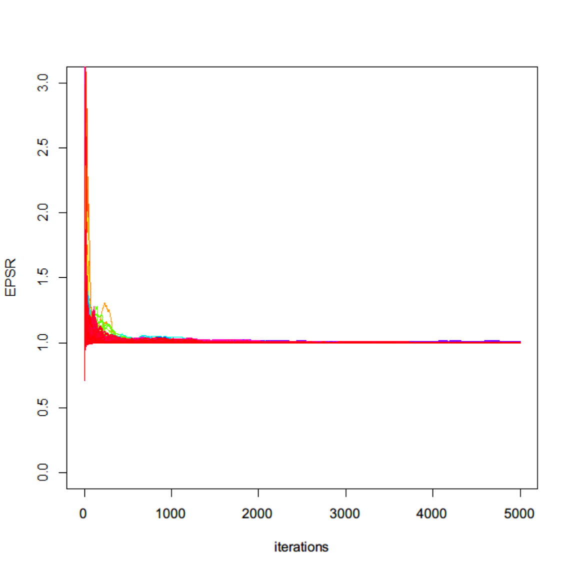
After running this function:

The program is running. See 'log.txt' for details.

Gibbs sampling ended up, specific results are being calculated.

('log.txt' records the process of parallel computing of two MCMC chains)

Convergence: Estimated Potential Scale Reduction Value



Results

```
TITLE: Bayesian Lasso CFA
DATA: FILE =  ss.txt ;
VARIABLE:
NAMES = gender y1 y2 y3 y4 y5 y6 y7 y8 y9
        y10 y11 y12 y13 y14 y15 y16 y17 ;
USEV = y1 y2 y3 y4 y5 y6 y7 y8 y9
        y10 y11 y12 y13 y14 y15 y16 y17 ;
ANALYSIS:
         ESTIMATOR = BAYES;
         PROC = 2;
         BITERATIONS = (10000);
MODEL:
        f1 by   y1  y2  y3  y4  y5  y17  ;
        f2 by   y6  y7  y8  y9  y10  y11  y13  y14  ;
        f3 by   y12  y5  y13  y14  y15  y16  y17  ;

        y11  with  y13 ;
        y11  with  y14 ;
        y13  with  y14 ;

 OUTPUT: TECH1  TECH8  STDY;
 PLOT: TYPE= PLOT2;
```

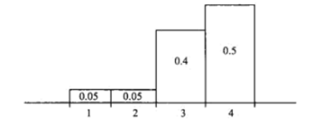
Ordinal Data and Adaptive Lasso

Ordinal Data

- Most analyses of the model have been carried out under the framework of confirmatory factor analysis with the assumption that the observed variables are continuous and have normal distribution.
- To satisfy the assumption, most subjects are required to select intermediate options from all options.
- However, in practical applications, the histogram of most variable is biased.

Ordinal Data

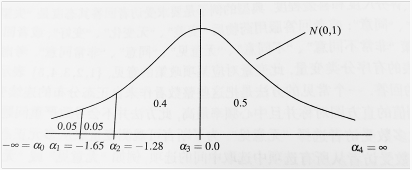
Assume a data set has such a biased histogram and the continuous measurements $y_j (y_j \sim N[0,1])$ are unobservable



The relationship between y_j and the observable variable:

For $l = 0, 1, 2, 3, \alpha_{jl} < y_j < \alpha_{j,l+1}$

$$-\infty = \alpha_0 < \alpha_1 < \alpha_2 < \alpha_3 < \alpha_4 = +\infty$$



Model Identification

The random vector of latent continuous variables $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{ip})^T$ satisfies the following factor analysis model:

$$\mathbf{y}_i = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\omega}_i + \boldsymbol{\epsilon}_i \tag{6}$$

Because intercpets, residual variances, and thresholds are uncertain, models with ordered categorical variables are not identified without imposing identification conditions.

Model Identification

The variances of the measurement errors ϵ_{ij} can be identified by using one of two types of constraints (Forero & Maydeu-Olivares, 2009). One is to fix the variances of ϵ_{ij} to 1.0, and the other is to fix the variances of latent continuous variables y_{ij} to 1.0. The latter one is adopted in our model.

Parameters μ_j and $\alpha_{j,l}$ can not be simultaneously estimable. One common method is to constrain the first threshold $\alpha_{j,1}$ to a fixed value. For example, to fix $\alpha_{j,1} = \phi^{*-1}(f^*_{j,1})$, where $\phi^*(\cdot)$ is the standard normal distribution function, $f^*_{j,1}$ is the frequency of the first category (see Lee, 2007).

Adaptive Lasso

As an extension of lasso, Zou (2006) proposed the adaptive lasso (alasso) by imposing different penalty strengths on parameters from different scales.

The frequentist alasso adds weights to rescale the penalty parameter as: $\gamma/|\beta_p^0|$, where β_p^0 is the preliminary estimates of β_p (e.g., ML estimate).

The Bayesian alternative to adaptive lasso can be obtained by including a coefficient-specific penalty parameter to impose unique shrinkage on each parameter.

Simulation Study

Four non-zero residual covariances $(\psi_{16}, \psi_{27}, \psi_{35}, \psi_{8,10})$ were set at 0.2.

The continuous measurements $(y_1, y_2, \cdots, y_{10})^T$ were transformed to ordered categorical observations $(z_1, z_2, \cdots, z_{10})^T$ via the following thresholds: $\boldsymbol{\alpha}_j = (\alpha_{j1}, \alpha_{j2}) = (0.0, 1.0)$ for $j = 1, 2, \cdots, 10$, where the α_{j1}'s were fixed to identify the ordered categorical variables.

$$\boldsymbol{\Lambda}^T = \begin{pmatrix} 0.8 & 0.8 & 0.5 & 0.5 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.8 & 0.8 & 0.8 & 0.5 & 0.5 \end{pmatrix}, \quad \boldsymbol{\Phi} = \begin{pmatrix} 1.0 & 0.3 \\ 0.3 & 1.0 \end{pmatrix},$$

Adaptive Lasso vs Lasso

		ALasso				Lasso			
Par	True	Bias	SE	RMS	RR	Bias	SE	RMS	RR
μ_1	0.5	-0.051	0.047	0.004	1.000	-0.051	0.047	0.004	1.000
μ_2	0.5	-0.057	0.047	0.005	1.000	-0.057	0.047	0.005	1.000
μ_3	0.5	-0.058	0.048	0.005	1.000	-0.058	0.048	0.005	1.000
μ_4	0.5	-0.062	0.048	0.006	1.000	-0.062	0.048	0.006	1.000
μ_5	0.5	-0.056	0.048	0.005	1.000	-0.056	0.048	0.005	1.000
μ_6	0.5	-0.052	0.047	0.004	1.000	-0.052	0.047	0.004	1.000
μ_7	0.5	-0.058	0.048	0.005	1.000	-0.058	0.047	0.005	1.000
μ_8	0.5	-0.051	0.047	0.004	1.000	-0.052	0.047	0.004	1.000
μ_9	0.5	-0.059	0.048	0.005	1.000	-0.059	0.048	0.005	1.000
μ_{10}	0.5	-0.050	0.048	0.004	1.000	-0.051	0.048	0.004	1.000
λ_{11}	0.8	-0.099	0.050	0.011	1.000	-0.148	0.081	0.024	1.000
λ_{21}	0.8	-0.094	0.049	0.010	1.000	-0.143	0.081	0.023	1.000
λ_{31}	0.5	-0.036	0.057	0.004	1.000	0.001	0.098	0.003	1.000
λ_{41}	0.5	-0.047	0.054	0.005	1.000	-0.043	0.098	0.006	0.990
λ_{51}	0.5	-0.034	0.058	0.004	1.000	0.005	0.099	0.003	1.000
λ_{62}	0.8	-0.094	0.045	0.010	1.000	-0.122	0.062	0.016	1.000
λ_{72}	0.8	-0.090	0.046	0.010	1.000	-0.118	0.062	0.015	1.000
λ_{82}	0.8	-0.075	0.048	0.007	1.000	-0.073	0.060	0.006	1.000
λ_{92}	0.5	-0.053	0.051	0.005	1.000	-0.054	0.080	0.005	1.000
$\lambda_{10,2}$	0.5	-0.036	0.059	0.004	1.000	0.017	0.089	0.002	1.000

Adaptive Lasso vs Lasso

ϕ_{12}	0.3	-0.011	0.057	0.002	1.000	-0.013	0.076	0.003	0.970
ψ_{11}	0.36	-0.053	0.056	0.005	1.000	0.012	0.094	0.003	1.000
ψ_{22}	0.36	-0.055	0.056	0.005	1.000	0.008	0.094	0.003	1.000
ψ_{33}	0.75	-0.163	0.057	0.029	1.000	-0.201	0.097	0.043	1.000
ψ_{44}	0.75	-0.149	0.054	0.025	1.000	-0.158	0.090	0.028	1.000
ψ_{55}	0.75	-0.166	0.057	0.030	1.000	-0.207	0.098	0.045	1.000
ψ_{66}	0.36	-0.057	0.048	0.005	1.000	-0.018	0.074	0.002	1.000
ψ_{77}	0.36	-0.057	0.049	0.005	1.000	-0.019	0.074	0.002	1.000
ψ_{88}	0.36	-0.084	0.051	0.009	1.000	-0.083	0.072	0.008	1.000
ψ_{99}	0.75	-0.145	0.051	0.023	1.000	-0.145	0.074	0.023	1.000
$\psi_{10,10}$	0.75	-0.170	0.058	0.032	1.000	-0.221	0.089	0.051	1.000
ψ_{16}	0.2	-0.045	0.037	0.003	0.980	-0.044	0.053	0.003	0.950
ψ_{27}	0.2	-0.040	0.036	0.003	0.990	-0.043	0.053	0.004	0.930
ψ_{35}	0.2	-0.067	0.051	0.008	0.620	-0.093	0.077	0.010	0.060
$\psi_{8,10}$	0.2	-0.067	0.050	0.007	0.680	-0.104	0.072	0.012	0.030
ψ_0	0	0.001	0.019	0.000	0.000	0.005	0.048	0.001	0.004

Discussion

Discussion

- We proposed Bayesian Lasso CFA for relaxing the exact zero constrains on residual covariances and cross-loadings.
- Accurate estimates and Acceptable power in most conditions.
- The performance of the thresholding rule is less sensitive to the change of sample size.
- An R package blcfa was developed to facilitate the usage.
- The Bayesian Lasso CFA method can detect significant residual covariances and cross-loadings in one estimation and circumvent the limitations of post-hoc model modifications.

Future Research

- We extended the lasso method to adaptive Lasso method and ordinal data.
- Comparison between Ridge, Lasso, and Alasso under the frequentist and Bayesian framework in factor analysis (will be presented at IMPS 2022).

Thanks for listening!

slides: `https://lijinzhang.com/share/220523_blcfa.pdf`