Modelos longitudinais de um único grupo sob uma abordagem multivariada

Prof. Caio Azevedo

Outline of the presentation

- Motivation.
- Literature review.
- Main contributions of this work.
- A longitudinal IRT model with restricted covariance matrices and time-heterogenous variances.
- Reparameterization of the model and Bayesian inference.
- Simulation studies.
- Real data anaylsis.
- Final conclusions.

Motivation

- Longitudinal item response data occur when students are assessed at several time points.
- This kind of data consist of response patterns of different examinees responding to different tests at different measurement occasions (e.g. grades).
- This leads to a complex dependence structure that arises from the fact that measurements (i.e., the latent traits and item reponses) from the same student are typically correlated.

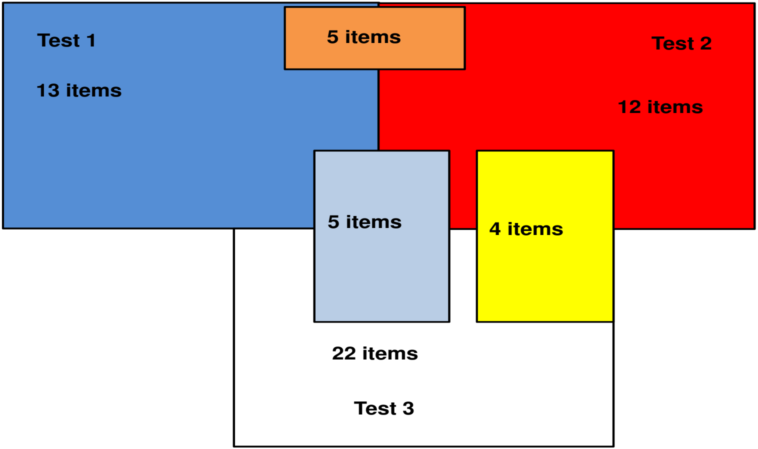
Motivation: real data set

- The data set analyzed stems from a major study initiated by the Brazilian Federal Government known as the School Development Program.
- The aim of the program is to improve the teaching quality and the general structure (classrooms, libraries, laboratory informatics etc) in Brazilian public schools.
- A total of 400 schools in different Brazilian states joined the program. Achievements in mathematics and Portuguese language were measured over five years (from fourth to eight grade of primary school) from students of schools selected and not selected for the program.

Motivation: real data set

- In the present study, mathematic performances of 1,500 randomly selected students, who were assessed in the fourth, fifth, and sixth grade, were considered.
- A total of 72 test items was used, where 23, 26, and 31 items were used in the test in grade four (Test 1), grade five (Test 2), and grade six (Test 3), respectively. Five anchor items were used in all three tests.
- Another common set of five items was used in the test in grade four and five. Furthermore, four common items were used in the tests in grades five and six.

Test design


Test 1
13 items
5 items
Test 2
12 items
5 items
4 items
22 items
Test 3

Motivation: real data set

- In an exploratory analysis, the Multiple Group Model (MGM), described in Azevedo et al. (2012), was used to estimate the latent student achievements given the response data.
- The MGM for cross-sectional data assumes that students are nested in groups and latent traits are assumed to be independent given the mean level of the group.
- Pearson's correlations, variances, and covariances were estimated among the vectors of estimated latent traits corresponding to grade four to six. The estimates are represented in the next slide.

Within-student correlation structure of the latent traits estimated by the MGM

	Grade four	Grade five	Grade six
Grade four	1.000	**.723**	**.629**
Grade five	.659	1.152	**.681**
Grade six	.540	.641	1.071

Estimated posterior variances, covariances, and correlations among estimated latent traits are given in the diagonal, lower and upper triangle, respectively.

Literature review

- Conoway (1990): longitudinal Rasch model (one-parameter model):
  - Uniform covariance matrix.
  - Complete test design.
- Dunson (2003): a general modeling framework that allows mixtures of count, categorical and continuous response:
  - Complete test design.
  - Did not explore sepecific covariance structures.

Literature review

- Andrade and Tavares (2005): longitudinal three-parameter model (population latent traits parameters estimation.
- Tavares and Andrade (2006): longitudinal three-parameter model (item and population latent traits parameters estimation):
  - Three-parameter model. time-homogeneous covariance matrices.
  - Numerical problems in handling many time-points.
  - All covariance matrices have closed expressions for their inverses and determinants.

Main contributions of this work

- General modeling of the time-heterogenous dependency between student achievements (some restricted covariance matrices never used for longitudinal IRT data).
- Bayesian inference which handles identification rules, restricted parametric covariance structures and situations with several time-points (there is no limit, numerically speaking).
- Simulation study: parameter recovery and model selection.
- Bayesian model fit assessment and model comparison tools.
- All developments were made considering the two-parameter model but can be straightforwardly extended to the three-parameter, polytomous or continuous IRT models.

A longitudinal IRT Model

$$
\begin{aligned}
Y_{ijt} \mid (\theta_{jt}, \zeta_i) &\sim \text{Bernoulli}(P_{ijt}) \\
P_{ijt} = P(Y_{ijt} = 1 \mid \theta_{jt}, \zeta_i) &= \Phi(a_i\theta_{jt} - b_i) \\
\boldsymbol{\theta}_{j.} | \boldsymbol{\eta}_{\boldsymbol{\theta}} &\sim N_T(\boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Psi}_{\boldsymbol{\theta}}),
\end{aligned}
$$

- Y_{ijt} : is the answer of subject (student) j, to item i, in time-point t. It is equal to 1 if the subject answers the item correctly and 0 otherwise, ($\Phi(.)$ is the cdf of the standard normal distribution).
- $\boldsymbol{\theta}_{j.} = (\theta_j, ..., \theta_{jT})'$: is the vector of the latent traits of the subject j.
- $\zeta_i = (a_i, b_i)'$; a_i is the discrimination parameter of item i; b_i : is the parameter related to the difficulty of item i. The true difficulty parameter is given by $b_i^* = b_i / a_i$.

Population (latent traits) parameters

$$\boldsymbol{\mu_\theta} = \begin{bmatrix} \mu_{\theta_1} \\ \mu_{\theta_2} \\ \vdots \\ \mu_{\boldsymbol{\theta}_T} \end{bmatrix} \text{ and } \boldsymbol{\Psi_\theta} = \begin{bmatrix} \psi_{\theta_1} & \psi_{\theta_{12}} & \dots & \psi_{\theta_{1T}} \\ \psi_{\theta_{12}} & \psi_{\theta_2} & \dots & \psi_{\theta_{2T}} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{\theta_{1T}} & \psi_{\theta_{2T}} & \dots & \psi_{\theta_T} \end{bmatrix},$$

$\boldsymbol{\eta_\theta}$ is a vector consisting of $\boldsymbol{\mu_\theta}$ and the non-repeated elements of $\boldsymbol{\Psi_\theta}$.

Modeling the covariance matrix

Heteroscedastic uniform model - HU

$$\boldsymbol{\Psi}_{\boldsymbol{\theta}} = \begin{bmatrix} \psi_{\theta_1} & \sqrt{\psi_{\theta_1}}\sqrt{\psi_{\theta_2}}\rho_\theta & \dots & \sqrt{\psi_{\theta_1}}\sqrt{\psi_{\theta_T}}\rho_\theta \\ \sqrt{\psi_{\theta_1}}\sqrt{\psi_{\theta_2}}\rho_\theta & \psi_{\theta_2} & \dots & \sqrt{\psi_{\theta_2}}\sqrt{\psi_{\theta_T}}\rho_\theta \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\psi_{\theta_1}}\sqrt{\psi_{\theta_T}}\rho_\theta & \sqrt{\psi_{\theta_2}}\sqrt{\psi_{\theta_T}}\rho_\theta & \dots & \psi_{\theta_T} \end{bmatrix},$$

Heteroscedastic Toeplitz model - HT

$$\boldsymbol{\Psi}_{\boldsymbol{\theta}} = \begin{bmatrix} \psi_{\theta_1} & \sqrt{\psi_{\theta_1}}\sqrt{\psi_{\theta_2}}\rho_\theta & 0 & \dots & 0 \\ \sqrt{\psi_{\theta_1}}\sqrt{\psi_{\theta_2}}\rho_\theta & \psi_{\theta_2} & \sqrt{\psi_{\theta_2}}\sqrt{\psi_{\theta_3}}\rho_\theta & \dots & 0 \\ 0 & \sqrt{\psi_{\theta_2}}\sqrt{\psi_{\theta_3}}\rho_\theta & \psi_{\theta_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \psi_{\theta_T} \end{bmatrix}.$$

Modeling the covariance matrix

Heteroscedastic covariance model - HC

$$\mathbf{\Psi}_{\boldsymbol{\theta}} = \begin{bmatrix} \psi_{\theta_1} & \rho_\theta & \ldots & \rho_\theta \\ \rho_\theta & \psi_{\theta_2} & \ldots & \rho_\theta \\ \vdots & \vdots & \ddots & \vdots \\ \rho_\theta & \rho_\theta & \ldots & \psi_{\theta_T} \end{bmatrix},$$

First-order autoregressive moving-average model - ARMAH

$$\mathbf{\Psi}_{\boldsymbol{\theta}} = \begin{bmatrix} \psi_{\theta_1} & \sqrt{\psi_{\theta_1}\psi_{\theta_2}}\gamma_\theta & \ldots & \sqrt{\psi_{\theta_1}\psi_{\theta_T}}\gamma_\theta\rho_\theta^{T-2} \\ \sqrt{\psi_{\theta_1}\psi_{\theta_2}}\gamma_\theta & \psi_{\theta_2} & \ldots & \sqrt{\psi_{\theta_2}\psi_{\theta_T}}\gamma_\theta\rho_\theta^{T-3} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\psi_{\theta_1}\psi_{\theta_T}}\gamma_\theta\rho_\theta^{T-2} & \sqrt{\psi_{\theta_2}\psi_{\theta_T}}\gamma_\theta\rho_\theta^{T-3} & \ldots & \psi_{\theta_T} \end{bmatrix}.$$

Modeling the covariance matrix

Ante-dependence model - AD

$$\boldsymbol{\Psi}_{\boldsymbol{\theta}} = \begin{bmatrix} \psi_{\theta_1} & \sqrt{\psi_{\theta_1}\psi_{\theta_2}}\rho_{\theta_1} & \cdots & \sqrt{\psi_{\theta_1}\psi_{\theta_T}}\prod_{t=1}^{T-1}\rho_{\theta_t} \\ \sqrt{\psi_{\theta_1}\psi_{\theta_2}}\rho_{\theta_1} & \psi_{\theta_2} & \cdots & \sqrt{\psi_{\theta_2}\psi_{\theta_T}}\prod_{t=2}^{T-1}\rho_{\theta_t} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\psi_{\theta_1}\psi_{\theta_T}}\prod_{t=1}^{T-1}\rho_{\theta_t} & \sqrt{\psi_{\theta_2}\psi_{\theta_T}}\prod_{t=2}^{T-1}\rho_{\theta_t} & \cdots & \psi_{\theta_T} \end{bmatrix},$$

Reparameterization of the model

- We have adapted the approach presented in McCulloh et al. (2000) which was developed for multivariate probit models with restricted covariance matrices.
- Let us consider the following partition of the latent traits structure:

$$\begin{aligned}\boldsymbol{\theta}_{j.} &= (\theta_{j1}, \theta_{j2}, \ldots, \theta_{jT})^t = (\theta_{j1}, \boldsymbol{\theta}_{j(1)})^t, \\ \boldsymbol{\mu}_{\boldsymbol{\theta}} &= (\mu_{\theta_1}, \mu_{\theta_2}, \ldots, \mu_{\theta_T})^t = (\mu_{\theta_1}, \boldsymbol{\mu}_{\boldsymbol{\theta}(1)})^t,\end{aligned}$$

where, $\boldsymbol{\theta}_{j(1)} = (\theta_{j2}, \ldots, \theta_{jT})^t$, $\boldsymbol{\mu}_{\boldsymbol{\theta}(1)} = (\mu_{\theta_2}, \ldots, \mu_{\theta_T})^t$.

Reparameterization of the model

- It follows that the covariance structure can be partitioned as,

$$\mathbf{\Psi}_{\boldsymbol{\theta}} = \begin{bmatrix} \psi_{\theta_1} & \boldsymbol{\psi}^t_{\boldsymbol{\theta}(1)} \\ \boldsymbol{\psi}_{\boldsymbol{\theta}(1)} & \mathbf{\Psi}_{\boldsymbol{\theta}(1)} \end{bmatrix}, \qquad (1)$$

where $\boldsymbol{\psi}_{\boldsymbol{\theta}(1)} = (\psi_{\theta_{12}}, \ldots, \psi_{\theta_{1T}})^t$ and

$$\mathbf{\Psi}_{\boldsymbol{\theta}(1)} = \begin{bmatrix} \psi_{\theta_2} & \ldots & \psi_{\theta_{2T}} \\ \vdots & \ddots & \vdots \\ \psi_{\theta_{2T}} & \ldots & \psi_{\theta_T} \end{bmatrix}.$$

Reparameterization of the model

- From properties of the multivariate normal distribution it follows that

$$\boldsymbol{\theta}_{j(1)}|\theta_{j1} \sim N_{(T-1)}\left(\boldsymbol{\mu}^*, \boldsymbol{\Psi}^*\right),$$

where

$$\boldsymbol{\mu}^* = \boldsymbol{\mu}_{\boldsymbol{\theta}(1)} + \psi_{\theta_1}^{-1}\boldsymbol{\psi}_{\boldsymbol{\theta}(1)}\left(\theta_{j1} - \mu_{\theta_1}\right),$$

and

$$\begin{aligned}\boldsymbol{\Psi}^* &= \boldsymbol{\Psi}_{\boldsymbol{\theta}(1)} - \psi_{\theta_1}^{-1}\boldsymbol{\psi}_{\boldsymbol{\theta}(1)}\boldsymbol{\psi}_{\boldsymbol{\theta}(1)}^t. \\ &= \boldsymbol{\Psi}_{\boldsymbol{\theta}(1)} - \boldsymbol{\psi}^*\boldsymbol{\psi}^{*t}\end{aligned}$$

$$\boldsymbol{\psi}^* = \boldsymbol{\psi}_{\boldsymbol{\theta}(1)}/\sqrt{\psi_{\theta_1}}$$

Reparameterization of the model

- As a result, when conditioning on the restricted first-time point parameter, θ_{j1}, the remaining $\boldsymbol{\theta}_{j(1)}$ are conditionally multivariate normally distributed given θ_{j1}, with an unrestricted covariance matrix.
- The matrix $\boldsymbol{\Psi}^*$ is an unstructured covariance matrix without any identifiability restrictions. As a result, the common modeling (e.g., using an Inverse-Wishart prior) and estimation approaches can be applied for Bayesian inference.

Reparameterization of the model

- The variance/correlation parameters,

$$\psi^* \text{ and } \boldsymbol{\Psi}^*, \tag{2}$$

  define an one-to-one relation with the free parameters of the original covariance matrix $\boldsymbol{\Psi}_{\boldsymbol{\theta}}$, since the parameter ψ_{θ_1} is restricted to 1.
- As a result, the estimates of the population variances and covariances can be obtained from the estimates of Equation (2).
- Therefore we estimate ψ^* and $\boldsymbol{\Psi}^*$ (based on a general model) and, according to the covariance matrix of interest, we calculate specific parameters (restricted version of the general model).

Reparameterization of the model

- Working parameters : $(\boldsymbol{\theta}^t, \boldsymbol{\zeta}^t, \boldsymbol{\mu}_{\boldsymbol{\theta}}^t, \psi^{*t})$ and $\boldsymbol{\Psi}^*$.
- Parameters of interest: $(\boldsymbol{\theta}^t, \boldsymbol{\zeta}^t, \boldsymbol{\mu}_{\boldsymbol{\theta}}^t, \boldsymbol{\psi}^t)$, where $\boldsymbol{\psi}$ depends on the the covariance matrix of interest. For example, in the case of ARH structure, $\boldsymbol{\psi} = (\psi_{\theta_2}, ..., \psi_{\theta_T}, \rho_\theta)^t$.
- Transformation: $\boldsymbol{\psi}^*, \boldsymbol{\Psi}^* \rightarrow \boldsymbol{\psi}$.

Reparameterization of the model

- Therefore, (considering the ARH structure, for example): once we have the simulated values of $\boldsymbol{\psi}^*$ and $\boldsymbol{\Psi}^*$, we can calculate $\boldsymbol{\psi}$ by using:

$$\boldsymbol{\psi}^*_{\boldsymbol{\theta}(1)} = (\boldsymbol{\psi}_{\theta_2}, ..., \psi_{\theta_T})^t = \text{Diag}(\boldsymbol{\Psi}^*)(\text{general formula})$$

$$\rho_\theta = \frac{1}{T-1}\mathbf{1}^t_{T-1}\left(\boldsymbol{\psi}^* \bullet (\boldsymbol{\psi}^*_{\boldsymbol{\theta}(1)})^{-1/2}\right)$$

if we reescale the latent traits of the first time point.

Reparameterization of the model

- Model identification:
  - The latent variable distribution of the first measurement occasion will be restricted to identify the model. This is done by re-scaling the vector of latent variable values of the first measurement occasion to a pre-specified scale in each MCMC iteration.
  - An incomplete test design is used such that common items are administered at different measurement occasions (time-points). The common items, also known as anchors, make it possible to measure the latent traits on one common scale.

Prior distributions

- Latent traits: $\boldsymbol{\theta}_{j.}|\boldsymbol{\eta}_{\boldsymbol{\theta}} \sim N_T(\boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Psi}_{\boldsymbol{\theta}})$.
- Population (working) parameters:

$$
\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{\theta}} &\sim N_T(\boldsymbol{\mu}_0, \boldsymbol{\Psi}_0)\,,\\
\psi_{\theta_1} &\sim IG(\nu_0, \kappa_0)\,,\\
\boldsymbol{\psi}^* &\sim N_{T-1}(\boldsymbol{\mu}_{\boldsymbol{\psi}}, \boldsymbol{\Psi}_{\boldsymbol{\psi}})\,,\\
\boldsymbol{\Psi}^* &\sim IW_{T-1}(\nu_{\boldsymbol{\Psi}}, \boldsymbol{\Psi}_{\boldsymbol{\Psi}})\,,
\end{aligned}
$$

where $IG(\nu_0, \kappa_0)$ stands for the inverse-gamma distribution with shape parameter ν_0 and scale parameter κ_0, and $IW_{T-1}(\nu_{\boldsymbol{\Psi}}, \boldsymbol{\Psi}_{\boldsymbol{\Psi}})$ for the inverse-Wishart distribution with degrees of freedom $\nu_{\boldsymbol{\Psi}}$ and dispersion matrix $\boldsymbol{\Psi}_{\boldsymbol{\Psi}}$.

Prior distributions

- Item parameters :

$$
\begin{aligned}
p\left(\zeta_i = (a_i, b_i) \mid \boldsymbol{\mu}_\zeta, \boldsymbol{\Psi}_\zeta\right) & \propto & \exp\Bigg(-0.5\left(\zeta_i - \boldsymbol{\mu}_\zeta\right)^t \boldsymbol{\Psi}_\zeta^{-1} \\
& \times & \left(\zeta_i - \boldsymbol{\mu}_\zeta\right)\Bigg)\mathbb{1}_{(a_i>0)},
\end{aligned}
$$

Augmented data and indicator variables

- Augmented data:

$$Z_{ijt}|(\theta_{jt}, \zeta_i, Y_{ijt}) \sim N(a_i\theta_{jt} - b_i, 1),$$

where Y_{ijt} is the indicator of Z_{ijt} being greater than zero.

- Indicator variables:

$$I_{ijt} = \begin{cases} 1, & \text{item } i\text{, administered for examinee } j\text{, at time point } t \\ 0, & \text{missing by design.} \end{cases}$$

$$V_{ijt} = \begin{cases} 1, & \text{observed response of examinee } j\text{, at time point } t\text{, on item } i \\ 0, & \text{otherwise,} \end{cases}$$

Posterior distribution

Combining the prior distributions, the augmented data and the indicator variables the posterior distributions is given by

$$
\begin{aligned}
p(\boldsymbol{\theta}_{..}, \zeta, \boldsymbol{\mu}_{\boldsymbol{\theta}}, \psi_{\theta_1}, \boldsymbol{\psi}^*, \boldsymbol{\Psi}^* | \boldsymbol{z}_{...}, \boldsymbol{y}_{...}) &\propto p(\boldsymbol{z}_{...} | \boldsymbol{\theta}_{..}, \zeta, \boldsymbol{y}_{...}) p(\boldsymbol{\theta}_{..} | \boldsymbol{\eta}_{\boldsymbol{\theta}}) \\
&\times p(\zeta | \boldsymbol{\mu}_{\zeta}, \boldsymbol{\Psi}_{\zeta}) p(\boldsymbol{\eta}_{\boldsymbol{\theta}}).
\end{aligned}
$$

where

$$
p(\boldsymbol{\theta}_{..} | \boldsymbol{\eta}_{\boldsymbol{\theta}}) = \prod_{j=1}^{n} p(\boldsymbol{\theta}_{j.} | \boldsymbol{\eta}_{\boldsymbol{\theta}}), = \prod_{j=1}^{n} p(\boldsymbol{\theta}_{j(1)} | \boldsymbol{\eta}_{\boldsymbol{\theta}}, \theta_{j1}) p(\theta_{j1} | \boldsymbol{\eta}_{\theta_1})
$$

and

$$
p(\boldsymbol{\eta}_{\boldsymbol{\theta}}) = p(\boldsymbol{\mu}_{\boldsymbol{\theta}}) p(\psi_{\theta_1}) p(\boldsymbol{\psi}^*) p(\boldsymbol{\Psi}^*) .
$$

Let (.) denote the set of all necessary parameters. The full Gibbs sampling algorithm is defined as follows:

1. Start the algorithm by choosing suitable initial values.

   Repeat steps 2–10.

2. Simulate Z_{ijt} from $Z_{ijt} \mid (.), t = 1, .., T, i = 1, ..., I_t, j = 1, ..., n$.
3. Simulate $\boldsymbol{\theta}_{j.}$ from $\boldsymbol{\theta}_{j.} \mid (.), j = 1, ..., n$.
4. Simulate $\boldsymbol{\zeta}_i$ from $\boldsymbol{\zeta}_i \mid (.)$, i =1,...,I.
5. Simulate $\boldsymbol{\mu_\theta}$ from $\boldsymbol{\mu_\theta} \mid (.)$.
6. Simulate ψ_{θ_1} from $\psi_{\theta_1} \mid (.)$.
7. Simulate $\boldsymbol{\psi}^*$ from $\boldsymbol{\psi}^* \mid (.)$.
8. Simulate $\boldsymbol{\Psi}^*$ from $\boldsymbol{\Psi}^* \mid (.)$.
9. Compute the unstructured covariance matrix using the sampled covariance components from Steps 6-8 and equations (1) and (2).
10. Through a parameter transformation method using sampled unstructured covariance parameters, compute restricted covariance components of interest. The sampled restricted covariance structure $\boldsymbol{\Psi}$ is used when repeating steps 2–9.

Convergence and autocorrelation assessment and parameter recovery

- The Geweke diagnostic, based on a burn-in period of 16,000 iterations, indicated convergence of the chains of all model parameters.
- Furthermore, the Gelman-Rubin diagnostic were close to one, for all parameters. Convergence was established easily without requiring informative initial parameter values or long burn-in periods.
- Therefore, the burn-in was set to be 16,000, and a total of 46,000 values were simulated, and samples were collected at a spacing of 30 iterations producing a valid sample with 1,000 values.

Parameter recovery and model selection assessment

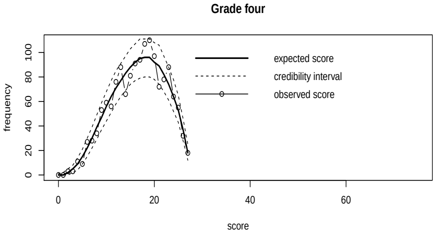
- The results of simulation studies indicated :
  - That all parameters were properly recovered by the MCMC algorithm.
  - The model selection procedure using AIC, BIC and DIC did not present very satisfactory results since the proportion of times that the true underlying model was selected is around 50%. This is probably due to the number of time-points (three) considered in the simulation study and to the fact that we are using the original likelihood instead of the marginal one (integrating out the latent traits).
- A more efficient method of model selection, based on Reversible Jump MCMC algorithms, is presented in Azevedo et al (2015).

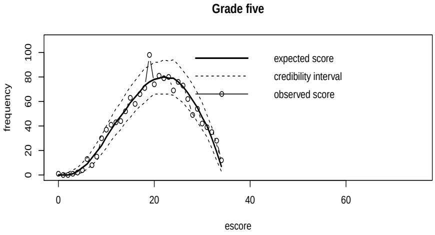
Real data anaylsis

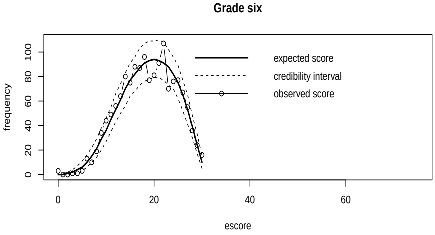
Selecting the optimal covariance structure for the real data set

Model	LL	AIC_2	BIC_2	DIC_2
HU	-71980	147477	**148941**	150398
ARH	-72164	147693	149157	150462
ARMAH	-72179	147727	149201	150496
HC	-72840	148707	150171	151139
AD	-72184	147723	149196	150477
Unst.	-71984	**147470**	148954	**150368**

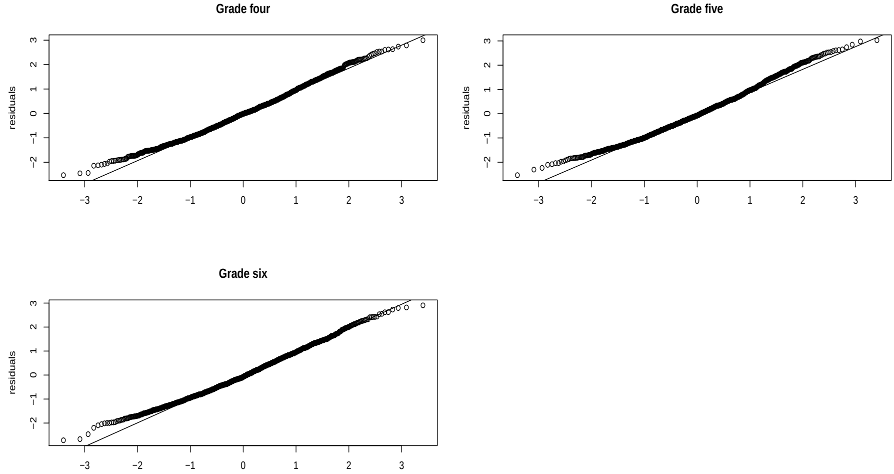
Real data anaylsis







Real data anaylsis

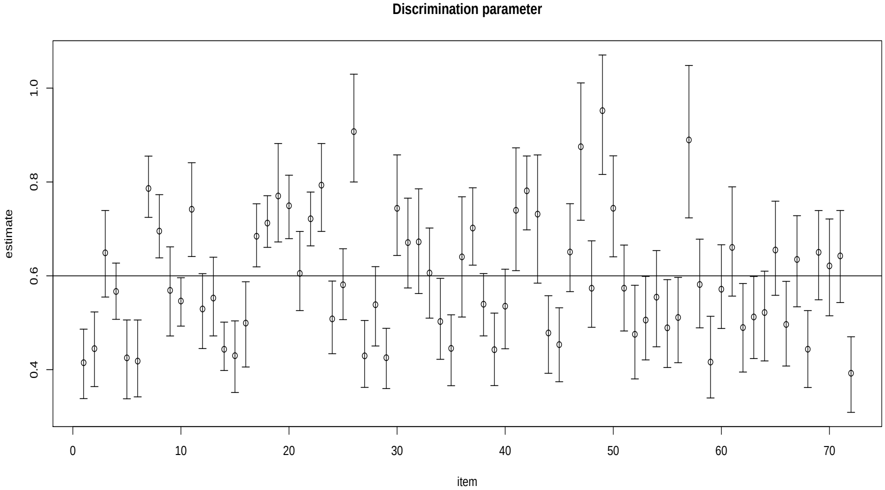


	Mean		
Grade	**Mean**	**SD**	**HPD 95%**
four (Reference)	0	-	-
five	.240	.040	[.170 , .319]
six	.763	.048	[.680 , .862]

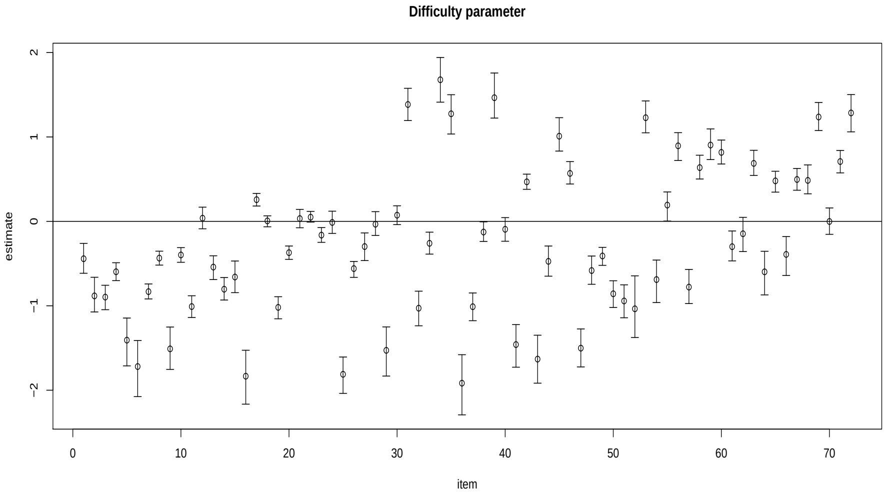
	Variance		
Grade	**Mean**	**SD**	**HPD 95%**
four (Reference)	1	-	-
five	1.032	.081	[.876 , 1.183]
six	.969	.087	[.794 , 1.131]

Correlations			
Grades	**Mean**	**SD**	**HPD 95%**
four and five	.857	.012	[.832 , .879]
four and six	.759	.017	[.724 , .790]
five and six	.810	.015	[.784 , .840]

Real data anaylsis



Real data anaylsis



Final conclusions

- The proposed general modeling allows for the time-heterogenous dependency between student achievements and accommodates some restricted covariance matrices never used for longitudinal IRT data analysis.
- Bayesian inference through MCMC algorithms was able to handle the identification rules and restricted parametric covariance structures.
- Simulation study: parameter recovery (very good performance) and model selection (not so good performance).

Final conclusions

- Bayesian model fit assessment tools shown to be very useful mechanisms for model validation.
- Extensions in order to consider other types of item reponse, latent traits distributions, multilevel structures, among other possibilities, can be straightforwardly developed.

Some references

- Azevedo, C. L. N., Fox,J.-P. and Andrade, D. F. (2016), Bayesian longitudinal item response modeling with restricted covariance pattern structures, *Statistics & Computing, 26, 1-2, 443-460*.
- Azevedo, C. L. N., Fox,J.-P. and Andrade, D. F. (2015), Longitudinal Multiple-Group IRT Modeling: Covariance pattern selection using MCMC, *Int. J. of Quantitative Research in Education, special issue: Bayesian Statistics in Psychometrics, 2, 3-4, 213-243*.

Some references

- Andrade, D.F. and Tavares, H.R. (2005), Item response theory for longitudinal data: population parameter estimation. J. Multivar. Anal. **95**, 1–22.
- Tavares, H.R. and Andrade, D.F. (2006), Item response theory for longitudinal data: item and population ability parameters estimation. Test **15**, 97–123.
- Azevedo, C.L.N., Andrade, D.F. and Fox, J.-P. (2012), A bayesian generalized multiple group IRT model with model-fit assessment tools. Comput. Stat. Data Anal. **56**, 4399-4412.