

Multidimensional LC models in the Item Response Theory (IRT) setting

Silvia Bacci

silvia.bacci@unipg.it

Dipartimento di Economia - Università degli Studi di Perugia (IT)

Outline

- 1 Introduction
- 2 Multidimensional LC models
- 3 Multidimensional LC-IRT models
 - Model formulation
 - Likelihood inference
- 4 Applications
 - Example 1: Hierarchical clustering for LSAT7 data items
 - Example 2: LC-IRT model for anxiety and depression assessment
- 5 Software implementation
 - Function `est_multi_poly`
 - Function `class_item`
 - Function `test_dim`

Motivation

- Basic LC model: manifest distribution of response vector \mathbf{Y}_i

$$p(\mathbf{y}) = p(\mathbf{Y}_i = \mathbf{y}) = \sum_{u=1}^k \pi_u \prod_{j=1}^{J_i} p(Y_{ij} = y | U_i = \xi_u)$$

- $g[p(Y_{ij} = y | U_i = \xi_u)] = \xi_u$, where $g(\cdot)$ is a suitable link function (e.g., logit in case of binary items, global or local logit in case of ordinal items)
- We may extend the basic LC model to allow for
 - more than one latent trait (**multidimensionality**)
 - item characteristics (**item response theory, IRT**, parameterization)

Notation

- s : number of different latent traits measured by the J items
- $\mathbf{U}_i = (U_{1i}, \dots, U_{si})'$: vector of latent variables corresponding to the latent traits
- $\mathbf{u}_i = (u_{1i}, \dots, u_{si})'$: one of the possible realisations of \mathbf{U}_i
- $\boldsymbol{\xi}_u = \{\xi_{ud}\}$: vector of values (support points) assumed by \mathbf{U}_i , with $u = 1, \dots, k$ and $d = 1, \dots, s$
- $\pi_u = p(\mathbf{U}_i = \boldsymbol{\xi}_u)$
- δ_{jd} : dummy variable equal to 1 if item j measures latent trait of type d and to 0 otherwise ($d = 1, \dots, s$)
 → **between-item multidimensionality** (vs within-item multidimensionality)

Multidimensional LC model

- General formulation for the conditional probability of answering y , given the latent class u

$$g[p(Y_{ij} = y | U_i = \xi_u)] = \sum_{d=1}^s \delta_{jd} \xi_{ud}$$

- specification of link function $g(\cdot)$ depends on the nature of items
 - **binary items** ($y = 0, 1$)

$$g[p(Y_{ij} = 1 | U_i = \xi_u)] = \log \frac{p(Y_{ij} = 1 | U_i = \xi_u)}{p(Y_{ij} = 0 | U_i = \xi_u)}$$

- **ordinal items** ($y = 0, 1, \dots, r_j - 1$)
 - 1 global (or cumulative) logits
 - 2 local (or adjacent categories) logits
 - 3 continuation ratio logits

Type of logits for ordinal items

1 global (or cumulative) logits

$$g[p(Y_{ij} = y | U_i = \xi_u)] = \log \frac{p(Y_{ij} \geq y | U_i = U_i = \xi_u)}{p(Y_{ij} < y | U_i = U_i = \xi_u)}, \quad y = 1, \dots, r_j - 1$$

2 local (or adjacent categories) logits

$$g[p(Y_{ij} = y | U_i = \xi_u)] = \log \frac{p(Y_{ij} = y | U_i = U_i = \xi_u)}{p(Y_{ij} = y - 1 | U_i = U_i = \xi_u)}, \quad y = 1, \dots, r_j - 1$$

3 continuation ratio logits

$$g[p(Y_{ij} = y | U_i = \xi_u)] = \log \frac{p(Y_{ij} \geq y | U_i = U_i = \xi_u)}{p(Y_{ij} = y - 1 | U_i = U_i = \xi_u)}, \quad y = 1, \dots, r_j - 1$$

Note that all types of logits coincide when $r_j = 2$ (binary items)

Item response theory parameterization

- In the basic (unidimensional and multidimensional) LC model we account neither for individual nor item characteristics
- However, it is reasonable assume that the item responses patterns y_i depend not only on the latent variable U_i but also on one or more characteristics of items
- This approach relies on the Item Response Theory (IRT) setting, which represents a model-based theory for the assessment of latent traits on the basis of observed item responses
- Formulation of IRT models is based on the introduction of some **item parameters**, describing the **difficulty** and the **discrimination power** of each item

Latent class formulation of IRT models

- In their original formulation, some IRT models (i.e., Rasch-type models) do not introduce any particular assumption on the distribution of the latent trait, whereas other IRT models typically assume the normality of the latent trait
- In both cases, the clustering of individuals in a finite number of latent classes with homogeneous **ability** (i.e., latent trait) levels is barred
- **LC-IRT models** (or finite-mixture IRT models) represent a combination of IRT and LC analysis based on the idea that the same IRT model holds for the subjects in the same class, whereas possibly different IRT models hold among different classes
- LC-IRT models allow for the **discreteness of the ability distribution** with a number of support points equal to the number of latent classes
- We consider a class of LC-IRT models for binary items and ordinal polytomous items, which assume **fixed item parameters** among latent classes (Bartolucci, 2007; Bacci, Bartolucci and Gnaldi, 2014) and take into account **multidimensional latent traits** (Reckase, 2009)

Multidimensional LC-IRT formulation

- General IRT formulation for the conditional probability of answering y , given the latent class u

$$g[p(Y_{ij} = y | U_i = \boldsymbol{\xi}_u)] = \gamma_j \left(\sum_{d=1}^s \delta_{jd} \xi_{ud} - \beta_{jy} \right) \quad y = 1, \dots, r_j - 1$$

- γ_j : discrimination index of item j
- β_{jy} : difficulty of threshold y of item j
- different model formulations are obtained according to the specification of
 - 1 type of link function (see above)
 - 2 constraints on the item discriminating parameters
 - 3 formulation of the item difficulty parameters

Constraints on the item discriminating parameters

- general situation in which each item may discriminate differently from the others
→ with binary items, we obtain a **two-parameters logistic (2PL) formulation**
- special case in which all the items discriminate in the same way

$$\gamma_j = 1, \quad j = 1, \dots, J$$

- with binary items, we obtain a **Rasch-type formulation**
- In both cases we assume that all the $r_j > 2$ response categories share the same discriminating power, in order to keep the conditional probabilities away from crossing and so avoiding degenerate conditional response probabilities

Constraints on the item difficulty parameters

- general situation in which the parameters β_{jy} are unconstrained
- special case in which the parameters β_{jy} are constrained in an additive way (**rating scale parameterization**)

$$\beta_{jy} = \beta_j + \tau_y, \quad j = 1, \dots, J, y = 0, \dots, r - 1$$

- β_j : difficulty of item j
 - τ_y : difficulty of response category y for all j
- with binary items $\beta_{jy} = \beta_j$ for all y

discrimination indices	difficulty levels	type of logit		
		global	local	continuation
free	free	GRM	GPCM	SM
free	constr.	RS-GRM	RS-GPCM	RS-SM
constr.	free	1P-GRM	PCM	SRM
constr.	constr.	1P-RS-GRM	RSM	SRSM

- GRM: Graded Response Model (Samejima, 1969)
- RS-GRM: Rating Scale GRM (Muraki, 1990)
- 1P-GRM: One-parameter GRM (Van der Ark, 2001)
- 1P-RS-GRM: One-parameter rating scale GRM (Van der Ark, 2001)
- GPCM: Generalized Partial Credit Model (Muraki, 1992)
- RS-GPCM: Rating Scale GPCM
- PCM: Partial Credit Model (Masters, 1982)
- RSM: Rating Scale Model (Andrich, 1978)
- SM: Sequential Model
- RS-SM: Rating scale SM
- SRM: Sequential Rasch model (Tutz, 1990)
- Sequential RSM (Tutz, 1990)

Model estimation

- Number of free parameters:

Table : *Number of free parameters*

discriminating indices	difficulty levels	number of free parameters			
		$\#\pi_u$	$\#\xi_{ud}$	$\#\beta_{jy}$	$\#\gamma_j$
free	free	$(k-1)$	sk	$[\sum_{j=1}^J (r_j - 1) - s]$	$(J-s)$
free	constrained	$(k-1)$	sk	$[(J-s) + (r-2)]$	$(J-s)$
constrained	free	$(k-1)$	sk	$[\sum_{j=1}^r (r_j - 1) - s]$	
constrained	constrained	$(k-1)$	sk	$[(J-s) + (r-2)]$	

- An EM algorithm is implemented to maximise the log-likelihood function along the same lines of the basic LC models, where the M-step is modified relying on standard algorithms for the maximisation of the likelihood of a binary/ordinal logit model

Model selection

- The model selection consists in assessing, according to some suitable criterion, the following elements
 - number of latent classes k
 - link function $g(\cdot)$ in case of ordinal items
 - constraints on the item parameters γ_j and β_{jy}
 - number of latent dimensions s and corresponding allocation of items within each dimension δ_{jd}
- We suggest to rely on the **likelihood ratio (LR) test** to **compare nested models** → selection of the number of latent classes and type of link function
and on the **Bayesian information criterion (BIC)** (or other information criteria) to obtain a relative measure of lost information to **compare non-nested models** → selection of item parameters and dimensional structure

Dimensionality and item allocation

- **Case 1: comparison between two alternative formulations**

- Null hypothesis: **items in \mathcal{J}_{d_1} and \mathcal{J}_{d_2}** measure the same latent trait, where \mathcal{J}_d denotes the subset of $\mathcal{J} = \{1, \dots, J\}$ containing the indices of the items measuring the d -th latent trait
- Hypothesis test: an LR test may be performed that compares a **general model with s dimensions** with a **restricted model with $s - 1$ dimensions**, where items in \mathcal{J}_{d_1} and \mathcal{J}_{d_2} are collapsed in the same group, being equal all the other elements of the model (number of latent classes, constraints on item parameters, type of logit)
- LR test statistic: $-2(\hat{\ell}_0 - \hat{\ell}_1)$, where $\hat{\ell}_0$ and $\hat{\ell}_1$ denote the maximum of log-likelihood of the restricted model and of the general model, respectively
- under the null hypothesis, the LR statistics is **asymptotically distributed as a χ_q^2** , where q is given by the difference in the number of parameters between the two nested models being compared

- **Case 2: hierarchical clustering of items**

- An algorithm based on the above described LR test for dimensionality may be repeated in a suitable way so that items measuring the same ability are clustered in a same group
- Similarly to a hierarchical clustering, at each step of the procedure the LR test statistic is computed for **every pair of possible aggregations of items (or groups of items)**
- The aggregation with **the minimum value of the statistic (or equivalently the highest p -value)** is then adopted and the corresponding model is fitted before moving to the next step
- The clustering procedure performs $J - 1$ steps

- This algorithm builds a **sequence of nested models**: the most general one uses a separate dimension for each item and the most restrictive one uses one dimension common to all items
- The output of the clustering algorithm may be displayed through a **dendrogram** that shows the deviance between the initial (J -dimensional) LC model and the model selected at each step of the clustering procedure
- Rule to **cut the dendrogram**: we suggest to choose $\hat{s} = J - h + 1$, where h is the first step for that the p -value corresponding to the LR statistic is smaller than a suitable threshold, such as 0.05

- Data is composed by responses of a sample of 1000 examinees to 5 binary items on the Law School Admissions Tests
- Our aim is to investigate on the dimensionality of the 5 items
- We assume an LC-Rasch model with $k = 3$ latent classes

Table : Hierarchical clustering output

	items	deviance	df	p -value	new group
Step 1	-1 -4	0.194	2	0.908	1 4 0 0 0
Step 2	-5 1	1.346	4	0.854	1 4 5 0 0
Step 3	-2 -3	10.823	6	0.094	2 3 0 0 0
Step 4	2 3	24.529	8	0.002	1 2 3 4 5

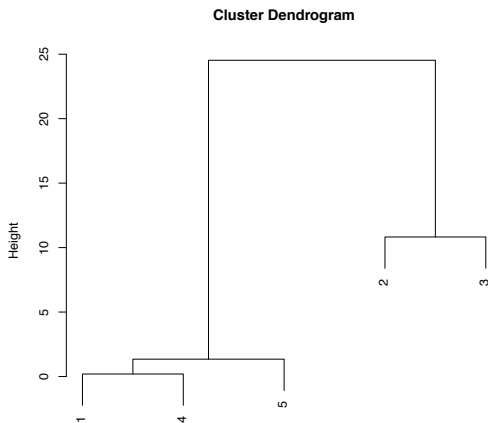


Figure : Dendrogram for the LSAT7 data

Data source

- Data is composed by responses of 201 oncological patients to 14 items about the assessment of anxiety and depression
- Items are equally divided between 2 dimensions: anxiety and depression
- All items have four response categories from the minimum value 0, corresponding to a low level of anxiety or depression, to the maximum value 3, corresponding to a high level of anxiety or depression
- Our aim is to fit an LC-IRT model to explain the data and estimate anxiety and depression

Number of latent class

- to select the optimal value of k , we adopt the **basic LC model**, so that no choice on the link function and the item parameterization is requested and restrictive assumptions on the item dimensionality are avoided
- a comparison among models which differ in the number of latent classes is performed for increasing values of k

Table : Selection of the number of latent classes: maximum log-likelihood value, number of parameters, and BIC index

k	$\hat{\ell}$	# par	BIC
1	-3153.151	42	6529.040
2	-2814.635	85	6080.051
3	-2677.822	128	6034.468
4	-2645.435	171	6197.736

Logit link function

- to select the logit link function, we adopt $\hat{k} = 3$ latent classes and, in order to avoid restrictive assumptions, free item discriminating and difficulties parameters and a completely general multidimensional structure for the data
- a comparison between a model with global logit link and a model with local logit link is carried out on the basis of the BIC index

Table : Selection of the logit link function: maximum log-likelihood value, number of parameters, and BIC index

link	$\hat{\ell}$	# par	BIC
global logit	-2726.348	72	5834.534
local logit	-2741.321	72	5864.479

Test of multidimensionality

- to select the dimensionality structure, we adopt $\hat{k} = 3$ latent classes and a global logit link and, in order to avoid restrictive assumptions, free item discriminating and difficulties parameters
- a comparison between **nested models** is performed, which are characterized by: (i) a J -dimensional structure (M1), (ii) a bidimensional structure (M2) (i.e., anxiety and depression), as suggested by the structure of the questionnaire, (iii) a unidimensional structure (M3)
- Null hypothesis (H_0): models M1 (M2) and M2 (M3) do not significantly differ one other

Table : Test of multidimensionality: maximum log-likelihood value, number of parameters, and BIC index

models	deviance	df	p -value	decision
M1 vs M2	10.722	12	0.553	not reject H_0
M2 vs M3	0.369	1	0.544	not reject H_0

- The hypothesis of unidimensionality cannot be rejected

Item parameterization

- to select the item parameterization, we adopt $\hat{k} = 3$ latent classes, a global logit link, and a unidimensional structure
- a comparison between four nested graded response-type models is performed: (i) GRM (M1), (ii) GRM with free discriminating item parameters and constrained difficulty item parameters (RS-GRM; M2), (iii) GRM with constrained discriminating parameters and free difficulty parameters (1P-GRM; M3), (iv) GRM with both constrained discriminating and difficulty parameters (1P-RS-GRM; M4)

Table : Item parameterization: maximum log-likelihood value, number of parameters, and BIC index

model	$\hat{\ell}$	# par.	BIC	deviance	df	<i>p</i> -value
GRM	-2731.9	59	5777			
RS-GRM	-2795.6	33	5766	127.4 (vs GRM)	26	0.000
1P-GRM	-2741.3	46	5727	18.8 (vs GRM)	13	0.130
1P-RS-GRM	-2844.5	20	5795	206.5 (vs 1P-GRM)	26	0.000

Conclusions

Table : *Unidimensional 1P-GRM: estimates of ξ_{ud} ($d = 1$) and $\hat{\pi}_u$*

	$u = 1$	$u = 2$	$u = 3$
$\hat{\xi}_{u1}$	-0.775	1.184	3.419
$\hat{\pi}_u$	0.342	0.491	0.168

- Item responses can be explained by a **graded response type model** with items having the **same discriminating power and different distances between consecutive response categories**
- Bidimensionality assumption is rejected in favor of the assumption of **unidimensionality**, so that all the items of the questionnaire measure the same latent psychopathological disturb
- Subjects can be classified in **three latent classes**: patients belonging to class 1 present the least severe conditions, whereas patients in class 3 present the worst conditions
- Patients who suffer from psychopathological disturbs are mostly represented in the first two classes, whereas only 16.7% of subjects belong to the third class

R package MultiLCIRT

- Multidimensional LC-IRT models may be estimated through R package `MultiLCIRT`
- The main function for the model estimation is `est_multi_poly`
- The main functions for the dimensionality analysis are `class_item` and `test_dim`
- Other useful functions: `aggr_data` to aggregate original data according to the item response configurations and relative frequencies; `compare_models` to visualise the results of LR tests and BIC values in a unified table

Specification and estimation of multidimensional LC-IRT models

- Model estimation

```
> # 1P-GRM
> out <- est_multi_poly(S, yv, k=3, start=0, link=1,
disc=0, difl=0, multi=1:ncol(S))
```

- `S`: data matrix
- `yv`: vector of weights of response patterns
- `k`: number of latent classes
- `start`: method of initialization of the algorithm (0 = deterministic, 1 = random, 2 = arguments given as input)
- `link`: type of link function: 0 is for basic LC model (default), 1 is for global logits, 2 is for local logits; with binary items `link=1` is the same as `link=2`
- `disc`: indicator of constraints on the discriminating indices: 0 = all equal to one (default), 1 = free; with binary items `disc=0` results in a Rasch model and `disc=1` in a 2PL model

- `dif1`: indicator of constraints on the difficulty levels: 0 = free (default), 1 = rating scale parameterization
- `multi`: matrix with a number of rows equal to the number of dimensions and elements in each row equal to the indices of the items measuring the dimension corresponding to that row (default is one dimension for each item)

→ Examples of `multi` matrices

```
> multi2 <- rbind(c(2,6,7,8,10,11,12),c(1,3,4,5,9,13,14))
```

```
> multi2 # 2 dimensions with 7 items
```

```
      [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,]    2    6    7    8   10   11   12
[2,]    1    3    4    5    9   13   14
```

```
> multi3 <- rbind(c(2,6,7,8,10,0,0,0,0),
c(1,3,4,5,9,13,14,11,12))
```

```
> multi3 # 1 dimension with 5 items; 1 dimension with 9 items
```

```
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
[1,]    2    6    7    8   10    0    0    0    0
[2,]    1    3    4    5    9   13   14   11   12
```

• Output

- `out$Th`: estimated matrix of latent trait levels for each dimension and latent class
- `out$gac`: estimated vector of discriminating indices for each item
- `out$Bec`: estimated vector of difficulty levels for each item; it is split in two vectors if `difl=1`
- All outputs shown for the LC models (with and without covariates) are still valid

Hierarchical clustering of items

- Data structure (Example 1)

```
> library(mirt)
```

```
> data(LSAT7)
```

```
> head(LSAT7)
```

	Item.1	Item.2	Item.3	Item.4	Item.5	freq
1	0	0	0	0	0	12
2	0	0	0	0	1	19
3	0	0	0	1	0	1
4	0	0	0	1	1	7
5	0	0	1	0	0	3
6	0	0	1	0	1	19

```
> S <- as.matrix(LSAT7[,1:5]) # matrix of item responses
```

```
> yv <- as.vector(LSAT7[,6]) # matrix of weights
```

• Clustering of items

```
> out <- class_item(S, yv, k=3, link=1, disc=0)
```

- `S`: data matrix
- `yv`: vector of weights of response patterns
- `k`: number of latent classes
- `link`: type of link function: 1 is for global logits, 2 is for local logits; with binary items `link=1` is the same as `link=2`
- `disc`: indicator of constraints on the discriminating indices: 0 = all equal to one, 1 = free; with binary items `disc=0` results in a Rasch model and `disc=1` in a 2PL model

• Output

```
> plot(out$dend) # dendrogram
> summary(out) # it gives details about the clustering process
```

Clustering output:

	items	deviance	df	p-value	newgroup
Step 1	-1 -4	0.194	2	0.908	1 4 0 0 0
Step 2	-5 1	1.346	4	0.854	1 4 5 0 0
Step 3	-2 -3	10.823	6	0.094	2 3 0 0 0
Step 4	2 3	24.529	8	0.002	1 2 3 4 5

Test of multidimensionality

- The function tests a certain multidimensional model (restricted model) against a larger multidimensional model based on a higher number of dimensions (general model), all other elements being constant
- The function relies on `est_multi_poly`

```
> out <- test_dim(S, yv, k=3, link=1, disc=1, difl=0,
multi0=multi2, multi1=cbind(1:ncol(S)))
```

- `multi0`: matrix specifying the multidimensional structure of the restricted model
 - `multi1`: matrix specifying the multidimensional structure of the general model
- Output

```
> out0 # output for the restricted model obtained from est_multi_poly
> out1 # output for the general model obtained from est_multi_poly
> dev # test statistic
> df # number of degrees of freedom of the test
> pv # p-value for the test
```


Main references

- Bartolucci, F. (2007). A class of multidimensional IRT models for testing unidimensionality and clustering items. *Psychometrika*, 72:141?157
- Bartolucci, F., Bacci, S., Gnaldi, M. (forthcoming 2015). *Statistical analysis of questionnaires: a unified approach based on Stata and R*, Chapman Hall/CRC
- Hambleton, R. K., Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Kluwer Nijhoff, Boston