A class of Multidimensional Latent Class IRT models for ordinal polytomous item responses

Silvia Bacci*[1], Francesco Bartolucci*, Michela Gnaldi*

*Dipartimento di Economia, Finanza e Statistica - Università di Perugia

Università di Pavia, 7-9 September 2011, Pavia

[1] silvia.bacci@stat.unipg.it

Outline

Starting point

Item Response Theory (IRT) models (Van der Linden and Hambleton, 1997) are increasingly used to the assessment of individuals' **latent traits**

They allow us to translate the qualitative information coming from the questionnaire in a quantitative measurement of the latent trait

Main assumptions of traditional IRT models

- Local independence
- Unidimensionality of latent trait
- Normality of latent trait

  Note that as far as the Rasch type models (Rasch, 1960) no special distributive assumption is need, to the detriment of restrictions on items' parameters

Limits of traditional IRT models

- A same questionnaire is usually used to measure **several** latent traits
- We are interested in assessing and testing the **correlation between latent traits**
- Often, normality of latent trait is not a realistic assumption
- In some contexts (e.g. health care) can be not only more realistic, but also more convenient for the decisional process, to assume that population is composed by **homogeneous classes of individuals** who have very similar latent characteristics (Lazarsfeld and Henry, 1968; Goodman, 1974), so that individuals in the same class will receive the same kind of decision (e.g. clinical treatment).

To take into account these elements, several extensions and generalizations of traditional IRT models have been proposed in the literature (see, for instance, Wilson and De Boeck, 2004; Von Davier and Carstensen, 2007)

Multidimensional latent class IRT models

Bartolucci (2007) proposes a class of multidimensional latent class (LC) IRT models characterized by these main features:

1. more latent traits are simultaneously considered (**multidimensionality**)
2. these latent traits are represented by a random vector with a **discrete distribution** common to all subjects (each support point of such a distribution identifies a different latent class of individuals)
3. either a Rasch or a two-parameter logistic (Birnbaum, 1968) parameterisation may be adopted for the probability of a correct response to each **binary item**
4. the conditional probability of a correct response to a given item is constant for subjects belonging to different known groups (e.g. males and females), i.e. **covariates are not included**

Aim of the contribution

Class of multidimensional LC IRT models can be extended in several ways. We are mainly interested in:

1. taking into account **ordinal polytomously-scored items**,
2. including covariates to detect items with differential functioning

In this contribution we treat the first point, whereas the second one is object of the contribution of Gnaldi, Bartolucci and Bacci.

Basic notation (1)

- X_j: response variable for the j-th item, with $j = 1, \ldots, r$
- r: number of items
- l_j: number of categories of item j, from 0 to $l_j - 1$
- $\phi^{(j)}_{x|\theta} = p(X_j = x|\Theta = \theta)$: probability that a subject with ability level θ responds by category x to item j ($x = 1, \ldots, l_j$)
- $\boldsymbol{\phi}^{(j)}_{\theta}$: probability vector $(\phi^{(j)}_{0|\theta}, \ldots, \phi^{(j)}_{l_j-1|\theta})'$
- γ_j: discrimination index of item j
- β_{jx}: difficulty parameter of item j and category x
- $g_x(\cdot)$: link function specific of category x

Classification criteria (1)

On the basis of the specification of the link function $g_x(\cdot)$ and on the basis of the adopted constraints on the item parameters γ_j and β_{jx}, different IRT models for polytomous responses result.

Three classification criteria may be specified:

- *Type of link function*
  - global (or cumulative) logits

$$g(\phi_\theta^{(j)}) = \log \frac{\phi_{x|\theta}^{(j)} + \cdots + \phi_{l_j|\theta}^{(j)}}{\phi_{0|\theta}^{(j)} + \cdots + \phi_{x-1|\theta}^{(j)}} = \log \frac{p(X_j \geq x|\theta)}{p(X_j < x|\theta)}, \quad x = 1, \ldots, l_j - 1,$$

  - local (or adjacent category) logits

$$g_x(\phi_\theta^{(j)}) = \log \frac{\phi_{x|\theta}^{(j)}}{\phi_{x-1|\theta}^{(j)}} = \log \frac{p(X_j = x|\theta)}{p(X_j = x-1|\theta)}, \quad x = 1, \ldots, l_j - 1,$$

Classification criteria (2)

- *Constraints on discrimination parameters*
  - each item may discriminate differently from the others
  - all the items discriminate in the same way: $\gamma_j = 1, j = 1, \ldots, r$.
- *Constraints on items and thresholds difficulty parameters*
  - each item differs from the others for different distances between consecutive response categories
  - the distance between difficulty levels from category to category within each item is the same across all items (rating scale parameterisation): $\beta_{jx} = \beta_j + \tau_x$, where β_j indicates the difficulty of item j and τ_x is the difficulty of response category x, independently of item j

Types of IRT models for ordinal polytomous items

Table 1: List of IRT models for polytomous responses

discrimination indices	difficulty levels	resulting parameterisation	resulting model Global logits	Local logits
free	free	$\gamma_j(\theta - \beta_{jx})$	GRM	GPCM
free	constrained	$\gamma_j[\theta - (\beta_j + \tau_x)]$	RS-GRM	GRSM
constrained	free	$\theta - \beta_{jx}$	1P-GRM	PCM
constrained	constrained	$\theta - (\beta_j + \tau_x)$	1P-RS-GRM	RSM

GRM: Graded Response Model (Samejima, 1969)
RS-GRM: Graded Response Model with a Rating Scale parameterisation
1P-GRM: Graded Response Model with fixed γ_j
1P-RS-GRM: Graded Response Model with a Rating Scale param. and fixed γ_j
GPCM: Generalized Partial Credit Model (Muraki, 1990)
GRSM: Generalized Rating Scale Model
PCM: Partial Credit Model (Masters, 1982)
RSM: Rating Scale Model (Andrich, 1978)

Basic notation (2)

- s: number of latent variables corresponding to the different traits measured by the items
- $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_s)$: vector of latent variables
- $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_s)$: one of the possible realizations of $\boldsymbol{\Theta}$
- δ_{jd}: dummy variable equal to 1 if item j measures latent trait of type d, $d = 1, \ldots, s$
- k: number of latent classes of individuals

Assumptions

- Items are ordinal polytomously-scored
- The parameterisation is one of those illustrated in Table 1
- The set of items measures s **different latent traits**
- Each item measures only one latent trait
- The random vector $\boldsymbol{\Theta}$ has a **discrete distribution** with support points $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_k\}$ and weights $\{\pi_1, \dots, \pi_k\}$
- The number k of latent classes is the same for each latent trait
- Manifest distribution of the full response vector $\mathbf{X} = (X_1, \dots, X_k)'$:

$$p(\boldsymbol{X} = \boldsymbol{x}) = \sum_{c=1}^{C} p(\boldsymbol{X} = \boldsymbol{x} | \boldsymbol{\Theta} = \boldsymbol{\xi}_c)\pi_c$$

where $\pi_c = p(\boldsymbol{\Theta} = \boldsymbol{\zeta}_c)$ and (assumption of *local independence*)

$$p(\boldsymbol{X} = \boldsymbol{x} | \boldsymbol{\Theta} = \boldsymbol{\xi}_c) = \prod_{j=1}^{r} p(X_j = x | \boldsymbol{\Theta} = \boldsymbol{\xi}_c)$$

Some examples of models

- Multidimensional LC GRM model (the most general model with global logit link):

$$\log \frac{P(X_j \geq x|\boldsymbol{\theta})}{P(X_j < x|\boldsymbol{\theta})} = \gamma_j(\sum_{d=1}^{s} \delta_{jd}\theta_d - \beta_{jx}) \quad x = 1, \ldots, l_j - 1$$

- Multidimensional LC GPCM model (the most general model with local logit link):

$$\log \frac{P(X_j = x|\boldsymbol{\theta})}{P(X_j = x-1|\boldsymbol{\theta})} = \gamma_j(\sum_{d=1}^{s} \delta_{jd}\theta_d - \beta_{jx}) \quad x = 1, \ldots, l_j - 1$$

- Multidimensional LC RSM model (the most special model with local logit link):

$$\log \frac{P(X_j = x|\boldsymbol{\theta})}{P(X_j = x-1|\boldsymbol{\theta})} = \sum_{d=1}^{s} \delta_{jd}\theta_d - (\beta_j + \tau_x) \quad x = 1, \ldots, l - 1$$

Maximum log-likelihood estimation

Let i denote a generic subject and let $\boldsymbol{\eta}$ the vector containing all the free parameters. The log-likelihood may be expressed as

$$\ell(\boldsymbol{\eta}) = \sum_i \log[p(\boldsymbol{X}_i = \boldsymbol{x}_i))]$$

- Estimation of $\boldsymbol{\eta}$ may be obtained by the **discrete (or LC) MML approach** (Bartolucci, 2007)
- $\ell(\boldsymbol{\eta})$ may be efficiently maximize by the **EM algorithm** (Dempster et al., 1977)
- The software for the model estimation has been implemented in MATLAB
- Number of free parameters is given by:

$$\#\text{par} = (k-1) + sk + \Big[\sum_{j=1}^{r}(l_j - 1) - s\Big] + a(r-s), \quad a = 0, 1,$$

where $a = 0$ when $\gamma_j = 1, \forall j = 1, \ldots, r$, and $a = 1$ otherwise

A strategy for the model selection

1. selection of the optimal number k of latent classes
2. selection of the type of link function
3. selection of constraints on the item discrimination and difficulty parameters
4. selection of the number of latent traits and detection of the item allocation within each dimension

Some remarks

- As concerns the ordering of the mentioned steps, the last three steps may be considered flexible, being their inversion acceptable and formally correct, since it leads to identical results
- As concerns the choice of k, to avoid problems with multimodality of log-likelihood function we suggest to repeat the step with different random and deterministic starting values
- Comparison between models at each step of the selection process may be driven by an information criterion, such as BIC index, or, whereas compared models are nested, by a likelihood ratio (LR) or a Wald test
- Models compared at each step of the selection process differ only by one type of element (k, link function, constraints on item parameters, or s), all other elements being equal
- To avoid too restrictive assumptions, at each step of the selection process we suggest to adopt the most general parameterisations as concerns the choice of those elements selected in subsequent steps (e.g. we should base the selection of k on the basis of standard LC model).

The data

- A set of 200 oncological Italian patients investigated about anxiety and depression
- Anxiety and depression assessed by the "Hospital Anxiety and Depression Scale" (HADS) (Zigmond and Snaith, 1983)
    - 2 latent traits: anxiety and depression
    - 14 polytomous items: minimum 0 indicates a low level of anxiety or depression; maximum 3 indicates a high level of anxiety or depression
    - Mean raw score for anxiety = 7.11 ($\sigma = 4.15$); mean raw score for depression = 7.17 ($\sigma = 4.17$)
    - Correlation between raw scores on anxiety and on depression results very high and equal to 0.98

Choice of k

Table 2: BIC values and log-likelihood (ℓ) for $k = 1, \dots, 10$ latent classes

	Deterministic start		Random start	
k	BIC	ℓ	BIC (min)	ℓ (max)
1	6529,040	-3153,151	6529,040	-3153,151
2	6080,051	-2814,635	6080,051	-2814,635
3	**6034,468**	-2677,822	**6027,791**	-2674,484
4	6197,736	-2645,435	6104,805	-2598,970
5	6415,568	-2640,330	6226,510	-2545,801
6	6610,982	-2624,016	6350,194	-2493,622
7	6823,831	-2616,420	6521,221	-2465,115
8	7040,847	-2610,907	6673,266	-2427,116
9	7274,232	-2613,578	6852,946	-2402,935
10	7467,897	-2596,389	7025,803	-2375,342

Choice of link function and constraints on item parameters

Table 3: Link function selection: BIC values and log-likelihood (ℓ) for the graded response- and the partial credit-type models

	Global logit	Local logit
BIC	**5780,696**	5817,877
ℓ	-2731,249	-2749,839

Table 4: Item parameters selection: log-likelihood, BIC values and LR test results (deviance and p-value) between nested models

Model	ℓ	BIC	Deviance		p-value
GRM	-2731,249	5780,696	–		–
RS-GRM	-2798,959	5778,230	135.4195	(vs GRM)	0.011
1P-GRM	-2740,658	**5735,875**	18.8185	(vs GRM)	0.093
1P-RS-GRM	-2843,227	5803,127	205.1375	(vs 1P-GRM)	0.000

Choice of dimensionality

Table 5: Bidimensional 1P-GRM and unidimensional 1P-GRM: log-likelihood, BIC value and LR test results (deviance and p-value)

1P-GRM	ℓ	BIC	Deviance	p-value
Bidimensional	-2740,658	5735,875	–	–
Unidimensional	-2741,285	**5726,521**	1.2529	0.5345

The selected model

In conclusion, we select a model based on a parameterisation of type 1P-GRM, with 3 latent classes and only one latent trait

$$\log \frac{P(X_j \geq x|\theta = \xi_c)}{P(X_j < x|\theta = \xi_c)} = \theta - \beta_{jx} \quad x = 1, \ldots, l_j - 1 \quad c = 1, 2, 3$$

Table 6: Estimated support points $\hat{\xi}_c$ and weights $\hat{\pi}_c$ of latent classes for the unidimensional 1P-GRM.

	Latent class c		
Latent trait	1	2	3
Psychopatological disturbs	-0.776	1.183	3.418
Probability	0.342	0.491	0.167

- Patients who suffer from psychopatological disturbs are mostly represented in the first two classes
- Patients in class 1 present the least severe conditions; patients in class 3 present the worst conditions

Conclusions

- In this contribution we extended the class of Multidimensional LC IRT models of Bartolucci (2007) to ordinal polytomous items
- The proposed class of models is flexible because it allows us several different types of parameterisations
- It is based on two main assumptions: (i) multidimensionality and (ii) discreteness of latent trait
- Further developments:
    - studying the problem of multimodality of the log-likelihood function
    - extending the proposed approach to take into account hierarchical data structures

Main references

ANDRICH, D. (1978). A rating formulation for ordered resopnse categories. Psychometrika, 43:561Ð573.

BARTOLUCCI F., A class of multidimensional IRT models for testing unidimensionality and clustering items, *Psychometrika*, vol. 72, 2007, p. 141-157.

BIRNBAUM A., Some latent trait models and their use in inferring an examinee's ability, Lord, F. M. and Novick, M. R. (eds.), *Statistical Theories of Mental Test Scores*,1968, Addison-Wesley, Reading (MA), p. 395-479.

DEMPSTER A.P., LAIRD N.M., RUBIN D.B., Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, vol. 39, 1977, p. 1-38.

GOODMAN L.A., Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, vol. 61, 1974, p. 215-231.

LAZARSFELD P.F., HENRY N.W., *Latent structure analysis*, Boston, Houghton Mifflin, 1968.

MASTERS, G. (1982). A Rasch model for partial credit scoring. Psychometrika, 47:149Ð174.

MURAKI, E. (1990). Fitting a polytomons item response model to Likert-type data. Applied Psychological Measurement, 14:59Ð71.

Main references

RASCH G., On general laws and the meaning of measurement in psychology, *Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability*, vol. 4, 1961, p. 321-333.

SAMEJIMA, F. (1969). Estimation of ability using a response pattern of graded scores. Psychometrika Monograph, 17.

VAN DER LINDEN W.J. and HAMBLETON R.K., Handbook of modern item response theory, Springer, 1997.

VON DAVIER M. and CARSTENSEN C.H., Multivariate and mixture distribution Rasch models, Springer, New York, 2007.

WILSON M. and DE BOECK P., Explanatory item response models: a generalized linear and nonlinear approach, Springer-Verlag, New York, 2004.

ZIGMOND A.S., SNAITH R.P., The hospital anxiety and depression scale, *Acta Psychiatrika Scandinavica*, vol. 67 no. 6, 1983 p. 361-370.