Rapporto n. 1 2008

dmsia unibg.it

A. Punzo

Remarks on the Existence of CML Estimates for the PCM by means of the R Package eRm*

Serie Ricerca

Dipartimento di Matematica, Statistica, Informatica e Applicazioni "Lorenzo Mascheroni"

UNIVERSITÀ DEGLI STUDI DI BERGAMO

Remarks on the Existence of CML Estimates for the PCM by means of the R Package eRm[*]

Antonio Punzo

Abstract

Mair and Hatzinger (2007b) have recently proposed in the Journal of Statistical Software the R package **eRm** (extended Rasch models) for computing Rasch models and several extensions. Undoubtedly, in the eRm class the partial credit model (PCM) – for practical testing purposes – is one of the best known. The package, using a unitary conditional maximum likelihood (CML) procedure, estimates the item parameters of the above-mentioned models.

Although the eRm belong to the Rasch family of models and share their distinguishing characteristics, they suffer from the problem of possible non-existence of estimates. In literature, both in the joint and in the conditional ML approach, the configurations and the conditions of non-existence for the RM are well-known (Fischer 1981). The **eRm** package performs a preliminary data check only for the RM. The conditions of non-existence are known for the PCM only in the joint case (Bertoli-Barsotti 2005).

In this article, the main focus is on the PCM; the above-mentioned JML non-existence configurations for this model will be the starting point. A class of counter examples is illustrated, which leads to "false" CML estimates with the **eRm** package, i.e., values that appear to be estimates but, through a more accurate analysis of the maximization function, they are rather a clear signal of non-existence. Moreover, the obtained results emphasize the presence of additional CML non-existence configurations, compared to those valid in the JML case.

Keywords: Rasch model, partial credit model, conditional maximum likelihood estimate, R package **eRm**.

1. Introduction

More generally, the analysis of the relation between latent continuous variables and observed categorical variables – which can be either dichotomous or (ordered) polytomous – is known as Latent Trait Analysis (LTA). In psychometrics and educational testing, LTA is called Item Response Theory (IRT).

[*]Lavoro svolto nell'ambito del progetto di ricerca ex60%, fondi di Ateneo 2006, dal titolo: "Effetti Di Disegni A Blocchi Sulla Stima Del Modello Di Rasch".

In this paper the attention will be focused on *polytomous* IRT models (see, just to cite a few, Ostini and Nering 2006, or van der Linden and Hambleton 1997a, Chap. I). Polytomous items are categorical items in the same way as dichotomous items; they simply have more than two possible response categories. The simplest and most obvious reason for the development of polytomous IRT models is the fact that polytomous items exist and are commonly used in applied psychological measurement. The need for polytomous response formats may be most acute in the measurement of personality and social variables. Kamakura and Balasubramanian (1989, p. 514) suggest that dichotomous distinctions are often less clear in this context than in ability measurement settings and that "more subtle nuances of agreement/disagreement" are needed than dichotomous items permit. Similarly, Cox (1980) argues that items with only two response alternatives are inadequate in this context because they cannot transmit much information and they frustrate respondents. Psychometric issues also exist that make polytomous items attractive in comparison to dichotomous items. At a general level, such an issue is that polytomous items measure across a wider range of the trait *continuum* than the dichotomous items. This simply occurs by virtue of the fact that polytomous items contain more response categories than dichotomous items do. Masters (1988) and Bejar (1977) note that the purpose of using more than two categories per item is to try to obtain more information about the trait level of the people being measured so that more precise trait-level estimates can be obtained. Samejima (1976, 1979) demonstrates the increase in statistical information that is available from a polytomous IRT model in comparison to a dichotomous model. Conversely, though in a different context, Cohen (1983) demonstrates that reducing continuous or multiple category data to the dichotomous level leads to a systematic loss of measurement information.

The discussion will concentrate on the case of polytomous models for items with *ordered categories*. Ordered polytomous items are simply those where the response categories have an explicit rank ordering with respect to the trait of interest. In this context, the partial credit model (PCM; Masters 1982) is one of the best known and simple unidimensional parametric IRT models for polytomous items with ordered categories. The PCM differs from the other IRT models of this category in that it belongs to the Rasch family of models (i.e., models that attempt to conform to the fundamental measurement theory) and shares the distinguishing characteristics of the family: separable person and item parameters, sufficient statistics, and, hence, conjoint additivity. These features enable "specifically objective" comparisons of persons and items (Rasch 1960, 1977; Fisher Jr 1992) and allow each set of model parameters to be conditioned out of the estimation procedure. The PCM can be considered – conceptually and structurally – as the basis for almost all other unidimensional polytomous Rasch models. This includes the rating scale, Binomial trials, and Poisson counts models (cf. Wright and Masters 1982), as well as other models for rating data developed by Andrich (1982) and by Rost (1988). Furthermore, in the special case of an item with two ordered categories, the PCM becomes the simple dichotomous Rasch model (RM; Rasch 1960). The importance of the model arises from these characteristics.

The PCM contains only *two* sets of parameters: one for persons and one for items. All parameters in the model are *locations* on an underlying variable. This feature distinguishes the PCM from measurement models including item "discrimination" or "dispersion" parameters, which qualify locations thus confounding the interpretation of variables.

The simplicity of the model formulation makes it easy to implement, and a range of software packages are dedicated to it (see, e.g., Adams and Khoo 1991; Sheridan, Andrich, and Luo

2002; Linacre and Wright 2004; Mair and Hatzinger 2007b). Successful applications of the PCM to a wide variety of measurement problems have been reported in the literature (see, e.g., Andersen 1995b; van der Linden and Hambleton 1997a, pp. 101–102). This model and its main characteristics are presented in Section 2.

Parameter estimates are the only realizations of the ideas embodied in measurement models. A variety of estimation approaches are available for IRT models in general and for Rasch model in particular (see, e.g., Baker and Kim 2004). For Rasch models, the commonly used approaches are: *joint maximum likelihood* (JML; Wright and Panchapakesan 1969), *conditional maximum likelihood* (CML) and *marginal maximum likelihood* (MML; Bock and Aitkin 1981). With the term "Rasch-based measures" one refers to item/person parameter estimates arising from one of these approaches. In the implementation of partial credit analysis, all parameters can be estimated for an item only if observations occur in each of the available response categories (if one or more response categories for an item are unused, these categories are said "null").

The joint approach estimates person and item parameters simultaneously. This method suffers from a serious drawback. Martin-Löf (1973), see also Andersen (1970, 1973); Haberman (1977); Andersen (1980, Chap. 6), showed that the JML estimates for the item parameters are inconsistent when n (the number of subjects) approaches infinity whereas k (the number of items) and m (the number of categories) are fixed. At the same time the individual parameter estimates can only take a finite number of values and cannot accordingly approach the true value. To go into detail of the dichotomous case see, for example, Haberman (1977) or Ghosh (1995).

As an alternative, Rasch (1960) suggested estimating the item parameters by the CML method, where the conditioning is with respect to the sufficient statistics for the individual parameters. Rasch supports CML estimation because it is the only estimation method within the Rasch measurement context fulfilling the requirement of *person-free item calibration* and, thus, it maps the epistemological theory of specific objectivity to a statistical ML framework. Under suitable conditions on the variability of the ability θ in the population, these CML estimates are consistent, as shown in Andersen (1970, 1973).

MML approach treats person parameters as "nuisance" parameters and removes them (by integration) from the likelihood function by assuming that persons are sampled randomly from a population in which ability θ is distributed according to some (parametric or nonparametric) density function $f(\theta)$. The MML estimates are asymptotically equivalent to the CML ones and provide, consequently, consistent estimators (Pfanzagl 1994); however, if the density specification is inadequate, MML is inferior to CML.

For the sake of completeness, some other methods for the estimation of the item parameters are present in literature. The pairwise conditional ML method, in which items are taken in pairs to eliminate the person parameter, is an adaptation of the CML. The estimates for the dichotomous case have been shown to be consistent (Zwinderman 1995); the same result may holds in the polytomous case as well (Wright and Masters 1982; Garner and Engelhard Jr 2002). This estimation method has been implemented in the RUMM2020 program (Sheridan *et al.* 2002). Moreover, Anderson, Li, and Vermunt (2007) propose a pseudo-ML approach, Linacre (2004a) and Molenaar (1995) give an overview of various (heuristic) non-ML methods, and Bayesian techniques can be found in Baker and Kim (2004, Chap. 7).

In this paper CML is used for the above-mentioned desirable epistemological and mathematical properties. This conditional approach, for the PCM, is described in Section 3. The discussion is focused on the existence of the CML estimates of the PCM item parameters under the assumption of a complete data matrix. Section 4 is devoted to the presentation of the state of art on this issue for both JML and CML case.

In Section 5, some configurations of non-existence of JML estimates has been analyzed, observing that they are still problematic in the CML setting. They indeed produce false critical points detection in the maximization process. This issue has been the point of departure of a systematic analysis concerning fixed small-dimensional datasets, developed in Section 6. The results have been obtained using the R (R Development Core Team 2007) package **eRm** (Mair and Hatzinger 2007a,b).

2. The model

Consider the responses of a n-dimensional set $\mathcal{S} = \{S_1, \ldots, S_v, \ldots, S_n\}$ of subjects to a k-dimensional sequence $\mathcal{I} = \{I_1, \ldots, I_i, \ldots, I_k\}$ of items. Each subject may respond to item I_i in $m+1$ ($m \geq 1$) ordered categories, $C_0, C_1, \ldots, C_h, \ldots, C_m$. The PCM can be applied in any situation in which the item is conceptualized as a series of ordered steps and the respondent receives a unitary credit for each successfully completed step (Wright and Masters 1982). The response *scores* are chosen to be $0, 1, \ldots, m$ in correspondence to $C_0, C_1, \ldots, C_m$, respectively. It is convenient to write the actual response for individual S_v to I_i as the selection vector $\boldsymbol{x}'_{vi} = (x_{vi0}, x_{vi1}, \ldots, x_{vim})$ (as usual, the transposition is indicated with a prime; all vectors in this paper are to be regarded as column vectors), where $\boldsymbol{x}_{vi}$ is an observation from the random variable $\boldsymbol{X}_{vi}$ and $x_{vih} = 1$ if the response is in category C_h, and 0 otherwise. Let $\boldsymbol{x}_{vi}$ be also a single element in the $n \times k$ data matrix $\boldsymbol{x}$. The model assumes that, for each item, the subject chooses one and only one of the $m+1$ categories. Moreover, let

$$s_{vi} = \max_{h \in \{0,1,\ldots,m\}} h x_{vih} = \sum_{h=0}^{m} h x_{vih}$$

be the *score* of a respondent S_v to an item I_i. Naturally, $s_{vi} \in \{0, 1, \ldots, m\}$. Let $\boldsymbol{s} = (s_{vi})$ be the *score matrix*. It is to be noted that there is a one-to-one correspondence between the 3-dimensional data matrix $\boldsymbol{x}$ and the 2-dimensional score matrix $\boldsymbol{s}$.

Usually, the model is introduced through the specification of the probability – the so-called probability function (pf) – that a subject S_v, with parameter θ_v, will respond to item I_i in category C_h. There are several equivalent parameterizations of the pf (cf. Masters and Wright 1997). The original one, used by Masters (1982), is the following

$$\mathsf{P}\left(X_{vih} = 1 \,|\, \theta_v, \boldsymbol{\delta}_i\right) = \frac{\exp\left(\theta_v h - \sum_{l=0}^{h} \delta_{il}\right)}{\sum_{t=0}^{m} \exp\left(\theta_v t - \sum_{l=0}^{t} \delta_{il}\right)}, \qquad h = 0, 1, \ldots, m, \tag{1}$$

where θ_v are *person parameters*, δ_{ih} are called *uncentralized threshold parameters* and $\boldsymbol{\delta}'_i = (\delta_{i0}, \delta_{i1}, \ldots, \delta_{ih}, \ldots, \delta_{im})$ is the item parameter vector related to I_i. It is easy to note that, while θ_v represent the position of S_v on the underlying unidimensional latent trait, the single item I_i is characterized by a $(m+1)$-dimensional vector $\boldsymbol{\delta}_i$ so that the model is capable of accommodating each possible interaction between items and categories. This fact may be considered cumbersome for the interpretation of the parameters (Hemker 2001; Verhelst, Glas, and De Vries 1997). For example, Sijtsma and Hemker (2000, p. 395) conclude that "no *item* difficulty exists" for the PCM. Nevertheless, a scalar parameter determining the "item location" is usually obtained quite naturally through the following reparameterization

$$\mathsf{P}\left(X_{vih}=1 \mid \theta_v, \alpha_i, \boldsymbol{\tau}_i\right)=\frac{\exp\left[\left(\theta_v-\alpha_i\right)h-\sum_{l=0}^{h}\tau_{il}\right]}{\sum_{t=0}^{m}\exp\left[\left(\theta_v-\alpha_i\right)t-\sum_{l=0}^{t}\tau_{il}\right]}, \qquad h=0,1,\ldots,m, \tag{2}$$

where the τs are called *centralized threshold parameters* whereas the parameter $\alpha_i=\delta_{ih}-\tau_{ih}$ is said *mean difficulty* and it can be understood as "item location". By definition of α_i, it holds $\sum_{h=0}^{m}\tau_{ih}=0$. The parameterization (2) allows, besides, the following consideration: when $\tau_{ih}=\tau_h$, the PCM gives the rating scale model (RSM; Andrich 1978) (in applications, the PCM is often preferred to the RSM because it allows different thresholds for different items).

In this paper the PCM is defined, conforming to Mair and Hatzinger (2007a,b), using the parameterization due to Andersen (1983):

$$\mathsf{P}\left(X_{vih}=1 \mid \theta_v, \boldsymbol{\beta}_i\right)=\frac{\exp\left(\theta_v h-\beta_{ih}\right)}{\sum_{t=0}^{m}\exp\left(\theta_v t-\beta_{it}\right)}, \qquad h=0,1,\ldots,m, \tag{3}$$

where $\beta_{ih}=\sum_{l=0}^{h}\delta_{il}$ are *item-category parameters* and $\boldsymbol{\beta}_i'=\left(\beta_{i0},\beta_{i1},\ldots,\beta_{ih},\ldots,\beta_{im}\right)$, is the item parameter vector related to I_i (see, also, Andersen 1980; Fischer and Ponocny 1994; Andersen 1995a). The only difference between this parameterization and one given in Mair and Hatzinger (2007a,b), is the minus sign that link the two sets of parameters (due to a better interpretation of the item parameters). The choice of this parameterization is in view of the use of (R Development Core Team 2007) package **eRm** in the following.

It is to be noted that if $\theta_v^*=\theta_v+c$ and $\beta_{ih}^*=\beta_{ih}+hc$ for any constant $c\in$, then $h\theta_v^*-\beta_{ih}^*=h\theta_v-\beta_{ih}$. In order to avoid overparameterization, and to allow the identifiability of the model, the parameters must be normalized, for instance, with the following $k+1$ constraints:

$$\beta_{i0}=\delta_{i0}=0, \qquad i=1,\ldots,k, \tag{4}$$

and

$$\sum_{i=1}^{k}\sum_{t=1}^{m}\beta_{it}=0. \tag{5}$$

Really, one could replace the condition (4), alternatively, with $\beta_{ih^*}=0$ for any other category C_{h^*}, $i=1,\ldots,k$ (cf. Fischer and Ponocny 1995; Andersen 1995a, p. 278); in the same way, one could substitute the condition (5) with, for example: $\sum_{v=1}^{n}\theta_v=0$, $\beta_{11}=0$, $\theta_1=0$, $\sum_{i=1}^{k}\alpha_i=0$, etc.. (cf. Linacre and Wright 2004; Wright and Masters 1982, p. 89). There are thus $n+k\left(m+1\right)-k-1=n+km-1$ unconstrained parameters to be estimated from (3).

It is worthwhile to recall that the PCM, like all polytomous Rasch models, is built on the successive dichotomization of adjacent categories. In particular, it can be defined by means of *adjacent-category logits*:

$$\ln\left\{\frac{\mathsf{P}\left[X_{vih}=1\right]}{\mathsf{P}\left[X_{vi(h-1)}=0\right]}\right\}=\theta_v-\delta_{ih},\qquad h=1,\ldots,m.$$

Moreover, it is straightforward to show that the model defines an exponential family (Andersen 1983).

3. CML approach to estimate the item parameters

The essentials of the CML procedure are here summarized. The usual "dot" notation is adopted (e.g., $x_{v\bullet h}$ stands for $\sum_{i=1}^{k} x_{vih}$, and so on).

As mentioned briefly, the separability of parameters, stressed by Rasch as the main justification for the model, is connected with the use of CML procedure. The item parameters are thus estimated based on the conditional likelihood, given the person's raw scores

$$r_v=\sum_{i=1}^{k}\sum_{h=0}^{m} h x_{vih}=\sum_{h=0}^{m} h x_{v\bullet h}$$

which are, according to standard results for the exponential family, minimal sufficient statistics for the person parameters θ_v, $v=1,\ldots,n$. Thus, by conditioning the likelihood onto $\boldsymbol{r}'=(r_1,\ldots,r_v,\ldots,r_n)$, the person parameters $\boldsymbol{\theta}'=(\theta_1,\ldots,\theta_v,\ldots,\theta_n)$, which in this context are nuisance parameters, vanish from the likelihood equation, thus, leading to the estimated item parameters $\widehat{\boldsymbol{\beta}}=\left(\widehat{\boldsymbol{\beta}}_1,\ldots,\widehat{\boldsymbol{\beta}}_i,\ldots,\widehat{\boldsymbol{\beta}}_k\right)$.

Given the nk independent observations $\boldsymbol{x}_{vi}$, $v=1,\ldots,n$, $i=1,\ldots,k$, the CML estimate of the parameters $\boldsymbol{\beta}$ is obtained by maximizing the following *conditional likelihood function*

$$L_C(\boldsymbol{\beta})=\frac{\exp\left[-\sum_{v=1}^{n}\sum_{i=1}^{k}\sum_{h=0}^{m}\beta_{ih}x_{vih}\right]}{\prod_{r=1}^{mk-1}\left[\gamma_r(\boldsymbol{\beta})\right]^{n_r}}=\frac{\exp\left[-\sum_{i=1}^{k}\sum_{h=0}^{m}\beta_{ih}x_{\bullet ih}\right]}{\prod_{r=1}^{mk-1}\left[\gamma_r(\boldsymbol{\beta})\right]^{n_r}},\tag{6}$$

where n_r is the number of persons with a particular score r and

$$\gamma_r(\boldsymbol{\beta})=\sum_{(r)}\exp\left(-\sum_{i=1}^{k}\sum_{h=0}^{m}\beta_{ih}x_{vih}\right),\qquad r=0,1,\ldots,m.\tag{7}$$

Equation (7) defines the known *elementary symmetric functions* and $\sum_{(r)}$ is the sum over all response vectors that produce the score r. Based on (6), it is straightforward to note that the item/category totals $x_{\bullet ih}$ are minimally sufficient for β_{ih}. Moreover, the product $\prod_{r=1}^{mk-1}$ in equation (6) implicitly excludes extreme person's raw score of 0 or mk because these subjects do not affect the conditioned procedure.

The $\gamma_r(\boldsymbol{\beta})$ functions are crucial for the CML approach. An elaborated derivation of these terms for ordinary RM can be found in Fischer (1974) and an overview of various computation algorithms is given in Liou (1994). For example, in the **eRm** package the numerically stable *summation algorithm*, as suggested by Andersen (1972), is implemented. In the past, computational problems (speed, numerical accuracy) involved in calculating the elementary symmetric function limited the practical usage of the CML approach (see, e.g., Gustafsson 1980). Nowadays, these issues are less crucial due to increased computer power.

Alternatively, starting from relation (6), the CML estimate of the parameters $\boldsymbol{\beta}$ can be obtained by maximizing the *conditional log-likelihood function*

$$l_C(\boldsymbol{\beta}) = \ln\left[L_C(\boldsymbol{\beta})\right] = -\sum_{i=1}^{k}\sum_{h=0}^{m}\beta_{ih}x_{\bullet ih} - \sum_{r=1}^{mk-1} n_r \ln\left[\gamma_r(\boldsymbol{\beta})\right]. \tag{8}$$

Based on (8), it is straightforward to realize that the conditional pf belongs to the exponential family with minimal representation. Then, it is known that the conditional log-likelihood function (8) is strictly concave (this issue is important in phase of maximization).

Maximization of $l_C(\boldsymbol{\beta})$ follows by equating to 0 all derivatives with respect to β_{ih}. To do this, it is useful to note that

$$\frac{\partial\ \gamma_r(\boldsymbol{\beta})}{\partial\ \beta_{ih}} = -\gamma_{r-h}^{(i)}(\boldsymbol{\beta}), \qquad i=1,\dots,k, \quad h=0,1,\dots,m, \tag{9}$$

where $\gamma_{r-h}^{(i)}(\boldsymbol{\beta})$ denotes the elementary symmetric function evaluated by omitting item I_i. Some algebra leads to the following set of CML equations:

$$x_{\bullet ih} = \exp(-\beta_{ih})\sum_r n_r \frac{\gamma_{r-h}^{(i)}(\boldsymbol{\beta})}{\gamma_r(\boldsymbol{\beta})}, \qquad i=1,\dots,k, \quad h=0,1,\dots,m. \tag{10}$$

To solve the CML equations (10) most computer algorithms (among them the algorithm of the **eRm** package) use a Newton-Raphson procedure, which is fast in the sense that it usually requires few iterations, and in the sense that the quantities involved are ratios of elementary symmetric functions, easily obtained from convenient recurrence relations (cf. Andersen 1995a, relation (15.26), p. 279).

4. Existence of ML estimates

The aim of this section is to resume the state of the art about the existence of a finite solution to ML estimation equations. Generally speaking, the n.s. conditions for the existence and uniqueness of the ML estimates are known in literature (cf. Barndorff-Nielsen 1978, Theorem 8.2, p. 117 and Corollary 9.6, p. 153); for an alternative n.s. condition, see Bertoli-Barsotti (2002).

For the simple RM, Fischer (1981) elaborates, both for complete and incomplete designs, the n.s. conditions for the existence and uniqueness of a solution of the joint and conditional ML estimation equations. The basic critical condition (cf. Fischer and Molenaar 1995, p. 43) is essentially the same in both cases and it occurs when the set $\mathcal{I}$ of the items can be subdivided into two non-empty subclasses, $\mathcal{I}_1$ and $\mathcal{I}_2$, such that all persons have either positive (or missing) answers to all items in the class $\mathcal{I}_1$, or negative (or missing) answers to all items in the class $\mathcal{I}_2$. Then the items from the class $\mathcal{I}_2$ appear to be “infinitely more difficult” than those from the class $\mathcal{I}_1$, and no comparison of item parameters from the two classes is possible.

Such data are called *ill-conditioned*. If no such subclasses exist, the data are said to be *well-conditioned*. The latter property is n.s. for the existence of a finite, unique normalized CML solution (in this context, the R package **eRm** performs a preliminary data check to confirm the presence/absence of ill-conditioning); the same property is only necessary in the JML approach. It becomes also sufficient, provided that the subjects with zero and perfect scores have been removed. Fischer (1981) also presents useful methods for establishing well-conditioning, both for complete and incomplete designs. For the complete data case, Pfanzagl (1994) gives asymptotic results indicating that Fisher's conditions are almost always fulfilled for large enough samples of persons.

Unfortunately, more complicated conditions are necessary for the existence of the ML estimate in the PCM context. Undoubtedly, a source of trouble is the presence of a null category in the data matrix. Clearly, as a result of algebraic simplifications occurring with binary data, in the special case of the simple RM the condition of a null category, for an item I_i, coincides with the presence of a perfect item total score or a null item total score. When $m > 1$ the situation is more complex because a perfect (or null) item total score implies the existence of a null category, but the reverse is not necessarily true: a null category may exist without a perfect (or null) item total score (the condition of presence/absence of null categories is routinely checked by some estimation programs, see e.g. LPCM-WIN 1.0, Fischer and Ponocny-Seliger 1998). Wilson and Masters (1993) have developed a procedure for automatically reparameterizing the model to provide JML estimates of a smaller number of item parameters when one or more response categories for an item are null.Moreover, in the JML case, Wright and Masters (1982, p. 61) recommended removing each person with a perfect (or null) total score. Bertoli-Barsotti (2005), in the complete case, gives a n.s. condition for the existence and uniqueness of the JML estimate, in the form of a method simply verifiable on the basis of the total scores of $\boldsymbol{x}$. The result is based on the redefinition of the concept of ill-conditioned matrix introduced by Fischer (1981). Roughly speaking, the author defines a dataset as *ill-conditioned* if there exists at least a partition of the respondent set $\mathcal{S}$ into (at least) two non-empty subsets, $\mathcal{S}_1$ and $\mathcal{S}_2$, such that if a subject belongs to $\mathcal{S}_2$, his response score on I_i is not better than the response score on I_i, $i = 1, \ldots, k$, of any other subject in $\mathcal{S}_1$. If no such subsets exist, the data are said to be *well-conditioned*. The latter property is necessary for the existence of a finite, unique normalized JML solution for the PCM; it is also sufficient, provided that the items with at least a null category have been removed. Only for complete designs, Bertoli-Barsotti (2005, p. 523) also presents a useful method for establishing ill-conditioning.

The definition of an ill-conditioned matrix given by Bertoli-Barsotti is more general than Fischer's. For this reason and for the sake of clearness, the expression "JML-ill-conditioned" will be referred to the definition given by Bertoli-Barsotti (2005, p. 522). In fact, it may be noted that if $m = 1$ (dichotomous case) the JML conditions of Bertoli-Barsotti (2005, Theorem 3, p. 523) are equivalent to those given by Fischer (1981, Theorem 1, p. 64), except for trivial details that, in Fischer's formulation, the partition leading to an ill-conditioned score matrix is defined by the set of items (instead of persons); accordingly, Fischer's additional condition $0 < r_v < k$, $v = 1, \ldots, n$, is formulated with respect to *persons*, while Bertoli-Barsotti's additional condition is referred to the *category/item* total scores.

For the CML approach, the conditions are in principle known either in the original Barndorff-Nielsen form (Barndorff-Nielsen 1978; cf. Andersen 1991, p. 45), or in the form given by Jacobsen (1989). It is not known yet, however, whether these conditions can be brought into such a form that the existence of solutions can be verified from the item totals as easily in the polytomous case as in the dichotomous case (see Andersen 1995a, p. 278; van der Linden and Hambleton 1997b, p. 23; Andersen 1997, p. 72; Fischer and Parzer 1991, p. 650; Fischer and Ponocny 1994, p. 182).

5. False critical points detection

Given that the n.s. conditions for the existence of a finite solution for the PCM, described in Bertoli-Barsotti (2005), are just referred to the JML procedure, and given that in this work the attention is directed to the same problem but in the CML approach, it may be interesting to analyze the effect of JML-ill-conditioned configurations in this new context. To do this, a JML-ill-conditioned complete artificial score matrix $\boldsymbol{s}$ with $n = 8$ persons and $k = 4$ items, each of them with $m + 1 = 3$ categories, is considered in Table 1.

		I_1	I_2	I_3	I_4	r_v
$\mathcal{S}_1$	S_1	2	1	2	2	7
	S_2	2	2	1	1	6
	S_3	2	2	1	0	5
	S_4	2	1	0	1	4
	S_5	2	1	0	0	3
$\mathcal{S}_2$	S_6	2	0	0	0	2
	S_7	1	1	0	0	2
	S_8	0	1	0	0	1
	c_i	13	9	4	4	

Table 1: *JML-ill-conditioned score matrix. The horizontal dotted line delimits one of the possible partitions of $\mathcal{S}$ into $\mathcal{S}_1$ and $\mathcal{S}_2$.*

From the JML method point of view, a finite solution for the item parameters does not exist. From the CML method point of view, instead, this matrix could be "estimable" (for what is so far known in literature) since it does not contain null categories.

The CML estimates of the item parameters, obtained by means of the R package **eRm**, are summarized in Table 2. The β's estimates are obtained with the constraint (5). δ's, α's and τ's estimates are obtained thanks to their relationships with the β's.

$\mathcal{I}$	$\widehat{\beta}_{i1}$	$\widehat{\beta}_{i2}$	$\widehat{\delta}_{i1}$	$\widehat{\delta}_{i2}$	$\widehat{\alpha}_i$	$\widehat{\tau}_{i1}$	$\widehat{\tau}_{i2}$
I_1	-8,2599	-41.9281	-8,2599	-33.6682	-20.9641	12.7042	-12.7042
I_2	-33.6682	-16.6808	-33.6682	16.9873	-8.3404	-25.3277	25.3277
I_3	16.2940	33.9745	16.2940	17.6804	16.9872	-0.6932	0.6932
I_4	16.2940	33.9745	16.2940	17.6804	16.9872	-0.6932	0.6932

Table 2: *CML estimates of the item parameters related to score matrix in Table 1, obtained with the R package* **eRm**.

It is straightforward to note that the estimates in Table 2 do not appear to be "meaningful" if one considers that the results are expressed in logits. This inkling of concern is confirmed by an exhaustive numerical analysis of the conditional log-likelihood function accomplished in Mathematica environment (Wolfram Research 2007). Indeed, increasing the numerical accuracy of the maximization algorithm, also the estimates of the parameters (in absolute value) increase. Some examples of the obtained results, in relation to several setting of the accuracy of the numeric maximization algorithm, are given in Table 3.

Working precision	$\mathcal{I}$	$\widehat{\beta}_{i1}$	$\widehat{\beta}_{i2}$	$\widehat{\delta}_{i1}$	$\widehat{\delta}_{i2}$	$\widehat{\alpha}_i$	$\widehat{\tau}_{i1}$	$\widehat{\tau}_{i2}$
40	I_1	-10.8717	-52.3071	-10.8717	-41.4354	-26.1536	15.2819	-15.2819
	I_2	-41.4354	-20.3731	-41.4354	21.0623	-10.1866	-31.2488	31.2488
	I_3	20.3691	42.1245	20.3691	21.7554	21.0623	-0.6932	0.6932
	I_4	20.3691	42.1245	20.3691	21.7554	21.0623	-0.6932	0.6932
50	I_1	-13.2050	-63.9738	-13.2050	-50.7687	-31.9869	18.7019	-18.7019
	I_2	-50.7687	-25.0398	-50.7687	25.7289	-12.5199	-38.2488	38.2488
	I_3	25.0358	51.4579	25.0358	26.4221	25.7290	-0.6932	0.6932
	I_4	25.0358	51.4579	25.0358	26.4221	25.7290	-0.6932	0.6932
60	I_1	-16.2050	-78.9738	-16.2050	-62.7687	-39.4869	23.2819	-23.2819
	I_2	-62.7687	-31.0398	-62.7687	31.7289	-15.5199	-47.2488	47.2488
	I_3	31.0358	63.4579	31.0358	32.4221	31.7290	-0.6932	0.6932
	I_4	31.0358	63.4579	31.0358	32.4221	31.7290	-0.6932	0.6932

Table 3: *CML estimates of the item parameters related to score matrix in Table 1. The estimates are obtained in Mathematica environment with 3 setting of the working precision (digits of precision maintained in internal computations).*

As a matter of fact, from a simple analytic study of the maximization function, it is possible to realize that a finite maximum point does not exist. This problem is not revealed by the **eRm** package because the Newton-Raphson approach fails in the convergence and, consequently, it detects a false critical point. From a geometrical point of view, this could be attributed to the fact that the curve to be maximized presents a kind of "plateau". Consequently, small variations (in this "flat" zone) of the maximization function value correspond to large variations in the position of the point.

With reference to the same dimension, in Table 4, further 4 examples of JML-ill-conditioned score matrices are added. The CML estimates performed on these matrices present the

	I_1	I_2	I_3	I_4	r_v
S_1	2	2	1	2	7
S_2	2	2	2	0	6
S_3	2	2	1	1	6
S_4	2	2	1	1	6
S_5	2	2	0	0	4
S_6	2	1	0	0	3
S_7	0	1	0	0	1
S_8	1	0	0	0	1
c_i	13	12	5	4	

	I_1	I_2	I_3	I_4	r_v
S_1	2	2	2	1	7
S_2	2	2	1	1	6
S_3	2	1	1	2	6
S_4	2	1	1	1	5
S_5	1	1	0	0	2
S_6	2	0	0	0	2
S_7	2	0	0	0	2
S_8	0	1	0	0	1
c_i	13	8	5	5	

	I_1	I_2	I_3	I_4	r_v
S_1	2	2	2	1	7
S_2	2	2	1	2	7
S_3	2	2	1	1	6
S_4	2	1	2	0	5
S_5	1	1	1	0	3
S_6	1	1	1	0	3
S_7	1	1	0	0	2
S_8	0	0	1	0	1
c_i	11	10	9	4	

	I_1	I_2	I_3	I_4	r_v
S_1	2	2	2	1	7
S_2	2	2	1	2	7
S_3	2	2	2	1	7
S_4	2	2	1	0	5
S_5	2	2	0	0	4
S_6	2	0	0	0	2
S_7	1	0	0	0	1
S_8	0	1	0	0	1
c_i	13	11	6	4	

Table 4: *JML-ill-conditioned score matrices. The horizontal dotted line delimits one of the possible partitions of $\mathcal{S}$ into $\mathcal{S}_1$ and $\mathcal{S}_2$.*

analogous problem compared with the ones obtained for score matrix in Table 1.

To sum up, the item parameters related to datasets in Table 1 and in Table 4 turn out to be inestimable with both JML and CML method even if, in the CML case, the **eRm** package provides false estimates (a false critical point). All matrices so far considered can be seen as CML-ill-conditioned (that is to say, the "ill-conditioned expression" characterizes a "defect" in the dataset). For the sake of uniformity, from now on, when a score matrix presents the problem of a false critical point with the CML procedure, it will be defined as CML-ill-conditioned.

6. A systematic analysis on fixed small-dimensional datasets

In the light of the issue advanced in the previous section, and following the Linacre (2004b, p. 60) setting, an in-depth analysis on fixed small-dimensional datasets with $m + 1 = 3$ categories has been performed. To tell the truth, these kind of matrices have not practical interest but they turn out interesting from a theoretical point of view.

It is convenient to look at a fixed matrix dimension to introduce the analysis. Consider, for example, the 4×3 score matrices with response categories 0, 1 or 2. An automatized procedure has been implemented in R environment; initially, it generates all possible $3^{12} = 531441$ matrices of this kind. Let A be the set of all matrices with at least a null category for at least an item in $\mathcal{I} = \{I_1, I_2, I_3\}$; these matrices are 484785. Let B be the set of all matrices with at least an extreme person's raw score of 0 or $mk = 6$; these matrices are 140816. As said before, these persons do not affect the conditioning. In the first step the procedure removes from the analysis the set $C = A \cup B$; this set has dimension 497037 (a data matrix could have both null category and extreme person's raw scores, i.e., $A \cap B \neq \emptyset$). In reality, one could consider the matrices belonging to B but, from a practical point of view, their "real" row dimension should be equal to the number of rows without extreme total scores whereas the analysis is restricted on 4×3 score matrices.

The remaining 34404 matrices have been ordered according to a nonincreasing sorting with respect to both row and column totals. After this step, only 1333 matrices turn out to be different among them ("different" means in at least one entry of the matrix). It is to be noted that these matrices are not yet "really" different; they are different up to row/column permutations. To realize the issue, it is useful to consider the 2 score matrices in Table 5.

(a) Score matrix

	I_1	I_2	I_3	r_v
S_1	2	2	0	4
S_2	1	1	1	3
S_3	2	0	1	3
S_4	0	1	2	3
c_i	5	4	4	

(b) Score matrix

	I_1	I_2	I_3	r_v
S_1	2	2	0	4
S_2	0	1	2	3
S_3	2	0	1	3
S_4	1	1	1	3
c_i	5	4	4	

Table 5: *Example of two* 4×3 *score matrices different up to permutations.*

With a simple permutation between the second and the fourth row of the matrix in Table 5(b), it is easy to note that one obtains the matrix in Table 5(a).

According to this example, it is possible to assert that the "really" different matrices are only 273; each of them can be considered as representative of an equivalence class in which the equality is meant entry by entry up to row/column permutations. These score matrices are shown in Table 6.

After an analysis with the **eRm** package, the 68 matrices shown in roman bold can be considered as CML-ill-conditioned because they have manifested the problem of false critical point detection (it is to be noted that an in-depth analysis in Mathematica could produce other CML-ill-conditioned score matrices). Therefore, the presence of anomalous configurations is confirmed. Among the 68 matrices CML-ill-conditioned, it is possible to identify only 7 matrices (shown in italics) that are also JML-ill-conditioned; these matrices are the only of this kind. Consequently, there exist at least 61 JML-well-conditioned matrices (the overwhelming majority) that result CML-ill-conditioned.

1 1 1 2 0 0 0 2 0 0 0 2	2 1 0 1 0 1 0 2 0 0 0 2	1 2 0 1 0 2 0 1 1 2 0 0	**2 1 0** **1 0 2** **0 2 0** **0 0 1**	1 2 0 0 1 2 2 0 0 1 0 1	**2 1 0** **1 0 2** **0 2 1** **1 0 0**	1 2 0 2 0 1 1 1 0 0 0 2	2 1 0 1 0 2 0 2 1 1 1 0	1 2 0 1 1 1 2 0 0 0 0 2	2 1 0 1 1 1 1 0 2 0 2 0	2 1 0 1 1 1 1 0 2 0 2 1	**2 1 0** **1 2 0** **0 0 2** **0 0 1**	2 1 0 1 2 0 1 0 1 0 0 2	**2 1 0** **1 2 0** **1 0 2** **0 0 1**
2 1 0 1 2 0 1 0 2 0 1 1	2 1 0 1 2 0 1 1 1 0 0 2	**2 1 0** **1 0 2** **0 2 1** **2 0 0**	**2 1 0** **2 0 1** **1 2 0** **0 0 2**	**2 1 0** **2 0 1** **1 2 0** **0 1 2**	**2 1 0** **2 1 0** **1 0 2** **0 2 1**	2 1 1 0 2 0 0 0 2 1 0 0	1 2 1 1 0 2 2 0 0 0 1 0	1 1 2 2 1 0 1 2 0 0 0 1	**1 2 1** **1 1 0** **2 0 0** **0 0 2**	1 1 2 2 1 0 1 0 1 0 2 0	1 1 2 1 1 1 2 0 0 0 2 0	1 2 1 0 1 2 2 0 0 1 0 0	1 2 1 0 1 2 2 0 0 2 0 0
1 2 1 2 1 0 0 0 2 1 0 0	1 1 2 2 0 1 1 1 0 0 2 0	1 1 2 2 0 1 1 2 0 0 1 0	1 1 2 0 2 1 2 1 0 2 0 0	1 2 1 2 1 0 2 0 0 0 0 2	1 1 2 2 0 1 2 1 0 0 2 0	2 1 1 1 2 0 0 0 2 1 0 0	2 1 1 1 2 0 0 1 2 1 0 0	2 1 1 1 0 2 1 1 0 0 2 0	2 1 1 1 2 0 0 1 2 1 0 1	2 1 1 1 2 0 1 0 2 0 1 0	2 1 1 1 2 0 1 0 2 0 1 1	1 2 1 2 1 0 0 1 2 2 0 0	1 2 1 2 1 0 0 1 2 2 0 1
2 1 1 1 2 0 1 1 0 0 0 2	2 1 1 1 2 0 1 1 1 0 0 2	1 2 1 2 1 0 2 1 0 0 0 2	2 1 1 1 1 2 0 2 0 1 0 0	2 1 1 1 2 1 1 0 2 0 1 0	2 1 1 1 2 1 1 1 0 0 0 2	2 1 1 1 2 0 2 0 0 0 0 2	2 1 1 1 2 0 0 1 2 2 0 0	**2 1 1** **1 1 2** **2 0 0** **0 2 0**	2 1 1 1 2 1 0 1 2 2 0 0	2 1 1 2 1 0 1 0 2 0 2 0	2 1 1 2 1 0 1 0 2 0 2 1	2 1 1 2 1 0 1 2 0 0 0 2	2 1 1 1 2 1 2 1 0 0 0 2
2 1 1 2 1 1 1 2 0 0 0 2	2 2 0 1 0 1 0 0 2 0 1 0	**0 2 2** **1 1 0** **1 0 1** **2 0 0**	**2 2 0** **1 0 2** **0 1 0** **0 0 1**	0 2 2 2 1 0 1 0 1 1 0 0	0 2 2 2 1 0 1 0 1 2 0 0	2 2 0 0 1 2 1 0 1 1 0 0	2 0 2 0 2 1 1 1 0 1 1 0	**0 2 2** **2 1 0** **2 0 1** **1 0 0**	2 2 0 1 0 2 0 1 1 1 0 0	2 2 0 1 0 2 1 0 1 0 1 0	2 0 2 1 2 0 1 1 0 0 1 1	0 2 2 2 1 0 2 0 1 1 1 0	0 2 2 2 1 0 2 1 0 1 0 1
2 2 0 1 1 0 0 0 2 0 0 1	**2 2 0** **1 1 0** **1 0 1** **0 0 2**	2 2 0 1 0 2 1 1 0 0 0 1	2 2 0 1 0 2 1 1 0 0 1 1	0 2 2 1 1 1 2 0 0 1 0 0	0 2 2 1 1 1 2 0 0 2 0 0	2 2 0 1 1 1 0 0 2 1 0 0	2 0 2 1 1 1 1 1 0 0 2 0	2 0 2 1 1 1 1 2 0 0 1 0	0 2 2 1 1 1 2 1 0 2 0 0	2 2 0 1 1 1 1 0 2 0 1 0	2 2 0 1 1 1 1 0 2 0 1 1	0 2 2 1 1 1 2 1 0 2 0 1	2 2 0 1 1 1 1 1 0 0 0 2
2 2 0 1 1 1 1 1 1 0 0 2	0 2 2 2 1 1 1 0 0 1 0 0	0 2 2 2 1 1 2 0 0 1 0 0	2 2 0 1 1 2 1 0 0 0 0 1	2 0 2 1 2 1 1 1 0 0 1 0	**0 2 2** **2 1 1** **1 1 0** **2 0 0**	2 2 0 1 1 2 1 0 1 0 1 0	2 2 0 1 1 2 1 0 1 0 1 1	0 2 2 2 1 1 2 1 0 1 0 0	0 2 2 2 1 1 2 1 0 1 0 1	2 2 0 1 1 2 1 1 0 0 0 1	2 2 0 1 1 2 1 1 1 0 0 1	0 2 2 2 1 1 1 1 1 2 0 0	0 2 2 2 1 1 2 1 1 1 0 0
2 2 0 0 1 2 2 0 0 1 0 1	2 2 0 1 0 2 2 0 0 0 1 1	2 2 0 1 1 1 2 0 0 0 0 2	2 2 0 1 1 1 0 1 2 2 0 0	2 2 0 1 1 2 2 0 0 0 0 1	**2 2 0** **1 1 2** **2 0 0** **0 1 1**	**2 2 0** **2 0 1** **0 1 2** **1 0 0**	2 0 2 2 1 0 0 2 1 1 1 0	**2 2 0** **2 0 1** **1 0 2** **0 1 0**	2 0 2 2 1 0 1 2 0 0 1 1	**2 0 2** **2 1 0** **1 2 0** **0 2 1**	2 2 0 2 0 1 1 1 0 0 0 2	2 2 0 2 0 1 0 1 2 1 1 0	2 0 2 2 1 0 1 1 1 0 2 0
2 2 0 2 0 1 1 1 1 0 1 2	2 0 2 1 2 1 2 1 0 0 1 0	2 0 2 1 2 1 2 1 0 0 2 0	2 2 0 1 1 2 2 0 1 0 1 0	2 2 0 1 1 2 2 0 1 0 1 1	2 0 2 1 2 1 2 1 0 0 2 1	2 2 0 1 1 2 2 0 1 0 2 0	2 2 0 1 1 2 2 0 1 0 2 1	2 2 0 2 0 2 0 1 1 1 0 0	2 2 0 2 0 2 1 1 0 0 0 1	**2 2 0** **2 0 2** **1 1 0** **0 1 1**	2 2 0 2 0 2 0 2 1 1 1 0	2 2 0 2 0 2 1 1 1 0 1 0	2 2 0 2 0 2 1 1 1 0 1 1
2 2 0 2 0 2 1 1 1 0 2 0	2 2 0 2 0 2 1 1 1 0 2 1	2 2 0 2 0 2 0 2 2 1 1 1	2 2 0 2 1 0 1 0 1 0 0 2	2 2 0 2 1 0 0 1 2 1 0 1	2 2 0 0 2 2 2 1 0 1 0 1	**2 2 0** **2 1 0** **1 0 2** **0 0 1**	2 2 0 2 1 0 1 0 2 0 1 1	2 2 0 2 1 0 1 0 2 0 2 1	2 2 0 2 1 0 1 1 1 0 0 2	2 2 0 1 1 2 2 1 0 0 0 1	2 2 0 2 1 1 0 0 2 1 0 0	2 2 0 2 1 1 0 1 2 1 0 0	2 2 0 2 1 1 0 2 2 1 0 0
2 0 2 **2 1 1** **1 1 0** **0 2 0**	2 2 0 2 1 1 0 1 2 1 0 1	**2 2 0** **2 1 1** **0 2 2** **1 0 1**	2 0 2 2 1 1 1 2 0 0 1 0	2 0 2 2 1 1 1 2 0 0 2 0	2 2 0 2 1 1 1 0 2 0 1 0	2 2 0 2 1 1 1 0 2 0 1 1	2 0 2 2 1 1 1 2 0 0 2 1	2 2 0 2 1 1 1 0 2 0 2 0	2 2 0 2 1 1 1 0 2 0 2 1	2 2 0 2 1 1 0 2 2 1 0 2	**2 2 0** **2 1 1** **1 1 0** **0 0 2**	2 2 0 2 1 1 1 1 1 0 0 2	2 2 0 2 1 1 1 1 2 0 0 1
2 0 2 **2 1 1** **1 2 1** **0 2 0**	2 2 0 2 1 1 1 2 0 0 0 2	**2 2 0** **2 1 1** **1 2 1** **0 0 2**	2 2 0 2 2 0 1 0 2 0 1 1	2 2 0 2 2 0 1 1 1 0 0 2	**2 2 0** **2 2 0** **1 1 2** **0 0 1**	**2 2 1** **0 0 2** **1 0 0** **0 1 0**	1 2 2 1 1 0 2 0 0 0 0 1	**1 2 2** **2 1 0** **0 0 1** **1 0 0**	**1 2 2** **2 1 0** **2 0 0** **0 0 1**	**2 2 1** **0 1 2** **1 0 0** **1 0 0**	2 1 2 0 2 1 1 1 0 1 0 0	**2 2 1** **1 0 2** **1 0 0** **0 1 0**	2 1 2 1 2 0 0 1 1 1 0 0
1 2 2 **2 1 0** **2 0 1** **0 1 0**	2 1 2 1 2 0 1 1 0 0 0 1	**1 2 2** **2 1 0** **2 1 0** **0 0 1**	1 2 2 0 1 1 2 0 0 1 0 0	1 2 2 0 1 1 2 0 0 2 0 0	2 2 1 1 1 0 0 0 2 1 0 0	2 1 2 1 0 1 1 1 0 0 2 0	2 1 2 1 2 0 1 0 1 0 1 0	1 2 2 2 1 0 0 1 1 2 0 0	2 2 1 1 0 2 1 1 0 0 1 0	2 2 1 1 0 2 1 1 0 0 1 1	1 2 2 2 1 0 2 0 1 0 1 1	2 2 1 1 1 0 1 1 0 0 0 2	2 1 2 1 1 1 0 2 0 1 0 0
2 2 1 1 1 1 1 0 2 0 1 0	2 2 1 1 1 1 1 1 0 0 0 2	*2 1 2* *1 2 1* *1 0 0* *0 1 0*	1 2 2 2 1 1 2 0 0 0 1 0	*2 2 1* *1 1 2* *1 0 0* *0 1 0*	*2 2 1* *1 1 2* *1 0 1* *0 1 0*	1 2 2 2 1 1 2 1 0 0 0 1	*2 2 1* *1 1 2* *1 1 0* *0 0 1*	1 2 2 2 1 1 0 1 1 2 0 0	**2 2 1** **0 1 2** **2 0 0** **1 0 0**	2 1 2 0 2 1 2 0 0 1 1 0	**2 2 1** **1 0 2** **2 0 0** **0 1 0**	2 1 2 1 2 0 2 0 0 0 1 1	**2 1 2** **1 2 0** **0 2 1** **2 0 0**
2 2 1 2 0 0 1 1 0 0 0 2	2 2 1 0 1 2 2 0 0 1 1 0	2 1 2 1 1 1 2 0 0 0 2 0	2 2 1 1 1 1 0 1 2 2 0 0	2 1 2 1 2 1 2 0 0 0 1 0	2 1 2 1 2 1 2 0 0 0 2 0	2 2 1 1 1 2 2 0 0 0 1 0	2 2 1 1 1 2 2 0 0 0 1 1	2 1 2 1 2 1 0 2 1 2 0 0	2 2 1 1 1 2 2 0 0 0 2 0	**2 1 2** **2 1 0** **0 2 1** **1 0 0**	**2 1 2** **2 1 0** **1 2 0** **0 0 1**	2 1 2 2 1 0 1 0 1 0 2 0	2 2 1 2 0 1 0 1 2 1 1 0
2 2 1 1 1 2 2 0 1 0 1 0	2 2 1 1 1 2 2 0 1 0 2 0	**2 2 1** **2 0 2** **1 0 0** **0 1 0**	2 1 2 2 2 0 0 1 1 1 0 0	**2 1 2** **2 2 0** **0 2 1** **1 0 0**	2 1 2 2 2 0 1 1 0 0 0 1	**2 1 2** **2 2 0** **1 2 0** **0 0 1**	2 1 2 2 2 0 1 0 1 0 1 0	2 1 2 2 2 0 1 0 1 0 2 0	2 2 1 2 0 2 1 1 0 0 1 0	2 2 1 2 0 2 1 1 0 0 1 1	2 1 2 2 2 0 0 2 1 1 0 1	2 2 1 2 0 2 1 1 0 0 2 0	2 2 1 2 0 2 0 2 1 1 1 0
2 2 1 2 0 2 0 2 2 1 1 0	2 2 1 2 0 2 1 1 1 0 1 0	2 2 1 2 0 2 1 1 1 0 2 0	2 1 2 2 2 0 1 2 1 0 0 1	2 2 1 2 0 2 1 1 2 0 2 0	**2 2 1** **2 1 0** **0 0 2** **1 0 0**	**2 2 1** **2 1 0** **0 1 2** **1 0 0**	**2 2 1** **0 2 2** **2 1 0** **1 0 0**	2 1 2 2 0 1 1 1 0 0 2 0	2 2 1 2 1 0 0 1 2 1 0 1	2 2 1 0 2 2 2 1 0 1 0 1	**2 1 2** **2 0 1** **1 2 0** **0 1 0**	**2 1 2** **2 0 1** **1 2 0** **0 2 0**	**2 2 1** **2 1 0** **1 0 2** **0 1 0**
2 2 1 2 1 0 1 0 2 0 1 1	**2 1 2** **2 0 1** **1 2 0** **0 2 1**	**2 2 1** **2 1 0** **1 0 2** **0 2 0**	**2 2 1** **2 1 0** **1 0 2** **0 2 1**	**2 2 1** **0 2 2** **2 1 0** **1 0 2**	2 2 1 2 1 0 1 1 0 0 0 2	2 2 1 2 1 0 1 1 1 0 0 2	2 2 1 1 1 2 2 1 0 0 0 1	2 1 2 1 2 1 2 0 1 0 2 0	**2 2 1** **2 1 0** **1 2 0** **0 0 2**	2 1 2 2 1 1 0 2 0 1 0 0	2 2 1 2 1 1 0 1 2 1 0 0	2 2 1 2 1 1 0 2 2 1 0 0	2 2 1 2 1 1 1 0 2 0 1 0
2 2 1 2 1 1 1 0 2 0 2 0	2 2 1 2 1 1 1 1 0 0 0 2	2 2 1 2 1 1 1 2 0 0 0 2	*2 2 1* *2 1 2* *1 0 0* *0 1 0*	*2 2 1* *2 1 2* *0 1 1* *1 0 0*	**2 2 1** **2 1 2** **0 2 0** **1 0 0**	**2 2 1** **2 1 2** **0 2 1** **1 0 0**	**2 2 1** **2 1 2** **0 2 2** **1 0 0**	*2 2 1* *2 1 2* *1 1 0* *0 0 1*	2 1 2 2 2 1 1 0 1 0 2 0	**2 2 1** **2 1 2** **1 2 0** **0 0 1**	**2 2 1** **2 1 2** **1 2 0** **0 0 2**	**2 2 1** **2 2 0** **1 0 2** **0 1 0**	2 2 1 2 2 0 1 0 2 0 1 1
2 2 1 **2 2 0** **1 0 2** **0 1 2**	2 2 1 2 2 0 1 1 0 0 0 2	2 2 1 2 2 0 1 1 1 0 0 2	2 2 1 2 2 0 1 1 2 0 0 1	2 2 1 2 2 0 1 1 2 0 0 2	**2 2 1** **2 2 1** **1 0 2** **0 1 0**	2 2 1 2 2 1 1 1 0 0 0 2							

Table 6: *All* 273 *"substantially different" data matrices, each comprising* 3 *polytomous items (columns), with* 3 *response categories, administered to* 4 *persons (rows).*

From a similar analysis implemented in R on 3×3, 4×2, 5×2 and 6×2 matrices with $m+1=3$ categories, a false critical point, systematically enough, whenever a JML-ill-conditioned score matrix has been detected. Consequently, the JML-ill-conditioning should be seen (whether by chance or because of problematic issues) as a sort of sufficient condition for the CML-ill-conditioning. Moreover, from the same analysis, the existence of several CML-anomalous configurations (leading to false critical points, i.e., CML-ill-conditioned) that are different from JML-ill-conditioned ones, stands out.

7. Concluding remarks

The existence of the CML estimates for the PCM item parameters has been here investigated by giving several examples of score matrices that, with the R package **eRm**, lead to detect false critical points in the log-likelihood function maximization. A systematic analysis on small-dimensional score matrices, with 3 response categories, has highlighted that JML-ill-conditioned score matrices are a little subset of CML-ill-conditioned ones; JML-ill-conditioned score matrices, instead, are CML-ill-conditioned as well, which means that the JML-ill-conditioning should be a sufficient condition for the CML-ill-conditioning.

The R package **eRm** – as well as all other packages or softwares performing a CML estimation procedure – should detect these "anomalous" datasets (by means of an initial data check) or, at least, the package should check the convergence of the maximization algorithm results. In the light of the first suggestion, the still open question is to analyze the structure characterizing these "CML-ill-conditioned" configurations from a theoretical point of view. According to this, it may be useful to systematically consider matrices of small dimension and by means of a few number of categories to emphasize the need for specific constraints.

Realizing that further questions can still be raised, the research needs to be carried on. How, for example, these false critical points influence person parameter estimates, is an important topic. Also the case of incomplete datasets needs to be furtherly analyzed; according to the analogy with the dichotomous case (cf. Fischer 1981) and owing to the recent results of Bertoli-Barsotti and Bacci (2007a,b), the presence of missing data is expected to be a source of a greater incidence of false critical points. Other important issues are the study of the non-existence of the CML estimates in the case of response alternatives that are free to vary in number from item to item (i.e. m is not constant) and in the case of models defined by Mair and Hatzinger (2007a,b) as extended Rasch models (eRm).

Acknowledgments

Many thanks go to L. Bertoli-Barsotti for his useful and valuable discussions during the writing of this article.

References

Adams RA, Khoo ST (1991). "QUEST [Computer Program]." *Hawthorn, Australia: Australian Council for Educational Research.*

Andersen EB (1970). "Asymptotic Properties of Conditional Maximum-likelihood Estimators." *Journal of the Royal Statistical Society, Series B: Methodological*, **32**, 283–301.

Andersen EB (1972). "The Numerical Solution of a Set of Conditional Estimation Equations." *Journal of the Royal Statistical Society, Series B: Methodological*, **34**, 42–54.

Andersen EB (1973). "Conditional Inference for Multiple-Choice Questionnaires." *British Journal of Mathematical and Statistical Psychology*, **26**(1), 31–44.

Andersen EB (1980). *Discrete Statistical Models with Social Science Applications.* Elsevier/North-Holland [Elsevier Science Publishing Co., New York; North-Holland Publishing Co., Amsterdam].

Andersen EB (1983). "A General Latent Structure Model for Contingency Table Data." *Principals of Modern Psychological Measurement. Lawrence Erlbaum: NJ*, pp. 117–138.

Andersen EB (1991). *The Statistical Analysis of Categorical Data (Second Edition).* Springer-Verlag Inc.

Andersen EB (1995a). "Polytomous Rasch Models and their Estimation." *Rasch Models: Foundations, Recent Developments, and Applications.*

Andersen EB (1995b). "Residual Analysis in the Polytomous Rasch Model." *Psychometrika*, **60**, 375–393.

Andersen EB (1997). "The Rating Scale Model." *Handbook of Modern Item Response Theory.*

Anderson CJ, Li Z, Vermunt JK (2007). "Estimation of Models in a Rasch Family for Polytomous Items and Multiple Latent Variables." *Journal of Statistical Software.*

Andrich D (1978). "A Rating Formulation for Ordered Response Categories." *Psychometrika*, **43**, 561–574.

Andrich D (1982). "An Extension of the Rasch Model for Ratings Providing Both Location and Dispersion Parameters." *Psychometrika*, **47**, 105–113.

Baker FB, Kim S (2004). *Item Response Theory: Parameter Estimation Techniques.* Marcel Dekker Inc.

Barndorff-Nielsen O (1978). *Information and Exponential Families in Statistical Theory.* John Wiley & Sons.

Bejar II (1977). "An Application of the Continuous Response Level Model to Personality Measurement." *Applied Psychological Measurement*, **1**(4), 509.

Bertoli-Barsotti L (2002). "On a Condition for the Existence of the Maximum Likelihood Estimate for Concave log-likelihood functions." *Studi in onore di Angelo Zanella*, pp. 25–42.

Bertoli-Barsotti L (2005). "On the Existence and Uniqueness of JML Estimates for the Partial Credit Model." *Psychometrika*, **70**(3), 517–531.

Bertoli-Barsotti L, Bacci S (2007a). "Identifiability of the Partial Credit Model with Missing Data." In "Convegno: Advanced in Rasch Modelling," Vannes (Fra), 4–8 giugno.

Bertoli-Barsotti L, Bacci S (2007b). "Inferenza in Analisi di Rasch nel caso di Dati Mancanti." In "Atti della Riunione Scientifica: Valutazione e Customer Satisfaction per la Qualità dei Servizi," pp. 25–28. Roma Tor Vergata, 12–13 aprile.

Bock RD, Aitkin M (1981). "Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM Algorithm (Corr: V47 P369)." *Psychometrika*, **46**, 443–459.

Cohen J (1983). "The Cost of Dichotomization." *Applied Psychological Measurement*, **7**(3), 249.

Cox EPC (1980). "The Optimal Number of Response Alternatives for a Scale: A Review." *Journal of Marketing Research*, **17**(4), 407–422.

Fischer GH (1974). *Einführung in die Theorie psychologischer Tests: Grundlagen und Anwendungen*. H. Huber.

Fischer GH (1981). "On the Existence and Uniqueness of Maximum-Likelihood Estimates in the Rasch Model." *Psychometrika*, **46**(1), 59–77.

Fischer GH, Molenaar IW (1995). *Rasch Models: Foundations, Recent Developments, and Applications*. Springer-Verlag Inc.

Fischer GH, Parzer P (1991). "An Extension of the Rating Scale Model with an Application to the Measurement of Change." *Psychometrika*, **56**, 637–651.

Fischer GH, Ponocny I (1994). "An Extension of the Partial Credit Model with an Application to the Measurement of Change." *Psychometrika*, **59**, 177–192.

Fischer GH, Ponocny I (1995). "Extended Rating Scale and Partial Credit Models for Assessing Change." *Rasch Models. Foundations, Recent Developments, and Applications*, pp. 353–370.

Fischer GH, Ponocny-Seliger E (1998). *Structural Rasch Modeling: Handbook of the Usage of LPCM-WIN 1.0*. ProGAMMA and sciencePlus.

Fisher Jr WP (1992). "Objectivity in Measurement: A Philosophical History of Rasch's Separability Theorem." *Objective Measurement: Theory into Practice*, **1**, 29–60.

Garner M, Engelhard Jr G (2002). "An Eigenvector Method for Estimating Item Parameters of the Dichotomous and Polytomous Rasch Models." *Journal of Applied Measurement*, **3**(2), 107–128.

Ghosh M (1995). "Inconsistent Maximum Likelihood Estimators for the Rasch Model." *Statistics & Probability Letters*, **23**, 165–170.

Gustafsson JE (1980). "Testing and Obtaining Fit of Data to the Rasch Model." *British Journal of Mathematical and Statistical Psychology*, **33**, 205–233.

Haberman SJ (1977). "Maximum Likelihood Estimates in Exponential Response Models." *The Annals of Statistics*, **5**, 815–841.

Hemker BT (2001). "Reversibility Revisited and Other Comparisons of Three Types of Polytomous IRT Models." In A Boomsma, MAJ van Duijn, TAB Snijders (eds.), "Essays on Item Response Theory," pp. 277–296. Springer-Verlag Inc.

Jacobsen M (1989). "Existence and Unicity of MLEs in Discrete Exponential Family Distributions." *Scandinavian Journal of Statistics*, **16**, 335–349.

Kamakura WA, Balasubramanian SK (1989). "Tailored Interviewing: An Application of Item Response Theory for Personality Measurement." *Journal of Personality Assessment*, **53**(3), 502–519.

Linacre JM (2004a). "Estimation Methods for Rasch Measures." *Introduction to Rasch Measurement, JAM Press, Maple Grove (MN)*, pp. 25–47.

Linacre JM (2004b). "Rasch Model Estimation: Further Topics." *Introduction to Rasch Measurement*, pp. 48–72.

Linacre JM, Wright BD (2004). "A User's Guide to WINSTEPS." *Linacre JM. Ministeps Rasch Model Computer Programs. Chicago: Winsteps*, pp. 32–45.

Liou M (1994). "More on the Computation of Higher-order Derivatives of the Elementary Symmetric Functions in the Rasch Model." *Applied Psychological Measurement*, **18**, 53–62.

Mair P, Hatzinger R (2007a). "CML Based Estimation of Extended Rasch Models with the **eRm** Package in R." *Psychology Science*, **49**(1), 26.

Mair P, Hatzinger R (2007b). "Extended Rasch Modeling: The **eRm** Package for the Application of IRT Models in R." *Journal of Statistical Software*, **20**(9), 1–20. ISSN 1548-7660. URL http://www.jstatsoft.org/v20/i09.

Martin-Löf P (1973). "Statistika Modeller. Anteckningarfrån seminarier Lasåret 1969-1970, utarbetade av Rolf Sundberg. Obetydligt ändrat nytryck, Oktober 1973."

Masters GN (1982). "A Rasch Model for Partial Credit Scoring." *Psychometrika*, **47**, 149–174.

Masters GN (1988). "Measurement Models for Ordered Response Categories." In R Langeheine, J Rost (eds.), "Latent Trait and Latent Class Models," pp. 11–29. Plenum Publishing Corporation.

Masters GN, Wright BD (1997). "The Partial Credit Model." *Handbook of Modern Item Response Theory*, pp. 101–121.

Molenaar IW (1995). "Estimation of Item Parameters." *Rasch Models: Foundations, Recent Developments, and Applications.*

Ostini R, Nering ML (2006). *Polytomous Item Response Theory Models*. Sage Publications.

Pfanzagl J (1994). "On Item Parameter Estimation in Certain Latent Trait Models." *Contributions to Mathematical Psychology, Psychometrics, and Methodology*, pp. 249–263.

Rasch G (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. University of Chicago Press.

Rasch G (1977). "On Specific Objectivity: An Attempt at Formalizing the Request for Generality and Validity of Scientific Statements." *The Danish Yearbook of Philosophy*, **14**(4), 58–93.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rost J (1988). "Measuring Attitudes With a Threshold Model Drawing on a Traditional Scaling Concept." *Applied Psychological Measurement*, **12**(4), 397.

Samejima F (1976). "Graded Response Model of the Latent Trait Theory and Tailored Testing." *Proceedings of the First Conference on Computerized Adaptive Testing (US Civil Service Commission, Personnel Research and Development Center, PS-75-6). Washington, DC: US Government Printing Office.*

Samejima F (1979). "A New Family of Models for the Multiple Choice Item (Research Report No. 79–4)." *Knoxville, TN: University of Tennessee, Department of Psychology.*

Sheridan BS, Andrich D, Luo G (2002). "RUMM2020: Rasch Unidimensional Models for Measurement."

Sijtsma K, Hemker BT (2000). "A Taxonomy of IRT Models for Ordering Persons and Items Using Simple Sum Scores." *Journal of Educational and Behavioral Statistics*, **25**(4), 391–415.

van der Linden WJ, Hambleton RK (1997a). *Handbook of Modern Item Response Theory.* Springer-Verlag Inc.

van der Linden WJ, Hambleton RK (1997b). "Item Response Theory: Brief History, Common Models, and Extensions." *Handbook of Modern Item Response Theory.*

Verhelst ND, Glas CAW, De Vries HH (1997). "A Steps Model to Analyze Partial Credit." *Handbook of Modern Item Response Theory*, pp. 123–138.

Wilson M, Masters GN (1993). "The Partial Credit Model and Null Categories." *Psychometrika*, **58**, 87–99.

Wolfram Research (2007). *Mathematica, Version 6.* Wolfram Research Inc.

Wright BD, Masters GN (1982). *Rating Scale Analysis.* MESA Press.

Wright BD, Panchapakesan N (1969). "A Procedure for Sample-Free Item Analysis." *Educational and Psychological Measurement*, **29**(1), 23–48.

Zwinderman AH (1995). "Pairwise Parameter Estimation in Rasch Models." *Applied Psychological Measurement*, **19**, 369–375.

Redazione

Dipartimento di Matematica, Statistica, Informatica ed Applicazioni
Università degli Studi di Bergamo
Via dei Caniana, 2
24127 Bergamo
Tel. 0039-035-2052536
Fax 0039-035-2052549

La Redazione ottempera agli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche

Stampato nel 2008
presso la Cooperativa
Studium Bergomense a r.l.
di Bergamo