Research Report

Monotone Properties of a General Diagnostic Model

Xueli Xu

Research & Development

June 2007
RR-07-25

Monotone Properties of a General Diagnostic Model

Xueli Xu

ETS, Princeton, NJ

June 2007

Abstract

Monotonicity properties of a general diagnostic model (GDM) are considered in this paper. Simple data summaries are identified to inform about the ordered categories of latent traits. The findings are very much in accordance with the statements made about the GPCM (Hemker, Sijtsma, Molenaar, & Junker, 1996, 1997). On the one hand, by fitting a GDM with equal slopes across items, the observed total score $X_+ = \sum_{j:q_{jk}=1} x_j$ demonstrates the monotone likelihood ratio (MLR) property in individual coordinate at the cost of losing model fit. On the other hand, fitting a GDM without slope restriction increases the model fit but by sacrificing MLR property. Trade-offs between these two situations should be considered in practice.

Key words: General diagnostic model, monotone likelihood ratio, stochastic ordering of the manifest variables, stochastic ordering of the latent trait

Acknowledgements

The author would like to thank Matthias von Davier and Sandip Sinharay for their useful suggestions and comments.

1 Introduction

Educational tests are often used to measure the position of students on a latent trait θ. Suppose that a test consists of J items each with $m+1$ ordered categories. The score on item j is denoted as X_j and its realization is denoted as $x_j (x_j = 0, 1, 2, \ldots, m)$. If $m = 1$, items are dichotomous; if $m > 1$ the items are polytomous. Under the very mild conditions of latent trait unidimensionality (UD), local independence (LI) and item response functions that are nondecreasing functions of θ, the most widely used statistic for estimating θ probably is the sum score, denoted as $X_+ = \sum_{j=1}^{J} X_j$. Note that $0 \leq X_+ \leq mJ$. Although there are some critiques concerning the use of X_+ to estimate the latent trait (Samejima, 2001), X_+ is still useful in communications between professionals and the general public, and it is also useful in the interpretation of item response models according to simple, monotone relationships between model components (van der Ark, 2005; Junker & Sijtsma, 2001). For example, Hemker, Sijtsma, Molenaar, and Junker (1996, 1997) considered in detail the monotone likelihood ratio (MLR) in the latent trait, stochastic ordering of the manifest sum score (SOM) by the latent trait and stochastic ordering of the latent trait(SOL) by the manifest sum score. For binary item responses, it was shown by Grayson (1988) and Huynh (1994) that under the conditions of UD, LI, and nondecreasing item response functions, the sum score X_+ demonstrates the MLR in θ. The MLR implies SOM and SOL (Lehmann, 1959, p. 74). For polytomous data, Hemker et al. (1997) demonstrated that the only two models that satisfy the MLR are the partial credit model (PCM; Masters, 1982) and the rating scale model (RSM; Andrich, 1978), that is, the generalization of the Rasch model (Rasch, 1960) to polytomous, ordinal data. SOM and SOL are two weaker monotonicity conditions, and neither of them nor their combinations imply MLR (Junker, 1993). SOM is satisfied by all unidimensional polytomous models, and SOL usually does not hold for them in general. Little is known about the monotonicity properties for multidimensional models for polytomous data. In this paper, all three properties are considered for a multidimensional polytomous model. Additionally, the monotonicity condition proposed by Junker and Sijtsma (2001) is also considered.

The organization of this paper is as follows: a general multidimensional polytomous model is introduced in Section 2. The monotonicity properties are discussed for this model in Section 3. Finally, a brief discussion follows to direct subsequent data analysis activities.

2 General Diagnostic Model

The general diagnostic model (GDM; von Davier, 2005; von Davier & Yamamoto, 2004) is a multidimensional extension of the PCM (Masters, 1982). The PCM has the following response function

$$P_j(X_j = x|\theta) = \frac{\exp[x\theta + \beta_{xj}]}{1 + \sum_{y=1}^{m} \exp[y\theta + \beta_{yj}]}.$$

von Davier and Yamamoto decomposed the unidimensional latent trait into an item-dependent linear combination of K underlying traits $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_K)$,

$$P(X_j = x|\boldsymbol{\alpha}) = \frac{\exp[\beta_{xj} + \sum_{k=1}^{K} x\gamma_{jk}q_{jk}\alpha_k]}{1 + \sum_{y=1}^{m} \exp[\beta_{yj} + \sum_{k=1}^{K} y\gamma_{jk}q_{jk}\alpha_k]},$$

with dichotomous design matrix q_{jk}, $j = 1, \ldots, J$ and $k = 1, \ldots, K$. This model makes it clear that $P(X_j = x|\boldsymbol{\alpha})$ is coordinate-wise monotone in $\boldsymbol{\alpha}$, if $\gamma_{jk} > 0$. If each underlying trait only takes several discrete values specified in advance, the model becomes a cognitive diagnostic model that can be used to identify students' skill mastery status. In some of the following sections, the continuous underlying traits are used to study the monotonicity properties. Because the cognitive diagnostic model is a special case for this general model, the statements from continuous underlying traits remain valid for the cognitive diagnostic model as well.

3 Stochastic Ordering Properties

3.1 Monotone Likelihood Ratio

Let X_+ be the observed total score on the test. The likelihood ratio in the latent traits $\boldsymbol{\alpha}$ is defined as

$$g(D, C; \boldsymbol{\alpha}) = \frac{P(X_+ = D|\boldsymbol{\alpha})}{P(X_+ = C|\boldsymbol{\alpha})},$$

where $0 \leq C < D \leq mJ$. The MLR of X_+ in $\boldsymbol{\alpha}$ holds for the GDM when the likelihood ratio is an increasing function of each coordinate α_k. Little is known about it. A weaker property related to the MLR is that this likelihood ratio is a nondecreasing function of individual latent trait α_k. Analogous to the derivations in Hemker et al. (1996), the MLR of X_+ in α_k holds if and only if

$$\sum_{\mu=1}^{R_D}\sum_{\nu=1}^{R_C}\Big(\sum_{j=1}^{J}\Big[\frac{\pi'_{jx(\mu)}}{\pi_{jx(\mu)}} - \frac{\pi'_{jx(\nu)}}{\pi_{jx(\nu)}}\Big]\prod_{j=1}^{J}\pi_{jx(\mu)}\pi_{jx(\nu)}\Big) \geq 0,$$

where R_D and R_C are the numbers of score vectors that yield $X_+ = D$ and $X_+ = C$ respectively. Here $x(\mu)$ and $x(\nu)$ denote realizations of vectors that yield $X_+ = D$ and $X_+ = C$, respectively.

Denote $\pi_{jx} = P(X_j = x|\boldsymbol{\alpha})$, then $\pi_{jx(\mu)}$ is the probability function of item j in vector $x(\mu)$, and $\pi'_{jx(\mu)}$ is the derivative of this function with respect to α_k. The same explanation applies to $\pi_{jx(\nu)}$ and its prime $\pi'_{jx(\nu)}$. Straightforward algebra shows that:

$$\frac{\partial \ln \pi_{jx}}{\partial \alpha_k} = \frac{\pi'_{jx}}{\pi_{jx}} = x_j \gamma_{jk} q_{jk} - \frac{\sum_{y=1}^{m} \exp[\beta_{yj} + \sum_{k=1}^{K} y\gamma_{jk} q_{jk} \alpha_k] y\gamma_{jk} q_{jk}}{1 + \sum_{y=1}^{m} \exp[\beta_{yj} + \sum_{k=1}^{K} y\gamma_{jk} q_{jk} \alpha_k]}.$$

The summation of this ratio across all J items produces

$$\sum_{j=1}^{J} \frac{\pi'_{jx}}{\pi_{jx}} = \sum_{j=1}^{J} x_j \gamma_{jk} q_{jk} - \sum_{j=1}^{J} \frac{\sum_{y=1}^{m} \exp[\beta_{yj} + \sum_{k=1}^{K} y\gamma_{jk} q_{jk} \alpha_k] y\gamma_{jk} q_{jk}}{1 + \sum_{y=1}^{m} \exp[\beta_{yj} + \sum_{k=1}^{K} y\gamma_{jk} q_{jk} \alpha_k]}.$$

Since the second term of the right side does not depend on data, we have

$$\sum_{j=1}^{J} [\frac{\pi'_{jx(\mu)}}{\pi_{jx(\mu)}} - \frac{\pi'_{jx(\nu)}}{\pi_{jx(\nu)}}] = \sum_{j=1}^{J} \gamma_{jk} q_{jk} (x_{j(\mu)} - x_{j(\nu)}).$$

It is observed that only the items that require the attribute k contribute to this summation. Because the derivative of $g(D, C; \boldsymbol{\alpha})$ with respect to α_k depends on D and C only through the scores of those items who require the attribute k, the total score in the likelihood ratio should be the total score for those items (i.e. $X_+ = \sum_{j:Q_{jk}=1} X_j$). If γ_{jk} has the same value for all items (i.e. $\gamma_{jk} = \gamma_k > 0$), then the MLR of X_+ in α_k holds. This can be easily seen that the equation above becomes

$$\sum_{j=1}^{J} [\frac{\pi'_{jx(\mu)}}{\pi_{jx(\mu)}} - \frac{\pi'_{jx(\nu)}}{\pi_{jx(\nu)}}] = \gamma_k \sum_{j:q_{jk}=1} (x_{j(\mu)} - x_{j(\nu)}) > 0$$

for all $R_D * R_C$ possible combination of $x(\mu)$ and $x(\nu)$. If γ_{jk} varies over items, the GDM does not in general imply the MLR of X_+ in α_k. In practice, the inferences towards each individual latent trait are often of interest, so we will focus on discussions of monotone properties in individual latent trait. Three weaker monotone properties are considered in the following.

3.2 Stochastic Ordering of Manifest Variables

Since the GDM satisfies local independence, monotonicity and low dimensionality, it follows immediately from Lemma 2 of Holland and Rosenbaum (1986) that SOM (Hemker et al., 1997) holds for the GDM. That is, $P(X_+ \geq x_+|\boldsymbol{\alpha})$ is nondecreasing in each coordinate α_k. Note that X_+ is the total score of the test. Direct derivation makes this clear. Analogous to the derivation in Hemker et al. (1997), the derivative of $P(X_+ \geq x_+|\boldsymbol{\alpha})$ with respect to individual α_k can

also be expressed as a sum of positive products where each product consists of one derivative $P^{'}(X_j \geq x|\boldsymbol{\alpha})$ with respect to α_k and $J-1$ probabilities of the form $\pi_{ix}, i \neq j$:

$$P^{'}(X_j \geq x|\boldsymbol{\alpha}) \prod_{i \neq j}^{J} \pi_{ix}.$$

For the GDM model, the derivative $P^{'}(X_j \geq x|\boldsymbol{\alpha}) \geq 0$ needs to be shown. For notational convenience, let $U_x = \exp[\beta_{xj} + \sum_k x\gamma_{jk}q_{jk}\alpha_k]$, and then the cumulative response function for item j in the GDM model can be expressed as

$$P(X_j \geq x|\boldsymbol{\alpha}) = \frac{U_x + \ldots + U_m}{1 + U_1 + U_2 + \ldots + U_m}, \qquad x = 0, 1, \ldots, m.$$

Direct algebra shows that

$$\begin{aligned}
\frac{\partial \ln P(X_j \geq x|\boldsymbol{\alpha})}{\partial \alpha_k} &= \frac{P^{'}(X_j \geq x|\boldsymbol{\alpha})}{P(X_j \geq x|\boldsymbol{\alpha})} \\
&= \frac{U_x^{'} + U_{x+1}^{'} + \ldots + U_m^{'}}{U_x + U_{x+1} + \ldots + U_m} - \frac{\sum_{s=1}^{m} U_s^{'}}{\sum_{s=1}^{m} U_s} \\
&= \gamma_{jk}q_{jk}[\frac{\sum_{t=1}^{m-x} U_{x+t}t}{U_x + \ldots + U_m} + x - \frac{\sum_{s=1}^{m} U_s s}{1 + \sum_{s=1}^{m} U_s}] \\
&= \gamma_{jk}q_{jk}[\frac{\sum_{t=1}^{m-x} U_{x+t}t}{U_x + \ldots + U_m} + \frac{x - \sum_{t=1-x}^{m-x} U_{t+x}t}{1 + \sum_{s=1}^{m} U_s}] \\
&= \gamma_{jk}q_{jk}[\frac{x - \sum_{t=1-x}^{0} U_{x+t}t}{1 + \sum_{s=1} U_s} + (\frac{1}{U_x + \ldots + U_m} - \frac{1}{1 + \sum_{s=1}^{m} U_s}) \sum_{t=1}^{m-x} U_{x+t}t].
\end{aligned}$$

The first term in the bracket is positive, and the second term is nonnegative. So $\frac{\partial \ln P(X_j \geq x|\boldsymbol{\alpha})}{\partial \alpha_k} > 0$, and consequently $\frac{\partial P(X_j \geq x|\boldsymbol{\alpha})}{\partial \alpha_k} > 0$. Thus, SOM holds for the coordinate α_k. In fact, this is true for every α_k since the summation in the bracket is always positive. This is also true for a more restrictive sum score $X_+ = \sum_{j:Q_{jk}X_j}$. It is observed that item j will not be counted in this derivative if $q_{jk} = 0$, thus indicating that only those items with $q_{jk} \neq 0$ are taken into the calculation of X_+.

3.3 Stochastic Ordering of Latent Variables

Little is known about the SOL (Hemker, et. al, 1997) – $P(\alpha_1 > c_1, \ldots, \alpha_K > c_K|X_+ = x)$ when the latent trait is multidimensional. A weaker property related to SOL is that for any two values of α_k, $z_1 > z_2$, the ratio

$$r(z_1, z_2) = \frac{P(\alpha_k = z_1|data, \alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_K)}{P(\alpha_k = z_2|data, \alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_K)}$$

is nondecreasing in $X_+ = \sum_{j:q_{jk}=1} X_j$, with all other parameters fixed. First, let π_{01} and π_{02} be the prior distributions $\alpha_k = z_1$ and $\alpha_k = z_2$, respectively. The posterior odds of $\alpha_k = z_1$ to $\alpha_k = z_2$ is

$$r(z_1, z_2) = \frac{\pi_{01}}{\pi_{02}} \prod_{j=1}^{J} \prod_{t=1}^{m} \left\{ \exp[t\gamma_{jk} q_{jk}(z_1 - z_2)] \frac{1 + \sum_{y=1}^{m} C_k^y \exp(y\gamma_{jk} q_{jk} z_2)}{1 + \sum_{y=1}^{m} C_k^y \exp(y\gamma_{jk} q_{jk} z_1)} \right\}^{I_{[x_j=t]}}.$$

Here, $C_k^y = \exp[\beta_{yj} + \sum_{l \neq k}^{K} y\gamma_{jl} q_{jl} \alpha_l]$. For an item with $q_{jk} = 0$, the score on that item does not count in the odds ratio since the ratio is 1 no matter what score it gets. Since the second part in the bracket does not depend on data, it will be cancelled out in the odds ratio. Let $X(S)$ and $X(T)$ be the response vectors to yield total score S and T with $S > T$. $X(S)_j$ and $X(T)_j$ are the responses to the item j in the vector $X(S)$ and $X(T)$, respectively. The odds ratio of these two response vector is

$$\frac{r(z_1, z_2; S)}{r(z_1, z_2; T)} = \exp(z_1 - z_2) \sum_{j:q_{jk}=1} [\gamma_{jk}(X(S)_j - X(T)_j)].$$

If $\gamma_{jk} = \gamma_k$, the odds ratio depends completely on the total score on those item who require the kth attribute. That is, the posterior odds ratio for each coordinate α_k is nondecreasing by this total score, and consequently, SOL holds. If γ_{jk} varies over items, SOL does not in general hold for this model.

3.4 Junker and Sijtsman's Property

Finally, a new type of monotonicity condition seems worthy of studying for the GDM. In a standard unidimensional IRT model, a higher level of the latent trait is associated with a higher probability of correct response. A corresponding property for the a cognitive diagnostic model could be the relationship between the number of task-relevant attributes the examinee masters and the probability of correct task performance (Junker & Sijtsma, 2001). This might require that the item response function in GDM model be nondecreasing in $m_{ij} = \sum_{k=1}^{K} \alpha_k q_{jk} \gamma_{jk}$. The odds for scores $x_1 > x_2$ is calculated as

$$h(x_1, x_2) = \frac{P(X_j = x_1|\boldsymbol{\alpha})}{P(X_j = x_2|\boldsymbol{\alpha})} = exp[\beta_{x_1 j} - \beta_{x_2 j} + (x_1 - x_2(\sum_k \gamma_{jk} q_{jk} \alpha_k))].$$

Since the first term in the exponent will cancel out in the odds ratio of two different realizations of $\boldsymbol{\alpha}$, the odds ratio over different realizations of $\boldsymbol{\alpha}$: $\boldsymbol{\alpha}(1)$ and $\boldsymbol{\alpha}(2)$, is

$$\frac{h(x_1, x_2; \boldsymbol{\alpha}(1))}{h(x_1, x_2; \boldsymbol{\alpha}(2))} = \exp(x_1 - x_2) \sum_{k:q_{jk}=1} \gamma_{jk}(\alpha(1)_k - \alpha(2)_k),$$

where $\boldsymbol{\alpha}(1)$ and $\boldsymbol{\alpha}(2)$ represent two vector of possible realization of $\boldsymbol{\alpha}$. If $\gamma_{jk} = \gamma_j$, the odds ratio is nondecreasing as the difference on the scores of task-relevant attributes increases. However, when γ_{jk} varies over the task-relevant attributes, this monotonicity does not hold in general.

4 Discussion

Relating the latent traits to simple and useful data summaries is important when communicating to the public. For example, SOL (Hemker et al., 1997) was considered, which asserted that the higher score on the items that require latent attribute k, the easier it is to get a higher category in this attribute. For inference on the latent traits, SOL by the sum score is more useful than SOM of the sum score. The MLR is more stringent than these two. When $\gamma_{jk} = \gamma_k$, the MLR holds and this implies that both SOM and SOL hold for GDM model. In general, MLR does not hold for GDM because γ_{jk} varies over items. The same is true for the SOL property. These statements are very much in accordance with those of the GPCM (Hemker et al., 1996, 1997; see Table 1). Under the GDM, similar to the GPCM, some weighted sum score $\sum_{j:q_{jk}=1} \gamma_{jk} x_j$, not $\sum_{j:q_{jk}=1} x_j$, has the MLR property. By fitting a stringent GDM with equal slope across items, the observed total score $X_+ = \sum_{j:q_{jk}=1} x_j$ has the MLR property. However, the model fit might be lost. By fitting the GDM without the slope restriction,the model fit increases, but the MLR property is sacrificed. In practice, a trade-off is necessary between these two choices.

Table 1

Monotonicity Properties of Models

Properties	Rasch, PCM, RSM,	2PL/3PL, GPCM	GDM with equal slope	GDM
SOM	X	X	X	X
SOL	X		X	
MLR	X		X	

As mentioned above, SOL generally does not hold for the GDM. However, it is not clear how seriously the violation of SOL will affect the inference of the latent traits in a fitted model. It is incumbent on GDM model users to check that SOL holds in the fitted model before asserting that a higher sum score corresponds to a higher category in the latent traits. A step in future research might be a thorough simulation study to show the severeness of this violation.

References

Andrich, D. (1978). A rating scale formulation for ordered response categories. *Psychometrika, 43*, 561–573.

Grayson, D. A. (1988). Two-group classification in latent trait theory: Scores with monotone likelihood ratio. *Psychometrika, 53*, 383–392.

Hemker, B. T., Sijtsma, K., Molenaar, I. W., & Junker, B. W. (1996). Polytomous IRT models and monotone likelihood ratio of the total score. *Psychometrika, 61*, 679–693.

Hemker, B. T., Sijtsma, K., Molenaar, I. W., & Junker, B. W. (1997). Stochastic ordering using the latent trait and the sum score in polytomous IRT models. *Psychometrika, 62*, 331–347.

Holland, P. W., & Rosenbaum, P. R. (1986). Conditional association and unidimensionality in monotone latent trait models. *Annals of Statistics, 14*, 1523–1543.

Huynh, H. (1994). A new proof for monotone likelihood ratio for the sum of independent bernoulli random variables. *Psychometrika, 59*, 77–79.

Lehmann, E. L. (1959). *Testing statistical hypothesis.* New York: Wiley.

Junker, B. W. (1993). Conditional association essential independence and monotone unidimensional item response models. *The Annals of Statistics, 21*, 1359–1378.

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*, 258–272.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149–174.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen, Denmark: Nielsen & Lydiche.

Samejima, F. (2001). *Substantiveness in psychometrics.* Paper presented at the first international meeting of the Psychometric Society, Osaka, Japan.

van der Ark, L. A. (2005). Stochastic ordering of the latent trait by the sum score under various polytomous IRT models. *Psychometrika, 70*, 283–304.

von Davier, M. (2005). *A general diagnostic model applied to language testing data* (ETS Research Rep. No. RR-05-16). Princeton, NJ: ETS.

von Davier, M., & Yamamoto, K. (2004). Partially observed mixtures of IRT models: An extension of the generalized partial credit model. *Applied Psychological Measurement, 28*, 389–406.