数理解析研究所講究録
第 1879 巻 2014 年 180-191

ノンパラメトリック項目反応理論のための数理最適化モデル

東京工業大学 大学院・社会理工学研究科　高野 祐一
Yuichi Takano
Graduate School of Decision Science and Technology,
Tokyo Institute of Technology

東京工業大学 大学院・社会理工学研究科　角田 信太郎
Shintaro Tsunoda
Graduate School of Decision Science and Technology,
Tokyo Institute of Technology

東京工業大学 大学院・社会理工学研究科　村木 正昭
Masaaki Muraki
Graduate School of Decision Science and Technology,
Tokyo Institute of Technology

Abstract

項目反応理論とは, テストの回答結果から「被験者の能力」と「設問の正答確率を表す項目特性曲線」を推定するテスト理論であり, TOEFL や IT パスポート試験などで実際に利用されている. 項目特性曲線の形状をロジスティック曲線に限定したパラメトリック項目反応理論が広く使われているが, ロジスティック曲線では当てはまりの悪い設問が数多く存在することがしばしば問題になる. 本論文では, 項目特性曲線の形状に強い仮定を置かないノンパラメトリック項目反応理論に着目し, 被験者の能力と項目特性曲線を同時に推定するための定式化と発見的解法を提案する.

Keywords: 項目反応理論, 混合整数非線形計画, ノンパラメトリック推定, 発見的解法

1 Introduction

Item response theory (IRT) [1, 17] is a modern test theory for the design, analysis, and scoring of tests. The key component of IRT is the item characteristic curve (ICC), which shows the relationship between the examinee's latent ability and the probability of correct answer. On the basis of the item response data of examinees, IRT models estimate the ICCs of question items and the latent abilities of examinees. IRT methodologies enable one to closely examine item characteristics, such as difficulty and discrimination, and to investigate not the test score but the latent (i.e., not directly observable) ability of each examinee. IRT models can be divided into two categories according to approaches to ICC estimation. Parametric item response theory (PIRT) models typically force ICCs to be parametric functions (e.g., logistic curves or normal ogives). On the other hand, this paper focuses on nonparametric item response theory (NIRT) models, which do not assume any particular parametric forms for ICCs.

NIRT has its origin in Meredith's work [10] and Mokken scale analysis [11], and it has achieved steady development in both the theory and applications (see, e.g., [15, 18, 19, 20, 22, 23]). The greatest benefit of NIRT models is being able to estimate various forms of ICCs on mild assumptions. Indeed, it has been demonstrated, e.g., in [3, 4, 16], that PIRT models do not always fit the data well. In this case, NIRT models, which provide a more flexible framework, are particularly useful. NIRT models are also useful to examine whether model assumptions of PIRT are valid or not (see, e.g., [6]). However, greater flexibility in nonparametric ICCs sometimes makes a model overly fit to the data. As pointed out by [15], consequently, estimation results obtained by NIRT models can be unstable especially when using small-sized item response data.

There are several estimation methods for nonparametric ICCs. The most commonly-used approach is kernel smoothing, which was first applied by Ramsay [16] to nonparametric ICC estimation. Although the usefulness of kernel smoothing methods has been shown, e.g., in [4], it may be that some estimated ICCs are decreasing with respect to the latent ability. Meanwhile, isotonic regression methods can always provide nondecreasing ICCs. Lee [7] compared the performance of three estimation procedures: isotonic regression, smoothed isotonic regression and kernel smoothing, and demonstrated that the smoothed isotonic regression yielded better results than the kernel smoothing did. A number of studies have assessed the goodness of fit of PIRT models by means of these estimation procedures for nonparametric ICCs (see, e.g., [4, 8, 9, 24, 25]). These procedures, however, estimate nonparametric ICCs under the assumption that latent abilities of examinees are predetermined.

The purpose of the present paper is to build a new computational framework for estimating the nonparametric ICCs and the latent abilities of examinees simultaneously. To accomplish this, we formulate mathematical optimization models for NIRT as mixed integer nonlinear programming (MINLP) problems. Mathematical optimization methodology makes it possible to place various restrictions on excessively flexible ICCs. In addition to the existing constraints, i.e., monotone homogeneity and double monotonicity, we propose slope smoothing constraints to prevent ICCs from overfitting the data. Although it is very hard to obtain an exact solution to the resulting optimization problems, we develop a heuristic optimization algorithm to find a good-quality solution in a reasonable amount of time.

2 Nonparametric Item Response Theory

Let us suppose that examinees $i = 1, 2, \ldots, I$ took a test consisting of dichotomously scored question items $j = 1, 2, \ldots, J$. More specifically, we are given the binary item response data,

$$\boldsymbol{U} = (u_{i,j};\ i = 1, 2, \ldots, I,\ j = 1, 2, \ldots, J) \in \{0, 1\}^{I \times J},$$

where $u_{i,j} = 1$ if examinee i gave a correct answer to question item j, otherwise $u_{i,j} = 0$. The main objective of the item response theory (IRT) is to estimate the item characteristic curves (ICCs) and the latent abilities of examinees on the basis of the item response data, $\boldsymbol{U}$.

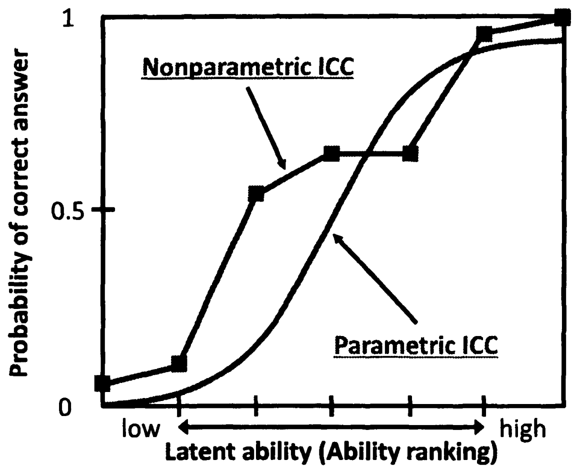

Figure 1: Parametric/Nonparametric item characteristic curves

In particular, this paper explores the nonparametric item response theory (NIRT) that employs nonparametric ICCs. In a conventional way, we assume throughout the present paper that

Unidimensionality : latent abilities of all examinees can be evaluated unidimensionally.

Local Independence : item responses are conditionally independent of each other given an individual latent ability.

In what follows, we shall consider ability rankings; that is, we evaluate the latent abilities of examinees on a discrete scale of $t = 1, 2, \ldots, T$. To estimate nonparametric ICCs, we introduce the decision variables:

$$\boldsymbol{X} = (x_{j,t};\ j = 1, 2, \ldots, J,\ t = 1, 2, \ldots, T) \in \mathbb{R}^{J \times T},$$

where $x_{j,t}$ is the probability of question item j answered correctly by examinees of ability ranking t. Figure 1 illustrates a nonparametric ICC which is represented as a piecewise linear function.

The fundamental property required for ICCs is monotone homogeneity (MH) [10, 11]. This requires that all ICCs are nondecreasing with a latent ability. This means that the probability of correct answer does not decrease with the ability ranking of examinee. Thus, the following constraints must be imposed on $\boldsymbol{X}$:

$$\textsf{Monotone Homogeneity}: \quad 0 \le x_{j,1} \le x_{j,2} \le \cdots \le x_{j,T} \le 1 \quad (\forall j = 1, 2, \ldots, J). \tag{1}$$

An additional assumption of nonparametric ICC is double monotonicity (DM) [11, 13]. This implies that the ICC of one item does not intersect with the other. In other words, for all levels of examinees, the difficulties of two question items are never reversed. To formulate a clear definition, we suppose that there is a permutation:

$$\sigma : \{1, 2, \ldots, J\} \to \{1, 2, \ldots, J\},$$

where $\sigma(k) = j$ means that the k-th most difficult item is question item j. We refer to σ as a difficulty ranking function. Then, the DM constraints are written as follows:

$$\text{Double Monotonicity}: \quad x_{\sigma(1),t} \leq x_{\sigma(2),t} \leq \cdots \leq x_{\sigma(J),t} \quad (\forall t = 1, 2, \ldots, T). \tag{2}$$

This means that, for all examinees, the probability of answering a high-ranking item correctly is lower than that of a low-ranking one.

To estimate ability rankings of examinees, we further introduce the decision variables,

$$\boldsymbol{Y} = (y_{i,t};\ i = 1, 2, \ldots, I,\ t = 1, 2, \ldots, T) \in \{0, 1\}^{I \times T},$$

where $y_{i,t} = 1$ if the ability ranking of examinee i is estimated to t, otherwise $y_{i,t} = 0$. Since only one ability ranking should be assigned to each examinee, $\boldsymbol{Y}$ must satisfy the following constraints:

$$\sum_{t=1}^{T} y_{i,t} = 1 \quad (\forall i = 1, 2, \ldots, I), \tag{3}$$

$$y_{i,t} \in \{0, 1\} \quad (\forall i = 1, 2, \ldots, I,\ \forall t = 1, 2, \ldots, T). \tag{4}$$

In what follows, we define a log likelihood function to be maximized. Given $\boldsymbol{x}_j := (x_{j,1}, x_{j,2}, \ldots, x_{j,T})$ and $\boldsymbol{y}_i := (y_{i,1}, y_{i,2}, \ldots, y_{i,T})$, the probability of having the response $u_{i,j}$ can be written as follows:

$$\Pr(u_{i,j} | \boldsymbol{x}_j, \boldsymbol{y}_i) = \sum_{t=1}^{T} y_{i,t} (x_{j,t})^{u_{i,j}} (1 - x_{j,t})^{1 - u_{i,j}}.$$

Under the local independence assumption, the probability of having the response $\boldsymbol{u}_i := (u_{i,1}, u_{i,2}, \ldots, u_{i,J})$ of examinee i becomes

$$\Pr(\boldsymbol{u}_i | \boldsymbol{X}, \boldsymbol{y}_i) = \prod_{j=1}^{J} \Pr(u_{i,j} | \boldsymbol{x}_j, \boldsymbol{y}_i).$$

Considering that the responses of different examinees are independent, we can see that the overall item response $\boldsymbol{U}$ occurs with the probability:

$$\Pr(\boldsymbol{U} | \boldsymbol{X}, \boldsymbol{Y}) = \prod_{i=1}^{I} \Pr(\boldsymbol{u}_i | \boldsymbol{X}, \boldsymbol{y}_i) = \prod_{i=1}^{I} \prod_{j=1}^{J} \left(\sum_{t=1}^{T} y_{i,t} (x_{j,t})^{u_{i,j}} (1 - x_{j,t})^{1 - u_{i,j}} \right).$$

Finally, by treating $\boldsymbol{X}$ and $\boldsymbol{Y}$ as decision variables, the log likelihood function is defined as follows:

$$\ell(\boldsymbol{X}, \boldsymbol{Y} | \boldsymbol{U}) = \log \Pr(\boldsymbol{U} | \boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=1}^{I} \sum_{j=1}^{J} \log \left(\sum_{t=1}^{T} y_{i,t} (x_{j,t})^{u_{i,j}} (1 - x_{j,t})^{1 - u_{i,j}} \right).$$

3 Mathematical Optimization Models

This section presents several mathematical optimization models for NIRT.

3.1 Monotone homogeneity model

In view of the constraints (3) and (4), the log likelihood function can be rewritten as follows:

$$\begin{aligned}\ell(\boldsymbol{X},\boldsymbol{Y}|\boldsymbol{U}) &\overset{(3),(4)}{=} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} y_{i,t}\log\left((x_{j,t})^{u_{i,j}}(1-x_{j,t})^{1-u_{i,j}}\right)\\ &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} y_{i,t}\left(u_{i,j}\log(x_{j,t}) + (1-u_{i,j})\log(1-x_{j,t})\right). \end{aligned} \tag{5}$$

The monotone homogeneity (MH) model estimates $\boldsymbol{X}$ and $\boldsymbol{Y}$ so that the log likelihood function, $\ell(\boldsymbol{X},\boldsymbol{Y}|\boldsymbol{U})$, is maximized under the conditions (1), (3) and (4). Consequently, the MH model can be framed as the following mixed integer nonlinear programming (MINLP) problem:

$$(\text{MHM})\quad\left|\begin{array}{ll} \underset{\boldsymbol{X},\boldsymbol{Y}}{\text{maximize}} & \displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} y_{i,t}\left(u_{i,j}\log(x_{j,t}) + (1-u_{i,j})\log(1-x_{j,t})\right)\\ \text{subject to} & 0 \le x_{j,1} \le x_{j,2} \le \cdots \le x_{j,T} \le 1 \quad (\forall j = 1,2,\ldots,J),\\ & \displaystyle\sum_{t=1}^{T} y_{i,t} = 1 \quad (\forall i = 1,2,\ldots,I),\\ & y_{i,t} \in \{0,1\} \quad (\forall i = 1,2,\ldots,I,\ \forall t = 1,2,\ldots,T). \end{array}\right.$$

3.2 Double monotonicity model

Next, we ponder a mathematical optimization problem with the double monotonicity (DM) constraints (2).

Let us recall that $\sigma(k) = j$ means that the k-th most difficult item is question item j. In the sequel, we shall represent a difficulty ranking function σ by using the following permutation matrix:

$$\boldsymbol{Z} = (z_{j,k};\ j = 1,2,\ldots,J,\ k = 1,2,\ldots,J) \in \{0,1\}^{J\times J}, \tag{6}$$

$$z_{j,k} = 1 \iff \sigma(k) = j. \tag{7}$$

The optimization model presented below finds an appropriate difficulty ranking by treating $\boldsymbol{Z}$ as the decision variable. Note that the permutation matrix needs to satisfy the following conditions:

$$\sum_{k=1}^{J} z_{j,k} = 1 \quad (\forall j = 1, 2, \ldots, J), \tag{8}$$

$$\sum_{j=1}^{J} z_{j,k} = 1 \quad (\forall k = 1, 2, \ldots, J), \tag{9}$$

$$z_{j,k} \in \{0, 1\} \quad (\forall j = 1, 2, \ldots, J,\ \forall k = 1, 2, \ldots, J). \tag{10}$$

To estimate ICCs under the DM constraints, we use new decision variables:

$$\boldsymbol{W} = (w_{k,t};\ k = 1, 2, \ldots, J,\ t = 1, 2, \ldots, T) \in \mathbb{R}^{J \times T},$$

which represents the probability of the k-th most difficult item answered correctly by examinees of ability ranking t. In this case, the monotone homogeneity and double monotonicity constraints on $\boldsymbol{W}$ can be expressed as follows:

$$\text{Monotone Homogeneity}: \quad 0 \leq w_{k,1} \leq w_{k,2} \leq \cdots \leq w_{k,T} \leq 1 \quad (\forall k = 1, 2, \ldots, J), \tag{11}$$

$$\text{Double Monotonicity}: \quad w_{1,t} \leq w_{2,t} \leq \cdots \leq w_{J,t} \quad (\forall t = 1, 2, \ldots, T). \tag{12}$$

The associated log likelihood function becomes

$$\begin{aligned}
\ell(\boldsymbol{W}, \boldsymbol{Y}, \boldsymbol{Z} | \boldsymbol{U}) &\overset{(5)}{=} \sum_{i=1}^{I} \sum_{k=1}^{J} \sum_{t=1}^{T} y_{i,t} \left(u_{i,\sigma(k)} \log(w_{k,t}) + (1 - u_{i,\sigma(k)}) \log(1 - w_{k,t}) \right) \\
&\overset{(6),(7)}{=} \sum_{i=1}^{I} \sum_{k=1}^{J} \sum_{t=1}^{T} y_{i,t} \left(\sum_{j=1}^{J} z_{j,k} \left(u_{i,j} \log(w_{k,t}) + (1 - u_{i,j}) \log(1 - w_{k,t}) \right) \right) \\
&= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{J} \sum_{t=1}^{T} y_{i,t} z_{j,k} \left(u_{i,j} \log(w_{k,t}) + (1 - u_{i,j}) \log(1 - w_{k,t}) \right).
\end{aligned}$$

We are now in a position to formulate a DM model, i.e., the problem of maximizing the likelihood function, $\ell(\boldsymbol{W}, \boldsymbol{Y}, \boldsymbol{Z} | \boldsymbol{U})$, subject to the constraints (3), (4), (8)–(10), (11) and (12) as the following MIMLP problem:

$$
(\text{DMM}) \quad \left|
\begin{array}{ll}
\underset{\boldsymbol{W},\boldsymbol{Y},\boldsymbol{Z}}{\text{maximize}} & \displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{J}\sum_{t=1}^{T} y_{i,t}z_{j,k}\left(u_{i,j}\log(w_{k,t}) + (1-u_{i,j})\log(1-w_{k,t})\right) \\
\text{subject to} & 0 \le w_{k,1} \le w_{k,2} \le \cdots \le w_{k,T} \le 1 \quad (\forall k = 1,2,\ldots,J), \\
& w_{1,t} \le w_{2,t} \le \cdots \le w_{J,t} \quad (\forall t = 1,2,\ldots,T), \\
& \displaystyle\sum_{k=1}^{J} z_{j,k} = 1 \quad (\forall j = 1,2,\ldots,J), \\
& \displaystyle\sum_{j=1}^{J} z_{j,k} = 1 \quad (\forall k = 1,2,\ldots,J), \\
& z_{j,k} \in \{0,1\} \quad (\forall j = 1,2,\ldots,J,\ \forall k = 1,2,\ldots,J), \\
& \displaystyle\sum_{t=1}^{T} y_{i,t} = 1 \quad (\forall i = 1,2,\ldots,I), \\
& y_{i,t} \in \{0,1\} \quad (\forall i = 1,2,\ldots,I,\ \forall t = 1,2,\ldots,T).
\end{array}
\right.
$$

3.3 Slope smoothing model

It has been pointed out, e.g., in [15], that estimated results can be unstable especially for small-sized item response data. This instability is caused by the enhanced flexibility of nonparametric ICCs. To overcome this drawback, it is effective to decrease flexibility of nonparametric ICCs moderately. This sort of approach is frequently utilized to enhance the generalization capability in statistical learning methods (see, e.g., [5]). For this reason, we propose additional constraints to force the slope of each ICC to vary smoothly. We shall call them "slope smoothing (SS) constraints", which are expressed as follows:

$$
\text{Slope Smoothing}: \quad \sum_{t=2}^{T-1} |(x_{j,t+1} - x_{j,t}) - (x_{j,t} - x_{j,t-1})| \le \gamma \quad (\forall j = 1,2,\ldots,J), \tag{13}
$$

where $\gamma \ge 0$ is an user-defined parameter. If γ is sufficiently large, the SS constraints (13) are invalidated. By contrast, $\gamma = 0$ forces all ICCs to be straight lines.

By placing the SS constraints (13) on ICCs of problem (MHM), we can pose the SS model as follows:

$$
(\text{SSM}) \quad \left|
\begin{array}{ll}
\underset{\boldsymbol{X},\boldsymbol{Y}}{\text{maximize}} & \displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} y_{i,t}\left(u_{i,j}\log(x_{j,t}) + (1-u_{i,j})\log(1-x_{j,t})\right) \\
\text{subject to} & 0 \le x_{j,1} \le x_{j,2} \le \cdots \le x_{j,T} \le 1 \quad (\forall j = 1,2,\ldots,J), \\
& \displaystyle\sum_{t=2}^{T-1} |x_{j,t+1} - 2x_{j,t} + x_{j,t-1}| \le \gamma \quad (\forall j = 1,2,\ldots,J), \\
& \displaystyle\sum_{t=1}^{T} y_{i,t} = 1 \quad (\forall i = 1,2,\ldots,I), \\
& y_{i,t} \in \{0,1\} \quad (\forall i = 1,2,\ldots,I,\ \forall t = 1,2,\ldots,T).
\end{array}
\right.
$$

4 Heuristic Optimization Algorithm

The optimization models presented in the previous section are mixed integer nonlinear programming (MINLP) problems, which are very hard to solve exactly. To efficiently compute a good-quality solution, we develop a heuristic optimization algorithm to the problems. In this section, we describe an algorithm for solving the slope smoothing model (SSM). We should notice that this algorithm can be readily applied to the monotone homogeneity model (MHM) because problem (MHM) is equivalent to problem (SSM) with $\gamma = \infty$.

We begin by giving an ability ranking to each examinee as an initial solution. To set an examinee's ability, one may use the number of question items that s/he answered correctly. Then, we denote by

$$\bar{\boldsymbol{Y}} = (\bar{y}_{i,t};\ i = 1, 2, \ldots, I,\ t = 1, 2, \ldots, T)$$

the determined ability rankings.

Next, we solve problem (SSM) in which the decision variable $\boldsymbol{Y}$ is fixed to $\bar{\boldsymbol{Y}}$. This problem can be decomposed into ones of each ICC ($j = 1, 2, \ldots, J$):

$$(\mathrm{SSM}(j|\bar{\boldsymbol{Y}}))\quad \left|\begin{array}{ll} \underset{\boldsymbol{x}_j}{\text{maximize}} & \displaystyle\sum_{i=1}^{I}\sum_{t=1}^{T} \bar{y}_{i,t}\,(u_{i,j}\log(x_{j,t}) + (1-u_{i,j})\log(1-x_{j,t})) \\ \text{subject to} & 0 \le x_{j,1} \le x_{j,2} \le \cdots \le x_{j,T} \le 1, \\ & \displaystyle\sum_{t=2}^{T-1} |x_{j,t+1} - 2x_{j,t} + x_{j,t-1}| \le \gamma. \end{array}\right.$$

Although the SS constraints (13) are nonlinear and nondifferentiable, it is well known that this sort of constraints can be converted into linear ones. Specifically, we can reformulate problem ($\mathrm{SSM}(j|\bar{\boldsymbol{Y}})$) as follows:

$$(\mathrm{SSM}(j|\bar{\boldsymbol{Y}}))\quad \left|\begin{array}{ll} \underset{\boldsymbol{s}_j,\boldsymbol{v}_j,\boldsymbol{x}_j}{\text{maximize}} & \displaystyle\sum_{i=1}^{I}\sum_{t=1}^{T} \bar{y}_{i,t}\,(u_{i,j}\log(x_{j,t}) + (1-u_{i,j})\log(1-x_{j,t})) \\ \text{subject to} & 0 \le x_{j,1} \le x_{j,2} \le \cdots \le x_{j,T} \le 1, \\ & \displaystyle\sum_{t=2}^{T-1} (s_{j,t} + v_{j,t}) \le \gamma, \\ & s_{j,t} - v_{j,t} = x_{j,t+1} - 2x_{j,t} + x_{j,t-1} \quad (\forall t = 2, 3, \ldots, T-1), \\ & s_{j,t} \ge 0,\ v_{j,t} \ge 0 \quad (\forall t = 2, 3, \ldots, T-1), \end{array}\right.$$

where $\boldsymbol{s}_j = (s_{j,2}, s_{j,3}, \ldots, s_{j,T-1})$ and $\boldsymbol{v}_j = (v_{j,2}, v_{j,3}, \ldots, v_{j,T-1})$ for $j = 1, 2, \ldots, J$ are auxiliary decision variables. When the SS constraint (13) of ICC j is tight, $s_{j,t}$ and $v_{j,t}$ correspond to positive and negative parts of $x_{j,t+1} - 2x_{j,t} + x_{j,t-1}$, respectively; therefore, $s_{j,t} + v_{j,t}$ coincides with $|x_{j,t+1} - 2x_{j,t} + x_{j,t-1}|$. Since problem ($\mathrm{SSM}(j|\bar{\boldsymbol{Y}})$) is concave function maximization with the linear constraints, it can be solved exactly with a standard nonlinear optimization solver.

Let

$$\bar{\boldsymbol{X}} = (\bar{x}_{j,t};\ j = 1, 2, \ldots, J,\ t = 1, 2, \ldots, T)$$

be optimal solutions to problems (SSM($j|\bar{\boldsymbol{Y}}$)) for $j = 1, 2, \ldots, J$. Now, we solve problem (SSM) in which the decision variable $\boldsymbol{X}$ is fixed to $\bar{\boldsymbol{X}}$. This problem can be decomposed into ones of each examinee ($i = 1, 2, \ldots, I$):

$$(\mathrm{SSM}(i|\bar{\boldsymbol{X}})) \quad \left| \begin{array}{ll} \underset{\boldsymbol{y}_i}{\text{maximize}} & \displaystyle\sum_{j=1}^{J}\sum_{t=1}^{T} y_{i,t}\left(u_{i,j}\log(\bar{x}_{j,t}) + (1-u_{i,j})\log(1-\bar{x}_{j,t})\right) \\ \text{subject to} & \displaystyle\sum_{t=1}^{T} y_{i,t} = 1, \\ & y_{i,t} \in \{0,1\} \quad (\forall t = 1, 2, \ldots, T). \end{array} \right.$$

Here, the objective function can be rewritten as follows:

$$\sum_{t=1}^{T} y_{i,t} \underbrace{\sum_{j=1}^{J}\left(u_{i,j}\log(\bar{x}_{j,t}) + (1-u_{i,j})\log(1-\bar{x}_{j,t})\right)}_{\ell(i,t)}.$$

Therefore, to determine an ability ranking of examinee i, it is only necessary to select t such that $\ell(i,t)$ is maximized. It follows that problem (SSM($i|\bar{\boldsymbol{X}}$)) can be easily solved by sorting $\ell(i,t)$. In this manner, we update $\bar{\boldsymbol{Y}}$ and return to the first step to find better $\boldsymbol{X}$. By repeating this procedure, the objective, $\ell(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}|\boldsymbol{U})$, monotonically increases. We terminate this algorithm when the solutions are unchanged. Our heuristic optimization algorithm is summarized in Algorithm 1.

Algorithm 1: Heuristic Optimization Algorithm for Solving Problem (SSM)

Step 0. (Initialization) Set an initial ability ranking, $\bar{\boldsymbol{Y}}$.

Step 1. (ICC Estimation) Solve problems (SSM($j|\bar{\boldsymbol{Y}}$)) for all $j = 1, 2, ..., J$. Let $\bar{\boldsymbol{X}}$ be an optimal solution.

Step 2. (Ability Estimation) Solve problems (SSM($i|\bar{\boldsymbol{X}}$)) for all $i = 1, 2, ..., I$. Let $\bar{\boldsymbol{Y}}$ be an optimal solution.

Step 3. (Termination Condition) If $\bar{\boldsymbol{Y}}$ remains the same as the previous one, terminate the algorithm with the solution: $(\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}})$. Otherwise, return to Step 1.

A search strategy of Algorithm 1 is similar to that of the well-known expectation-maximization (EM) algorithm [2]. In contrast to the standard EM algorithm, however, Algorithm 1 estimates the discrete variable $\boldsymbol{Y}$.

5 Conclusion

We dealt with mathematical optimization models and a heuristic optimization algorithm for nonparametric item response theory (NIRT). A future direction of study will be to extend our formulation to polytomous NIRT models [14, 20, 21].

References

[1] A. Birnbaum, "Some Latent Trait Models and Their Use in Inferring an Examinee's Ability," In F. Lord & M. Novick (eds.), *Statistical Theories of Mental Test Scores* (Addison-Wesley, 1968), pp.397–492.

[2] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, Vol.39, No.1, pp.1–38 (1977).

[3] J. Douglas, "Joint Consistency of Nonparametric Item Characteristic Curve and Ability Estimation," *Psychometrika*, Vol.62, No.1, pp.7–28 (1997).

[4] J. Douglas and A. Cohen, "Nonparametric Item Response Function Estimation for Assessing Parametric Model Fit," *Applied Psychological Measurement*, Vol.25, No.3, pp.234–243 (2001).

[5] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd Edition (Springer, 2009).

[6] B. Junker and K. Sijtsma, "Nonparametric Item Response Theory in Action: An Overview of the Special Issue," *Applied Psychological Measurement*, Vol.25, No.3, pp.211–220 (2001).

[7] Y.-S. Lee, "A Comparison of Methods for Nonparametric Estimation of Item Characteristic Curves for Binary Items," *Applied Psychological Measurement*, Vol.31, No.2, pp.121–134 (2007).

[8] Y.-S. Lee, J.A. Wollack, and J. Douglas, "On the Use of Nonparametric Item Characteristic Curve Estimation Techniques for Checking Parametric Model Fit," *Educational and Psychological Measurement*, Vol.69, No.2, pp.181–197 (2009).

[9] T. Liang and C.S. Wells, "A Model Fit Statistic for Generalized Partial Credit Model," *Educational and Psychological Measurement*, Vol.69, No.6, pp.913–928 (2009).

[10] W. Meredith, "Some Results based on a General Stochastic Model for Mental Tests," *Psychometrika*, Vol.30, No.4, pp.419–440 (1965).

[11] R.J. Mokken, *A Theory and Procedure of Scale Analysis with Applications in Political Research* (Walter de Gruyter, 1971).

[12] R.J. Mokken, "Nonparametric Models for Dichotomous Responses," In W.J. van der Linden and R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (Springer, 1997), pp.351–367.

[13] R.J. Mokken and C. Lewis, "A Nonparametric Approach to the Analysis of Dichotomous Item Responses," *Applied Psychological Measurement*, Vol.6, No.4, pp.417–430 (1982).

[14] I.W. Molenaar, "Nonparametric Models for Polytomous Responses," In W.J. van der Linden and R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (Springer, 1997), pp.369–380.

[15] I.W. Molenaar, "Thirty Years of Nonparametric Item Response Theory," *Applied Psychological Measurement*, Vol.25, No.3, pp.295–299 (2001).

[16] J.O. Ramsay, "Kernel Smoothing Approaches to Nonparametric Item Characteristic Curve Estimation," *Psychometrika*, Vol.56, No.4, pp.611–630 (1991).

[17] G. Rasch, *Probabilistic Models for Some Intelligence and Attainment Tests* (Danish Institute for Educational Research, 1960). Expanded Edition (University of Chicago Press, 1980).

[18] K. Sijtsma, "Methodology Review: Nonparametric IRT Approaches to the Analysis of Dichotomous Item Scores," *Applied Psychological Measurement*, Vol.22, No.1, pp.3–31 (1998).

[19] K. Sijtsma, "Developments in Measurement of Persons and Items by Means of Item Response Models," *Behaviormetrika*, Vol.28, No.1, pp.65–94 (2001).

[20] K. Sijtsma and I.W. Molenaar (eds.), *Introduction to Nonparametric Item Response Theory* (Saga, 2002).

[21] K. Sijtsma and L.A. van der Ark, "NIRT Analysis of Polytomous Item Scores: Dilemmas and Practical Solutions," In A.J. van Duijn and T.A.B. Snijders (eds.), *Essays on Item Response Theory* (Springer, 2001), pp.297–318.

[22] W.F. Stout, "A Nonparametric Approach for Assessing Latent Trait Unidimensionality," *Psychometrika*, Vol.52, No.4, pp.589–617 (1987).

[23] W.F. Stout, "Nonparametric Item Response Theory: A Maturing and Applicable Measurement Modeling Approach," *Applied Psychological Measurement*, Vol.25, No.3, pp.300–306 (2001).

[24] M.J. Sueiro and F.J. Abad, "Assessing Goodness of Fit in Item Response Theory With Nonparametric Models: A Comparison of Posterior Probabilities and Kernel-Smoothing Approaches," *Educational and Psychological Measurement*, Vol.71, No.5, pp.834–848 (2011).

[25] C.S. Wells and D.M. Bolt, "Investigation of a Nonparametric Procedure for Assessing Goodness-of-Fit in Item Response Theory," *Applied Measurement in Education*, Vol.21, No.1, pp.22–40 (2008).