Ali Ünlü

Knowledge Space Theory and Item Response Theory

Institut für Mathematik, Universitätsstraße, D-86135 Augsburg

http://www.math.uni-augsburg.de/
Impressum:

Herausgeber:
Institut für Mathematik
Universität Augsburg
86135 Augsburg
http://www.math.uni-augsburg.de/pages/de/forschung/preprints.shtml

ViSdP:
Ali Ünlü
Institut für Mathematik
Universität Augsburg
86135 Augsburg

Preprint: Sämtliche Rechte verbleiben den Autoren © 2009
1

Knowledge Space Theory and Item Response Theory

A. Ünlü¹

1.1 Introduction

The Guttman model (Guttman, 1944) can be viewed as a common origin of item response theory (IRT; e.g., Boomsma, Van Duijn, and Snijders, 2001; Fischer and Molenaar, 1995; Van der Linden and Hambleton, 1997) and knowledge space theory (KST; e.g., Doignon and Falmagne, 1985, 1999). They generalize the Guttman model in probabilistic, statistical and deterministic, combinatorial directions, respectively.

In KST, persons are represented by collections of items (of a representative and fully comprehensive domain) they are capable of mastering. Persons can be incomparable, with respect to set-inclusion. Items are assumed to be ordered, for instance, with respect to a hierarchy of mastery dependencies. Items can be incomparable, with respect to that hierarchy. In IRT, on the other hand, persons and items are, for instance, represented by single real numbers, ability and difficulty parameters, respectively. Persons and items are linearly ordered, with respect to the natural ordering of the real numbers. Conceptually speaking, KST may be viewed as a more ‘qualitative, behavioral’ approach (mainly based on combinatorics and stochastic processes), unlike IRT, as a ‘quantitative, statistical’ approach (mainly based on calculus and statistics). Further technical and philosophical differences between the two theories are discussed in Falmagne, Cosyn, Doble, Thiéry, and Uzun (2008, see also Chapter 2 of this book) and Ünlü (2007).

¹ Institute of Mathematics, University of Augsburg, Germany
A nonparametric, as opposed to a parametric, approach is pursued. Nonparametric IRT includes a broad range of parametric IRT models. Nevertheless, parametric IRT-type modeling strategies in KST are important directions for future research. For a logistic approach, see Stefanutti (2006); for a generalized normal ogive approach, see Ünlü (2006).

1.2 Nonparametric IRT: Axioms and Properties

This section reviews the axioms underlying the Mokken (1971) nonparametric IRT models of monotone homogeneity (MHM) and double monotonicity (DMM) (see also Mokken, 1997; Sijtsma, 1998; Sijtsma and Molenaar, 2002). The properties of monotone likelihood ratio (MLR) and stochastic ordering (SO) justifying the use of Mokken’s models as measurement models for persons are also reviewed.

Axioms. Let \( X_l \) with realization \( x_l \in \{0,1\} \) be the item score variable for item \( I_l \) (1 ≤ \( l \) ≤ \( m \)), and let \( X_+ = \sum_{l=1}^{m} X_l \) with realization \( x_+ \in \{0,1,\ldots,m\} \) denote the total score variable. A function \( f : \{0,1,\ldots,m\} \rightarrow \mathbb{R} \) is nondecreasing iff

\[
\forall x, y \in \{0,1,\ldots,m\}, x \leq y : f(x) \leq f(y).
\]

Let the latent trait be denoted by \( \theta, \theta \in \Theta \subseteq \mathbb{R} \); this is referred to as the axiom of unidimensionality. A function \( f : \Theta \rightarrow \mathbb{R} \) is nondecreasing iff it satisfies an obvious analog of the above condition. Let the conditional positive response probability \( P(X_l = 1|\theta) \) as a function of \( \theta \in \Theta \) be the item response function (IRF) of the item \( I_l \). The axiom of local independence states that

\[
P(X_1 = x_1, \ldots, X_m = x_m|\theta) = \prod_{l=1}^{m} P(X_l = x_l|\theta)
\]

for any \( x_l \in \{0,1\} \) and \( \theta \in \Theta \). The axiom of monotonicity holds iff any IRF \( P(X_l = 1|\theta) \) is nondecreasing. The axiom of invariant item ordering states that the IRFs \( P(X_l = 1|\theta) \) can be ordered such that

\[ ^2 \text{Throughout, only dichotomous items are considered.} \]
\[ \forall \theta \in \Theta : P(X_{l_1} = 1|\theta) \leq \cdots \leq P(X_{l_m} = 1|\theta) \]

where \(1 \leq l_i \leq m\) (\(1 \leq i \leq m\)).

Mokken’s MHM is based on the axioms of unidimensionality, local independence, and monotonicity. His DMM further adds the axiom of invariant item ordering.

Properties. MLR for the total score variable and latent trait plays an important role in IRT. It implies SO properties that can be interpreted in an IRT context (e.g., Hemker, Van der Ark, and Sijtsma, 2001; Ünlü, 2007; Van der Ark, 2001, 2005).

The total score variable \(X_+\) has MLR in \(\theta\) iff, for any \(0 \leq x_{+,1} \leq x_{+,2} \leq m\),

\[
\frac{P(X_+ = x_{+,2}|\theta)}{P(X_+ = x_{+,1}|\theta)}
\]

is a nondecreasing function of (unidimensional) \(\theta \in \Theta\). Similarly, the latent trait \(\theta\) has MLR in \(X_+\) iff, for any \(\theta_1 \leq \theta_2\),

\[
\frac{P(\theta_2|X_+ = x_+)}{P(\theta_1|X_+ = x_+)}
\]

is a nondecreasing function of \(0 \leq x_+ \leq m\).

The fundamental result (Ghurye and Wallace, 1959; Grayson, 1988; Huynh, 1994; Ünlü, 2008) states that under the axioms of unidimensionality, local independence, and monotonicity, the total score variable has MLR in the (unidimensional) latent trait.

The property of MLR implies that \(X_+\) is stochastically ordered by \(\theta\). The stochastic ordering of the manifest variable \(X_+\) by \(\theta\) (SOM) means that, for any \(0 \leq x_+ \leq m\),

\[
P(X_+ \geq x_+|\theta)
\]

is a nondecreasing function of (unidimensional) \(\theta \in \Theta\). The MLR property also implies that \(\theta\) is stochastically ordered by \(X_+\). The stochastic ordering of the latent trait \(\theta\) by \(X_+\) (SOL) means that, for any \(\theta_0 \in \Theta\),

\[
P(\theta \geq \theta_0|X_+ = x_+)
\]

is a nondecreasing function of \(0 \leq x_+ \leq m\). The property of SOL is very important for practical measurement, because it justifies the use of the total score variable to estimate the ordering of subjects on the latent trait. This is the key result that justifies the use of the MHM and DMM as measurement models for persons.
1.3 Application of Nonparametric IRT in KST

Ünlü (2007) generalizes the unidimensional nonparametric IRT axioms and properties to quasi-ordered person and indicator spaces, and applies the extended IRT concepts in KST.

**Axioms.** Let \( Q = \{I_l : 1 \leq l \leq m\} \). Let \( \mathcal{K} \) be a knowledge structure on \( Q \), partially ordered with respect to set-inclusion \( \subseteq \). The IRT concepts can be formulated for \((\mathcal{K}, \subseteq)\). For instance, a function \( f : \mathcal{K} \rightarrow \mathbb{R} \) is isotonic iff
\[
\forall K_1, K_2 \in \mathcal{K}, K_1 \subseteq K_2 : f(K_1) \leq f(K_2).
\]
The item response function (IRF) of an item \( I_l \in Q \) is
\[
P(X_l = 1 | .) : \mathcal{K} \rightarrow [0, 1], K \mapsto P(X_l = 1 | K).
\]
The axioms of local independence and isotonicity are obviously defined.

Let \( S \) be a surmise relation on \( Q \). Let \( \mathcal{K}_S \) be the quasi-ordinal knowledge space derived from it according to Birkhoff’s theorem (e.g., Doignon and Falmagne, 1999, Theorem 1.49). The axiom of invariant item ordering states that the IRFs \( P(X_l = 1 | .) \) can be ordered such that
\[
\forall K \in \mathcal{K}_S : P(X_{l_2} = 1 | K) \leq P(X_{l_1} = 1 | K)
\]
for any \((I_{l_1}, I_{l_2}) \in S \) \((1 \leq l_1, l_2 \leq m)\).

**Properties.** The properties of MLR, SOM, and SOL can be formulated for \((\mathcal{K}, \subseteq)\). For instance, the stochastic ordering of the latent ‘trait’ \( K \in \mathcal{K} \) by \( X_+ \) (SOL) means that, for any \( K_0 \in \mathcal{K} \),
\[
P(K \supseteq K_0 | X_+ = x_+)
\]
is a nondecreasing function of \( 0 \leq x_+ \leq m \).

As presented in Ünlü (2008), the fundamental result on MLR of the total score variable in unidimensional IRT is extended to quasi-ordered latent trait spaces, including, as special cases, partially ordered knowledge structures. In particular, for \((\mathcal{K}, \subseteq)\), under the axioms of local independence and isotonicity, the total score variable has MLR in the (discrete-dimensional) latent trait \( K \in \mathcal{K} \).
The generalized MLR property implies the generalized SOM property, but may fail to imply the generalized SOL property. The reason for this is the order-theoretic completeness property. Conditions can be specified under which the MLR property implies the SOL property, in the framework of the Mokken-type nonparametric KST formulation.

1.4 Parametric Versus Nonparametric KST

The nonparametric KST axioms and properties are compared with the assumptions underlying the parametric basic local independence model (BLIM; Doignon and Falmagne, 1999, pp. 144–145). The BLIM satisfies the axiom of local independence by definition. Since the BLIM assumes item-specific, state-independent careless error and lucky guess probabilities, respectively, $\beta_l$ and $\eta_l$, at any item $I_l \in Q$, the IRF of an item $I_l \in Q$ is (as a function of $K \in \mathcal{K}$)

$$P(X_l = 1|K) = \begin{cases} 1 - \beta_l & : \text{if } I_l \in K, \\ \eta_l & : \text{if } I_l \notin K. \end{cases}$$

A characterization of the axiom of isotonicity under the BLIM is as follows.

1.4.1 Theorem. Let $\mathcal{K}$ be a knowledge structure on $Q$. In general, a set $Q$ of BLIM IRFs does not satisfy the axiom of isotonicity. A set $Q$ of BLIM IRFs satisfies the axiom of isotonicity if, and only if, $\eta_l \leq 1 - \beta_l$ for any $I_l \in Q$ ($1 \leq l \leq m$).

Proof. See Ünlü (2007, Theorem 6) \hfill \Box

The axiom of invariant item ordering can be characterized as follows.

1.4.2 Theorem. Let $S$ be a surmise relation on $Q$, and let $\mathcal{K}$ be the corresponding quasi ordinal knowledge space. In general, a set $Q$ of BLIM IRFs does not satisfy the axiom of invariant item ordering. A set $Q$ of BLIM IRFs exhibits an invariant item ordering if, and only if, for any $(I_{l_1}, I_{l_2}) \in S$ ($1 \leq l_1, l_2 \leq m$),

$$\beta_{l_1} \leq \beta_{l_2},$$
$$\eta_{l_1} \geq \eta_{l_2},$$
and if \((I_{l_2}, I_{l_1}) \not\in S\), in addition,

\[ 1 - \beta_l \geq \eta_{l_2}. \]

**Proof.** See Ünlü (2007, Theorem 7)

A set of BLIM IRFs may not possess the SOL property (for a counterexample, see Ünlü, 2007, p. 397). Under a knowledge structure, the property of MLR does not in general imply the SOL property; even in case of a restrictive set of parametric BLIM IRFs satisfying the axioms of local independence and isotonicity. However, simulations demonstrate that violations of the SOL property occur only for extreme (unrealistic) values for some of the BLIM parameters; for non-extreme and thus practical parameter vectors the BLIM seems to satisfy the property of SOL. Therefore, if at all of interest, it is necessary to check for the SOL property in any fitted BLIM.

### 1.5 Conclusion

Statistical and probabilistic contributions to KST are presented generalizing the theory of knowledge spaces in parametric as well as nonparametric directions (Ünlü, 2006, 2007, 2008).

The proposed nonparametric Mokken-type formulation in KST is new. It must be further elaborated in research, as a necessary prerequisite for the development of a superior probabilistic test theory, with corresponding statistical inference methodology. Such a theory could include most of the existing IRT and KST models as special cases. For example, the elaboration of a Mokken-type scale analysis for the surmise relation or even surmise system model would be an important contribution.
A. Boomsma, M.A.J. Van Duijn, and T.A.B. Snijders, editors. 


Index

basic local independence model, 5, 6
Birkhoff’s theorem, 4
Birkhoff, G., 4
BLIM, see basic local independence model
Boomsma, A., 1
careless error, 5, 6
characterization
  invariant item ordering, 5, 6
  isotonicity, 5
completeness, 5
Cosyn, E., 1
discrete-dimensional, 4
DMM, see double monotonicity model
Doble, C., 1
Doignon, J.-P., 1, 4
double monotonicity model, 2, 3
Falmagne, J.-Cl., 1, 4
Fischer, G.H., 1
fundamental result, 3, 4
generalized normal ogive, 2
Ghurye, S.G., 3
Grayson, D.A., 3
Guttman model, 1
Guttman, L., 1
Hambleton, R.K., 1
Hemker, B.T., 3
Huyhn, H., 3
invariant item ordering, 3, 4
IRF, see item response function
IRT, see item response theory
isotonic, 4
isotonicity, 4, 6
item response function, 3–6
item response theory, 1, 4, 6
  nonparametric, 2, 4
  parametric, 2
item score, 3
knowledge space
  quasi-ordinal, 4, 6
knowledge space theory, 1, 6
knowledge structure, 4
KST, see knowledge space theory
latent trait, 3, 4
local independence, 3–6
logistic, 2
lucky guess, 5, 6
measurement model, 2, 3
MHM, see monotone homogeneity model
MLR, see monotone likelihood ratio
Mokken, R.J., 2, 3, 5, 6
Molenaar, I.W., 1, 2
monotone homogeneity model, 2, 3
monotone likelihood ratio, 2–6
monotonicity, 3
nondecreasing, 3
Index

order restriction
  invariant item ordering, 6
  isotonicity, 5

partially ordered, 4

quasi-ordered, 4

Sijtsma, K., 2, 3
Snijders, T.A.B., 1
SO, see stochastic ordering
SOL, see stochastic ordering
SOM, see stochastic ordering
Stefanutti, L., 2
stochastic ordering, 2, 3
  latent trait, 3–6

  total score, 3–5
  surmise relation, 4, 6
  surmise system, 6

  Thiéry, N., 1
  total score, 3, 4

  unidimensional, 3, 4
  unidimensionality, 3
  Unlu, A., 1–6
  Uzun, H.B., 1

  Van der Ark, L.A., 3
  Van der Linden, W.J., 1
  Van Duijn, M.A.J., 1
  Wallace, D.L., 3
Item response theory. Questionnaire of God Representations: multi-dimensional operationalization from a relational perspective. Aims of the study. The Dutch Questionnaire of God Representations (QGR) was investigated by means of item response theory (IRT) modeling in a clinical (n = 329) and a nonclinical sample (n = 792). Through a graded response model and IRT-based differential functioning techniques, detailed item-level analyses and information about measurement invariance between the clinical and nonclinical sample were obtained. On the basis of the results of the IRT analyses, a shortened version of the QGR (S-QGR) was constructed, consisting of 22 items, which functions in the same way in both the clinical and the nonclinical sample. Knowledge Space Theory applies concepts from Combinatorics and stochastic processes to the modeling and empirical description of particular fields of knowledge. Within this theory, a mathematical language has been developed to delineate the ways in which particular elements of knowledge (concepts in Algebra, for example) can be gathered to form distinct knowledge states of individuals. This framework enables the creation of computer algorithms for the construction and application of discipline-specific knowledge structures (known as "Knowledge Spaces"). For example, Algebra 1 is rega