Journal of Mathematical Psychology 75 (2016) 110–117

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Regular choice systems: A general technique to represent them by random variables

Reinhard Suck

University of Osnabrück, FB 8, D–49069 Osnabrück, FRG

HIGHLIGHTS

- The possibility of random utility representations for incomplete regular choice systems is explored.
- The theory of convex polytopes is utilized.
- Conditions similar to the Block/Marschak conditions for a complete choice system are derived.
- The proposed technique depends on the Möbius function and Möbius inversion.

ARTICLE INFO

Article history:
Available online 7 May 2016

Keywords:
Random utility representations
Incomplete regular choice systems
Ordering polytopes
Block–Marschak conditions
Möbius function

ABSTRACT

Regular choice systems and their random utility representations are investigated. A generalization of the derivation of the Block–Marschak conditions, based on the Möbius function of a partial order is presented. The technique is demonstrated in connection with two examples. The first is similar to complete choice data. In the second example a complete characterization of the ensuing polytope is obtained including a procedure to explicitly derive a convex representation of a data matrix if it is in the polytope.

1. Introduction

When studying human choice behavior it is obvious that people are rarely consistent. If a decision can be in favor of alternative x from a set of possible alternatives it often can be observed that at another occasion which is quite similar to the first situation alternative y is chosen with $y \neq x$. The reasons for this behavior can be manyfold. The alternatives or stimuli can be almost indistinguishable for the decider, or their value (utility) can vary within the time elapsed between the decisions. This fact must not be regarded as a nuisance or an error of the subject. Rather it can be utilized when it comes to assigning numerical values to the utilities or other attributes of the alternatives. Luce (1959) presents an approach to this enterprise. Here, the probability of choosing x is made proportional to the scale value assigned to x. Clearly, this probability is not directly observable. However, it can be estimated from paired comparison data, given the so called 'choice axiom', (cf. Luce, 1959). The famous BTL-model can be derived in this way, (Bradley & Terry, 1952 and Zermelo, 1929 exploited this idea before Duncan Luce).

The random utility approach – going back to Thurstone (1927) – has a different starting point. Instead of assigning numbers to the stimuli one associates random variable to them. These random variables are thought of as reflecting the variability in the utility a stimulus has for a subject or within a population of subjects. Once this model is accepted the expectation of the random variable is a natural candidate for a scale value. In the heyday of this model a lot of effort was put into estimating the parameters of the respective random variables, mostly under the assumption of normal distributions. Torgerson (1958) is but one example of a comprehensive account of this approach.

Data which can be utilized to carry through this program mostly arise in paired comparison experiments. In this context one often speaks of binary choice data. But there are other experimental paradigms. We mention a few in later sections. A more modern treatment of random utility focuses on necessary and sufficient conditions data must fulfill to be representable by random variables. The older literature focused on algorithms to estimate the expected values under normality conditions. It came as a surprise that even without distributional assumptions data must satisfy severe and testable restrictions to render this approach feasible. In the case of binary choice data these investigations lead to very difficult mathematical problems. However, it seems that these difficulties are not only restricted to binary choice. It took some time to recognize the geometric flavor of this way of thinking. The first seems to be Cohen and Falmagne (1990), who formulated this program (the paper was written in the

E-mail address: rsuck@uos.de.

http://dx.doi.org/10.1016/j.jmp.2016.04.003
0022-2496/

late seventies long before its publication cf. Marley, 1990). Suck (1992) brought the full power of the theory of polytopes to bear on this problem. Since then several related problems have been discussed along these lines, see e.g., approval voting (Doignon & Regenwetter, 1997), interval orders, Regenwetter and Marley (2001), best worst choice (Marley & Louviere, 2005) and many more. Marley (1990) and Suppes, Krantz, Luce, and Tversky (1989, ch. 17) give comprehensive overviews of the state of the art before the geometric impact was fully recognized, while Fishburn's (1992) overview focuses on the geometry of the (binary) choice problem. Some results can be found in Koppen (1995). The geometric description of the so called 'linear ordering polytope' was already an issue in other branches of mathematics such as polyhedral combinatorics. In this literature the random utility aspect is not in the focus. These papers focus on applications to optimization problems.

The literature on random utility in general is scattered over very different special fields such as economics (McFadden, 1981), decision theory (Luce, 2000), psychology (Colonius, 1984), and others. Block and Marschak (1960) and Marschak (1960) seem to have initiated the search for necessary and sufficient conditions which eventually lead to the geometric formulation now prevailing.

Various generalizations have been regarded, see Niederée and Heyer (1997). Suck (2002) investigates the construction of random variables and addresses the question of characterizing binary choice data which allow the construction of independent representing variables. Furthermore, instead of forced choices, (which means the decider must choose exactly one alternative from an actual choice set) other means of data collection have been used. Suck (2005) deals with the random utility aspects of categorical judgments which are in practice a lot more parsimonious.

In Section 2 we introduce the basic concepts of choice systems and give a few definitions from geometry, in particular from the theory of polytopes. Section 3 deals with the derivation of distributions with the help of a partial order associated with an incomplete choice system. From this we are able to formulate in Section 4 analogues of the Block–Marschak conditions. This is however only an intermediate step in finding the final random utility representation. Therefore, in Section 5 a further step is described which finally provides the desired result. In Section 6 the whole procedure is applied to two examples.

2. Basic properties of regular choice systems

The structure we are investigating in this paper consists of a finite set A of stimuli, or objects, commodities, items of a questionnaire, situations, events which are more or less desirable, more or less easy to perceive, remember, recognize, or more or less likely to happen, etc. They have these features with respect to the members of some population or the subjects of an experiment. These persons are presented with some subset $X \subseteq A$ and choose $x \in X$ with probability $p(x, X)$. Usually the choice set X is varied within a particular investigation over some subset $\mathcal{X}$ of the power set of A, i.e., $\mathcal{X} \subseteq 2^A$. Furthermore, in experiments of this kind often the subjects are forced to choose exactly one element of the set X.

In the context of the present investigation we also assume that the probability for choosing x decreases when the choice set is extended. Such choice systems are called regular. There may be rare cases of violations of this property. They are not considered here.

We summarize these properties in the following definition.

Definition 1. A system $(A, \mathcal{X}, p)$ is called a regular choice system if A is a nonempty finite set, $\mathcal{X}$ a nonempty subset of subsets of A with at least two elements and $p : A \times \mathcal{X} \to [0, 1]$ satisfying

(i) $\sum_{x \in X} p(x, X) = 1$
(ii) If $x \in X \subseteq Y$, then $p(x, X) \geq p(x, Y)$.

$\mathcal{X}$ is called the set of choice sets. When $\mathcal{X}$ consists of all subsets with at least two elements, $(A, \mathcal{X}, p)$ is called complete, otherwise incomplete. In the 'binary' case $\mathcal{X}$ consists of the two-element subsets of A.

An experiment utilizing regular choice systems consists of determining or estimating the probabilities $p(x, X)$ for all $X \in \mathcal{X}$. In practice the choice sets $\mathcal{X}$ are in most cases the two element subsets of A. This technique of data collection is known under the name 'paired comparisons'. One way of evaluating such data replaces the stimuli $a \in A$ by random variables U_a with the property

$$p(a, \{a, b\}) = P(U_a > U_b). \quad (1)$$

This equation is called the 'binary case' of a random utility representation. In analogy to the binary case a random utility representation for an arbitrary choice system consists of random variables U_a for each $a \in A$ fulfilling

$$p(x, X) = P\left(U_x > \max_{y \in X - \{x\}} \{U_y\}\right) \quad \text{for all } X \in \mathcal{X} \quad (2)$$

instead of Eq. (1).

We summarize this approach in a formal definition:

Definition 2. Given a regular choice system $(A, \mathcal{X}, p)$. A set $\{U_a;\ a \in A\}$ of jointly distributed random variables is called a random utility representation (abbreviated RU-representation) when it satisfies Eq. (2).

Although rather general, the requirement of the existence of the random variables imposes severe restrictions on the choice probabilities $p(x, X)$. In other words, there are many data sets which do not admit a random utility representation. It is the aim of many theoretical investigations on random utility theory – including the present one – to formulate conditions which render this approach feasible, i.e., to find necessary and sufficient conditions for a choice system to be RU-representable in the sense of Eq. (2). Nowadays this question is handled by looking for facets of a polytope which is naturally associated with the structure $(A, \mathcal{X}, p)$. The basis for this geometric reinterpretation of choice data is the following theorem essentially due to Block and Marschak (1960):

Theorem 3. *A structure of choice probabilities is RU-representable if and only if there exists a probability Q on the set S_n of rankings (or permutations) of n objects such that for all $X \in \mathcal{X}$ and all $x \in X$*

$$p(x, X) = \sum_{\sigma \in S(x,X)} Q(\sigma),$$

where $S(x, X)$ is the set of rankings with x preceding all elements in $X - \{x\}$.

Proof. The proof for complete choice systems (which is the part due to Block/Marschak) can be found in Suppes et al. (1989). The general case follows immediately when we observe that $(A, \mathcal{X}, p)$ is RU-representable if and only if it can be extended to a complete choice system. ■

An immediate consequence of Theorem 3 is that RU-representability implies regularity. Therefore we focus on regular choice systems.

The characterization of random utility models by the possibility to assign probabilities to the rankings in S_n has produced conditions on the choice probabilities. However, it is by no means easy to find some let alone to find a minimal set of conditions. But a