

On L-decomposability of random orderings

Villő Csiszár

September 1, 2009

Abstract

In the setting of random orderings, we study conditional independence properties related to L-decomposability. We show that if a random ordering satisfies L-decomposability for any labelling of the ranks, then it is quasi-independent, provided the number of alternatives is at least 4, and each ordering has positive probability.

1 Introduction

Many sample surveys are designed and conducted to study the preference of consumers for a set of products. Preference can be expressed by ratings, rankings or choices. In the first case, each consumer is required to give a liking score to each of the products. However, consumers are often not able to accurately report their degrees of preferences on a linear scale. This suggests considering non-metric approaches. In the ranking situation, each consumer reports his complete ordering of the products, from “most liked” to “least liked.” If there are a large number of products, this task can still be difficult, and thus can be replaced by more simple decision making tasks. For example, pairs of products are presented to the consumer, according to some design, and he is required to choose the more preferred one from each pair. Alternatively, the consumer may be asked to choose the k best preferred products from the whole set, or to choose the best and the least preferred products.

There is a substantial literature on the methodology of modelling preference data. In this paper, we will be concerned with rankings (equivalently, orderings) of n alternatives. Let the alternatives be labelled by the integers $1, \dots, n$. Then an ordering π is a permutation of these integers, where $\pi(k)$ is the k th best preferred alternative. Equivalently, π^{-1} is the corresponding ranking, i.e. $\pi^{-1}(i)$ is the rank given to alternative i . The simplest probability model on rankings is the uniform distribution, but for most real data, it is inadequate. Therefore, it is necessary to find meaningful non-uniform ranking models. Many approaches of constructing such models have been extensively explored in the literature (Critchlow et al. (1991); Doignon et al. (2004); Fligner and Verducci (1986, 1988, 1993); Mallows (1957); Marden (1995)).

Our aim in this paper is to explore some ranking models based on conditional independence relations. Such models are very popular in modelling factorial

data, where they are called graphical models, see e.g. Lauritzen (1996). The success of graphical models motivates the study of similar models for ordering data.

In Section 2, we review L-decomposability (also called Luce-decomposability), introduced by Critchlow et al. (1991), and show that this property is equivalent to the following set of conditional independence relations: For all k , given the set of alternatives receiving the first k ranks, the ordering of these alternatives and the ordering of the remaining alternatives is independent. We briefly discuss a related property, bi-decomposability, which we introduced in Csiszár (2008).

In Section 3, we define total L-decomposability (or TL-decomposability), which is a natural extension of L-decomposability. A random ordering is TL-decomposable, if for all subsets C of the ranks, given the set of alternatives receiving the ranks in C , the ordering of these alternatives and the ordering of the remaining alternatives is independent. We show that for $n \geq 4$, a TL-decomposable random ordering (provided that each ordering has positive probability) is quasi-independent, i.e the probability of ordering π is $\prod_{k=1}^n c_k(\pi(k))$ for some parameters $c_k(x) > 0$. The quasi-independent model is attractive due to its simple analytic form and convenient number of parameters. Moreover, the maximum likelihood estimate of the parameters is readily obtained by iterative scaling. However, quasi-independence is not directly interpretable (it is “like independence, but not quite”). Our result gives a direct interpretation of quasi-independence in terms of conditional independence relations.

The assumptions of L-, bi- and TL-decomposability are quite strong. However, as we show in Section 4, many models considered in the literature of random orderings have these decomposability properties. Studying these properties therefore helps us to better understand the assumptions implied by the widely used ordering models.

2 L-decomposability and bi-decomposability

Let S_n stand for the set of all orderings π of the alternatives $[n] = \{1, \dots, n\}$. We denote a probability distribution on S_n by $p = (p(\pi) : \pi \in S_n)$. Denote by Π a random ordering with distribution p , that is $P(\Pi = \pi) = p(\pi)$. How can we develop conditional independence models for random orderings? Since the coordinates of a random ordering Π must be distinct, no two subsets of coordinates can be independent, or even conditionally independent on the values of a third subset of coordinates. However, they can be (conditionally) independent, if we condition their values on belonging to disjoint sets. For example, take random orderings Π of $[4]$. It is possible that $(\Pi(1), \Pi(2))$ is independent of $(\Pi(3), \Pi(4))$, conditional on the event that $\{\Pi(1), \Pi(2)\} = \{2, 4\}$ and $\{\Pi(3), \Pi(4)\} = \{1, 3\}$.

First we introduce L-decomposability. A random ordering is L-decomposable, if it satisfies Luce’s ranking postulate given in Luce (1959). This postulates that the ordering is created in n steps: in the k th step, the “judge” chooses $\pi(k)$ as his best preferred alternative from the set of remaining alternatives.

Definition 1. Let Π be a random ordering with probability distribution p on S_n . We call Π or p *L-decomposable*, if there are choice probabilities $p_C(x)$, for all $C \subseteq [n]$ and $x \in C$, such that

$$P(\Pi = \pi) = p(\pi) = \prod_{k=1}^n p_{\{\pi(k), \dots, \pi(n)\}}(\pi(k)) \quad \forall \pi \in S_n.$$

The choice probability $p_C(x)$ is the probability that alternative x is chosen as best preferred from the set C of alternatives. It is easy to show that Π is L-decomposable, if and only if the random sets $X_k = \{\Pi(1), \dots, \Pi(k)\}$ form a Markov chain for $k = 1, \dots, n$, that is

$$P(X_k = C_k | X_1 = C_1, \dots, X_{k-1} = C_{k-1}) = P(X_k = C_k | X_{k-1} = C_{k-1})$$

for every k and for all sets $C_i \subseteq [n]$ such that the lefthandside is defined. The fact that the random sets X_k form a Markov chain is equivalent to the independence of the “past” and the “future”, conditional on the “present”. This means that the random ordering Π is L-decomposable, if and only if for all k , the coordinates $(\Pi(1), \dots, \Pi(k))$ are independent of the coordinates $(\Pi(k+1), \dots, \Pi(n))$, conditional on the set of coordinates $\{\Pi(1), \dots, \Pi(k)\}$. See Critchlow et al. (1991) or Csiszár (2008) for more details.

An L-decomposable random ordering is parametrized by the choice probabilities $p_C(x)$, which can be arbitrary non-negative numbers such that for all $C \subseteq [n]$, their sum is $\sum_{x \in C} p_C(x) = 1$. Thus $\sum_{i=1}^n \binom{n}{i} (i-1) = 2^n(n/2 - 1) + 1$ choice probabilities can be chosen freely. So if we think of the family of all L-decomposable random orderings as a model, then it has many parameters. In Csiszár (2008) we studied those L-decomposable random orderings Π , for which the corresponding random ranking Π^{-1} is also L-decomposable. We called these random orderings bi-decomposable, and showed that they can be parametrized by only $\sum_{k=1}^{n-1} k^2$ parameters. In the next section, we consider an even smaller family of random orderings, the TL-decomposable random orderings.

3 Total L-decomposability

Let $\sigma \in S_n$ be an arbitrary permutation, and let p be a distribution on S_n . Define a new distribution p_σ by

$$p_\sigma(\pi) = p(\pi\sigma) \quad \pi \in S_n.$$

Definition 2. Let Π be a random ordering with probability distribution p on S_n . We call Π or p *totally L-decomposable* (or *TL-decomposable*), if p_σ is L-decomposable for all $\sigma \in S_n$.

An equivalent formulation is the following. A random ordering Π is TL-decomposable, if and only if for every $C \subset [n]$, the coordinates $(\Pi(k) : k \in C)$ and $(\Pi(k) : k \notin C)$ are conditionally independent, given the set of coordinates

$\{\Pi(k) : k \in C\}$. Thus we defined TL-decomposability in terms of conditional independence relations. We also know that for a TL-decomposable random ordering Π , there are choice probabilities $p_C^\sigma(x)$ such that for all $\sigma \in S_n$,

$$P(\Pi = \pi) = p(\pi) = \prod_{k=1}^n p_{\{\pi(\sigma(k)), \dots, \pi(\sigma(n))\}}^\sigma(\pi(\sigma(k))) \quad \forall \pi \in S_n.$$

How many of the choice probabilities $p_C^\sigma(x)$ can be chosen freely, i.e. how many free parameters does a TL-decomposable distribution have? In the next theorem, we answer this question for strictly positive TL-decomposable distributions.

Theorem 1. *If $n \leq 3$, then all distributions on S_n are TL-decomposable. If $n \geq 4$, then a strictly positive distribution p on S_n is TL-decomposable, if and only if it is quasi-independent, i.e. there exist constants $c_k(x) > 0$ (where $k, x = 1, \dots, n$) such that $p(\pi) = \prod_{k=1}^n c_k(\pi(k))$ for all $\pi \in S_n$.*

Thus, for $n \geq 4$, a strictly positive TL-decomposable distribution has $(n-1)^2$ free parameters, since in the above theorem, we can take $c_n(x) = c_k(n) = 1$ for all k and x , without loss of generality. Before proving the theorem, a few more remarks. For $n \leq 2$, all distributions on S_n are TL-decomposable, as well as quasi-independent. For $n = 3$, every distribution on S_3 is still TL-decomposable, however, it is well-known that a distribution p on S_3 is quasi-independent if and only if $p(123)p(231)p(312) = p(132)p(321)p(213)$. From this it also follows that the $n \geq 4$ case of Theorem 1 is false without the requirement of strict positivity, as the next example shows.

Example 1. Let $n \geq 4$, and let q denote any distribution on S_3 . Fix an arbitrary permutation σ of the numbers $\{4, \dots, n\}$. Define a probability distribution p on S_n by $p(\pi) = q(\rho)$ if $\pi = (\rho, \sigma)$, and $p(\pi) = 0$ otherwise, where (ρ, σ) denotes the concatenation of the two permutations. Then p is TL-decomposable, but not quasi-independent, unless q is.

We will need one more proposition. Let p be any distribution on S_n , and $\sigma \in S_n$ arbitrary. Then one can define the new distributions p_σ , p^σ and p' as

$$p_\sigma(\pi) = p(\pi\sigma), \quad p^\sigma(\pi) = p(\sigma\pi), \quad p'(\pi) = p(\pi^{-1}), \quad \pi \in S_n. \quad (1)$$

The following proposition is immediate from the definitions, so we omit its proof.

Proposition 1. *If the distribution p is TL-decomposable, then the distributions p_σ , p^σ and p' given by (1) are also TL-decomposable, for all $\sigma \in S_n$.*

Note that it follows that TL-decomposability implies bi-decomposability. Now we prove Theorem 1. The main part of the proof will rely on four lemmas, which we state and prove afterwards.

Proof. The claim about the $n \leq 3$ case is straightforward from Definition 1. Turning to the $n \geq 4$ case, one direction is trivial: a quasi-independent distribution is clearly TL-decomposable. For the other direction, let p be a strictly

positive TL-decomposable distribution on S_n . By Lemma 1, there exist constants $d_{ij}(x, y)$ satisfying (2). By Lemma 2, these constants also satisfy (3). Therefore, by Lemma 3, there exist constants $c_i(x)$ such that (6) holds. Finally, by Lemma 4, the distribution p satisfies (8), so p is indeed quasi-independent. ■ □

Lemma 1. *Let p be a strictly positive TL-decomposable distribution on S_n . Then for $1 \leq i < j \leq n$ and $1 \leq x < y \leq n$ there exist constants $d_{ij}(x, y)$, such that*

$$p(\pi)/p(\sigma) = d_{ij}(x, y), \quad (2)$$

whenever $\pi(i) = x$, $\pi(j) = y$, $\sigma(i) = y$, $\sigma(j) = x$, and $\pi(k) = \sigma(k)$ for $k \neq i, j$.

Proof. Let π, σ be as in the statement of the lemma. We have to show that the quotient $p(\pi)/p(\sigma)$ depends only on i, j, x, y . In the following, we use simplifying notations such as $p_i(x) = P(\Pi(i) = x)$ or $p_{j|i}(y|x) = P(\Pi(j) = y | \Pi(i) = x)$. Using TL-decomposability, we have

$$p(\pi) = p_i(x) \cdot p_{j|i}(y|x) \cdot p_{k_1, \dots, k_{n-2}|i, j}(\pi(k_1), \dots, \pi(k_{n-2}) | \{x, y\}),$$

where k_1, \dots, k_{n-2} is any listing of the set $[n] \setminus \{i, j\}$, and TL-decomposability was used when writing $\{x, y\}$ instead of x, y in the condition of the last conditional probability. Writing $p(\sigma)$ in similar form, we get

$$\frac{p(\pi)}{p(\sigma)} = \frac{p_i(x)p_{j|i}(y|x)}{p_i(y)p_{j|i}(x|y)},$$

proving our lemma. ■ □

Lemma 2. *Let p be a strictly positive TL-decomposable distribution on S_n , where $n \geq 4$. Then the quantities $d_{ij}(x, y)$ defined in (2) satisfy*

$$\begin{aligned} d_{ij}(x, y)d_{jk}(x, y) &= d_{ik}(x, y) \quad \forall i < j < k, x < y \\ d_{ij}(x, y)d_{ij}(y, z) &= d_{ij}(x, z) \quad \forall x < y < z, i < j. \end{aligned} \quad (3)$$

Proof. First we note that by Proposition 1, it suffices to prove

$$d_{12}(1, 2)d_{23}(1, 2) = d_{13}(1, 2), \quad (4)$$

since the other equalities in the first line of (3) follow by right and left multiplications, while the equalities in the second line of (3) follow by inversion. Writing out (4), we have to prove

$$p(123\sigma')p(231\sigma')p(312\sigma') = p(132\sigma')p(321\sigma')p(213\sigma'), \quad (5)$$

where σ' is an arbitrary fixed permutation of the numbers $4, \dots, n$. Introduce the notation $q(\pi) = \log p(\pi)$. The logarithms of TL-decomposable distributions form a linear subspace, characterized by the collection of linear constraints expressing the conditional independence relations. One then has to check that

Table 1: Some relations holding for q .

$q(\mathbf{3412}\sigma)$	+	$q(\mathbf{2143}\sigma)$	-	$q(\mathbf{3142}\sigma)$	-	$q(\mathbf{2413}\sigma)$	=	0
$q(3241\sigma)$	+	$q(1423\sigma)$	-	$q(3421\sigma)$	-	$q(1243\sigma)$	=	0
$q(2431\sigma)$	+	$q(1342\sigma)$	-	$q(2341\sigma)$	-	$q(1432\sigma)$	=	0
$q(\mathbf{4312}\sigma)$	+	$q(\mathbf{1243}\sigma)$	-	$q(\mathbf{1342}\sigma)$	-	$q(\mathbf{4213}\sigma)$	=	0
$q(2341\sigma)$	+	$q(4123\sigma)$	-	$q(4321\sigma)$	-	$q(2143\sigma)$	=	0
$q(4231\sigma)$	+	$q(3142\sigma)$	-	$q(3241\sigma)$	-	$q(4132\sigma)$	=	0
$q(\mathbf{1432}\sigma)$	+	$q(\mathbf{4123}\sigma)$	-	$q(\mathbf{4132}\sigma)$	-	$q(\mathbf{1423}\sigma)$	=	0
$q(4231\sigma)$	+	$q(2413\sigma)$	-	$q(2431\sigma)$	-	$q(4213\sigma)$	=	0
$q(3421\sigma)$	+	$q(4312\sigma)$	-	$q(4321\sigma)$	-	$q(3412\sigma)$	=	0
$2q(\mathbf{3124}\sigma)$	+	$2q(\mathbf{4213}\sigma)$	-	$2q(\mathbf{3214}\sigma)$	-	$2q(\mathbf{4123}\sigma)$	=	0
$2q(2314\sigma)$	+	$2q(4132\sigma)$	-	$2q(2134\sigma)$	-	$2q(4312\sigma)$	=	0
$2q(1234\sigma)$	+	$2q(4321\sigma)$	-	$2q(1324\sigma)$	-	$2q(4231\sigma)$	=	0

the logarithm of (5) is a consequence of these constraints. Let now $\sigma' = (4, \sigma)$, where σ is any fixed permutation of the numbers $5, \dots, n$. Table 1 contains a set of relations, which hold for q by TL-decomposability. For example, the first three equations follow from the fact that $(\Pi(2), \Pi(3), \Pi(5), \dots)$ are independent of $(\Pi(1), \Pi(4))$ (indicated in bold), given $\{\Pi(1), \Pi(4)\}$. Summing the equalities in Table 1, and dividing by 2, we get

$$q(3124\sigma) + q(2314\sigma) + q(1234\sigma) - q(3214\sigma) - q(2134\sigma) - q(1324\sigma) = 0$$

as required. ■ □

Lemma 3. *Let the positive numbers $d_{ij}(x, y)$ be defined for $1 \leq i < j \leq n$ and $1 \leq x < y \leq n$, and suppose that (3) holds. Then there exist constants $c_i(x)$ for $1 \leq i, x \leq n$ such that*

$$d_{ij}(x, y) = \frac{c_i(x)c_j(y)}{c_i(y)c_j(x)}. \quad (6)$$

Proof. Define $c_i(x) = 1$ if $\min(i, x) = 1$, and $c_i(x) = d_{1i}(1, x)$ if $i, x > 1$. It is easy to check that with this definition, (6) holds. ■ □

Lemma 4. *Let p be a strictly positive distribution on S_n . Suppose there exist constants $c_i(x)$ for $1 \leq i, x \leq n$ such that*

$$\frac{p(\pi)}{p(\sigma)} = \frac{c_i(x)c_j(y)}{c_i(y)c_j(x)} \quad (7)$$

whenever $\pi(i) = x, \pi(j) = y, \sigma(i) = y, \sigma(j) = x$, and $\pi(k) = \sigma(k)$ for $k \neq i, j$. Then

$$p(\pi) = K \prod_{i=1}^n c_i(\pi(i)) \quad \forall \pi \in S_n, \quad (8)$$

with some constant K .

Proof. Let $id = (12 \dots n)$ be the identity permutation, and suppose (7) holds. Starting from id , we can obtain π by a series of transpositions. Each time we remove x from position i , the probability is divided by $c_i(x)$, and each time we move x to position i , the probability is multiplied by $c_i(x)$. These terms cancel each other in the end, except for terms $c_i(i)$ and $c_i(\pi(i))$ such that $\pi(i) \neq i$. Thus we get

$$p(\pi) = \frac{\prod_{i:\pi(i) \neq i} c_i(\pi(i))}{\prod_{i:\pi(i) \neq i} c_i(i)} p(id) = \frac{p(id)}{\prod_{i=1}^n c_i(i)} \prod_{i=1}^n c_i(\pi(i)). \quad \blacksquare$$

□

4 Conclusion

In this paper, we presented three nested models for random orderings, each of which can be defined via conditional independence relations. The largest model is the family of L-decomposable distributions, a subset of which is the family of bi-decomposable distributions, which finally contains the family of totally L-decomposable distributions. As our main result, we showed that for $n \geq 4$, the strictly positive TL-decomposable distributions are exactly the quasi-independent ones.

To end this paper, we briefly summarize the decomposability properties of some popular ordering models. The distance-based models based on the Cayley and Ulam distance are not L-decomposable (for the definition of distances and distance-based models, see Marden (1995)). The Plackett-Luce (Daniels (1950); Thurstone (1927)) and the Babington-Smith (Babington-Smith (1950)) models are L-decomposable, but not bi-decomposable. The distance-based model based on Kendall's τ distance is also L-decomposable, but it is bi-decomposable only for some central orderings π_0 . The multistage ranking model (Fligner and Verducci (1988)) and the repeated insertion model (Doignon et al. (2004)) are bi-decomposable, but not TL-decomposable. The Mallows-Bradley-Terry (Bradley and Terry (1952); Mallows (1957)) model is TL-decomposable. Of the distance-based models, the ones based on the d_p , Hamming, and Hoeffding distances are also TL-decomposable.

References

- Babington-Smith, B. (1950). Discussion of Professor Ross' paper. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **12** 153–162.
- Bradley, R. A. and Terry, M. A. (1952). Rank analysis of incomplete block designs I. *Biometrika* **39** 324–345.

- Chung, L. and Marden, J. I. (1993). Extensions of Mallows' ϕ Model. In *Probability Models and Statistical Analyses for Ranking Data* (Fligner, M.A. and Verducci, J.S, eds.) 108–139. Springer-Verlag, New York.
- Critchlow, D. E., Fligner, M. A. and Verducci, J. S. (1991). Probability models on rankings. *J. Math. Psych.* **35** 294–318.
- Csiszár, V. (2008). Conditional independence relations and log-linear models for random matchings. *Acta Mathematica Hungarica*, Online First.
- Daniels, H. E. (1950) Rank correlation and population models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **12** 171–181.
- Doignon, J.-P., Pekeč, A. and Regenwetter, M. (2004). The repeated insertion model for rankings: missing link between two subset choice models. *Psychometrika* **69** 33–54.
- Fligner, M. A. and Verducci, J. S. (1986). Distance-based ranking models. *J. R. Stat. Soc. Ser. B* **83** 859–869.
- Fligner, M. A. and Verducci, J. S. (1988). Multistage ranking models. *J. Amer. Statist. Assoc.* **83** 892–901.
- Fligner, M. A. and Verducci, J. S. (eds.) (1993). *Probability Models and Statistical Analyses for Ranking Data*, Springer-Verlag, New York.
- Lauritzen, S. (1996). *Graphical Models*, Clarendon Press, Oxford.
- Luce, R. D. (1959). *Individual Choice Behavior*, Wiley, New York.
- Mallows, C. L. (1957). Non-null ranking models I. *Biometrika* **44** 114–130.
- Marden, J. I. (1995). *Analyzing and Modeling Rank Data*, Chapman&Hall, London.
- Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review* **34** 273–286.