On Transitivity Properties of Probability Distributions on Rankings

Björn Haddenhorst [1], **Eyke Hüllermeier** [2], **Martin Kolb** [3]

Abstract. We give a short overview of the notions of *stochastic* and *t-transitivity*, and analyse corresponding transitivity properties of reciprocal pairwise preference relations induced by parametrised probability distributions on the set S_n of rankings/permutations on $\{1, \ldots, n\}$, including the *Plackett-Luce*, the *Babington-Smith*, and the *Mallows model*.

Introduction

The analysis of ranking data has played an important role in various fields of application, such as psychology and the social sciences [8], and more recently also in machine learning [4]. In this regard, the notion of transitivity is of often of specific interest. For example, in the setting of *preference-based multi-armed bandits* [10, 3], consistency assumptions such as transitivity of pairwise comparisons allow for learning rankings of n alternatives efficiently. The aim of this paper is to shed some light on the notion of transitivity in such settings.

We write $[n] := \{1, \ldots, n\}$ and call a map $Q : [n] \times [n] \to [0, 1]$ a *reciprocal relation* on $[n]$ if $q_{i,j} := Q(i, j) = 1 - Q(j, i)$ holds for every $i, j \in [n]$. We denote the set of reciprocal relations on $[n]$ by $\mathcal{Q}_n$ and define different notions of transitivity for these in the subsequent sections. Afterwards, we will in particular focus on the transitivity of reciprocal relations of the type $Q^{\mathbb{P}} := (\mathbb{P}(i \succ j))_{i,j}$ given by

$$\mathbb{P}(i \succ j) := \sum_{\sigma \in S_n : \sigma(i) < \sigma(j)} \mathbb{P}(\sigma)$$

for $i < j$ and $\mathbb{P}(i \succ i) := \frac{1}{2}$, wherein $\mathbb{P}$ is some probability distribution on S_n. Thus, $q_{i,j}$ is the probability that i precedes (is preferred to) j in a ranking randomly drawn from $\mathbb{P}$.

We start by introducing different notions of transitivity in Section 2 and then analyse three parametrised probability distributions with regard to these transitivities in Section 3.

Generalised Transitivity

In this section, we recall two frameworks of generalised transitivity, namely *stochastic transitivity* and *t-transitivity*.

Stochastic Transitivity

Given a symmetric function $g : [\frac{1}{2}, 1]^2 \to [0, 1]$, which is increasing in each component, some $Q \in \mathcal{Q}_n$ is called *g-stochastic transitive* if

$$q_{i,j}, q_{j,k} \geq \frac{1}{2} \quad \Rightarrow \quad q_{i,k} \geq g(q_{i,j}, q_{j,k})$$

holds for all pairwise distinct $i, j, k \in [n]$. We write

$$M_n(g) := \{Q \in \mathcal{Q}_n \,|\, Q \text{ is } g\text{-stochastic transitive}\}.$$

As special types of transitivity we obtain

- *strong stochastic transitivity* with $g_{\mathrm{sst}}(x, y) := \max(x, y)$
- *moderate stochastic transitivity* with $g_{\mathrm{mst}}(x, y) := \min(x, y)$
- *weak stochastic transitivity* with $g_{\mathrm{wst}}(x, y) := \frac{1}{2}$
- *λ-stochastic transitivity* for any $\lambda \in (0, 1)$ with the function $g_{\lambda\text{-st}}(x, y) := \lambda \max(x, y) + (1 - \lambda) \min(x, y)$.

The inequalities $g_{\mathrm{sst}} \geq g_{\lambda\text{-st}} \geq g_{\mathrm{mst}} \geq g_{\mathrm{wst}}$ directly imply

$$M_n(g_{\mathrm{sst}}) \subset M_n(g_{\lambda\text{-st}}) \subset M_n(g_{\mathrm{mst}}) \subset M_n(g_{\mathrm{wst}}).$$

t-transitivity

The notion of *t-transitivity* exists in the context of fuzzy logic: Given a t-norm T, a reciprocal relation $Q \in \mathcal{Q}_n$ is called *T-transitive* if for all distinct $i, j, k \in [n]$ the inequality

$$q_{i,k} \geq T(q_{i,j}, q_{j,k})$$

holds. We write $M_n(T) := \{Q \in \mathcal{Q}_n \,|\, Q \text{ is } T\text{-transitive}\}$. Frequently used t-norms in this setting (cf. [2]) include

- the *minimum t-norm* $T_M(x, y) := \min(x, y)$,
- the *product t-norm* $T_P(x, y) := xy$,
- the *Lukasiewicz t-norm* $T_L(x, y) := \max(0, x + y - 1)$.

For every t-norm T, we have $T \leq T_M$, which directly implies $M_n(T_M) \subset M_n(T)$.

Transitivity of Probability Models on Rankings

In [5] it has already been shown that $Q^{\mathbb{P}}$ is T_L-transitive for every probability distribution $\mathbb{P}$ on S_n. The other types of transitivity mentioned above depend on the particular choice of $\mathbb{P}$. In the following, we recall and analyse the *Plackett-Luce*, the *Babington-Smith*, and the *Mallows* model.

The Plackett-Luce Model

The *Plackett-Luce model* [9, 6] is a family of probability distributions on S_n, which is parametrised by some "skill parameter" $v \in (0, 1)^n$ and defined as

$$\mathbb{P}^{\mathrm{PL}(v)}(\sigma) := \prod_{i=1}^{n} \frac{v_{\sigma(i)}}{v_{\sigma(i)} + \cdots + v_{\sigma(n)}} \tag{1}$$

for $\sigma \in S_n$. An induction over n yields that the marginals of $\mathbb{P}^{\mathrm{PL}(v)}$ are given by

$$\mathbb{P}^{\mathrm{PL}(v)}(i \succ j) = \frac{v_i}{v_i + v_j}.$$

Using this identity we obtain the following result, wherein we use the abbreviation $Q^{\mathrm{PL}(v)} := Q^{\mathbb{P}^{\mathrm{PL}(v)}}$.

[1] Paderborn University, email: bjoernha@mail.uni-paderborn.de
[2] Paderborn University, email: eyke@upb.de
[3] Paderborn University, email: kolb@math.uni-paderborn.de

Proposition 3.1. *(i) For every $v \in (0,\infty)^n$, the relation $Q^{\mathrm{PL}(v)}$ is strongly stochastic transitive.*
(ii) For every $v \in (0,\infty)^n$, the relation $Q^{\mathrm{PL}(v)}$ is T_P-transitive.

In contrast to this, the Plackett-Luce model does not ensure T_M-transitivity in general, as the following example shows.

Example 3.2. *$Q^{\mathrm{PL}((5,2,1))} \in \mathcal{Q}_3$ is not T_M-transitive.*

The Babington-Smith Model

The *Babington-Smith model* [1] is parametrised by $\binom{n}{2}$ parameters $(\pi_{i,j})_{1\le i<j\le n}$, which we may extend to $\pi := (\pi_{i,j})_{1\le i,j\le n}$ with $\pi_{j,i} = 1-\pi_{i,j}$, and assigns each $\sigma \in S_n$ the probability

$$\mathbb{P}^{\mathrm{BS}(\pi)}(\sigma) := \frac{1}{C(\pi)} \prod_{1\le i<j\le n} \pi_{i,j}^{I_{i,j}(\sigma)}(1-\pi_{i,j})^{1-I_{i,j}(\sigma)}, \qquad (2)$$

wherein we use the notation $I_{i,j}(\sigma) := \mathbf{1}_{\{\sigma(i)<\sigma(j)\}}$ and choose $C(\pi)$ such that $\mathbb{P}^{\mathrm{BS}(\pi)}(S_n) = 1$ holds. The transitivity of $Q^{\mathrm{BS}(\pi)} := Q^{\mathbb{P}^{\mathrm{BS}(\pi)}}$ highly depends on these parameters, and the question arises whether transitivity of π might imply some transitivity of $Q^{\mathrm{BS}(\pi)}$. In [5] it has been proven that $\pi \in M_n(g_{\mathrm{sst}})$ suffices to ensure $Q^{\mathrm{BS}(\pi)} \in M_n(g_{\mathrm{sst}})$. The next proposition shows that this result can not be expected to hold for any other of the above mentioned types of stochastic transitivity.

Proposition 3.3. *The mapping $\pi \mapsto Q^{\mathrm{BS}(\pi)}$ does neither preserve moderate nor weak nor λ-stochastic transitivity for any $n \ge 4$. More specifically, for each $n \ge 4$, there is a $\lambda(n) \in (0,1)$ such that λ-stochastic transitivity fails to hold for all $\lambda \le \lambda(n)$.*

Proof. We have

$$\pi = \begin{pmatrix} 0.5 & 0.88 & 0.92 & 9.7 \\ & 0.5 & 0.51 & 0.515 \\ & & 0.5 & 0.55 \\ & & & 0.5 \end{pmatrix} \in M_4(g_{0.125\text{-st}})$$

and a calculation shows $Q^{\mathrm{BS}(\pi)} \notin M_4(g_{\mathrm{wst}})$, which proves the case $n=4$. Due to the continuity of the right-hand side of (2) in the parameters $\pi_{i,j}$, similar counterexamples can be constructed for the case $n \ge 5$. ∎

In the case $n=3$, the distribution $\mathbb{P}^{\mathrm{BS}(\pi)}$ is fully determined by the three parameters $\pi_{1,2}, \pi_{2,3}$ and $\pi_{1,3}$, whence a straight-forward approach including several case distinctions is reasonable and yields the following result.

Proposition 3.4. *For every $\tilde{g} \in \{g_{\mathrm{wst}}, g_{\lambda\text{-st}}, g_{\mathrm{mst}}\}$ and $\pi \in M_3(\tilde{g})$ we have $Q^{\mathrm{BS}(\pi)} \in M_3(\tilde{g})$.*

In the context of t-transitivity, the next example shows that the mapping $\pi \mapsto Q^{\mathrm{BS}(\pi)}$ does neither preserve T_P- nor T_M-transitivity in general.

Example 3.5. *For $\pi^{(1)}$ and $\pi^{(2)}$ given by*

$$\begin{pmatrix} 0.5 & 0.8519 & 0.9812 \\ & 0.5 & 0.9056 \\ & & 0.5 \end{pmatrix} \text{ and } \begin{pmatrix} 0.5 & 0.6495 & 0.8732 \\ & 0.5 & 0.8732 \\ & & 0.5 \end{pmatrix},$$

respectively, we observe $\pi^{(1)} \in M_3(T_P)$ and $Q^{\mathrm{BS}(\pi^{(1)})} \notin M_3(T_P)$ as well as $\pi^{(2)} \in M_3(T_M)$ and $Q^{\mathrm{BS}(\pi^{(2)})} \notin M_3(T_M)$.

The Mallows Model

The *Mallows model* [7] is a two-parameter family of probability distributions $\{\mathbb{P}^{\mathrm{Mal}(\theta,\pi)}\}_{\theta\in(0,1),\pi\in S_n}$ given by

$$\mathbb{P}^{\mathrm{Mal}(\theta,\pi)}(\sigma) = \frac{1}{C(\theta)}\theta^{\Delta(\sigma,\pi)} \text{ for } \sigma \in S_n,$$

in which $\Delta(\sigma,\pi)$ is the Kendall distance (the number of pairwise inversions between σ and π) and $C(\theta)$ a constant chosen such that $\mathbb{P}^{\mathrm{Mal}(\theta,\pi)}(S_n) = 1$ holds. Since the equality

$$\Delta(\sigma \circ \nu, \pi \circ \nu) = \Delta(\sigma, \nu)$$

holds for all $\nu \in S_n$, and every type of transitivity mentioned in Sections 2.1 and 2.2 is invariant under permuting the elements according to ν (i.e., some $Q \in \mathcal{Q}_n$ is transitive if and only if $(q_{\nu(i),\nu(j)})_{1\le i,j\le n}$ is transitive), we may assume $\pi = \mathrm{id}_{S_n}$ without loss of generality and simply write $\mathbb{P}^{\mathrm{Mal}(\theta,n)} := \mathbb{P}^{\mathrm{Mal}(\theta,\mathrm{id}_{S_n})}$ and $Q^{\mathrm{Mal}(\theta,n)} := Q^{\mathbb{P}^{\mathrm{Mal}(\theta,n)}}$. Then we get for $i<j$ that

$$\mathbb{P}^{\mathrm{Mal}(\theta,n)}(i \succ j) = h_\theta(j-i+1) - h_\theta(j-i)$$

holds with $h_\theta(k) := \frac{k}{1-\exp(-k\theta)}$. Hence, $\mathbb{P}^{\mathrm{Mal}(\theta,n)}(i \succ j)$ only depends on $j-i$ but not on i or j itself, that is, the relation $(\mathbb{P}^{\mathrm{Mal}(\theta,n)}(i \succ j))_{1\le i,j\le n}$ has a Toeplitz structure and its entries are increasing in every row. This structure is sufficient to prove the following result.

Proposition 3.6. *For every $\theta \in (0,1)$ we have*

$$Q^{\mathrm{Mal}(\theta,n)} \in M_3(g_{\mathrm{sst}}) \quad \text{and} \quad Q^{\mathrm{Mal}(\theta,n)} \notin M_3(T_M).$$

Conclusion

In this paper, we investigated some transitivity properties of the Plackett-Luce, the Babington-Smith, and the Mallows model. We have not analysed every type of transitivity for every model, for example, we have not verified whether the Mallows model fulfills T_P-transitivity. Filling these gaps is left for future work. Furthermore, it might be interesting to derive concrete conditions on the parameters of the models for fulfilling different types of transitivity.

REFERENCES

[1] B. Babington-Smith, 'Discussion of Professor Ross' paper', *Journal of the Royal Statistical Society B*, **12**, 153–162, (1950).
[2] B. De Baets and H. De Meyer, 'Cycle-transitivity is all around', in *NAFIPS 2008, Annual Meeting of the North American Fuzzy Information Processing Society*, pp. 1–5, (May 2008).
[3] R. Busa-Fekete and E. Hüllermeier, 'A survey of preference-based online learning with bandit algorithms', in *Proc. ALT, 25th International Conference on Algorithmic Learning Theory*, pp. 18–39, Bled, Slovenia, (2014). Springer-Verlag.
[4] Johannes Fürnkranz and Eyke Hüllermeier, *Preference Learning*, Springer Berlin Heidelberg: Berlin, Heidelberg, 2011.
[5] Harry Joe and Joseph S. Verducci, 'On the Babington Smith class of models for rankings', in *Probability Models and Statistical Analyses for Ranking Data*, eds., Michael A. Fligner and Joseph S. Verducci, (1993).
[6] Robert Duncan Luce, *Individual Choice Behavior : A Theoretical Analysis*, 1959.
[7] C. L. Mallows, 'Non-null ranking models. i', *Biometrika*, **44**(1/2), 114–130, (1957).
[8] John I. Marden, *Analyzing and Modeling Rank Data*, Monographs on statistics and applied probability, Chapman and Hall, 1 edn., 1995.
[9] R. L. Plackett, 'The analysis of permutations', *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **24**(2), 193–202, (1975).
[10] Yisong Yue and Thorsten Joachims, 'Beat the mean bandit', in *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, pp. 241–248, USA, (2011). Omnipress.