Learning from Pairwise Comparisons Under Preference Reversals

Abdul Bakey Mir [1] **Arun Rajkumar** [1]

Abstract

We consider the problem of learning to rank from pairwise comparisons in the presence of contextual preference reversals. Preference reversal is a phenomenon well studied in the social psychology and cognitive science literature where users are known to reverse their preference over a pair of alternatives when a carefully chosen third alternative is also presented in the list from which they are required to make a choice. This pertinent effect has been largely ignored in standard representation learning models for pairwise comparisons. In this work, we propose a flexible pairwise comparison model capable of modeling the preference reversal effect. We show that the model is rich enough to capture intransitive preference relations that arise due to reversals. We develop a coupled interpretable neural network based algorithm that learns embeddings for the items from pairwise comparisons. Our network is interpretable as one part of the network learns the standard transitive score based Bradley-Terry-Luce (BTL) Model while the other part explicitly learns the preference reversal effect. We perform experiments to show the efficacy of the proposed network on synthetic datasets against a standard spectral ranking based algorithm and a standard deep network in terms of prediction accuracy on a held-out dataset and the ability of the model to capture intransitive relationships.

1. Introduction

The problem of ranking a set of items from pairwise comparisons is a classical problem that has been studied in many applications. It is used in sports (Cattelan et al., 2013) to predict which team/player is more likely to win the tournament, aggregating social opinions, machine translation for ranking a set of hypothesized translations(Duh & Kirchhoff, 2008), in recommendation systems (Gleich & Lim, 2011) and many others.

Formally, given a set of binary comparisons among n items where for every pair (i, j) that is compared, item i is chosen as the winner with probability $\mathbf{P}_{ij}$ and the goal is to infer a ranking that aggregates the given comparisons. The matrix $\mathbf{P} \in [0, 1]^{n \times n}$ is the underlying probability preference matrix. This matrix is generally unknown. The task of obtaining optimal rankings is NP-Hard for general preference matrices, even in the case when the true underlying preference matrix is known (Charbit et al., 2006). Being an NP-Hard problem for a general $\mathbf{P}$, previous works have tried to impose various parametric assumptions on the preference matrix and then develop algorithms tailor-made for such models.

Popular parametric pairwise comparison models include the BTL Model (Bradley & Terry, 1952; Luce, 1959) and Thurstone Model (Thurstone, 1927) where $\mathbf{P}_{ij}$ solely depends on items i and j without taking into account the rest of the items in the set. However, this may not always be true, and it has been studied that the presence of more items in a set indeed does affect the comparison between two particular items. One such effect is the contextual preference reversal effect, wherein a human choice maker reverses her preference between two items when further options are added to the set (Huber et al., 1982; Huber & Puto, 1983; Pettibone & Wedell, 2007; Soltani et al., 2012).

The general setting considered for a preference reversal involves every item having a two-dimensional embedding $[p, v]$ interpreted as the item having a value v with probability p so that the expected value of the item is pv. There is a taxonomy of contextual preference reversal effects based on the relative positioning of the items on the $p - v$ curve.

Attraction Effect(Howes et al., 2016a) occurs when there is a relative increase in the selection of item A over item B, in comparisons between A and B, despite both items having the same expected value, when a third item C is introduced that is dominated by item A in both attributes but not by item B. Apparently, the presence of item C exerts an attractive influence on the choice of item A.

Similarity Effect(Howes et al., 2016a) occurs when there

[1]Department of Computer Science and Engineering, IIT Madras, India. Correspondence to: Abdul Bakey Mir <abdulbakeymir@smail.iitm.ac.in>.

The Many Facets of Preference Learning Workshop at the International Conference on Machine Learning (ICML), Honolulu, Hawaii, USA, 2023.

is a relative increase in the selection of item B over item A, despite both items having the same expected value, when a third item C, also with the same expected value, is introduced but positioned closer to the item A.

2. Related Work

Preference Reversals: In cognitive science, there have been many attempts to explain the phenomenon of preference reversal from a psychological point of view. (Tversky & Simonson, 1993) proposed an explanation by delineating two psychological processes: a background process that influences decision-making by considering the overall context and a comparison process that accounts for the immediate local context.

An alternative theory explaining preference reversals is provided by Decision Field Theory (Busemeyer & Townsend, 1993), which was initially developed as a framework for explaining the process of deliberation and the gradual formation of preferences over time during decision-making. This model was subsequently extended by (Roe et al., 2001) to account for contextual preference reversals. The model utilizes a connectionist network that progressively aggregates preferences for each option while the decision maker's focus probabilistically transitions between the various options and their respective attributes.

(Bhatia, 2013) introduces the associative accumulation model, a process-oriented framework for decision-making that accounts for several effects, including contextual preference reversals. This model attributes this to different degrees of association between a feature and an option. More precisely, the model proposes that features closely tied to options enhance their accessibility, exert a greater impact on the decision-making process, and subsequently shape the choices made.

(Howes et al., 2016a) explain the preference reversal as a rational computational choice by showing that the choices made are actually the ones that maximize the expected value of a certain computational model when the observations made are noisy.

Learning to Rank from Pairwise Comparisons:

(Negahban et al., 2017a) proposed Rank Centrality to learn the parameters of the BTL Model. It is a spectral algorithm that constructs a Markov chain based on empirical pairwise comparison probabilities and outputs the stationary distribution of this Markov chain as the estimate for the true BTL scores. They show that with $O(n \log n)$ pairs with each pair being compared $O(\log n)$ times, the estmated score is close to the true scores of with high probability. However, the algorithm has guarantees only for BTL Model, which is a stochastically transitive model that too when all the comparisons are faithful to underlying scores.

Several attempts have been made to build parametric intransitive models. (Chen & Joachims, 2016) introduce a parametric preference learning model called the Blade-Chest Model that represents each item as a multi-dimensional vector where different dimensions can be considered as accounting for different aspects of each item. Each item is represented using two n-dimensional vectors called the blade vector and the chest vector. A notion of matchup functions is introduced, which takes in the blade and chest vectors of two items and outputs a number representing the advantage of one item over the other. This number is then turned into pairwise probabilities using the sigmoid function. They perform real-world experimentations to show the model indeed captures intransitivity.

(Makhijani & Ugander, 2019) introduce a simple model called the Majority-Vote model that is capable of exhibiting intransitivity by also inferring a multidimensional embedding for each object. Specifically, their investigation focuses on the 3-dimensional majority vote model and demonstrates its capability in effectively modeling long cycles and arbitrary triplets. Here, the pairwise probability of an item i beating an item j in comparison is the probability that the difference vector of the attributes of the items plus a nose term is positive in at least two dimensions. In addition, they establish the non-concavity of the log-likelihood for any pairwise comparison model that assumes a parametric representation for each item with the power to exhibit an intransitive cycle.

(Rajkumar & Agarwal, 2016a) show that preference matrices that lead to low-rank matrices when transformed using a link function can give to intransitive tournaments with BTL preference matrices being a strict subset of rank 2 preference matrices under logit transformation. Building on this, (Veerathu & Rajkumar, 2021) characterize the rank 2 tournaments in terms of the nature of the tournaments it contains and develop a polynomial time algorithm to obtain optimal rankings for this class of tournaments. In addition, they introduce a pairwise comparison model called the Block-Rank2(BR2) model. This model basically divides items into a certain number of partitions. Within each partition, the preference matrix is rank 2 under logit transformation, and between the partitions, it behaves as a standard BTL model.

The cognitive science literature has aimed to offer psychological explanations for the phenomenon of preference reversal. Rank centrality fails for intransitive models. While some of the models discussed here demonstrate intransitivity, none of them adequately explain the preference reversal phenomenon. In this study, our objective is to develop a pairwise comparison model that provides an explanation for this phenomenon and propose an algorithm for learning this effect from empirical data.

Table 1. Performance for Different Algorithms under Decoy Model (Equation 1)

	PDA	λ	% of recovered 3-Cycles	False 3-Cycles Induced
Our Coupled Network	98	3.25	0.42	29
Rank Centrality (Negahban et al., 2017b)	240	NA	0	0
Standard Deep Network	752	NA	0.17	5094

Table 2. Performance for Different Algorithms under BTL Model

	Uniform Distribution		Gaussian Distribution	
	PDA	λ	PDA	λ
Our Coupled Network	130	0.0015	63	0.026
Rank Centrality (Negahban et al., 2017b)	110	NA	91	NA
Standard Deep Network	224	NA	271	NA

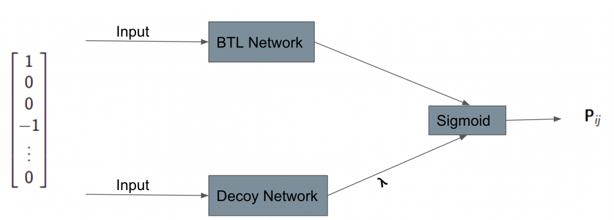

Figure 1. A Coupled Neural Network for learning the preference Model

3. Problem Setting

Given a set of binary comparisons among n items, the task is to infer a ranking that aggregates the given comparisons. As discussed, all the models in the literature assume that whenever a pair of items (i, j) is compared, item i is chosen as the winner with probability $\mathbf{P}_{ij}$ where matrix $\mathbf{P} \in [0, 1]^{n \times n}$ is the underlying probability preference matrix. This setting implicitly assumes that the result of a comparison is only dependent on the two items being compared. But we consider a general setting where the comparison between two items may potentially depend on all other items in the set.

Decoy: An item i is a decoy for item j if $\mathbf{x}_i < \mathbf{x}_j$(Howes et al., 2016a). Here the inequality is to be interpreted coordinate wise.

We propose the following parametric pairwise comparison model to capture the contextual preference reversal effect.

$$\mathbf{P}_{ij} = \sigma\Bigg(\underbrace{\log(S_i/S_j)}_{1} + \underbrace{\sum_{S_k=S_i} \log\left(\frac{||\mathbf{x}_i - \mathbf{x}_k||^2}{||\mathbf{x}_j - \mathbf{x}_k||^2}\right)\mathbb{1}(S_i = S_j)}_{2} + \underbrace{\sum_{\text{k is a decoy for i}} ||\mathbf{x}_i - \mathbf{x}_k||^2}_{3} - \underbrace{\sum_{\text{k is a decoy for j}} ||\mathbf{x}_j - \mathbf{x}_k||^2}_{4}\Bigg) \tag{1}$$

Here each item i has an attribute vector $\mathbf{x}_i \in \mathbb{R}^n$ accounting for different strengths of an item. S_i is the overall score of an item which we define as the product of coordinates of the vector $\mathbf{x}_i$. One can also work with the sum of coordinates as the overall score of an item, but here we use the product interpretation. The first term in the above equation is the standard BTL Model (Bradley & Terry, 1952; Luce, 1959), and the second term in the equation takes into account the similarity effect. The third term counters for the positive attraction effect while as the 4^{th} term handles the negative attraction effect (Howes et al., 2016b).

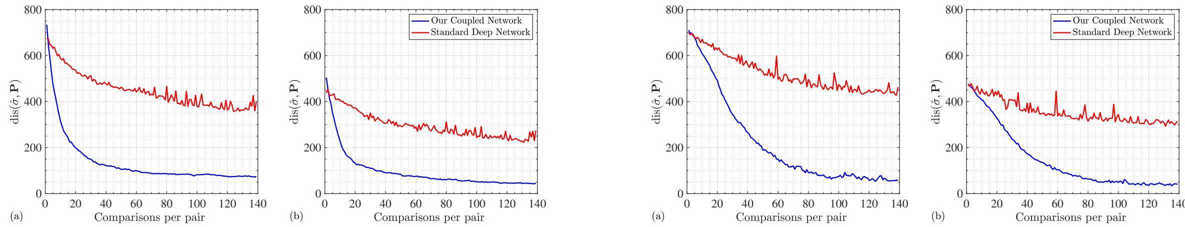

Figure 2. Pairwise Disagreement Error: (**Left**) When the underlying model is BTL ($\mathcal{U}(0,1)$) (**Right**) When the underlying model is Equation 1 (a) Train (b) Validation

3.1. Evaluation Metric

We use the standard Pairwise Disagreement Error (PDA) (Rajkumar et al., 2015) w.r.t underlying preference matrix $\mathbf{P}$ to measure the quality of the ranking $\hat{\sigma}$ outputted by an algorithm defined as follows:

$$\begin{aligned} \text{dis}(\hat{\sigma}, \mathbf{P}) = &\sum_{i<j} \mathbb{1}\Big((i \succ_{\mathbf{P}} j) \wedge (\hat{\sigma}(i) > \hat{\sigma}(j)\Big) \\ &+ \mathbb{1}\Big((j \succ_{\mathbf{P}} i) \wedge (\hat{\sigma}(i) < \hat{\sigma}(j))\Big) \\ &\text{where } i \succ_{\mathbf{P}} j \iff P_{ij} > \frac{1}{2} \end{aligned} \tag{2}$$

The above measure essentially computes the number of pairs where the ranking $\hat{\sigma}$ and the preference matrix $\mathbf{P}$ differ. A pair (i, j) differs in $\hat{\sigma}$ and $\mathbf{P}$ if i is ranked higher than j by $\hat{\sigma}$ but $\mathbf{P}_{ij} < \frac{1}{2}$ or the other way around. The goal of any ranking algorithm is to minimize this. For stochastically transitive models, this can be as small as 0.

In general, the task of identifying an optimal ranking with respect to the above metric is computationally challenging, even when the underlying probability distribution P is known. This problem corresponds to the NP-hard minimum feedback arc set problem. Therefore, achieving an optimal ranking based on the above metric can only be expected under specific restrictive conditions on $\mathbf{P}$. Notably, all the conditions examined in the study by (Rajkumar & Agarwal, 2016b) assumed that the underlying graph is acyclic.

4. Algorithm

We consider a neural network based approach as shown in Figure 1. The input to our network is a pairwise comparison between two items i and j, represented as an n-dimensional vector with 1 and -1 at indices i and j respectively if i is the winner and vice versa. Rest of the co-ordinates of the input vector are 0. The network is split into two components which we call as the BTL Network and Decoy Network. Rather than using a standard deep network for learning the preference model, the idea of using a coupled network is to make the network results interpretable. The BTL network will learn the first term of Equation 1 i.e. the standard BTL part, and the Decoy Network will learn the rest of the Equation 1. The Decoy Network is left flexible to infer from the data whether to give importance to other items while comparing two particular items or not. This is controlled by the λ parameter by which we are weighting the output of the Decoy Network. Here λ is a learnable parameter and the overall network is expected to learn a small lambda when the underlying comparisons depend only on the items being compared and in the case when the comparisons are indeed affected by the presence of other items in the set, the network will learn a significant lambda.

5. Results

The results presented here correspond to the following configurations. In the case of our coupled neural network, the BTL network was implemented as a single-layer fully connected network without a bias term to replicate the BTL Model (Bradley & Terry, 1952; Luce, 1959). On the other hand, the Decoy network was constructed as a two-layer fully connected network. The standard deep network utilized in our experiments consisted of three fully connected layers. Cross-entropy was employed as the loss function for both networks and Stochastic Gradient Descent with a batch size of 8 as the optimizer. The learning rate was set to 0.045 for our coupled network and 0.015 for the standard network. These specific configurations were selected based on the minimal error achieved after hyperparameter tuning.

We evaluated our algorithm through experiments conducted on two synthetic datasets. In the first dataset, the pairwise comparison data was generated following the standard Bradley-Terry-Luce (BTL) model (Bradley & Terry, 1952; Luce, 1959). We considered 2-dimensional embeddings for the items. Both coordinates of the embeddings were generated from a uniform distribution $\mathcal{U}(0,1)$ for one experiment, while for another, they were generated from a normal distribution $\mathcal{N}(0,1)$. To ensure positive scores in the Gaussian case, the exponential function was applied to transform the scores into BTL scores. In the second dataset, data was generated according to the preference reversal model defined by Equation 1 with embeddings from a uniform distribution $\mathcal{U}(0, 1)$. In both cases, the dataset consisted of 100 items. When data was generated based on Equation 1, the parameter λ was set to 5. The coupled network and a standard deep network were trained using a $30 - 20 - 50$ train-validation-test split. This split was chosen to ensure we only $O(n \log n)$ pairs for the training. Both the models were trained for $O(\log n)$ epochs amounting to $O(\log n)$ comparisons per pair. We consider the product interpretation of scores with the score S_i of an item as the product of the coordinates of its embeddings.

As we can see in Figure 2, the coupled network is flexible enough to learn both types of pairwise comparisons with better performance as compared to a standard deep network. A similar trend was seen for Gaussian distribution, and plots have been excluded for redundancy. Also, for the BTL case, the λ value learned by the model was of the order -2, while in the case of Preference Reversal, the λ value was of the order 1. We note that a low value of λ indicates a diminished weight assigned to the Decoy Network, while conversely, a high value of λ signifies an increased weight on the Decoy Network. In addition, we note that the Standard Network was also able to decrease the PDA for the preference model, but it was able to recover only 127, 3-cycles out of 735 cycles compared to 312 by the coupled network. Detailed results are shown in Tables 1 and 2. We note that the input to Rank Centrality were all $\binom{n}{2}$ pairs, and the PDA in Tables 1 and 2 is on the held-out test-set for all algorithms, including Rank Centrality.

6. Conclusion

The relative proportion of selection of an item with respect to another item has been assumed to be independent of the rest of the items in the set in current pairwise comparison models. But there are many effects including the preference reversal effect that defy this assumption. So there is a need for learnable parametric pairwise comparison models that rely on the context of other items as well while comparing a pair of items. In this work, we introduce one such model. We also built a coupled neural network that learns the parameters of this model and is interpretable. We have tested it on synthetic datasets. We were not able to test it on real-world data because of the non-availability of the data.

References

Bhatia, S. Associations and the accumulation of preference. *Psychological review*, 120(3):522, 2013.

Bradley, R. A. and Terry, M. E. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

Busemeyer, J. R. and Townsend, J. T. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review*, 100(3): 432, 1993.

Cattelan, M., Varin, C., and Firth, D. Dynamic bradley–terry modelling of sports tournaments. *Journal of The Royal Statistical Society Series C-applied Statistics*, 62(1):135–150, 2013.

Charbit, P., Thomassé, S., and Yeo, A. The minimum feedback arc set problem is np-hard for tournaments. *Combinatorics, Probability and Computing*, 16:1 – 4, 2006.

Chen, S. and Joachims, T. Modeling intransitivity in matchup and comparison data. In *Proceedings of the ninth acm international conference on web search and data mining*, pp. 227–236, 2016.

Duh, K. and Kirchhoff, K. Learning to rank with partially-labeled data. In *SIGIR '08*, 2008.

Gleich, D. and Lim, L.-H. Rank aggregation via nuclear norm minimization. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 60–68. ACM, 2011.

Howes, A., Warren, P. A., Farmer, G., El-Deredy, W., and Lewis, R. L. Why contextual preference reversals maximize expected value. *Psychological Review*, 123:368 – 391, 2016a.

Howes, A., Warren, P. A., Farmer, G., El-Deredy, W., and Lewis, R. L. Why contextual preference reversals maximize expected value. *Psychological Review*, 123:368 – 391, 2016b.

Huber, J. and Puto, C. Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of consumer research*, 10(1):31–44, 1983.

Huber, J., Payne, J. W., and Puto, C. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, 9 (1):90–98, 1982.

Luce, R. D. *Individual Choice Behaviour a Theoretical Analysis*. Wiley, 1959.

Makhijani, R. and Ugander, J. Parametric models for intransitivity in pairwise rankings. In *The World Wide Web Conference*, pp. 3056–3062. ACM, 2019.

Negahban, S., Oh, S., and Shah, D. Rank centrality: Ranking from pairwise comparisons. *Operations Research*, 65 (1):266–287, 2017a.

Negahban, S., Oh, S., and Shah, D. Rank centrality: Ranking from pairwise comparisons. *Operations Research*, 65 (1):266–287, 2017b.

Pettibone, J. C. and Wedell, D. H. Testing alternative explanations of phantom decoy effects. *Journal of Behavioral Decision Making*, 20(3):323–341, 2007.

Rajkumar, A. and Agarwal, S. When can we rank well from comparisons of $o(n\log(n))$ non-actively chosen pairs? In *COLT*, pp. 1376–1401, 2016a.

Rajkumar, A. and Agarwal, S. When can we rank well from comparisons of $O(n\log(n))$ non-actively chosen pairs? In *Conference on Learning Theory*, pp. 1376–1401. PMLR, 2016b.

Rajkumar, A., Ghoshal, S., Lim, L.-H., and Agarwal, S. Ranking from stochastic pairwise preferences: Recovering condorcet winners and tournament solution sets at the top. In Bach, F. and Blei, D. (eds.), *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 665–673, Lille, France, 07–09 Jul 2015. PMLR.

Roe, R. M., Busemeyer, J. R., and Townsend, J. T. Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological review*, 108(2): 370, 2001.

Soltani, A., De Martino, B., and Camerer, C. A range-normalization model of context-dependent choice: a new model and evidence. *PLoS computational biology*, 8(7): e1002607, 2012.

Thurstone, L. L. A law of comparative judgment. *Psychological Review*, 34(4):273–286, 1927.

Tversky, A. and Simonson, I. Context-dependent preferences. *Management science*, 39(10):1179–1189, 1993.

Veerathu, V. and Rajkumar, A. On the structure of parametric tournaments with application to ranking from pairwise comparisons. *Advances in Neural Information Processing Systems*, 34, 2021.