Recommendation from intransitive pairwise comparisons

Marta Crispino
DEC, Bocconi University
Milano, Italy
marta.crispino@phd.unibocconi.it

Elja Arjas
University of Helsinki and
OCBE, University of Oslo
earjas@mappi.helsinki.fi

Valeria Vitelli
OCBE, University of Oslo
Oslo, Norway
valeria.vitelli@medisin.uio.no

Arnoldo Frigessi
OCBE, University of Oslo
Oslo, Norway
arnoldo.frigessi@medisin.uio.no

ABSTRACT

In this paper we propose a full Bayesian probabilistic method to learn preferences from non-transitive pairwise comparison data. Such lack of transitivity easily arises when the number of pairwise comparisons is large, and they are given sequentially without allowing for consistency check. We develop a Bayesian Mallows model able to handle such data through a latent layer of uncertainty which captures the generation of preference misreporting. We then construct an MCMC algorithm, and test the procedure on simulated data.

1. INTRODUCTION

We consider pairwise preference data of the form "*x is preferred to y*", denoted $x \prec y$, where x and y are some items of interest. The challenge with this kind of data, is that pairwise preferences are not always transitive. For instance the data coming from a single user may contain preferences of the form $\{x \prec y,\ y \prec z,\ z \prec x\}$. Such non-transitivity of preferences arises for many reasons, including the user's inattentiveness, actual ambiguity in the user's preference, multiple users with the same account and varying framing of the data collection. These situations are so common that most pairwise comparison data are in fact non-transitive, thus creating the need for methods able to learn users' preferences from data that lack logical transitivity.

In this paper we assume that non-transitive data arise because users make mistakes, i.e. switch the order between two compared items, either simply by mistake or because the items compared have a rather similar rank for the user. The developed Bayesian methodology provides the posterior distribution of the consensus ranking of a homogeneous group of users, as well as of the estimated individual rankings for each user. The consensus ranking can be seen as a compromise which is formed from the individual pairwise preferences. The estimated individual rankings can be used for personalized recommendation. Inference is based on an adapted version of the Markov Chain Monte Carlo algorithm in [3].

Previous work on pairwise preferences includes the Bradley Terry model [1] (but no personalized preferences are estimated) and, within the framework of Mallows model, [2] and [3], where transitivity is assumed.

RecSys 2016 Poster Proceedings, Boston, MA, USA, 15th-19th September 2016.

2. METHODS

Let N users independently express their preferences between pairs of items in $\mathcal{O} = \{O_1, ..., O_n\}$. We assume that each user j receives a possibly different subset $\mathcal{C}_j = \{\mathcal{C}_{j,1}, ..., \mathcal{C}_{j,M_j}\}$ of pairs of dimension $M_j \leq n(n-1)/2$. Let $\mathcal{B}_j = \{\mathcal{B}_{j,1}, ..., \mathcal{B}_{j,M_j}\}$ be the set of preferences given by user j, where $\mathcal{B}_{j,m}$ is the order she gives to the pair $\mathcal{C}_{j,m}$. For example $\mathcal{B}_{j,m} = \{O_1 \prec O_2\}$ means that O_1 is preferred to O_2 if $\mathcal{C}_{j,m} = (O_1, O_2)$. The considered data are incomplete since not all items are observed by each user. We assume no ties in the data: users are forced to express their preference for all pairs in the list $\mathcal{C}_j$ assigned to them, and indifference is not permitted.

The main assumption is that each user j has a personal latent ranking, $\tilde{\mathbf{R}}_j = (\tilde{R}_{j1}, ..., \tilde{R}_{jn}) \in \mathcal{P}_n$ (the space of n-dimensional permutations), distributed according to the Mallows model

$$\pi(\tilde{\mathbf{R}}_j \mid \alpha, \rho) = \frac{\exp\{-(\alpha/n) d(\tilde{\mathbf{R}}_j, \rho)\} \mathbb{1}_{\mathcal{P}_n}(\tilde{\mathbf{R}}_j)}{Z_n(\alpha)} \quad \forall j \quad .$$

We assume conditional independence between $\tilde{\mathbf{R}}_1, ..., \tilde{\mathbf{R}}_N$ given ρ and α. Here $\rho \in \mathcal{P}_n$ is the shared consensus ranking, $\alpha > 0$ is the scale parameter (describing the concentration around the shared consensus), $d(\cdot,\cdot) : \mathcal{P}_n \times \mathcal{P}_n \to [0, \infty)$ is any right-invariant distance function (e.g. Kendall, Spearman or footrule), and $Z_n(\alpha)$ is the normalizing constant. See [3] for more details.

We model the situation where each user j, when announcing her pairwise preferences, mentally checks where the items under comparison are in her latent ranking $\tilde{\mathbf{R}}_j$. Then, if the user is consistent with $\tilde{\mathbf{R}}_j$, the pairwise orderings in $\mathcal{B}_j$ are induced by $\tilde{\mathbf{R}}_j$ according to the rule: $(O_k \prec O_i) \iff \tilde{R}_{jk} < \tilde{R}_{ji}$. In this case $\mathcal{B}_j$ contains only transitive preferences. However, if the user is not fully consistent with her own latent ranking, the pairwise orderings in $\mathcal{B}_j$ can be non-transitive. In order to deal with this situation, we propose a probabilistic model based on the assumption that non-transitivities are due to mistakes in deriving the pair order from the latent raking $\tilde{\mathbf{R}}_j$. The likelihood assumed for a set of preferences $\mathcal{B}_j$ is

$$\pi(\mathcal{B}_j|\alpha,\rho) = \sum_{\mathbf{R}_j \in \mathcal{P}_n} \pi(\mathcal{B}_j|\tilde{\mathbf{R}}_j = \mathbf{R}_j)\pi(\tilde{\mathbf{R}}_j = \mathbf{R}_j|\,\alpha,\rho),$$

where $\pi(\mathcal{B}_j|\tilde{\mathbf{R}}_j = \mathbf{R}_j)$ is the probability of ordering the pairs in $\mathcal{C}_j$ as in $\mathcal{B}_j$ (possibly generating non-transitivities), when the latent ranking for user j is $\mathbf{R}_j$. It is therefore the probability of making mistakes instead of just following $\mathbf{R}_j$.
The posterior density is then:

$$\pi(\alpha,\rho|\,\mathcal{B}_{1:N}) \propto \pi(\alpha)\pi(\rho)\prod_{j=1}^{N}\left[\sum_{\mathbf{R}_j \in \mathcal{P}_n} \pi(\mathcal{B}_j|\mathbf{R}_j)\pi(\mathbf{R}_j|\,\alpha,\rho)\right],$$

where we assume a gamma prior, $\pi(\alpha)$, for α and a uniform prior on $\mathcal{P}_n$, $\pi(\rho)$, for ρ. We suggest two models for the probability of making a mistake: the Bernoulli model (BM) to handle random mistakes, and the logistic model (LM) for mistakes that are due to difficulty in ordering similar items. In both models we assume that any two pair comparisons made by a user are conditionally independent given her latent ranking: $(\mathcal{B}_{j,m_1} \perp\!\!\!\perp \mathcal{B}_{j,m_2}) \mid \tilde{\mathbf{R}}_j, \forall m_1, m_2 = 1, ...M_j$. In BM we assume the following Bernoulli type model for the probability of making a mistake:

$$\mathbb{P}(\mathcal{B}_{j,m} = -\tilde{\mathcal{B}}_{j,m}(\tilde{\mathbf{R}}_j)\,|\,\theta,\tilde{\mathbf{R}}_j) = \theta\ , \qquad \theta \in [0, 0.5) \quad ,$$

where $-\tilde{\mathcal{B}}_{j,m}(\tilde{\mathbf{R}}_j)$ is the reversed preference order, i.e. a mistake. We assign to θ the truncated Beta distribution on the interval $[0, 0.5)$ as prior. In LM we assume the following logistic type model for the probability of making a mistake:

$$\text{logit}\left(\mathbb{P}\left(\mathcal{B}_{j,m} = -\tilde{\mathcal{B}}_{j,m}(\tilde{\mathbf{R}}_j)\middle|\tilde{\mathbf{R}}_j,\beta_0,\beta_1\right)\right) = \beta_0 + \beta_1 d_{\tilde{\mathbf{R}}_j,m},$$

where $d_{\tilde{\mathbf{R}}_j,m}$ is the footrule or l_1 distance of the individual ranks of the items compared: if $\mathcal{B}_{j,m} = (O_1 \prec O_2)$, $d_{\tilde{\mathbf{R}}_j,m} = |\tilde{R}_{j1} - \tilde{R}_{j2}|$. We assign to β_1 an exponential prior on the negative support[1] and to β_0, conditioned on β_1, an exponential prior on the shifted support $[-\infty, -\beta_1]^2$. These choices are motivated by the fact that we want to model a negative dependence between the distance of the items and the probability of making a mistake. Also, we want to force the probability of making a mistake when the items have ranks differing by 1 to be less than 0.5.

3. EXPERIMENTAL RESULTS

We performed a number of experiments on simulated data, generated according to the model in Section 2 with BM noise, a fixed ρ_{TRUE} and $n = 10$ items. For various values of α we sampled latent rankings $\tilde{\mathbf{R}}_{j,\text{TRUE}}$ from the Mallows density centered at ρ_{TRUE}, using which we generated the individual non-transitive pair comparisons. In order to assess the performance of our methods, in Figure 1 we plotted the posterior CDF of the footrule distance of the estimated consensus ranking and the true generating one, $d_f(\rho,\rho_{\text{TRUE}}) = \sum_{i=1}^{n} |\rho_i - \rho_{i,\text{TRUE}}|$, for varying parameters N, α, θ and λ_M (the parameter of a Poisson from which the number of comparisons assigned to each user is sampled). As expected, the performance of the method improves as the number of users N increases (Figure 1, top right), becomes worse as the probability of doing mistakes θ increases (top left), improves as the dispersion of the individual latent rankings $\tilde{\mathbf{R}}_{j,\text{TRUE}}$ around ρ_{TRUE} decreases, i.e. when α increases, (bottom left), and becomes worse when the number of pairwise comparisons diminishes, i.e. when λ_M decreases, (bottom right). The new method performs generally well also if the number of pair comparisons per user is around one third of all possible pair comparisons ($\lambda_M = 15$). We then studied the precision of the estimated individual preferences by comparing $\tilde{\mathbf{R}}_j$ with $\tilde{\mathbf{R}}_{j,\text{TRUE}}$ in terms of top-3 detection. For each user j, we found the triplet of items $D_3^j = \{O_{i_1}, O_{i_2} O_{i_3}\}$ with the maximal estimated posterior probability of being ranked jointly among the top-3 items. Let H_5^j be the set of 5 items with the 5 highest ranks in $\tilde{\mathbf{R}}_{j,\text{TRUE}}$. We computed the percentage of users for which $D_3^j \subset H_5^j$. This success-percentage was 60% - 85% when the data were generated with $\alpha = 2$ and $M_j = 15\ \forall j$, which is a hard problem. For easier experiments, with larger α (ranging from 3 to 4) and larger M (up to 30), the success-percentages increased to 90%-100%.

[1] $\pi(\beta_1) = \lambda_1 e^{\lambda_1\beta_1}\mathbb{1}(\beta_1 < 0)$, $\lambda_1 > 0$.

[2] $\pi(\beta_0|\beta_1) = \lambda_0 e^{\lambda_0(\beta_0+\beta_1)}\mathbb{1}(\beta_0 < -\beta_1)$, $\lambda_0 > 0$.

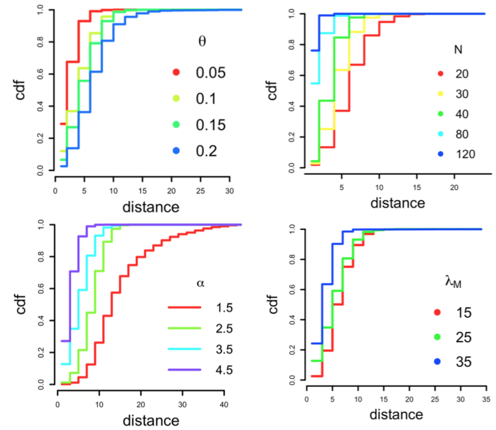

Figure 1: Posterior CDFs of $d_f(\rho,\rho_{\mathbf{TRUE}})$ for varying parameters.

4. CONCLUSIONS

We extended the Bayesian method for learning preferences in [3] to non-transitive pairwise comparison data. We produce estimates of the consensus ranking of all items under considerations, as well as a personal estimated ranking of all items for each user. This can be directly used for personalized recommendations. We also obtain posterior distributions for these preferences, allowing quantification of the uncertainty of recommendations of interest.

5. REFERENCES

[1] F. Caron, A. Doucet, "Efficient Bayesian Inference for generalized Bradley-Terry Models", *Journal of Computational and Graphical Statistics*, 21(1), 2012.

[2] T. Lu and C. Boutilier, "Effective sampling and learning for Mallows models with pairwise preference data", *Journal of Machine Learning Research*, 2014.

[3] V. Vitelli, Ø. Sørensen, A. Frigessi and E. Arjas, "Probabilistic preference learning with the Mallows rank model", *ArXiv e-prints*, 2016.