Bayesian Inference for the Mallows Rank Model

Øystein Sørensen[1], **Valeria Vitelli**[1], **Arnoldo Frigessi**[1], **and Elja Arjas**[1]
[1] Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

A large number of data types have ranks as their natural scale. Students applying to secondary or post-secondary education rank prospective schools or colleges, and are usually admitted based on a ranking of their present results. The manufacturing industry recruits panels to rank novel products, like food with new flavours. Market studies are often based on interviews where competing products are compared and ranked. In the era of big data, analysing preference data (e.g., movie preferences, restaurant rankings) also receives much attention. In addition to these examples, where rankings or preferences are the natural units, converting quantitative data to ranks can lead to more robust inference when absolute scales are not easily comparable, as in the case of genomic data.

Due to the many novel applications, interest in ranks has increased lately. Two of the classical models are the Placket-Luce (Luce, 1959; Placket, 1975), which models the ranking process as sequential pairwise comparisons of items, and the distance based Mallows model (Mallows, 1957), which is an exponential probability model peaked at some underlying true ranking. There is a growing literature on the inference and methodology of these models and their applications. For example, Caron and Teh (2012) do Bayesian inference in the Placket-Luce model with time-dependent preferences, and further develop the framework in Caron et al. (2014), where a Dirichlet process mixture is used to cluster assessors based on their rankings. Concerning the Mallows models, Meila and Chen (2010) perform clustering of assessors into groups with similar preferences. Lu and Boutilier (2011) also consider clustering with the Mallows model, and allow for observations in the form of pairwise comparisons of items.

Remarkably, the work on the Mallows models referenced above has been limited to *one particular form* of the Mallows model which uses the Kendall distance. For this metric, the normalizing constant of the Mallows distribution can be easily computed analytically. This model and its extensions has been studied under the name *Generalized Mallows Model* (Fligner and Verducci, 1986; Meila and Chen, 2010), and has convenient sufficient statistics and conjugacy properties. However, many other metrics, like the footrule and Spearman distance, can be used. Except for computational considerations, there is no reason to assume that these metrics are less useful than the Kendall distance.

In this paper, we develop a Bayesian framework for inference in Mallows models with any right-invariant metric. In any given problem, the normalizing constant is computed offline once and for all, using an efficient importance sampling scheme. Full Bayesian inference is based on an MCMC scheme designed for rank data. In addition, using data augmentation techniques with appropriate proposal distributions, our methods allow for incomplete data, like top-t rankings, pairwise comparisons, and ranks which are missing at random. We also perform clustering of assessors via mixtures, and study ranks which are varying over time. The main advantage of our method is the flexibility with which it handels different metrics, while at the same time performing clustering and allowing for various types of partial data. By being fully Bayesian, all uncertainty in the model is taken into account in the posterior distribution, and relevant prior knowledge can be represented in the posterior distribution.

We present many real world examples which illustrate the use of the model. We study a full season of the Premier League, England's top football league; using each match as a pairwise comparison between two teams, we compute a ranking of the teams, and compare it to the official league table. We study a sushi dataset, and find useful clusters of people with similar taste preferences. Finally, we study test scores in mathematics for a class of Italian high school students over four years, and investigate how each student's rank in the class develops over time.

Keywords: Distance-based ranking model, Mallows model, Pairwise comparison, Ranking.

References:

Caron, F. and Y. W. Teh (2012). Bayesian nonparametric models for ranked data. In *Advances in Neural Information Processing Systems, 25*, 1529–1537.

Caron, F., Y.W. Teh and T.B. Murphy (2014). Bayesian nonparametric Plackett-Luce models for the analysis of preferences for college degree programmes. To appear in *Ann. Appl. Statist.*

Fligner, M.A. and J.S. Verducci (1986). Distance based ranking models. *J.R. Statist. Soc. B, 48,* 359–369.

Lu, T. and C. Boutilier (2011). Learning Mallows models with pairwise preferences. In *Proceedings of the 28th International Conference on Machine Learning,* 145–152.

Luce, R.D. (1959). *Individual choice behavior: A theoretical analysis.* Wiley.

Mallows, C.L. (1957). Non-null ranking models. *Biometrika, 44,* 114–130.

Meila, M. and H. Chen (2010). Dirichlet process mixtures of generalized Mallows models. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artifical Intelligence.*

Placket, R.L. (1975). The analysis of permutations. *Appl. Statist., 24,* 193–202.