A Mallows Model for Coxeter Groups and Buildings

Paige E. Rinker

Abstract

Rank data is comprised of a set of complete and partial rankings which reflect preference or standing (examples of such data can be generated by voters, search engines, market survey responses, etc.). Complete rankings (viewed as bijective maps from a set of n items to n ranks) can be identified with the symmetric group S_n, and this identification gives rise to a variety of analytic techniques.

One such method is a probabilistic soft clustering technique based on the Mallows probability distribution for the symmetric group. This distribution relies on Kendall's tau (a distance on S_n) and is normalized by a sum over the entire space of $n!$ rankings. It can be extended to partial rankings (which differ from complete rankings by allowing multi-way ties at varying ranks), theoretically facilitating the extension of this probabilistic method. Unfortunately, the generalization of the Mallows distribution requires the evaluation of even more complex sums, which quickly become prohibitively large.

In this thesis, we present a closed form of Mallows distribution for partial rankings which allows us to extend the clustering method to accommodate arbitrary sets of rank data. In addition, our combinatorial proof of this closed form allows Mallows model and its accompanying closed forms to be generalized to several natural extensions of the symmetric group, namely the flag variety of the finite general linear group and all finite Coxeter groups.