Bayesian Nonparametric Mixture Modeling with Unimodal Kernels.

Carlos E. Rodríguez and Stephen G. Walker

Within the context of mixture modeling, the normal distribution is typically used as the components distribution. However, if a cluster is skewed or heavy tailed, then the normal distribution will be inefficient and many may be needed to model a single cluster. In this paper, we present an attempt to solve this problem. We define a cluster, in the absence of further information, to be a group of data which can be modeled by a unimodal density function.

Hence, our intention is to use a family of distribution functions, to replace the normal, for which the only constraint is unimodality. With this aim, we device a new family of semi-parametric unimodal distributions, which has big support over the space of unimodal distributions.

The difficult aspect of the Bayesian model is to construct a suitable MCMC algorithm to sample from the correct posterior distribution. The key will be the introduction of strategic latent variables and the use of the Product Space view of Reversible Jump methodology.

Keywords: Cluster, Dirichlet Process, Mixture model, Slice sampler, Product space, Reversible jump and Label switching

*Contact author: cerh2@kent.ac.uk