BayesMix: An R package for Bayesian Mixture Modelling

Bettina Grün[*], Friedrich Leisch
Institut für Statistik und Wahrscheinlichkeitstheorie
Technische Universität Wien
Wiedner Hauptstraße 8-10, 1040 Wien, Österreich
{Bettina.Gruen, Friedrich.Leisch}@ci.tuwien.ac.at

Finite mixture models are a popular method for modeling unobserved heterogeneity as well as for parametric approximations of multimodal distributions. Areas of application are, e.g., biology, medicine, economics and engineering among many others. For ML estimation the EM algorithm is most frequently used which can be done in R, e.g., with the package `mclust`.

Bayesian estimation has become feasible with the advent of Markov Chain Monte Carlo (MCMC) simulation and the R package `BayesMix` provides facilities for estimating univariate Gaussian finite mixtures with MCMC methods. It has been developed as accompanying material to the forthcoming book Frühwirth-Schnatter (2005).

The model class which can be estimated with `BayesMix` is a special case of a graphical model where the nodes and their distributions are fixed and the user only needs to specify the values of the constant nodes, the data and the initial values. Small variations of the model are allowed with respect to the segment specific priors. The MCMC sampling is done by JAGS (Just Another Gibbs Sampler; Plummer, 2003) and its output can be analyzed in R using functionality from the package `coda`. In addition to the visualization of the MCMC chains there are diagnostic plots implemented which can be used for determining the appropriate number of segments or a suitable variable for ordering the segments as in Bayesian mixture modelling it makes in general a difference which constraint is imposed for ordering the segments due to label switching.

`BayesMix` can be seen as a prototype for a special purpose interface to the software JAGS. Its advantage is that a user who "only" wants to estimate finite mixtures of Gaussian distributions can use JAGS as sampling engine, but does not need to know the BUGS syntax which is used by JAGS for specifying general Bayesian hierarchical models. `BayesMix` offers the opportunity to be a starting point for learning the BUGS syntax as the model specifications are written into separate files and can be inspected or modified by the user.

Keywords: Finite mixture models, Bayesian modelling

References

S. Frühwirth-Schnatter. *Bayesian Mixture Modelling.* Springer, 2005. Forthcoming.

M. Plummer. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, and A. Zeileis, editors, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, Technische Universität Wien, Vienna, Austria, 2003. URL `http://www.ci.tuwien.ac.at/Conferences/DSC.html`. ISSN 1609-395X.

[*]supported by the Austrian Science Foundation (FWF) under grant SFB#010 ('Adaptive Information Systems and Modeling in Economics and Management Science').