Summary 1: The Infinite Gaussian Mixture Model

Susan Cherry

This paper introduces the infinite Gaussian mixture model, which removes the problem of determining the correct number of mixture components. The paper begins by describing the finite hierarchical mixture. The finite Gaussian mixture model is introduced, with its means, precisions, and mixing proportions. Stochastic indicator variables are introduced (c_i) for each observation. These indicator variables contain information on which class generated each observation (the indicators are often referred to as missing data in the mixture model context). Next, the priors on the parameters and hyperparameters are specified. The conditional posterior distributions are also derived. The model is then extended to the infinite limit. The conditional posteriors of the infinite limit are obtained by simply replacing k (the number of classes) with k_{rep}. Final derivations of the conditional posteriors are derived for the infinite limit. The model can also easily be generalized to the multivariate case by replacing means and precisions with vectors and matrices.

Then the paper begins to describe inference using the 3 dimensional spirals dataset. The mixture model is started with a single component and then a large number of Gibbs sweeps are performed. This updates the parameters and hyperparameters by sampling from the conditional posterior distributions that were derived earlier in the paper. Finally, the paper discusses the predictive distribution. The predictive distribution has two parts: both the represented classes which are Gaussian and the unrepresented classes. The unrepresented classes are approximated by a finite mixture model of Gaussians with parameters drawn from the prior. Results from the inference and predictive distribution sections can be explored in more detail in Figures 1 and 2 on pages 558 and 559.

In conclusion, the infinite hierarchical Bayesian mixture model is a practical method (with good performance and without overfitting) that can be used on multidimensional data. Tests show that the infinite mixture model produces densities that are competitive with other common methods. The four main advantages of the infinite model over the finite are as follows:

1. For many applications it is not appropriate to limit the number of classes
2. The number of represented classes is automatically determined
3. Using MCMC avoids local minima
4. It is simpler to work with the infinite limit than to handle the finite models with unknown sizes.