Dirichlet or Potts ?

Ali Mohammad-Djafari
Laboratoire des signaux et systmes (L2S)
Suplec, Plateau de Moulon, 3 rue Joliot Curie
91192 Gif-sur-Yvette, France

Abstract

When modeling the distribution of a set of data $\{x_i, i = 1, \cdots, n\}$ by a mixture of Gaussians (MoG), there are two possibilities: i) the classical one is using a set of parameters which are the proportions α_k, the means μ_k and the variances σ_k^2; ii) the second is to consider the proportions α_k as the probabilities of a hidden variable z whith $\alpha_k = P(z = k)$ and assignining a prior law for z. In the first case a usual prior distribution for α_k is the Dirichlet which account for the fact that $\sum_k \alpha_k = 1$. In the second case, to each data x_i we associate a hidden variable z_i. Then, we have two possibilities: either assuming the variables z_i to be i.i.d. or assigning them a Potts distribution. In this paper we give some details on these models and different algorithms used for their simulation and the estimation of their parameters.

More precisely, in the first case, the assumption is that the data are i.i.d samples from $p(x) = \sum_{k=1} \alpha_k \mathcal{N}(\mu_k, \sigma_k^2)$ and the objective is the estimation of $\theta = \{K, (\alpha_k, \mu_k, \sigma_k^2), k = 1, \cdots, K\}$. In the second case, the assumption is that the data x_i is a sample from $p(x_i|z_i = k) = \mathcal{N}(\mu_k, \sigma_k^2), \forall i$ where the z_i can only take the values $k = 1, \cdots, K$. Then if we assume z_i i.i.d., then the two models become equivalent with $\alpha_k = \frac{1}{n} \sum_{i=1}^{n} \delta(z_i - k)$. But if we assume that there some structure in the hidden variables, we can use the Potts model $p(z_i|z_j, j \neq i) \propto \exp\left\{\gamma \sum_{j \in \mathcal{V}(i)} \delta(z_i - z_j)\right\}$ where $\mathcal{V}(i)$ represents the neighboring elements of i, for example $\mathcal{V}(i) = i - 1$ or $\mathcal{V}(i) = \{i - 1, i + 1\}$ or in cases where i represents the index of a pixel in an image, then $\mathcal{V}(i)$ represents the four nearest neigbors of that pixel. γ is the Potts parameter.

These two models have been used in many data classification or image segmentation where the x_i represents either the grey level or the color components of the pixel i and z_i its class labels. The main objective of an image segmentation algorithm is the estimation of z_i. When the hyperparameters K, $\theta = (\alpha_k, \mu_k, \sigma_k^2), k = 1, \cdots, K$ and *gamma* are not known and have also to be estimated, we say that we are in *totally unsupervised* mode, when are known we are in *totally supervised* mode and we say that we are in *partially supervised* mode when some of those hyperparameters are fixed. In the following, we present some of these methods.

Key Words: Mixture of Gaussians, Dirichlet, Potts, Classification, Segmentation.