

Identifying Bayesian Mixture Models (Betancourt, 2017)

Satvik Saha (ss7228)

October 21, 2025

A common problem in Bayesian mixture models is identifiability of components: in a θ -mixture of distributions where the labels k for each component have no effect, there is a fundamental ambiguity in associating parameters α_k with components. This means that permuting the labels of the component parameters α_k to $\alpha_{\sigma(k)}$ has no effect on the model likelihood. If in addition the priors on α and θ are exchangeable with $\pi(\alpha_1, \dots, \alpha_K) = \pi(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(K)})$ and an analogous form for $\pi(\theta)$, the posterior also remains invariant under permutations of labels. These often leads to a multimodal posterior; the space of (α, θ) partitioned into $K!$ regions with effectively the same posterior distribution $\pi(\alpha, \theta | y)$ on them. Algorithms such as Monte-Carlo Markov Chain methods often deal poorly with such multimodal distributions, as chains may get stuck at one mode and fail to explore the others.

One way to resolve this is to use non-exchangeable priors: by forcing each component to have certain characteristics individual to its label, we can break the symmetry in the posterior and restore a unimodal form.

Another option is simply identify the space of (α, θ) up to permutations, ‘gluing’ together all points $(\alpha, \theta) \sim (\sigma(\alpha), \sigma(\theta))$ for permutations $\sigma \in S_K$. By restricting ourselves to a ‘canonical’ subspace, say a region where the component parameters $\alpha_1 \leq \dots \leq \alpha_K$ are ordered (this forms a pyramid with its apex at 0), we can recover a unique ‘standard’ labeling; the full posterior can also be recovered by mirroring the inference on this subspace to the remaining $K! - 1$ pyramids. This removes the need for sampling algorithms to explore the full space, and saves us from wasting samples.