1 Hierarchical Dirichlet Process (HDP)

We begin with the model-based treatment of the HDP. The extra layer of Dirichlet process in the HDP as compared to the Dirichlet Process Mixing Model (DPMM) allows for groups to be assigned to multiple clusters, and also lets the groups share information about those clusters.

\[
\begin{align*}
\pi_0 & \sim \text{Dir}\left(\frac{\gamma}{K}, \ldots, \frac{\gamma}{K}\right) \\
\pi_j & \sim \text{Dir}(\alpha \pi_0) \\
z_{jn} & \sim \text{Cat}(\pi_j) \\
y_{jn} & \sim f(\phi_{z_{jn}}) \\
\phi_k & \sim H
\end{align*}
\]

Note here that \( K \) is taken to go to infinity, so that there is no theoretical limit to the number of clusters. A slightly more compact treatment of the HDP is the measure-theoretic one, which makes use of the Dirichlet process (DP) and a base measure \( H \).

\[
\begin{align*}
G_0 & \sim DP(\gamma, H) \\
G_j & \sim DP(\alpha, G_0) \\
\phi_{jn} & \sim G_j \\
y_{jn} & \sim f(\phi_{jn})
\end{align*}
\]

2 Impromptu Discussion of “Nonparametric” Part of BNP

The gist is that “nonparametric” refers to not being limited by a finite set of parameters, i.e. there is no limit to the number of clusters. That being said, there aren’t really fewer parameters being chosen, but hopefully the parameter choices are more robust. That is to say, for instance, the quality of the clustering (deliberately vaguely defined) ought to be less sensitive to the choice of \( \alpha \) than to an explicit choice of the number of clusters.
3 Impromptu Discussion of MCMC Convergence

The gist is that it may be unrealistic to expect convergence for high dimensional / complex hierarchical models. However, the algorithm will often get around the right area of the state space and is fairly robust to local optima. In low dimensions or for simple models, convergence is not unrealistic.

4 Review of DPMM

Similarly as for the HDP above, the DPMM can be written in both model-based form (where again $K$ is going to infinity),

\[
\pi \sim \text{Dir}(\frac{\gamma_1}{K}, \cdots, \frac{\gamma_K}{K})
\]

\[z_n \sim \text{Cat}(\pi)\]

\[y_n \sim f(\phi_{z_n})\]

\[\phi_k \sim H\]

and measure-theoretic form,

\[G \sim \text{DP}(\gamma, H)\]

\[\phi_n \sim G\]

\[y_n \sim f(\phi_n)\]