Lecture Summary 16

Here we consider the infinite mixture model

$$k(x|\theta) = \sum_{j=1}^{\infty} w_j \, N(x|\mu_j, \sigma^2).$$

For simplicity assume the weights are geometric, so for some τ we have

$$w_j = \tau(1-\tau)^{j-1}, \quad j = 1, 2, \ldots,$$

and τ has prior $f(\tau)$; e.g. a beta distribution.

We need a way to solve the problem that the choice of d becomes infinite. To simplify but accurately describe the problem, consider us wanting to sample d where $P(d) = w_d$ and $d \in \{1, 2, \ldots\}$. We can introduce the latent variable u for which we get a joint density

$$k(u, d) = \mathbf{1}(u < w_d).$$

Now $f(u|d)$ is got by taking u as uniform on $(0, w_d)$. And $f(d|u)$ is uniform from the set $\{1, \ldots, J\}$ where $J = \max\{j : w_j > u\}$.

If we now have d_i for $i = 1, \ldots, n$ and include w as part of the problem, we now have the following Gibbs framework: u_i is uniform on w_{d_i}; $P(d_i = j)$ is uniform from the set $\{1, \ldots, J_i\}$, where $J_i = \max\{j : w_j > u_i\}$; and

$$f(\tau|d_1, \ldots, d_n) \propto f(\tau) \prod_{i=1}^{n} \tau(1-\tau)^{d_i - 1}.$$

This last conditional has excluded the u_i; we will discuss how this works later in the course. Note that we would only need to compute the weights $(w_1, \ldots, w_J)$ where $J = \max_i\{J_i\}$.

To now deal with the infinite mixture model, we use the joint posterior based on the full likelihood

$$f(u, d, \mu, \lambda, \tau) \propto f(\mu)\, f(\tau)\, f(\lambda) \prod_{i=1}^{n} \mathbf{1}(u_i < w_{d_i})\, N(x_i|\mu_{d_i}, \sigma^2).$$

We would know how to deal with the (μ_j), and how many we need to sample, as we do the weights (w_j) and τ, and λ and the (u_i). The d_i are given by given by

$$P(d_i = j|\cdots) \propto N(x_i|\mu_j, \sigma^2)\ \mathbf{1}(j \in \{1, \ldots, J_i\}).$$