Lecture Summary 15

Following on from Lecture 14, we will look at a more general mixture model;

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

where now $\theta = (w, \mu, \sigma)$ and the w and μ are vectors of size M. Here we assume a common σ for each component.

The priors are a common $f(\mu_j) = N(\nu, \phi^2)$ for each j and $f(\lambda) = \mathrm{Ga}(a, b)$ for $\lambda = 1/\sigma^2$. The prior for w is Dirichlet; i.e.

$$f(w) \propto \prod_{j=1}^{M} w_j^{\alpha_j - 1}$$

subject to the constraint $w_1 + \cdots + w_M = 1$. This can also be written as

$$f(w) \propto \prod_{j=1}^{M-1} w_j^{\alpha_j - 1} \times (1 - w_1 - \cdots - w_{M-1})^{\alpha_M - 1}.$$

This generalizes the beta distribution, which arises when $M = 2$, and can be sampled via independent gamma variables. So if $V_j \sim \mathrm{Ga}(\alpha_j, 1)$ we can take

$$w_j = \frac{V_j}{\sum_{j=1}^{M} V_j}.$$

As we did previously, we consider the full likelihood

$$k(x, d|\theta) = \prod_{i=1}^{n} w_{d_i} \, N(x_i|\mu_{d_i}, \sigma^2)$$

and we can check that

$$\sum_{d_1=1}^{M} \cdots \sum_{d_n=1}^{M} k(x, d|\theta) = k(x|\theta),$$

with each $d_i \in \{1, \ldots, M\}$.

In class we will look at the conditionals required for a Gibbs sampler framework and see how we can get the guaranteed density estimate of k.