Lecture Summary 18

Here we consider the infinite mixture model

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

where the parameters are $(M, \mu_1, \mu_2, \ldots, \mu_M, \mu_{M+1}, \ldots)$, as there really are an infinite number of μ's as M can be arbitrarily large. Even though for each M we only need $(\mu_1, \ldots, \mu_M)$ for the model.

The prior for M is $f(M)$ and the prior for λ is the usual gamma, the prior for the $f(\mu_1, \ldots, \mu_M|M)$ will be independent $N(\nu, \phi^2)$. We also need to specify $f(\mu_{M+1}, \mu_{M+2}, \ldots |M, \mu_1, \ldots, \mu_M)$, but we will leave this to one side for now. For simplicity we assume the weights $w_{j,M}$ are known, e.g. $w_{j,M} = 1/M$ for $j = 1, \ldots, M$.

Given the chain arrives at $(M, \lambda, \mu_1, \ldots, \mu_M)$ we can sample a new set of $(\lambda, \mu_1, \ldots, \mu_M)$ by introducing the (d_i) and sampling as in the fixed M case.

The real question is how to move between one M and another, and for simplicity we will only consider, for now, the move $M \to M+1$. The proposal is to move $(M, \mu) \to (M+1, \mu)$ with probability $\frac{1}{2}$ and (M, μ) to $(M-1, \mu)$ with probability $\frac{1}{2}$, except when $M = 1$ in which case we make the former proposal with probability 1.

The proposal is accepted with probability

$$\alpha = \min\left\{1, \frac{\frac{1}{2}\, f(M+1, \mu)}{\frac{1}{2} f(M, \mu)}\right\},$$

the $\frac{1}{2}$ being present, but cancel out, as these form the proposal. They do not cancel when we are involving $M = 1$.

The acceptance probability depends on an infinite set of μ but there are cancellations. Now

$$\frac{f(M+1, \mu)}{f(M, \mu)} = \frac{\prod_{i=1}^{n} k(x_i|M+1, \mu_{1:M+1}) f(\mu_{1:M+1}|M+1)\, f(\mu_{M+2:\infty}|M+1, \mu_{1:M+1})}{\prod_{i=1}^{n} k(x_i|M, \mu_{1:M}) f(\mu_{1:M}|M)\, f(\mu_{M+1:\infty}|M, \mu_{1:M})}.$$

If all the priors are independent $N(\nu, \phi^2)$, regardless of M, then this all cancels to

$$\frac{\prod_{i=1}^{n} k(x_i|M+1, \mu_{1:M+1})}{\prod_{i=1}^{n} k(x_i|M, \mu_{1:M})}.$$

More general types of prior are available, and cancellations still occur from

$$\frac{f(\mu_{M+2:\infty}|M+1, \mu_{1:M+1})}{f(\mu_{M+1:\infty}|M, \mu_{1:M})}.$$