Nota-1. Deducción del algoritmo EM para mezcla de gausianas

$$LL = \log P(X|\pi,\theta) = \sum_{i=1}^{M} \log \left\{ \sum_{k=1}^{K} \pi_k N(x_i|\mu_k,\sigma_k^2) \right\} =$$

$$\frac{\partial LL}{\partial \mu_k} = \sum_{i=1}^{M} \frac{\pi_k}{\sum_{l=1}^{K} \pi_l N(x|\mu_l,\sigma_l^2)} \frac{1}{\sqrt{2\pi}\sigma_k} \left(\frac{2(x_i-\mu_k)}{2\sigma_k^2} \right) N(x_i|\mu_k,\sigma_k^2) = 0$$

$$\sum_{i=1}^{M} \gamma_k(x_i) \frac{(x_i-\mu_k)}{\sigma_k^2} = 0 \Rightarrow \sum_{i=1}^{M} \gamma_k(x_i) x_i = \sum_{i=1}^{M} \gamma_k(x_i) \mu_k$$

$$\mu_k = \frac{\sum_{k=1}^{K} \gamma_k(x_i) x_i}{\sum_{i=1}^{M} \gamma_k(x_i)}$$

$$\frac{\partial LL}{\partial \sigma_k^2} = \sum_{i=1}^{M} \frac{\pi_k}{\sum_{l=1}^{K} \pi_l N(x|\mu_l,\sigma_l^2)} \left[\frac{-\sigma_k^2}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu_k)^2}{2\sigma_k^2}} + \frac{\sigma_k^{-1}}{\sqrt{2\pi}} \left(-\frac{1}{2}(x_i-\mu_k)^2 (-2\sigma_k^{-3}) \right) e^{-\frac{(x_i-\mu_k)^2}{2\sigma^2}} \right]$$

$$= \sum_{i=1}^{M} \gamma_k(x_i) \left[(-\sigma_k^{-1}) + \sigma_k^{-3} (x_i-\mu_k)^2 \right] = 0$$

$$\sum_{i=1}^{M} \gamma_k(x_i) \sigma_k^{-1} = \sum_{i=1}^{M} \gamma_k(x_i) \sigma_k^{\cancel{-3}\,-2} (x_i-\mu_k)^2$$

$$\sigma_k^2 = \frac{\sum_{i=1}^{M} \gamma_k(x_i) (x_i-\mu_k)^2}{\sum_{i=1}^{M} \gamma_k(x_i)}$$

$$L = LL + \lambda \left(\sum_{k=1}^{K} \pi_k - 1 \right)$$

$$\frac{\partial L}{\partial \pi_k} = \sum_{i=1}^{M} \frac{N(x_i|\mu_k,\sigma_k^2)}{\sum_{l=1}^{K} \pi_l N(x_i|\mu_k,\sigma_k^2)} + \lambda = 0 \Rightarrow \sum_{i=1}^{M} \frac{\sum_{k=1}^{K} \pi_k N(x_i|\mu_k,\sigma_k^2)}{\sum_{l=1}^{K} \pi_l N(x_i|\mu_k,\sigma_k^2)} + \lambda \sum_{k=1}^{K} \pi_k = 0$$

$$\Rightarrow \lambda = -\sum_{i=1}^{M} 1 = M \Rightarrow \sum_{i=1}^{M} \pi_k \gamma_k(x_i) - M \pi_k = 0 \Rightarrow \pi_k = \frac{1}{M} \sum_{i=1}^{M} \gamma_k(x_i) \pi_k$$