EM for Mixture of Gaussians (by hand)

$$p(x_n \mid \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \boldsymbol{\pi}) = \sum_m p(x_n, z_n = m \mid \mu_m, \sigma_m{}^2, \pi_m) = \sum_m \frac{1}{\sqrt{2\pi\sigma_m{}^2}} \exp\left(-\frac{(x_n - \mu_m)^2}{\sigma_m{}^2} \right) \cdot \pi_m$$

1. E-step: $\mathrm{E}_{z \mid x, \theta^{(t)}} \left[\log(p(\mathbf{x}, \mathbf{z} \mid \theta)) \right] = \sum_{\mathbf{z}} \log(p(\mathbf{x}, \mathbf{z} \mid \theta)) p(\mathbf{z} \mid \mathbf{x}, \theta^{(t)})$

2. M-step: $\theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}} \sum_{\mathbf{z}} \log(p(\mathbf{x}, \mathbf{z} \mid \theta)) p(\mathbf{z} \mid \mathbf{x}, \theta^{(t)})$

$$\sum_{\underline{z}} \sum_n \log(p(x_n \mid z_n, \theta) P(z_n \mid \theta)) \; p(\underline{z} \mid \underline{x}, \theta^t)$$

$$\sum_{n=1}^{N} \sum_{z_1} \cdots \sum_{z_n} \cdots \sum_{z_N} \left(\log P(x_n \mid z_n, \theta) + \log P(z_n \mid \theta) \right) P(\underline{z} \mid x, \theta^t)$$

$$\sum_{n=1}^{N} \sum_{z_n = m} \left(\log P(x_n \mid z_n, \theta) + \log P(z_n \mid \theta) \right) P(z_n = m \mid x_n, \theta^t)$$

$$\alpha_{nm} = P(z_n = m \mid x_n, \theta^t) = \frac{P(x_n \mid z_n = m, \theta^t) P(z_n = m, \theta^t)}{\sum_{k=1}^{M} \underbrace{P(x_n \mid z_n = k, \theta^t)}_{N(x_n; \mu_k^t, \sigma_k^t)} \; \underbrace{P(z_n = k, \theta^t)}_{\pi_k^t}}$$

$$\underline{\mu}^{t+1}, \underline{\sigma}^{t+1}, \underline{\pi}^{t+1} = \underset{\mu,\sigma,\pi}{\arg\max} \sum_{n=1}^{N} \sum_{m=1}^{M} \left(\log P(x_n | z_n = m, \theta) + \log P(z_n = m | \theta) \right) \alpha_{nm} \qquad \text{s.t.} \sum_{k} \pi_k = 1$$

$$\sum_{n} \sum_{m} \left(\log \frac{1}{\sqrt{2\pi}} - \log \sigma_m - \frac{1}{2} \frac{(x_n - \mu_m)^2}{\sigma_m^2} + \log \pi_m \right) \alpha_{nm} + \lambda \left(\sum_{k} \pi_k - 1 \right)$$

$$\partial / \partial \mu_m = 0 \Rightarrow \sum_{n} \cancel{\sum_{m}} \left(\frac{(x_n - \mu_m)}{\sigma_m^2} \right) \alpha_{nm} = 0$$

$$\Rightarrow \hat{\mu}_m = \frac{1}{\sum_{n} \cancel{\sum_{m}} \alpha_{nm}} \sum_{n} \alpha_{nm} x_n$$

$$\partial / \partial \sigma_m = 0 \Rightarrow \frac{1}{\sum_{n} \alpha_{nm}} \sum_{n=1}^{N} (x_n - \hat{\mu}_m)^2 \alpha_{nm}$$

$$\partial / \partial \pi_m = 0 \Rightarrow \sum_{n} -\frac{1}{\pi_m} \alpha_{nm} + \lambda = 0$$

$$\cancel{\lambda = \sum_{n}} \qquad \pi_m = \frac{1}{\lambda} \sum_{n} \alpha_{nm}$$

$$\sum_{k} \pi_k = 1$$

$$\sum_{k} \frac{1}{\lambda} \sum_{n} \alpha_{nk} = 1$$

$$\lambda = \sum_{n} \sum_{k} \alpha_{nk}$$

$$-\sum_{n} \frac{1}{\pi_m} \alpha_{nm} + \underbrace{\sum_{n} \sum_{k} \alpha_{nk}}_{N} = 0$$

$$\hat{\pi}_m = \frac{1}{N} \sum_{n} \alpha_{nm}$$