1) $P(x;\theta) = \sum_{k=1}^{K} \pi_k N(x : \mu_k, \sigma_k^2)$ ①

where π_k is mixture weight such that

$\pi_k \geq 0$, $\sum_{k=1}^{K} \pi_k = 1$, μ_k, σ_k are the mean & variance of the Gaussian distribution for cluster 'k'.

For E-step, we've to find

$$Z_{ik} = P(Z_i = k \mid X, \mu, \sigma, \pi)$$

From Bayes' Theorem,

$$Z_{ik} = \frac{P(X, \mu, \sigma, \pi \mid Z_i = k) \cdot P(Z_i = k)}{P(X, \mu, \sigma, \pi)}$$

where $P(Z_i = k)$ is the probability of k^{th} cluster $= \pi_k$

$$P(X, \mu, \sigma, \pi \mid Z_i = k) = N(x_i ; \mu_k, \sigma_k^2)$$

$$P(X, \mu, \sigma, \pi) = \sum_{k=1}^{K} N(x_i ; \mu_k, \sigma_k^2)$$

$$\therefore \boxed{Z_{ik} = \frac{\mathcal{N}(x_i ; \mu_k, \sigma_k^2) . \pi_k}{\sum_{k=1}^{K} \pi_k \mathcal{N}(x_i ; \mu_k, \sigma_k^2)}}$$

For M-step, we need to maximize the log-likelihood of all the datapoints.

~~So~~ Here, we'll try to maximize the expected value of log-likelihood rather than just maximizing the peak. The optimization statement can be rewritten as,

$$\underset{\pi, \mu, \sigma}{\operatorname{argmax}} \underbrace{\sum_{z} P(z / x, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1}) . \log(P(z, x | \pi, \mu, \sigma))}_{\Omega}$$

Lets call the term marked above as Ω. (ie)

$$\Omega = \sum_{z} P(z / x, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1}) . \log(P(z, x | \pi, \mu, \sigma))$$

$$\therefore P(Z, X \mid \pi, \mu, \sigma) = P(Z / \pi, X, \mu, \sigma) \cdot P(X \mid \pi, \mu, \sigma)$$

$$P(Z, X \mid \pi, \mu, \sigma) = P(Z/\pi) \cdot P(X/\mu, \sigma)$$ [Since Z is independent of X, μ, σ & X is independent of π)

Substituting this in the equation for Ω, we can write

$$\cancel{P(Z,\pi)\, P(X/\mu,\sigma) = \prod_{i=1}^{N} P(Z_i/\pi) \cdot P(X_i/\mu,\sigma)}$$

$$\Omega = \sum_{Z} P(Z/X, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1}) \log\left(P(Z/\pi) \cdot P(X/\mu, \sigma)\right)$$

If we consider ~~Since~~ datapoints are independent,

$$P(Z/\pi) \cdot P(X/\mu, \sigma) = \prod_{i=1}^{N} P(Z_i/\pi) \cdot P(X_i/\mu, \sigma)$$

$$\Omega = \sum_{Z} P(Z/X, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1}) \cdot \log\left(\prod_{i=1}^{N} P(Z_i/\pi) \cdot P(X_i/\mu, \sigma)\right)$$

$$= \sum_{Z} P(Z/X, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1}) \cdot \sum_{i=1}^{N} \log\left[P(Z_i/\pi) \cdot P(X_i/\mu, \sigma)\right]$$

$$= \sum_{Z} \sum_{i=1}^{N} P(Z / X, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1}).$$

$$\log \{ P(z_i / \pi) . P(x_i / \mu, \sigma)$$

$$\sum_{Z} P(Z / X, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1}) = \sum_{z_1, z_2 \dots z_n} P(z_1, z_2 \dots z_n / X_1 \dots X_n, \pi^{t-1}, \mu^{t-1}, \sigma^{t-1})$$

Also $z_1, z_2 \dots z_n$ are independent so, we can further break the above expression down as

$$\sum_{z_1 z_2 \dots z_n} P(z_1 | x_1, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) . P(z_2 | x_2, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) \dots P(z_n / x_n, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1})$$

$$\text{So, } \Omega = \sum_{i=1}^{N} \sum_{z_1, z_2 \dots z_n} P(z_1 / x_1, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) \dots P(z_n / x_n, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) \log P(z_i / \pi) . P(x_i / \mu, \sigma)$$

$$= \sum_{i=1}^{N} \sum_{z_i} \log P(z_i / \pi) . P(x_i / \mu, \sigma) . P(z_i / x_i, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1})$$

(since $\sum_{z_1..z_n} P(z_1 / x_1, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) \ldots P(z_n / x_n, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) = 1$

So, $\mu_k = \text{argmax } \Omega(\mu)$

$$\Rightarrow \frac{\partial}{\partial \mu_k}(\Omega(\mu)) = 0$$

$$\frac{\partial}{\partial \mu_k}\left[\sum_{i=1}^{N}\sum_{z_i} P(z_i / x_i, \mu_k^{t-1}, \sigma^{t-1}, \pi^{t-1}) \log P(z_i / \pi) . P(x_i / \mu, \sigma)\right] = 0$$

$$\sum_{i=1}^{N}\sum_{z_i} P(z_i / x_i, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) \frac{\partial}{\partial \mu_k}\left[\log(P(z_i/\pi) . P(x_i/\mu, \sigma))\right] = 0$$

$$\frac{\partial}{\partial \mu_k}\left[\log(P(z_i/\pi)) + \log(P(x_i/\mu, \sigma))\right] = 0$$

$$0 + \frac{\partial}{\partial \mu_k}(\log(P(x_i/\mu, \sigma))) = 0$$

$$P(x_i / \mu_k, \sigma_k) = \mathcal{N}(x_i / \mu_k, \sigma_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(x_i - \mu_k)^2}{2\sigma_k^2}}$$

$$\log(P(x_i / \mu_k, \sigma_k)) = \log\left(\frac{1}{\sqrt{2\pi\sigma_k^2}}\right) + \log\left(e^{-\frac{(x_i - \mu_k)^2}{2\sigma_k^2}}\right)$$

$$= \log\left(\frac{1}{\sqrt{2\pi\sigma_k^2}}\right) - \frac{(x_i - \mu_k)^2}{2\sigma_k^2}$$

$$\frac{\partial}{\partial \mu_k}\left[\log(P(x_i / \mu_k, \sigma_k))\right] = 0 \quad \frac{+2(x_i - \mu_k)(+1)}{2\sigma_k^2} = \left(\frac{x_i - \mu_k}{\sigma_k^2}\right)$$

$$\therefore \frac{\partial}{\partial \mu_k}\left(\Omega(\mu)\right) = \sum_{i=1}^{n} p(z_i = k / x_i, \mu^{t-1}) \cdot \frac{(x_i - \mu_k)}{\sigma_k^2}$$

$$\frac{\partial}{\partial \mu_k}\left(\Omega(\mu)\right) = \sum_{i=1}^{N} z_{ik}\,(x_i - \mu_k) = 0$$

$$\therefore \sum_{i=1}^{N} z_{ik}\, x_i - \sum_{i=1}^{N} z_{ik}\, \mu_k = 0$$

$$\sum_{i=1}^{N} z_{ik}\, x_i = \mu_k \sum_{i=1}^{n} z_{ik}$$

$$\mu_k = \frac{\sum_{i=1}^{N} z_{ik}\, x_i}{\sum_{i=1}^{N} z_{ik}}$$

To find π_k, we need to solve the equation

$$\pi_k = \operatorname{argmax}(\Omega(\pi_k))$$

$$\Rightarrow \frac{\partial}{\partial \pi_k}\Omega(\pi_k) = 0$$

$$\frac{\partial}{\partial \pi_k} = \frac{\partial}{\partial \pi_k}\left(\sum_{i=1}^{N}\sum_{z_i} P(z_i | x_i, \mu^{t-1}, \sigma^{t-1}, \pi^{t-1}) \cdot \log P(z_i/\pi)\cdot P(x_i/\mu, \sigma)\right)$$

$$= \sum_{i=1}^{N} z_{ik}\frac{\partial}{\partial \pi_k}\left[\log P(z_i/\pi) + \log P(x_i/\mu, \sigma)\right]$$

$$= \sum_{i=1}^{N} z_{ik}\left[\cancel{\log}\frac{\partial}{\partial \pi_k}\log P(z_i/\pi) + 0\right]$$

$$= \sum_{i=1}^{N} z_{ik} - \pi_k$$

$$\frac{\partial}{\partial \pi_k} = 0 \Rightarrow \sum_{i=1}^{N}(z_{ik} - \pi_k) = 0$$

$$\cancel{\pi_k = \sum_{i=1}^{N} z_{ik}}$$

$$\sum_{i=1}^{N} z_{ik} = \sum_{i=1}^{N}\pi_k$$

$$= \pi_k \sum_{i=1}^{N} 1$$

$$\sum_{i=1}^{N} z_{ik} = N\pi_k \Rightarrow \boxed{\pi_k = \frac{\prod_{i=1}^{N} z_{ik}}{N}}$$