(Derivations for the update equations for GMM)
MLE for μ_K, Σ_K, π_K

log-likelihood:
$$\mathcal{L} = \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma_{nK} \left[\log \pi_K + \log N(x_n | \mu_K, \Sigma_K) \right]$$

Consider a single μ_K:
$$\frac{\partial \mathcal{L}}{\partial \mu_K} = \sum_{n=1}^{N} \gamma_{nK} \frac{\partial}{\partial \mu_K} \log N(x_n | \mu_K, \Sigma_K)$$

Now, $\log N(x_n | \mu_K, \Sigma_K) \propto -\frac{1}{2} (x_n - \mu_K)^T \Sigma_K^{-1} (x_n - \mu_K)$
(other terms don't depend on μ_K and can be ignored)

Thus
$$\frac{\partial L}{\partial \mu_K} = \sum_{n=1}^{N} \gamma_{nK} \frac{\partial}{\partial \mu_K} \left[-\frac{1}{2} (x_n - \mu_K)^T \Sigma_K^{-1} (x_n - \mu_K) \right]$$
$$= \sum_{n=1}^{N} \gamma_{nK} \left[+ \cancel{\frac{1}{2}} \times \cancel{2} \times \Sigma_K^{-1} (x_n - \mu_K) \right]$$

$$\frac{\partial L}{\partial \mu_K} = \sum_{n=1}^{N} \gamma_{nK} \Sigma_K^{-1} (x_n - \mu_K) = 0$$

Multiplying both sides by Σ_K, we get

$\cancel{\frac{\partial L}{\partial \mu_K}} =$ [illegible]

$$\frac{\partial L}{\partial \mu_K} = \sum_{n=1}^{N} \gamma_{nK} (x_n - \mu_K) = 0$$

$$\Rightarrow \boxed{\mu_K = \frac{\sum_{n=1}^{N} \gamma_{nK} x_n}{\sum_{n=1}^{N} \gamma_{nK}}}$$

Now consider a single Σ_K:
$$\frac{\partial L}{\partial \Sigma_K} = \sum_{n=1}^{N} \gamma_{nK} \frac{\partial}{\partial \Sigma_K} \log N(x_n | \mu_K, \Sigma_K)$$

$$\log N(x_n | \mu_K, \Sigma_K) \propto -\frac{1}{2} \log |\Sigma_K| - \frac{1}{2} (x_n - \mu_K)^T \Sigma_K^{-1} (x_n - \mu_K)$$
$$= \frac{1}{2} \log |\Sigma_K^{-1}| - \frac{1}{2} (x_n - \mu_K)^T \Sigma_K^{-1} (x_n - \mu_K)$$
$$= \frac{1}{2} \log |\Sigma_K^{-1}| - \frac{1}{2} tr\left[(x_n - \mu_K)(x_n - \mu_K)^T \Sigma_K^{-1} \right]$$

Taking derivatives w.r.t. Σ_k^{-1} (instead of Σ_k; it doesn't matter)

$$\frac{\partial L}{\partial \Sigma_k^{-1}} = \sum_{n=1}^{N} \gamma_{nk} \frac{\partial}{\partial \Sigma_k^{-1}} \left[\frac{1}{2}\log|\Sigma_k^{-1}| - \frac{1}{2} tr\left[(x_n-\mu_k)(x_n-\mu_k)^T \Sigma_k^{-1}\right]\right]$$

$$= \frac{1}{2}\sum_{n=1}^{N} \gamma_{nk} \left[(\Sigma_k^{-1})^{-1} - (x_n-\mu_k)(x_n-\mu_k)^T \right] = 0$$

$$\Rightarrow \sum_{n=1}^{N} \gamma_{nk} \Sigma_k = \sum_{n=1}^{N} \gamma_{nk} (x_n-\mu_k)(x_n-\mu_k)^T$$

$$\Rightarrow \boxed{\Sigma_k = \frac{1}{\sum_{n=1}^{N}\gamma_{nk}} \left[\sum_{n=1}^{N} \gamma_{nk}(x_n-\mu_k)(x_n-\mu_k)^T\right]}$$

Note: Σ_k is actually P.S.D., so this constraint should be imposed while optimizing w.r.t Σ_k, but I ignored that above for simplicity.

Finally, let's find $\{\pi_k\}_{k=1}^{K}$. We know that $\sum_{k=1}^{K} \pi_k = 1$

The constrained optimization problem's objective:

$$\mathcal{L} = \sum_{n=1}^{N}\sum_{k=1}^{K} \gamma_{nk} \log \pi_k + \underset{\substack{\downarrow \\ \text{Lagrange} \\ \text{Multiplier}}}{\lambda} \left[1 - \sum_{k=1}^{K}\pi_k\right]$$

Taking derivative w.r.t. a single π_k:

$$\frac{\partial \mathcal{L}}{\partial \pi_k} = \sum_{n=1}^{N} \gamma_{nk} \frac{1}{\pi_k} - \lambda = 0$$

$$\Rightarrow \pi_k = \frac{\sum_{n=1}^{N}\gamma_{nk}}{\lambda}$$

Now, what's λ? Since $\sum_{k=1}^{K}\pi_k = 1$, $\frac{1}{\lambda} \underbrace{\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma_{nk}}_{\text{equal to } N} = 1$

thus $\lambda = N$

Therefore $\boxed{\pi_k = \frac{\sum_{n=1}^{N}\gamma_{nk}}{N}}$