1, 1)

$$r_{n,k} \triangleq P(z_n = k \mid x_n ; \theta)$$

$$= \frac{P(x_n, z_n = k ; \theta)}{P(x_n ; \theta)}$$

$$= \frac{P(x_n \mid z_n = k ; \theta)\, P(z_n = k ; \theta)}{\sum_{k=1}^{K} P(x_n \mid z_n = k ; \theta)\, P(z_n = k ; \theta)}$$

$$\boxed{= \frac{\cancel{P(x_n \mid z_n = k ; \theta)\, \pi_k}\ \mathcal{N}(x_n ; \mu_k, \Sigma_k)\, \pi_k}{\sum_{k=1}^{K} \mathcal{N}(x_n ; \mu_k, \Sigma_k)\, \pi_k}}$$

1, 2). The responsability $r_{n,k}$ indicates ~~that~~ the probability that, given the observed stone x_n, the student z_n threw it.

1, 3). $\ln P(\vec{x} ; \theta) = \cancel{\ln P(\vec{x}}\ \ln \prod_{n=1}^{N} P(x_n ; \theta)$

$$= \ln \prod_{n=1}^{N} \sum_{k=1}^{K} P(x_n \mid z_n = k ; \theta)\, P(z_n = k ; \theta)$$

$$= \boxed{\sum_{n=1}^{N} \ln \left(\sum_{k=1}^{K} \mathcal{N}(x_n ; \mu_k, \Sigma_k)\, \pi_k \right)} \rightarrow LML.$$

2, 1). $Q(\theta, \hat{\theta}) = \mathbb{E}_{P(Z \mid X ; \hat{\theta})} \left[\ln P(\vec{x}, \vec{z} ; \theta) \right]$

$$= \sum_{n=1}^{N} \sum_{k=1}^{K} \underset{r_{n,k,\hat{\theta}}}{P(Z \mid X ; \hat{\theta})}\, \mathbb{1}\{z_n = k\} \left\{ \left(\ln \pi_k + \ln \left(\frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2} (\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu}) \right) \right) \right) \right\}$$

$$\boxed{= \sum_{n=1}^{N} \sum_{k=1}^{K} r_{n,k,\hat{\theta}} \left\{ \ln \pi_k - \frac{1}{2} \ln(\det(\Sigma_k)) - \frac{1}{2} \ln 2\pi - \frac{1}{2} (\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k) \right\}}$$

2, 2).

a. $$\frac{\partial Q(\theta, \hat{\theta})}{\partial \mu_k} = \sum_{n=1}^{N} r_{n,k,\hat{\theta}} \left\{ -\frac{1}{2} \frac{\partial (\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k)}{\partial \mu_k} \right\}$$

$$\frac{\partial X^T \Sigma X}{\partial X} = 2\Sigma X.$$

$$= \sum_{n=1}^{N} r_{n,k,\hat{\theta}} \left\{ \frac{1}{2} \frac{\partial (\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k)}{\partial (x_n - 2\mu_k)} \frac{\partial (x_n - \mu_k)}{\partial \mu_k} \right\}$$

$$= \sum_{n=1}^{N} r_{n,k,\hat{\theta}} \left\{ +2\, \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k) \right\}$$

$$= 0.$$

$$\Rightarrow \sum_{n=1}^{N} r_{n,k,\hat{\theta}}\, \Sigma_k^{-1} \underline{x}_n = \sum_{n=1}^{N} r_{n,k,\hat{\theta}}\, \Sigma_k^{-1} \mu_k = N_k \Sigma_k^{-1} \mu_k.$$

$$\Rightarrow \boxed{\mu_k = \frac{\sum_{n=1}^{N} r_{n,k,\hat{\theta}}\, \underline{x}_n}{N_k}.}$$

b). $\nabla_{\Sigma_k} Q(\theta,\tilde{\theta}) = \nabla_{\Sigma_k} \sum_{n=1}^{N} \sum_{k=1}^{K} r_{n,k,\tilde{\theta}} \left(\cancel{\ln \pi_k} - \frac{1}{2} \cancel{\ln 2\pi} - \frac{1}{2}\ln(\det \Sigma_k) - \frac{1}{2}(\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k) \right)$

$$= -\frac{1}{2}\sum_{n=1}^{N} r_{n,k,\tilde{\theta}} \left(\boxed{\nabla_{\Sigma_k} \ln(\det \Sigma_k)} + \boxed{\nabla_{\Sigma_k} (\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k)} \right)$$

K, D.

$$= -\frac{1}{2}\sum_{n=1}^{N} r_{n,k,\tilde{\theta}} \left([(\Sigma_k^T)^{-1}]^T + \nabla_{\Sigma_k} tr(\Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k)(\underline{x}_n - \underline{\mu}_k)^T) \right)$$

N. K. N×D

$$= -\frac{1}{2}\sum_{n=1}^{N} r_{n,k,\tilde{\theta}} \left(\Sigma_k^{-1} - (\Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k)(\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-1})^T \right)$$

N×D, N×K

$$= -\frac{1}{2}\sum_{n=1}^{N} r_{n,k,\tilde{\theta}} \left(\Sigma_k^{-1} - \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k)^T (\underline{x}_n - \underline{\mu}_k) \Sigma_k^{-2} \right)$$

⇒ D×K? K×D?

$= 0$

N. D. - K. D

$$\Rightarrow \sum_{n=1}^{N} r_{n,k,\tilde{\theta}} \Sigma_k^{-1} = \sum_{n=1}^{N} r_{n,k,\tilde{\theta}} \Sigma_k^{-1} (\underline{x}_n - \underline{\mu}_k)^T (\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-1}$$

$$\Rightarrow N_k = \sum_{n=1}^{N} r_{nk} (\underline{x}_n - \underline{\mu}_k)^T (\underline{x}_n - \underline{\mu}_k)^T \Sigma_k^{-2}$$

means: K. D.

N×K.

K×N. N×D.

$$\Rightarrow \boxed{\Sigma_k = \frac{\sum_{n=1}^{N} r_{nk} (\underline{x}_n - \underline{\mu}_k)^T (\underline{x}_n - \underline{\mu}_k)^T}{N_k}}$$

500×2

K. D.

N. D

K. N. K. D.

~~N×D - D~~

~~D×N - D.~~

N×D - D.

c). $\frac{\partial L(\theta,\hat{\theta},\lambda)}{\partial \pi_k} = \frac{\partial Q(\theta,\hat{\theta})}{\partial \pi_k} + \lambda = 0.$

$$\frac{\partial Q(\theta,\hat{\theta})}{\partial \pi_k} = \sum_{n=1}^{N} r_{n,k,\tilde{\theta}} \frac{1}{\pi_k} = \frac{N_k}{\pi_k}.$$

D. K. K. N

D.

N. D.

N. K.

N. D. N_k K. N.

K. D K. D.

K. N

N. K N. D

$$\Rightarrow \frac{N_k}{\pi_k} + \lambda = 0.$$

$$\Rightarrow \boxed{N + \lambda = 0}$$

$$\begin{cases} \lambda = -N \\ \pi_k = \frac{N_k}{N}. \end{cases}$$