2.16

$$
\begin{aligned}
Q &= \sum_{i=1}^{N}\sum_{k=1}^{K} -\gamma_{k,i} \log\left[p_{X|Z}(x_i|k)\cdot p_Z(k)\right] \\
&= \sum_{i=1}^{N}\sum_{k=1}^{K} \gamma_{k,i}\left(\frac{1}{2}\log[2\pi\det(\Sigma_k)] + \frac{1}{2}(x_i-\mu_k)^T\Sigma_k^{-1}(x_i-\mu_k) - \log[p_Z(k)]\right)
\end{aligned}
$$

Assuming that Q is convex, find the optimal values for μ_k and Σ_k

HINT: You may use the following general matrix/vector gradient equations (refer to the [matrix cook book by Peterson and Pedersen (2012), p.10-11](https://www.math.uwaterloo.ca/ hwolkowi/matrix-cookbook.pdf) :

$$
\begin{aligned}
\nabla_x (x-y)^T W (x-y) &= 2W(x-y) && (1)\\
\nabla_W (x-y)^T W (x-y) &= (x-y)(x-y)^T && (2)\\
\nabla_{W^{-1}} \log[\det(W)] &= -W \qquad \text{if } W \text{ is symmetric and positive semi-definite} && (3)
\end{aligned}
$$

$$
\begin{aligned}
\nabla_{\mu_k}\sum_{i=1}^{N}\sum_{k=1}^{K}\gamma_{k,i}\left(\frac{1}{2}\log[2\pi\det(\Sigma_k)] + \frac{1}{2}(x_i-\mu_k)^T\Sigma_k^{-1}(x_i-\mu_k) - \log[p_Z(k)]\right) &\overset{!}{=} \vec{0} \\
\Leftrightarrow \sum_{i=1}^{N}\gamma_{k,i}\left(\nabla_{\mu_k}\frac{1}{2}(x_i-\mu_k)^T\Sigma_k^{-1}(x_i-\mu_k)\right) &= 0 \\
\Leftrightarrow \sum_{i=1}^{N}\gamma_{k,i}\left(\Sigma_k^{-1}(x_i-\mu_k)\right) &= 0 \\
\Leftrightarrow \sum_{i=1}^{N}\gamma_{k,i}\left(\Sigma_k^{-1}x_i-\Sigma_k^{-1}\mu_k\right) &= 0 \\
\Leftrightarrow \sum_{i=1}^{N}\gamma_{k,i}\Sigma_k^{-1}x_i &= \sum_{i=1}^{N}\gamma_{k,i}\Sigma_k^{-1}\mu_k \\
\Leftrightarrow \Sigma_k^{-1}\sum_{i=1}^{N}\gamma_{k,i}x_i &= \Sigma_k^{-1}\mu_k\sum_{i=1}^{N}\gamma_{k,i} \\
\Leftrightarrow \frac{\sum_{i=1}^{N}\gamma_{k,i}x_i}{\sum_{i=1}^{N}\gamma_{k,i}} &= \mu_k
\end{aligned}
$$

Lorem ipsum

number 57 141

$$
\begin{aligned}
\nabla_{\Sigma_k^{-1}} \sum_{i=1}^{N} \sum_{k=1}^{K} \gamma_{k,i} \left(\frac{1}{2} \log[2\pi \det(\Sigma_k)] + \frac{1}{2}(x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) - \log[p_Z(k)] \right) &\overset{!}{=} \vec{0} \\
\Leftrightarrow \sum_{i=1}^{N} \gamma_{k,i} \left(-\frac{1}{2}\Sigma_k + \frac{1}{2}(x_i - \mu_k)^T (x_i - \mu_k) \right) &= 0 \\
\Leftarrow \sum_{i=1}^{N} \gamma_{k,i} \Sigma_k &= \sum_{i=1}^{N} \gamma_{k,i} (x_i - \mu_k)^T (x_i - \mu_k) \\
\Leftrightarrow \Sigma_k \sum_{i=1}^{N} \gamma_{k,i} &= \sum_{i=1}^{N} \gamma_{k,i} (x_i - \mu_k)^T (x_i - \mu_k) \\
\Leftrightarrow \Sigma_k &= \frac{\sum_{i=1}^{N} \gamma_{k,i} (x_i - \mu_k)^T (x_i - \mu_k)}{\sum_{i=1}^{N} \gamma_{k,i}}
\end{aligned}
$$