Gaussian mixture Model.

$$q(x;\mu,\Sigma) = \sum_{l=1}^{m} N(x;\mu_l,\Sigma_l)\, w_l.$$

$$N(x,\mu_l,\Sigma_l) = \frac{1}{(2\pi)^{\frac{d}{2}}\det(\Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_l)^T\Sigma^{-1}(x-\mu_l)}$$

注意到. $\int q(x;\mu,\Sigma) = 1 = \sum_{l=1}^{m} w_l$

objective: maximize $\ln L(\theta)$

s.t: $\sum_{l=1}^{m} w_l = 1$.

令 $w_l = \frac{y_l}{\sum_{i=1}^{m} y_i}$

$$\begin{aligned}
\ln L(\theta) &= \sum_{i=1}^{n} \ln q(x_i;\theta) \\
&= \sum_{i=1}^{n} \ln\left(\sum_{l=1}^{m} w_l N(x_i;\mu_l,\Sigma_l)\right) \\
&= \sum_{i=1}^{n} \ln\left(\sum_{l=1}^{m} \frac{y_l}{\sum_{j=1}^{m} y_j} N(x_i;\mu_l,\Sigma_l)\right) \\
&= \sum_{i=1}^{n} \ln \frac{\sum_{l=1}^{m} y_l N(x_i;\mu_l,\Sigma_l)}{\sum_{j=1}^{m} y_j} \\
&= \sum_{i=1}^{n} \ln\left(\sum_{l=1}^{m} y_l N(x_i;\mu_l,\Sigma_l)\right) - \sum_{i=1}^{n} \ln \sum_{j=1}^{m} y_j \\
&= \sum_{i=1}^{n} \ln\left(\sum_{l=1}^{m} y_l N(x_i;\mu_l,\Sigma_l)\right) - n \ln \sum_{j=1}^{m} y_j
\end{aligned}$$

$$\therefore \frac{\partial \ln L(\theta)}{\partial y_k} = \sum_{i=1}^{n} \frac{N(x_i;\mu_k,\Sigma_k)}{\sum_{l=1}^{m} y_l N(x_i;\mu_l,\Sigma_l)} - \frac{1}{\sum_{j=1}^{m} y_j}\times n.$$

如果令 $y_l = \exp(y_l)$

$$\text{则 } \frac{\partial \ln L(\theta)}{\partial y_k} = \sum_{i=1}^{n} \frac{\exp(y_k) N(x_i;\mu_k,\Sigma_k)}{\sum_{l=1}^{m} \exp(y_l) N(x_i;\mu_l,\Sigma_l)} - \frac{\exp(y_k)}{\sum_{j=1}^{m}\exp(y_l)}\times n$$

$$= \sum_{i=1}^{n} \frac{w_k N(x_i;\mu_k,\Sigma_k)}{\sum_{l=1}^{m} w_l N(x_i;\mu_l,\Sigma_l)} - n w_k.$$

令 $\eta_{il} = \frac{w_l N(x_i;\mu_l,\Sigma_l)}{\sum_{l'=1}^{m} w_{l'} N(x_i;\mu_{l'},\Sigma_{l'})}$

$$\text{则 } \frac{\partial \ln L(\theta)}{\partial y_l} = \sum_{i=1}^{n} \eta_{il} - n w_l.$$

$$p(z_i=j \mid x_i) \propto p(x_i, z_j^{(}=j)$$

$$= p(z_i=j)\, p(x_i \mid z_i=j).$$

$$= w_j N(x_i; \Sigma_j, \mu_j)$$

$$\ln L(\theta) = \sum_{i=1}^{n} \ln q(x_i \mid \theta)$$

$$= \sum_{i=1}^{n} \ln\left(\sum_{l=1}^{m} w_l N(x_i; \mu_l, \Sigma_l)\right)$$

$$= \sum_{i=1}^{n} \ln \sum_{l=1}^{m} \hat{n}_{il} \times \frac{w_l}{\hat{n}_{il}} N(x_i; \mu_l, \Sigma_l)$$

$$\geq \sum_{i=1}^{n} \sum_{l=1}^{m} \hat{n}_{il} \ln\left(\frac{w_l N(x_i; \mu_l, \Sigma_l)}{n_{il}}\right) = b(\theta)$$

$$= \sum_{i=1}^{n} \sum_{l=1}^{m} \hat{n}_{il} \Big[\ln w_l - \ln (2\pi)^{\frac{d}{2}} \det(\Sigma)^{\frac{1}{2}} - \frac{1}{2}(x-\mu_l)^T (\Sigma_l)^{-1} (x-\mu_l) - \ln \hat{n}_{il} (\ )\ b(\theta)$$

当 $\theta = \hat{\theta}$ 两个相等.

$$\frac{\partial b(\theta)}{\partial w_l} = \frac{\sum_{i=1}^{n} \sum_{l=1}^{m} \hat{n}_{il}}{w_l} = 0.$$