Programming 2

等协方差矩阵情形 M-step 的推导

参考上课时张长水老师推荐的材料《A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixtures and HMM》(Jeff A. Bilmes)，我们延用其上下文，继续 M-step 之推导

将 Q-Function 中含有 μ, Σ 的项拿出来单独求极值

令 $B = \sum_{l=1}^{M} \sum_{i=1}^{N} \log p_l(x_i|\mu_l, \Sigma_l)\, p(l|x_i, \Theta^g)$

$$= \sum_{l=1}^{M} \left[\frac{1}{2}\ln|\Sigma^{-1}| \cdot \sum_{i=1}^{N} p(l|x_i, \Theta^g) - \frac{1}{2}\sum_{i=1}^{N} p(l|x_i, \Theta^g)\, tr(\Sigma^{-1} N_{l,i}) \right]$$

其中 $N_{l,i} = (x_i-\mu)(x_i-\mu)^T$

$$\frac{\partial B}{\partial \Sigma} = \sum_{l=1}^{M} \left[\frac{1}{2}\frac{1}{|\Sigma^{-1}|} \cdot \frac{\partial}{\partial \Sigma}|\Sigma^{-1}| \cdot \sum_{i=1}^{N} p(l|x_i, \Theta^g) - \frac{1}{2}\sum_{i=1}^{N} p(l|x_i, \Theta^g) \cdot \frac{\partial}{\partial \Sigma} tr(\Sigma^{-1} N_{l,i}) \right]$$

$$= \sum_{l=1}^{M} \left[-\frac{1}{2} I \sum_{i=1}^{N} p(l|x_i, \Theta^g) + \frac{\Sigma}{2} \sum_{i=1}^{N} p(l|x_i, \Theta^g)\, (N_{l,i}) \right]$$

$$\frac{\partial B}{\partial \Sigma} = 0 \Rightarrow \sum_{l=1}^{M} \left[-I \cdot \sum_{i=1}^{N} p(l|x_i, \Theta^g) + \Sigma \sum_{i=1}^{N} p(l|x_i, \Theta^g)\, N_{l,i} \right] = 0$$

$$\Rightarrow \Sigma = \frac{\sum_{l=1}^{M} \sum_{i=1}^{N} p(l|x_i, \Theta^g)}{\sum_{l=1}^{M} \sum_{i=1}^{N} p(l|x_i, \Theta^g) \cdot (x_i-\mu^g)(x_i-\mu^g)^T}$$