第7次作业

自硕21 崔晨菲 2021210976

1解：

(1) N个样本点的联合概率为

$$L(\alpha,\mu,\Sigma)=\prod_{i=1}^{N}\sum_{j=1}^{k}\alpha_j N(x|\mu_j,\Sigma_j)$$ 取对数为：

$$l(\alpha,\mu,\Sigma)=\sum_{i=1}^{N}\ln\left(\sum_{j=1}^{k}\alpha_j N(x|\mu_j,\Sigma_j)\right)$$

引入隐变量 $\gamma=(\gamma_1,\gamma_2,\cdots,\gamma_k)$，若x属于第j类，则 $\gamma_j=1$，其他都为0.

则 $$p(x,\gamma|\alpha,\mu,\Sigma)=\prod_{i=1}^{N}\prod_{j=1}^{k}\left(\alpha_j N(x_i|\mu_j,\Sigma_j)\right)^{\gamma_{i,j}}$$

$$=\prod_{j=1}^{k}\left(\alpha_j^{\sum_{i=1}^{N}\gamma_{i,j}}\prod_{i=1}^{N}\left(N(x_i|\mu_j,\Sigma_j)\right)^{\gamma_{ij}}\right)$$ 取对数有

$$\ln p(x,\gamma|\alpha,\mu,\Sigma)=\sum_{j=1}^{k}\left(\left(\sum_{i=1}^{N}\gamma_{i,j}\right)\ln\alpha_j+\sum_{i=1}^{N}\gamma_{i,j}\left(-\ln2\pi-\frac{1}{2}\ln|\Sigma_j|-\frac{1}{2}(x_i-\mu_j)^T(\Sigma_j)^{-1}(x_i-\mu_j)\right)\right)$$

后验期望为

$$E_\gamma\left(\ln p(x,\gamma|\alpha,\mu,\Sigma)\,|\,X,\alpha,\mu,\Sigma\right)$$

$$=\sum_{j=1}^{k}\left(\sum_{i=1}^{N}E(\gamma_{i,j}|x_i,\alpha,\mu,\Sigma)\ln\alpha_j+\sum_{i=1}^{N}E(\gamma_{i,j}|x_i,\alpha,\mu,\Sigma)\left(-\ln2\pi-\frac{1}{2}\ln|\Sigma_j|-\frac{1}{2}(x_i-\mu_j)^T(\Sigma_j)^{-1}(x_i-\mu_j)\right)\right)$$

其中 $$E(\gamma_{i,j}|x_i,\alpha,\mu,\Sigma)=p(\gamma_{i,j}=1|x_i,\alpha,\mu,\Sigma)=\frac{P(\gamma_{ij}=1,x_i|\alpha,\mu,\Sigma)}{p(x_i)}$$

$$=\frac{\alpha_j N(x_i|\mu_j,\Sigma_j)}{\sum_{j=1}^{k}\alpha_j N(x_i|\mu_j,\Sigma_j)}$$

(2) 为了将上式后验期望最大，我们随机初始化一批 (α,μ,Σ) 记为 $(\alpha^0,\mu^0,\Sigma^0)$

则我们有 $$\alpha^{t+1},\mu^{t+1},\Sigma^{t+1}=\arg\max E_\gamma\left(\ln P(X,\gamma|\alpha,\mu,\Sigma)\,|\,X,\alpha^t,\mu^t,\Sigma^t\right)$$

令导数为0，我们有

$$\alpha_j^{t+1} = \frac{\alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}{K \sum_{j=1}^{K} \alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}$$

$$\mu_j^{t+1} = \frac{\sum_{i=1}^{N} \frac{\alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}{\sum_{j=1}^{K} \alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)} x_i}{\frac{\alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}{\sum_{j=1}^{K} \alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}}$$

$$\Sigma_j^{t+1} = \frac{\sum_{i=1}^{N} \frac{\alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}{\sum_{j=1}^{K} \alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)} (x_i - \mu_j^t)^T (x_i - \mu_j^t)}{\frac{\alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}{\sum_{j=1}^{K} \alpha_j^t N(x_i | \mu_j^t, \Sigma_j^t)}}$$