

## 统计学习：第八章

**1** 如例 9.1 的三硬币模型。假设观测数据不变，试选择不同的初值，例如， $\pi^0 = 0.46, p^{(0)} = 0.55, q^{(0)} = 0.67$ ，求模型参数  $\theta = (\pi, p, q)$  的极大似然估计。

答：(0.4619, 0.5346, 0.6561)

**2** 证明引理 9.2。

答：

$$\begin{aligned} F(\tilde{P}, \theta) &= \sum_Z P(Z|Y, \theta) \log P(Y, Z|\theta) - \sum_Z P(Z|Y, \theta) \log P(Z|Y, \theta) \\ &= \sum_Z P(Z|Y, \theta) \log \frac{P(Y, Z|\theta)}{P(Z|Y, \theta)} \\ &= \sum_Z P(Z|Y, \theta) \log P(Y|\theta) \\ &= \log P(Y|\theta) \end{aligned}$$

**3** 已知观测数据 -67, -48, 6, 8, 14, 16, 23, 24, 28, 29, 41, 49, 56, 60, 75，试估计两个分量的高斯混合模型的 5 个参数。

答： $\alpha = 0.1332, \mu_1 = -57.5111, \sigma_1^2 = 90.2499, \mu_2 = 32.9849, \sigma_2^2 = 429.4583$ 。

**4** EM 算法可以用到朴素贝叶斯法的无监督学习，试写出其算法。

答：对于无监督学习，样本类别为隐变量  $y_i = 1, 2, \dots, K$ ，模型参数为  $\pi_k, \theta_{jk}, k = 1, 2, \dots, K, j = 1, 2, \dots, m, P(y_j) = \pi_k, P(x^j = a_{js}|y = k) = \theta_{jks}$ ，故有

1. 初始化参数

2. E 步骤，计算  $P(y = k|x, \pi, \theta)$  有

$$\begin{aligned} \gamma_{ik} &= P(y_i = k|x_i, \pi, \theta) \\ &= \frac{\pi_k \prod_{j=1}^m p(x_i^j|y_i = k, \theta)}{\sum_{k'=1}^K \pi_{k'} \prod_{j=1}^m p(x_i^j|y_i = k', \theta)} \end{aligned}$$

3. M 步骤, 更新  $\pi, \theta$

$$\pi_k^* = \frac{1}{N} \sum_{i=1}^N \gamma_{ik}$$
$$\theta_{jks}^* = \frac{\sum_{i=1}^N \gamma_{ik} I(x_i^j = a_j s)}{\sum_{s'} \sum_{i=1}^N \gamma_{ik} I(x_i^j = a_j s)}$$

3. 重复 2 和 3 步, 直至收敛。