Quiz 2 12110644 周[illegible]

1. $p(x) = \pi N(x|\mu_1, \Sigma_1) + (1-\pi) N(x|\mu_2, \Sigma_2)$

Given $\begin{cases} X = [x_1 \dots x_N] \\ T = [t_1 \dots t_N] \end{cases}$

1° What is the MLE of $\pi, \mu_1, \mu_2, \Sigma_1, \Sigma_2$?

Likelihood $p(t, X|\pi, \mu_1, \mu_2, \Sigma) = \prod_{n=1}^{N} [\pi N(x_n|\mu_1, \Sigma)]^{t_n} [(1-\pi) N(x_n|\mu_2, \Sigma)]^{1-t_n}$

$\Rightarrow \pi_{MLE} = \frac{1}{N} \sum_{n=1}^{N} t_n = \frac{N_1}{N_1+N_2}$

$\mu_{1MLE} = \frac{1}{N_1} \sum_{n=1}^{N} t_n x_n$

$\mu_{2MLE} = \frac{1}{N_2} \sum_{n=1}^{N} (1-t_n) x_n$

$\Sigma_{iMLE} = \frac{1}{N_i} \sum_{x_n \in X} (x_n - \mu_i)(x_n - \mu_i)^T \quad i = 1, 2$

2° $\pi \sim \text{Beta}(a_0, b_0)$, $\mu_1 \sim N(\mu_1|m_{10}, \Sigma_{10})$, $\mu_2 \sim N(\mu_2|m_{20}, \Sigma_{20})$

What is the MAP estimation of $\pi, \mu_1, \mu_2, \Sigma_1, \Sigma_2$?

$\Rightarrow \pi_{MAP} = \frac{N_1 + N_{10}}{N + N_0} = \frac{N_1 + N_{10}}{N_1 + N_2 + N_{10} + N_{20}}$

$\begin{cases} \Sigma_{iMAP}^{-1} = \Sigma_{iML}^{-1} + \Sigma_{i0}^{-1} \\ \Sigma_{iMAP}^{-1} \mu_{iMAP} = \Sigma_{iML}^{-1} \mu_{iML} + \Sigma_{i0}^{-1} \mu_{i0} \end{cases}$

$\Sigma = \pi \Sigma_1 + (1-\pi) \Sigma_2$

3° $p(C_1|x)$ for ML and MAP models, respectively?

$$P_{ML}(C_1|x) = \frac{P_{ML}(x, C_1)}{P_{ML}(x)} = \frac{\pi_{ML} N(x|\mu_{1ML}, \Sigma_{1ML})}{\pi_{ML} N(x|\mu_{1ML}, \Sigma_{1ML}) + (1-\pi_{ML}) N(x|\mu_{2ML}, \Sigma_{2ML})}$$

$$P_{MAP}(C_1|x) = \frac{P_{MAP}(x, C_1)}{P_{MAP}(x)} = \frac{\pi_{MAP} N(x|\mu_{1MAP}, \Sigma_{1MAP})}{\pi_{MAP} N(x|\mu_{1MAP}, \Sigma_{1MAP}) + (1-\pi_{MAP}) N(x|\mu_{2MAP}, \Sigma_{2MAP})}$$

2. $y = \sigma(w^T \phi(x))$. Given $\begin{cases} X = [x_1 \dots x_N] \\ T = [t_1 \dots t_N] \end{cases}$

1° What is the MLE of $q(w)$?

$p(C_0|\phi) = y(\phi) = \sigma(w^T \phi)$

$p(C_1|\phi) = 1 - p(C_0|\phi)$

$\frac{d\sigma(a)}{da} = \sigma(1-\sigma)$

$p(t|w) = \prod_{n=1}^{N} y_n^{t_n} (1-y_n)^{1-t_n}$

Likelihood $E(w) = -\ln p(t|w) = -\sum_{n=1}^{N} [t_n \ln y_n + (1-t_n) \ln(1-y_n)]$

$\nabla E(w) = \sum_{n=1}^{N} (y_n - t_n) \phi_n$

$H = \nabla\nabla E(w) = \sum_{n=1}^{N} y_n (1-y_n) \phi_n \phi_n^T$

$w_{ML} \leftarrow \quad w^{new} = w^{old} - H^{-1} \nabla E(w)$

$q(w) = N(w|w_{ML}, H^{-1})$

2° If $w \sim N(m_0, \Sigma_0)$, what is the MAP estimation of $q(w)$?

$$p(w|t) \propto p(w)\, p(t|w)$$

$$E(w) = -\ln p(w|t) = \frac{1}{2}(w-m_0)^T S_0^{-1}(w-m_0) - \sum_{n=1}^{N}[t_n \ln y_n + (1-t_n)\ln(1-y_n)]$$

$$\nabla E(w) = S_0^{-1}(w-m_0) + \sum_{n=1}^{N}(y_n - t_n)\phi_n$$

$$H = \nabla\nabla E(w) = S_0^{-1} + \sum_{n=1}^{N} y_n(1-y_n)\,\phi_n \phi_n^T$$

$$w_{MAP} \leftarrow w^{new} = w^{old} - H^{-1}\nabla E(w)$$

$$q(w) = N(w | w_{MAP}, H^{-1})$$

3° $p(t|y(w,x))$ for ML and MAP estimation, respectively?

$$p_{ML}(t|y(w,x)) = \prod_{i=1}^{N} y_i^{t_i}(1-y_i)^{1-t_i}, \quad y_i = 6(w_{ML}^T \phi_i)$$

$$p_{MAP}(t|y(w,x)) = \prod_{i=1}^{N} y_i^{t_i}(1-y_i)^{1-t_i}, \quad y_i = 6(w_{MAP}^T \phi_i)$$