

Homework 3

Your Name

October 31, 2018

-
- **Acknowledgments:** This template takes some materials from course CSE 547/Stat 548 of Washington University:
<https://courses.cs.washington.edu/courses/cse547/17sp/index.html>.
 - **Collaborators:** I finish my homework all by myself.
-

3.1. **Solution:** when $K = 2$, the conditional distribution of data x is

$$Pr(x^i|y^i = j; \theta) = \text{Norm}_x(\mu_j, \Sigma_j), \quad j \in \{1, \dots, K\} \quad (1)$$

the log likelihood function is

$$\begin{aligned} \mathcal{L} &= \log \prod_{i=1}^m Pr(x^i|y^i; \theta) Pr(y^i|\theta) \\ &= \sum_{i=1}^m \log Pr(x^i|y^i; \theta) + \log Pr(y^i|\theta) \end{aligned} \quad (2)$$

the derivative of \mathcal{L} by Σ_j is

$$\frac{\partial \mathcal{L}}{\partial \Sigma_j} = \sum_{i=1}^m \frac{1}{Pr(x^i|y^i; \theta)} \frac{\partial Pr(x^i|y^i; \theta)}{\partial \Sigma_j} \mathbb{1}\{y^i = j\} \quad (3)$$

as

$$\begin{aligned} \frac{\partial Pr(x^i|y^i; \theta)}{\partial \Sigma_j} &= \frac{1}{(2\pi)^{n/2}} e^{-0.5(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)} \\ &\quad * \left\{ \frac{\partial \frac{1}{|\Sigma_j|^{1/2}}}{\partial \Sigma_j} + \frac{1}{|\Sigma_j|^{1/2}} \frac{\partial (-0.5(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j))}{\partial \Sigma_j} \right\} \\ &= \frac{1}{(2\pi)^{n/2}} e^{-0.5(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)} \\ &\quad * \left\{ -\frac{1}{2} |\Sigma_j|^{-1/2} \Sigma_j^{-1} + \frac{1}{2} |\Sigma_j|^{-1/2} \Sigma_j^{-1} (x-\mu_j)(x-\mu_j)^T \Sigma_j^{-1} \right\} \\ &= \frac{1}{(2\pi)^{n/2}} e^{-0.5(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)} \\ &\quad * -\frac{1}{2} |\Sigma_j|^{-1/2} \Sigma_j^{-1} \{I - (x-\mu_j)(x-\mu_j)^T \Sigma_j^{-1}\} \end{aligned} \quad (4)$$

combine the (??) and (??), then let $\frac{\partial \mathcal{L}}{\partial \Sigma_j} = 0$, we get

$$\begin{aligned}
-\frac{1}{2} \sum_{i=1}^m \mathbb{1}\{y^i = j\} \Sigma_j^{-1} [I - (x - \mu_j)(x - \mu_j)^T \Sigma_j^{-1}] &= 0 \\
\Sigma_j \sum_{i=1}^m \mathbb{1}\{y^i = j\} &= \sum_{i=1}^m \mathbb{1}\{y^i = j\} (x - \mu_j)(x - \mu_j)^T
\end{aligned} \tag{5}$$

therefore the maximum likelihood estimate of Σ_j for the case of $K = 2$ in QDA model is

$$\hat{\Sigma}_j = \frac{\sum_{i=1}^m \mathbb{1}\{y^i = j\} (x^i - \mu_j)(x^i - \mu_j)^T}{\sum_{i=1}^m \mathbb{1}\{y^i = j\}} \tag{6}$$