

Let $z_{il} = I(y_i = l)$, $i = 1, \dots, N$, $l = 1, \dots, K$

The log-likelihood function is given by:

$$\begin{aligned}
 l(\hat{p}_1, \dots, \hat{p}_{K+1}) &= \log\left(\prod_{i=1}^N \prod_{l=1}^K \hat{p}_l(x_i)^{z_{il}}\right) \\
 &= \sum_{i=1}^N \sum_{l=1}^K z_{il} \log(\hat{p}_l(x_i)) \quad , \quad \hat{p}_l(x_i) = \sum_{m=1}^{|\mathcal{T}|} \hat{p}_{m,l} I(\underline{x}_i \in R_m) \\
 &= \sum_{i=1}^N \sum_{l=1}^K z_{il} \sum_{m=1}^{|\mathcal{T}|} I(\underline{x}_i \in R_m) \log(\hat{p}_{m,l}) \\
 &= \sum_{m=1}^{|\mathcal{T}|} \sum_{l=1}^K \log(\hat{p}_{m,l}) \sum_{i=1}^N z_{il} I(\underline{x}_i \in R_m) \\
 &= \sum_{m=1}^{|\mathcal{T}|} \sum_{l=1}^K \log(\hat{p}_{m,l}) \sum_{\{i: \underline{x}_i \in R_m\}} z_{il}
 \end{aligned}$$

Recall that

$$\hat{p}_{m,l} = \frac{1}{N_m} \sum_{\{i: \underline{x}_i \in R_m\}} \frac{I(y_i = l)}{z_{il}} = \frac{1}{N_m} \sum_{\{i: \underline{x}_i \in R_m\}} z_{il}$$

$$\longrightarrow \sum_{\{i: \underline{x}_i \in R_m\}} z_{il} = N_m \hat{p}_{m,l}$$

$$\begin{aligned}
 \longrightarrow l(\hat{p}_1, \dots, \hat{p}_{K+1}) &= \sum_{m=1}^{|\mathcal{T}|} \sum_{l=1}^K N_m \hat{p}_{m,l} \log(\hat{p}_{m,l}) = \sum_{m=1}^{|\mathcal{T}|} N_m \underbrace{\sum_{l=1}^K \hat{p}_{m,l} \log(\hat{p}_{m,l})}_{Q_m(T)} \\
 &= \sum_{m=1}^{|\mathcal{T}|} N_m Q_m(T) \\
 Q_m(T) &= \sum_{l=1}^K \hat{p}_{m,l} \log(\hat{p}_{m,l})
 \end{aligned}$$

Let $\hat{p}_{m, h(m)} = \max_u \hat{p}_{m, u}$, so y_i is classified as class $h(m)$, i.e. $\hat{y}_i = h(m)$, in region R_m . We have

$$\begin{aligned}
 Q_n(T) &= \frac{1}{N_m \{i: \underline{x}_i \in R_m\}} \sum \mathbb{I}(y_i \neq \hat{y}_i) = \frac{1}{N_m \{i: \underline{x}_i \in R_m\}} \sum (1 - \mathbb{I}(y_i = \hat{y}_i)) \\
 &= \underbrace{\frac{1}{N_m \{i: \underline{x}_i \in R_m\}} \sum_{i=1}^{N_m} 1}_{=1} - \frac{1}{N_m \{i: \underline{x}_i \in R_m\}} \sum \mathbb{I}(y_i = \hat{y}_i) \\
 &= 1 - \frac{1}{N_m \{i: \underline{x}_i \in R_m\}} \sum \mathbb{I}(y_i = h(m)) \\
 &= 1 - \hat{p}_{m, h(m)}
 \end{aligned}$$