

Task 4) $L(w) = \frac{1}{n} \sum_{i=1}^n l(f_w(x_i), y_i) \quad \textcircled{I}$

$$l(f_w(x_i), y_i) = -\log P(y=y_i | x_i, w)$$

$$= \sum_{j=1}^k \left[\log \left(\frac{e^{f_w(x_i)_j}}{\sum_{l=1}^k e^{f_w(x_i)_l}} \right) \right] \delta_{y_i, j}$$

For each $1 \leq m \leq k \implies l(f_w(x_i)_j, y_i) = -\delta_{y_i, m} \log \frac{e^{f_w(x_i)_m}}{\sum_{l=1}^k e^{f_w(x_i)_l}} \quad \textcircled{II}$

$$f_w(x) = w x \quad \textcircled{III}$$

$$\begin{aligned} \textcircled{I}, \textcircled{II}, \textcircled{III} \implies L(w_m) &= \frac{1}{n} \sum_{i=1}^n \delta_{y_i, m} \log \frac{e^{w_m x_i}}{\sum_{l=1}^k e^{w_l x_i}} \\ &= \frac{1}{n} \sum_{i=1}^n \delta_{y_i, m} (w_m x_i - \log \sum_{l=1}^k e^{w_l x_i}) \end{aligned}$$

$$\begin{aligned} \implies \frac{dL(w_m)}{dw_m} &= \frac{1}{n} \sum_{i=1}^n \delta_{y_i, m} x_i - \frac{(x_i) e^{w_m x_i}}{\sum_{l=1}^k e^{w_l x_i}} \\ &= \frac{1}{n} \sum_{i=1}^n \left(\delta_{y_i, m} - \frac{e^{w_m x_i}}{\sum_{l=1}^k e^{w_l x_i}} \right) x_i \end{aligned}$$