Assignment 1 TDT4265 Fredrik Almås

1\. a) $$\frac{\partial C^n(w)}{\partial w_i} = \frac{\partial C^n(w)}{\partial f(x^n)} \frac{\partial f(x^n)}{\partial w_i}$$

$$C^n(w) = -\left(y^n \ln(\hat{y}^n) + (1-y^n)\ln(1-\hat{y}^n)\right)$$

$$\hat{y}^n = f(x^n)$$

$$C^n(w) = -\left(y^n \ln(f(x^n)) + (1-y^n)\ln(1-f(x^n))\right)$$

$$\frac{\partial C^n(w)}{\partial f(x^n)} = -\left(\frac{y^n}{f(x^n)} + (1-y^n)\left(-\frac{1}{1-f(x^n)}\right)\right)$$

$$= -\frac{y^n}{f(x^n)} + \frac{1-y^n}{1-f(x^n)}$$

$$\frac{\partial f(x^n)}{\partial w_i} = x_i^n f(x^n)(1-f(x^n))$$

$$\frac{\partial C^n(w)}{\partial w_i} = \left(-\frac{y^n}{f(x^n)} + \frac{1-y^n}{1-f(x^n)}\right) x_i^n f(x^n)(1-f(x^n))$$

$$= \left(-y^n(1-f(x^n)) + (1-y^n)f(x^n)\right) x_i^n$$

$$= \left(-y^n + \cancel{y^n f(x^n)} + f(x^n) - \cancel{y^n f(x^n)}\right) x_i^n$$

$$= \left(-y^n + f(x^n)\right) x_i^n$$

$$\overset{\hat{y}^n = f(x^n)}{=} \underline{-(y^n - \hat{y}^n)x_i^n}$$

1. b) $$\frac{\partial C^n(w)}{\partial w_{kj}} = \frac{\partial C^n(w)}{\partial z_l} \frac{\partial z_l}{\partial w_{kj}}$$

$$z_l = \sum_i^I w_{ki} \cdot x_i^n$$

$$\frac{\partial z_l}{\partial w_{kj}} = \frac{\partial \sum_i^I w_{ki} x_i^n}{\partial w_{kj}} = \frac{\partial \left(\sum_{i \neq j}^I w_{ki} x_i^n + w_{kj} x_j^n \right)}{\partial w_{kj}} = x_j^n$$

$$C^n(w) = -\sum_{k=1}^K y_k^n \ln(\hat{y}_k^n)$$

$$\hat{y}_k^n = \frac{e^{z_k}}{\sum_{k'}^K e^{z_{k'}}}$$

$$\frac{\partial C^n(w)}{\partial z_l} = \frac{\partial \left(-\sum_{k=1}^K y_k^n \ln(\hat{y}_k^n) \right)}{\partial z_l}$$

$$= -\sum_{k=1}^K y_k^n \frac{\partial \ln(\hat{y}_k^n)}{\partial z_l} = -\sum_{k=1}^K y_k^n \frac{1}{\hat{y}_k^n} \frac{\partial \hat{y}_k^n}{\partial z_l}$$

$$\frac{\partial \hat{y}_k^n}{\partial z_l} = \frac{\partial \frac{e^{z_k}}{\sum_{k'}^K e^{z_{k'}}}}{\partial z_l}$$

for $l = k$:

$$\frac{\partial \frac{e^{z_k}}{\sum_{k'}^K e^{z_{k'}}}}{\partial z_k} = \frac{e^{z_k} \sum_{k'}^K e^{z_{k'}} - \left(e^{z_k}\right)^2}{\left(\sum_{k'}^K e^{z_{k'}}\right)^2}$$

$$= \frac{e^{z_k}}{\sum_{k'}^{K} e^{z_{k'}}} \left(\frac{\sum_{k'}^{K} e^{z_{k'}} - e^{z_k}}{\sum_{k'}^{K} e^{z_k}} \right) = \hat{y}_k^n (1 - \hat{y}_k^n)$$

for $l \neq k$

$$\frac{\partial \frac{e^{z_k}}{\sum_{k'}^{K} e^{z_{k'}}}}{\partial z_l} = \frac{0 \cdot \sum_{k'}^{K} e^{z_{k'}} - e^{z_k} e^{z_l}}{\left(\sum_{k'}^{K} e^{z_{k'}}\right)^2}$$

$$= - \frac{e^{z_k}}{\sum_{k'}^{K} e^{z_{k'}}} \cdot \frac{e^{z_l}}{\sum_{k'}^{K} e^{z_{k'}}} = -\hat{y}_k^n \hat{y}_l^n$$

$$\frac{\partial \hat{y}_k^n}{\partial z_l} = \begin{cases} \hat{y}_k^n (1 - \hat{y}_k^n) & \text{for } l = k \\ -\hat{y}_k^n \hat{y}_l^n & \text{for } l \neq k \end{cases}$$

$$\frac{\partial C^n(w)}{\partial z_l} = - \sum_{k=1}^{K} \frac{y_k^n}{\hat{y}_k^n} \frac{\partial \hat{y}_k^n}{\partial z_l}$$

$$= - \sum_{k \neq l}^{K} \frac{y_k^n}{\hat{y}_k^n} (-\hat{y}_k^n \hat{y}_l^n) - \frac{y_l^n}{\hat{y}_l^n} \hat{y}_l^n (1 - \hat{y}_l^n)$$

$$= \sum_{k \neq l}^{K} y_k^n \hat{y}_l^n - y_l^n + y_l^n \hat{y}_l^n$$

$$= -y_l^n + \sum_{k=1}^{K} y_k^n \hat{y}_l^n$$

$$= -y_l^n + \hat{y}_l^n \sum_{k=1}^{K} y_k^n = -y_l^n + \hat{y}_l^n$$

$$\frac{\partial C^n(w)}{\partial w_{kj}} = (-y_l^n + \hat{y}_l^n) x_j^n = -x_j^n (y_l^n - \hat{y}_l^n)$$

$$= -x_j^n (y_k^n - \hat{y}_k^n)$$

as both k and j are arbitiary integers between 1 and K.