CS 7150 Homework #2

1.

1.1 gradient of loss w.r.t weight vector w_i

$$\frac{\partial l}{\partial w_i} = \frac{\partial l}{\partial p_k} \cdot \frac{\partial p_k}{\partial z_i} \cdot \frac{\partial z_i}{\partial w_i} \quad \text{(Chain Rule)}$$

First term $\frac{\partial l}{\partial p_k}$

Since $l = -\log p_c \Rightarrow \frac{\partial l}{\partial p_k} = \frac{-1}{p_k}$

Second term $\frac{\partial p_k}{\partial z_i}$

$$p_c = \frac{e^{z_k}}{\sum_{j=1}^{c} e^{z_j}}$$

$$\Rightarrow \frac{\partial p_k}{\partial z_i} = \frac{\partial \left(\frac{e^{z_k}}{\sum_{j=1}^{c} e^{z_j}} \right)}{\partial z_i}$$

By quotient rule of derivative for $f(x) = \frac{g(x)}{h(x)}$:

$$f'(x) = \frac{g'(x)\,h(x) - h'(x)\,g(x)}{[h(x)]^2}$$

In our case we have

$$g(k) = e^{z_k}$$

$$h(k) = \sum_{j=1}^{c} e^{z_j}$$

No matter which z_i we compute the derivative of $h(k)$ for the answer will always be e^{z_i} (differentiate each term individually, $\frac{\partial e^{z_j}}{\partial z_i} = 0$ when $j \neq i$ and $\frac{\partial e^{z_j}}{\partial z_i} = e^{z_i}$ when $j = i$)

However with $g'(x) = \frac{\partial e^{z_k}}{\partial z_i}$, it will be equal to e^{z_k} when $k=i$ and 0 when $k \neq i$

$$\Rightarrow \frac{\partial e^{z_k}}{\partial z_i} = e^{z_i} \quad \text{when } k=i$$

$$= 0 \quad \text{when } k \neq i$$

For simplicity let Σ stand for $\sum_{j=1}^{c} e^{z_j}$

(*) Case $k = i$

$$\frac{\partial p_k}{\partial z_i} = \frac{e^{z_i} . \Sigma - e^{z_k} . e^{z_i}}{\Sigma^2}$$

$$= \frac{e^{z_i} (\Sigma - e^{z_k})}{\Sigma . \Sigma} = \frac{e^{z_i}}{\Sigma} . \frac{\Sigma - e^{z_k}}{\Sigma}$$

$$= \frac{e^{z_i}}{\Sigma} \left(1 - \frac{e^{z_k}}{\Sigma}\right)$$

We know $p_i = \frac{e^{z_i}}{\Sigma}$

$$\Rightarrow \frac{\partial p_k}{\partial z_i} = p_i (1 - p_k)$$

(*) Case $k \neq i$

$$\frac{\partial p_k}{\partial z_i} = \frac{0 . \Sigma - e^{z_k} . e^{z_i}}{\Sigma^2}$$

$$= \frac{-e^{z_k} . e^{z_i}}{\Sigma^2} = \frac{-e^{z_k}}{\Sigma} . \frac{e^{z_i}}{\Sigma} = -p_k . p_i$$

$z_i = w_i^T x + b_i$

$\Rightarrow \dfrac{\partial z_i}{\partial w_i} = x$

Combine all 3 terms

$\Rightarrow \dfrac{\partial \ell}{\partial w_i} = \dfrac{-1}{p_k} \cdot p_i(1-p_k) \cdot x$

$= -(1-p_k) \cdot x = (p_k - 1) \cdot x$ when $k = i$

$\dfrac{\partial \ell}{\partial w_i} = \dfrac{-1}{p_k} \cdot (-p_k p_i) \cdot x = -p_i \cdot x$ when $k \neq i$

1.2 gradient of ℓ w.r.t b_i

$$\frac{\partial \ell}{\partial b_i} = \frac{\partial \ell}{\partial p_k} \cdot \frac{\partial p_k}{\partial z_i} \cdot \frac{\partial z_i}{\partial b_i}$$

Similarly to 1.1 we have computed $\dfrac{\partial \ell}{\partial p_k}$ and $\dfrac{\partial p_k}{\partial z_i}$

$z_i = w_i^T x + b_i \Rightarrow \dfrac{\partial z_i}{\partial b_i} = 1$.

$\Rightarrow \dfrac{\partial \ell}{\partial b_i} = \dfrac{-1}{p_k} \cdot p_i(1-p_k) \cdot 1 = p_k - 1$ when $k = i$

$\dfrac{\partial \ell}{\partial b_i} = \dfrac{-1}{p_k} \cdot (-p_k p_i) = p_i$ when $k \neq i$

2.

2.1

$$\frac{\partial \tilde{\ell}}{\partial w_i} = \frac{\partial \tilde{\ell}}{\partial \tilde{p}_c} \cdot \frac{\partial \tilde{p}_c}{\partial z_i} \cdot \frac{\partial z_i}{\partial w_i}$$

$$\tilde{\ell} = -\log(\tilde{p}_c) \Rightarrow \frac{\partial \tilde{\ell}}{\partial \tilde{p}_c} = \frac{-1}{\tilde{p}_c}$$

$$\tilde{p}_c = \frac{e^{z_c}}{e^{z_c} + \frac{1}{K}\sum_{j \in S} q_j e^{z_j}}$$

$$\frac{\partial \tilde{p}_c}{\partial z_i} = \frac{\partial}{\partial z_i}\left(\frac{e^{z_c}}{e^{z_c} + \frac{1}{K}\sum_{j \in S} q_j e^{z_j}}\right)$$

Quotient rule of derivative $f(x) = \frac{g(x)}{h(x)}$

$$f'(x) = \frac{\frac{\partial g(x)}{\partial f(x)} h(x) - \frac{\partial h(x)}{\partial f(x)} g(x)}{[h(x)]^2}$$

When $i = c$ $\quad \frac{\partial g(x)}{\partial f(x)} = \frac{\partial(e^{z_c})}{\partial z_i} = e^{z_c}$

$$\frac{\partial h(x)}{\partial f(x)} = \frac{\partial\left(e^{z_c} + \frac{1}{K}\sum_{j \in S} q_j e^{z_j}\right)}{\partial z_i}$$

$$= e^{z_c}$$

When $i \neq c$ $\quad \frac{\partial(e^{z_c})}{\partial z_i} = 0$

$$\frac{\partial h(x)}{\partial f(x)} = \frac{1}{K} \frac{e^{z_i}}{q_i}$$ (where $j = i$ only ~~e^{z_i}~~ ~~...~~, the rest will be 0)

when $i = c$, simplify $\sum_{j \in S} \frac{e^{z_j}}{q_j}$ as Σ

$$\Rightarrow \frac{\partial \tilde{p}_c}{\partial z_i} = \frac{e^{z_c} \cdot \left(e^{z_c} + \frac{1}{k}\Sigma\right) - e^{z_c} \cdot e^{z_c}}{\left(e^{z_c} + \frac{1}{k}\Sigma\right)^2}$$

$$= \underbrace{\frac{e^{z_c}}{e^{z_c} + \frac{1}{k}\Sigma}}_{= \tilde{p}_c} \cdot \left(1 - \underbrace{\frac{e^{z_c}}{e^{z_c} + \frac{1}{k}\Sigma}}_{= \tilde{p}_c}\right)$$

$$= \tilde{p}_c (1 - \tilde{p}_c)$$

when $i \neq c$

$$\Rightarrow \frac{\partial \tilde{p}_c}{\partial z_i} = \frac{0 - e^{z_c} \cdot \frac{1}{k} \frac{e^{z_j}}{q_i}}{\left(e^{z_c} + \frac{1}{k}\Sigma\right)^2}$$

$$= - \underbrace{\frac{e^{z_c}}{\left(e^{z_c} + \frac{1}{k}\Sigma\right)}}_{= \tilde{p}_c} \cdot \frac{\frac{1}{k} \frac{e^{z_i}}{q_i}}{\left(e^{z_c} + \frac{1}{k}\Sigma\right)}$$

$$= \tilde{p}_c \cdot \frac{1}{k} \cdot \frac{e^{z_i}}{q_i \left(e^{z_c} + \frac{1}{k}\Sigma\right)}$$

$$= \tilde{p}_c \cdot \frac{\tilde{p}_i}{k \cdot q_i}$$

when $i=c$ $\quad \frac{\partial z_i}{\partial w_i} = x$

$i \neq c$ $\quad \frac{\partial z_i}{\partial w_i} = 1$

$$\Rightarrow \frac{\partial \tilde{\ell}}{\partial w_i} = -\frac{1}{\tilde{p}_c} \cdot \tilde{p}_c(1-\tilde{p}_c) \cdot x$$

$$= (\tilde{p}_c - 1).x \quad \text{when \& } i=c$$

$$\frac{\partial \tilde{\ell}}{\partial w_i} = -\frac{1}{\tilde{p}_c} \cdot \tilde{p}_c \cdot \frac{\tilde{p}_i}{K.q_i} . x = \frac{-\tilde{p}_i}{K.q_i} . x \quad \text{when } i \neq c$$

2.2

$$\frac{\partial \tilde{\ell}}{\partial b_i} = \cancel{\frac{\partial \tilde{\ell}}{\partial b_i}} \; \frac{\partial \tilde{\ell}}{\partial \tilde{p}_c} \cdot \frac{\partial \tilde{p}_c}{\partial z_i} \cdot \underbrace{\frac{\partial z_i}{\partial b_i}}_{=1}$$

$$\Rightarrow \frac{\partial \tilde{\ell}}{\partial b_i} = -\frac{1}{\tilde{p}_c} \cdot \tilde{p}_c(1-\tilde{p}_c) \cdot 1 = \tilde{p}_c - 1 \quad \text{when } i=c$$

$$\frac{\partial \tilde{\ell}}{\partial b_i} = -\frac{1}{\tilde{p}_c} \cdot \tilde{p}_c \cdot \frac{\tilde{p}_i}{K.q_i} \cdot 1 = \frac{-\tilde{p}_i}{K.q_i}$$