최윤서

<assignment 2>

1-(a) $-\sum_{w \in Vocab} y_w \log(\hat{y}_w) = -\log(\hat{y}_o)$. $y \rightarrow$ 실제 확률분포
$\hat{y} \rightarrow$ 모델 확률분포.

y는 context word에 해당하는 element가 1인 one-hot vector

$\boxed{\therefore \text{true.}}$

1-(b) $J(v_c, o, U) = -\log P(O=o \mid C=c)$

$P(O=o \mid C=c) = \dfrac{\exp(u_o^T v_c)}{\sum_w \exp(u_w^T v_c)}$ the probability ~~that~~ of when center word is given

given ⇒ center w
| outside w

$$J(v_c, o, U) = -\log P(O=o \mid C=c)$$
$$= -\log \frac{\exp(u_o^T v_c)}{\sum_w \exp(u_w^T v_c)}$$
$$= -u_o^T v_c + \log \sum_w \exp(u_w^T v_c)$$

reminders are all zeros while y equals to 1 (merely for the answer)

$$\frac{dJ}{dv_c} = -u_o + \sum_w \frac{\exp(u_w^T v_c)\, u_w}{\sum \exp(u_w^T v_c)} = -u_o + \sum_w P(O=w \mid C=c)\, u_w$$
$$= -u_o + \sum_w \hat{y}_w u_w$$
$$\boxed{\therefore = u(\hat{y} - y)}$$

1-(c) i) When the weight ans is outside w. $(W=0)$

$$\frac{\partial J}{\partial w} = -v_c + \frac{d}{du_w}\left(\log \sum_w \exp(u_w^T v_c)\right) = -v_c + \frac{(u_w^T v_c)}{\sum \exp(u_w^T v_c)} \cdot \frac{d(u_w^T v_c)}{du_w}$$

$$= v_c + P(O=w \mid C=c) \cdot v_c$$
$$\boxed{\therefore = v_c(\hat{y} - y)}$$

ii) When outside W is wrong ans.

$$\frac{dJ}{du_w} = -\frac{d}{du_w}(u_o^T v_c) + \frac{d}{du_w}\left(\log \sum_w \exp(u_w^T v_c)\right)$$
$$0 + \frac{u_w^T v_c}{\sum_w \exp(u_w^T v_c)} \cdot \frac{d}{du_w}(u_w^T v_c)$$
$$= P(O=w \mid C=c) \cdot v_c$$
$$\boxed{\therefore = v_c \cdot \hat{y}}$$

1-(d) $\sigma(x) = \frac{1}{1+e^{-x}} = \frac{e^{x}}{e^{x}+1}$

$$\frac{d}{d(x)}\sigma(x) = \frac{d}{dx}(1+e^{-x})^{-1} = (-1)\cdot(1+e^{-x})^{-2}\cdot\frac{d}{dx}(1+e^{-x})$$

$$= -1\cdot\frac{1}{(1+e^{-x})^2}\cdot(-e^{-x})$$

$$= \frac{e^{-x}}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}}\cdot\frac{e^{-x}}{1+e^{-x}}$$

$$= \underbrace{\frac{1}{1+e^{-x}}}_{=\sigma(x)}\cdot\Big(1-\underbrace{\frac{1}{1+e^{-x}}}_{=\sigma(x)}\Big)$$

$$= \sigma(x)\cdot(1-\sigma(x))$$

$$\boxed{\therefore\ \sigma(x)\cdot(1-\sigma(x))}$$

☆ 1-(e) $J_{neg\text{-}sample}(v_c, o, U) = -\log(\sigma(u_o^T v_c)) - \sum_{k=1}^{K}\log(\sigma(-u_k^T v_c))$

※ $\sigma(\cdot) \Rightarrow$ sigmoid func.

풀이 참고함.

이해하는 X.

i) $\frac{dJ}{dv_c} = -\frac{1}{\sigma(u_o^T v_c)}\cdot\frac{d}{dv_c}\sigma(u_o^T v_c) - \sum_{k=1}^{K}\frac{1}{\sigma(-u_k^T v_c)}\frac{d}{dv_c}\sigma(-u_k^T v_c)$

$\therefore = -(1-\sigma(u_o^T v_c))\cdot u_o - \sum_{k=1}^{K}(1-\sigma(u_k^T v_c))\,u_k$

ii) $\frac{dJ}{du_o} = -\frac{1}{\sigma(u_o^T v_c)}\frac{d}{du_o}(u_o^T v_c) - \underbrace{\sum_{k=1}^{K}\frac{d}{du_o}\log\sigma(-u_k^T v_c)}$

neg 이므로 모든 k에 대해 0

$= -(1-\sigma(u_o^T v_c))\cdot v_c$

$= (\sigma(u_o^T v_c)-1)\,v_c$

☆ 1-(f) i) $\frac{d}{dU}J_{skipgram}(v_c, w_{t-m}, \ldots w_{t+m}, U) = \frac{d}{dU}\sum_{\substack{-m\le j\le m\\ j\ne 0}} J(v_c, w_{t+j}, U)$

ii) $\frac{d}{dv_c}J_{skipgram}(v_c, w_{t-m}, \ldots w_{t+m}, U) = \frac{d}{dv_c}\sum_{\substack{-m\le j\le m\\ j\ne 0}} J(v_c, w_{t+j}, U)$

iii) $\frac{d}{dv_w}J_{skipgram}(v_c, w_{t-m} \ldots w_{t+m}, U) = \frac{d}{dv_w}\sum_{\substack{-m\le j\le m\\ j\ne 0}} J(v_c, w_{t+j}, U),\ w\ne$