CS 224N : Assignment 2

1.a) $y_w = 0 \quad \forall\ w \in \text{Vocab} / \{o\}$

b) $J = -\log\left(\frac{e^{u_o' v_c}}{\sum_w e^{u_w' v_c}}\right)$

$= -\log(e^{u_o' v_c}) + \log\left(\sum_w e^{u_w' v_c}\right)$

$\frac{\partial J}{\partial v_c} = \frac{-1}{e^{u_o' v_c}} \cdot e^{u_o' v_c} \cdot u_o$

$+$

$\frac{1}{\sum_w e^{u_w' v_c}} \cdot \sum_w e^{u_w' v_c} \cdot u_w$

$= -u_o + \frac{1}{\sum_w e^{u_w' v_c}} \cdot \sum_w u_w \cdot e^{u_w' v_c}$

c) $$\frac{\partial J}{\partial u_o} = \frac{-1}{e^{u_o' v_c}} \cdot e^{u_o' v_c} \cdot v_c$$

$$+$$

$$\frac{1}{\sum_w e^{u_w' v_c}} \cdot e^{u_o' v_c} \cdot v_c$$

$$= \left(\frac{e^{u_o' v_c}}{\sum_w e^{u_w' v_c}} - 1 \right) \cdot v_c$$

$$\frac{\partial J}{\partial u_2} = \left(\frac{e^{u_2' v_c}}{\sum_w e^{u_w' v_c}} - 1 \right) \cdot v_c \quad (\text{where } z \neq o)$$

d) $$\frac{\partial J}{\partial U} = \begin{bmatrix} [\partial J / \partial u_1]^T \\ [\partial J / \partial u_2]^T \\ \vdots \\ [\partial J / \partial u_{|vocab|}]^T \end{bmatrix}$$

Where $\partial J / \partial U_k$ is given in c).

e) $\sigma(x) = \dfrac{e^x}{1+e^x}$

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

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

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

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