Homework_2:

Lv_Xinpeng

Part_1:

a) if w=o, y_w=1, else : y_w=0 so:

$$-\sum_{w\in Vocab} y_w \log(\hat{y}_w) = -y_o \log(\hat{y}_o) = -\log(\hat{y}_o)$$

b)

$$\frac{\partial J}{\partial v_c} = \frac{\partial\left(-\log\frac{\exp(u_o^T v_c)}{\sum_{w\in Vocab}\exp(u_w^T v_c)}\right)}{\partial v_c}$$

$$\frac{\partial J}{\partial v_c} = \frac{\partial(-u_o^T v_c)}{\partial v_c} - \frac{\partial(-\log(\sum_{w\in Vocab}\exp(u_w^T v_c)))}{\partial v_c}$$

$$\frac{\partial J}{\partial v_c} = -u_o + \frac{1}{\sum_{w\in Vocab}\exp(u_w^T v_c)} \bullet \sum_{i\in Vocab}\exp(u_i^T v_c) u_i$$

$$\frac{\partial J}{\partial v_c} = U(\hat{y} - y)$$

c) if w=o, we have:

$$\frac{\partial J}{\partial u_w} = \frac{\partial\left(-\log\frac{\exp(u_o^T v_c)}{\sum_{w\in Vocab}\exp(u_o^T v_c)}\right)}{\partial u_o}$$

$$\frac{\partial J}{\partial u_w} = -v_c + \frac{\exp(u_o^T v_c)\cdot v_c}{\sum_{w\in Vocab}\exp(u_o^T v_c)}$$

$$\frac{\partial J}{\partial u_w} = (-y + \hat{y}) v_c$$

if $w \neq o$, y = 0 , so we have:

$$\frac{\partial J}{\partial u_w} = \hat{y} v_c$$

d) in sigmoid fuction： y’=y(1-y), so:

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