Natural language Processing

Assignment - 2

a) y is the true distribution and the predicted distribution is $\hat{y}$.

$$\therefore (\hat{y})_k = \frac{\exp(u_k^T v_c)}{\sum_{i=\phi}^{V} \exp(u_i^T v_c)}$$

where u_k is the kth row of U. or also called the output vector for word with index k.

Suppose the true outside word is o

$$\therefore (y)_o = 1$$

$$(y)_w = 0 \quad \text{if } w \neq o$$

$$\therefore \text{Cross entropy} = -\sum_{w \in \text{Vocab}} y_w \log(\hat{y}_w)$$

$$= -\sum_{w \neq 0} y_w \log(\hat{y}_w) - y_0 \log(\hat{y}_0)$$

$\therefore \quad y_w = 0 \quad \forall \; w \neq 0$

$\Rightarrow \quad = -y_0 \log(\hat{y}_0)$

$\therefore \quad y_0 = 1$

$\Rightarrow \quad = -\log(\hat{y}_0)$

Hence proved.

b) $\quad J = -\log P(O=o \mid C=c)$

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

$$= -u_o^T v_c + \log\left(\sum_w \exp(u_w^T . v_c)\right)$$

$$\frac{\partial J}{\partial v_c} = -u_o + \frac{1}{\sum_w \exp(u_w^T v_c)} \times \sum_w u_w \exp(u_w^T v_c)$$

$$\frac{\partial J}{\partial v_c} = -u_o + \sum_w u_w \frac{\exp(u_w^T v_c)}{\sum_{w1} \exp(u_{w1}^T v_c)}$$

$$\frac{\partial J}{\partial v_{dc}} = \sum_w u_w \hat{y}_w - u_o$$

$$\boxed{\frac{\partial J}{\partial v_{dc}} = E[u_w] - u_o}$$

2) (case 1) $w \neq 0$

$$J = -u_o^T v_c + \log\left(\sum_k \exp u_k^T v_c\right)$$

$$\frac{\partial J}{\partial u_w} = 0 + \frac{1}{\quad}$$

$$J = -u_o^T v_c + \log\left(\sum_k \exp(u_k^T v_c)\right)$$

$$\frac{\partial J}{\partial u_w} = 0 + \frac{1}{\sum_k \exp(u_k^T v_c)} \times \frac{\partial}{\partial u_w} v_c \exp(u_w^T v_c)$$

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

Case 2 : $w = 0$

$$J = -\mu_o^T v_c + \log\left(\sum_w \exp(\mu_w^T v_c)\right)$$

$$\frac{\partial J}{\partial \mu_o} = -v_c + \frac{1}{\sum_w \exp(\mu_w^T v_c)} \times \sum_w \frac{\partial}{\partial \mu_o} \exp(\mu_w^T v_c)$$

$$= -v_c + \frac{1}{\sum_w \exp(\mu_w^T v_c)} \times v_c \exp(\mu_o^T v_c)$$

$$\boxed{\frac{\partial J}{\partial \mu_o} = -v_c(1 - \hat{y}_o)}$$

Coding ~~Part~~ Part

c) In the plot I can see several clusters of words that make sense, such as tea & coffee, enjoyable & annoying, woman & female, etc.

However many things don't make sense, like how male is away from man, earth is close to man.

Also snow & woman are close, which does not make sense. We can also see the bias of the model, by observing that the word dumb is closer to queen than king.