Yiwen Zhang

YZ3310

Hw 2.

Question 1

① $L = \sum_i v \cdot f(x_i, y_i) - \sum_i \log \sum_{y' \in V} e^{v \cdot f(x_i, y')} - C \sum_k v_k^2$

$$\frac{\partial L}{\partial v_1} = \sum_i f_1(x_i, y_i) - \sum_i \frac{\sum_y f(x_i, y)\, e^{v \cdot f(x_i, y)}}{\sum_{y' \in V} e^{v \cdot f(x, y')}} - 2Cv_1 = \sum_i f_1(x_i, y_i) - \sum_i \sum_y f(x_i, y) P(y | x_i, v) - 2Cv_1$$

$$\frac{\partial L}{\partial v_2} = \sum_i f_2(x_i, y_i) - \sum_i \sum_{y \in V} f_2(x_i, y) P(y | x_i, v) - 2Cv_2$$

for optimum parameters v^*, we know $\frac{\partial L}{\partial v_1}\Big|_{v^*} = \frac{\partial L}{\partial v_2}\Big|_{v_2^*} = 0$

$$\therefore \sum_i f_1(x_i, y_i) - \sum_i \sum_{y \in V} f_1(x_i, y) P(y | x_i, v^*) - 2Cv_1^*$$
$$= \sum_i f_2(x_i, y_i) - \sum_i \sum_{y \in V} f_2(x_i, y) P(y | x_i, v^*) - 2Cv_2^*$$

$\because f_1, f_2$ are identical, $\therefore \sum_i f_1(x_i, y_i) = \sum_i f_2(x_i, y_i)$ and $\sum_i \sum_{y \in V} f_1(x_i, y) P(y | x_i, v^*) = \sum_i \sum_{y \in V} f_2(x_i, y) P(y | x_i, v^*)$

$\therefore$ we can get $-2Cv_1^* = -2Cv_2^*$

$\therefore v_1^* = v_2^*$

② $\because \frac{\partial L}{\partial v_1}\Big|_{v^*} = \frac{\partial L}{\partial v_2}\Big|_{v_2^*} = 0$ Similarly with ①, we have :

$$\therefore \sum_i f_1(x_i, y_i) - \sum_i \sum_{y \in V} f_1(x_i, y) P(y | x_i, v^*) - 2C \cdot \frac{|v_1^*|}{\partial v_1^*} = \sum_i f_2(x_i, y_i) - \sum_i \sum_{y \in V} f_2(x_i, y) P(y | x_i, v^*) - 2C \cdot \frac{|v_2^*|}{\partial v_2^*}$$

$$-2C \cdot \frac{|v_1^*|}{\partial v_1^*} = -2C \cdot \frac{|v_2^*|}{\partial v_2^*}$$

$$\frac{|v_1^*|}{\partial v_1^*} = \frac{|v_2^*|}{\partial v_2^*}$$

$\therefore v_1^*$ and v_2^* are both positive
or both negative
or both $\neq 0$.

$\therefore$ Not necessarily $v_1^* = v_2^*$ hold, but v_1^* and v_2^* are both positive or both negative or both zero.

Question 2.

Assume the featre vector is $f_{u,k}(x,y) = \begin{cases} 1 & \text{if } x = w_u \text{ and } y = w_k \\ 0 & \text{o.w.} \end{cases}$

$$L(v) = \sum_i v \cdot f(x_i, y_i) - \sum_i \log \sum_y e^{v \cdot f(x_i, y)}$$

$$\frac{\partial L}{\partial v_{u,k}} = \sum_i f_{u,k}(x_i, y_i) - \sum_i \sum_y f_{u,k}(x_i, y) \cdot P(y \mid x_i, v)$$

$\sum_i f_{u,k}(x_i, y_i)$ is equivalent to $Count(w_u, w_k)$.

$f_{u,k}(x_i, y) \cdot P(y \mid x_i, v)$ is 0 if $y \neq w_k$, and is nozero if $y = w_k$.

$$\therefore \frac{\partial L}{\partial v_{u,k}} = Count(w_u, w_k) - \sum_i f_{u,k}(x_i, w_k) \cdot P(w_k \mid x_i, v)$$

$\because$ $f_{u,k}(x_i, w_k) = 0$ unless $x_i = w_u$

$$\frac{\partial L}{\partial v_{u,k}} = Count(w_u, w_k) - \sum_{i: x_i = w_u} f_{u,k}(w_u, w_k) P(w_k \mid w_u, v)$$

$$= Count(w_u, w_k) - P(w_k \mid w_u, v) \sum_{i: x_i = w_u} 1$$

$$= Count(w_u, w_k) - P(w_k \mid w_u, v) \cdot Count(w_u)$$

Finally, we know that for the optimal v^*, ~~the~~ we should have $\frac{\partial L}{\partial v_{u,k}}\Big|_{v^*_{u,k}} = 0$,

$$\therefore \frac{\partial L}{\partial v_{u,k}}\Big|_{v^*_{u,k}} = 0 = Count(w_u, w_k) - P(w_k \mid w_u, v^*) \cdot Count(w_u)$$

$$\therefore P(w_k \mid w_u, v^*) = \frac{Count(w_u, w_k)}{Count(w_u)}$$

Question 3.

① $$p(y|x;v) = \frac{e^{v \cdot f(x,y)}}{\sum_{y' \in \mathcal{Y}} e^{v \cdot f(x,y')}}$$

$\because$ $$f_1(x,y) = \begin{cases} 1 & \text{if } x=y \\ 0 & \text{otherwise} \end{cases}$$

$$f_2(x,y) = \begin{cases} 1 & \text{if } x = \text{reverse}(y) \\ 0 & \text{otherwise} \end{cases}$$

② $$P(\text{the}|\text{the}) = 0.4 = \frac{e^{v_1}}{e^{v_1}+e^{v_2}+e^{0}\cdot(|V'|-2)} = \frac{e^{v_1}}{e^{v_1}+e^{v_2}+|V'|-2}$$

$$P(\text{eht}|\text{the}) = 0.3 = \frac{e^{v_2}}{e^{v_1}+e^{v_2}+|V'|-2}$$

$$P(\text{dog}|\text{the}) = \frac{0.3}{|V'|-2} = \frac{e^{0}}{e^{v_1}+e^{v_2}+(|V'|-2)\times e^{0}} = \frac{1}{e^{v_1}+e^{v_2}+|V'|-2}$$

③ solve the equations above in ②.

we could get $$\begin{cases} v_1 = \ln\left(\frac{4}{3}(|V'|-2)\right) \\ v_2 = \ln(|V'|-2) \end{cases}$$