Assignment2

Taekang Eom

April 28, 2019

(a) $y_w = 0$ if $w \neq o$ and $y_o = 1$, so

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

(b)

$$\frac{\partial J_{naive}}{\partial \boldsymbol{v}_c} = \frac{\partial}{\partial \boldsymbol{v}_c} \left(-\log\left(\exp(\boldsymbol{u}_o^{\mathrm{T}} \boldsymbol{v}_c)\right) + \log\left(\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^{\mathrm{T}} \boldsymbol{v}_c) \right) \right)$$

$$= -\boldsymbol{u}_o + \frac{\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^{\mathrm{T}} v_c)}{\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^{\mathrm{T}} \boldsymbol{v}_c)} \boldsymbol{u}_w = - \sum_{\boldsymbol{w} \in Vocab} y_w \boldsymbol{u}_w + \sum_{\boldsymbol{w} \in Vocab} \hat{y}_w \boldsymbol{u}_w$$

$$= \boldsymbol{U}(\hat{\boldsymbol{y}} - \boldsymbol{y})$$

(c)

i)$w = o$

$$\frac{\partial J_{naive}}{\partial \boldsymbol{u}_o} = -\boldsymbol{v}_c + \frac{\exp(\boldsymbol{u}_o^{\mathrm{T}} v_c)}{\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^{\mathrm{T}} \boldsymbol{v}_c)} \boldsymbol{v}_c = (\hat{y_o} - y_o) \boldsymbol{v}_c$$

ii)$w \neq o$

$$\frac{\partial J_{naive}}{\partial \boldsymbol{u}_w} = \frac{\exp(\boldsymbol{u}_w^{\mathrm{T}} v_c)}{\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^{\mathrm{T}} \boldsymbol{v}_c)} \boldsymbol{v}_c = (\hat{y_w} - y_w) \boldsymbol{v}_c$$

(d)

$$\frac{d\sigma}{dX} = \sigma(X) \odot [\mathbf{1} - \sigma(X)]$$

(e)