CS 224n Assignment #2: Written Assignment

Yoshihiro Kumazawa

May 23, 2020

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

(b)

$$\begin{split} \frac{\partial}{\partial \boldsymbol{v}_c} \boldsymbol{J}_{\text{naive-softmax}}(\boldsymbol{v}_c, o, \boldsymbol{U}) &= \frac{\partial}{\partial \boldsymbol{v}_c} \left(-\log \left(\frac{\exp(\boldsymbol{u}_o^\top \boldsymbol{v}_c)}{\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^\top \boldsymbol{v}_c)} \right) \right) \\ &= \frac{\partial}{\partial \boldsymbol{v}_c} \left(\log \left(\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^\top \boldsymbol{v}_c) \right) - \log \left(\exp(\boldsymbol{u}_o^\top \boldsymbol{v}_c) \right) \right) \\ &= \frac{\partial}{\partial \boldsymbol{v}_c} \log \left(\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^\top \boldsymbol{v}_c) \right) - \frac{\partial}{\partial \boldsymbol{v}_c} \boldsymbol{u}_o^\top \boldsymbol{v}_c \\ &= \frac{\sum_{x \in Vocab} \exp(\boldsymbol{u}_x^\top \boldsymbol{v}_c) \boldsymbol{u}_x}{\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^\top \boldsymbol{v}_c)} - \boldsymbol{u}_o \\ &= \sum_{x \in Vocab} \frac{\exp(\boldsymbol{u}_x^\top \boldsymbol{v}_c)}{\sum_{w \in Vocab} \exp(\boldsymbol{u}_w^\top \boldsymbol{v}_c)} \boldsymbol{u}_x - \boldsymbol{u}_o \\ &= \sum_{x \in Vocab} \hat{\boldsymbol{y}}_x \boldsymbol{u}_x - \sum_{x \in Vocab} \boldsymbol{y}_x \boldsymbol{u}_x \\ &= \sum_{x \in Vocab} \boldsymbol{u}_x(\hat{\boldsymbol{y}}_x - \boldsymbol{y}_x) \\ &= \boldsymbol{U}(\hat{\boldsymbol{y}} - \boldsymbol{y}). \end{split}$$

(c)

$$egin{aligned} rac{\partial}{\partial oldsymbol{u}_w} oldsymbol{J}_{ ext{naive-softmax}}(oldsymbol{v}_c, o, oldsymbol{U}) &= rac{\partial}{\partial oldsymbol{u}_w} \log \left(\sum_{w \in Vocab} \exp(oldsymbol{u}_w^ op oldsymbol{v}_c)
ight) - rac{\partial}{\partial oldsymbol{u}_w} oldsymbol{u}_o^ op oldsymbol{v}_c
ight. \\ &= rac{\exp(oldsymbol{u}_w^ op oldsymbol{v}_c)}{\sum_{w \in Vocab} \exp(oldsymbol{u}_w^ op oldsymbol{v}_c)} oldsymbol{v}_c - oldsymbol{y}_w oldsymbol{v}_c \\ &= oldsymbol{v}_c(oldsymbol{\hat{y}}_w - oldsymbol{y}_w). \end{aligned}$$