

Assignment 2: word2vec solutions

Shubham Gupta

January 1, 2020

1 Understanding word2vec

1.1 a. Derivation

From the question, we know the following:

$$J(\nu_c, o, \mathbf{U}) = -\log P(O = o | C = c).$$

We know cross entropy loss is given as:

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

Since \mathbf{y} is a one hot vector, it implies that y will be equal to 1 only when $i == o$. Hence, the LHS can be written as:

$$\begin{aligned} & -\sum_{w \in Vocab} y_w \log(\hat{y}_w) \\ &= -[y_0 \log(y_0) + y_1 \log(y_1) + \dots + y_o \log(y_o) + \dots + y_w \log(y_w)] \\ &= -y_o \log(y_o) \\ &= -\log(y_o) \\ &= -\log(P(O = o | C = c)) \end{aligned} \tag{1}$$

1.2 a. Derivation