3. Word2vec (2)

2019년 1월 15일 화요일 오전 1:25

(c) $J = -\log(\sigma(u_o^T v_c)) - \sum_{k=1}^{K} \log(\sigma(-u_k^T v_c))$, $o \notin \{1, \cdots, K\}$

(c)-(1)

$$\frac{\partial J}{\partial v_c} = ?$$

$$\frac{\partial J}{\partial v_c} = -\frac{1}{\sigma(u_o^T v_c)} \cdot \sigma(u_o^T v_c) \cdot (1 - \sigma(u_o^T v_c)) \cdot u_o$$

$$-\sum_{k=1}^{K} \left\{ \frac{1}{\sigma(-u_k^T v_c)} \cdot \sigma(-u_k^T v_c) \cdot (1 - \sigma(-u_k^T v_c)) \cdot (-u_k) \right\}$$

$$= (\sigma(u_o^T v_c) - 1) u_o - \sum_{k=1}^{K} (\sigma(-u_k^T v_c) - 1) u_k$$

(c)-(2)

$$\frac{\partial J}{\partial u_i} = \frac{\partial}{\partial u_i} \left[-\log(\sigma(u_o^T v_c)) - \sum_{k=1}^{K} \log(\sigma(-u_k^T v_c)) \right] = ?$$

(i) $i = o$ case

$$\frac{\partial J}{\partial u_o} = -\frac{1}{\sigma(u_o^T v_c)} \cdot \sigma(u_o^T v_c) \cdot (1 - \sigma(u_o^T v_c)) \cdot v_c$$

$$= (\sigma(u_o^T v_c) - 1) v_c$$

(ii) $i = 1, \cdots, K$ case

$$\frac{\partial J}{\partial u_k} = -\frac{1}{\sigma(-u_k^T v_c)} \cdot \sigma(-u_k^T v_c) \cdot (1 - \sigma(-u_k^T v_c)) \cdot (-v_c)$$

$$= -(\sigma(-u_k^T v_c) - 1) v_c \quad \text{for } k = 1, 2, \cdots, K$$