1.a

$$\frac{\partial J}{\partial v_c} = \frac{\partial}{\partial v_c}\left(-u_o^T v_c + \log \sum_{w=1}^{V} \exp(u_w^T v_c)\right)$$

$$= -u_o + \frac{\partial}{\partial v_c} \log \sum_{w=1}^{V} \exp(u_w^T v_c)$$

$$= -u_o + \frac{1}{\sum_{w=1}^{V} \exp(u_w^T v_c)} \cdot \frac{\partial}{\partial v_c} \sum_{x=1}^{V} \exp(u_x^T v_c)$$

$$= -u_o + \frac{1}{\sum_{w=1}^{V} \exp(u_w^T v_c)} \cdot \sum_{x=1}^{V} \exp(u_x^T v_c) \frac{\partial}{\partial v_c} u_x^T v_c$$

$$= -u_o + \frac{1}{\sum_{w=1}^{V} \exp(u_w^T v_c)} \cdot \sum_{x=1}^{V} \exp(u_x^T v_c) \cdot u_x$$

$$= -u_o + \sum_{x=1}^{V} \frac{\exp(u_x^T v_c)}{\sum_{w=1}^{V} \exp(u_w^T v_c)} \cdot u_x$$

$$= -u_o + \sum_{w=1}^{V} \hat{y}_w \cdot u_x$$

1.b

$$\frac{\partial J}{\partial U} = \frac{\partial J}{\partial u_w} \cdot \left(-u_o^T v_c + \log \sum_{w' vocab} \exp(u_{w'}^T \cdot v_c) \right)$$

$$= \underbrace{\frac{\partial J}{\partial u_w} \cdot (-u_o^T v_c)}_{①} + \underbrace{\frac{\partial J}{\partial u_w} \log \sum_{w' \in vocab} \exp(u_{w'}^T, v_c)}_{②}$$

$$① = \begin{cases} -v_c & \text{if } w = o \\ 0 & \text{otherwise} \end{cases}$$

$$② = \frac{\partial J}{\partial u_w} \log \sum_{w' \in vocab} \exp(u_{w'}^T, v_c)$$

$$= \frac{1}{\sum_{w' \in vocab} (u_{w'}^T, v_c)} \cdot \frac{\partial J}{\partial u_w} \sum_{w' \in vocab} \exp(u_{w'}^T, v_c)$$

$$= \frac{1}{\sum_{w' \in vocab} (u_{w'}^T \cdot v_c)} \cdot \sum_{x \in vocab} \exp(u_x^T, v_c) \cdot \underbrace{\frac{\partial J}{\partial u_w}(u_x^T \cdot v_c)}_{\begin{cases} v_c & \text{if } x = w \\ 0 & \text{otherwise} \end{cases}}$$

$$= \frac{\exp(u_w^T v_c)}{\sum_{w' \in vocab} (u_{w'}^T v_c)} \cdot v_c$$

$$= \hat{y}_w \cdot v_c$$

combine ① and ②

$$\therefore \frac{\partial J}{\partial U} = \begin{cases} (\hat{y}_w - 1) v_c & \text{if } w = o \\ \hat{y}_w v_c & \text{otherwise} \end{cases}$$