(a):
$$-\sum_{w \in vocab} y_w \log(\hat{y}_w) = -(0+0+...+y_o \log(\hat{y}_o)+0+...+0) = -\log(\hat{y}_o)$$

(b):

$$rac{\partial J}{\partial v_c} = -u_o + \sum_{x \in v} p(x|c)u_x = -u_o + \sum_{x \in v} p(x|c)u_x = U^T(\hat{y} - y)$$

(c):

$$rac{\partial J}{\partial u_w} = -rac{\partial}{\partial u_w}\lograc{\exp(u_o^Tv_c)}{\sum_{w\in V}\exp(u_w^Tv_c)} = -rac{\partial}{\partial u_w}u_o^Tv_c + rac{\partial}{\partial u_w}\log\sum_{w\in V}\exp(u_w^Tv_c)$$

Then we first calculate the later part, we give the annotation l as the value of the later part, so we can deduct that:

$$l = rac{1}{\sum_{w \in V} \exp(u_w^T v_c)} \cdot rac{\partial}{\partial u_w} \exp(u_w^T v_c) = rac{\exp(u_w^T v_c)}{\sum_{w \in V} \exp(u_w^T v_c)} \cdot v_c = \hat{y_w} \cdot v_c$$

So we can get the final solution:

$$egin{aligned} rac{\partial J}{\partial u_w} &= \left\{egin{array}{cc} (\hat{y_w}-1) \cdot v_c & ext{if } w = o \ \hat{y_w} \cdot v_c & ext{if } w
eq o \end{aligned}
ight. \ rac{\partial J}{\partial U} &= (\hat{y}-y)^T \cdot v_c \end{aligned}$$

(d):

$$rac{\partial \sigma(x)}{\partial x} = rac{e^x}{(e^x + 1)^2} = \sigma(x)(1 - \sigma(x))$$

(e):

1):

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

2):

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

3):

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

we give the annotations l as the left part, r as the right part, then we get following:

$$egin{aligned} l &= -rac{1}{\sigma(u_o^T v_c)} \cdot rac{\partial}{\partial v_c} \sigma(u_o^T v_c) = (\sigma(u_o^T v_c) - 1) u_o \ & r &= \sum_{k=1}^K (1 - \sigma(-u_k^T v_c)) \cdot u_k \end{aligned}$$

So we can get the final derivative:

$$rac{\partial J}{\partial v_c} = (\sigma(u_o^T v_c) - 1)u_o + \sum_{k=1}^K (1 - \sigma(-u_k^T v_c)) \cdot u_k$$

(f):

$$egin{aligned} \partial oldsymbol{J}_{ ext{skip-gram}} & \left(oldsymbol{v}_c, w_{t-m}, \ldots w_{t+m}, oldsymbol{U}
ight) / \partial oldsymbol{U} = \sum_{-m \leq j \leq m, j
eq 0} rac{\partial oldsymbol{J}(oldsymbol{v}_c, w_{t+j}, oldsymbol{U})}{\partial oldsymbol{U}} \ \partial oldsymbol{J}_{ ext{skip-gram}} & \left(oldsymbol{v}_c, w_{t-m}, \ldots w_{t+m}, oldsymbol{U}
ight) / \partial oldsymbol{v}_c = \sum_{-m \leq j \leq m, j
eq 0} \partial oldsymbol{J} \left(oldsymbol{v}_c, w_{t+j}, oldsymbol{U}
ight) / \partial oldsymbol{v}_c \ \partial oldsymbol{J}_{ ext{skip-gram}} & \left(oldsymbol{v}_c, w_{t-m}, \ldots w_{t+m}, oldsymbol{U}
ight) / \partial oldsymbol{v}_w = 0 ext{ when } w
eq c \end{aligned}$$