CSC311 A2 Non-programming

David Pham: 1002597825

Nov 15 2020

Question 4(a).
SEE NON-PROG PDF FOR FIGURE

Question 4(b).
We learned that

$$\left[\frac{\partial C}{\partial \tilde{G}}\right]_n = [O - T]_n$$

Coming back to our question we have that

$$\frac{\partial C}{\partial \tilde{G}} = \frac{-tlogo - (1-t)log(1-0)}{\partial o}\frac{\partial o}{\partial \tilde{g}}$$

Deriving on o gives us

$$= \frac{-(o-t)}{o(o-1)}\frac{\partial o}{\partial \tilde{g}}$$

Using (4) we have that

$$= \frac{-(o-t)}{o(o-1)}\frac{\sigma(\tilde{g})}{\partial \tilde{g}}$$

Using (8),

$$= \frac{-(o-t)}{o(o-1)}\sigma(\tilde{g})[1-\sigma(\tilde{g})$$

Simplifying we get

$$= o - t$$

$$= O - T$$

as needed

Question 4(c).

$$\left[\frac{\partial C}{\partial \tilde{H}}\right]_{nk} = \frac{\partial C}{\partial \tilde{H}_{nk}}$$

By definition:

$$= \frac{\partial C}{\partial \tilde{h}_k^{(n)}}$$

Using chain rule:

$$= \frac{\partial C}{\partial g_k^{(n)}} \frac{\partial g_k^{(n)}}{\partial \tilde{h}_k^{(n)}}$$

Using (10)

$$= \frac{\partial C}{\partial g_k^{(n)}} \frac{\partial tanh(\tilde{h}_k^{(n)})}{\partial \tilde{h}_k^{(n)}}$$

Using (7)

$$= \frac{\partial C}{\partial g_k^{(n)}} (1 - [tanh(\tilde{h}_k^{(n)})]^2)$$

Using (10)

$$= \frac{\partial C}{\partial G_{nk}} (1 - [g_k^{(n)}]^2)$$

Using definition and rearranging

$$= (1 - G_{nk}^2) \left[\frac{\partial C}{\partial G}\right]_{nk}$$

Question 4(d).

$$\frac{\partial C}{\partial V} = H^T \frac{\partial C}{\partial \tilde{H}}$$

By definition we have that:

$$\left[\frac{\partial C}{\partial V}\right]_{mk} = \frac{\partial C}{\partial V_{mk}}$$

By (9) we have that:

$$= \frac{\partial \sum_n c(t^{(n)}, o^{(n)})}{\partial V_{mk}}$$

Using chain rule and moving the sum out:

$$= \sum_n \frac{\partial c(t^{(n)}, o^{(n)})}{\partial \tilde{h}_k^{(n)}} \frac{\partial \tilde{h}_k^{(n)}}{\partial V_{mk}}$$

Using (5)

$$= \sum_n \frac{\partial c(t^{(n)}, o^{(n)})}{\partial \tilde{h}_k^{(n)}} \frac{\partial hV + v_0}{\partial V_{mk}}$$

From (5), we have that $\tilde{h}_k = \sum_m h_m V_{mk} + v_{0k}$. Therefore we have that $\frac{\partial \tilde{h}_k}{\partial V_{mk}} = h_m$

$$= \sum_n \frac{\partial c(t^{(n)}, o^{(n)})}{\partial \tilde{h}_k^{(n)}} h_m^{(n)}$$

Rearranging and definition of matrices, we have that

$$= \sum_n \left[H^T\right]_{mn} \left[\frac{\partial C}{\partial \tilde{H}}\right]_{mk}$$

$$= \left[H^T \frac{\partial C}{\partial \tilde{H}} \right]_{mk}$$

Therefore,

$$\frac{\partial C}{\partial V} = H^T \frac{\partial C}{\partial \tilde{H}}$$

Question 4(e).

$$\frac{\partial C}{\partial v_0} = \vec{1} \frac{\partial C}{\partial \tilde{H}}$$

Similar to the previous question, we prove that

$$\left[\frac{\partial C}{\partial v_0} \right]_m = \left[\vec{1} \frac{\partial C}{\partial \tilde{H}} \right]_m$$

Using chain rule and using the previous proof we can get that

$$\left[\frac{\partial C}{\partial v_0} \right]_m = \sum_n \left[\frac{\partial C}{\partial \tilde{H}} \right] \frac{\partial \tilde{h}_m^{(n)}}{\partial v_{0m}}$$

By (5) we can see that $\frac{\partial \tilde{h}_m}{\partial v_{0m}} = 1$

$$= \sum_n \left[\frac{\partial C}{\partial \tilde{H}} \right] [\vec{1}]_n$$

Rearranging we can see that

$$\left[\vec{1} \frac{\partial C}{\partial \tilde{H}} \right]_m$$

as needed

Question 4(f).

$$\frac{\partial C}{\partial H} = \frac{\partial C}{\partial \tilde{H}} V^T$$

Similar as before we prove

$$\left[\frac{\partial C}{\partial H} \right]_{nk} = \left[\frac{\partial C}{\partial \tilde{H}} V^T \right]_{nk}$$

Using the proofs from before and chain rule we have that

$$\left[\frac{\partial C}{\partial H} \right]_{nk} = \sum_m \left[\frac{\partial C}{\partial \tilde{H}} \right]_{nm} \frac{\partial \tilde{h}_m^{(n)}}{\partial h_k^{(n)}}$$

From (5) we can see that $\frac{\partial \tilde{h}_m}{\partial h_k} = V_{km}$

$$= \sum_m \left[\frac{\partial C}{\partial \tilde{H}} \right]_{nm} V_{km}$$

Using matrix laws we get that

$$= \sum_m \left[\frac{\partial C}{\partial \tilde{H}}\right]_{nm} [V^T]_{mk}$$

$$= \left[\frac{\partial C}{\partial \tilde{H}} V^T\right]_{nk}$$

as needed