HW-6

Consider the function $l(w) = y\log(\sigma(w^Tx)) + (1-y)\log(1-\sigma(w^Tx))$ where w and x are k^{th} dimensional vectors. Assume 1 training example.
Find $\nabla_{w_j} l(w)$, that is, the partial derivative of $l(w)$ with respect to the j^{th} element of vector w.

$\rightarrow$ $$\sigma(w^Tx) = \frac{1}{1+\exp(-w^Tx)} = \frac{\exp(w^Tx)}{1+\exp(w^Tx)}$$

On substituting

$$l(w) = y\log\left(\frac{\exp(w^Tx)}{1+\exp(w^Tx)}\right) + (1-y)\log\left(\frac{1}{1+\exp(w^Tx)}\right)$$

$$= y\log(\exp(w^Tx)) - \log(1+\exp(w^Tx)) + (1-y)(\log(\exp(w^Tx)+1))$$

$$= y(w^Tx) - y\log(1+\exp(w^Tx)) + y\log(1+\exp(w^Tx)) - \log(\exp(w^Tx)+1)$$

$$= y(w^Tx) - \log(\exp(w^Tx)+1)$$

$$\therefore \nabla_{w_j} l(w) = \frac{\delta}{\delta w_j}(y(w^Tx)) - \frac{\delta}{\delta w_j}\left[\log[\exp(w^Tx)+1]\right]$$

$$= yx_j - \frac{\exp(w^Tx)}{1+\exp(w^Tx)}x_j \quad \left(\because \log(x+y) = \frac{x}{y+x}\right)$$

$$= yx_j - \sigma(w^Tx)x_j$$

$$\therefore \boxed{\nabla_{w_j} l(w) = x_j(y - \sigma(w^Tx))}$$