Homework - 6
COEN 240-Machine Learning

Pujitha Kallu
ID : W1653660
pkallu@scu.edu

Consider the function $\ell(\mathbf{w}) = y\log\big(\sigma(\mathbf{w}^T\mathbf{x})\big) + (1-y)\log\big(1-\sigma(\mathbf{w}^T\mathbf{x})\big)$

Where **w** and **x** are k^{th} dimensional vectors. Assume 1 training example.

Find $\nabla_{w_j}\ell(\mathbf{w})$, that is, the partial derivative of $\ell(\mathbf{w})$ with respect to the j^{th} element of vector **w**.

- Recall if $\sigma(x) = \frac{1}{1+e^{-x}}$ then $\sigma'(x) = \sigma(x)\big(1-\sigma(x)\big)$
- Show your work!

Sol: $\nabla_{w_j} l(w) = \nabla_{w_j}\left(y log\left(\sigma\left(w^T x\right)\right) + (1-y)\log log\left(1-\sigma\left(w^T x\right)\right)\right)$

$$\nabla_{w_j} l(w) = \frac{y}{\sigma\left(w^T x\right)}\nabla_{w_j}\sigma\left(w^T x\right) + \frac{(1-y)}{-(1-\sigma\left(w^T x\right))}\nabla_{w_j}\sigma\left(w^T x\right)$$

where,

$$\nabla_{w_j}\sigma\left(w^T x\right) = \left(\sigma\left(w^T x\right)\right)\left(1-\sigma\left(w^T x\right)\right) * \nabla_{w_j}\left(w^T x\right) = \left(\sigma\left(w^T x\right)\right)\left(1-\sigma\left(w^T x\right)\right) * x_j$$

Hence,

$$\nabla_{w_j} l(w) = \frac{y}{\sigma\left(w^T x\right)}\left(\sigma\left(w^T x\right)\right)\left(1-\sigma\left(w^T x\right)\right) * x_j - \frac{(1-y)}{(1-\sigma\left(w^T x\right))}\left(\sigma\left(w^T x\right)\right)\left(1-\sigma\left(w^T x\right)\right) * x_j$$

$$\nabla_{w_j} l(w) = y\left(1-\sigma\left(w^T x\right)\right) * x_j - (1-y)\left(1-\sigma\left(w^T x\right)\right) * x_j$$

$$\boxed{\nabla_{w_j} l(w) = \left(y-\sigma\left(w^T x\right)\right) * x_j}$$

References:

1. Class notes - Logistic Regression
2. https://web.stanford.edu/~jurafsky/slp3/5.pdf