Nicholas Wilson

Answers to homework 2 question 1:

HW2 Q1

1.1) $\frac{d}{dz}\left(\frac{1}{1+e^{-z}}\right) = \frac{e^{-z}}{(1+e^{-z})^2}$

$= \frac{1}{(1+e^{-z})} \cdot \frac{e^{-z}}{(1+e^{-z})} = \frac{1}{1+e^{-z}} \cdot \left(\frac{1+e^{-z}-1}{1+e^{-z}}\right)$

$= \frac{1}{1+e^{-z}} \cdot \left(\frac{1}{1} - \frac{1}{1+e^{-z}}\right) = g(z)(1-g(z))$ QED

1.2) $1-g(z) = \frac{e^{-z}}{1+e^{-z}} = \frac{1}{e^{z}(1+e^{-z})} = \frac{1}{e^{z}+1} = g(-z)$ QED

1.3) $\cancel{J(\theta)} \; \frac{1}{1+e^{-\theta^T x}} = \frac{1}{1+e^{-(\theta_0+\theta_1 x_1+\dots+\theta_n x_n)}} = h_\theta(x)$

$J(\theta) = -y\ln\left(\frac{1}{1+e^{-(\theta_0+\theta_1 x_1+\dots+\theta_n x_n)}}\right) - (1-y)\ln\left(1 - \frac{1}{1+e^{-\theta^T x}}\right)$

$\frac{\partial}{\partial \theta_j} J(\theta) = \frac{\partial}{\partial \theta_j}\Big((-y^{(i)}\ln(h_\theta(x^{(i)}))$

$- (1-y^{(i)})\ln(1-h_\theta(x^{(i)}))\Big)$

$\frac{\partial}{\partial \theta_j}\left((-y^{(i)}\ln(h_\theta(x^{(i)}))\right) = -y^{(i)} \cdot \frac{1}{h_\theta(x^{(i)})} \cdot \frac{\partial}{\partial \theta_j}\left(h_\theta(x^{(i)})\right)$

$= \cancel{-y^{(i)} \cdot (1-g(\theta^T x^{(i)}))x_j^{(i)}}$

$-y^{(i)}$

HW2 Q1 Continued

1.3 $\frac{\partial}{\partial \theta_j}(-y \ln(h_\theta(x))) = -y \cdot \frac{1}{h_\theta(x)} \cdot \frac{\partial}{\partial \theta_j}(h_\theta(x))$

Substitute $h_\theta(x) \Rightarrow g(\theta^T x)$

$= \cancel{-y(1+e^{-\theta^T x})}$

$-y\,(1+e^{\theta^T x})\, g(\theta^T x)(1-g(\theta^T x)) \cdot \frac{\partial \theta^T x}{\partial \theta_j}$

$= -y(1-g(\theta^T x))x_j$

$\frac{\partial}{\partial \theta_j}(-(1-y)\ln(1-g(\theta^T x))) = (1-y)g(\theta^T x)x_j$

(by analogy)

So $\frac{\partial J(\theta)}{\partial \theta_j} = -y(1-g(\theta^T x))x_j + (1-y)g(\theta^T x)x_j$

$= (-y x_j + g(\theta^T x)x_j)$

$= (h_\theta(x)-y)x_j \quad$ QED