Logistic Regression Cost Function and Gradient Derivation

Cost function:

Identity:

Useful partial derivative:

Using the identity above:

Gradient of w.r.t. . Let :

Using the partial derivative above:

Using the identity to sub in for :

Now bring the -1 into the sum:

Since can only ever be a 0 or a 1: