# CS416 – HW2

David Montenegro

## Introduction

All problems are designed by Dan Sheldon.

Please complete all exercises and problems below.

- All the files can be found in

  `http://www.cs.wm.edu/~liqun/teaching/cs416/hw2/`

- You can also copy to your directory on a department machine by:

  ```
  cp ~liqun/public_html/teaching/cs416/hw2/* .
  ```

Your submission consists of three steps:

1. Create hw2.pdf with your solutions to the following problems. Put your name in the file.
2. You'll need to create or edit these files in the directory hw2. Complete the requested code in these files.
   - exercise_2.ipynb
   - logistic_regression.py
   - one_vs_all.py
3. Compile your **exercise_2.ipynb** to a pdf named **exercise_2.pdf**.
4. Submit on Gradescope:
   - **hw2.pdf**
   - **exercise_2.pdf**
   - **exercise_2.ipynb**
   - **logistic_regression.py**
   - **one_vs_all.py**

## Problem 1 (15 points).

**Logistic Regression**

Let $g(z) = \frac{1}{1+e^{-z}}$ be the logistic function.

**1.1 (5 points).** Show that $\frac{d}{dz}g(z) = g(z)(1 - g(z))$.

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

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

**1.2 (5 points).** Show that $1 - g(z) = g(-z)$.

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

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

**1.3 (5 points).** Consider the log loss function for logistic regression simplified so there is only one training example:

$$J(\theta) = -y \log h_\theta(x) - (1 - y)\log(1 - h_\theta(x)),\ h_\theta(x) = g(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}}$$

Show that the partial derivative with respect to $\theta_j$ is:

$$\frac{\partial}{\partial\theta_j}J(\theta) = (h_\theta(x) - y)x_j$$

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

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

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

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

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

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

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

## Problem 2 (10 points).

**Logistic regression for book classification**. In this problem, you will implement logistic regression for book classification.

Open **exercise_2.ipynb** and follow the instructions to complete the problem.

## Problem 3 (10 points).

**SMS spam classification**. In this problem you will use your implementation of logistic regression to create a spam classifier for SMS messages.

Open **exercise_2.ipynb** and follow the instructions to complete the problem.

## Problem 4 (34 points).

**Hand-Written Digit Classification.** In this assignment you will implement one-vs-all multiclass logistic regression to classify images of hand-written digits. Then you will explore the effects of regularization and training set size on training and test accuracy.

Open **exercise_2.ipynb** and follow the instructions to complete the problems.