MAT 345 - Homework 5
Due Wednesday, October 31, 2018, in class

1. (4 points) For each of the data sets in the attached file *HW5data.xls*, assume $\mathbf{x}$ denotes the input and y the output. Run through the Linear Regression Algorithm to find $\mathbf{w}$ that gives the best approximation

$$y \approx \mathbf{w}^T \mathbf{x}.$$

Do not use Excel's regression feature and use the algorithm discussed in class instead, showing your steps (include the computed matrix A and weight vector $\mathbf{w}$ in your answer).

2. (4 points) Use the Gradient Descent Algorithm to estimate the value of $\mathbf{x}$ where the function f attains its minimum, by using the given fixed step η and the starting point $\mathbf{x}_0$:

(a) $f(x) = (x-1)(x-2)(x-5)$ with

(i) $\eta = 0.1$, $x_0 = 3$
(iii) $\eta = 0.05$, $x_0 = 6$
(iii) $\eta = 0.01$, $x_0 = 1$

(b) $f(x, y) = x^2 - 2xy^2 + y^4$ with

(i) $\eta = 0.1$, $x_0 = 1$, $y_0 = 2$
(ii) $\eta = 0.1$, $x_0 = 4$, $y_0 = -2$
(iii) $\eta = 0.05$, $x_0 = 0$, $y_0 = 3$

3. (2 points) Consider the following selling data from sample of 10 Corvette cars, aged 1-6 years, available from the *Kelley Blue Book*. Here x denotes the age of the car and y denotes the selling price, in hundreds of dollars:

x	6	6	6	2	2	5	4	5	1	4
y	270	260	275	405	364	295	335	308	405	305

Use the gradient descent algorithm to solve this linear regression question. That is, instead of inverting matrices to find the optimal $\mathbf{w}_{\text{lin}}$, as was done in the group worksheet, use the sample data set and the gradient descent algorithm to minimize the error function

$$f(\mathbf{w}) = \frac{1}{N} \sum_{k=1}^{N} (\mathbf{w}^T \mathbf{x}_k - y_k)^2.$$

Recall that we found the gradient of the error function to be

$$\nabla f(\mathbf{w}) = \frac{2}{N} [X^T X \mathbf{w} - X^T \mathbf{y}].$$

Start the gradient descent algorithm with $\mathbf{w}(0) = \mathbf{0}$ and use a step $\eta = 0.04$. Stop when neither w_0, nor w_1 change by more than 10^{-3} in a given iteration.