# Written Assignment 2

Isobel Smith

05/10/2016

I worked on this assignment with Wendy and Elise

## 1 Question 1: Vectorization

$$\theta = (\theta_0, \theta_1, ...\theta_n)^T$$
$$x^{(i)} = (x_0, x_1 ... x_n)^T$$
$$x_0 = 1$$

### 1.1 a) What is the vectorial expression for the multivareate linear regression hypothesis function?

the hypothesis function for multivareate linear regression is:

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + ... + \theta_n x_n$$

the parameters are n+1 feature vectors:

$$x = \begin{vmatrix} x_0 \\ x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{vmatrix}$$

$$\theta = \begin{vmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ . \\ . \\ . \\ \theta_n \end{vmatrix}$$

We can rewrite the hypothesis as:

$$h_\theta(x) = \theta^T X$$

because the hypthesis function is a matrix multiplication:

$$\begin{vmatrix} \theta_0 & \theta_1 & \theta_2 & . & . & . & \theta_n \end{vmatrix}$$

$$\begin{vmatrix} x_0 \\ x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{vmatrix}$$

### 1.2 b)What is the vectorized expression for the cost function?

the cost function is:

$$J(\theta) = 1/2m \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

but as the hypothesis has now changed to be:

$$h_\theta(x) = \theta^T X$$

rewrite the cost function to be:

$$J(\theta) = 1/2msum_{(i=1)}^{m} (\theta^T x^{(i)} - y^{(i)})^2$$

### 1.3 c) What is the vectorized expression for the gradient of the cost function?

To calculate the gradient:

$$\frac{\delta J(\theta)}{\delta \theta} = \begin{vmatrix} \frac{\delta J(\theta)}{\delta \theta_0} \\ . \\ . \\ . \\ \frac{\delta J(\theta)}{\delta \theta_n} \end{vmatrix}$$

which is:

$$\frac{\delta J(\theta)}{\delta\theta} = \begin{vmatrix} \frac{1}{m}\sum_{i=1}^{m}((h_\theta(x^{(i)} - y^{(i)}) \\ \frac{1}{m}\sum_{i=1}^{m}((h_\theta(x^{(i)} - y^{(i)}) \cdot x_1 \\ . \\ . \\ . \\ \frac{1}{m}\sum_{i=1}^{m}((h_\theta(x^{(i)} - y^{(i)}) \cdot x_n \end{vmatrix}$$

### 1.4 d) What is the vectorized expression for the $\theta$ update rule in the gradient descent procedure?

the update rule is now:

$$\theta_0 = \theta_0 - \alpha\frac{\delta J(\theta)}{\delta\theta_0} = \theta_0 - \alpha \cdot \frac{1}{m} \cdot \sum_{(i=1)}^{m}(h_\theta(x^{(i)} - y^{(i)})x_0^{(i)}$$

where j $\geq$ 1:

$$\theta_j = \theta - \alpha \cdot \frac{1}{m} \cdot \left[ \begin{array}{c} \sum_{(i=1)}^{m}\theta^T(x^{(i)} - y^{(i)}) \cdot x_2^{(i)} \\ \sum_{(i=1)}^{m}\theta^T(x^{(i)} - y^{(i)}) \cdot x_1^{(i)} \\ . \\ . \\ . \\ \sum_{(i=1)}^{m}\theta^T(x^{(i)} - y^{(i)}) \cdot x_n^{(i)} \end{array} \right|$$

we simply add in the new hypothesis function

## 2 Question 3: probability

### 2.1 a. estimate mean and variance from data

the data set is: 2, 5, 7, 7, 9, 25

to calculate the mean:

$$\mu = \frac{1}{m}\sum_{i=1}^{m}x^{(i)}$$
$$\mu = \frac{1}{6}(2 + 5 + 7 + 7 + 9 + 25)$$
$$\mu = 9.17$$

to calculate the variance:

$$\sigma^2 = \frac{1}{m}\sum_{(i=1)}^{m}(x^{(i)} - \mu)^2$$

$$\sigma^2 = \frac{1}{6} \cdot (2 - 9.17)^2 + (5 - 9.17)^2 + (7 - 9.17)^2 + (7 - 9.17)^2 + (9 - 9.17)^2 + (25 - 9.17)^2$$

$$\sigma^2 = 54.81$$

## 2.2 b) calculate the probability density function

$$X \sim N(\mu, \sigma^2)$$

is a random variable.
calculate the probability density function:

$$fx^{(20)}$$

the probablilty density function is: CHECK THIS

$$fx^{(x)} = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-(\frac{(x-\mu)^2}{2 \cdot \sigma^2})}$$
$$fx^{(20)} = \frac{1}{7.4 \cdot \sqrt{2\pi}} \cdot e^{-(\frac{(20-9.17)^2}{2 \cdot 54.81})}$$
$$fx^{(20)} = 0.054 \cdot 0.343$$
$$fx^{(20)} = 0.0185$$

## 2.3 c) calculate $fx_1...x_6(2, 5, 7, 7, 9, 25)$

The variables are independant of each other and identically and normally distributed with $\mu = 9.17$ and $\sigma^2 = 54.81$
$fx_1...x_6(x_1...x_6)$ is the joint probability distribution function

## 2.4 e) estimate the covariance

Table 1: Table of values

| x | y |
|---|---|
| 2 | 4 |
| 5 | 4 |
| 7 | 5 |
| 7 | 6 |
| 9 | 7 |
| 25 | 10 |

$$Cov(X, Y) = \frac{\sum_{(i=1)}^{m} (x_i - \mu_x)(y_i - \mu_y)}{m - 1}$$
$$\mu_x = 9.16$$
$$\mu_y = 6$$
$$Cov(X, Y) = \frac{((2 - 9.16) \cdot (5 - 9.16) \cdot (7 - 9.16) \cdot (7 - 9.16) \cdot (9 - 9.16) \cdot (25 - 9.16)) \cdot ((4 - 6) \cdot (4 - 6) \cdot (5 - 6}{5}$$
$$Cov(X, Y) = \frac{(-352) \cdot (0)}{5}$$
$$Cov(X, Y) = 0$$