Amsterdam University College

Written assignment 2

Yuri Tax

October 7, 2016

q.1)

a) The vectorized expression for the cost function $h_\theta(x)$ is: $h_\theta(x) = \theta^T x$, or the Dot product of the vectors θ and x.

b) The vectorized form of the cost function $J(\theta)$ can be written as: $J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}((\theta^T x^{(i)} - y^{(i)}))^2$.

c) The partial derivative of the cost function with respect to θ, in its vectorized form can be written as:

$$\frac{\partial J(\theta)}{\partial \theta} = \frac{1}{m}\sum_{i=1}^{m}\begin{bmatrix}(\theta^T x^{(i)} - y^{(i)})x_0^{(i)} \\ (\theta^T x^{(i)} - y^{(i)})x_1^{(i)} \\ \vdots \\ (\theta^T x^{(i)} - y^{(i)})x_n^{(i)}\end{bmatrix}$$

d) The vecotized expression of the gradient update rule is $\vec{\theta} := \vec{\theta} - \alpha\frac{\partial}{\partial\theta}J(\vec{\theta})$. Filling in the partial derivate of the Cost function gives:

$$\vec{\theta} := \vec{\theta} - \alpha\frac{1}{m}\sum_{i=1}^{m}\begin{bmatrix}(\theta^T x^{(i)} - y^{(i)})x_0^{(i)} \\ (\theta^T x^{(i)} - y^{(i)})x_1^{(i)} \\ \vdots \\ (\theta^T x^{(i)} - y^{(i)})x_n^{(i)}\end{bmatrix}$$

q.3)
a) X = 2,5,7,7,25, n = 6, and the values are assumed to be randomly distributed.

The mean is:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} i = \frac{(2+5+7+7+9+25)}{6} = 9.1\overline{6}$$

The standard deviation is:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = 54.81$$

The variance therefore is

$$\sigma = \sqrt{54.81} = 7.40$$

b) $p(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$
Using the mean and variance calculated before and filling in x = 20 this gives:
$p(20) = \frac{1}{7.40\sqrt{2\pi}}e^{\frac{-(20-9.167)^2}{2*54.81}} \approx 0.018$

c) The fact that the probabilities of the values are independent of each other and equally distributed means that the probabilities can be multiplied with each other. Therefore $fx_1 \dots x_6(x_1, \dots, x_6) = p(2) * p(5) * p(7) * p(7) * p(9) * p(25) \approx 0$

d) The most important difference between the LMS in the cost function and the covariance is that the covariance uses with the mean of the sample set whereas the Cost function applies the Hypothesis function that is updated every iteration. Other than that the covariance multiplies the measurements minus the respective means and than adds those numbers. Whereas the LMS does not multiplies these measurements but only adds them up. The covariance therefore explains the relation between two data sets.

e) The formula for the covariance is $cov(X, Y) = \frac{\sum_{i=1}^{n}((x_i - \overline{x})(y_i - \overline{y})}{n-1}$. With the mean of X being 9.167 and with the mean of Y equal to 6.167, filling in the formula returns $cov(X, Y) = 87.83/5 = 17.57$.