Leren Written assignment 1

Micha de Groot

November 2016

1

(a)

A supervised learning task: Matching age groups of customers to certain products.

An unsupervised learning task: Classifying customers in groups of certain shopping behaviour. This can than be used for specific marketing

(b)

1. Age of a customer combined with their purchases.
2. Customers and their shopping history.

(c)

1. Age, times a certain item was purchased, customer ID.
2. Customer ID, all items purchases by a customer.

(d)

1. A model of what products a customer of a certain age will most likely buy
2. A model of a way to cluster segments of customers

2

(a)

Cost function: $J(\theta) = \frac{1}{2*3}\sum_{i=1}^{3}(0 + 1 * x^{(i)} - y^{(i)})^2$

The cost is: $\frac{1}{2*3}((0+1*6-5)^2 + (0+1*5-6)^2 + (0+1*3-10)^2) = \frac{161}{6} = 8.5$

(b)

Update rule:
$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 * x^{(i)} - y^{(i)})$
$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 * x^{(i)} - y^{(i)}) * x$

Updated parameters:
$\theta_0 := 0 - 0.01 * \frac{1}{3}((0 + 1 * 6 - 5) + (0 + 1 * 5 - 6) + (0 + 1 * 3 - 10)) = -0.02$
$\theta_1 := 1 - 0.01 * \frac{1}{3}((0 + 1 * 6 - 5) * 6 + (0 + 1 * 5 - 6) * 5 + (0 + 1 * 3 - 10) * 3) = 1.06$

New cost function:
$J(\theta) = \frac{1}{2*3}((-0.02 + 1.06 * 6 - 5)^2 + (-0.02 + 1.06 * 5 - 6)^2 + (-0.02 + 1.06 * 3 - 10)^2) = 8.18$

(c)

Updated parameters:
$\theta_0 := -0.02 - 0.01 * \frac{1}{3}((-0.02 + 1.06 * 6 - 5) + (-0.02 + 1.06 * 5 - 6) + (-0.02 + 1.06 * 3 - 10)) = 0.001$
$\theta_1 := 1.06 - 0.01 * \frac{1}{3}((-0.02 + 1.06 * 6 - 5) * 6 + (-0.02 + 1.06 * 5 - 6) * 5 + (-0.02 + 1.06 * 3 - 10) * 3) = 1.114$

New cost function:
$J(\theta) = \frac{1}{2*3}((0.001 + 1.114 * 6 - 5)^2 + (0.001 + 1.114 * 5 - 6)^2 + (0.001 + 1.114 * 3 - 10)^2) = 7.89$

The cost has decreased so the two steps have improved the theta values.

3

It is possible to extend the gradient and cost with a quadratic term. the new update rules will be:
$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 * x^{(i)} + \theta_2 * x^{2(i)} - y^{(i)})$
$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 * x^{(i)} + \theta_2 * x^{2(i)} - y^{(i)}) * x^{(i)}$
$\theta_2 := \theta_2 - \alpha \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 * x^{(i)} + \theta_2 * x^{2(i)} - y^{(i)}) * x^{2(i)}$

4

To calculate the optimal value for θ_1 directly we set the derivative of the cost to 0 and computer for θ_1:
$\theta_0 = 0$
$\frac{1}{m} * \sum_{i=1}^{m} (\theta_1 * x^{(i)} - y^{(i)}) = 0$