**Linear Regression: A Solved Examples**

Consider a dataset with two attributes $x_1$ and $x_2$ a target variable $y$:

| $x_1$ | $x_2$ | $y$ |
|---|---|---|
| 1 | 2 | 3 |
| 2 | 1 | 2 |
| 2 | 3 | 5 |
| 3 | 2 | 4 |
| 3 | 3 | 6 |
| 4 | 1 | 4 |

**Step 1: Hypothesis Function**

The linear regression hypothesis for multiple variables is:

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

**Step 2: Cost Function (Mean Squared Error):**

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Where, m =6

**Step 3:** We need to determine the parameters using (Batch) Gradient descent algorithm.

*Repeat until convergence*

*{*

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

*}*

Initialize Parameters: $\boldsymbol{\theta_0=0}$, $\boldsymbol{\theta_1=0}$, $\boldsymbol{\theta_2=0}$, Learning rate α=0.01

**With initial parameters, estimate the Cost function for all the 6 training examples:**

$$J(\theta) = \frac{1}{2*6} * (9 + 4 + 25 + 16 + 36 + 16)$$

$$J(\theta) = 8.833$$

**Step 4: Gradient Calculation:** We need to estimate the gradient for all the individual parameters ($\boldsymbol{\theta}_0,\ \boldsymbol{\theta}_1, \boldsymbol{\theta}_2$)

$$\frac{\partial J}{\partial \theta_0} = \sum_{i=0}^{6} (h_\theta(x^i) - y^i)$$

$$= \frac{1}{6} * (-3 - 2 - 5 - 4 - 6 - 4)$$

$$= -4$$

$$\frac{\partial J}{\partial \theta_1} = \sum_{i=0}^{6} (h_\theta(x^i) - y^i)x_1^i)$$

$$= \frac{1}{6} * \big((-3 * 1) + (-2 * 2) + (-5 * 2) + (-4 * 3) + (-6 * 3) + (-4 * 4)\big)$$

$$= -10.5$$

$$\frac{\partial J}{\partial \theta_2} = \sum_{i=0}^{6} (h_\theta(x^i) - y^i)x_2^i)$$

$$= \frac{1}{6} * \big((-3 * 2) + (-2 * 1) + (-5 * 3) + (-4 * 2) + (-6 * 3) + (-4 * 1)\big)$$

$$= -8.83$$

**Step 5: Update Parameters:**

$$\theta_0 = \theta_0 - \alpha \frac{\partial J}{\partial \theta_0}$$

$$\theta_0 = 0 - (0.01) * (-4)$$

$$\theta_0 = 0.04$$

$$\theta_1 = \theta_1 - \alpha \frac{\partial J}{\partial \theta_1}$$

$$\theta_1 = 0 - (0.01) * (-10.5)$$

$$\theta_1 = 0.105$$

$$\theta_2 = \theta_2 - \alpha \frac{\partial J}{\partial \theta_2}$$

$\theta_2 = 0 - (0.01) * (-8.83)$

$\theta_2 = 0.0883$

**Step 6: Updated Hypothesis After 1 Iteration:**

$$\boldsymbol{h_\theta(x)} = 0.04 + 0.105 \boldsymbol{x}_1 + 0.0883 \boldsymbol{x}_2$$

**Step 7: Cost after 1st iteration:**

$$J(\theta) = \frac{1}{2} * \frac{1}{6}(7.084 + 2.761 + 20.116 + 12.029 + 28.945 + 11.91)$$

$$\boldsymbol{J(\theta)} = 6.9042$$

**Logistic regression solved example:**

•Consider the following dataset with 3 training samples and two features:

| X1 | X2 | Y |
|---|---|---|
| 1 | 2 | 1 |
| 2 | 1 | 0 |
| 3 | 2 | 1 |

We have used batch gradient descent in this example.

- Initialize Parameters: $\boldsymbol{\theta_0 = 0.1}$, $\boldsymbol{\theta_1 = 0.2}$, $\boldsymbol{\theta_2 = 0.3}$ , Learning rate α=0.1
- Our hypothesis is a logistic function:

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

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

- Initially, compute the error for the given dataset and assumed parameters:

$$z = \theta_0 + x_1\theta_1 + x_2\theta_2$$

- For the first training sample:
  - 0.1+1*(0.2)+2*(0.3)=0.9
  - Substitute this for the sigmoid equation: $\frac{1}{1+e^{-z}}$=0.7109

- For the second training sample:
  - 0.1+2*(0.2)+1*(0.3)=0.8
  - Substitute this for the sigmoid equation: $\frac{1}{1+e^{-z}}$= 0.6899
- For the third training sample:
  - 0.1+3*(0.2)+2*(0.3)=1.3
  - Substitute this for the sigmoid equation: $\frac{1}{1+e^{-z}}$= 0.7858
- Next we have to compute the error on all training data:

$$Error = \ y - h_{\theta}(x)$$

Error=1-0.7109 + 0-0.6899+1-0.7858=-0.1947

This is the error observed on training data using initial parameters.

We need to update the parameters using gradient ascent.

- Update rule is given as follows:

$$\theta_0 := \theta_0 + \alpha \sum_{i=1}^{3} (y^{(i)} - h_\theta(x^{(i)}))$$

$$\theta_0 = 0.1 + (0.1) * (1 - 0.7109{+}0{-}0.6899{+}1{-}0.7858)$$
$$\theta_1 = 0.2 + (0.1) * ((1 - 0.7109)1{+}(0{-}0.6899)2{+}(1{-}0.7858)3)$$
$$\theta_2 = 0.3 + (0.1) * ((1 - 0.7109)2{+}(0{-}0.6899)1{+}(1{-}0.7858)2)$$

We need to recalculate the error by using the updated parameters.