**ASSIGNMENT – 6**

**Name – Urmi Adak**

**Reg no – SIRSS1125**

**Q1) Calculate/ derive the gradients used to update the parameters in cost function optimization for simple linear regression.**

The equation for simple linear regression is

Y=a1 \* x+a0

Error= y^-y

Mean squared loss=y1-y^i

Mean squared error=1/N ∑i=N1 (yi-y^i)2

Y=actual value b=bias wi=weight or coefficient

Xi= feature

Y=b+w1x1+w2x2+……. + wnxn

Math behind Gradient Descent

W=(w0,w1,w2,w3,………wn)

C(w,b)= Cost function involving parameters W and b

dw=d(w,b)/dw[Partial differentiation of cost function wrt weights]

db=d(w,b)/dw[Partial differentiation of cost function wrt bias]

Update parameters w and b.

W=w-(a\*dw)

B=b-(a\*db)

Cost function

J( Ɵ0 , Ɵ1)= 1/2m ∑i=1 [h0(xi)-yi]2

Gradient Descent

Ɵj= Ɵj-αd/dƟj J(Ɵ0 , Ɵi)

**Q2) What does the sign of gradient say about the relationship between the parameters and cost function?**

Gradient descent is an iterative optimization algorithm for finding the local minimum of a function. To find the local minimum using we must take steps proportional to the negative of gradient at the current point.

* When the sign of gradient is positive then the step will decrease
* When the sign of gradient is negative then the step will increase

**Q3) Why Mean squared error is taken as the cost function for regression problems.**

Mean Squared Error is used to check how close predictions made by the model are to actual values. It calculates the error as actual prediction and squared the difference to eliminate the negative values.

The lower the MSE, the closer is the prediction to actual. In Regression model a lower MSE indicates a better fit.

**Q4) What is the effect of learning rate on optimization, discuss all the cases.**

* It can be seen that for an optimal value of the learning rate, the cost function value is minimized.
* If learning rate selected is very high, the cost function could continue to increase.
* If we take lower than optimal learning rate then even after substantial iterations the cost function will not minimize sufficiently and will take a long time.