KEDARNATH_ Assignment - Linear Regression PART-2

Question - 1

List down at least three main assumptions of Linear Regression and explain them in your own words. To Explain an assumption take an example or a specific use case to show why assumption makes sense.

Answer:

1.) Linearity:
2) Outlies
3.) Auto correlation
4.) Multicollinearity
5.) Heteroskedasticity

=> Linearity => In the above assignment of carprice The catigoical variables -> aspiration, fueltyp etc ~~doe~~ Dont have linear relationship with dependent variable price.
Hence we need to create Scatter plot/ boxplot to check data points fit or not. So, we need to Transform Such charateristics of variable.

=> Multi collinearity => It is to find variable with Similar VIF and that will impact your dependent variable. Hence ther independent variables with more correlation make model difficult to predict the Significance.
In the above assignment citympg and highwayMPG are good examples of multicollinearity.

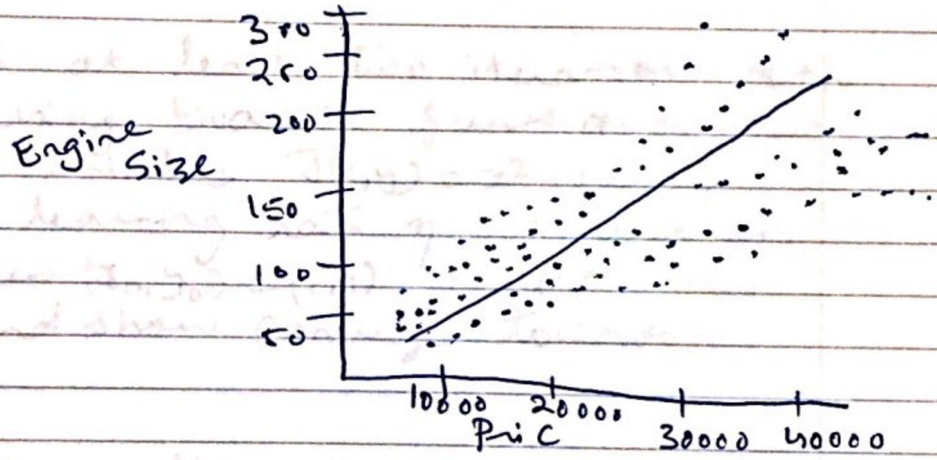
VIF

carlength	10.23 } Show no multicolineaity
whellbase	10.00
car-body	- 1.54 } -> Show multicolinearity
car-heigh	- 1.60

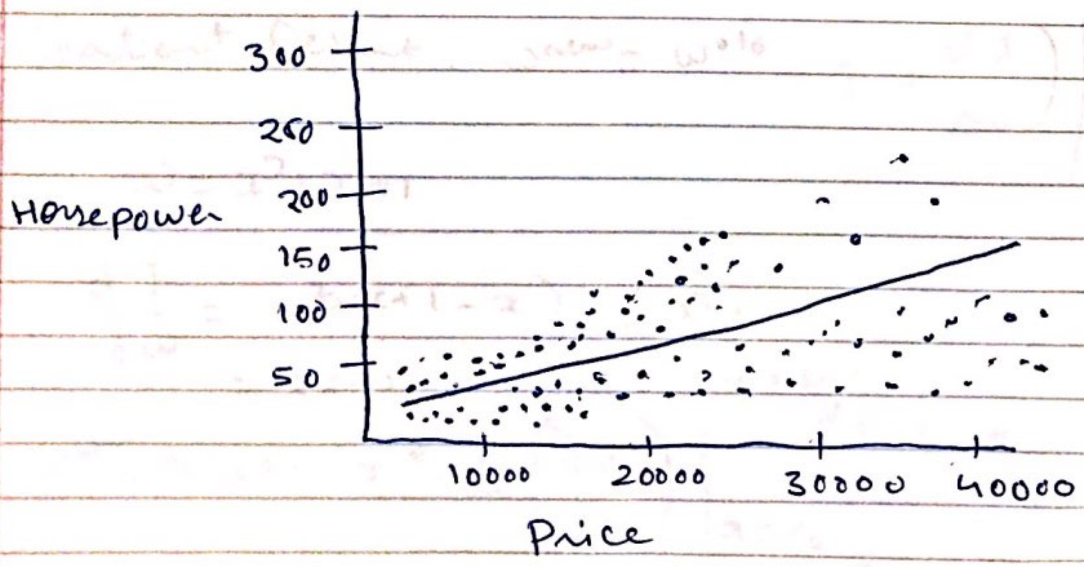
⇒ Heteroskedasticity:

When scatterplot is mapped between independent variable and dependent variable. The plot must look like cone-like shape/ funnel shape from Origin.

In the Part I - Assignment we can see heteroskedasticity for 1.) Engine Size Vs Price



2.) Horsepower Vs Price



Question 2:

Explain the gradient Descent algorithm in following

1.) Illustrate at least two iterations of the algothim using univariate function $J(x) = x^2 + x + 1$. Assume a learing rate $\eta = 0.1$ and an initial guess $x_0 = 1$ and demonstrate the iterations converge towards the minima. Also report the minima (which can compute using the closed form function)

2.) Illustrate at least two iterations of the alogorithm using bivariate funsion of two independent variables $J(x,y) = x^2 + 2xy + y^2$. Assume a learning rate $\eta = 0.1$ and an initial guess $(x_0, y_0) = (1,1)$. Report the minima and show converge towards it.

Answer:

1.) Power Rule $f(x) = x^n \Rightarrow nx^{n-1}$

Gradient Desent : $w^{new} = w^{old} - \eta \left(\frac{dJ}{\partial w}\right)\Big|_{w = w^{old}}$

$J(x) = x^2 + x + 1$

$\frac{\partial J}{\partial w} = (x^2 + x + 1 - x) = x^2 + 1$

$= 2x + 1$

where $x_0 = 1$, $\eta = 0.1$

$w^{new} = J'(x) = x^0 - \eta \left(\frac{\partial J}{\partial x}\right)\Big|_{x = x_0}$

$= 1 - (0.1)(2(1)+1) = 1 - (0.1)(3) = 0.7$

$\boxed{1^{st} \text{ Itration} = 0.7}$

1.) 2^{nd} Iteration

$$w^2 = w^1 - \eta\left(\frac{\partial J}{\partial w}\right)\Bigg|_{w=0.7}$$

$$= 0.7 - (0.1)\,(2(0.7)+1)$$

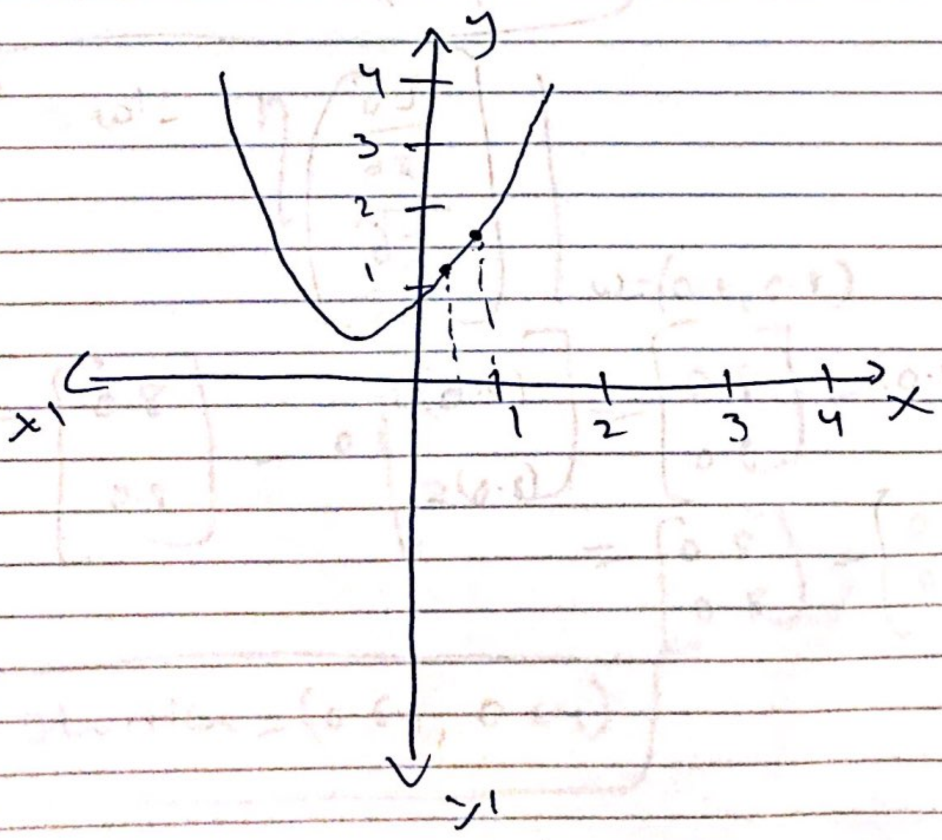
$$= 0.7 - (0.1)(1.4+1)$$

$$= 0.7 - (0.1)(2.4)$$

$$= 0.7 - 0.24$$

$$= 0.46$$

$$\boxed{2^{nd}\ \text{Iteration} = 0.46}$$



2.) $J(x,y) = x^2 + 2xy + y^2$

$$\frac{\partial J}{\partial x} = 2x \qquad \frac{\partial J}{\partial y} = 2y$$

$$\Rightarrow \quad w^1 = w^0 - \eta \begin{bmatrix} \frac{\partial J}{\partial x} \\ \frac{\partial J}{\partial y} \end{bmatrix} \Bigg|$$

$\eta = 0.1$

$(x,y) = (1,1)$

$$w^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - (0.1)\begin{bmatrix} 2(1) \\ 2(1) \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.8 \end{bmatrix}$$

$$\boxed{1^{st} \text{ Iteration} = (0.8, 0.8)}$$

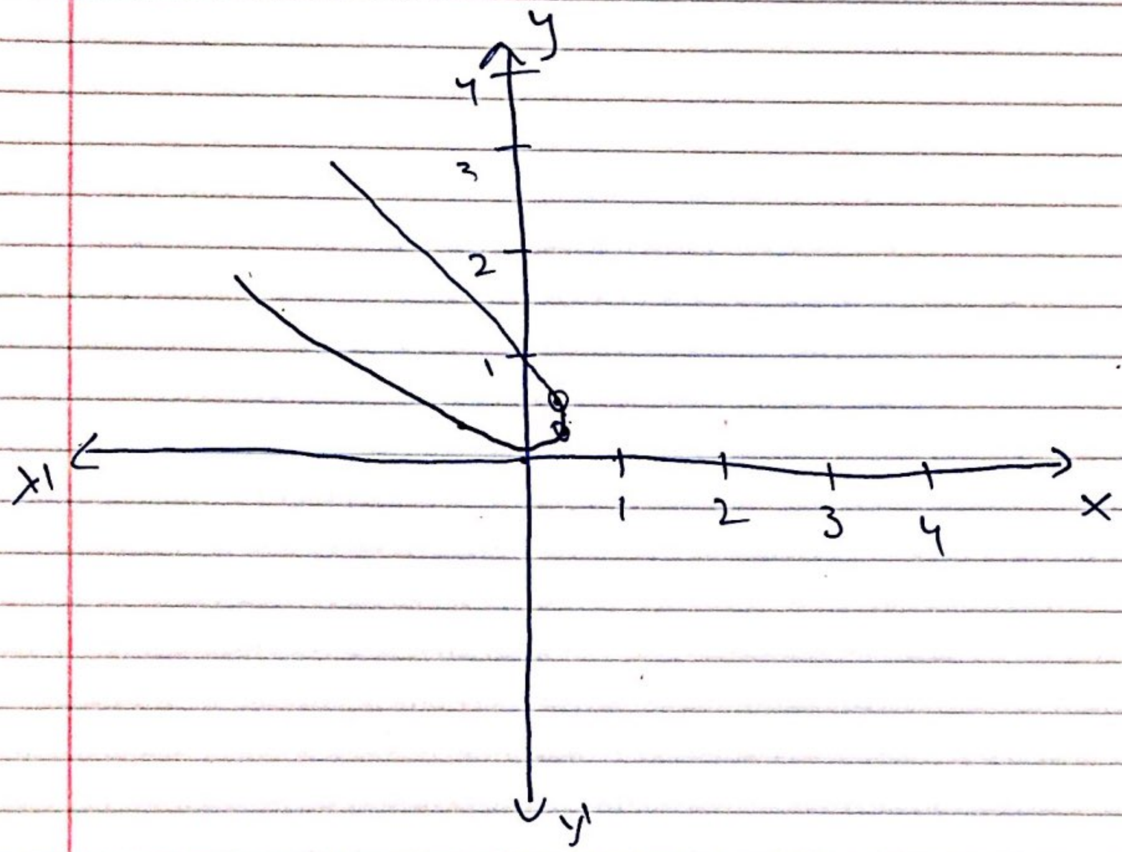
$$\Rightarrow \quad w^2 = w^1 - \eta \begin{pmatrix} \frac{\partial J}{\partial x} \\ \frac{\partial J}{\partial y} \end{pmatrix} \Bigg|_{w=(0.8,\,0.8)}$$

$$= \begin{bmatrix} 0.8 \\ 0.8 \end{bmatrix} - 0.1\begin{bmatrix} 2(0.8) \\ 2(0.8) \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.8 \end{bmatrix} - (0.1)\begin{bmatrix} 1.6 \\ 1.6 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8 \\ 0.8 \end{bmatrix} - \begin{bmatrix} 0.16 \\ 0.16 \end{bmatrix} = \begin{bmatrix} 0.64 \\ 0.64 \end{bmatrix}$$

$$\boxed{2^{nd} \text{ Iteration} = (0.64, 0.64)}$$

5 4


y
4
3
2
1
x'
1
2
3
4
x
y'