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 the assumption makes sense.

Answer:

As mentioned above let us consider a case study of Event management company, Every event management company Takes up responsibilities of different events like Marriages, Birthday parties and many more in our case study we will limit ourselves to Marriages. .so let us consider a Event Management company which takes responsibility of Marriages. So whenever these people accept request from client Event managers will consider their previous records and estimate the overall price for the new request. So we need to build a Regression Model for this .Lets make assumptions on Linear Regression model using Event management dataset.

Event Management Dataset consists : 1)Total budget aff: This tells how much budget the costumer can afford.

2) Invitations : Number of people who are invited for the Marriage.

3) Catering Expenses : Amount of money estimated for food.

4) Other expenses : this includes other expense like decoration Marriage hall bookings and others.

5) Total estimated budget : Total budget that is estimated taking this is dependent variable.

In order to carry out statistical inferences it is mandatory for us to make some good assumptions on our model.

Assumptions in Linear Regression:

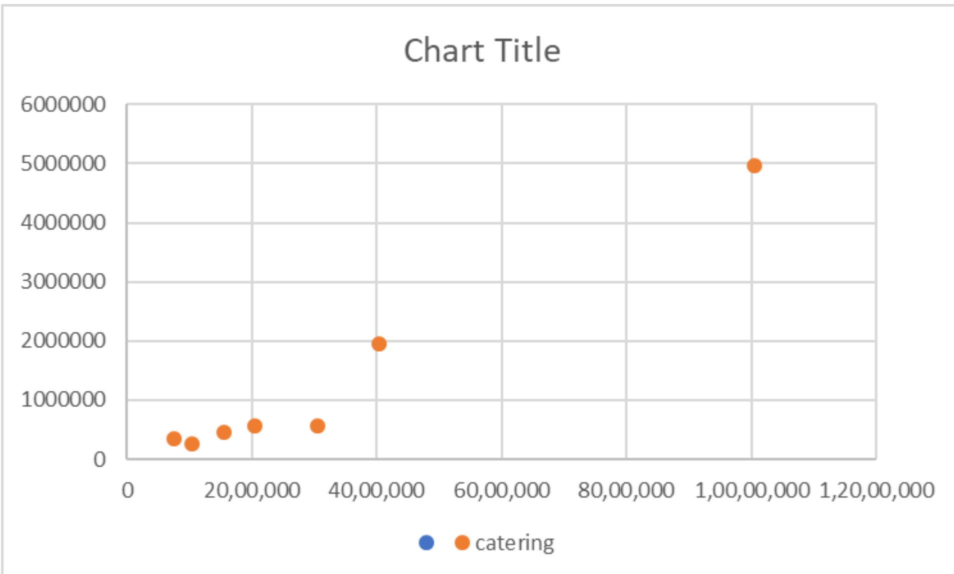
-Linear Regression Model can be applied only on **Numerical data** not on Categorical data In our case study of Event Management for Marriages all of the columns are Numerical Not categorical if we have categorical columns we need to change it into Numerical data.

invitations	total budget aff	catering	others	estimated
200	10,00,000	3,00,000	2,00,000	6,00,000
500	30,00,000	6,00,000	6,00,000	20,00,000
450	15,00,000	5,00,000	5,00,000	11,00,000
600	20,00,000	6,00,000	6,00,000	18,00,000
10000	1,00,00,000	50,00,000	30.00.000	90.00,000

600	40,00,000		20,00,000	15,00,000	37,00,000
300	7,00,000		4,00,000	2,00,000	7,00,000

-Linearity: Relationship between an independent and dependent variable should be Linear This can be seen using scatter plot in our case study dependent variable is total estimated budget and independent variable is invitations.

This is the sample data for our case study



This is the scatter plot between catering and estimated we can see that it follows some linearity as own dataset is small it is not clear but if dataset increases the view can be more clear

OUTLIER: outliers are those points which deviate from the path of fitted line so in our case study we have only one outlier where as if we have more outliers we need to build models with and without the outliers so that we can get the efficient model. we can use boxplot for knowing Outliers.

Multicollinearity: Multicollinearity is seen between independent variables if there is more collinearity then we cannot get good model to predict multicollinearity we need to use VIF values if VIF value is less than 4 then the model has no multicollinearity.

We can drop the columns which has greater multicollinearity.

Question-2:

Explain the gradient descent algorithm in the following two parts:

1. **Illustrate at least two iterations of the algorithm using the univariate function $J(x)=x^2+x+1$. Assume a learning rate $\eta=0.1$ and an initial guess $x_0=1$ and demonstrate that the iterations converge towards the minima. Also, report the minima (which you can compute using the closed form solution).**
2. **Illustrate at least two iterations of the algorithm using the bivariate function 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 that the solution converges towards it.**

3.

Gradient descent Method

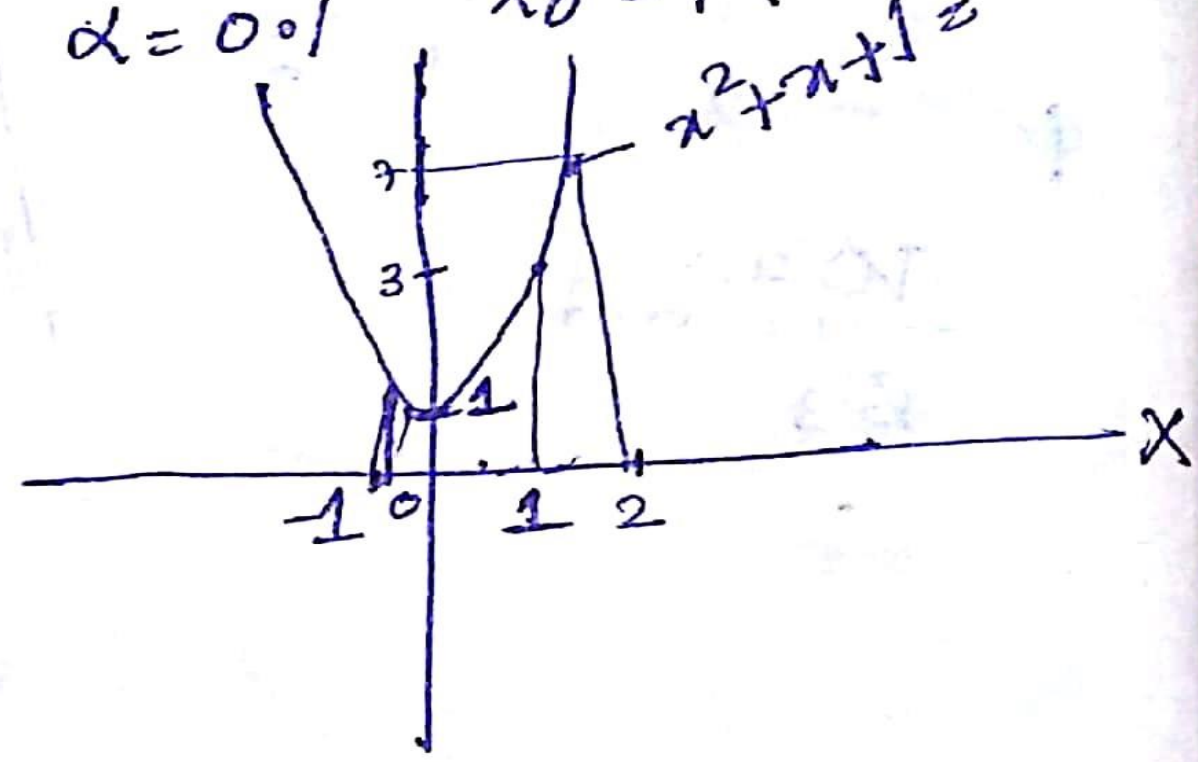
It is a method which is used to find the minimum value of y using the value of x.

We use formula: $X_{i+1} = X_i - \alpha f'(X_i)$ —— (1)

X_{i+1} = Next Guess α = learning rate.

X_0 = initial guess.

(1) $J(x) = x^2 + x + 1$ $\alpha = 0.1$ $X_0 = 1$.



$$J'(x) = \frac{d(J(x))}{dx}$$

$$= \frac{d(x^2+x+1)}{dx}$$

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

$J'(X_0=1) = 3$.

Subs in (1) $X_1 = X_0 - \alpha J'(X_0=1)$

$= 1 - (0.1)(3)$

$= 1 - 0.3$.

$X_1 = \underline{0.7}$

$= 2(0.7) + 1$

$= 1.4 + 1$

$= 2.4$.

Next guess $\Rightarrow X_1 = 0.7$.

$X_2 = X_1 - (0.1)(J'(X_1 = 0.7))$

$= 0.7 - (0.1)(2.7)$

$= 0.7 - \cancel{0.07}\ 0.24$

$X_2 = \cancel{0.63}.\ \underline{0.46}$

~~$X_3 = X_2 - (0.1)(0.63)$~~

~~$= 0.63 - 0.063 = \underline{0.567}$.~~

$$x_3 = x_2 - (0.1)(J'(x_2 = 0.46))$$

$$J'(x_2) = 2(0.46) + 1$$

$$= 0.92 + 1 = 1.92.$$

$$(0.1)(J'(x_2)) = 0.192.$$

$$x_3 = 0.46 - 0.192.$$

$$x_3 = 0.268.$$

$$x_4 = x_3 - (0.1)(J'(x_3 = 0.268))$$

$$J'(x_3) = 2(0.268) + 1$$

$$= 0.536 + 1 = 1.536.$$

$$x_4 = 0.268 - 0.1536. = 0.1144$$

$$x_5 = x_4 - (0.1)(J'(x_4 = 0.1144))$$

$$J'(x_4) = 2(0.1144) + 1 = 1.2288.$$

$$x_5 = 0.1144 - 0.12288 = -0.00848.$$

So Till x_4 almost $x_4 \cong 0$.

for $x_5 = -0.00848$

~~$x_5 = -0.00848 -$~~ (0.1) it the follow the Same process for Consecutively guess Value is Constant

It is the Value where y is min.

2)

(2) If Cost function $J(\theta_0, \theta_1, \ldots \theta_n)$

$$= \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

Gradient descent : (Algorithm)

Repeat {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0 \ldots \theta_n)$$

} $(j = 0, 1)$ (Simultaneously update θ_0, θ_1)

Given Question :

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

$$n = 0.1 \quad (x_0, y_0) = (1, 1)$$

I-Iteration

$$x_1 = x_0 - \alpha \frac{\partial}{\partial x} [J(x, y)]$$

$$= 1 - (0.1) \left[\frac{\partial}{\partial x} (x^2 + 2xy + y^2) \right]$$

$$= 1 - (0.1) [2x + 2y]$$

$$= 1 - (0.1) (2x_0 + 2y_0]$$

$$= 1 - (0.1) (2 + 2) = 1 - (0.4)$$

$$X_1 = 0.6$$

$$Y_1 = Y_0 - \alpha \frac{\partial}{\partial y} [J(x, y)]$$

$$= 1 - (0.1) \left[\frac{\partial}{\partial y} (x^2 + 2xy + y^2) \right]$$

$$= 1 - (0.1) [0 + 2x + 2y]$$

$$Y_1 = 1 - (0.1) [4] = 0.6$$

II Iteration

$$X_2 = X_1 - (0.1)[(2x_1 + 2y_1)].$$

$$= (0.6) - (0.1)(2[0.12])$$

$$= (0.6) - (0.1)[0.24] = (0.6) - (0.024)$$

$$\underline{X_2 = 0.576.}$$

$$Y_2 = Y_1 - (0.1)[(2x_1 + 2y_1)]$$

$$= \overset{(0.6)}{\cancel{0.576}} - (0.1)[2(\overset{0.12}{\cancel{1.152}})]$$

$$= \overset{(0.6)}{\cancel{0.576}} - \cancel{0.0304}\ 0.024$$

$$Y_2 = \underline{\underline{0.576}}.$$

$\therefore$ We can say that ~~minima~~ solution converges towards minima.