Linear Regression

Supervised Machine Learning
Regression Algorithm

Course Topics

- ✓ What is Linear Regression?
- ✓ How the algorithm works
- ✓ Linear Regression Assumptions
- ✓ Pros and Cons
- ✓ Applications
- ✓ Modeling Linear Regression

SKILLOGIC

Datamites

What is Linear Regression?

- Linear Regression is a Supervised Machine Learning Algorithm.
- Linear Regression is used to predict the continuous value.
- Linear Regression is used to identify the linear relationship between dependent and independent variables.

1	Height (in cm)	Weight (in Kg)
2	151	63
3	174	81
4	138	56
5	186	91
6	128	47
7	136	57
8	179	76
9	163	72
10	152	62
11	131	48

SKILLOGIC

Datamites

How Linear Regression works?

$$Y = X_1 + X_2 + X_3$$

Dependent Variable	Independent Variable
Outcome Variable	Predictor Variable
Response Variable	Explanatory Variable

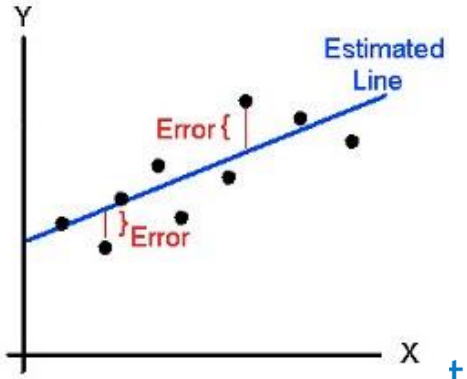
Estimated (or predicted) Y value for observation i

Estimate of the regression intercept

Estimate of the regression slope

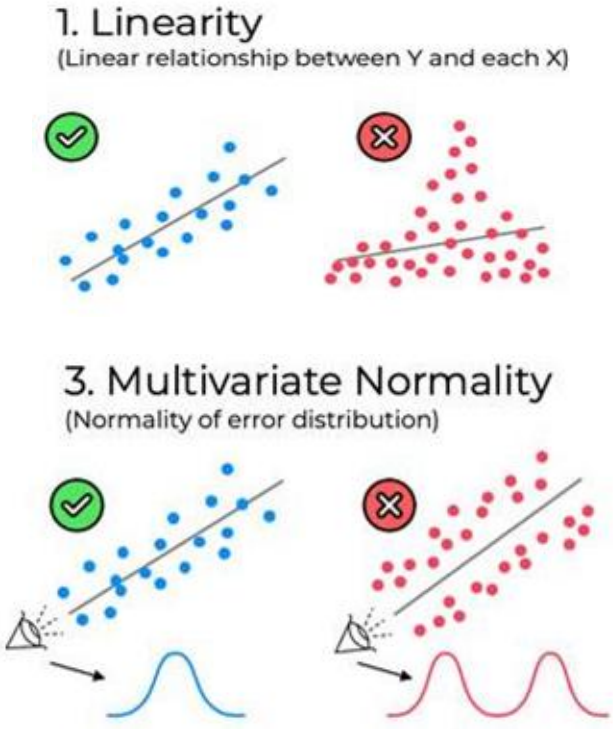
Value of X for observation i

$$\hat{Y}_i = b_0 + b_1 X_i$$



Assumptions of Linear Regression

- **Linear Relationship**: There should be a linear relationship between dependent and independent variables.
- **Multivariate normality :** Linear regression analysis requires all variables to be multivariate normal**. Residual (Error)** should be following normal distribution.
- **No or little collinearity:** Linear regression assume that there is no multi-collinearity in the data, i.e. the independent variables are not interrelated.



SKILLOGIC

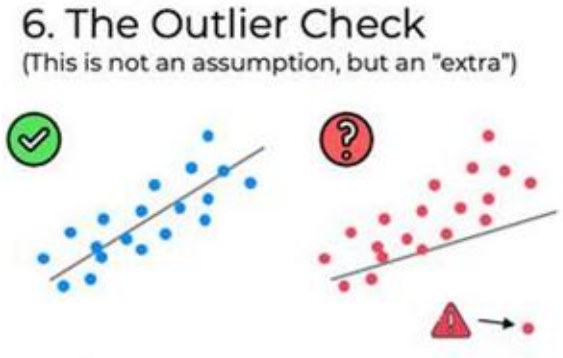
Datamites

Pros of Linear Regression

- Linear regression is an extremely simple method. It is very easy and intuitive to use and understand.
- Widely used, as many problems can be transformed into linear regression problems.
- With regularization techniques, handling overfitting issue quite well.

Cons of Linear Regression

- Assumes there is a straight-line (Linear) relationship between them. If your data are intrinsically nonlinear then linear regression may not yield best results.
- Linear regression is very sensitive to the outliers in the data.



SKILLOGIC

Datamites

Applications of Linear Regression

Majority continuous variable predictions are through Linear Regression in business applications. Here are some applications--

- Stock Exchange – Price predictions
- Weather forecast - temperatures
- Flight price prediction
- House price predictions

SKILLOGIC

Modeling Linear Regression

Regression in action