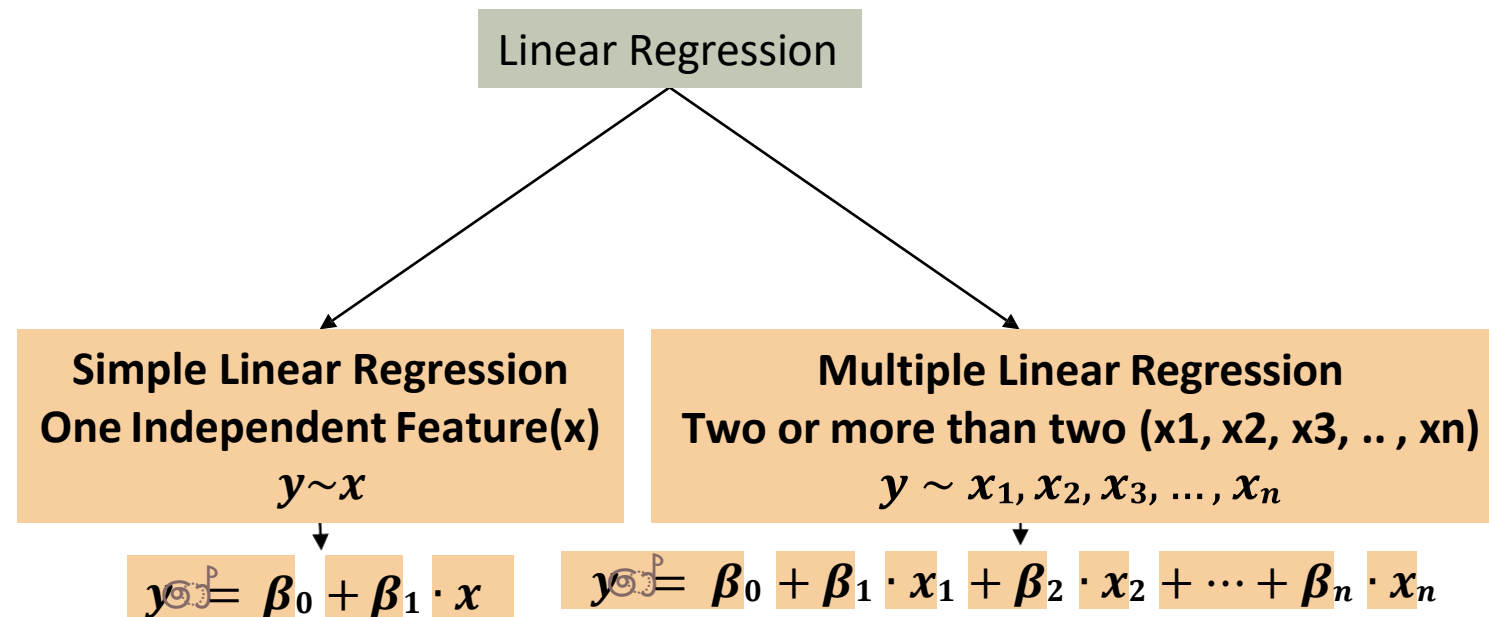


Multiple Linear Regression

What is Multiple linear regression



Example of MLR

Age	Income	Purchase
25	50,000	1,000
30	60,000	2,000
35	70,000	3,000
40	125,000	4,000
45	150,000	5,000

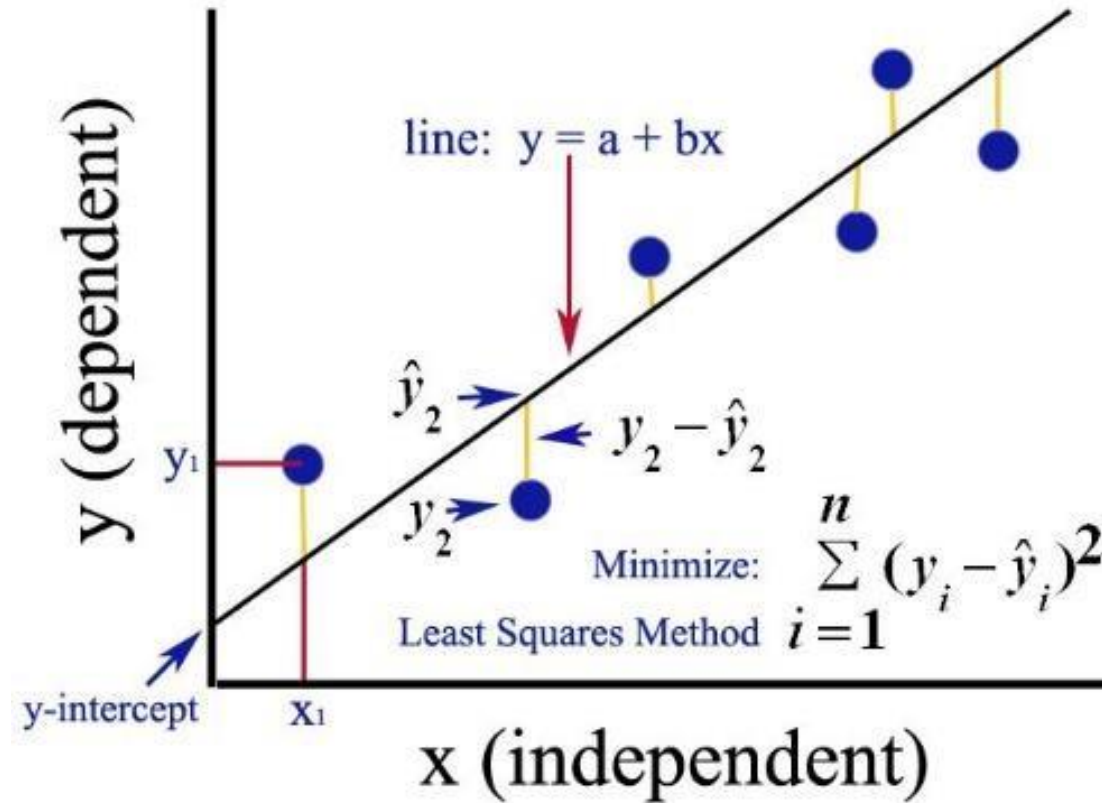
$Purchase \sim Age, Income$

$$Purchase_cap = \beta_0 + \beta_1 \cdot Age + \beta_2 \cdot Income$$

Cost Function : Mean Squared Error
i.e. Minimize The Squared Error in model
Least Squared Error Model

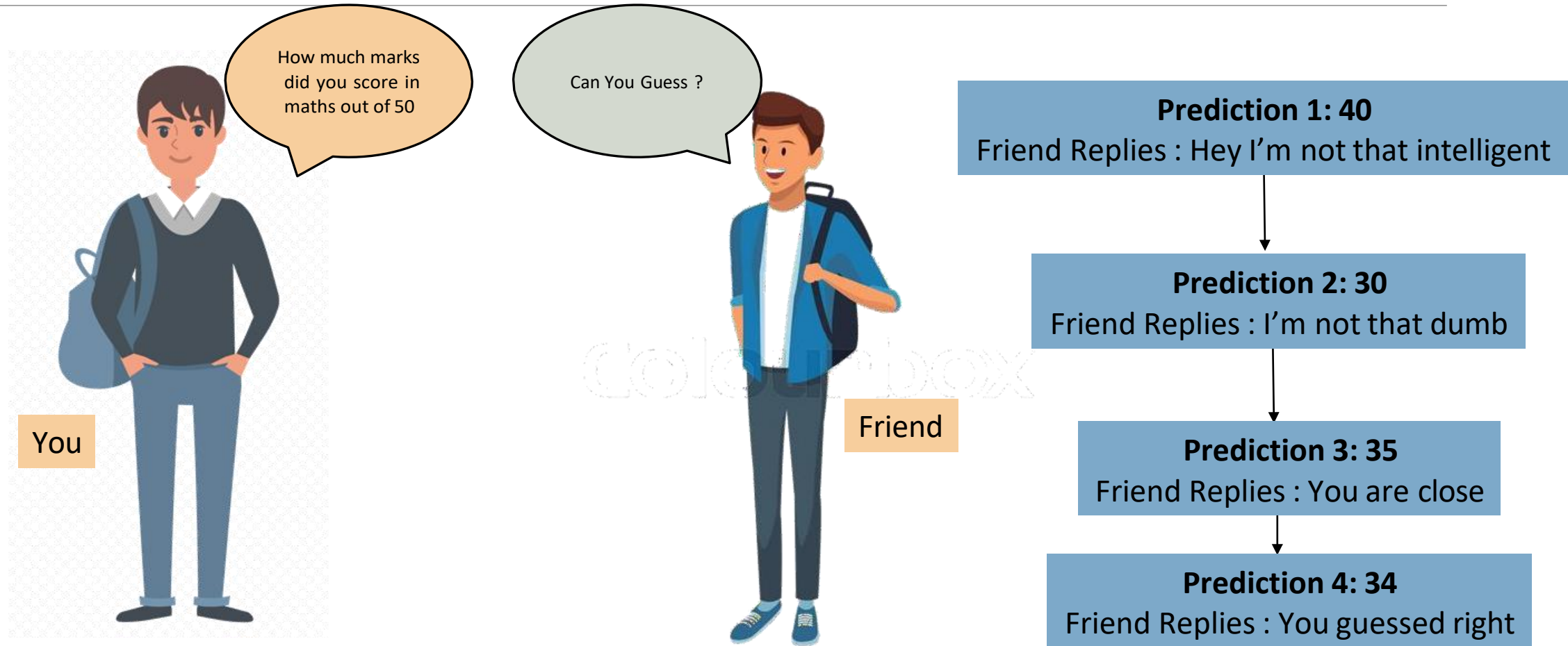
Dependent Variable
(Which is to be predicted)

Least Squared Error model

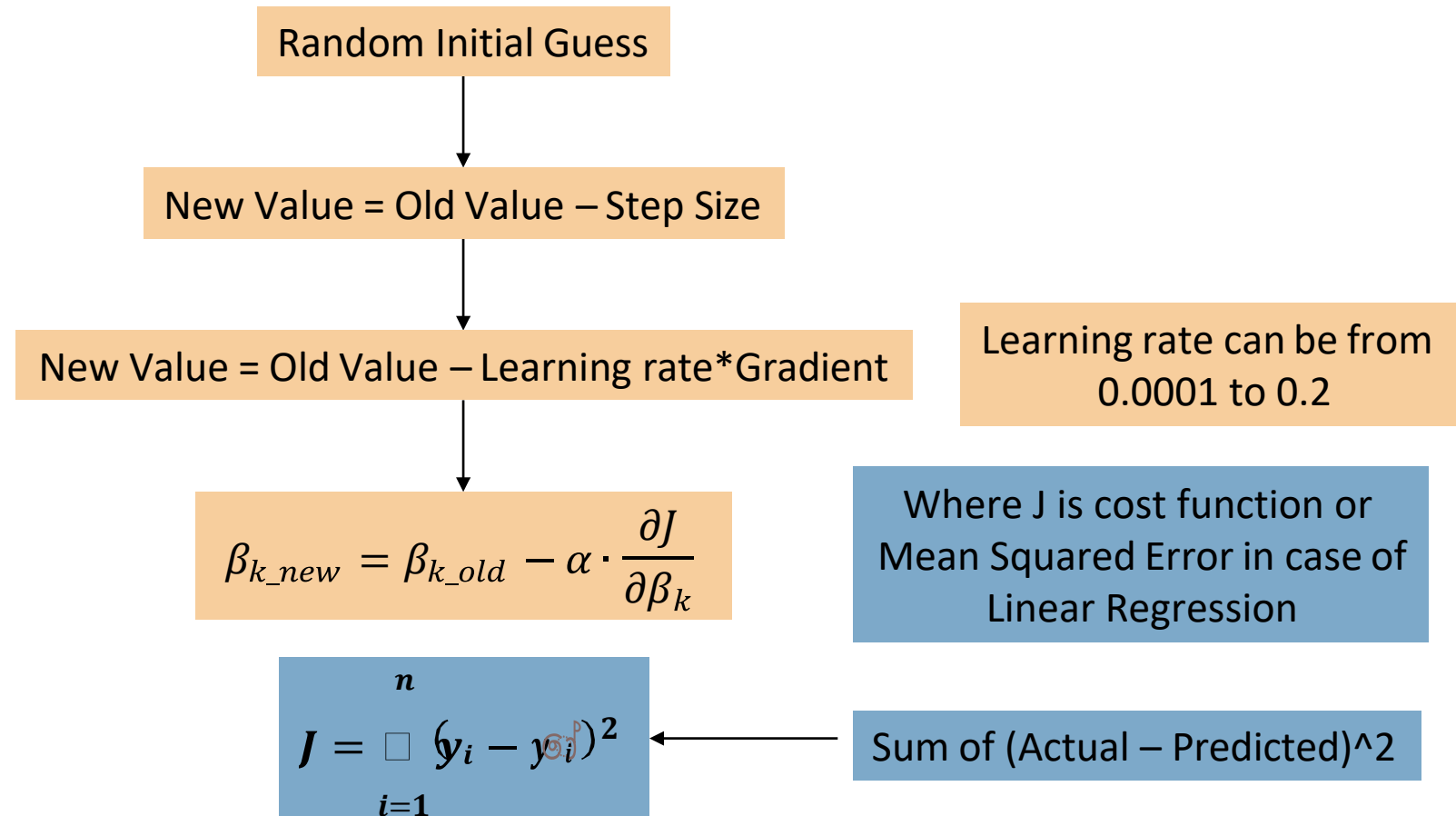


How are coefficients obtained in MLR?

Gradient Descent Algorithm

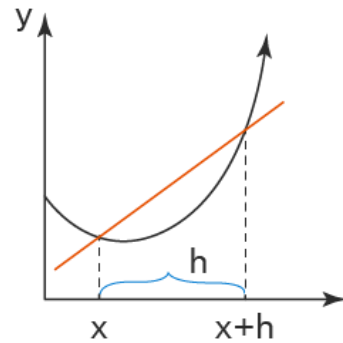


Gradient descent algorithm



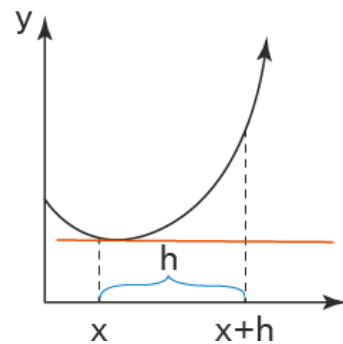
What is a derivative?

Definition of Derivative



$$\text{Slope of Secant} = \frac{f(x+h) - f(x)}{h}$$

("Difference quotient")



$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

(if limit exists)

Example

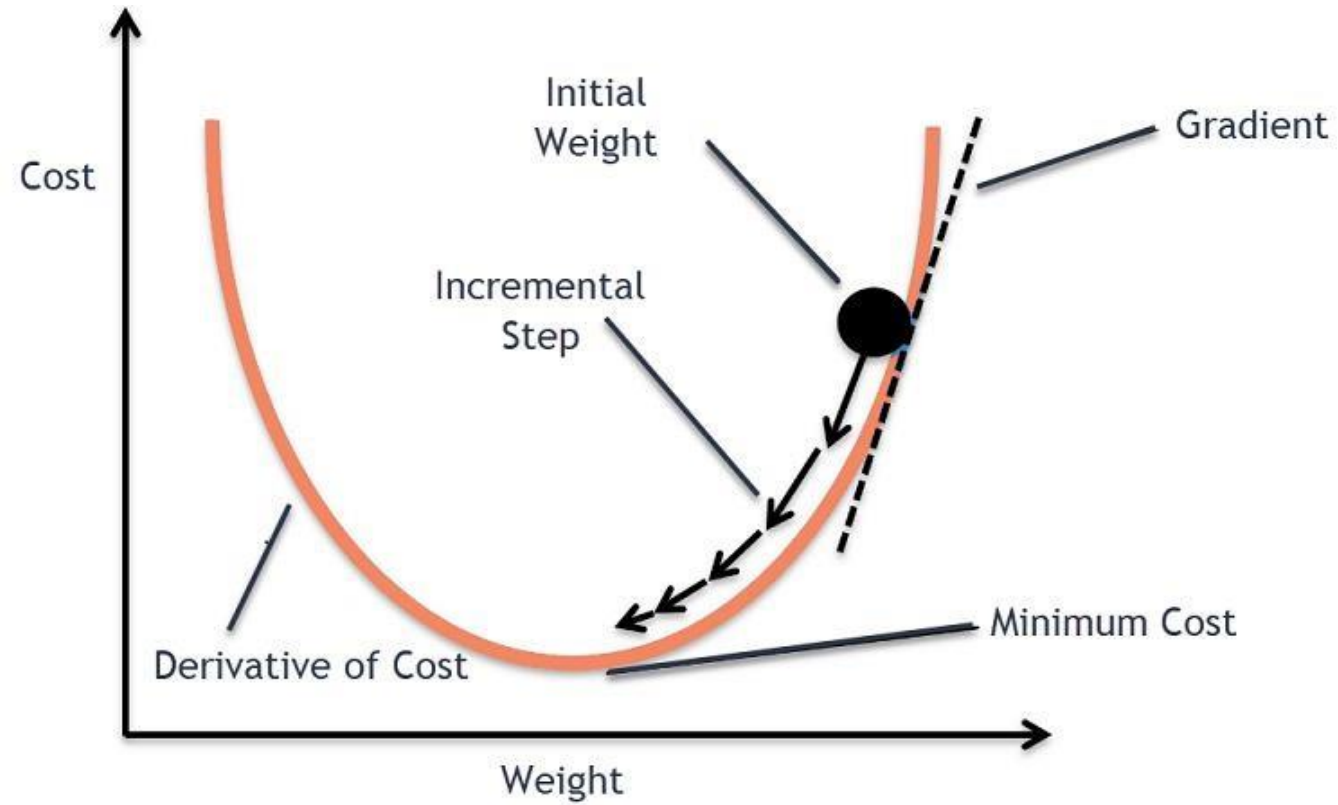
Take a function $y = x^2$

If x increases by small amount at any point
How much will y increase ?

$$x_1 = 2, x_2 = 2.001$$
$$y_1 = 4, y_2 = 4.004001$$

$$\frac{dy}{dx} = \frac{y_2 - y_1}{x_2 - x_1} = 4.001$$

Gradient Descent working



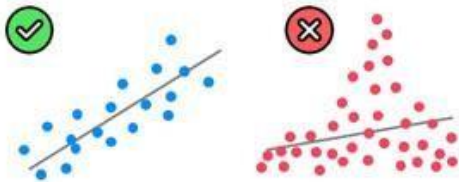
Assumptions in Linear Regression

Assumptions of Linear Regression



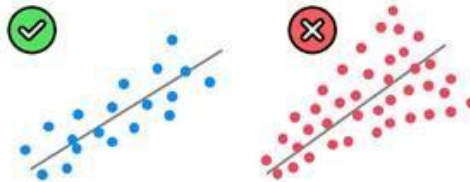
1. Linearity

(Linear relationship between Y and each X)



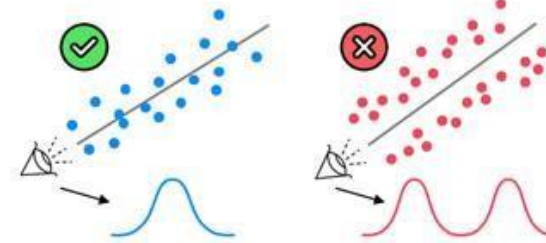
2. Homoscedasticity

(Equal variance)



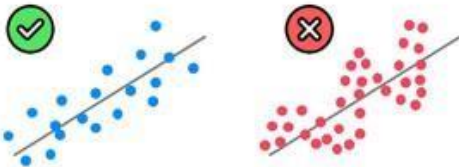
3. Multivariate Normality

(Normality of error distribution)



4. Independence

(of observations. Includes "no autocorrelation")



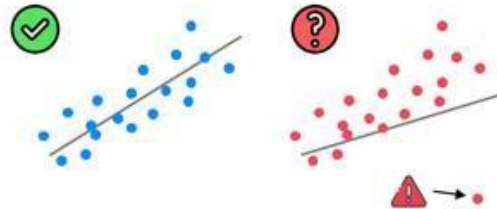
5. Lack of Multicollinearity

(Predictors are not correlated with each other)



6. The Outlier Check

(This is not an assumption, but an "extra")



Thank you
