Linear Regression

Linear regression was developed in the field of statistics

↓ **Borrowed by**

Machine Learning

Linear Regression Is Used For what

Understanding The Relationship Between Input And Output Numerical Variables

Predict an observation's value (Y) based on the relationship

Types of Regression

Linear Regression

Single Input Variable (x)
and
Y variable

Multiple Linear Regression

Multiple Input Variables (X's)
and
Y variable

Assumptions of Linear Regression

- Variables should be **continuous numeric variables**
- No significant outliers or Missing values
- **Linear relationship** between the dependent and independent variables
- Predictor variables are **independent** of each other
- Residuals (aka prediction errors) are **normally distributed**
- **Homoscedasticity**

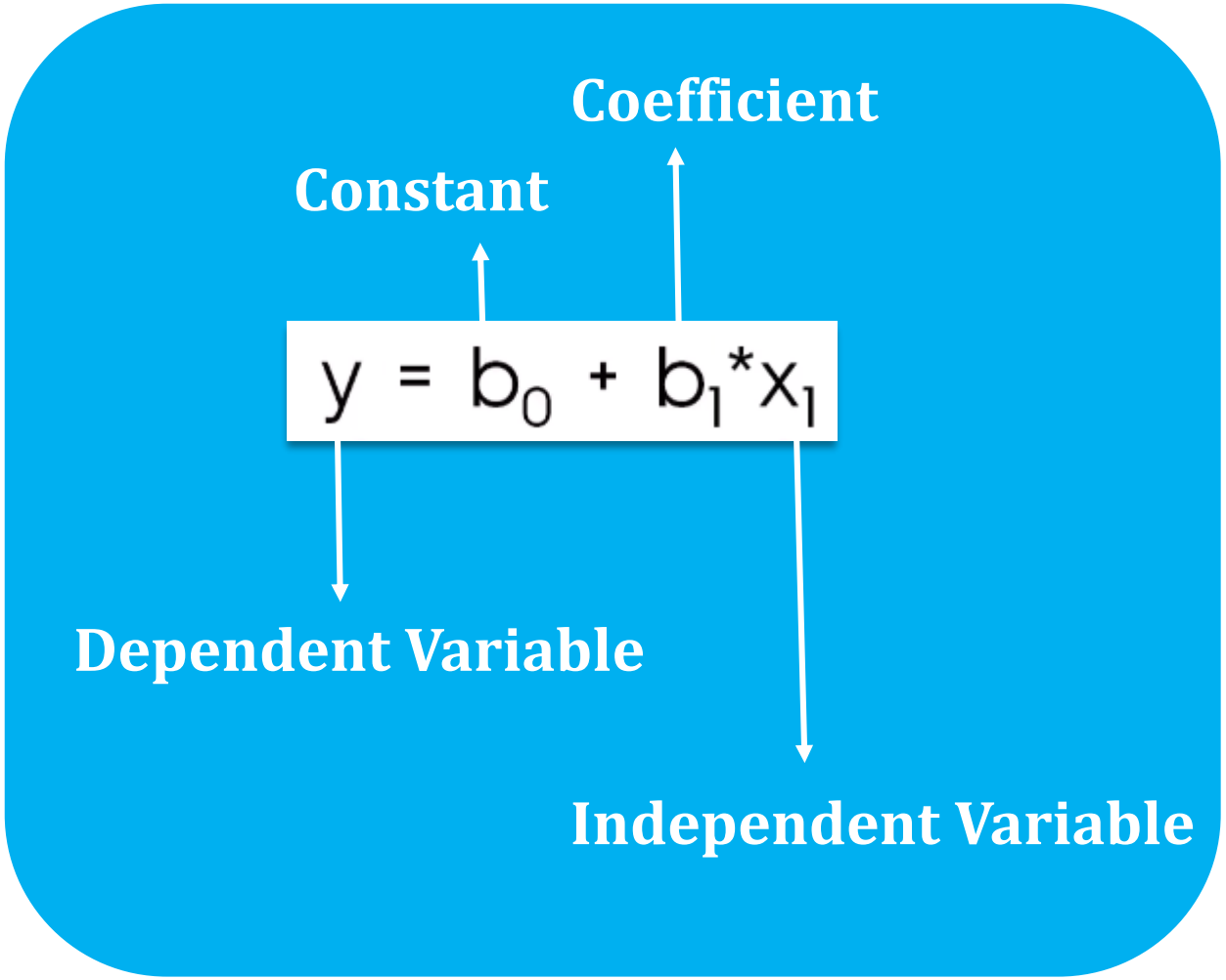
Example: Linear Regression

A	B	C
No_Bedrooms	Total_Sqft	Price
4	1859	270897
3	2002	302404
3	1578	2519996
4	2277	197193
4	1749	207897
3	1672	196559
3	2365	434697
5	1741	64887
5	1745	143636

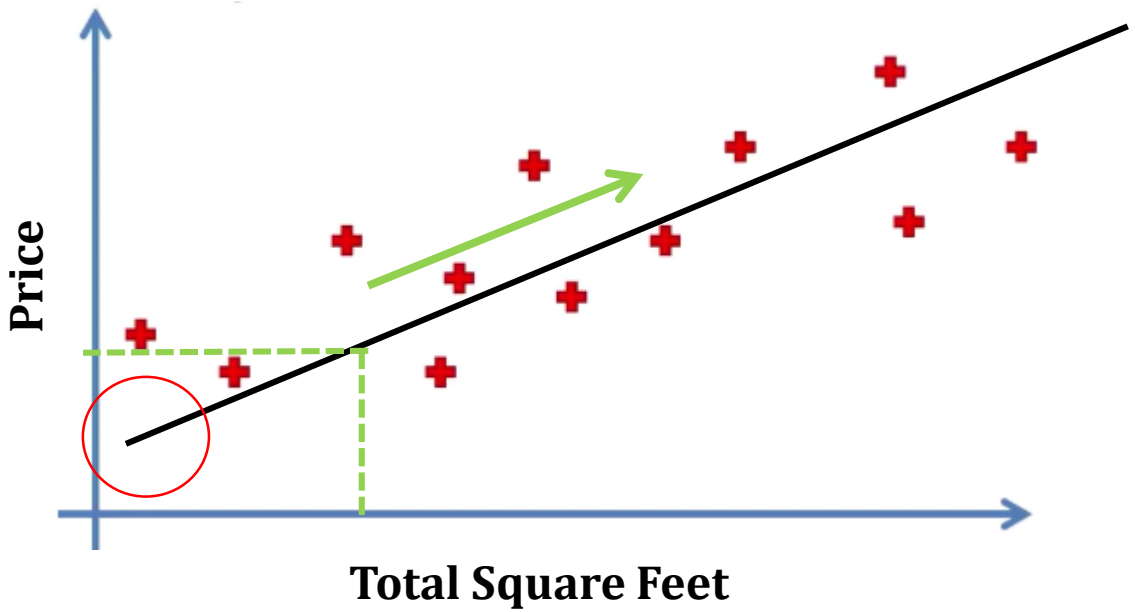
Imagine we are predicting **Price of a flat (Y)** depending upon **Number of bedrooms (X1)** and **Square feet area (X2)**

Example: Lets Work

Linear Regression Works on a Line Equation



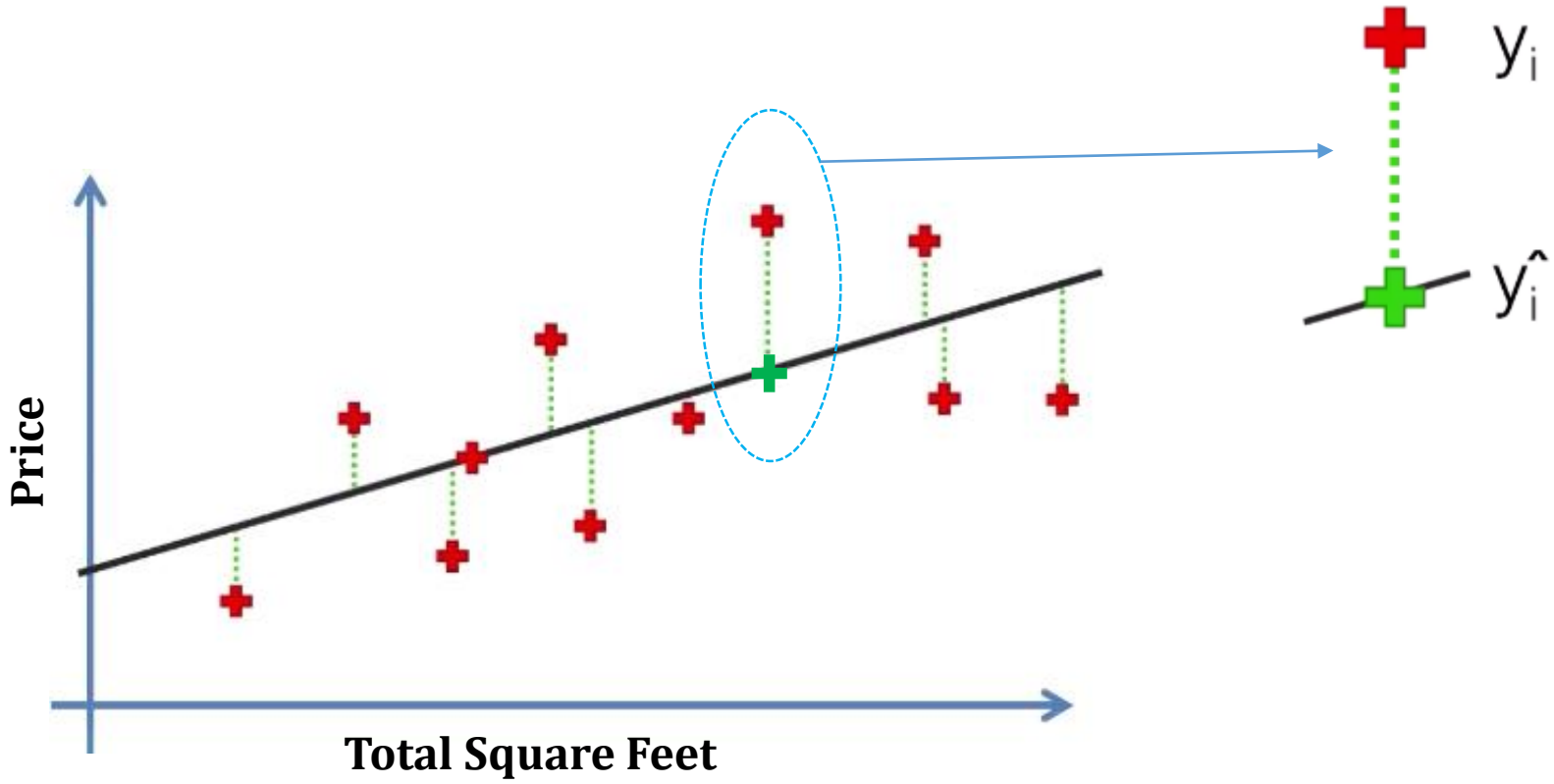
Linear Regression



$$y = b_0 + b_1 * x_1$$

Price = B0 + B1 * Square Feet

Linear Regression



$$SUM\ (y - \hat{y})^2 \rightarrow min$$

Multiple Linear Regression

Simple Linear Regression

$$y = b_0 + b_1 * x_1$$

Multiple Linear Regression

$$y = b_0 + b_1 * x_1 + b_2 * x_2 + \ldots + b_n * x_n$$

Techniques To Train The Linear Regression

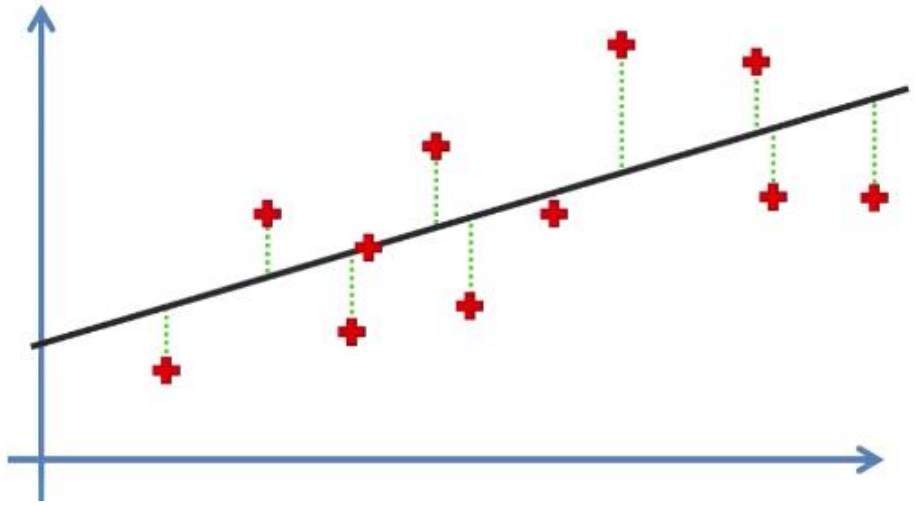
- **Ordinary Least Squares**
- **Gradient Descent**
- **Regularized Linear Regression**

Linear Regression Learning the Model

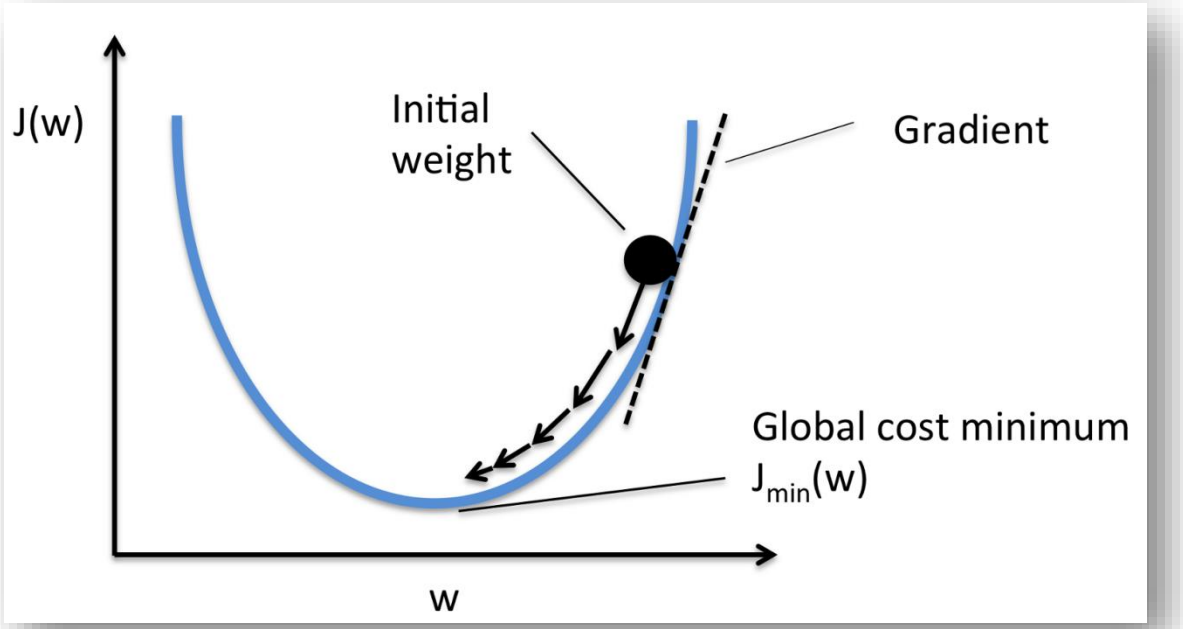
Simple Linear Regression

Ordinary Least Squares

- More than one input we can use Ordinary Least Squares
- Ordinary Least Squares procedure seeks to minimize the sum of the squared residuals



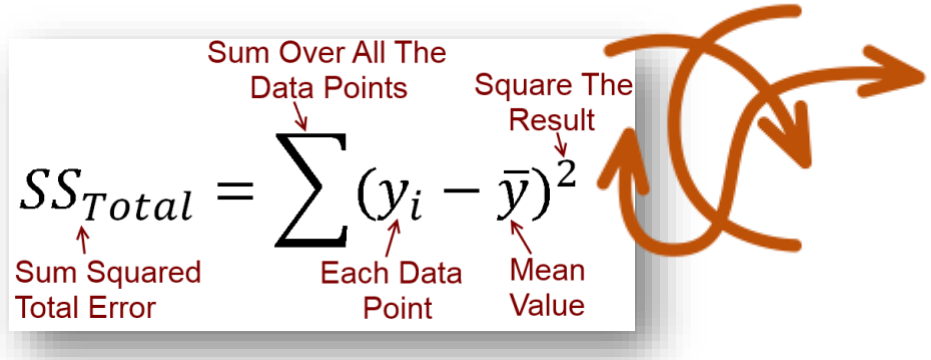
Gradient Descent



Learning rate (alpha) parameter that determines the size of the improvement step to take on each iteration

Regularized Linear Regression

Extensions of linear model called regularization methods.



Minimize error and also reduces complexity of the model

Types	Methodology
Lasso Regression: L1 regularization	OLS is modified to minimize the absolute sum of the coefficients
Ridge Regression: L2 regularization	**OLS is modified to minimize the squared absolute sum of the coefficients**

Making Predictions with Linear Regression

Predicting weight (y) from height (x)
Our linear regression model representation for this problem would be:

X variable is " Height "
Y variable is " Weight "

y = B0 + B1 X1
weight = B0 + B1 height

weight = 0.1 + 0.05 182
weight = 91.1