Introduction to Data Science CS61 June 12 - July 12, 2018

Dr. Ash Pahwa

Lesson 5: Regression

Lesson 5.1: Regression 2 Variables

Outline

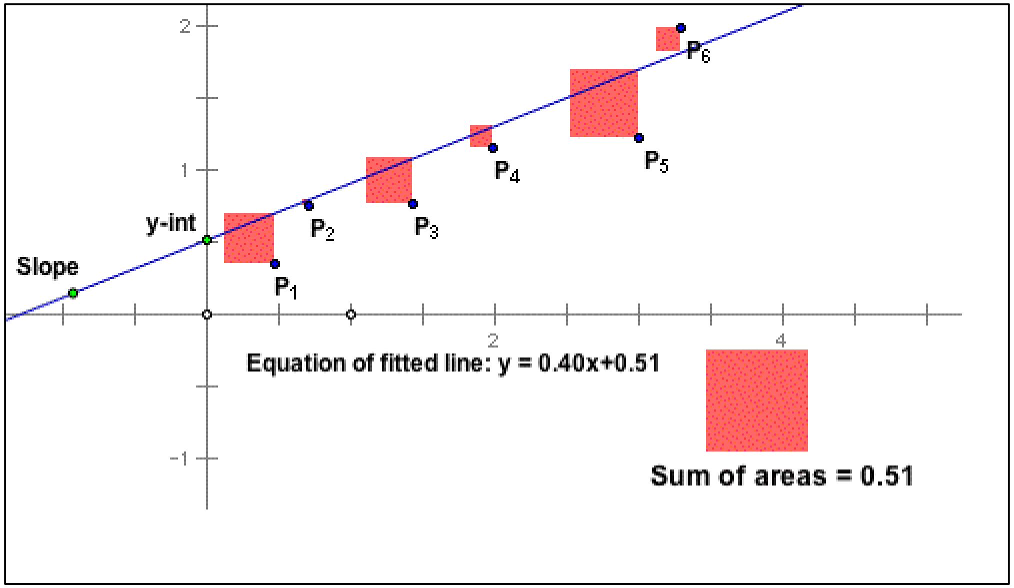
- Finding the Least Square Line
- Approach#1: Closed Form Solution: Calculus
  - Statistics: Correlation + Standard Deviation
  - Matrix Solution
- Approach#2: Gradient Decent Algorithm
- Regression
  - R
  - Excel (Data Analysis / Regression)

Regression
Finding the Least Square Line

Regression Strategy
Ordinary Least Squares (OLS)

The least-squares regression line of y and x is the line that makes

- the sum of the squares of the vertical distances
- of the data points from the line
- <u>**As small as possible**</u>



Solution for Regression Line

- Approach#1
  - Closed Form Solution
    - Compute Gradient
      - Vector of Partial Derivatives
      - Set gradient to zero
      - Compute slope and intercept
    - Matrix Approach
- Approach#2
  - Gradient Decent Algorithm
  - Gradually change slope and intercept till we reach at the optimum solution
  - Can be computed by Excel Solver

Approach#1

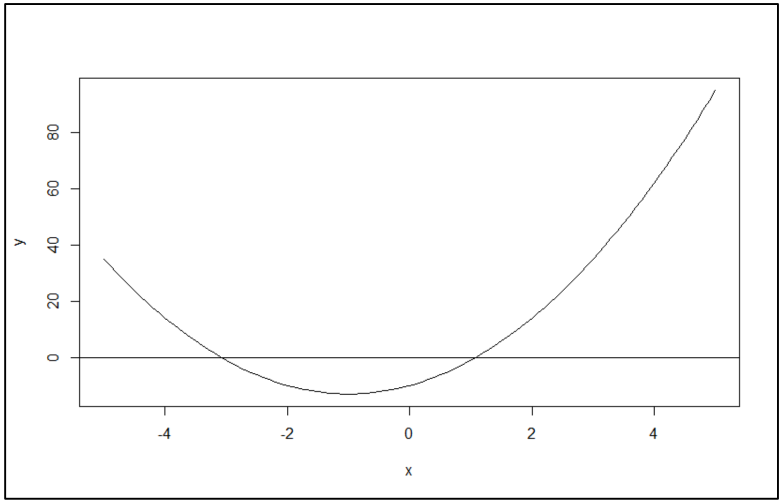
Closed Form Solution

Closed Form Solution Definition

- In mathematics, a closed-form expression is a mathematical expression that can be evaluated in a finite number of operations.
- It may contain constants, variables, certain "well-known" operations (e.g., + − × ÷), and functions (e.g., nth root, exponent, logarithm, trigonometric functions, and inverse hyperbolic functions), but usually no limit.
- The set of operations and functions admitted in a closed-form expression may vary with author and context.

Closed Form Solution

- Example
  - Quadratic Equation: $ax^2 + bx + c = 0$
  - Closed Form Solution
    - $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
    - $3x^2 + 6x - 10 = 0$
    - $x = 1.08, x = -3.08$
  - Iterative Solution
    - $x = 0, y = -10$
    - $x = 1, y = -1$
    - $x = 2, y = 14$
    - Solution must be between 1 and 2
    - Increase ‘x’ value by 0.1 and keep on checking if the y value becomes zero.

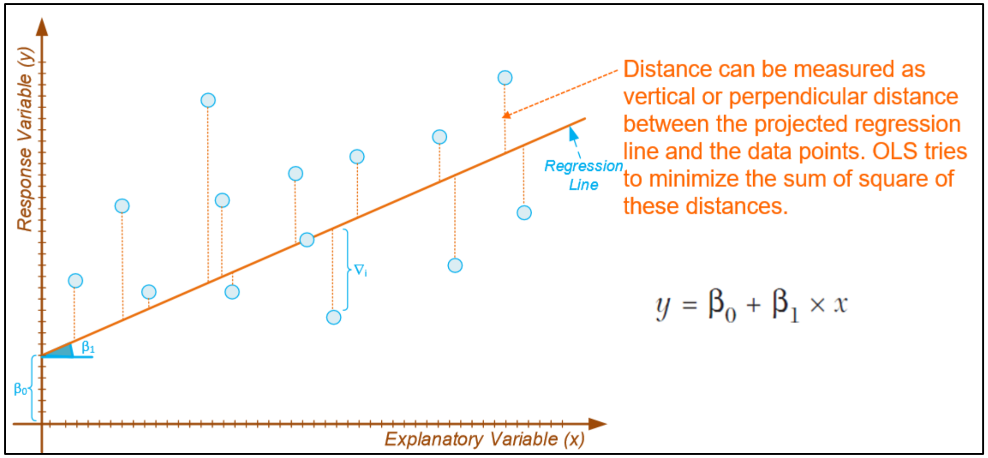


Solution for Regression Line Closed Form Solution

- Approach#1
    - Closed Form Solution
        - Compute Gradient
            - Vector of Partial Derivatives of RSS (Residual Sum of Squares) w.r.t. predictor variables
            - Set gradient to zero
            - Compute slope and intercept
        - Matrix Approach

Computing the Regression Line Compute: Intercept and Slope

- Residual = Observed value – Computed Value
- Suppose regression equation is
  - $y = mx + b$
  - y *is the explanatory variable*
  - x *is the pedictor variable*
  - m *is the slope of the line*
  - b *is the intercept*
- $Residual = \; y_i - (mx_i + b)$
- $Residual^2 = \; (y_i - (mx_i + b))^2$
- $Residuals\ Sum\ of\ Squares = (RSS) = \sum_{i=1}^{N}(y_i - (mx_i + b))^2$



Partial Derivatives of the RSS w.r.t. Intercept and Slope

- $Residuals\ Sum\ of\ Squares = (RSS) = \sum_{i=1}^{N}(y_i - (mx_i + b))^2$
- To find the minimum point of this function,
  - we will take the partial derivative of RSS with respect to 'm' and 'b' and set that to zero.

- $RSS = \sum_{i=1}^{N}(y_i - (mx_i + b))^2$
- $\frac{\partial RSS(m,b)}{\partial b} = \sum_{i=1}^{N} \frac{\partial}{\partial b}(y_i - (mx_i + b))^2$
- $\frac{\partial RSS(m,b)}{\partial b} = -2\sum_{i=1}^{N}(y_i - (mx_i + b))$

- $RSS = \sum_{i=1}^{N}(y_i - (mx_i + b))^2$
- $\frac{\partial RSS(m,b)}{\partial m} = \sum_{i=1}^{N} \frac{\partial}{\partial m}(y_i - (mx_i + b))^2$
- $\frac{\partial RSS(m,b)}{\partial m} = -2\sum_{i=1}^{N}(y_i - (mx_i + b))x_i$

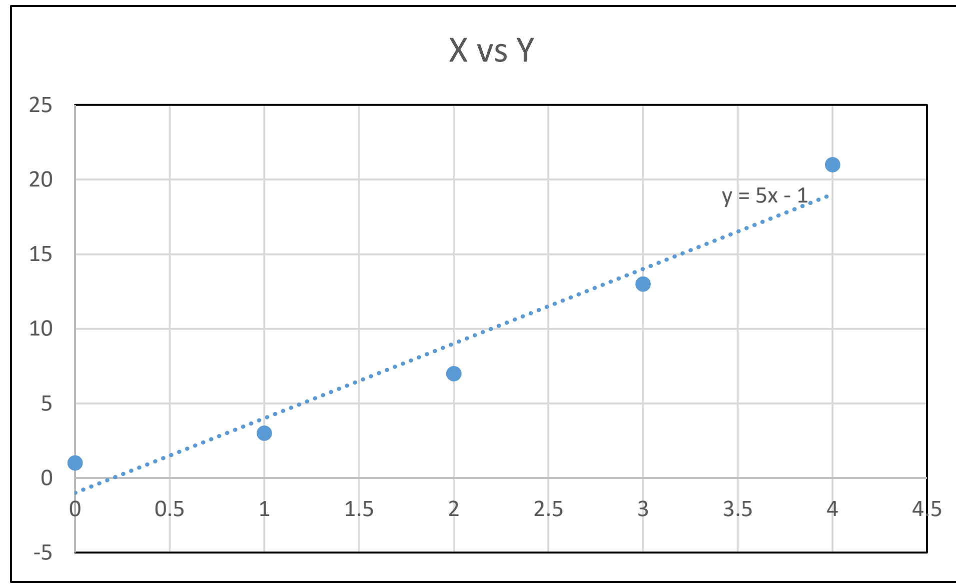
$$\nabla RSS(b,m) = \begin{vmatrix} \frac{\partial RSS(m,b)}{\partial b} \\ \frac{\partial RSS(m,b)}{\partial m} \end{vmatrix} = \begin{vmatrix} -2\sum_{i=1}^{N}(y_i - (mx_i + b)) \\ -2\sum_{i=1}^{N}(y_i - (mx_i + b))x_i \end{vmatrix} = 0$$

Gradient
Vector of Partial Derivatives

- To Compute 'm' and 'b'
  - SET GRADIENT = 0
- $\nabla\, RSS(b,m) = \begin{vmatrix} \frac{\partial RSS(m,b)}{\partial b} \\ \frac{\partial RSS(m,b)}{\partial m} \end{vmatrix} = \begin{vmatrix} -2\sum_{i=1}^{N}(y_i - (mx_i + b)) \\ -2\sum_{i=1}^{N}(y_i - (mx_i + b))x_i \end{vmatrix} = 0$
- ----------------------------------
- Top term
- $b = \left(\frac{\sum y_i}{N} - m\frac{\sum x_i}{N}\right) = \mu_y - m\mu_x$
- -------------------------------------
- Bottom term
- $m = \frac{\sum y_i x_i - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{\sum x_i \sum x_i}{N}} = r\frac{\sigma_y}{\sigma_x} = Correlation\,\frac{Std\ Dev\ of\ y}{Std\ Dev\ of\ x}$
- --

Example
Sample Data

X	Y
0	1
1	3
2	7
3	13
4	21



Basic Computations

$$m = \frac{\sum y_i x_i - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{\sum x_i \sum x_i}{N}}$$

$$b = \left(\frac{\sum y_i}{N} - m\frac{\sum x_i}{N}\right)$$

	A	B	C	D	E	F	G
1							
2							
3		X	Y		X*Y		X^2
4		0	1		0		0
5		1	3		3		1
6		2	7		14		4
7		3	13		39		9
8		4	21		84		16
9							
10	SUM	10	45		140		30
11	AVERAGE	2	9		28		6
12	StdDev	1.58113883	8.124038405				
13	Correlation	0.97312368					
14							

Method#1
Compute Slope

$$m = \frac{\sum y_i x_i - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{\sum x_i \sum x_i}{N}}$$

$$b = \left(\frac{\sum y_i}{N} - m\frac{\sum x_i}{N}\right)$$

Regression Equation
$y = 5x - 1$

	A	B	C	D	E	F	G
1							
2							
3		X	Y		X*Y		X^2
4		0	1		0		0
5		1	3		3		1
6		2	7		14		4
7		3	13		39		9
8		4	21		84		16
9							
10	SUM	10	45		140		30
11	AVERAGE	2	9		28		6
12	StdDev	1.58113883	8.124038405				
13	Correlation	0.97312368					
14							

C16 =E10-(B10*C10/5)

	A	B	C	D	E	F	G
14							
15	Closed Form	Slope : Using SUM					
16		Numerator	50		(Sum of X*Y) - (1/N)*((Sum of X) * (Sum of Y))		
17		Denominator	10		(Sum of X^2) - (1/N)*((Sum of X * Sum of X))		
18		Slope	5				
19							
20		Intercept	-1		(Mean of Y) - slope * (Mean of X)		
21							

$$m = \frac{\sum y_i x_i - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{\sum x_i \sum x_i}{N}}$$

$$b = \left(\frac{\sum y_i}{N} - m \frac{\sum x_i}{N} \right)$$

Method#2
Compute Slope

	A	B	C	D	E	F	G
1							
2							
3		X	Y		X*Y		X^2
4		0	1		0		0
5		1	3		3		1
6		2	7		14		4
7		3	13		39		9
8		4	21		84		16
9							
10	SUM	10	45		140		30
11	AVERAGE	2	9		28		6
12	StdDev	1.58113883	8.124038405				
13	Correlation	0.97312368					
14							

- $m = \frac{\sum y_i x_i - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{\sum x_i \sum x_i}{N}}$
- Divide both numerator and denominator by N
- $\frac{\frac{\sum y_i x_i}{N} - \frac{\sum y_i \sum x_i}{N.N}}{\frac{\sum x_i^2}{N} - \frac{\sum x_i \sum x_i}{N.N}} = \frac{Mean\ of\ X*Y - (Mean\ of\ X)*(Mean\ of\ Y)}{Mean\ of\ x^2 - (Mean\ of\ X)*(Mean\ of\ X)}$

Regression Equation
$y = 5x - 1$

	A	B	C	D	E	F	G
22							
23		Slope : Using Average					
24		Numerator	10		(Mean of X * Y) - (Mean of X)*(Mean of Y)		
25		Denominator	2		(Mean of X^2) - (Mean of X)*(Mean of X)		
26		Slope	5				
27							
28							

$$m = \frac{\sum y_i x_i - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{\sum x_i \sum x_i}{N}}$$

$$b = \left(\frac{\sum y_i}{N} - m \frac{\sum x_i}{N} \right)$$

Method#3 Compute Slope

- $m = \frac{\sum y_i x_i - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{\sum x_i \sum x_i}{N}}$
- $= r \frac{\sigma_y}{\sigma_x} = Correlation \frac{Std\ Dev\ of\ y}{Std\ Dev\ of\ x}$

	A	B	C	D	E	F	G
1							
2							
3		X	Y		X*Y		X^2
4		0	1		0		0
5		1	3		3		1
6		2	7		14		4
7		3	13		39		9
8		4	21		84		16
9							
10	SUM	10	45		140		30
11	AVERAGE	2	9		28		6
12	StdDev	1.58113883	8.124038405				
13	Correlation	0.97312368					
14							

Regression Equation
y = 5x − 1

C29 =B13*C12/B12

	A	B	C	D	E	F	G
27							
28							
29	Statistics	Slope	5		(Correlation * StdDev of Y) / StdDev of X		
30		Intercept	-1		(Mean of Y) - slope * (Mean of X)		
31							

Closed Form Solution

Linear Regression Using Matrices

2 Variable Regression

X	Y
x_1	y_1
x_2	y_2
x_3	y_3
…	…
x_n	y_n

- Given values of X and Y
- Suppose the least square regression line is given by the following function
  - $f(x) = b + mx$
- Residual Sum of Squares = Error term
  - $(y_1 - f(x_1))^2 + (y_2 - f(x_2))^2 + \cdots + (y_n - f(x_n))^2$

Systems of Linear Equations

- Systems of Linear Equations
  - $y_1 = (b + mx_1) + e_1$
  - $y_2 = (b + mx_2) + e_2$
  - …
  - $y_n = (b + mx_n) + e_n$
- $Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} \quad A = \begin{bmatrix} b \\ m \end{bmatrix} \quad E = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix}$
- Matrix Equation: $Y = XA + E$
- $E = Y - XA$
- Solve this matrix equation for
  - $Matrix\ A\ where\ square\ of\ residuals\ is\ least$

Objective Function: Minimize Sum of Square Errors

- $Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$ $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix}$ $A = \begin{bmatrix} b \\ m \end{bmatrix}$ $E = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix}$
- Matrix Equation: $Y = XA + E$
- $E = Y - XA$

- $RSS\ (Residual\ Sum\ of\ Squares) = e_1^2 + e_2^2 + \cdots + e_n^2$
- $RSS = \begin{bmatrix} e_1 & e_2 & \dots & e_n \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix} = e_1^2 + e_2^2 + \cdots + e_n^2$
- $RSS = \boldsymbol{E^T E}$
- We have already seen that $\boldsymbol{E = Y - XA}$
- $RSS = (\boldsymbol{Y - XA})^{\boldsymbol{T}}(\boldsymbol{Y - XA})$

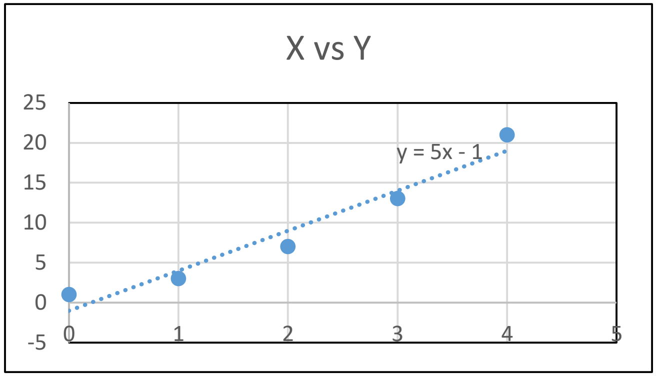
Compute Gradient of RSS
Set Gradient = 0

- $RSS = (\boldsymbol{Y} - \boldsymbol{XA})^{\boldsymbol{T}}(\boldsymbol{Y} - \boldsymbol{XA})$
- $RSS = \boldsymbol{Y^TY} - \boldsymbol{2Y^TXA} + \boldsymbol{AX^TXA}$
- $Set\ \frac{\partial RSS}{\partial A} = 0,$
  - $to\ compute\ the\ value\ of\ A\ for\ minimum\ RSS$
- $\nabla RSS = \nabla[(\boldsymbol{Y} - \boldsymbol{XA})^{\boldsymbol{T}}(\boldsymbol{Y} - \boldsymbol{XA})]$
- $\nabla RSS = -2\boldsymbol{X}^T(\boldsymbol{Y} - \boldsymbol{XA}) = 0$
- $-2\boldsymbol{X}^T\boldsymbol{Y} + 2\boldsymbol{X}^T\boldsymbol{XA} = 0$
- $\boldsymbol{X}^T\boldsymbol{XA} = \boldsymbol{X}^T\boldsymbol{Y}$
- Multiply both sides with $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$
- $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{XA} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$
- $Since: (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{I} \quad \& \quad \boldsymbol{IA} = \boldsymbol{A}$
- $\boldsymbol{A} = (\boldsymbol{X^TX})^{-1}\boldsymbol{X^TY}$

- By Analogy 1D case
- $\frac{d}{dA}(Y - XA)(Y - XA)$
- $= \frac{d}{dA}(Y - XA)^2$
- $= 2(Y - XA)(-X)$
- $= -2X(Y - XA)$

Example-1

X	Y
0	1
1	3
2	7
3	13
4	21



- $Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$ $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix}$ $A = \begin{bmatrix} b \\ m \end{bmatrix}$ $E = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix}$
- $Y = \begin{bmatrix} 1 \\ 3 \\ 7 \\ 13 \\ 21 \end{bmatrix}$ $X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$
- $Y = XA + E$
- Solution for Least RSS = $A = (X^T X)^{-1} X^T Y$

Example-1

- $Y = \begin{bmatrix} 1 \\ 3 \\ 7 \\ 13 \\ 21 \end{bmatrix}$ $X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$
- $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

- Solution for Least RSS = $A = (X^T X)^{-1} X^T Y$
- $X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 30 \end{bmatrix}$
- $(X^T X)^{-1} = \frac{1}{50}\begin{bmatrix} 30 & -10 \\ -10 & 5 \end{bmatrix} = \begin{bmatrix} 0.6 & -0.2 \\ -0.2 & 0.1 \end{bmatrix}$
- $X^T Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 7 \\ 13 \\ 21 \end{bmatrix} = \begin{bmatrix} 45 \\ 140 \end{bmatrix}$
- $A = (X^T X)^{-1} X^T Y = \begin{bmatrix} 0.6 & -0.2 \\ -0.2 & 0.1 \end{bmatrix} \begin{bmatrix} 45 \\ 140 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$

Example-1: Final Answer

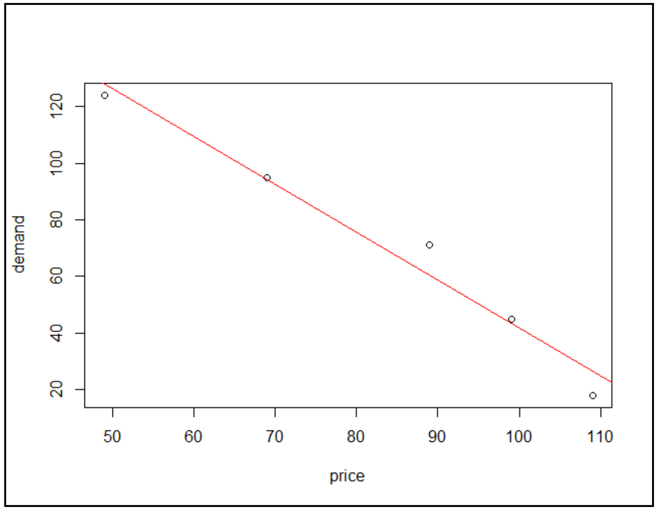
- $A = (X^T X)^{-1} X^T Y = \begin{bmatrix} 0.6 & -0.2 \\ -0.2 & 0.1 \end{bmatrix} \begin{bmatrix} 45 \\ 140 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$
- $f(x) = -1 + 5x$

Regression Equation
$y = 5x - 1$

Example-2

Price	Demand
$49	124
$69	95
$89	71
$99	45
$109	18

- $Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} \quad A = \begin{bmatrix} b \\ m \end{bmatrix} \quad E = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix}$
- $Y = \begin{bmatrix} 124 \\ 95 \\ 71 \\ 45 \\ 18 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 49 \\ 1 & 69 \\ 1 & 89 \\ 1 & 99 \\ 1 & 109 \end{bmatrix}$
- $Y = XA + E$
- Solution for Least RSS = $A = (X^T X)^{-1} X^T Y$



Example-2

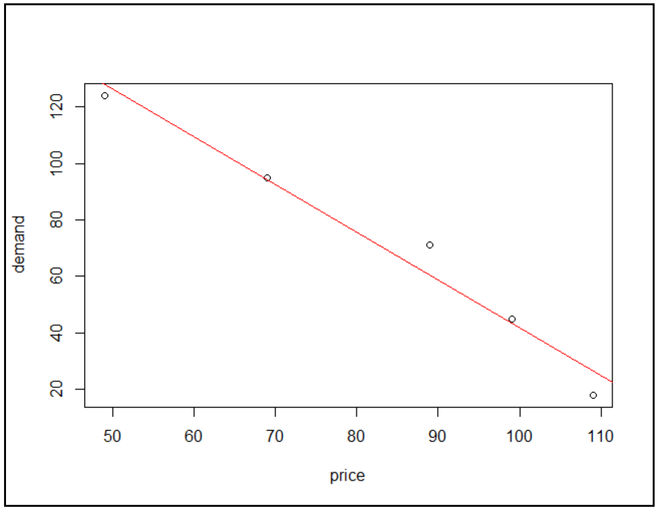
- $Y = \begin{bmatrix} 124 \\ 95 \\ 71 \\ 45 \\ 18 \end{bmatrix}$ $X = \begin{bmatrix} 1 & 49 \\ 1 & 69 \\ 1 & 89 \\ 1 & 99 \\ 1 & 109 \end{bmatrix}$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$
- $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

- Solution for Least RSS = $A = (X^TX)^{-1}X^TY$
- $X^TX = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 49 & 69 & 89 & 99 & 109 \end{bmatrix}\begin{bmatrix} 1 & 49 \\ 1 & 69 \\ 1 & 89 \\ 1 & 99 \\ 1 & 109 \end{bmatrix} = \begin{bmatrix} 5 & 415 \\ 415 & 36765 \end{bmatrix}$
- $(X^TX)^{-1} = \frac{1}{11600}\begin{bmatrix} 36765 & -415 \\ -415 & 5 \end{bmatrix}$
- $X^TY = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 49 & 69 & 89 & 99 & 109 \end{bmatrix}\begin{bmatrix} 124 \\ 95 \\ 71 \\ 45 \\ 18 \end{bmatrix} = \begin{bmatrix} 353 \\ 25367 \end{bmatrix}$
- $A = (X^TX)^{-1}X^TY = \frac{1}{11600}\begin{bmatrix} 36765 & -415 \\ -415 & 5 \end{bmatrix}\begin{bmatrix} 353 \\ 25367 \end{bmatrix} = \begin{bmatrix} 211.2707 \\ -1.6948 \end{bmatrix}$

Example-2: Final Answer

- $A = (X^T X)^{-1} X^T Y = \frac{1}{11600}\begin{bmatrix} 36765 & -415 \\ -415 & 5 \end{bmatrix}\begin{bmatrix} 353 \\ 25367 \end{bmatrix} = \begin{bmatrix} 211.2707 \\ -1.6948 \end{bmatrix}$
- $f(x) = 211.27 - 1.6948x$



Regression Equation
$y = -1.6948x + 211.2707$

Approach#2

Gradient Decent Algorithm

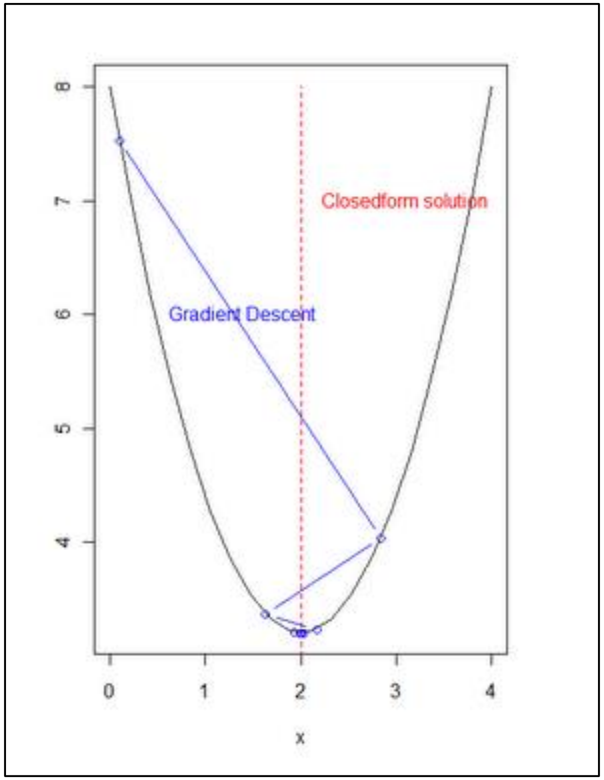
Solution for Regression Line Gradient Decent

- Approach#2
  - Gradient Decent Algorithm
  - Gradually change slope and intercept till we reach at the optimum solution (minimum value of RSS)
  - Can be computed by Excel Solver

Gradient Decent Algorithm
2 Variables

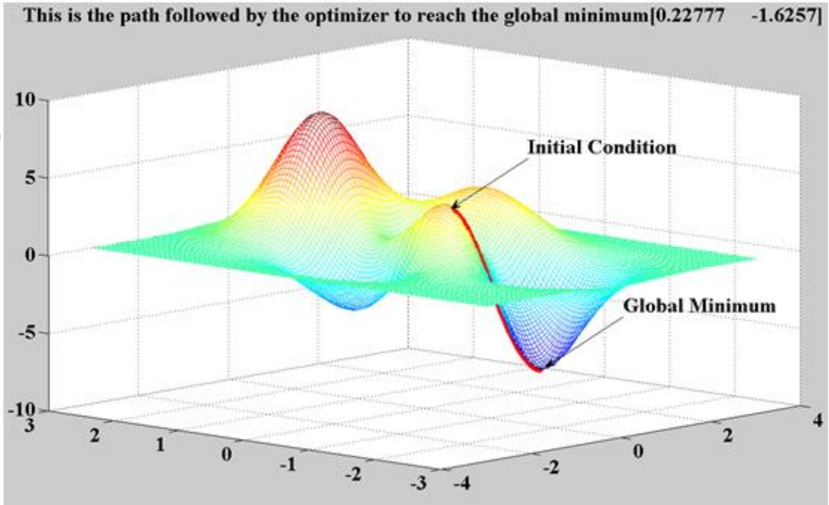
$$\nabla RSS(b,m) = \begin{vmatrix} \dfrac{\partial RSS(m,b)}{\partial b} \\ \dfrac{\partial RSS(m,b)}{\partial m} \end{vmatrix} = \begin{vmatrix} -2\sum_{i=1}^{N}(y_i-(mx_i+b)) \\ -2\sum_{i=1}^{N}(y_i-(mx_i+b))x_i \end{vmatrix} = 0$$

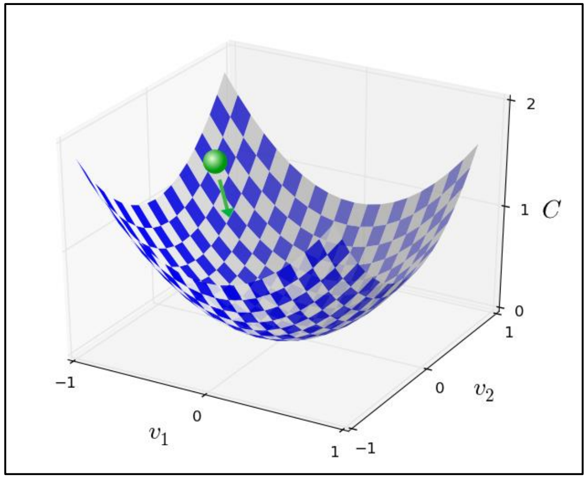
- $g(x) = mx + b$
- Algorithm
- ------------------------------
- $While\ NOT\ Converged$
  - $m^{t+1} \leftarrow m^t - \eta \frac{\partial g}{\partial m} \|_{m^t}$
- ------------------------------
- Convergence Test
- $\left|\frac{\partial g(w)}{\partial m}\right| < \varepsilon\ (threshold\ to\ be\ set)$



Gradient Decent Algorithm 3 Variables

- Algorithm
- -------------------------------
- $While\ NOT\ Converged$
  - $w^{t+1} \leftarrow w^t - \eta \nabla g(w^t)$
- -------------------------------
- Convergence Test
- $|\nabla g(w)| < \varepsilon\ (threshold\ to\ be\ set)$





Regression in R and Excel

Example-2
Solution in R – Using Matrix

Regression Equation
$$y = -1.6948x + 211.2707$$

```
> price = c(49, 69, 89, 99, 109)
> demand = c(124, 95, 71, 45, 18)
> plot(price,demand)
>
> ##################################
> Oneprice = c(1,1,1,1,1,price)
> Xprice = matrix(Oneprice,nrow=5)
> Xprice
     [,1] [,2]
[1,]    1   49
[2,]    1   69
[3,]    1   89
[4,]    1   99
[5,]    1  109
>
> #################################
> demand = c(124, 95, 71, 45, 18)
> Ydemand = matrix(demand,nrow=5)
> Ydemand
     [,1]
[1,]  124
[2,]   95
[3,]   71
[4,]   45
[5,]   18
```

```
> ###########################
> z1 = t(Xprice)%*%Xprice
> z1
     [,1]  [,2]
[1,]    5   415
[2,]  415 36765
> # comput the inverse of z
> det(z1)
[1] 11600
> invz1 = solve(z1)
>
> ############################
> z2 = t(Xprice)%*%Ydemand
> z2
      [,1]
[1,]   353
[2,] 25367
> ############################
> ans = invz1 %*% z2
> ans
           [,1]
[1,] 211.270690
[2,]  -1.694828
>
```

Example-2
Solution in R – Using 'lm' command

```
> summary(lm(demand~price))

Call:
lm(formula = demand ~ price)

Residuals:
      1       2       3       4       5 
-4.2241  0.6724 10.5690  1.5172 -8.5345 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 211.2707    14.7215  14.351 0.000733 ***
price        -1.6948     0.1717  -9.872 0.002210 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.269 on 3 degrees of freedom
Multiple R-squared:  0.9701,	Adjusted R-squared:  0.9602 
F-statistic: 97.46 on 1 and 3 DF,  p-value: 0.00221
```

Regression Equation

$y = -1.6948x + 211.2707$

Summary

- Finding the Least Square Line
- Approach#1: Closed Form Solution: Calculus
  - Statistics: Correlation + Standard Deviation
  - Matrix Solution
- Approach#2: Gradient Decent Algorithm
- Regression
  - R
  - Excel (Data Analysis / Regression)