Linear Regression with Linear Algebra

Manual Computation we did

$$\hat{y} = slope * x + intercept$$

$$slope = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2}$$

$$intercept = \bar{y} - slope \cdot \bar{x}$$

	x	y	xy	x^2
	Year	Population		
	1980	2.1		
	1985	2.9		
	1990	3.2		
	1995	4.1		
	2000	4.9		
Sum				
Average				
Count (n) =				

Slope	
Intercept	

Manual Computation we did

$$\hat{y} = slope * x + intercept$$

$$slope = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2}$$

$$intercept = \bar{y} - slope \cdot \bar{x}$$

	x	y	xy	x^2
	Year	Population		
	1980	2.1	4158	3920400
	1985	2.9	5756.5	3940225
	1990	3.2	6368	3960100
	1995	4.1	8179.5	3980025
	2000	4.9	9800	4000000
Sum	9950	17.2	34262	19800750
Average	1990	3.44		
Count (n) =	5			

Slope	0.136
Intercept	-267.2

Linear Regression with Linear Algebra

$$y = m \cdot x + c$$

$$\begin{aligned} m \cdot x_1 + c &= y_1 \\ m \cdot x_2 + c &= y_2 \\ &\vdots \\ m \cdot x_n + c &= y_n \end{aligned} \qquad \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \cdot \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{x} & \mathbf{1} \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} m \\ c \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Linear Regression with Linear Algebra

$$y = m \cdot x + c$$

$$\begin{aligned} m \cdot x_1 + c &= y_1 \\ m \cdot x_2 + c &= y_2 \\ &\vdots \\ m \cdot x_n + c &= y_n \end{aligned}$$

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \cdot \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \Rightarrow \quad \mathbf{A} \cdot \mathbf{b} = \mathbf{y}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{x} & \mathbf{1} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} m \\ c \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Linear Regression with Linear Algebra

$$y = m \cdot x + c$$

$$\begin{aligned} m \cdot x_1 + c &= y_1 \\ m \cdot x_2 + c &= y_2 \\ &\vdots \\ m \cdot x_n + c &= y_n \end{aligned}$$

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \cdot \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\mathbf{A} \cdot \mathbf{b} = \mathbf{y}$$

$$\left(\mathbf{A}^t \cdot \mathbf{A}\right) \cdot \mathbf{b} = \mathbf{A}^t \cdot \mathbf{y}$$

$$\left(\mathbf{A}^t \cdot \mathbf{A}\right)^{-1} \cdot \left(\mathbf{A}^t \cdot \mathbf{A}\right) \cdot \mathbf{b} = \left(\mathbf{A}^t \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^t \cdot \mathbf{y}$$

$$\boxed{\mathbf{b} = \left(\mathbf{A}^t \cdot \mathbf{A}\right)^{-1} \cdot \mathbf{A}^t \cdot \mathbf{y}}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{x} & \mathbf{1} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} m \\ c \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\mathbf{A}^t \cdot \mathbf{A} = \begin{bmatrix} \sum x^2 & \sum x \\ \sum x & n \end{bmatrix}$$

$$\mathbf{A}^t \cdot \mathbf{y} = \begin{bmatrix} \sum xy \\ \sum y \end{bmatrix}$$

Linear Regression with Linear Algebra

A = [x 1]		y
1980	1	2.1
1985	1	2.9
1990	1	3.2
1995	1	4.1
2000	1	4.9

A'				
1980	1985	1990	1995	2000
1	1	1	1	1
A'A				
Inverse of A'A				
Inv(A'A).A'				
b = Inv(A'A).A' * y				

Linear Regression with Linear Algebra

A = [x 1]		y
1980	1	2.1
1985	1	2.9
1990	1	3.2
1995	1	4.1
2000	1	4.9

A'

1980	1985	1990	1995	2000
1	1	1	1	1

A'A

19800750	9950
9950	5

Inverse of A'A

0.004	-7.96
-7.96	15840.6

Inv(A'A).A'

-0.04	-0.02	8.88E-16	0.02	0.04
79.8	40	0.2	-39.6	-79.4

b = Inv(A'A).A' * y

0.136
-267.2

THANK YOU!