Linear Regression

Closed Form Solution

Notation

- Let the given data is $D = \{(X_1, y_1), \ldots, (X_n, y_n)\}$.
- Let $X_i \in R^d$ and $y_i \in R$
- Further, $X_i = [x_{i1} \quad \cdots \quad x_{id}]^t$

We believe that

- $\theta_0 + \theta_1 x_{i1} + \cdots + \theta_d x_{id} = y_i$

We like to solve

- $$\begin{gathered}\theta_0 + \theta_1 x_{11} + \cdots + \theta_d x_{1d} = y_1 \\ \theta_0 + \theta_1 x_{21} + \cdots + \theta_d x_{2d} = y_2 \\ \vdots \\ \theta_0 + \theta_1 x_{n1} + \cdots + \theta_d x_{nd} = y_n\end{gathered}$$

In matrix form $Z\Theta = Y$

Solve this to find Θ

Z may not be invertible !!

$$Z\Theta = Y$$

- $$\begin{bmatrix} 1 & x_{11} & \cdots & x_{1d} \\ 1 & x_{21} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Closest Solution: by minimizing the squared error

- Obtained by minimizing the criterion
  $J(\Theta) = \|Z\Theta - Y\|^2 = (Z\Theta - Y)^t(Z\Theta - Y)$

$$J(\Theta) = \Theta^t Z^t Z\Theta - 2(Z\Theta)^t Y - Y^t Y$$

$$J(\Theta) = \Theta^t Z^t Z\Theta - 2(Z\Theta)^t Y - Y^t Y$$

- Equating gradient of J to zero. So

$$\nabla_\Theta(J) = 2(Z^t Z)\Theta - 2\mathrm{Z}^{\mathrm{t}}\mathrm{Y} = 0$$

We get $$\Theta = (Z^t Z)^{-1} Z^t Y$$

Example: 1D problem

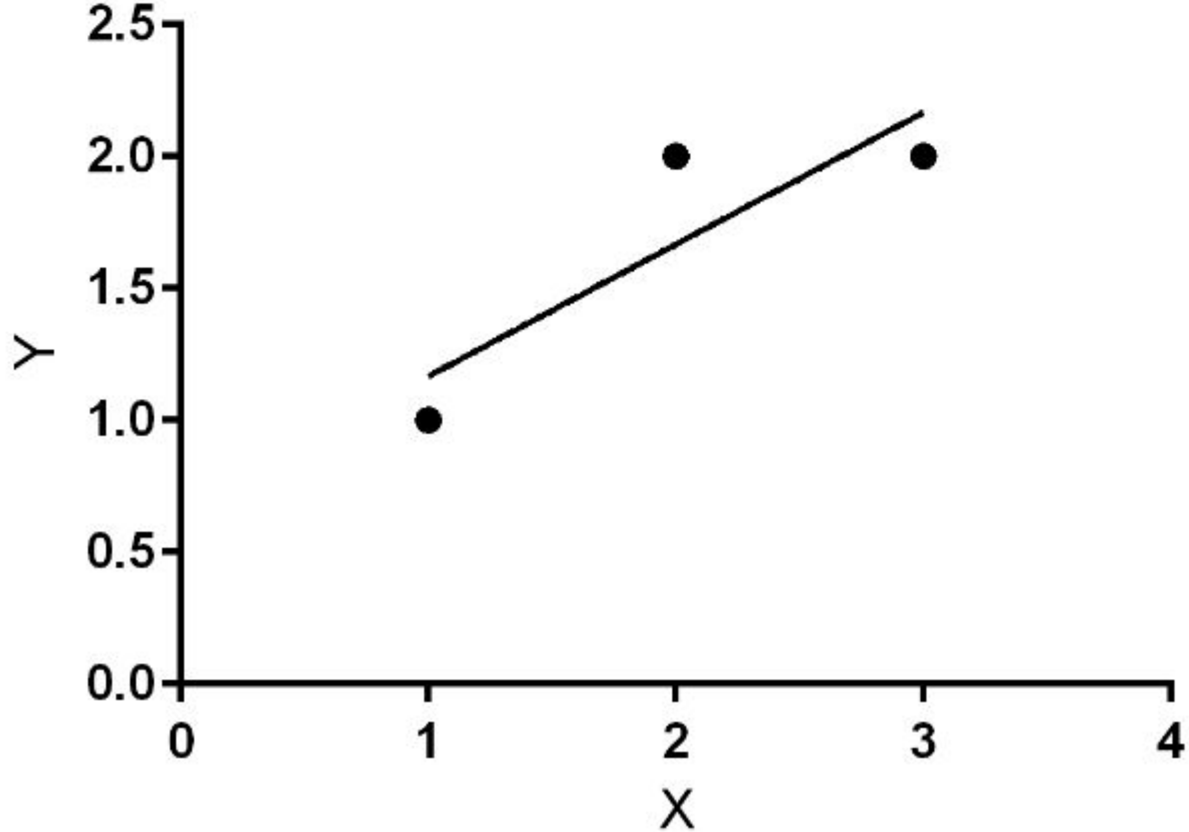
- $D = \{(1,1),(2,2),(3,2)\}$

- $Z = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \Theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}, Y = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

- $Z^t Z = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \quad (Z^t Z)^{-1} = \frac{1}{6}\begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix}$

- $\Theta = (Z^t Z)^{-1} Z^t Y$
- $\Theta = \frac{1}{6}\begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 11 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$

- So the line we fitted is $y = \frac{2}{3} + \frac{x}{2}$


2.5
2.0
1.5
1.0
0.5
0.0
Y
0
1
2
3
4
X