10 Sept 2021 Data Analytics

Last Time:

Almost finished our mathematics fundamentals

Math Quiz

Today:

Set Theory

Intro to Linear Regression our first algorithm to predict

Monday

Finish Linear Regression

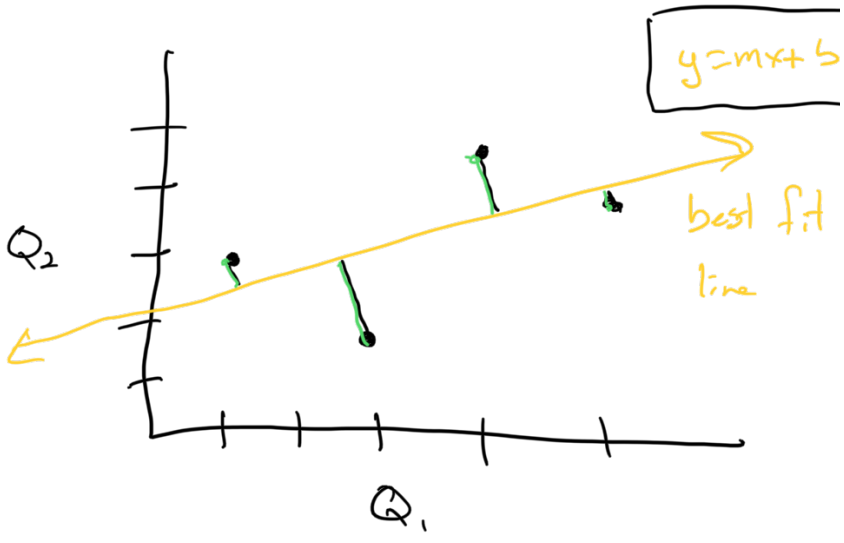
Review HW 1

Chap 2 Reading extended to 15 Sept

Chalkboard Example - Linear Regression with one variable

name	Q_1	Q_2
a	1	3
b	3	2
c	4	5
d	5	4

e 2 ?



$$\left.\begin{aligned} 3 &= m + b \\ 2 &= 3m + b \\ 5 &= 4m + b \end{aligned}\right\} \text{rearrange} \Rightarrow \begin{aligned} b + m &= 3 \\ b + 3m &= 2 \\ b + 4m &= 5 \end{aligned}$$

$y = mx + b$ $\qquad$ $b + 5m = 4$

$$\underset{X}{\begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}} \begin{bmatrix} b \\ m \end{bmatrix} = \underset{y}{\begin{bmatrix} 3 \\ 2 \\ 5 \\ 4 \end{bmatrix}}$$

Solve for $\begin{bmatrix} b \\ m \end{bmatrix}$

pseudo inverse

$$\begin{bmatrix} b \\ m \end{bmatrix} = \boxed{(X^T X)^{-1} X^T} y$$

2×1 $\qquad$ 2×4 $\qquad$ 4×1

$$\begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 2.2 \\ 0.4 \end{bmatrix}$$

my model

Find my prediction of Q_2 score given Q_1 is 2

$$Q = 0.4 Q_1 + 2.2 \Rightarrow 3$$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2.2 \\ 0.4 \end{bmatrix} = 1 \cdot 2.2 + 2 * 0.4$$

↑ y intercept term

model

$$\begin{bmatrix} 1 & 2 \\ 1 & 2.5 \\ & \end{bmatrix} \begin{bmatrix} 2.2 \\ 0.4 \end{bmatrix} = \begin{bmatrix} \square \\ \square \\ \end{bmatrix}$$