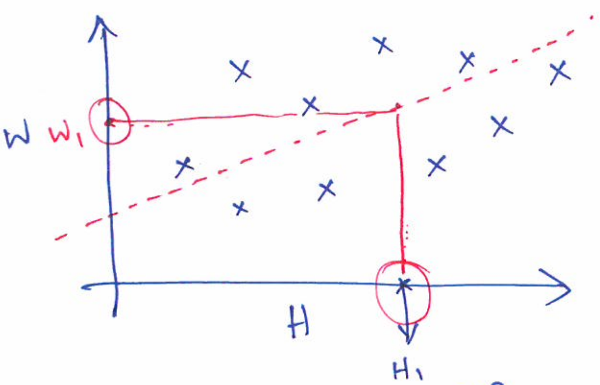
Linear Regression



$w_1 \to$?

Best fit line for the given data.

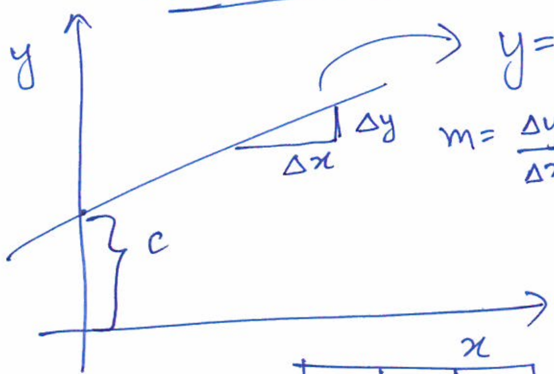
Sup. Learning

- Prediction (Regression) → real-values: 100.27, 257, etc.
- Classification → categorical values: A, B, C, ...

(given) (i/p) x y ? (output)

	Ht	Wt ?
1	—	—
2	—	—
⋮	⋮	⋮
100	—	—
	108	?

Best-fit line

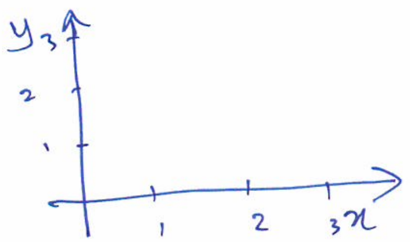


$y = \underbrace{c}_{\text{int.}} + x\,\underbrace{m}_{\text{slope}}$

$m = \dfrac{\Delta y}{\Delta x}$

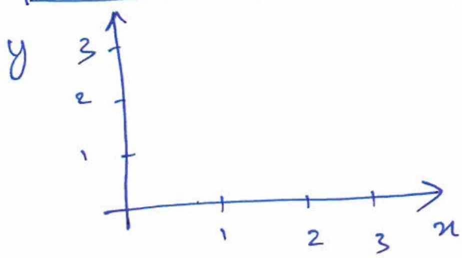
$\Rightarrow$ given training data, we need to estimate m, c for best-fit line. → ? least error y_i and $\hat{y}_i$

$\boxed{y = b_0 + b_1 x}$



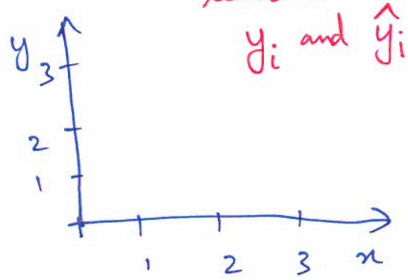
$b_0 = 1.5$

$b_1 = 0$



$b_0 = 0$

$b_1 = 0.5$



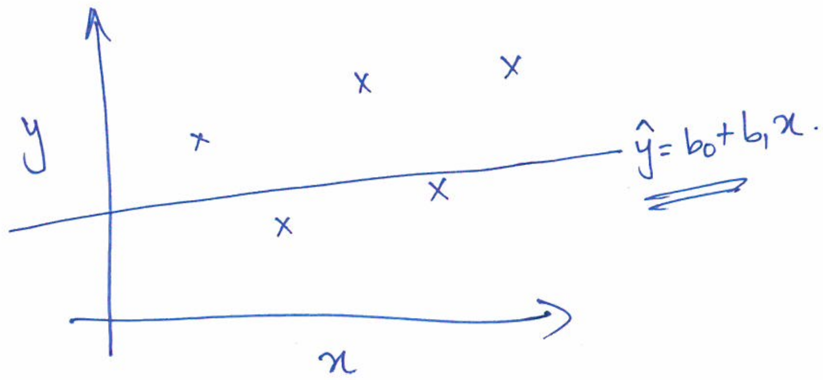
$b_0 = 1$

$b_1 = 0.5$

Sum of Squared Error (SSE) → $\hat{y}_i$ (predicted from line (b_0, b_1))

→ y_i (actual given data)

	W	H	
	x	y	$\hat{y}$
1	5	100	.
2	6	200	.
3	7	300	.
4	8	150	.
5	9	200	



Eg: $$\sum_{i=1}^{N=5} \left(y_i - \hat{y}_i\right)^2 \Rightarrow SSE$$

find b_0 & b_1 ? $$\Rightarrow \min_{b_0, b_1} \left[\sum_{i=1}^{N} \overbrace{\left(y_i - \hat{y}_i\right)^2}^{\text{error}} \right]$$

$\boxed{\text{Least Square Estimates}}$

$$SSE = \sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2$$

1. Partially differentiate SSE w. r. to b_0 and equate to zero.

2. " " " w. r. to b_1 and equate to zero.

(1.) Partially diff. SSE w.r.to b_0

$$SSE = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^{N} \big(\underbrace{y_i - (b_0 + b_1 x_i)}_{(a-b)^2}\big)^2$$

$$= \sum_{i=1}^{N} \Big[y_i^2 + \underbrace{(b_0 + b_1 x_i)^2}_{(a+b)^2} - 2 y_i (b_0 + b_1 x_i) \Big]$$

$$SSE = \sum_{i=1}^{N} \Big[y_i^2 + b_0^2 + b_1^2 x_i^2 + 2 b_0 b_1 x_i - 2 y_i b_0 - 2 y_i b_1 x_i \Big]$$

$$\frac{\partial SSE}{\partial b_0} = \sum_{i=1}^{N} \Big[0 + 2b_0 + 0 + 2 b_1 x_i - 2 y_i - 0 \Big]$$

$$0 = \sum_{i=1}^{N} \Big[2b_0 + 2b_1 x_i - 2 y_i \Big]$$

$$\times \frac{1}{n}) \quad 0 = \frac{2 \cancel{n}\, b_0}{\cancel{n}} + \frac{2 \sum_{i=1}^{n} b_1 x_i}{n} - 2 \frac{\sum_{i=1}^{n} y_i}{n}$$

$$0 = 2 b_0 + 2 b_1 \frac{\sum_{i=1}^{n} x_i}{n} - 2 \frac{\sum_{i=1}^{n} y_i}{n}$$

$$0 = \cancel{2} \overset{?}{b_0} + \cancel{2} b_1 \bar{x} - \cancel{2} \bar{y}$$

$$\bar{x} = \sum_{i=1}^{n} x_i$$

$$\bar{y} = \sum_{i=1}^{n} y_i$$

$$\boxed{b_0 = (\bar{y} - b_1 \bar{x})}$$