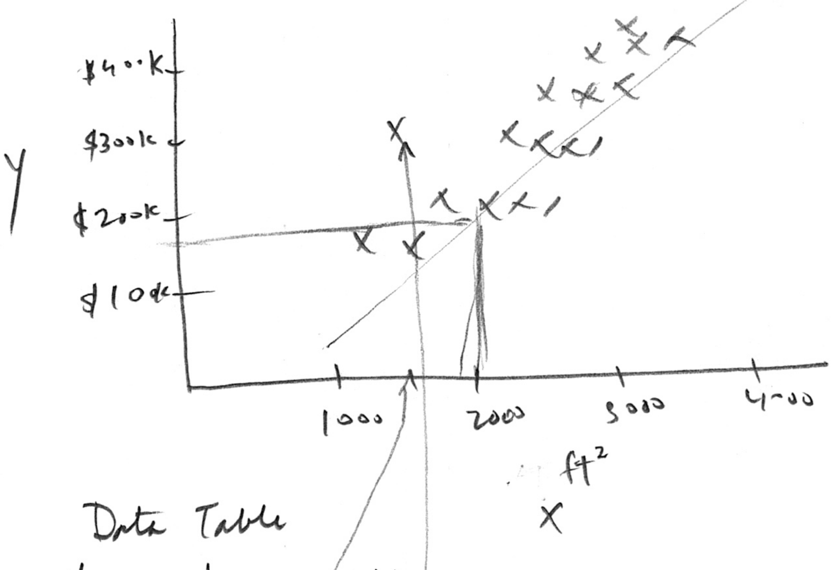
Linear Regression



- Supervised learning model
  We have the answers
- Regression: predict #s (of ∞ possibilities)
- not classification (small # of outputs)

Data Table

	X size (ft²)	Y price ($1000)
(1)	2104	400
(2)	1416	232
(3)	1534	315
⋮	...	...
(47)	3210	870

$x^{(1)} = 2104 \qquad y^{(1)} = 400$

$(x^{(1)}, y^{(1)}) = (2104, 400)$

$x^{(2)}$ is not x^2 ← squared

Terms:

x = "input" variable

y = "output" variable
"target" variable

m = # of training examples

eg $m = 47$

$(x^{(i)}, y^{(i)})$ = i^{th} single training example

Task:

training set

$\downarrow$

learning algorithm

$\downarrow$

$x \rightarrow \boxed{f} \rightarrow \hat{y}$ ← "y-hat"

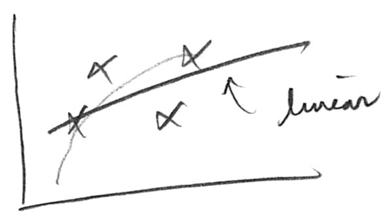
"feature" "model" "prediction" ← this is not y (the answer) [target → true]
it is the estimate of y

size $\rightarrow \boxed{f} \rightarrow$ price

How do we represent f?

$$f_{w,b}(x) = wx + b$$

$$f(x)$$



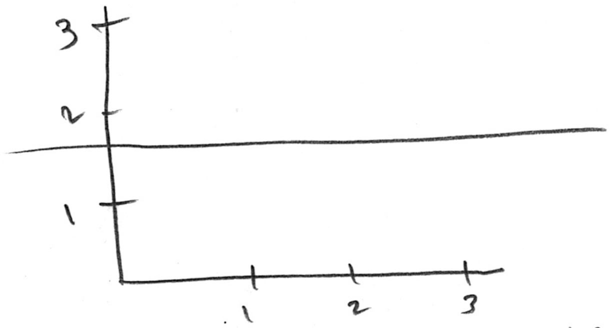
⇒ Linear regression w/ $\underbrace{\text{one variable}}$ ← single feature

aka univariate linear regression

w, b : parameters
coefficients
weights

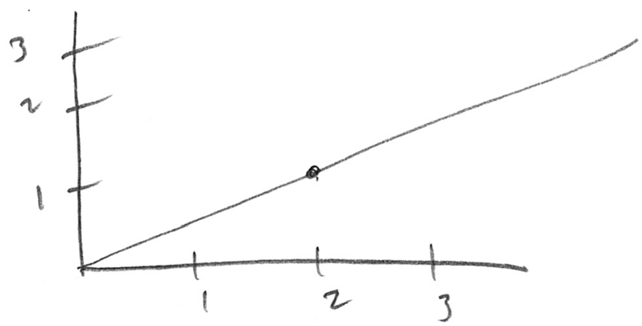
Examples:

$$f_{w,b}(x) = wx + b$$

- 

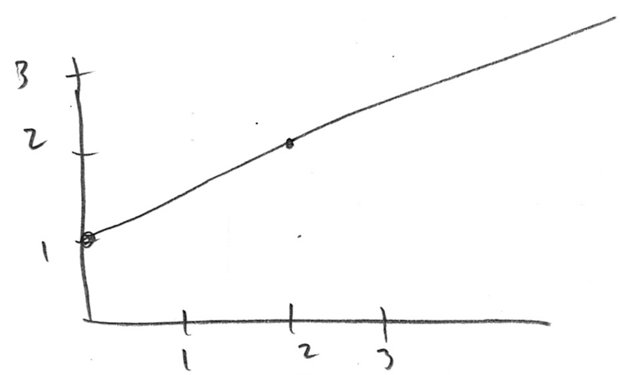
  $\Rightarrow$ b is the intercept of line w/ y axis

  $w = 0$
  $b = 1.5$ $\rightarrow$ $f_{0,1.5}(x) = \cancel{0w} + 1.5$

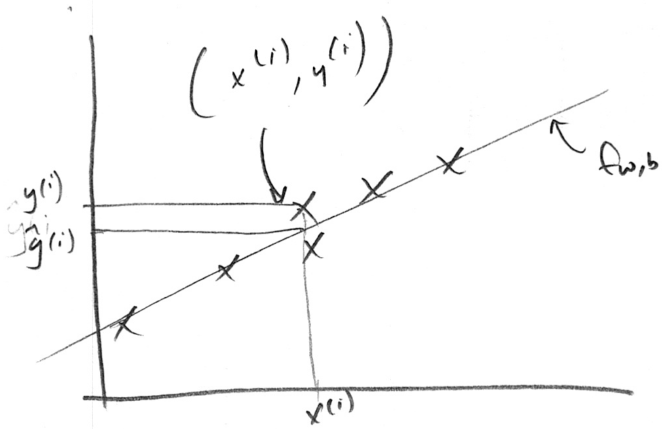
- 

  $w = 0.5$
  $b = 0$ $\Rightarrow$ $f_{0.5,0}(x) = 0.5x$

  slope

- 

  $w = 0.5$
  $b = 1$ $\Rightarrow$ $f_{0.5,1} = 0.5x + 1$



$$\hat{y}^{(i)} = f_{w,b}(x^{(i)})$$

$$f_{w,b}(x^{(i)}) = w x^{(i)} + b$$

Cost Function: $J(w,b) = \frac{1}{2m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)} \right)^2$ ← # of training samples (m)

$$= \frac{1}{2m} \sum_{i=1}^{m} \left(f_{w,b}(x^{(i)}) - y^{(i)} \right)^2$$

Training: Find w, b where $\hat{y}^{(i)}$ is close to $y^{(i)}$ for all $(x^{(i)}, y^{(i)})$

$\Rightarrow$ minimize $J(w,b)$

// Simplified Example

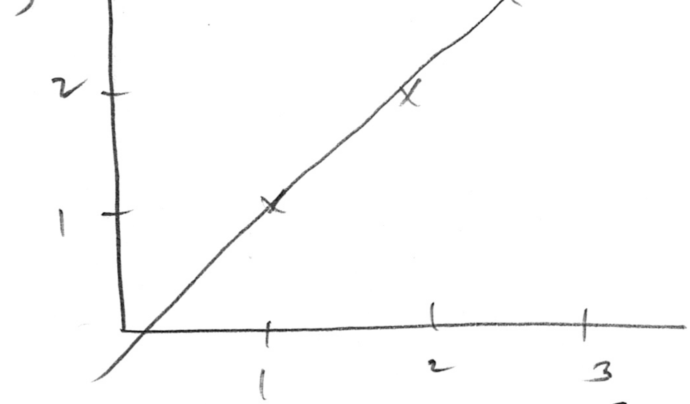
$f_w(x) = wx \qquad b = 0$

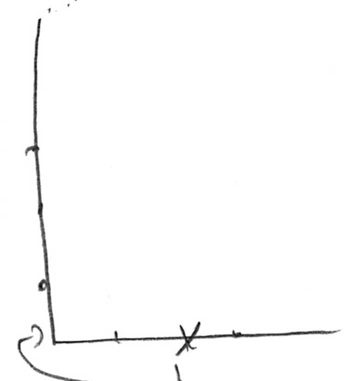
$$\Rightarrow J(w) = \frac{1}{2m} \sum_{i=1}^{m} \left(\underbrace{f_w(x^{(i)})}_{w x^{(i)}} - y^{(i)} \right)^2$$

$f_w(x)$: for fixed $w \Rightarrow$ value of x

$J(w)$ = function of w

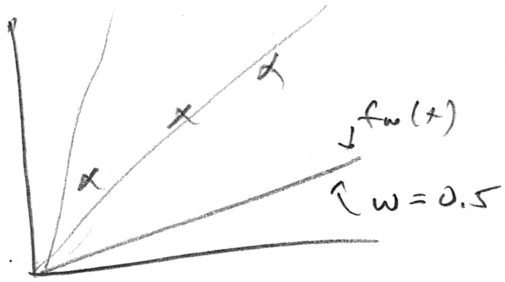
minimize $J(w)$ (w)

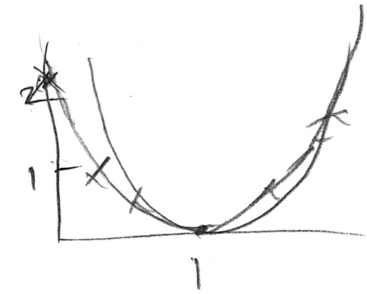




$\Rightarrow \underset{w=1}{J(w)} \; \left(f_w(x^{(i)}) - y^{(i)} \right)^2 = \frac{1}{2m} \sum \left(w x^{(i)} - y^{(i)} \right)^2$

$$= \frac{1}{2m} \left(0^2 + 0^2 + 0^2 \right) = 0$$





$$J(0.5) = \frac{1}{\underset{3}{2m}}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$$

$$= \frac{1}{2 \times 3}\left[3.5\right]$$

$$J(0) = \frac{1}{2m}\left(1^2 + 2^2 + 3^2\right) = \frac{1}{6}\left[14\right] \approx 2.3$$