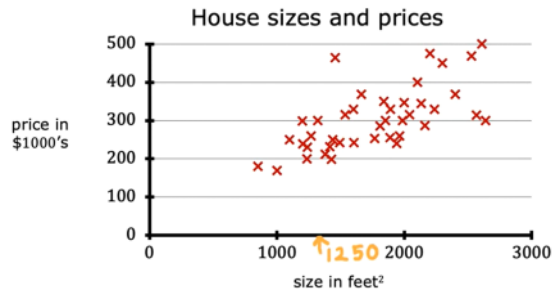
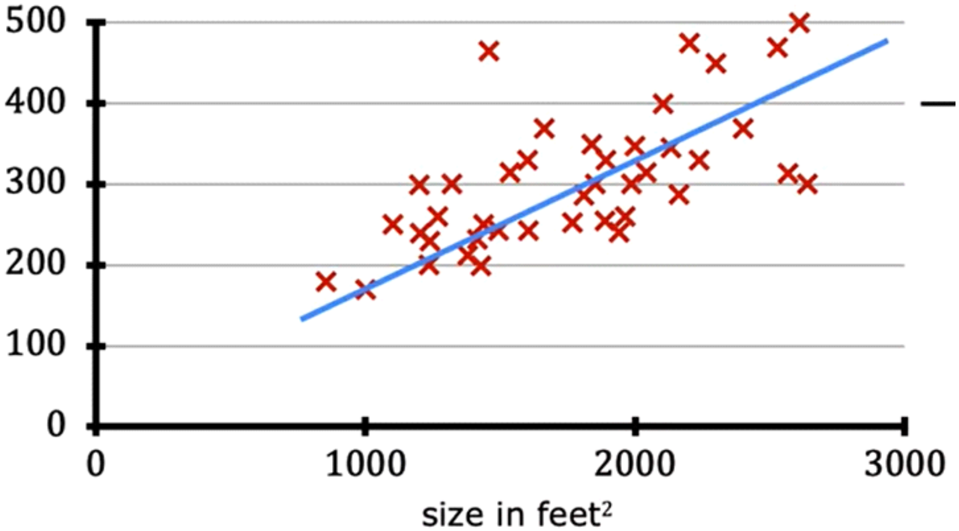
Linear Regression with One Variable

Data table

size in feet²	price in $1000's
2104	400
1416	232
1534	315
852	178
...	...
3210	870





Terminology

Training set: Data used to train the model

	x size in feet²	y price in $1000's
(1)	2104	400
(2)	1416	232
(3)	1534	315
(4)	852	178
...	...	...
(47)	3210	870

m = 47

x = 2104 y = 400

(x , y) = (2104, 400)

Notation:

x = "input" variable
feature

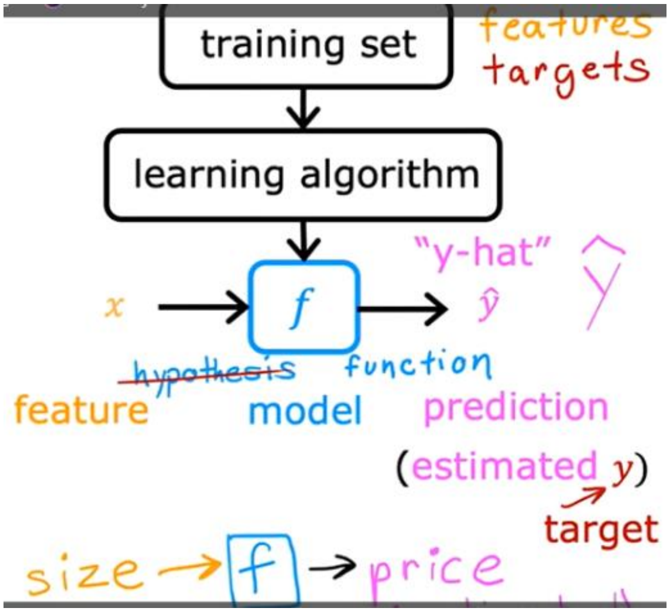
y = "output" variable
"target" variable

m = number of training examples

(x, y) = single training example

$(x^{(i)}, y^{(i)})$

$(x^{(i)}, y^{(i)})$ = i^th training example
(1^st, 2^nd, 3^rd ...)



How to represent f?

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

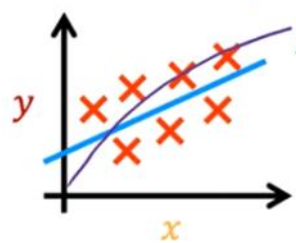
$f_{w,b}(x) = wx + b$

w, b: parameters

coefficients

weight

$f(x)$



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

$f(x) = wx + b$

linear

single feature x

Linear regression with one variable.

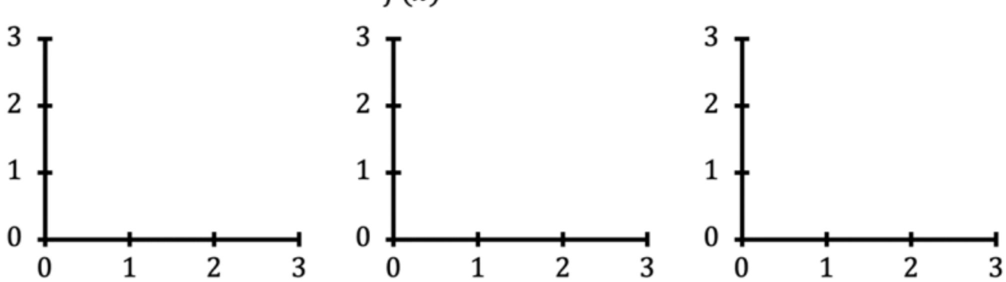
size

Univariate linear regression.

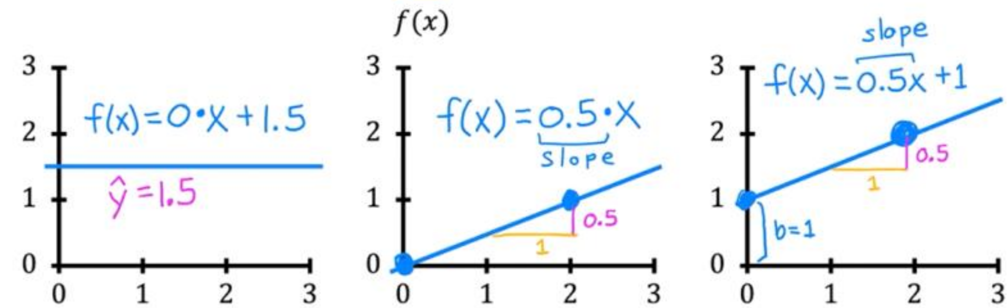
one variable

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

$f(x)$

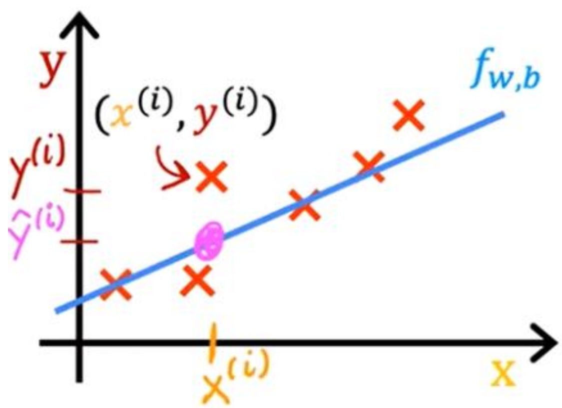


$$\underline{f_{w,b}}(x) = wx + b$$



→ $w = 0$	→ $w = 0.5$	→ $w = 0.5$
→ $b = 1.5$ y-intercept	→ $\mathrm{b} = 0$	→ $b = 1$

$$\hat{y} = \overbrace{f}^{\text{St. Line}}(\omega, b)(x^{i})$$



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

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

Cost function: Squared error cost function

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

error

m = number of training examples

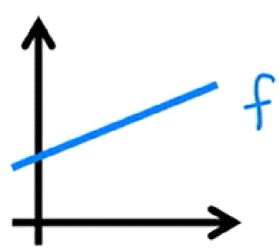
$$J(w,b) = \frac{1}{2m}\sum_{i=1}^{m}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)^2$$

intuition

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

model:

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



parameters:

$$w, b$$

cost function:

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

goal:

$$\underset{w,b}{\text{minimize}}\, J(w,b)$$

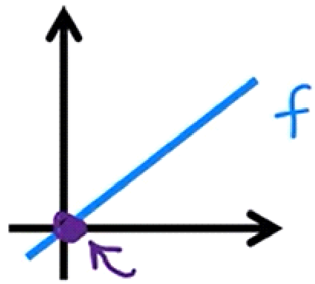
simplified

$f_w(x) = \underline{wx}$ $\quad b = \emptyset$

$\uparrow \qquad \uparrow$

$0 \qquad 0$

w



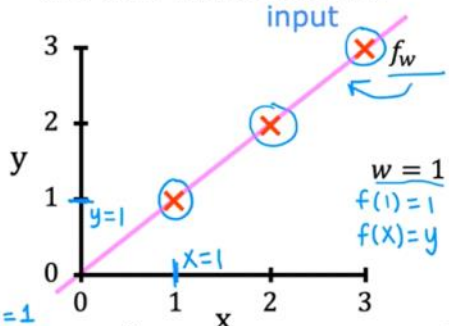
$$J(w) = \frac{1}{2m}\sum_{i=1}^{m}(\underbrace{f_w(x^{(i)})}_{wx^{(i)}} - y^{(i)})^2$$

$$\underset{w}{\text{minimize}}\, J(w)$$

$f_w(x)$

(for fixed w, function of x)

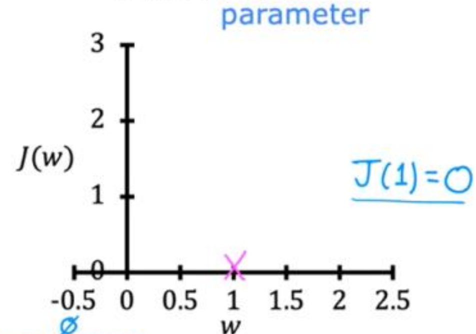
input



$J(w)$

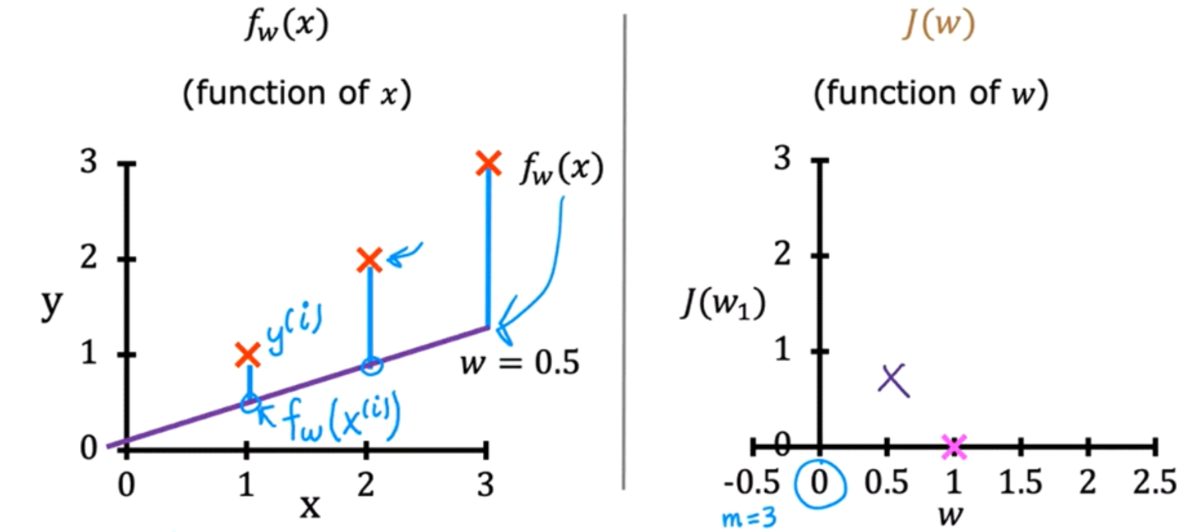
(function of w)

parameter

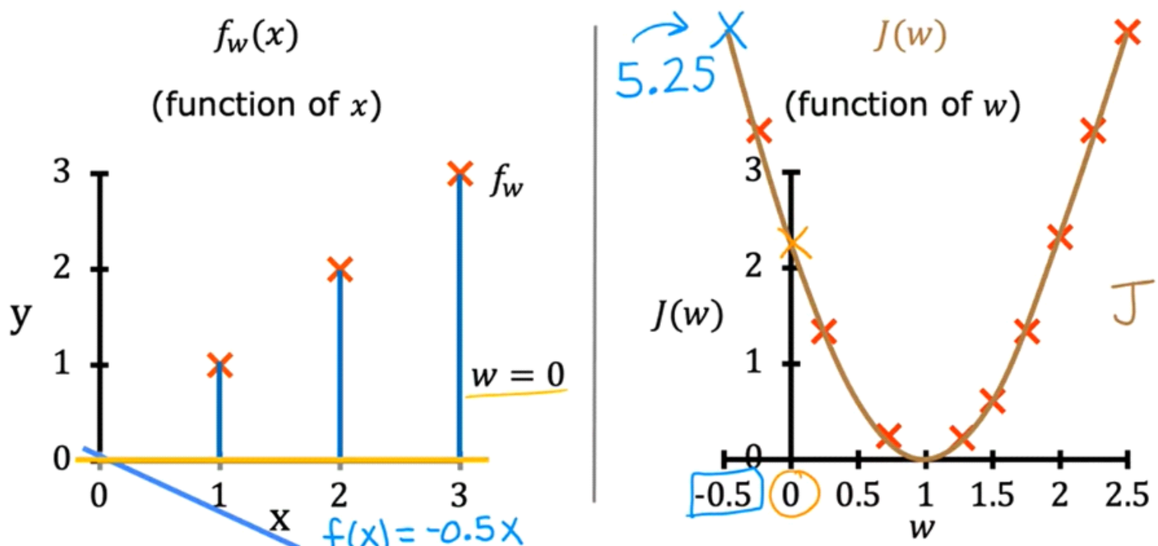


$w=1$

$$\underset{\downarrow}{J(w)} = \frac{1}{2m}\sum_{i=1}^{m}(\underbrace{f_w(x^{(i)})}_{wx^{(i)}} - y^{(i)})^2 = \frac{1}{2m}\sum_{i=1}^{m}(\overbrace{wx^{(i)} - y^{(i)}}^{\emptyset})^2 = \frac{1}{2m}(0^2 + 0^2 + 0^2) = 0$$



$J(0.5) = \frac{1}{2m}\left[(0.5-1)^2+(1-2)^2+(1.5-3)^2\right] = \frac{1}{2\times3}[3.5] = \frac{3.5}{6} \approx 0.58$



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

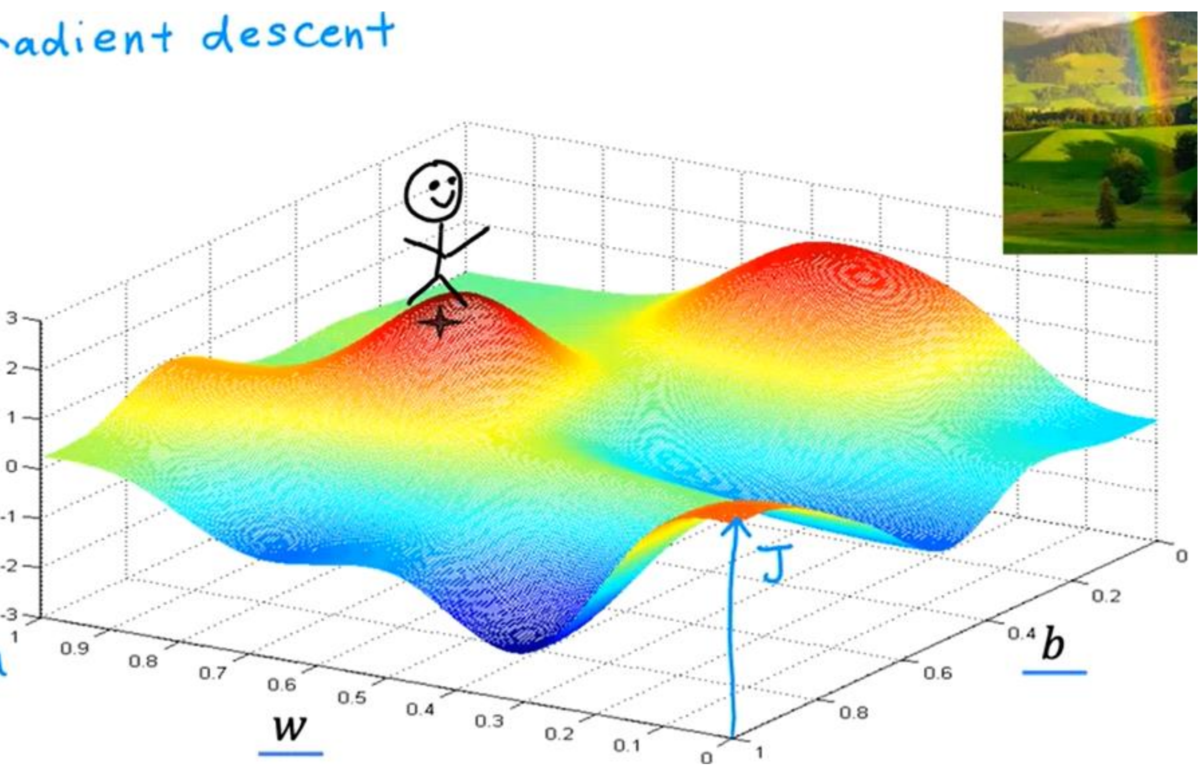
goal of linear regression:

$$\underset{w}{\text{minimize}}\, J(w)$$

general case:

$$\underset{w,b}{\text{minimize}}\, J(w, b)$$

gradient descent



$J(w, b)$

not squared error cost

not linear regression