Cost Function

Wednesday, February 26, 2025 1:45 PM

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

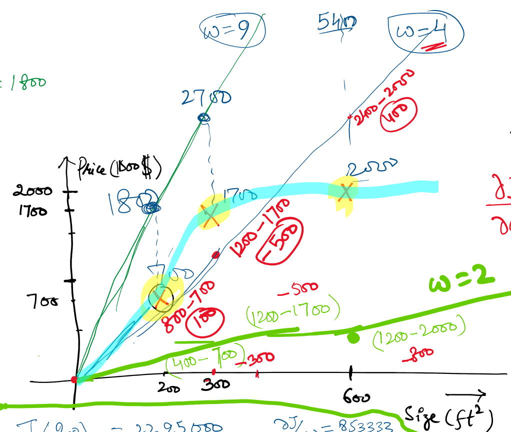
Find w,b:
$\hat{y}^{(i)}$ is close to $y^{(i)}$ for all $(x^{(i)},y^{(i)})$.

$f = wx + b$
$= 9(200)+0 = 1800$

$w, \boxed{b=0}$
$\boxed{w = 9}$

Prediction

$$J = \frac{1}{2m}\sum_{1}^{3}\left(\hat{y}-y^{i}\right)^2$$
$$= \frac{1}{6}\left[(1800-700)^2 + (2700-1700)^2 + (5400-2000)^2\right]$$
$$= 2295000$$



$$J = \frac{1}{2m}\sum\left(\hat{y}-y^{i}\right)^2$$

$$J = \frac{1}{2m}\sum\left(wx^{i}+b-y^{i}\right)^2$$

$$\frac{\partial J}{\partial w} = \frac{2}{2m}\sum\left[(wx^{i}+\cancel{b}-y^{i})x^{i}\right]$$

$$= \frac{1}{3}\left[9(200)-700)200 + \cdots\right]$$

$$= \boxed{853333}$$

$$\frac{\partial J}{\partial w} = +ve$$

$$\boxed{w_{new} = w_{old} \pm \Delta}$$

$$w = w \pm \Delta$$

$$w = w - \alpha\frac{\partial J}{\partial w}$$

5

$$w = 9-5$$

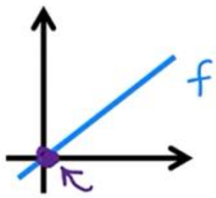
$$\boxed{w=4}$$

$J(9,0) = 2295000$	$\partial J/\partial w = 853333$
$J(4,0) = 70,000$	$\frac{\partial J}{\partial w} = 36666$
$J(2,0) = 323,333$	$\frac{\partial J}{\partial w} = -290000$

simplified

$f_w(x) = wx$ $\quad b = \emptyset$

w



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

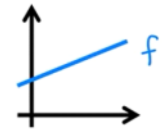
$$\underset{w}{\text{minimize}}\, J(w)$$

model:

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

parameters:

w, b



cost function:

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

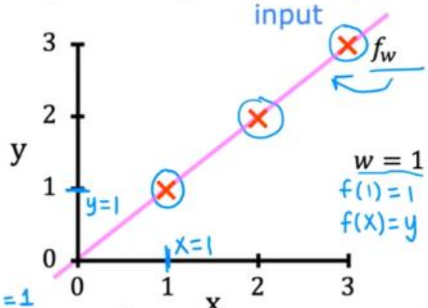
goal:

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

$f_w(x)$

(for fixed w, function of x)

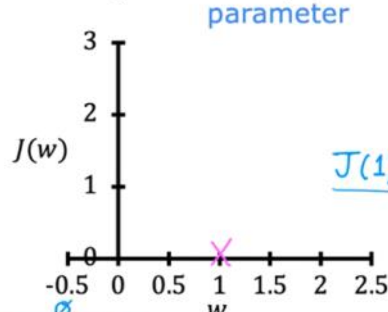
input



$J(w)$

(function of w)

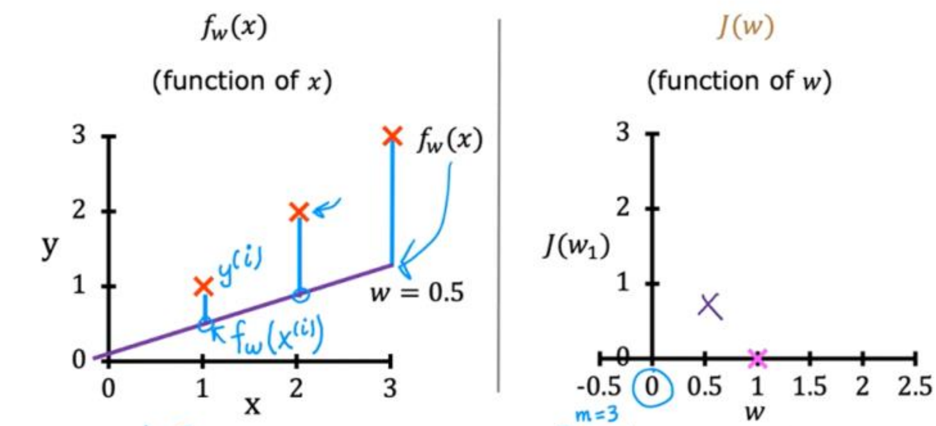
parameter



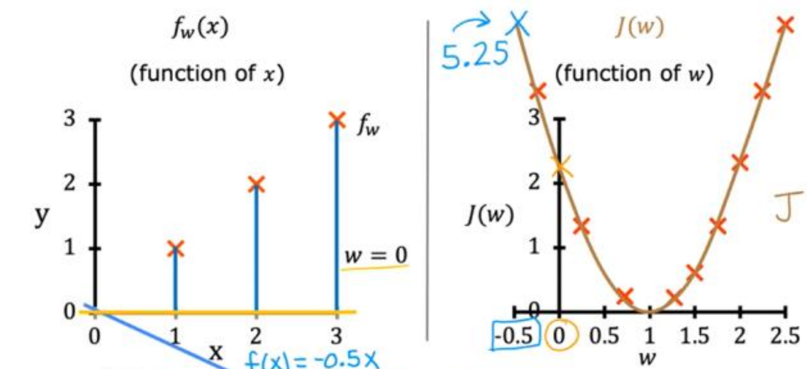
J(1)=0

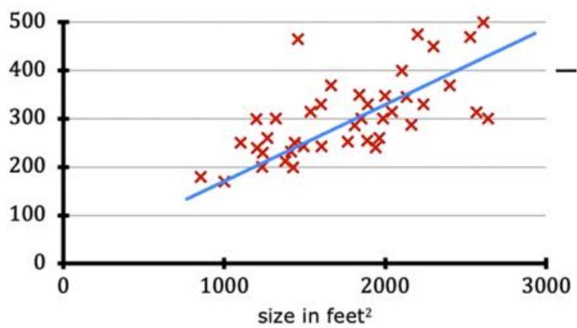
$w=1$

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



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





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

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

goal of linear regression:

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

general case:

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