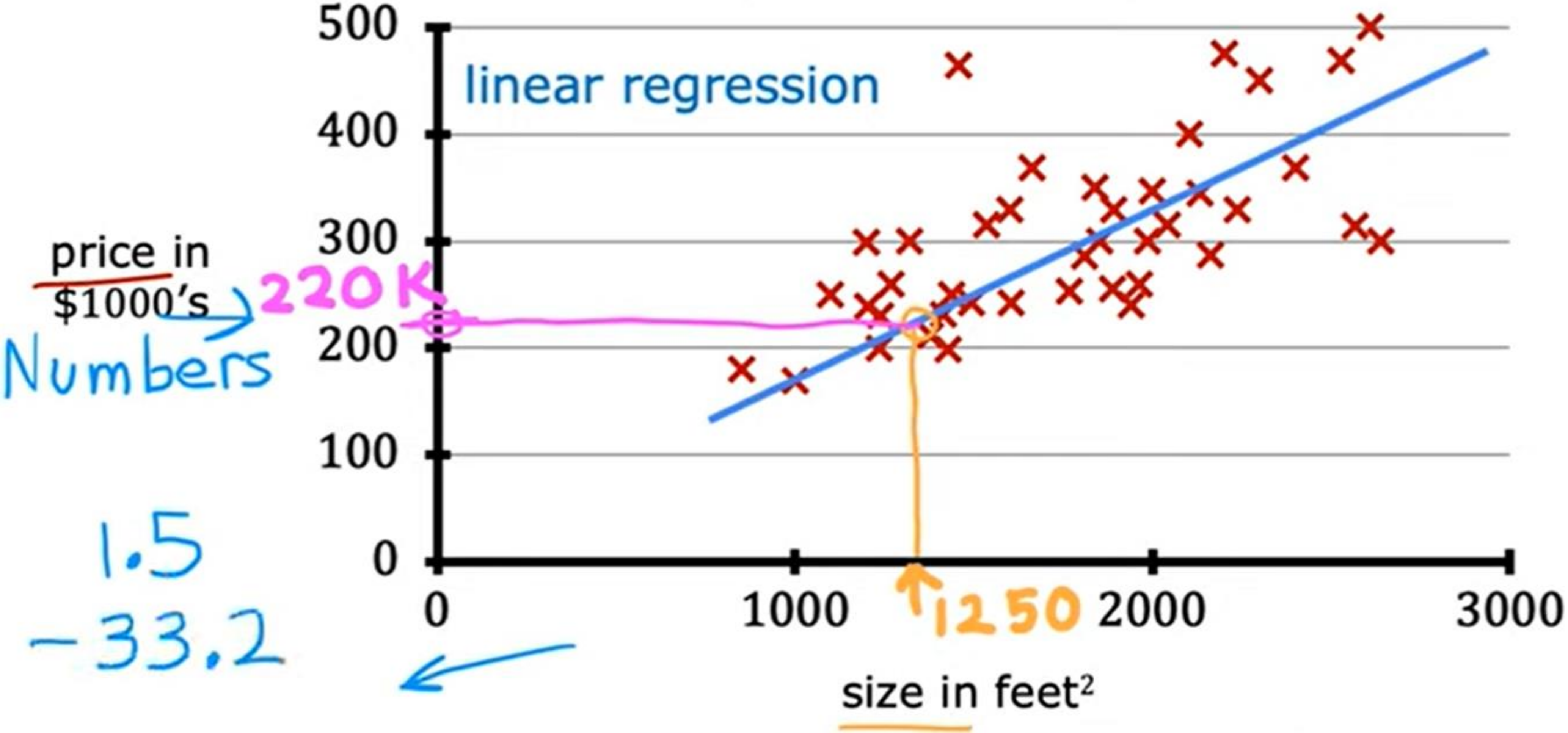
Linear Regression with One Variable

Linear Regression Model Part 1

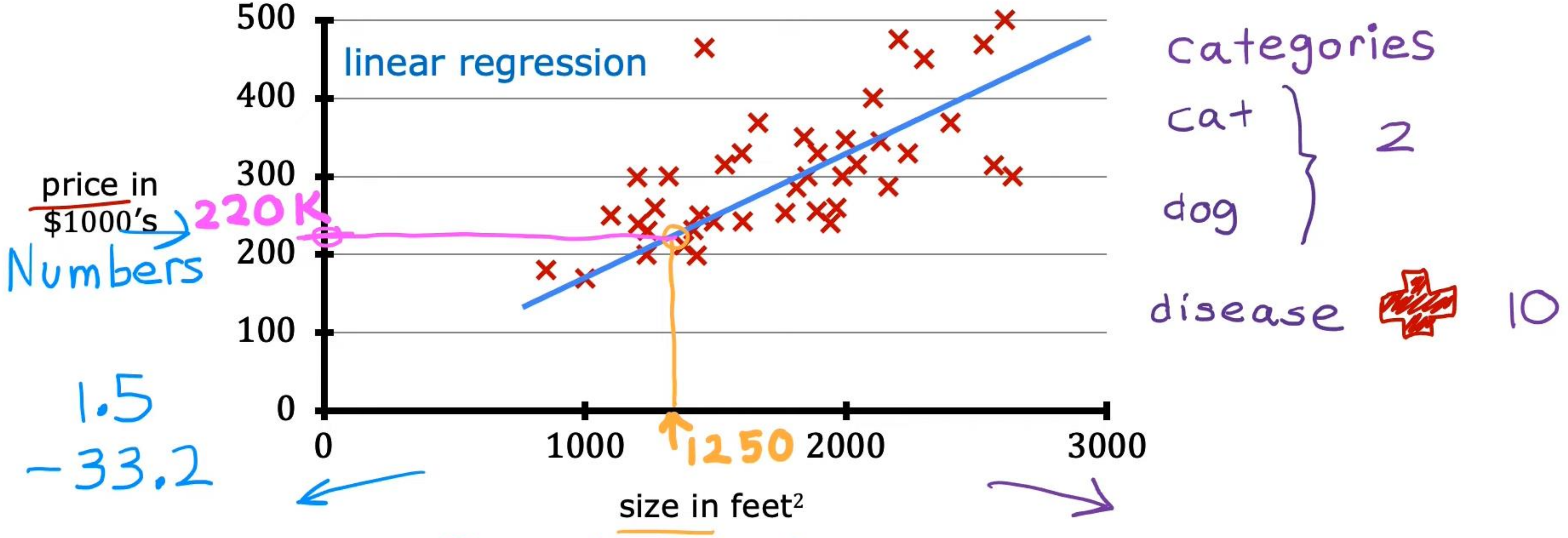
House sizes and prices



Regression model
Predicts numbers

Supervised learning model
Data has "right answers"

House sizes and prices



Regression model
Predicts numbers
Infinitely many possible outputs

Supervised learning model
Data has "right answers"

Classification model
Predicts categories
Small number of possible outputs

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

$m = 47$

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

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

$x^{(2)} = 1416$ $x^{(2)} \neq x^2$ not exponent

Notation:

x = "input" variable
feature

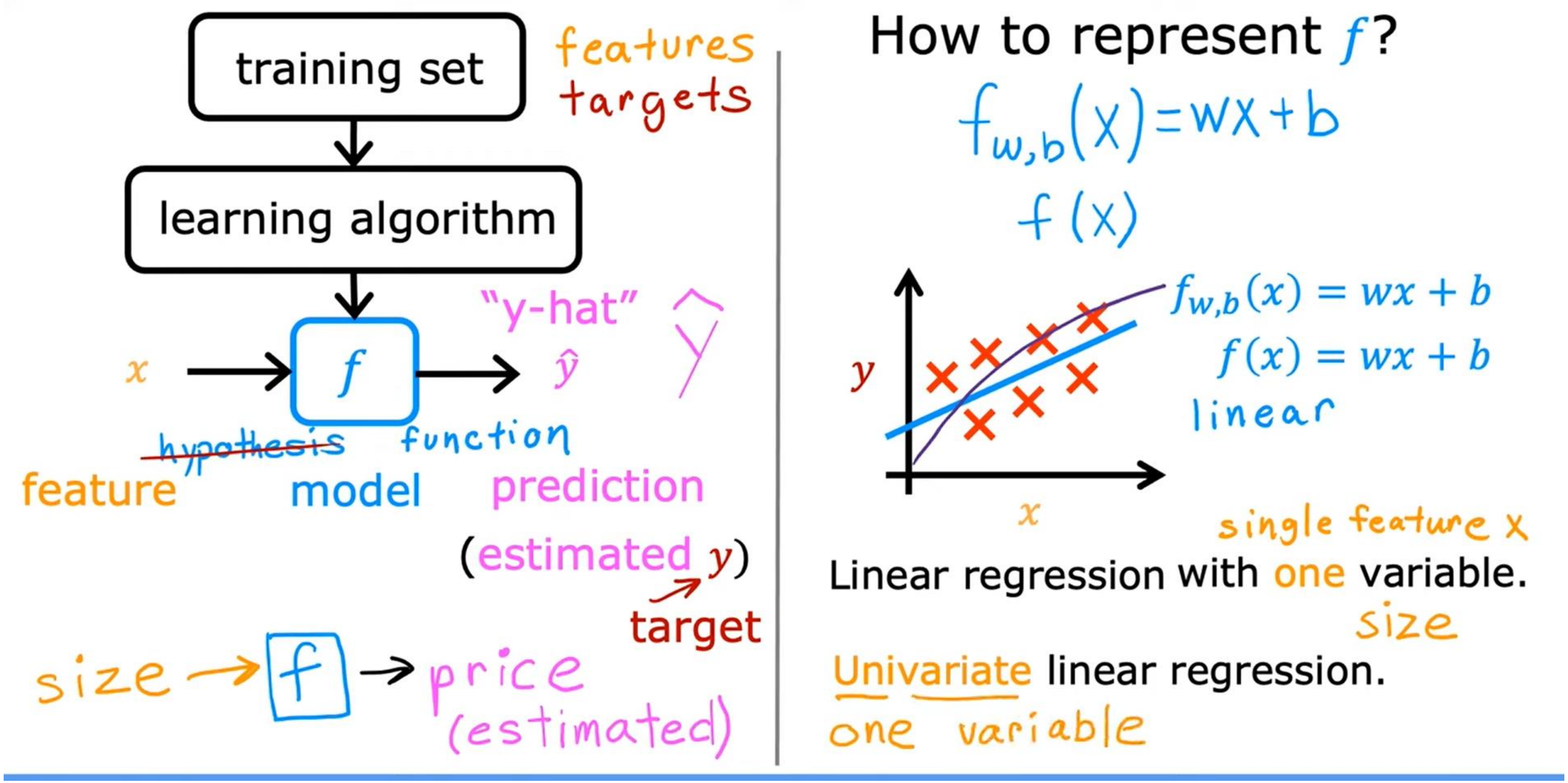
y = "output" variable
"target" variable

m = number of training examples

(x, y) = single training example

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

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


training set
features
targets
learning algorithm
"y-hat"
x
f
ŷ
hypothesis
function
feature
model
prediction
(estimated y)
target
size
f
price
(estimated)
How to represent f?
f_w,b(x) = wx + b
f(x)
y
f_w,b(x) = wx + b
f(x) = wx + b
linear
x
single feature x
Linear regression with one variable.
size
Univariate linear regression.
one variable

Linear Regression with One Variable

Cost Function

Training set

features size in feet^2 (x)	targets price \$1000's (y)
2104	460
1416	232
1534	315
852	178
...	...

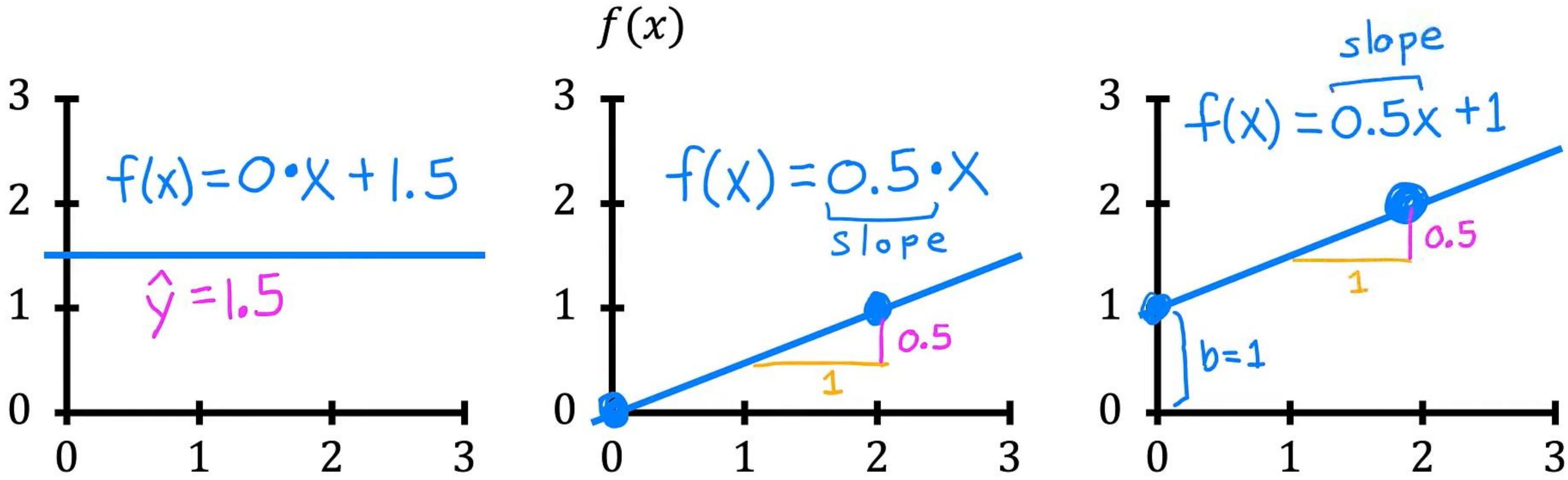
Model: $f_{w,b}(x) = wx + b$

w, b: parameters

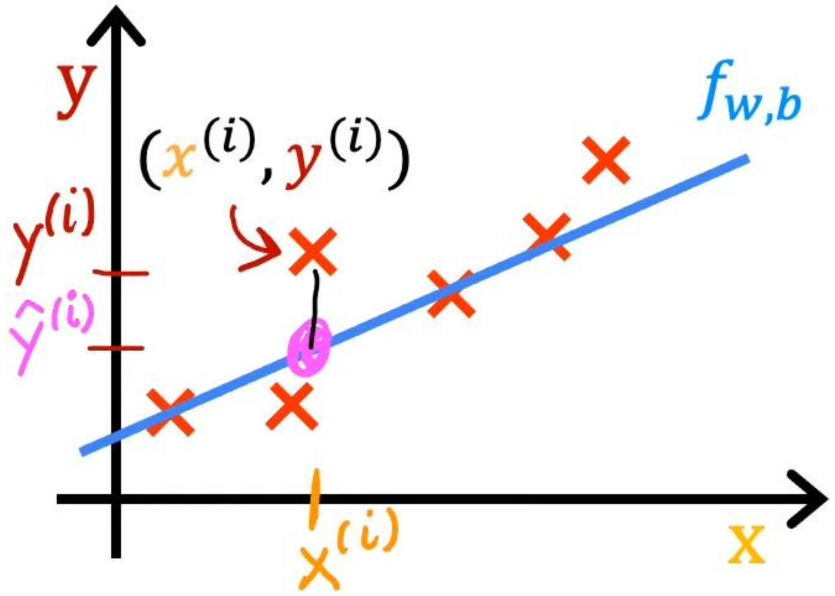
coefficients
weights

What do w, b do?

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



$w = 0$	$w = 0.5$	$w = 0.5$
$b = 1.5$	$b = 0$	$b = 1$
y-intercept		



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

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

Cost function: Squared error cost function

$$J(w,b) = \frac{1}{2m} \sum_{i=1}^{m} (\underbrace{\hat{y}^{(i)} - y^{(i)}}_{\text{error}})^2$$

m = number of training examples

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

Find w, b:

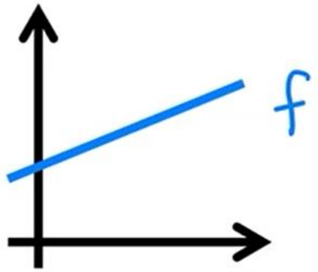
$\hat{y}^{(i)}$ is close to $y^{(i)}$ for all $(x^{(i)}, y^{(i)})$.

Linear Regression with One Variable

Cost Function Intuition

model:

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



parameters:

w, b

cost function:

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

goal:

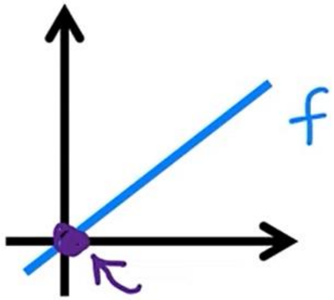
$$\underset{w,b}{\text{minimize}}\, J(w,b)$$

simplified

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

0 0

w



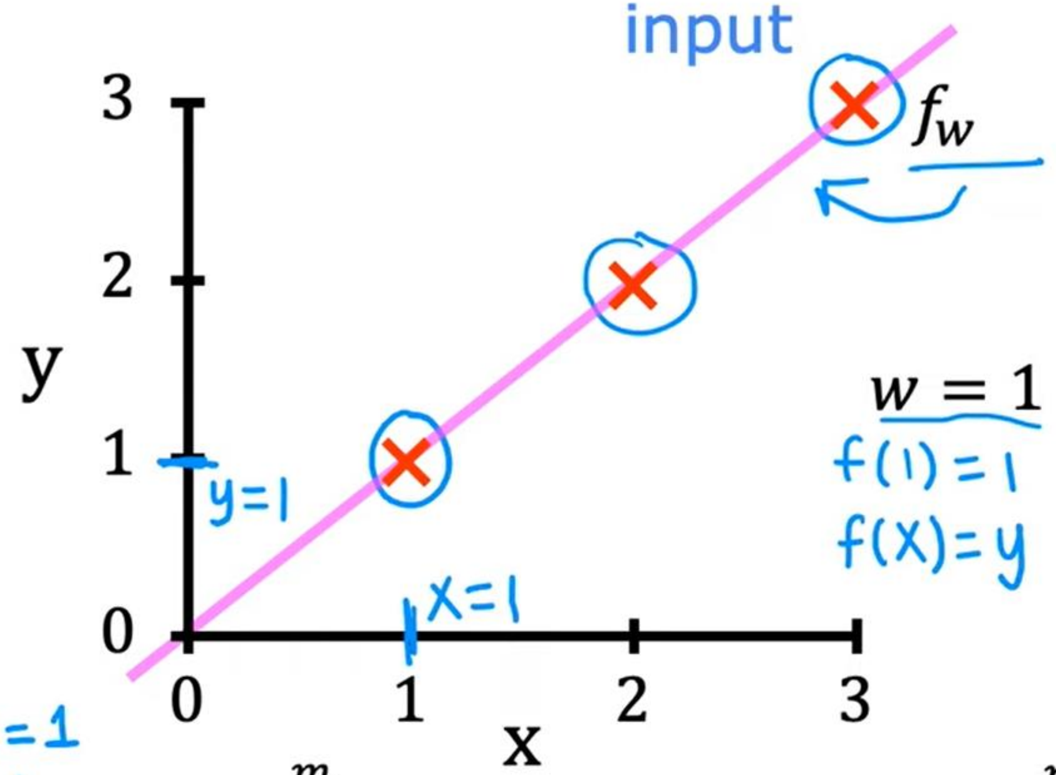
$$J(w) = \frac{1}{2m}\sum_{i=1}^{m}\left(\underbrace{f_w\left(x^{(i)}\right)}_{wx^{(i)}} - y^{(i)}\right)^2$$

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

$f_w(x)$

(for fixed w, function of x)

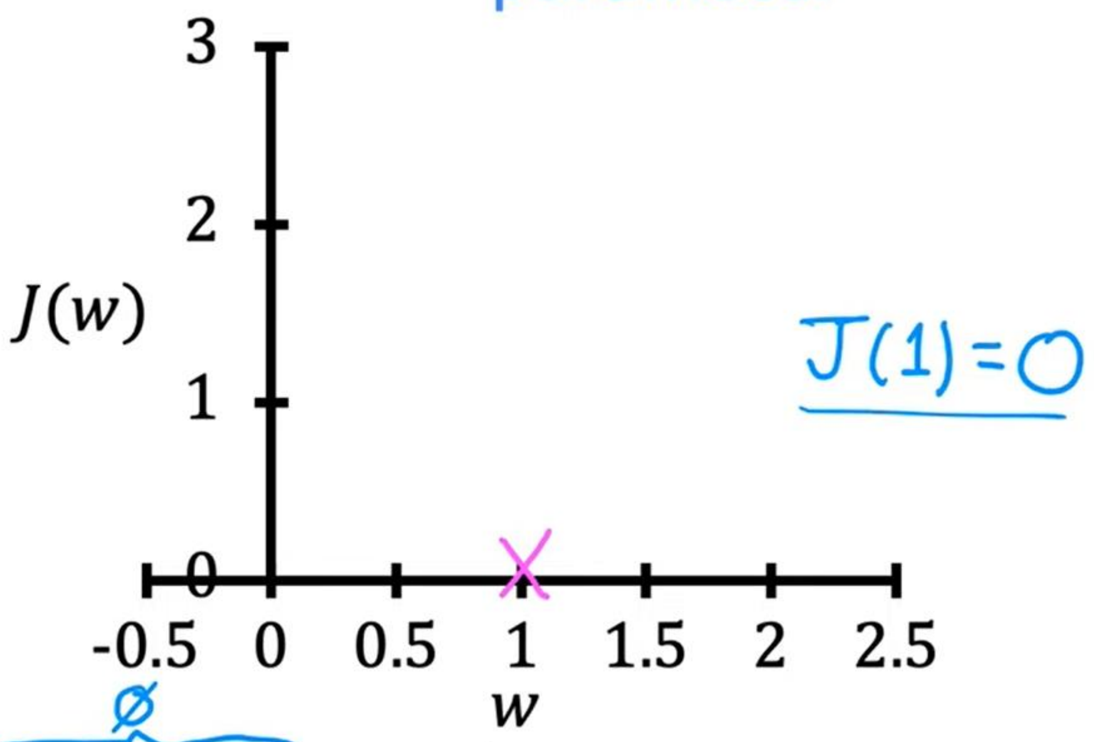
input



$J(w)$

(function of w)

parameter

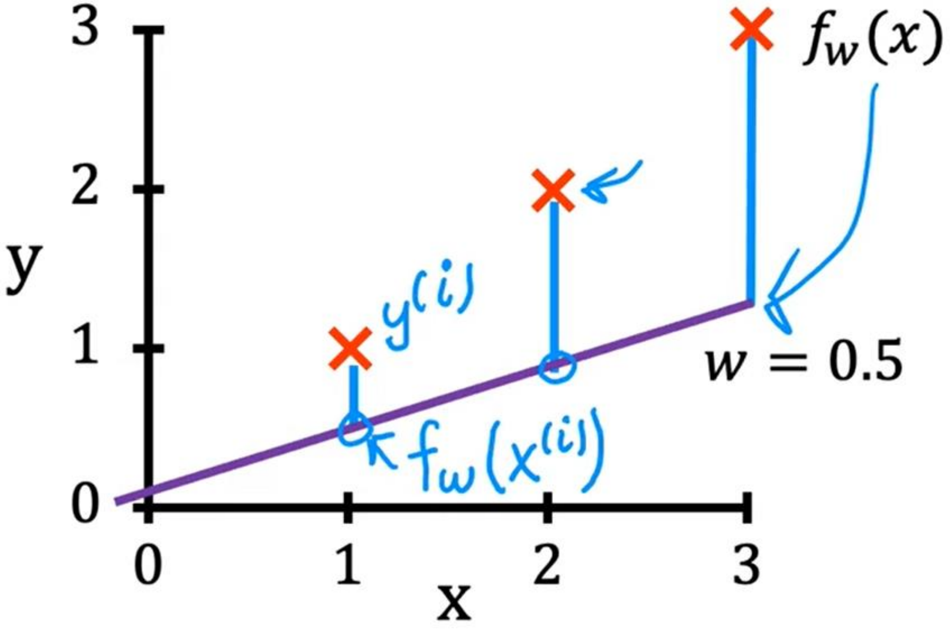


$w=1$

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

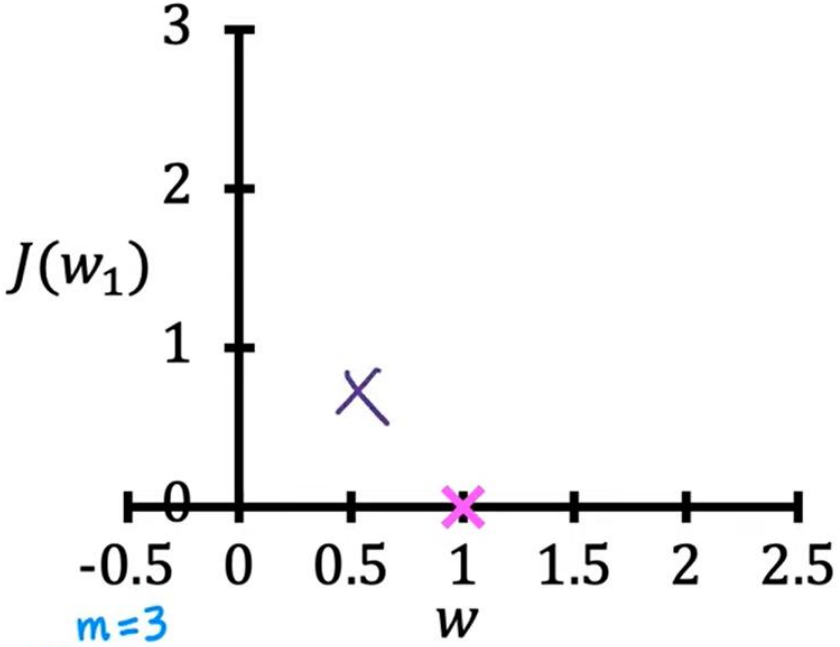
$f_w(x)$

(function of x)



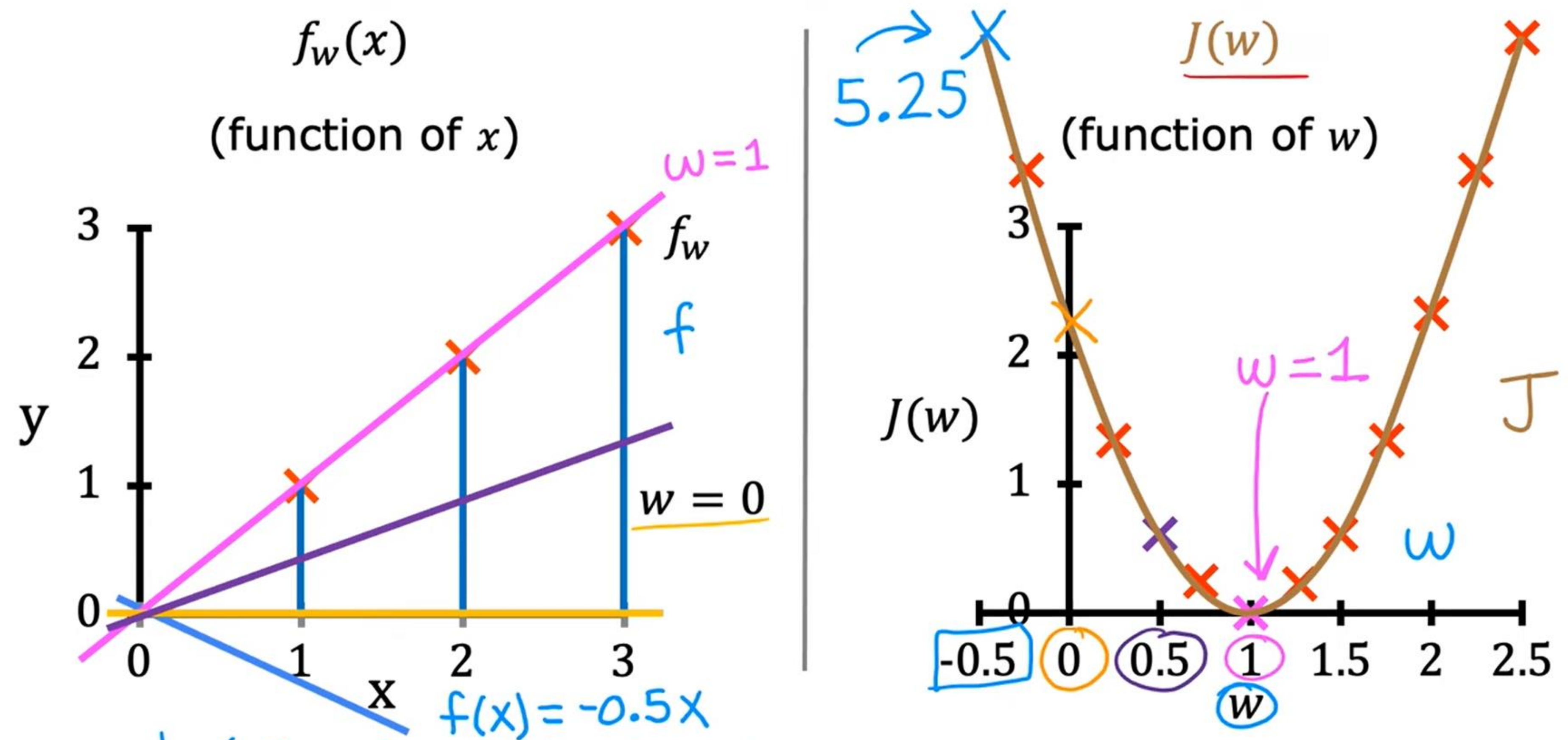
$J(w)$

(function of w)



$m=3$

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



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

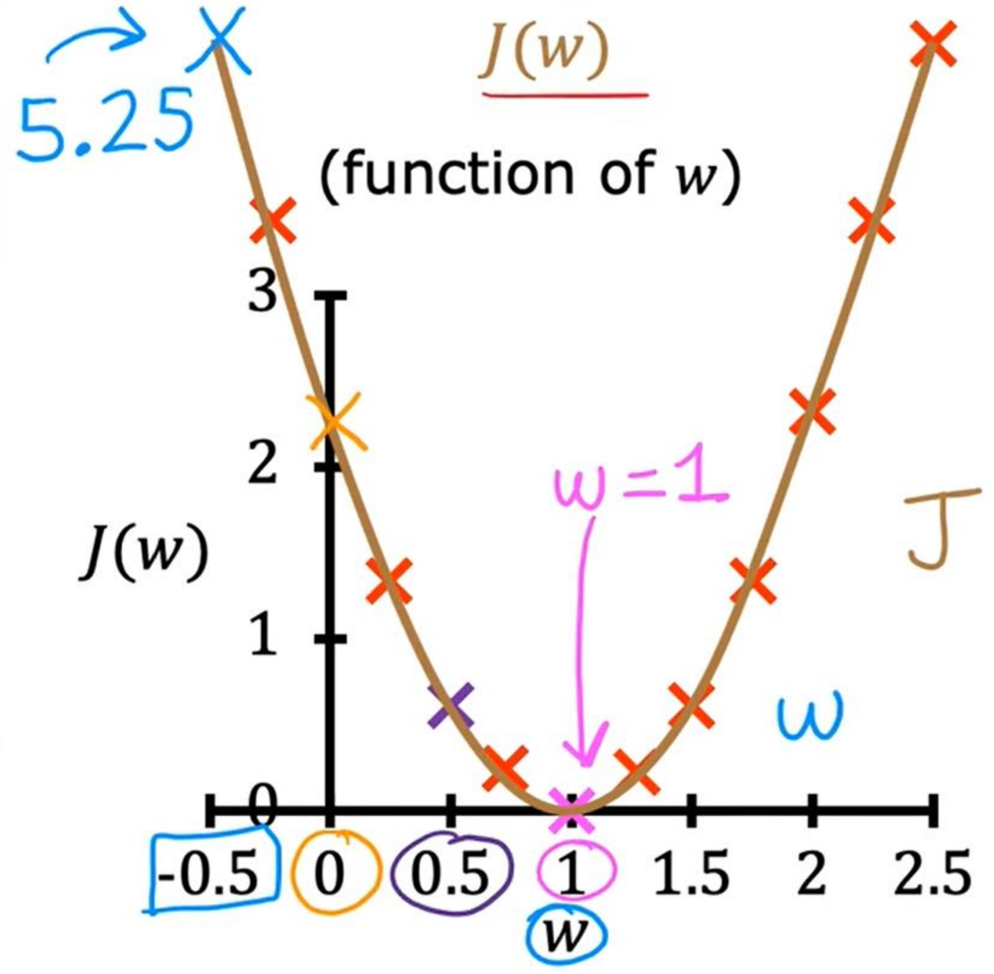
how to choose w?

goal of linear regression:

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

general case:

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



choose w to minimize J(w)

Linear Regression with One Variable

Visualizing the Cost Function

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

~~before: b=0~~

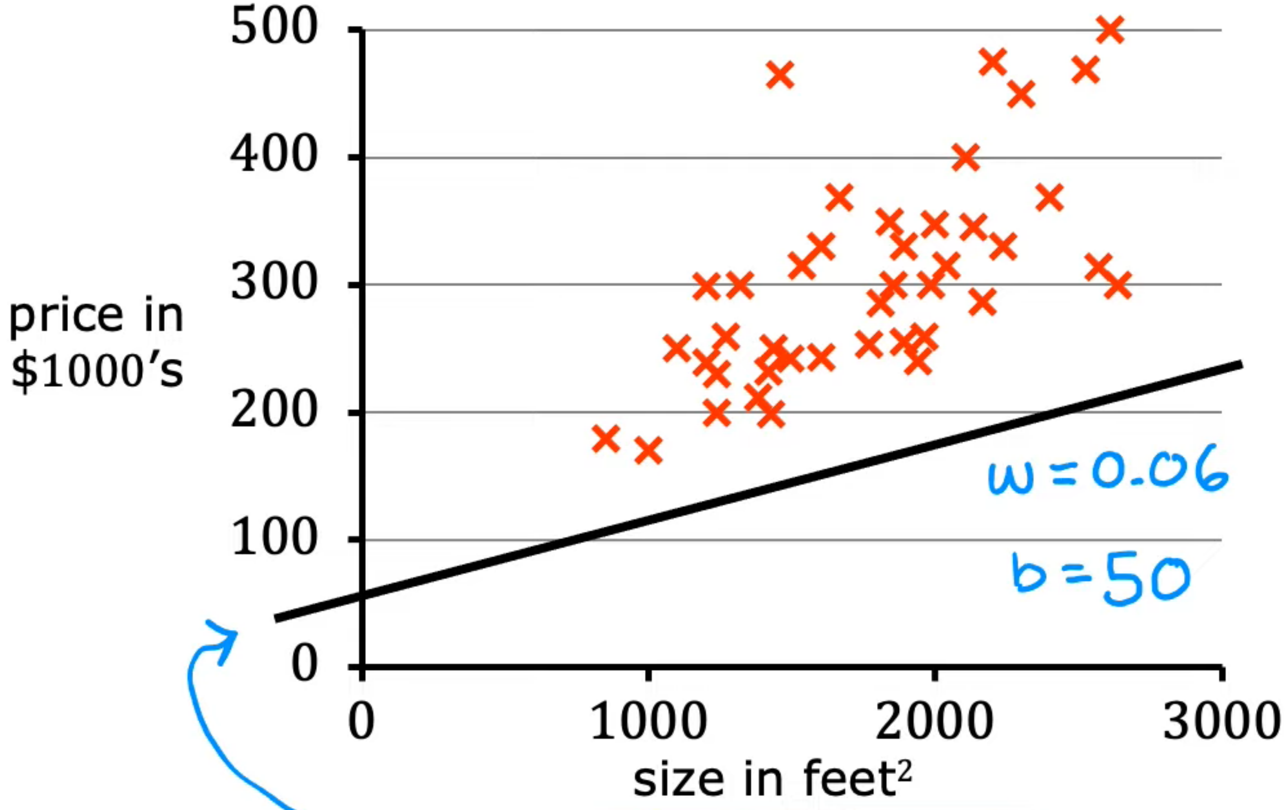
Parameters w, b

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

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

$f_{w,b}$

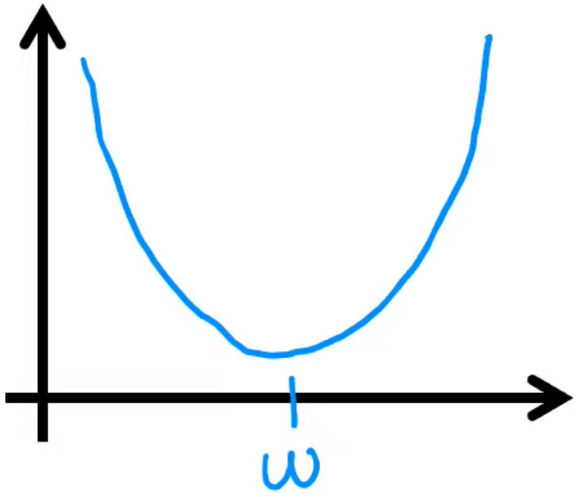
(function of x)



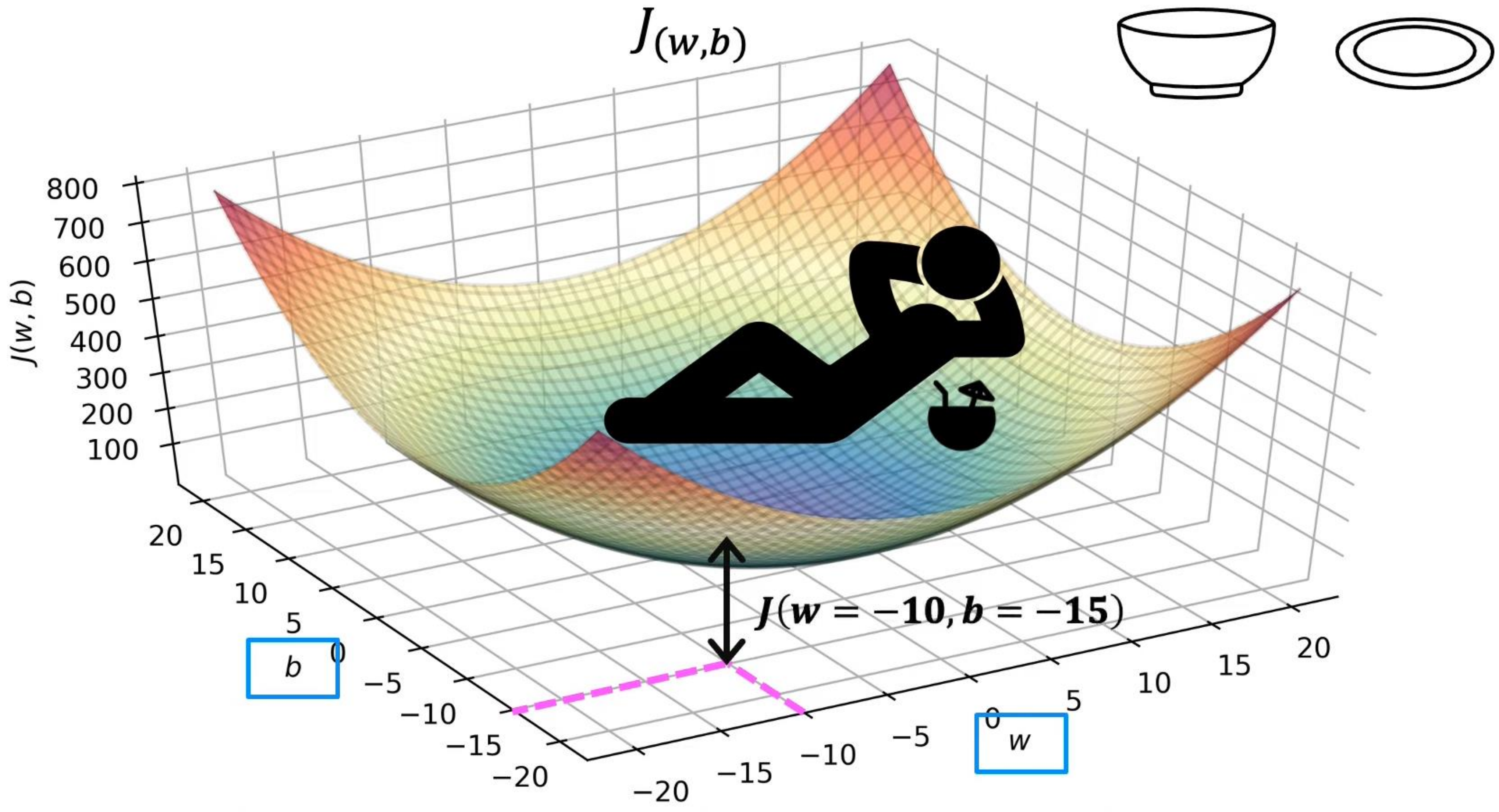
$$\boxed{f_{w,b}(x) = 0.06x + 50}$$

J

(function of w, b)



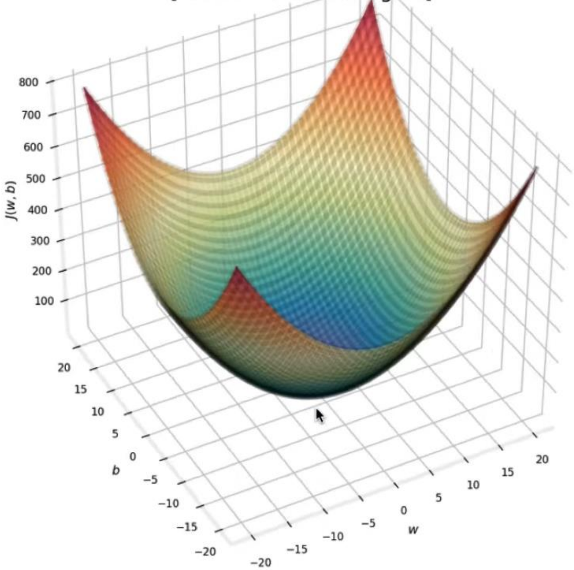
$\boxed{w, b}$



$J_{(w,b)}$
800
700
600
500
400
300
200
100
$J(w,b)$
20
15
10
5
0
−5
−10
−15
−20
b
−20
−15
−10
−5
0
5
10
15
20
w
$J(\boldsymbol{w} = \mathbf{-10}, \boldsymbol{b} = \mathbf{-15})$

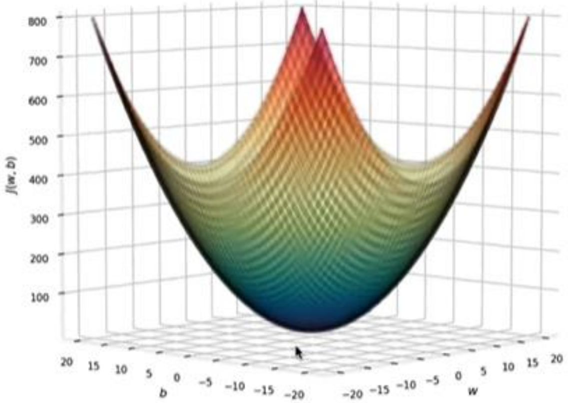
3D surface plot

$J(w, b)$
[You can rotate this figure]



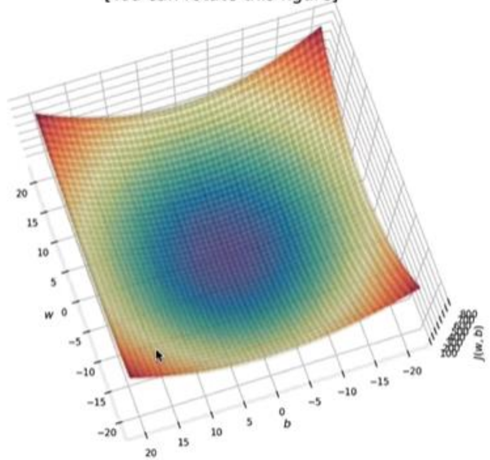
3D surface plot

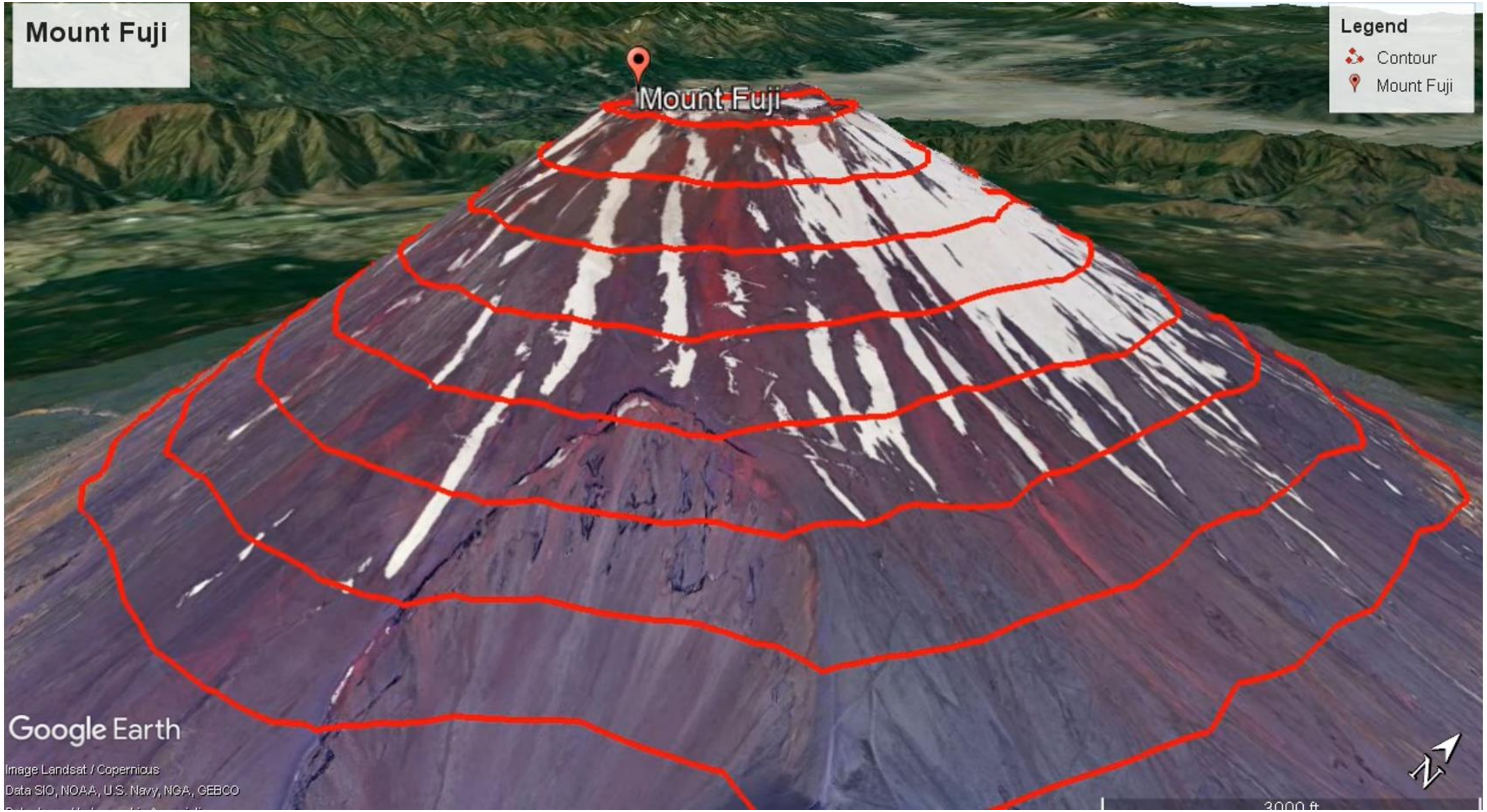
$J(w, b)$
[You can rotate this figure]

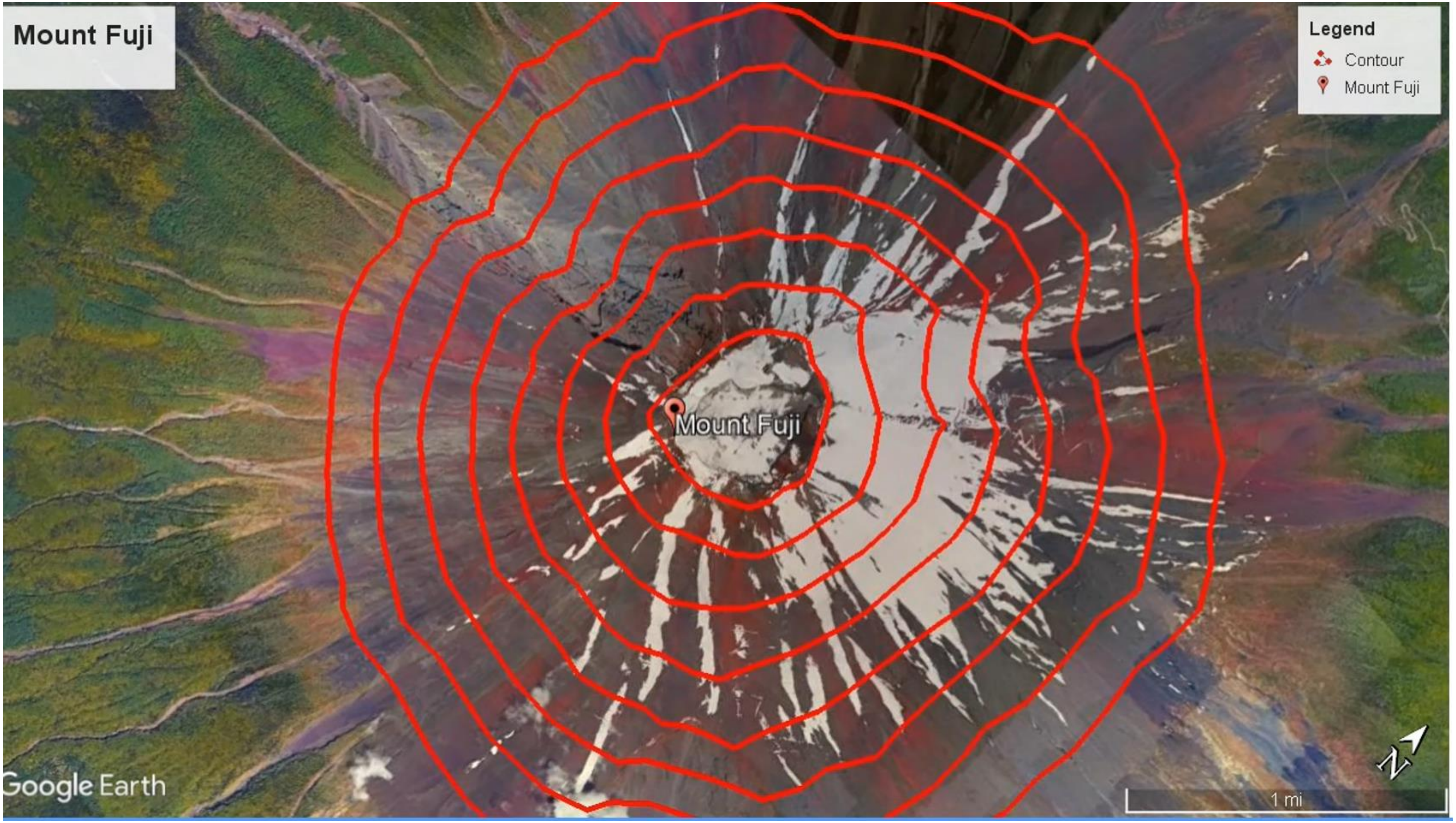


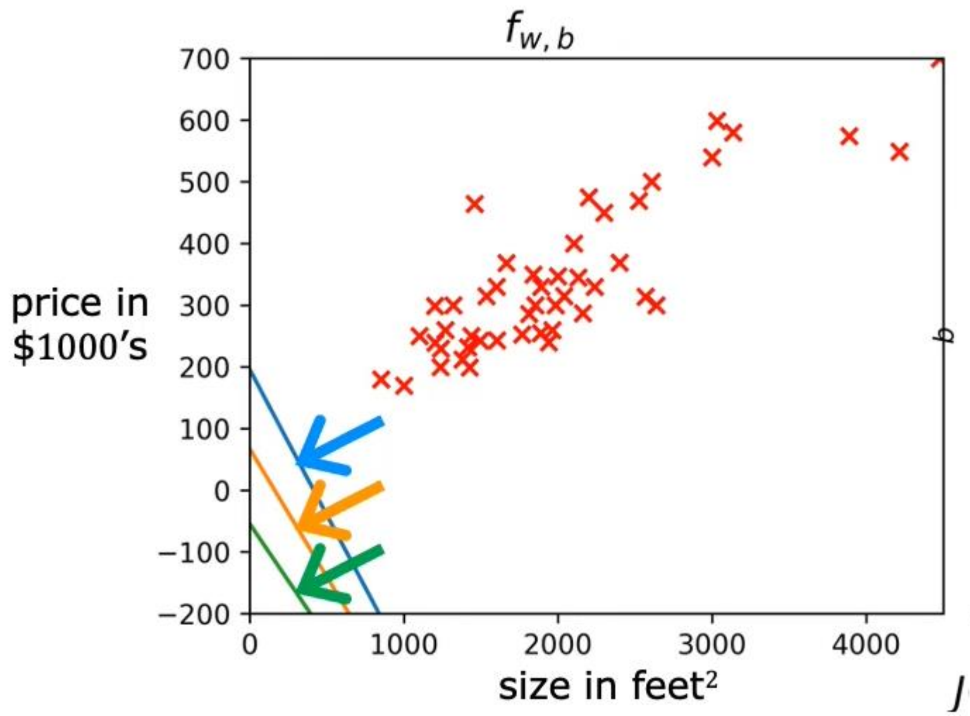
3D surface plot

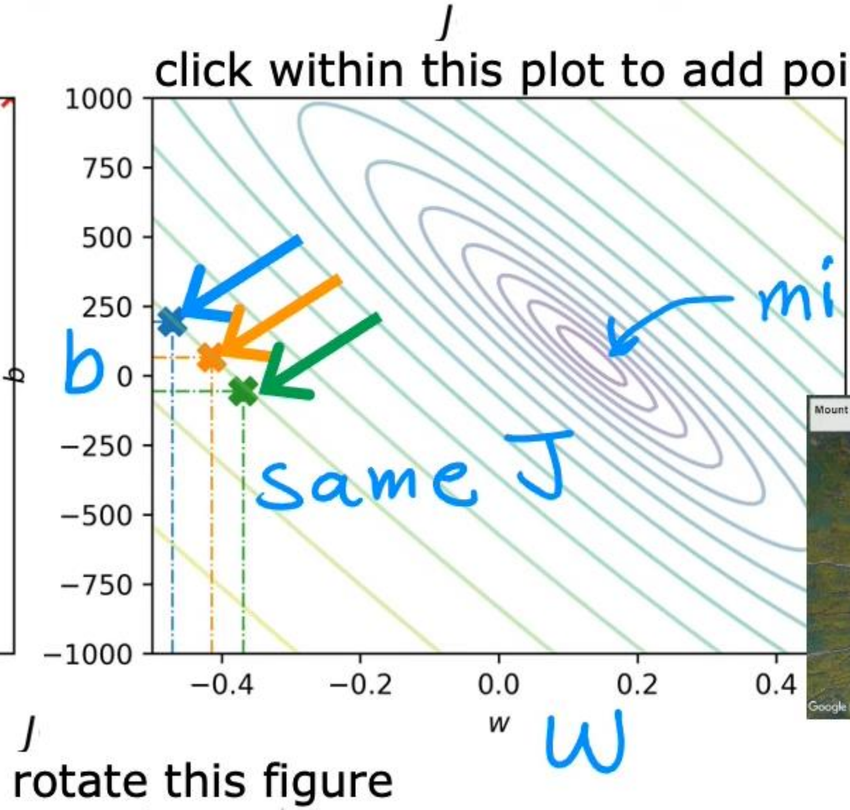
$J(w, b)$
[You can rotate this figure]

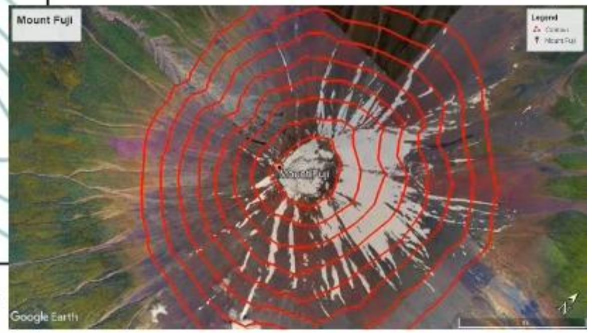


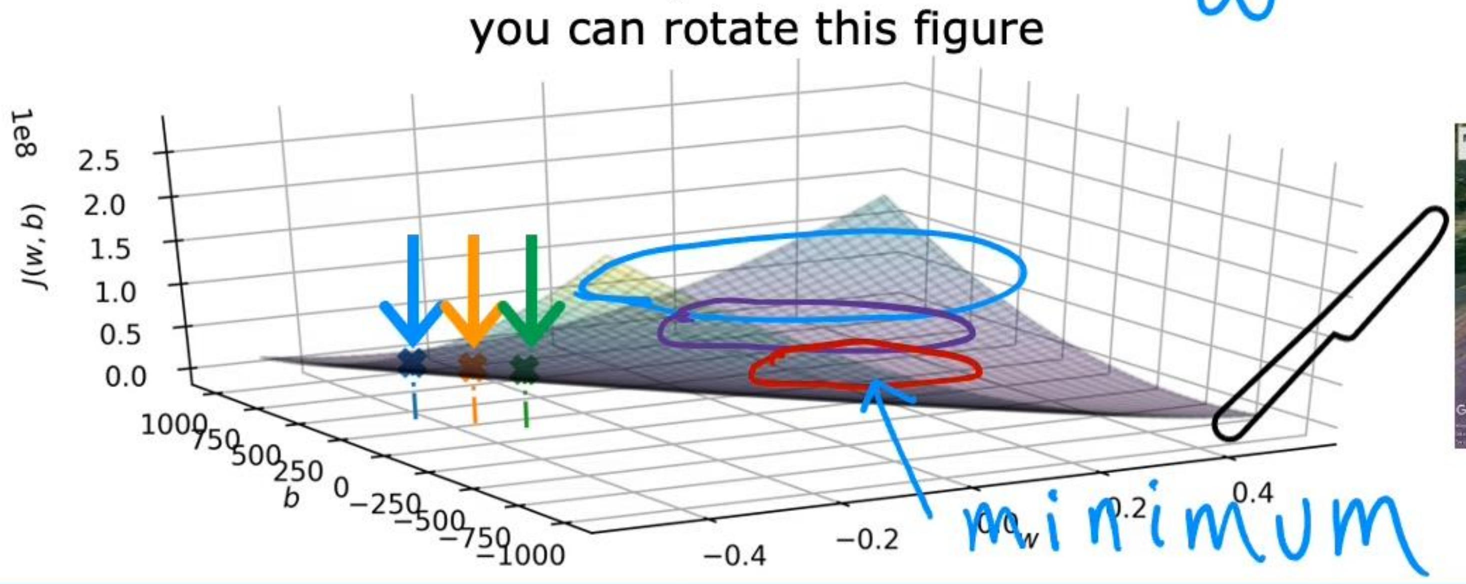

Mount Fuji
Legend
Contour
Mount Fuji
Mount Fuji
Google Earth
Image Landsat / Copernicus
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
3000 ft

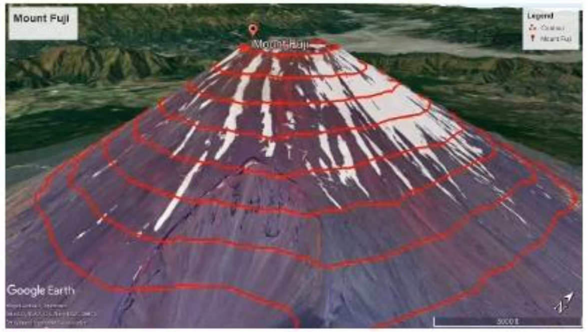

Mount Fuji
Legend
Contour
Mount Fuji
Mount Fuji
Google Earth
1 mi
N

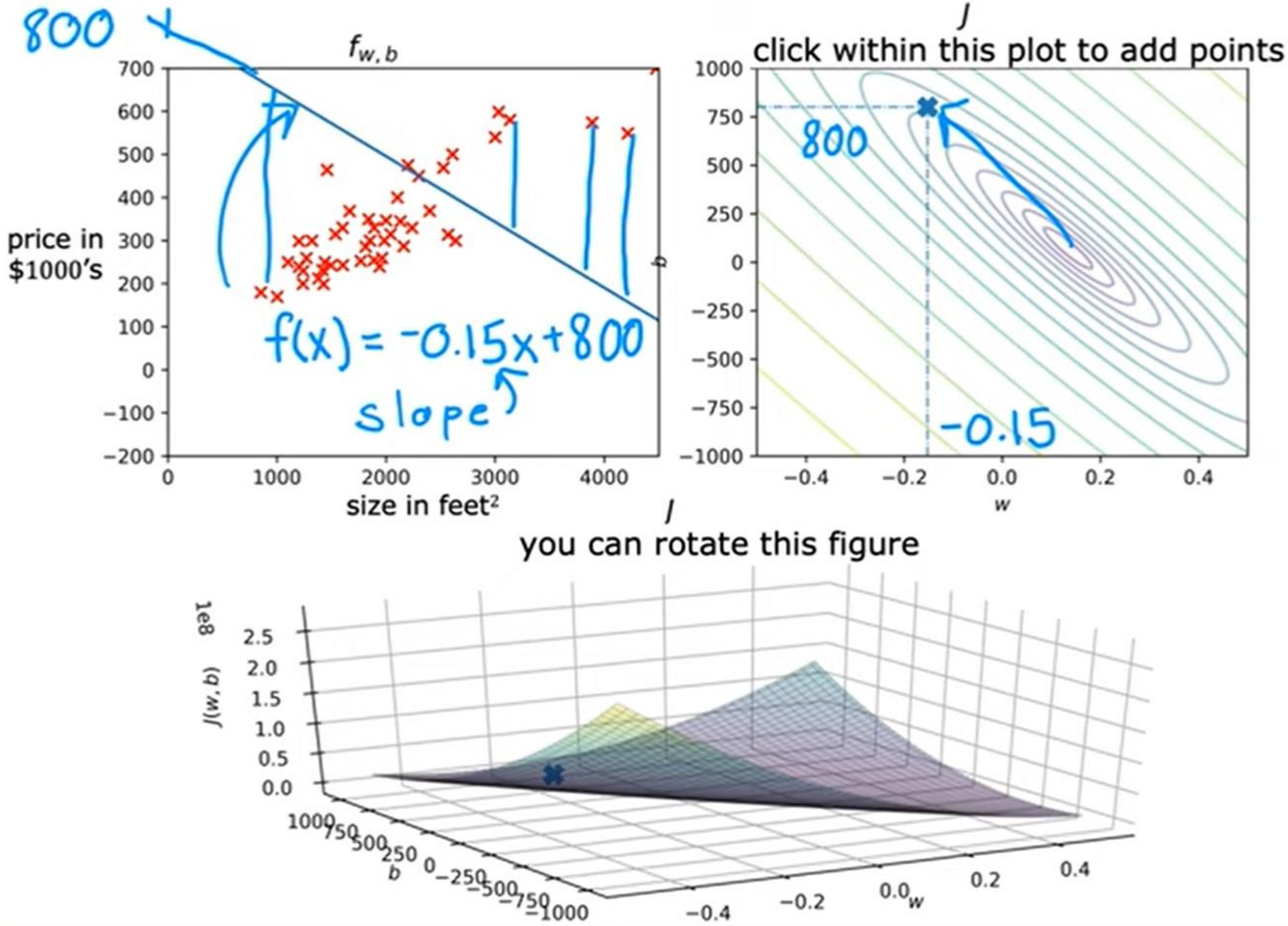

$f_{w,b}$
price in $1000's
size in feet²
700
600
500
400
300
200
100
0
−100
−200
0
1000
2000
3000
4000


J
click within this plot to add points
b
w
1000
750
500
250
0
−250
−500
−750
−1000
−0.4
−0.2
0.0
0.2
0.4
minimum
b
same J
w

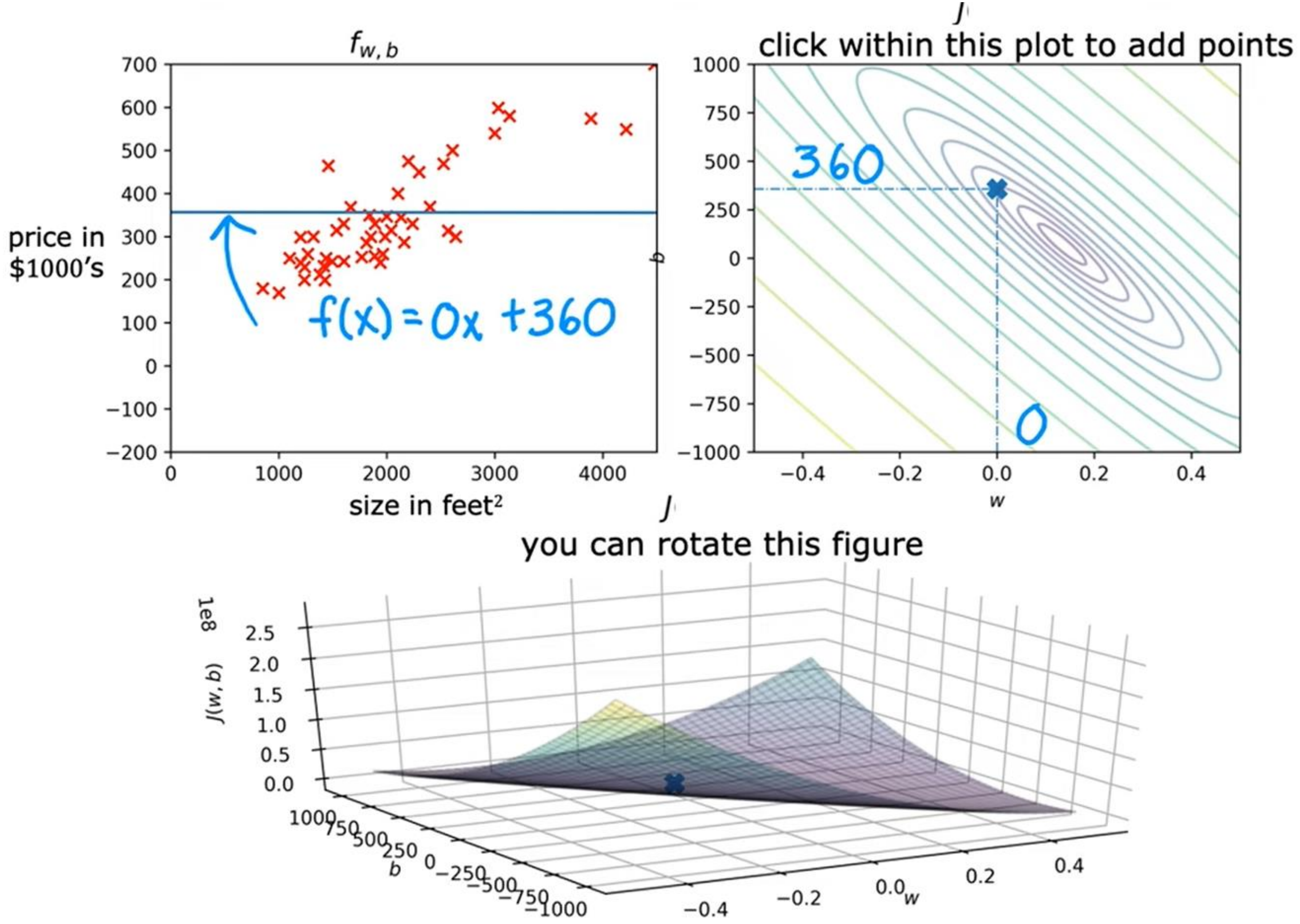

Mount Fuji
Legend
Google Earth


J
you can rotate this figure
1e8
J(w, b)
2.5
2.0
1.5
1.0
0.5
0.0
b
1000
750
500
250
0
−250
−500
−750
−1000
w
−0.4
−0.2
0.0
0.2
0.4
minimum

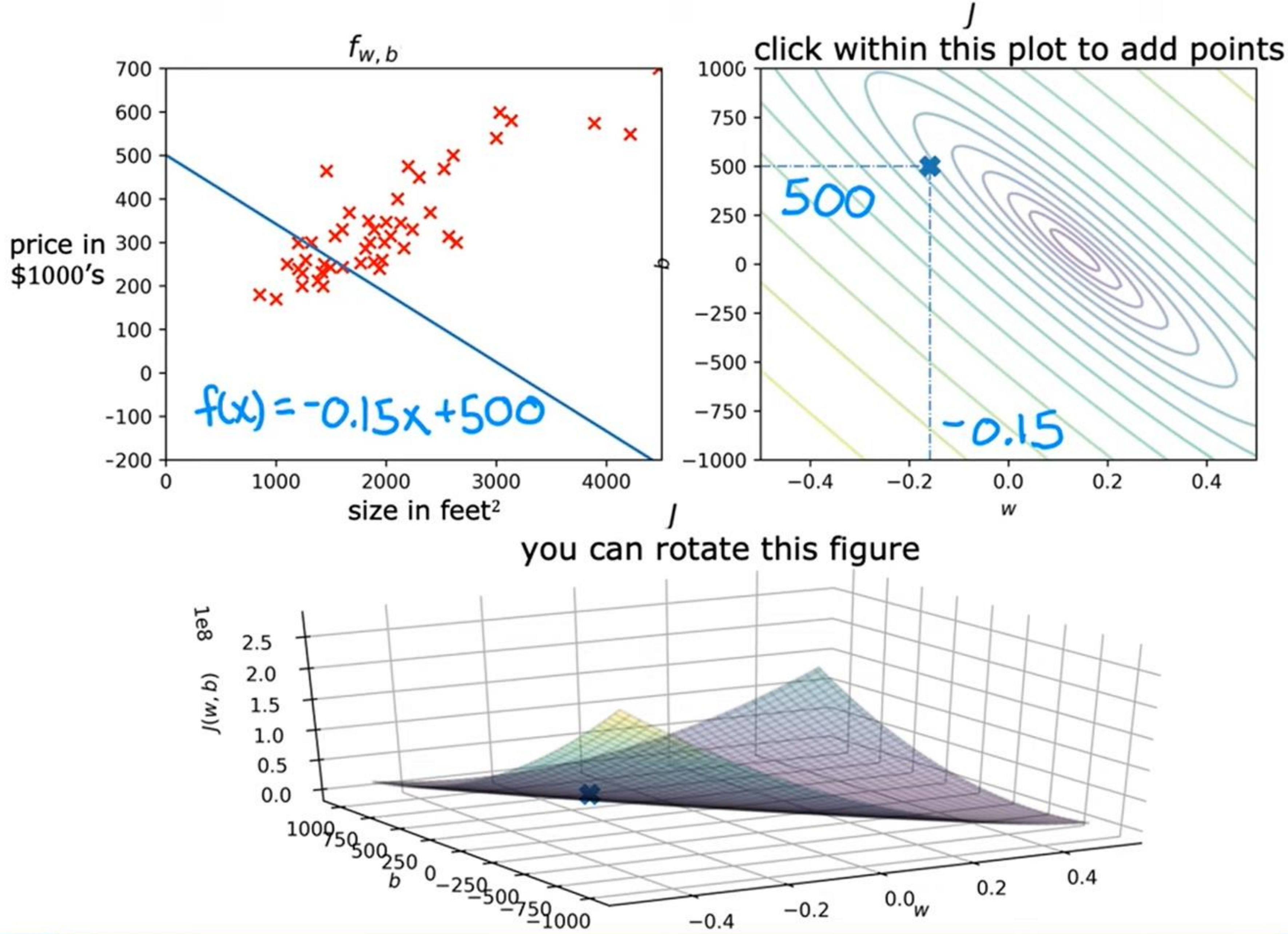

Mount Fuji
Legend
Mount Fuji
Google Earth



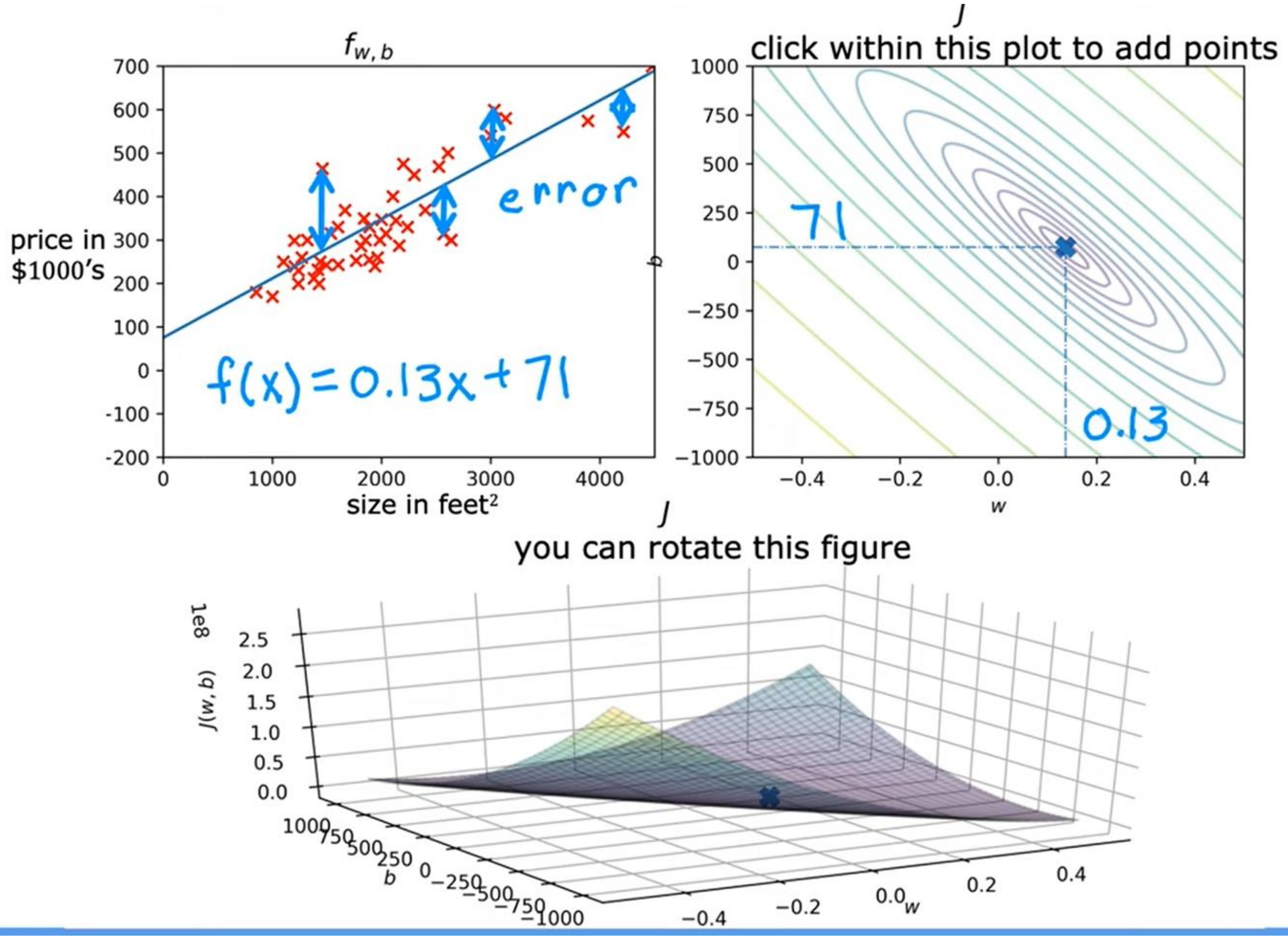
800
$f_{w,b}$
price in $1000's
size in feet²
f(x) = -0.15x + 800
slope
J
click within this plot to add points
800
-0.15
b
w
J
you can rotate this figure
1e8
J(w, b)
b
w



$f_{w,b}$
price in $1000's
size in feet²
f(x) = 0x +360
J
click within this plot to add points
360
0
b
w
J
you can rotate this figure
1e8
J(w, b)
b
w



$f_{w,b}$
price in $1000's
size in feet²
f(x) = -0.15x + 500
J
click within this plot to add points
500
-0.15
b
w
J
you can rotate this figure
1e8
J(w, b)
b
w



$f_{w,b}$
price in $1000's
size in feet^2
error
f(x)=0.13x+71
J
click within this plot to add points
71
0.13
b
w
J
you can rotate this figure
1e8
J(w, b)
b
w