California State University, Dominguez Hills
Department of Computer Science
CSC 595

Professor: Dr. Benyamin Ahmadnia
bahmadniayebosari@csudh.edu

Spring 2025

Copyright Notice

Table of Contents

Gradient Descent

Gradient Descent

Have some function $J(w,b)$ for linear regression or any function

Want $\min_{w,b} J(w,b)$ $\qquad \min_{w_1,\ldots,w_n,b} J(w_1,w_2,\ldots,w_n,b)$

Outline:

- Start with some w,b (set $w{=}0$, $b{=}0$)
- Keep changing w,b to reduce $J(w,b)$
- Until we settle at or near a minimum

may have >1 minimum

J not always

Implementing Gradient Descent

Gradient descent algorithm

$$w = w - \alpha \frac{\partial}{\partial w} J(w,b)$$

Implementing Gradient Descent

Gradient descent algorithm

$$w = w - \alpha \frac{d}{dw} J(w,b)$$

Assignment

$a = c$

$a = a + 1$

Code

Implementing Gradient Descent

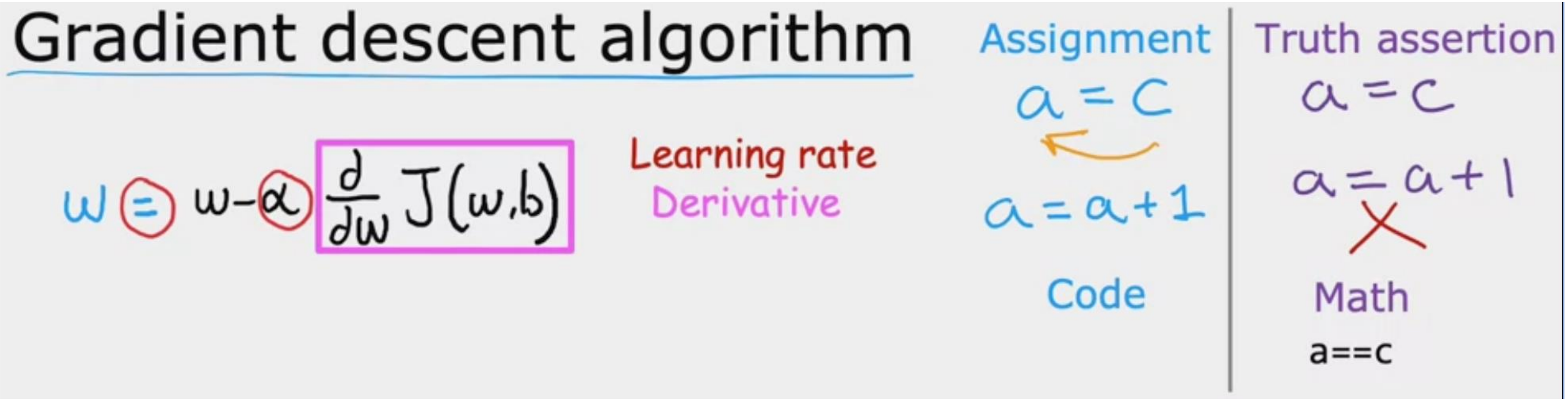
Gradient descent algorithm

$$w = w - \alpha \frac{\partial}{\partial w} J(w,b)$$

Learning rate

Assignment	Truth assertion
$a = c$	$a = c$
$a = a + 1$	$a = a + 1$
Code	Math
	a==c

Implementing Gradient Descent



Implementing Gradient Descent

Gradient descent algorithm

Repeat until convergence

$$w = w - \alpha \boxed{\frac{\partial}{\partial w} J(w,b)}$$

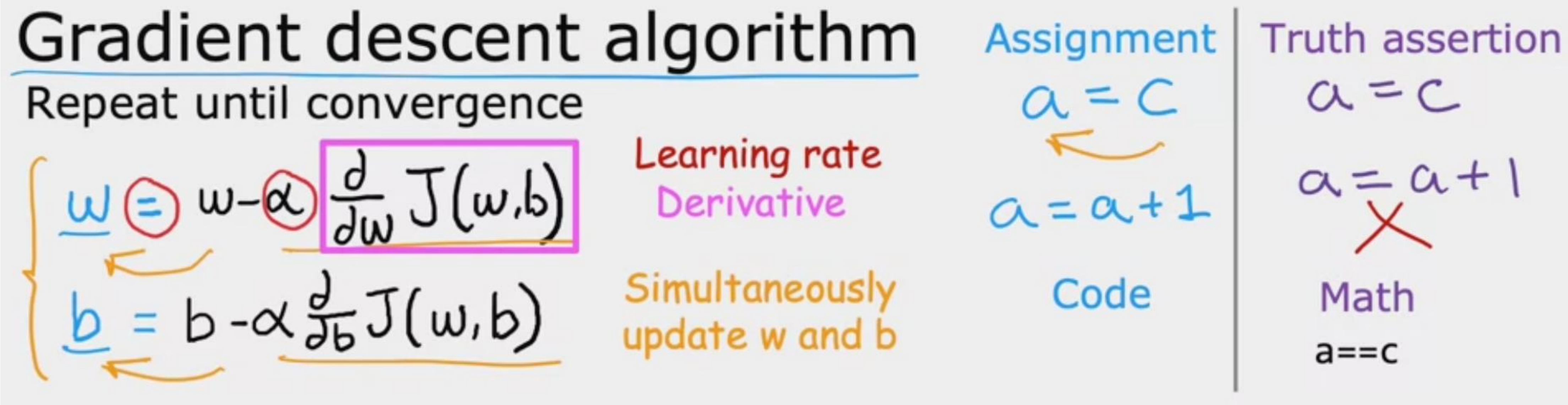
Learning rate

Derivative

$$b = b - \alpha \frac{\partial}{\partial b} J(w,b)$$

Assignment	Truth assertion
$a = c$	$a = c$
$a = a + 1$	$a = a + 1$
Code	Math
	a==c

Implementing Gradient Descent



Implementing Gradient Descent

Gradient descent algorithm

Repeat until convergence

$$
\begin{cases}
w = w - \alpha \boxed{\dfrac{\partial}{\partial w} J(w,b)} \\
b = b - \alpha \dfrac{\partial}{\partial b} J(w,b)
\end{cases}
$$

Learning rate

Derivative

Simultaneously update w and b

Assignment	Truth assertion
$a = c$	$a = c$
$a = a + 1$	$\cancel{a = a + 1}$
Code	Math
	a==c

Correct: Simultaneous update

$$
\left.
\begin{aligned}
tmp\_w &= w - \alpha \frac{\partial}{\partial w} J(w,b) \\
tmp\_b &= b - \alpha \frac{\partial}{\partial b} J(w,b)
\end{aligned}
\right\}
$$

$$w = tmp\_w$$

$$b = tmp\_b$$

Implementing Gradient Descent

Gradient descent algorithm

Repeat until convergence

$$\left\{ \begin{aligned} w &= w - \alpha \boxed{\frac{\partial}{\partial w} J(w,b)} \\ b &= b - \alpha \frac{\partial}{\partial b} J(w,b) \end{aligned} \right.$$

Learning rate

Derivative

Simultaneously update w and b

Assignment	Truth assertion
$a = c$	$a = c$
$a = a + 1$	$\cancel{a = a + 1}$
Code	Math
	a==c

Correct: Simultaneous update

$$\left. \begin{aligned} tmp\_w &= w - \alpha \frac{\partial}{\partial w} J(w,b) \\ tmp\_b &= b - \alpha \frac{\partial}{\partial b} J(w,b) \end{aligned} \right\}$$

$w = tmp\_w$

$b = tmp\_b$

Gradient Descent Intuition

Gradient descent algorithm

repeat until convergence {

learning rate

$$w = w - \alpha \boxed{\frac{\partial}{\partial w} J(w,b)}$$

derivative

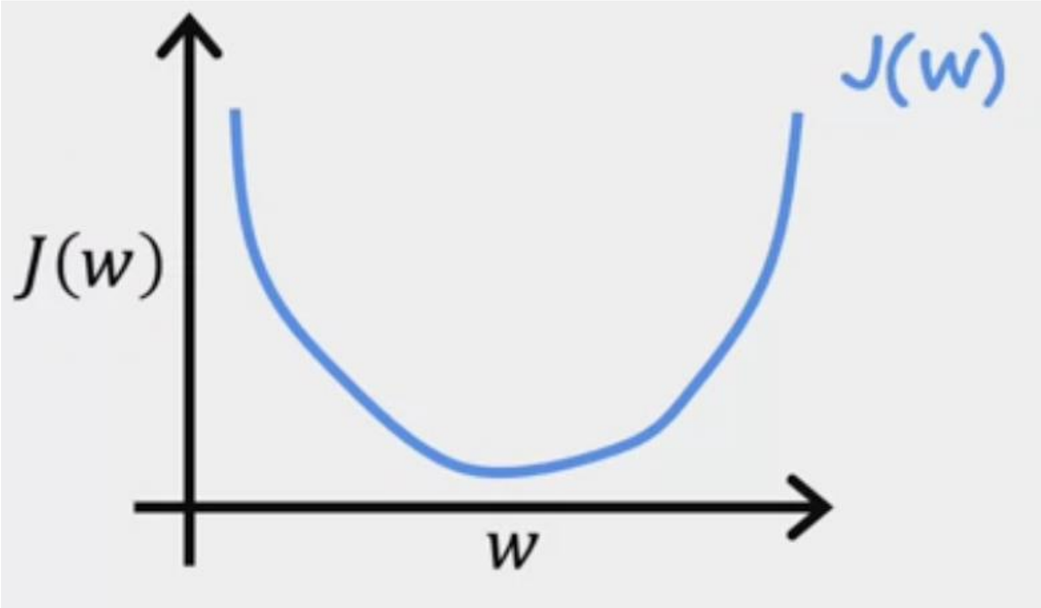
$$b = b - \alpha \frac{\partial}{\partial b} J(w,b)$$

$J(w)$

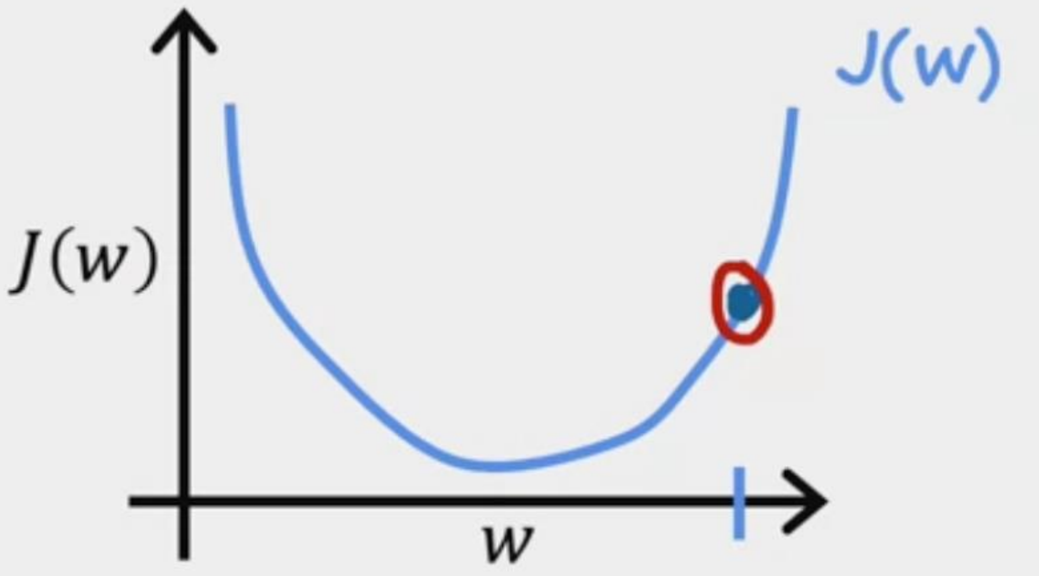
$$w = w - \alpha \frac{\partial}{\partial w} J(w)$$

$$\min_{w} J(w)$$

Gradient Descent Intuition

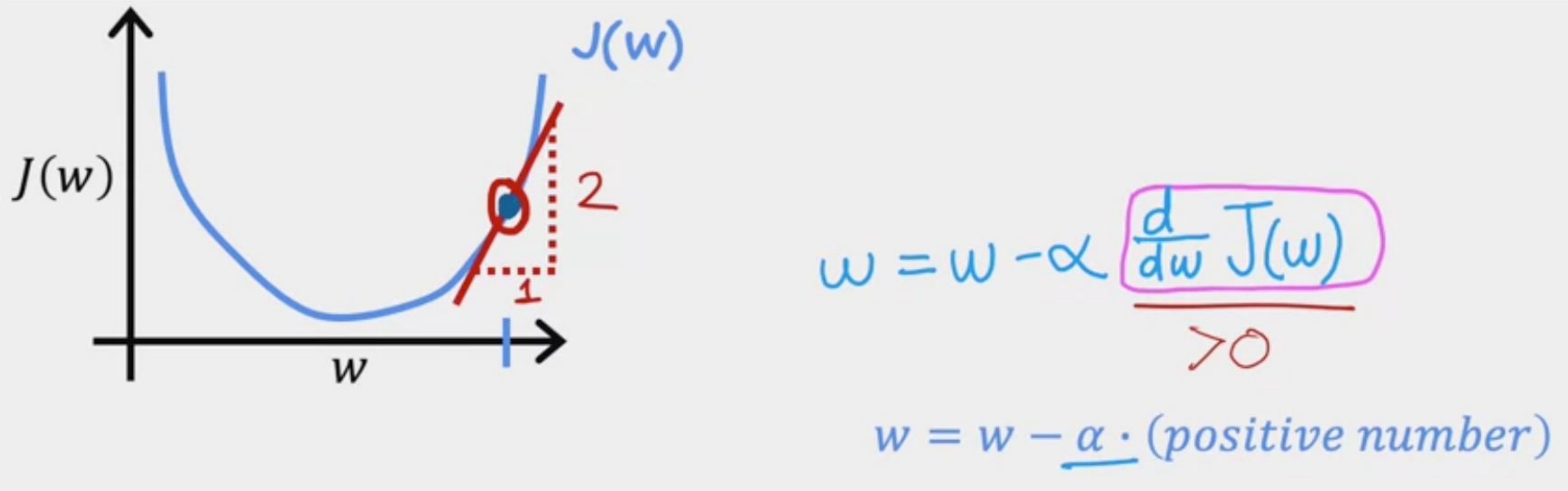


Gradient Descent Intuition

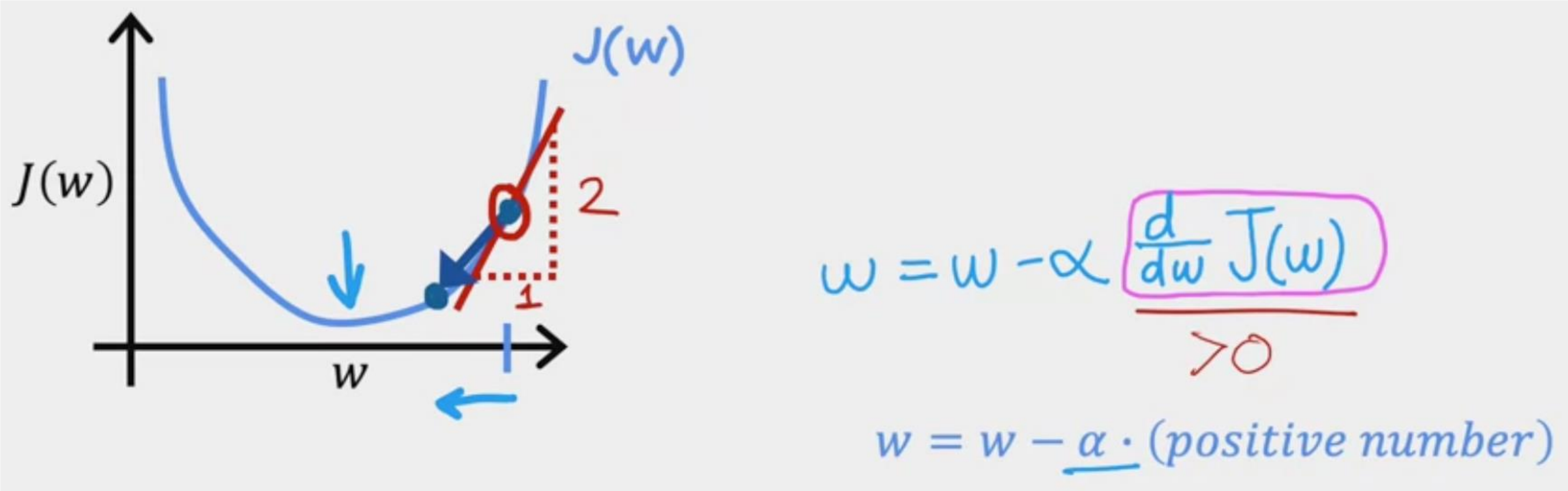


$$w = w - \alpha \boxed{\frac{d}{dw} J(w)}$$

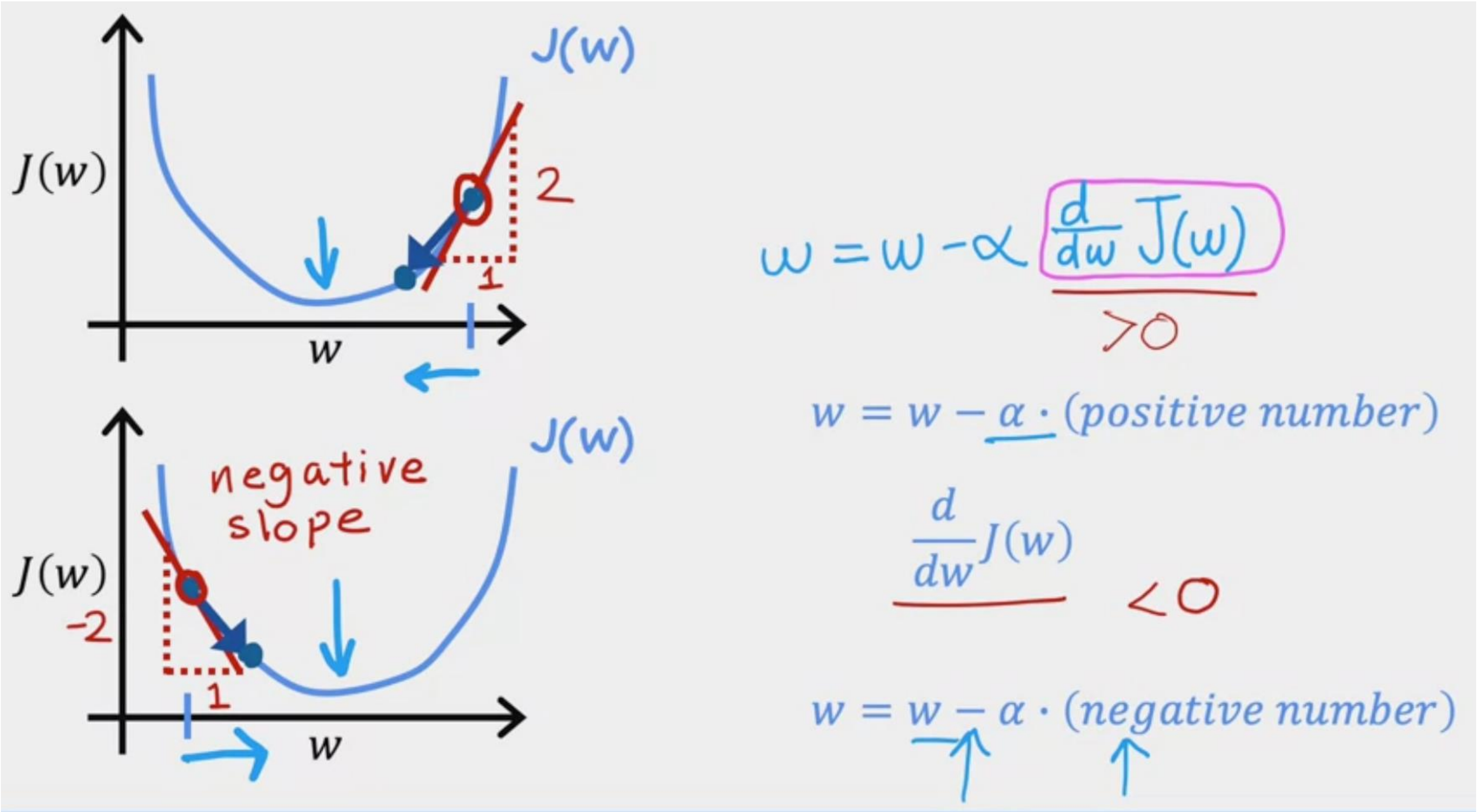
Gradient Descent Intuition



Gradient Descent Intuition



Gradient Descent Intuition



Classification with Logistic Regression

Classification – Motivation

Question	Answer "y"	
Is this email spam?	no	yes
Is the transaction fraudulent?	no	yes
Is the tumor malignant?	no	yes
	false	true
	0	1

useful for classification

y can only be one of two values

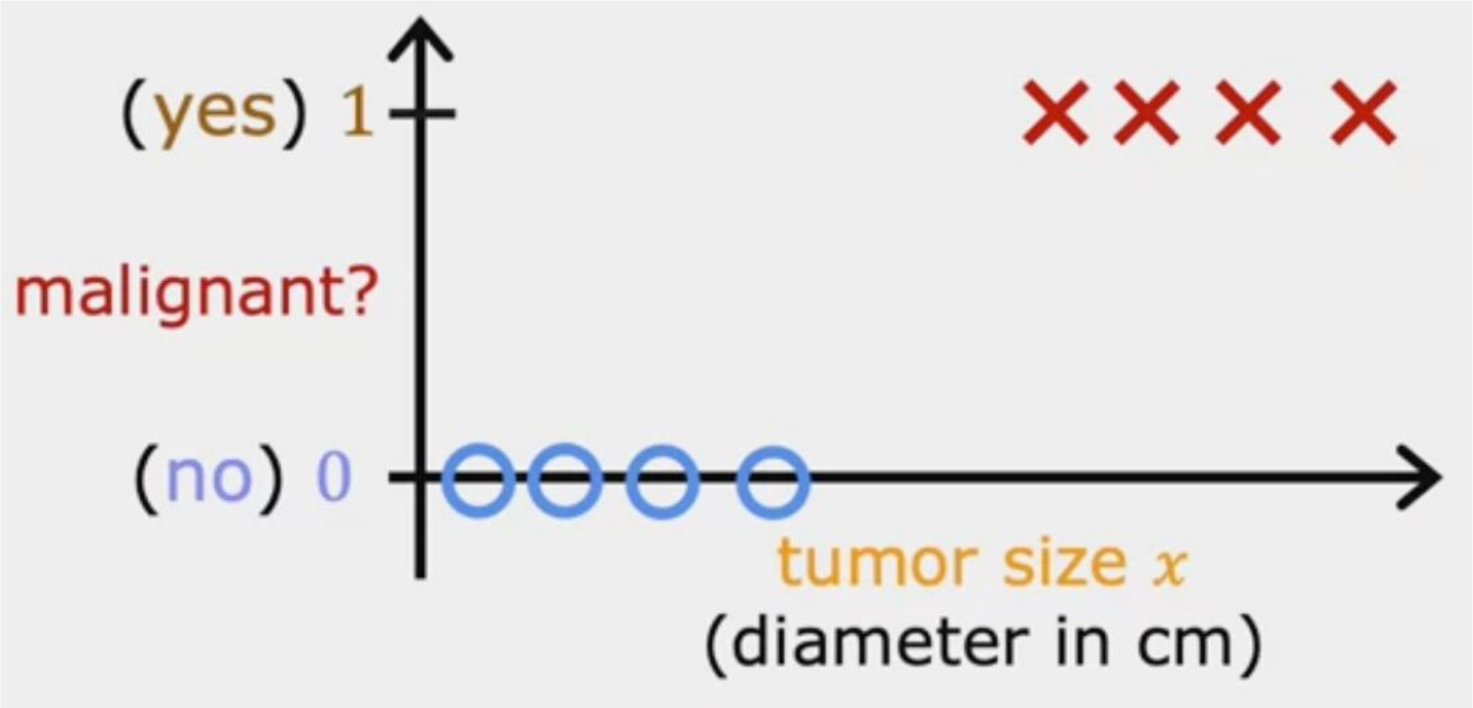
"binary classification"

class = category

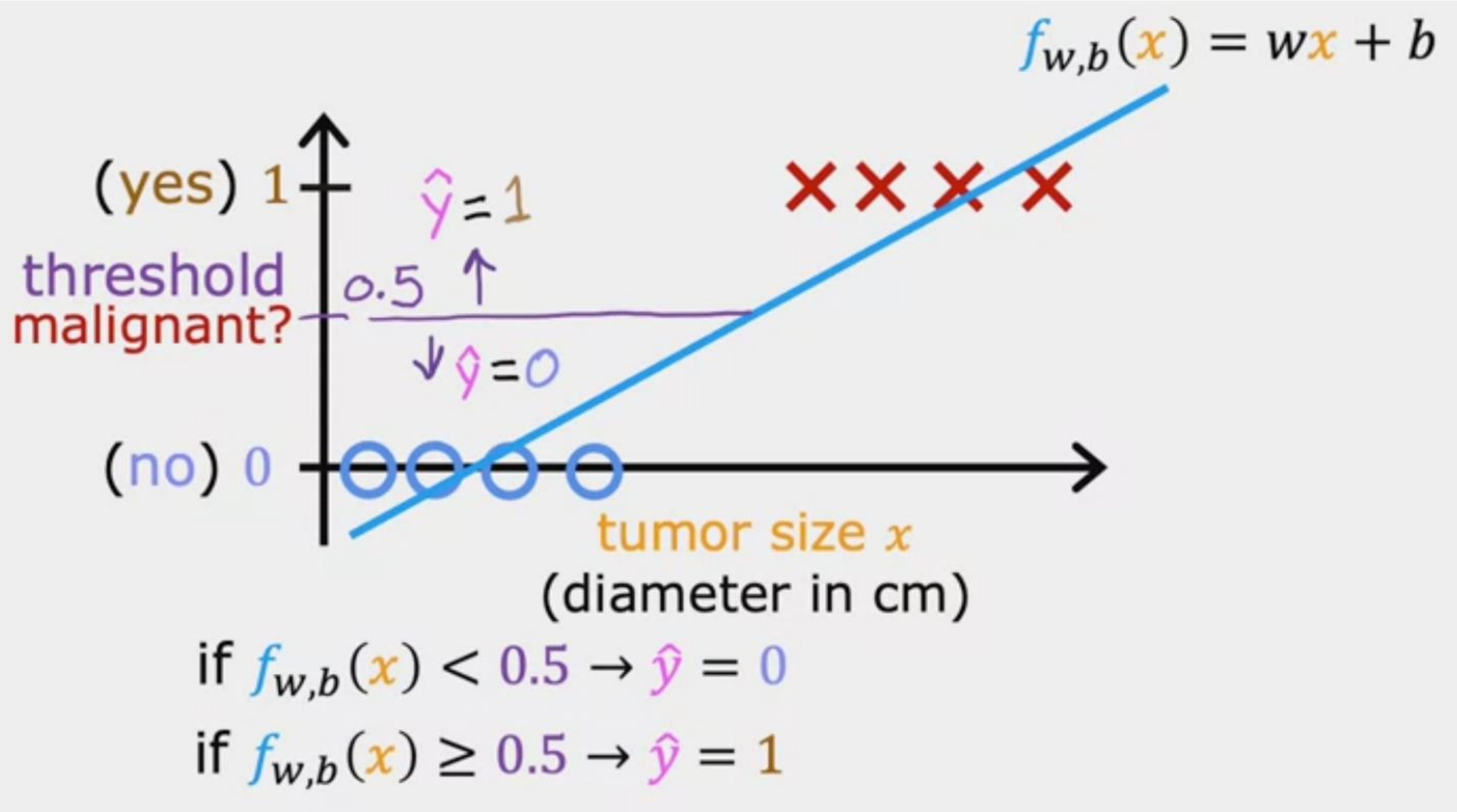
"negative class" ≠ "bad"
absence

"positive class" ≠ "good"
presence

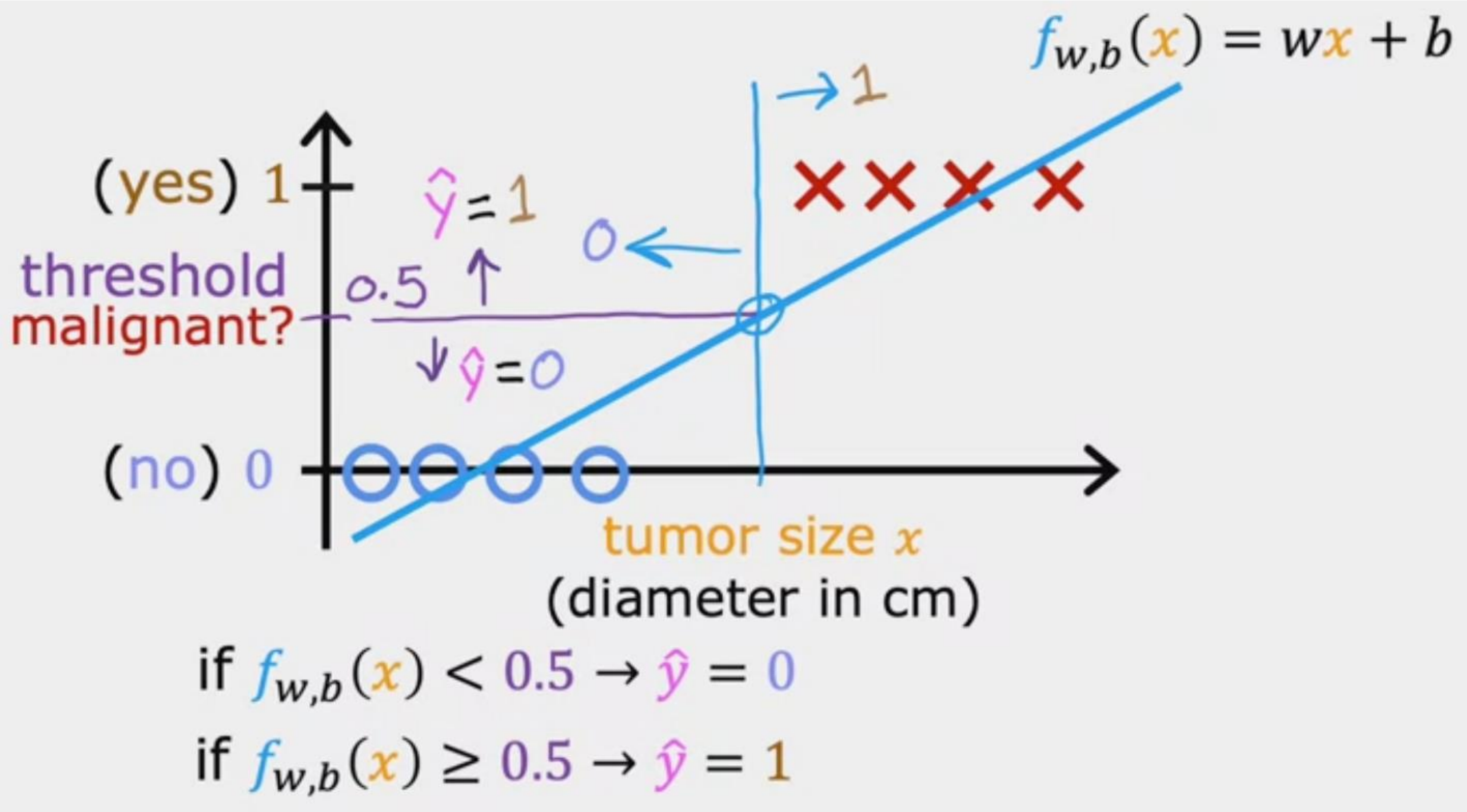
Classification – Motivation



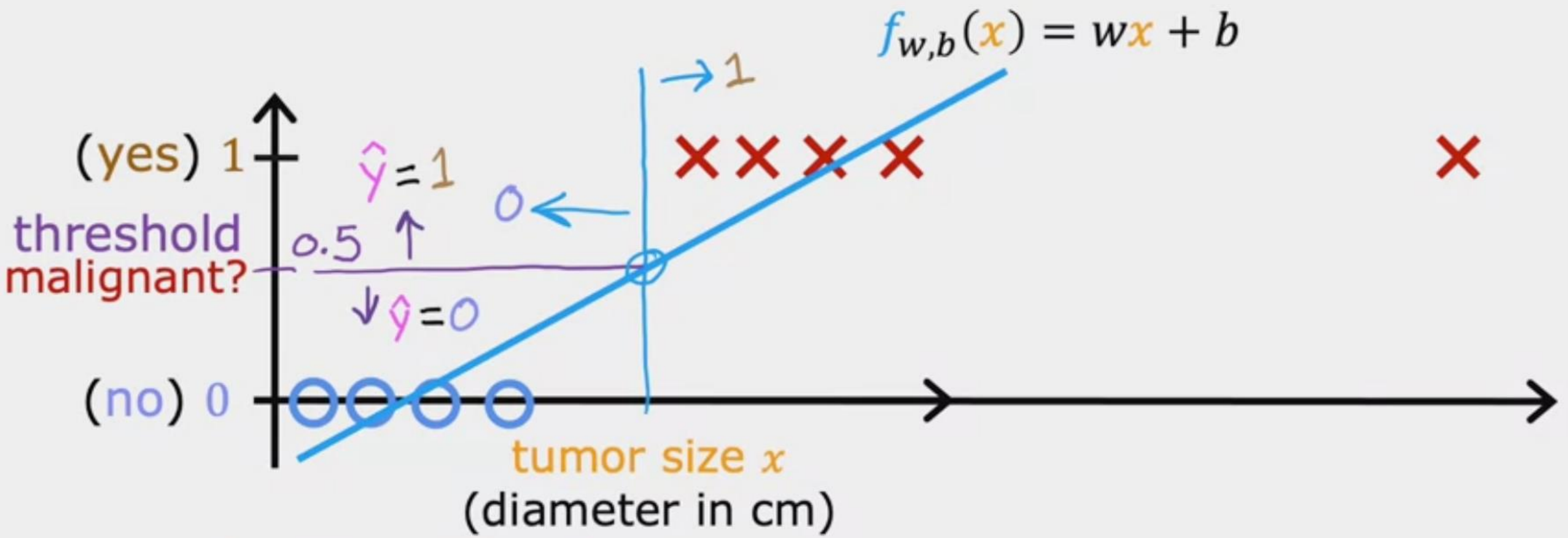
Classification – Motivation



Classification – Motivation



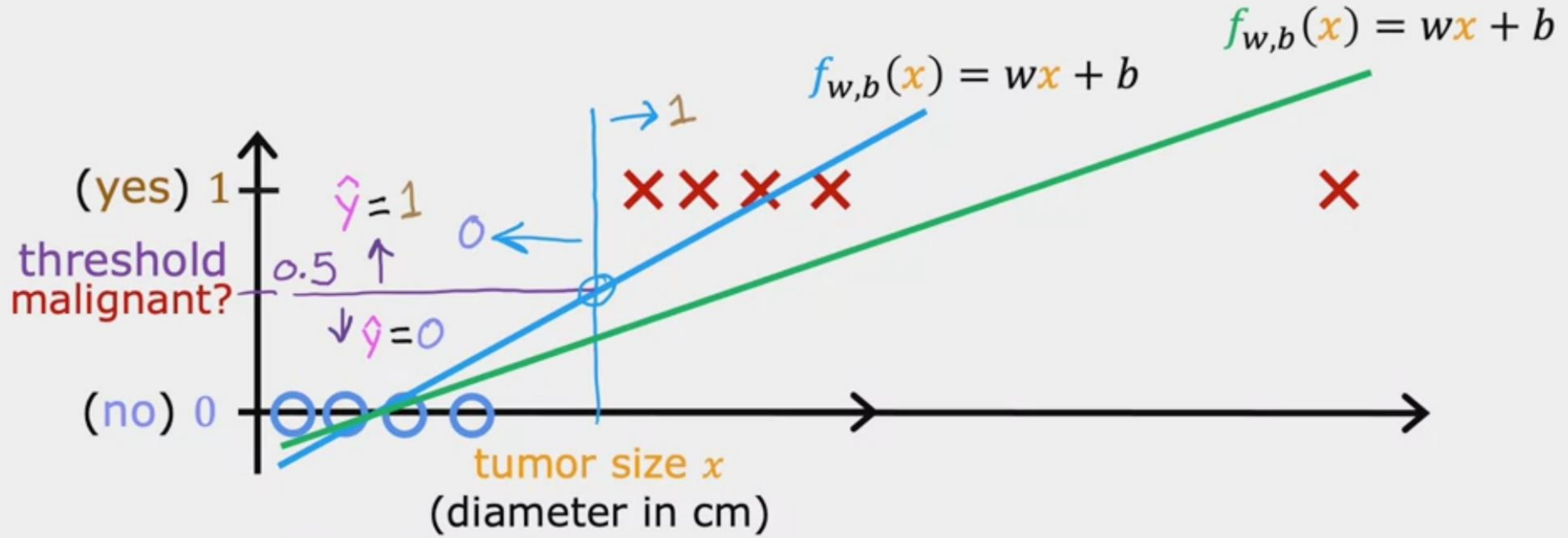
Classification – Motivation



if $f_{w,b}(x) < 0.5 \rightarrow \hat{y} = 0$

if $f_{w,b}(x) \geq 0.5 \rightarrow \hat{y} = 1$

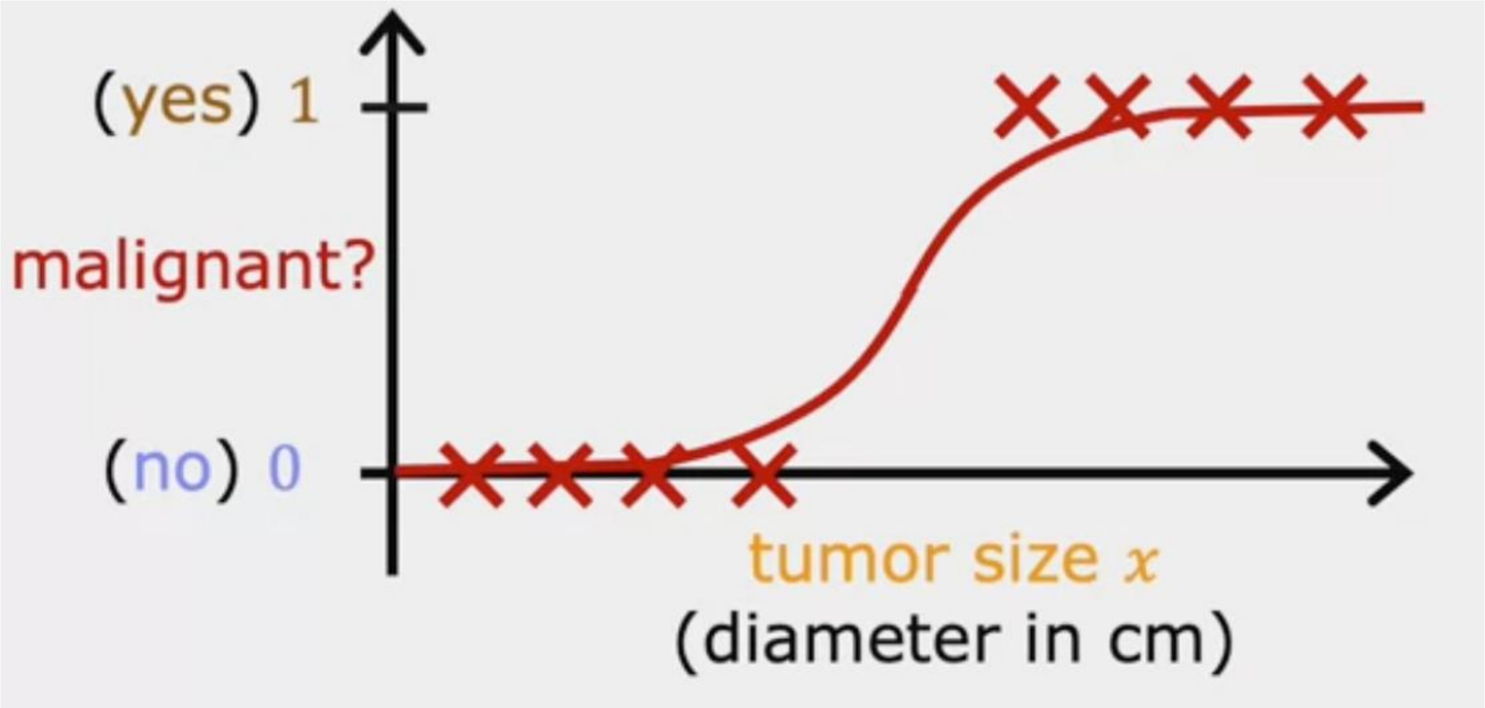
Classification – Motivation



if $f_{w,b}(x) < 0.5 \rightarrow \hat{y} = 0$

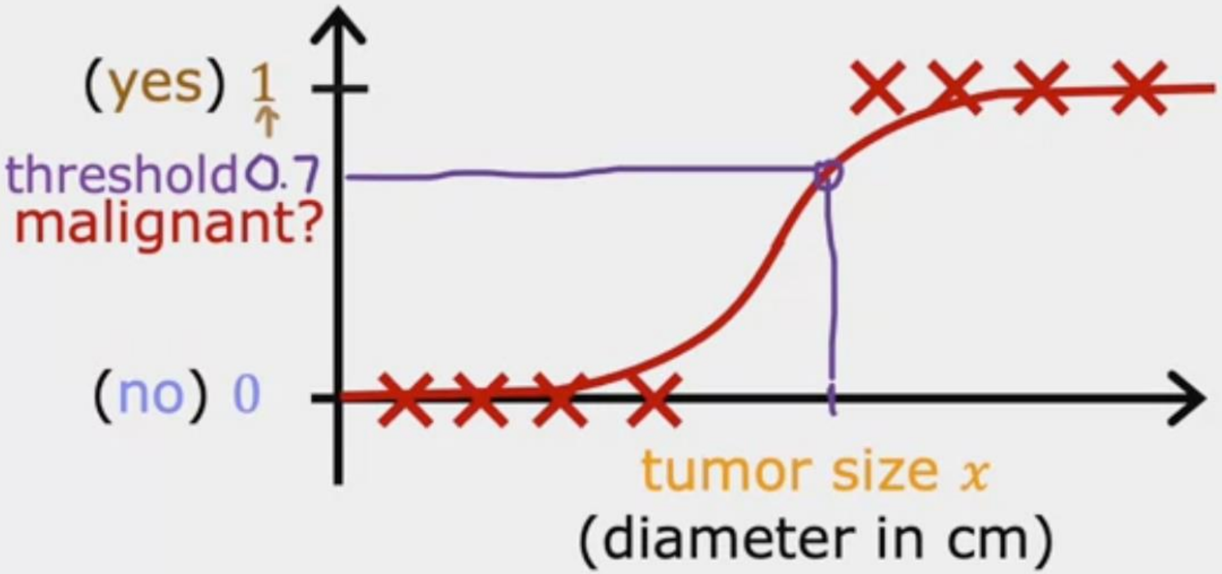
if $f_{w,b}(x) \geq 0.5 \rightarrow \hat{y} = 1$

Classification – Logistic Regression



Classification – Logistic Regression

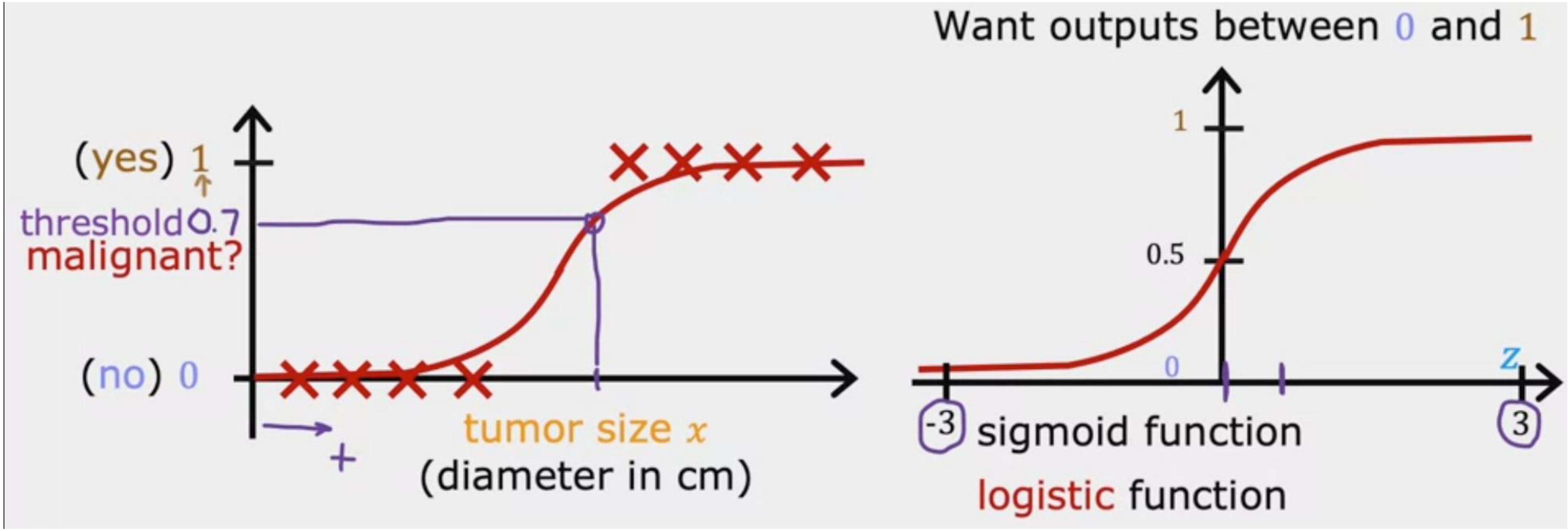
Want outputs between 0 and 1



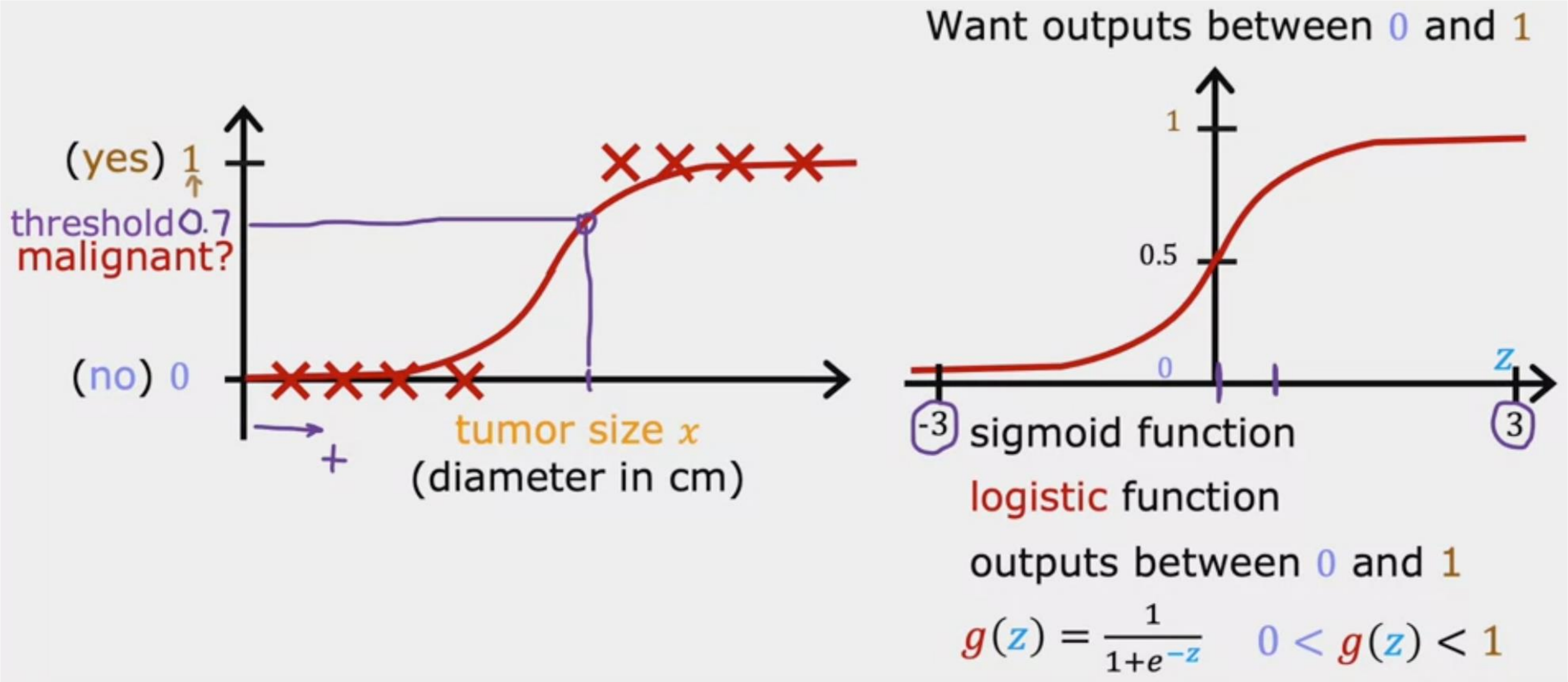
sigmoid function

logistic function

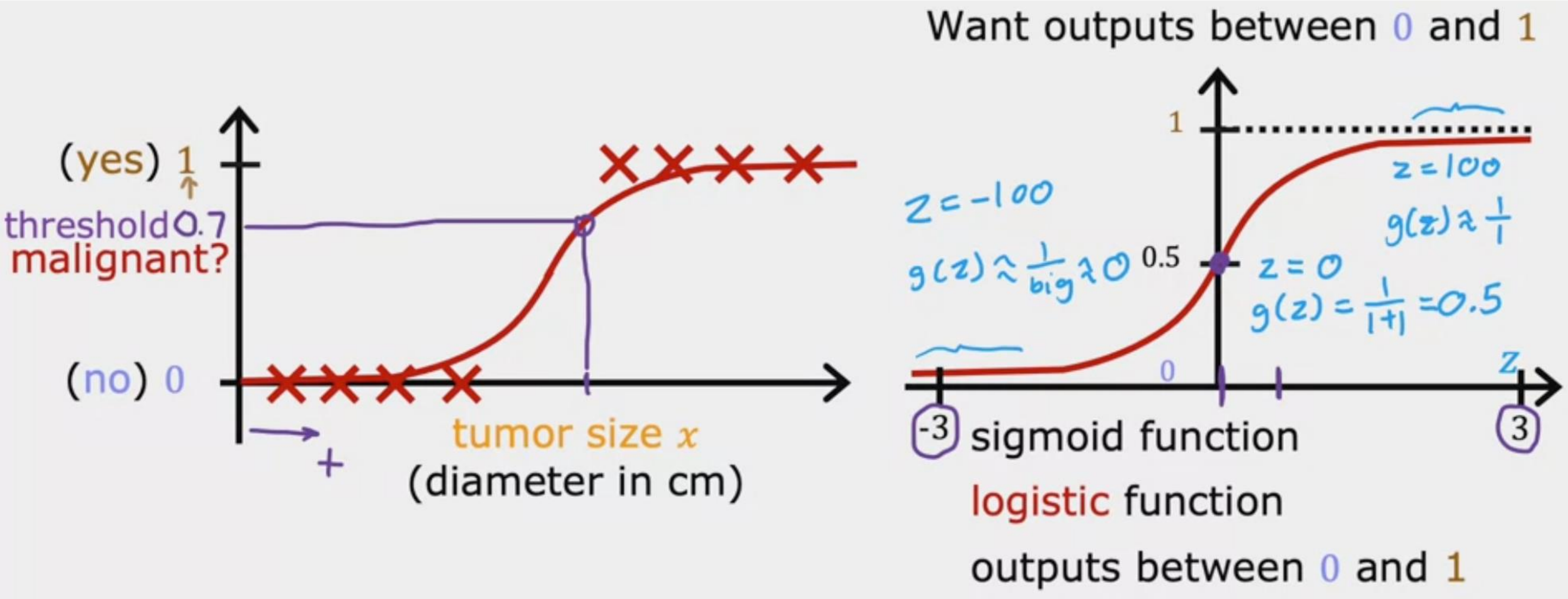
Classification – Logistic Regression



Classification – Logistic Regression

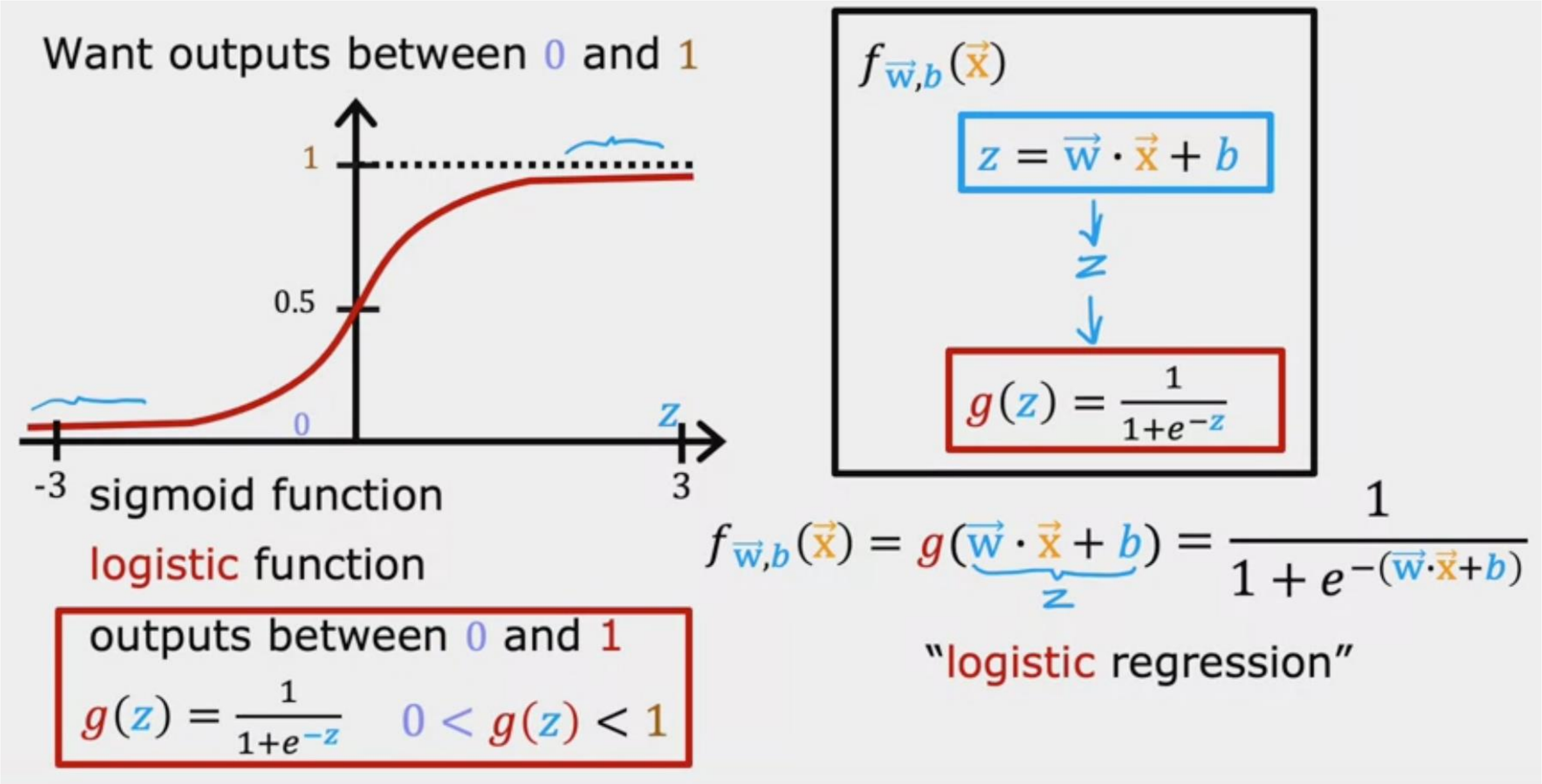


Classification – Logistic Regression



$$g(z) = \frac{1}{1+e^{-z}} \qquad 0 < g(z) < 1$$

Classification – Logistic Regression



Classification – Logistic Regression

Interpretation of logistic regression output

$$f_{\vec{w},b}(\vec{x}) = \frac{1}{1 + e^{-(\vec{w}\cdot\vec{x}+b)}}$$

"probability" that class is 1

Example:

x is "tumor size"
y is 0 (not malignant)
  or 1 (malignant)

$$f_{\vec{w},b}(\vec{x}) = 0.7$$

70% chance that y is 1

P(1)

$$f_{\vec{w},b}(\vec{x}) = P(y = 1|\vec{x}; \vec{w}, b)$$

Probability that y is 1,
given input $\vec{x}$, parameters $\vec{w}, b$

$$P(y = 0) + P(y = 1) = 1$$