

California State University, Dominguez Hills

Department of Computer Science

CSC 595

Professor: Dr. Benyamin Ahmadnia
bahmadniayebosari@csudh.edu

Spring 2025

Copyright Notice

These slides are distributed under the Creative Commons License.

[DeepLearning.AI](#) makes these slides available for educational purposes. You may not use or distribute these slides for commercial purposes. You may make copies of these slides and use or distribute them for educational purposes as long as you cite [DeepLearning.AI](#) as the source of the slides.

For the rest of the details of the license, see <https://creativecommons.org/licenses/by-sa/2.0/legalcode>

Table of Contents

- Cost Function
- Gradient Descent
- Classification with Logistic Regression
- Cost Function for Logistic Regression
- Gradient Descent for Logistic Regression

Cost Function for Logistic Regression

Training Set

	tumor size (cm) x_1	...	patient's age x_n	malignant? y	$i = 1, \dots, m \leftarrow$ training examples $j = 1, \dots, n \leftarrow$ features
$i=1$	10		52	1	<div>target y is 0 or 1</div> $f_{\vec{w}, b}(\vec{x}) = \frac{1}{1 + e^{-(\vec{w} \cdot \vec{x} + b)}}$
\vdots	2		73	0	
\vdots	5		55	0	
	12		49	1	
$i=m$	

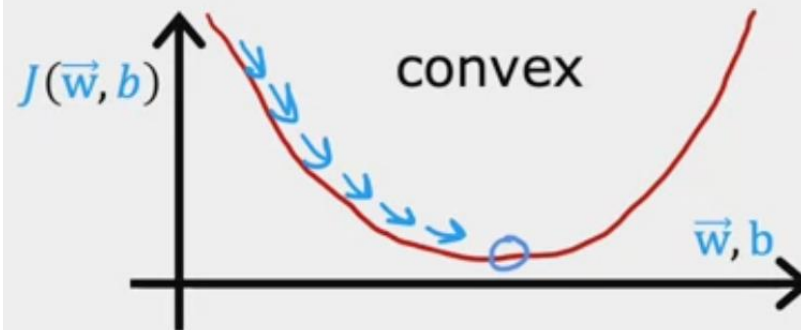
How to choose $\vec{w} = [w_1 \ w_2 \ \dots \ w_n]$ and b ?

Squared Error Cost

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

linear regression

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

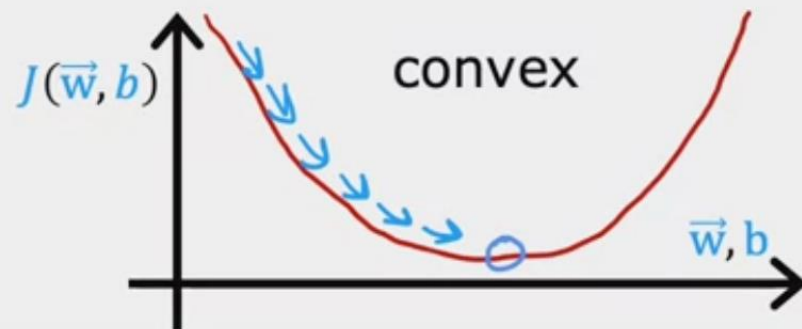


Squared Error Cost

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

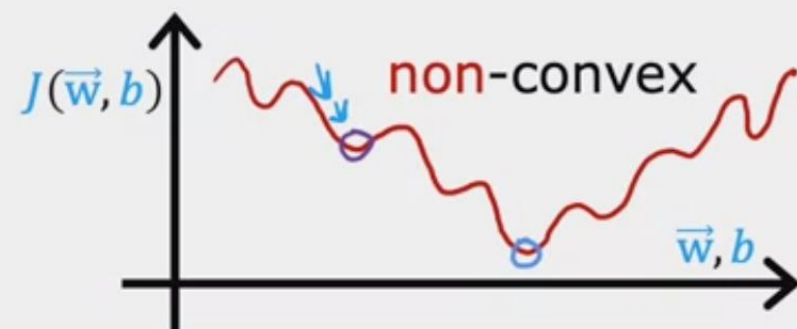
linear regression

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



logistic regression

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

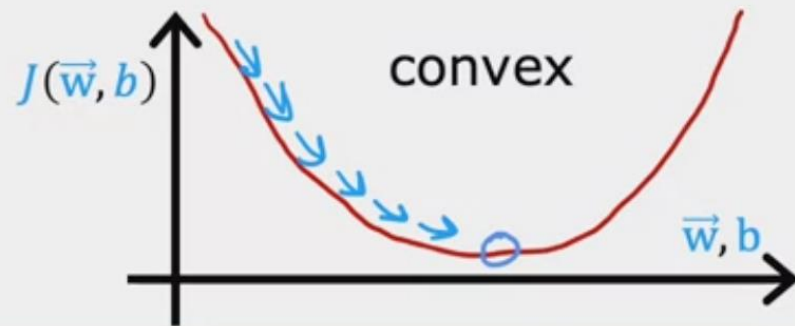


Squared Error Cost

$$J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^m \underbrace{\frac{1}{2} (f_{\vec{w}, b}(\vec{x}^{(i)}) - y^{(i)})^2}_{\text{loss } L(f_{\vec{w}, b}(\vec{x}^{(i)}), y^{(i)})}$$

linear regression

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



logistic regression

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

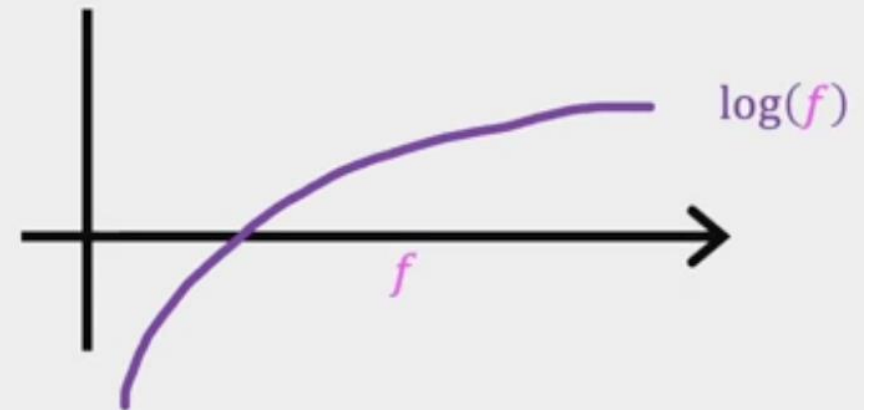


Logistic Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

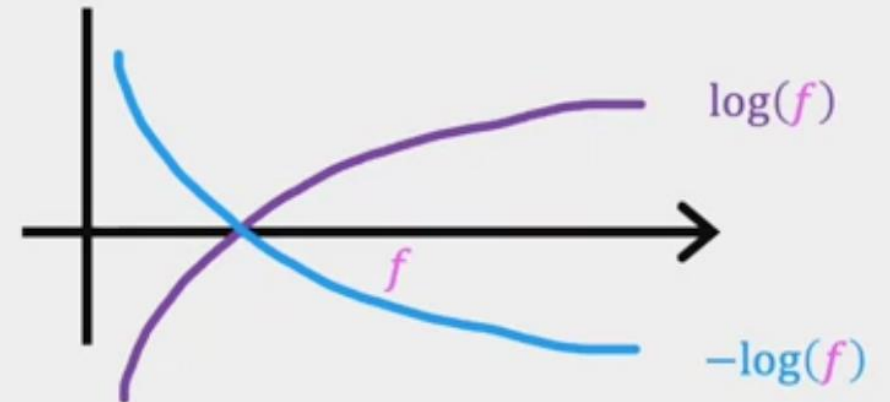
Logistic Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$



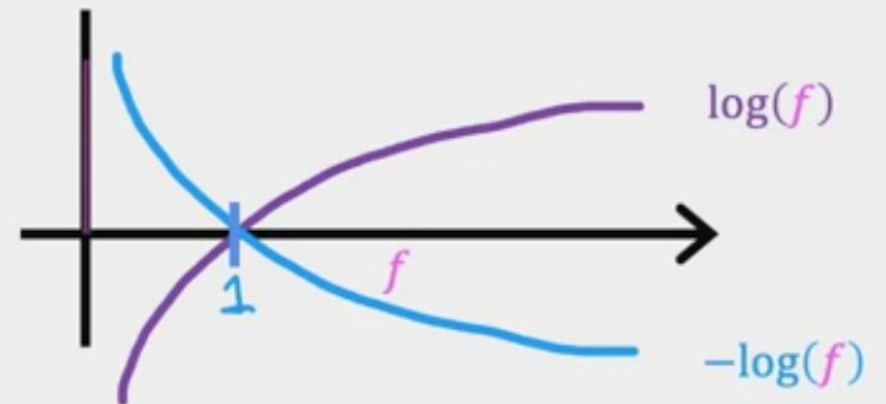
Logistic Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$



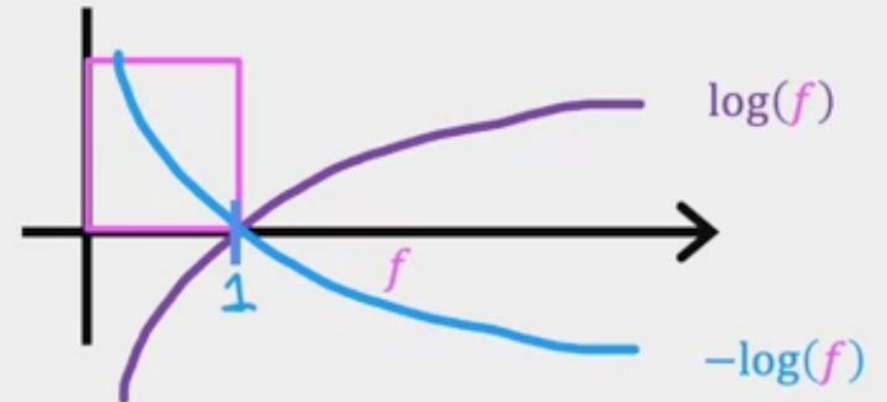
Logistic Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$



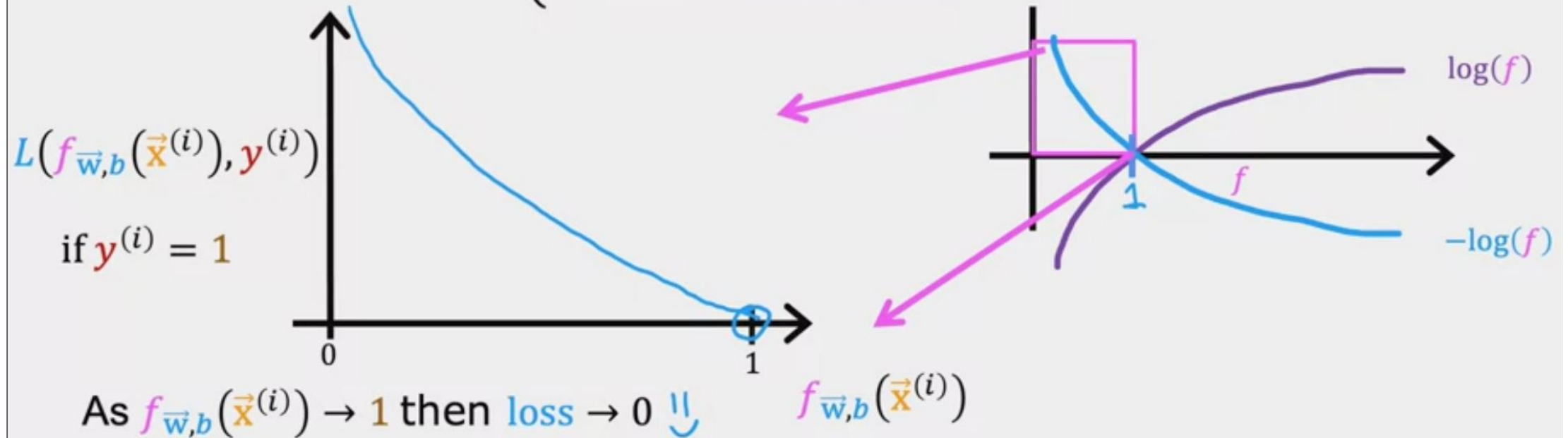
Logistic Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$



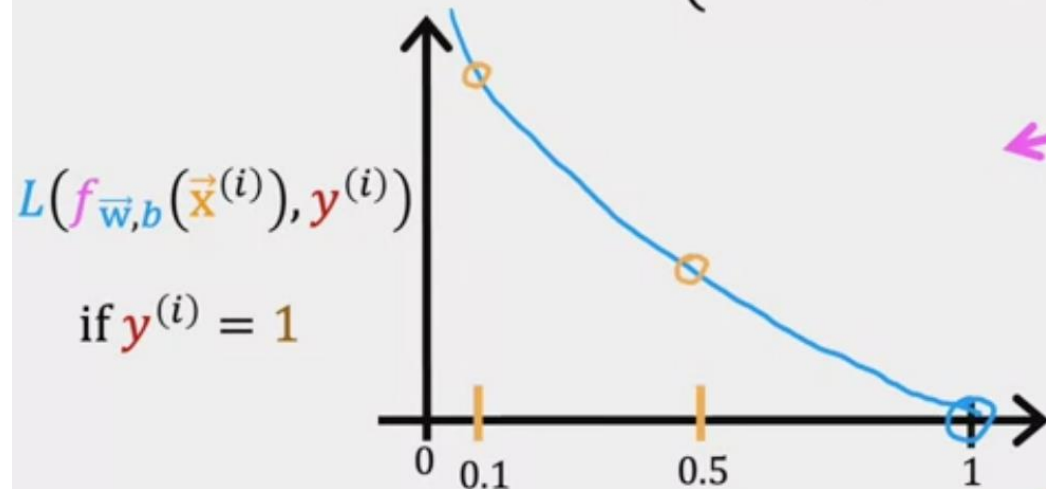
Logistic Loss Function


$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$




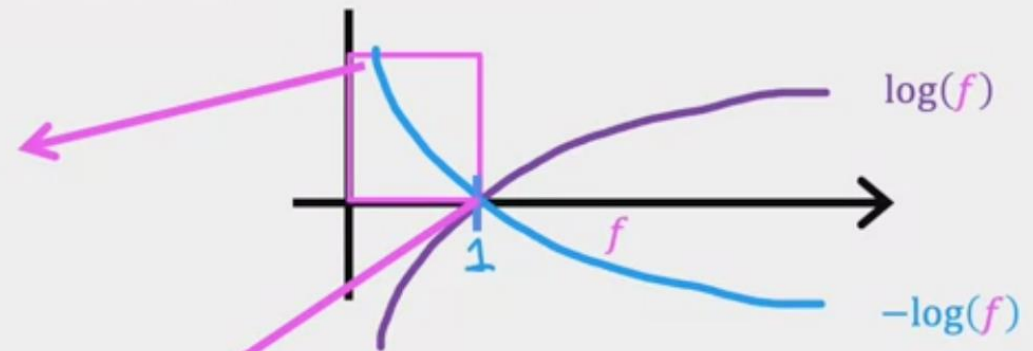
Logistic Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$



As $f_{\vec{w},b}(\vec{x}^{(i)}) \rightarrow 1$ then loss $\rightarrow 0$ 

As $f_{\vec{w},b}(\vec{x}^{(i)}) \rightarrow 0$ then loss $\rightarrow \infty$ 



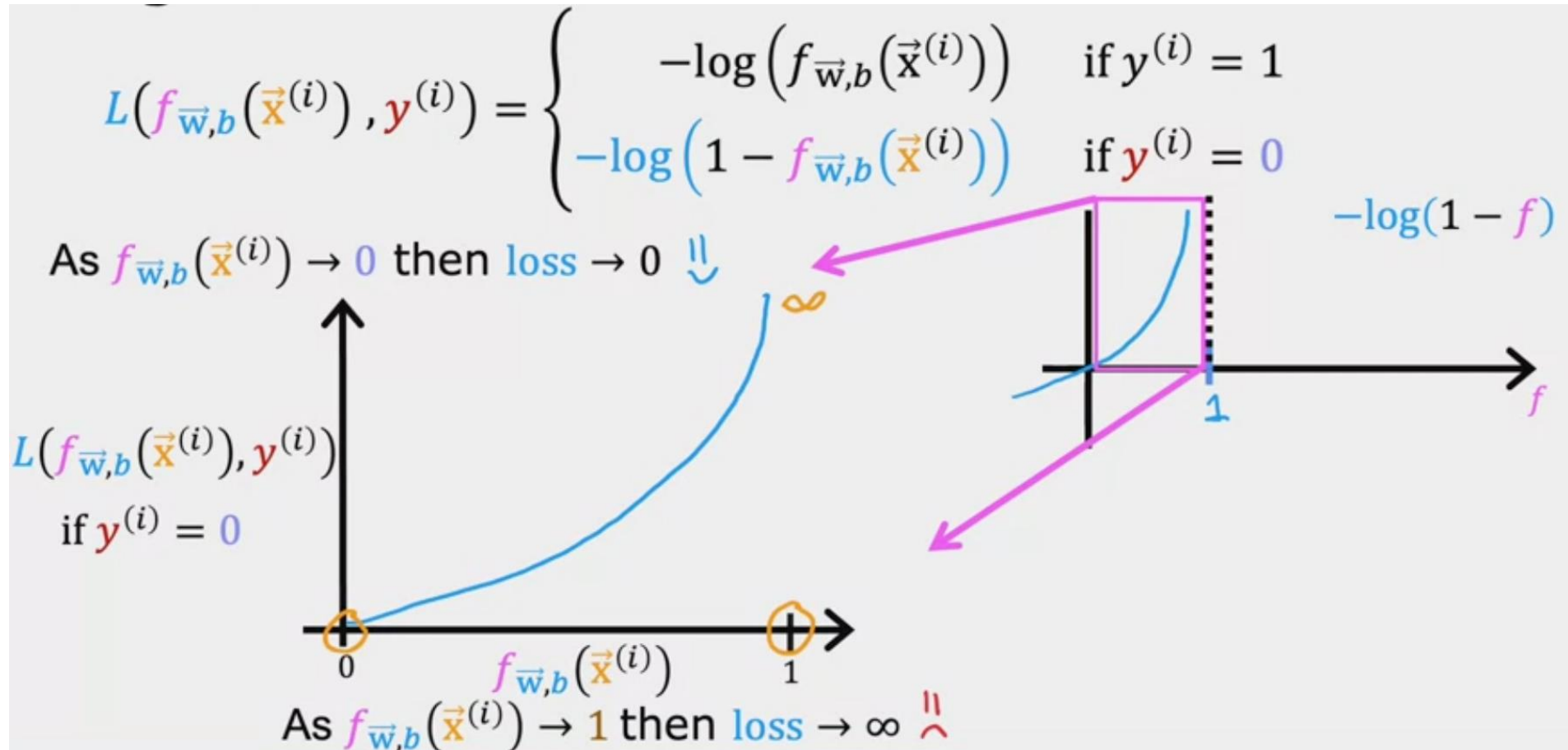
Loss is lowest when $f_{\vec{w},b}(\vec{x}^{(i)})$ predicts close to true label $y^{(i)}$.

Logistic Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

A graph illustrating the component of the logistic loss function, $-\log(1-f)$, plotted against f . The horizontal axis is labeled f and the vertical axis is unlabeled. A blue curve represents the function $-\log(1-f)$, which is increasing and concave up. A vertical dashed line is drawn at $f=1$, where the function approaches infinity. The label $-\log(1-f)$ is placed near the curve.

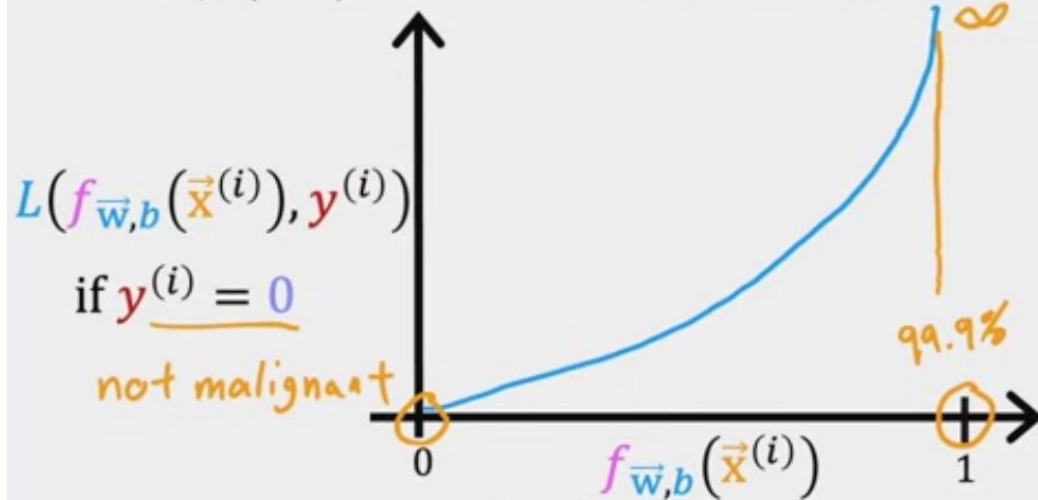
Logistic Loss Function



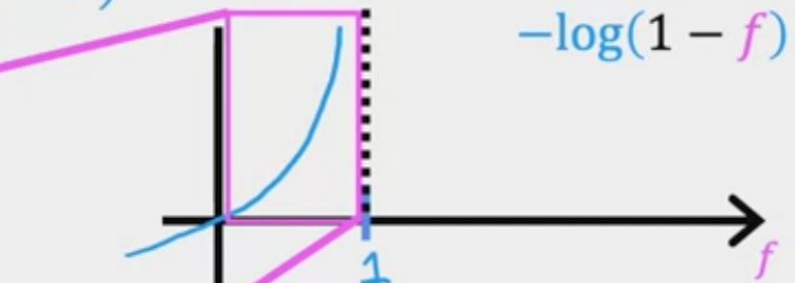
Logistic Loss Function

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\bar{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\bar{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

As $f_{\bar{w},b}(\vec{x}^{(i)}) \rightarrow 0$ then $\text{loss} \rightarrow 0$ \Downarrow



As $f_{\bar{w},b}(\vec{x}^{(i)}) \rightarrow 1$ then $\text{loss} \rightarrow \infty$ \Uparrow



The further prediction $f_{\bar{w},b}(\vec{x}^{(i)})$ is from target $y^{(i)}$, the higher the loss.

Cost

cost

$$J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^m L(\underbrace{f_{\vec{w}, b}(\vec{x}^{(i)})}_{\text{loss}}, y^{(i)})$$

$$= \begin{cases} -\log(f_{\vec{w}, b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w}, b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

if $y^{(i)} = 1$ convex \rightarrow can reach a global minimum
if $y^{(i)} = 0$ global minimum

find w, b that minimize cost J

Simplified Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = -y^{(i)}\log(f_{\vec{w},b}(\vec{x}^{(i)})) - (1 - y^{(i)})\log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))$$

Simplified Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = - \underbrace{y^{(i)}}_1 \log(f_{\vec{w},b}(\vec{x}^{(i)})) - \underbrace{(1 - y^{(i)})}_0 \log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))$$

if $y^{(i)} = 1$:

Simplified Loss Function

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\bar{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\bar{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = -\underbrace{y^{(i)}}_1 \log(f_{\bar{w},b}(\vec{x}^{(i)})) - \underbrace{(1 - y^{(i)})}_0 \log(1 - f_{\bar{w},b}(\vec{x}^{(i)}))$$

if $y^{(i)} = 1$:

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = -1 \log(f(\hat{x}))$$

Simplified Loss Function

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\bar{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\bar{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = -y^{(i)} \log(f_{\bar{w},b}(\vec{x}^{(i)})) - (1 - y^{(i)}) \log(1 - f_{\bar{w},b}(\vec{x}^{(i)}))$$

if $y^{(i)} = 1$:

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = \underbrace{-1}_{1} \log(f_{\bar{w},b}(\vec{x}^{(i)})) - \underbrace{(1-1)}_0 \log(1 - f_{\bar{w},b}(\vec{x}^{(i)}))$$

$$L(f_{\bar{w},b}(\vec{x}^{(i)}), y^{(i)}) = -1 \log(f_{\bar{w},b}(\vec{x}^{(i)}))$$

Simplified Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = - \underbrace{y^{(i)}}_0 \log(\cancel{f_{\vec{w},b}(\vec{x}^{(i)})}) - (1 - \underbrace{y^{(i)}}_{(1-0)}) \log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))$$

if $y^{(i)} = 1$:

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = -1 \log(f(\hat{x}))$$

if $y^{(i)} = 0$:

Simplified Loss Function

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{\vec{w},b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \underbrace{-y^{(i)}\log(f_{\vec{w},b}(\vec{x}^{(i)}))}_0 - \underbrace{(1 - y^{(i)})\log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))}_{(1-0)}$$

if $y^{(i)} = 1$:

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = -1 \log(f(\vec{x}))$$

if $y^{(i)} = 0$:

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = \underline{-(1-0)\log(1-f(\vec{x}))}$$

Simplified Loss Function

loss

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = -y^{(i)}\log(f_{\vec{w},b}(\vec{x}^{(i)})) - (1 - y^{(i)})\log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))$$

cost

$$J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^m [L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)})]$$

Simplified Loss Function

loss

$$L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = -y^{(i)}\log(f_{\vec{w},b}(\vec{x}^{(i)})) - (1 - y^{(i)})\log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))$$

cost

$$J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^m [L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)})]$$

Simplified Loss Function

$$\text{loss} \\ L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = - \underbrace{y^{(i)} \log(f_{\vec{w},b}(\vec{x}^{(i)})) + (1 - y^{(i)}) \log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))}_{\text{cost}}$$

$$\text{cost} \\ J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^m [L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)})]$$

$$= - \frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(f_{\vec{w},b}(\vec{x}^{(i)})) + (1 - y^{(i)}) \log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))]$$

Simplified Loss Function

$$\text{loss} \\ L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)}) = - \underbrace{y^{(i)} \log(f_{\vec{w},b}(\vec{x}^{(i)})) + (1 - y^{(i)}) \log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))}_{\text{cost}}$$

$$\text{cost} \\ J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^m [L(f_{\vec{w},b}(\vec{x}^{(i)}), y^{(i)})]$$

$$= - \frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(f_{\vec{w},b}(\vec{x}^{(i)})) + (1 - y^{(i)}) \log(1 - f_{\vec{w},b}(\vec{x}^{(i)}))]$$

maximum likelihood

Evaluate Yourself!

For the simplified loss function:

$$L(f_{\mathbf{w},b}(\mathbf{x}^{(i)}), y^{(i)}) = -y^{(i)}\log(f_{\mathbf{w},b}(\mathbf{x}^{(i)})) - (1 - y^{(i)})\log(1 - f_{\mathbf{w},b}(\mathbf{x}^{(i)}))$$

if the target $y^{(i)} = 1$, then what does this expression simplify to?

- ☐ $-\log(f_{\vec{w},b}(\mathbf{x}^{(i)}))$
- ☐ $-\log(1 - f_{\vec{w},b}(\mathbf{x}^{(i)}))$