Day-21, Oct-23, 2024.

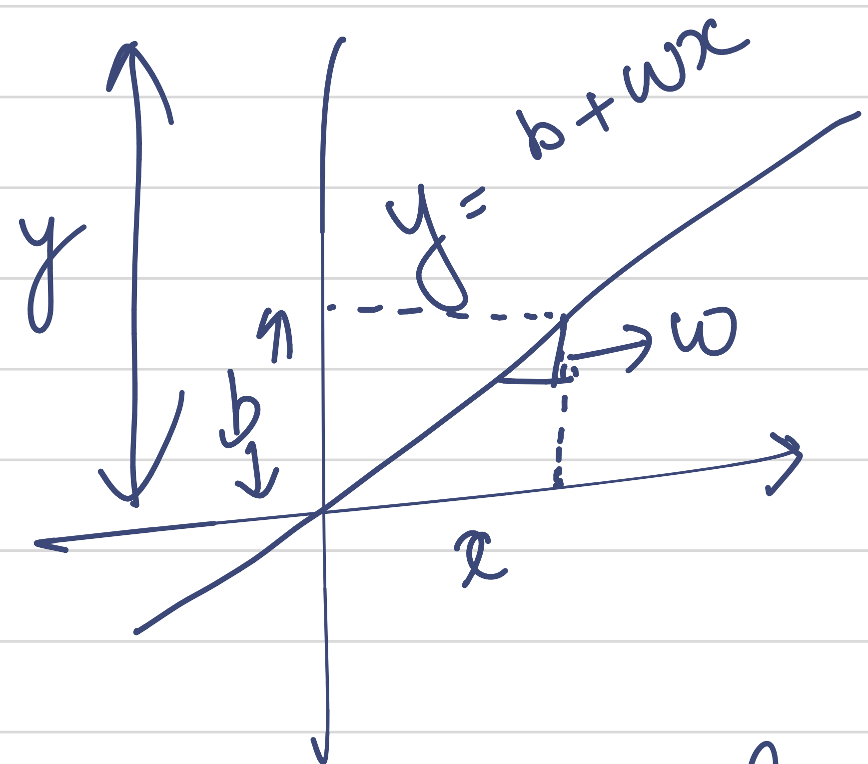
Logistic Regression

$$z = w * x + b$$

where w = weight

x = input

$b \Rightarrow$ bias



But Logistic Regression gives binary classification so for that we use Sigmoid function

$$\sigma(z) \Rightarrow \frac{1}{1+e^{-z}}$$

$$\Rightarrow \frac{1}{1+e^{-(wx+b)}}$$

So, Loss function for Logistic Regression is not Squared Error but -log approach so that output [0,1]

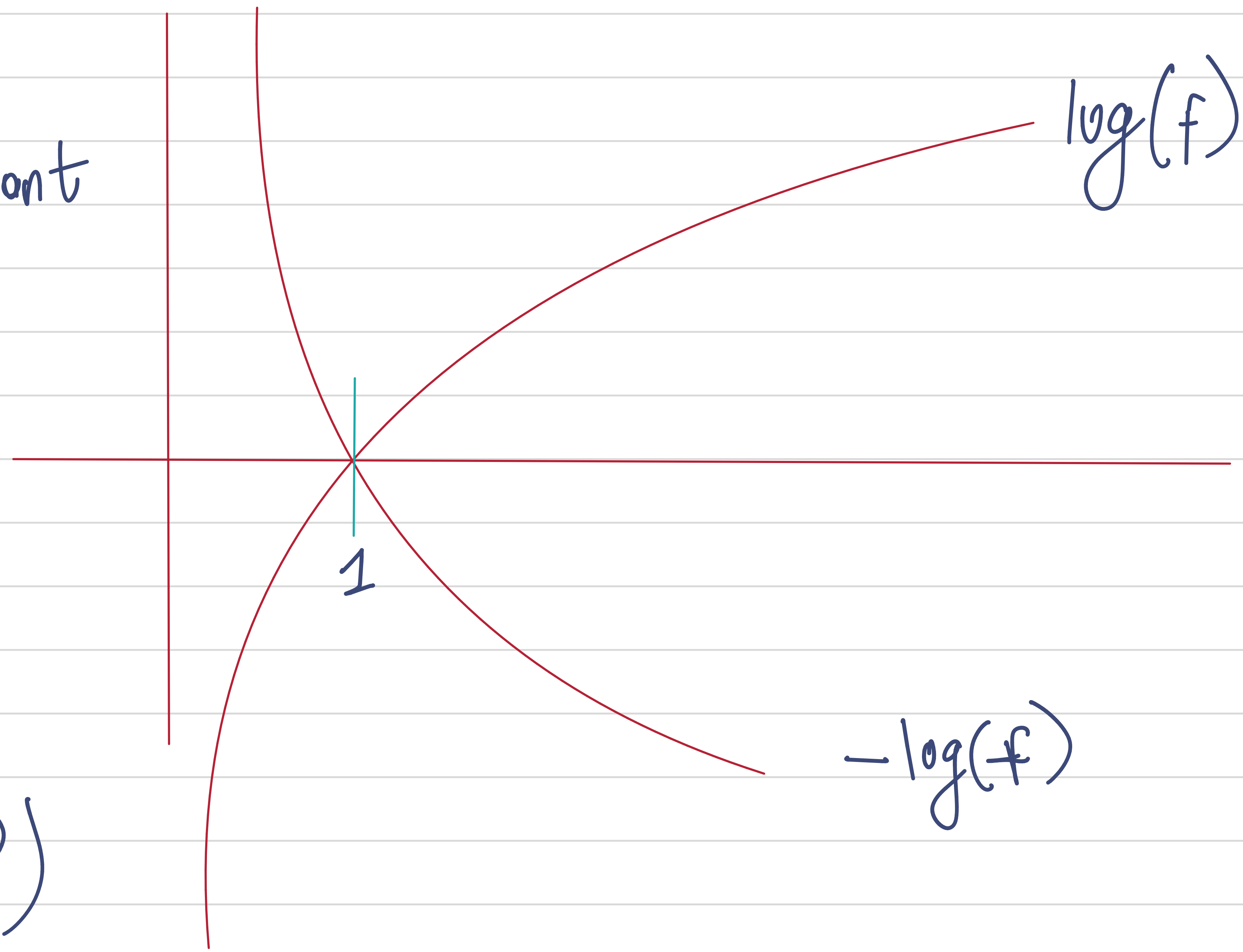
So,

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

Visualization part

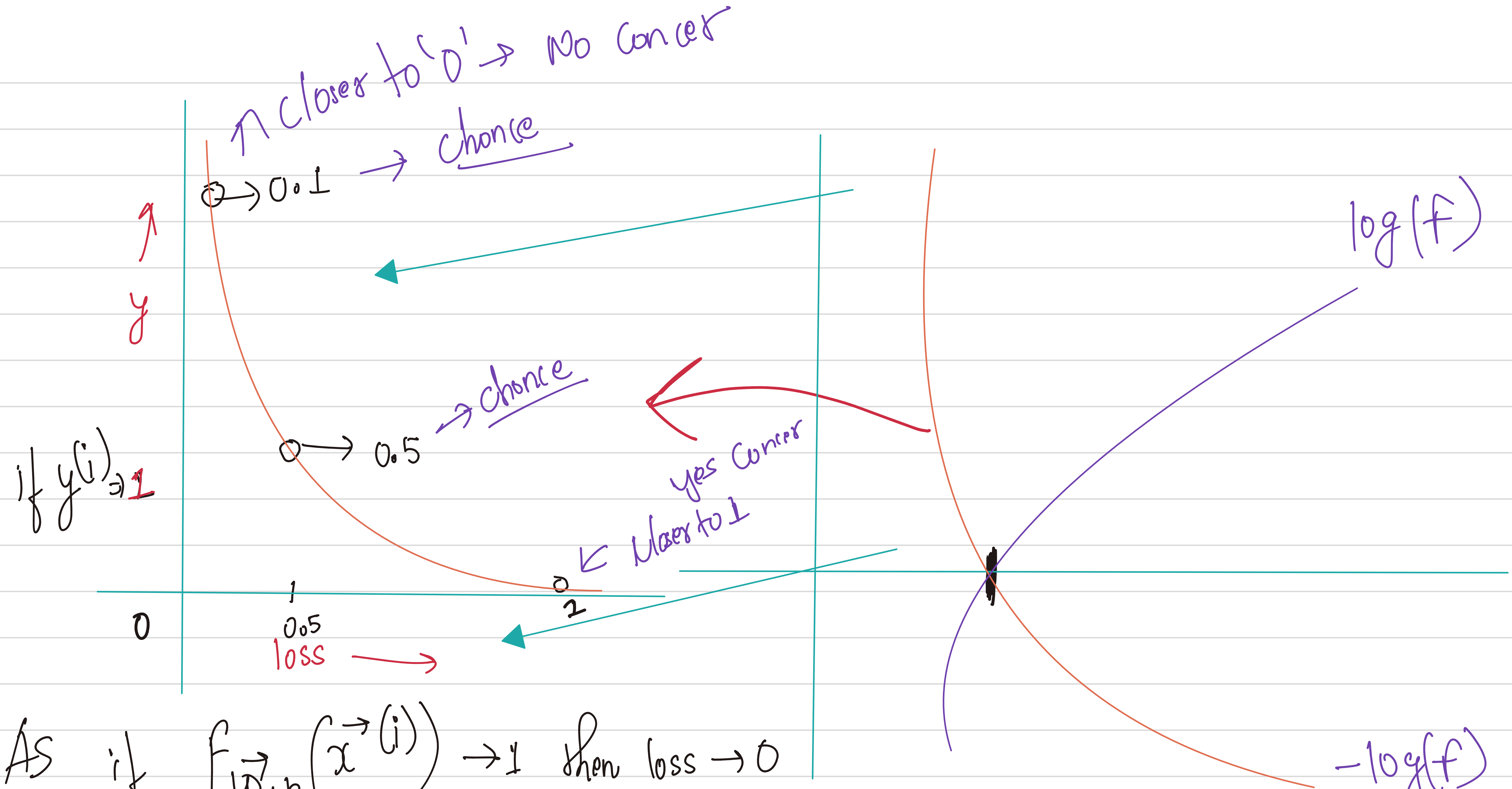
Dataset

tumor size (cm)	...	patient age	malignant
10		55	0
2		42	1
3		23	0
5		⋮	⋮
⋮		25	1
2			



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

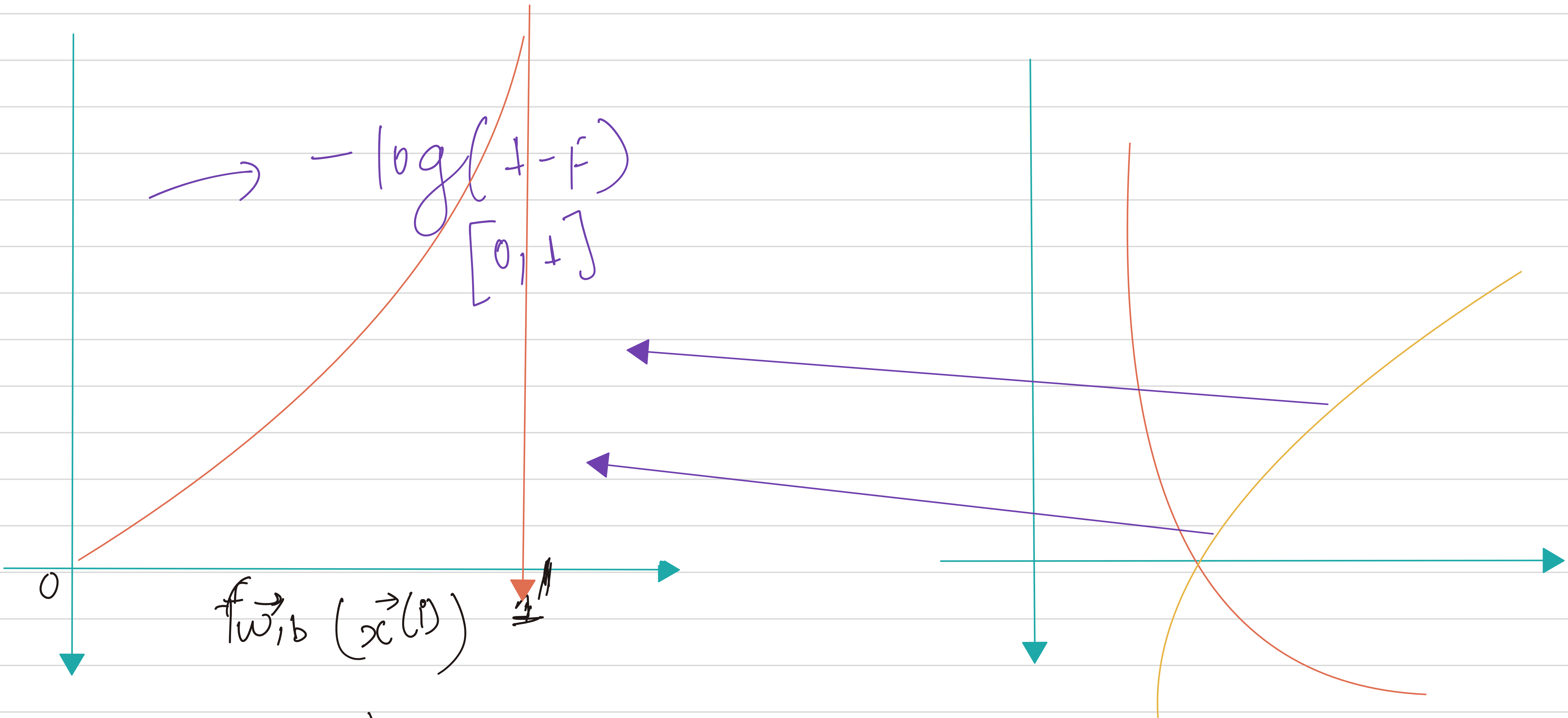
$y=0, \quad -\log\left(1 - f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right)$



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

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

So, $f_{\vec{w},b}(\vec{x}^{(i)})$ predicts close to y lables if the loss is lowest.



if $f_{\vec{w},b}(\vec{x}(i)) \to 1$ then loss $\to \infty$

if $f_{\vec{w},b}(\vec{x}(i)) \to 0$ then loss $\to 0$

Difference between $y(i)$ and $f_{\vec{w},b}(\vec{x}(i))$ is high then the loss is higher

So the mathematically we can write, Cost function as:-

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

Loss

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

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

So,

$$\mathcal{L}\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right), y^{(i)}\right) \Rightarrow -y^{(i)} \cancel{\log\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right)} - \cancel{\left(1 - y^{(i)}\right)\log\left(1 - f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right)}$$

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

if $y^{(i)} \Rightarrow 0 \rightarrow \mathcal{L}\left(f_{\vec{w},b}\left(\vec{x}^{(i)}, y^{(i)}\right)\right) \Rightarrow -\log\left(1 - f(\vec{x})\right)$

Simplified Cost 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

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

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

Parameters of Logistic Regression can be derived for MLE and it is Convex-function Cost function.