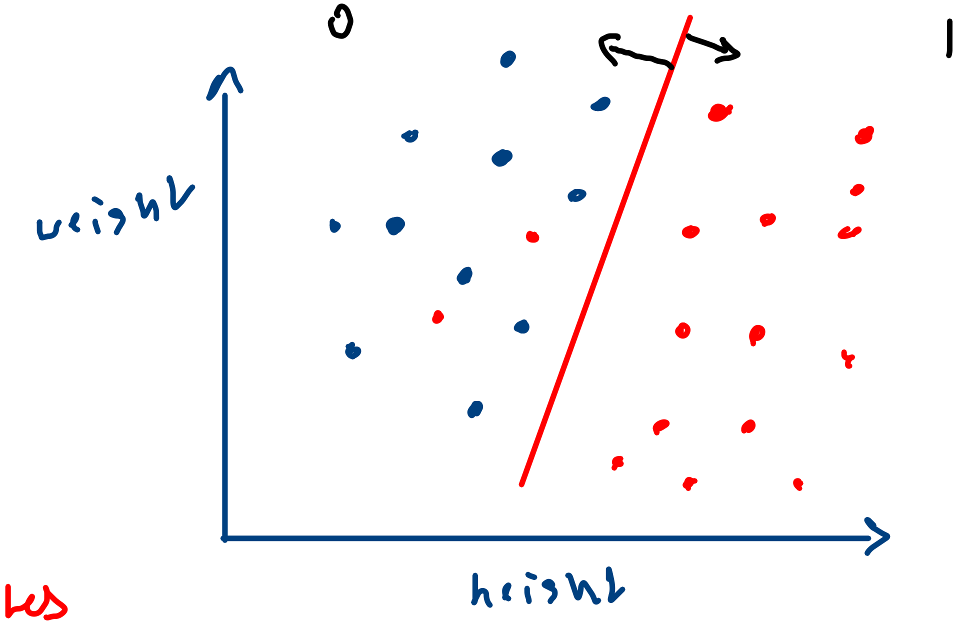
Linear ~ feature

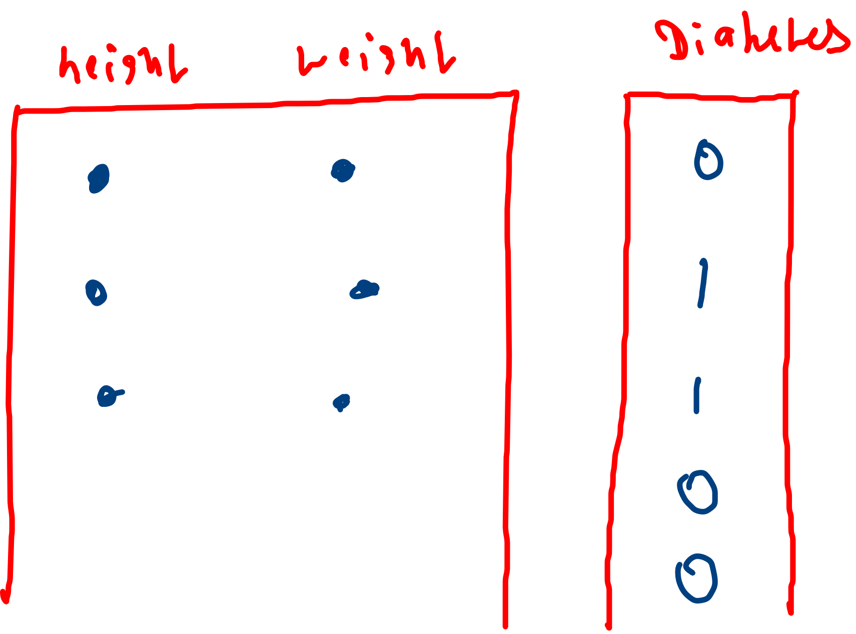
Logistic Regression → Binary Classification

Assumption → linearly seperable

Best Fit Line

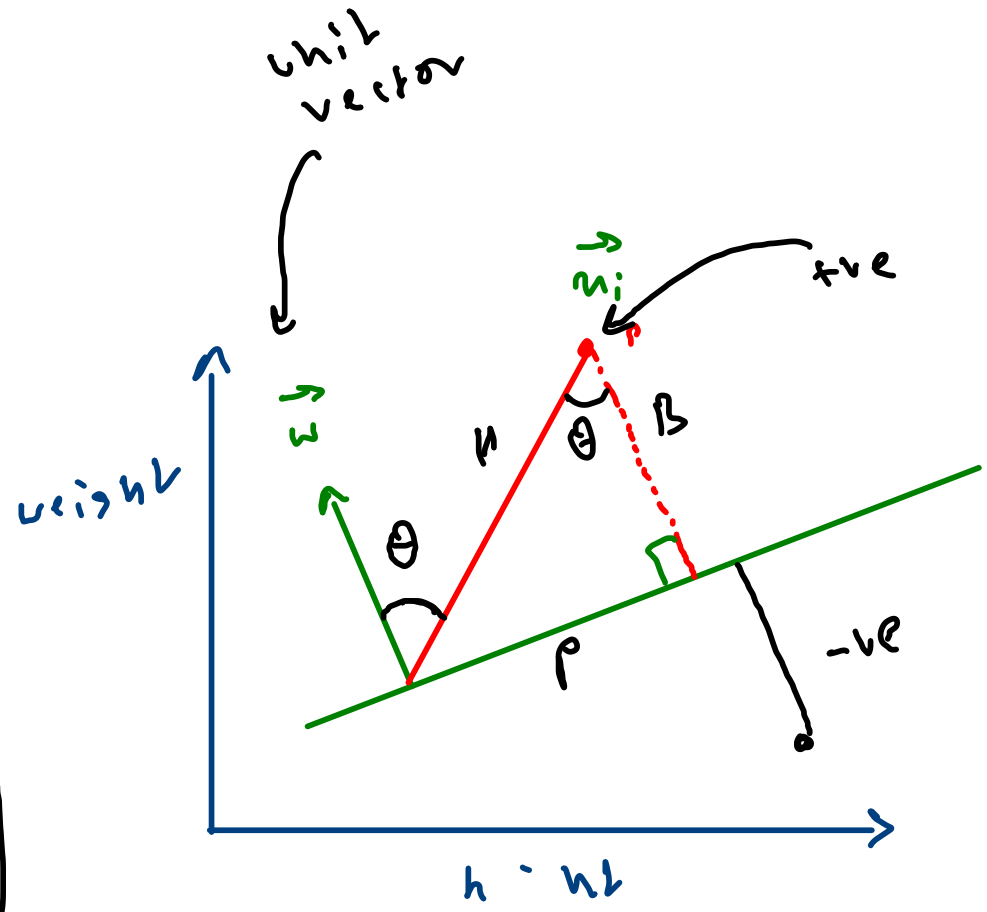


- Best fit line
- distance
- probability
- Cost function
- Gradient Descent
- Softmax



0 $\boxed{0.3 \quad 0.5 \quad 0.2}$ 1

Distance



-ve +ve

0 . - - - - - - . 1

$$\vec{w}\cdot\vec{x} = w\,x\cos\theta$$

$$= w\,\cancel{x}\,\frac{B}{\cancel{H}}$$

$$\vec{w}\cdot\vec{x} = \|w\|\,B$$

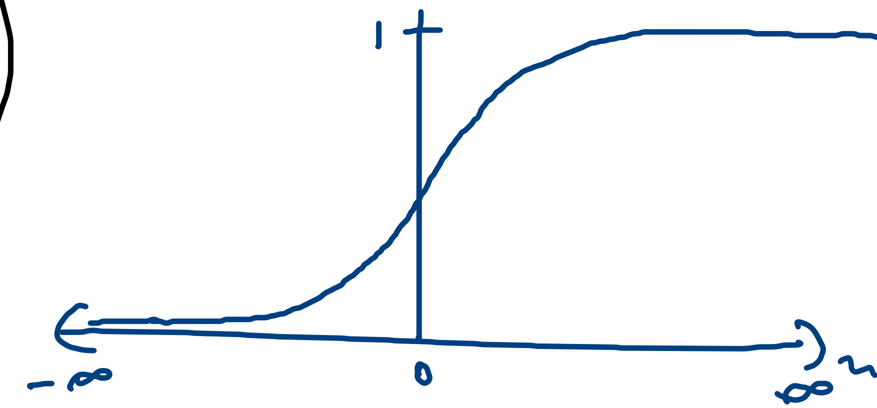
$$B = \frac{\vec{w}\cdot\vec{x}}{\|\vec{w}\|} \rightarrow 1$$

$$\boxed{B = \vec{w}\cdot\vec{x}}$$

$$B = w\,x\cos\theta$$

Probability

sigmoid $\sigma(x) = \frac{1}{1 + e^{-x}}$



$-\infty$ 0		∞
0	0.5	1

distance	$\sigma(x)$
∞	1
30	0.7
3	0.6
0	0.5
-3	0.4
-30	0.3
$-\infty$	0

Cost function

Y_{actual}	$\sigma(w^T x + b)$	cost
0	0	0
0	1	∞
$\{0,1\}$	$[0,1]$	
1	0	∞
1	1	0

cost function

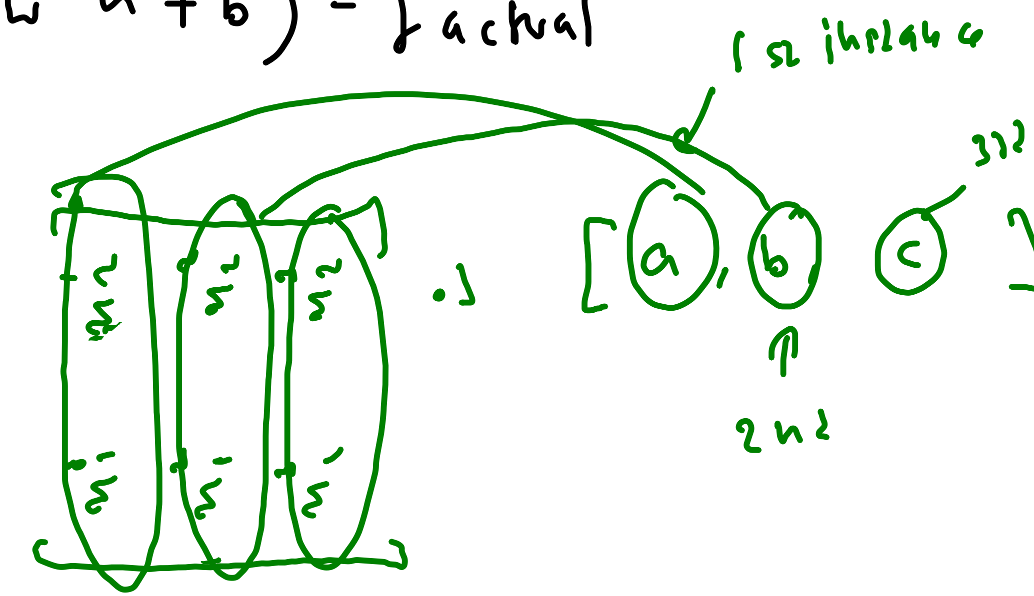
$$\begin{cases} -\log\left[1-(\sigma(wx))\right] & y=0 \\ -\log\left[\sigma(wx)\right] & y=1 \end{cases}$$

number of instance

$$\text{cost function} = \frac{1}{m}\sum_{i=1}^{m} -y\left(\log\left(\sigma(wx)\right)\right) - (1-y)\left(\log\left(1-\sigma(wx)\right)\right)$$

$$\frac{d\ \text{costfunction}}{dw} = \frac{1}{m} \sum \Big[\underbrace{\sigma(w^T x + b) - y_{actual}}_{\text{diff}} \Big] x$$

$$\frac{d\ \text{costfunction}}{db} = \frac{1}{m} \sum \sigma(w^T x + b) - y_{actual}$$



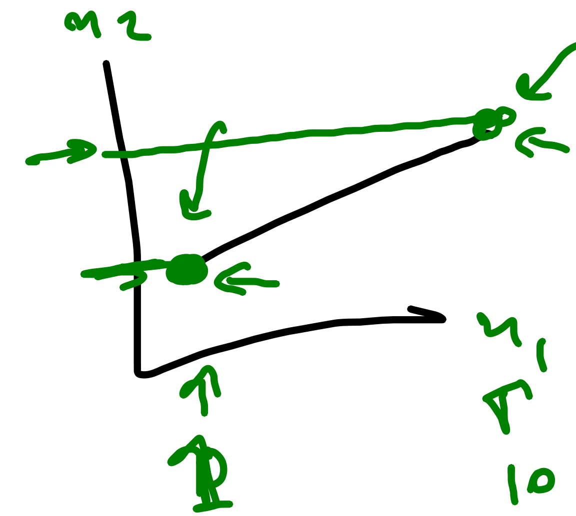
distance $= w^T x + b$

probability $= \sigma($distance

distance = $w^T \cdot x + b$

$$0 = m_1 x_1 + m_2 x_2 + b$$

$$\frac{-(m_1 x_1 + b)}{m_2} = x_2$$



solve solution

M ✓

✓ ← Linear + Logistic

W ✗

F

Titanic

Boston housing

ML

connectivity

LinkedIn, Twitter

Facebook

● ← saturday

● ← sunday

Softmax

weight weight

	x_1	x_2	y	y_1	y_2	y_3
1 –	x_{11}	x_{12}	1	1	0	0
2 –	x_{21}	x_{22}	2	0	1	0
3 –	x_{31}	x_{32}	3	0	0	1
4 –	x_{41}	x_{42}	1	1	0	0

One vs rest

overfit

underfit

fit

overfit

$w^1 = [w_1, w_2, w_0]$ $w^2 = [w_2, w_1, w_3]$ w^3

$$\sigma\begin{pmatrix}\text{under}\\ \text{fit}\end{pmatrix} = \frac{e^{w^1 x}}{e^{w^1 x} + e^{w^2 x} + e^{w^3 x}}$$

$$\sigma\begin{pmatrix}\text{over}\\ \text{fit}\end{pmatrix} = \frac{e^{w^3 x}}{e^{w^1 x} + e^{w^2 x} + e^{w^3 x}}$$

$$\sigma(\text{fit}) = \frac{e^{w^2 x}}{e^{w^1 x} + e^{w^2 x} + e^{w^3 x}}$$

1 → underfit

2 → fit

3 → overfit

Training Softmax

$$\Rightarrow \frac{1}{m}\sum_{i=1}^{m} -y\left(\log\left(\sigma(wx)\right)\right) - (1-y)\left(\log\left(1-\sigma(wx)\right)\right)$$

$$\Rightarrow \frac{1}{m}\sum_{i=1}^{m}\sum_{k=1}^{K} y^i_k \log\Big(\underbrace{w_k x_i}_{w_k^1 x_i^1 + w_k^2 x_i^2 \cdots + b_k}\Big)$$

For all instances

For all classes