linear classification دسته بند خطی

linear: تنها توان یک پارامترها استفاده می‌شود

classification: $y \in \{class1, class2, \ldots, class_K\}$

$y \in \{$ داخل ، بیرون $\}$

$X \in 2D$ ، تصاویر

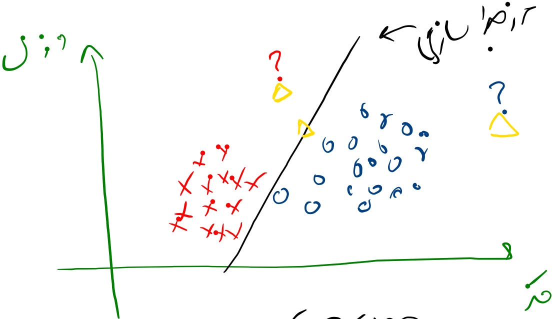
$w_0 + w_1x_1 + \ldots + w_dx_d$

$y \in \{$ زن ، مرد $\}$

$X \in \mathbb{R}^{n\times 2}$ قد وزن

بیرون

داخل

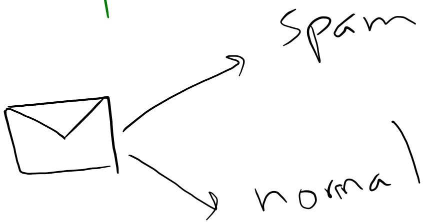


X زن

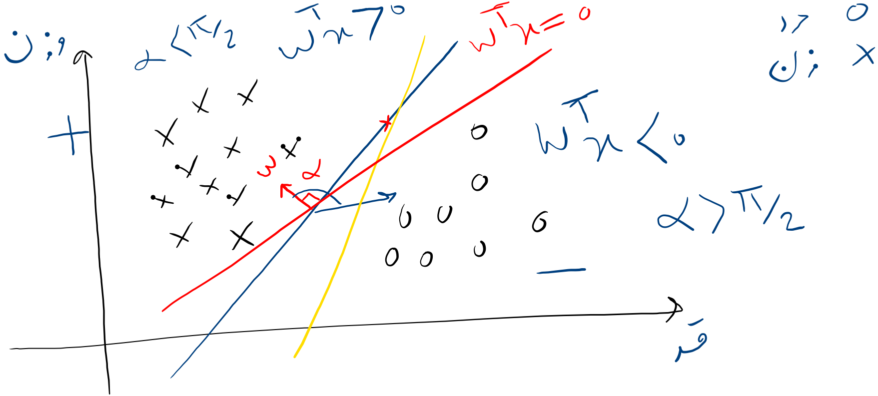
O مرد

تشخیص هرزنامه

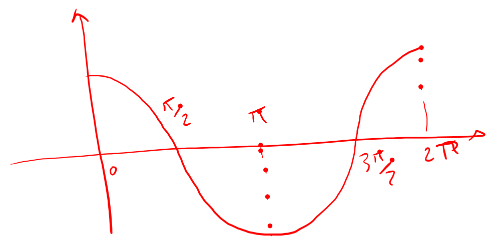
Spam Email



$y \in \{spam, normal\}$



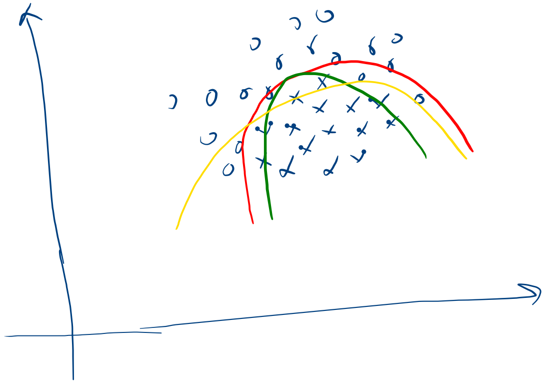
$$w^T x = \|w\| \, \|x\| \cos\alpha$$



linearly separable

linear Classification

داده‌ها بصورت خطی جداپذیر باشند



$$W = \begin{bmatrix} 0.5 \\ 0.5 \\ -0.5 \end{bmatrix} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} \qquad x^{(i)} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} \qquad w^T x^{(i)} = w_0 + w_1 x_1 + w_2 x_2$$

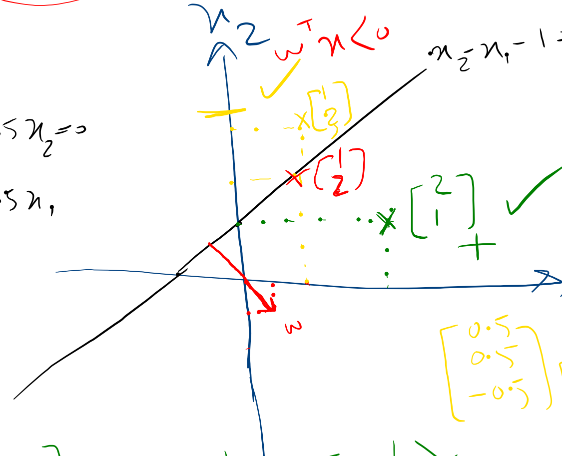
$x_2 - x_1 - 1 = 0 \quad (w^T x = 0)$

$w^T x = 0$

$$0.5 + 0.5 x_1 - 0.5 x_2 = 0$$

$$-0.5 x_2 = -0.5 - 0.5 x_1$$

$$\boxed{x_2 = x_1 + 1}$$



$w^T x > 0$

$\boxed{w^T x}$

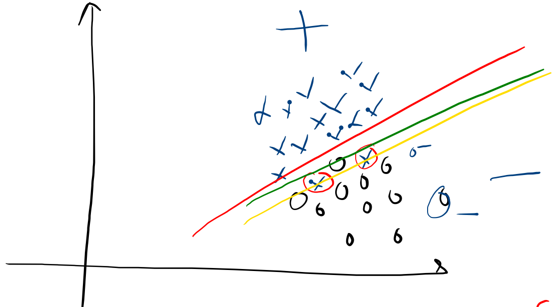
حاصل
[illegible]
ضرب

$$\begin{bmatrix} 0.5 \\ 0.5 \\ -0.5 \end{bmatrix} [1 \ 1 \ 3] = 0.5 + 0.5 - 1.5 = -0.5 < 0$$

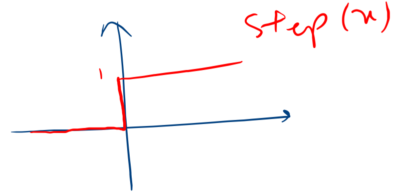
$$\begin{bmatrix} 0.5 \\ 0.5 \\ -0.5 \end{bmatrix} [1 \ 2 \ 1] = 0.5 + 1 - 0.5 = 1 > 0$$

$w^* = ?$

w را طوری پیدا کنیم که همه را درست
دسته‌بندی کند



$$\text{step}(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$



$$\text{step}(w^T x) = \begin{cases} \tilde{y} = 1 & w^T x > 0 \\ \tilde{y} = 0 & w^T x < 0 \end{cases}$$

$y = \{$ آری , نه $\}$

$\Downarrow$ 1 $\quad \Downarrow$ 0

$$\operatorname{step}(w^T x) = \begin{cases} \tilde{y} = 1 & w^T x > 0 \\ \tilde{y} = 0 & w^T x < 0 \end{cases}$$

$$w^* \Rightarrow \min J(w) = \frac{1}{n}\sum_{i=1}^{n}\left[\operatorname{step}(w^T x^i) \neq y^i\right]$$

کمینه

minimum

تابع هزینه $\equiv$ loss function $\Rightarrow$ 0-1 Loss

$$w^* = \min_{w} J(w) = \frac{1}{n}\sum_{i=1}^{n}\Big[\underbrace{\operatorname{step}(w^T x^{(i)})}_{\tilde{y}^{(i)}} \neq y^{(i)}\Big]$$

$\in \{$True , false$\}$

1 $\quad$ 0

$$J(w) = \frac{1}{n}\sum_{i=1}^{n}\left[\operatorname{step}(w^T x^{(i)}) \neq y^{(i)}\right]$$

$$J(w) = \frac{1}{n}\sum_{i=1}^{n}\left[y^{(i)}\left(1-\operatorname{step}(w^T x^{(i)})\right) + (1-y^{(i)})\left(\operatorname{step}(w^T x^{(i)})\right)\right]$$

$\neq \Rightarrow 1$

$= \Rightarrow 0$

$y^{(i)} = \begin{matrix}0\\1\end{matrix} \Leftrightarrow \operatorname{step} = \begin{matrix}1\\0\end{matrix}$

$y^{i} = \begin{matrix}0\\1\end{matrix} \Rightarrow \operatorname{step} \begin{matrix}0\\1\end{matrix}$

$w^* = ?$ $\qquad \frac{\partial J}{\partial w} = 0$

(1)

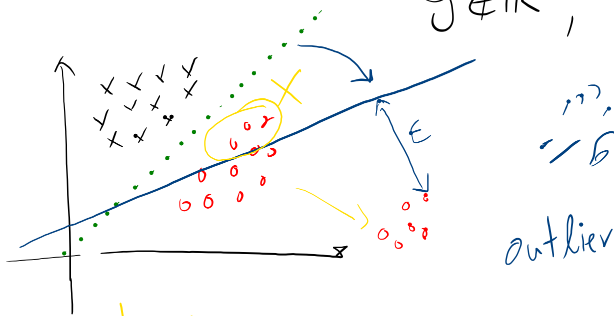
$$J(w)=\frac{1}{n}\sum_{i=1}^{n}\left[y^{(i)}\left(1-\mathrm{Step}(w^Tx^{(i)})\right)+(1-y^{(i)})\,\mathrm{Step}(w^Tx^{(i)})\right]$$

0 - 1 loss

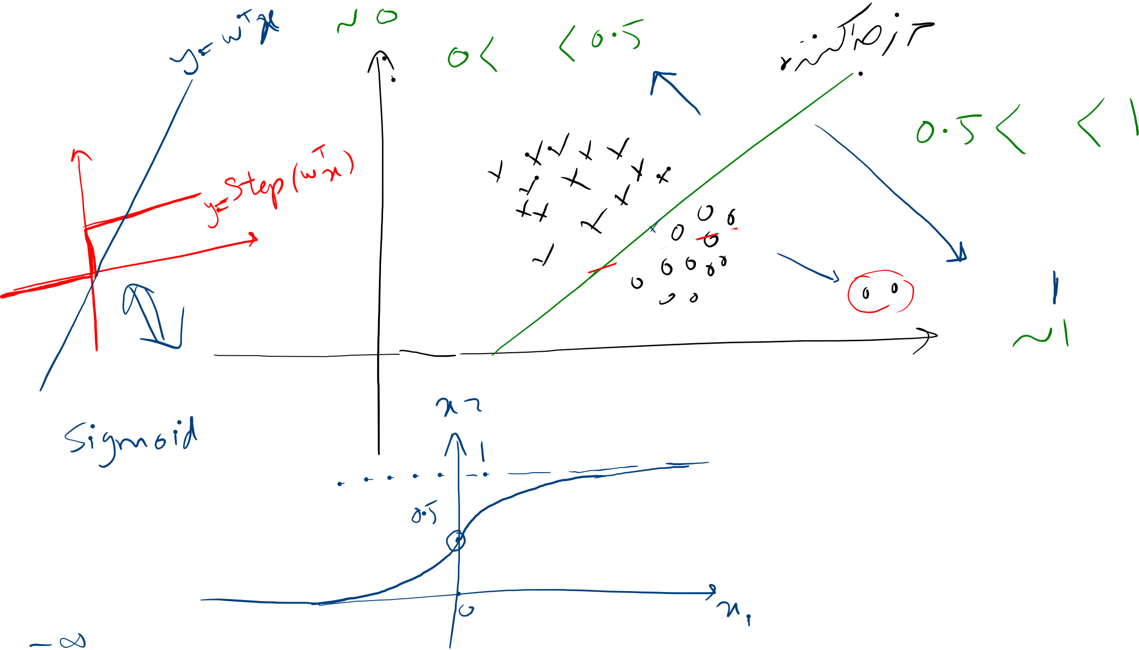
X $$J(w)=\frac{1}{n}\sum_{i=1}^{n}\left(w^Tx^{(i)}-y^{(i)}\right)^2$$

LMSE

$y\notin\mathbb{R}$, $y\in\{0,1\}$

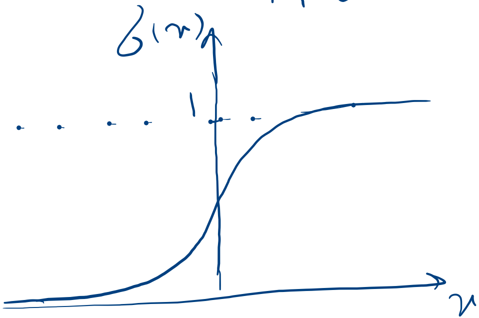


dense



$y = w^T x$
~ 0
$0 < \quad < 0.5$
جداکننده
$0.5 < \quad < 1$
$y = \mathrm{Step}(w^T x)$
~ 1
Sigmoid
x_2
1
0.5
0
x_1
$-\infty$

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



$$x \to +\infty \Rightarrow \sigma(x) \to 1$$

$$x \to -\infty \Rightarrow \sigma(x) \to 0$$

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

$$J(w) = \frac{1}{n}\sum_{i=1}^{n}\left(\sigma(w^T x^{(i)}) - y^{(i)}\right)^2$$

1

$$\frac{\partial \sigma(x)}{\partial x} = \frac{\partial}{\partial x}\left(\frac{1}{1+e^{-x}}\right) = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{e^{-x}}{1+e^{-x}} \times \frac{1}{1+e^{-x}}$$

$$= \sigma(x)\left[\frac{e^{-x}+1-1}{1+e^{-x}}\right] \Rightarrow \boxed{\frac{\partial \sigma(x)}{\partial x} = \left(1-\sigma(x)\right)\sigma(x)}$$

$$\Downarrow$$

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

2

$$1-\sigma(x) \equiv \sigma(-x)$$

$$\frac{\sigma(x)}{\sigma(x)} - \frac{\sigma^2(x)}{\sigma(x)} = \frac{\sigma(x)(1-\sigma(x))}{\sigma(x)} = \frac{e^{-x}}{1+e^{-x}} = \frac{1}{1+e^{x}}$$

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

$y \in \{0,1\}$

3

$\sigma(x) \in (0,1)$

$$J(w) = \frac{1}{n}\sum_{i=1}^{n}\left(\sigma(w^T x^{(i)}) - y^{(i)}\right)^2$$

$$\boxed{w^* = \frac{\partial J(w)}{\partial w} = 0}$$