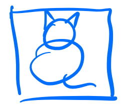
Classification (clf)

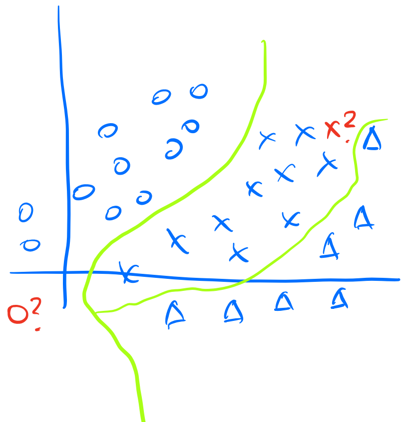
- Regression $\rightarrow x_i \rightarrow$ target value $y_i \in \mathbb{R} = (-\infty, \infty)$
- clf $\rightarrow x_i \rightarrow$ target class $c_i \in \{1, 2, \dots, k\} = [k]$

e.g. x_i = images $c_i \in \{$cat, dog$\}$

$\{0,1\}$
$\{-1,1\}$



$\rightarrow \boxed{\text{clf}} \rightarrow c_i$

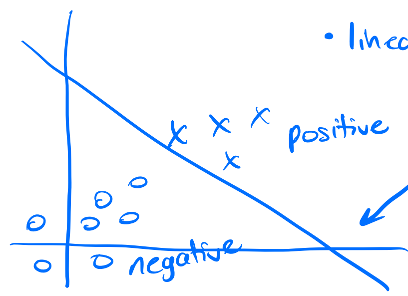


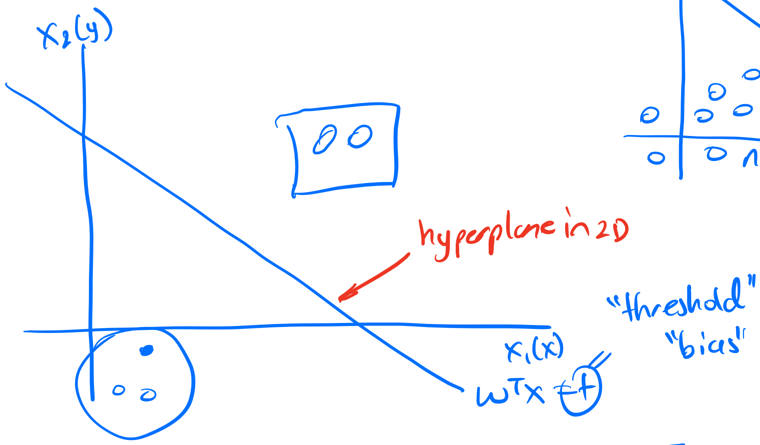
$D = \{(x_i, c_i) \; ; i = 1, \dots, n\}$

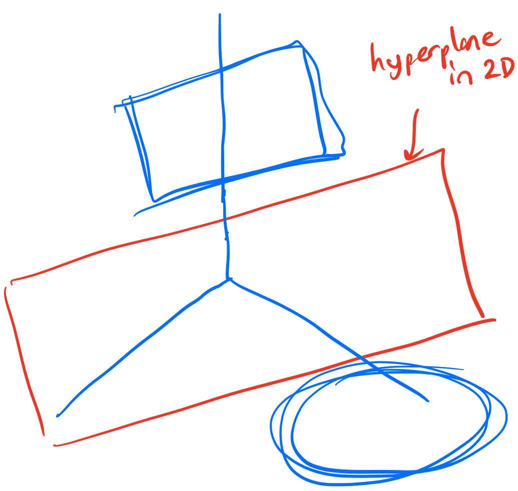
$c_i \in \{o, x, \Delta\}$

$x_i \in \mathbb{R}^p$

- Binary clf $k = 2$
- linear clf







$w^Tx = t \Rightarrow$ equation of a hyperplane (line)

in $\mathbb{R}^p$: $w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{pmatrix}$ $\quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ $\quad t \in \mathbb{R}^1$

in $\mathbb{R}^2$: $w^Tx = t \Rightarrow \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = t$

$\Rightarrow w_1 x_1 + w_2 x_2 = t$

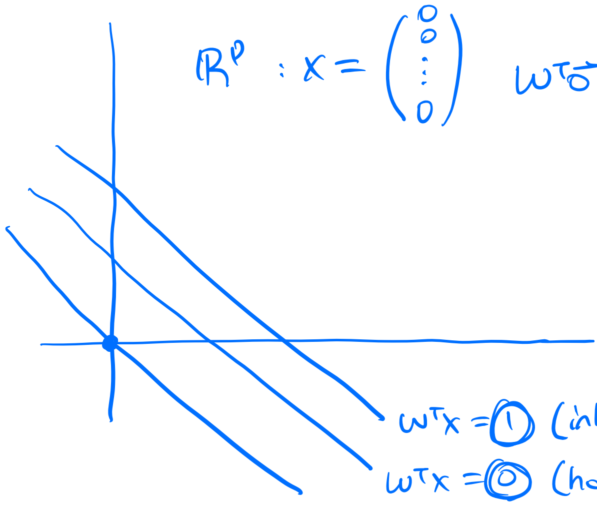
$\Rightarrow w_2 x_2 = t - w_1 x_1$

$\Rightarrow \boxed{x_2 = \frac{t}{w_2} - \frac{w_1}{w_2} x_1}$

$\mathbb{R}^p : x = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ $w^T\vec{0} = 0$

slope/intercept form $\rightarrow$ $y = a + bx$

$$\boxed{\begin{aligned} a &= \frac{t}{w_2} \\ b &= -\frac{w_1}{w_2} \end{aligned}}$$



$w^Tx = 1$ (inhomogenous)

$w^Tx = 0$ (homogenous)

$w^Tx = -1$

$$\vec{w} = \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_p \end{pmatrix} = \underbrace{\begin{pmatrix} -t \\ w \end{pmatrix}}_{\in \mathbb{R}^{p+1}} \qquad \tilde{x} = \begin{pmatrix} 1 \\ x \end{pmatrix} \in \mathbb{R}^{p+1}$$

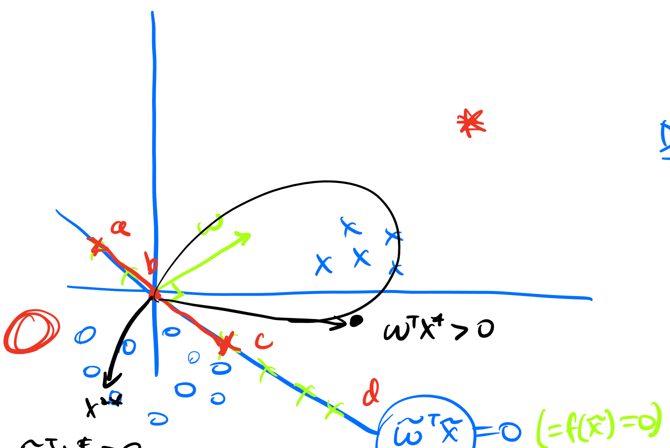
$w^Tx = t \Leftrightarrow \tilde{w}^T\tilde{x} = 0$

$w_1x_1 + \dots + w_px_p = t$

$-t + w^Tx = 0$

$w^Tx = t$

$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$ $x = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_p \end{pmatrix}$



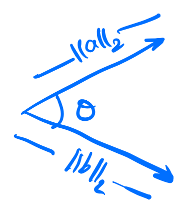
$\tilde{w}^Tx^* > 0 \rightarrow$

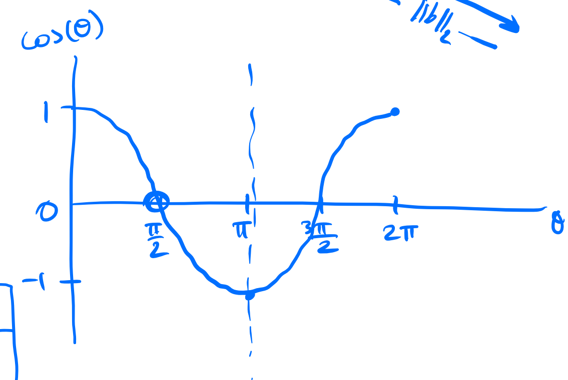
$\tilde{w}^Tx^{**} < 0 \rightarrow$

$\tilde{w}^Ta = \tilde{w}^Tb = \tilde{w}^Tc = 0$

Dot Products

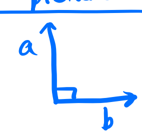
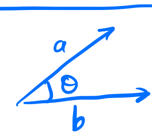
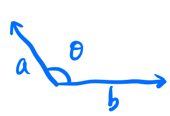
$a^Tb = \|a\|_2 \|b\|_2 \cos(\theta)$





$\tilde{w}^Tx^* > 0 \rightarrow$ clf 1

$\tilde{w}^Tx^{**} < 0 \rightarrow$ clf 0

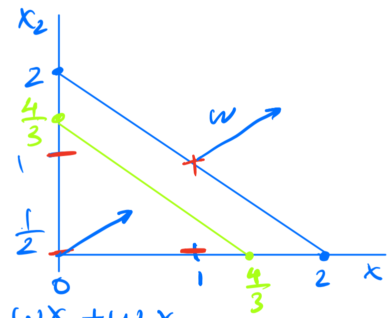
case	θ	angle type	picture
$a^Tb = 0$	$\cos(\theta) = 0$ $\theta = \frac{\pi}{2} = 90°$	right angle	
$a^Tb > 0$	$\cos(\theta) > 0$ $\theta \in (0, \frac{\pi}{2})$ $\theta \in (0, 90°)$	acute	
$a^Tb < 0$	$(90°, 180°)$	obtuse	

picking w?

1. Guess (Brute force) ✓
2. Basic linear clf (centroid)
3. Perceptron

① Guess (Brute force)

x_1	x_2	AND (x_1,x_2)
0	0	0
0	1	0
1	0	0
1	1	1



Rule:

$\tilde{w}^T\tilde{x} \geq 0 \rightarrow \tilde{x}$ is '+'

$\tilde{w}^T\tilde{x} < 0 \rightarrow \tilde{x}$ is '−'

$w^Tx \geq t$

$w^Tx < t$

$w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ $\quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $\quad w_1x_1 + w_2x_2$

$w_1x_1 + w_2x_2 = t$ $\qquad$ find (w_1, w_2, t)

- $w_1(0) + w_2(0) < t \Rightarrow t > 0$
- $w_1(0) + w_2(1) < t \Rightarrow t > w_2$
- $w_1(1) + w_2(0) < t \Rightarrow t > w_1$
- $w_1(1) + w_2(1) \geq t \Rightarrow w_1 + w_2 \geq t$

$\boxed{t = 1 \quad w_1 = w_2 = \frac{1}{2}}$

$x_2 = \frac{t}{w_2} - \frac{w_1}{w_2}x_1$

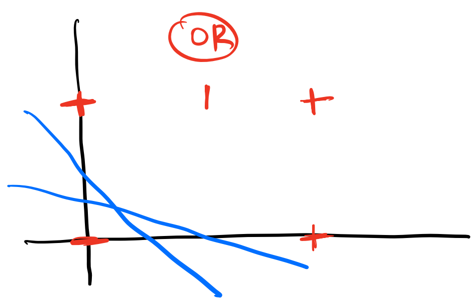
$= \frac{1}{1/2} - \frac{1/2}{1/2}x_1$

$= 2 - x_1$

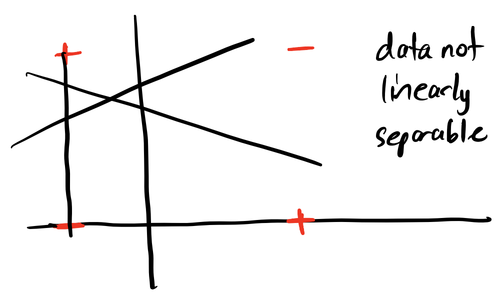
$t = 2, w_1 = w_2 = \frac{3}{2}$

$x_2 = \frac{4}{3} - x_1$

x_1	x_2	OR (x_1, x_2)	XOR (x_1, x_2)
0	0	0	0
0	1	1	1
1	0	1	1
1	1	1	0



XOR



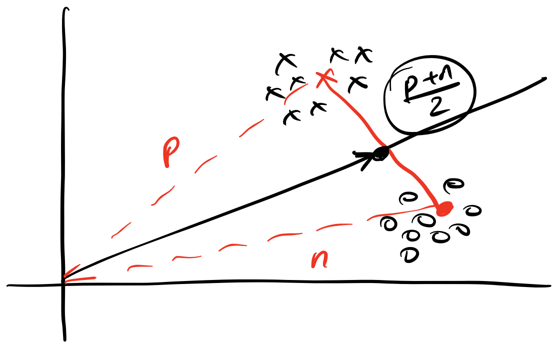
data not linearly separable

$$\left.\begin{cases} t > 0 \\ w_2 \geq t > 0 \\ w_1 \geq t > 0 \\ (w_1 + w_2) < t \end{cases}\right\}$$ inconsistent

$w_1 + w_2 \geq t + t = 2t > 0$

$(w_1, w_2, t)^*$

"Basic" Centroid linear clf



$\boxed{\begin{array}{l} \bullet\ w = p - n \\ \bullet\ w^T x = t \end{array}}$

$w^T x = t$

when $x = \frac{p+n}{2}$ $\qquad$ $w^T x = t$

$$w^T x = (p-n)^T \left(\frac{p+n}{2}\right)$$

$$= \frac{1}{2}(p^T - n^T)(p+n)$$

$$= \frac{1}{2}(p^T p + p^T n - n^T p - n^T n)$$

$$= \frac{1}{2}(p^T p - n^T n)$$

$$= \frac{1}{2}(\|p\|^2 - \|n\|^2)$$

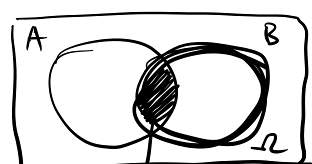
$$z^T z = [z_1, \ldots, z_p] \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix}$$

$$= z_1^2 + z_2^2 + \ldots + z_p^2$$

$$\|z\|_2^2 = \sqrt{\sum_{i=1}^{p} z_i^2}$$

Bayes Rule

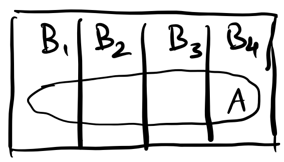
$$\boxed{P(A|B) = \frac{P(AB)}{P(B)}} = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$$



$\frac{P(AB)}{P(B)}$

$$P(A|B)\,P(B) = P(AB) = P(B|A)\,P(A)$$

$$P(B|A) = \frac{P(AB)}{P(A)}$$



Law of total probability:

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + P(A \cap B_4)$$

$$= P(A|B_1)\,P(B_1) + P(A|B_2)\,P(B_2) + \ldots$$

$$= \sum_{i=1}^{4} P(A|B_i)\,P(B_i)$$

Assume the probability of a certain disease is 0.01. The probability of testing positive given that a person is infected with the disease is 0.95 and the probability of testing positive given the person is not infected with the disease is 0.05.

False positive

(a) Calculate the probability of testing positive. [5pt]

(b) Use Bayes' Rule to calculate the probability of being infected with the disease given that the test is positive. [5pt]

posterior $\propto$ likelihood x prior

$$P(D|T) \propto \underbrace{P(T|D)}_{0.95} \underbrace{P(D)}_{\times 0.01}$$

T = test is positive

D = has disease

(a) $P(T) = P(T|D)P(D) + P(T|\bar{D})P(\bar{D})$

$= 0.95 \times 0.01 + 0.05 \times 0.99 = 0.059$

(b) $P(D|T) = \dfrac{P(T|D)P(D)}{P(T)} = \dfrac{0.95 \times 0.01}{0.059} = 0.16$

Assume that we tested twice, positive twice. Assume tests are conditionally independent, given the disease:

$P(TT) = P(TT|D)P(D) + P(TT|\bar{D})P(\bar{D})$

$P(T)P(T) = P(T|D)^2 P(D) + P(T|\bar{D}^2)P(\bar{D})$

$= (0.95)^2 \times 0.01 + (0.05)^2 \times 0.99$

$= \boxed{}$

$P(D|TT) = \dfrac{P(TT|D)P(D)}{P(TT)}$

≈ 0.7