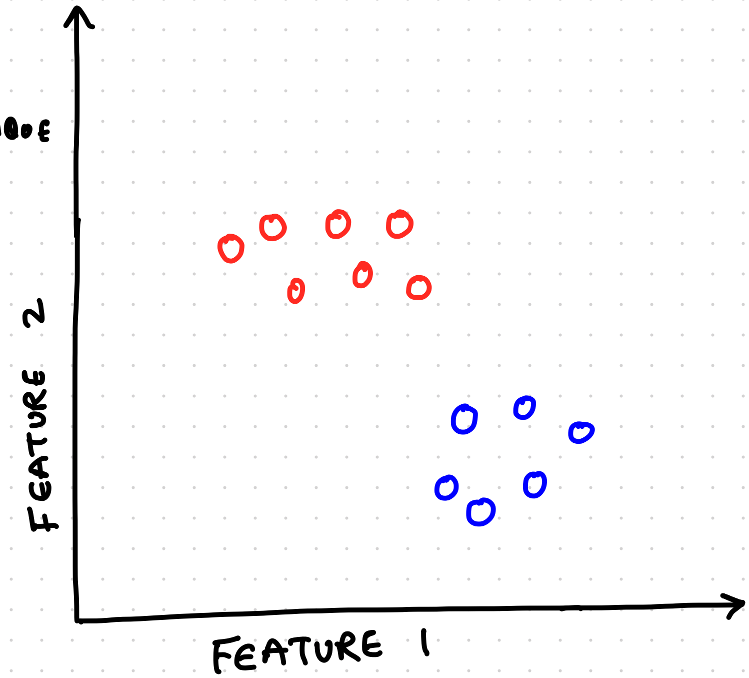
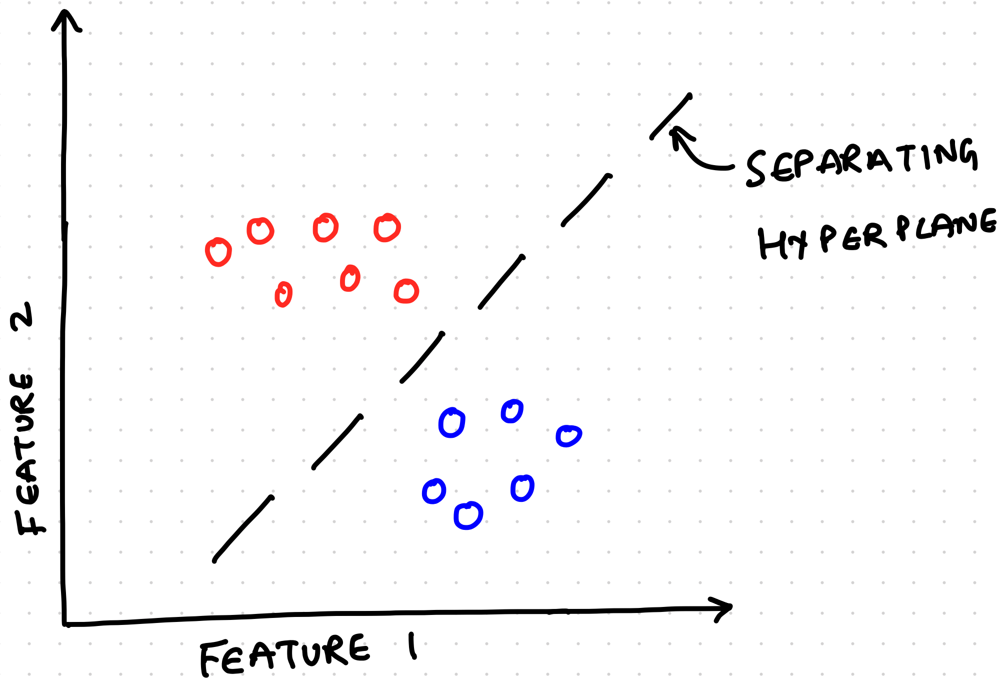
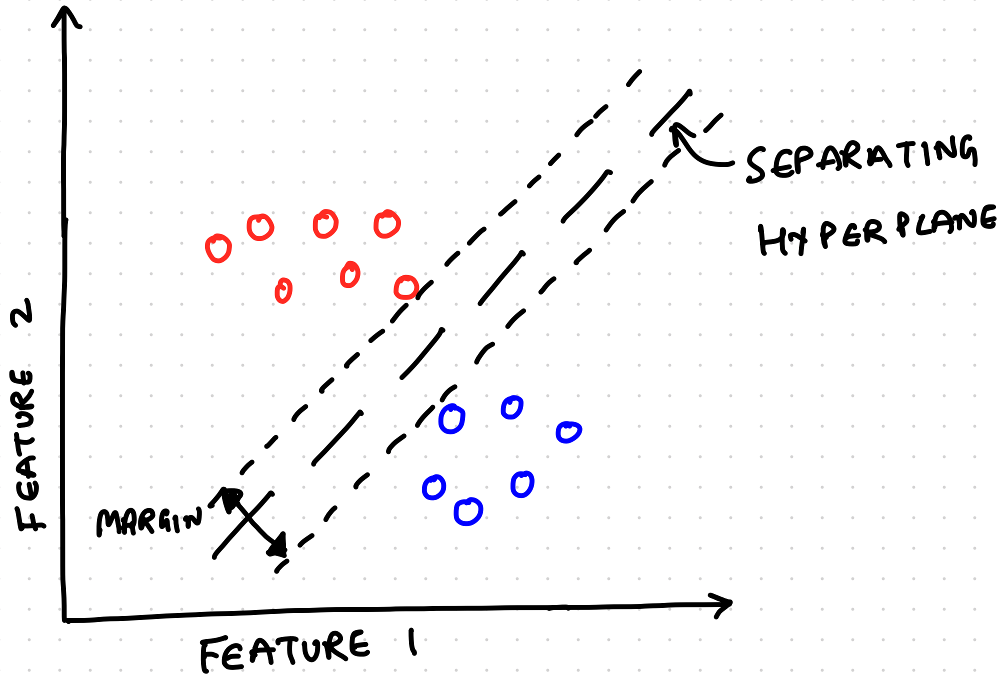
SUPPORT VECTOR MACHINES

POPULAR BINARY CLASSIFICATION TECHNIQUE

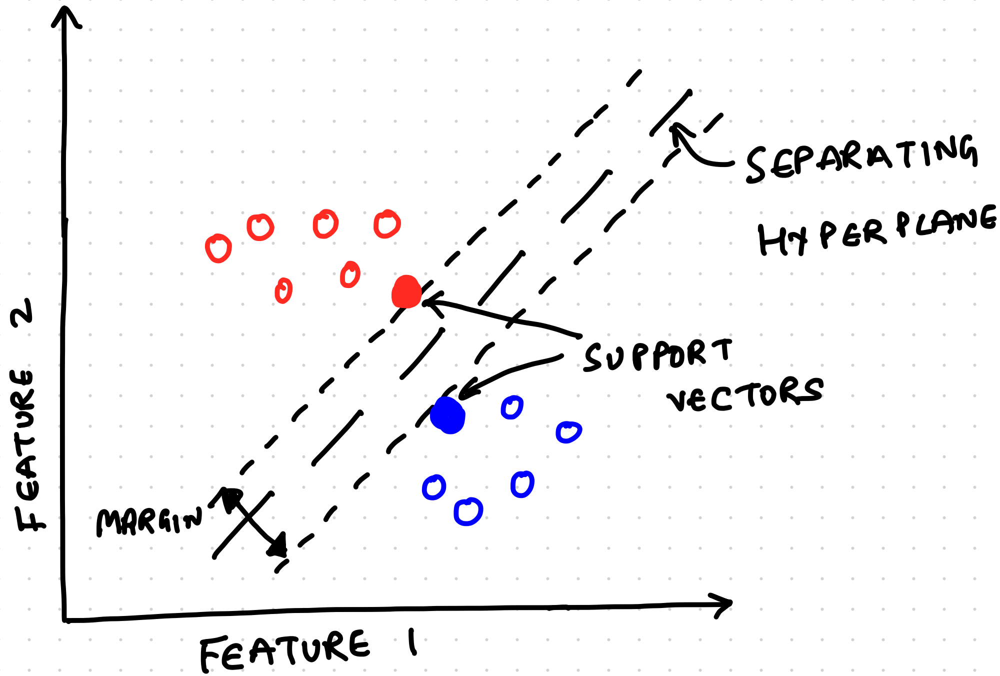




IDEA: DRAW A SEPARATING HYPERPLANE



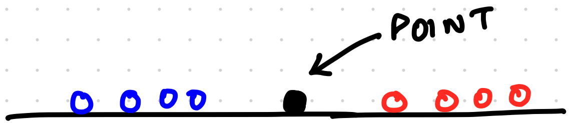
IDEA: MAXIMIZE THE MARGIN



SUPPORT VECTORS: POINTS ON BOUNDARY | MARGIN

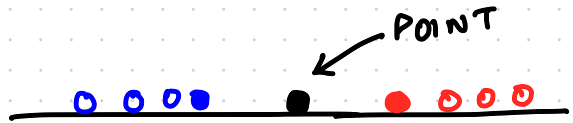
HYPERPLANE V/S # DIMENSIONS

1D

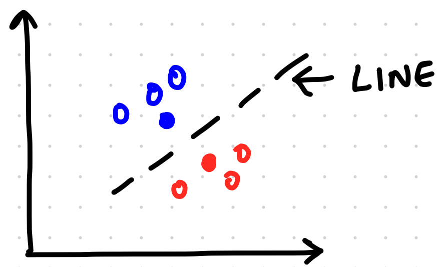


HYPERPLANE V/S # DIMENSIONS

1D

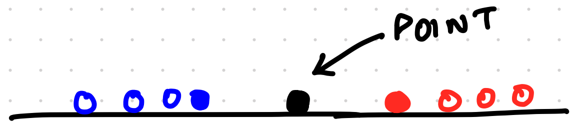


2D

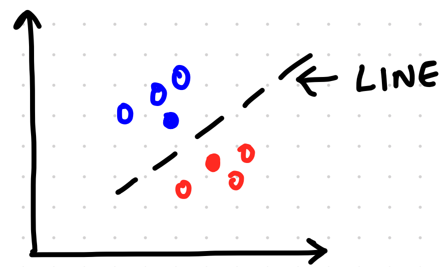


HYPERPLANE V/S # DIMENSIONS

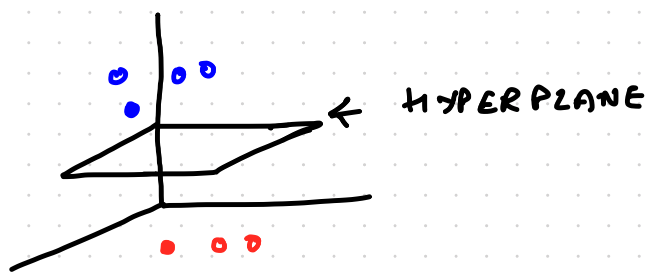
1D



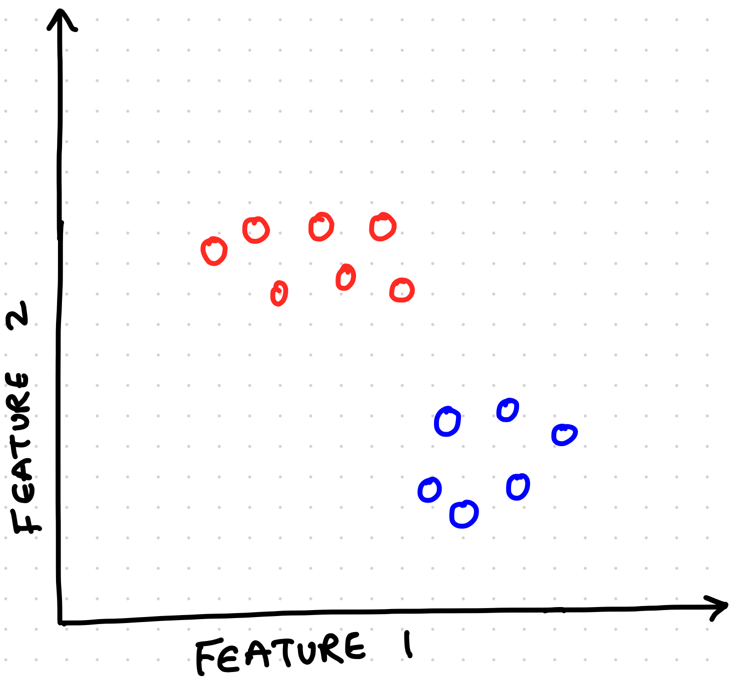
2D



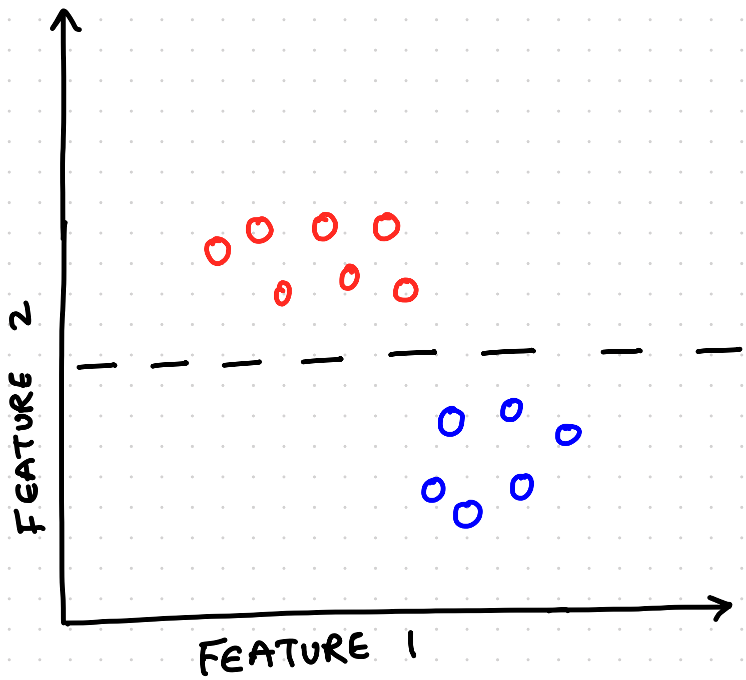
3D
(AND
MORE)



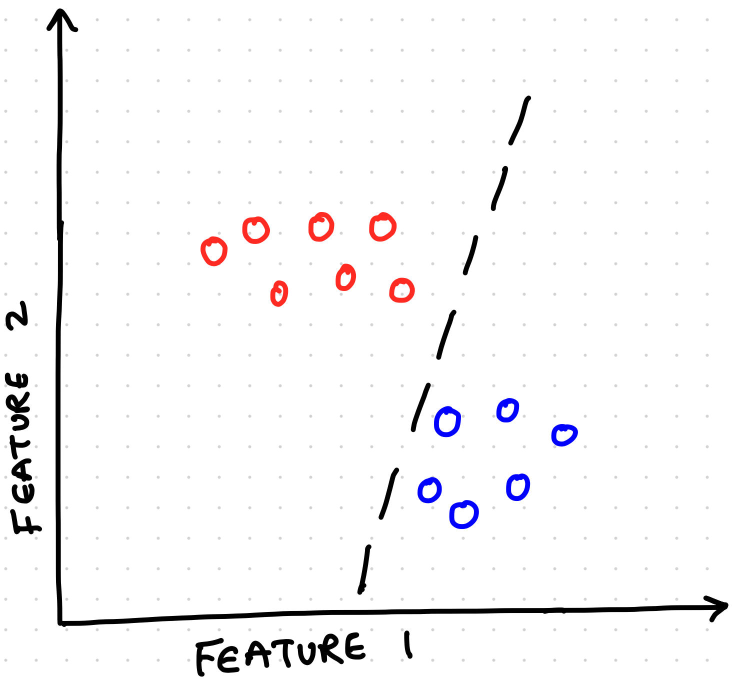
WHICH HYPERPLANE?



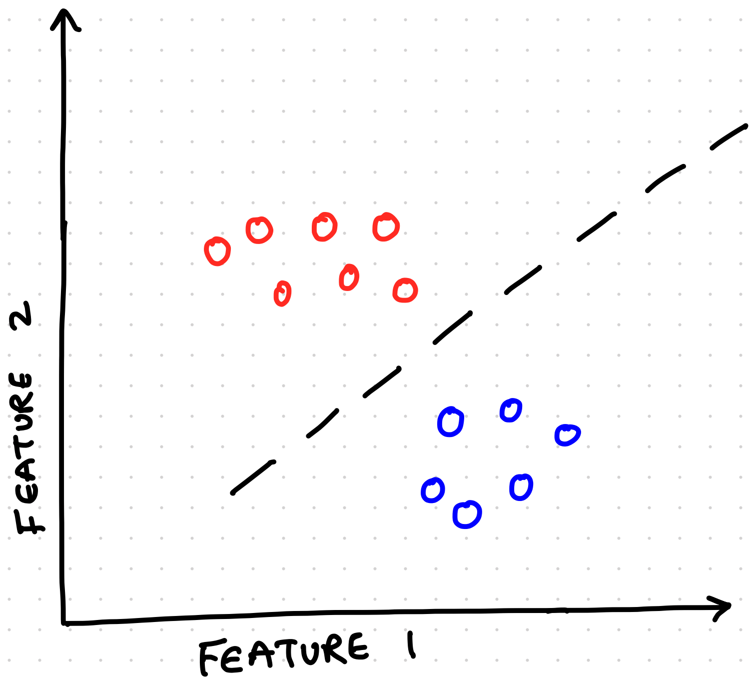
WHICH HYPERPLANE?



WHICH HYPERPLANE?

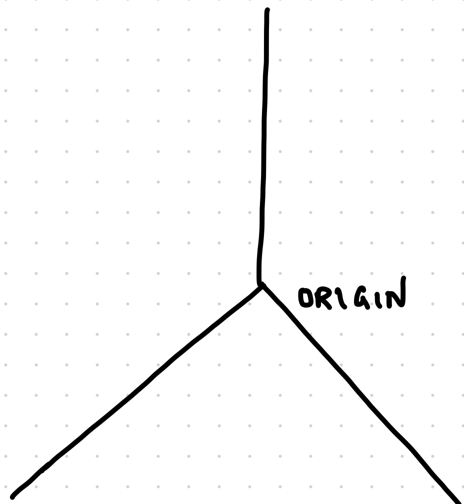


WHICH HYPERPLANE?

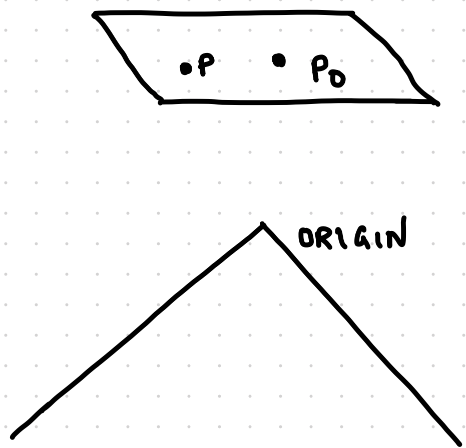


EQUATION OF HYPERPLANE

HOW TO DEFINE?



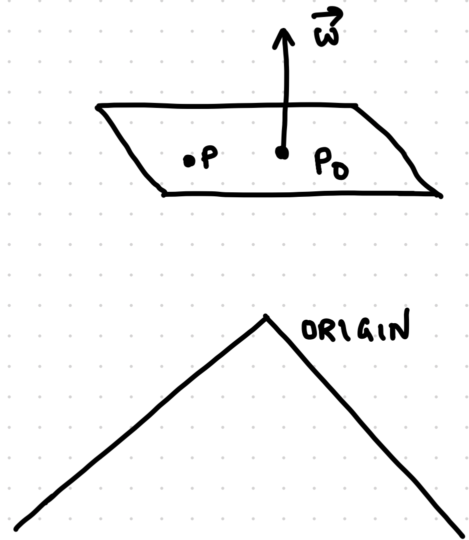
EQUATION OF HYPERPLANE



P : Any point on plane

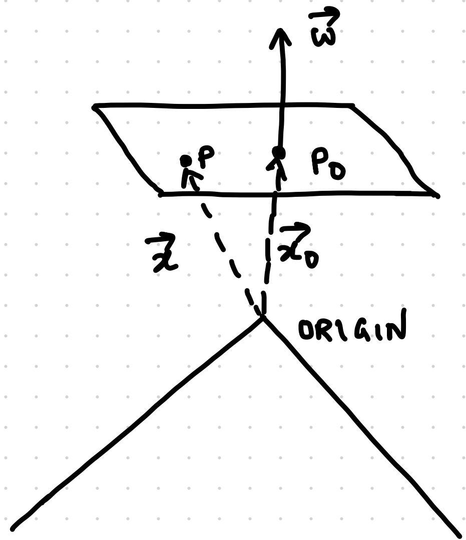
P_0 : One point on plane

EQUATION OF HYPERPLANE



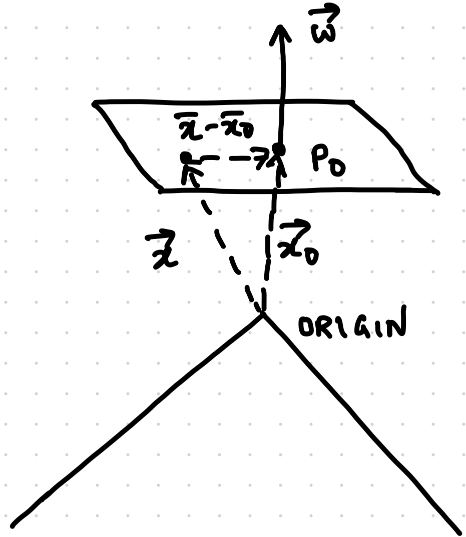
$\vec{w}$: $\perp$ vector to plane at P_0

EQUATION OF HYPERPLANE



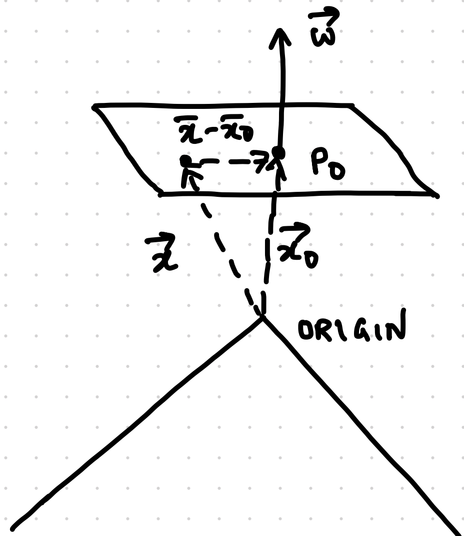
P and P_0 lie on plane

EQUATION OF HYPERPLANE



$\overrightarrow{P P_0} = \vec{x} - \vec{x}_0$ lies on plane

EQUATION OF HYPERPLANE



$\overrightarrow{PP_0} = \vec{x} - \vec{x}_0$ lies on plane

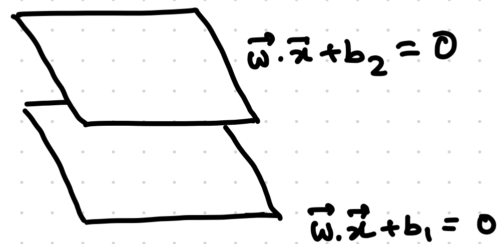
$\Rightarrow \vec{w} \perp (\vec{x} - \vec{x}_0)$

or, $\vec{w} \cdot (\vec{x} - \vec{x}_0) = 0$

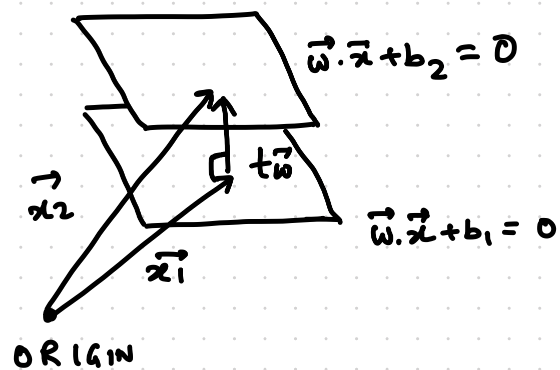
or, $\vec{w} \cdot \vec{x} - \vec{w} \cdot \vec{x}_0 = 0$

or, $\boxed{\vec{w} \cdot \vec{x} + b = 0}$

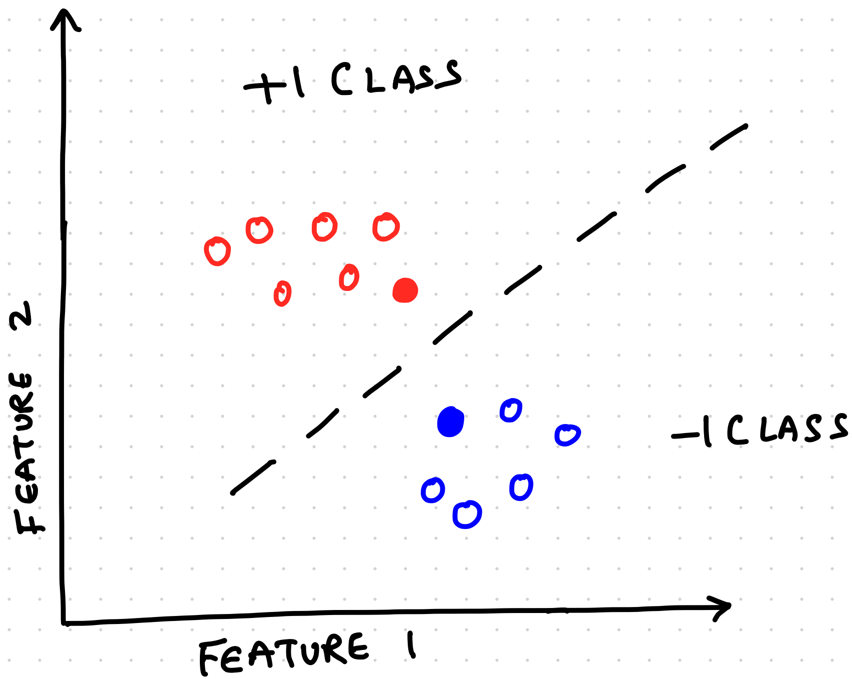
DISTANCE B/W || HYPERPLANES



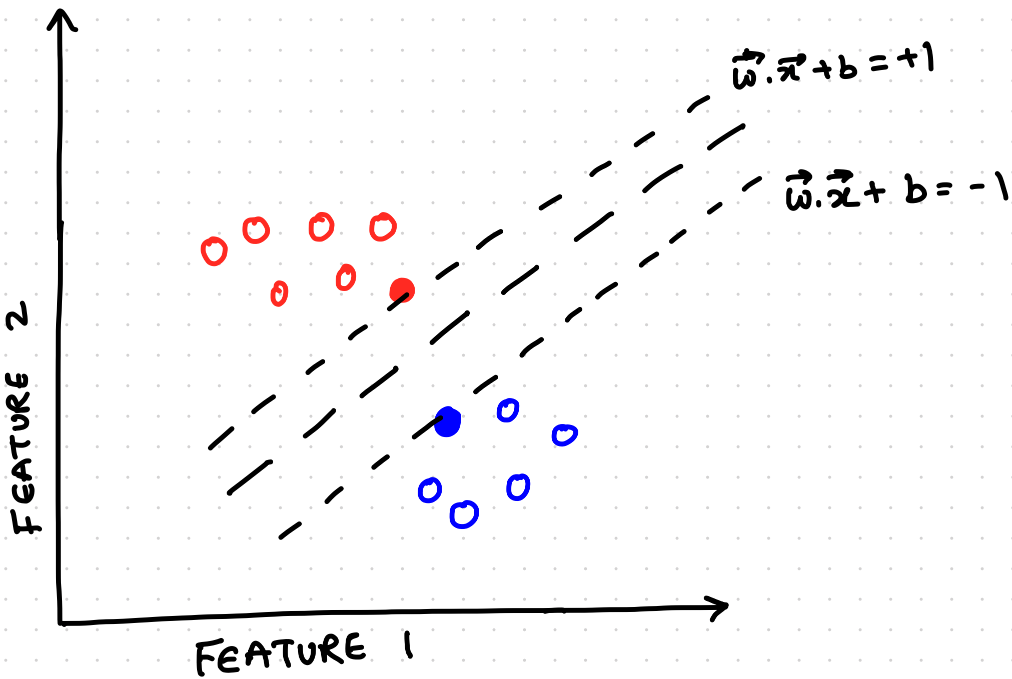
DISTANCE B/W || HYPERPLANES



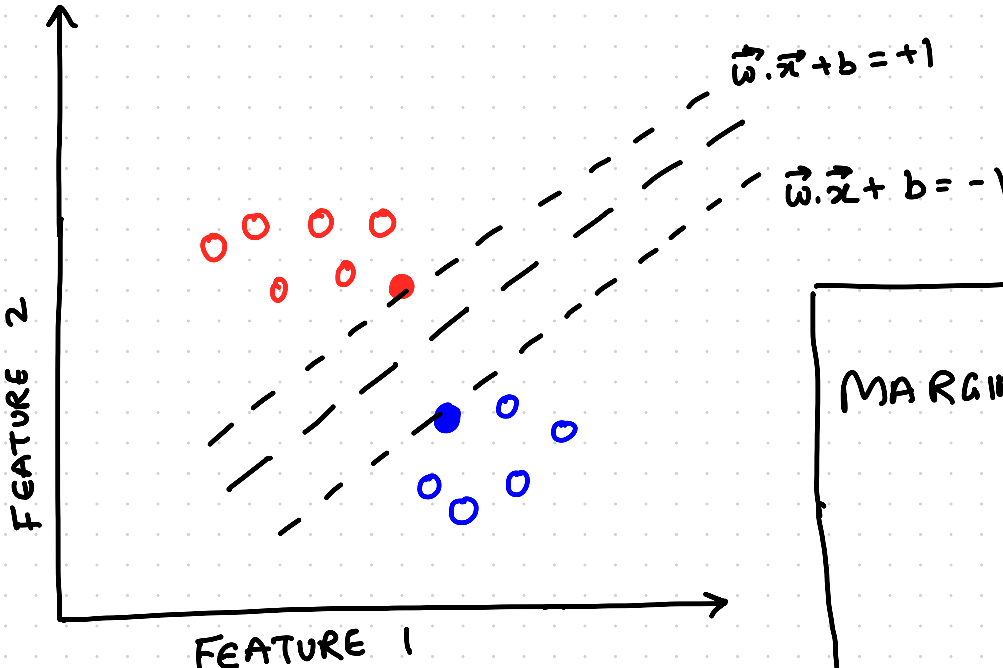
FORMULATION



FORMULATION



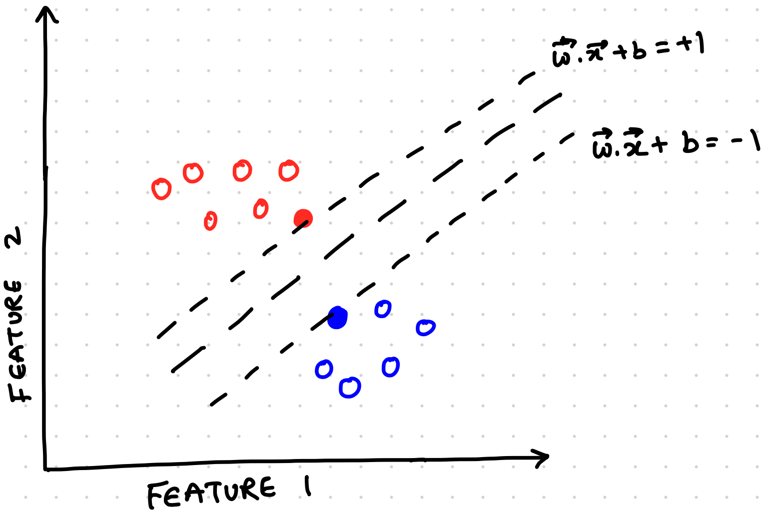
FORMULATION



$$\text{MARGIN} = \frac{(b+1)-(b-1)}{\|\vec{w}\|}$$

$$= \frac{2}{\|\vec{w}\|}$$

FORMULATION



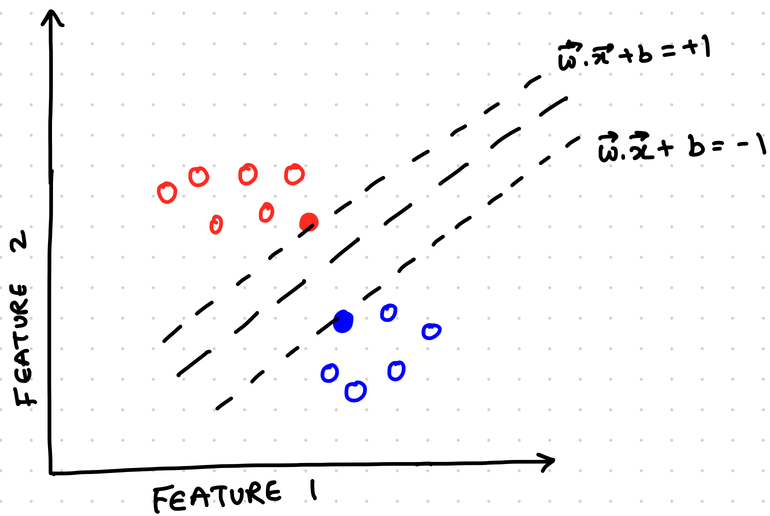
GOAL: MAXIMIZE MARGIN

$\Rightarrow$ MAXIMIZE $\dfrac{2}{||\vec{w}||}$

$\Rightarrow$ MINIMIZE $||\vec{w}||$

S.T. Correctly label points

FORMULATION



GOAL: MAXIMIZE MARGIN

$\Rightarrow$ MAXIMIZE $\dfrac{2}{||\vec{w}||}$

$\Rightarrow$ MINIMIZE $||\vec{w}||$

S.T. Correctly label points

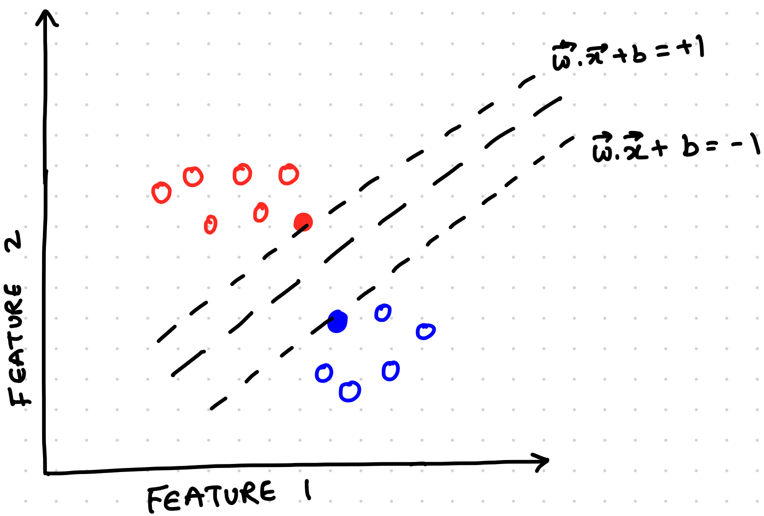
i.e. if $y_i = -1$

$\vec{w}.\vec{x}+b \leq -1$

if $y_i = +1$

$\vec{w}.\vec{x}+b \geq 1$

FORMULATION



GOAL: MAXIMIZE MARGIN

$\Rightarrow$ MAXIMIZE $\dfrac{2}{||\vec{w}||}$

$\Rightarrow$ $\boxed{\text{MINIMIZE } ||\vec{w}||}$

S.T. Correctly label points

i.e. if $y_i = -1$

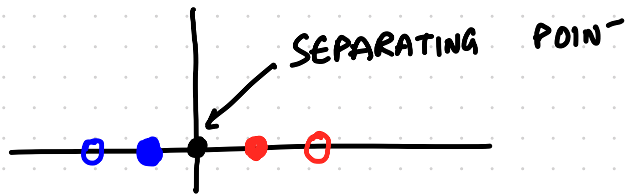
$\vec{w}.\vec{x}+b \le -1$

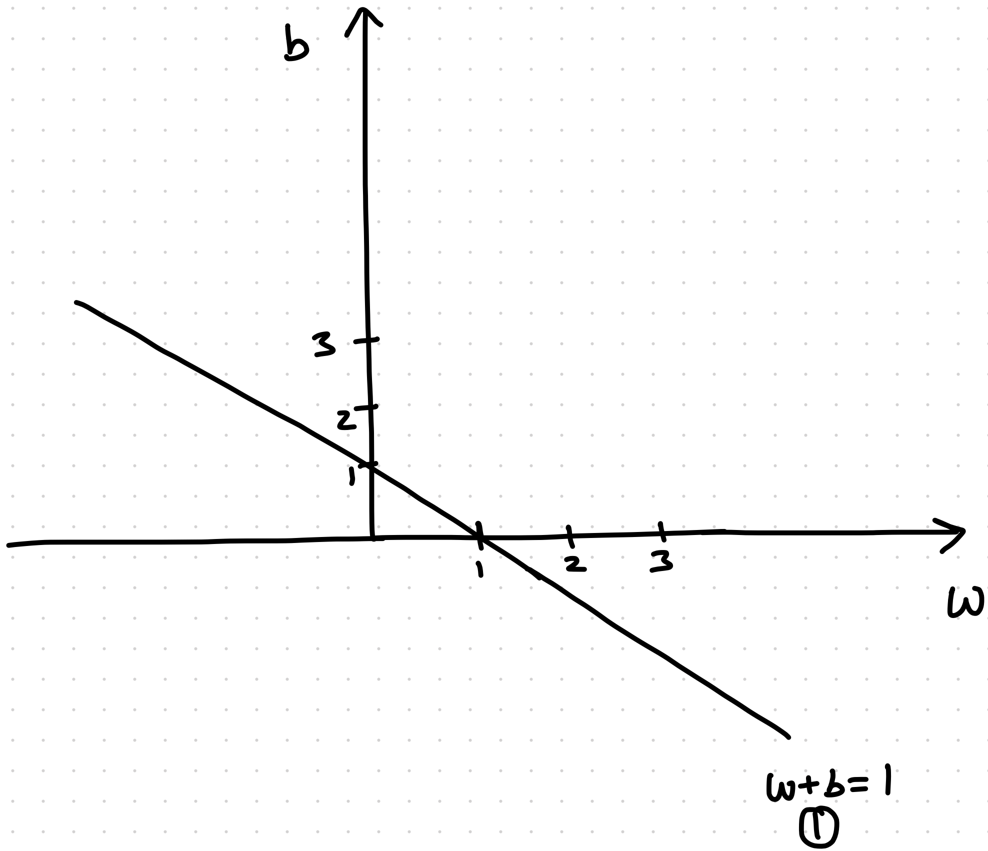
if $y_i = +1$

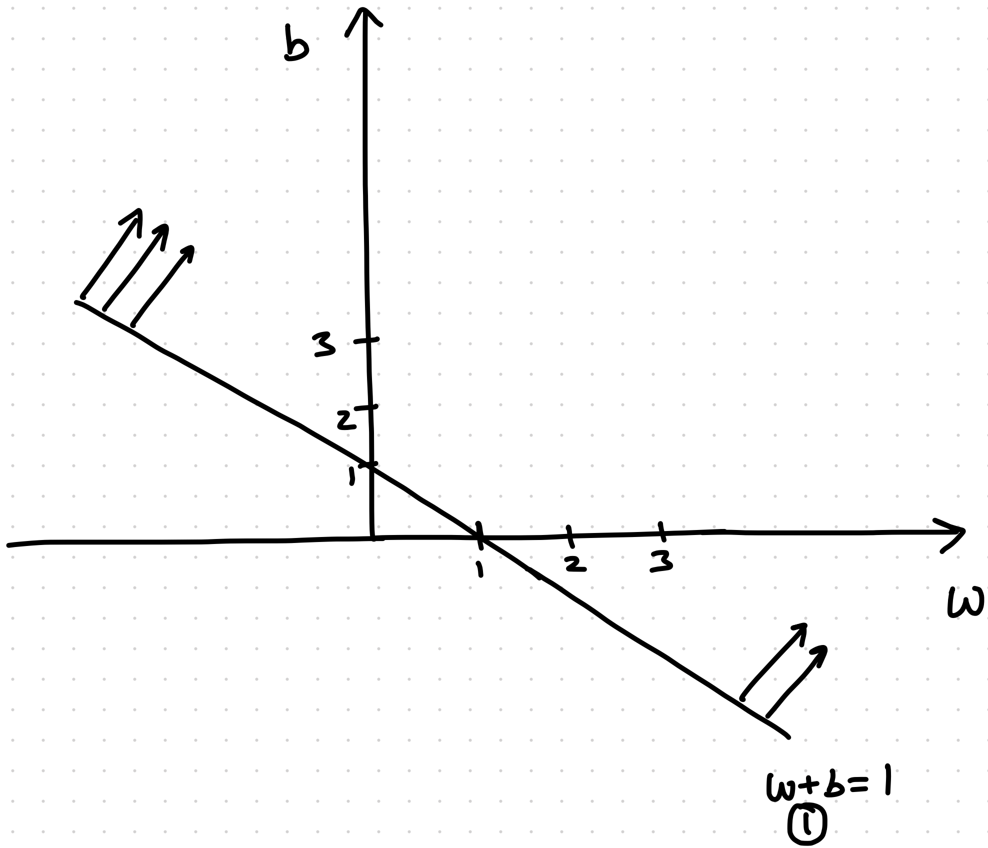
$\vec{w}.\vec{x}+b \geq 1$

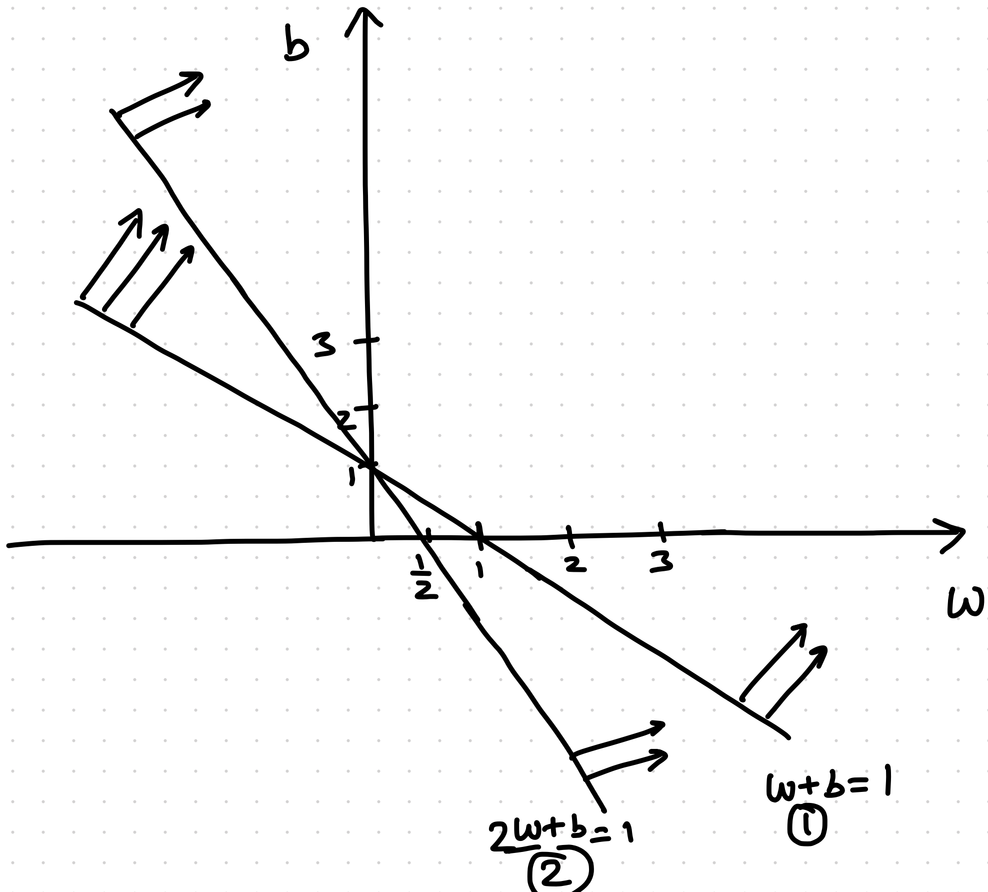
$$\boxed{y_i(\vec{w}.\vec{x}+b) \geq 1}$$

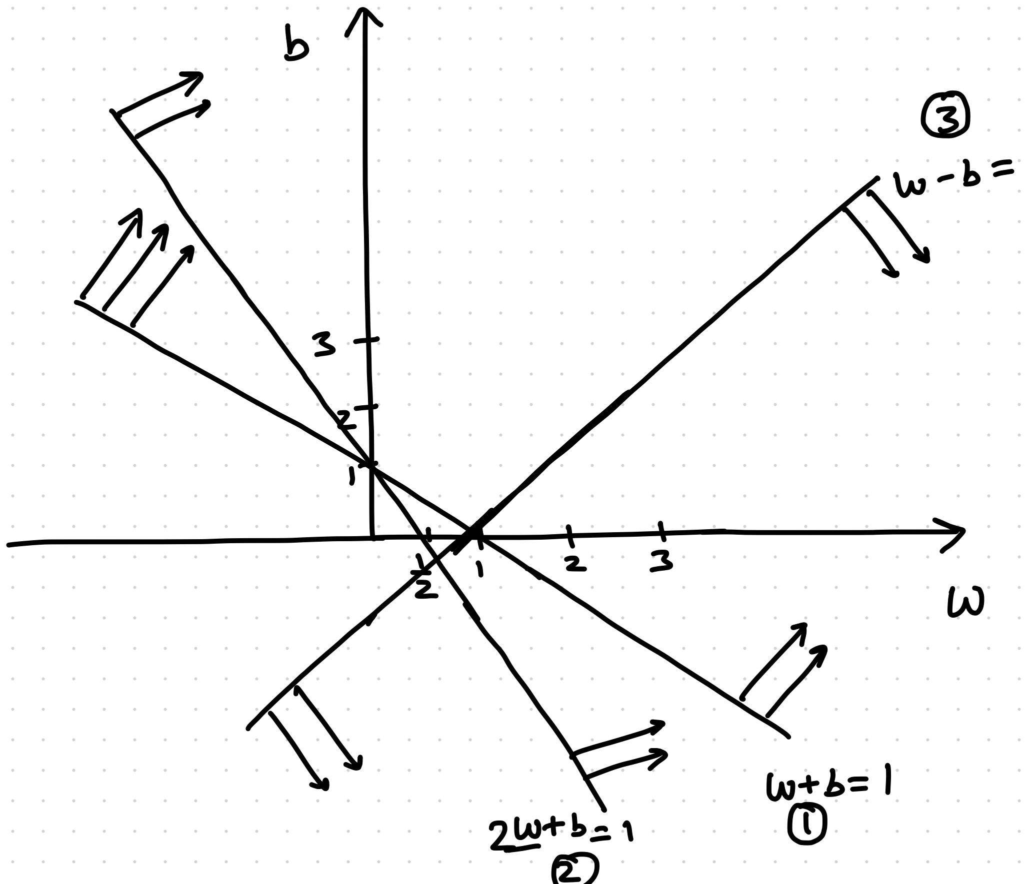
EXAMPLE (IN 1D)



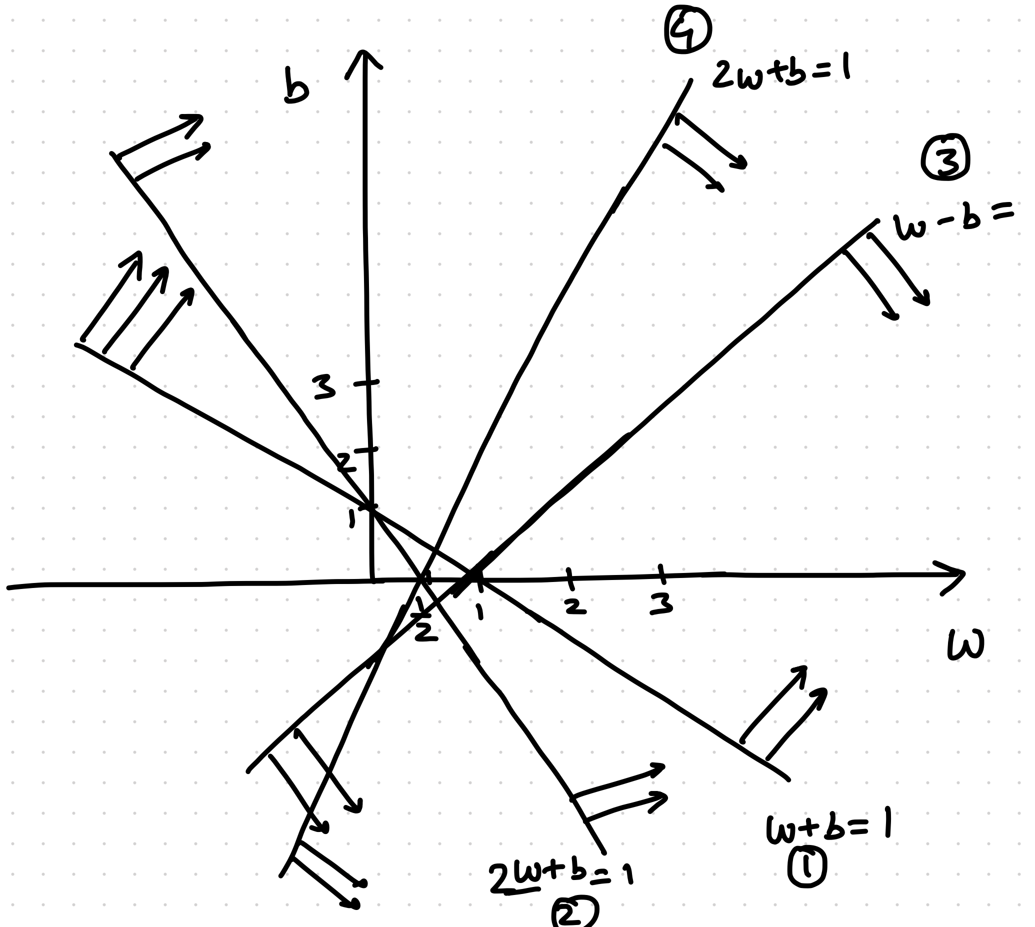

b
3
2
1
1
2
3
w
w+b=1
1


b
3
2
1
1
2
3
ω
w+b=1
1

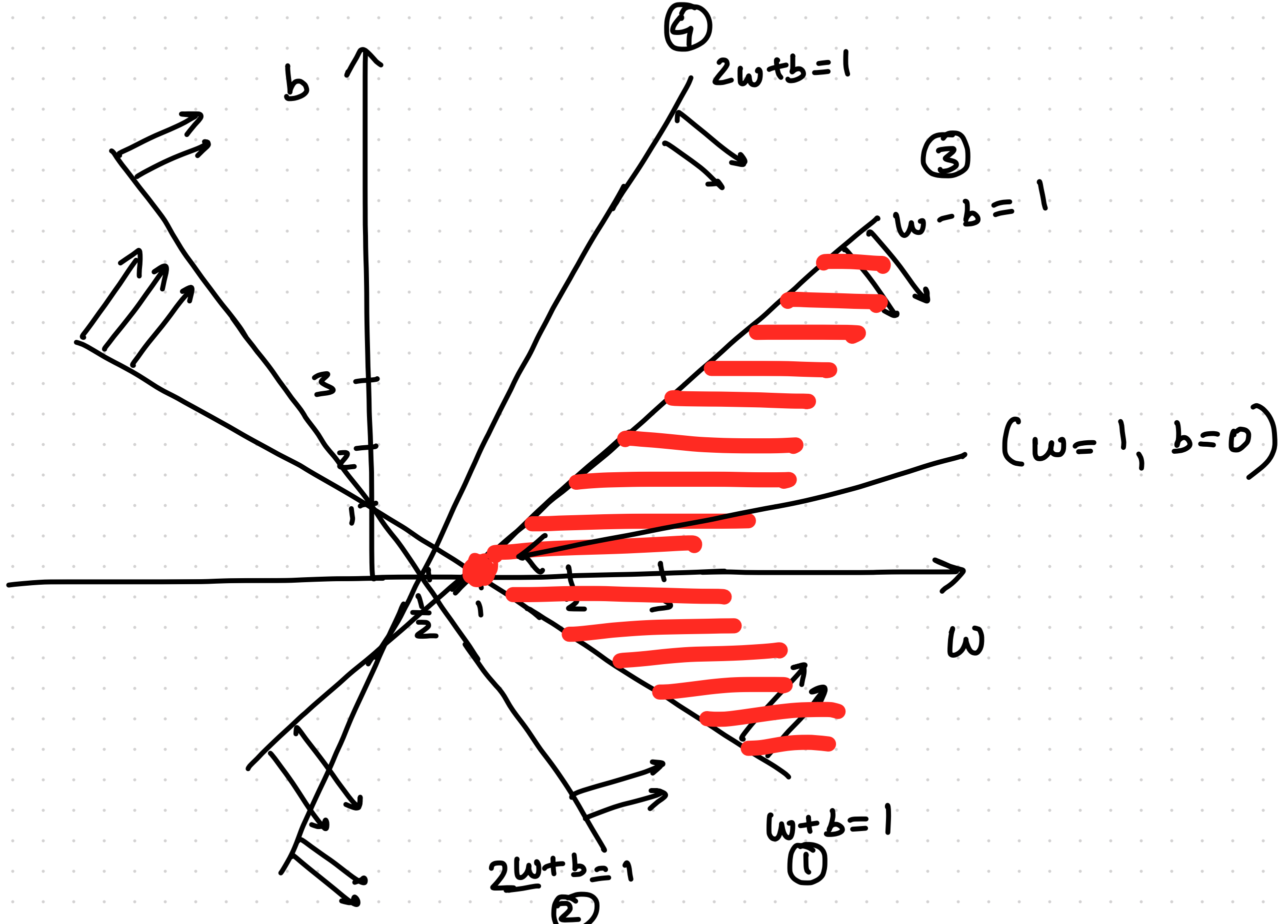

b
3
2
1
1/2
1
2
3
w
w+b=1
1
2w+b=1
2



b
w
3
2
1
1
2
3
1/2
③
w - b = 1
w+b=1
①
2w+b=1
②



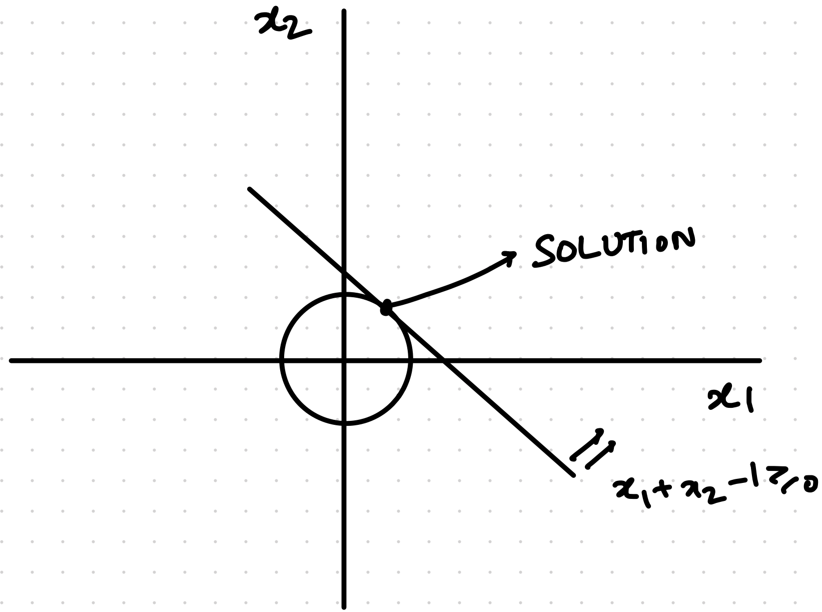
b
w
1
2
3
1
2
3
$\frac{1}{2}$
① $w+b=1$
② $2w+b=1$
③ $w-b=1$
④ $2w+b=1$

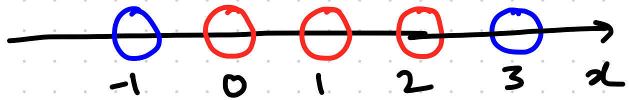


b
w
④
2w+b=1
③
w-b=1
(w=1, b=0)
w+b=1
①
2w+b=1
②
1
2
3
1
2
3
1/2

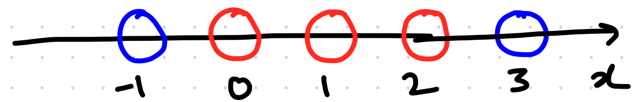
MINIMIZE QUADRATIC

s.t. LINEAR

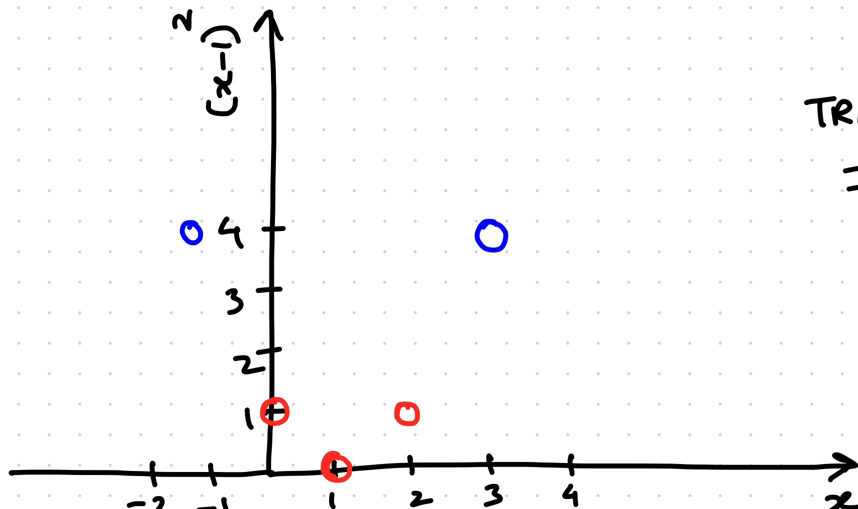




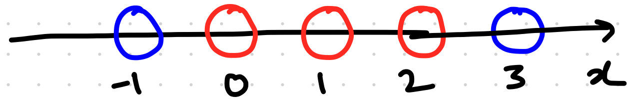
ORIGINAL DATA
IN R



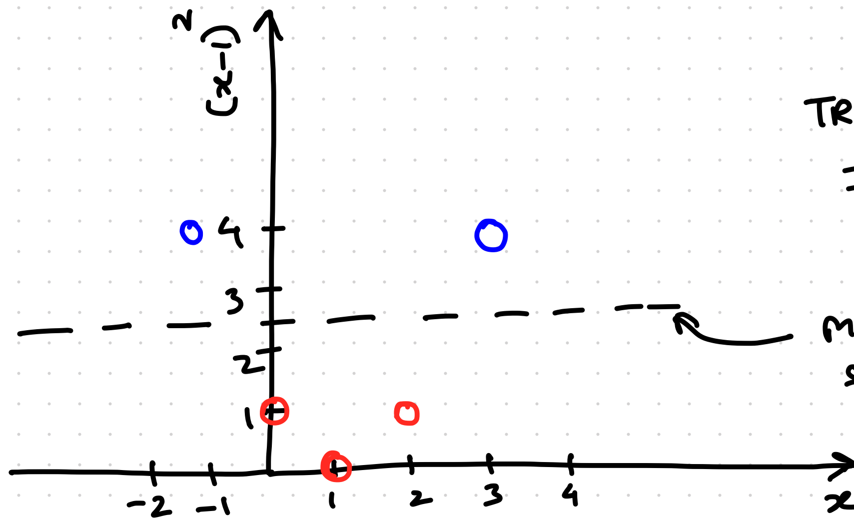
ORIGINAL DATA IN R



TRANSFORMED DATA IN R^2



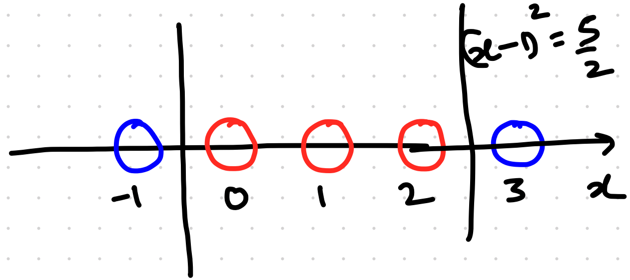
ORIGINAL DATA IN R



TRANSFORMED DATA IN R^2

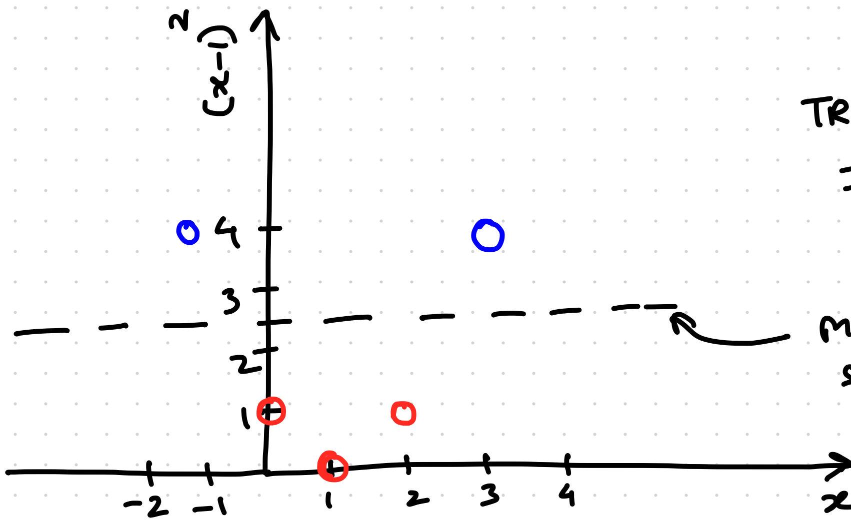
MAX MARGIN SEPARATING HYPERPLANE

$$(x-1)^2 = \frac{5}{2}$$



[DECISION BOUNDARY IN R]

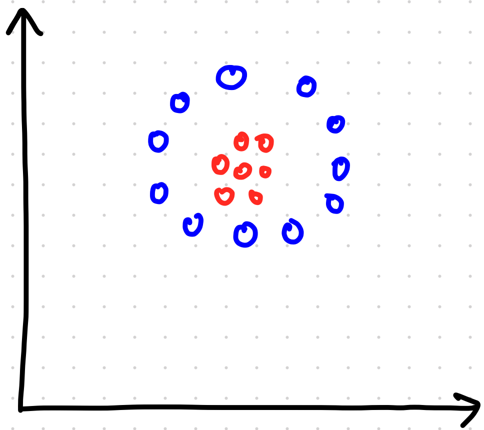
ORIGINAL DATA IN R

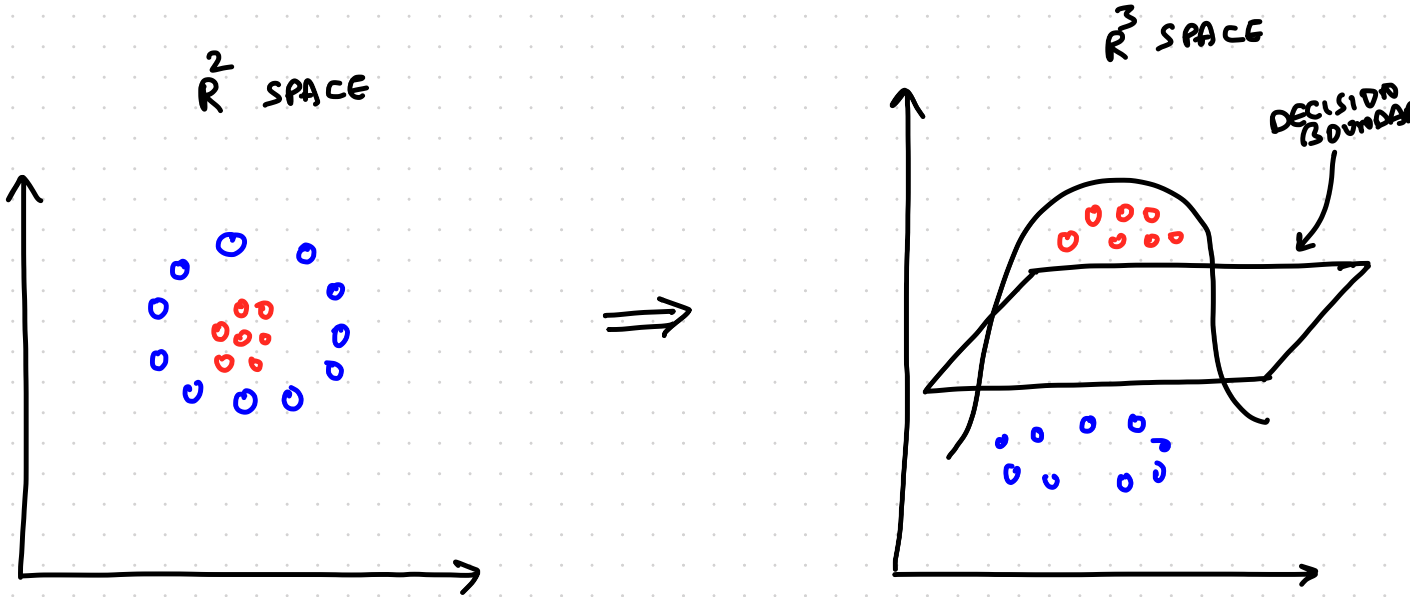


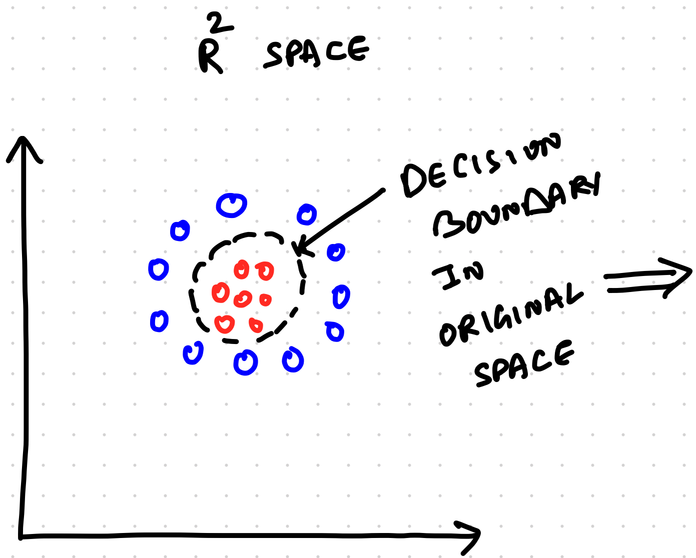
TRANSFORMED DATA IN R^2

MAX MARGIN SEPARATING HYPERPLANE $(x-1)^2 = \frac{5}{2}$

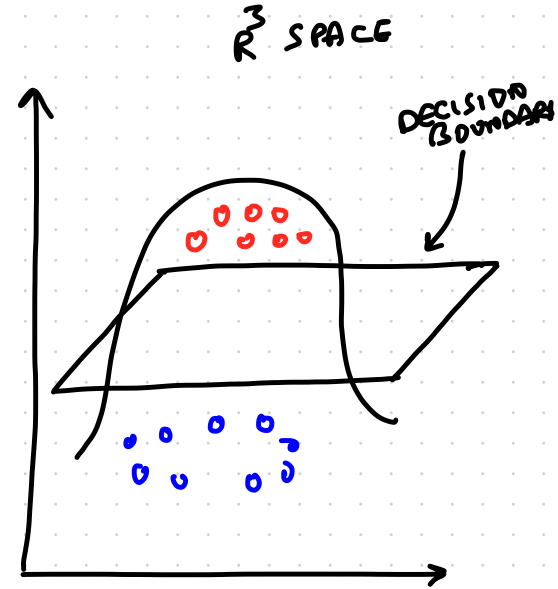
$\mathbb{R}^2$ SPACE



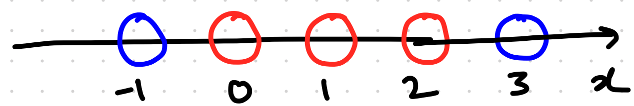

R^2 SPACE
R^3 SPACE
DECISION BOUNDARY



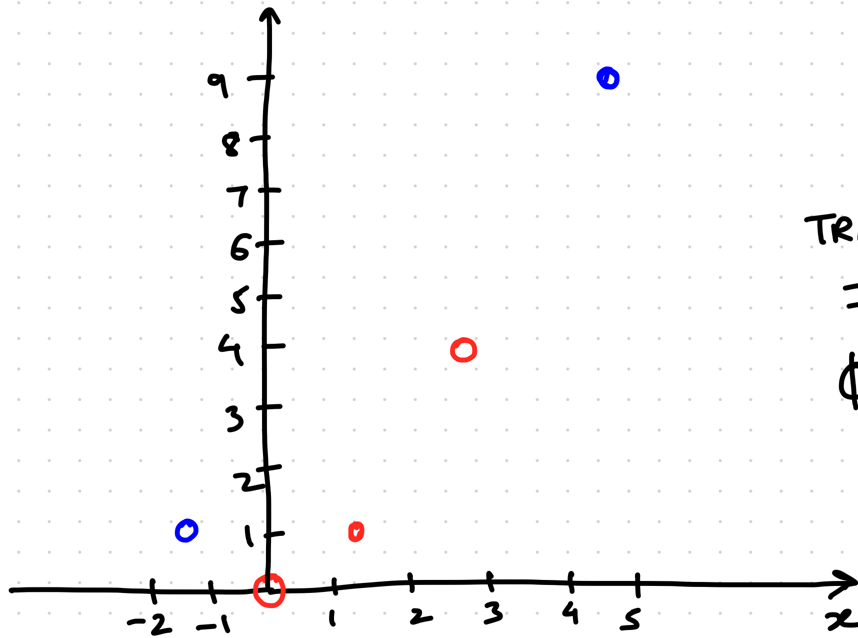
R^2 SPACE
DECISION BOUNDARY IN ORIGINAL SPACE



R^3 SPACE
DECISION BOUNDARY

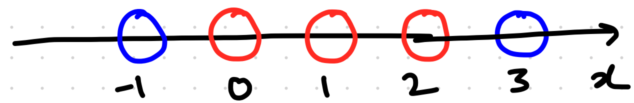


ORIGINAL DATA IN R

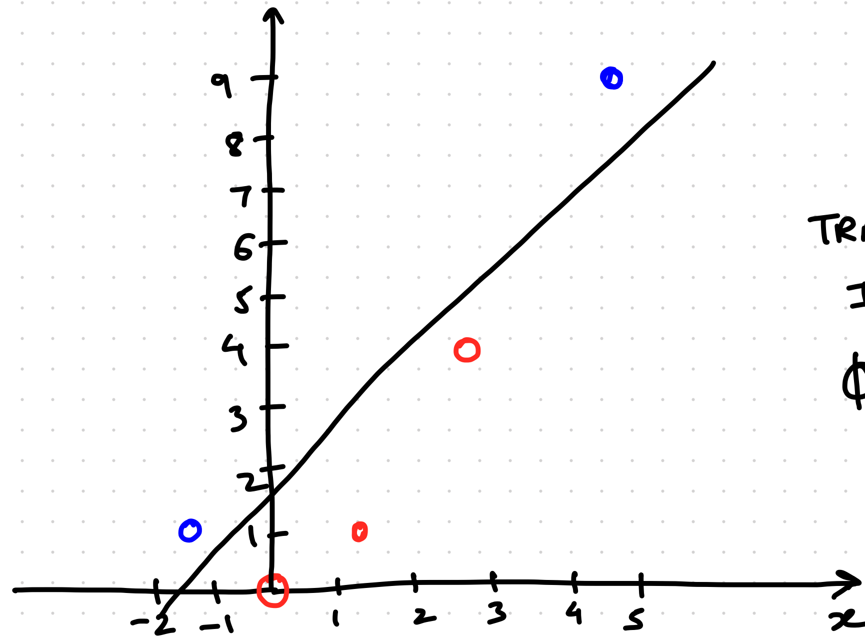


TRANSFORMED DATA IN R^2

$$\phi(x) = \begin{bmatrix} \sqrt{2}\,x \\ x^2 \end{bmatrix}$$

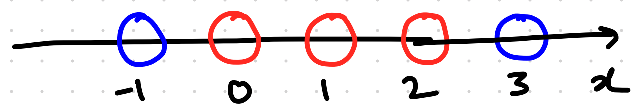


ORIGINAL DATA IN $\mathbb{R}$

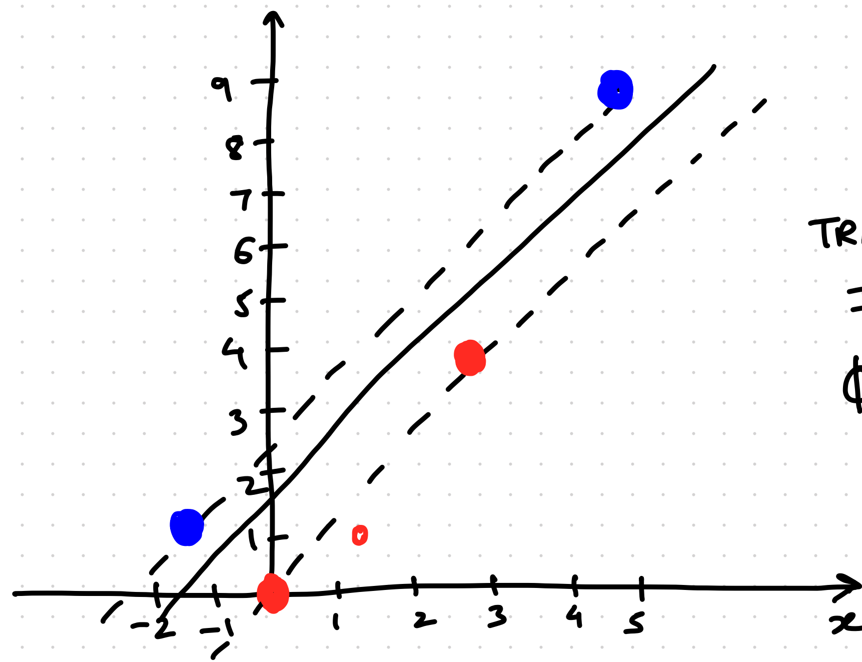


TRANSFORMED DATA IN $\mathbb{R}^2$

$$\phi(x) = \begin{bmatrix} \sqrt{2}\,x \\ x^2 \end{bmatrix}$$

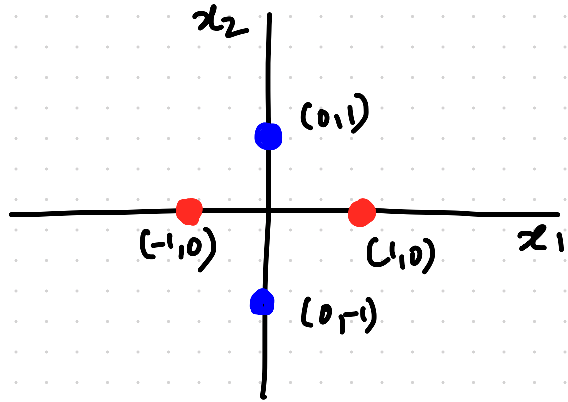


ORIGINAL DATA IN R

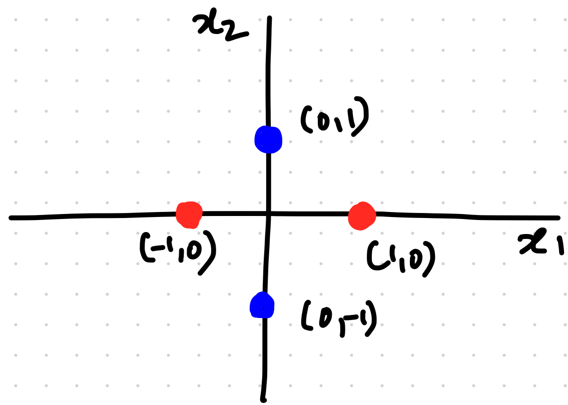


TRANSFORMED DATA IN R^2

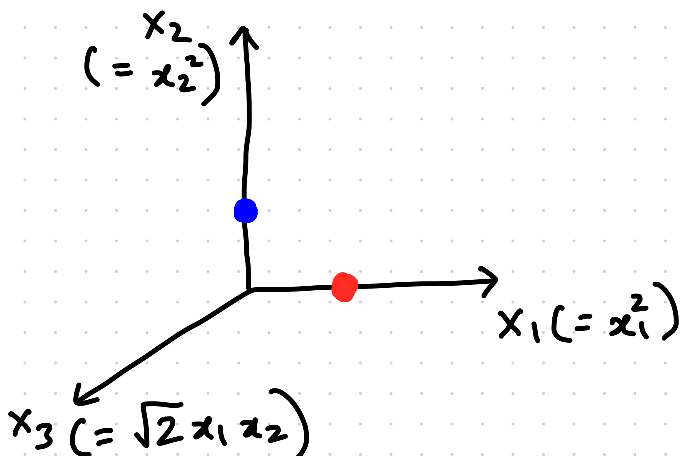
$$\phi(x) = \begin{bmatrix} \sqrt{2}\, x \\ x^2 \end{bmatrix}$$



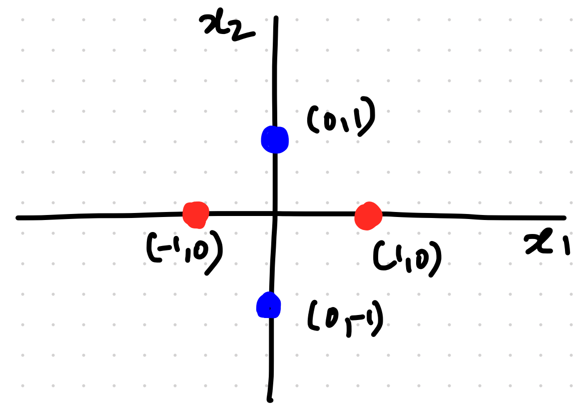
x_2
(0,1)
(-1,0)
(1,0)
x_1
(0,-1)

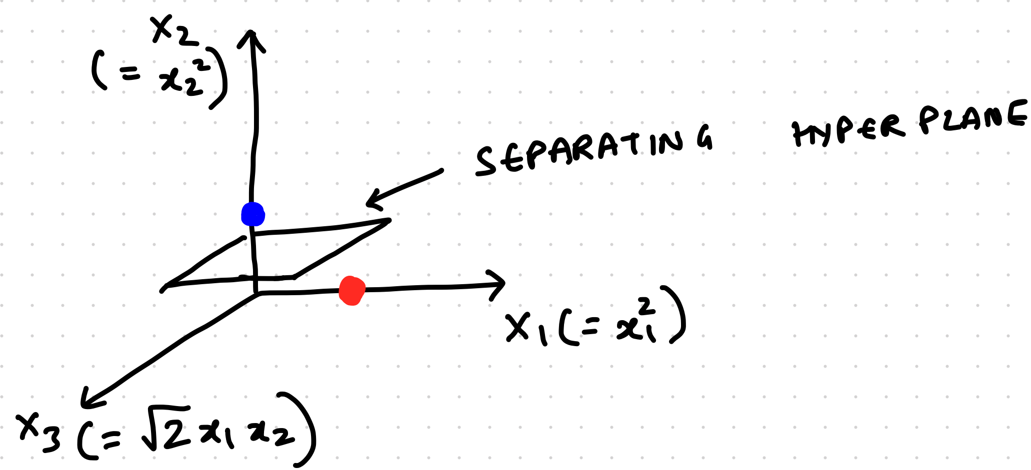


x_2
$(0,1)$
$(-1,0)$
$(1,0)$
x_1
$(0,-1)$

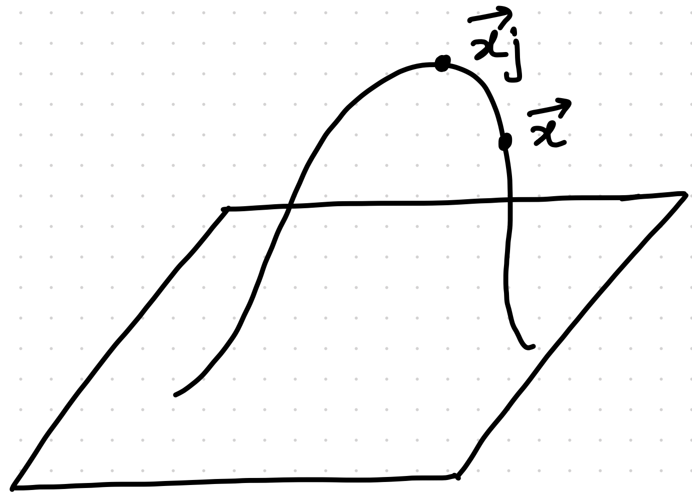


X_2
$(= x_2^2)$
$X_1 (= x_1^2)$
$X_3 (= \sqrt{2} x_1 x_2)$

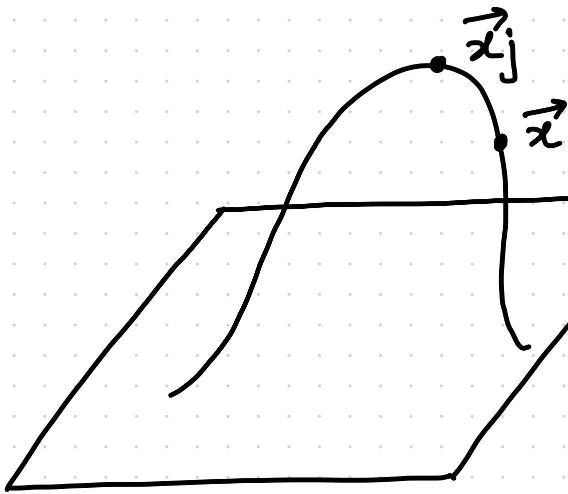

x_2
(0,1)
(-1,0)
(1,0)
x_1
(0,-1)


X_2 $(= x_2^2)$
SEPARATING HYPERPLANE
$X_1 (= x_1^2)$
$X_3 (= \sqrt{2} x_1 x_2)$

RBF INTERPRETATION

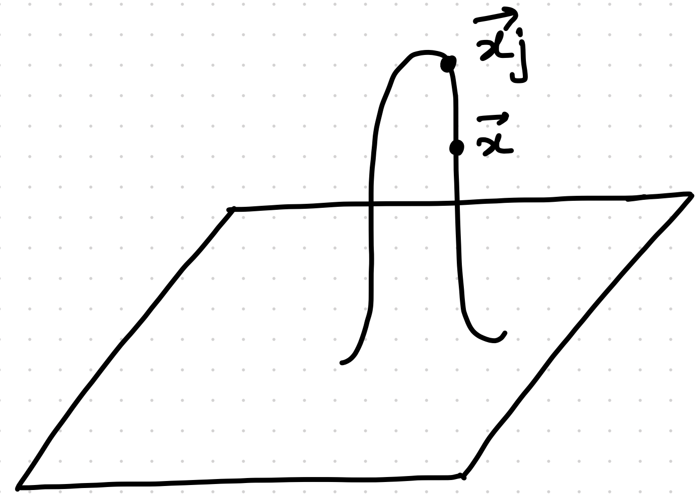


RBF INTERPRETATION



LOW γ

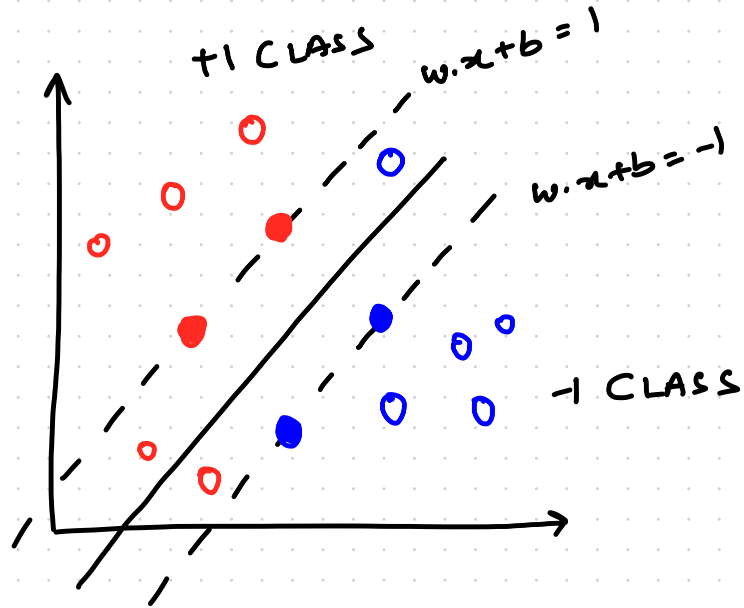
HIGH INFLUENCE OF $\vec{x}_j$



HIGH γ

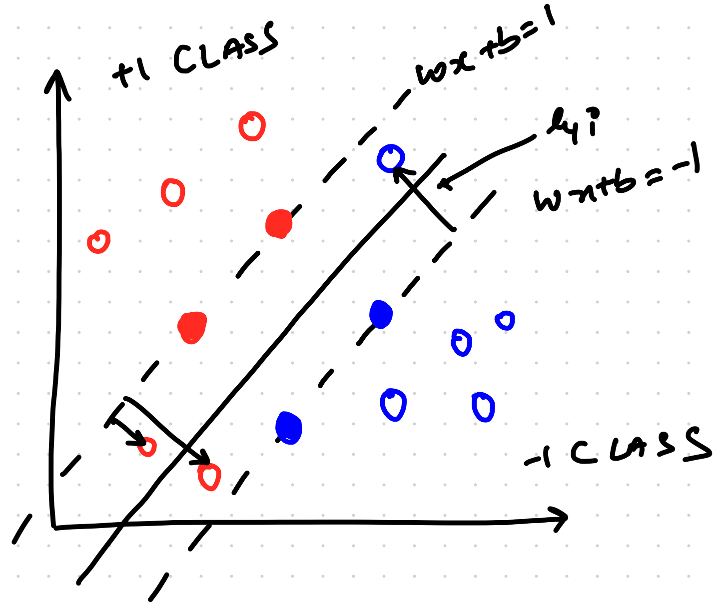
LOW INFLUENCE OF x_j

"SLIGHTLY" NON-SEPARABLE DATA



"SLIGHTLY" NON-SEPARABLE DATA

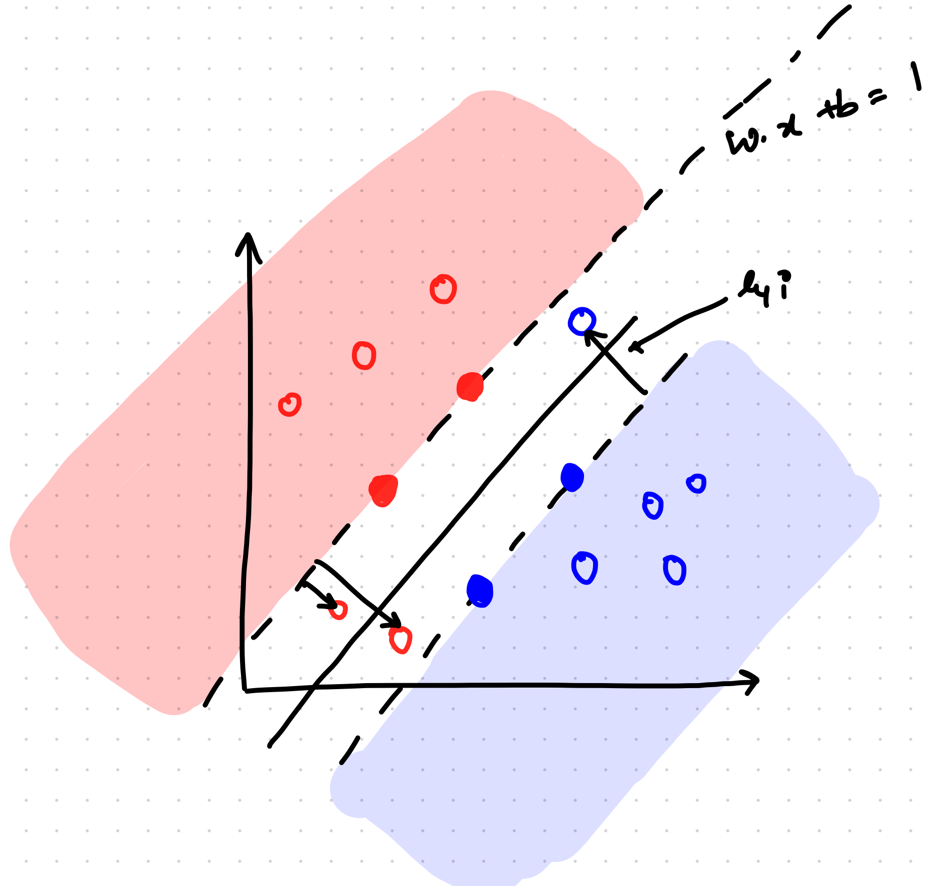
ϵ_i: Distance from margin



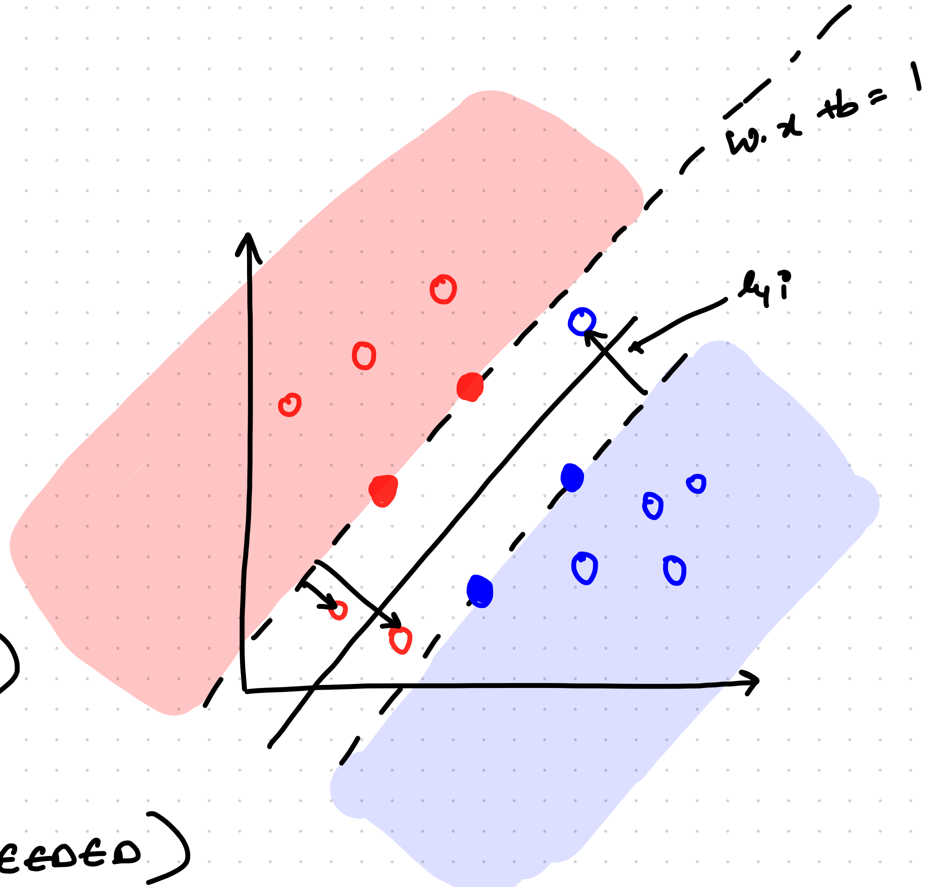
ZONE 1

$y_i = +1$

$w \cdot x_i + b \geq 1$



ZONE 1



$y_i (\vec{w} \cdot \vec{x_i} + b) > 1$

$\text{Loss}_i = 0 \quad (\xi_i = 0)$

$\Downarrow$

NO PENALTY

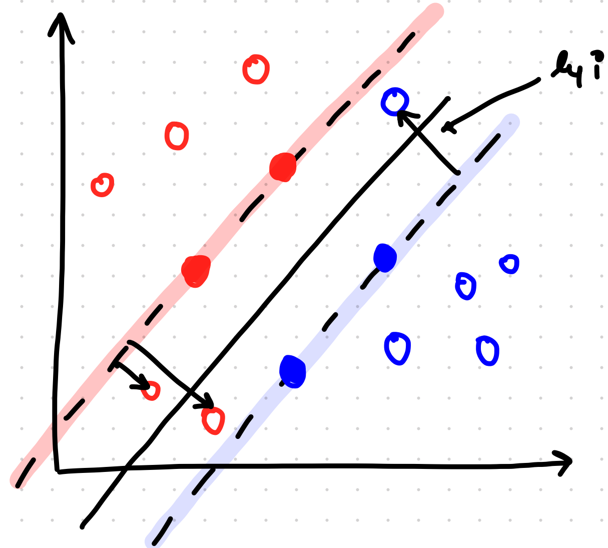
(OR SLACK NEEDED)

ZONE 2

$y_i(\vec{w}.\vec{x}_i + b) = 1$

$LOSS_i = 0$

$(\xi_i = 0)$



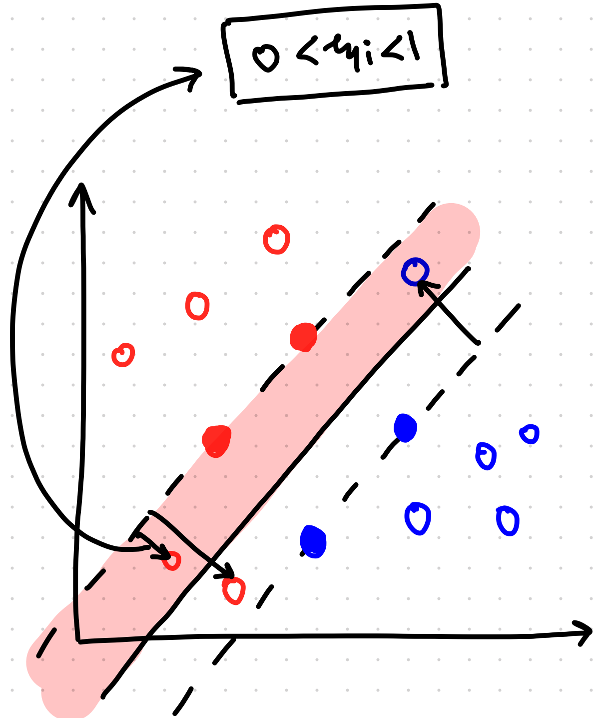
ZONE 3

$y_i(\vec{w}\cdot\vec{x}_i + b) < 1$

$LOSS_i \neq 0 \quad (0 < \xi_i < 1)$

POINT CORRECTLY CLASSIFIED

(BUT WRONG SIDE OF MARGIN)



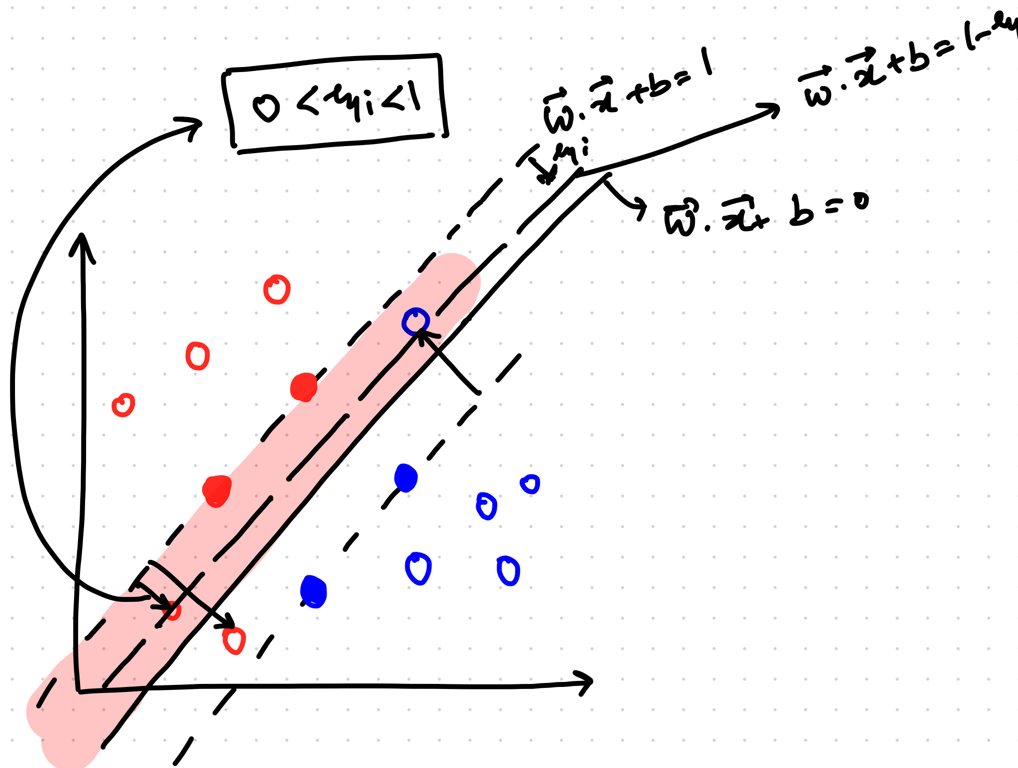
ZONE 3

$y_i(\vec{w}\cdot\vec{x}_i+b)<1$

$LOSS_i \neq 0 \quad (0<\xi_i<1)$

POINT CORRECTLY CLASSIFIED

(BUT WRONG SIDE OF MARGIN)



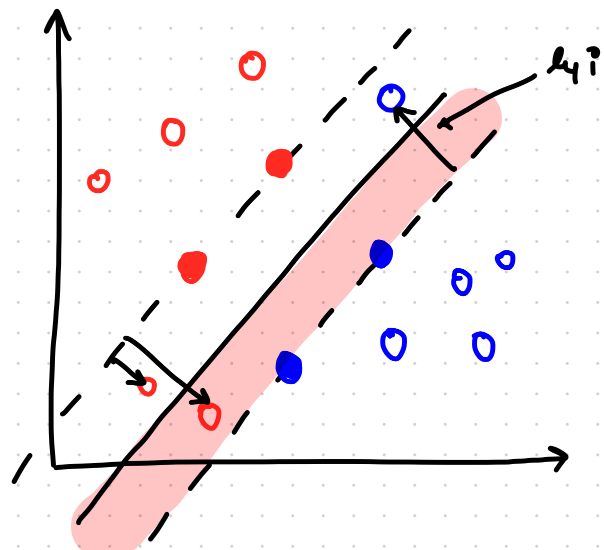
ZONE 4

$y_i(\vec{w} \cdot \vec{x_i} + b) < 1$

POINT INCORRECTLY CLASSIFIED

$LOSS_i \neq 0$

$\xi_i > 1$



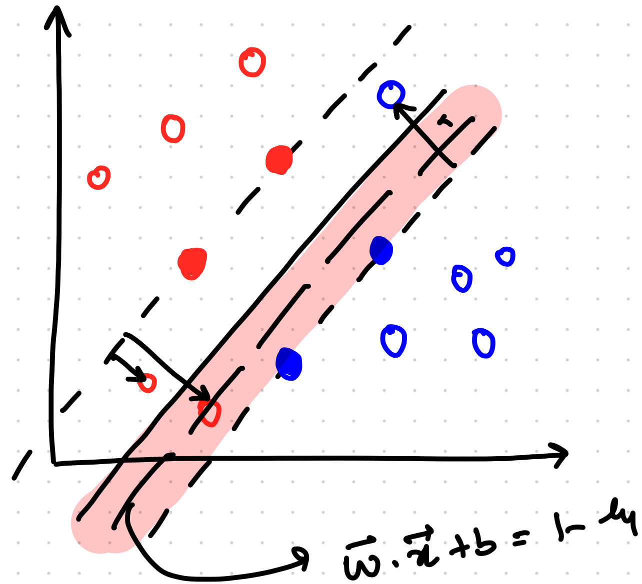
ZONE 4

$$y_i(\vec{w}\cdot\vec{x}_i+b) < 1$$

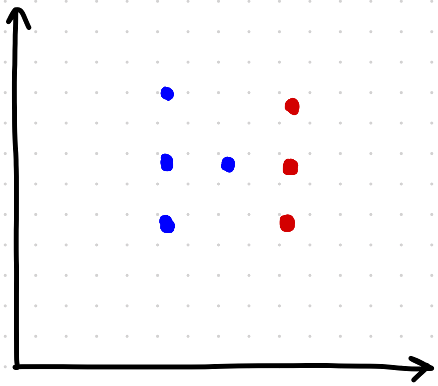
POINT INCORRECTLY CLASSIFIED

$LOSS_i \neq 0$

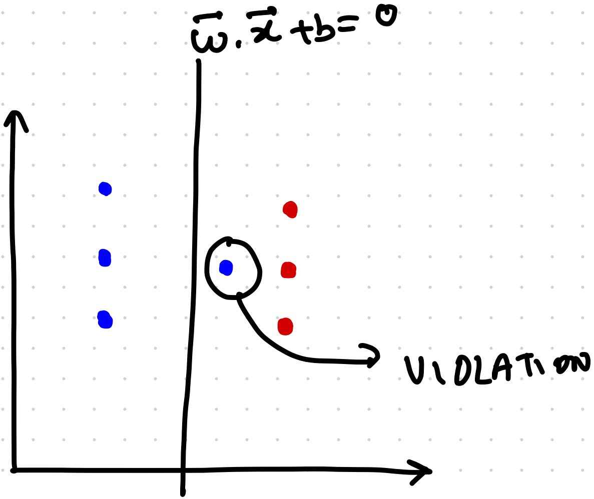
$\xi_i > 1$

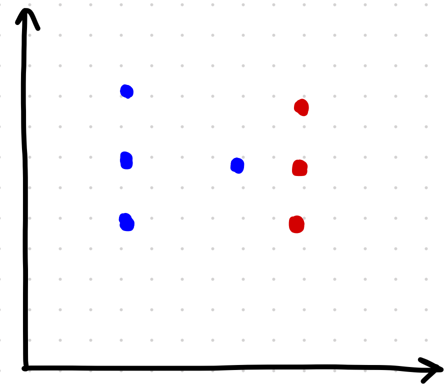


BIAS- VARIANCE TRADE-OFF



BIAS-VARIANCE TRADE-OFF





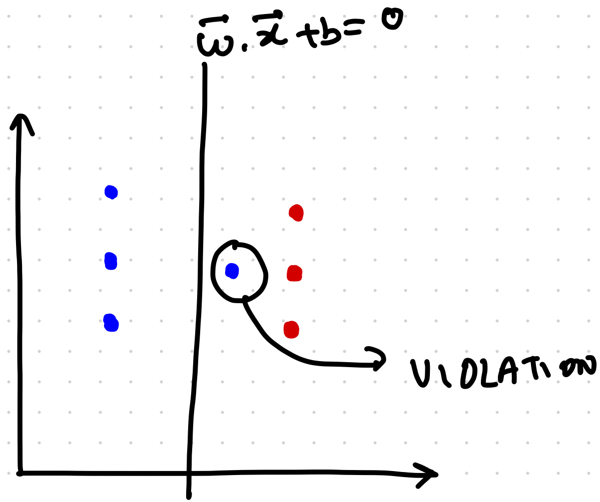
LOW C

LOW PENALTY FOR VIOLATION

HIGH TRAIN ERROR

HIGH BIAS

BIAS-VARIANCE TRADE-OFF



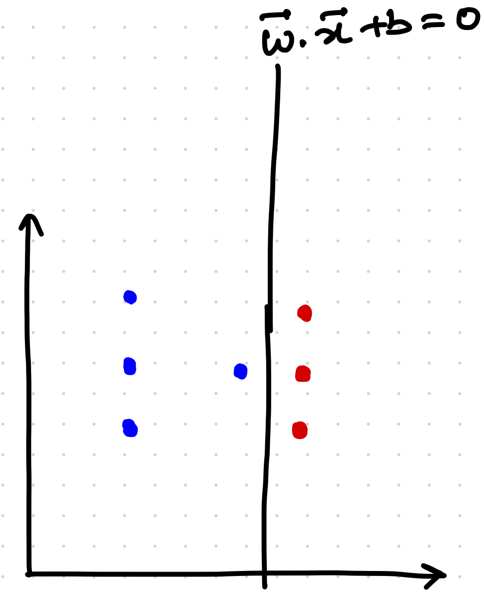
LOW C

LOW PENALTY FOR VIOLATION

HIGH TRAIN ERROR

HIGH BIAS

BIG MARGIN



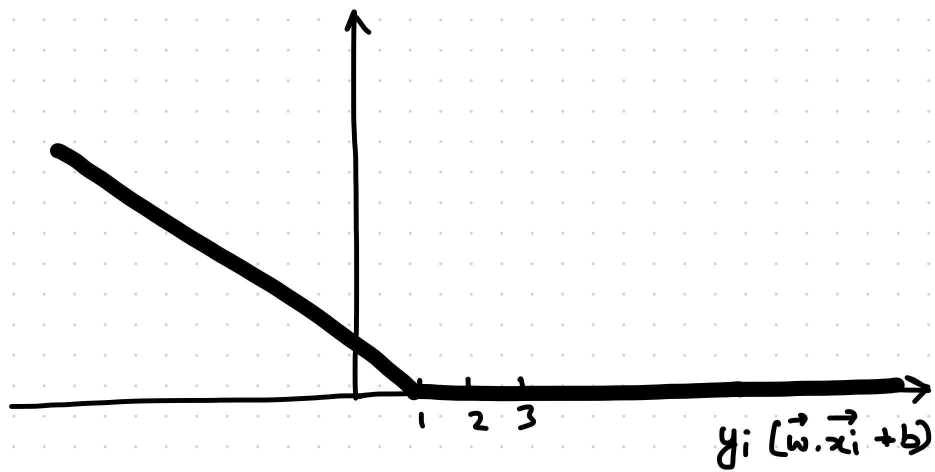
HIGH C

HIGH PENALTY FOR VIOLATION

HIGH VARIANCE

SMALL MARGIN

HINGE LOSS



ℓ_1- SVR