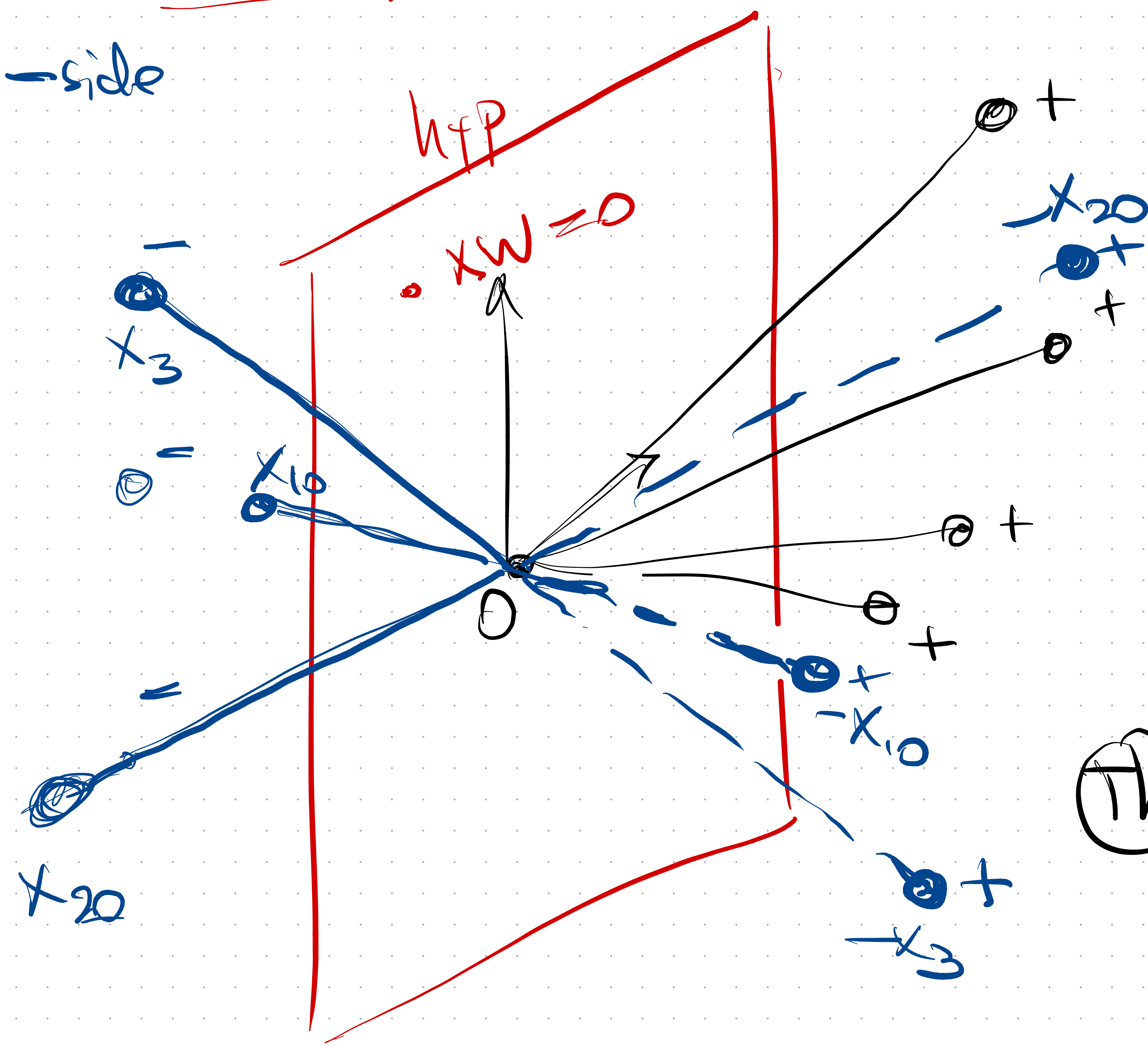
5/28: Perceptron: A linear model (simple) $h(x_i) = x_i w$ label $y_i \to +, -$

w = coef.

New objective (geom) we want hyplane [$xw = 0$] to separate +/−



Step 1 (before training): convert neg datapoints ($y = -$) into positive examples:

$$\begin{cases} x_i = -x_i \\ y_i = -y_i \end{cases}$$ for negatives

⇒ All points pos $y_i = +$

new OBJ: have all points on positive side

Th: hyperplane 100% sep original data ⟺ hypplane has 100% new data on ⊕ side.

Perceptron problem (modified data)

- Given N vectors $x_1, x_2, \ldots x_N$
- find hyplane w s.t. $x_i w \geq 0 \ \forall i$

(puts all points on pos side)

ERROR $J(w) = \sum_{x_i} \underbrace{-x_i w}_{\text{pos}} \geq 0$

$\boxed{\begin{array}{c}\text{mistake} \\ x_i w < 0\end{array}}$

how far from hyplane x_i is

not mistake (correct) $x_i w \geq 0$

OBS: no error $\Rightarrow$ no mistakes

$\Rightarrow J(w) = 0$

$x_i w < 0$

update: Every mistake x_i gets added to w = normal of hyplane

Gradient Desc update

$$\frac{\partial J}{\partial w^d} = -\frac{\partial x_i w}{\partial w^d} = -\frac{\partial\left[\sum_d x_i^d w^d\right]}{\partial w^d}$$

$$= -x_i^d$$

if x_i = mistake

$$\frac{\partial J}{\partial w} = -x_i^T$$

Vector

GD update for datap x_i

$$w_{new} = w_{old} + \lambda' x_i^T$$

all datapoints x_i = mistakes

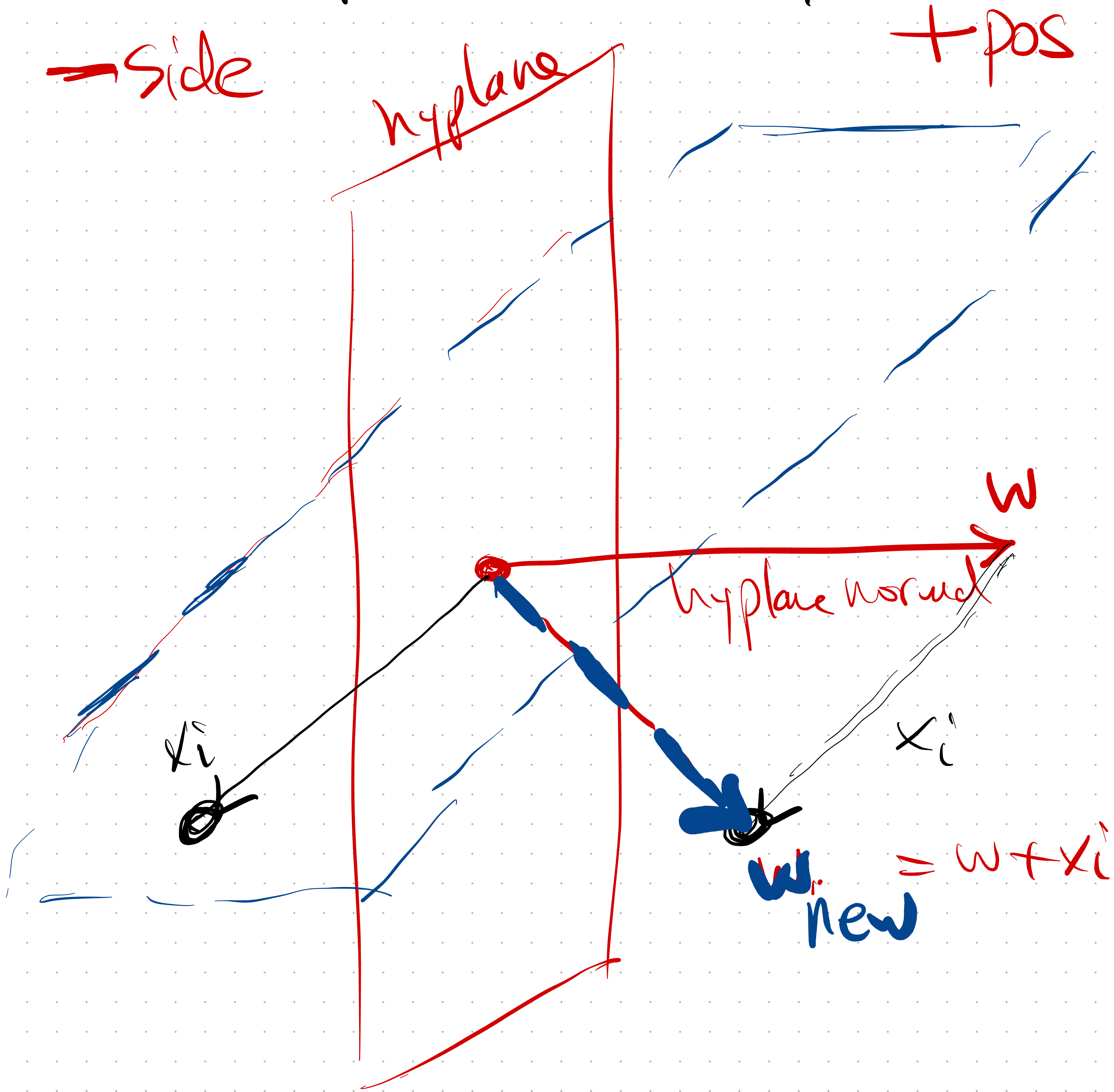
$$w_{new} = w_{old} + \lambda \sum_{\substack{x_i \text{ mistakes} \\ \{x_i w < 0\}}} x_i^T$$

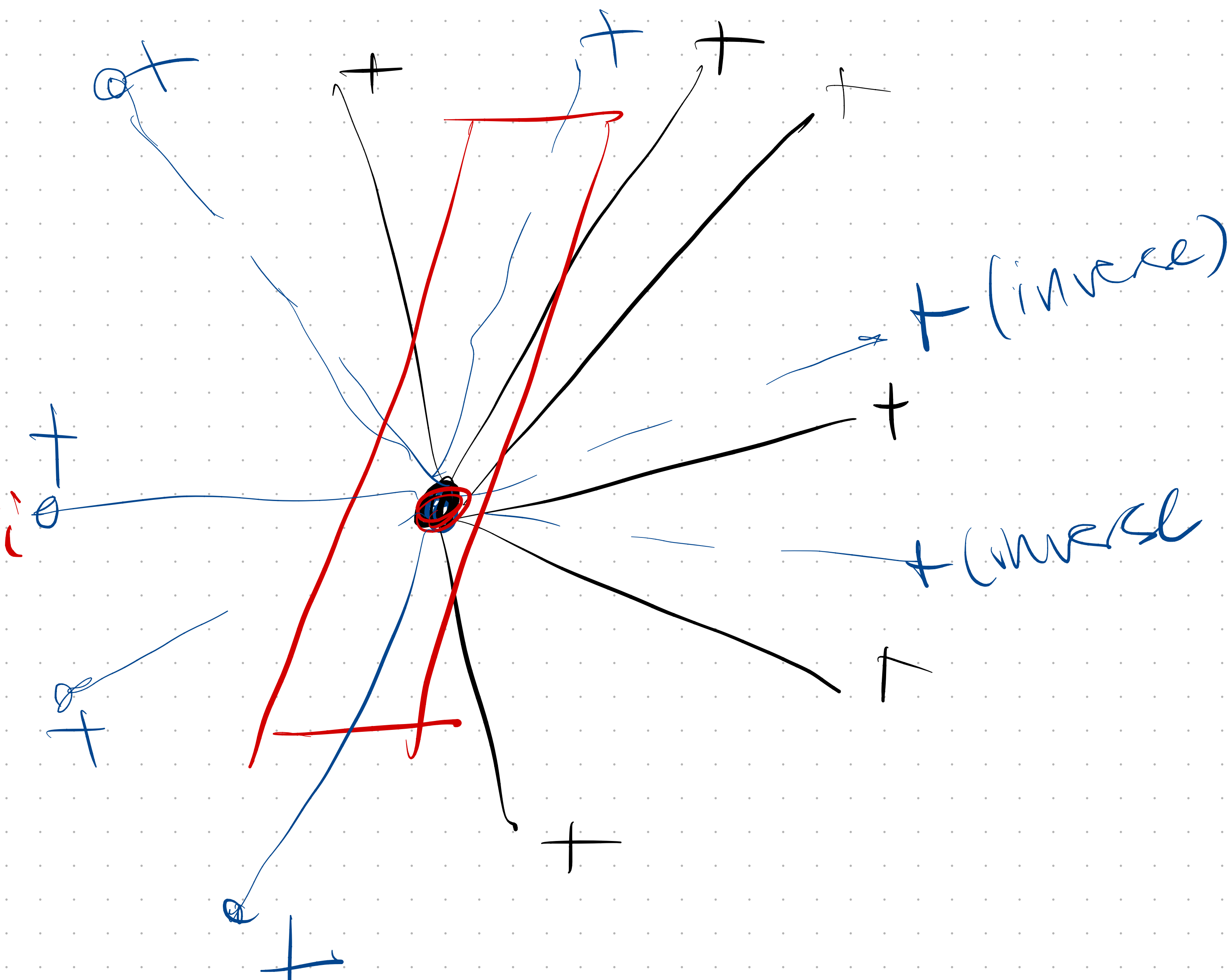
GD. upd. Geometrically

x_i = mistake $(x_i w < 0)$

add x_i to w $(w_{new} = w + x_i)$

Geom: hyplane rotated **TOWARDS** x_i





Th $x_1 x_2 x_3 \ldots x_N$ are linearly separable

$\Rightarrow$ (after step 1) all points can be on $\oplus$ side of hyplane $\bar{w}$ $\downarrow$ 0% error separator

Then Perceptron Alg will find a separator (0 mistakes) after finite # iterations

proof: w_k = hypplane after k iterations.

update: one mistake at a time

x_k = datapoint (mistake) that updated w_k $\rightarrow$ $x_k w_k < 0$

α = scalar/const to be decided later

$$\underset{\text{(new)}}{w_{k+1}} = w_k + x_k^T \qquad (\lambda = 1 \text{ simple})$$

$\exists$ separator

$$w_{k+1} - \alpha \boxed{\bar{w}} = \cancel{w}_k - \alpha \bar{w} + x_k^T$$

$$\|w_{k+1} - \alpha\bar{w}\|^2 = \|w_k - \alpha\bar{w} + x_k^T\|^2 = \|w_k - \alpha\bar{w}\|^2 + 2x_k(w_k - \alpha\bar{w}) + \|x_k\|^2$$

$$= \|w_k - \alpha\bar{w}\|^2 + \underbrace{2x_k w_k}_{\text{negative}} - 2x_k\alpha\bar{w} + \|x_k\|^2$$

x_k mistake for w_k

$$\le \|w_k - \alpha\bar{w}\|^2 \underbrace{- 2x_k\alpha\bar{w} + \|x_k\|^2}$$

choose $\alpha > 0$ large enough such that

strict pos $> \varepsilon$ fixed

use $x_k\bar{w} > 0$

$\bar{w}$ hyp all points pos side

$$\le \|w_k - \alpha\bar{w}\|^2 - \varepsilon$$

$$\|w_{k+1} - \alpha\bar{w}\|^2 \le \|w_k - \alpha\bar{w}\|^2 - \varepsilon$$

$\Rightarrow$ this cannot go on forever!

- has to stop: no x_k mistake $\Rightarrow$ 0% error
- $\bar{w}$ does not exist: there is no hyperplane separator