Linear models for classification

$D = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ $\quad x_i \in \mathbb{R}^d$

$y_i \in \{0, 1\}$ or $y_i \in \{\pm 1\}$.

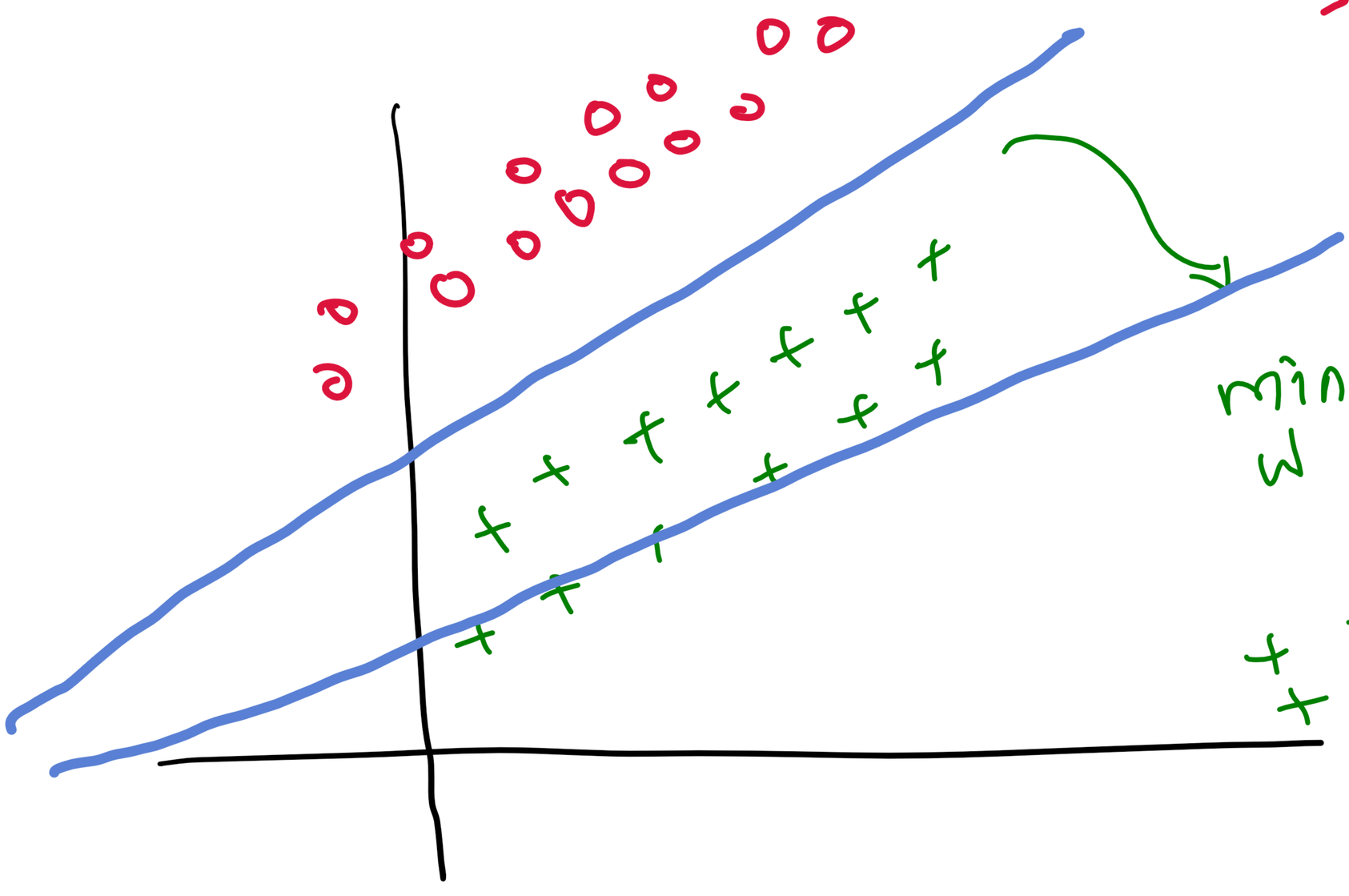
$x_i \rightarrow$ email i

$y_i \rightarrow$ spam / not-spam

In regression $y_i \in \mathbb{R}$

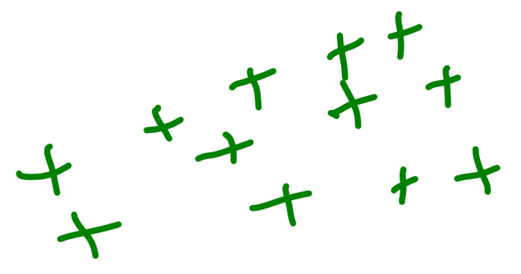
$h(x) = \boxed{\text{sign}(w^T x)}$ → Linear regression

$\text{sign}(z) = \begin{cases} +1 & \text{if } z \geq \\ -1 & \text{o/w} \end{cases}$



$\min_w \sum_{i=1}^{n} (\vec{w} x_i - y_i)^2$

$\in \mathbb{R}$.



Regression

$$\min_{w} \sum_{i=1}^{n} (w^T x_i - y_i)^2$$

Classification

$$\min_{w \in \mathbb{R}^d} \sum_{i=1}^{n} \mathbb{1}\left(\text{sign}(w^T x_i) \neq y_i\right)$$

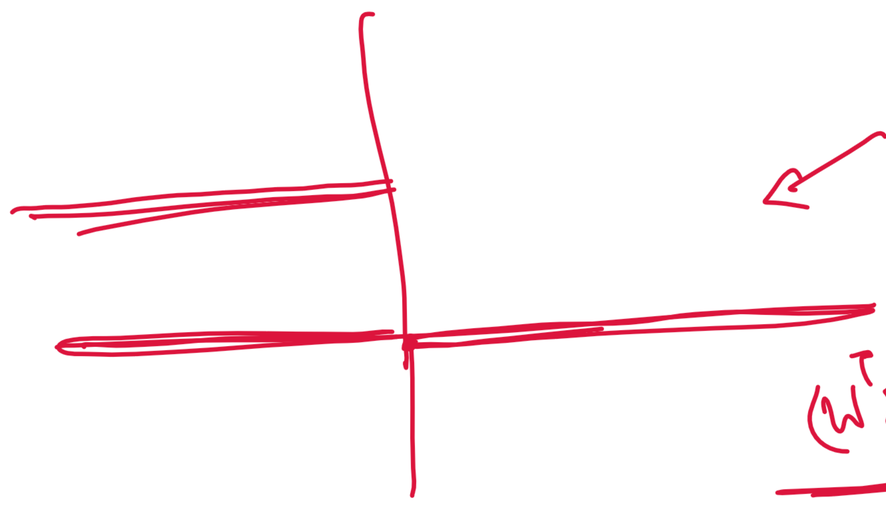
Prediction

$$\mathbb{1}(z) = \begin{cases} 1 & \text{if } z \text{ is true} \\ 0 & \text{o/w} \end{cases}$$

$$\min_{w \in \mathbb{R}^d} \sum_{i=1}^{n} \mathbb{1}\left((w^T x_i) y_i < 0\right)$$

$y_i \in \{\pm 1\}$

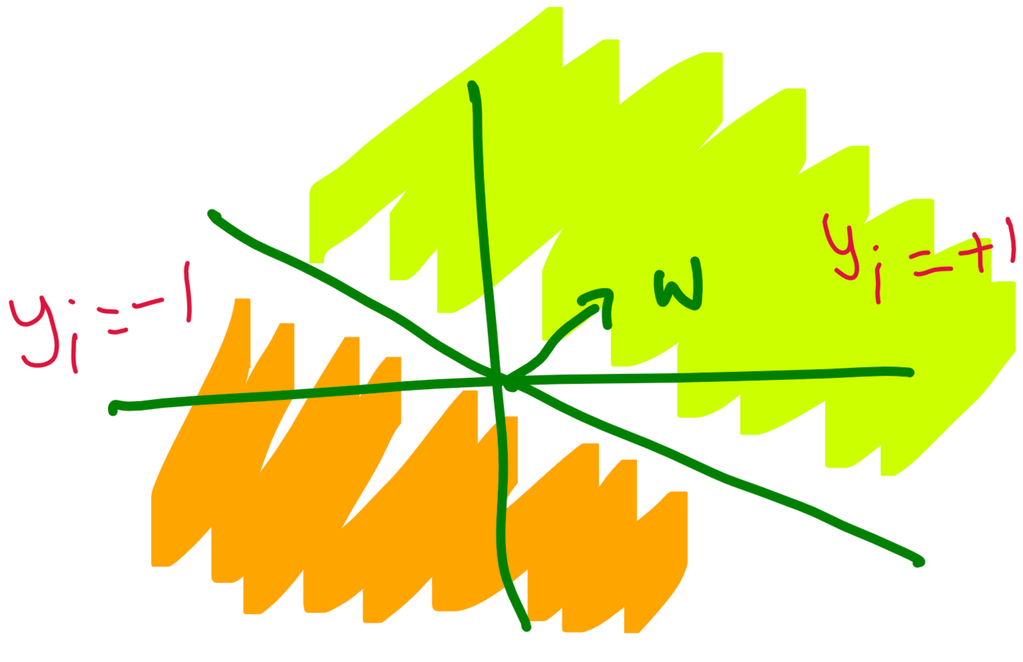
NP-hard.



"Linear separability" assumption

$$\exists w \in \mathbb{R}^d \quad \text{s.t} \quad \forall x \in \mathbb{R}$$

$$y_x = \text{sign}(w^T x)$$



PERCEPTRON [Rosen Blatt 1960's]

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

$y_i \in \{+1, -1\}$

Init: $w_0 = 0 \in \mathbb{R}^d$ ← iteration.

until Convergence

- Pick (x_i, y_i) from dataset
- if $(\text{sign}(w_t^T x_i) = y_i)$

  do nothing

  else

$$w_{t+1} = w_t + x_i y_i$$ ← UPDATE STEP

end.

$\forall x_i, \text{sign}(w_t^T x_i) = y_i$

$$w_{t+1} = w_t + x_i y_i$$

A data point where a mistake was made.

Mistake 1

$w_t^T x_i > 0$

$y_i = -1$

Mistake 2

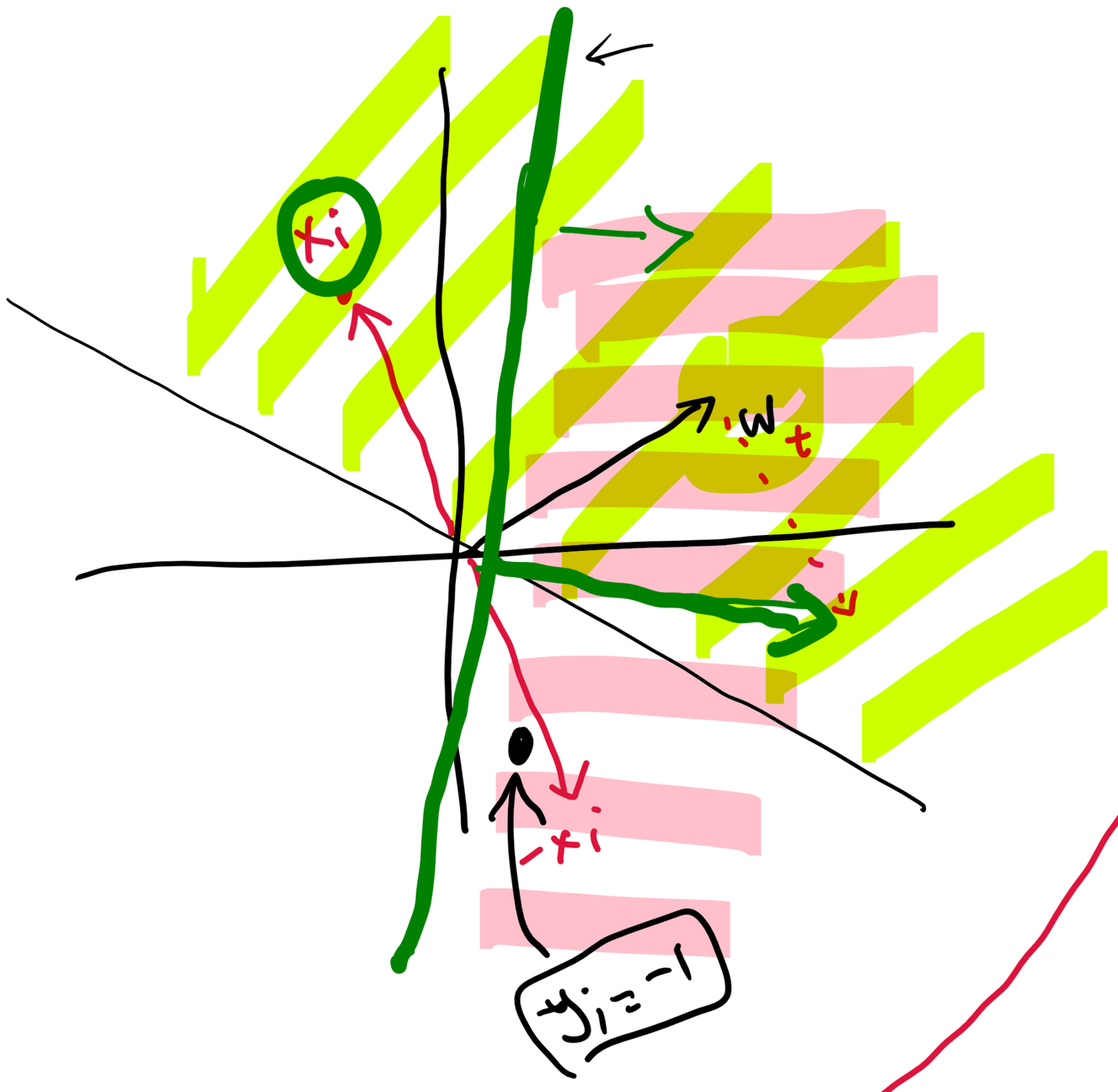
$w_t^T x_i < 0$

$y_i = +1$

$$w_{t+1}^T x_i = (w_t + x_i y_i)^T x_i$$

$$= \underbrace{w_t^T x_i}_{>0} + \underbrace{y_i}_{-1} \underbrace{\|x_i\|_2^2}_{>0} \leftarrow \text{negative.}$$

< 0 + Positive.

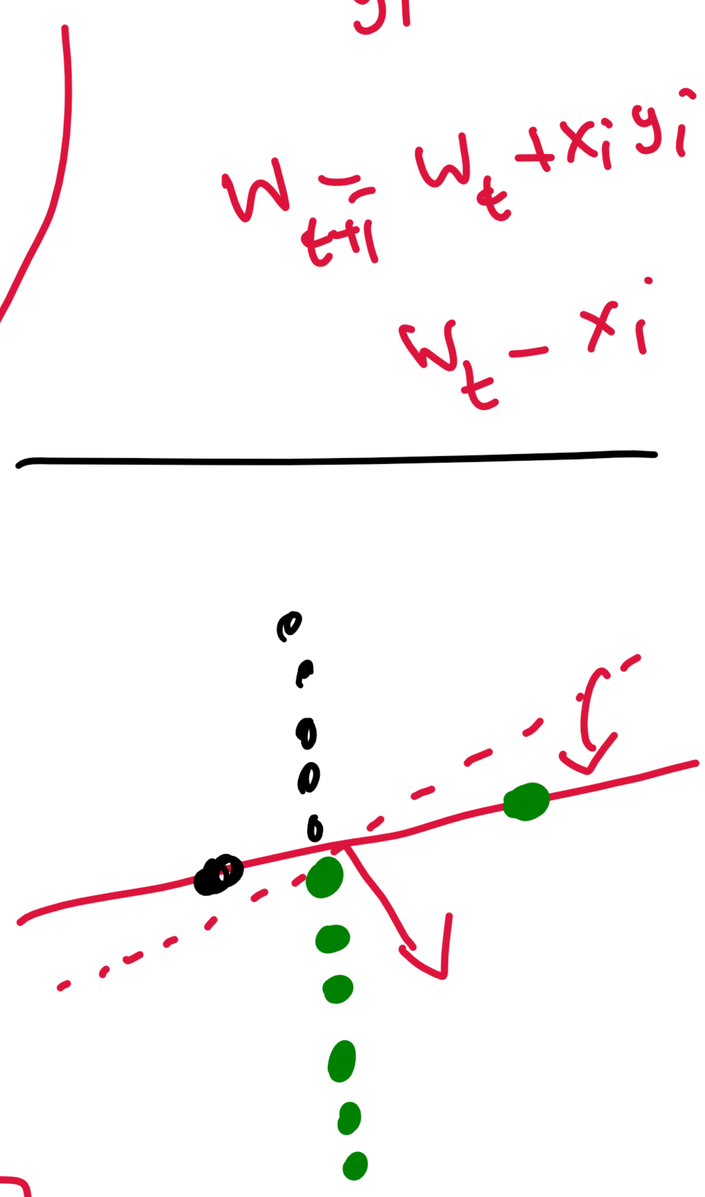


$w^T x_i > 0$ $+1$

$y_i = -1$

$w_{t+1} = w_t + x_i y_i$

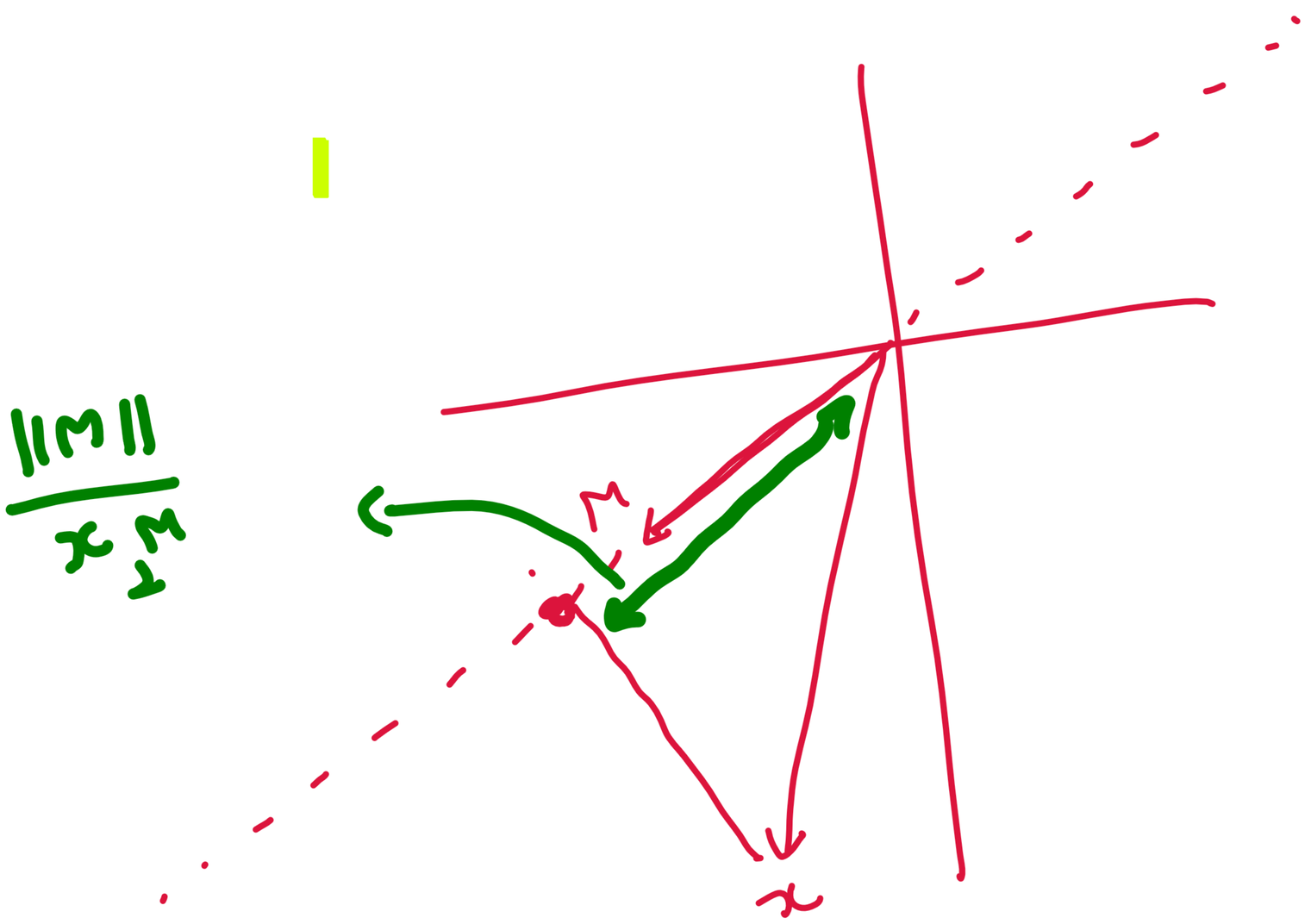
$w_t - x_i$

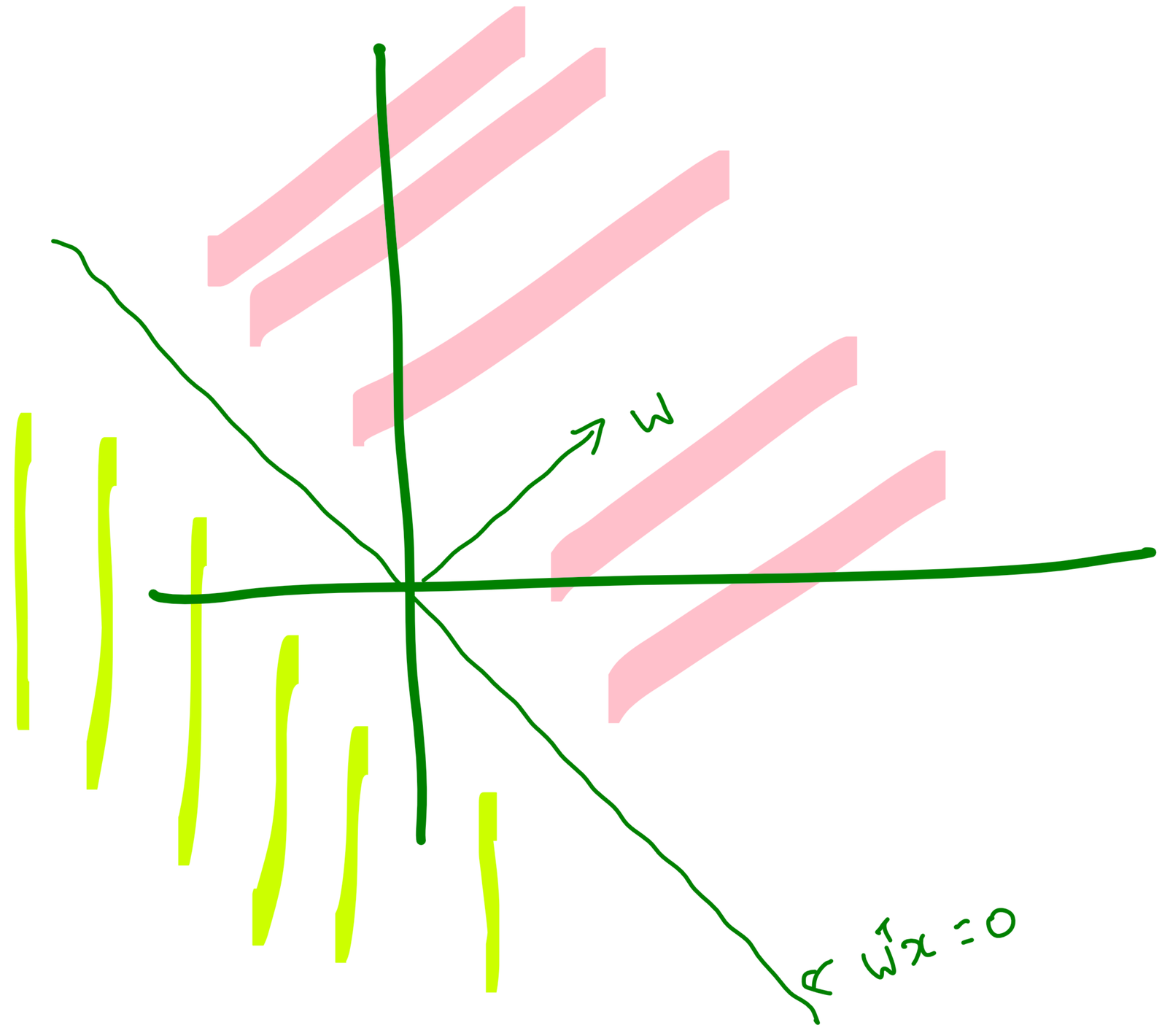


$$w = [w_1 \; w_2 \; \cdots \; w_d]$$

$$x = [x_1 \; x_2 \; \cdots \; x_d]$$

$$w^T x = \sum_{i=1}^{d} w_i x_i$$


$\frac{\|M\|}{x_1^M}$
M
x



$\{x : \vec{w}^T x = 0\}$

$= ?$

$\{x : \vec{w}^T x > 0\}$