Supervised Learning :

$X \longrightarrow y \Rightarrow f(x) = y$

entradas salidas

targets $f(\ ) = ?$

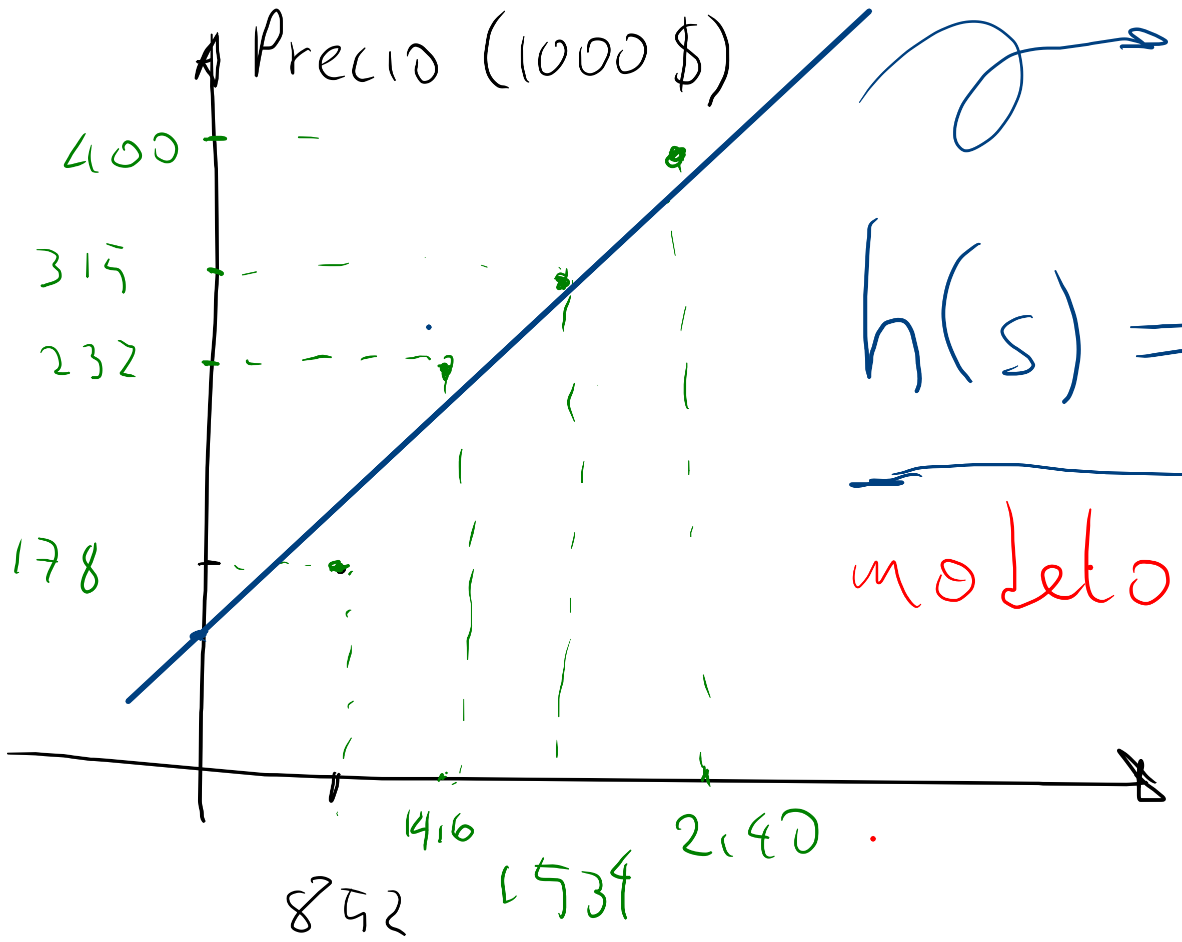
Labels

Aproximar la función $f(\ )$ usando un algo. de aprendizaje

Regresión Lineal

Sup(m²)	Precio (1K $)
2140	400
1416	232
1534	315
852	178
1200	???

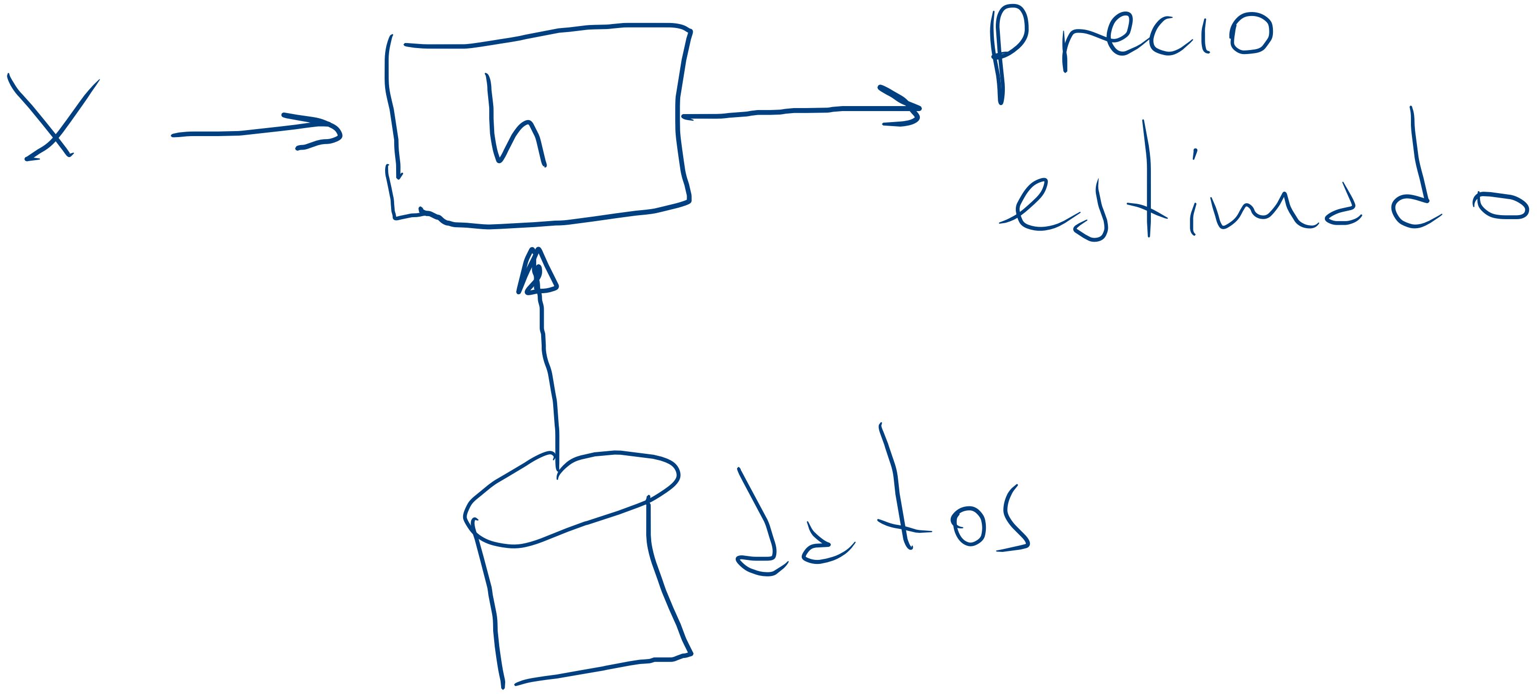
datos



h(sup) ax+b

$$h(s) = \theta_0 \cdot s + \theta_1$$

modelo: parametros ??



$$h(x) = \theta_0 + \theta_1 x$$

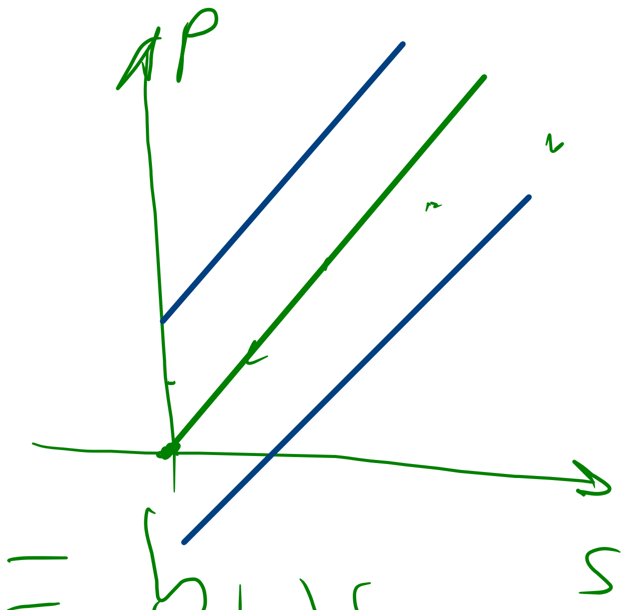
$$h(x) = \sum_{j=0}^{n} \theta_j \cdot x_j = \theta^T \cdot x$$

3 car. $\begin{bmatrix} x_0 = 1 \\ x_1 = x \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$

$$x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

1	2	3	
sup	baños	garaje	Precio
1400	3	NO	230

Características / features — target



$x \in \mathbb{R}^4$ $\theta_0 = b$ = bias → sesgo

$\theta \in \mathbb{R}^4$

$$x = \begin{pmatrix} x_0 = 1 \\ x_1 = \text{sup} \\ x_2 = \text{baños} \\ x_3 = \text{garaje} \end{pmatrix}$$

$$h(x) = \Theta^T \cdot x$$

$\Theta \in \mathbb{R}^n$

$n = \#$ parametros

$K =$ caracteristicas

$$\Theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} \longrightarrow \text{Parámetros} \longrightarrow \text{desconocido}$$

$n = K + 1$

$X =$ "Entradas" $\to$ características

$y =$ "salidas" $\to$ precio

conocidas

training set

$(X^{(i)}, y^{(i)}) =$ ejemplo, muestra

sample.

X	y

$m = \#$ ejemplos de entrenamiento

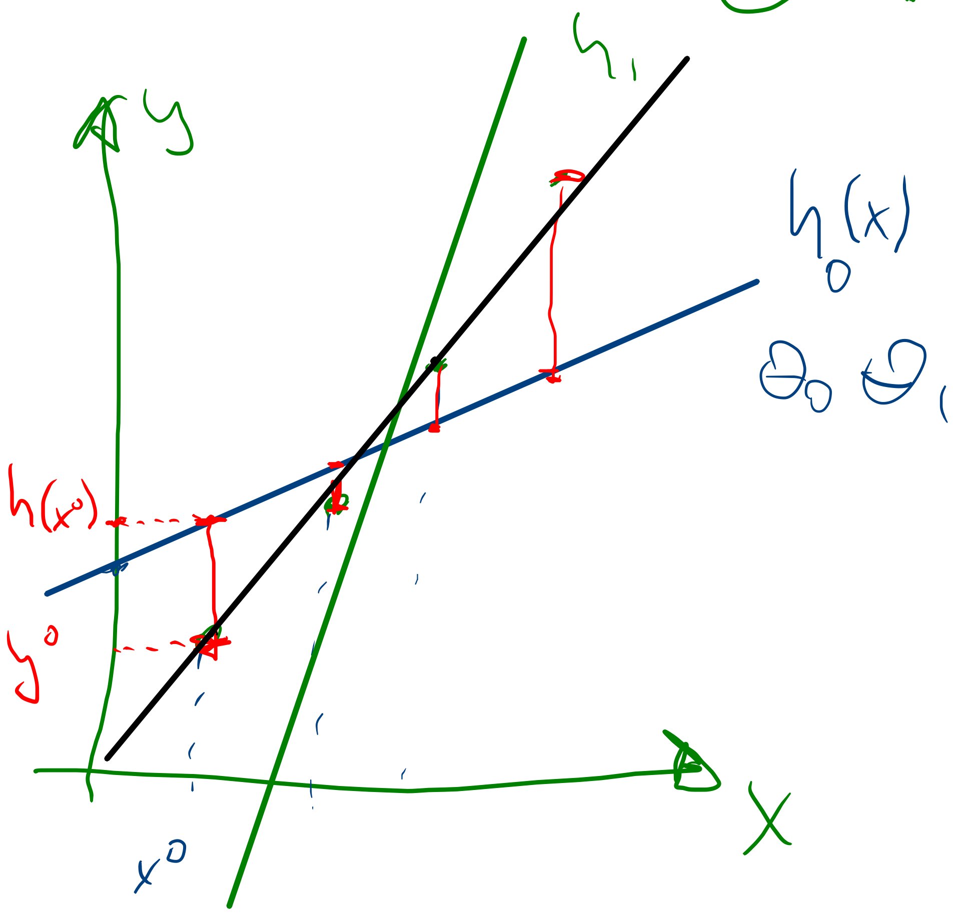
i

$$h(x) = \Theta^T \cdot X$$

$$h_\Theta(x)$$

encontrar Θ :

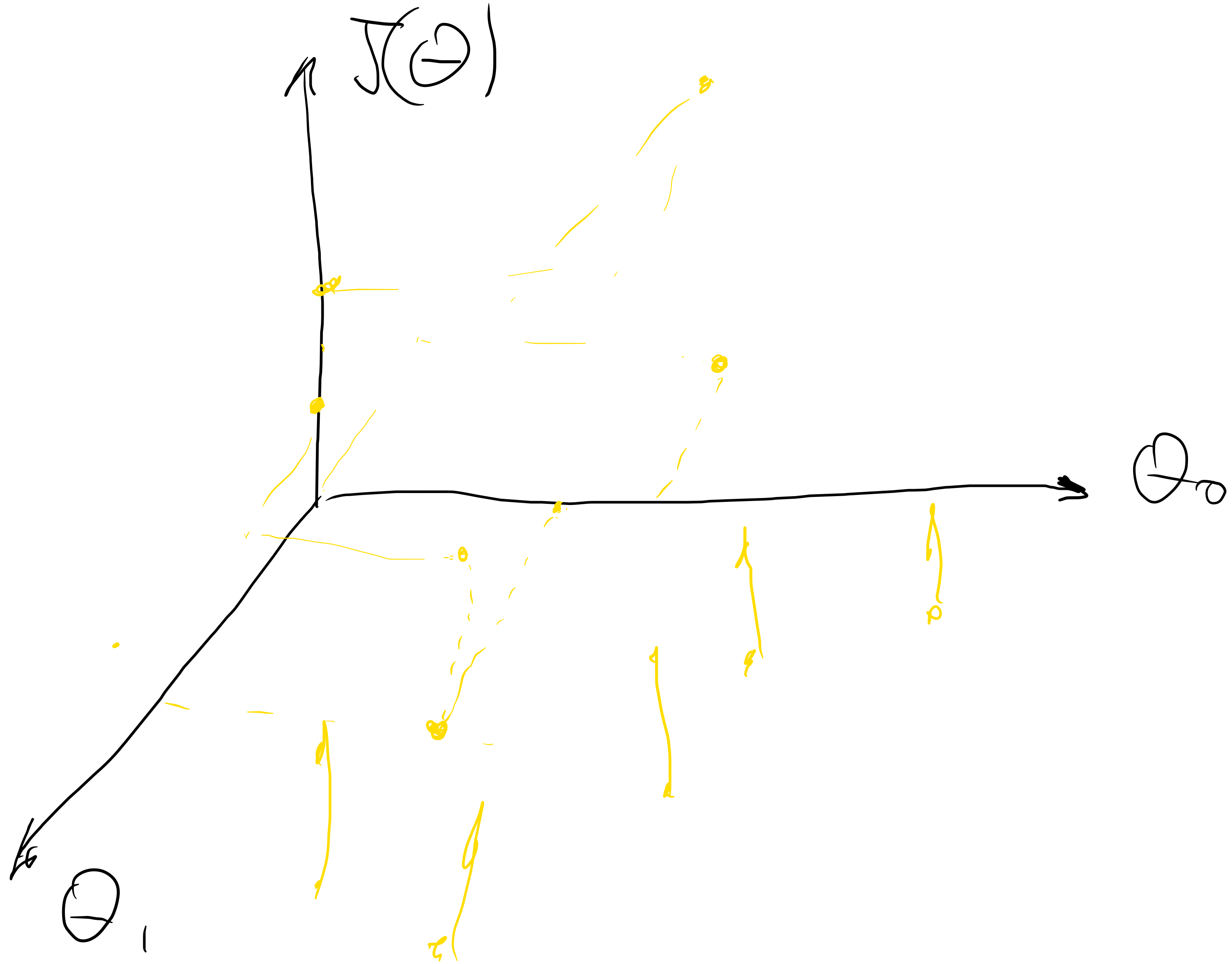
$h(x) \approx y$ para todos los ejemplos de entrenamiento



$$\sum_{i=1}^{m} \left(h_\Theta(x^{(i)}) - y^{(i)} \right)^2 = J(\Theta)$$

$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$J(\theta)$$



Descenso de Gradiente

minimizar: $J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \left(h(x^{(i)}) - y^{(i)} \right)^2$
θ

$$\nabla J(\theta) = \begin{bmatrix} \frac{\partial J}{\partial \theta_0} \\ \vdots \\ \frac{\partial J}{\partial \theta_n} \end{bmatrix}$$

Algoritmo:

dados X, y, h, J;

repetir hasta la convergencia:

$$\theta := \theta - \alpha \cdot \underline{\nabla J(\theta)};$$

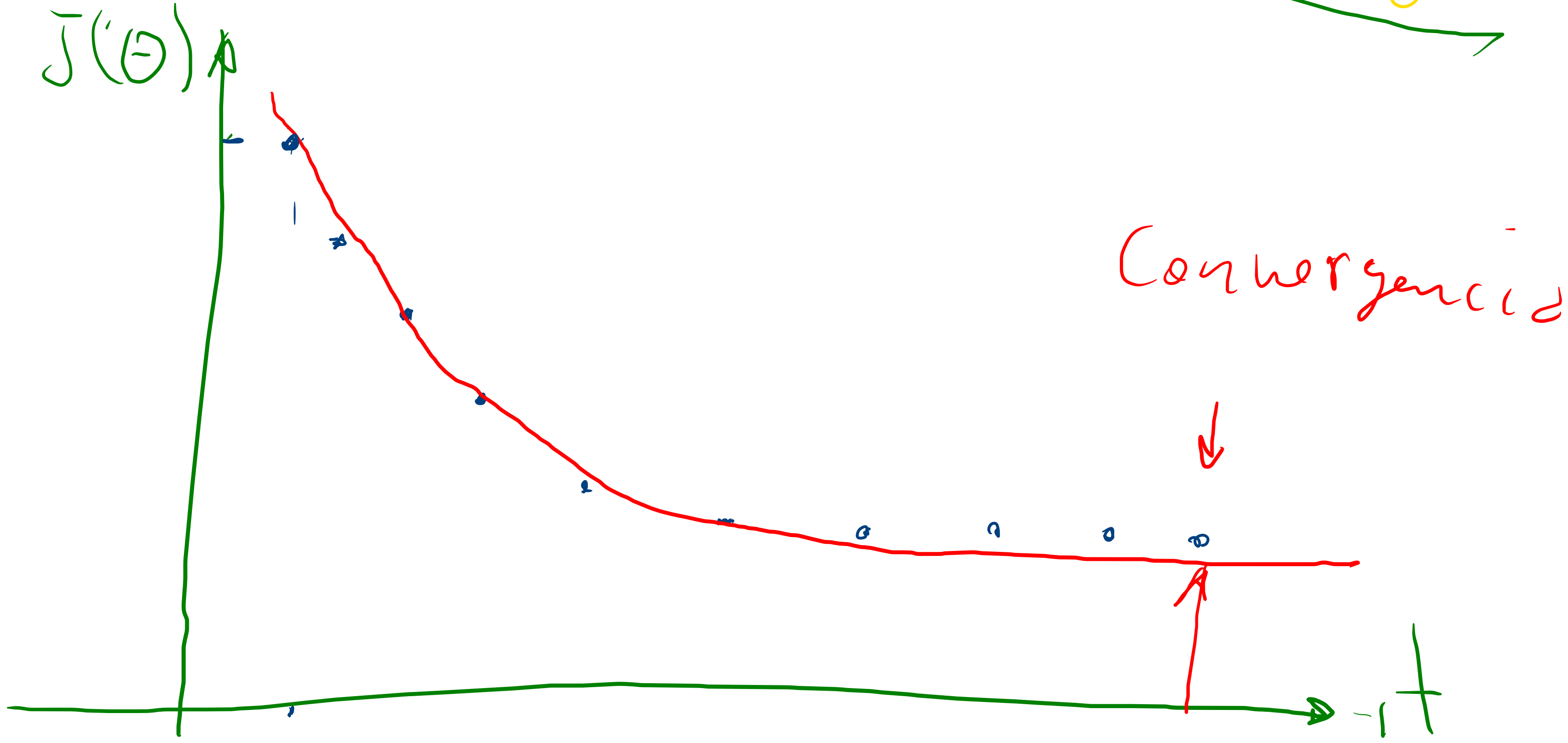
learning rate

Algoritmo:

dados X, y, h, J;

repetir hasta la convergencia,

$$\theta_j := \theta_j - \alpha \frac{\partial J(\theta)}{\partial \theta_j}$$



$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \qquad |\nabla J(\theta)| - \begin{bmatrix} g\theta_0 \\ g\theta_1 \\ \vdots \end{bmatrix}$$

t	$J(\theta)$
1	3.5
2	2.0
3	1.0
4	0,9
5	0.88
6	0.8799

1.5

1

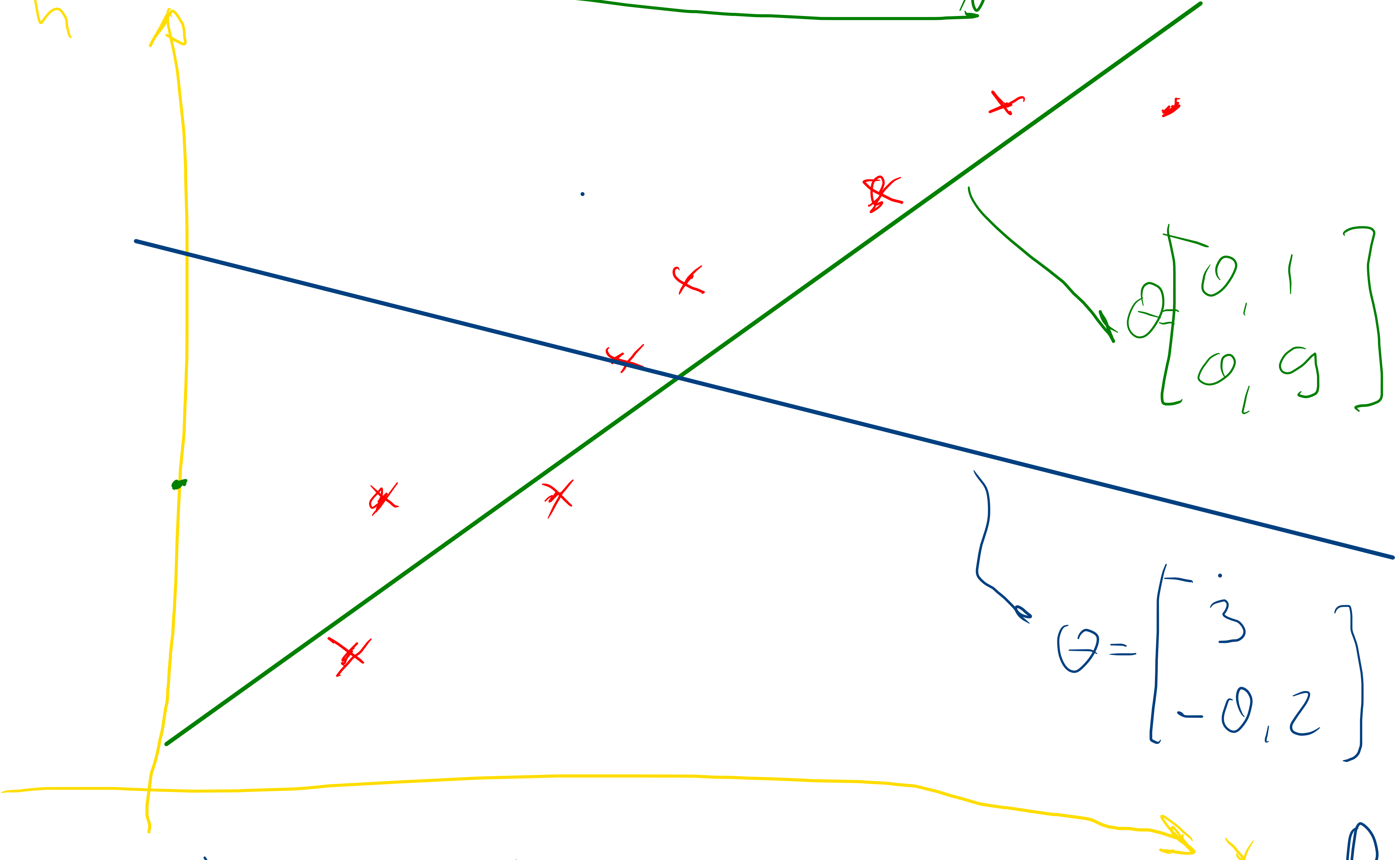
0,1

0,02

0,0

$$h_\theta(x) = \underbrace{\theta_0 + \theta_1 x_1}_{} + \theta_2 x_2 + \theta_3 \cdot x_1^3 + \theta_4 \cdot x_1 \cdot x_2$$

$$\Theta = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad y = 2 - x$$



$$\theta = \begin{bmatrix} 0,1 \\ 0,9 \end{bmatrix}$$

$$\theta = \begin{bmatrix} 3 \\ -0,2 \end{bmatrix}$$

$$J(\theta) < J(\Theta)$$

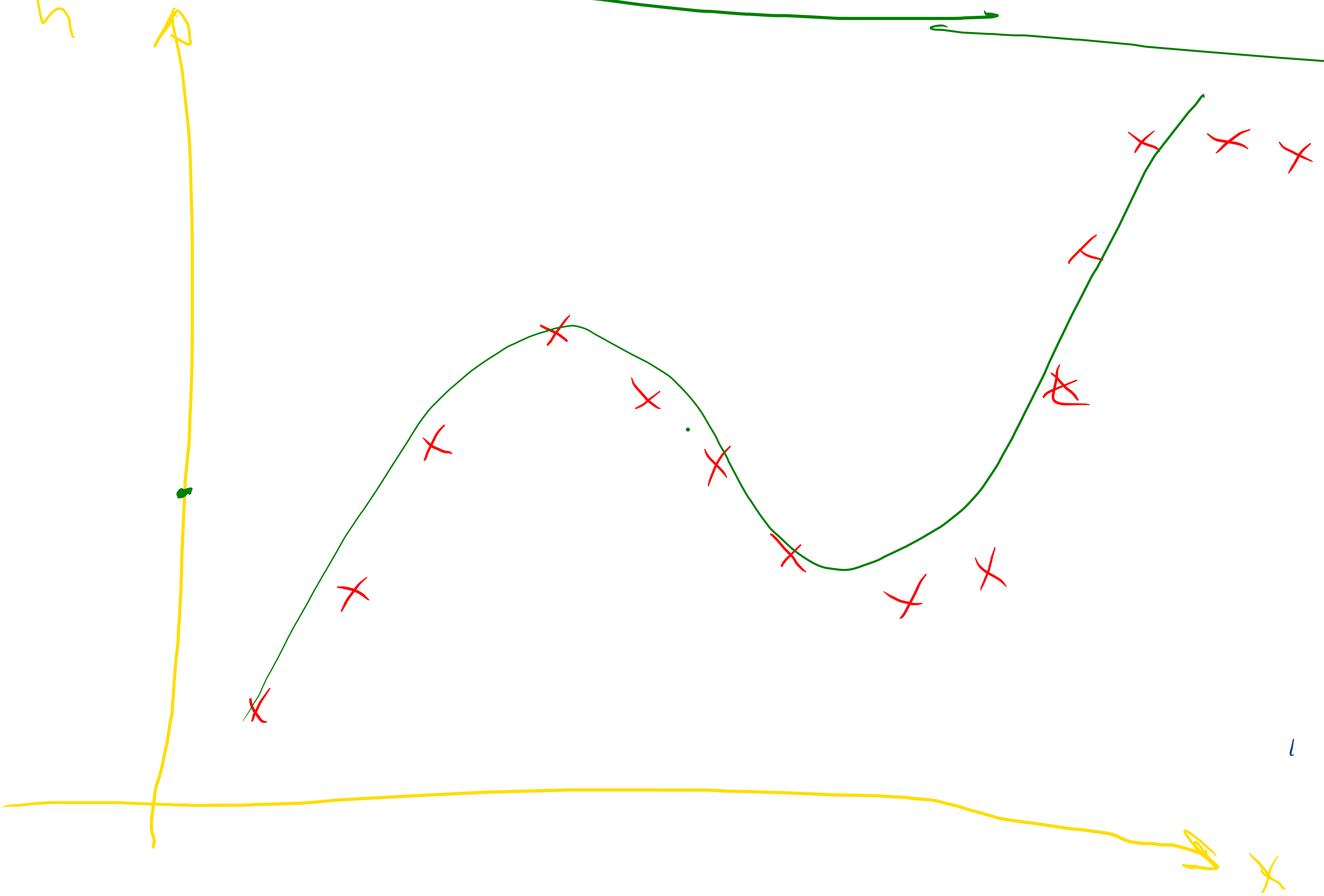
Regresion

$$h_\theta(x) : \Theta \in \mathbb{R}^4$$

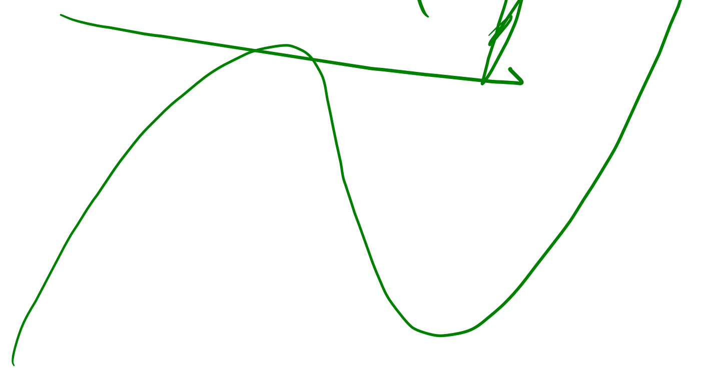
hiperplano

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_1^2 + \theta_3 x^3$$

h



SVM



$$\theta_j := \theta_j - \alpha \boxed{\frac{\partial J(\theta)}{\partial \theta_j}}$$

$$h_\theta(\theta) = \sum^{n} \theta_i \cdot x_i$$

$\sum \hat{=} \int$ = linealcs

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \left(h(x^{(i)}) - y^{(i)} \right)^2$$

$$= \frac{1}{m} \sum_{i=1}^{m} \left(\left(\sum_{j=0}^{n-1} \theta_j \cdot x_j^{(i)} \right) - y^{(i)} \right)^2$$

$$\frac{\partial J(\theta)}{\partial \theta_j} = \underbrace{\left(h(x^{(i)}) - y^{(i)} \right) \cdot x^{(i)}}$$

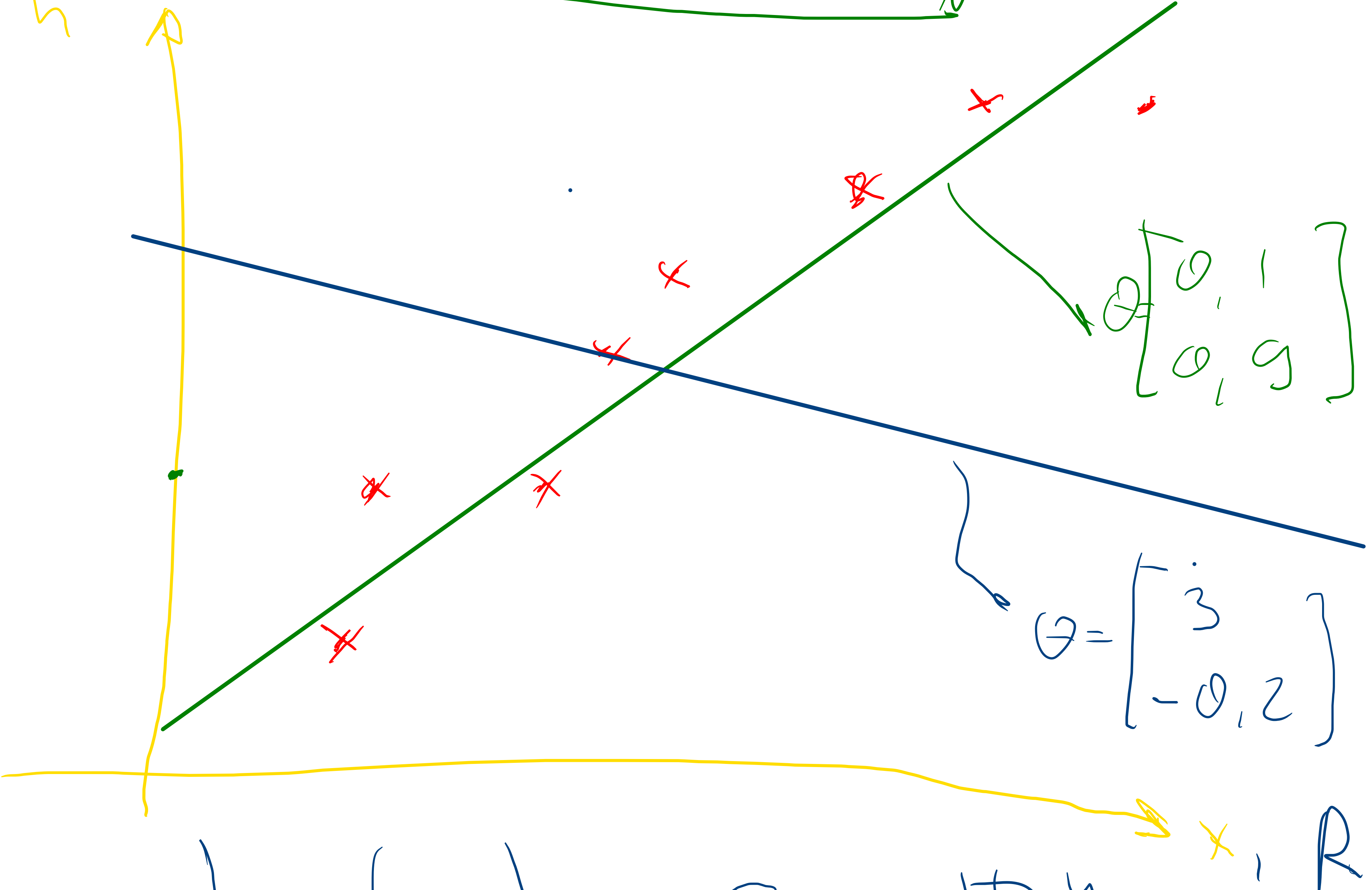
$$h_\theta(\theta) = \sum^{n} \theta_i \cdot x_i$$

Repeat:

$$\theta_j := \theta_j - \alpha \left(h(x^{(i)}) - y^{(i)}\right) \cdot x^{(i)}$$

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 \cdot x_1^3 + \theta_4 \cdot x_1 \cdot x_2$$

h



$$\Theta = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad h = 2 - x$$

$$\underline{J(\theta) < J(\Theta)}$$

Regresion

$$h_\theta(x) : \Theta \in \mathbb{R}^4$$

hiperplano