QT4

1. Conceptual
2. Ahead of time (learnability, ...)
3. Details / specific points
4. Practical aspects

Conceptual issues

(1) Perceptron, Linear regression, Logistic regression

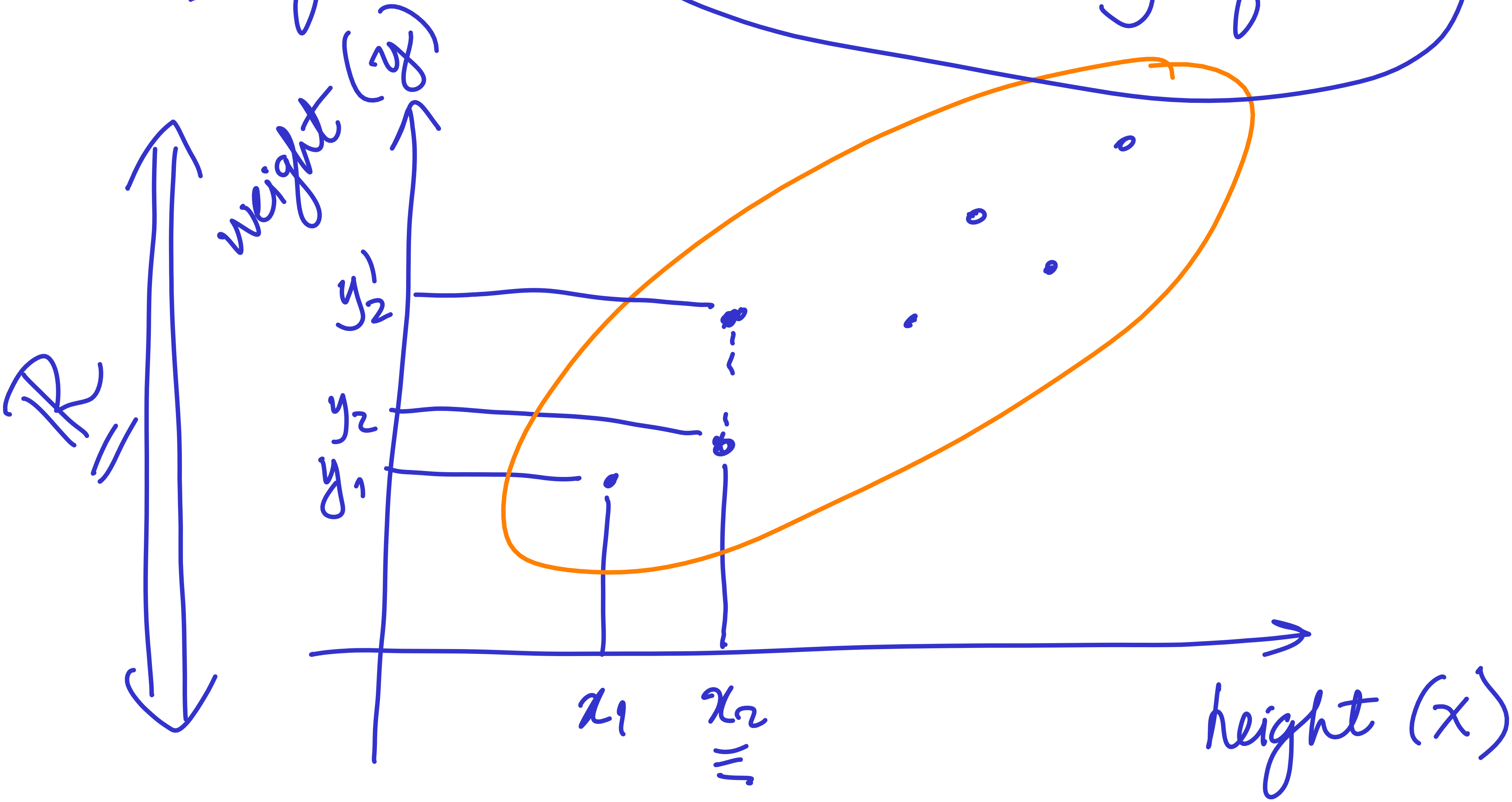
Regression Problem

×

Classification Problem.

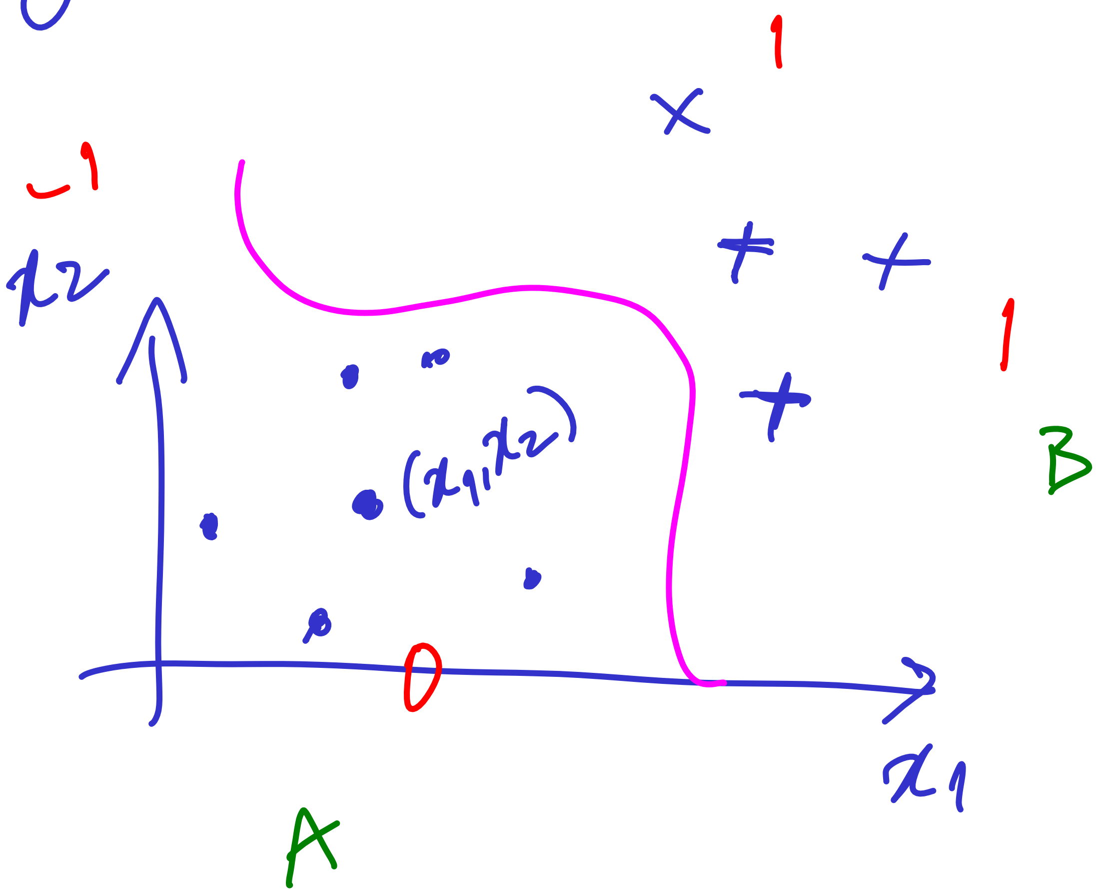
Regression

Find f s.t. $y \approx f(x)$



Classification.

y is class label.



Cost Function

$$\boxed{J(w) = \frac{1}{N}\sum_{n=1}^{N}\big(\underbrace{y^{(n)} - \hat{y}^{(n)}}\big)^2} \quad (\text{Reg})$$

binary classif.

$$\boxed{J(w) = \frac{1}{N}\sum_{n=1}^{N} y^{(n)} \ln\left(\hat{y}^{(n)}\right) + \left(1 - y^{(n)}\right)\ln\left(1 - \hat{y}^{(n)}\right)}$$

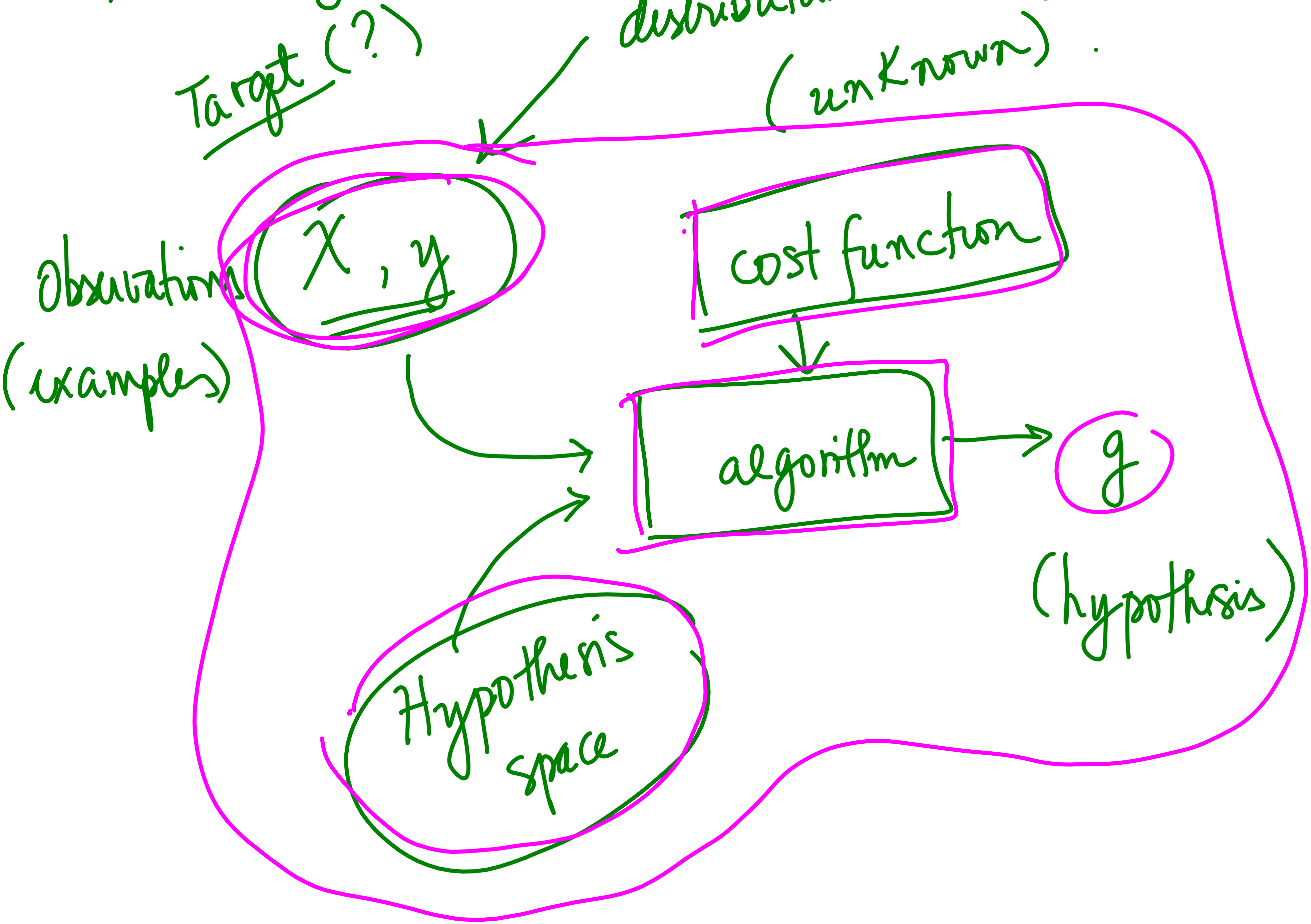
2 Function optimization

w^* ← optimal weight.

Gradient descent → a tool

* for optimizing functions.

Learning Model



1. Perceptron

2. Linear regression

3. Logistic regression

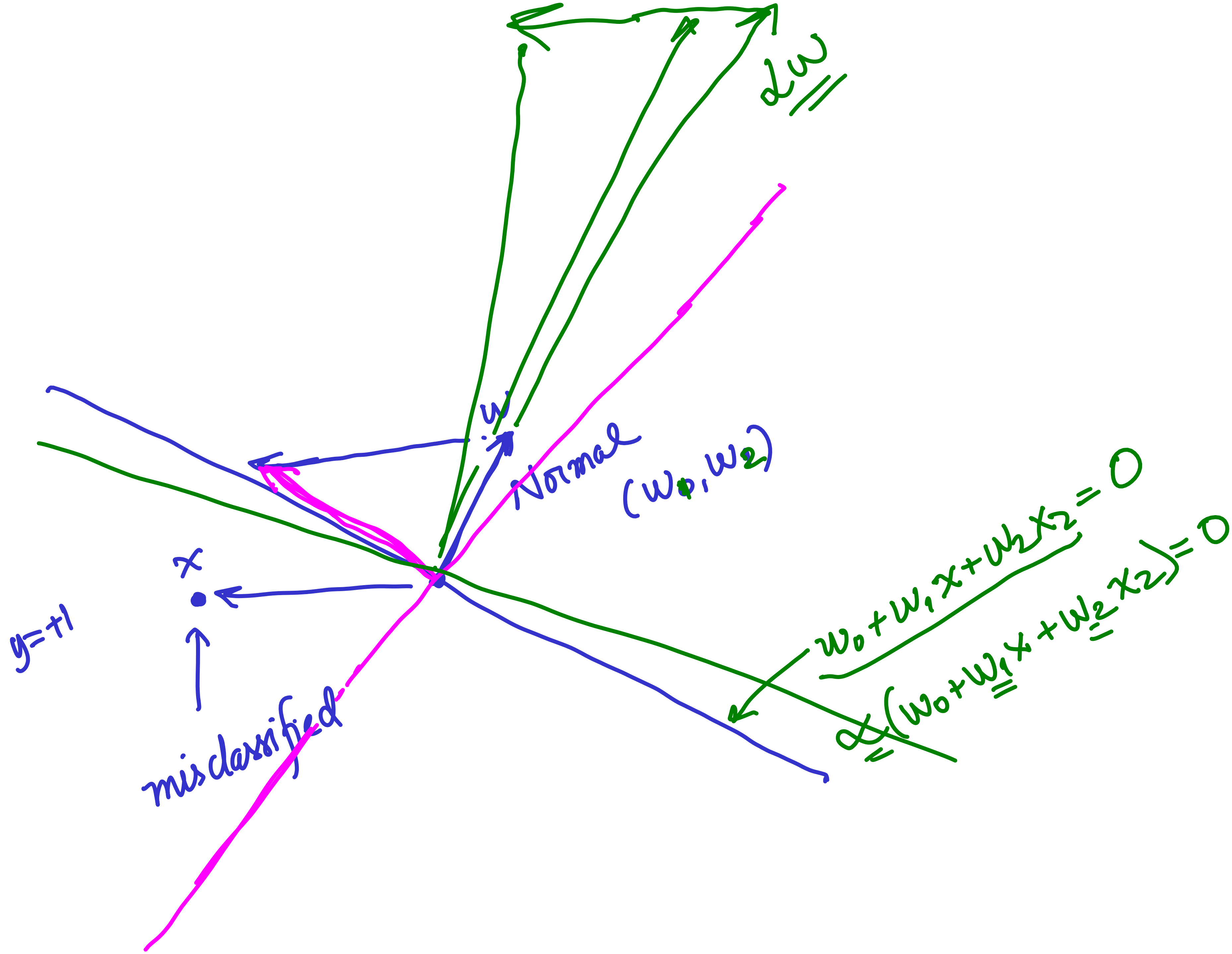
regression

classifica-
tion
(binary)

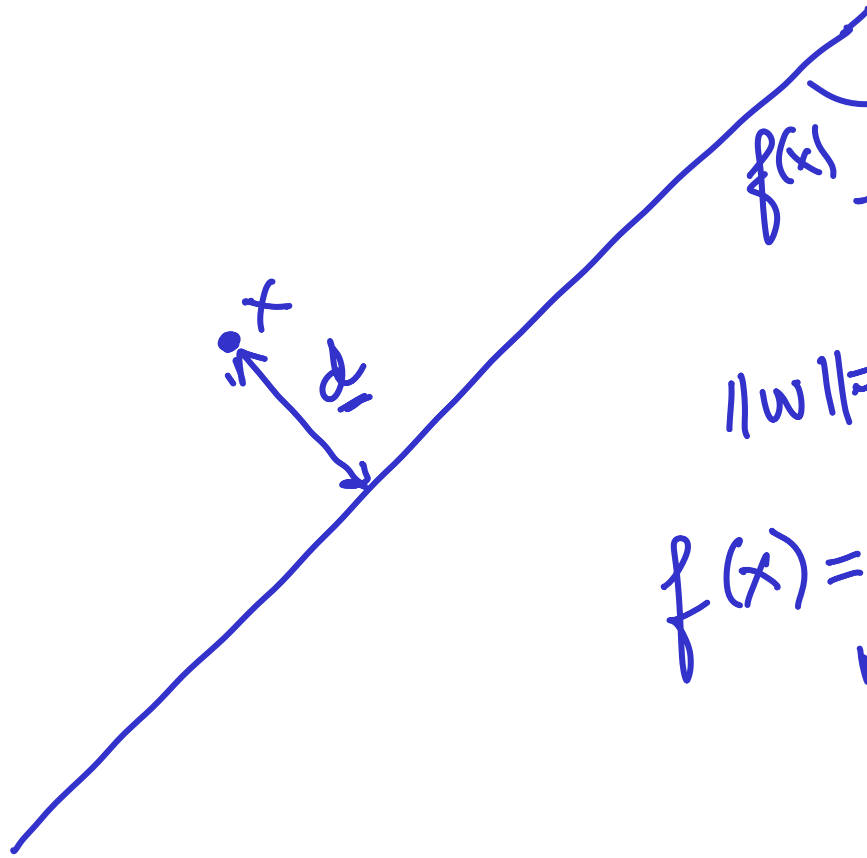
Perceptron

$$w \leftarrow w + y\,x$$

misclassified



αw
w
Normal
(w1, w2)
w0 + w1x + w2x2 = 0
α(w0 + w1x + w2x2) = 0
x
y=+1
misclassified



$f(x) \to w_0 + w_1 x_1 + w_2 x_2 = 0$

$\|w\| = 1 \to$

$f(x) =$ distância da reta a x

$$w \leftarrow w + \eta y x$$

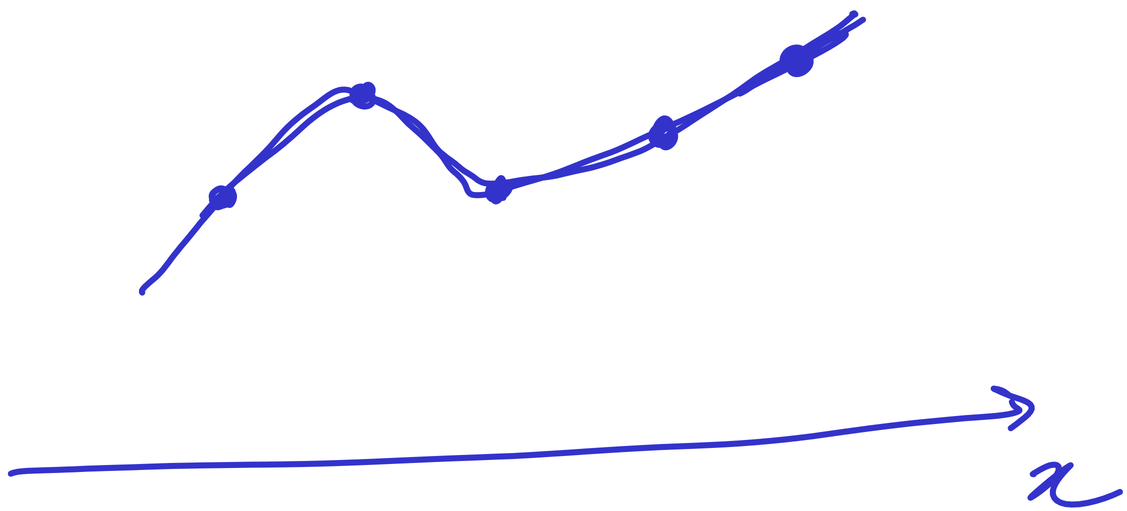
learning rate.

accelerate convergence?

η

Linear regression.

interpolation x regression



$$h(x) = w^T x$$

$$= \underline{w_0 + w_1 x}$$

$$\cancel{f}\ h(x) = \underline{w_0 + w_1 x + w_2 x^2}$$

$$\Rightarrow \boxed{J(w) = \frac{1}{N} \sum \Big(y^{(n)} - \underbrace{h(x^{(n)})}_{\hat{y}^{(n)}} \Big)^2}$$

Logistic Regression

(x, y)

$\{0, +1\}$

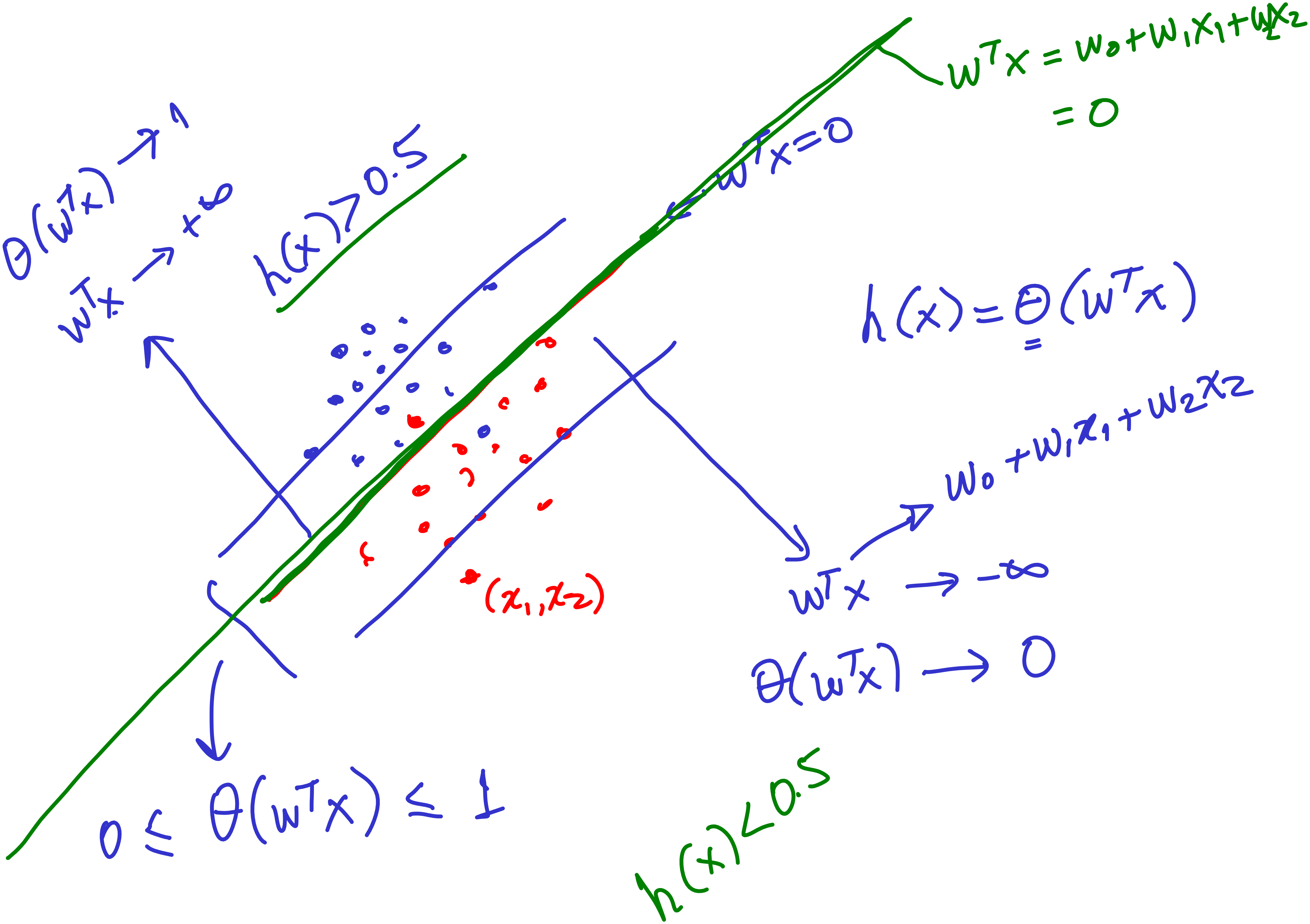
Target

$f(x) = \boxed{P(y=+1 / x)}$

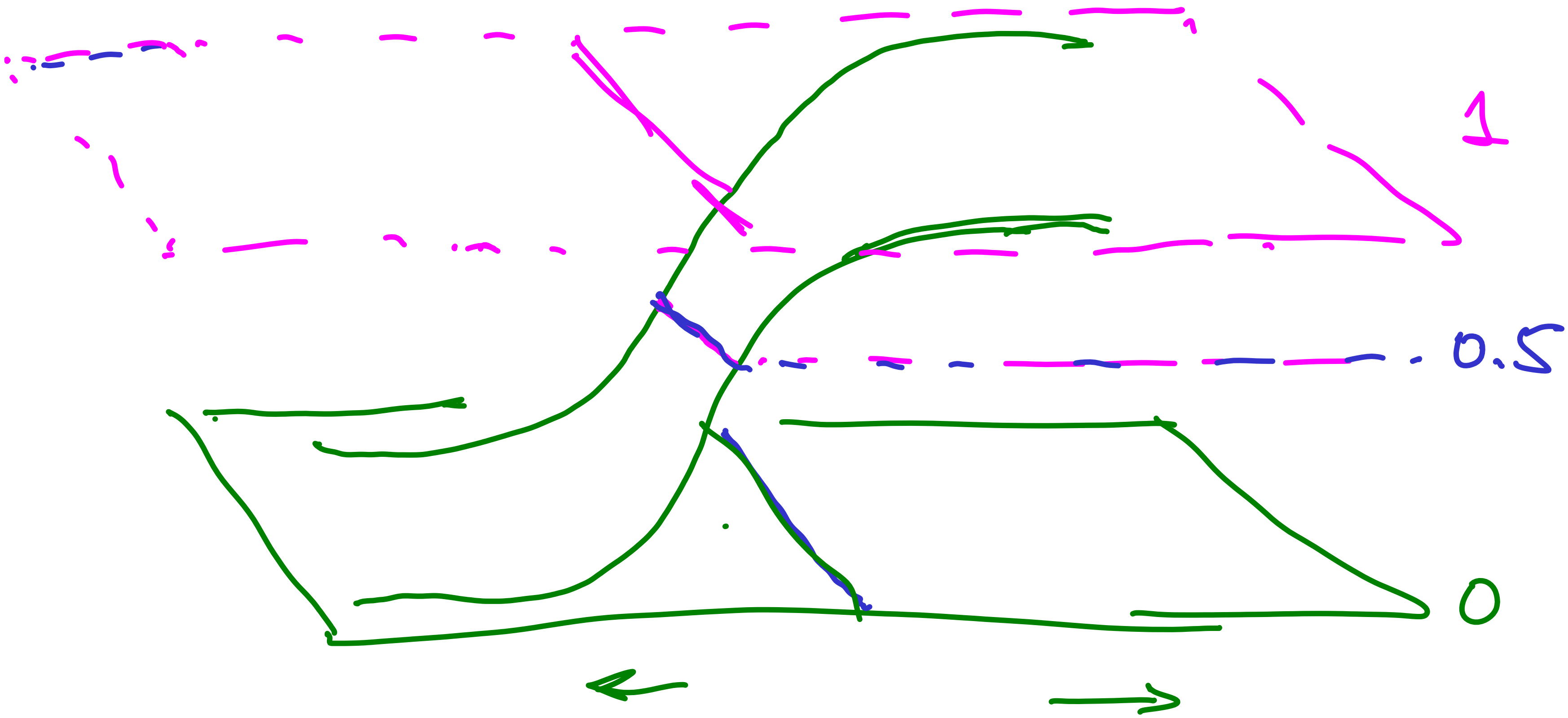
Cross-entropy loss

$$J(w) = \frac{1}{N}\sum y \ln(\hat{y}) + (1-y)\ln(1-\hat{y})$$

$$h(x) = \theta(w^T x)$$

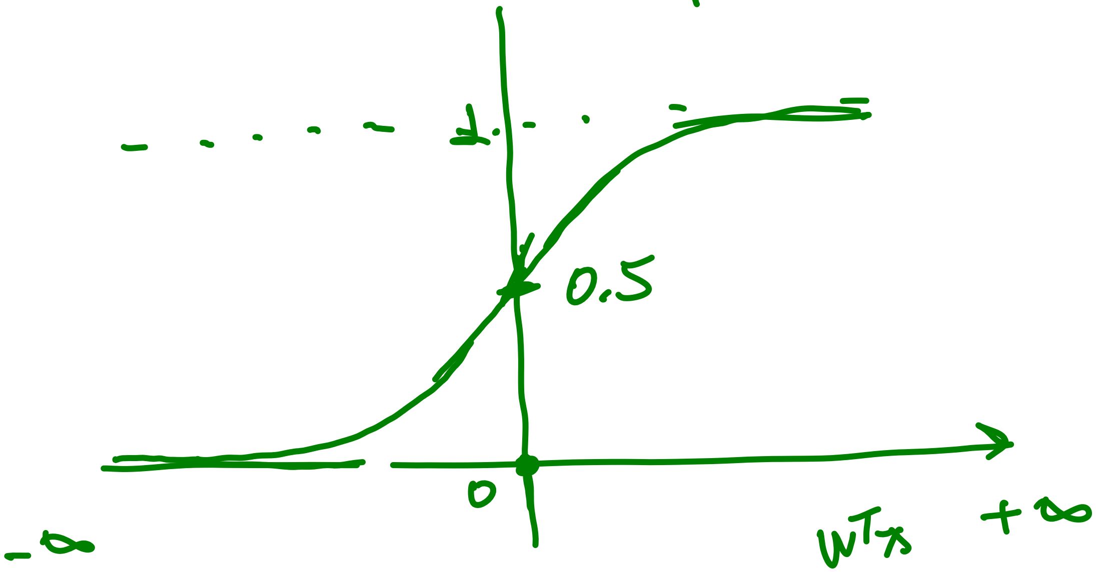


$w^T x = w_0 + w_1 x_1 + w_2 x_2 = 0$
$w^T x = 0$
$\theta(w^T x) \to 1$
$w^T x \to +\infty$
$h(x) > 0.5$
$h(x) = \theta(w^T x)$
(x_1, x_2)
$w_0 + w_1 x_1 + w_2 x_2$
$w^T x \to -\infty$
$\theta(w^T x) \to 0$
$0 \leq \theta(w^T x) \leq 1$
$h(x) < 0.5$



negativos → y=0 positivos → y=1

$$\theta(z) = \frac{1}{1+e^{-z}} = 0.5 \qquad 0 \leq \theta(z) \leq 1$$



$$\boxed{\hat{y} = \begin{cases} 1, & \text{se } \underline{\theta(w^T x) \geq 0.5} \\ 0, & \text{c.c} \end{cases}}$$

$$\underline{\underline{\theta(\overline{w^T x})}}$$

%

$$\underline{\underline{h(x)}} = \theta(w^T x) \simeq P(y=+1 \mid x)$$