MATH 462
Lecture 1

HIGH LEVEL INTRO:

supervised ML

~ regression

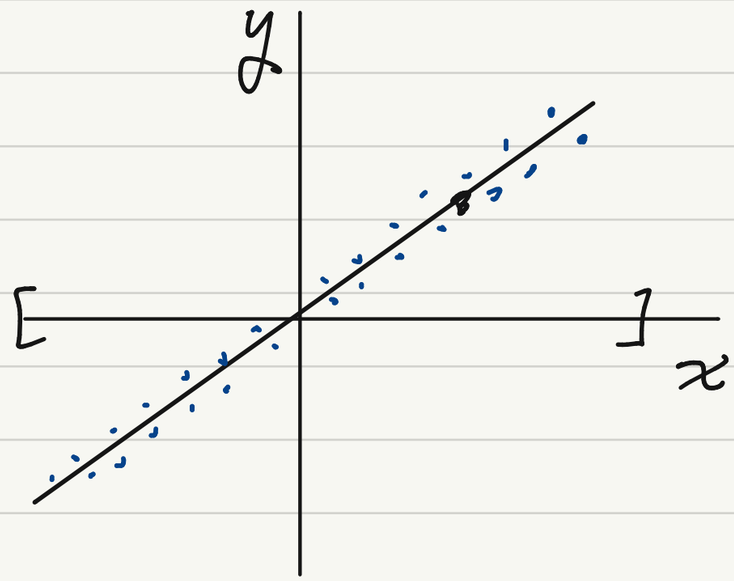
~ class ~ multiclass

~ Futer: more regression

binary class. imbalanced

ex. COVID test FP/FN

Regression



$x \in \mathbb{R}^1$, $y \in \mathbb{R}^1$

Inputs

$S_m \{(x_1, y_1), \ldots, (x_m, y_m)\}$

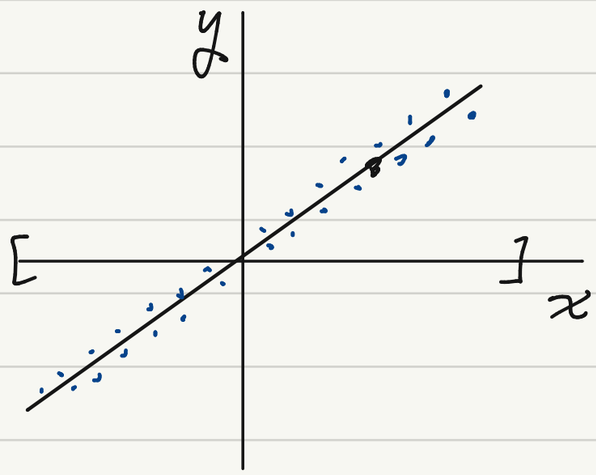
Model parameter w

$h(x; w)$ $\quad h(\cdot; w): x \to y$

$\in \mathbb{R}^1 \quad \in \mathbb{R}^1$

Goal: use S_m to find w^* so that:

Regression



$x \in \mathbb{R}^1$, $y \in \mathbb{R}^1$ Inputs

$S_m \{(x_1, y_1), \ldots, (x_m, y_m)\}$

Model parameter w

$h(x; w)$ $\quad h(\cdot; w): x \to y$, $x \in \mathbb{R}^1$, $y \in \mathbb{R}^1$

Goal: use S_m to find w^* so that:

"first time"

loss $\ell(y_1, y_2)$ measures the misfit between $y_1, y_2 \in \mathbb{R}$

defined $\ell(y_1, y_2) = (y_1 - y_2)^2$

"ERM"

$$w = \arg\min_w \hat{L}(w) \equiv \frac{1}{m} \sum_{i=1}^{m} \ell(h(x_i; w), y_i)$$

Regression:

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

$$\min_w \frac{1}{m} \sum_{i=1}^{m} (wx_i - y_i)^2$$

- SOLVE: analytically explicit
- Future ML: prove soln w^* using convex analysis.
  & give algorithm to find it.

- Also: geometric soln

Calculus: $\hat{L}(w) = \frac{1}{m} \sum_{i=1}^{m} (wx_i - y_i)^2$

$$0 = \frac{\partial}{\partial w} \hat{L}(w) = \frac{2}{m} \sum_{i=1}^{m} (wx_i - y_i) x_i = 0$$

$$w \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} w x_i^2 = \sum_{i=1}^{m} y_i x_i \quad \Rightarrow \quad w = \frac{\sum_{i=1}^{m} y_i x_i}{\sum_{i=1}^{m} x_i^2}$$

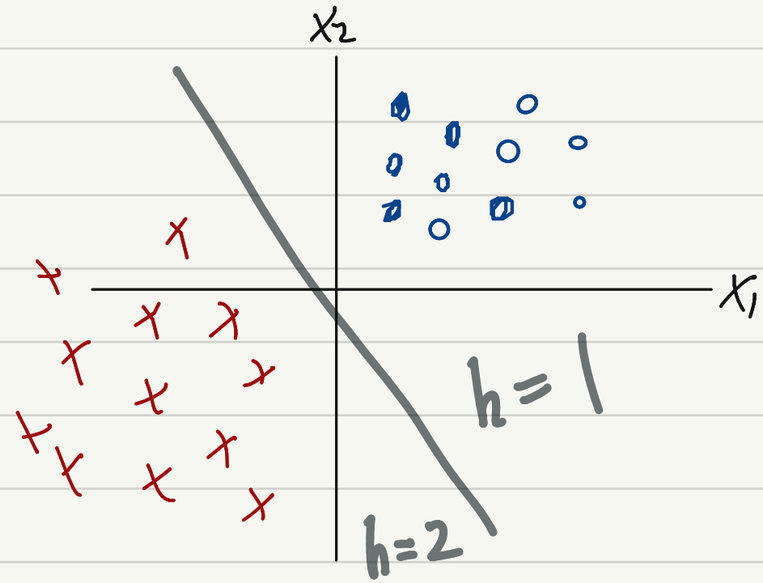
Classification

inputs $x \in \mathbb{R}^2 = \mathcal{X}$

$\Rightarrow (x_1, x_2)$

outputs $y \in \{o, x\}$

$\in \{1, 2\} = \mathcal{Y}$



model: $h : \mathcal{X} \to \mathcal{Y}$ discontinuous function No!

Loss 0-1.

1st Approach

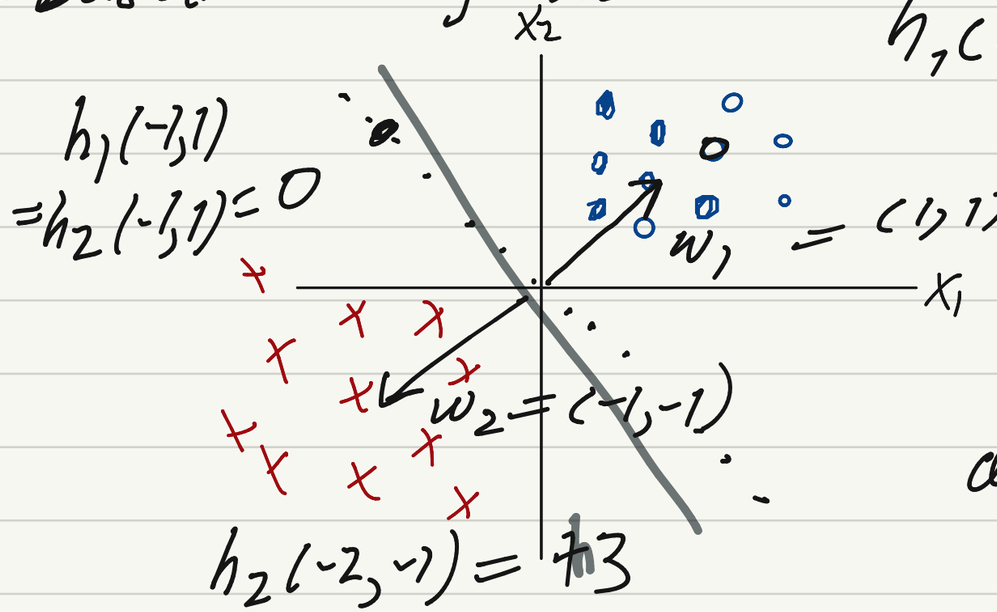
score-based classification

$h_1(x) = w_1 \cdot x$

$= x_1 + x_2$

$h_2(x) = w_2 \cdot x$

$= -x_1 - x_2$



$h_1(1.5, 1.5) = 3$

$c(x) =$

$\text{argmax}(h_1(x), h_2(x))$

C+ Approach

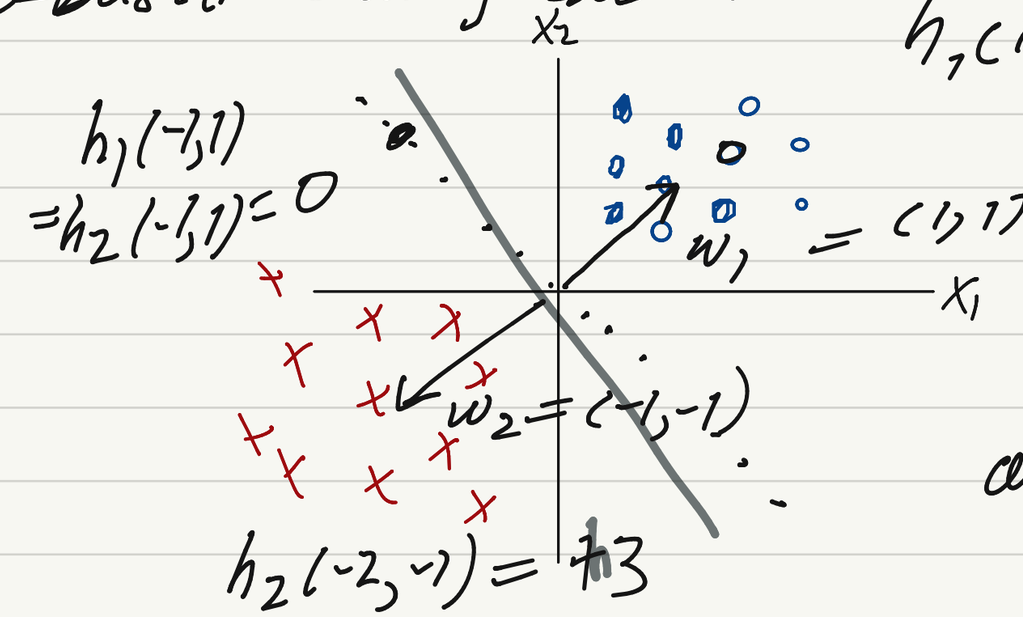
score-based classification

$h_1(x) = w_1 \cdot x$

$= x_1 + x_2$

$h_2(x) = w_2 \cdot x$

$= -x_1 - x_2$



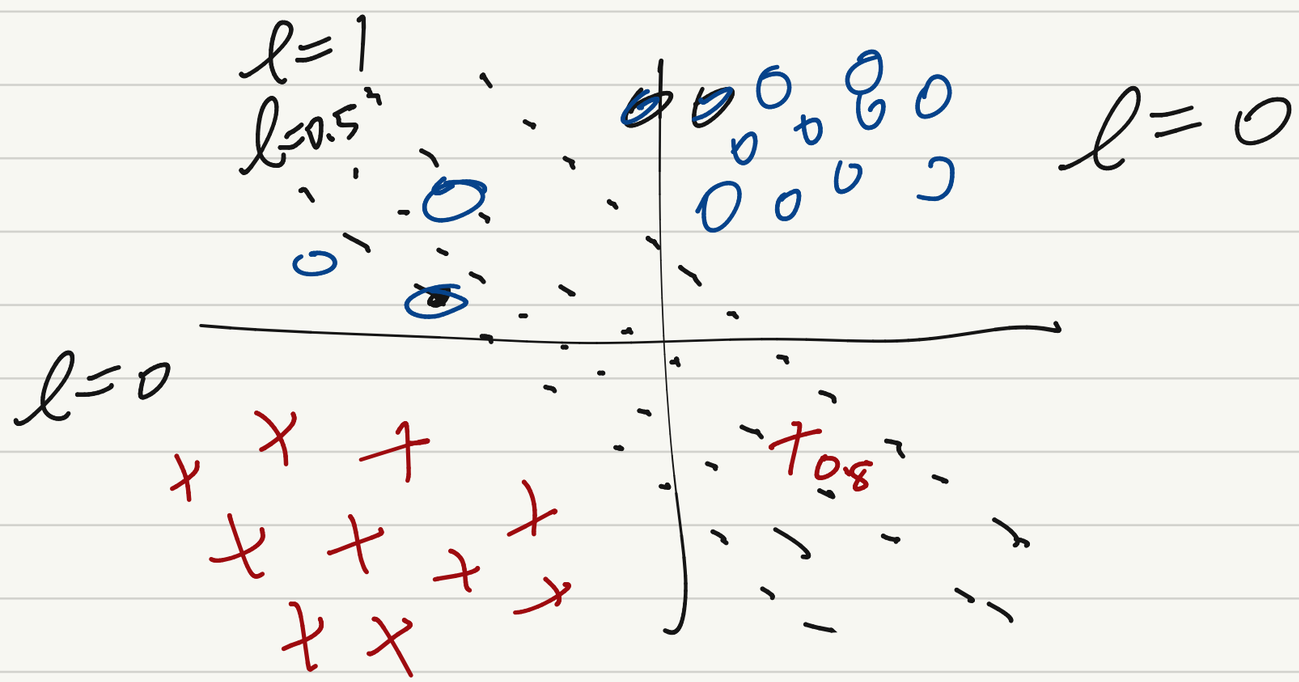
$h_1(1.5, 1.5) = 3$

$C(x) =$

$\text{argmax}(h_1(x), h_2(x))$

simplify

$h_2(x) = -h_1(x)$



Best Classifier.
will have
wide margin.