- Practice "R" (Practice lecture: Explain Codes using Projector.)

- Training Data

$$\mathcal{D} = \{(\vec{x}_1, y_1), (\vec{x}_2, y_2), \ldots, (\vec{x}_n, y_n)\}$$

$\vec{x}_1$ is Bill's characteristics, $y_1 = 1$ means he paid his loan

$\vec{x}_2$ is Jill's " " $y_2 = 1$ " she "

$\vec{x}_3$ is Tony's " " $y_3 = 0$ " he did not pay his loans.

$$\mathcal{D} = (\mathbf{X}, \vec{y}), \quad \mathbf{X} \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_p$$

$$\vec{y} \in \mathcal{Y}^n$$

y was either 0, 1, 2, 3.

$$\begin{bmatrix} \cdots \vec{x}_1 \cdots \\ \cdots \vec{x}_2 \cdots \\ \vdots \\ \cdots \vec{x}_n \cdots \end{bmatrix} \quad \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

n by p

OFFLINE

- We will focus on SUPERVISED LEARNING

You need 3 ingredients.

1. $\mathcal{D}$, the training data
2. $\mathcal{H} = \{$all candidate functions for $f\}$
3. $\mathcal{A}$, the algorithm which produces $g = A(\mathcal{D}, \mathcal{H})$

$Y = t(Z)$

↑ true (continuous).

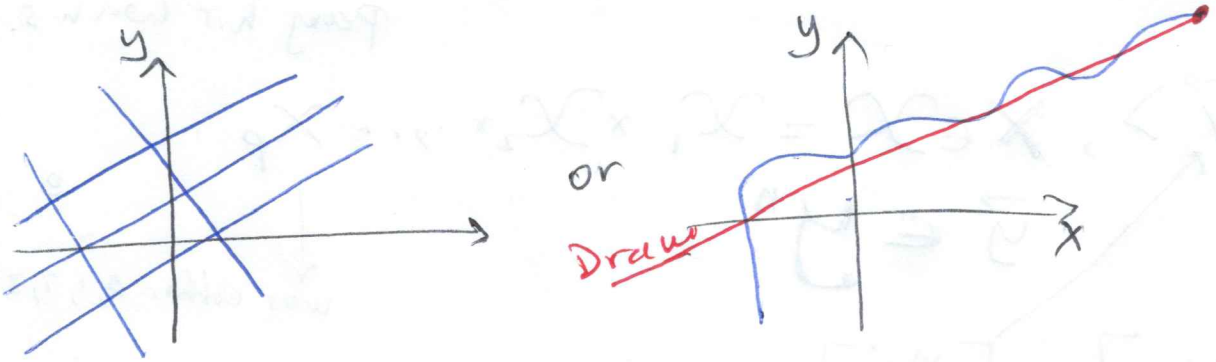
phenomenon.
Best $\vec{x}$ is observable

$Y = f(\vec{x}) + \int$, where $\int = t(z) - f(\vec{x})$

↑ Approx to g.

Goal: estimate f

g is best approximate to f in H

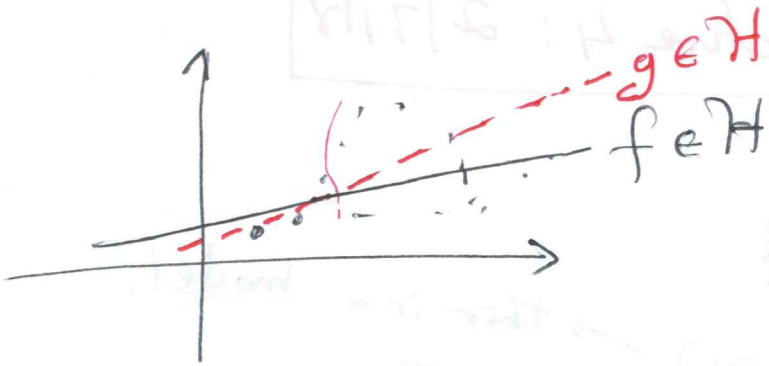


H is all functions from $X \longrightarrow Y$

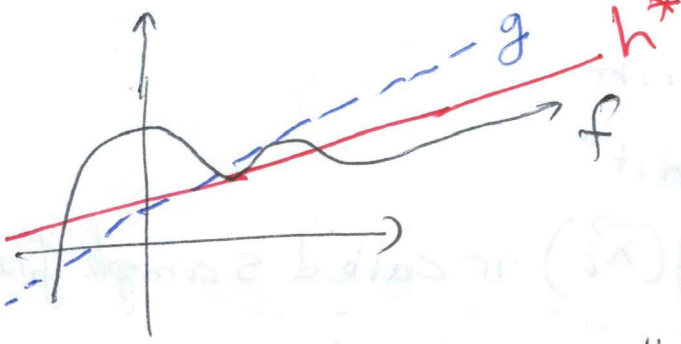
Let say f is in H.

So If f ~~$\in$~~ H

$$Y = g(\vec{x}) + \underbrace{\left(f(x) - g(\vec{x})\right)}_{\text{estimation error or primary estimate error.}} + \underbrace{\left(t(\vec{z}) - f(\vec{x})\right)}_{\int: \text{ error due to ignorance.}}$$



If $f \notin \mathcal{H}$



h^* is the best approx of $f \in$

Your algorithm will be:

$$y = g(\vec{x}) - \underbrace{(h(\vec{x}) - g(\vec{x}))}_{\text{estimation error}} + \underbrace{(f(\vec{x}) - h^*(\vec{x}))}_{\text{mis-specification error}} + \underbrace{(t(\vec{z}) - f(\vec{x}))}_{\text{error due to ignorance}}$$

↖ Three sources of error. To minimize, increase $p \nearrow$, more useful

How to minimize $(f(\vec{x}) - h^*(\vec{x}))$

Ans: need better algorithm.