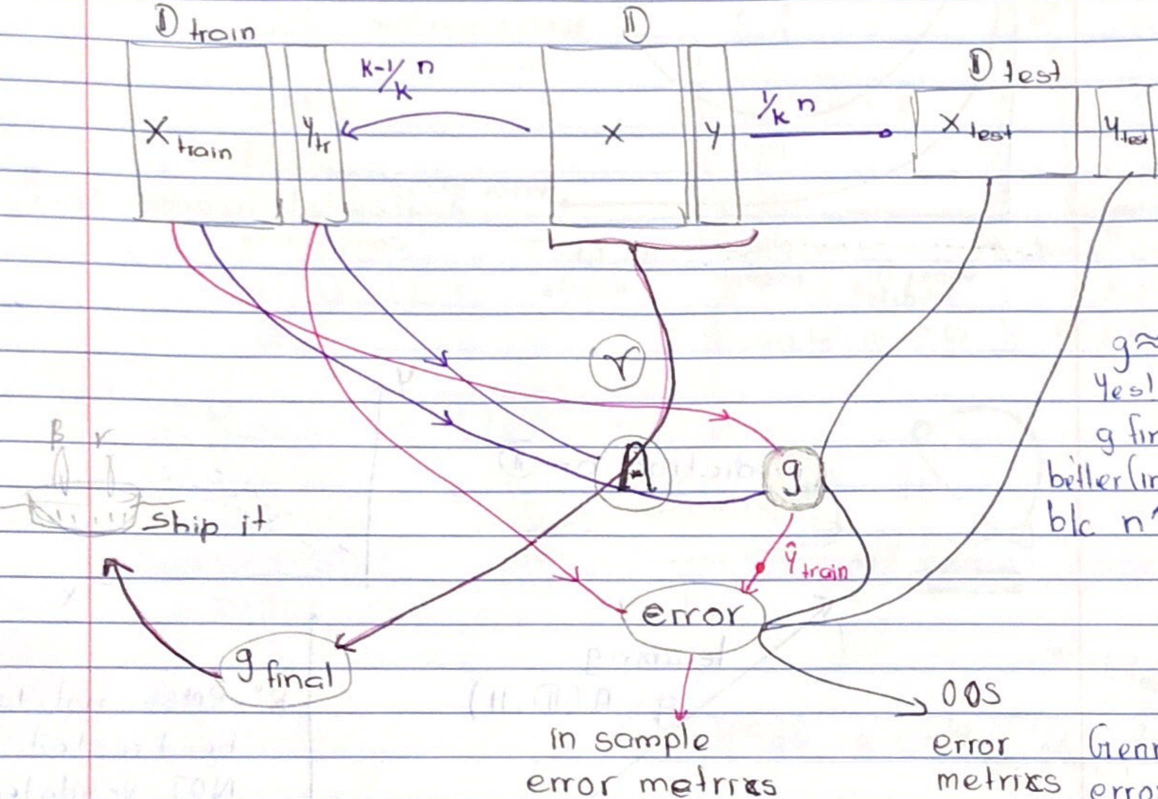
Lecture - 13 03/19/2020



$g \approx g_{final}$? Yes! But g_{final} is better (in expectation) b/c $n \uparrow$

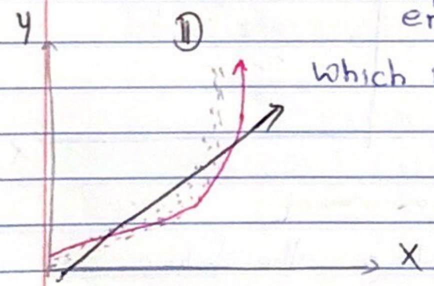
Genralization error is smaller than g's

(minimize making h^* more complexity)

$$Y = g + (h^* - g) + \underbrace{(f - h^*)}_{\text{misspecification error}} + (t - f)$$

misspecification error — which is limited by $\mathcal{H}$

nothing we can do



$H_1 = \{ w_0 + w_1 x : \vec{w} \in \mathbb{R}^2 \}$ linear models

$H_2 = \{ w_0 + w_1 x + w_2 x^2 : \vec{w} \in \mathbb{R}^3 \}$ linear non-linear model

A = OLS polynomial regression

Why are polynomials a good idea to use as bases for arbitrary function

Weierstrass approximation theorem: For any cont. function f, $\exists$ polynomial function p s.t $\forall_\varepsilon > 0$, $\forall \vec{x} \in \mathcal{X}$ where $|f(\vec{x}) - p(\vec{x})| < \varepsilon$

(Probably a bad idea in the real world - Polynomials of high degree are very bad, mostly past squares.)

H_2

A

$$X_{\text{original}} = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ \vdots & \vdots \\ 1 & x_{1n} \end{bmatrix} \longrightarrow X = \begin{bmatrix} 1 & x_{11} & x_{11}^2 \\ 1 & x_{12} & x_{12}^2 \\ 1 & x_{13} & x_{13}^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{1n}^2 \end{bmatrix}$$

Is this full rank?

$A: (X^T X)^{-1} X^T \vec{y} = \vec{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$

$p_{raw} + 1 = 2$

$p_{raw} = 1$

Original measurements

$p + 1 = 3$

$p_n = 2$

raw + derived

$H_3 = \{ w_0 + w_1 x + w_2 x^2 + w_3 x^3 : \vec{w} \in \mathbb{R}^4 \}$ (highest is $n-1$)

(Go up to H_{n-1}, but we need to be careful of overfitting bad points)

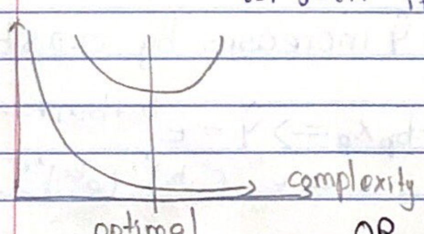
$p + 1 = n \implies R^2 = 100\%$; $S_e = 0$ Perfect fit and Over fit

$X = [\vec{1} \mid \vec{x} \mid \vec{x}^2 \mid \ldots \mid \vec{x}^{n-1}]$ full rank?

$n \times n$ matrix

Thm. Vandermonde matrix

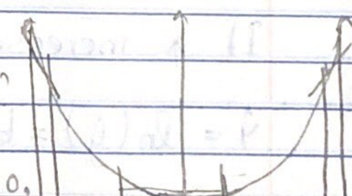
$\det[X] = \prod_{i=1}^{n} \prod_{j=i+1}^{n} (x_j - x_i) \neq 0$ if all x's are unique



Correlation$[X, X^2]$?

It depends on your domain

e.g. close to domain 0, corr ≈ 0



$y = b_0 + \underline{b_1} X + \underline{b_2} X^2 = b_0 + (b_1 + b_2 x) X$

$$\begin{array}{ccc} X & \xrightarrow{QR} & R \\ \downarrow & & \downarrow \text{ortho. proj} \\ H & = & H \end{array}$$

$\hat{y} = \hat{y} \implies R^2 = R^2$