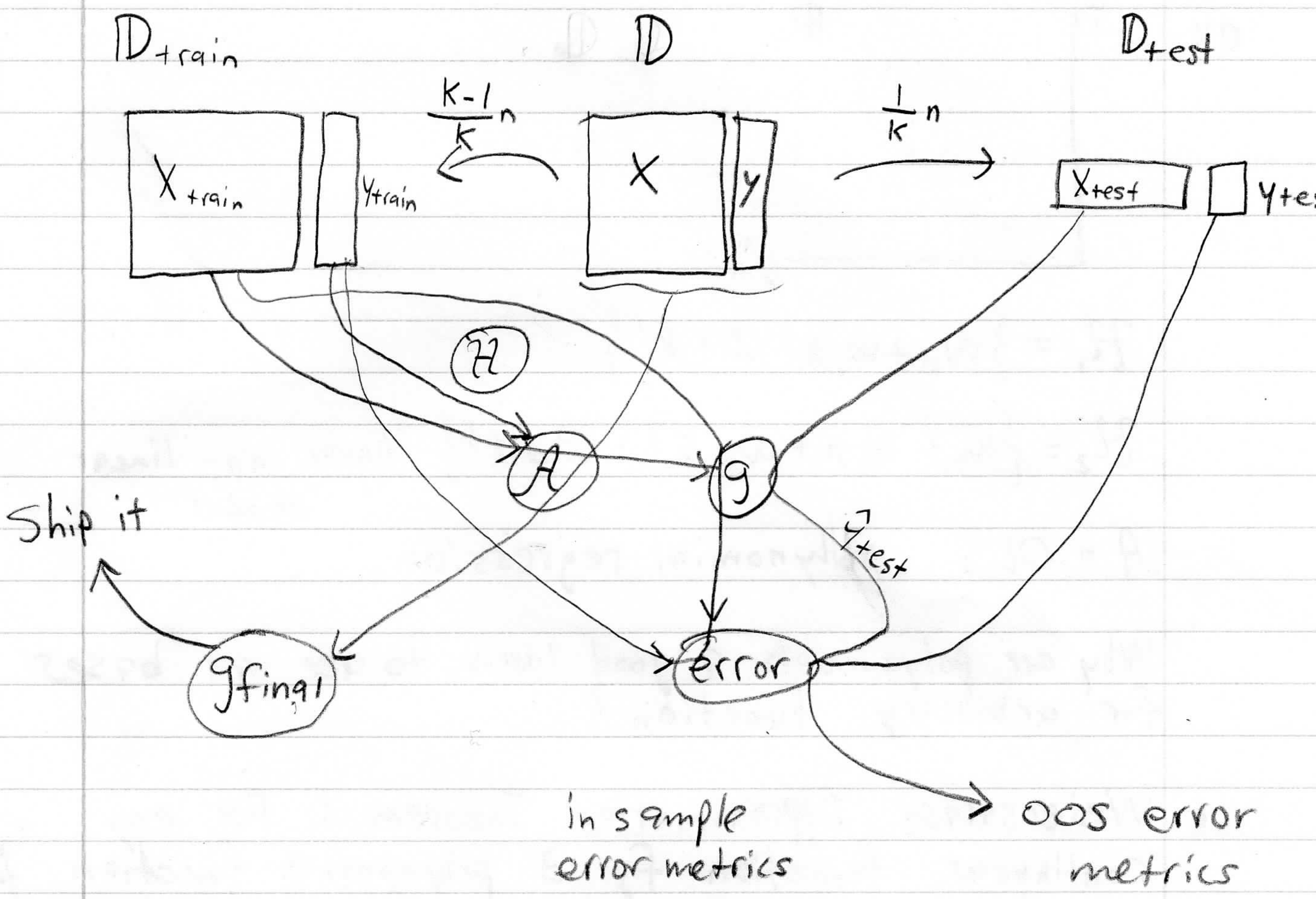
*Start of Midterm 2 Material

3/19

- R

set.seed(a number)

- Data Science



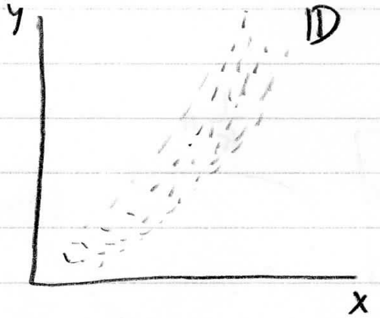
$g \approx g_{final}$ but g_{final} is better (in expectation) because $n \uparrow$.

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

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

$(t-f)$: nothing we can do

ex.



$$\mathcal{H}_1 = \{w_0 + w_1 x : \vec{w} \in \mathbb{R}^2\}$$

$$\mathcal{H}_2 = \{w_0 + w_1 x + w_2 x^2 : \vec{w} \in \mathbb{R}^3\}$$ "linear non-linear model"

$\mathcal{A}$ = OLS polynomial regression

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

- Weierstrass Approximation Theorem: For any continuous 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.

$$X_{orig} = \begin{bmatrix} 1 & X_{11} \\ 1 & X_{12} \\ \vdots & \vdots \\ 1 & X_{1n} \end{bmatrix} \longrightarrow X = \begin{bmatrix} 1 & X_{11} & X_{11}^2 \\ 1 & X_{12} & X_{12}^2 \\ \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{1n}^2 \end{bmatrix}$$

$p_{raw} + 1 = 2$ $\qquad\qquad$ $p+1 = 3$

$p_{raw} = 1$ $\qquad\qquad$ $p = 2$

original measurements $\qquad\qquad$ raw + derived

X is full rank since 3rd column can't be a lin. comb. of the first two, since they're squared terms

$$A: \ (X^T X)^{-1} X^T \vec{y} = \vec{b}$$

$$= \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$$

We can have $\mathcal{H}_3 = \{ w_0 + w_1 x + w_2 x^2 + w_3 x^3 : \vec{w} \in \mathbb{R}^4 \}$ + go up to $\mathcal{H}_{n-1}$, but we need to be careful of overfitting bad points.

$p+1 = n \Rightarrow R^2 = 100\%$,

$S_e = 0$

Perfect Model + overfit.

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

$n \times n$ matrix

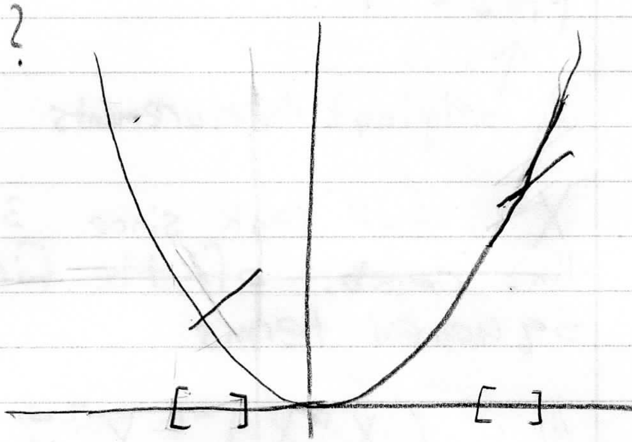
Thm (Vandermonde Matrix)

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

What is $\mathrm{Corr}[X, X^2]$?

It depends on your domain.

ex. Close to domain 0, corr ≈ 0.



$$X \xrightarrow{QR} Q$$

$X \downarrow H$, $Q \downarrow$ orthog. proj. H

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