POLYNOMIAL REGRESSION:

Was a type of "non-linear" modeling

$\mathcal{H} = \{ w_0 + w_1 x + w_2 x^2 \cdot \vec{w} \in \mathbb{R}^3 \}$

Another type of non-linear model is interaction model

Consider $p_0 = 2$

$$X = \begin{bmatrix} \downarrow & \downarrow \\ \vec{x}_1 & \vec{x}_2 \\ \downarrow & \downarrow \end{bmatrix}$$

$\mathcal{H} = \{ w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1 x_2 : \vec{w} \in \mathbb{R}^4 \}$

$g = \mathcal{A}(\mathcal{H}, \mathbb{D}) = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2$

$= b_0 + (b_1 + b_3 x_2) x_1 + b_2 x_2$

$$\Longrightarrow X = \begin{bmatrix} 1 & x_{11} & x_{12} & (x_{11})(x_{12}) \\ 1 & x_{21} & x_{22} & (x_{21})(x_{22}) \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

using $\mathcal{H} = \{ w_0 + w_1 x + w_2 x^2 : \vec{w} \in \mathbb{R}^3$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12}^2 \\ \vdots & x_{21} & x_{21}^2 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Given the same X, there are many features $\mathcal{H}$'s (and $\mathcal{A}$'s) to produce different g's (models)

for exple
$$\begin{cases} g_1 = b_0 + b_1 x_1 \\ g_2 = b_0 + b_1 x_1 + b_2 x_2 \\ g_3 = b_0 + b_1 \log(x_1) + b_2 x_2 \end{cases}$$

$$\mathcal{M}\ \text{model} \left\{ \begin{array}{l} g_4 = b_0 + b_1 x_1 + b_2 x_1^2 \\ \vdots \\ g_m = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2 \end{array} \right.$$

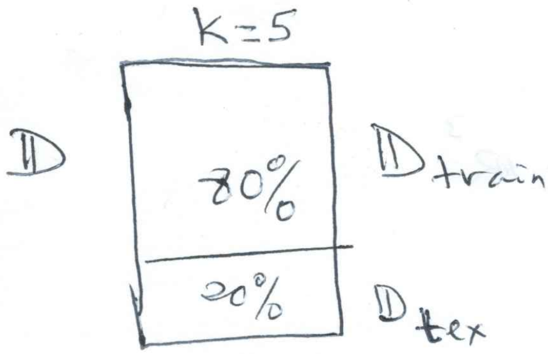
We have many models but the problem you need to choose only 1 model.

<u>Problem</u> is How to choose model?

⇒ <u>Model selection</u>

One of the fundamental question is statistics (if not all of science).

Newton Model : $a = \frac{F}{m}$



Build $g_1, \dots, g_m$ on $\mathbb{D}_{train}$

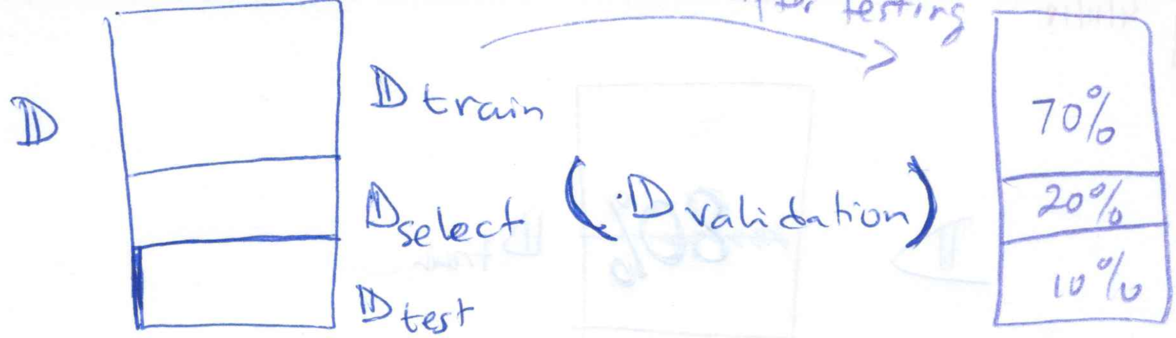
Validate $g_1, g_2, \dots, g_m$ on $\mathbb{D}_{test}$

select g_j where g_i give basic oose

Then ... fix

$$g_{final} = A_j(\mathcal{H}, \underbrace{\mathbb{D}}_{full})$$

How do we validate g_{final}?



Model Selection & Validation Protocol

① $g_j = \mathcal{A}_j(\mathcal{H}_j, \mathbb{D}_{train})$

② compare $oose_j = error(Y_{select}, \overbrace{g_j(X_{select})})$

③ Respeat step ① & ② forall $j = \{1, \dots, M\}$ models

④ $j^* = \arg\min\{oose_1, \dots, oose_M\}$

⑤ $oose_{j^*} = error(Y_{test}, \underbrace{g_j(X_{select})}_{\hat{Y}_{test}})$

⑥ Do steps $1 \sim 4$ on $\mathbb{D}$ to produce g_{final}

- Other model selections Protocol are "analytic" (based on probability & Statistics), e.xple AIC, BIC, C_p

Review
- Rank is the nber of columns dimension.
- Degree of freedom is how big dimensions.
- If you have 2 degree of freedom => 2 dimensions => you're not free you're stucked.