Lecture - 12 03/05/2020

Q-R Decomposition

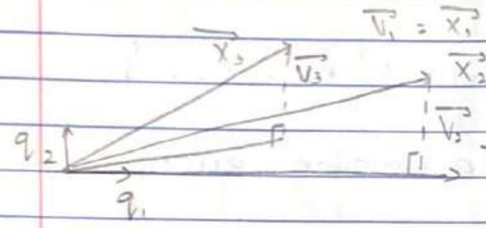
$X = QR$

i. Use to derive Q from X.

ii. Compute entries of R from X & Q

Let $X = [\vec{x_1} | \vec{x_2} | \vec{x_3} | \ldots]$

We first want an orthogonal basis for colsp[X]



$\vec{v_1} = \vec{x_1}$

$\vec{v_2} = \vec{x_2} - \text{proj}_{\vec{x_1}}(\vec{x_2})$

$\vec{v_3} = \vec{x_3} - \text{proj}_{[\vec{x_1}|\vec{x_2}]}(\vec{x_3}) = \vec{x_3} - \text{proj}_{[\vec{v_1}|\vec{v_2}]}(\vec{x_3})$

sum of proj. v_1 + proj v_2.

$= \vec{x_3} - \text{proj}_{\vec{v_1}}(\vec{x_3}) - \text{proj}_{\vec{v_2}}(\vec{x_3})$

$\vec{v_j} = \vec{x_j} - \sum_{i=1}^{i=(j-1)} \text{pro}_{\vec{v_i}}(\vec{x_j})$

$$\vec{q_1} = \frac{1}{\|\vec{v_1}\|} \vec{v_1}$$

$$\vec{q_2} = \frac{1}{\|\vec{v_2}\|} \vec{v_2} \implies Q := [\vec{q_1} \mid \vec{q_2} \mid \cdots]$$

$$\vdots$$

$$\vec{q_j} = \frac{1}{\|\vec{v_j}\|} \vec{v_j}$$

$$\begin{bmatrix} a & b & c & \cdots \\ 0 & d & e & \cdots \\ 0 & 0 & f & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & g \end{bmatrix}$$

$X = QR$

$[\vec{x_1} \mid \vec{x_2} \mid \vec{x_3} \mid \cdots] = [\vec{q_1} \mid \vec{q_2} \mid \vec{q_3} \mid \cdots]$

$$\vec{x_1} = a\vec{q_1} = \|\vec{x_1}\| \vec{q_1}$$

$$\vec{x_2} = b\vec{q_1} + d\vec{q_2}$$

$$\vec{q_1}\vec{q}^T\vec{x_1} = \vec{q_1} \frac{1}{\|x_1\|} \vec{x_1}^T \vec{x_1} = \frac{\vec{q} \, \|x_1\|^2}{\|x_1\|}$$

$$\vec{x_2} \in \text{span}[\vec{q_1}, \vec{q_2}]$$

$$\vec{x_2} = \text{proj}_{[\vec{q_1} \mid \vec{q_2}]}(\vec{x_2}) = \text{proj}_{\vec{q_1}}(\vec{x_2}) + \text{proj}_{\vec{q_2}}(\vec{x_2})$$

$$= H_1\vec{x_2} + H_2\vec{x_2}$$

$$= \vec{q_1}\underbrace{\vec{q_1}^T\vec{x_2}}_{b \text{ inner product}} + \vec{q_2}\underbrace{\vec{q_2}^T\vec{x_2}}_{d}$$

$$= b\vec{q_1} + d\vec{q_2}$$

$$X = QR$$

$$R = \begin{cases} \vec{q_i}^T \vec{x_j} & \text{if } i \leq j \\ 0 & \text{o/t.} \end{cases}$$

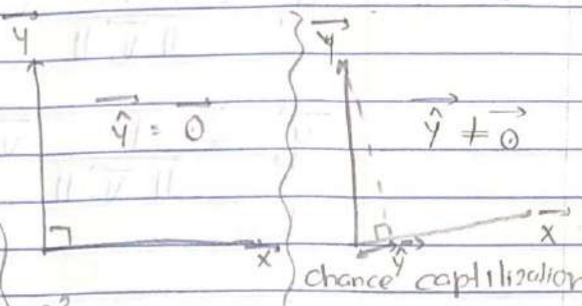
$H = X(X^T X)^{-1} X^T$
$= Q Q^T$

$\vec{\hat{y}} = H \vec{y}$

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

$\vec{b} = (Q^T Q)^{-1} Q^T \vec{y}$
$\vec{b} = Q^T \vec{y}$

$SSR = \| proj_X (\vec{y}) \| - n\bar{y}^2$



$R^2 = 0 \Rightarrow r = 0$
correlation $\Rightarrow 0$

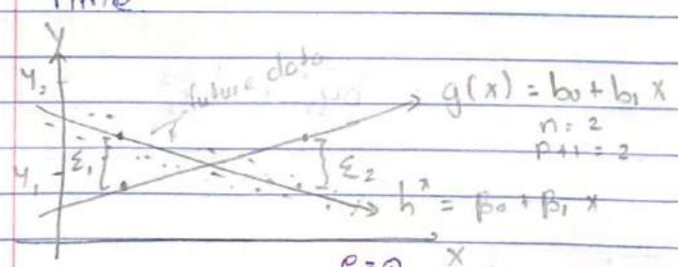
$r \neq 0$
$R^2 \neq 0$, $b \neq 0$

Generalization Error := Error when using your model on $\vec{x}_*$'s $\notin \mathbb{D}$
i.e. in the future.

$\mathbb{D} \xrightarrow{A} g \xrightarrow{\mathbb{D}} R^2$, RMSE etc (not trustworthy)
$g \xrightarrow{\mathbb{D}_*} R^2$, RMSE are trustworthy

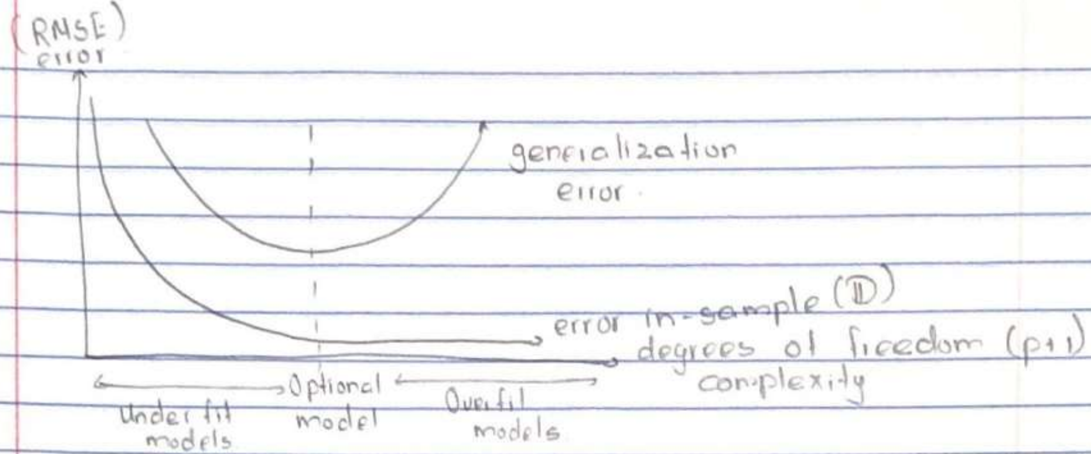
$\frac{\mathbb{D} \text{ past}}{\mathbb{D}_* \text{ future}}$ $\downarrow$ time.

Assume "stationarity" which means z's, t, f, h^*, measurement protocol of $\vec{x}_j$'s, y do not change with time.



$y = g + \overbrace{(h^* - g) + (f - h^*)}^{e = 0} + (t - f)$

fixed ε



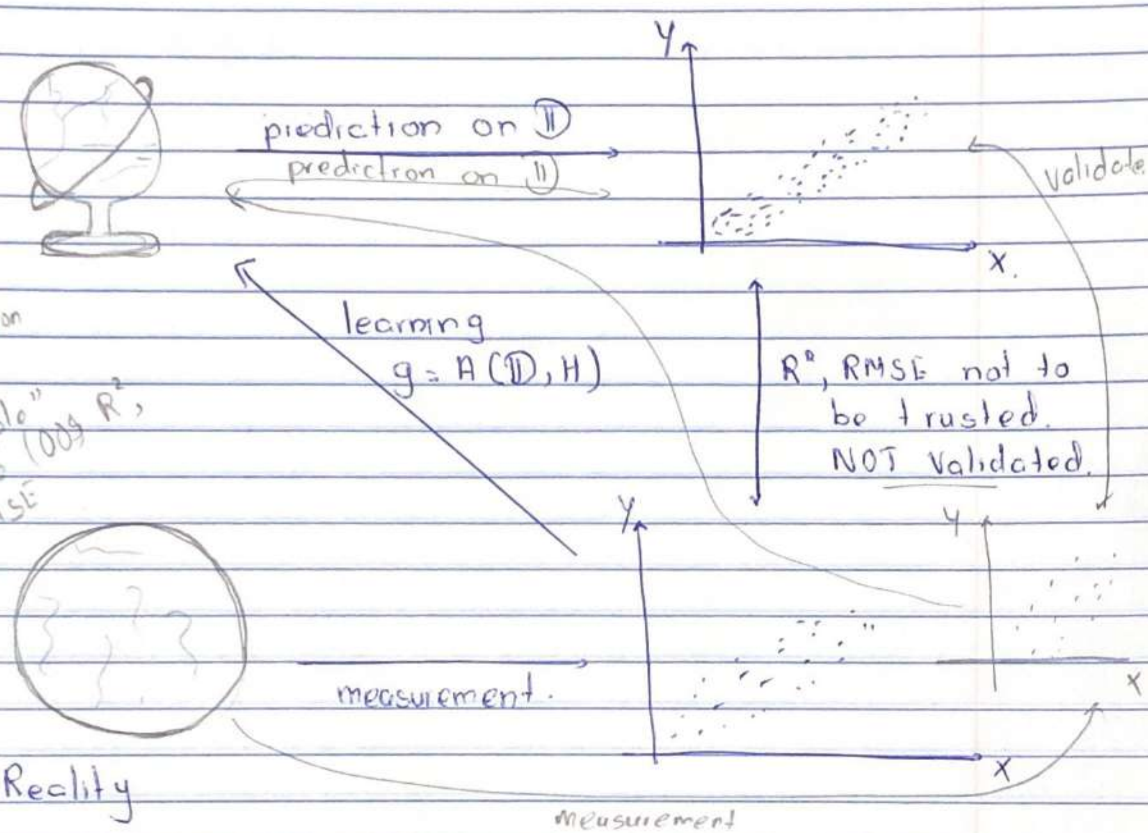
Honest validation

R^2, RMSE

"Out of sample" (OOS) R^2,

(OOS) Validate

(OOS) RMSE



$\mathbb{D} = \underbrace{\mathbb{D}_{train}} \cup \mathbb{D}_{test}$ → Simulating the future and disjoint.

→ estimate generalization error.

$g = \mathcal{A}(\mathbb{D}_{train}, H)$ — don't let A peek in this set use it to "validate"