03/05/20

Q − R Decomposition

$X = QR$

① Use Gram-Schimit to derive Q from X.
② ~~Colspace~~ Compute entries of R from X & Q

Let $X = [\,\vec{x}_1 \,|\, \vec{x}_2 \,|\, \vec{x}_3 \,|\, \dots\,]$

We first want an orthogonal basis from Colsp[X]

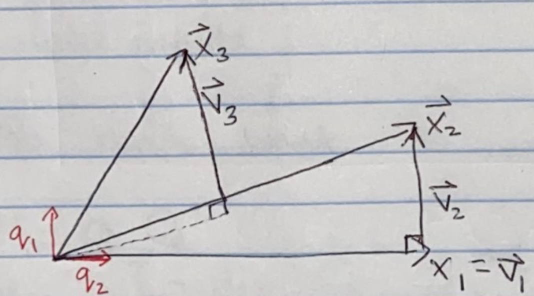
$\vec{v}_1 = \vec{x}_1$

$\vec{v}_2 = \vec{x}_2 - \text{proj}_{\vec{x}_1}(\vec{x}_2)$

$\vec{v}_3 = \vec{x}_3 - \text{proj}_{[x_1, x_2]}(\vec{x}_3) = \vec{x}_3 - \text{proj}_{[x_1, x_2]}(\vec{x}_3) = \vec{x}_3 - \text{proj}_{\vec{v}_1}(\vec{x}_3) - \text{proj}_{\vec{v}_2}(\vec{x}_3)$

$\vdots$

$\vec{v}_j = \vec{x}_j - \sum_{i=1}^{j-1} \text{proj}_{\vec{v}_i}(\vec{x}_j)$



$X = QR$

$[\vec{x}_1|\vec{x}_2|\vec{x}_3|\dots] = [\vec{q}_1|\vec{q}_2|\vec{q}_3|\dots] \underbrace{\begin{bmatrix} a & b & c & \dots \\ 0 & d & e & \dots \\ 0 & 0 & f & \dots \\ 0 & 0 & 0 & \dots \\ & & & \ddots \end{bmatrix}}_{\text{upper triangular}}$

$\vec{x}_1 = a\vec{q}_1 = \|\vec{x}_1\|\,\vec{q}_1$

$\vec{x}_2 = b\vec{q}_1 + d\vec{q}_2$

$\vec{q}_1 \vec{q}_1^T \vec{x}_1 = \vec{q}_1 \frac{1}{\|x_1\|} \vec{x}_1^T \vec{x}_2$

$= \vec{q}_1 \cancel{\frac{\|x_1\|^2}{\|x_1\|}}$

$\begin{bmatrix} \vec{q}_1^T \vec{x}_1 & \vec{q}_1^T \vec{x}_2 \\ 0 & \vec{q}_2^T x_2 \\ \vdots & 0 \\ 0 & 0 \end{bmatrix}$

$\vec{q}_1 = \frac{1}{\|v_1\|} \vec{v}_1$

$\vec{q}_2 = \frac{1}{\|v_2\|} \vec{v}_2 \implies Q = [\,q_1 \,|\, \vec{q}_2 \,|\, \dots\,]$

$\vdots$

$\vec{q}_j = \frac{1}{\|\vec{v}_j\|} \vec{v}_j$

$\vec{x}_2 \in \text{Span}\{\vec{q}_1, \vec{q}_2\}$

$$\vec{x}_3 = \text{proj}_{[q_1 | q_2]}(x) = \text{proj}_{\vec{q}_1}(\vec{x}_2) + \text{proj}_{\vec{q}_2}(\vec{x}_2)$$
$$= H_1 \vec{x}_2 + H_3 \vec{x}_2$$
$$= \vec{q}_1 \vec{q}_1^T x_2 + \vec{q}_2 q_2^T \vec{x}_2$$
$$= b\vec{q}_1 + d\vec{q}_2$$

Recall:

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

$H = QQ^T$

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

not equal to each other

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

$\vec{b}' = (Q^T Q)^{-1} Q^T \vec{y}$

$\vec{b}' = Q^T \vec{y}$

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

No corollation means they are orthogonal

If two vectors are orthogonal then you get back exactly 0

Generalization Error:

error when using your model on $\vec{x}_Q' \notin \mathbb{D}$

i.e. in the future.

Overffiting monotonichly increases your error

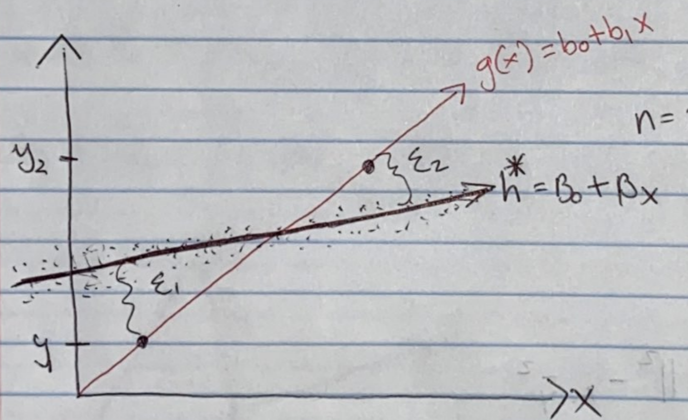
We learned that R^2, RMSE can't be trusted by looking at a demo.

$D \xrightarrow{A} g \xrightarrow{D} R^2$, RMSE, etc not trusted

$g \xrightarrow{D_*} R^2$, RMSE, etc are trusted

D past
D_* future
↓ time

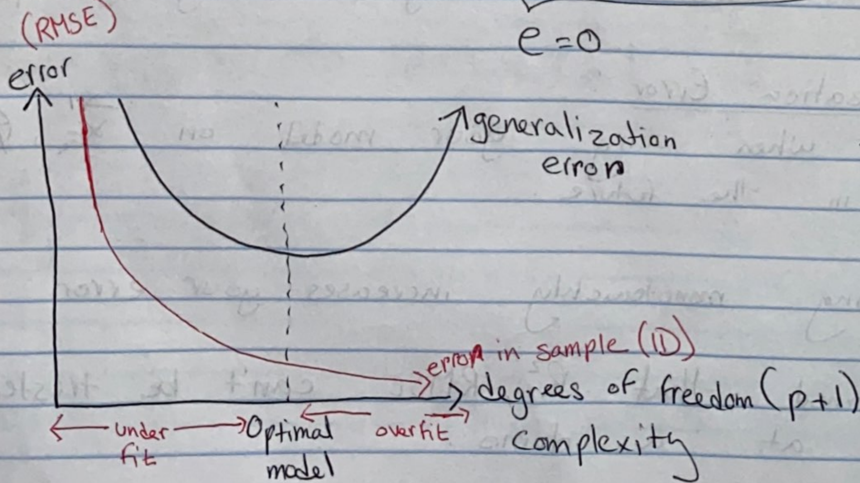
Assume "Stationarity" which mean z's, t, f, h^*, measure protocals of $\vec{x}$'s and y do not change with time.



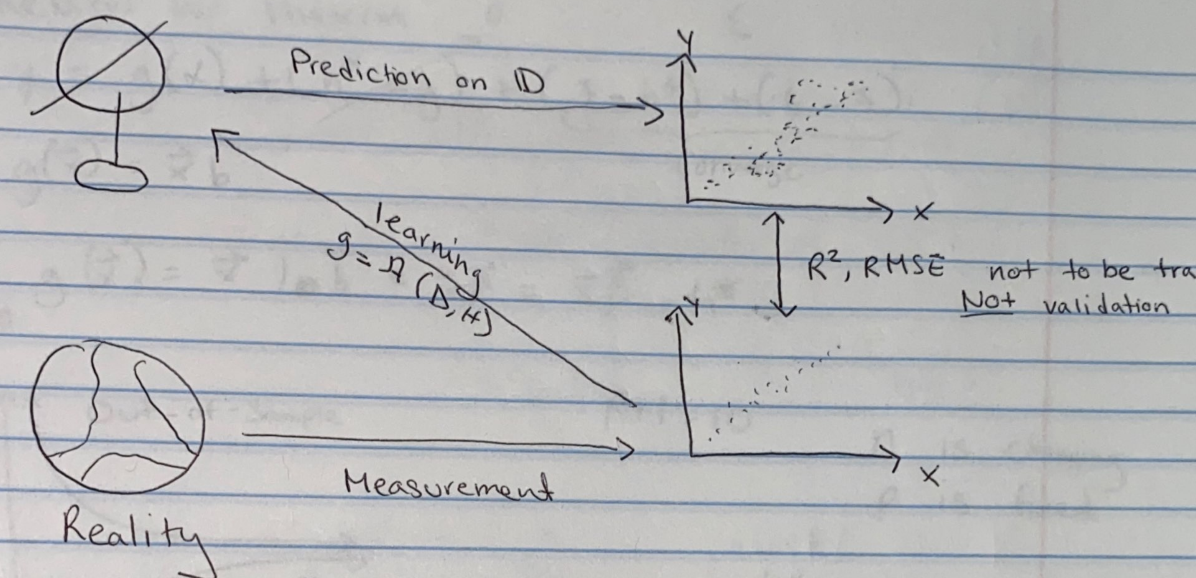
$n = 2$, $p + 1 = 2$

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

where h^*, f, t are fixed



* Overfitting increases generalization error



$\mathbb{D} = \underbrace{\mathbb{D}_{tran}}_{g = \mathcal{A}(\mathbb{D}_{tran}, \mathcal{H})} \cup \mathbb{D}_{test}$

$\mathbb{D}_{test}$ → Simulating the future and disjoint

$\mathbb{D}_{test}$ → don't let $\mathcal{A}$ peek in this set. Use it to "validate"
↓
estimates generalization error