Agenda:

1. method of moments under exogeneity
2. OLS properties under exogeneity
3. t-statistic

---

classical linear model:

(A1) (linearity): There is a vector of regression coeff's $\beta^* \in \mathbb{R}^p$ s.t. $y = X\beta^* + \boxed{\epsilon}$ vector of error terms

$$\epsilon = y - X\beta^*$$

$$\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} - x_1^T - \\ \vdots \\ - x_n^T - \end{bmatrix} \beta^*$$

(A2) (exogeneity) $E[\epsilon | X] = 0$

(A3) (spherical error variance) $E[\epsilon\epsilon^T | X] = \sigma^2 I_n$

Method of moments:

Define $\underbrace{m(\beta)}_{\mathbb{R}^p} \triangleq E[\underset{p\times 1}{x_i}(\underset{1\times 1}{y_i} - \underset{1\times p}{x_i^T}\underset{p\times 1}{\beta})]$, $\beta \in \mathbb{R}^p$

Notice $m(\beta^*) = E[x_i(y_i - x_i^T\beta^*)]$ (plug in β^*)

$= E[x_i \epsilon_i]$ (linearity)

$= E[E[x_i\epsilon_i | X]]$ (tower property / law of total exp)

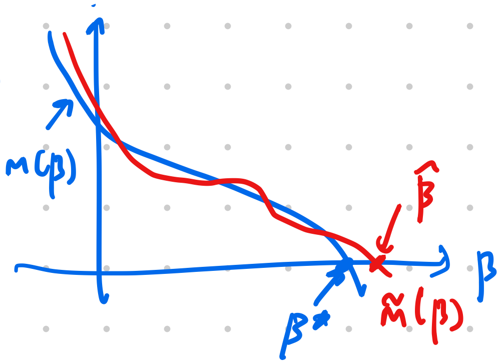
$= E[x_i \underbrace{E[\epsilon_i | X]}_{=0 \text{ by exogeneity}}]$

$= 0_p$

Approximate m with a sum:

$$m(\beta) \approx \tilde{m}(\beta) \triangleq \frac{1}{n} \sum_{i=1}^{n} x_i (y_i - x_i^T \beta)$$

Estimate β^* with $\hat{\beta}$ s.t. $\tilde{m}(\hat{\beta}) = 0$



$$\tilde{m}(\beta) = \frac{1}{n} \sum_{i=1}^{n} x_i (y_i - x_i^T \beta)$$

$$= \frac{1}{n} \underbrace{\begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}}_{p \times n} \underbrace{\begin{bmatrix} y_1 - x_1^T \beta \\ \vdots \\ y_n - x_n^T \beta \end{bmatrix}}_{n \times 1} \qquad X = \underbrace{\begin{bmatrix} - x_1^T - \\ \vdots \\ - x_n^T - \end{bmatrix}}_{n \times p}$$

$$= \frac{1}{n} X^T (y - X\beta)$$

Note $\tilde{m}(\hat{\beta}) = 0$ is the normal eq's

H: hat matrix

$$H \triangleq \underset{n\times p}{X} \underset{p\times n \; n\times p}{(X^T X)^{-1}} \underset{p\times n}{X^T}$$

$Hv = \text{argmin}_{u \in \text{span}(X)} \|u - v\|_2$

$H^2 = H$

$H^T = H$

$$H^2 = \underbrace{X \cancel{(X^T X)^{-1}}}_{H} \underbrace{\cancel{X^T} X (X^T X)^{-1} X^T}_{H}$$

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

Notice H is a projector

vector of fitted values: $\hat{y} = X\hat{\beta}$

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

$$= Hy$$

vector of OLS residuals: $\hat{\epsilon} = y - \hat{y}$

$$= \underbrace{X\beta^* + \epsilon}_{y} - \underbrace{X\hat\beta}_{\hat y}$$

$$= X\beta^* + \epsilon - X\underbrace{(X^T X)^{-1} X^T y}_{\hat\beta}$$

$$= X\beta^* + \epsilon - X(X^T X)^{-1} X^T \underbrace{(X\beta^* + \epsilon)}_{y}$$

$$= X\beta^* + \epsilon - X\cancel{(X^T X)^{-1}}\cancel{X^T X}\beta^* - X(X^T X)^{-1} X^T \epsilon$$

$$= \cancel{X\beta^*} + \epsilon - \cancel{X\beta^*} - \underbrace{X(X^T X)^{-1} X^T}_{H}\epsilon$$

$$= \epsilon - H\epsilon$$

$$= (I_n - H)\epsilon$$

(uncentered) $R^2 \doteq 1 - \frac{\|\hat\epsilon\|_2^2}{\|y\|_2^2}$ "fraction of explained variance"

Notice: $\|y\|_2^2 = \|\hat y + \hat\epsilon\|_2^2$

$$= \|\hat y\|_2^2 + 2\cancel{\hat y^T \hat\epsilon}^{\,0} + \|\hat\epsilon\|_2^2$$

$$= \|\hat y\|_2^2 + \|\hat\epsilon\|_2^2$$

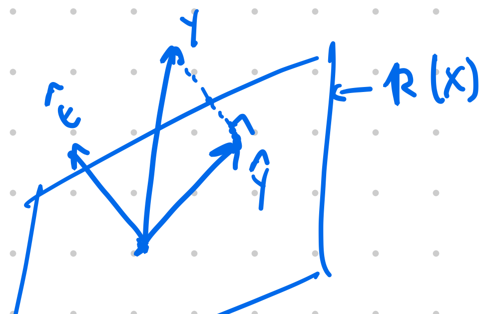
Claim: $\hat y^T \hat\epsilon = 0$

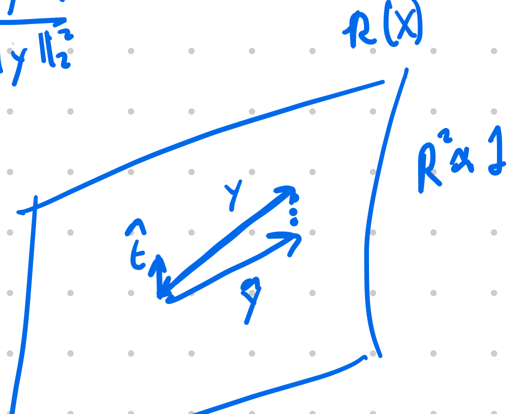
$$\underbrace{(X\hat\beta)^T}_{\hat y} \underbrace{(I_n - H)\epsilon}_{\hat\epsilon} = \hat\beta^T X^T (I_n - H)\epsilon$$

$$= \epsilon^T \underbrace{(I_n - H)X}_{=0}\hat\beta$$

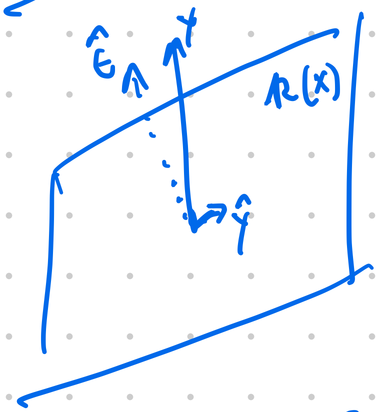
$$X - \boxed{HX}, \quad HX = X$$

uncentered $R^2 = 1 - \frac{\|\hat\epsilon\|_2^2}{\|\hat y\|_2^2 + \|\hat\epsilon\|_2^2} = \frac{\|\hat y\|_2^2}{\|y\|_2^2}$





$R^2 \approx 0$



$\frac{1}{n}\|y\|_2^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2$

centered $R^2 = \frac{\|\hat{y} - 1_n\bar{y}\|_2^2}{\|y - 1_n\bar{y}\|_2^2}$, when $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$