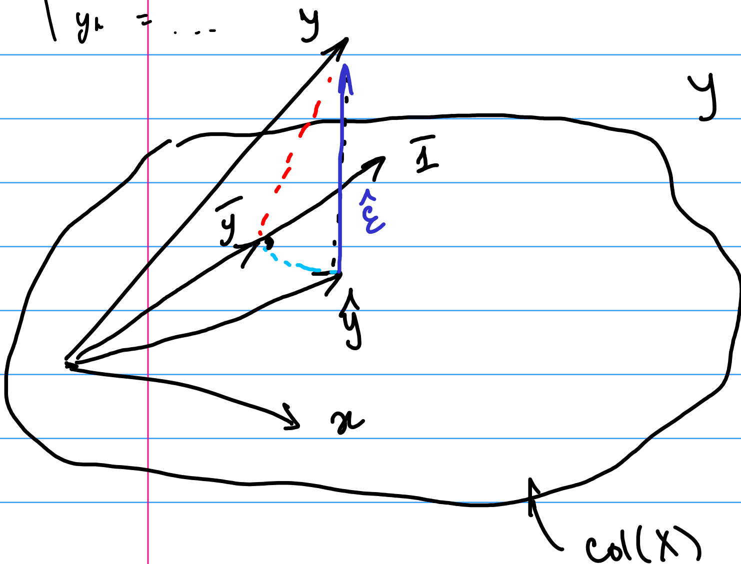
$$\begin{cases} y_1 = 1\cdot\beta_0 + \beta_{11} + \dots + \beta_{k1} + \varepsilon_1 \\ \vdots \\ y_n = \dots \end{cases}$$

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{k1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{1n} & \cdots & x_{kn} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix}$$

$$y = X\beta + \varepsilon \qquad \varepsilon \sim \mathrm{Norm}(0, \sigma_\varepsilon^2 \cdot \mathbb{1})$$



$\hat{\beta}_{OLS}$

$$\hat{\varepsilon} = y - \hat{y}$$

$$\hat{y} = \underset{\hat{\beta}}{\arg\min} \|\hat{\varepsilon}\|^2 = \underset{\hat{\beta}}{\arg\min} \|y - \hat{y}\|^2$$

$$X^T \hat{\varepsilon} = 0$$

$$X^T (y - \hat{y}) = 0$$

$$X^T (y - X\hat{\beta}) = 0$$

$$X^T y = X^T X \hat{\beta} \Rightarrow \hat{\beta} = \underline{(X^T X)^{-1} X^T y}$$

- $y_i = \beta x_i + \varepsilon_i \qquad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$$\hat{\beta} = \left((x_1 \dots x_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)^{-1} (x_1 \dots x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} =$$

$$= \frac{\sum x_i y_i}{\sum x_i^2}$$

- $y_i = \alpha + \varepsilon_i \qquad X = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

$$\hat{\alpha} = \frac{\sum y_i}{n} = \bar{y}$$

Problem 1. $y_i = \underbrace{\alpha + \beta x_i}_{\parallel} + \varepsilon_i$ Geom:

(a) $\frac{1}{n}\sum \hat{\varepsilon}_i = 0$ $\quad y_i = \underbrace{\hat{\alpha} + \hat{\beta} x_i}_{\hat{y}_i} + \hat{\varepsilon}_i$ (a) $1'\hat{\varepsilon} = 0 \Rightarrow \sum \varepsilon_i = 0$

(c) $X'\hat{\varepsilon} = 0$

(d) $\hat{y}'\hat{\varepsilon} = 0$

$$\sum (y_i - (\hat{\alpha} + \hat{\beta} x_i)) = \{\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}\} =$$

$$= \sum (y_i - (\bar{y} - \hat{\beta}\bar{x} + \hat{\beta} x_i)) =$$

$$= \underbrace{\sum (y_i - \bar{y})}_{=0} - \hat{\beta} \cdot \underbrace{\sum (x_i - \bar{x})}_{=0} = 0$$

$$\bar{y} = \frac{\sum y_i}{n}$$

(b) $\frac{1}{n}\sum \hat{y}_i = \frac{1}{n}\sum y_i$

$$\sum y_i = \sum (\hat{y}_i + \hat{\varepsilon}_i) = \sum \hat{y}_i + \underbrace{\sum \hat{\varepsilon}_i}_{0} = \sum \hat{y}_i$$

(c) $\frac{1}{n}\sum x_i \hat{\varepsilon}_i = 0 \iff s_{x,\hat{\varepsilon}} = 0$

$$\sum \hat{\varepsilon}_i x_i = \underbrace{\sum \hat{\varepsilon}_i x_i - \sum \hat{\varepsilon}_i \bar{x}}_{\{\bar{x} \cdot \sum \varepsilon_i = 0\}} = \sum \hat{\varepsilon}_i (x_i - \bar{x}) =$$

$$= \sum ((y_i - \bar{y}) - \hat{\beta}(x_i - \bar{x}))(x_i - \bar{x}) =$$

$$= \sum (y_i - \bar{y})(x_i - \bar{x}) - \hat{\beta} \cdot \sum (x_i - \bar{x})^2 =$$

$$= \sum (y_i - \bar{y})(x_i - \bar{x}) - \sum (y_i - \bar{y})(x_i - \bar{x}) = 0$$

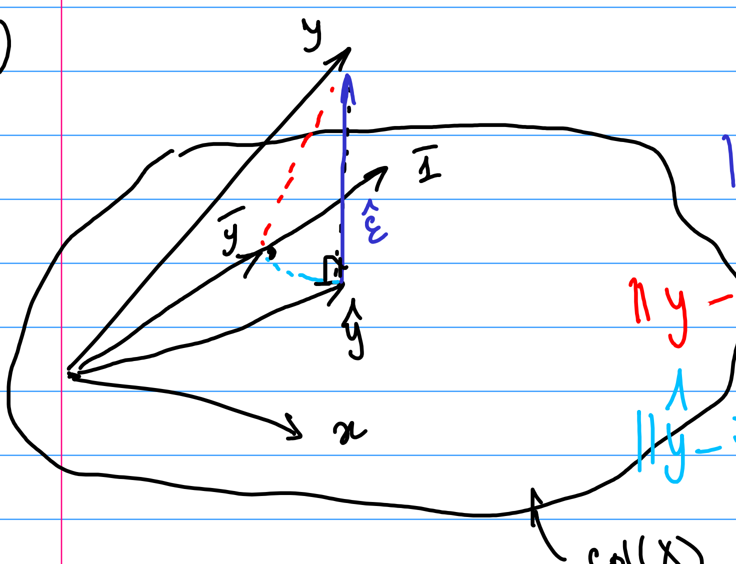
$$\hat{\beta}\sum(x_i-\bar{x})^2 = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}\cdot\sum(x_i-\bar{x})^2$$

$$= \sum(x_i-\bar{x})(y_i-\bar{y})$$

(d) $\frac{1}{n}\sum \hat{y}_i\hat{\varepsilon}_i = 0$

$$\sum \hat{y}_i\hat{\varepsilon}_i = \sum(\hat{\alpha}+\hat{\beta}x_i)\hat{\varepsilon}_i = \hat{\alpha}\underbrace{\sum\hat{\varepsilon}_i}_{=0} + \hat{\beta}\underbrace{\sum x_i\hat{\varepsilon}_i}_{=0} = 0$$

(e)



$\|y-\bar{y}\|^2 = \|\hat{y}-\bar{y}\|^2 + \|\hat{\varepsilon}\|^2$

TSS = ESS + RSS

Thm Pythagor.

$\|\hat{\varepsilon}\|^2 = RSS = \sum\hat{\varepsilon}_i^2 = \sum(y_i-\hat{y}_i)^2$

$\|y-\bar{y}\|^2 = TSS = \sum(y_i-\bar{y})^2$

$\|\hat{y}-\bar{y}\|^2 = ESS = \sum(\hat{y}_i-\bar{y})^2$

$$TSS = \sum(y_i-\bar{y})^2 = \sum((y_i-\hat{y}_i)+(\hat{y}_i-\bar{y}))^2 =$$

$$= \underbrace{\sum(y_i-\hat{y}_i)^2}_{RSS} + \underbrace{\sum(\hat{y}_i-\bar{y})^2}_{ESS} + 2\sum(y_i-\hat{y}_i)(\hat{y}_i-\bar{y}) =$$

$$= RSS + ESS + 2\underbrace{\sum\hat{\varepsilon}_i\cdot\hat{y}_i}_{=0} = RSS + ESS$$

$$R^2 = \frac{ESS / n-1}{TSS / n-1} \quad \begin{array}{l} - \text{ Explained variance} \\ - \text{ Total variance} \end{array}$$

$$R^2 = \frac{ESS}{TSS} = \hat{\rho}^2_{x,y}$$

a) $TSS = ESS + RSS$

$$R^2 = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS}$$

b) $$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{\sum(\hat{\alpha} + \hat{\beta}x_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} =$$

$$= \frac{\sum(\bar{y} - \hat{\beta}\bar{x} + \hat{\beta}x_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{\hat{\beta}^2 \sum(x_i - \bar{x})^2}{\sum(y_i - \bar{y})^2} =$$

$$= \frac{\left(S_{x,y} / S_x^2\right)^2 S_x^2}{S_y^2} = \frac{S_{x,y}^2}{S_x^2 S_y^2} = \hat{\rho}^2_{x,y}$$

$$\Rightarrow R^2 \in (0,1)$$

Problem 3.

$y_i = \alpha + \beta X_i + \varepsilon_i$

$\hat{\alpha}, \hat{\beta}$ - OLS

$\tilde{\beta}$ - OLS (Assume: $\alpha = 0$)