slide 17 · Derive the OLS estimators

First Order Conditions　　　　rem $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$$\frac{\partial J}{\partial \theta_0} = \sum (y_i - \hat{\theta}_0 - \hat{\theta}_1 x_i) = 0$$

$$n\bar{y} - n\hat{\theta}_0 - n\hat{\theta}_1 \bar{x} = 0$$

$$\bar{y} - \hat{\theta}_0 - \hat{\theta}_1\bar{x} = 0$$

$$\boxed{\hat{\theta}_0 = \bar{y} - \hat{\theta}_1 \bar{x}} \qquad (1)$$

$$\frac{\partial J}{\partial \theta_1} = \sum (y_i - \hat{\theta}_0 - \hat{\theta}_1 x_i) x_i = 0$$

$$\sum x_i y_i - n\hat{\theta}_0 \bar{x} - \hat{\theta}_1 \sum x_i^2 = 0$$

by (1)

$$\sum x_i y_i - (\bar{y} - \hat{\theta}_1 \bar{x})\bar{x} n - \hat{\theta}_1 \sum x_i^2 = 0$$

$$\sum x_i y_i - \bar{x}\bar{y} n + \hat{\theta}_1 \bar{x}^2 n - \hat{\theta}_1 \sum x_i^2 = 0$$

$$\sum x_i y_i - \bar{x}\bar{y} n - \hat{\theta}_1 \left(\sum x_i^2 - \bar{x}^2 n\right) = 0$$

$$\boxed{\hat{\theta}_1 = \frac{\sum x_i y_i - \bar{x}\bar{y} n}{\sum x_i^2 - \bar{x}^2 n} = \frac{\mathrm{cov}(x,y)}{\mathrm{var}(x)}}$$

~~Rem~~ $Var(x) = \frac{1}{n} \sum (x_i - \bar{x})^2 = \frac{1}{n} \sum x_i^2 - \bar{x}^2$

$cov(x,y) = \frac{1}{n} (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum x_i y_i - \bar{x}\bar{y}$

$$corr(x,y) = \frac{cov(x,y)}{\sqrt{var(x) \cdot var(y)}}$$

- Is the solution unique?

$$\nabla^2 f(\theta_0, \theta_1) = \begin{bmatrix} -n & -\sum x_i \\ -\sum x_i & -\sum x_i^2 \end{bmatrix}$$

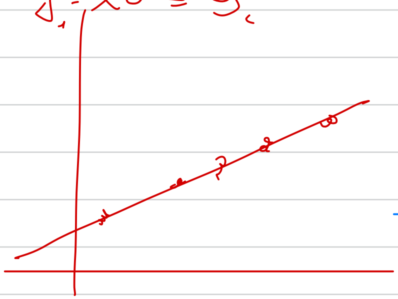
is it positive definite?

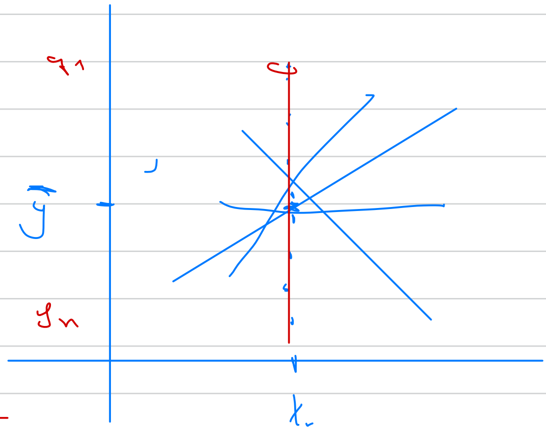
$$\det\left(\nabla^2 J(\theta_0, \theta_1)\right) = n \sum x_i^2 - \left(\sum x_i\right)^2 > 0$$

$$\iff \frac{1}{n} \sum x_i^2 - \bar{x}^2 > 0 \iff Var(x) > 0$$

$$\frac{1}{n} \sum (x_i - \bar{x})^2$$

$\hat{y}_i = x\hat{\theta} = y_i$





For the 1D linear regression

$$\hat{\theta}_0 = \bar{y} - \hat{\theta}_1 \bar{x}$$

$$\hat{\theta}_1 = cov(x,y) / var(x)$$

The solution is unique $\iff Var(X) > 0$

Variance analysis correlation, R^2

__recall__ $\hat{\theta}_1 = \frac{cov(x,y)}{var(x,y)}$

centered data (x', y') $\Big\}$ $x_i' = x_i - \bar{x}$

$\hat{\theta}_0 = 0$; $\hat{\theta}_1 = \frac{\sum x_i y_i}{\sum (x_i')^2}$. $\Rightarrow \hat{\theta}_1 \sum (x_i')^2 - \sum x_i' y_i' = 0$ (3)

predicted value at x_i' : $\hat{y}_i' = \hat{\theta}_1 x_i'$ (2)

residual $\hat{\varepsilon}_i = y_i' - \hat{y}_i'$ $\rightarrow$ $y_i' = \hat{\varepsilon}_i + \hat{y}_i'$ (1)

__Exer 1__ Show that $\sum_{i=1}^{n} (y_i')^2 = \sum_{i=1}^{n} (\hat{y}_i')^2 + \sum \hat{\varepsilon}_i^2$

Note that from (1)

$\sum (y_i')^2 = \sum_{i=1}^{n} (\hat{y}_i') + \sum \hat{\varepsilon}_i^2 + 2 \underbrace{\sum \hat{\varepsilon}_i \hat{y}_i'}$

We show that $\sum \hat{\varepsilon}_i \hat{y}_i' = 0$

$\sum \hat{\varepsilon}_i \hat{y}_i' = \sum (y_i' - \hat{y}_i') \hat{y}_i' = \sum (y_i' - \hat{\theta}_1 x_i) \hat{\theta}_1 x_i =$

$= \hat{\theta}_1 \sum (y_i' - \hat{\theta}_1 x_i') x_i' =$

$= \hat{\theta}_1 \sum (y_i' x_i' - \hat{\theta}_1 \sum (x_i')^2) \overset{(3)}{=} 0$

Interpretation

$\sum (y_i')^2 = \sum (y_i - \bar{y})^2$ variance in the data

$\sum (\hat{y}_i')^2 = \sum (\hat{y}_i - \bar{y})^2$ variance on the predict.

$\sum \hat{\varepsilon}_i^2 = \sum (y_i - \hat{y}_i)^2$ variance the residuals

definition R^2 determination coefl

$$R^2 = \frac{\sum (\hat{y}_i')^2}{\sum (y_i')^2}$$

exer. 2 Show that $R^2 = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum (y_i')^2}$

$$\frac{\sum (y_i')^2}{\sum (y_i')^2} = \frac{\sum \hat{\varepsilon}_i^2}{\sum (y_i')^2} + \frac{\sum \hat{y}_i'}{\sum (y_i')^2} \rightarrow \text{exer } 1$$

$$1 = \frac{\sum \hat{\varepsilon}_i^2}{\sum (y_i')^2} + R^2$$

by reordering we get the result

exer 3 Show that $R^2 = corr(x,y)^2$

$$R^2 = \frac{\sum (\hat{y}_i')^2}{\sum (y_i')^2} = \frac{\sum (\hat{\theta}_1 x_i')^2}{\sum (y_i')^2} = \hat{\theta}_1^2 \frac{\sum (x_i')^2}{\sum (y_i')^2} =$$

$$= corr(x,y)^2 \frac{var(y)}{var(x)} \frac{\sum (x_i')^2}{\sum (y_i')^2} = corr(x,y)^2$$

$[-1, 1]$

Interpretation $R^2 \in [0, 1]$