EC402 (2023/24)

Raguir's Speaking Notes for Class

Vassilis PS2 (WK3, AT)

PLAN

(Q1) Let's chat about it in class.

(Q2) Vassilis has provided a very detailed explanation indeed – I'll leave you to read it.

My only additional comment is that Dataset 3 helps demonstrate that OLS is more susceptible to outlying observations than is LAD.

Dataset 4 helps demonstrate that both methods are equally sensitive to influential observations.

[Tip: Keep an eye out for the story behind these sorts of observations in empirical work.]

Q3, Q4, Q5 SEE MY EQUATIONS BELOW

(Q3) $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t \;,\; t = 1, \dots, T$

Given $\hat{\beta} = (X'X)^{-1}X'y$, find explicit expressions for $\hat{\beta}_1, \hat{\beta}_2$.

· let $X = [i \;\; x]$ where $\underset{T\times 1}{i} = (1, \dots, 1)'$ and $x = (x_1, \dots, x_T)'$; $y = (y_1, \dots, y_T)'$

· $(X'X)^{-1} = \begin{bmatrix} i'i & i'x \\ x'i & x'x \end{bmatrix}^{-1} = \begin{bmatrix} x'x & -i'x \\ -x'i & i'i \end{bmatrix} \dfrac{1}{\Delta}$, where $\Delta = i'i\, x'x - i'x\, x'i$.

· $(X'y) = \begin{pmatrix} i'y \\ x'y \end{pmatrix}$ $\quad\therefore \hat{\beta} = \begin{bmatrix} x'x & -i'x \\ -x'i & i'i \end{bmatrix} \begin{bmatrix} i'y \\ x'y \end{bmatrix} \dfrac{1}{\Delta}$

$$\Rightarrow \hat{\beta}_2 = \left[T \sum_{t=1}^{T} x_t y_t - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} y_t \right] \Big/ \left[T \sum_{t=1}^{T} x_t^2 - \left(\sum_{t=1}^{T} x_t \right)^2 \right]$$

$$\& \quad \hat{\beta}_1 = \left[\sum_{t=1}^{T} x_t^2 \sum_{t=1}^{T} y_t - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} x_t y_t \right] \Big/ \left[T \sum_{t=1}^{T} x_t^2 - \left(\sum_{t=1}^{T} x_t \right)^2 \right]$$

REWRITING IN A MORE FAMILIAR WAY:

Note that $T\sum_{t=1}^{T}(x_t-\bar{x})(y_t-\bar{y}) = \ldots = T\left[\sum_{t=1}^{T}x_t y_t - T\bar{x}\bar{y}\right]$

[Please do try to derive if not obvious. The only trick is to realise $\sum_{t=1}^{T} z_t = T\bar{z}$]

Clearly then, $T\sum_{t=1}^{T}(x_t-\bar{x})^2 = T\left[\sum_{t=1}^{T}x_t^2 - T\bar{x}^2\right]$

Thus, $\hat{\beta}_2 = \sum_{t=1}^{T}(x_t-\bar{x})(y_t-\bar{y}) \Big/ \sum_{t=1}^{T}(x_t-\bar{x})^2$

and $\hat{\beta}_1$ ... on next slide.

$$\hat{\beta}_1 = \left[\sum_{t=1}^{T} x_t^2 \sum_{t=1}^{T} y_t - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} x_t y_t\right] \Big/ \Delta$$

What if we +/- $T^2 \bar{x}^2 \bar{y}$ in the numerator?

i.e. $$\Delta \hat{\beta}_1 = \sum_{t=1}^{T} x_t^2 \sum_{t=1}^{T} y_t \underline{- T^2 \bar{x}^2 \bar{y} + T^2 \bar{x}^2 \bar{y}} - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} x_t y_t$$

$$= \bar{y} \left[T \sum_{t=1}^{T} x_t^2 - T^2 \bar{x}^2\right] - \bar{x} \left[T \sum_{t=1}^{T} x_t y_t - T^2 \bar{x} \bar{y}\right]$$

$$\therefore \hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}$$

(Q4) - Given $y = X\beta + \varepsilon$ we know that $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$, where $\begin{cases} y \text{ is } T\times 1 \\ X \text{ is } T\times k, \text{ and} \\ \text{rank}(X) = k. \end{cases}$

- The usual convention is to define associated (symmetric and idempotent) $T\times T$ matrices:

$P_X := X(X'X)^{-1}X'$, referred to as a projection (or prediction) matrix,

and $M_X := I_T - X(X'X)^{-1}X'$, referred to as a residual-maker matrix.

- We define $\hat{y} := P_X y$ as our predicted/fitted y's,

and $\hat{\varepsilon} := M_X y$ as the residual from the fitting operation.

In other words, $y = \hat{y} + \hat{\varepsilon}$, represents our decomposition of y into two (orthogonal) additive components, the first representing movements in y explained by X, and the second represents whatever remains.

- In this question, we have $X = \mathbb{1}$ where $\mathbb{1}$ is a T-dimensional vector of ones.

- So $\hat{\beta}_{OLS} := (X'X)^{-1} X'y$

$$= (\mathbb{1}'\mathbb{1})^{-1} \mathbb{1}' y$$

$$= \left(T^{-1} \right) \sum_{t=1}^{T} y_t$$

$$= \bar{y}$$

- $P_X := X(X'X)^{-1}X'$

$$= \mathbb{1}(\mathbb{1}'\mathbb{1})^{-1}\mathbb{1}'$$

$$= \begin{bmatrix} 1/T & 1/T & \cdots & 1/T \\ 1/T & 1/T & \cdots & 1/T \\ \vdots & \vdots & \ddots & \vdots \\ 1/T & 1/T & \cdots & 1/T \end{bmatrix}$$, a $T \times T$ matrix where every element is $(1/T)$.

- $M_X := I_T - X(X'X)^{-1}X'$

$$= I_T - P_X$$

$$= \begin{bmatrix} 1-(1/T) & -1/T & \cdots & -1/T \\ -1/T & 1-(1/T) & \cdots & -1/T \\ \vdots & \vdots & \ddots & \vdots \\ -1/T & -1/T & \cdots & 1-1/T \end{bmatrix},$$ another $T \times T$ matrix.

Of course, $P_X y = \mathbb{1}\hat{\beta}_{OLS} = \mathbb{1}\,\bar{y} = (\bar{y}, \ldots, \bar{y})'$, a $T \times 1$ vector with every elem. equal to the sample mean,

and $M_X y = y - \mathbb{1}\,\bar{y} = (y_1 - \bar{y}, y_2 - \bar{y}, \ldots, y_T - \bar{y})'$ a $T \times 1$ vector containing demeaned data.

Q5(a) We have $y = X\hat{\beta} + \hat{\varepsilon} = \hat{y} + \hat{\varepsilon}$. Show that TSS = ESS + RSS.

Step 1: $y'y = (\hat{y}+\hat{\varepsilon})'(\hat{y}+\hat{\varepsilon}) = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{y} + \hat{y}'\hat{\varepsilon} + \hat{\varepsilon}'\hat{\varepsilon}$.

Note that $\hat{y}'\hat{\varepsilon} \overset{\text{why?}}{=} \hat{\varepsilon}'\hat{y} = \varepsilon' M_X X \hat{\beta} \overset{\text{why?}}{=} 0$,

$\therefore\ y'y = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon}$.

Step 2: Note that $i'\hat{\varepsilon} = i' M_X \varepsilon = (M_X i)'\varepsilon = 0$ if X contains a constant.
Further, if $\frac{i'\hat{\varepsilon}}{T} = 0$, then $\frac{i'y}{T} = \frac{i'\hat{y}}{T}$ so $\bar{y} = \bar{\hat{y}}$.

Step 3: $y'y - T\bar{y}^2 = \hat{y}'\hat{y} - T\bar{\hat{y}}^2 + \hat{\varepsilon}'\hat{\varepsilon}$

$$\Rightarrow \sum_{t=1}^{T} y_t^2 - 2\sum_{t=1}^{T} y_t \bar{y} + T\bar{y}^2 = \sum_{t=1}^{T} \hat{y}_t^2 - 2\sum_{t=1}^{T} \hat{y}_t \bar{\hat{y}} + T\bar{\hat{y}}^2 + \sum_{t=1}^{T} \hat{\varepsilon}_t^2$$

$$\Rightarrow \sum_{t=1}^{T} (y_t - \bar{y})^2 = \sum_{t=1}^{T} (\hat{y}_t - \bar{y})^2 + \sum_{t=1}^{T} \hat{\varepsilon}_t^2,$$ or TSS = ESS + RSS.

Q5(b) Show: $R^2 = \left[\widehat{\text{Corr}}(y_t, \hat{y}_t) \right]^2$

let $y^* := (y - \bar{y})$, and $\hat{y}^* := (\hat{y} - \bar{y})$, still assuming X contains a constant.

Step 1. $R^2 := \frac{ESS}{TSS} = \frac{\hat{y}^{*\prime}\hat{y}^*}{y^{*\prime}y^*} = \frac{\hat{y}^{*\prime}\hat{y}^*}{\hat{y}^{*\prime}\hat{y}^*}\frac{\hat{y}^{*\prime}\hat{y}^*}{y^{*\prime}y^*}$.

Step 2. $\hat{y}^{*\prime}\hat{y}^* = \hat{y}^{*\prime}[\hat{y} - \bar{y}] = \hat{y}^{*\prime}[y - \hat{\varepsilon} - \bar{y}] = \hat{y}^{*\prime}[y^* - \hat{\varepsilon}]$.

Step 3. $\hat{y}^{*\prime}\hat{\varepsilon} = [X(X'X)^{-1}X'y^*]' M_X \varepsilon = y^{*\prime} X (X'X)^{-1} X' (I_T - X(X'X)^{-1}X')\varepsilon = 0$.

Steps 1, 2, 3 $\Rightarrow R^2 = \frac{(\hat{y}^{*\prime} y^*)^2}{(\hat{y}^{*\prime}\hat{y}^*)(y^{*\prime}y^*)} = \widehat{\text{Corr}}(y_t, \hat{y}_t)^2$.