Estatística 2 - Lista 2

Luciano Fabio Busatto Venturim
26/07/21

(1) a) (3.4) We have that $X'\hat{e} =$
$$X'(y - X\hat{\beta}) = X'y - X'X(X'X)^{-1}X'y = 0.$$
Thus, in particular, since $X = [X_1 \; X_2]$, $X_2'\hat{e} = 0$:
$X_2 = X\Gamma$, where $\Gamma = [0 \; I]'$, and I is an identity matrix with the same dimension as the number of columns of X_2.
So, $X_2'(y - X\hat{\beta}) = \Gamma'X'(y - X\hat{\beta}) = \Gamma 0 = 0$.

b) (3.5) Let $\hat{e} = X\hat{\varphi} + \hat{\mu}$ be the regression model. We have that
$$\hat{\varphi} = (X'X)^{-1}X'\hat{e} = (X'X)^{-1}X'(y - X\hat{\beta}) =$$
$$= (X'X)^{-1}X'(y - X(X'X)^{-1}X'y) =$$
$$= (X'X)^{-1}X'y - (X'X)^{-1}X'y = 0$$
where $\hat{\beta}$ is the coefficient vector from the regression on y on X.

c) (3.10) $P = X(X'X)^{-1}X'$
$$P_i = X_i(X_i'X_i)^{-1}X_i' \quad i = 1, 2.$$

Now, $X'X = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} =$
$\begin{bmatrix} X_1'X_1 & 0 \\ 0 & X_2'X_2 \end{bmatrix}$, thus, $(X'X)^{-1} = \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_1)^{-1} \end{bmatrix}$

and $P = [X_1\ X_2]\begin{bmatrix}(X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1}\end{bmatrix}\begin{bmatrix}X_1' \\ X_2'\end{bmatrix} =$

$$= X_1(X_1'X_1)^{-1}X_1' + X_2(X_2'X_2)^{-1}X_2' = P_1 + P_2$$

d) (3.12) The first equation (3.53) cannot be estimated be OLS. See that if, define $X = [1\ d_1\ d_2]$ and $\beta = (\mu\ \alpha_1\ \alpha_2)'$, we can rewrite 3.5.3 as

$$y = X\beta + e$$

However, $(X'X)$ is not invertible, since $d_2 = (-1)(d_1 - 1)$, i.e., X does not have full rank. Thus, $\hat{\beta}$ is not well-defined in this model.

The last two equations can be estimated since $X_1 = [d_1\ d_2]$ and $X_2 = [1\ d_1]$ have full rank.

(ii) The equation (3.55) is more general, since it includes an intercept and d_1 and d_2 convey the same information.

Note that, in (3.54)

$$E(y_i|d_{1i}=0) = E(y_i|d_{2i}=1) = \alpha_2,$$
$$E(y_i|d_{1i}=1) = E(y_i|d_{2i}=0) = \alpha_1$$

and in (3.55)

$$E(y_i|d_{1i}=0) = \mu,$$
$$E(y_i|d_{1i}=1) = \mu + \phi,$$

so

$$\alpha_2 = \mu,$$
$$\alpha_1 = \mu + \phi$$

(ii) $1_n' d_1 = n_1$ and $1_n' d_2 = n_2$

(iii) Letting $\alpha = (\alpha_1, \alpha_2)'$ and $X = [d_1 \; d_2]$,
$y = X\alpha + e$.

The OLS estimator is

$$\hat{\alpha} = (X'X)^{-1}X'y = (1/n\, X'X)^{-1}(1/n\, X'y) =$$
$$= (1/n\, X'X)^{-1}(1/n\, X'(X\alpha + e)) = \alpha + (1/n\, X'X)^{-1}(1/n\, X'e)$$

If $E(x_i e_i) = 0$, as $n \to \infty$, $(1/n\, X'X)^{-1} \xrightarrow{p} E(x_i x_i')$
and $(1/n\, x'e) \xrightarrow{p} E(x_i e_i) = 0$, so
$\hat{\alpha} \xrightarrow{p} \alpha$

e) (3.16) See that

$$\hat{e}'\hat{e} = \min_{\beta_1, \beta_2} (y - X_1\beta_1 - X_2\beta_2)'(y - X_1\beta_1 - X_2\beta_2)$$

and
$$\tilde{e}'\tilde{e} = \min_{\beta_1} (y - X_1\beta_1)'(y - X_1\beta_1)$$
$$= \min_{\beta_1, \beta_2 = 0} (y - X_1\beta_1 - X_2\beta_2)'(y - X_1\beta_1 - X_2\beta_2)$$

Since $\tilde{e}'\tilde{e}$ is the restricted minimum,
$\tilde{e}'\tilde{e} \geq \hat{e}'\hat{e}$

Now, from the definition of R^2,

$$R_1^2 = 1 - \frac{\tilde{e}'\tilde{e}}{(y - 1_n\bar{y})'(y - 1_n\bar{y})} \leq 1 - \frac{\hat{e}'\hat{e}}{(y - 1_n\bar{y})(y - 1_n\bar{y})} = R_2^2.$$

Moreover, $R_1^2 = R_2^2 \iff \tilde{e}'\tilde{e} = \hat{e}'\hat{e} \iff \hat{\beta}_2 = 0$

(2) Let $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $\beta = \begin{pmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{pmatrix}$, $e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$, and

$X = (1_n\ x_1\ x_2)$. Thus, the model can be written as $y = X\beta + e$ and the OLS estimator of β is $\hat{\beta} = (X'X)^{-1}X'y$ with residuals
$\hat{e} = y - X\hat{\beta}$.

a) If $y - 1_n\bar{y}$ is substituded for y, the new estimator will be

$$\hat{\beta}_a = (X'X)^{-1}X'(y - 1_n\bar{y}) = \hat{\beta} - (X'X)^{-1}X'1_n\bar{y}$$

As $X = (1_n\ x_1\ x_2)$, $1_n = X(1\ 0\ 0)'$. Thus,

$$\hat{\beta}_a = \hat{\beta} - (X'X)^{-1}X'X(1\ 0\ 0)'\bar{y} = \hat{\beta} - \begin{pmatrix} \bar{y} \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \hat{\alpha} - \bar{y} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$$

The residuals are

$$\hat{e}_a = y - 1_n\bar{y} - X\hat{\beta}_a = y - 1_n\bar{y} - X(\hat{\beta} - (1\ 0\ 0)'\bar{y})$$
$$= \hat{e} - (1_n - X(1\ 0\ 0)')\bar{y} = \hat{e}$$

b) If we use σy instead of y,

$$\hat{\beta}_b = (X'X)^{-1}X'\sigma y = \sigma\hat{\beta} = \begin{pmatrix} \sigma\hat{\alpha} \\ \sigma\hat{\beta}_1 \\ \sigma\hat{\beta}_2 \end{pmatrix}$$

and $\hat{e}_b = \sigma y - X\hat{\beta}_b = \sigma y - \sigma X\hat{\beta} = \sigma\hat{e}$.

c) If we use $X_c = (1 \ \sigma x_1 \ x_2)$ instead of X,

$$\hat{\beta}_c = (X_c' X_c)^{-1} X_c' y$$

See that $X_c = X \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 1 \end{bmatrix} = X\Gamma$ where

$\Gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Thus,

$$\begin{aligned} \hat{\beta}_c &= (\Gamma' X' X \Gamma)^{-1} \Gamma' X' y \\ &= \Gamma^{-1} (X'X)^{-1} \Gamma^{-1} \Gamma y \\ &= \Gamma^{-1} \hat{\beta} \\ &= \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_1 / \sigma \\ \hat{\beta}_2 \end{pmatrix} \end{aligned}$$

and $\hat{e}_c = y - X_c \hat{\beta}_c = y - X \Gamma \Gamma^{-1} \hat{\beta} = \hat{e}$

③ e) For each $j \in \{1, \ldots, 5\}$, let $G_{i,-j} = G_{i,1}, \ldots, G_{i,j-1}, G_{i,j+1}, \ldots, G_{i,5}$, i.e., the variables different from j.

For $j = 1, \ldots, 4$

$$E(y_i \mid x_i, G_{i,j} = 1, G_{i,-j} = 0) = w_j + \lambda_j x_i \quad \text{and}$$
$$E(y_i \mid x_i, G_{i,j} = 1, G_{i,-j} = 0) = \beta_0 + \beta_1 x_i + \delta_j + \alpha_j x_i$$

Since this holds for all x_i,

$$w_j = \beta_0 + \delta_j \qquad \delta_j = w_j - \beta_0$$
$$\lambda_j = \beta_1 + \alpha_j \quad \Rightarrow \quad \alpha_j = \lambda_j - \beta_1 \quad \forall j = 1, \ldots, 4$$

For $j = 5$, $G_{i,j} = 0 \quad \forall j = 1, \ldots, 4$, thus,

$$E(y_i \mid x_i, G_{i,5} = 1, G_{i,-5} = 0) = w_5 + \lambda_5 x_i \quad \text{and}$$
$$E(y_i \mid x_i, G_{i,5} = 1, G_{i,-5} = 0) = \beta_0 + \beta_1 x_i$$

Thus, $\beta_0 = w_5$ and $\beta_1 = \lambda_5$, and

$$\delta_j = w_j - w_5 \quad \text{and} \quad \alpha_j = \lambda_j - \lambda_5 \quad \forall j = 1, \ldots, 4$$