Gregory Kordas
Last update: June 6, 2023

ECONOMETRICS
Problem Set 4
Due on: June 13, 2023, in class.

(1) (Least Squares). Prove that in the linear regression model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$ where $\boldsymbol{X}$ includes an intercept (a column of 1's as the first regressor), the OLS plane $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$ has the following properties:

(i)

$$\bar{\boldsymbol{x}}^{\mathsf{T}}\hat{\boldsymbol{\beta}} = \bar{y}.$$

where $\bar{\boldsymbol{x}} = (1, \bar{x}_1, ..., \bar{x}_k)^{\mathsf{T}}$ is the $k \times 1$ vector of means of the independent variables x_j, $j = 1, ..., k$. This means that the point $(\bar{y}, \bar{\boldsymbol{x}}) \in \mathbb{R}^{k+1}$ satisfies the normal equations, and therefore the OLS plane always passes through the sample means when the regression includes a conscant term. We say that OLS passes through the "center-of-gravity" $(\bar{y}, \bar{\boldsymbol{x}})$ of the sample.

(ii)

$$\bar{\hat{y}} = \bar{y},$$

that is, the mean of the fitted values $\hat{\boldsymbol{y}}$ equals the mean of $\boldsymbol{y}$.

(iii)

$$\boldsymbol{1}^{\mathsf{T}}\hat{\boldsymbol{u}} = \bar{u} = 0,$$

that is, the sum and the mean of the OLS residuals is zero.

(iv)

$$\hat{\boldsymbol{y}}^{\mathsf{T}}\hat{\boldsymbol{u}} = 0 \qquad \text{or} \qquad \hat{\boldsymbol{y}} \perp \hat{\boldsymbol{u}}. \tag{0.1}$$

that is, the OLS fitted values $\hat{\boldsymbol{y}}$ and the OLS residulas $\hat{\boldsymbol{u}}$ are orthogonal vectors.

(2) (Restricted Least Squares – Theory). Consider the *restricted* regression model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$$
$$\text{such that} \qquad \boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r}$$

where $\boldsymbol{R}$ is $q \times k$ matrix and $\boldsymbol{r}$ is a $q \times 1$ vector.

(i) Derive the RLS estimator $\hat{\boldsymbol{\beta}}_R$.

(ii) Derive its expectation, $E(\hat{\boldsymbol{\beta}}_R|X)$, and variance-covariance matrix, $V(\hat{\boldsymbol{\beta}}_R|X)$, under the null hypothesis $H_0 : \boldsymbol{R\beta} = \boldsymbol{r}$.

(iii) Compare the LS and the RLS estimator under the null, that is, comment on the definiteness of the matrix $V(\hat{\boldsymbol{\beta}}|X) - V(\hat{\boldsymbol{\beta}}_R|X)$.

(3) (Restricted Least Squares – Application). Handsaker and Douglas (1937) [1] considered the Cobb-Douglas production model

$$Q_t = A\,K_t^{\alpha}\,L_t^{\beta}\,u_t$$

where Q_t is the index of the product of the manufacturing sector in Victoria, Australia for the period 1907-1928, K_t is the index of the capital and L_t is the index of the labor used as inputs, and u_t is a multiplicative error. The great success of the Cobb-Douglas production function in modeling such data contributed greatly in making it very popular in applied and theoretical work later in the 20th century.

(i) Give an economic interpretation of the parameters A, α and β.

(ii) Argue for taking logarithms to linearize the model. What is the distribution of u_t if we are to assume that $\log u_t$ is normally distributed? Also explain why there is no loss of generality in taking natural logs (ln's) instead of some other logarithm, for example $\log_{10}$. (We write log to mean ln, and write $\log_b$ if we need to take logs in some other base).

(iii) Using the dataset given below, estimate the log-log linear model

$$\log Q_t = \log A + \alpha \log K_t + \beta \log L_t + \log u_t$$

by OLS and interpret your results. In particular, comment on the overall statistical significance of the regression and the statistical significance of individual parameters.

(iv) Argue that, from an economic standpoint, in this particular case, it is reasonable to expect constant returns to scale, that is, $\alpha+\beta = 1$. Compute a 95% confidence interval for $\alpha + \beta$, and test the hypothesis of constant returns to scale.

(v) Now estimate the model under the null $H_0 : \alpha + \beta = 1$, that is estimate the RLS estimator. What is $\boldsymbol{R}$ here? How do your estimates change?

(vi) Show that in the constant-returns-to-scale case we can write the model in terms of the per-capita output $q_t = Q_t/L_t$ and the per-capita capital $k_t = K_t/L_t$ as a simple

[1] Handsaker M.L., Douglas P.H. (1937), "The theory of marginal productivity tested by data for manufacturing in Victoria, I", *Quarterly Journal of Economics*, **52:1**, pp. 1–36.

regression. Estimate this model and graph the sample points and the regression. How is this related to the restricted LS model above?

```
. list year prod lab cap
```

	year	prod	lab	cap
1.	1907	81.6	82.8	82.7
2.	1908	82.2	85.5	84.6
3.	1909	87	88.5	85.5
4.	1910	93.1	92	89.2
5.	1911	100	100	100
6.	1912	103.5	103	106.3
7.	1913	108.5	104.2	112.9
8.	1914	110.9	103.6	118.3
9.	1915	99.1	98.2	120.1
10.	1916	96.9	101.6	123.6
11.	1917	98.8	102.4	128.2
12.	1918	103.4	106.4	132.9
13.	1919	117.5	118.7	141.7
14.	1920	116.4	121	153.1
15.	1921	128.7	126.4	168.5
16.	1922	138.1	133	182.1
17.	1923	141.2	134.9	195.6
18.	1924	134.2	132.6	210.2
19.	1925	147.3	133.8	215.9
20.	1926	170.6	141.9	226.6
21.	1927	177	141.1	237.5
22.	1928	168.6	138.3	241.5