1st set assignments Introductory Econometrics

Task 1.

Consider the OLS estimator $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Write extensively the matrices $\mathbf{X}'\mathbf{X}$ and $\mathbf{X}'\mathbf{y}$ using $\mathbf{x}_i = (x_{i1}, \ldots, x_{iK})'$.

Task 2.

Show, that in the simple regression model

$$y_i = \beta_1 + \beta_2 x_{i2} + \varepsilon_i$$

the OLS estimator b_2 can be written as

$$b_2 = \frac{s_{\mathbf{x_2 y}}}{s^2_{\mathbf{x_2}}}$$

where $s_{\mathbf{x_2 y}}$ is the sample covariance of $\mathbf{x}_2$ and $\mathbf{y}$ and $s^2_{\mathbf{x_2}}$ is the sample variance of $\mathbf{x}_2$.

Show, that $b_1 = \bar{y} - b_2\bar{x}$.

Task 3.

Show, that in a semi-log model

$$\ln y_i = \beta_1 + \beta_2 x_{i2} + \varepsilon_i$$

β_2 can be interpreted as a percentage change of y_i if x_{i2} increases one unit.

Task 4.

Show the validity of the law of iterated expectations by showing that

$$E_{Z|X}[E_{Y|X,Z}(Y|X,Z)|X] = E_{Y|X}(Y|X)$$

Task 5.

Show that $E(\varepsilon_i|1, x_{i1}, x_{i2}) = 0$ implies

$$\begin{aligned} E(\varepsilon_i) &= 0 \\ E(\varepsilon_i x_{i1}) &= 0 \\ E(\varepsilon_i x_{i2}) &= 0 \end{aligned}$$

Show that $E(\varepsilon_i) = 0$ implies that $Cov(\varepsilon_i, x_{i2}) = 0$.

Task 6.

Show that for

$$y_i = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_K x_{iK} + \varepsilon_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i$$

the OLS estimate $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ implies that

a) the regression passes through the means of the dependent variable and the regressors

$$\bar{y} = \beta_1 + \beta_2 \bar{x}_{i2} + \ldots + \beta_K \bar{x}_{iK}$$

where $\bar{y} = \frac{1}{n}\sum y_i$ and $\bar{x}_{ik} = \frac{1}{n}\sum x_{ik}$.

b) $\bar{y} = \bar{\hat{y}}$ where $\hat{y}_i = \mathbf{x}_i'\mathbf{b}$ and $\bar{\hat{y}} = \frac{1}{n}\sum \hat{y}_i$

Task 7.

Estimate the Glosten/Harris(1988) model with OLS:

$$\Delta p_t = \beta_1 + \beta_2 \Delta Q_t + \beta_3 Q_t + \beta_4 Q_t V_t + \varepsilon_t$$

Names for the variables in the data set:

- $\Delta p_t \hat{=}$ `delta_p`
- $\Delta Q_t \hat{=}$ `delta_q`
- $Q_t \hat{=}$ `q`
- $Q_t V_t \hat{=}$ `qv`

Use the Excel spreadsheet `dcxft_tim.xls` to compute the OLS estimator $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ where $\mathbf{y}$ is the vector `delta_p` found in column A and $\mathbf{X}$ is the data matrix found in column B to column F. First, name the matrices, then determine your output fields where the vector $\mathbf{b}$ should be written to and use the Excel matrix functions (`mmult`, `mtrans` and `minv`, *read Excel help to learn how those functions are used*) to compute the vector $\mathbf{b}$.

As a second alternative use the Excel Add-In `Solver` to compute $\mathbf{b}$. Therefore, define an empty column vector in your spreadsheet and name it (for convenience b). Then, choose a cell where you write down the objective function (sum of squared residuals). Open `Solver` and minimize your objective function. The cells which can be varied by `Solver` to find a minimum are those constituting your pre-specified vector b.