University of Essex — Department of Economics
Session: 2011/12 — João Santos Silva

EC501 Econometric Methods and Applications

Problem Set 1: Sketch Solutions

The Classical Linear Regression Model

NB: all summations run from $i = 1, \ldots, n$, i.e., $\sum y_i = \sum_{i=1}^{n} y_i$.

1. Model 1: $y_i = \beta_1 + \epsilon_i$.

   (a) Let $S(b_{01}) = \sum (y_i - b_{01})^2$, then $b_1 = \arg\min_{b_{01}} S(b_{01})$.

$$\begin{aligned} \frac{dS(b_{01})}{db_{01}} &= -2\sum(y_i - b_{01}) \\ \left.\frac{dS(b_{01})}{db_{01}}\right|_{b_1} &= -2\sum(y_i - b_1) = 0 \\ \sum(y_i - b_1) &= 0 \quad \Rightarrow \quad \sum y_i = nb_1 \quad \Rightarrow \quad b_1 = \frac{1}{n}\sum y_i = \bar{y}. \end{aligned}$$

   (b) $\dfrac{d^2 S(b_{01})}{db_{01}^2} = 2n > 0$ which corresponds to a minimum.

   (c) $E(b_1|x_1, x_2) = n^{-1}E(\sum y_i|x_1, x_2) = n^{-1}E(n\beta_1|x_1, x_2) = \beta_1$. Then, $E(b_1) = E_{x_1,x_2}(E(b_1|x_1, x_2)) = E_{x_1,x_2}(\beta_1) = \beta_1$. Hence b_1 is an unbiased estimator of β_1.

   Model 2: $y_i = \beta_2 x_{i2} + \epsilon_i$.

   (a) Let $S(b_{02}) = \sum(y_i - b_{02}x_{i2})^2$, then $b_2 = \arg\min_{b_{02}} S(b_{02})$.

$$\begin{aligned} \frac{dS(b_{02})}{db_{02}} &= -2\sum(y_i - b_{02}x_{i2})x_{i2} \\ \left.\frac{dS(b_{02})}{db_{02}}\right|_{b_2} &= -2\sum(y_i - b_2 x_{i2})x_{i2} = 0 \\ \sum(y_i - b_2 x_{i2})x_{i2} &= 0 \Rightarrow \sum y_i x_{i2} = b_2 \sum x_{i2}^2 \quad \Rightarrow \quad b_2 = \frac{\sum y_i x_{i2}}{\sum x_{i2}^2}. \end{aligned}$$

   (b) $\dfrac{d^2 S(b_{02})}{db_{02}^2} = 2\sum x_{i2}^2 > 0$ which corresponds to a minimum.

   (c) $E(b_2|x_1, x_2) = \frac{1}{\sum x_{i2}^2}\sum E(y_i x_{i2}|x_1, x_2) = \frac{1}{\sum x_{i2}^2}\sum(\beta_2 x_{i2}^2|x_1, x_2) = \beta_2$. Then, $E(b_2) = E_{x_1,x_2}(E(b_2|x_1, x_2)) = E_{x_1,x_2}(\beta_2) = \beta_2$. Hence b_2 is an unbiased estimator of β_2.

Model 3: $y_i = \beta_1 + \beta_2 x_{i2} + \epsilon_i$.

(a) Let $S(\mathbf{b_0}) = \sum (y_i - b_{01} - b_{02}x_{i2})^2$, then $b = \arg\min_{\mathbf{b_0}} S(\mathbf{b_0})$. The first-order conditions are:

$$\begin{aligned} \left.\frac{\partial S(\mathbf{b_0})}{\partial b_{01}}\right|_{b_1,b_2} &= -2\sum (y_i - b_1 - b_2 x_{i2}) = 0, \\ \left.\frac{\partial S(\mathbf{b_0})}{\partial b_{02}}\right|_{b_1,b_2} &= -2\sum (y_i - b_1 - b_2 x_{i2})x_{i2} = 0. \end{aligned}$$

Solving these equations gives

$$b_1 = \bar{y} - b_2\bar{x}_2, \quad b_2 = \frac{n\sum x_{i2}y_i - \sum x_{i2}\sum y_i}{n\sum x_{i2}^2 - (\sum x_{i2})^2},$$

where $\bar{y}$ and $\bar{x}_2$ are the sample means of y and x_2 respectively.

(b) The Hessian matrix is

$$H = \begin{bmatrix} \dfrac{\partial^2 S}{\partial b_{01}^2} & \dfrac{\partial^2 S}{\partial b_{02}\partial b_{01}} \\ \dfrac{\partial^2 S}{\partial b_{02}\partial b_{01}} & \dfrac{\partial^2 S}{\partial b_{02}^2} \end{bmatrix} = \begin{bmatrix} 2n & 2\sum x_{i2} \\ 2\sum x_{i2} & 2\sum x_{i2}^2 \end{bmatrix}$$

Recall that for H to be positive definite, the following matrices must have a positive determinant (Sylvester's criterion):

- the upper left 1-by-1 corner of H
- the upper left 2-by-2 corner of H
- ...
- H itself.

  The first element of H is clearly positive and the determinant of H is

  $$|H| = 4n\sum x_{i2}^2 - 4\left(\sum x_{i2}\right)^2 = 4n\sum (x_{i2} - \bar{x}_2)^2 > 0,$$

  corresponding to a minimum.

(c) First note that

$$E\left(\sum x_{i2}y_i | x_1, x_2\right) = \sum x_{i2}(\beta_1 + \beta_2 x_{i2}) = \beta_1 \sum x_{i2} + \beta_2 \sum x_{i2}^2.$$

Also,

$$E\left(\sum x_{i2}\sum y_i | x_1, x_2\right) = \sum x_{i2}\sum (\beta_1 + \beta_2 x_{i2}) = \sum x_{i2}\left(n\beta_1 + \beta_2\sum x_{i2}\right).$$

Hence the conditional expectation of the numerator of b_2 is

$$n\beta_1\sum x_{i2} + n\beta_2\sum x_{i2}^2 - n\beta_1\sum x_{i2} - \beta_2\left(\sum x_{i2}\right)^2 = \beta_2\left(n\sum x_{i2}^2 - \left(\sum x_{i2}\right)^2\right)$$

which, when combined with the denominator, implies that $E(b_2|x_1, x_2) = \beta_2$. Then, $E(b_2) = E_{x_1,x_2}(E(b_2|x_1, x_2)) = E_{x_1,x_2}(\beta_2) = \beta_2$, implying that b_2 is an unbiased estimator of β_2. For b_1, note that $E(\bar{y}|x_1, x_2) = \beta_1 + \beta_2\bar{x}_2$ from above, and so we have $E(b_1|x_1, x_2) = \beta_1$ upon substitution. Then, $E(b_1) = E_{x_1,x_2}(E(b_1|x_1, x_2)) = E_{x_1,x_2}(\beta_1) = \beta_1$, which means that b_1 is an unbiased estimator of β_1.

2. Model 1: $y_i = \beta_1 + \epsilon_i$.

   (a) Here $\mathbf{X} = \ell$ where $\ell = [1, 1, \ldots, 1]'$ is an $n \times 1$ vector of ones. Hence $\mathbf{X}'\mathbf{X} = \ell'\ell = n$ and $\mathbf{X}'\mathbf{y} = \ell'\mathbf{y} = \sum y_i$.

   (b) $b_1 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\ell'\mathbf{y})/(\ell'\ell) = \sum y_i/n = \bar{y}$.

   (c) Note that $E(y_i) = \beta_1$ since $E(\epsilon_i) = 0$, and so y_i is modelled as IID variation around its mean (β_1).

   (d) The residuals are simply $e_i = y_i - \bar{y}$ which can be interpreted as deviations from the mean or as a demeaned series.

   Model 2: $y_i = \beta_1 + \beta_2 x_{i2} + \epsilon_i$.

   (a) $\mathbf{X} = [\ell, x_2]$ where x_2 denotes the $n \times 1$ vector of observations on the variable x_2. Hence

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \ell' \\ x_2' \end{bmatrix} [\ell, x_2] = \begin{bmatrix} \ell'\ell & \ell' x_2 \\ x_2'\ell & x_2' x_2 \end{bmatrix} = \begin{bmatrix} n & \sum x_{i2} \\ \sum x_{i2} & \sum x_{i2}^2 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} \ell' \\ x_2' \end{bmatrix} y = \begin{bmatrix} \ell' y \\ x_2' y \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_{i2} y_i \end{bmatrix}.$$

   Note that $|\mathbf{X}'\mathbf{X}| = n \sum x_{i2}^2 - (\sum x_{i2})^2$ and so

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{|\mathbf{X}'\mathbf{X}|} \begin{bmatrix} \sum x_{i2}^2 & -\sum x_{i2} \\ -\sum x_{i2} & n \end{bmatrix}.$$

   Hence

$$\begin{aligned} \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} &= \frac{1}{|\mathbf{X}'\mathbf{X}|} \begin{bmatrix} \sum x_{i2}^2 & -\sum x_{i2} \\ -\sum x_{i2} & n \end{bmatrix} \begin{bmatrix} \sum y_i \\ \sum x_{i2} y_i \end{bmatrix} \\ &= \frac{1}{|\mathbf{X}'\mathbf{X}|} \begin{bmatrix} \sum x_{i2}^2 \sum y_i - \sum x_{i2} \sum x_{i2} y_i \\ n \sum x_{i2} y_i - \sum x_{i2} \sum y_i \end{bmatrix} \end{aligned}$$

   which gives

$$b_2 = \frac{n \sum x_{i2} y_i - \sum x_{i2} \sum y_i}{n \sum x_{i2}^2 - (\sum x_{i2})^2},$$

   which is the expression for b_2 in Model 3 of Question 1.

   (b) Here $E(y_i | x_{i1}, x_{i2}) = \beta_1 + \beta_2 x_{i2}$ which is not constant (it depends on i). The parameter β_2 is a slope parameter that represents the responsiveness of the conditional mean of y_i to changes in x_{i2}.

Model 3: $y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$.

(a) $\mathbf{X} = [x_1, x_2]$ where x_1 and x_2 denote, respectively, $n \times 1$ vectors of observations on x_1 and x_2. Hence

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} [x_1, x_2] = \begin{bmatrix} x_1'x_1 & x_1'x_2 \\ x_2'x_1 & x_2'x_2 \end{bmatrix} = \begin{bmatrix} \sum x_{i1}^2 & \sum x_{i1}x_{i2} \\ \sum x_{i2}x_{i1} & \sum x_{i2}^2 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} y = \begin{bmatrix} x_1'y \\ x_2'y \end{bmatrix} = \begin{bmatrix} \sum x_{i1}y_i \\ \sum x_{i2}y_i \end{bmatrix}.$$

(b) Note that $|\mathbf{X}'\mathbf{X}| = \sum x_{i1}^2 \sum x_{i2}^2 - (\sum x_{i1}x_{i2})^2$ and so

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{|\mathbf{X}'\mathbf{X}|} \begin{bmatrix} \sum x_{i2}^2 & -\sum x_{i1}x_{i2} \\ -\sum x_{i1}x_{i2} & \sum x_{i1}^2 \end{bmatrix}.$$

Hence

$$\begin{aligned} \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} &= \frac{1}{|\mathbf{X}'\mathbf{X}|} \begin{bmatrix} \sum x_{i2}^2 & -\sum x_{i1}x_{i2} \\ -\sum x_{i1}x_{i2} & \sum x_{i1}^2 \end{bmatrix} \begin{bmatrix} \sum x_{i1}y_i \\ \sum x_{i2}y_i \end{bmatrix} \\ &= \frac{1}{|\mathbf{X}'\mathbf{X}|} \begin{bmatrix} \sum x_{i2}^2 \sum x_{i1}y_i - \sum x_{i1}x_{i2} \sum x_{i2}y_i \\ \sum x_{i1}^2 \sum x_{i2}y_i - \sum x_{i1}x_{i2} \sum x_{i1}y_i \end{bmatrix} \end{aligned}$$

which gives

$$b_1 = \frac{\sum x_{i2}^2 \sum x_{i1}y_i - \sum x_{i1}x_{i2} \sum x_{i2}y_i}{\sum x_{i1}^2 \sum x_{i2}^2 - (\sum x_{i1}x_{i2})^2}.$$

Here $E(y_i|x_{i1}, x_{i2}) = \beta_1 x_{i1} + \beta_2 x_{i2}$ which depends on both x_1 and x_2 and has a zero intercept.

Model 4: $y_i = \beta_1 + \beta_2 i + \epsilon_i$.

(a) $\mathbf{X} = [\ell, j]$ where $j = [1, 2, \ldots, n]'$, which is a special case of Model 2 in which $x_{i2} = i$ and $x_2 = j$. Thus

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \sum i \\ \sum i & \sum i^2 \end{bmatrix} = \begin{bmatrix} n & n(n+1)/2 \\ n(n+1)/2 & n(n+1)(2n+1)/6 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} \sum y_i \\ \sum i y_i \end{bmatrix}.$$

(b) Using the results for Model 2:

$$b_2 = \frac{n \sum i y_i - \sum i \sum y_i}{n \sum i^2 - (\sum i)^2}.$$

(c) This type of model is most typically found in a time series setting where the index i represents time and $\beta_1 + \beta_2 i$ represents a linear trend. Here the regressors are non-stochastic and $E(y_i) = \beta_1 + \beta_2 i$; that is, if $\beta_2 > 0$ the expected value of y grows linearly over time.

(d) The residuals are $e_i = y_i - b_1 - b_2 i$ which represent deviations from the trend or a detrended series.