LECTURE 10.1

Ordinary least squares estimator (OLS) for multiple LR

MATRIX APPROACH TO MULTIPLE LINEAR REGRESSION

- Suppose the model relating the regressors to the response is

$$\mathrm{y}_{\mathrm{i}} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots\ldots + + \beta_k x_{ik} + \varepsilon_i$$

In matrix notation this model can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$ Formula (1)

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1k} \\ 1 & x_{21} & x_{22} & \ldots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{nk} \end{bmatrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \text{and} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

$$y = Xb + \epsilon$$

Formula (2)

$$\begin{bmatrix} y_1 \\ \cdots\cdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{bmatrix} * \begin{bmatrix} b_1 \\ \cdots\cdots \\ b_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \cdots\cdots \\ \varepsilon_k \end{bmatrix}$$

Formula (3)

$$\sum_{i=1}^{n}(y_i - b_0 - b_1 x_{i1} - \cdots - b_k x_{ik})^2$$

$$= \sum_{i=1}^{n}\left(y_i - b_0 - \Sigma b_i x_{ij}\right)^2 = \boxed{\sum_{i=1}^{n}(e_i)^2 = \min!}$$

$$\boldsymbol{\epsilon}'\boldsymbol{\epsilon} = \left[e_1 \ldots\ldots. e_n\right]\begin{bmatrix} e_1 \\ e_2 \\ \ldots\ldots \\ e_n \end{bmatrix} = \sum_{i=1}^{n} e_i^2$$

$$\begin{bmatrix} e_1 & e_2 & \ldots & \ldots & e_n \end{bmatrix}_{1\times n}\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} e_1 \times e_1 + e_2 \times e_2 + \ldots + e_n \times e_n \end{bmatrix}_{1\times 1}$$

Formula (4)

$$\min_b e'e = (y - Xb)'(y - Xb)$$ Formula (5)

$$\min_b e'e = (y' - b'X')(y - Xb)$$ Formula (6)

$$\min_b e'e = y'y - b'X'y - y'Xb + b'X'Xb$$

Formula (7)

$$\boxed{b'X'y = (b'X'y)' = y'Xb}$$

$$\min_b e'e = y'y - 2b'X'y + b'X'Xb$$ Formula (8)

$$\frac{\partial \mathrm{b'X'y}}{\partial \boldsymbol{b}} = \frac{\partial \mathrm{b'(X'y)}}{\partial \boldsymbol{b}} = \mathrm{X'y}$$

$$\frac{\partial \mathrm{b'X'Xb}}{\partial \boldsymbol{b}} = 2\mathrm{X'Xb}$$

$$\min_b \mathbf{e'e} = \mathbf{y'y - 2b'X'y + b'X'Xb}$$

Formula (8)

$$\frac{\partial(\mathbf{e'e})}{\partial \boldsymbol{b}} = -2\mathbf{X'y} + 2\mathbf{X'Xb} \overset{!}{=} 0$$

Formula (9)

$$\mathbf{X'Xb} = \mathbf{X'y}$$

Formula (10) , normal equation

$$\mathbf{b} = \boxed{(\mathbf{X'X})^{-1}}\mathbf{X'y}$$

Formula (11)

Nonsingular
THE inverse of (X'X) exist

The inverse of the **X'X** matrix, and **X'** is the transpose of the **X** matrix

$$X'X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & x_n \\ 1 & \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$$

PROPERTIES OF THE OLS ESTIMATORS

$$\mathbf{X'Xb} = \mathbf{X'y}$$

Now substitute in y = Xb+ e to get

$$\mathbf{X'Xb} = \mathbf{X'(Xb + e)}$$

$$= \mathbf{X'Xb + X'e}$$

$$\text{therefore, } \mathbf{X'e} = 0$$

$$\begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ X_{k1} & X_{k2} & \dots & X_{kn} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} X_{11} \times e_1 + X_{12} \times e_2 + \dots + X_{1n} \times e_n \\ X_{21} \times e_1 + X_{22} \times e_2 + \dots + X_{2n} \times e_n \\ \vdots \\ \vdots \\ X_{k1} \times e_1 + X_{k2} \times e_2 + \dots + X_{kn} \times e_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

PROPERTIES OF THE OLS ESTIMATORS

- The observed values of X are uncorrelated with the residuals.

X ‘**e** = 0 implies that for every column x_k of X, $x_k^{'}e = 0$. In other words, each predictor has zero sample correlation with the residuals. Note that this does not mean that X is uncorrelated with the disturbances; we’ll have to assume this.

The sum of the residuals is zero.

The sample mean of the residuals is zero.

- The regression hyperplane passes through the means of the observed values ($\bar{x}$ and $\bar{y}$).
- The predicted values of y are uncorrelated with the residuals.

$$\hat{y} = \boldsymbol{Xb}$$

$$\text{then } \hat{y}'e = (\boldsymbol{Xb})'\boldsymbol{e} = \boldsymbol{b'X'e} = \boldsymbol{0}$$

ASSUMPTIONS

1. Random: Regressors $(X_{1i}, X_{2i}, \ldots, X_{ki}, Y_i)$, i=1,…,n, are drawn such that the i.i.d. assumption holds, is unrelated to ε.
2. e_I is an error term with conditional mean zero given the regressors, i.e., $E(e_i | X_{1i}, X_{2i}, \ldots, X_{ki}) = 0$. (zero coitional mean of the e).
3. Large outliers are unlikely, formally X1i,…,Xki and Yi have finite fourth moments.
4. No perfect collinearity.

ASSUMPTIONS

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

1. Linear:
2. Random: Regressors ($X_{1i}, X_{2i}, \ldots, X_{ki}, Y_i$) , i=1,…,n, are drawn such that the i.i.d. assumption holds, is unrelated to ε.
3. e_I is an error term with conditional mean zero given the regressors, i.e., $E(e_i | X_{1i}, X_{2i}, \ldots, X_{ki}) = 0$. (zero coitional mean of the e).
4. Large outliers are unlikely, formally X1i,…,Xki and Yi have finite fourth moments.

ASSUMPTIONS

- No perfect collinearity: **X** is an n × k matrix of full rank.
- **Homoscedasticity:**

$$E(\boldsymbol{\epsilon}'\boldsymbol{\epsilon} \mid X) = \sigma^2 I$$

Proof can be found in supporting material

ASSUMPTIONS

- No perfect collinearity: **X** is an n × k matrix of full rank.
- **Homoscedasticity:**

$$E(\boldsymbol{\epsilon}'\boldsymbol{\epsilon} \mid X) = \sigma^2 I$$

Proof can be found in supporting material