Econometrics

Natalia Nehrebecka

Part 1

Agenda

- Organisational issues
- Recommended books
- Plan of the lecture
- Ordinary Least Squares
- Function form

Organisational issues

- **Office hours:** nnehrebecka@wne.uw.edu.pl
- **Website of the course:** www.ekonometria.wne.uw.edu.pl
- **Contact:** nnehrebecka@wne.uw.edu.pl

- **Written exam**
  - exam takes 90 minutes (4 theoretical questions, 3 exercises)
  - **June, 25 starting from 4:00 PM,** room A
  - Re-take of final exam: september

- Final mark:
  - 100% exam

- **Tutorial**
  - During the tutorials we solve exercises

Recommended books

- Introductory Econometrics - A Modern Approach, J.M. Wooldridge
- Econometric Analysis, W.H. Green

Plan of the lecture

1. **May, 24 from 1.15 PM to 4.45 PM, room G**
   - Ordinary Least Squares
   - Function form (part 1)
2. **June, 1 from 1.15 PM to 4.45 PM, room G**
   - Function form (part 2)
   - Classical linear regression model
3. **June, 07 from 1.15 PM to 4.45 PM, room G**
   - Hypothesis testing
4. **June, 14 from 1.15 PM to 4.45 PM, room G**
   - Diagnostic tests
5. **June, 21 from 1.15 PM to 4.45 PM, room G**
   - Problems with the data

May, 24 from 1.15 PM to 4.45 PM

Explain the difference between time series and cross-sectional data

- **Cross-sectional data**
  - Consists of a sample of individuals, households, firms, cities, states, countries, or other units, taken at a given point in time

- **Time-series data**
  - A time series data set consists of observations on a variable or several variables over time

May, 24 from 1.15 PM to 4.45 PM

1. The Simple Regression Model
2. Partial Effects
3. Elasticity
4. Semi-Elasticity

The Simple Regression Model

$$y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_K X_{Ki} + \varepsilon_i$$

- y - dependent variable, explained variable
- $X_2, X_3, \ldots X_K$ - independent variables, explanatory variables
- ε - error term
- $\beta_1, \beta_2, \ldots \beta_K$ - parameters, coefficients
- β_1 - intercept, constant term

- Matrix notation:

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

The Simple Regression Model

$$\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} x_{11} & x_{21} & \dots & x_{K1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1N} & x_{2N} & \dots & x_{KN} \end{bmatrix}}_{\boldsymbol{X}} \underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}}_{\boldsymbol{\varepsilon}}$$

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

The Simple Regression Model

$$\hat{y}_i = b_1 + b_2 X_{2i} + \cdots + b_K X_{Ki}$$

$$e_i = y_i - \hat{y}_i$$

- $\hat{y}_i$ –fitted value,
- $b_1, \ldots b_K$ – estimators,
- e_i – resuduals.

$$y_i = b_1 + b_2 X_{2i} + \cdots + b_K X_{Ki} + e_i$$

- Matrix notation:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{e}$$

The Simple Regression Model

- $$y_i = \beta_1 + \beta_2 X_{2i} + \varepsilon_i$$

- the Ordinary Least Squares Estimates:

$$b_2 = \frac{S_{yX}}{S_X^2}$$

$$b_1 = \bar{y} - b_2 \bar{X}$$

The Simple Regression Model

- OLS for many explanatory variables:

$$y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_K X_{Ki} + \varepsilon_i$$

1. system of normal equations:

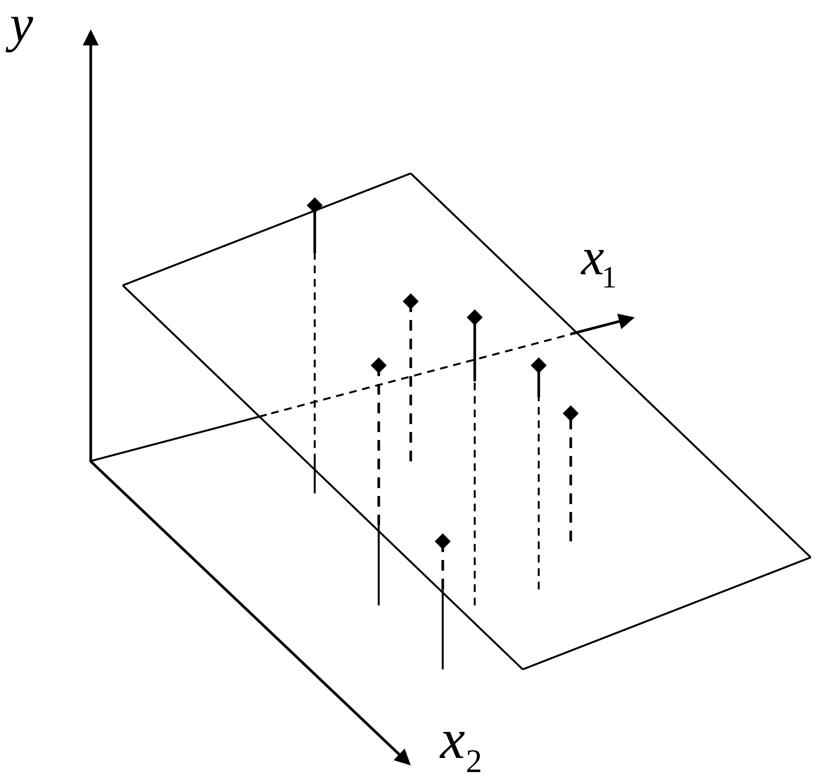
$$X'Xb = X'y$$

2. Estimator OLS:

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

The Simple Regression Model

- Regression hyperplane properties:
  - $X'e = 0$
  - $\hat{y}'e = 0$



- For models with constant:
  - $\sum_{i=1}^{N} e_i = 0$
  - $\bar{\boldsymbol{y}} = \bar{\hat{\boldsymbol{y}}}$

The Simple Regression Model

- Decomposition of the total sum of square

$$\boldsymbol{TSS = ESS + RSS}$$

- Total sum of squares

$$\boldsymbol{TSS} = \sum_{i=1}^{N} (y_i - \bar{y})^2 = (y - \bar{y})'(y - \bar{y})$$

- Explained sum of squares

$$\boldsymbol{ESS} = \sum_{i=1}^{N} (\hat{y}_i - \bar{\hat{y}})^2 = (\hat{y} - \bar{\hat{y}})'(\hat{y} - \bar{\hat{y}})$$

- Residual sum of squares

$$\boldsymbol{RSS} = \sum_{i=1}^{N} {e_i}^2 = e'e$$

The Simple Regression Model

- Goodnes of fit

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

  - never decreases and usually increases when another independent variable is added to a regression
  - The fact that R^2 never decreases when any variable is added to a regression makes it a poor tool for deciding whether one variable or several variables should be added to a model

The Simple Regression Model

- Goodnes of fit

$$\overline{R}^2 = 1 - \frac{N-1}{N-K}(1-R^2)$$

Case

- Two models were estimated on the sample with 12 observations and following results were obtained:

a) $\hat{y}_i = 2 + 1.5X_{1i} + 3X_{2i},\ R^2 = 0{,}8$

b) $\hat{y}_i = 1 + 0.8X_{1i} + 4X_{2i} + 6X_{3i},\ R^2 = 0{,}82$

- Which model should you choose and why?

Units of Measurement and Functional Form

- **Partial Effect:** $\frac{\Delta E(y)}{\Delta x_k}$
  - For $\Delta x_k = 1$ partial effect is equal to β_k
  - β_k describes the impact of a unit change in explanatory variable x_k in expected value of dependent variable holding all other regressors fixed

- **Elasticity:** $\frac{\partial lnE(y)}{\partial lnx_k}$
  - percentage change in expected value of dependent variable in response to 1% change in x_k

- **Semi-Elasticity**
  - multiplied by 100% - percentage change in expected value of dependent variable in response to 1 unit change in x_k

May, 24 from 1.15 PM to 4.45 PM

1. Explain the difference between time series and cross-secional data.
2. Write the function form of linear regression model and interpret its element.
3. Explain the relation between dependent variable, fitted values, coefficients and residuals.
4. Explain the difference between coefficients and estimators, random term and residuals.
5. Explain from where the name ”Ordinary Least Squares” is coming.
6. Derive the OLS estimator for the model with constant and one explanatory variable.
7. What is the system of normal equations?
8. Derive the OLS estimator for the model with multiple explanatory variables.
9. Why it is not possible to calculate OLS estimator in the case when the number of parameters is larger than the number of observations?
10. Prove that in the model with constant term the sum of residual is equal to zero.
11. Show that in the model with constant term the mean value of dependent variable is equal to mean value of fitted variable.
12. Prove that in the model with constant $TSS = ESS + RSS$.
13. Explain what it the interpretation of $R2$.
14. Explain why $R2$ should not be used as a criterion for comparing models.

May, 24 from 1.15 PM to 4.45 PM

15. Give definition of partial effect
16. Give definition of elasticity
17. Give definition of semi-elasticity

Dziękuję za uwagę