Tutorial 4

Hong Xiang Yue

20/03/2019

Admin

Mid semester test

- Friday 5th of April, 6pm
- Covers content from week 1 to 4
- Types of data in econometrics
- Predictive vs prescriptive analytics
- Probability and statistics
- Derivation of the OLS estimator
- Properties of the OLS estimator
- Interpret regressions

Group assignment

- Due end of week 7
- Form groups of 3-4
- Will make use of the WDI2019 file

Part B

Linear combinations of vectors

Multiplying a matrix by a column vector will produce a linear combination of the columns of the matrix.

For example:

$$\begin{bmatrix} 7 \\ 1 \\ 11 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 4 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \times -2 + 1 \times 3 \\ 4 \times -2 + 3 \times 3 \\ 2 \times -2 + 5 \times 3 \end{bmatrix}$$

is the same as

$$\begin{bmatrix} 7 \\ 1 \\ 11 \end{bmatrix} = -2 \times \begin{bmatrix} -2 \\ 4 \\ 2 \end{bmatrix} + 3 \times \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

Question 1

Compute $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$ where

$$\mathbf{X} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}, \hat{\beta} = \begin{bmatrix} 0.7 \\ 0.2 \end{bmatrix}$$

And show that $\hat{\mathbf{y}}$ is a linear combination of the columns of $\mathbf{X}$.

Question 2

Now, let

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix}, \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$$

And once again,show that $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$ is a linear combination of the columns of $\mathbf{X}$.

Note, this can be even further generalised by allowing $\mathbf{X}$ to be $n \times k$ matrix and $\hat{\beta}$ tp be a $k \times 1$ column vector

The column space of X and the need for a geometric interpretation of OLS

- The point of the first two questions was to show that $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$ is a particular linear combination of the columns of $\mathbf{X}$
- It can also be interpreted as a weighted sum of the columns of $\mathbf{X}$ where the weights are given by β
- Knowing that the column space of $\mathbf{X}$ refers to all possible linear combinations of the columns of $\mathbf{X}$, we can say that $\hat{\mathbf{y}}$ exists in the column space of $\mathbf{X}$

The column space of X and the need for a geometric interpretation of OLS

- However it is usually not possible to arrange the columns of $\mathbf{X}$ in such a way to produce a $\hat{\mathbf{y}}$ that is exactly equal to $\mathbf{y}$
- This is because we have n equations but only k variables
- Hence, the possible combinations of $\mathbf{X}$ are restricted to a k dimensional subset of the full n dimensional space
- This results in some mismatch in $\mathbf{y}$ vs $\hat{\mathbf{y}}$ leaving us with some leftover or *resdiual*

Question 3

Consider a regression model with an intercept but no explanatory variables

$$y_i = \beta_0 + u_i, \quad i = 1, .., n$$

and which has an $\mathbf{X}$ matrix

$$\mathbf{X} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

.

Show that $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ is equivalent to the sample average of $\mathbf{y}$

Question 3 continued

1. What is $\hat{u}_i$?
2. Show that $\mathbf{X}'\hat{\mathbf{u}} = \sum_{i=1}^{n}(y_i - \bar{y}) = 0$
3. What is $\hat{\sigma}^2$, the sample variance of the residuals?

Question 4: Verification of question 3

Question 4a

- Obtain the histogram and descriptive statistics of IQ and Wage
- Graph IQ and Wage on a scatter plot, is the relationship linear?
- Run a regression of wage on a constant only. What is the coefficient estimate of the intercept? What is the standard error of regression?

Question 4b

- Estimate a simple linear regression model of Wage on a constant and IQ
- If IQ increases by 15 points, how much does the model predict Wage will increase by?
- What is the R^2 of the regression? What is the sample correlation between Wage and IQ?
- Is there a relationship between R^2 and the sample correlation?
- How can we interpret the intercept, is it meaningful?
- Save the equation as **eq01** and the residuals as **uhat01**
- https://flux.qa/LDPPHD

Question 4c

- Create a new variable IQ-100 and estimate a new regression of Wage on a constant and IQ-100
- Save the new equation and its residuals, are the residuals of the first equation different from the second?
- What is the interpretation of the intercept in the second equation, is it meaningful?
- What is the relationship between the coefficient estimates of the first and second equation?

Deriving the OLS estimator in summation form

- Find $\hat{\beta}$ which minimises the sum of squared residuals
- Residual:
$$\hat{u}_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$$
- Squared residuals:
$$\hat{u}_i^2 = (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$
- Sum of squared residuals:
$$\sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$
- Why do we minimise the sum of *squared* residuals and not just the sum of residuals?

Deriving the OLS estimator in matrix form

- There will be a mismatch between $\hat{\mathbf{y}}$ and $\mathbf{y}$ leaving us with some leftovers or *residuals*

$$\mathbf{y} = \hat{\mathbf{y}} + \hat{\mathbf{u}}$$

- OLS comes from minimising the length (or magnitude) of $\hat{\mathbf{u}}$
- This minimisation occurs when we choose β so that $\hat{\mathbf{u}}$ is orthogonal to the columns of $\mathbf{X}$ i.e. $\mathbf{X}'\hat{\mathbf{u}} = 0$