**Universidade de São Paulo / Faculdade de Filosofia, Letras e Ciências Humanas**

Departamento de Ciência Política

FLP-406 & FLS-6183

2º semestre / 2015

**Lab Class 4**

**In this lab, we will focus on understanding the basics of how matrix algebra and how the operations that are used to derive the estimates in a simple regression problem with one explanatory variable. We will first do the matrix operations by hand. Next, we will explore how to undertake the same operations with Stata, and explore additional insights that can be obtained after executing the same operations.**

**The example we will use is a small data set that examines the vote for the incumbent party in Brazilian presidential elections since 1985. The data set below contains information for the vote for the incumbent party in the current election, and rate of growth of GDP per capita in the same year.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Election | President | Incumbent Party | Per Capita GDP Growth | First Round Incumbent Vote (%) |
| 1985 | Sarney |  |  |  |
| 1989 | Collor (PRN) /Itamar Franco | PMDB | 1.341 | 4.7 |
| 1994 | FHC I | PRN or PSDB | 4.201 | 54.3 |
| 1998 | FHC II | PSDB | -1.431 | 53.1 |
| 2002 | Lula I | PSDB | 1.275 | 23.2 |
| 2006 | Lula II | PT | 2.804 | 48.61 |
| 2010 | Dilma I | PT | 6.537 | 46.91 |
| 2014 | Dilma II | PT | -0.352 | 41.59 |
| 2018 | Bolsonaro | PMDB | 0.328 | 1.20 |

Sources: GDP per capita data was retrieved from the World Bank’s World Development Indicators (2019). The data for election results were compiled from several sources.

**Part I. In class written exercise.**

**Exercise 1.** Based on Table 1, show how you would stack this regression equation into vector and matrix form. (Note: I want you to insert the particular observations provided in the table into vector and matrix form). Please be sure to specify the number of rows and columns for each vector and matrix. Hint: Remember the regression model equation we would like to estimate: . (where denotes the error term associated with observation i. These quantities define y, X, and u, as discussed in class.)

**Exercise 2.** Recall, the matrix X. What is its transpose X’? What are the dimensions of these two matrices?

**Exercise 3**. Find X’X. What are its dimensions?

**Exercise 4**. In general terms, we know that is defined as

**What will be the dimensions of** ?

**Exercise 5.** Find What are its dimensions?

**Exercise 6.** Obtain the OLS estimates of and using the formula:

**Exercise 7.** Confirm that these are the estimators for and by estimating the same regression in Stata using the Stata data file: brazil.dta. To do so, use the command regress y x. Can you summarize what the estimate effect is of per capita economic growth on the vote for the incumbent president?

**Exercise 8.** Now, using Matrix Algebra please find the residuals () obtained from the model. Please provide the formula. What are the dimensions of ?

**Exercise 9.** Now, let`s find the Residual Sums of Squares (RSS) using matrix algebra. What is the formula?

**Exercise 10.** Once the Residual Sums of Squares is obtained, RMS () can easily obtained because = . Calculate .

**Exercise 11.** In the Stata regression output from the regression you estimated earlier, do you find the same ?

**Exercise 12.** We can use the results obtained above to calculate the Variance-Covariance Matrix of Given what you know thus far, how many dimensions will this matrix have?

Gujarati summarizes this matrix:

**Part II. Stata Routine**

**Exercise 1**. Using Stata, find X’X.

**Exercise 2**. Using Stata, find .

**Exercise 3.** Using Stata, obtain

**Exercise 4.** In Stata, we can use the command invsym to calculate the inverse of a symmetric matrix.

Using Stata, obtain the OLS estimates of and using the formula:

.

**Exercise 5.** Going back to matrices and Stata, please find the residuals obtained from the model.

What are the dimensions of ?

**Exercise 6.** Now, let`s find the Residual Sums of Squares. Formally, Gujarti shows that . In this example, what will be the dimensions of ? Can you calculate ?

**Exercise 7.** We can use Stata to obtain the same Variance-Covariance Matrix of found in Exercise 12.

**Exercise 8.** Now, let`s use Stata to find this same matrix from a regression.

(Hint: You can obtain the same matrix after estimating a regression by issuing the command estat vce. Use this command to double-check your work in Exercise 12, Part I)

**Exercise 9.** Based on the results you found above, what do you conclude about economic voting?

**EXTRA CREDIT:**

**Exercise 10.** In the table below, we added the elections for the earlier “*Segunda República”* period. Based on this new data, what changes on your conclusions about economic voting?

**(FOR THIS ANSWER SEND A DO-FILE WITH YOUR WORK AND A .DOCX FILE WITH YOUR ANSWER. THERE IS A PLACE ON MOODLE FOR THIS).**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Election | President | Incumbent Party | Per Capita GDP Growth | First Round Incumbent Vote (%) |
| 1945 | Eurico Gaspar Dutra | \_ | \_ | \_ |
| 1950 | Getúlio Vargas | PSD | 0.751 | 20.5 |
| 1955 | Juscelino Kubitschek | PTB | 4.230 | 33.8 |
| 1960 | Jânio Quadros | PSD-PTB | 5.126 | 30.5 |
| 1985 | Sarney |  |  |  |
| 1989 | Collor (PRN) /Itamar Franco | PMDB | 1.341 | 4.7 |
| 1994 | FHC I | PRN or PSDB | 4.201 | 54.3 |
| 1998 | FHC II | PSDB | -1.431 | 53.1 |
| 2002 | Lula I | PSDB | 1.275 | 23.2 |
| 2006 | Lula II | PT | 2.804 | 48.61 |
| 2010 | Dilma I | PT | 6.537 | 46.91 |
| 2014 | Dilma II | PT | -0.352 | 41.59 |
| 2018 | Bolsonaro | PMDB | 0.328 | 1.20 |