

PLSC 503: “Multivariate Analysis for Political Research”

Exercise Two

February 3, 2021

Introduction

In this exercise, you’ll use linear/matrix algebra software – that is, the matrix algebra operators in R – to estimate a linear regression, and then cross-check those results using `lm`. The point is to demonstrate both a basic grasp of the software and your understanding of the matrix-algebra representation of the classical linear regression model. The data consist of the following:

| Observation | Y | X_1 | X_2 | X_3 |
|-------------|-----|-------|-------|-------|
| 1 | 39 | 0.3 | 8 | 0 |
| 2 | -18 | 0.1 | 111 | 1 |
| 3 | 13 | 0.7 | 79 | 1 |
| 4 | 19 | 1.1 | 13 | 1 |
| 5 | 6 | 0 | 91 | 0 |
| 6 | -2 | 0.1 | 43 | 1 |
| 7 | 15 | 1.5 | 52 | 0 |
| 8 | -11 | 0 | 98 | 1 |
| 9 | 17 | 0.9 | 22 | 0 |
| 10 | 3 | 0.2 | 106 | 1 |

Exercise

Using the linear algebra functions in R (or, e.g., the `mata/matrix` commands in Stata), do the following:

1. Estimate the $\hat{\beta}$ s and their variances and covariances $\widehat{\text{Var}}(\hat{\beta})$ for the OLS equation:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i.$$

2. Calculate the predicted \hat{Y}_i s, and the estimated residuals (the \hat{u}_i s).
3. Calculate the estimated root mean squared error (RMSE), the R^2 , and the R^2_{adj} .
4. Calculate the standard errors of the estimated $\hat{\beta}$ s, as well as their covariances.
5. Calculate a t -test for the hypothesis that $\beta_2 = 0$.
6. Calculate an F -test for the joint hypothesis that $\beta_1 = \beta_2 = \beta_3 = 0$.
7. Calculate an F -test for the hypothesis that $\beta_1 = \beta_3$.
8. Check all your results using `lm` / `linearHypothesis` / etc.

Be sure to include *all* code that you used to complete the various steps above.

This assignment is due no later than Friday, February 12, 2021 at 11:59 p.m. EDT. You can submit your homework by emailing copies **both** to Dr. Zorn (zorn@psu.edu) and Mr. Bolte (b1b72@psu.edu). In addition to your responses to the items above, please include all code used to fit models, conduct diagnostics, generate plots, and so forth. This assignment is worth 50 possible points.