Home-work # 5 (generating data with multi-collineartity and specification errors)

1. Generate data with collinearity problems (follow the definition of quasi-collinearity) in the following way:

i) Generate a matrix of regressors dimenssions (100, 5) including the column of ones. Designe some pattern of multicollinearity between X2 and X3. Make X4 not correlated with X2 and X3. Designe some pattern of collinearity for X5 with all of them.

ii) Consider two extrem cases, one, say case A, where the variance of the error term in the generation of the multicollinearity is very small, and another, say case B, where that variance is very big.

2. Generate the dependent variable following a model with the independent term, X2, X3 and X4. Consider three models:

- the true model.
- a model missing a rellevant regressor.
- a model including an irrellevant regressor, so X5.

Estimate all the models you need to answer part 3. Note that this is not a Monte Carlo study, you generate data with problems and you see what happens when you estimate.

3. Do from i) to v) for the well specified model and for cases A and B. And do parts vi) and vii) for the three models and for cases A and B.

i) Compare the values of the sum of the squares of the residuals when you drop 20 observations.

ii) Compare the values of the coefficients in i).

iii) Compare the values of the determinant of X'X in case A and B.

iv) Calculate the correlation matrix for case A and B.

v) Calculate the artificial regressions for case A and B.

vi) Do t tests on the true values of the parameters for ALL the cases.

vii) Do t tests of significance of the coefficients for ALL the cases.