HSE and University of London Double Degree Programme in Data Science and Business Analytics

Elements of Econometrics, 2023-2024

Classteacher: Ksenia Kasianova.

Class 4: Multiple linear regresion.

Problem 1

Regress $Y_i|1, X_i$.

(i) Using matrix notations and geometric intuition answer True or False:

(a) $\frac{1}{n}\sum_{i=1}^{n}\widehat{u}_i = 0$,

(b) $\frac{1}{n}\sum_{i=1}^{n}X_i\widehat{u}_i = 0$,

(c) $\frac{1}{n}\sum_{i=1}^{n}\widehat{Y}_i\widehat{u}_i = 0$,

(d) $\frac{1}{n}\sum_{i=1}^{n}\widehat{Y}_i = \frac{1}{n}\sum_{i=1}^{n}Y_i$,

(e) $TSS = ESS + RSS$.

(f) How the above analysis changes if you regress $Y_i|1$? $Y_i|X_i$?

(ii) Define R^2. How it is related to

(a) the residual sum of squares,

(b) the correlation between the actual and fitted values of the dependent variable, $r_{Y,\widehat{Y}}$.

(c) How would you measure goodness of fit, if you had to choose among RSS, R^2 and $r_{Y,\widehat{Y}}$? Why?

(iii) Derive regression coefficients in a simple regression model using matrix notation

(a) on constant (naive model)

(b) without intercept

What properties of linear regression are violated for the regression without intercept?

Problem 2

(a) Recall GMT in matrix notation

(b) Derive variance of $\hat{\beta}$ in matrix form

(c) Find variance-covariance matrix for a pair linear regression

(d) Consider a formula for variance of $\hat{\beta}_j$:

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{TSS_j(1 - R_j^2)}$$

What factors lead to the inflation of s.e. of the estimator of the coefficients?