Laboratory 2

Least Squares.

Exercise 1 - LS estimator of a linear model

Use the data file datalab2 from the web site. The first column of the data set is the dependent variable Y, whereas the remaining columns describe the explanatory variables X. Let's assume that

$$y_n = x_n\alpha + u_n.$$

Suppose, the loss function is

$$L = \sum_{n=1}^{N} u_n^2.$$

- Express the loss function, L as a function of Y, X and α.
- Derive the first derivative of the function L with respect to the vector of parameters α as a function of Y, X and α.
- Since the estimator minimize L, write down the relevant FOC. What is the LS estimator of the model parameters?
- Estimate the variance of the residuals u_n.
- If the residuals are uncorrelated and homoscedastic, what is the variance-covariance matrix of the estimator $\sqrt{N}\hat{\alpha}$, where N is the sample size?
- Compute the t-rations of the parameters.

Exercise 2 - a linear model with restrictions

Use the data and model specification from Exercise 1. Suppose, you know that $\alpha_1+\alpha_2+\alpha_3 = 0$ and $\alpha_2 - \alpha_3 = 0$.

- Write down the restriction matrix. What is the rank of this matrix?
- Express the vector α and the loss function L as functions of α_3.
- Estimate α_3 with the LS method. Compute its variance-covariance matrix.
- What is the estimator of α? Compute its variance-covariance matrix and t-rations.

Exercise 3 - partial LS

Lets consider a model

$$y_n = \alpha_1 x_{1n} + \alpha_2 x_{2n} + \varepsilon_n,$$

with $\alpha_1 = \alpha_2 = 1$ and $\varepsilon_n \sim N(0, 1)$.

- Generate a sample of Y for (a) X_1 and X_2 independent: $X \sim N(0, I_2)$, (b) for X_1 and X_2 dependent: $X \sim N(0, \Sigma)$
- Estimate the parameter α_1 in the true model: $y_n = \alpha_1 x_{1n} + \alpha_2 x_{2n} + \varepsilon_n$ and a corresponding parameter β in a miss-specified model: $y_n = \beta x_{1n} + u_n$.
- Compare the true value of α_1 with its estimators: $\hat{\alpha_1}$ and $\hat{\beta}$ for two cases (a) and (b) and different sample sizes: $N = 10, 100$ and 1000.
- What can you say about consistency of the estimators? Why in one case the estimator $\hat{\beta}$ is consistent and in other case not? Comment.