ECON220B Discussion Section 4

Midterm Review

Lapo Bini

Roadmap

1. Question 1: Algebraic Properties of OLS
2. Question 2: Properties Empirical CDF Estimator

Exercise 1

Consider the following linear model:

$$y_i = \alpha + \mathbf{x}_i^T \beta + u_i$$

where $y_i \in \mathbb{R}$, $\mathbf{x}_i \in \mathbb{R}^d$ and $u_i \in \mathbb{R}$. Note that the intercept is captured by α and it is not included in $\mathbf{x}_i$. Suppose that we have an iid sample $(y_i, \mathbf{x}_i)$ for $i = 1, \dots, n$.

We will assume $E[u_i] = 0$ and $E[\mathbf{x}_i u_i] = 0$.

Exercise 1 - Question 1

Write down the sample moment conditions for $\hat{\alpha}$ and $\hat{\beta}$

Exercise 1 - Question 2

Let $\hat{u}_i$ be the regression residual, write down its expression in terms of y_i, $\mathbf{x}_i$ and the estimated coefficient

Exercise 1 - Question 3

Now regress $\hat{u}_i$ on an intercept and $\mathbf{x}_i$. Find the estimated coefficients.

Exercise 1 - Question 4

Let $\bar{x}$ be the sample mean of the regressors, and define $\check{\mathbf{x}}_i \equiv \mathbf{x}_i - \bar{x}$. Find the estimated coefficient. Now regress y_i on an intercept and $\check{\mathbf{x}}_i$. Find the estimated coefficients $\check{\alpha}$ and $\check{\beta}$. How are they related to the estimates $\hat{\alpha}$ and $\hat{\beta}$?

Exercise 1 - Question 5

Claim: $\breve{\alpha} = \frac{1}{n}\sum_{i=1}^{n} y_i$. True or false?

Exercise 2

Assume $\{x_i\}_{i=1}^n$ iid sample from a univariate distribution, $x_i \in \mathbb{R}$. Denote by $F(\cdot)$ the cumulative distribution which is defined as:

$$F(x) = \mathbb{P}(x_i \leq x)$$

For simplicity we will assume that x_i is continuously distributed on the unit interval such that:

- $F(x) = 0$ for all $x \leq 0$.
- $F(x) = 1$ for all $x \geq 1$.
- $F(x)$ continuous and strictly increasing for all $x \in (0, 1)$.

Define the empirical CDF as:

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{x_i \leq x\}$$

Exercise 2 - Question 1

Assume $0 < x < 1$, find the asymptotic distribution of $\hat{F}(x)$

Exercise 2 - Question 2

Assume $0 < x \neq x' < 1$, find the asymptotic distribution of $\hat{F}(x)$ and $\hat{F}(x')$

Exercise 2 - Question 3

Consider the hypothesis $\mathsf{H}_0 : F(x) = G(x)$ and define the following statistic:

$$KS = \sup_{x \in [0,1]} \sqrt{n} |\hat{F}(x) - G(x)|$$

Show that under the null, the KS statistic can be rewritten as:

$$KS = \sup_{x \in [0,1]} \sqrt{n} \left| \frac{1}{n} \sum_{i=1}^{n} \big(\mathbb{1}\{F(x_i) \leq x\} - x \big) \right|$$

Exercise 2 - Question 4

Consider the hypothesis $\mathsf{H}_0 : F(x) = G(x)$ and show that the KS statistic does not depend on the underlying distribution $F(\cdot)$.

Exercise 2 - Question 5

Discuss how you can modify the KS statistic to test FOSD.