Econ 704 Discussion Section-Handout 8

Emilio Cuilty,

Handout desingned by: Jian Zhang

cardenascuil@wisc.edu . Office: 7143, Office Hours: Thursdays 9:00-11:00 A.M.

---

1. Review:

1.1. OLS estimation

- The Model: Suppose that we observe a random sample $\{y_i, x_i\}$ (y_i is a scalar and x_i is a $k \times 1$ vector). We usually establish the linear model:

$$\begin{aligned} y_i &= x_i'\beta + \epsilon_i \\ &= \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \epsilon_i \end{aligned}$$

  Here, β is a $k \times 1$ vector that represents partial effects of x_i on y_i. y_i is called dependent variable and x_i are called independent variables.

- Assumptions:

  (OLS 0) $\{(y_i, x_i')\}$ is an i.i.d. random sample.
  (OLS 1) $E(x_i x_i')$ ($k \times k$ matrix) is finite and nonsingular. ("rank condition")
  (OLS 2) $E(\epsilon_i | x_i) = 0$. ("mean independent")
  (OLS 2') $E(x_i \epsilon_i) = 0$.
  (OLS 3) $E(\epsilon_i^2 | x_i) = \sigma^2$ is finite. ("Homoskedasticity")

- Identification of β:
  1. $\beta = (E x_i x_i')^{-1} E x_i y_i$.
  2. $b = \beta$ minimizes $E(y_i - x_i' b)^2$.
  3. $b = \beta$ minimizes $E[\big(E(y_i|x_i) - x_i' b\big)^2]$.

- OLS Estimator of β:

$$\hat{\beta} = \left(\frac{1}{n} \sum_{i=1}^{n} x_i x_i' \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} x_i y_i \right)$$

  or

$$b = \hat{\beta} \text{ minimizes } \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i' b)^2$$

- Consistency of $\hat{\beta}$: Under assumptions OLS 0–1–2', then $\hat{\beta} \xrightarrow{p} \beta$.

- Consistent estimator of σ^2:

  – Consistent but biased: $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \hat{\epsilon}_i^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i' \hat{\beta})^2$;
  – Consistent and unbiased: $s^2 = \frac{1}{n-k} \sum_{i=1}^{n} \hat{\epsilon}_i^2 = \frac{1}{n-k} \sum_{i=1}^{n} (y_i - x_i' \hat{\beta})^2$.

1.2. Theorems that you may need to prove consistency and asymptotic normality

1. In Weak Law of Large Number (WLLN), we allow random variables to be vector or matrix of the same dimension. In Central Limit Theorem (CLT), we allow random variables to be vector of the same dimension, then their convergent distribution is a multivariate normal distribution.
2. Suppose X and Y are two random variables, and $g(\cdot)$ is a function, then $E[g(X)Y|X] = g(X)E[Y|X]$.
3. (Continuous Mapping theorem) If $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y$, $g(\cdot,\cdot)$ is a continuous function (map), then $g(X_n, Y_n) \xrightarrow{p} g(X, Y)$.
   Example: If $n^{-1}\sum_{i=1}^n x_i x_i' \xrightarrow{p} E(x_i x_i')$ is invertible and $n^{-1}\sum_{i=1}^n x_i y_i \xrightarrow{p} E(x_i y_i)$, $g(A, B) = A^{-1}B$ is continuous when A is invertible, then
   $$\hat\beta = \big(n^{-1}\sum_{i=1}^n x_i x_i'\big)^{-1}\big(n^{-1}\sum_{i=1}^n x_i y_i\big) \xrightarrow{p} (Ex_i x_i')^{-1}Ex_i y_i = \beta.$$
4. (Generalized Slutsky Theorem) If $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{p} c$, c is nonrandom and $g(\cdot,\cdot)$ is continuous, then $g(X_n, Y_n) \xrightarrow{d} g(X, c)$.
   Example: If $n^{-1}\sum_{i=1}^n x_i x_i' \xrightarrow{p} E(x_i x_i')$ and $n^{-1/2}\sum_{i=1}^n x_i\epsilon_i \xrightarrow{d} Z \sim \mathcal{N}(0, \sigma^2 E(x_i x_i'))$, then
   $$\sqrt{n}(\hat\beta - \beta) = \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n x_i\epsilon_i\right) \xrightarrow{d} [E(x_i x_i')]^{-1}Z \sim \mathcal{N}(0, \sigma^2[E(x_i x_i')]^{-1})$$

2. Problems

1. Suppose that x_i, y_i and ϵ_i are random sample, show that
   (a) If ϵ_i and x_i are independent, and $E(\epsilon_i) = 0$, then OLS 2 holds, i.e., $E(\epsilon_i|x_i) = 0$;
   Since ϵ_i and x_i are independent, then $E(\epsilon_i|x_i) = E(\epsilon_i) = 0$
   (b) If OLS 2 holds, then OLS 2' also holds, i.e., $E(x_i\epsilon_i) = 0$.
   $E(x_i\epsilon_i) = E[E(x_i\epsilon_i|x_i)] = E[x_i E(\epsilon_i|x_i)] = E[x_i \cdot 0] = 0$.
   (c) If OLS 2 holds and $y_i = x_i'\beta + \epsilon_i$, then $E(y_i|x_i) = x_i'\beta$.
   $E(y_i|x_i) = E(x_i'\beta + \epsilon_i|x_i) = x_i'\beta + E(\epsilon_i|x_i) = x_i'\beta$
   (d) If OLS 0–1–2 holds and $y_i = x_i'\beta + \epsilon_i$, then OLS estimator $\hat\beta$ is unbiased.

$$\begin{aligned}\hat\beta - \beta &= \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i y_i\right) - \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\beta\right)\\ &= \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i(y_i - x_i'\beta\right)\\ &= \left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i\epsilon_i\right) = \sum_{i=1}^n f_i(x)\epsilon_i.\end{aligned}$$

Here, $f_i(x) = n^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1} x_i$ and x is the collection of all x_i. Note that $E(\epsilon_i|x) = E(\epsilon_i|x_i) = 0$. Therefore,

$$E(\hat\beta - \beta|x) = \sum_{i=1}^n f_i(x)E(\epsilon_i|x) = 0$$

By law of iterated expectations,

$$E(\hat\beta) - \beta = E[E(\hat\beta - \beta|x)] = 0.$$

(e) Show that $\sum_{i=1}^n x_i\hat{y}_i = \sum_{i=1}^n x_i y_i$ where $\hat{y}_i = x_i'\hat{\beta}$.
Note that

$$\frac{1}{n}\sum_{i=1}^n x_i\hat{y}_i = \frac{1}{n}\sum_{i=1}^n x_i x_i'\hat{\beta} = \frac{1}{n}\sum_{i=1}^n x_i x_i'\left(\frac{1}{n}\sum_{i=1}^n x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n x_i y_i\right) = \frac{1}{n}\sum_{i=1}^n x_i y_i.$$

2. Suppose that (y_i, x_i) is a random sample and y_i, x_i are scalars. We are interested in a regression model $y_i = \alpha + \beta x_i + \epsilon_i$. Derive OLS estimators of α and β.
Denote $w_i = (1, x_i)'$ and $\hat{\gamma} = (\hat{\alpha}, \hat{\beta})'$ the OLS estimator of $\gamma = (\alpha, \beta)'$. Note that

$$\frac{1}{n}\sum_{i=1}^n w_i w_i' = \frac{1}{n}\begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix} \text{ and } \frac{1}{n}\sum_{i=1}^n w_i y_i = \frac{1}{n}\begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix}$$

Therefore,

$$\hat{\gamma} = \left(\frac{1}{n}\sum_{i=1}^n w_i w_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n w_i y_i\right) = \begin{pmatrix} \frac{\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n x_i y_i}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \\ \frac{\sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i - n\sum_{i=1}^n x_i y_i}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \end{pmatrix}$$

3. Suppose that (y_i, d_i) is a random sample and y_i is a scalar, d_i can only be 0 or 1. Consider a regression model $y_i = \beta d_i + \epsilon_i$.
   (a) Show that OLS 1 holds if and only if $\Pr(d_1 = 1) > 0$.
   The OLS 1 is that $E(d_1 d_1')$ finite and nonsingular. Note that $E(d_i d_i') = E(d_i^2) = E(d_i) = \Pr(d_i = 1)$. Therefore, OLS 1 is equivalent to

$$\Pr(d_i = 1) \text{ invertible.}$$

   That is equivalent to $\Pr(d_i = 1) > 0$.
   (b) Under OLS 0–1–2', show that $\beta = E(y_i d_i)/\Pr(d_i = 1)$.
   Note that

$$E[y_i d_i] = E[\beta d_i + d_i\epsilon_i] = \beta E[d_i] + 0 = \beta\Pr(d_i = 1).$$

   (c) Under OLS 0–1–2, show that $\beta = E(y_i | d_i = 1)$.
   Note that $E[y_i|d_i = 1] = \beta E[d_i|d_i = 1] + E[\epsilon_i|d_i] = \beta \cdot 1 + 0 = \beta$.
   (d) Derive OLS estimator $\hat{\beta}$ of β.
   The OLS estimator of β is given by

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^n d_i^2\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n y_i d_i\right) = \frac{\sum_{i=1}^n y_i d_i}{\sum_{i=1}^n d_i} = \frac{1}{\#\{d_i = 1\}}\sum_{i:d_i=1} y_i.$$

   Here , $\#\{d_i = 1\}$ represents the number of d_i that has value 1
4. (2017 Spring, ECON 705 Exam) Suppose (y_i, x_i) is an i.i.d. sequence, where x_i is a scalar. Suppose $E(y_i|x_i) = 2 + 3x_i$. Let $\mu_k = E[x_i^k]$ (use this notation to simply expressions). Be sure to clearly define any new variables that you introduce and state any additional assumptions you use.
   (a) Consider the OLS coefficient estimator where y_i is the dependent variable and $(1, x_i)'$ are the independent variables (so a constant term is included). Let $\hat{\alpha}$ denote the resulting 2×1 estimator. Let $\hat{\alpha} \xrightarrow{p} \alpha$. What is α?
   Note that OLS 0 holds and $E(\epsilon_i|x_i) = E(y_i - 2 - 3x_i|x_i) = 0$, that is, OLS 2 holds. If OLS 3 is also true, that is, $det(E(1, x_i)'(1, x_i)) = Ex_i^2 - (Ex_i)^2 = Var(x_i) > 0$, i.e. $\mu_2 - \mu_1^2 > 0$, then we know that this OLS estimator is consistent. Therefore, $\alpha = (2, 3)'$ if $\mu_2 - \mu_1^2 > 0$.
   (b) Consider the scalar OLS coefficient estimator where y_i is the independent variable and x_i is the independent variable (no constant term). Let $\hat{\beta}$ denote the resulting scalar estimator. Let $\hat{\beta} \xrightarrow{p} \beta$. What is β?

Note that

$$\hat{\beta} \xrightarrow{p} \frac{E(x_i y_i)}{E(x_i^2)} = \frac{2E(x_i) + 3E(x_i^2)}{E(x_i^2)} = \frac{2\mu_1 + 3\mu_2}{\mu_2}.$$

5. (Modified from Jangsu's Discussion 2 Exercise 6, Spring 2017 ECON 705) Consider a regression model $y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$ where y_i, x_{1i} and x_{2i} are scalars. Under assumptions OLS 0–1–2', consider a scalar OLS estimator where x_{1i} is the only independent variable, denoted by $\hat{\beta}_1$. Find the conditions that is needed for $\hat{\beta}_1 \xrightarrow{p} \beta_1$.

   Note that

$$\hat{\beta}_1 \xrightarrow{p} \frac{E(x_{1i} y_i)}{E(x_{1i}^2)} = \frac{\beta_1 E(x_{1i}^2) + \beta_2 E(x_{1i} x_{2i})}{E(x_{1i}^2)} = \beta_1 + \beta_2 \frac{E(x_{1i} x_{2i})}{E(x_{1i}^2)}.$$

   Therefore, $\hat{\beta}_1$ is consistent if $\beta_2 = 0$ or $E(x_{1i} x_{2i}) = 0$.