Econometrics 2 (2018) TA session 9*

Kenji Hatakenaka †

6 December 2018

目次

1 Method of Moments(積率法)

In the method of moment, we estimate parameters by using the moment conditions. The generalized method of moments (GMM) estimator is robust to some variations in the underlying data generating process. However, in most cases, method of moments are not efficient. The exception is in random sampling from exponential families of distributions*.

* All comments welcome!
† E-mail: u626530i@ecs.osaka-u.ac.jp, Room 501
* See Greene.

1.1 Review

As $n \to \infty$, we have the result:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow E[X] = \mu,$$

where $X_1, X_2, \ldots X_n$ are n realizations(実現値) of X.

Review: Chebyshev's inequality (チェビシェフの不等式)

Chebyshev's inequality (チェビシェフの不等式) is given by

$$P[|X-\mu| > \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2} \quad \text{or} \quad P[|X-\mu| \leq \varepsilon] > 1 - \frac{\sigma^2}{\varepsilon^2}$$

where $\mu = E[X]$, $\sigma^2 = V(X)$ and any $\varepsilon > 0$.

Proof. The following Markov's Inequality gives the straight forward proof. By putting $u(X) = (X-\mu)^2$ and $c = k^2\sigma^2$. we obtain

$$P[(X-\mu)^2 \geq k^2\sigma^2] \leq \frac{E[(X-\mu)^2]}{k^2\sigma^2} = \frac{1}{k^2}.$$

Then we can rewrite the inequality as follows:

$$P[|(X-\mu)| \geq k\sigma] \leq \frac{1}{k^2}.$$

Putting $k = \varepsilon/\sigma$, we obtain the inequality.

Review: Markov's inequality (マルコフの不等式)

Let $u(X)$ be a nonnegative function of the random variable X. If $E[u(X)]$ exists, then for every positive constant c,

$$P[u(X) \geq c] \leq \frac{E[u(X)]}{c}.$$

Proof. We consider the case when the random variable is continuous. Let $A = \{x : u(x) \geq c\}$ and let $f(x)$ denote the probability distribution function of X. Then we have

$$\begin{aligned} E[u(X)] &= \int_{-\infty}^{\infty} u(x)f(x)\mathrm{d}x = \int_A u(x)f(x)\mathrm{d}x + \int_{A^c} u(x)f(x)\mathrm{d}x \\ &\geq \int_A u(x)f(x)\mathrm{d}x \geq \int_A cf(x)\mathrm{d}x = cP[X \in A] = cP[u(X) \geq c]. \end{aligned}$$

Rearranging this, we have the inequality.

Replace X, $E[X]$ and $V(X)$ by $\overline{X}$, $E[\overline{X}] = \mu$ and $V[\overline{X}] = \sigma^2/n$ in the Chebyshev's Inequality. Then, as $n \to \infty$, we obtain

$$P[|\overline{X} - \mu| \leq \varepsilon] \geq 1 - \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 1.$$

This implies that $\overline{X} \to \mu$ as $n \to \infty$.

$\overline{X}$ is an approximation of $E[X] = \mu$. Therefore, $\overline{X} = \sum_{i=1}^n X_i/n$ is taken as an estimator of μ.

2 Application: Regression model

We consider an application of the MM method to regression model. Consider the regression model:

$$y_i = x_i\beta + u_i,$$

We place familiar assumption: $E[u|x] = 0 (E[x'u] = 0)$, where x is a $1 \times k$ vector and u is a scalar. This is called the **orthogonality condition (直交条件)**.

From the law of large numbers, we have the following equality:

$$\frac{1}{n}X'u = \frac{1}{n}\sum_{i=1}^n \underbrace{x_i'}_{k\times 1}\underbrace{u_i}_{1\times 1} = \frac{1}{n}\sum_{i=1}^n x_i'(y_i - x_i\beta) \quad \longrightarrow \quad E[x'u] = 0.$$

Thus, the MM estimator of β, denoted by β_{MM}, satisfies:

$$\frac{1}{n}\sum_{i=1}^{n} x_i'(y_i - x_i\beta) = 0.$$

Rearranging this equation, we obtain β_{MM} as follows:

$$\beta_{MM} = \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i'x_i\right)^{-1}}_{k\times k}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i'y_i\right)}_{k\times 1} = (X'X)^{-1}X'y$$

where $X' = (x_1', \cdots, x_n')$ and $y = (y_1, \cdots, y_n)'$.

This formula implies that β_{MM} is equivalent to OLSE.

3 Instrumental Variables

Note that $\hat{\beta} = (X'X)^{-1}X'y$ is inconsistent when $E[x'u] \neq 0$, i.e.,

$$\hat{\beta} = (X'X)^{-1}X'y = \beta + (X'X)^{-1}Xu = \beta + \left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'u\right) \not\to \beta,$$

where

$$\frac{1}{n}X'u = \frac{1}{n}\sum_{i=1}^{n} x_i'u_i \longrightarrow E[x'u] \neq 0.$$

In order to obtain a consistent estimator of β, we find an instrumental variable z which satisfies $E[z'u] = 0$. Let z_i be the ith realization of z, where z_i is a $1 \times k$ vector[†]. Letting $Z' = (z_1', \cdots, z_n')$, we have the following formula:

$$\frac{1}{n}Z'u = \frac{1}{n}\sum_{i=1}^{n} z_i'u_i \longrightarrow E[z'u] = 0.$$

We denote β_{IV} which satisfies:

$$\frac{1}{n}\sum_{i=1}^{n} z_i'(y_i - x_i\beta_{IV}) = 0.$$

[†] Important assumption!

Rearranging this equation, we obtain

$$\beta_{IV} = \left(\frac{1}{n}\sum_{i=1}^{n} z_i' x_i\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} z_i' y_i\right) = (Z'X)^{-1}Z'y.$$

Note that $Z'X$ is a $k \times k$ square matrix, where we assume that the inverse matrix of $Z'X$ exists. This implies that the matrix has rank k. Assume that, as $n \to \infty$, there exist the following moment matrices:

$$\frac{1}{n}\sum_{i=1}^{n} z_i' x_i = \frac{1}{n}Z'X \to M_{zx}, \quad \frac{1}{n}\sum_{i=1}^{n} z_i' z_i \to M_{zz}.$$

As n goes to infinity, β_{IV} is rewritten as:

$$\begin{aligned}\beta_{IV} &= (Z'X)^{-1}Z'y = (Z'X)^{-1}Z'(X\beta + u) = \beta + (Z'X)^{-1}Z'u \\ &= \beta + \left(\frac{1}{n}Z'X\right)^{-1}\left(\frac{1}{n}Z'u\right) \quad \longrightarrow \quad \beta + M_{zx} \times 0 = \beta.\end{aligned}$$

Thus, β_{IV} is a consistent estimator of β.

3.1 Asymptotic distribution

Now, we consider the asymptotic distribution of β_{IV}. By the Central Limit Theorem,

$$\sqrt{n}\left(\frac{1}{n}Z'u\right) = \frac{1}{\sqrt{n}}Z'u \longrightarrow N(0, \sigma^2 M_{zz})$$

Note that

$$\begin{aligned}V\left[\frac{1}{n}Z'u\right] &= nV(Z'u) = \frac{1}{n}E[Z'uu'Z] = \frac{1}{n}E[E[Z'uuZ \mid Z]] \\ &= \frac{1}{n}E[Z'E[uu' \mid Z]Z] = \frac{1}{n}E[\sigma^2 Z'Z] = E[\sigma^2\frac{1}{n}Z'Z] \longrightarrow E[\sigma^2 M_{zz}] = \sigma M_{zz}.\end{aligned}$$

We obtain the following asymptotic distribution:

$$\sqrt{n}(\beta_{IV} - \beta) = \left(\frac{1}{n}Z'X\right)^{-1}\left(\frac{1}{\sqrt{n}}Z'u\right) \longrightarrow N(0, \sigma^2 M_{zx}^{-1} M_{zz} M_{zx}^{-1}).$$

Practically, for large n, we use the following distribution:

$$\beta_{IV} \sim N(\beta, s^2(Z'X)^{-1}Z'Z(Z'X)^{-1}), \quad \text{where} \quad s^2 = \frac{1}{n-k}(y - X\beta_{IV})'(y - X\beta_{IV}).$$

In the case where z_i is a $1 \times r$ vector for $r > k$, $Z'X$ is a $r \times k$ matrix, which is not a square matrix. This implies that the matrix has no inverse matrix. In order to apply this method to $r > k$ case, we consider **Generalized Method of Moments (GMM, 一般化積率法)** in the next lecture.

4 Example

I will introduce an example of GMM estimation (Hall(1978)).The original form of the model is derived by optimization problem below.

$$\text{Maximize } E_t\left[\sum_{\tau=0}^{T-t}\left(\frac{1}{1+\delta}\right)^{\tau} U(c_{t+\tau})\right]$$
$$\text{subject to } \sum_{\tau=0}^{T-t}\left(\frac{1}{1+r}\right)^{\tau}(c_{t+1\tau} - w_{t+\tau}) = A_t$$

The solution to the optimization problem is

$$E_t[U'(c_{t+1})|\Omega_t] = \frac{1+\delta}{1+r}U'(c_t).$$

When $U(c_t) = \frac{c_t^{1-\alpha}-1}{1-\alpha}$ (power utility),

$$E_t\left[(1+r)\left(\frac{1}{1+\delta}\right)\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha} - 1|\Omega_t\right] = E_t\left[\beta(1+r)R_{t+1}^{\lambda} - 1|\Omega_t\right] = 0$$

Where $R_{t+1} = c_{t+1}/c_t$ and $\lambda = -\alpha$. Assume that r is not constant, the estimator can be estimated by using

$$E_t\left[\underbrace{\begin{pmatrix} 1 \\ R_t \end{pmatrix}}_{\mathbf{z}'}\underbrace{(\beta(1+r_{t+1})R_{t+1}^{\lambda} - 1)}_{\mathbf{u}}\right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$