TWIERDZENIE GAUSSA MARKOVA

EKONOMETRIA WNE 2022/23

Sebastian Zalas

8 grudnia 2022

WPROWADZENIE

- Dotychczas uzyskiwaliśmy oszacowania parametrów metodą Najmniejszych Kwadratów
- Dziś przeanalizujemy własności estymatora
  - założenia i ich implikacje oraz sposoby niespełnienia założeń
  - Tw. Gaussa-Markova - jest to jedno z najważniejszych twierdzeń w teorii ekonometrii. Zapewnia ono klasyczne uzasadnienie dla estymatora Najmniejszych Kwadratów pokazując, że ma on najniższą wariancję pośród wszystkich nieobciążonych estymatorów.

ZAŁOŻENIA

Założenie 1

Próba jest losowa

$$(x_{i,1}, x_{i,2}, \ldots, x_{i,k}, y_i) \sim i.i.d. \ \forall_{i=1,\ldots,n}$$

- To założenie oznacza że dla dwóch wybranych jednostek i oraz j ($i \neq j$) z próby, wartości $(x_{i,1}, \ldots, x_{i,k}, y_i)$ są niezależne od wartości $(x_{j,1}, \ldots, x_{j,k}, y_j)$ lecz mają one ten sam rozkład. Niezależność oznacza że decyzje jednostki i nie wpływają na decyzje jednostki j oraz odwrotnie.
- Np.: dane panelowe nie spełniają tego założenia.

Założenie 2 - Brak dokładnej współliniowości

W próbie (i co za tym idzie także w populacji), żadna ze zmiennych objaśniających nie jest stała oraz nie ma *dokładnie* liniowej zależności między zmiennymi objaśniającymi.

- W przypadku dokładnej współliniowości, macierz $\boldsymbol{X}'\boldsymbol{X}$ nie jest odwracalna, przez co nie można uzyskać oszacowań MNK.
- To założenie nie wyklucza istnienia pewnej korelacji między zmiennymi objaśniającymi, natomiast nie mogą one być dokładnie skorelowane.

ZAŁOŻENIA

Założenie 3

Warunkowa wartość oczekiwana błędu losowego jest równa zero:

$$\mathbb{E}[u \mid x_1, x_2, \ldots, x_k] = 0$$

- To założenie wyklucza istnienie korelacji między zmiennymi objaśniającymi a składnikiem losowym.

To założenie będzie złamane gdy:

- zostanie przyjęta niewłaściwa forma funkcyjna modelu
- zostanie pominięta zmienna która objaśniająca y
- wystąpi błąd pomiaru

NIEOBCIĄŻONOŚĆ

Twierdzenie 1 (Nieobciążoność MNK)

Jeśli założenia 1-3 są spełnione, to zachodzi:

$$\mathbb{E}[u \mid x_1, x_2, \ldots, x_k] = 0 \ \forall_{j=1,\ldots,k}$$

- Wektor $\hat{\beta}$ jest nieobciążonym estymatorem wektora prawdziwych parametrów modelu β

Znaczenie nieobciążoności:

- Konkretne oszacowanie nie może być nieobciążone, ponieważ zostało ono uzyskane z danej próby, i zwykle nie jest równe prawdziwemu parametrowi.
- Nieobciążoność estymatora w tym przypadku oznacza, że procedura MNK dzięki której uzyskano $\hat{\beta}$ jest nieobciążona gdy jest zastosowana do wszystkich możliwych losowych prób z populacji.

Założenie 4 (Sferyczność wariancji)

Wariancja błędu losowego spełnia

$$\mathrm{Var}[\varepsilon \mid \boldsymbol{x}] = \sigma^2 \boldsymbol{I}$$

- Innymi słowy, składnik losowy ma taką samą wariancję przy przy danych wartościach zmiennych objaśniających.
- Składniki losowe związane z poszczególnymi obserwacjami nie są ze sobą skorelowane i mają stałą wariancję.

Założenie 5

Składnik losowy ma rozkład normalny:

$$\varepsilon \sim N(0, \sigma^2 \boldsymbol{I})$$

- Założenie to nie jest potrzebne do sformułowania tw. Gaussa-Markowa, ale ułatwia weryfikację modelu

TWIERDZENIE GAUSSA-MARKOWA

Twierdzenie Gaussa-Markowa

Przy spełnionych założeniach 1-4, estymator $\hat{\beta}$ uzyskany Metodą Najmniejszych Kwadratów jest estymatorem BLUE [best linear unbiased estimator], tj. zgodnym, nieobciążonym i najefektywniejszy w klasie liniowych estymatorów wektora β.

- liniowy - estymator (MNK) jest liniowy wtedy, gdy może być wyrażony jako liniowa funkcja danych i zm. zależnej.
- najlepszy - estymator MNK ma najmniejszą wariancję spośród wszystkich liniowych estymatorów.

Pytania? Wątpliwości?

Dziękuję!

e: s.zalas@uw.edu.pl