Econometrics with Python

Olivér Nagy

John von Neumann University - MNB Institute
Central Bank of Hungary

2023. október 13.

MNB INTÉZET
FENNTARTHATÓ PÉNZÜGYEK KÖZPONT

Review: alapszakon / kvantitatív alapokon elsajátított ismeret

- Többváltozós lineáris regressziós modell (MLR) definíciója:

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki} + \epsilon_i \tag{1}$$

  ahol $i = 1, ..., n$
- A model y változó alakulásást próbálja leírni $x_1, ..., x_k$ függvényében
- y_i függő változó (hivatkozunk rá LHS, vagy outcome változóként)
- $x_1, ..., x_k$ független változók (RHS, prediktor változók)
- ϵ_i a hibatag

Review: OLS elvű becslés

- A hagyományos legkisebb négyzetek módszere (OLS) egy becslési elv, ami lehetővé teszi lineáris regressziós modellek modell paramétereinek ($\hat{\beta}$) becslését, a négyzetes távolságok minimalizálásán keresztül:

$$\hat{\beta} = \arg\min_{\beta \in \Theta} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - ... - \beta_k x_{ki})^2$$

- Legyen $Y = [y_1, ..., y_n]^T, \beta = [\beta_0, \beta_1, ..., \beta_n]^T, \epsilon = [\epsilon_1, ..., \epsilon_n]^T$, és

$$X = \begin{bmatrix} 1 & x_{11} & ... & x_{1k} \\ ... & & & \\ 1 & x_{n1} & ... & x_n k \end{bmatrix}.$$

- Ebben az esetben (1) egyenlet mátrixos formában átírható a következő módon:

$$Y = X\beta + \epsilon \tag{2}$$

Review: Lineáris regresszió OLS elvű becslőjének levezetése

- Az OLS elvű becslő a maradéktagok ($residuals$) négyzetének az összegét (SSR/RSS) minimalizálja, amit mátrixos formában a következő módon tudunk felírni:

$$SSR = \epsilon^T \epsilon = (Y - X\beta)^T (Y - X\beta).$$

- SSR konkáv β függvényében, ezért az elsőrendű feltétel (FOC) elégséges az egyenlet minimalizálásához.
- Alkalmazva a következő azonosságot $(Y^T X\beta)^T = \beta^T X^T Y$ azt kapjuk, hogy

$$\epsilon^T \epsilon = Y^T Y - 2\beta^T X^T Y + \beta^T X^T X\beta.$$

- Az FOC alapján

$$\frac{\partial \epsilon^T \epsilon}{\partial \beta} = -2X^T Y + 2X^T X\hat{\beta} = 0$$

$$\implies \hat{\beta}_{OLS} = (X^T X)^{-1} X^T Y$$

Gauss-Markov feltevések

Ahhoz, hogy az OLS elven becsült lineáris regressziónak fennálljanak bizonyos előnyös tulajdonságai, meghatározott feltevéseknek teljesülniük kell:

1. **Modell linearitás**: Feltesszük, hogy az Y és az X közötti kapcsolat lineárisan leírható
2. Az adatok a sokaság **véletlen mintái**. A legtöbbször használt i.i.d. ennél szigorúbb, tehát teljesíti a véletlen mintás feltevést.
3. X mátrix **teljes oszlop rangú**. Ezen feltevés alpján tehát nincs egzakt multikolinearitás. Ha ezt nem teszük fel, akkor $(X^T X)$ nem lenne invertálható tetszőleges X esetén.
4. $\mathbb{E}[\epsilon|X] = 0$. Ez a feltevés (0 feltételes átlag) alapján a hibatagok átlaga 0. Ebből következik, hogy $\mathbb{E}(Y) = X\beta$.
5. $\mathbb{E}[\epsilon\epsilon^T|X] = \sigma^2 I$. Ezen feltevés szerint a hibatagok homoszkedasztikusak és autokorrelálatlanok.

OLS torzítatlanság

A Gauss-Markov feltevések (1.-4.) alapján $\hat{\beta}_{OLS}$ torzítatlan becslője β-nak. $\mathbb{E}[\hat{\beta}_{OLS}] - \beta = 0$

Bizonyítás

$$\mathbb{E}[\hat{\beta}_{OLS}] = \mathbb{E}[(X^TX)^{-1}X^TY] = \mathbb{E}[(X^TX)^{-1}X^T(X\beta + \epsilon)]$$

$$= \mathbb{E}[(X^TX)^{-1}X^TX\beta] + \mathbb{E}[(X^TX)^{-1}X^T\epsilon]$$

- Ismerjük fel, hogy $(X^TX)^{-1}X^TX = I$, és $\mathbb{E}[\beta] = \beta$ így

$$= \beta + \mathbb{E}[(X^TX)^{-1}X^T\epsilon]$$

- A 4. Gauss-Markov feltevés ($\mathbb{E}(\epsilon|X) = 0$) alapján

$$= \beta$$

$$\implies \mathbb{E}[\hat{\beta}_{OLS}] - \beta = 0$$

OLS kovariancia mátrix

A Gauss-Markov feltevések (1.-5.) alapján $\mathbb{V}[\hat{\beta}_{OLS}] = \sigma^2(X^TX)^{-1}$

Bizonyítás

$$\mathbb{V}[\hat{\beta}_{OLS}] = \mathbb{E}[(\hat{\beta}_{OLS} - \mathbb{E}[\hat{\beta}_{OLS}])(\hat{\beta}_{OLS} - \mathbb{E}[\hat{\beta}_{OLS}])^T]$$

$$= \mathbb{E}[(\hat{\beta}_{OLS} - \beta)(\hat{\beta}_{OLS} - \beta)^T]$$

$$= \mathbb{E}[[(X^TX)^{-1}X^T\epsilon] * [(X^TX)^{-1}X^T\epsilon)^T]]$$

$$= \mathbb{E}[(X^TX)^{-1}X^T\epsilon\epsilon^TX(X^TX)^{-1}]$$

$$= (\mathsf{X}^TX)^{-1}X^T\,\mathbb{E}[\epsilon\epsilon^T]X(X^TX)^{-1}$$

- Az 5. Gauss-Markov feltevés ($\mathbb{E}[\epsilon\epsilon^T|X] = \sigma^2 I$) alapján

$$= (\mathsf{X}^TX)^{-1}X^T\sigma^2IX(X^TX)^{-1}$$

$$= \sigma^2(X^TX)^{-1}X^TX(X^TX)^{-1} = \sigma^2(X^TX)^{-1}$$

OLS konzisztencia

A Gauss-Markov feltevések (1.-3.) alapján az OLS-becslő konzisztens:
$\hat{\beta}_{OLS} \xrightarrow{p} \beta$

Konvergencia (pongyola áttekintés)

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T Y = (\frac{1}{N} X^T X)^{-1} (\frac{1}{N} X^T Y)$$

Felismerve, hogy $X^T X = \sum_{i=1}^{N} \underline{x_i} * \underline{x_i}^T$, a WLLN segítségével belátható, hogy $\frac{1}{N} \sum_{i=1}^{N} \underline{x_i} * \underline{x_i}^T \xrightarrow{p} \mathbb{E}[X^T X]$ (analóg módon $X^T Y$ esetén is belátható). A konvergencia azonosságokat felhasználva:

$$\hat{\beta}_{OLS} \xrightarrow{p} \mathbb{E}[(X^T X)^{-1} X^T Y]$$

Mivel tudjuk, hogy,

$$\mathbb{E}[\beta] = \mathbb{E}[(X^T X)^{-1} X^T Y]$$

ezért

$$\hat{\beta}_{OLS} \xrightarrow{p} \beta$$

Gauss-Markov tétel

Gauss-Markov tétel

Ha teljesül mindegyik feltétel (1.-5.), akkor a **lineáris, torzítatlan becslők körében** az OLS-becslő **minimális varianciájú** (azaz hatásos).

Más szavakkal elmondva, az OLS-becslő **B**est, **L**inear, **U**nbiased, **E**fficient becslő, tehát **BLUE** estimator.