4EK211 Základy ekonometrie

Zobecněná metoda nejmenších čtverců (ZMNČ)

Cvičení 8

Gaussovy-Markovovy předpoklady

Náhodná složka: G.M. předpoklady

1. $E(\boldsymbol{u}) = 0$

2. $E(\boldsymbol{u}\,\boldsymbol{u}^T) = \boldsymbol{\sigma}^2 \boldsymbol{I}_n$
   - konečný a konstantní rozptyl = homoskedasticita
     - → porušení: heteroskedasticita
   - náhodné složky jsou sériově nezávislé
     - → porušení: autokorelace

3. ***X*** je nestochastická matice – $E(\boldsymbol{X}^T\boldsymbol{u}) = 0$
   - veškerá náhodnost je obsažena v náhodné složce

4. ***X*** má plnou hodnost *k*
   - matice ***X*** neobsahuje žádné perfektně lineárně závislé sloupce pozorování vysvětlujících proměnných
     - → porušení: multikolinearita

Zobecněná metoda nejmenších čtverců – ZMNČ

- Při heteroskedastické náhodné složce platí:
  - $E(\boldsymbol{u}) = \mathbf{0}$,
  - $E(\boldsymbol{u}\,\boldsymbol{u}^{\mathrm{T}}) = \sigma^2\mathbf{V}$ (tj. ne $\sigma^2\mathbf{I_n}$)
  - tzv. zobecněný lineární regresní model
- Podstatou MZNČ je **transformace** LRM tak, aby bylo splněno:

$$E(\boldsymbol{u}\,\boldsymbol{u}^{\mathrm{T}}) = \sigma^2\mathbf{I_n}$$

- odhad modifikovaného modelu MNČ
- pomocí transformační matice ***T***
- pomocí matice ***T*** „posouváme“ regresní nadrovinu s cílem zachovat stabilitu regresních koeficientů
- matice ***T*** (postup ZMNČ) se liší u heteroskedasticity a autokorelace

KLRM: $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$ odhadová funkce: $\boldsymbol{b} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$

ZLRM: $\mathbf{T}\boldsymbol{y} = \mathbf{T}\boldsymbol{X\beta} + \mathbf{T}\boldsymbol{u}$ odhadová funkce: $\boldsymbol{b}^* = (\boldsymbol{X}^{\mathrm{T}}\,\mathbf{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\mathbf{V}^{-1}\boldsymbol{y}$,

kde $\mathbf{V}^{-1} = \mathbf{T}^{-1}\mathbf{T}$ → **Aitkenův odhadový postup**

Zobecněná metoda nejmenších čtverců – ZMNČ

Postup ZMNČ:

1. Model odhadneme MNČ.
2. Vyhodnotíme, zda se v modelu vyskytuje heteroskedasticita.
3. Nalezneme/určíme vhodnou transformační matici ***T***.
4. Maticí ***T*** pronásobíme proměnné modelu – získáme upravené proměnné.
5. Odhadneme model složený z upravených proměnných pomocí MNČ.
6. Podle potřeby provedeme zpětnou transformaci modelu, abychom získali odhady regresních parametrů, odpovídajících původní specifikaci LRM.

ZMNČ – heteroskedasticita – pro známé rozptyly $\sigma_i^{\,2}$

- známe rozptyl jednotlivých náhodných složek VAR(u_i) = $\sigma_i^{\,2}$
- transformační matice ***T***:

$$\begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \frac{1}{\sigma_n} \end{bmatrix}$$

- tj. vydělíme původní model σ_i:

$$\frac{y_i}{\sigma_i} = \beta_0\left(\frac{1}{\sigma_i}\right) + \beta_1\left(\frac{x_i}{\sigma_i}\right) + \frac{u_i}{\sigma_i} \;=\; \beta_0\, x_{0i}^* + \beta_1\, x_{1i}^* + u_i^*$$

$$\Downarrow$$

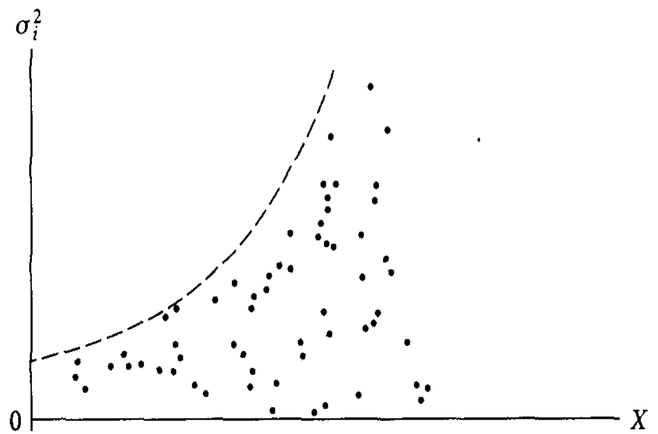
$$E\left(\left(u_i^*\right)^2\right) = 1$$

- spíše teoretická varianta, $\sigma_i^{\,2}$ v praktických úlohách většinou neznáme

ZMNČ – heteroskedasticita – kvadratická závislost

- kvadratická závislost: $\sigma^2 = k^2 x_i^2$
- transformační matice $\boldsymbol{T}$:

$$\begin{bmatrix} \frac{1}{x_1} & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \frac{1}{x_n} \end{bmatrix}$$



- tj. vydělíme původní model x_i :

$$\frac{y_i}{x_i} = \frac{\beta_0}{x_i} + \beta_1 + \frac{u_i}{x_i} = \beta_0 \frac{1}{x_i} + \beta_1 + v_i$$

$$\Downarrow$$

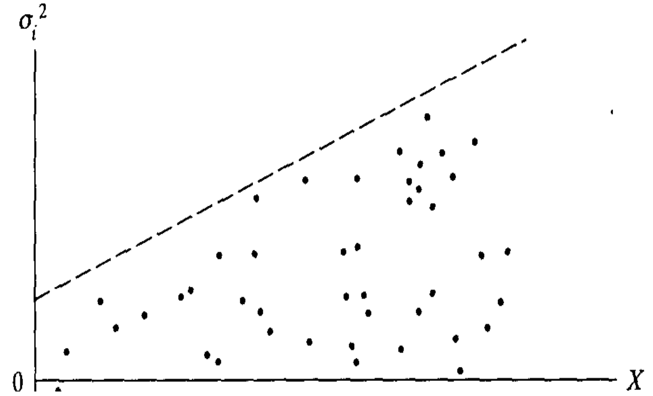
$$E\left(v_i^2\right) = \sigma^2$$

- praktický dopad do odhadu pomocí PcGive/GRETLu:
  - upravím proměnné ***y*** a ***x*** v Algebra editoru
  - konstanta je nyní β_1

ZMNČ – heteroskedasticita – lineární závislost

- lineární závislost: $\sigma^2 = k^2 x_i$
- transformační matice $\boldsymbol{T}$:

$$\begin{bmatrix} \frac{1}{\sqrt{x_1}} & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \frac{1}{\sqrt{x_n}} \end{bmatrix}$$



- vydělíme původní model $\sqrt{x_i}$:

$$\frac{y_i}{\sqrt{x_i}} = \frac{\beta_0}{\sqrt{x_i}} + \beta_1\sqrt{x_i} + \frac{u_i}{\sqrt{x_i}} = \beta_0\frac{1}{\sqrt{x_i}} + \beta_1\sqrt{x_i} + v_i$$

$$\Downarrow$$

$$E\left(v_i^2\right) = \sigma^2$$

- praktický dopad do odhadu pomocí PcGive/GRETLu:
  - upravím proměnné ***y*** a ***x*** v Algebra editoru
  - musím nadefinovat konstantu taky v Algebra editoru!!!

ZMNČ – heteroskedasticita – příklad 1

Soubor: CV8_PR1.xls

Data: ***y*** = výdaje obyvatelstva na zboží v běžných cenách (mld. Kč)
x = disponibilní příjmy obyvatelstva (mld. Kč)
p = index cen zboží

Zadání: Odhadněte závislost výdajů (y) na příjmech (x) a indexu cen zboží (p).
Vyhodnoťte autokorelaci v modelu pro $\alpha = 0,05$.
Vyhodnoťte heteroskedasticitu Whiteovým testem $\alpha = 0,05$.
Vytvořte transformační matici pro ZMNČ.
Transformujte data maticí T a odhadněte model MNČ na transformovaných datech.
Vypište regresní nadrovinu na datech

- transformovaných
- původních.

$y_i = \beta_0 + \beta_1 x_i + \beta_2 p_i + u_i, \qquad i = 1, 2, ..., 15$

ZMNČ – heteroskedasticita – příklad 2

Soubor: CV8_PR2.xls

Data: ***RD*** = výdavky na výskum a rozvoj (mil. USD)
SALES = predaj (mil. USD)
PROFITS = zisk (mil. USD)

Zadání: Odhadněte závislost *RD* na *SALES*.

Vyhodnoťte heteroskedasticitu Whiteovým testem (pozor na $\alpha = 0{,}10$).
Vytvořte transformační matici pro ZMNČ.
Transformujte data maticí T a odhadněte model MNČ na transformovaných datech.
Vypište regresní nadrovinu na datech

- transformovaných
- původních.

$RD_i = \beta_0 + \beta_1 SALES_i + u_i, \quad i = 1, 2, \ldots, 18$

White's Heteroskedasticity consistent estimator (HCE)

White's HCE: základní informace

- Při heteroskedasticitě běžný odhad MNČ zůstává nestranný a konzistentní (není vydatný).
- ALE: odhad rozptylu náhodné složky s^2 , odhad $VAR(\boldsymbol{b})$, resp. standardní chyby bodových odhadů $\boldsymbol{s}_{b_j}$ jsou vychýlené.
- White's HCE: metoda konzistentního odhadu $VAR(\boldsymbol{b})$ při heteroskedasticitě náhodné složky:
  - způsob „ošetření" heteroskedasticity při odhadu $VAR(\boldsymbol{b})$:
  - získáme „konzistentní" odhady $VAR(\boldsymbol{b})$
  - *t*-poměry lze interpretovat „obvyklým způsobem" (i při heteroskedasticitě)
  - lze konstruovat intervaly spolehlivosti bodových odhadů (i při heteroskedasticitě).

White's Heteroskedasticity consistent estimator (HCE)

White's HCE: odvození

Pro homoskedastickou náhodnou složku platí:

$$\text{odh.}VAR(\boldsymbol{b}) = \boldsymbol{S}(\boldsymbol{b}) = s^2(\boldsymbol{X}^T\boldsymbol{X})^{-1} \qquad (1)$$

tento vztah je odvozen z:

$$VAR(\boldsymbol{b}) = VAR\,[\boldsymbol{\beta} + (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{u}] \qquad (2)$$

kde jedním z mezikroků odvození je:

$$VAR(\boldsymbol{b}) = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\sigma^2\boldsymbol{I}_n\,\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1} \qquad (3)$$

a další „zjednodušování“ výrazu (3) je podmíněno homoskedasticitiou náhodné složky, tj. $\mathrm{VAR}(\boldsymbol{u}) = \sigma^2\boldsymbol{I}_n$.

Při heteroskedasticitě platí: $\mathrm{VAR}(\boldsymbol{u}) = \boldsymbol{V}_n$ a vztah (3) přepíšeme jako:

$$\mathrm{VAR}(\boldsymbol{b}) = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\,\boldsymbol{V}_n\,\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1} \qquad (4)$$

kde $\boldsymbol{V}_n$ má plnou hodnost (a při neexistenci autokorelace je diagonální)

White's Heteroskedasticity consistent estimator (HCE)

<u>White's HCE:</u> HCE$_0$

Pokud platí:

$$\mathrm{VAR}(\boldsymbol{b}) = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}_n\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1} \qquad (4)$$

lze rozptyl vektoru ***b*** odhadnout pomocí

$$\mathrm{VAR}(\boldsymbol{b}) = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\hat{\boldsymbol{\Sigma}}_n\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1} \qquad (5)$$

kde $\hat{\boldsymbol{\Sigma}}_n$ je odhadem $\boldsymbol{V}_n$,
je $n_x n$ diagonální matice,
na diagonále $\boldsymbol{V}_n$ rozptyly σ_i^2 nepozorovatelné náhodné složky u_i
na diagonále $\hat{\boldsymbol{\Sigma}}_n$ jejich odhady: $\hat{\sigma}_i^2$

Odhady $\hat{\sigma}_i^2$ jsou vypočteny takto:

$$\hat{\sigma}_i^2 = e_i^2 \qquad (6)$$

Vztah (6) znamená, že jako odhad rozptylu u v i-tém pozorování slouží čtverec rezidua e_i, získaného odhadem LRM pomocí MNČ pro dané pozorování.

White's Heteroskedasticity consistent estimator (HCE)

<u>White's HCE:</u> HCE$_0$

Například, pro n = 20 (20 pozorování) má matice $\hat{\Sigma}_n$ tvar:

$$\hat{\Sigma}_{20} = \begin{bmatrix} e_1^2 & . & . & . & . & 0 \\ . & e_2^2 & & & & . \\ . & & \ddots & & & . \\ . & & & \ddots & & . \\ . & & & & \ddots & . \\ 0 & . & . & . & . & e_{20}^2 \end{bmatrix}$$

Výsledkem výpočtu podle (5) je matice konzistentních odhadů rozptylu vektoru $\boldsymbol{b}$ (bodových odhadů) při heteroskedasticitě náhodné složky, značíme $\boldsymbol{S}_{HCE}(\boldsymbol{b})$. Z diagonálních prvků $\boldsymbol{S}_{HCE}(\boldsymbol{b})$ odečteme odhadnuté rozptyly parametrů b_j. Na základě $\boldsymbol{S}_{HCE}(\boldsymbol{b})$ lze konstruovat konzistentní t-poměry a intervalové odhady („platí“ i při heteroskedasticitě).

White's Heteroskedasticity consistent estimator (HCE)

<u>White's HCE:</u> HCE$_1$

Jeden z řady alternativních postupů odhadu $\boldsymbol{S}_{HCE}(\boldsymbol{b})$, kdy do výpočtu $\hat{\Sigma}_n$ zahrneme stupně volnosti:

$$\hat{\Sigma}_n = \frac{n}{n-k-1}\begin{bmatrix} e_1^2 & . & . & . & . & 0 \\ . & e_2^2 & & & & . \\ . & & \ddots & & & . \\ . & & & \ddots & & . \\ . & & & & \ddots & . \\ 0 & . & . & . & . & e_n^2 \end{bmatrix}$$

neboli: $\hat{\sigma}_i^2 = \dfrac{n}{n-k-1}\, e_i^2$

HCE$_0$, HCE$_1$ a další HCE odhady s_{b_j} jsou implementovány v GRETLu.

ZMNČ – autokorelace

- závislost: $u_t = \rho^* u_{t-1} + \varepsilon_t$; ρ je koeficient autokorelace ρ = (-1; 1)
- Praisova-Winstenova metoda (transformace)
- transformační matice ***T***, transformovaný vektor ***y*** a matice ***X***:

$$T = \frac{1}{\sqrt{1-\rho^2}} \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & 0 \\ -\rho & 1 & 0 & 0 \\ 0 & -\rho & 1 & 0 \\ 0 & 0 & -\rho & 1 \end{bmatrix} \qquad y^* = \begin{bmatrix} y_1\sqrt{1-\rho^2} \\ y_2 - \rho y_1 \\ y_3 - \rho y_2 \end{bmatrix}$$

$$X^* = \begin{bmatrix} \sqrt{1-\rho^2} & x_{11}\sqrt{1-\rho^2} & x_{21}\sqrt{1-\rho^2} \\ 1-\rho & x_{12} - \rho x_{11} & x_{22} - \rho x_{21} \\ 1-\rho & x_{13} - \rho x_{12} & x_{23} - \rho x_{22} \end{bmatrix}$$

- P-W transformace založena na částečných (zobecněných) diferencích všech proměnných LRM + aproximace pro první období.
- Při transformaci vynecháváme zlomek před maticí – jde o konstantu, takže výsledek regrese není ovlivněn

ZMNČ – autokorelace

Transformace Cochrane-Orcutt

- Založena na zobecněných diferencích
  (opakovaně postup P-W, bez aproximace prvního pozorování)
- Iteracemi zpřesňujeme odhad autokorelačního koeficientu i odhad regresních parametrů modelu.
- Metoda zabudována do PcGive, GRETLu i Eviews

- Postup: viz samostatný soubor .pdf (detaily této metody přesahují rámec kurzu Základy ekonometrie)

ZMNČ – autokorelace – příklad

Soubor: CV8_PR3.xls

Data: ***CONS*** = spotřební výdaje
INC = disponibilní příjmy

Zadání: Odhadněte závislost výdajů (*CONS*) na příjmech (*INC*).
Vyhodnoťte autokorelaci v modelu pro $\alpha = 0{,}05$.
Vytvořte transformační matici pro ZMNČ.
Transformujte data maticí T a odhadněte model MNČ na transformovaných datech.
Vypište regresní nadrovinu na datech

- transformovaných
- původních.

$$CONS_i = \beta_0 + \beta_1 INC_i + u_i, \qquad i = 1, 2, \ldots, 159$$