

# Cvičení z Teorie ekonometrie I – 12.3.2008, 19.3.2008

- **Obsah:** Metoda nejmenších čtverců - vybrané otázky a ilustrace.
- **Změna v součtu čtverců.** Předpokládejme, že  $\mathbf{b}$  je vektor parametrů získaný metodou nejmenších čtverců regrese  $\mathbf{y}$  na  $\mathbf{X}$  a  $\mathbf{c}$  je jiný vektor rozměru  $K \times 1$ . Dokažte, že rozdíl dvou součtů čtverců reziduí je

$$(\mathbf{y} - \mathbf{Xc})'(\mathbf{y} - \mathbf{Xc}) - (\mathbf{y} - \mathbf{Xb})'(\mathbf{y} - \mathbf{Xb}) = (\mathbf{c} - \mathbf{b})' \mathbf{X}' \mathbf{X} (\mathbf{c} - \mathbf{b})$$

Dokažte, že tento rozdíl je kladný.

- **Lineární transformace dat.** Předpokládejme regresi metodou nejmenších čtverců  $y$  na  $K$  proměnných (s konstantním členem)  $X$ . Předpokládejme alternativní sadu regresorů  $Z = \mathbf{XP}$ , kdy  $P$  je nesingulární matice. Každý sloupec matice  $Z$  je tedy mixem některých sloupců  $X$ . Dokažte, že vektor reziduí v regresi  $y$  na  $X$  a  $y$  na  $Z$  jsou identické. Jaký význam to má pro otázku kvality (vystižení) regrese změnou měřítek u nezávislých proměnných?
- **Frisch and Waugh.** V regresi pomocí metody nejmenších čtverců  $y$  na konstantu a  $X$  můžeme spočítat regresní koeficienty i tak, že nejdříve transformujeme  $y$  na své odchylky od střední hodnoty (průměru)  $\bar{y}$  a stejně tak i upravíme sloupce matice  $X$ . Po té provedeme regresi takto centrovaných hodnot na transformované hodnoty matice  $X$  (bez konstanty). Získáme stejné výsledky pokud takto budeme transformovat jen  $y$ ? A co když transformujeme pouze  $X$ ? Zkuste si tento postup i na empirických datech.
- Předpokládejme, že  $E_d$ ,  $E_n$ ,  $E_s$  jsou výdaje na tři kategorie zboží (consumer durables, non-durables and services). Celkový příjem (důchod) je pak dán jako  $Y = E_d + E_n + E_s$ . Předpokládejme dále, že je dán výdajový systém:

$$\begin{aligned} E_d &= \alpha_d + \beta_d Y + \gamma_{dd} P_d + \gamma_{dn} P_n + \gamma_{ds} P_s + \epsilon_d \\ E_n &= \alpha_n + \beta_n Y + \gamma_{nd} P_d + \gamma_{nn} P_n + \gamma_{ns} P_s + \epsilon_n \\ E_s &= \alpha_s + \beta_s Y + \gamma_{sd} P_d + \gamma_{sn} P_n + \gamma_{ss} P_s + \epsilon_s \end{aligned}$$

– Jestliže všechny rovnice odhadneme metodou nejmenších čtverců, dokažte, že součet důchodových koeficientů bude jednička a součet ostatních koeficientů (po sloupcích) bude nulový.

- Využijte data v matlabovském datovém souboru `wage2.mat` k odhadu jednoduché regrese vysvětlující měsíční plat (*wage*) na dosaženém počtu bodů IQ (*IQ*). Datový soubor je nahrán a "zpracován" v m-fajlu `cv02_wage2.m`.

- Nalezněte průměrnou mzdu a průměrné IQ ve vzorku. Vykreslete datové vzorky (se svými průměry). Jaká je standardní odchylka IQ? (IQ je standardizováno tak, že průměr populace e 100 a standardní odchylka 15)
  - Odhadněte jednoduchý regresní model kde jednobodové zvýšení IQ změní mzdu o konstantní výši (v dolarech). Využijte tento moel k predikci zvýšení mzdy pokud by IQ vzrostlo o 15 bodů. Vysvětluje *IQ* většinu variability ve mzdě?
  - Odhadněte model, kde každé zvýšení IQ o jeden bod má podobný procentní efekt na mzdu. Pokud se *IQ* zvýší o 15 bodů, jaké bude přibližné procentní zvýšení predikované mzdy?
  - K výpočtu si zkuste vytvořit jednak svou vlastní funkci s názvem např. `moje_ols.m` popř. pak využijte funkci `ols.m` z ekonometrického toolboxu.
- Soubor `USGas_Greene.dat` obsahuje údaje o spotřebě benzínu v USA v letech 1960-1995 a další časové řady, které jsou popsány v souboru `usgas_greene.m`.
    - Spočítejte regresi spotřeby benzínu na hlavu na všechny ostatní vysvětlující proměnné, včetně časového trendu. Jsou znaménka odhadnutých parametrů v souladu s vaším očekáváním?
    - Testujte hypotézu, že alespoň z pohledu poptávky po benzínu spotřebitel nerozlišuje mezi změnami v cenách nových a použitých vozů.
    - Odhadněte cenovou elasticitu poptávky, důchodovou elasticitu poptávky a křížovou cenovou elasticitu s ohledem na cenu veřejné dopravy.
    - Odhadněte předchozí regresi v logaritmech, kdy koeficienty budou přímo odhady elasticit (samozřejmě časový trend nelogaritmujte). Porovnejte tyto odhady s odhady předchozí podotázky. Jakou specifikaci byste upřednostnili?
    - Cenové indexy trhu s automobily jsou normalizovány vzhledem k roku 1967, přičemž agregátní cenové indexy jsou ukotveny vzhledem k roku 1982? Ovlivňuje tento nesoulad výsledky? Jak? Pokud znormalizujete veškeré indexy tak, aby v roce 1982 měli hodnotu 1, jak to ovlivní výsledky?
    - Upravte svůj model tak, aby neobsahoval většinu statisticky nevýznamných proměnných a zopakujte úlohu na test hypotézy o vlivu cen nových a ojetých automobilů, případně porovnejte své nové výsledky elasticit s původními odhady.