Cvičení z Teorie ekonometrie I – 18.3.2009, 25.3.2009

- **Obsah:** Metoda maximální věrohodnosti - vybrané otázky a ilustrace.
- **ML odhad.** Vytvořte si vlastní umělý model (např. s dvěma vysvětlujícími proměnnými a úrovňovou konstantou) a odhadněte parametry tohoto modelu metodou maximální věrohodnosti. Srovnejte výsledky s odhady metodou nejmenších čtverců pro malé a velké vzorky. K odhadu parametrů využijte Matlabovskou funkci `fminunc`.
- Využijte data v matlabovském datovém souboru `hprice1.mat` k odhadu modelu prodejních cen domů v jednom městečku. Vyjděte z m-fajlu `cv04_hprice1.m`. Jako odhadové techniky opět využijte OLS popř. ML.
  - Zapište výsledky v rovnicovém vyjádření a interpretujte je.
  - Jaké je odhadované zvýšení ceny domu s dodatečnou ložnicí, přičemž rozloha domu se nemění?
  - Jaké je odhadované zvýšení ceny domu s dodatečnou ložnicí o rozloze 140 čtverečních stop? (porovnejte svou odpověď s předchozí otázkou)
  - Modifikujte svůj model pro analýzu toho, jestli koloniální styl domu ovlivňuje cenu za metr čtvereční.
  - Kolik procent variability v ceně domu je vysvětleno modelem?
  - Porovnejte skutečnou prodejní cenu prvního domu a cenou, kterou predikuje váš model. Jaké je příslušné reziduum? Znamená to, že kupec dal více než by měl?
  - Vymyslete si i další hypotézy a tomu odpovídající specifikace modelu, které lze na datovém vzorku testovat.

- **Regresní model s dvěma vysvětlujícími proměnnými.** Pro regresní model $y = \alpha + \beta x + \epsilon$:
    - Ukažte, že normální rovnice pro metodu nejmenších čtverců implikují $\sum_i e_i = 0$ a $\sum_i x_i e_i = 0$.
    - Ukažte, že řešení pro úrovňovou konstantu je $a = \overline{y} - b\overline{x}$.
    - Ukažte, že řešení pro b je $b = [\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})]/[\sum_{i=1}^{n}(x_i - \overline{x})^2]$.
    - Dokažte, že tyto dvě hodnoty jednoznažně minimalizují součet čterců. Ukažte tedy, že diagonální prvky matice druhých derivací sumy čtverců podle jednotlivých parametrů jsou oba pozitivní a že determinant je roven $4n[(\sum_{i=1}^{n} x_i^2) - n\overline{x}^2] = 4n[\sum_{i=1}^{n}(x_i - \overline{x})^2]$ a je kladný pokud nejsou všechny hodnoty x stejné.
- **Změna v součtu čtverců.** Předpokládejme, že **b** je vektor parametrů získaný metodou nejmenších čtverců regresí **y** na **X** a **c** je jiný vektor rozměru $K \times 1$. Dokažte, že rozdíl dvou součtů čtverců reziduí je

  $$(y - Xc)'(y - Xc) - (y - Xb)'(y - Xb) = (c - b)'X'X(c - b)$$

  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 $Z = XP$, 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 příslušející proměnným v matici X tak, že nejdříve transformujeme y na své odchylky od střední hodnoty (průměru) $\overline{y}$ a stejně tak upravíme sloupce matice X. Po té provedeme regresi takto centrovaných hodnot na transformované hodnoty matice X (již 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 nejměnších čtverců, dokažte, že součet důchodových koeficientů bude jednička a součet ostatních koeficientů (po sloupcích) bude nulový.