

Cvičení z Teorie ekonometrie I – 27.2.2008, 5.2.2008

- **Obsah:** Lineární regrese. Normální rovnice. Ekonometrický toolbox.
- **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 úrovnovou konstantu je $a = \bar{y} - b\bar{x}$.
 - Ukažte, že řešení pro b je $b = [\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})] / [\sum_{i=1}^n (x_i - \bar{x})^2]$.
 - Dokažte, že tyto dvě hodnoty jednoznačně minimalizují součet čtverců. Ukažte tedy, že diagonální prvky matice druhých derivací sumy čtverců podle jednotlivých parametrů jsou oba pozitivní a že determinant je roven $4n[(\sum_{i=1}^n x_i^2) - n\bar{x}^2] = 4n[\sum_{i=1}^n (x_i - \bar{x})^2]$ a je kladný pokud nejsou všechny hodnoty x stejné.
- **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

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

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 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.

- **Simulace dat a odhad OLS** Vytvořte dva vektory (vysvětlujících proměnných) x_1 a x_2 (využijte např. generátory náhodných čísel) o délce např. $N = 100$. Zvolte parametry a_0 , a_1 a a_2 . Vygenerujte si vektor náhodných složek ϵ z normálního rozdělení se střední hodnotou nula a nějakým rozptylem σ^2 . Na tomto základě vygenerujte vektor vysvětlující proměnné y :

$$y = a_0 + a_1 x_1 + a_2 x_2 + \epsilon$$

Odhadnět pomocí metody nejmenších čtverců parametry výše uvedeného modelu. Využijte funkci `ols.m` z ekonometrického toolboxu. Jaký vliv na přesnost odhadu bude mít velikost zvoleného vzorku N ?

- 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.