

## Cvičení z Teorie ekonometrie I – 9.4.2008, 16.4.2008

- **Obsah:** Metoda nejmenších čtverců - dokončení teoretických otázek. Metoda maximální věrohodnosti - teorie a praktická aplikace.
- **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  $Z = XP$ , 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.
- **ML odhad.** Vytvořte si vlastní umělý model (např. s dvěma vysvětlujícími proměnnými a úrovnovou konstantou) a odhadněte parametry tohoto modelu metodou maximální věrohodnosti. Srovnajte výsledky s odhadem metodou nejmenších čtverců pro malé a velké vzorky. K odhadu parametrů využijte Matlabovskou funkci `fminunc`.

- Využijte data v matlabovském datovém souboru `wage2.mat` k odhadu mzdové rovnice (tedy rovnice měsíčního platu) obsahující vámi specifikované proměnné (abyste byli schopni zodpovědět níže uvedené otázky). Datový soubor je "zpracován" v m-fajlu `cv04_wage2.m`. Jako odhadové metody zvolte metodu nejmenších čtverců nebo metodu maximální věrohodnosti (popř. obě pro srovnání výsledků). Dle potřeby formulujte více alternativních modelů.
  - Vypiště výslednou odhadnutou rovnici a omentujte výsledky.
  - Ověřte hypotézu, že dodatečný rok obecných pracovních zkušeností má tentýž vliv na procentní zvýšení mzdy jako dodatečný rok setrvání u stávajícího zaměstnavatele.
  - Ověřte hypotézu, zda-li existuje ve firmě mzdová nebo rasová diskriminace, a to z hlediska absolutní výše mezd, tak i z hlediska tempa růstu mezd.
  - Která z vašich vysvětlujících proměnných vysvětluje největší procento variability v absolutním růstu mezd a která v rámci variability tempa růstu mezd?
- 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)
  - 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?