Cvičenia z ekonometrie 2005/2006

Domáca úloha 12 - riešenie

1. *Náhodná premenná X má Laplaceovo rozdelenie s parametrom $a > 0$, ak jej hustota je*

$$f(x) = \frac{1}{2a}\, e^{-\frac{1}{a}|x|}.$$

Uvažujme teraz regresný model

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

kde ε_i sú nezávislé náhodné premenné s Laplaceovym rozdelením a odhad tohto modelu metódou maximálnej vierohodnosti. Dokážte, že odhady parametrov β_0 a β_1 z dát $(x_1, y_1), \ldots, (x_n, y_n)$ sú riešením minimalizačnej úlohy

$$\sum_{i=1}^{n} |y_i - (\beta_0 + \beta_1 x_i)| \to \min \tag{1}$$

Ako odhadneme parameter a?

Riešenie:
Funkcia vierohodnosti pre tento model je

$$L = \prod_{i=1}^{n} f(y_i - (\beta_0 + \beta_1 x_i)) = \left(\frac{1}{2a}\right)^n e^{-\frac{1}{a}\sum_{i=1}^{n}|y_i - (\beta_0 + \beta_1 x_i)|}$$

a je prirodzený logaritmus je

$$\ell = -n \log 2a - \frac{1}{a}\sum_{i=1}^{n} |y_i - (\beta_0 + \beta_1 x_i)|\,. \tag{2}$$

Pre ľubovoľné fixované a má funkcia ℓ maximum v bode, ktoré je riešením úlohy

$$\sum_{i=1}^{n} |y_i - (\beta_0 + \beta_1 x_i)| \to \min_{\beta_0, \beta_1}. \tag{3}$$

Táto úloha (a teda ani optimálne riešenie) nezávisí od a, takže $\hat{\beta}_0, \hat{\beta}_1$ dostaneme ako optimálne riešenia úlohy (3).

Dosaďme teraz tieto optimálne $\hat{\beta}_0, \hat{\beta}_1$ do funkcie (2). Dostaneme

$$\ell = -n \log 2a - \frac{1}{a}\sum_{i=1}^{n} \left|y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right|.$$

Nájdeme maximum tejto funkcie vzhľadom na a. Z podmienky

$$\left.\frac{d\ell}{da}\right|_{a=\hat{a}} = -\frac{n}{\hat{a}} + \frac{1}{\hat{a}^2}\sum_{i=1}^{n} \left|y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right| = 0$$

dostaneme

$$\hat{a} = \frac{1}{n}\sum_{i=1}^{n} \left|y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right|.$$

Poznámka: Výsledok nemôže byť $\hat{a} = \frac{1}{n}\sum_{i=1}^{n} |y_i - (\beta_0 + \beta_1 x_i)|$. Hodnoty β_0 a β_1 sú skutočné hodnoty parametrov a sú neznáme. Preto ich nemôžeme dosadzovať.

Bonus:
Navrhnite postup, ako vyriešiť optimalizačnú úlohu (1). (Absolútna hodnota nie je diferencovateľná, takže napríklad gradientná alebo Newtonova metóda sa nedá použiť.)

Riešenie:
Úloha (1) je ekvivalentná s úlohou

$$\sum_{i=1}^{n} u_i \to \min$$

$$-u_i \leq y_i - (\beta_0 + \beta_1 x_i) \leq u_i \quad (i = 1, \ldots, n)$$

s premennými β_0, β_1, u_i $(i = 1, \ldots, n)$. To je úloha lineárneho programovania a dá sa vyriešiť niektorým z algoritmov LP (napríklad simplexovou metódou).

2. *Uvažujme regresný model*

$$Y = X\beta + \varepsilon$$

kde vektor ε má normálne rozdelenie $N(0, \sigma^2 I)$ a jeho odhad metódou maximálnej vierohodnosti. Na cvičení sme ukázali, že odhad parametra β sa zhoduje s odhadom metódou najmenších štvorcov. Nájdite odhad parametra σ^2.

Riešenie:
Vieme, že logartimus funkcie vierohodnosti je

$$\ell = -\frac{n}{2} \log 2\pi - \frac{1}{2} \log \det(\sigma^2 I) - \frac{1}{2\sigma^2} (Y - X\beta)^T (Y - X\beta)$$

Za β dosadíme $\hat{\beta}$ (odhad maximálnej vierohodnosti tohto parametra modelu už poznáme) a maximalizujeme funkciu

$$\begin{aligned} \ell &= -\frac{n}{2} \log 2\pi - \frac{1}{2} \log \det(\sigma^2 I) - \frac{1}{2\sigma^2} (Y - X\hat{\beta})^T (Y - X\hat{\beta}) \\ &= -\frac{n}{2} \log 2\pi - \frac{1}{2} \log(\sigma^2)^n - \frac{1}{2\sigma^2} (Y - X\hat{\beta})^T (Y - X\hat{\beta}) = \\ &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} (Y - X\hat{\beta})^T (Y - X\hat{\beta}) \end{aligned}$$

vzhľadom na β. Z podmienky

$$\left.\frac{d\ell}{d\sigma^2}\right|_{\sigma_2 = \hat{\sigma}^2} = -\frac{n}{2\hat{\sigma}^2} + \frac{1}{2(\sigma^2)^2} (Y - X\hat{\beta})^T (Y - X\hat{\beta}) = 0$$

vyplýva, že

$$\hat{\sigma}^2 = \frac{1}{n} (Y - X\hat{\beta})^T (Y - X\hat{\beta}),$$

resp., ak označíme e vektor rezíduí, tak $\hat{\sigma}^2 = \frac{e^T e}{n}$.

3. *Uvažujme regresný model*

$$Y = \beta_0 + \beta_1 x_1 + \ldots \beta_k x_k + \varepsilon,$$

kde vektor ε má nulovú strednú hodnotu a jeho kovariančná matica je diagonálna - s prvkami $\sigma_1^2, \ldots, \sigma_n^2$ na diagonále. Dokážte, že odhad zovšeobecnenou metódou najmenších štvorcov je riešením úlohy

$$\sum_{i=1}^{n} w_i (y_i - (\beta_0 + \beta_1 x_{1i} + \ldots \beta_k x_{ki}))^2 \to \min,$$

kde $w_i = \frac{1}{\sigma_i^2}$.

Poznámka:
Tvrdenie tohto príkladu znamená, že minimalizujme váženú sumu štvorcov rezíduí. Preto sa tento prípad nazýva vážená metóda najmenších štvorcov

Riešenie:
Vieme, že zovšeobecnenou metódou najmenších štvorcov dostaneme ako odhad modelu $Y = X\beta + \varepsilon$ optimálne riešenie úlohy

$$(Y - X\beta)^T \Omega^{-1} (Y - X\beta) \to \min . \tag{4}$$

V našom prípade

$$\begin{aligned} (Y - X\beta)_i &= y_i - (\beta_0 + \beta_1 x_{1i} + \dots \beta_k x_{ki}) \\ \Omega &= diag(\sigma_1^2, \dots, \sigma_n^2) \text{ a teda } \Omega^{-1} = diag\left(\frac{1}{\sigma_1^2}, \dots, \frac{1}{\sigma_n^2}\right) \end{aligned}$$

Rozpísaním minimalizovanej funkcie v (4) dostaneme

$$(Y - X\beta)^T \Omega^{-1} (Y - X\beta) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} (y_i - (\beta_0 + \beta_1 x_{1i} + \dots \beta_k x_{ki}))^2 .$$

To znamená, že minimalizujeme funkciu $\sum_{i=1}^{n} w_i (y_i - (\beta_0 + \beta_1 x_{1i} + \dots \beta_k x_{ki}))^2$, kde $w_i = \frac{1}{\sigma_i^2}$, čo bolo treba dokázať.