Cvičenia z ekonometrie 2005/2006

Riešenie príkladu 2a z DÚ7

Dôkaz z prednášky a z cvičenia:
Uvažujme model

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

pričom X je matica $k \times n$ s lineárne nezávislými stĺpcami a platí $E[\varepsilon] = 0, Var[\varepsilon] = \sigma^2 I$. Označme $\hat{\beta}$ MNŠ odhad parametra $\hat{\beta}$ v tomto modeli, t.j. $\hat{\beta} = (X^T X)^{-1} X^T Y$. Potom

$$\hat{\sigma}^2 = \frac{e^T e}{n-k},$$

kde e sú rezíduá z regresie, je nevychýlený odhad parametra σ^2.

Vektor rezíduí e sa dá zapísať ako

$$e = Y - \hat{Y} = Y - X(X^T X)^{-1} X^T Y = (I - X(X^T X)^{-1} X^T) Y = M_X Y,$$

pričom $M_X = I - X(X^T X)^{-1} X^T$ je matica projekcie na $S(X)^\perp$. Preto

$$\hat{\sigma}^2 = \frac{e^T e}{n-k} = \frac{(M_X Y)^T (M_X Y)}{n-k}.$$

Do výrazu $M_X Y$ dosadíme $Y = X\beta + \varepsilon$, a využijeme, že $M_X X = 0$:

$$M_X Y = M_X (X\beta + \varepsilon) = M_X X \beta + M_X \varepsilon = M_X \varepsilon.$$

To znamená, že

$$\hat{\sigma}^2 = \frac{(M_X Y)^T (M_X Y)}{n-k} = \frac{(M_X \varepsilon)^T (M_X \varepsilon)}{n-k} = \frac{\varepsilon^T M_X \varepsilon}{n-k},$$

lebo $M_X^T M_X = M_X M_X = M_X$ (matica M_X je symetrická a idempotentná). Počítajme teraz strednú hodnotu:

$$E\left[\hat{\sigma}^2\right] = E\left[\frac{\varepsilon^T M_X \varepsilon}{n-k}\right] = \frac{E\left[\varepsilon^T M_X \varepsilon\right]}{n-k} = \frac{E\left[Tr\left(\varepsilon^T M_X \varepsilon\right)\right]}{n-k}.$$

Posledná rovnosť vyplýva z toho, že $\varepsilon^T M_X \varepsilon$ je číslo, a preto sa rovná $Tr\left(\varepsilon^T M_X \varepsilon\right)$, ak ho chápeme ako maticu 1×1. Teraz použijeme vlastnosť stopy: $Tr(AB) = Tr(BA)$, zameníme poriadie stopy a strednej hodnoty (to sa dá, lebo stredná hodnota je lineárna) a využijeme nenáhodnosť matice M_X:

$$E\left[\hat{\sigma}^2\right] = \frac{E\left[Tr\left(M_X \varepsilon \varepsilon^T\right)\right]}{n-k} = \frac{Tr\left(E\left[M_X \varepsilon \varepsilon^T\right]\right)}{n-k} = \frac{Tr\left(M_X E\left[\varepsilon \varepsilon^T\right]\right)}{n-k}.$$

Podľa predpokladu je $E\left[\varepsilon \varepsilon^T\right] = \sigma^2 I$, a teda

$$E\left[\hat{\sigma}^2\right] = \frac{Tr\left(M_X \sigma^2 I\right)}{n-k} = \frac{\sigma^2 Tr\left(M_X\right)}{n-k}.$$

Na cvičení sme dokázali, že pre symetrickú idempotentnú maticu sa jej hodnosť rovná stope. Pretože hodnosť matice M_X je $n-k$ (dimenzia priestoru $S(X)^\perp$), máme

$$E\left[\hat{\sigma}^2\right] = \sigma^2.$$

Domáca úloha:

Nech platí

$$Y = X\beta + \varepsilon, \tag{1}$$

pričom štandardné predpoklady sú splnené (X je typu $n \times k$, má lineárne nezávislé stĺpce, $E[\varepsilon] = 0$, $Var[\varepsilon] = \sigma^2 I$). Vieme, že odhadom z n dát metódou najmenších štvorcov dostaneme

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

Predpokladajme, že namiesto správneho modelu odhadujeme model

$$Y = Z\gamma + \eta, \tag{2}$$

kde Z je typu $n \times r$, má lineárne nezávislé stĺpce a predpokladáme, že $E[\eta] = 0$, $Var[\eta] = s^2 I$. To znamená, že na vysvetlenie Y sme zvolili nesprávny model. Odhadom tohto modelu metódou najmenších štvorcov dostaneme

$$\hat{\gamma} = (Z^T Z)^{-1} Z^T Y, \hat{s}^2 = \frac{(Y - Z\hat{\gamma})^T (Y - Z\hat{\gamma})}{n-r}.$$

Dokážte, že $E[\hat{s}^2] \geq E[\hat{\sigma}^2]$

Vieme, že $E[\hat{\sigma}^2] = \sigma^2$. Vypočítame teraz $E[\hat{s}^2]$. Pretože odhadujeme model (2), vektor rezíduí z regresie je $Y - Z\hat{\gamma}$ a sa dá zapísať ako

$$Y - Z\hat{\gamma} = Y - Z(Z^T Z)^{-1} Z^T Y = (I - Z(Z^T Z)^{-1} Z^T) Y = M_Z Y,$$

pričom $M_Z = I - X(Z^T Z)^{-1} Z^T$ je matica projekcie na $S(Z)^{\perp}$. Preto

$$\hat{s}^2 = \frac{(Y - Z\hat{\gamma})^T (Y - Z\hat{\gamma})}{n-r} = \frac{(M_Z Y)^T (M_Z Y)}{n-r}.$$

Do výrazu $M_Z Y$ dosadíme Y. Treba si uvedomiť, že skutočná hodnota Y je daná správnym modelom (1), teda $Y = X\beta + \varepsilon$:

$$M_Z Y = M_Z (X\beta + \varepsilon) = M_Z X\beta + M_Z \varepsilon.$$

To znamená, že

$$\begin{aligned}
\hat{s}^2 &= \frac{(M_Z Y)^T (M_Z Y)}{n-r} = \frac{(M_Z X\beta + M_X \varepsilon)^T (M_Z X\beta + M_Z \varepsilon)}{n-r} = \\
&= \frac{\beta^T X^T M_Z^T M_Z X\beta + 2\beta^T X^T M_Z^T M_Z \varepsilon + \varepsilon^T M_Z^T M_Z \varepsilon}{n-r} = \\
&= \frac{\beta^T X^T M_Z X\beta + \varepsilon^T M_Z \varepsilon}{n-r},
\end{aligned}$$

lebo $M_Z^T M_Z = M_Z M_Z = M_Z$ (matica M_Z je symetrická a idempotentná), riadok $\beta^T X^T M_Z^T M_Z$ je nenáhodný a $E[\varepsilon] = 0$. Počítajme teraz strednú hodnotu:

$$\begin{aligned}
E\left[s^2\right] &= E\left[\frac{\beta^T X^T M_Z X\beta + \varepsilon^T M_Z \varepsilon}{n-r}\right] = \frac{E\left[\beta^T X^T M_Z X\beta + \varepsilon^T M_Z \varepsilon\right]}{n-r} = \\
&= \frac{\beta^T X^T M_Z X\beta}{n-r} + \frac{E\left[\varepsilon^T M_Z \varepsilon\right]}{n-r} = \frac{\beta^T X^T M_Z X\beta}{n-r} + \frac{E\left[Tr\left(\varepsilon^T M_Z \varepsilon\right)\right]}{n-r}.
\end{aligned}$$

Posledná rovnosť vyplýva z toho, že $\varepsilon^T M_Z \varepsilon$ je číslo, a preto sa rovná $Tr\left(\varepsilon^T M_Z \varepsilon\right)$, ak ho chápeme ako maticu 1×1. Teraz použijeme vlastnosť stopy: $Tr(AB) = Tr(BA)$, zameníme poriadie stopy a strednej hodnoty (to sa dá, lebo stredná hodnota je lineárna) a využijeme nenáhodnosť matice M_Z:

$$\begin{aligned} E\left[\hat{s}^2\right] &= \frac{\beta^T X^T M_Z X \beta}{n-r} + \frac{E\left[Tr\left(M_Z \varepsilon \varepsilon^T\right)\right]}{n-r} = \frac{\beta^T X^T M_Z X \beta}{n-r} + \\ &+ \frac{Tr\left(E\left[M_Z \varepsilon \varepsilon^T\right]\right)}{n-r} = \frac{\beta^T X^T M_Z X \beta}{n-r} + \frac{Tr\left(M_Z E\left[\varepsilon \varepsilon^T\right]\right)}{n-r}. \end{aligned}$$

Podľa predpokladu je $E\left[\varepsilon \varepsilon^T\right] = \sigma^2 I$, a teda

$$E\left[\hat{s}^2\right] = \frac{\beta^T X^T M_Z X \beta}{n-r} + \frac{Tr\left(M_Z \sigma^2 I\right)}{n-r} = \frac{\beta^T X^T M_Z X \beta}{n-r} + \frac{\sigma^2 Tr\left(M_Z\right)}{n-r}.$$

Na cvičení sme dokázali, že pre symetrickú idempotentnú maticu sa jej hodnosť rovná stope. Pretože hodnosť matice M_Z je $n-r$ (dimenzia priestoru $S(Z)^{\perp}$), máme

$$E\left[\hat{s}^2\right] = \sigma^2 + \frac{\beta^T X^T M_Z X \beta}{n-r},$$

t.j.

$$E\left[\hat{s}^2\right] = E\left[\hat{\sigma}^2\right] + \frac{\beta^T X^T M_Z X \beta}{n-r}.$$

Zostáva teda dokázať, že

$$\frac{\beta^T X^T M_Z X \beta}{n-r} \geq 0.$$

Táto nerovnosť však platí, lebo matica M_Z je kladne semidefinitná. Pre každý vektor v platí:

$$v^T M_Z v = v^T M_Z M_Z v = v^T M_Z^T M_Z v = (M_Z v)^T (M_Z v) \geq 0.$$

Rovnosť nastáva práve vtedy, keď $M_Z v$ je nulový vektor. Keďže M_Z je matica projekcie na $S(Z)^{\perp}$, znamená to, že projekcia vektora v na $S(Z)^{\perp}$ je nulový vektor. To je práve vtedy, keď vektor v je z priestoru $S(Z)$ Aby sme dokázali naše tvrdenie, zoberieme $v = X\beta$. Ďalej vidíme, že rovnosť medzi $E\left[\hat{s}^2\right]$ a $E\left[\hat{\sigma}^2\right]$ je práve vtedy, keď vektor $X\beta$ je z priestoru $S(Z)$. To znamená, že stredná hodnota vektora Y sa dá napísať ako lineárna kombinácia stĺpcov matice Z (t.j. vysvetľujúcich premenných v modeli (2)).