Lineaariset mallit, kevät 2014
Ratkaisut 2
Juho Leppänen

1. Palautetaan mieleen, että $\boldsymbol{y} = \hat{\boldsymbol{y}} + \hat{\boldsymbol{\epsilon}}$, missä sovite $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$. Kun kirjoitetaan $\hat{\boldsymbol{y}} = \boldsymbol{y} - \hat{\boldsymbol{\epsilon}}$ ja sijoitetaan tämä lausekkeeseen $\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}}$ ensimmäisen $\hat{\boldsymbol{y}}$:n paikalle, saadaan

$$\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}} = (\boldsymbol{y} - \hat{\boldsymbol{\epsilon}})'\hat{\boldsymbol{y}} = (\boldsymbol{y}' - \hat{\boldsymbol{\epsilon}}')\hat{\boldsymbol{y}} = \boldsymbol{y}'\hat{\boldsymbol{y}} + \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{y}} = \boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{y}}.$$

Koska mallin sovite ja residuaali ovat ortogonaaliset (ks. materiaalin sivu 9) eli $\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{y}} = 0$, seuraa tehtävän ensimmäinen väite.

Siirrytään tarkastelemaan mallia, jossa on vakio, jolloin matriisin $\boldsymbol{X}$ ensimmäinen sarake on ykkösvektori $\mathbf{1}_n$. Matriisin $\boldsymbol{X}$ i. saraketta on seuraavassa merkitty symbolilla $\boldsymbol{x}_i$, kun $i \in \{1, \ldots, p\}$. Toisen väitteen todistamiseksi muistetaan, että PNS-estimaatille $\hat{\boldsymbol{\beta}}$ pätee

$$(*) \qquad \boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y}.$$

Tarkastellaan tässä ensimmäistä yhtälöä. Matriisin $\boldsymbol{X}'$ ensimmäinen rivi on matriisin $\boldsymbol{X}$ ensimmäisen sarakkeen transpoosi eli $\boldsymbol{x}_1' = \mathbf{1}_n'$. Täten tulomatriisin $\boldsymbol{X}'\boldsymbol{X}$ ensimmäinen rivi on

$$\begin{bmatrix} \boldsymbol{x}_1'\boldsymbol{x}_1 & \cdots & \boldsymbol{x}_1'\boldsymbol{x}_p \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n'\mathbf{1}_n & \mathbf{1}_n'\boldsymbol{x}_2 & \cdots & \mathbf{1}_n'\boldsymbol{x}_p \end{bmatrix} = \begin{bmatrix} n & n\overline{\boldsymbol{x}}_2 & \cdots & n\overline{\boldsymbol{x}}_p \end{bmatrix}$$

Siis nähdään, että vektorin $\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$ ensimmäinen komponentti on

$$\begin{bmatrix} n & n\overline{\boldsymbol{x}}_2 & \cdots & n\overline{\boldsymbol{x}}_p \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{bmatrix} = n\hat{\beta}_1 + n\overline{\boldsymbol{x}}_2\hat{\beta}_2 + \ldots + n\overline{\boldsymbol{x}}_p\hat{\beta}_p.$$

Yhtälön (*) nojalla tämä on sama kuin vektorin $\boldsymbol{X}'\boldsymbol{y}$ ensimmäinen komponentti, mikä puolestaan on $\mathbf{1}_n'\boldsymbol{y} = n\overline{\boldsymbol{y}}$. Toinen väite seuraa.

2. Olkoot $\epsilon_{11}, \ldots, \epsilon_{1n_1}, \epsilon_{21}, \ldots, \epsilon_{2n_2}, \ldots, \epsilon_{p1}, \ldots, \epsilon_{pp_n} \backsim N(0, \sigma^2) \perp\!\!\!\perp$. Tällöin $Y_{ji} = \mu_j + \epsilon_{ji}$, kun $j \in \{1, \ldots, p\}$ ja $i \in \{1, \ldots, n_j\}$. Kun $i \in \{1, \ldots, p\}$, merkitään $\boldsymbol{Y}_i = \begin{bmatrix} Y_{i1} & \cdots & Y_{in_i} \end{bmatrix}'$ ja vastaavasti $\boldsymbol{\epsilon}_i = \begin{bmatrix} \epsilon_{i1} & \cdots & \epsilon_{in_i} \end{bmatrix}'$. Tällöin kaikilla $i \in \{1, \ldots, p\}$ pätee

$$\boldsymbol{Y}_i = \begin{bmatrix} Y_{i1} \\ \vdots \\ Y_{in_i} \end{bmatrix} = \begin{bmatrix} \mu_i + \epsilon_{i1} \\ \vdots \\ \mu_i + \epsilon_{in_i} \end{bmatrix} = \mu_i\mathbf{1}_{n_i} + \boldsymbol{\epsilon}_i,$$

ja edelleen

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \vdots \\ \boldsymbol{Y}_p \end{bmatrix} = \begin{bmatrix} \mu_1 \mathbf{1}_{n_1} + \boldsymbol{\epsilon}_1 \\ \vdots \\ \mu_p \mathbf{1}_{n_p} + \boldsymbol{\epsilon}_p \end{bmatrix} = \mu_1 \begin{bmatrix} \mathbf{1}_{n_1} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} + \ldots + \mu_p \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{1}_{n_p} \end{bmatrix} + \boldsymbol{\epsilon} = \underbrace{\begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{n_2} & \mathbf{0} & \cdots & \mathbf{0} \\ & & \vdots & & \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{1}_{n_p} \end{bmatrix}}_{=\boldsymbol{X}} \boldsymbol{\mu} + \boldsymbol{\epsilon},$$

missä $\boldsymbol{\epsilon} = \begin{bmatrix} \boldsymbol{\epsilon}_1 & \cdots & \boldsymbol{\epsilon}_p \end{bmatrix}'$ ja $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 & \cdots & \mu_p \end{bmatrix}'$. Matriisin $\boldsymbol{X}$ sarakkeet ovat lineaarisesti riippumattomat, joten $r(\boldsymbol{X}) = p$. Toisin sanoen matriisi $\boldsymbol{X}$ on täyttä sarakeastetta.

3. Suoraviivaisella laskulla nähdään, että

$$\begin{aligned}(n-1)s^2 &= \sum_{i=1}^{n} (y_i - \overline{y})^2 = \sum_{i=1}^{n} (y_i^2 - 2y_i\overline{y} + \overline{y}^2) = \sum_{i=1}^{n} y_i^2 - 2\overline{y} \sum_{i=1}^{n} y_i + n\overline{y}^2 \\ &= \sum_{i=1}^{n} y_i^2 - 2n\overline{y}^2 + n\overline{y}^2 = \sum_{i=1}^{n} y_i^2 - n\overline{y}^2.\end{aligned}$$

Toisen yhtälön näkemiseksi muokataan tässä esiintyvää viimeistä lauseketta, pitäen mielessä, että $\boldsymbol{y}'\mathbf{1}_n = (\boldsymbol{y}'\mathbf{1}_n)' = \mathbf{1}_n'\boldsymbol{y} = n\overline{y}$. Huomaa myös, että $1/n = (\mathbf{1}_n'\mathbf{1}_n)^{-1}$.

$$\begin{aligned}\sum_{i=1}^{n} y_i^2 - n\overline{y}^2 &= \boldsymbol{y}'\boldsymbol{y} - \frac{1}{n}(n\overline{y})^2 = \boldsymbol{y}'\boldsymbol{y} - \frac{1}{n}(\boldsymbol{y}'\mathbf{1}_n)(\mathbf{1}_n'\boldsymbol{y}) = \boldsymbol{y}'\boldsymbol{y} - (\boldsymbol{y}'\mathbf{1}_n)(\mathbf{1}_n'\mathbf{1}_n)^{-1}(\mathbf{1}_n'\boldsymbol{y}) \\ &= \boldsymbol{y}'(\boldsymbol{I}_n - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n')\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{J})\boldsymbol{y}.\end{aligned}$$

Mille hyvänsä kääntyvälle matriisille A pätee $(A^{-1})' = (A')^{-1}$ (koska $A'(A^{-1})' = (A^{-1}A)' = I_n = (AA^{-1})' = (A^{-1})'A'$). Täten

$$\boldsymbol{J}' = (\mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n')' = (\mathbf{1}_n')'((\mathbf{1}_n'\mathbf{1}_n)^{-1})'\mathbf{1}_n' = \mathbf{1}_n((\mathbf{1}_n'\mathbf{1}_n)')^{-1}\mathbf{1}_n' = \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n' = \boldsymbol{J}$$

eli $\boldsymbol{J}$ on symmetrinen. Lisäksi $\boldsymbol{J}$ on idempotentti:

$$\boldsymbol{J}^2 = (\mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n')(\mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n') = \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\underbrace{(\mathbf{1}_n'\mathbf{1}_n)(\mathbf{1}_n'\mathbf{1}_n)^{-1}}_{=\boldsymbol{I}_n}\mathbf{1}_n' = \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n' = \boldsymbol{J}.$$

4. Materiaalin sivulla 8 esiteltyjen normaaliyhtälöiden mukaan on siis $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{y}$. Nyt tarkasteltavassa mallissa matriisi $\boldsymbol{X}$ on muotoa $\boldsymbol{X} = \begin{bmatrix} \mathbf{1}_n & \boldsymbol{x} \end{bmatrix}$, missä $\boldsymbol{x} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}'$. Siis:

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \begin{bmatrix} \mathbf{1}_n' \\ \boldsymbol{x}' \end{bmatrix} \begin{bmatrix} \mathbf{1}_n & \boldsymbol{x} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n'\mathbf{1}_n & \mathbf{1}_n'\boldsymbol{x} \\ \boldsymbol{x}'\mathbf{1}_n & \boldsymbol{x}'\boldsymbol{x} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} n & n\overline{x} \\ n\overline{x} & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} n\beta_1 + n\overline{x}\beta_2 \\ n\overline{x}\beta_1 + \sum_{i=1}^n x_i^2\beta_2 \end{bmatrix};$$

$$\boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} \mathbf{1}'_n \\ \boldsymbol{x}' \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} \mathbf{1}'_n \boldsymbol{y} \\ \boldsymbol{x}'\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} n\overline{y} \\ \sum_{i=1}^n x_i y_i \end{bmatrix}.$$

Normaaliyhtälöt voidaan siis esittää komponenttimuodossa

$$\beta_1 + \overline{x}\beta_2 = \overline{y} \tag{1}$$

$$n\overline{x}\beta_1 + \sum_{i=1}^n x_i^2 \beta_2 = \sum_{i=1}^n x_i y_i \tag{2}$$

Kun yhtälöstä (1) ratkaistaan $\beta_1 = \overline{y} - \overline{x}\beta_2$ ja sijoitetaan tämä yhtälöön (2), saadaan

$$\begin{aligned} n\overline{x}(\overline{y} - \overline{x}\beta_2) + \sum_{i=1}^n x_i^2 \beta_2 = \sum_{i=1}^n x_i y_i &\implies (\sum_{i=1}^n x_i^2 - n\overline{x}^2)\beta_2 + n\overline{x}\,\overline{y} = \sum_{i=1}^n x_i y_i \\ &\implies \beta_2 = \frac{\sum_{i=1}^n x_i y_i - n\overline{x}\,\overline{y}}{\sum_{i=1}^n x_i^2 - n\overline{x}^2}. \end{aligned}$$

Tehtävän 3 perusteella $\sum_{i=1}^n x_i^2 - n\overline{x}^2 = \sum_{i=1}^n (x_i - \overline{x})^2$, joten yllä nimittäjä on haluttua muotoa. Muokataan osoittajaa:

$$\begin{aligned} \sum_{i=1}^n x_i y_i - n\overline{x}\,\overline{y} &= \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \overline{y} - \sum_{i=1}^n y_i \overline{x} + n\overline{x}\,\overline{y} = \sum_{i=1}^n (x_i y_i - x_i\overline{y} - y_i\overline{x} + \overline{x}\,\overline{y}) \\ &= \sum_{i=1}^n (x_i(y_i - \overline{y}) - \overline{x}(y_i - \overline{y})) = \sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y}). \end{aligned}$$

Koska PNS-estimaatit $\hat{\beta}_1$ ja $\hat{\beta}_2$ saadaan normaaliyhtälöiden ratkaisuina, on yllä todetun perusteella voimassa

$$\begin{aligned} \hat{\beta}_2 &= \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^n (x_i - \overline{x})^2} \quad ; \\ \hat{\beta}_1 &= \overline{y} - \overline{x}\hat{\beta}_2 \end{aligned}$$

Vektoreiden $\boldsymbol{x}$ ja $\boldsymbol{y}$ korrelaatiokerroin on

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^n (x_i - \overline{x})^2}\sqrt{\sum_{i=1}^n (y_i - \overline{y})^2}} = \frac{\sum_{i=1}^n (x_i - \overline{x})(x_i - \overline{y})}{\sqrt{n-1}s_x\sqrt{n-1}s_y},$$

joten nähdään heti, että

$$\hat{\beta}_2 = \frac{r_{xy}\sqrt{n-1}s_x\sqrt{n-1}s_y}{(n-1)s_x^2} = r_{xy}\frac{s_y}{s_x}.$$

5. Tehtävässä 3 osoitetun perusteella $n\overline{y}^2 = \boldsymbol{y}'\boldsymbol{J}\boldsymbol{y}$ (tämä käy ilmi tehtävässä 3 tehdyistä laskuista, mutta se näkyy myös kyseisen tehtävän toisesta yhtälöstä). Tehtävässä

3 näytettiin myös, että $\boldsymbol{J} = \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'$ on projektio (symmetrinen ja idempotentti). Jos tässä vektori $\mathbf{1}_n$ korvataan täyttä sarakeastetta olevalla $(n \times p)$-matriisilla $\boldsymbol{X}$, saadaan matriisi $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ joka edelleenkin on projektio (tämän näkee aivan vastaavasti kuin tehtävässä 3). Lisäksi $\boldsymbol{P}\boldsymbol{y} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{X}\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{y}}$. Koska mallissa on vakio, pätee $\overline{y} = (1/n)\sum_{i=1}^n \hat{y}_i$ kuten materiaalin sivulla 10 on todettu. Pitäen nämä tiedot mielessä havaitaan, että

$$
\begin{aligned}
SSR &= \sum_{i=1}^n (\hat{y}_i - \overline{y})^2 = \sum_{i=1}^n \hat{y}_i^2 - \sum_{i=1}^n 2\hat{y}_i\overline{y} + n\overline{y}^2 = \sum_{i=1}^n \hat{y}_i^2 - 2n\overline{y} + n\overline{y}^2 = \sum_{i=1}^n \hat{y}_i^2 - n\overline{y} \\
&= \hat{\boldsymbol{y}}'\hat{\boldsymbol{y}} - \boldsymbol{y}'\boldsymbol{J}\boldsymbol{y} = (\boldsymbol{P}\boldsymbol{y})'\boldsymbol{P}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{J}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}'\boldsymbol{P}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{J}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{J}\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{J})\boldsymbol{y}.
\end{aligned}
$$

Residuaali $\hat{\boldsymbol{\epsilon}}$ voidaan kirjoittaa muodossa $\hat{\boldsymbol{\epsilon}} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{P}\boldsymbol{y} = (\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{y}$. Koska $\boldsymbol{P}$ on projektio, on myös $\boldsymbol{I}_n - \boldsymbol{P}$ projektio. Täten

$$
SSE = \sum_{i=1}^n \hat{\epsilon}_i = \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}} = ((\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{y})'(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P})'(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{y}.
$$

Nyt saadaan vaihtoehtoinen perustelu vakiomallin yhtälölle $SST = SSR + SSE$:

$$
\begin{aligned}
SSR + SSE &= \boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{J})\boldsymbol{y} + \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{P} - \boldsymbol{J} + \boldsymbol{I}_n - \boldsymbol{P})\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{J})\boldsymbol{y} \\
&= \sum_{i=1}^n (y_i - \overline{y}) = SST.
\end{aligned}
$$