Lineaariset mallit, kevät 2014
Ratkaisut 1
Juho Leppänen

1. Ensimmäinen yhtälö

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

ilmaisee sen väittämän tai olettamuksen, että satunnaisvektori $\boldsymbol{Y}$ saadaan lisäämällä satunnaisvektoriin $\boldsymbol{\epsilon}$ kiinteä vektori $\boldsymbol{X}\boldsymbol{\beta}$ (jossa tosin $\boldsymbol{\beta}$ on parametrina tuntematon). Koska $\boldsymbol{X}$ on $(n \times p)$-matriisi ja $\boldsymbol{\beta}$ on $(p \times 1)$ vektori, on niiden tulo $\boldsymbol{X}\boldsymbol{\beta}$ $(n \times 1)$-vektori. Myös $\boldsymbol{Y}$ on kahden $(n \times 1)$-vektorin summana dimensiota $(n \times 1)$.

Kun lausekkeessa $\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ kirjoitetaan kunkin esiintyvän matriisin komponentit näkyviin, saadaan toinen yhtälö eli

$$\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}.$$

Merkitään $\boldsymbol{x}_k = [x_{k1} \cdots x_{kp}]'$, kun $k = 1, \ldots, n$. Tällöin $\boldsymbol{x}_i'$ on siis matriisin $\boldsymbol{X}$ i. rivi. Kun $(n \times p)$-matriisi $\boldsymbol{X}$ ja $(p \times 1)$-vektori $\boldsymbol{\beta}$ kerrotaan keskenään, syntyy $(n \times 1)$-vektori, jonka k. rivillä esiintyy matriisin $\boldsymbol{X}$ k. rivin ja vektorin $\boldsymbol{\beta}$ välinen sisätulo

$$\boldsymbol{x}_k'\boldsymbol{\beta} = [x_{k1} \cdots x_{kp}] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} = \sum_{i=1}^{p} x_{ki}\beta_i.$$

Tämä tieto on perusteena seuraaville kahdelle yhtälölle

$$\begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{p} x_{1i}\beta_i \\ \vdots \\ \sum_{i=1}^{p} x_{ni}\beta_i \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1'\boldsymbol{\beta} \\ \vdots \\ \boldsymbol{x}_n'\boldsymbol{\beta} \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Koska yllä esiintyvän matriisin $[\boldsymbol{x}_1'\boldsymbol{\beta} \cdots \boldsymbol{x}_n'\boldsymbol{\beta}]'$ i. rivillä esiintyy vektoreiden $\boldsymbol{x}_i$ ja $\boldsymbol{\beta}$ välinen sisätulo, on yllä todetun matriisitulon karakterisaation nojalla siis

$$\begin{bmatrix} \boldsymbol{x}_1'\boldsymbol{\beta} \\ \vdots \\ \boldsymbol{x}_n'\boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1' \\ \vdots \\ \boldsymbol{x}_n' \end{bmatrix} \boldsymbol{\beta},$$

mistä saadaan tehtävänannon kolmanneksi viimeinen yhtälö. Merkitään edelleen matriisin $\boldsymbol{X}$ i. saraketta symbolilla $\boldsymbol{x}_{(i)}$. Tällöin $\boldsymbol{X}$:lle saadaan esitys sarakkeiden avulla: $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_{(1)} & \cdots & \boldsymbol{x}_{(p)} \end{bmatrix}$. Kun X:n riviesitys vaihdetaan tähän sarake-esitykseen, saadaan toiseksi viimeinen yhtälö:

$$\begin{bmatrix} \boldsymbol{x}_1' \\ \vdots \\ \boldsymbol{x}_n' \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\epsilon} = \begin{bmatrix} \boldsymbol{x}_{(1)} & \cdots & \boldsymbol{x}_{(p)} \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Viimeinen yhtälö ilmaisee matriisin $\boldsymbol{X}$ ja vektorin $\boldsymbol{\beta}$ tulon toisella tavalla: matriisin $\boldsymbol{X}$ ja vektorin $\boldsymbol{\beta}$ tulo saadaan kertomalla $\boldsymbol{X}$:n i. sarake vektorin $\boldsymbol{\beta}$ i. skalaarikomponentilla β_i ja summaamalla yli kaikkien näiden:

$$\begin{aligned}\boldsymbol{X\beta} &= \begin{bmatrix} \sum_{i=1}^p x_{1i}\beta_i \\ \vdots \\ \sum_{i=1}^p x_{ni}\beta_i \end{bmatrix} = \begin{bmatrix} x_{11}\beta_1 \\ \vdots \\ x_{n1}\beta_1 \end{bmatrix} + \cdots + \begin{bmatrix} x_{1p}\beta_p \\ \vdots \\ x_{np}\beta_p \end{bmatrix} = \begin{bmatrix} x_{11} \\ \vdots \\ x_{n1} \end{bmatrix} \beta_1 + \cdots + \begin{bmatrix} x_{1p} \\ \vdots \\ x_{np} \end{bmatrix} \beta_p \\ &= \boldsymbol{x}_{(1)}\beta_1 + \cdots + \boldsymbol{x}_{(p)}\beta_p.\end{aligned}$$

2. Vektorit $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ ovat lineaarisesti riippumattomat jos ja vain jos kaikilla $c_1, \ldots, c_n \in \mathbb{R}$:

$$\sum_{i=1}^n c_i \boldsymbol{x}_1 = 0 \Rightarrow c_1 = \cdots = c_n = 0$$

Palautetaan mieleen lineaarialgebrasta se seikka, että jos A ja B ovat kaksi matriisia, ja jos B saadaan A:sta soveltamalla peräkkäin äärellinen määrä alkeisrivioperaatioita, niin matriiseja A ja B vastaavilla yhtälöryhmillä on täsmälleen samat ratkaisut.

(a) Tavoitteena on löytää luvut c_1, c_2, c_3, joille

$$0 = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + c_3\boldsymbol{x}_3 = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 & -13 \\ 2 & 3 & -1 \\ 1 & 2 & 2 \end{bmatrix}}_{=\boldsymbol{X}} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

ja joista ainakin yksi on nollasta poikkeava. Muokataan matriisia $\boldsymbol{X}$ alkeisrivioperaatioilla sopivaan muotoon:

$$\begin{aligned}&\begin{bmatrix} 1 & -1 & -13 \\ 2 & 3 & -1 \\ 1 & 2 & 2 \end{bmatrix} \overset{(1)}{\to} \begin{bmatrix} 1 & -1 & -13 \\ 2 & 3 & -1 \\ 0 & 3 & 15 \end{bmatrix} \overset{(2)}{\to} \begin{bmatrix} 1 & -1 & -13 \\ 0 & 5 & 25 \\ 0 & 3 & 15 \end{bmatrix} \overset{(3)}{\to} \begin{bmatrix} 1 & -1 & -13 \\ 0 & 1 & 5 \\ 0 & 3 & 15 \end{bmatrix} \\ &\overset{(4)}{\to} \underbrace{\begin{bmatrix} 1 & -1 & -13 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}}_{=A}\end{aligned}$$

Alkeisrivioperaatiot:
(1) Vähennetään rivistä 3 rivi 1.
(2) Vähennetään rivistä 2 rivi 1 kerrottuna skalaarilla 2.
(3) Kerrotaan riviä 2 skalaarilla 1/5.
(4) Vähennetään rivistä 3 rivi 2 kerrottuna skalaarilla 3.

Merkitään $c_3 = t$. Tällöin

$$0 = A \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}' \iff \begin{cases} c_1 - c_2 - 13c_3 & = 0 \\ c_2 + 5c_3 & = 0 \\ 0 & = 0 \end{cases} \iff \begin{cases} c_1 & = 8t \\ c_2 & = -5t \\ c_3 & = t \end{cases},$$

missä t voi olla mitä vain. Valitaan esim. $t = 1$, jolloin $c_1 = 8$, $c_2 = -5$, $c_3 = 1$ ja siis $0 = X \begin{bmatrix} 8 & -5 & 1 \end{bmatrix}' = 8\boldsymbol{x}_1 + -5\boldsymbol{x}_2 + \boldsymbol{x}_3$. Nähdään myös, että $\boldsymbol{x}_3 = 5\boldsymbol{x}_2 - 8\boldsymbol{x}_1$.

(b) Etsitään ensin luvut c_1, c_2, c_3, joille

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + c_3\boldsymbol{x}_4 = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Ratkaistavaksi tuleva yhtälöryhmä on matriisimuodossa

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & a \\ 2 & 3 & 1 & b \\ 1 & 2 & 0 & c \end{array}\right]. \tag{1}$$

Soveltamalla tähän alkeisrivioperaatioita saadaan:

$$\begin{aligned}
\left[\begin{array}{ccc|c} 1 & -1 & 1 & a \\ 2 & 3 & 1 & b \\ 1 & 2 & 0 & c \end{array}\right] &\overset{(1)}{\to} \left[\begin{array}{ccc|c} 1 & -1 & 1 & a \\ 2 & 3 & 1 & b \\ 0 & 3 & -1 & c-a \end{array}\right] \overset{(2)}{\to} \left[\begin{array}{ccc|c} 1 & -1 & 1 & a \\ 2 & 0 & 2 & b-(c-a) \\ 0 & 3 & -1 & c-a \end{array}\right] \\
&\overset{(3)}{\to} \left[\begin{array}{ccc|c} 0 & -1 & 0 & a - \frac{1}{2}(b-c+a) \\ 2 & 0 & 2 & b-c+a \\ 0 & 3 & -1 & c-a \end{array}\right] \overset{(4)}{\to} \left[\begin{array}{ccc|c} 0 & -1 & 0 & \frac{1}{2}(a-b+c) \\ 2 & 0 & 2 & b-c+a \\ 0 & 0 & -1 & c-a+\frac{3}{2}(a-b+c) \end{array}\right] \\
&\overset{(5)}{\to} \underbrace{\left[\begin{array}{ccc|c} 0 & -1 & 0 & \frac{1}{2}(a-b+c) \\ 2 & 0 & 0 & 2a+4c-2b \\ 0 & 0 & -1 & \frac{1}{2}(a-3b+5c) \end{array}\right]}_{=B}
\end{aligned}$$

Alkeisrivioperaatiot:
(1) Rivistä 3 vähennetään rivi 1
(2) Rivistä 2 vähennetään rivi 3
(3) Rivistä 1 vähennetään rivi 2 kerrottuna skalaarilla $\frac{1}{2}$
(4) Riviin 3 lisätään rivi 1 kerrottuna skalaarilla 3.

(5) Riviin 2 lisätään rivi 3 kerrottuna skalaarilla 2.

Tässä matriisia B vastaavalla yhtälöryhmällä on yksikäsitteinen ratkaisu $c_1 = a + 2c - b$, $c_2 = (1/2)(b - a - c)$ ja $c_3 = (1/2)(3b - a - 5c)$. Siis myös matriisia (1) vastaavalla yhtälöryhmällä on tämä ainoana ratkaisuna. Havaitaan, että saadaan seurauksena vektoreiden $\boldsymbol{x}_1, \boldsymbol{x}_2$ ja $\boldsymbol{x}_4$ lineaarinen riippumattomuus; nimittäin jos yllä valitaan $a = b = c = 0$, on edellä todetun nojalla yhtälöllä $0 = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + c_3\boldsymbol{x}_4$ vain yksi ratkaisu, jonka täytyy olla $c_1 = c_2 = c_3 = 0$.

3. Olkoon $\boldsymbol{\epsilon} \backsim N_5(0, \sigma^2\boldsymbol{I})$, jolloin siis $\epsilon_1, \ldots, \epsilon_5 \backsim N(0, \sigma^2) \perp\!\!\!\perp$ ja

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{bmatrix} = \begin{bmatrix} \beta_1 + \epsilon_1 \\ \beta_1 + \epsilon_2 \\ \beta_1 + 2\beta_2 + \epsilon_3 \\ \beta_1 - \beta_2 + \epsilon_4 \\ \beta_1 - \beta_2 + \epsilon_5 \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ 0 \\ 2 \\ -1 \\ -1 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 2 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\epsilon}.$$

Siis malli voidaan esittää muodossa $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$, kun valitaan

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & -1 & -1 \end{bmatrix}'$$

4. Olkoon $\boldsymbol{\epsilon} \backsim N_7(0, \sigma^2\boldsymbol{I})$, jolloin siis $\epsilon_1, \ldots, \epsilon_7 \backsim N(0, \sigma^2) \perp\!\!\!\perp$ ja

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_7 \end{bmatrix} = \begin{bmatrix} \beta_1\mathbf{1}_2 \\ 2\beta_1\mathbf{1}_2 \\ \beta_2\mathbf{1}_3 \end{bmatrix} + \boldsymbol{\epsilon} = \beta_1 \begin{bmatrix} \mathbf{1}_2 \\ 2\mathbf{1}_2 \\ \mathbf{0} \end{bmatrix} + \beta_2 \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1}_3 \end{bmatrix} + \boldsymbol{\epsilon} = \begin{bmatrix} \mathbf{1}_2 & \mathbf{0} \\ 2\mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\epsilon}.$$

Siis malli voidaan esittää muodossa $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$, kun valitaan

$$\boldsymbol{X} = \begin{bmatrix} \mathbf{1}_2 & \mathbf{0} \\ 2\mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}'$$

5. Olkoon $\boldsymbol{\epsilon} \backsim N_{15}(0, \sigma^2\boldsymbol{I})$, jolloin siis $\epsilon_1, \ldots, \epsilon_{15} \backsim N(0, \sigma^2) \perp\!\!\!\perp$ ja

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_{15} \end{bmatrix} = \begin{bmatrix} \beta_1\mathbf{1}_5 \\ (\beta_1 + \beta_2)\mathbf{1}_5 \\ (\beta_1 - \beta_2)\mathbf{1}_5 \end{bmatrix} + \boldsymbol{\epsilon} = \beta_1\mathbf{1}_{15} + \beta_2 \begin{bmatrix} \mathbf{0} \\ \mathbf{1}_5 \\ -\mathbf{1}_5 \end{bmatrix} + \boldsymbol{\epsilon} = \begin{bmatrix} \mathbf{1}_5 & \mathbf{0} \\ \mathbf{1}_5 & \mathbf{1}_5 \\ \mathbf{1}_5 & -\mathbf{1}_5 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \boldsymbol{\epsilon}.$$

Täten asettamalla

$$
\boldsymbol{X} = \begin{bmatrix} \mathbf{1}_5 & \mathbf{0} \\ \mathbf{1}_5 & \mathbf{1}_5 \\ \mathbf{1}_5 & -\mathbf{1}_5 \end{bmatrix},
$$

saadaan malli esitettyä muodossa $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$.