Jyrki Möttönen

Matriisilaskentaa tilastotieteilijöille

Helsingin yliopisto
Sosiaalitieteiden laitos
Kevät 2015

Sisältö

Luku 1

Matriisialgebran alkeita

1.1 Notaatioita

Matriiseja merkitään yleensä vahvennetuilla isoilla kirjaimilla. Esimerkiksi $m \times n$ matriisi

$$\boldsymbol{A} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

on matriisi, jolla on m riviä ja n saraketta.

Matriisin $\boldsymbol{A}$ i:nnen rivin j:nnen sarakkeen alkio on $a_{ij} = (\boldsymbol{A})_{ij}$.

Matriisin $\boldsymbol{A}$ i:s rivi on $(\boldsymbol{A})_{i\cdot}$ ja j:s sarake on $(\boldsymbol{A})_{\cdot j}$

Pystyvektori on matriisi, jolla on ainoastaan yksi sarake. Esim.

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Vaakavektori on matriisi, jolla on ainoastaan yksi rivi. Esim.

$$\begin{pmatrix} c_1 & c_2 & \cdots & c_n \end{pmatrix}.$$

Tällä kurssilla pystyvektoria merkitään vahvennetulla pienellä kirjaimella. Pystyvektorista saadaan vaakavektori transpoosin avulla (ks. kappale 1.3).

Yksittäistä numeroa kutsutaan *skalaariksi*. Voidaan ajatella myös, että skaalaari on 1×1 matriisi.

1.2 Matriisien yhteenlasku ja kertolasku

Määritelmä 1. *Olkoon* $\boldsymbol{A}$ *ja* $\boldsymbol{B}$ $m \times n$ *matriiseja. Tällöin matriisien* $\boldsymbol{A}$ *ja* $\boldsymbol{B}$ *summa on*

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij}).$$

Huom! Yhteenlaskettavat matriisit täytyy olla samankokoisia eli sekä $\boldsymbol{A}$ että $\boldsymbol{B}$ täytyy olla $m \times n$ matriisi.

Määritelmä 2. *Olkoon* $\boldsymbol{A}$ $m \times p$ *matriisi ja* $\boldsymbol{B}$ $p \times n$ *matriisi. Tällöin matriisien* $\boldsymbol{A}$ *ja* $\boldsymbol{B}$ *tulo on*

$$\boldsymbol{AB} = (c_{ij}),$$

jossa

$$c_{ij} = (\boldsymbol{A})_{i\cdot}(\boldsymbol{B})_{\cdot j} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Huom! Matriisien tulossa vasemmampuoleisella matriisilla täytyy olla sarakkeita yhtä paljon kuin oikeanpuoleisella matriisilla rivejä, jotta operaatio olisi määritelty.

Määritelmä 3. *Skalaarin* α *ja matriisin* $\boldsymbol{A}$ *tulo on*

$$\alpha \boldsymbol{A} = \boldsymbol{A} \alpha = (\alpha a_{ij}).$$

Lause 1. *Olkoon* α *ja* β *skalaareja ja* $\boldsymbol{A}$, $\boldsymbol{B}$ *ja* $\boldsymbol{C}$ *matriiseja. Tällöin, laskuoperaatioiden ollessa määriteltyjä,*

(a) $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$

(b) $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$

(c) $\alpha(\boldsymbol{A} + \boldsymbol{B}) = \alpha\boldsymbol{A} + \alpha\boldsymbol{B}$

(d) $(\alpha + \beta)\boldsymbol{A} = \alpha\boldsymbol{A} + \beta\boldsymbol{A}$

(e) $\boldsymbol{A} - \boldsymbol{A} = \boldsymbol{A} + (-\boldsymbol{A}) = (0)$

(f) $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$

(g) $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$

(h) $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$

Huom! Yleensä $\boldsymbol{AB} \neq \boldsymbol{BA}$ eli skalaaritulon tuttu ominaisuus $\alpha\beta = \beta\alpha$ ei ole yleisesti voimassa matriisien kertolaskussa!!

```
> A<-matrix(c(1,-1,2,3),2,2)
> B<-matrix(c(3,6,2,3),2,2)
> C<-matrix(c(1,2,3,-1),2,2)
> A
     [,1] [,2]
[1,]    1    2
[2,]   -1    3
> B
     [,1] [,2]
[1,]    3    2
[2,]    6    3
> C
     [,1] [,2]
[1,]    1    3
[2,]    2   -1
> A+B
     [,1] [,2]
[1,]    4    4
[2,]    5    6
> 2*(A+B)
     [,1] [,2]
[1,]    8    8
[2,]   10   12
> A%*%B
     [,1] [,2]
[1,]   15    8
[2,]   15    7
> A%*%B%*%C
     [,1] [,2]
[1,]   31   37
[2,]   29   38
> (A%*%B)%*%C
     [,1] [,2]
[1,]   31   37
[2,]   29   38
> A%*%(B%*%C)
     [,1] [,2]
[1,]   31   37
[2,]   29   38
> B%*%C%*%A
     [,1] [,2]
[1,]    0   35
[2,]   -3   69
```

1.3 Matriisin transpoosi

Olkoon $\boldsymbol{A}$ $m \times n$ matriisi. Tällöin matriisin $\boldsymbol{A}$ *transpoosi* $\boldsymbol{A}'$ on matriisi, jonka riveinä ovat matriisin $\boldsymbol{A}$ sarakkeet ("käännettynä 90 astetta vastapäivään") järjestys säilyttäen. Toisin sanottuna

$$\boldsymbol{A}' = (a_{ij})' = (a_{ji})$$

Matriisin $\boldsymbol{A}$ transpoosi on siis $n \times m$ matriisi eli rivien ja sarakkeiden lukumäärät vaihtavat paikkaa.

Lause 2. *Olkoon α ja β skalaareja ja $\boldsymbol{A}$ ja $\boldsymbol{B}$ matriiseja. Tällöin, laskuoperaatioiden ollessa määriteltyjä,*

(a) $(\alpha\boldsymbol{A})' = \alpha\boldsymbol{A}'$.

(b) $(\boldsymbol{A}')' = \boldsymbol{A}$.

(c) $(\alpha\boldsymbol{A} + \beta\boldsymbol{B})' = \alpha\boldsymbol{A}' + \beta\boldsymbol{B}'$.

(d) $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$.

Esimerkki 1.3.1
Matriisin

$$\boldsymbol{A} = \begin{pmatrix} 2 & 5 \\ 3 & 7 \\ 4 & 9 \end{pmatrix}$$

transpoosi on

$$\boldsymbol{A}' = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 7 & 9 \end{pmatrix}.$$

```
> A<-matrix(c(2,3,4,5,7,9),3,2)
> A
     [,1] [,2]
[1,]    2    5
[2,]    3    7
[3,]    4    9
> t(A)
     [,1] [,2] [,3]
[1,]    2    3    4
[2,]    5    7    9
```

Esimerkki 1.3.2
Pystyvektorin

$$\boldsymbol{b} = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}$$

transpoosi on vaakavektori

$$\boldsymbol{b}' = \begin{pmatrix} 1 & 4 & 7 \end{pmatrix}.$$

1.4 Matriisin jälki (trace)

Olkoon $A = (a_{ij})$ $n \times n$ neliömatriisi. Matriisin A jälki (eng. trace) on tällöin

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Matriisin jäljen ominaisuuksia:

Lause 3. *Olkoon α skalaari ja $\boldsymbol{A}$ ja $\boldsymbol{B}$ matriiseja. Tällöin, laskuoperaatioiden ollessa määriteltyjä,*

(a) $tr(\boldsymbol{A}') = tr(\boldsymbol{A})$.

(b) $tr(\alpha\boldsymbol{A}) = \alpha tr(\boldsymbol{A})$.

(c) $tr(\boldsymbol{A} + \boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$.

(d) $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$.

(e) $tr(\boldsymbol{A}'\boldsymbol{A}) = 0,$ *jos ja vain jos* $\boldsymbol{A} = (0)$.

1.5 Determinantti

Määritelmä 4. *Olkoon $\boldsymbol{A}$ $m \times m$ neliömatriisi. Tällöin matriisin $\boldsymbol{A}$ determinantti on*

$$\begin{aligned} |\boldsymbol{A}| = \det(\boldsymbol{A}) = &= \sum (-1)^{f(i_1,\ldots,i_m)} a_{1i_1} a_{2i_2} \cdots a_{mi_m} \\ &= \sum (-1)^{f(i_1,\ldots,i_m)} a_{i_1 1} a_{i_2 2} \cdots a_{i_m m}, \end{aligned}$$

jossa summaus otetaan yli kaikkien joukon $(1, \ldots, m)$ permutaatioiden $(i_1, \ldots, i_m)$ ja funktio $f(i_1, \ldots, i_m)$ on transpositioiden lukumäärä, joka tarvitaan lukujonon $(i_1, \ldots, i_m)$ muuttamiseksi lukujonoksi $(1, \ldots, m)$.

Transpositiolla tarkoitetaan kahden lukujonon luvun paikkojen vaihtamista. Esimerkiksi lukujonosta $(2, 3, 1)$ saadaan transpositioiden avulla $(1, 2, 3)$ seuraavasti:

$$(2, 3, 1) \longrightarrow (1, 3, 2) \longrightarrow (1, 2, 3).$$

Tarvittavien transpositioiden lukumäärä ei ole yksikäsitteinen mutta se on aina yksikäsitteisesti joko parillinen tai pariton.

Lause 4. *Olkoon $\boldsymbol{A}$ $m \times m$ neliömatriisi ja olkoon $A_{ij} = (-1)^{i+j} m_{ij}$ matriisin $\boldsymbol{A}$ alkiota a_{ij} vastaava kofaktori, jossa m_{ij} on determinantti $(m-1) \times (m-1)$ matriisista, joka saadaan poistamalla matriisin $\boldsymbol{A}$ i:s rivi ja j:s sarake.*

(a) Kun $m = 1$ (eli kyseessä on skalaari), niin

$$|\boldsymbol{A}| = |a_{11}| = a_{11}.$$

(b) Kun $m = 2$, niin

$$|\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(c) Kun $m \geq 2$, niin

$$|\boldsymbol{A}| = \sum_{j=1}^{m} a_{ij} A_{ij},$$

(d) Kun $m \geq 2$, niin

$$|\boldsymbol{A}| = \sum_{j=1}^{m} a_{ji} A_{ji},$$

Esimerkki 1.5.1

(a)

$$\begin{vmatrix} 4 & 6 \\ 3 & 7 \end{vmatrix} = 4 \cdot 7 - 6 \cdot 3 = 28 - 18 = 10.$$

```
> A<-matrix(c(4,3,6,7),2,2)
> A
     [,1] [,2]
[1,]    4    6
[2,]    3    7
> det(A)
[1]10
```

(b)

$$\begin{aligned}
\begin{vmatrix} 4 & 6 & 2 \\ 3 & 7 & 1 \\ 5 & 9 & 8 \end{vmatrix} &= 4(-1)^{1+1} \begin{vmatrix} 7 & 1 \\ 9 & 8 \end{vmatrix} + 6(-1)^{1+2} \begin{vmatrix} 3 & 1 \\ 5 & 8 \end{vmatrix} + 2(-1)^{1+3} \begin{vmatrix} 3 & 7 \\ 5 & 9 \end{vmatrix} \\
&= 4 \begin{vmatrix} 7 & 1 \\ 9 & 8 \end{vmatrix} - 6 \begin{vmatrix} 3 & 1 \\ 5 & 8 \end{vmatrix} + 2 \begin{vmatrix} 3 & 7 \\ 5 & 9 \end{vmatrix} \\
&= 4(7 \cdot 8 - 1 \cdot 9) - 6(3 \cdot 8 - 1 \cdot 5) + 2(3 \cdot 9 - 7 \cdot 5) = 58.
\end{aligned}$$

```
> A<-matrix(c(4,3,5,6,7,9,2,1,8),3,3)
> A
     [,1] [,2] [,3]
[1,]    4    6    2
[2,]    3    7    1
[3,]    5    9    8
> det(A)
[1] 58
```

Lause 5. *Olkoon α skalaari ja $\boldsymbol{A}$ $m \times m$ matriisi. Tällöin*

(a) $|\boldsymbol{A}'| = |\boldsymbol{A}|$.

(b) $|\alpha \boldsymbol{A}| = \alpha^m |\boldsymbol{A}|$.

(c) Jos $\boldsymbol{A}$ on diagonaalimatriisi, niin

$$|\boldsymbol{A}| = |diag(a_{11}, \ldots, a_{mm})| = \prod_{i=1}^{m} a_{ii}.$$

(d) Jos matriisin $\boldsymbol{A}$ jollain rivillä (tai sarakkeella) kaikki alkiot ovat nollia, niin $|\boldsymbol{A}| = 0$.

(e) Jos matriisin $\boldsymbol{A}$ kaksi riviä (tai saraketta) ovat lineaarisesti riippuvia, niin $|\boldsymbol{A}| = 0$.

(f) Matriisin $\boldsymbol{A}$ kahden rivin (tai sarakkeen) paikkojen vaihtaminen keskenään vaihtaa determinantin $|\boldsymbol{A}|$ etumerkin.

(g) Jos matriisin $\boldsymbol{A}$ yhden rivin (tai sarakkeen) alkiot kerrotaan vakiolla α, niin determinantti tulee kerrottua myös vakiolla α.

(h) Matriisin $\boldsymbol{A}$ determinantti pysyy muuttumattomana, jos matriisin riviin (tai sarakkeeseen) lisätään jonkin toisen rivin (tai sarakkeen) monikerta.

Lause 6. *Olkoon $\boldsymbol{A}$ ja $\boldsymbol{B}$ $m \times m$ neliömatriiseja. Tällöin*

$$|\boldsymbol{AB}| = |\boldsymbol{BA}| = |\boldsymbol{A}||\boldsymbol{B}|.$$

1.6 Käänteismatriisi

Esimerkki 1.6.1
Halutaan sovittaa aineistoon lineaarinen regressiomalli

$$y_i = x_{i1}\beta_1 + \ldots + x_{ip}\beta_p + \epsilon_i,$$

jossa y_i on havaittu arvo, x_{ij} ovat tunnettuja selittävän muuttujan j arvoja, β_j on j:nnen selittävän muuttujan regressiokerroin ja ϵ_j on virhetermi. Edellinen malli voidaan ilmaista matriisien avulla muodossa

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

jossa $\boldsymbol{y} = (y_1, \ldots, y_n)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$, $\boldsymbol{X} = (\boldsymbol{x}_1 \ \ldots \ \boldsymbol{x}_n)'$ ja $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)'$. Parametrien β_j estimoimiseen käytetään pienimmän neliösumman menetelmää, jolloin halutaan minimoida jäännösneliösumma

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2$$

parametrivektorin $\boldsymbol{\beta}$ suhteen. Derivoimalla $S(\boldsymbol{\beta})$ parametrivektorin $\boldsymbol{\beta}$ suhteen ja asettamalla derivaatta nollaksi, saadaan ns. normaaliyhtälö

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{y}.$$

Jos on olemassa matriisi $\boldsymbol{A}$, jolle pätee $\boldsymbol{A}(\boldsymbol{X}'\boldsymbol{X}) = \boldsymbol{I}_p$, niin kertomalla edellinen yhtälö vasemmalta puolittain matriisilla $\boldsymbol{A}$, saadaan

$$\boldsymbol{A}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{A}\boldsymbol{X}'\boldsymbol{y} \Leftrightarrow \boldsymbol{I}_p\boldsymbol{\beta} = \boldsymbol{A}\boldsymbol{X}'\boldsymbol{y} \Leftrightarrow \boldsymbol{\beta} = \boldsymbol{A}\boldsymbol{X}'\boldsymbol{y}.$$

Kyseistä matriisia $\boldsymbol{A}$ kutsutaan matriisin $\boldsymbol{B} = \boldsymbol{X}'\boldsymbol{X}$ käänteismatriisiksi.

Esimerkki 1.6.2
Halutaan ratkaista yhtälöryhmä

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = c_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = c_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = c_3 \end{cases}$$

muuttujien x_1, x_2, x_3 suhteen. Edellinen yhtälöryhmä voidaan kirjoittaa matriisien avulla seuraavassa muodossa:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c},$$

jossa

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ ja } \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Jos on olemassa matriisi $\boldsymbol{A}^{-1}$, jolle pätee $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_3$, niin kertomalla edellinen yhtälö vasemmalta puolittain matriisilla $\boldsymbol{A}^{-1}$, saadaan

$$\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{c} \Leftrightarrow \boldsymbol{I}_3\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{c} \Leftrightarrow \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{c}.$$

Määritelmä 5. *Neliömatriisi $\boldsymbol{A}$ on <u>ei-singulaarinen</u>, jos $|\boldsymbol{A}| \neq 0$ ja <u>singulaarinen</u>, jos $|\boldsymbol{A}| = 0$.*

Olkoon $\boldsymbol{A}$ ei-singulaarinen $m \times m$ neliömatriisi. Tällöin on olemassa ei-singulaarinen $m \times m$ neliömatriisi $\boldsymbol{A}^{-1}$, jolle pätee

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_m.$$

Matriisia $\boldsymbol{A}^{-1}$ kutsutaan matriisin $\boldsymbol{A}$ *käänteismatriisiksi.*

Lause 7. *Olkoon $\alpha \neq 0$ skalaari ja $\boldsymbol{A}$ ja $\boldsymbol{B}$ ei-singulaarisia $m \times m$ neliömatriiseja. Tällöin*

(a) $(\alpha\boldsymbol{A})^{-1} = \alpha^{-1}\boldsymbol{A}^{-1}$

(b) $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$

(c) $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$

(d) $|\boldsymbol{A}^{-1}| = |\boldsymbol{A}|^{-1}$

(e) Jos $\boldsymbol{A} = diag(a_{11}, \ldots, a_{mm})$, *niin* $\boldsymbol{A}^{-1} = diag(a_{11}^{-1}, \ldots, a_{mm}^{-1})$

(f) Jos $\boldsymbol{A} = \boldsymbol{A}'$, *niin* $\boldsymbol{A}^{-1} = (\boldsymbol{A}^{-1})'$

(g) $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

Käänteismatriisi voidaan laskea determinantin tapaan kofaktorien avulla seuraavasti:

$$\boldsymbol{A}^{-1} = |\boldsymbol{A}|^{-1} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{m1} \\ A_{12} & A_{22} & \cdots & A_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1m} & A_{2m} & \cdots & A_{mm} \end{pmatrix}.$$

Esimerkki 1.6.3
Matriisin

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 10 & 5 \end{pmatrix}$$

käänteismatriisi on

$$\boldsymbol{A}^{-1} = \frac{1}{1 \cdot 5 - 2 \cdot 10} \begin{pmatrix} 5 & -2 \\ -10 & 1 \end{pmatrix} = -\frac{1}{15} \begin{pmatrix} 5 & -2 \\ -10 & 1 \end{pmatrix} = \begin{pmatrix} -1/3 & 2/15 \\ 2/3 & -1/15 \end{pmatrix}.$$

```
> A<-matrix(c(1,10,2,5),2,2)
> A
     [,1] [,2]
[1,]    1    2
[2,]   10    5
> invA<-solve(A)
> invA
           [,1]        [,2]
[1,] -0.3333333  0.13333333
[2,]  0.6666667 -0.06666667
> -15*invA
     [,1] [,2]
[1,]    5   -2
[2,]  -10    1
> invA%*%A
              [,1]         [,2]
[1,]  1.000000e+00 2.775558e-17
[2,] -2.775558e-17 1.000000e+00
```

Esimerkki 1.6.4
Olkoon

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 10 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Tällöin

$$|\boldsymbol{A}| = 1\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2\begin{vmatrix} 10 & 6 \\ 7 & 9 \end{vmatrix} + 3\begin{vmatrix} 10 & 5 \\ 7 & 8 \end{vmatrix} = 1(-3) - 2 \cdot 48 + 3 \cdot 45 = 36$$

ja

$$\boldsymbol{A}^{-1} = \frac{1}{36}\begin{pmatrix} \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \\ -\begin{vmatrix} 10 & 6 \\ 7 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 10 & 6 \end{vmatrix} \\ \begin{vmatrix} 10 & 5 \\ 7 & 8 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 10 & 5 \end{vmatrix} \end{pmatrix} = \frac{1}{36}\begin{pmatrix} -3 & 6 & -3 \\ -48 & -12 & 24 \\ 45 & 6 & -15 \end{pmatrix}.$$

```
> A<-matrix(c(1,10,7,2,5,8,3,6,9),3,3)
> A
     [,1] [,2] [,3]
[1,]    1    2    3
```

```
[2,]   10    5    6
[3,]    7    8    9
> det(A)
[1] 36
> invA<-solve(A)
> invA
            [,1]       [,2]        [,3]
[1,] -0.08333333  0.1666667 -0.08333333
[2,] -1.33333333 -0.3333333  0.66666667
[3,]  1.25000000  0.1666667 -0.41666667
> invA%*%A
              [,1]         [,2]          [,3]
[1,]  1.000000e+00 3.885781e-16  4.163336e-16
[2,] -7.771561e-16 1.000000e+00 -9.992007e-16
[3,]  1.665335e-16 2.498002e-16  1.000000e+00
> det(A)*invA
     [,1] [,2] [,3]
[1,]   -3    6   -3
[2,]  -48  -12   24
[3,]   45    6  -15
```

1.7 Harjoituksia

1.1
Olkoon

$$\boldsymbol{A} = \begin{pmatrix} 2 & 4 & 3 \\ 8 & 1 & 4 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \\ 8 & 3 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{C} = \begin{pmatrix} 5 & 1 & 2 \\ 6 & 7 & 5 \end{pmatrix}.$$

Laske seuraavat summat

(a) $\boldsymbol{A} + \boldsymbol{B}$

(b) $\boldsymbol{A} + \boldsymbol{C}$

(c) $\boldsymbol{B} + \boldsymbol{C}$

1.2
Olkoon

$$\boldsymbol{A} = \begin{pmatrix} 2 & 4 & 3 \\ 8 & 1 & 4 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \\ 8 & 3 \end{pmatrix}, \boldsymbol{C} = \begin{pmatrix} 5 & 1 & 2 \\ 6 & 7 & 5 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{D} = \begin{pmatrix} 5 & 1 & 2 \\ 6 & 7 & 5 \\ 2 & 8 & 1 \end{pmatrix}.$$

Laske seuraavat matriisitulot

(a) $\boldsymbol{AB}$

(b) $\boldsymbol{AC}$

(c) $\boldsymbol{BC}$

(c) $\boldsymbol{AD}$

1.3
Olkoon

$$\boldsymbol{A} = \begin{pmatrix} 2 & 4 & 3 \\ 8 & 1 & 4 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \\ 8 & 3 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{D} = \begin{pmatrix} 5 & 1 & 2 \\ 6 & 7 & 5 \\ 2 & 8 & 1 \end{pmatrix}.$$

Laske

(a) $\operatorname{tr}(\boldsymbol{A})$

(b) $\operatorname{tr}(\boldsymbol{D})$

(c) $\operatorname{tr}(\boldsymbol{BA})$

(d) $\operatorname{tr}(\boldsymbol{AB})$

1.4
Olkoon

$$\boldsymbol{A} = \begin{pmatrix} 5 & 1 \\ 6 & 7 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}, \boldsymbol{C} = \begin{pmatrix} 5 & 1 & 2 \\ 6 & 7 & 5 \\ 2 & 8 & 1 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{D} = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 3 & 7 \\ 1 & 4 & 6 \end{pmatrix},$$
$$\boldsymbol{E} = \begin{pmatrix} 5 & 1 & 0 \\ 6 & 7 & 0 \\ 2 & 8 & 0 \end{pmatrix}, \boldsymbol{F} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 9 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{G} = \begin{pmatrix} 3 & 2 & 3 \\ 4 & 3 & 4 \\ 1 & 4 & 1 \end{pmatrix}.$$

Laske

(a) $|\boldsymbol{A}|$

(b) $|10\boldsymbol{A}|$

(b) $|\boldsymbol{B}|$

(c) $|\boldsymbol{AB}|$

(d) $|\boldsymbol{C}|$

(e) $|\boldsymbol{D}|$

(f) $|\boldsymbol{C}D|$

(g) $|\boldsymbol{E}|$

(h) $|\boldsymbol{F}|$

(i) $|\boldsymbol{G}|$

1.5
Olkoon

$$\boldsymbol{A} = \begin{pmatrix} 5 & 1 \\ 6 & 7 \\ 2 & 8 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{B} = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 3 & 7 \end{pmatrix}.$$

Laske

(a) $|\boldsymbol{AB}|$

(b) $|\boldsymbol{BA}|$

(c) $|\boldsymbol{B}'\boldsymbol{A}'|$

Luku 2

Vektoriavaruudet

2.1 Määritelmät

Määritelmä 6. *Olkoon $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ $p \times 1$ vektoreita. Vektoreiden $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ virittämä vektoriavaruus koostuu kaikista niiden muodostamista lineaarikombinaatioista, ts.*

$$\{\boldsymbol{y} \; : \; \boldsymbol{y} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i, \alpha_i \in \mathbb{R} \text{ kaikilla } i = 1, \ldots, n\}$$

Määritelmä 7. *Olkoon $\boldsymbol{X}$ $n \times p$ matriisi. Tällöin matriisin $\boldsymbol{X} = (\boldsymbol{x}_1 \ldots \boldsymbol{x}_p)$ sarakeavaruus $\mathcal{R}(\boldsymbol{X})$ muodostuu sarakkeiden $\boldsymbol{x}_1 \ldots \boldsymbol{x}_p$ lineaarikombinaatioista, ts.*

$$\mathcal{R}(\boldsymbol{X}) = \left\{\boldsymbol{y} \quad : \; \boldsymbol{y} = \sum_{i=1}^{p} \beta_i \boldsymbol{x}_i = \boldsymbol{X}\boldsymbol{\beta}, \; \boldsymbol{\beta} \in \mathbb{R}^p\right\}$$

Esimerkki 2.1.1
Olkoon

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Vektorien $\boldsymbol{x}_1$ ja $\boldsymbol{x}_2$ virittämä vektoriavaruus on

$$\{\boldsymbol{y} \; : \; \boldsymbol{y} = \alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2, \alpha_1 \in \mathbb{R}, \alpha_2 \in \mathbb{R}\} = \left\{\boldsymbol{y} \; : \; \boldsymbol{y} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \alpha_1 \in \mathbb{R}, \alpha_2 \in \mathbb{R}\right\} = \mathbb{R}^2.$$

2.2 Lineaarinen riippumattomuus ja riippuvuus

Määritelmä 8. *Vektorit $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ ovat lineaarisesti riippumattomia, jos yhtälön*

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i = \boldsymbol{0}$$

ainoa ratkaisu on $\alpha_1 = \cdots = \alpha_n = 0$. *Jos löytyy myös muita ratkaisuja, niin vektorit ovat lineaarisesti riippuvia.*

Lause 8. *Vektorit* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ *ovat lineaarisesti riippuvia, jos ja vain jos vähintään yksi vektori pystytään ilmaisemaan muiden vektorien lineaarikombinaationa.*

Lause 9. *Olkoon* $\boldsymbol{X}$ $m \times m$ *matriisi, jonka sarakkeet ovat* $\boldsymbol{x}_1, \ldots \boldsymbol{x}_m$. *Silloin* $|\boldsymbol{X}| \neq 0$ *jos ja vain jos vektorit* $\boldsymbol{x}_1, \ldots \boldsymbol{x}_m$ *ovat lineaarisesti rippumattomia.*

2.3 Vektoriavaruuden kannat ja dimensio

Määritelmä 9. *Vektoriavaruuden* S *osajoukko* $V = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_n\}$ *on vektoriavaruuden* S *kanta, jos*

(a) V *virittää vektoriavaruuden* S *ja*

(b) V *on lineaarisesti riippumaton.*

Määritelmä 10. *Jos vektoriavaruus* S *on* $\{\mathbf{0}\}$, *niin vektoriavaruuden dimensio on* $dim(S) = 0$. *Muussa tapauksessa vektoriavaruuden* S *dimensio on vektoriavaruuden* S *minkä tahansa kannan vektorien lukumäärä.*

2.4 Matriisin aste ja lineaarinen riippumattomuus

Määritelmä 11. *Olkoon* $\boldsymbol{X}$ $n \times p$ *matriisi. Matriisin* $\boldsymbol{X}$ *aste* $r(\boldsymbol{X})$ *on* $\boldsymbol{X}$*:n lineaarisesti riippumattomien sarakkeiden lukumäärä tai yhtäpitävästi lineaarisesti riippumattomien rivien lukumäärä.*

Lauseesta 9 ja määritelmästä 11 seuraa välittömästi seuraava lause.

Lause 10. *Olkoon* $\boldsymbol{X}$ $m \times m$ *matriisi. Silloin* $|\boldsymbol{X}| \neq 0$ *jos ja vain jos* $r(\boldsymbol{X}) = m$.

Lause 11. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi. Tällöin*

(a) Jos $\boldsymbol{B}$ *on* $n \times p$ *matriisi, niin* $r(\boldsymbol{AB}) \leq \min\{r(\boldsymbol{A}), r(\boldsymbol{B})\}$.

(b) $r(\boldsymbol{A}) = r(\boldsymbol{A}') = r(\boldsymbol{AA}') = r(\boldsymbol{A}'\boldsymbol{A})$.

2.5 Projektiomatriisit

Määritelmä 12. *Olkoon* $\boldsymbol{P}$ *symmetrinen* ($\boldsymbol{P}' = \boldsymbol{P}$) *ja idempotentti* ($\boldsymbol{P} = \boldsymbol{P}^2$) $n \times n$ *matriisi. Tällöin sanotaan, että matriisi* $\boldsymbol{P}$ *on ortogonaalinen projektio tai lyhyesti projektiomatriisi.*

Lause 12. *Olkoon $V = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_p\}$ vektoriavaruuden S kanta, jossa $\boldsymbol{x}_i \in \mathbb{R}^n$ ja $p \leq n$. Jos matriisin $\boldsymbol{X}$ sarakkeina ovat kantavektorit $\boldsymbol{x}_1, \ldots \boldsymbol{x}_p$, niin vektorin $\boldsymbol{y} \in \mathbb{R}^n$ projektio vektoriavaruudelle S on $\boldsymbol{P}\boldsymbol{y}$, jossa $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ on projektiomatriisi.*

Esimerkki 2.5.1
Olkoon

$$\boldsymbol{X} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

eli matriisi $\boldsymbol{X}$ on x-akselin kanta. Kun $\boldsymbol{z} \in \mathbb{R}^2$ projisoidaan x-akselille, niin projektiomatriisina on

$$\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \begin{pmatrix} 1 \\ 0 \end{pmatrix} 1^{-1} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Jos $\boldsymbol{z} = \begin{pmatrix} a \\ b \end{pmatrix}$, niin vektorin $\boldsymbol{z}$ projektio x-akselille on

$$\boldsymbol{P}\boldsymbol{z} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}$$

eli vektorin $\boldsymbol{z}$ x-koordinaatti.

Esimerkki 2.5.2
Tarkastellaan jälleen lineaarista regressiomallia (ks. esimerkki 1.6.1). Minimoitava jäännösneliösumma voidaan kirjoittaa matriisimuodossa seuraavalla tavalla:

$$S(\boldsymbol{\beta}) = ||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}||^2,$$

jossa $||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}||^2$ on vektorin $\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}$ pituuden neliö. Vektorin $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}$ pituus on pienimmillään silloin kun vektori $\boldsymbol{z}$ on kohtisuorassa matriisin $\boldsymbol{X}$ sarakeavaruutta vastaan eli halutaan löytää sellainen $\boldsymbol{\beta}$, jolla $\boldsymbol{X}\boldsymbol{\beta}$ on vektorin $\boldsymbol{y}$ ortogonaalinen projektio sarakeavaruudelle. Olkoon matriisi $\boldsymbol{X}$ täyttä sarakeastetta. Tällöin $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p\}$ on sarakeavaruuden $\mathcal{R}(\boldsymbol{X})$ kanta ja vektorin $\boldsymbol{y} \in \mathbb{R}^n$ projektio sarakeavaruudelle $\mathcal{R}(\boldsymbol{X})$ on lauseen 12 perusteella $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$. Tällöin siis $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$.

Määritelmä 13. *Olkoon $V = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_n\}$ vektoriavaruuden S kanta. Jos jokainen kantavektori on ortogonaalinen muita kantavektoreita vastaan, niin kantaa kutsutaan ortogonaaliseksi ja jos lisäksi kaikki kantavektorit ovat yksikkövektoreita, niin kantaa kutsutaan ortonormaaliseksi.*

Huom! Jos matriisin $\boldsymbol{X}$ sarakkeina ovat ortonormaalit vektorit $\boldsymbol{x}_1, \ldots \boldsymbol{x}_n$, niin $\boldsymbol{X}$ on ortogonaalinen matriisi ja

$$\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{I}_n.$$

Lause 13. *Olkoon $\boldsymbol{X}$ $n \times p$ matriisi, jonka sarakkeina ovat vektoriavaruuden S ortonormaalisen kannan vektorit. Jos $\boldsymbol{x} \in \mathbb{R}^n$, niin vektorin $\boldsymbol{x}$ ortogonaalinen projektio vektoriavaruudelle S on $\boldsymbol{P}\boldsymbol{x}$, jossa $\boldsymbol{P} = \boldsymbol{X}\boldsymbol{X}'$ on projektiomatriisi.*

2.6 Harjoituksia

Luku 3

Ominaisarvot ja ominaisvektorit

3.1 Johdanto

3.2 Määritelmät

Määritelmä 14. *Olkoon* $\boldsymbol{A}$ $m \times m$ *neliömatriisi. Skalaari* λ *on matriisin* $\boldsymbol{A}$ *ominaisarvo (eigenvalue), jos on olemassa vektori* $\boldsymbol{x} \neq \boldsymbol{0}$ *siten, että*

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}. \tag{3.1}$$

Yhtälön 3.1 toteuttavat vektorit $\boldsymbol{x} \neq \boldsymbol{0}$ *ovat ominaisarvoon* λ *liittyviä ominaisvektoreita (eigenvector).*

Yhtälö 3.1 voidaan kirjoittaa yhtäpitävästi muodossa (totea)

$$(\boldsymbol{A} - \lambda \boldsymbol{I}_m)\boldsymbol{x} = \boldsymbol{0}. \tag{3.2}$$

Jos matriisilla $\boldsymbol{A} - \lambda \boldsymbol{I}_m$ olisi käänteismatriisi, niin yhtälö 3.2 voitaisiin kertoa puolittain vasemmalta matriisilla $(\boldsymbol{A} - \lambda \boldsymbol{I}_m)^{-1}$, jolloin saataisiin

$$\begin{aligned}
&(\boldsymbol{A} - \lambda \boldsymbol{I}_m)^{-1}(\boldsymbol{A} - \lambda \boldsymbol{I}_m)\boldsymbol{x} = (\boldsymbol{A} - \lambda \boldsymbol{I}_m)^{-1}\boldsymbol{0} \\
&\Leftrightarrow \boldsymbol{I}_m \boldsymbol{x} = \boldsymbol{0} \\
&\Leftrightarrow \boldsymbol{x} = \boldsymbol{0}.
\end{aligned}$$

Seuraa ristiriita, koska aikaisemmin on sovittu, että $\boldsymbol{x} \neq \boldsymbol{0}$. Tästä seuraa, että matriisilla $\boldsymbol{A} - \lambda \boldsymbol{I}_m$ ei ole käänteismatriisia, joten sen determinantin täytyy olla nolla. Matriisin $\boldsymbol{A}$ ominaisarvot löydetään siis ratkaisemalla *karakteristinen yhtälö*

$$|\boldsymbol{A} - \lambda \boldsymbol{I}_m| = 0. \tag{3.3}$$

skalaarin λ suhteen. Ominaisarvoa λ vastaava ominaisvektori löydetään sen jälkeen ratkaisemalla yhtälö 3.2 vektorin $\boldsymbol{x}$ suhteen.

Esimerkki 3.2.1
Lasketaan matriisin $\boldsymbol{A} = \begin{pmatrix} 4 & 6 \\ 3 & 7 \end{pmatrix}$ ominaisarvot ja ominaisvektorit.

$$
\begin{aligned}
&|\boldsymbol{A} - \lambda \boldsymbol{I}_2| = 0 \Leftrightarrow \left| \begin{pmatrix} 4 & 6 \\ 3 & 7 \end{pmatrix} - \lambda \boldsymbol{I}_2 \right| = 0 \Leftrightarrow \begin{vmatrix} 4-\lambda & 6 \\ 3 & 7-\lambda \end{vmatrix} = 0 \\
&\Leftrightarrow (4-\lambda)(7-\lambda) - 6 \cdot 3 = 0 \Leftrightarrow \lambda^2 - 11\lambda + 10 = 0 \\
&\lambda = \frac{11 \pm \sqrt{(-11)^2 - 4 \cdot 1 \cdot 10}}{2 \cdot 1} \Leftrightarrow \lambda = \frac{11 \pm 9}{2} \Leftrightarrow \lambda = 1 \quad \text{tai} \quad \lambda = 10.
\end{aligned}
$$

Lasketaan ominaisarvoa $\lambda = 1$ vastaava ominaisvektori.

$$
\begin{aligned}
&(\boldsymbol{A} - \boldsymbol{I}_2)\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \begin{pmatrix} 4-1 & 6 \\ 3 & 7-1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 3 & 6 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&\Leftrightarrow \begin{pmatrix} 3x_1 + 6x_2 \\ 3x_1 + 6x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \frac{1}{3} \begin{pmatrix} 3x_1 + 6x_2 \\ 3x_1 + 6x_2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x_1 + 2x_2 \\ x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned}
$$

Nähdään, että yhtälön ratkaisuna on vektori

$$
\boldsymbol{x} = \begin{pmatrix} 2a \\ -a \end{pmatrix}, \quad a \neq 0.
$$

Ohjelmistot antavat yleensä normalisoidun ominaisvektorin eli ominaisvektorin, jonka pituus on yksi. Koska $|\boldsymbol{x}| = \sqrt{(2a)^2 + (-a)^2} = \sqrt{5a^2} = a\sqrt{5}$, niin normalisoiduksi ominaisvektoriksi saadaan

$$
\boldsymbol{x} = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix} \quad \text{tai} \quad \boldsymbol{x} = \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}.
$$

Lasketaan vielä ominaisarvoa $\lambda = 10$ vastaava ominaisvektori.

$$
\begin{aligned}
&(\boldsymbol{A} - \boldsymbol{I}_2)\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \begin{pmatrix} 4-10 & 6 \\ 3 & 7-10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -6 & 6 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&\Leftrightarrow \begin{pmatrix} -6x_1 + 6x_2 \\ 3x_1 - 3x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x_1 - x_2 \\ x_1 - x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned}
$$

Yhtälön ratkaisuna on nyt vektori

$$
\boldsymbol{x} = \begin{pmatrix} a \\ a \end{pmatrix}, \quad a \neq 0,
$$

josta saadaan normalisoiduksi ominaisvektoriksi

$$
\boldsymbol{x} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \text{tai} \quad \boldsymbol{x} = \begin{pmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}.
$$

```
> A<-matrix(c(4,3,6,7),2,2)
> A
     [,1] [,2]
[1,]    4    6
[2,]    3    7
> eigen(A)
$values
[1] 10  1

$vectors
           [,1]       [,2]
[1,] -0.7071068 -0.8944272
[2,] -0.7071068  0.4472136

> ev<-eigen(A)
> ev$values
[1] 10  1
> ev$vectors
           [,1]       [,2]
[1,] -0.7071068 -0.8944272
[2,] -0.7071068  0.4472136
> 2/sqrt(5)
[1] 0.8944272
> 1/sqrt(5)
[1] 0.4472136
> 1/sqrt(2)
[1] 0.7071068
```

3.3 Ominaisarvojen ja ominaisvektorien perusominaisuuksia

Lause 14. *Olkoon $\boldsymbol{A}$ neliömatriisi. Tällöin*

(a) matriiseilla $\boldsymbol{A}'$ ja $\boldsymbol{A}$ on samat ominaisarvot,

(b) matriisi $\boldsymbol{A}$ on singulaarinen jos ja vain jos vähintään yksi matriisin $\boldsymbol{A}$ ominaisarvoista on nolla,

(c) yläkolmiomatriisin (tai alakolmiomatriisin) $\boldsymbol{A}$ diagonaalialkiot ovat samalla myös matriisin $\boldsymbol{A}$ ominaisarvot,

(d) matriiseilla $\boldsymbol{BAB}^{-1}$ ja $\boldsymbol{A}$ on samat ominaisarvot, jos $\boldsymbol{B}$ on ei-singulaarinen neliömatriisi.

Määritelmä 15. *Olkoon $\boldsymbol{A}$ neliömatriisi. Tällöin matriisin $\boldsymbol{A}$ k:s potenssi on*

$$\boldsymbol{A}^k = \prod_{i=1}^{k} \boldsymbol{A}.$$

Toisin sanoen, $\boldsymbol{A}^2 = \boldsymbol{AA}$, $\boldsymbol{A}^3 = \boldsymbol{AAA}$, $\boldsymbol{A}^4 = \boldsymbol{AAAA}, \cdots$.

Lause 15. *Olkoon λ matriisin $\boldsymbol{A}$ ominaisarvo, $c \in \mathbb{R}$ ja $k \in \{1, 2, \ldots\}$. Tällöin*

(a) λ^k on matriisin $\boldsymbol{A}^k$ ominaisarvo ja

(b) $c\lambda$ on matriisin $c\boldsymbol{A}$ ominaisarvo.

Lause 16. *Olkoon $\boldsymbol{A}$ $n \times n$ matriisi, jonka ominaisarvot ovat $\lambda_1, \ldots, \lambda_n$. Tällöin*

(a) $\sum_{i=1}^{n} \lambda_i = tr(\boldsymbol{A})$ ja

(b) $\prod_{i=1}^{n} \lambda_i = |\boldsymbol{A}|$.

Esimerkki 3.3.1

Jatkoa esimerkkiin 3.2.1. Matriisin $\boldsymbol{A}$ ominaisarvojen summa on $10+1 = 11$, joka on sama kuin matriisin $\boldsymbol{A}$ jälki $4 + 7 = 11$. Matriisin $\boldsymbol{A}$ ominaisarvojen tulo on $10 \cdot 1 = 10$, joka on sama kuin matriisin $\boldsymbol{A}$ determinantti $|\boldsymbol{A}| = 4 \cdot 7 - 6 \cdot 3 = 28 - 18 = 10$.

Lause 17. *Olkoon $\boldsymbol{A}$ idempotentti $m \times m$ matriisi, ts. $\boldsymbol{A}^2 = \boldsymbol{A}$. Tällöin matriisin $\boldsymbol{A}$ ominaisarvot ovat ainoastaan nollia tai ykkösiä.*

Todistus. Olkoon $\boldsymbol{A}$ idempotentti matriisi. Tällöin ominaisarvot λ saadaan yhtälöstä

$$\boldsymbol{Ax} = \lambda \boldsymbol{x}, \quad \boldsymbol{x} \neq \boldsymbol{0}. \tag{3.4}$$

ja lauseen 15 perusteella

$$\boldsymbol{A}^2\boldsymbol{x} = \lambda^2 \boldsymbol{x}, \quad \boldsymbol{x} \neq \boldsymbol{0},$$

jossa λ on matriisin $\boldsymbol{A}$ ominaisarvo. Toisaalta, koska $\boldsymbol{A}$ on idempotentti, niin edellinen yhtälö saadaan muotoon

$$\boldsymbol{Ax} = \lambda^2 \boldsymbol{x}, \quad \boldsymbol{x} \neq \boldsymbol{0}. \tag{3.5}$$

Yhtälöitä 3.4 ja 3.5 vertaamalla huomataan, että

$$\lambda = \lambda^2 \Leftrightarrow \lambda(1 - \lambda) = 0 \Leftarrow \lambda = 0 \text{ tai } \lambda = 1.$$

$\square$

3.4 Symmetriset matriisit

Lause 18. *Olkoon* $\boldsymbol{A}$ *symmetrinen neliömatriisi. Tällöin*

(a) *matriisin* $\boldsymbol{A}$ *kaikki ominaisarvot ovat reaalisia.*

(b) *jokaiselle matriisin* $\boldsymbol{A}$ *ominaisarvolle löytyy vastaava reaalinen ominaisvektori.*

Lause 19. *Olkoon* $\boldsymbol{A}$ *symmetrinen* $n \times n$ *neliömatriisi. Tällöin matriisin* $\boldsymbol{A}$ *ominaisvektoreista on mahdollista muodostaa ortonormaali joukko* $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$.

Lause 20. *Olkoon* $\boldsymbol{A}$ *symmetrinen* $n \times n$ *neliömatriisi, jolla on* p *nollasta eroavaa ominaisarvoa. Tällöin matriisin* $\boldsymbol{A}$ *aste on* $r(\boldsymbol{A}) = p$.

Lause 21. *Olkoon* $\boldsymbol{A}$ *symmetrinen* $n \times n$ *neliömatriisi, jonka ominaisarvot ovat* $\lambda_1 \geq \cdots \geq \lambda_n$. *Jokaisella* $n \times 1$ *vektorilla* $\boldsymbol{x} \neq \boldsymbol{0}$,

$$\lambda_n \leq \frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} \leq \lambda_1,$$

ja

$$\lambda_n = \min_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} \quad \textit{ja} \quad \lambda_1 = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}}.$$

Minimi saavutetaan pienintä ominaisarvoa vastaavalla ominaisvektorilla ja maksimi saavutetaan suurinta ominaisarvoa vastaavalla ominaisvektorilla.

Esimerkki 3.4.1
Olkoon $\boldsymbol{x} = (x_1, \dots, x_n)'$ $n \times 1$ satunnaisvektori, jonka kovarianssimatriisi on $\boldsymbol{\Omega}$. Halutaan löytää $n \times 1$ vektori $\boldsymbol{a}_1$, jolla satunnaismuuttujan $\boldsymbol{a}_1'\boldsymbol{x}$ (eli satunnaismuuttujien $x_1, \dots, x_n$ lineaarikombinaation) varianssi saataisiin mahdollisimman suureksi. Ongelmana on nyt, että satunnaismuuttujan $\boldsymbol{a}_1'\boldsymbol{x}$ varianssi

$$Var(\boldsymbol{a}_1'\boldsymbol{x}) = \boldsymbol{a}_1' Var(\boldsymbol{x})\boldsymbol{a}_1 = \boldsymbol{a}_1'\Omega\boldsymbol{a}_1,$$

saadaan mielivaltaisen suureksi esimerkiksi asettamalla $\boldsymbol{a}_1 = \alpha\boldsymbol{c}$, $\boldsymbol{c} \neq 0$ ja antamalla kertoimen α kasvaa rajatta. Ratkaisuna on, että maksimoidaan varianssi käyttämällä rajoitetta $\boldsymbol{a}_1'\boldsymbol{a}_1 = 1$ eli yhtäpitävästi maksimoidaan osamäärä

$$\frac{\boldsymbol{a}_1'\Omega\boldsymbol{a}_1}{\boldsymbol{a}_1'\boldsymbol{a}_1}.$$

Lauseen 21 perusteella maksimointitehtävän ratkaisuna saadaan, että $\boldsymbol{a}_1 = \boldsymbol{u}_1$, jossa $\boldsymbol{u}_1$ on kovarianssimatriisin $\boldsymbol{\Omega}$ suurinta ominaisarvoa vastaava normeerattu ominaisvektori. Oletetaan edelleen, että me halutaan löytää yksikkövektori $\boldsymbol{a}_2$, joka on ortogonaalinen vektorille $\boldsymbol{a}_1$ ja joka maksimoi muuttujan $\boldsymbol{a}_2'\boldsymbol{x}$ varianssin. Voidaan osoittaa, että toiseksi suurinta ominaisarvo vastaava ominaisvektori maksimoi tällöin varianssin ja toiseksi suurin ominaisarvo antaa varianssin maksimiarvon. Samalla tavalla jatkamalla saadaan ortonormaali vektorijoukko $\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_n\}$ ja *pääkomponentit* $(\boldsymbol{a}_1'\boldsymbol{x}, \dots, \boldsymbol{a}_n'\boldsymbol{x})'$.

3.5 Positiivisesti definiitit ja semidefiniitit matriisit

Määritelmä 16. *Olkoon* $\boldsymbol{x}$ $m \times 1$ *vektori ja* $\boldsymbol{A}$ *symmetrinen* $m \times m$ *matriisi.*

(a) *Jos* $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0$ *kaikilla* $\boldsymbol{x} \neq 0$*, niin* $\boldsymbol{A}$ *on positiivisesti definiitti.*

(b) *Jos* $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0$ *kaikilla* $\boldsymbol{x} \in \mathbb{R}^m$ *ja* $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = 0$ *jollain* $\boldsymbol{x} \neq 0$*, niin* $\boldsymbol{A}$ *on positiivisesti semidefiniitti.*

(c) *Jos* $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} < 0$ *kaikilla* $\boldsymbol{x} \neq 0$*, niin* $\boldsymbol{A}$ *on negatiivisesti definiitti.*

(d) *Jos* $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \leq 0$ *kaikilla* $\boldsymbol{x} \in \mathbb{R}^m$ *ja* $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = 0$ *jollain* $\boldsymbol{x} \neq 0$*, niin* $\boldsymbol{A}$ *on negatiivisesti semidefiniitti.*

Esimerkki 3.5.1
Olkoon $\boldsymbol{y}$ $n \times 1$ satunnaisvektori, jonka kovarianssimatriisi on $\boldsymbol{\Sigma}$ ja $\boldsymbol{a}$ tunnettu $n \times 1$ vektori. Tällöin

$$\mathrm{Var}(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{a}'\mathrm{Var}(\boldsymbol{y})\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}.$$

Huomataan, että satunnaisvektorin kovarianssimatriisi on aina positiivisesti semidefiniitti, koska satunnaismuuttujan varianssi on aina ei-negatiivinen.

Lause 22. *Positiivisesti definiitti ja negatiivisesti definiitti matriisi on aina ei-singulaarinen.*

Todistus. Jos symmetrinen positiivisesti definiitti $m \times m$ matriisi $\boldsymbol{A}$ olisi singulaarinen, niin siinä tapauksessa olisi olemassa $m \times m$ vektori $\boldsymbol{x} \neq \boldsymbol{0}$ siten, että $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$. Tällöin saataisiin $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ mikä on ristiriidassa oletuksen $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0$ kanssa. Negatiivisesti definiitin matriisin $\boldsymbol{A}$ ei-singulaarisuus seuraa siitä, että matriisi $-\boldsymbol{A}$ on positiivisesti definiitti. □

Lause 23. *Olkoon* $\lambda_1, \ldots, \lambda_m$ *symmetrisen* $m \times m$ *matriisin* $\boldsymbol{A}$ *ominaisarvot. Tällöin*

(a) $\boldsymbol{A}$ *on positiivisesti definiitti jos ja vain jos* $\lambda_i > 0$ *kaikilla* $i = 1, \ldots, m$.

(b) $\boldsymbol{A}$ *on positiivisesti semidefiniitti jos ja vain jos* $\lambda_i \geq 0$ *kaikilla* $i = 1, \ldots, m$ *ja* $\lambda_i = 0$ *vähintään yhdellä* i.

Lause 24. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi, jonka aste on* r*. Tällöin matriiisilla* $\boldsymbol{A}'\boldsymbol{A}$ *on* r *positiivista ominaisarvoa. Matriisi* $\boldsymbol{A}'\boldsymbol{A}$ *on positiivisesti definiitti, jos* $r = n$ *ja positiivisesti semidefiniitti, jos* $r < n$.

Lause 25. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi, jonka aste on* r*. Tällöin matriiisin* $\boldsymbol{A}'\boldsymbol{A}$ *positiiviset ominaisarvot ovat samat kuin matriisin* $\boldsymbol{A}\boldsymbol{A}'$ *positiiviset ominaisarvot.*

Esimerkki 3.5.2
Tarkastellaan lineaarista regressiomallia

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

jossa $\boldsymbol{\beta}$ on $p \times 1$ vektori, $E(\boldsymbol{\epsilon}) = \boldsymbol{0}$ ja $\mathrm{Var}(\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{I}_n$. Parametrivektorin $\boldsymbol{\beta}$ pienimmän neliösumman estimaattori on tunnetusti

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},$$

Osoitetaan, että $\boldsymbol{c}'\hat{\boldsymbol{\beta}}$ on paras lineaarinen harhaton estimaattori suureelle $\boldsymbol{c}'\boldsymbol{\beta}$, jossa $\boldsymbol{c}$ on mielivaltainen $p \times 1$ vektori. Satunnaismuuttuja $\boldsymbol{c}'\hat{\boldsymbol{\beta}}$ on harhaton $\boldsymbol{c}'\boldsymbol{\beta}$:n estimaattori, sillä

$$\mathrm{E}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) = \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\mathrm{E}(\boldsymbol{y}) = \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{c}'\boldsymbol{\beta},$$

kaikilla $\boldsymbol{\beta} \in \mathbb{R}^p$. Näytetään vielä, että satunnaismuuttujan $\boldsymbol{c}'\hat{\boldsymbol{\beta}}$ varianssi vähintään yhtä pieni kuin millä tahansa muulla $\boldsymbol{c}'\boldsymbol{\beta}$:n lineaarisella harhattomalla estimaattorilla. Olkoon $\boldsymbol{a}'\boldsymbol{y}$ mielivaltainen $\boldsymbol{c}'\boldsymbol{\beta}$:n lineaarinen harhaton estimaattori. Tällöin

$$\boldsymbol{c}'\boldsymbol{\beta} = \mathrm{E}(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{a}'\mathrm{E}(\boldsymbol{y}) = \boldsymbol{a}'\boldsymbol{X}\beta,$$

josta seuraa, että

$$\boldsymbol{c}' = \boldsymbol{a}'\boldsymbol{X}.$$

Estimaattorin $\boldsymbol{c}'\hat{\boldsymbol{\beta}}$ varianssi on

$$\mathrm{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) = \boldsymbol{c}'\mathrm{Var}(\hat{\boldsymbol{\beta}})\boldsymbol{c} = \boldsymbol{c}'(\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1})\boldsymbol{c} = \sigma^2\boldsymbol{a}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a}$$

kun taas estimaattorin $\boldsymbol{a}'\boldsymbol{y}$ varianssi on

$$\mathrm{Var}(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{a}'\mathrm{Var}(y)\boldsymbol{a} = \boldsymbol{a}'(\sigma^2\boldsymbol{I}_n)\boldsymbol{a} = \sigma^2\boldsymbol{a}'\boldsymbol{a}.$$

Varianssien erotukseksi saadaan nyt

$$\begin{aligned}&\mathrm{Var}(\boldsymbol{a}'\boldsymbol{y}) - \mathrm{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{a}'\boldsymbol{a} - \sigma^2\boldsymbol{a}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{a}\\ &= \sigma^2\boldsymbol{a}'(\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{a}\end{aligned}$$

Matriisi $\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ on idempotentti, joten lauseen 17 perusteeella sen ominaisarvot ovat pelkästään nollia tai ykkösia. Edelleen lauseen 23(a) perusteella $\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ on positiivisesti semidefiniitti ja näin ollen

$$\mathrm{Var}(\boldsymbol{a}'\boldsymbol{y}) - \mathrm{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) \geq 0.$$

3.6 Harjoituksia

Luku 4

Matriisihajotelmat

4.1 Pääakseliesitys

Lause 26. *Olkoon $\boldsymbol{A}$ symmetrinen $m \times m$ matriisi, jonka ominaisarvot ovat $\lambda_1, \ldots, \lambda_m$ ja $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ ovat ominaisarvoja vastaavat ortonormaalit ominaisvektorit. Tällöin matriisille $\boldsymbol{A}$ saadaan pääakseliesitys*

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}',$$

jossa $\Lambda = diag(\lambda_1, \ldots, \lambda_m)$ ja $\boldsymbol{U} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$.

Esimerkki 4.1.1
Pääakseliesitystä voidaan käyttää positiivisesti semidefiniitin matriisin $\boldsymbol{A}$ neliöjuurimatriisin $\boldsymbol{A}^{1/2}$ laskemiseen. Neliöjuurimatriisilla $\boldsymbol{A}^{1/2}$ on ominaisuus $\boldsymbol{A}^{1/2}\boldsymbol{A}^{1/2} = \boldsymbol{A}$. Olkoon $\boldsymbol{\Lambda}$ ja $\boldsymbol{U}$ määritelty lauseen 26 mukaisesti. Olkoon $\boldsymbol{\Lambda}^{1/2} = \mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_m^{1/2})$ ja $\boldsymbol{A}^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}'$. Tällöin saadaan

$$\boldsymbol{A}^{1/2}\boldsymbol{A}^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}'\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}' = \boldsymbol{A}.$$

Matriisi $\boldsymbol{A}^{1/2}$ on symmetrinen sillä

$$(\boldsymbol{A}^{1/2})' = (\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}')' = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}' = \boldsymbol{A}^{1/2}.$$

Matriisin $\boldsymbol{A}^{1/2}$ käänteismatriisi $\boldsymbol{A}^{-1/2}$, on vastaavasti $\boldsymbol{A}^{-1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}'$, jossa $\boldsymbol{\Lambda}^{-1/2} = \mathrm{diag}(\lambda_1^{-1/2}, \ldots, \lambda_m^{-1/2})$, sillä

$$\boldsymbol{A}^{-1/2}\boldsymbol{A}^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}'\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{I}_m\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{U}' = \boldsymbol{I}_m.$$

Neliöjuurimatriisin määritelmää voidaan yleistää, jos tingitään symmetrisyysvaatimuksesta. Olkoon $\boldsymbol{A}^{1/2}$ mikä tahansa matriisi, jolle pätee $\boldsymbol{A} = \boldsymbol{A}^{1/2}(\boldsymbol{A}^{1/2})'$. Olkoon $\boldsymbol{Q}$ on mikä tahansa ortogonaalinen $m \times m$ matriisi ja $\boldsymbol{A}^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{Q}'$, niin

$$\boldsymbol{A}^{1/2}(\boldsymbol{A}^{1/2})' = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}' = \boldsymbol{A}.$$

Lause 27. *Olkoon $\boldsymbol{A}$ positiivisesti semidefiniitti matriisi. Tällöin on olemassa alakolmiomatriisi $\boldsymbol{T}$, jolla on ei-negatiiviset diagonaalialkiot, siten että*

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{T}'.$$

Edelleen, jos $\boldsymbol{A}$ on positiivisesti definiitti, niin matriisi $\boldsymbol{T}$ on yksikäsitteinen matriisi, jolla on positiiviset diagonaalialkiot.

Edellisen lauseen hajotelmaa $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{T}'$ kutsutaan matriisin $\boldsymbol{A}$ *Choleskyn hajotelmaksi*. Lause 27 ilmoittaa, että positiivisesti definiitille matriisille löytyy aina *yksikäsitteinen* Choleskyn hajotelma. Jos matriisi on positiivisesti semidefiniitti, niin Choleskyn hajotelma voi olla yksikäsitteinen tai ei-yksikäsitteinen. Harjoituksessa 4 tehtävässä 2 on esimerkki ei-yksikäsitteisestä hajotelmasta. Tarkastellaan nollamatriisia $\mathbf{0}_{n\times n}$. Se on selvästikin positiivisesti semidefiniitti sillä $\boldsymbol{x}'\mathbf{0}_{n\times n}\boldsymbol{x} = 0 \geq 0$ kaikilla $\boldsymbol{x} \in \mathbb{R}^n$ ja $\boldsymbol{x}'\mathbf{0}_{n\times n}\boldsymbol{x} = 0$ jollain (itse asiassa kaikilla) $\boldsymbol{x} \neq \mathbf{0}$. Nollamatriisille saadaan yksikäsitteinen Choleskyn hajotelma $\mathbf{0}_{n\times n} = \boldsymbol{T}\boldsymbol{T}'$, jossa $\boldsymbol{T} = \mathbf{0}_{n\times n}$.

Esimerkki 4.1.2
Olkoon $\boldsymbol{x}$ $m \times 1$ satunnaisvektori, jolla on odotusarvovektori $\boldsymbol{\mu}$ ja positiivisesti definiitti kovarianssimatriisi $\boldsymbol{\Omega}$. Kovarianssimatriisin neliöjuurimatriisin käänteismatriisia käyttämällä voidaan muodostaa satunnaisvektorin $\boldsymbol{x}$ lineaarinen transformaatio, jonka kovarianssimatriisi on $\boldsymbol{I}_m$. Olkoon matriisin $\boldsymbol{\Omega}$ pääakseliesitys $\boldsymbol{\Omega} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'$ ja $\boldsymbol{\Omega}^{-1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}'$. Lasketaan lineaarisen transformaation $\boldsymbol{z} = \boldsymbol{\Omega}^{-1/2}\boldsymbol{x}$ kovarianssimatriisi:

$$\begin{aligned}
\mathrm{Var}(\boldsymbol{z}) &= \mathrm{Var}(\boldsymbol{\Omega}^{-1/2}\boldsymbol{x}) = \boldsymbol{\Omega}^{-1/2}\mathrm{Var}(\boldsymbol{x})(\boldsymbol{\Omega}^{-1/2})' \\
&= [\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}'][\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'][(\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}')'] \\
&= \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{I}_m\boldsymbol{\Lambda}\boldsymbol{I}_m\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}' \\
&= \boldsymbol{U}\boldsymbol{I}_m\boldsymbol{U}' = \boldsymbol{U}\boldsymbol{U}' = \boldsymbol{I}_m.
\end{aligned}$$

Toinen tapa on käyttää kovarianssimatriisin $\boldsymbol{\Omega}$ Choleskyn hajotelman käänteismatriisia $\boldsymbol{T}^{-1}$. Tällöin

$$\begin{aligned}
\mathrm{Var}(\boldsymbol{z}) &= \mathrm{Var}(\boldsymbol{T}^{-1}\boldsymbol{x}) = \boldsymbol{T}^{-1}\mathrm{Var}(\boldsymbol{x})(\boldsymbol{T}^{-1})' \\
&= \boldsymbol{T}^{-1}\boldsymbol{T}\boldsymbol{T}'(\boldsymbol{T}^{-1})' = \boldsymbol{T}^{-1}\boldsymbol{T}\boldsymbol{T}'(\boldsymbol{T}')^{-1} = \boldsymbol{I}_m.
\end{aligned}$$

4.2 Singulaariarvohajotelma

Lause 28. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi, jonka aste on* $r > 0$. *Tällöin on olemassa ortogonaalinen* $m \times m$ *matriisi* $\boldsymbol{P}$ *ja ortogonaalinen* $n \times n$ *matriisi* $\boldsymbol{Q}$, *siten että* $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{Q}'$ *ja* $\boldsymbol{D} = \boldsymbol{P}'\boldsymbol{A}\boldsymbol{Q}$, *jossa* $m \times n$ *matriisi* $\boldsymbol{D}$ *on*

a)
$$\boldsymbol{\Delta}, \quad \text{jos } r = m = n$$

b)
$$\begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0}_{m\times(n-m)} \end{pmatrix} \quad \text{jos } r = m < n$$

c)
$$\begin{pmatrix} \boldsymbol{\Delta} \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix} \quad \text{jos } r = n < m$$

d)
$$\begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix} \quad \text{jos } r < m, r < n$$

ja $\boldsymbol{\Delta}$ *on* $r \times r$ *diagonaalimatriisi, jonka diagonaalialkiot (singulaariarvot) ovat positiivisia. Matriisin* $\boldsymbol{\Delta}^2$ *diagonaalialkiot ovat matriisien* $\boldsymbol{A}'\boldsymbol{A}$ *ja* $\boldsymbol{A}\boldsymbol{A}'$ *positiiviset ominaisarvot. Matrisin* $\boldsymbol{P}$ *sarakkeina ovat matriisin* $\boldsymbol{A}\boldsymbol{A}'$ *ominaisvektorit ja matriisin* $\boldsymbol{Q}$ *sarakkeina ovat matriisin* $\boldsymbol{A}'\boldsymbol{A}$ *ominaisvektorit.*

Huom! Singulaariarvohajotelman singulaariarvojen lukumäärä antaa yleisen $m\times n$ matriisin asteen. Aikaisemmin todettiin (Lause 20), että *symmetrisellä neliömatriisilla* nollasta eroavien ominaisarvojen lukumäärä ilmoittaa matriisin asteen.

Esimerkki 4.2.1
Olkoon
$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 1 \\ 3 & -1 & 1 \\ -2 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Lasketaan singulaariarvohajotelma R-ohjelmiston funktion `svd` avulla:

```
> A<-matrix(c(2,3,-2,1,0,-1,4,1,1,1,1,1),4,3)
> A
     [,1] [,2] [,3]
[1,]    2    0    1
[2,]    3   -1    1
[3,]   -2    4    1
[4,]    1    1    1
> svdA<-svd(A)
```

```
> svdA
$d
[1] 5.291503e+00 3.464102e+00 1.762781e-16


$u
              [,1] [,2]       [,3]
[1,] -2.672612e-01 -0.5 -0.1262763
[2,] -5.345225e-01 -0.5  0.5540196
[3,]  8.017837e-01 -0.5  0.3272543
[4,]  2.507895e-16 -0.5 -0.7549975

$v
              [,1]       [,2]       [,3]
[1,] -7.071068e-01 -0.5773503 -0.4082483
[2,]  7.071068e-01 -0.5773503 -0.4082483
[3,]  1.919580e-16 -0.5773503  0.8164966

> svdA$u%*%diag(svdA$d)%*%t(svdA$v)
     [,1]          [,2] [,3]
[1,]    2 -4.292554e-16    1
[2,]    3 -1.000000e+00    1
[3,]   -2  4.000000e+00    1
[4,]    1  1.000000e+00    1
```

Funktio `svd` antaa singulaariarvohajotelman tiivistetyssä muodossa. Itseasiassa sitä voi tiivistää enemmänkin, koska vektorin `svdA$d` on kolmas alkio on nolla (toisinsanottuna singulaariarvoja on kaksi):

```
> svdA$u[,1:2]
              [,1] [,2]
[1,] -2.672612e-01 -0.5
[2,] -5.345225e-01 -0.5
[3,]  8.017837e-01 -0.5
[4,]  2.507895e-16 -0.5
> diag(svdA$d[1:2])
         [,1]     [,2]
[1,] 5.291503 0.000000
[2,] 0.000000 3.464102
> svdA$v[,1:2]
              [,1]       [,2]
[1,] -7.071068e-01 -0.5773503
[2,]  7.071068e-01 -0.5773503
[3,]  1.919580e-16 -0.5773503
```

```
> svdA$u[,1:2]%*%diag(svdA$d[1:2])%*%t(svdA$v[,1:2])
     [,1]          [,2] [,3]
[1,]    2 -4.383429e-16    1
[2,]    3 -1.000000e+00    1
[3,]   -2  4.000000e+00    1
[4,]    1  1.000000e+00    1
```

Korollaari 28.1. *Olkoon $\boldsymbol{A}$ $m \times n$ matriisi, jonka aste on $r > 0$. Tällöin on olemassa $m \times r$ matriisi $\boldsymbol{P}_1$ ja $n \times r$ matriisi $\boldsymbol{Q}_1$, siten että $\boldsymbol{P}_1'\boldsymbol{P}_1 = \boldsymbol{Q}_1'\boldsymbol{Q}_1 = \boldsymbol{I}_r$ ja $\boldsymbol{A} = \boldsymbol{P}_1\boldsymbol{\Lambda}\boldsymbol{Q}_1'$, jossa $\boldsymbol{\Lambda}$ on $r \times r$ diagonaalimatriisi, jonka diagonaalialkiot ovat positiivisia.*

Edellinen korollaari seuraa suoraan Lauseesta 28 osittamalla matriisit $\boldsymbol{P}$ ja $\boldsymbol{Q}$ seuraavalla tavalla:

$$\boldsymbol{P} = (\boldsymbol{P}_1 \ \boldsymbol{P}_2) \ \text{ ja } \ \boldsymbol{Q} = (\boldsymbol{Q}_1 \ \boldsymbol{Q}_2),$$

jossa $\boldsymbol{P}_1$ on $m \times r$ matriisi ja $\boldsymbol{Q}_1$ on $n \times r$ matriisi. Matriisit $\boldsymbol{P}_1$ ja $\boldsymbol{Q}_1$ eivät voi olla ortogonaalisia matriiseja kun $m > r$ ja $n > r$. Ortogonaaliset matriisit ovat aina neliömatriiseja. Nyt $\boldsymbol{P}_1\boldsymbol{P}_1' \neq \boldsymbol{I}_m$ ja $\boldsymbol{Q}_1\boldsymbol{Q}_1' \neq \boldsymbol{I}_n$. Matriisin $\boldsymbol{A}$ singulaariarvohajotelma antaa paljon tietoa matriisin rakenteesta:

- Singulaariarvojen lukumäärä r on sama kuin matriisin $\boldsymbol{A}$ aste.
- Matriisin $\boldsymbol{P}_1$ sarakkeet antavat matriisin $\boldsymbol{A}$ sarakeavaruuden ortonormaalin kannan.
- Matriisin $\boldsymbol{Q}_1$ sarakkeet antavat matriisin $\boldsymbol{A}$ riviavaruuden ortonormaalin kannan.
- Matriisin $\boldsymbol{P}_2$ sarakkeet virittävät matriisin $\boldsymbol{A}'$ nolla-avaruuden $N(\boldsymbol{A}') = \{\boldsymbol{x} : \boldsymbol{A}'\boldsymbol{x} = \boldsymbol{0}, \boldsymbol{x} \in \mathbb{R}^m\}$.
- Matriisin $\boldsymbol{Q}_2$ sarakkeet virittävät matriisin $\boldsymbol{A}$ nolla-avaruuden $N(\boldsymbol{A}) = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}, \boldsymbol{x} \in \mathbb{R}^n\}$.

Singulaariarvohajotelmaa voidaan käyttää myös yleistetyn käänteismatriisin laskemiseen. Yleistetty käänteismatriisi määritellään seuraavasti

Määritelmä 17. *Olkoon $\boldsymbol{A}$ $m \times n$ matriisi. Matriisin $\boldsymbol{A}$ yleistetty $n \times m$ käänteismatriisi $\boldsymbol{A}^-$ toteuttaa ehdon*

$$\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}.$$

Huom! Edellisen ehdon toteuttavia yleistettyjä käänteismatriiseja voi olla ääretön määrä. Yleistetty käänteismatriisi ei ole siis välttämättä yksikäsitteinen. Jos vaaditaan määritelmän 17 ehdon lisäksi, että $\boldsymbol{A}^-\boldsymbol{A}\boldsymbol{A}^- = \boldsymbol{A}^-$, $(\boldsymbol{A}\boldsymbol{A}^-)' = \boldsymbol{A}\boldsymbol{A}^-$ ja $(\boldsymbol{A}^-\boldsymbol{A})' = \boldsymbol{A}^-\boldsymbol{A}$, niin saadaan yksikäsitteinen Mooren-Penrosen yleistetty käänteismatriisi $\boldsymbol{A}^+$.

Lause 29. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi, jonka aste on* $r > 0$ *ja jolla on singulaariarvohajotelma*

$$\boldsymbol{A} = \boldsymbol{P} \begin{pmatrix} \boldsymbol{\Delta} & \boldsymbol{0}_{r\times(n-r)} \\ \boldsymbol{0}_{(m-r)\times r} & \boldsymbol{0}_{(m-r)\times(n-r)} \end{pmatrix} \boldsymbol{Q}'.$$

Olkoon

$$\boldsymbol{B} = \boldsymbol{Q} \begin{pmatrix} \boldsymbol{\Delta}^{-1} & \boldsymbol{E} \\ \boldsymbol{F} & \boldsymbol{G} \end{pmatrix} \boldsymbol{P}',$$

jossa $\boldsymbol{E}$ *on* $r \times (m-r)$ *matriisi,* $\boldsymbol{F}$ *on* $(n-r) \times r$ *matriisi ja* $\boldsymbol{G}$ *on* $(n-r) \times (m-r)$ *matriisi. Tällöin kaikilla valinnoilla* $\boldsymbol{E}$, $\boldsymbol{F}$ *ja* $\boldsymbol{G}$, *matriisi* $\boldsymbol{B}$ *on matriisin* $\boldsymbol{A}$ *yleistetty käänteismatriisi ja mikä tahansa yleistetty käänteismatriisi voidaan ilmaista muodossa* $\boldsymbol{B}$ *joillain matriiseilla* $\boldsymbol{E}$, $\boldsymbol{F}$ *ja* $\boldsymbol{G}$.

Esimerkki 4.2.2
Lasketaan matriisille

$$\boldsymbol{A} = \begin{pmatrix} 3 & 6 \\ 1 & 2 \\ 2 & 4 \end{pmatrix}$$

yleistetty käänteismatriisi.

```
> A<-matrix(c(3,1,2,6,2,4),3,2)
> A
     [,1] [,2]
[1,]    3    6
[2,]    1    2
[3,]    2    4
> P<-eigen(A%*%t(A))$vectors
> Q<-eigen(t(A)%*%A)$vectors
> eigen(t(A)%*%A)$values
[1] 70  0
> eigen(A%*%t(A))$values
[1]  7.000000e+01  4.263256e-14 -2.428613e-15
> Delta<-sqrt(70)
> P
          [,1]       [,2]       [,3]
[1,] 0.8017837  0.5976143  0.0000000
[2,] 0.2672612 -0.3585686 -0.8944272
[3,] 0.5345225 -0.7171372  0.4472136
> Q
          [,1]       [,2]
[1,] 0.4472136 -0.8944272
[2,] 0.8944272  0.4472136
> D<-matrix(0,3,2)
```

```
> D[1,1]<-Delta
> D
       [,1] [,2]
[1,] 8.3666    0
[2,] 0.0000    0
[3,] 0.0000    0
> P%*%D%*%t(Q)
     [,1] [,2]
[1,]    3    6
[2,]    1    2
[3,]    2    4
> invD<-t(D)
> invD[1,1]<-1/invD[1,1]
> invD
          [,1] [,2] [,3]
[1,] 0.1195229    0    0
[2,] 0.0000000    0    0
> B<-Q%*%invD%*%t(P)
> B
           [,1]       [,2]       [,3]
[1,] 0.04285714 0.01428571 0.02857143
[2,] 0.08571429 0.02857143 0.05714286
> A%*%B%*%A
     [,1] [,2]
[1,]    3    6
[2,]    1    2
[3,]    2    4
```

4.3 Harjoituksia

Luku 5

Ositetut matriisit

5.1 Ositettujen matriisien kertolasku

Olkoon $\boldsymbol{A}$ $m \times n$ matriisi, joka on ositettu seuraavasti:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix},$$

jossa $\boldsymbol{A}_{11}$ on $m_1 \times n_1$ matriisi, $\boldsymbol{A}_{12}$ on $m_1 \times n_2$ matriisi, $\boldsymbol{A}_{21}$ on $m_2 \times n_1$ matriisi ja $\boldsymbol{A}_{22}$ on $m_2 \times n_2$ matriisi. Olkoon lisäksi $\boldsymbol{B}$ $n \times p$ matriisi, joka on ositettu seuraavalla tavalla:

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix},$$

jossa $\boldsymbol{B}_{11}$ on $n_1 \times p_1$ matriisi, $\boldsymbol{B}_{12}$ on $n_1 \times p_2$ matriisi, $\boldsymbol{B}_{21}$ on $n_2 \times p_1$ matriisi ja $\boldsymbol{B}_{22}$ on $n_2 \times p_2$ matriisi. Tällöin matriisien $\boldsymbol{A}$ ja $\boldsymbol{B}$ tulo voidaan ilmaista alimatriisien avulla seuraavasti:

$$\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{A}_{11}\boldsymbol{B}_{11} + \boldsymbol{A}_{12}\boldsymbol{B}_{21} & \boldsymbol{A}_{11}\boldsymbol{B}_{12} + \boldsymbol{A}_{12}\boldsymbol{B}_{22} \\ \boldsymbol{A}_{21}\boldsymbol{B}_{11} + \boldsymbol{A}_{22}\boldsymbol{B}_{21} & \boldsymbol{A}_{21}\boldsymbol{B}_{12} + \boldsymbol{A}_{22}\boldsymbol{B}_{22} \end{pmatrix}.$$

Huom! Ositettujen matriisien kertolasku suoritettiin merkinnällisesti täsmälleen samaan tapaan kuin 2×2 matriiseilla:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Ositettujen matriisien tapauksessa täytyy lisäksi alimatriisien kertolaskut olla määriteltyjä. Edellinen ositettujen matriisien kertolasku voidaan luonnollisesti yleistää ositetuille matriiseille

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1m} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rm} \end{pmatrix} \quad \text{ja} \quad \boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1c} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{B}_{m1} & \boldsymbol{B}_{m2} & \cdots & \boldsymbol{B}_{mc} \end{pmatrix}.$$

Kertolasku $\boldsymbol{AB}$ suoritetaan nyt samaan tapaan kuin $r \times m$ matriisin ja $m \times c$ matriisin kertolaskussa. Alimatriisien kertolaskut täytyy tietysti tässäkin tapauksessa olla määriteltyjä.

5.2 Ositetun matriisin käänteismatriisi

Tässä kappaleessa tarkastellaan ositettua $m \times m$ matriisia

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}, \tag{5.1}$$

jossa $\boldsymbol{A}_{11}$ on $m_1 \times m_1$ neliömatriisi, $\boldsymbol{A}_{12}$ on $m_1 \times m_2$ matriisi, $\boldsymbol{A}_{21}$ on $m_2 \times m_1$ matriisi ja $\boldsymbol{A}_{22}$ on $m_2 \times m_2$ neliömatriisi.

Lause 30. *Olkoon $\boldsymbol{A}$ ei-singulaarinen $m \times m$ matriisi, joka on ositettu kuten kaavassa (5.1) ja olkoon*

$$\boldsymbol{B} = \boldsymbol{A}^{-1} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix},$$

matriisin $\boldsymbol{A}$ ositettu käänteismatriisi, jossa matriisit $\boldsymbol{B}_{11}$, $\boldsymbol{B}_{12}$, $\boldsymbol{B}_{21}$ ja $\boldsymbol{B}_{22}$ ovat samaa kokoa vastaavien matriisin $\boldsymbol{A}$ alimatriisien kanssa.
Jos matriisit $\boldsymbol{A}_{11}$ ja $\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$ ovat ei-singulaarisia, niin

(a) $\boldsymbol{B}_{11} = \boldsymbol{A}_{11}^{-1} + \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{B}_{22}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}$,

(b) $\boldsymbol{B}_{22} = (\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1}$,

(c) $\boldsymbol{B}_{12} = -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{B}_{22}$,

(d) $\boldsymbol{B}_{21} = -\boldsymbol{B}_{22}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}$

Jos matriisit $\boldsymbol{A}_{22}$ ja $\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$ ovat ei-singulaarisia, niin

(e) $\boldsymbol{B}_{11} = (\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1}$,

(f) $\boldsymbol{B}_{22} = \boldsymbol{A}_{22}^{-1} + \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{B}_{11}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}$,

(g) $\boldsymbol{B}_{12} = -\boldsymbol{B}_{11}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}$,

(h) $\boldsymbol{B}_{21} = -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{B}_{11}$.

Esimerkki 5.2.1
Matriisien $\boldsymbol{A}_{11}$ ja $\boldsymbol{A}_{22}$ ei tarvitse olla ei-singulaarisisia, jotta ositettu matriisi $\boldsymbol{A}$ olisi ei-singulaarinen. Tarkastellaan matriisia

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{O} \end{pmatrix},$$

jossa $m_1 = m_2$ ja $\boldsymbol{A}_{12}$ ja $\boldsymbol{A}_{21}$ ovat ei-singulaarisia $m_1 \times m_1$ matriiseja. Matriisit $\boldsymbol{A}_{11}$ ja $\boldsymbol{A}_{22}$ ovat nyt $m_1 \times m_1$ nollamatriiseja ja siis luonnollisesti myös singulaarisia. Tällöin

$$\boldsymbol{A}^{-1} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{A}_{21}^{-1} \\ \boldsymbol{A}_{12}^{-1} & \boldsymbol{O} \end{pmatrix},$$

sillä

$$\begin{pmatrix} \boldsymbol{O} & \boldsymbol{A}_{21}^{-1} \\ \boldsymbol{A}_{12}^{-1} & \boldsymbol{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{O} \end{pmatrix} = \begin{pmatrix} \boldsymbol{O}\boldsymbol{O} + \boldsymbol{A}_{21}^{-1}\boldsymbol{A}_{21} & \boldsymbol{O}\boldsymbol{A}_{12} + \boldsymbol{A}_{21}^{-1}\boldsymbol{O} \\ \boldsymbol{A}_{12}^{-1}\boldsymbol{O} + \boldsymbol{O}\boldsymbol{A}_{21} & \boldsymbol{A}_{12}^{-1}\boldsymbol{A}_{12} + \boldsymbol{O}\boldsymbol{O} \end{pmatrix}$$
$$= \begin{pmatrix} \boldsymbol{I}_{m_1} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I}_{m_1} \end{pmatrix} = \boldsymbol{I}_m.$$

Esimerkki 5.2.2
Tarkastellaan lineaarista regressiomallia

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{5.2}$$

jossa $\boldsymbol{y}$ on $n \times 1$ vektori, $\boldsymbol{X}$ on $n \times p$ matriisi, $\boldsymbol{\beta}$ on $p \times 1$ vektori ja $\boldsymbol{\epsilon}$ on $n \times 1$ vektori. Olkoon

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix},$$

jossa $\boldsymbol{X}_1$ on $n \times p_1$ matriisi, $\boldsymbol{X}_2$ on $n \times p_2$ matriisi, $\boldsymbol{\beta}_1$ on $p_1 \times 1$ vektori ja $\boldsymbol{\beta}_2$ on $p_2 \times 1$ vektori ($p = p_1 + p_2$). Tällöin lineaarinen malli (5.2) voidaan kirjoittaa yhtäpitävästi muodossa

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}, \tag{5.3}$$

Oletetaan, että me halutaan verrata lineaarista mallia (5.2) rajoitettuun malliin

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\epsilon}, \tag{5.4}$$

(eli malliin jossa $\boldsymbol{\beta}_2 = \boldsymbol{0}$) residuaalineliösummien erotuksen avulla. Jos $\boldsymbol{X}$ on täyttä sarakeastetta, niin mallien (5.2) ja (5.4) regressiokerroinvektorien pienimmän neliösumman estimaattorit ovat

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \quad \text{ja} \quad \hat{\boldsymbol{\beta}}_1 = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y}.$$

ja residuaalineliösummien erotus on

$$\begin{aligned} &(\boldsymbol{y} - \boldsymbol{X}_1\hat{\boldsymbol{\beta}}_1)'(\boldsymbol{y} - \boldsymbol{X}_1\hat{\boldsymbol{\beta}}_1) - (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\ &= \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1')\boldsymbol{y} - \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y} \\ &= \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y}. \end{aligned}$$

Lasketaan seuraavaksi matriisin

$$\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{pmatrix}$$

käänteismatriisi lauseen 30 avulla. Olkoon $\boldsymbol{A}_{11} = \boldsymbol{X}_1'\boldsymbol{X}_1$, $\boldsymbol{A}_{12} = \boldsymbol{X}_1'\boldsymbol{X}_2$, $\boldsymbol{A}_{21} = \boldsymbol{X}_2'\boldsymbol{X}_1$, $\boldsymbol{A}_{22} = \boldsymbol{X}_2'\boldsymbol{X}_2$ ja $\boldsymbol{X}_{2*} = (\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1')\boldsymbol{X}_2$. Tällöin

$$\begin{aligned}
\boldsymbol{B}_{22} &= (\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1} \\
&= (\boldsymbol{X}_2'\boldsymbol{X}_2 - \boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2)^{-1} \\
&= [\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1')\boldsymbol{X}_2]^{-1} \\
&= [\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1')'(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1')\boldsymbol{X}_2]^{-1} \\
&= (\boldsymbol{X}_{2*}'\boldsymbol{X}_{2*})^{-1}, \\
\boldsymbol{B}_{11} &= \boldsymbol{A}_{11}^{-1} + \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{B}_{22}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} \\
&= (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} + (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} \\
\boldsymbol{B}_{12} &= -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{B}_{22} \\
&= -(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{B}_{22} \\
\boldsymbol{B}_{21} &= -\boldsymbol{B}_{22}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} \\
&= -\boldsymbol{B}_{22}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}
\end{aligned}$$

$$\begin{aligned}
&\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \begin{pmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{pmatrix} \\
&= \boldsymbol{X}_1\boldsymbol{B}_{11}\boldsymbol{X}_1' + \boldsymbol{X}_2\boldsymbol{B}_{21}\boldsymbol{X}_1' + \boldsymbol{X}_1\boldsymbol{B}_{12}\boldsymbol{X}_2' + \boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2' \\
&= \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' + \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \\
&\quad -\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \\
&\quad -\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2' \\
&\quad +\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2' \\
&= \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'[\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' - \boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'] \\
&\quad +\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1') \\
&\quad +\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \\
&= -\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'[\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'](\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1') \\
&\quad +\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'(\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1') \\
&\quad +\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \\
&= (\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'[\boldsymbol{X}_2\boldsymbol{B}_{22}\boldsymbol{X}_2'](\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1') \\
&\quad +\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \\
&= \boldsymbol{X}_{2*}\boldsymbol{B}_{22}\boldsymbol{X}_{2*}' + \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'
\end{aligned}$$

Residuaalineliösummien erotukseksi saadaan nyt

$$\begin{aligned}
&\boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y} \\
&= \boldsymbol{y}'\boldsymbol{X}_{2*}\boldsymbol{B}_{22}\boldsymbol{X}_{2*}'\boldsymbol{y} + \boldsymbol{y}'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y} \\
&= \boldsymbol{y}'\boldsymbol{X}_{2*}(\boldsymbol{X}_{2*}'\boldsymbol{X}_{2*})^{-1}\boldsymbol{X}_{2*}'\boldsymbol{y},
\end{aligned}$$

jossa $\boldsymbol{X}_{2*} = (\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1')\boldsymbol{X}_2$.

5.3 Ositetun matriisin determinantti

Lause 31. *Olkoon $\boldsymbol{A}$ ei-singulaarinen $m \times m$ matriisi, joka on ositettu kuten kaavassa (5.1). Jos $\boldsymbol{A}_{12} = \boldsymbol{O}$ tai $\boldsymbol{A}_{21} = \boldsymbol{O}$, niin $|\boldsymbol{A}| = |\boldsymbol{A}_{11}||\boldsymbol{A}_{22}|$.*

Lause 32. *Olkoon $\boldsymbol{A}$ ei-singulaarinen $m \times m$ matriisi, joka on ositettu kuten kaavassa (5.1).*

(a) $|\boldsymbol{A}| = |\boldsymbol{A}_{22}||\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}|$, *jos* $\boldsymbol{A}_{22}$ *on ei-singulaarinen.*

(b) $|\boldsymbol{A}| = |\boldsymbol{A}_{11}||\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}|$, *jos* $\boldsymbol{A}_{11}$ *on ei-singulaarinen.*

Lause 33. *Olkoon $\boldsymbol{A}$ $m \times n$ matriisi ja $\boldsymbol{B}$ $n \times m$ matriisi. Tällöin*

$$|\boldsymbol{I}_m + \boldsymbol{A}\boldsymbol{B}| = |\boldsymbol{I}_n + \boldsymbol{B}\boldsymbol{A}|.$$

Lause 34. *Olkoon $\boldsymbol{A}$ symmetrinen $m \times m$ matriisi ja olkoon $\boldsymbol{A}_k$ matriisi, joka saadaan poistamalla matriisista $\boldsymbol{A}$ viimeiset $m - k$ riviä ja saraketta. Matriisi $\boldsymbol{A}$ on positiivisesti definiitti, jos ja vain jos $|\boldsymbol{A}_1| > 0$, $|\boldsymbol{A}_2| > 0$, $\ldots$, $|\boldsymbol{A}_k| > 0$.*

Esimerkki 5.3.1
Lauseen 34 perusteella matriisi

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

on positiivisesti definiitti jos ja vain jos

$$|\boldsymbol{A}_1| = a_{11} > 0,\ |\boldsymbol{A}_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0 \text{ ja } |\boldsymbol{A}_3| = |\boldsymbol{A}| > 0.$$

5.4 Harjoituksia

Luku 6

Lineaariset yhtälöryhmät

6.1 Lineaarisen yhtälöryhmän ratkaisujen lukumäärä

Määritelmä 18. *Olkoon $\boldsymbol{A}$ $m \times n$ matriisi, $\boldsymbol{x}$ $n \times 1$ vektori ja $\boldsymbol{c}$ $m \times 1$ vektori. Halutaan ratkaista lineaarinen yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c} \tag{6.1}$$

vektorin $\boldsymbol{x}$ suhteen. Lineaarinen yhtälöryhmä 6.1 on **konsistentti**, *jos sillä on vähintään yksi ratkaisu.*

Lause 35. *Olkoon $\boldsymbol{A}$ $m \times n$ matriisi, $\boldsymbol{x}$ $n \times 1$ vektori ja $\boldsymbol{c}$ $m \times 1$ vektori. Yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$$

on konsistentti, jos ja vain jos $r(\boldsymbol{B}) = r(\boldsymbol{A})$, jossa $\boldsymbol{B} = \begin{pmatrix}\boldsymbol{A} & \boldsymbol{c}\end{pmatrix}$.

Esimerkki 6.1.1
Tarkastellaan lineaarista yhtälöryhmää $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$, jossa

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{ja} \quad \boldsymbol{c} = \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix}.$$

Huomataan että matriisin $\boldsymbol{A}$ aste on 2. Matriisin

$$\boldsymbol{B} = \begin{pmatrix}\boldsymbol{A} & \boldsymbol{c}\end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 5 \\ 1 & 0 & 3 \end{pmatrix}$$

determinantti on nolla, joten myös matriisin $\boldsymbol{B}$ aste on 2. Yhtälöryhmä on siis konsistentti eli silä on vähintään yksi ratkaisu.

Lause 36. *Lineaarinen yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$$

on konsistentti, jos ja vain jos matriisilla $\boldsymbol{A}$ *on yleistetty käänteismatriisi* $\boldsymbol{A}^-$*, jolle pätee* $\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{c} = \boldsymbol{c}$*.*

Korollaari 36.1. *Jos* $\boldsymbol{A}$ *on ei-singulaarinen* $m \times m$ *matriisi ja* $\boldsymbol{c}$ *on* $m \times 1$ *vektori, niin yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$$

on konsistentti.

Korollaari 36.2. *Jos* $\boldsymbol{A}$ *on* $m \times n$ *matriisi, jonka aste on* $r(\boldsymbol{A}) = m$*, niin yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$$

on konsistentti.

6.2 Lineaarisen yhtälöryhmän ratkaiseminen

Lause 37. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi,* $\boldsymbol{x}$ $n \times 1$ *vektori ja* $\boldsymbol{c}$ $m \times 1$ *vektori. Oletetaan, että yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$$

on konsistentti ja olkoon $\boldsymbol{A}^-$ *matriisin* $\boldsymbol{A}$ *yleistetty käänteismatriisi. Jokaiselle* $n \times 1$ *vektorille* $\boldsymbol{y}$*,*

$$\boldsymbol{x}_y = \boldsymbol{A}^-\boldsymbol{c} + (\boldsymbol{I}_n - \boldsymbol{A}^-\boldsymbol{A})\boldsymbol{y}$$

on yhtälöryhmän ratkaisu ja jokaiselle ratkaisulle $\boldsymbol{x}_*$ *on olemassa vektori* $\boldsymbol{y}$*, jolla* $\boldsymbol{x}_* = \boldsymbol{x}_y$*.*

Lause 38. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi,* $\boldsymbol{x}$ $n \times 1$ *vektori ja* $\boldsymbol{c}$ $m \times 1$ *vektori. Oletetaan, että yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$$

on konsistentti. Ratkaisu $\boldsymbol{x}_* = \boldsymbol{A}^-\boldsymbol{c}$ *on yhtälöryhmän yksikäsitteinen ratkaisu, jos ja vain jos* $\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{I}_n$*, jossa* $\boldsymbol{A}^-$ *on mikä tahansa* $m \times n$ *matriisin* $\boldsymbol{A}$ *yleistetty käänteismatriisi.*

Korollaari 38.1. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi,* $\boldsymbol{x}$ $n \times 1$ *vektori ja* $\boldsymbol{c}$ $m \times 1$ *vektori. Oletetaan, että yhtälöryhmä*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$$

on konsistentti. Ratkaisu $\boldsymbol{x}_* = \boldsymbol{A}^-\boldsymbol{c}$ *on yhtälöryhmän yksikäsitteinen ratkaisu, jos ja vain jos* $r(\boldsymbol{A}) = n$*.*

Lause 39. *Olkoon* $\boldsymbol{A}$ $m \times n$ *matriisi ja* $\boldsymbol{x}$ $n \times 1$ *vektori. Yhtälöryhmällä*

$$\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$$

on ei-triviaali ratkaisu $\boldsymbol{x}_* \neq 0$,

a) jos ja vain jos $\boldsymbol{A}^{-}\boldsymbol{A} \neq \boldsymbol{I}_n$.

b) jos ja vain jos $r(\boldsymbol{A}) < n$.

6.3 Harjoituksia

Luku 7

Matriisiderivaatat

7.1 Muutamia hyödyllisiä matriisiderivaattoja

Olkoon $f(\boldsymbol{x}) = f(x_1, \ldots, x_n)$ skalaarifunktio. Määritellään funktion f derivaatta vektorin $\boldsymbol{x}$ suhteen seuraavasti:

$$\frac{\partial}{\partial \boldsymbol{x}} f(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} f(\boldsymbol{x}) \\ \frac{\partial}{\partial x_2} f(\boldsymbol{x}) \\ \vdots \\ \frac{\partial}{\partial x_n} f(\boldsymbol{x}) \end{pmatrix}$$

Lause 40. *Olkoon* $f(\boldsymbol{x}) = \boldsymbol{a}'\boldsymbol{x} = \boldsymbol{x}'\boldsymbol{a}$*, jossa* $\boldsymbol{x} = (x_1 \ldots x_n)'$ *ja* $\boldsymbol{a} = (a_1 \ldots a_n)'$*. Tällöin*

$$\frac{\partial}{\partial \boldsymbol{x}} f(\boldsymbol{x}) = \boldsymbol{a}.$$

Lause 41. *Olkoon* $f(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$*, jossa* $\boldsymbol{x} = (x_1 \ldots x_n)'$ *ja* $\boldsymbol{A}$ *on symmetrinen* $n \times n$ *matriisi. Tällöin*

$$\frac{\partial}{\partial \boldsymbol{x}} f(\boldsymbol{x}) = 2\boldsymbol{A}\boldsymbol{x}.$$

Esimerkki 7.1.1
Tarkastellaan lineaarista regressiomallia

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}.$$

Regressiokerroinvektorin $\boldsymbol{\beta}$ pienimmän neliösumman estimaatti saadaan minimoimalla residuaalineliösumma

$$\begin{aligned} f(\boldsymbol{\beta}) &= (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \\ &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} \end{aligned}$$

vektorin $\boldsymbol{\beta}$ suhteen. Derivoidaan $f(\boldsymbol{\beta})$ vektorin $\boldsymbol{\beta}$ suhteen:

$$\frac{\partial}{\partial \boldsymbol{x}} f(\boldsymbol{x}) = \boldsymbol{0} - \boldsymbol{X}'\boldsymbol{y} - \boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = 2(\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{X}'\boldsymbol{y}).$$

Kun derivaatta asetetaan nollaksi, niin saadaan normaaliyhtälö

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{y},$$

jolla on yksikäsitteinen ratkaisu

$$\boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},$$

jos matriisi $\boldsymbol{X}$ on täyttä sarakeastetta.

7.2 Harjoituksia

Kirjallisuutta

[1] Abadir, K. M. and Magnus, J. R. (2005). *Matrix Algebra*. New York: Cambridge University Press.

[2] Basilevsky, A. (2005). *Applied Matrix Algebra in the Statistical Sciences*. New York: Dover.

[3] Schott, J. R. (2005). *Matrix Analysis for Statistics*, Second Edition. New York: Wiley.

[4] Searle, S. R. (1982). *Matrix Algebra Useful for Statistics*. New York: Wiley.