Matriisihajotelmia

Saaga Koskimäki

Matematiikan Pro Gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Syksy 2023

Tiivistelmä: Saaga Koskimäki, *Matriisihajotelmia* (engl. *Matrix decompositions*), matematiikan Pro Gradu -tutkielma, 49 s., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, syksy 2023.

Tämän tutkielman tarkoituksena on tarkastella matriisin kolmea erilaista hajotelmaa. Matriisihajotelmien avulla matriisi voidaan esittää hyödyllisessä muodossa muita tuloksia varten. Tutkielmassa perehdytään matriisin CR-hajotelmaan, symmetrisen matriisin diagonaalihajotelmaan sekä singulaariarvohajotelmaan. Lisäksi singulaariarvohajotelman sovelluksena käsitellään matriisin perturbaatiota.

Matriisin CR-hajotelmassa matriisi A esitetään matriisien C ja R avulla muodossa $A = CR$. Tässä hyödynnetään matriisin A astetta, lineaarisesti riippumattomia sarakkeita sekä redusoitua porrasmatriisia. Tutkielmassa tarkastellaan myös, miten symmetrinen matriisi A voidaan ortogonaalisesti diagonalisoida diagonaalimatriisin D ja ortogonaalisen matriisin V avulla muodossa $A = VDV^T$. Tätä kutsutaan symmetrisen matriisin diagonaalihajotelmaksi. Hajotelmassa matriisin D diagonaalialkiot koostuvat matriisin A ominaisarvoista ja matriisi V kyseisiä ominaisarvoja vastaavista ominaisvektoreista.

Kolmantena hajotelmana tutkielmassa käsitellään matriisin A singulaariarvohajotelmaa $A = U\Sigma V^T$, missä matriisi Σ on singulaariarvoista koostuva diagonaalimatriisi ja matriisit U ja V ovat ortogonaalisia. Lisäksi singulaariarvohajotelman avulla todistetaan matriisin perturbaatiolause, jossa tarkastellaan matriisin $A + cd^T$ astetta, kun tässä matriisin cd^T aste on yksi.

Sisällys

Johdanto 1

Luku 1. Esitietoja 3

Luku 2. Lohkomatriisit 7
- 2.1. Lohkomatriisien summa 7
- 2.2. Lohkomatriisien tulo 8
- 2.3. Lohkomatriisien transpoosi 8

Luku 3. Sarake- ja nolla-avaruus 11

Luku 4. CR-hajotelma 15

Luku 5. Ortogonaalikomplementti 23

Luku 6. Symmetrisen matriisin diagonalisointi 27

Luku 7. Singulaariarvohajotelma 31
- 7.1. Pseudokäänteismatriisi 36

Luku 8. Asteen yksi perturbaatio 39

Kirjallisuutta 49

Johdanto

Tässä tutkielmassa tarkastellaan yleisen $m \times n$ -matriisin A erilaisia hajotelmia sekä sen asteen yksi perturbaatiota. Työssä käsitellään erityisesti matriisin CR-hajotelmaa $A = CR$, symmetrisen matriisin diagonaalihajotelmaa $A = VDV^T$ sekä singulaariarvohajotelmaa $A = U\Sigma V^T$. Näissä hajotelmissa matriisi pystytään esittämään yksinkertaisemmassa muodossa, ja siten, että hajotelmia voidaan hyödyntää erilaisissa sovelluksissa ja todistuksissa.

Matriisin CR-hajotelmassa matriisi A esitetään muodossa $A = CR$, missä matriisi C muodostuu matriisin A lineaarisesti riippumattomista sarakkeista ja matriisi R muodostetaan redusoidun porrasmatriisin ja permutaatiomatriisin avulla. Matriisin A astetta, rank A, hyödynnetään CR-hajotelmassa, koska aste r vastaa $m \times r$ -matriisin C ja $r \times n$ -matriisin R dimensioiden lukua r. Lisäksi CR-hajotelman avulla saadaan todistettua matriisin asteeseen liittyvä hyödyllinen tulos

$$\text{rank } A = \text{rank } A^T,$$

jota tarvitaan esimerkiksi perturbaatiolausetta varten. Matriisin CR-hajotelmaa tarkastellaan Luvussa 4. CR-hajotelman tarkastelussa ja todistuksessa on käytetty pohjana Strangin ja Molerin artikkelia [**7**]. Artikkelissa esiintyvää CR-hajotelman johtamista ja todistusta on muokattu ja täydennetty merkittävästi tässä työssä.

Toisena hajotelmana tutkielmassa tarkastellaan symmetrisen matriisin diagonalisoituvuutta, ja saadaan symmetriselle matriisille hajotelma $A = VDV^T$. Tässä matriisi D on diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin A ominaisarvot, ja matriisi V on ortogonaalinen matriisi, jonka sarakkeet ovat matriisin A omainaisarvoja vastaavat ominaisvektorit. Symmetrisen matriisin diagonalisoituvuutta käsitellään Luvussa 6, ja sen tarkastelussa on käytetty lähteinä Saarimäen kurssikirjaa sekä Strangin kirjaa [**5**], [**6**].

Matriisin singulaariarvohajotelmassa $A = U\Sigma V^T$ mikä tahansa matriisi voidaan esittää ortogonaalisten matriisien U ja V sekä diagonaalimatriisin Σ avulla. Matriisin singulaariarvohajotelma ei ole yksikäsitteinen, mutta sitä voidaan käyttää useissa sovelluksissa ja todistuksissa apuna yksinkertaistamassa tuloksia. Tässä tutkielmassa on todistettu useita singulaariarvohajotelmaa varten tarvittavia tuloksia, sekä singulaariarvohajotelman todistuksen yksityiskohtia on perusteltu täsmällisemmin. Singulaariarvohajotelman matriisien V, Σ ja U avulla voidaan muodostaa matriisin pseudokäänteismatriisi $A^+ = V\Sigma^+U^T$. Jos matriisi A on kääntyvä, niin pseudokäänteismatriisi A^+ vastaa matriisin A käänteismatriisia. Singulaariarvohajotelmaa ja pseudokäänteismatriisia tarkastellaan Luvussa 7, ja niiden läpikäynnissä on käytetty lähteenä Leonin kirjaa [**3**].

Singulaariarvohajotelman sovelluksena työssä perehdytään myös perturbaatiolauseeseen, jossa käsitellään matriisin asteen yksi perturbaatiota. Perturbaatiolauseessa tarkastellaan astetta matriisille $A + cd^T$, missä matriisin cd^T aste on yksi. Perturbaation tarkastelussa Luvussa 8 lähteenä on hyödynnetty Meyerin artikkelia [**4**]. Perturbaatiolauseen Meyerin artikkelissa esiintyvän todistuksen useita yksityiskohtia on tässä tutkielmassa selkeytetty ja tarkennettu.

Keskeisenä apuvälineenä todistuksissa on matriisin A aste, rank A, jolle todistetaan useampi ominaisuus, joita tarvitaan apuna muissa tuloksissa. Asteelle osoitetaan aiemmin esitetyn tuloksen lisäksi ominaisuus

$$\operatorname{rank} A = \operatorname{rank}(A^T A),$$

jota tarvitaan singulaariarvohajotelman todistusta varten. Matriisin asteen tarkastelussa on käytetty lähteinä Leonin ja Strangin kirjoja [**3**], [**6**]. Toinen tutkielmassa käytettävä merkittävä apuväline on matriisien esittäminen lohkomuodossa, minkä avulla saadaan osoitettua suurten matriisien matriisitulo tai transpoosi yksinkertaisessa muodossa. Tätä varten lohkomatriisien laskutoimituksiin perehdytään Luvussa 2. Lohkomatriiseista löytyy lisätietoa Saarimäen kurssikirjasta [**5**].

LUKU 1

Esitietoja

Tässä luvussa kerrataan lineaarisesta algebrasta tuttuja vektoreihin, matriiseihin ja aliavaruuksiin liittyviä määritelmiä ja ominaisuuksia. Lukijalta oletetaan perustietoina Jyväskylän yliopiston Lineaarinen algebra ja geometria 1 ja 2 -kursseja vastaavien tietojen hallintaa. Lähteinä luvussa on käytetty Lineaarisen algebran luentomateriaaleja sekä Leonin kirjaa [**2**], [**1**], [**3**].

Määritellään aluksi symmetrinen matriisi lineaarisesta algebrasta tunnetulla tavalla. Symmetrinen matriisi on hyödyllinen jatkossa, koska sille saadaan osoitettua ortogonaalinen diagonalisoituvuus Luvussa 6, sekä symmetrisen matriisin avulla löydetään matriisin singulaariarvot, joita tarkastellaan luvussa 7. Muistetaan vielä, että $n \times n$ -matriisi A on kääntyvä, jos on olemassa matriisi B siten, että $AB = BA = I$, missä I on identtinen $n \times n$ -matriisi. Identtinen matriisi on yleisessä muodossa

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Merkitään matriisin A käänteismatriisia A^{-1}.

MÄÄRITELMÄ 1.1. Olkoon $n \times n$ -matriisi A. Tällöin A on *symmetrinen*, jos

$$A = A^T.$$

Tässä määritelmässä A^T on matriisin A transpoosi, jossa matriisin A rivit on muutettu sarakkeiksi ja sarakkeet riveiksi.

Määritellään seuraavaksi vektoreiden sisätulo sekä ortogonaalinen ja ortonormaali vektorijoukko. Lähteenä on käytetty Leonin kirjaa [**3**].

MÄÄRITELMÄ 1.2. Olkoot $x \in \mathbb{R}^n$ ja $y \in \mathbb{R}^n$ vektoreita. Vektoreiden x ja y *sisätulo* on

$$x \cdot y = x^T y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

MÄÄRITELMÄ 1.3. Vektoreiden joukko $\{v_1, v_2, \ldots, v_m\} \subset \mathbb{R}^n$ on

1. *ortogonaalinen*, jos $v_i \cdot v_j = 0$, eli yhtäpitävästi $v_i^T v_j = 0$, kun $i \neq j$.
2. *ortonormaali*, jos se ortogonaalinen, ja $\|v_i\| = 1$ kaikilla $i = 1, 2, \ldots, m$.

Seuraavaksi kerrataan ortogonaalinen matriisin määritelmä. Tässä ortogonaalisessa matriisissa rivit, ja toisaalta yhtäpitävästi sarakkeet, ovat keskenään ortonormaalit.

MÄÄRITELMÄ 1.4. Olkoon $n \times n$ -matriisi V. Tällöin V on ortogonaalinen matriisi, jos

$$V^T V = I,$$

missä I on identtinen matriisi. Ortogonaaliselle matriisille V pätee, että

$$V^T = V^{-1},$$

missä V^{-1} on matriisin V käänteismatriisi.

Tarkastellaan sitten vektorin normia ja normin ominaisuuksia.

MÄÄRITELMÄ 1.5. Olkoon vektori $x \in \mathbb{R}^n$. Vektorin x *normi* on

$$\|x\| = \sqrt{x \cdot x}.$$

LAUSE 1.6. *Olkoon vektori $x \in \mathbb{R}^n$. Tällöin vektorin normille $\|x\|$ pätee*

1. $\|x\| = 0$ *jos ja vain jos* $x = 0$,
2. $\|x\|^2 = x^T x$.

TODISTUS. 1. Jos $\|x\| = 0$, niin normin Määritelmän 1.5 nojalla pätee

$$\|x\| = \sqrt{x \cdot x} = 0.$$

Tällöin sisätulon Määritelmän 1.2 nojalla pätee

$$x \cdot x = x_1x_1 + x_2x_2 + \cdots + x_nx_n = (x_1)^2 + (x_2)^2 + \cdots + (x_n)^2 = 0.$$

Koska neliöt $(x_i)^2 \geq 0$, niin jokainen $x_i = 0$ kaikilla $i = 1, 2, \ldots, n$. Näin ollen täytyy olla $x = 0$.

Jos $x = 0$, niin normin Määritelmän 1.5 nojalla saadaan

$$\|x\| = \sqrt{x \cdot x} = \sqrt{0 \cdot 0} = 0.$$

Näin ollen kohta 1 on todistettu.

2. Normin Määritelmän 1.5 nojalla saadaan

$$\|x\|^2 = (\sqrt{x \cdot x})^2 = x \cdot x.$$

Tällöin sisätulon Määritelmän 1.2 nojalla

$$\|x\|^2 = x \cdot x = x^T x,$$

joten saadaan yhtäsuuruus $\|x\|^2 = x^T x$. $\square$

Kerrataan seuraavaksi aliavaruuden ja lineaarisen riippumattomuuden määritelmät. Näissä lähteinä on käytetty Leonin kirjaa [**3**] sekä Lineaarinen algebra ja geometria -kurssin Kilpeläisen luentomateriaalia [**2**]. Näiden avulla saadaan määriteltyä aliavaruuden kanta myöhemmin tässä luvussa.

MÄÄRITELMÄ 1.7. Olkoon $V \subset \mathbb{R}^n$. Joukko V on *aliavaruus*, jos se täyttää seuraavat ehdot:

1. $0_n \in V$,
2. jos avaruuden $\mathbb{R}^n$ vektori $v \in V$, niin $\lambda v \in V$ kaikilla $\lambda \in \mathbb{R}$,
3. jos vektorit $v, w \in V$, niin vektori $v + w \in V$.

MÄÄRITELMÄ 1.8. Olkoot vektorit $\{v_1, v_2, \ldots, v_k\} \subset \mathbb{R}^n$, missä $k \geq 1$. Vektorit ovat *lineaarisesti riippumattomat*, jos yhtälö

$$a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0$$

toteutuu vain, jos kaikille kertoimille pätee $a_i = 0$, missä $a_i \in \mathbb{R}$, ja $i = 1, 2, \ldots, k$. Muussa tapauksessa vektorit ovat *lineaarisesti riippuvat*.

Seuraavassa aliavaruuden kannan määritelmässä tarvitaan vektoreiden $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ lineaarista verhoa,

$$\langle v_1, v_2, \dots, v_k \rangle = \{a_1 v_1 + a_2 v_2 + \cdots + a_k v_k : a_1, a_2, \dots, a_k \in \mathbb{R}\}.$$

Aliavaruuden kanta määritellään seuraavasti:

MÄÄRITELMÄ 1.9. Olkoon V avaruuden $\mathbb{R}^n$ aliavaruus. Tällöin joukon V vektorit $v_1, v_2, \dots, v_k$ muodostavat aliavaruuden V *kannan*, jos

1. $v_1, v_2, \dots, v_k$ ovat lineaarisesti riippumattomat, ja
2. $v_1, v_2, \dots, v_k$ virittävät avaruuden V, eli

$$V = \langle v_1, v_2, \dots, v_k \rangle.$$

Tässä määritelmässä luku k, eli kantavektoreiden lukumäärä, on aliavaruuden V dimensio, jota merkitään $\dim V = k$.

LUKU 2

Lohkomatriisit

Tässä luvussa määritellään lohkomatriisien käsite, sekä tutustutaan lohkomatriisien perusominaisuuksiin. Tarkastellaan lohkomatriisien summaa, tuloa sekä transpoosia keskenään sopivasti lohkotuilla lohkomatriiseilla. Lohkomatriiseja hyödynnetään tässä työssä CR-hajotelman, diagonalisoinnin, singulaariarvohajotelman ja perturbaatiolauseen yhteydessä. Lohkomatriisien avulla saadaan esitettyä isoja matriiseja sekä matriisituloja huomattavasti yksinkertaisemmassa muodossa. Lohkomatriisien käsittelyssä on käytetty lähteenä Saarimäen lineaarisen algebran kurssikirjaa [**5**].

Matriisit voidaan esittää yksinkertaisemmassa muodossa pienempien lohkomatriisien avulla. Määritellään matriisien lohkomuoto seuraavasti:

MÄÄRITELMÄ 2.1. Olkoon $m \times n$ -matriisi A. Tällöin matriisi A esitettynä *lohkomuodossa* on

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1l} \\ A_{21} & A_{22} & \dots & A_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \dots & A_{kl} \end{bmatrix},$$

missä *lohkomatriisit* A_{ij} ovat kokoa $m_i \times n_j$, missä $m = m_1 + m_2 + \cdots + m_k$ ja $n = n_1 + n_2 + \cdots + n_l$, $i = 1, 2, \dots, k$, $j = 1, 2, \dots, l$.

Tarkastellaan seuraavissa alaluvuissa lohkomatriisien summaa, tuloa sekä transpoosia, sekä käydään läpi esimerkki näiden operaatioiden käytöstä. Kaavojen todistuksia ei käydä läpi, vaan havainnollistetaan niiden toimintaa. Todistuksiin voi tutustua Saarimäen kirjan [**5**] avulla.

2.1. Lohkomatriisien summa

Olkoon $A = [A_{ij}]$ $m \times n$ -matriisi ja $B = [B_{ij}]$ $m \times n$ -matriisi. Oletetaan, että matriisit A ja B on ositettu samalla tavalla, eli toisiaan vastaavat lohkot ovat saman kokoiset. Tällöin matriisin A ja B summa on muotoa $[A_{ij} + B_{ij}]$. Tätä havainnollistaa summa

$$A + B = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1l} \\ A_{21} & A_{22} & \dots & A_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \dots & A_{kl} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & \dots & B_{1l} \\ B_{21} & B_{22} & \dots & B_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ B_{k1} & B_{k2} & \dots & B_{kl} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \dots & A_{1l} + B_{1l} \\ A_{21} + B_{21} & A_{22} + B_{22} & \dots & A_{2l} + B_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} + B_{k1} & A_{k2} + B_{k2} & \dots & A_{kl} + B_{kl} \end{bmatrix},$$

missä lohkot A_{ij} ja B_{ij} ovat samankokoiset.

2.2. Lohkomatriisien tulo

Olkoon $A = [A_{ij}]$ $m \times n$ -matriisi ja $B = [B_{ij}]$ $n \times p$ -matriisi. Oletetaan, että matriisit on lohkottu siten, että matriisin A lohkot A_{ij} ovat $m_i \times n_j$ -matriiseja ja B_{ij} ovat $n_i \times p_j$ -matriiseja, missä $m = m_1 + m_2 + \cdots + m_k$, $n = n_1 + n_2 + \cdots + n_l$ ja $p = p_1 + p_1 + \cdots + p_r$.

Tällöin tulo $AB = C$ on $m \times p$ -matriisi, joka on lohkomatriisi $C = [C_{ij}]$, missä $C_{ij} = \sum_{s=1}^{l} A_{is}B_{sj}$ on $m_i \times p_j$ -matriisi kaikilla $i = 1, 2, \ldots, k$ ja $j = 1, 2, \ldots, r$. Näin ollen voidaan tarkastella tuloa

$$
\begin{aligned}
AB &= \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1l} \\ A_{21} & A_{22} & \ldots & A_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \ldots & A_{kl} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \ldots & B_{1r} \\ B_{21} & B_{22} & \ldots & B_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ B_{l1} & B_{l2} & \ldots & B_{lr} \end{bmatrix} \\
&= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} + \cdots + A_{1l}B_{l1} & \ldots & A_{11}B_{1r} + A_{12}B_{2r} + \cdots + A_{1l}B_{lr} \\ A_{21}B_{11} + A_{22}B_{21} + \cdots + A_{2l}B_{l1} & \ldots & A_{21}B_{1r} + A_{22}B_{2r} + \cdots + A_{2l}B_{lr} \\ \vdots & \ddots & \vdots \\ A_{k1}B_{11} + A_{k2}B_{22} + \cdots + A_{kl}B_{l1} & \ldots & A_{k1}B_{1r} + A_{k2}B_{2r} + \cdots + A_{kl}B_{lr} \end{bmatrix} \\
&= \begin{bmatrix} \sum A_{1s}B_{s1} & \sum A_{1s}B_{s2} & \ldots & \sum A_{1s}B_{sr} \\ \sum A_{2s}B_{s1} & \sum A_{2s}B_{s2} & \ldots & \sum A_{2s}B_{sr} \\ \vdots & \vdots & \ddots & \vdots \\ \sum A_{ks}B_{s1} & \sum A_{ks}B_{s2} & \ldots & \sum A_{ks}B_{sr} \end{bmatrix}.
\end{aligned}
$$

Tässä tapauksessa nähdään, että tulomatriisin jokainen alkio on aiemmin esitettyä muotoa.

2.3. Lohkomatriisien transpoosi

Olkoon $A = [A_{ij}]$ lohkomatriisi. Tällöin matriisin transpoosi on lohkomatriisi $C = [C_{ij}]$, missä $C_{ij} = A_{ji}^T$. Näin ollen matriisin A transpoosi A^T, eli matriisi C, on

$$
C = A^T = \begin{bmatrix} A_{11}^T & A_{21}^T & \cdots & A_{m1}^T \\ A_{12}^T & A_{22}^T & \cdots & A_{m2}^T \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n}^T & A_{2m}^T & \cdots & A_{mn}^T \end{bmatrix}.
$$

Käydään läpi esimerkki, jossa lasketaan kahden lohkomuotoisen matriisin tulo, sekä yhden lohkomuotoisen matriisin transpoosi.

Esimerkki 2.2. Olkoot 4×5 -matriisi A ja 5×3 -matriisi B siten, että

$$
A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 0 & 1 & 0 \\ 5 & 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 3 & 4 & 5 \\ 0 & 1 & 0 \end{bmatrix} \\ \begin{bmatrix} 5 & 1 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},
$$

ja

$$B = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 0 & 0 \\ 5 & 3 & 5 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} & \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 5 & 3 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 5 \\ 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Tällöin matriisien tulolle saadaan

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 4 & 5 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 5 & 3 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 3 & 4 & 5 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 5 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 5 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 5 & 3 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 5 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 5 \\ 2 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 31 & 18 \\ 6 & 5 \end{bmatrix} & \begin{bmatrix} 37 \\ 8 \end{bmatrix} \\ \begin{bmatrix} 27 & 21 \\ 6 & 5 \end{bmatrix} & \begin{bmatrix} 31 \\ 6 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 31 & 18 & 37 \\ 6 & 5 & 8 \\ 27 & 21 & 31 \\ 6 & 5 & 6 \end{bmatrix}, \end{aligned}$$

ja matriisin A transpoosille

$$\begin{aligned} A^T &= \begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} & \begin{bmatrix} 5 & 1 \\ 1 & 0 \end{bmatrix} \\ \begin{bmatrix} 3 & 0 \\ 4 & 1 \\ 5 & 0 \end{bmatrix} & \begin{bmatrix} 2 & 0 \\ 3 & 1 \\ 4 & 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 5 & 1 \\ 2 & 1 & 1 & 0 \\ 3 & 0 & 2 & 0 \\ 4 & 1 & 3 & 1 \\ 5 & 0 & 4 & 0 \end{bmatrix}. \end{aligned}$$

LUKU 3

Sarake- ja nolla-avaruus

Tässä osiossa määritellään matriisien nolla- ja sarakeavaruudet, sekä tarkastellaan niihin liittyviä lauseita, kuten dimensiolausetta. Määritellään myös matriisin aste ja todistetaan siihen liittyviä tuloksia. Dimensiolausetta sekä matriisin asteeseen liittyviä ominaisuuksia tarvitaan myöhemmissä luvuissa useammassa todistuksessa, kuten singulaariarvohajotelman ja perturbaatiolauseen yhteydessä. Tässä luvussa päälähteenä on käytetty Leonin kirjaa [**3**], sekä lisäksi lähteenä on ollut Strangin kirja [**6**].

Määritellään ensin matriisin sarake- ja riviavaruudet, sekä matriisin ydin eli nolla-avaruus.

MÄÄRITELMÄ 3.1. Olkoon $m \times n$ -matriisi A. Tällöin matriisin A *sarakeavaruus* on joukko

$$C(A) = \{y \in \mathbb{R}^m : y = Ax \text{ jollekin } x \in \mathbb{R}^n\}.$$

Sarakeavaruus on matriisin A sarakevektoreiden virittämä avaruuden $\mathbb{R}^m$ aliavaruus. Tämä on osoitettu Leonin kirjassa [**3**].

MÄÄRITELMÄ 3.2. Olkoon $m \times n$ -matriisi A. Tällöin matriisin A *nolla-avaruus*, eli *ydin* $N(A)$, on joukko

$$N(A) = \{x \in \mathbb{R}^n : Ax = 0\}.$$

Tällöin ydin $N(A)$ on avaruuden $\mathbb{R}^n$ aliavaruus, kuten Leonin kirjassa on osoitettu [**3**].

Esitellään seuraavaksi hyödyllinen matriisin Dimensiolause sarakeavaruuden ja ytimen avulla. Dimensiolausetta tarvitaan myöhemmin useassa todistuksessa Luvuissa 4, 5 ja 6. Dimensiolauseen todistus löytyy Leonin kirjasta [**3**, s. 175-176].

LAUSE 3.3. *Olkoon $m \times n$ -matriisi A. Tällöin sarakeavaruuden ja ytimen dimensioille pätee*

$$\dim C(A) + \dim N(A) = n.$$

Seuraavaksi tarkastellaan matriisien A ja A^TA ytimien välistä yhteyttä. Matriisi A^TA on symmetrinen, mistä on hyötyä esimerkiksi ortogonaalisessa diagonalisoituvuudessa Luvussa 6. Näiden matriisien ytimille pätee seuraava lause:

LAUSE 3.4. *Olkoon $m \times n$ -matriisi A. Tällöin matriisien A ja A^TA ytimille pätee*

$$N(A^TA) = N(A).$$

TODISTUS. Osoitetaan ensin, että $N(A) \subset N(A^TA)$. Olkoon tätä varten vektori $x \in N(A)$. Tällöin ytimen Määritelmän 3.2 nojalla

$$Ax = 0,$$

jolloin kertomalla vasemmalta matriisilla A^T saadaan

$$A^TAx = 0.$$

Siispä nyt vektori x kuuluu myös ytimeen $N(A^TA)$, joten $N(A) \subset N(A^TA)$.

Osoitetaan sitten, että $N(A^TA) \subset N(A)$. Olkoon vektori $x \in N(A^TA)$. Tällöin ytimen Määritelmän 3.2 nojalla

$$A^TAx = 0.$$

Tarkastellaan Lauseen 1.6 mukaisesti vektorin Ax normin neliötä

$$\|Ax\|^2 = (Ax)^TAx = x^TA^TAx = x^T0 = 0,$$

koska $A^TAx = 0$. Siispä koska vektorin Ax normin neliö on 0, niin Lauseen 1.6 nojalla myös $Ax = 0$. Tällöin vektori $x \in N(A)$, joten myös $N(A^TA) \subset N(A)$. Näin ollen $N(A^TA) = N(A)$. □

Määritellään seuraavaksi matriisin A aste rank A, ja tarkastellaan siihen liittyviä tuloksia. Matriisin aste ja sen ominaisuudet ovat tässä työssä keskeisessä osassa erityisesti singulaariarvohajotelman ja perturbaatiolauseen todistuksissa. Matriisin astetta hyödynnetään myös matriisin CR-hajotelman muodostamisessa Luvussa 4.

MÄÄRITELMÄ 3.5. Olkoon $m \times n$ -matriisi A, jonka sarakeavaruuden $C(A)$ dimensio on r. Tällöin matriisin A *aste* on

$$\operatorname{rank} A = r.$$

Neliömatriisi, eli jokin $n \times n$ -matriisi A, on kääntyvä, jos ja vain jos sen asteelle pätee rank $A = n$. Todistetaan seuraavaksi kaksi lausetta matriisin asteelle, joiden todistuksessa sovelletaan Dimensiolausetta 3.3.

LAUSE 3.6. *Olkoon $m \times n$ -matriisi A. Tällöin*

$$\operatorname{rank} A = \operatorname{rank} A^TA.$$

TODISTUS. Matriisin dimensiolauseen 3.3 nojalla matriisille A pätee

$$\dim C(A) + \dim N(A) = n,$$

sekä matriisille A^TA pätee

$$\dim C(A^TA) + \dim N(A^TA) = n.$$

Siispä saadaan yhtäsuuruus

$$\dim C(A) + \dim N(A) = \dim C(A^TA) + \dim N(A^TA), \tag{3.1}$$

jossa Lauseen 3.4 nojalla $\dim N(A) = \dim N(A^TA)$, koska $N(A) = N(A^TA)$. Näin ollen edellisen tuloksen ja yhtäsuuruuden (3.1) nojalla pätee

$$\dim C(A) = \dim C(A^TA),$$

missä asteen Määritelmän 3.5 nojalla $\dim C(A) = \operatorname{rank} A$ ja $\dim C(A^TA) = \operatorname{rank} A^TA$, jolloin

$$\operatorname{rank} A = \operatorname{rank} A^TA.$$

□

Seuraavan lauseen toinen osuus todistetaan Lauseessa 4.8.

LAUSE 3.7. *Olkoot $m \times n$ -matriisi A ja kääntyvä $m \times m$ -matriisi U. Tällöin*

$$\text{rank } A = \text{rank } UA.$$

TODISTUS. Tarkastellaan aluksi matriisien A ja UA ytimiä. Matriisin ytimen Määritelmän 3.2 nojalla matriisin A ydin on

$$N(A) = \{x \in \mathbb{R}^n : Ax = 0\},$$

ja $m \times n$ -matriisin UA ydin on

$$N(UA) = \{x \in \mathbb{R}^n : UAx = 0\}.$$

Olkoon nyt vektori $x \in N(A)$, jolloin $Ax = 0$. Tällöin pätee myös $UAx = 0$, joten vektorille x saadaan $x \in N(UA)$. Siispä

$$N(A) \subseteq N(UA).$$

Olkoon sitten vektori $x \in N(UA)$, jolloin $UAx = 0$. Tällöin koska matriisi U on kääntyvä, niin matriisille Ax saadaan $Ax = 0$, joten vektorille x pätee $x \in N(A)$. Siispä matriisien ytimille pätee myös

$$N(UA) \subseteq N(A).$$

Näin ollen ytimille saadaan yhtäsuuruus $N(A) = N(UA)$.

Dimensiolauseen 3.3 nojalla $m \times n$ -matriiseille A ja UA pätee

$$\dim C(A) + \dim N(A) = n,$$

ja

$$\dim C(UA) + \dim N(UA) = n.$$

Siispä saadaan yhtäsuuruus

$$\dim C(A) + \dim N(A) = \dim C(UA) + \dim N(UA),$$

joten koska $\dim N(A) = \dim N(UA)$, niin myös

$$\dim C(A) = \dim C(UA).$$

Näin ollen asteen Määritelmän 3.5 nojalla

$$\text{rank } A = \text{rank } UA.$$

□

LUKU 4

CR-hajotelma

Tässä luvussa muodostetaan ja todistetaan matriisin CR-hajotelma, eli löydetään $m \times n$ -matriisille A esitys $A = CR$. Tässä $m \times r$ -matriisi C saadaan matriisin A lineaarisesti riippumattomien sarakkeiden avulla, ja $r \times n$ -matriisi R johdetaan redusoidun porrasmatriisin ja permutaatiomatriisin avulla. Matriisien dimensioissa luku r on matriisin A aste. CR-hajotelman hajotelman avulla saadaan esitettyä matriisi A monikäyttöisessä muodossa. Luvun lopussa todistetaankin CR-hajotelmaa apuna käyttäen lause siitä, että matriisin ja sen transpoosin asteet ovat yhtä suuret. Luvussa on käytetty päälähteenä Strangin ja Molerin artikkelia [**7**], määritelmissä on käytetty myös Leonin kirjaa [**3**], ja redusoidun porrasmatriisin määritelmän kertauksessa pohjana Lineaarinen algebra ja geometria -kurssin Kilpeläisen luentomonistetta [**2**]. CR-hajotelman todistusta on muokattu ja täydennetty huomattavasti artikkelin todistukseen verrattuna.

Määritellään ensin CR-hajotelman muodostamista varten redusoitu porrasmatriisi. Redusoitua porrasmatriisimuotoa tarvitaan myös todistuksissa Luvussa 8.

Määritelmä 4.1. Olkoon A nollasta poikkeava $m \times n$ -matriisi. Tällöin matriisi R_0 on *redusoitu porrasmatriisi* matriisille A, jos sille pätevät seuraavat ehdot:

1. matriisin mahdolliset nollarivit ovat matriisissa alimpina,
2. rivi on nollarivi tai jokaisen rivin ensimmäinen nollasta poikkeava alkio eli *johtava alkio* on ylemmän rivin johtavan alkion oikealla puolella, ja johtavana alkiona on 1,
3. jokainen johtava alkio on ainoa nollasta poikkeava alkio sarakkeessaan,
4. matriisi A saadaan muokattua matriisiksi R_0 alkeisrivioperaatioiden avulla. Alkeisrivioperaatioita ovat rivien vaihto, rivin kertominen nollasta poikkeavalla vakiolla sekä rivien lisääminen toisiinsa vakiolla kerrottuna.

Merkitään redusoitua porrasmatriisia $\operatorname{rref} A = R_0$.

Lineaarisesta algebrasta tiedetään, että mikä tahansa $m \times n$ -matriisi A voidaan esittää kääntyvän matriisin E^{-1} ja redusoidun porrasmatriisin R_0 avulla. Tässä matriisi E koostuu kaikista matriisille A tehdyistä alkeisrivioperaatioista. Tarkastellaan tätä seuraavassa lauseessa.

Lause 4.2. *Olkoon $m \times n$ -matriisi A. Tällöin matriisin A redusoitu porrasmatriisi R_0 saadaan kääntyvän matriisin E ja matriisin A tulona seuraavasti:*

$$R_0 = EA.$$

Määritellään seuraavaksi permutaatiomatriisi P, jota tarvitaan CR-hajotelman matriisin R muodostamisessa. Lähteenä määritelmässä on käytetty Leonin kirjaa [**3**, s. 268]. Sanotaan, että $k_1, k_2, \ldots, k_n$ on lukujen $1, 2, \ldots, n$ permutaatio, jos se määrittää lukujen $1, 2, \ldots, n$ järjestyksen. Esimerkiksi järjestys 2, 4, 1, 3, 5 on lukujen 1, 2, 3, 4, 5 eräs permutaatio.

MÄÄRITELMÄ 4.3. Olkoon $n \times n$ -identiteettimatriisi I, jonka sarakevektorit ovat $e_1, e_2, \ldots, e_n$ oikeassa järjestyksessä. *Permutaatiomatriisi* P saadaan muodostettua vaihtamalla identiteettimatriisin sarakkeiden paikkoja siten, että

$$P = \begin{bmatrix} e_{k_1} & e_{k_2} & \cdots & e_{k_n} \end{bmatrix},$$

missä $k_1, k_2, \ldots, k_n$ on lukujen $1, 2, \ldots, n$ jokin permutaatio.

Koska permutaatiomatriisi on muodostettu ortogonaalisen identiteettimatriisin sarakkeiden permutaation avulla, niin permutaatiomatriisi on myös ortogonaalinen. Permutaatiomatriisin määritelmän perusteella sen jokaisella rivillä ja jokaisessa sarakkeessa yksi alkio on 1 ja muut alkiot ovat nollia. Siispä kun kerrotaan permutaatiomatriisilla mitä tahansa matriisia oikealta, se vaihtaa kyseisen matriisin sarakkeiden paikkoja, pitäen kuitenkin sarakkeet muuttumattomina. Tarkastellaan tätä $m \times n$ -matriisin A avulla, kun

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}.$$

Kertomalla matriisia A oikealta permutaatiomatriisilla P saadaan

$$AP = \begin{bmatrix} Ae_{k_1} & Ae_{k_2} & \cdots & Ae_{k_n} \end{bmatrix} = \begin{bmatrix} a_{k_1} & a_{k_2} & \cdots & a_{k_n} \end{bmatrix}.$$

Tästä nähdään, että permutaatiomatriisi järjestelee matriisin A sarakkeet uuteen järjestykseen.

Tarkastellaan redusoidun porrasmatriisin ja permutaatiomatriisin käyttöä seuraavan esimerkin avulla.

ESIMERKKI 4.4. Olkoon 4×3 -matriisi A siten, että

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}.$$

Tällöin matriisin A aste on rank $A = 2$, ja sen redusoitu porrasmatriisi on

$$\operatorname{rref} A = R_0 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Muodostetaan eräs 3×3 -permutaatiomatriisi, jota tarvitaan seuraavassa Esimerkissä 4.6. Olkoon siis permutaatiomatriisi

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Lasketaan seuraavaksi jatkoa ajatellen keskeinen matriisitulo redusoidulle porrasmatriisille R_0 ja permutaatiomatriisille P, jolloin

$$R_0P = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Tätä tarvitaan myöhemmin Esimerkissä 4.6 matriisin CR-hajotelmaa varten.

Seuraavassa lauseessa tarkastellaan CR-hajotelmaa, ja todistuksena johdetaan se vaiheittain. Pohjana tässä todistuksessa on käytetty Strangin ja Molerin artikkelia [**7**]. Lauseessa matriisin A ensimmäiset r kappaletta lineaarisesti riippumattomia sarakkeita saadaan redusoidun porrasmatriisimuodon R_0 avulla, kun redusoidun porrasmatriisin johtavien alkioiden paikat ilmaisevat nämä matriisin A sarakkeet.

Lause 4.5. *Olkoon $m \times n$ -matriisi A, jonka aste on* rank $A = r$. *Matriisilla* $A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$ *on CR-hajotelma*

$$A = CR,$$

missä C on $m \times r$ -matriisi ja R on $r \times n$ -matriisi. Matriisi C, missä

$$C = \begin{bmatrix} a_{k_1} & a_{k_2} & \cdots & a_{k_r} \end{bmatrix},$$

koostuu sarakkeista, jotka ovat matriisin A ensimmäiset r lineaarisesti riippumatonta saraketta, ja kyseiset sarakevektorit muodostavat matriisin A sarakeavaruuden kannan. Matriisi R on muotoa

$$R = \begin{bmatrix} I & F \end{bmatrix} P^T,$$

missä identiteettimatriisi I on $r \times r$ -matriisi, F on $r \times (n-r)$ -matriisi ja P on $n \times n$ -permutaatiomatriisi.

Todistus. Muodostetaan ensin CR-hajotelman matriisi R, joka on $r \times n$ -matriisi. Matriisi R saadaan muodostettua matriisin A redusoidun porrasmatriisin R_0 sekä permutaatiomatriisin P avulla. Muokataan ensin redusoitu porrasmatriisi R_0 haluttuun lohkomuotoon siirtämällä sen sarakkeita permutaatiomatriisin P avulla. Permutaatiomatriisi P muodostetaan sellaiseksi, että kerrottaessa matriisia R_0 oikealta permutaatiomatriisilla, saadaan matriisi, joka koostuu $r \times r$ -identiteettimatriisista I, $r \times (n-r)$ -matriisista F, sekä näiden alapuolella olevista mahdollisista nollariveistä. Nyt tässä tulomatriisissa R_0P identiteettimatriisi on sen vasemmassa yläkulmassa ja matriisi F oikeassa yläkulmassa, eli matriisi R_0P on lohkomuodossa

$$R_0P = \begin{bmatrix} I_{r\times r} & F_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}. \tag{4.1}$$

Permutaatiomatriisi P on sen määritelmän nojalla ortogonaalinen. Siispä kerrotaan matriisia R_0P oikealta permutaatiomatriisin transpoosilla P^T, jolloin saadaan

$$R_0 = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} P^T.$$

Tällöin matriisi R saadaan poistamalla matriisista R_0 nollarivit, jolloin

$$R = \begin{bmatrix} I & F \end{bmatrix} P^T, \tag{4.2}$$

missä R on $r \times n$ -matriisi, koska matriisi $\begin{bmatrix} I & F \end{bmatrix}$ on $r \times n$ -matriisi, ja P^T on $n \times n$ -matriisi.

Muodostetaan sitten CR-hajotelman matriisi C, joka on $m \times r$ -matriisi. Matriisin R_0 johtavat alkiot määräävät ne matriisin A sarakkeet, jotka tulevat matriisiin C. Osoitetaan vielä, että todella pätee yhtäsuuruus $A = CR$. Kaavan (4.2) avulla tiedetään, että tulomatriisille CR pätee

$$CR = C \begin{bmatrix} I & F \end{bmatrix} P^T = \begin{bmatrix} C & CF \end{bmatrix} P^T.$$

Täytyy siis osoittaa, että

$$A = \begin{bmatrix} C & CF \end{bmatrix} P^T.$$

Lauseen 4.2 nojalla redusoitu porrasmatriisi R_0 voidaan esittää matriisin E^{-1} ja redusoidun matriisin A avulla muodossa

$$A = E^{-1} R_0. \tag{4.3}$$

Tässä E^{-1} on $m \times m$ -matriisi siten, että

$$E^{-1} = \begin{bmatrix} E_1^{-1} & E_2^{-1} \end{bmatrix},$$

missä matriisin E_1^1 dimensio on $m \times r$ ja matriisin E_2^{-1} dimensio on $m \times (m - r)$. Lisäksi R_0 on $m \times n$ -matriisi ja A on $m \times n$ -matriisi. Tarkastellaan matriisin A esitystä näiden lohkomuotojen avulla. Tässä matriisi A kirjoitetaan kaavan (4.3) mukaisesti ja tulo $R_0 P$ saadaan kaavan (4.1) avulla seuraavasti

$$\begin{aligned} AP = E^{-1} R_0 P &= E^{-1} \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} E_1^{-1} & E_2^{-1} \end{bmatrix} \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} E_1^{-1} & E_1^{-1} F \end{bmatrix}. \end{aligned} \tag{4.4}$$

Tästä muodosta nähdään, että matriisi E_1^{-1} on itse asiassa

$$E_1^{-1} = \begin{bmatrix} a_{k_1} & a_{k_2} & \cdots & a_{k_r} \end{bmatrix},$$

eli siinä on matriisin A ensimmäiset r lineaarisesti riippumatonta saraketta, joten se vastaa haluttua matriisia C. Siispä nyt Kaavan (4.4) yhtälö voidaan vielä muokata kertomalla oikealta matriisilla P^T ja ottamalla matriisi E_1^{-1} kerroinmatriisiksi seuraavasti:

$$A = E_1^{-1} \begin{bmatrix} I & F \end{bmatrix} P^T = C \begin{bmatrix} I & F \end{bmatrix} P^T = CR.$$

Näin ollen todellakin $A = CR$, missä C koostuu matriisin A ensimmäisistä r lineaarisesti riippumattomasta sarakkeesta ja R on muotoa $R = \begin{bmatrix} I & F \end{bmatrix} P^T$. □

Käydään läpi hajotelman muodostamista esimerkin avulla muodostamalla ja ratkaisemalla matriisit C ja R.

Esimerkki 4.6. Jatketaan Esimerkkiä 4.4. Olkoon siis 4×3 -matriisi A ja 3×3-permutaatiomatriisi P siten, että

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \qquad \text{ja} \qquad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Esimerkissä 4.4 saatiin tulomatriisi

$$R_0P = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

mistä nähdään, että se on CR-hajotelman halutussa muodossa, kun nyt matriisit

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{ja} \qquad F = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Siispä

$$R_0P = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I & F \\ 0_{2\times 2} & 0_{2\times 1} \end{bmatrix},$$

jolloin CR-hajotelman matriisi R saadaan kertomalla oikealta puolelta permutaatiomatriisilla P^T ja poistamalla nollarivit:

$$\begin{aligned} R &= \begin{bmatrix} I & F \end{bmatrix} P^T \\ &= \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Ratkaistaan seuraavaksi hajotelman matriisi C. Nyt redusoidusta porrasmatriisista

$$R_0 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

nähdään, että sarakkeet 1 ja 3 ovat lineaarisesti riippumattomat, koska kyseisissä sarakkeissa on rivien 1 ja 2 johtavat alkiot. Siispä matriisi C koostuu matriisin A sarakkeista 1 ja 3 seuraavasti

$$C = \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ 2 & 6 \\ 3 & 9 \end{bmatrix}.$$

Näin ollen CR-hajotelma matriisille A on

$$CR = \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ 2 & 6 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} = A.$$

Seuraavaksi tutustutaan lauseeseen matriisin asteen sekä sen transpoosin asteen yhtäsuudesta, ja todistetaan lause CR-hajotelman avulla.

Lause 4.7. *Olkoon A nollasta poikkeava $m \times n$ -matriisi. Tällöin*

$$\text{rank } A = \text{rank } A^T.$$

Todistus. Riittää osoittaa, että rank $A \geq$ rank A^T, koska tällöin sovellettaessa tulosta matriisille A^T saadaan myös epäyhtälö

$$\text{rank } A^T \geq \text{rank}(A^T)^T = \text{rank } A, \tag{4.5}$$

mistä saadaan asteille yhtäsuuruus.

Oletetaan, että rank $A = r$. Lauseen 4.5 mukaisesti matriisi A voidaan esittää muodossa $A = CR$. Tarkastellaan $n \times m$ -matriisia A^T, joka voidaan esittää CR-hajotelman avulla muodossa

$$A^T = (CR)^T = R^T C^T,$$

missä R^T on $n \times r$ -matriisi ja C^T on $r \times m$ -matriisi. Koska asteen Määritelmän 3.5 nojalla on

$$\text{rank } A^T = \dim C(A^T),$$

niin tarkastellaan matriisin A^T sarakeavaruutta

$$C(A^T) = C(R^T C^T).$$

Osoitetaan, että tälle sarakeavaruudelle pätee

$$C(A^T) = C(R^T C^T) \subset C(R^T).$$

Olkoon siis vektori $y \in \mathbb{R}^n$ siten, että $y \in C(R^T C^T)$. Tällöin vektorille y pätee Määritelmän 3.1 nojalla

$$y = R^T C^T x, \qquad \text{jollekin } x \in \mathbb{R}^m.$$

Tässä matriisitulon $C^T x$ dimensio on $r \times 1$. Siispä kun merkitään $z = C^T x \in \mathbb{R}^r$, niin vektorille y pätee myös

$$y = R^T z.$$

Tällöin sarakeavaruuden Määritelmän 3.1 nojalla vektori y kuuluu myös matriisin R^T sarakeavaruuteen, eli $y \in C(R^T)$, jolloin todella on voimassa

$$C(A^T) = C(R^T C^T) \subset C(R^T).$$

Osoitetaan sitten haluttu tulos. Dimensiolauseen 3.3 nojalla $n \times r$ -matriisille R^T pätee tulos

$$\dim C(R^T) + \dim N(R^T) = r.$$

Koska ytimen $N(R^T)$ dimensiolle pätee

$$\dim N(R^T) \geq 0,$$

niin Dimensiolauseen 3.3 perusteella täytyy olla, että

$$\dim C(R^T) \leq r. \tag{4.6}$$

Koska edellä ollaan osoitettu, että $C(A^T) \subset C(R^T)$, niin tämän tuloksen, kaavan (4.6) ja asteen Määritelmän 3.5 nojalla saadaan

$$\text{rank } A^T = \dim C(A^T) \leq \dim C(R^T) \leq r = \text{rank } A.$$

Näin ollen saadaan haluttu tulos

$$\text{rank } A \geq \text{rank } A^T.$$

Epäyhtälö pätee myös toiseen suuntaan, kuten kaavassa (4.5). Siispä voimassa on myös yhtäsuuruus

$$\text{rank } A = \text{rank } A^T.$$

□

Todistetaan seuraavaksi lemma, joka on sovellus Dimensiolauseesta 3.3. Lemman toinen osuus on todistettu Lauseessa 3.7, joten todistetaan tässä vain toinen osa, jota varten tarvitaan edellistä tulosta. Tätä lemmaa tarvitaan myöhemmin Luvussa 8, kun todistetaan matriisin perturbaatiolausetta.

LEMMA 4.8. *Olkoot $m \times n$ -matriisi A, kääntyvä $m \times m$ -matriisi U sekä kääntyvä $n \times n$ -matriisi V. Tällöin*

$$\text{rank } A = \text{rank } UA = \text{rank } AV.$$

TODISTUS. Yhtäsuuruus rank $A =$ rank UA on todistettu Lauseessa 3.7

Todistetaan vielä yhtäsuuruus rank $A =$ rank AV. Lauseen 4.7 nojalla saadaan

$$\text{rank } AV = \text{rank}(AV)^T = \text{rank } V^T A^T. \tag{4.7}$$

Osoitetaan, että tässä matriisi V^T on kääntyvä. Matriisille V^T käänteismatriisiksi sopii matriisi $(V^{-1})^T$, koska

$$(V^T (V^{-1})^T)^T = V^{-1} V = I = I^T,$$

koska matriisi V on kääntyvä. Tällöin $V^T (V^{-1})^T = I$, joten myös matriisi V^T on kääntyvä. Näin ollen kaavan (4.7) sekä Lauseiden 3.7 ja 4.7 perusteella pätee

$$\text{rank } V^T A^T = \text{rank } A^T = \text{rank } A,$$

joten todella

$$\text{rank } A = \text{rank } AV.$$

□

LAUSE 4.9. *Olkoon $m \times n$ -matriisi A, jonka redusoitu porrasmatriisi on* rref A. *Tällöin redusoidun porrasmatriisin* rref A *johtavien alkioiden lukumäärä on yhtä suuri, kuin matriisin A aste* rank A.

TODISTUS. Väite seuraa Lauseista 3.7 ja 4.2. □

LUKU 5

Ortogonaalikomplementti

Tässä luvussa kerrataan ortogonaalikomplementin määritelmä lineaarisesta algebrasta, sekä tutustutaan tuloksiin, missä ortogonaalikomplementtia hyödynnetään. Luvussa tarkastellaan ytimen ja sarakeavaruuden yhteyttä ortogonaalikomplementin avulla sekä aliavaruuteen ja sen ortogonaalikomplementtiin liittyviä tuloksia. Luvussa on käytetty lähteenä Leonin kirjaa [**3**].

Määritellään ensin aliavaruuden ortogonaalikomplementti aliavaruuden vektoreiden avulla.

MÄÄRITELMÄ 5.1. Olkoon S avaruuden $\mathbb{R}^n$ alivaruus. Tällöin joukko

$$S^\perp = \{x \in \mathbb{R}^n : x^T y = 0 \text{ kaikille } y \in S\}$$

on aliavaruuden S *ortogonaalikomplementti.*

Tässä ortogonalikomplementti $S^\perp$ on avaruuden $\mathbb{R}^n$ aliavaruus, mikä on osoitettu Leonin kirjassa [**3**]. Määritelmän nojalla seuraa, että joukko $S^\perp$ sisältää näin ollen kaikki avaruuden $\mathbb{R}^n$ vektorit, jotka ovat ortogonaalisia avaruuden S vektoreiden kanssa. Muistetaan, että $x^T y = x \cdot y$.

Tarkastellaan seuraavaksi matriisin ytimen ja sarakeavaruuden yhteyttä ortogonaalikomplementtien avulla.

LAUSE 5.2. *Olkoon $m \times n$ -matriisi A. Tällöin pätee yhtäsuuruudet*

$$N(A) = C(A^T)^\perp$$

ja

$$N(A^T) = C(A)^\perp.$$

TODISTUS. Yhtälö $N(A^T) = C(A)^\perp$ saadaan osoitettua soveltamalla ensimmäistä yhtälöä matriisille A^T. Osoitetaan näin ollen ainoastaan yhtälö $N(A) = C(A^T)^\perp$. Osoitetaan ensin, että $N(A) \subset C(A^T)^\perp$. Olkoon siis vektori $x \in N(A)$, jolloin ytimen Määritelmän 3.2 nojalla $Ax = 0$. Lisäksi ortogonaalikomplementin Määritelmän 5.1 nojalla sarakeavaruuden $C(A^T)$ ortogonaalikomplementti on

$$C(A^T)^\perp = \{v \in \mathbb{R}^n : v^T y = 0 \text{ kaikille } y \in C(A^T)\}.$$

Olkoon siis avaruuden $C(A^T)$ vektori $y = A^T z$, missä vektori $z \in \mathbb{R}^m$. Vektori x kuuluu sarakeavaruuden ortogonaalikomplementtiin, jos $x^T y = 0$. Siispä tarkastellaan matriisia Ax, mikä voidaan kirjoittaa lohkomuodossa

$$Ax = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} x = \begin{bmatrix} a_1 x \\ a_2 x \\ \vdots \\ a_m x \end{bmatrix},$$

missä alkioiden $a_i, i = 1, 2, \dots, m$ dimensio on $1 \times n$, koska ne ovat matriisin A rivejä. Koska tulolle Ax pätee $Ax = 0$, niin saadaan kaikille

$$a_i x = 0, \text{ kun } i = 1, 2, \dots, m. \tag{5.1}$$

Lasketaan edellisen avulla tulo $y^T x = (x^T y)^T = x^T y$:

$$\begin{aligned} y^T x = (A^T z)^T x = z^T A x &= \begin{bmatrix} z_1 & z_2 & \cdots & z_m \end{bmatrix} \begin{bmatrix} a_1 x \\ a_2 x \\ \vdots \\ a_m x \end{bmatrix} \\ &= z_1 a_1 x + z_2 a_2 x + \cdots + z_m a_m x \\ &= z \cdot 0 + z \cdot 0 + \cdots + z \cdot 0 = 0. \end{aligned}$$

Näin ollen kaavan (5.1) nojalla saadaan $y^T x = x^T y = 0$, joten vektorille x pätee $x \in C(A^T)^\perp$. Siispä tämä osoittaa, että $N(A) \subset C(A^T)^\perp$.

Osoitetaan seuraavaksi, että joukoille pätee $C(A^T)^\perp \subset N(A)$. Olkoon vektori $x \in C(A^T)^\perp$. Tällöin ortogonaalikomplementin Määritelmän 5.1 nojalla $x^T y = 0$ kaikille $y \in C(A^T)$. Olkoon $y = A^T z$, missä vektorille z pätee $z \in \mathbb{R}^m$. Tällöin

$$\begin{aligned} x^T y &= x^T A^T z \\ &= (Ax)^T z = 0 \qquad \text{kaikilla} \qquad z \in \mathbb{R}^m. \end{aligned}$$

Valitaan $z = Ax$, jolloin edellisen kaavan ja Lauseen 1.6 nojalla saadaan

$$0 = (Ax)^T z = (Ax)^T (Ax) = \|Ax\|^2 .$$

Tällöin Lauseen 1.6 nojalla myös $Ax = 0$. Siispä vektorille x pätee $x \in N(A)$, joten sarakeavaruudelle ja ytimelle saadaan $C(A^T)^\perp \subset N(A)$.

□

Seuraavassa lauseessa tutustutaan vielä avaruuden $\mathbb{R}^n$ aliavaruuden ja sen ortogonaalikomplementin dimensioiden summaan, sekä muodostetaan niiden ortonormaaleista kannoista avaruuden $\mathbb{R}^n$ ortonormaali kanta. Lauseen todistuksessa on käytetty pohjana Leonin kirjaa [**3**].

Lause 5.3. *Olkoon S avaruuden $\mathbb{R}^n$ aliavaruus, Tällöin*

$$\dim S + \dim S^\perp = n.$$

Toisaalta, jos joukko $\{x_1, x_2, \dots, x_r\}$ on avaruuden S ortonormaali kanta ja joukko $\{x_{r+1}, x_{r+2}, \dots, x_n\}$ on avaruuden $S^\perp$ ortonormaali kanta, niin joukko

$$\{x_1, x_2, \dots, x_r, x_{r+1}, \dots, x_n\}$$

on avaruuden $\mathbb{R}^n$ ortonormaali kanta.

Todistus. Todistetaan ensin lauseen ensimmäinen osuus. Olkoon aliavaruuden $S \subset \mathbb{R}^n$ kanta $\{x_1, x_2, \dots, x_r\}$, jolloin $\dim S = r$. Muodostetaan matriisi A, joka koostuu näistä kantavektoreista:

$$A = \begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix},$$

jolloin A on $n \times r$ -matriisi. Nyt matriisin A sarakeavaruus on joukko

$$C(A) = \langle x_1, x_2, \dots, x_r \rangle = S,$$

koska joukko $\{x_1, x_2, \ldots, x_r\}$ on avaruuden S kanta. Tällöin sarakeavaruuden ortogonaalikomplementille pätee

$$C(A)^\perp = \langle x_1, x_2, \ldots, x_r \rangle^\perp = S^\perp.$$

Nyt Lauseen 5.2 nojalla saadaan yhtäsuuruus $C(A)^\perp = N(A^T)$, ja Määritelmän 3.5 nojalla pätee rank $A = \dim C(A)$, joten

$$\begin{aligned} \dim S + \dim S^\perp &= \dim C(A) + \dim C(A)^\perp \\ &= \text{rank } A + \dim N(A^T). \end{aligned} \tag{5.2}$$

Lisäksi Lauseen 4.7 perusteella pätee rank $A =$ rank A^T, ja Määritelmän 3.5 perusteella pätee rank $A^T = \dim C(A^T)$, joten Kaava (5.2) saadaan muotoon

$$\begin{aligned} \dim S + \dim S^\perp &= \text{rank } A^T + \dim N(A^T) \\ &= \dim C(A^T) + \dim N(A^T). \end{aligned}$$

Lisäksi nyt matriisin Dimensiolauseen 3.3 nojalla saadaan

$$\begin{aligned} \dim S + \dim S^\perp &= \dim C(A^T) + \dim N(A^T) \\ &= n. \end{aligned}$$

Näin ollen päästiin tulokseen $\dim S + \dim S^\perp = n$, joten lauseen alkuosa on osoitettu.

Todistetaan sitten lauseen jälkimmäinen osuus. Koska kannan Määritelmän 1.9 mukaisesti ortonormaali vektorijoukko, jossa on n kappaletta vektoreita, on avaruuden $\mathbb{R}^n$ kanta, niin riittää osoittaa, että joukko

$$\{x_1, x_2, \ldots, x_r, x_{r+1}, \ldots, x_n\}$$

on ortonormaali. Koska joukot $\{x_1, x_2, \ldots, x_r\}$ ja $\{x_{r+1}, x_{r+2}, \ldots, x_n\}$ ovat ortonormaaleja, niin kummassakin joukossa olevien vektoreiden pituus on 1. Siispä myös kaikkien joukon $\{x_1, x_2, \ldots, x_r, x_{r+1}, \ldots, x_n\}$ vektoreiden pituus on 1.

Tämän lisäksi joukon ortonormaalisuutta varten osoitetaan, että kaikki joukon $\{x_1, x_2, \ldots, x_r, x_{r+1}, \ldots, x_n\}$ vektorit ovat ortogonaalisia eli kohtisuorassa toisiaan vastaan Määritelmän 1.3 mukaisesti, eli tulee olla $x_i \cdot x_j = 0$ kaikille $1 \leq i, j \leq n$, $i \neq j$. Käydään läpi mahdolliset tapaukset:

1. Olkoon $1 \leq i, j \leq r, i \neq j$. Olkoot siis vektorit $x_i, x_j \in \{x_1, x_2, \ldots, x_r\}$. Koska joukko $\{x_1, x_2, \ldots, x_r\}$ on ortonormaali, niin vektoreiden sisätulolle pätee

$$x_i \cdot x_j = 0.$$

2. Olkoon $r + 1 \leq i, j \leq n, i \neq j$. Olkoot siis vektorit $x_i, x_j \in \{x_{r+1}, x_{r+2}, \ldots, x_n\}$. Koska joukko $\{x_{r+1}, x_{r+2}, \ldots, x_n\}$ on ortonormaali, niin vektoreiden sisätulolle pätee

$$x_i \cdot x_j = 0.$$

3. Olkoon $i \leq r < r + 1 \leq j$. Olkoot siis vektorit $x_i \in \{x_1, x_2, \ldots, x_r\}$ ja $x_j \in \{x_{r+1}, x_{r+2}, \ldots, x_n\}$. Tällöin vektori $x_j \in \langle x_1, x_2, \ldots, x_r \rangle^\perp$, joten sille pätee ortogonaalikomplementin Määritelmän 5.1 nojalla

$$x_j^T x_i = x_i^T x_j = x_i \cdot x_j = 0.$$

4. Olkoon $j \leq r < r+1 \leq i$. Olkoot siis vektorit $x_i \in \{x_{r+1}, x_{r+2}, \dots, x_n\}$ ja $x_j \in \{x_1, x_2, \dots, x_r\}$. Tällöin vektori $x_i \in \langle x_1, x_2, \dots, x_r \rangle^{\perp}$, joten sille pätee ortogonaalikomplementin Määritelmän 5.1 nojalla

$$x_i^T x_j = x_i \cdot x_j = 0.$$

Siispä nyt kaikissa tapauksissa pätee $x_i \cdot x_j = 0$. Näin ollen joukko

$$\{x_1, x_2, \dots, x_r, x_{r+1}, \dots, x_n\} \subset \mathbb{R}^n$$

on ortonormaali joukko, jossa on n kappaletta vektoreita, joten se on myös avaruuden $\mathbb{R}^n$ ortonormaali kanta. $\square$

LUKU 6

Symmetrisen matriisin diagonalisointi

Tässä luvussa tarkastellaan matriisin ominaisarvoa ja ominaisvektoria, sekä niiden avulla ortogonaalisesti diagonalisoituvia symmetrisiä matriiseja. Muodostetaan siis diagonalisoimalla symmetriselle matriisille A hajotelma $A = VDV^T$, missä V on ortogonaalinen matriisi ja D on diagonaalimatriisi. Käydään myös symmetrisen matriisin diagonalisoituvuutta läpi esimerkin avulla. Lähteinä tässä luvussa on käytetty Leonin, Strangin sekä Saarimäen kirjoja [**3**], [**6**], [**5**].

Määritellään aluksi matriisin ominaisarvo ja ominaisarvoa vastaava ominaisvektori.

MÄÄRITELMÄ 6.1. Olkoon $n \times n$ -matriisi A. Jos on olemassa vektori $x \in \mathbb{R}^n$, $x \neq 0$, ja luku $\lambda \in \mathbb{R}$ siten, että

$$Ax = \lambda x,$$

niin luku λ on matriisin A *ominaisarvo* ja vektori x on tätä ominaisarvoa vastaava *ominaisvektori*.

Määritellään seuraavaksi symmetrisen matriisin ortogonaalinen diagonalisoituvuus.

MÄÄRITELMÄ 6.2. Olkoon A symmetrinen $n \times n$ -matriisi. Tällöin matriisi A on *ortogonaalisesti diagonalisoituva*, jos on ortogonaalinen matriisi V ja diagonaalimatriisi D siten, että

$$A = VDV^T.$$

Esitellään seuraavaksi symmetrisille matriiseille lemma, jonka avulla todistetaan sen jälkeinen lause. Lemman todistus löytyy Saarimäen kirjasta [**5**, s. 34-35].

LEMMA 6.3. *Olkoon A symmetrinen matriisi. Tällöin sillä on ainakin yksi ominaisarvo.*

Seuraavassa lauseessa perehdytään symmetrisen matriisin diagonalisoituvuuteen. Lauseen todistuksessa on käytetty lähteenä Saarimäen kurssikirjaa [**5**].

LAUSE 6.4. *Olkoon A symmetrinen matriisi. Tällöin A on ortogonaalisesti diagonalisoituva.*

TODISTUS. Todistetaan väite induktion avulla.

Tapauksessa $n = 1$ matriisi $A = \begin{bmatrix} a \end{bmatrix}$ voidaan ajatella 1×1 -matriisina, jonka ainoa ominaisarvo on a ja eräs ominaisvektori 1. Tällöin matriisi A voidaan esittää muodossa

$$A = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} a \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = VDV^T,$$

jolloin matriisi $V = \begin{bmatrix} 1 \end{bmatrix}$ on ortogonaalinen matriisi ja matriisi $D = \begin{bmatrix} a \end{bmatrix}$ on diagonaalimatriisi. Näin ollen väite pätee tapauksessa $n = 1$.

Oletetaan, että väite pätee mille tahansa symmetriselle matriisille, jonka dimensio on

$$(n-1) \times (n-1), \qquad \text{kun } n \geq 2.$$

Olkoon sitten A symmetrinen $n \times n$ -matriisi. Tällöin sillä on Lemman 6.3 mukaan ainakin yksi ominaisarvo λ_1. Olkoon yksikkövektori v_1 ominaisarvoa λ_1 vastaava ominaisvektori. Laajennetaan v_1 avaruuden $\mathbb{R}^n$ ortonormaaliksi kannaksi $K = \{v_1, v_2, \ldots, v_n\}$, jolloin $\langle v_1 \rangle^\perp = \langle v_2, v_3, \ldots, v_n \rangle$. Valitaan tarkasteluun ortonormaaleista vektoreista koostuva ortogonaalinen matriisi

$$V_1 = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}.$$

Toisaalta ominaisarvojen Määritelmän 6.1 nojalla $Av_1 = \lambda_1 v_1$, joten tällöin saadaan

$$v_i^T A v_1 = v_i^T \lambda_1 v_1 = \lambda_1 v_i^T v_1 = 0, \qquad \text{kun } i = 2, 3, \ldots, n,$$

koska $v_i^T v_1 = 0$ ortonormaaleille vektoreille, kun $i \neq 1$. Sen sijaan koska reaaliluvun transpoosi on luku itse, ja symmetriselle matriisille $A = A^T$, niin

$$\begin{aligned} v_1^T A v_i &= (v_1^T A v_i)^T = v_i^T (v_1^T A)^T \\ &= v_i^T A^T v_1 = v_i^T A v_1 \\ &= v_i^T \lambda_1 v_1 = \lambda_1 v_i^T v_1 = 0, \qquad \text{kun } i = 2, 3, \ldots, n. \end{aligned}$$

Lisäksi pätee

$$v_1^T A v_1 = v_1^T \lambda_1 v_1 = \lambda_1 v_1^T v_1 = \lambda_1,$$

joten matriisi $V_1^T A V_1$ on edellisten nojalla lohkomuodossa esitettynä

$$\begin{aligned} V_1^T A V_1 &= \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} A \begin{bmatrix} v_1 & v_2 \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \begin{bmatrix} Av_1 & Av_2 \cdots & Av_n \end{bmatrix} \\ &= \begin{bmatrix} v_1^T A v_1 & v_1^T A v_2 & \cdots & v_1^T A v_n \\ v_2^T A v_1 & v_2^T A v_2 & \cdots & v_2^T A v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n^T A v_1 & v_n^T A v_2 & \cdots & v_n^T A v_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & v_2^T A v_2 & \cdots & v_2^T A v_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & v_n^T A v_2 & \cdots & v_n^T A v_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_2 \end{bmatrix}, \end{aligned} \tag{6.1}$$

missä

$$A_2 = \begin{bmatrix} v_2^T A v_2 & v_2^T A v_3 & \cdots & v_2^T A v_n \\ v_3^T A v_2 & v_3^T A v_3 & \cdots & v_3^T A v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n^T A v_2 & v_n^T A v_3 & \cdots & v_n^T A v_n \end{bmatrix}.$$

Matriisi A_2 on symmetrinen, koska lohkomatriisin laskusääntöjen nojalla sen transpoosi on

$$A_2^T = \begin{bmatrix} (v_2^T A v_2)^T & (v_3^T A v_2)^T & \cdots & (v_n^T A v_2)^T \\ (v_2^T A v_3)^T & (v_3^T A v_3)^T & \cdots & (v_n^T A v_3)^T \\ \vdots & \vdots & \ddots & \vdots \\ (v_2^T A v_n)^T & (v_3^T A v_n)^T & \cdots & (v_n^T A v_n)^T \end{bmatrix},$$

missä

$$(v_i^T A v_j)^T = v_j^T (v_i^T A)^T = v_j^T A^T v_i = v_j^T A v_i \qquad \text{kaikilla } 2 \leq i, j \leq n$$

Tässä $A^T = A$, koska matriisi A on oletuksen nojalla symmetrinen. Siispä nyt pätee $A_2^T = A_2$, joten myös matriisi A_2 on symmetrisen matriisin Määritelmän 1.1 nojalla symmetrinen.

Koska symmetrisen matriisin A_2 dimensio on $(n-1) \times (n-1)$, niin induktio-oletuksen nojalla se on ortogonaalisesti diagonalisoituva. Siten on olemassa ortogonaalinen $(n-1) \times (n-1)$ -matriisi V_2 siten, että $V_2^T A_2 V_2 = D_2$ on $(n-1) \times (n-1)$ -diagonaalimatriisi. Valitaan nyt matriisi

$$V = V_1 \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix}.$$

Matriisin V_1 symmetrisyyden nojalla $V_1^T V_1 = I_n$. Nyt matriisin V dimensio on $n \times n$, ja V on ortogonaalinen Määritelmän 1.4 nojalla, koska

$$\begin{aligned} V^T V &= \begin{bmatrix} 1 & 0 \\ 0 & V_2^T \end{bmatrix} V_1^T V_1 \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & V_2^T \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & V_2^T V_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I_{n-1} \end{bmatrix} = I_n. \end{aligned}$$

Tällöin lohkomatriisien tulon ominaisuuksien sekä kaavan (6.1) avulla saadaan myös

$$\begin{aligned} V^T A V &= \begin{bmatrix} 1 & 0 \\ 0 & V_2^T \end{bmatrix} V_1^T A V_1 \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & V_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & V_2^T A_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & V_2^T A_2 V_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & D_2 \end{bmatrix}, \end{aligned}$$

mikä on diagonaalimatriisi, koska matriisi D_2 on diagonaalimatriisi. Siispä koska matriisi V on ortogonaalinen matriisi ja $V^T A V$ on diagonaalimatriisi, ja lisäksi matriisille A pätee $A = V D V^T$, niin A on ortogonaalisesti diagonalisoituva. □

Kun tarkastellaan symmetrisen matriisin A diagonalisointia, diagonaalimatriisin D diagonaalialkioita ovat matriisin A ominaisarvot $\lambda_1, \lambda_2, \ldots, \lambda_n$ ja ortogonaalisen matriisin V sarakkeet koostuvat matriisin A ominaisarvoja vastaavista ominaisvektoreista $v_1, v_2, \ldots, v_n$. Tässä ominaisarvot $\lambda_i, i = 1, 2, \ldots, n$, ovat ei-negatiivisia. Siispä matriisit D ja V ovat muotoa

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \qquad ja \qquad V = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}.$$

Tarkastellaan vielä ortogonaalista diagonalisoituvuutta esimerkin avulla. Käytetään esimerkissä symmetristä matriisia, joka koostuu tulona matriiseista A^T ja A, joita käytetään myöhemmin myös Esimerkeissä 7.4, 7.6 ja 8.3.

Esimerkki 6.5. Diagonalisoidaan symmetrinen matriisi A^TA , missä

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

Tällöin matriisi A^TA on

$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix},$$

jonka ominaisarvot $\lambda \in \mathbb{R}$ saadaan Määritelmän 6.1 avulla. Siispä tarkastellaan determinanttia $\det(A - \lambda I)$ yhtälöä $A^TAx = \lambda x$ varten seuraavasti:

$$\begin{vmatrix} 2-\lambda & 4 \\ 4 & 8-\lambda \end{vmatrix} = (2-\lambda)(8-\lambda) - 16 = \lambda^2 - 10\lambda.$$

Tällöin karakteristisen polynomin $\lambda^2 - 10\lambda = 0$ ratkaisut ovat $\lambda_1 = 10$ ja $\lambda_2 = 0$, mitkä ovat matriisin A^TA ominaisarvot. Ominaisarvoja vastaavat eräät ominaisvektorit ovat

$$x_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{ja} \qquad x_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

joiden normit ovat

$$\|x_1\| = \sqrt{1^2 + 2^2} = \sqrt{5}$$

ja

$$\|x_2\| = \sqrt{(-2)^2 + 1^2} = \sqrt{5}.$$

Tällöin diagonalisointiin voidaan käyttää ortogonaalista matriisia, jossa ominaisvektorit on normeerattu ortonormaaleiksi, eli matriisiksi V voidaan valita

$$V = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

Matriisi V on ortogonaalinen Määritelmän 1.4 mukaisesti, koska

$$V^TV = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Siispä nyt kun diagonaalimatriisi D on muotoa

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix},$$

niin saadaan matriisituloksi

$$VDV^T = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} = A^TA.$$

Näin ollen ollaan saatu ortogonaalisesti diagonalisoitua matriisi A^TA Määritelmän 6.2 mukaisesti.

LUKU 7

Singulaariarvohajotelma

Tässä luvussa esitetään matriisin A singulaariarvohajotelma $A = U\Sigma V^T$ sekä todistetaan singulaariarvohajotelmalause kaikille $m \times n$ -matriiseille. Singulaariarvohajotelmassa ja sen lauseen todistuksessa on käytetty tärkeimpänä lähteenä Leonin kirjaa [**3**]. Leonin kirjassa esiintyvää todistusta on täsmennetty, ja lisäksi todistusta varten käytetään useita aiemmin osoitettuja tuloksia.

Huomataan, että Lauseen 6.4 seurauksena symmetrisellä matriisilla A^TA on n kappaletta ominaisarvoja kertaluvut huomioiden. Osoitetaankin aluksi lemma symmetrisen matriisin ominaisarvoista.

LEMMA 7.1. *Olkoon symmetrinen $n \times n$ -matriisi A^TA, jonka ominaisarvot ovat $\lambda_1, \lambda_2, \ldots, \lambda_n$. Tällöin ominaisarvot $\lambda_i, i = 1, 2, \ldots, n$ ovat epänegatiivisia.*

TODISTUS. Olkoon λ matriisin A^TA eräs ominaisarvo ja x sitä vastaava ominaisvektori. Tällöin Määritelmän 6.1 nojalla matriisille A^TA pätee

$$A^TAx = \lambda x.$$

Tällöin normin neliölle $\|Ax\|^2$ saadaan Lauseen 1.6 ja Määritelmän 6.1 avulla

$$\begin{aligned} \|Ax\|^2 &= (Ax)^T(Ax) = x^TA^TAx = x^T\lambda x \\ &= \lambda x^Tx = \lambda \|x\|^2 . \end{aligned} \tag{7.1}$$

Siispä ominaisarvoa λ voidaan arvioida

$$\lambda = \frac{\|Ax\|^2}{\|x\|^2} \geq 0, \tag{7.2}$$

koska normin neliöt ovat epänegatiivisia. Näin ollen kaikki matriisin A^TA ominaisarvot on osoitettu epänegatiivisiksi. □

Määritellään seuraavaksi mille tahansa $m \times n$ -matriisille A singulaariarvot matriisin A^TA ominaisarvojen avulla. Määritelmässä ominaisarvot on hyvin määriteltyjä Lemman 7.1 nojalla.

MÄÄRITELMÄ 7.2. Olkoon $m \times n$ -matriisi A ja olkoot symmetrisen matriisin A^TA ominaisarvot $\lambda_1, \lambda_2, \ldots, \lambda_n$. Matriisin A *singulaariarvot* $\sigma_1, \sigma_2, \ldots, \sigma_n$ ovat

$$\sigma_j = \sqrt{\lambda_j}, \ j = 1, \ldots, n.$$

Tarkastellaan seuraavaksi tulosta matriisin A singulaariarvohajotelmalle $A = U\Sigma V^T$. Lauseen todistuksessa on pääpiirteittäin käytetty lähteenä Leonin kirjan [**3**] todistusta, jota on täydennetty ja perusteltu täsmällisemmin.

Lause 7.3. *Olkoot $m\times n$ -matriisi A, jonka aste on r ja jonka singulaariarvot ovat $\sigma_1, \sigma_2, \ldots, \sigma_n$ laskevassa järjestyksessä. Tällöin matriisi A voidaan esittää muodossa*

$$A = U\Sigma V^T,$$

missä U on ortogonaalinen $m\times m$ -matriisi, V on ortogonaalinen $n\times n$ -matriisi ja Σ on $m\times n$ -diagonaalimatriisi. Matriisin Σ diagonaalialkiot ovat matriisin A singulaariarvot, joille pätee

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \qquad ja \qquad \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0.$$

Matriisi Σ on muotoa

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix} & 0 \\ 0 & 0 \end{bmatrix}.$$

Hajotelmaa kutsutaan matriisin A singulaariarvohajotelmaksi.

Todistus. Tarkastellaan matriisia A^TA, missä A on $m\times n$ -matriisi. Tällöin matriisi A^TA on on $n\times n$ -neliömatriisi. Matriisin transpoosin laskusääntöjen nojalla matriisin A^TA transpoosiksi saadaan

$$(A^TA)^T = A^T(A^T)^T = A^TA.$$

Siispä $A^TA = (A^TA)^T$, joten matriisi A^TA on symmetrinen. Tällöin Lauseen 6.2 nojalla matriisi A^TA on ortogonaalisesti diagonalisoituva. Lisäksi Lemmassa 7.1 on osoitettu, että matriisin A^TA ominaisarvot ovat epänegatiiviset.

Symmetrinen matriisi A^TA voidaan diagonalisoida Määritelmän 6.2 mukaisesti, eli matriisi A^TA on muotoa $A^TA = VDV^T$, missä ortogonaalisen matriisin V ja diagonaalimatriisin D sarakkeet on järjestelty siten, että ominaisvektorit ja ominaisarvot ovat vastaavassa järjestyksessä, ja siten, että ominaisarvoille pätee

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

Merkitään matriisin A astetta $\operatorname{rank} A = r$. Tällöin Lauseen 3.6 nojalla $\operatorname{rank} A^TA = \operatorname{rank} A = r$, eli myös matriisin A^TA aste on r. Koska matriisi A^TA on symmetrinen, niin sen aste on diagonaalimuodon $A^TA = VDV^T$ seurauksena Lemman 4.8 nojalla sama, kuin matriisin A^TA positiivisten ominaisarvojen määrä. Siispä

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$$

ja

$$\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0.$$

Singulaariarvojen Määritelmän 7.2 nojalla $\sigma_j = \sqrt{\lambda_j}$, joten myös niille pätee

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \tag{7.3}$$

ja

$$\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0.$$

Olkoot lohkomatriisit V_1, V_2 ja Σ_1 siten, että

$$V_1 = \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix}, \qquad V_2 = \begin{bmatrix} v_{r+1} & v_{r+2} & \cdots & v_n \end{bmatrix},$$

jolloin matriisi V voidaan kirjoittaa lohkomuodossa

$$V = \begin{bmatrix} V_1 & V_2 \end{bmatrix},$$

ja

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix}.$$

Tällöin matriisi Σ_1 on $r \times r$ -diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin A aidosti positiiviset singulaariarvot $\sigma_1, \sigma_2, \ldots, \sigma_r$. Näin ollen singulaariarvohajotelman $m \times n$ -matriisi Σ on lohkomatriisi, joka on muotoa

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix}.$$

Lisäksi matriisin V_2 sarakevektorit $v_{r+1}, v_{r+2}, \ldots, v_n$ ovat matriisin $A^T A$ ominaisarvoa 0 vastaavia ominaisvektoreita. Siispä ominaisarvojen Määritelmän 6.1 nojalla

$$A^T A v_j = 0 v_j = 0, \qquad \text{kun } j = r+1, r+2, \ldots, n,$$

joten vektorit $v_j, j = r+1, r+2, \ldots, n$, kuuluvat ytimen Määritelmän 3.2 nojalla matriisin $A^T A$ ytimeen. Lauseen 3.4 nojalla pätee yhtäsuuruus

$$N(A^T A) = N(A),$$

joten ytimen Määritelmän 3.2 nojalla myös matriisitulolle Av_i pätee $Av_i = 0$ kaikille $i = r+1, r+2, \ldots n$. Siispä matriisitulolle AV_2 saadaan lohkomatriisien avulla

$$\begin{aligned} AV_2 &= A \begin{bmatrix} v_{r+1} & v_{r+2} & \cdots & v_n \end{bmatrix} \\ &= \begin{bmatrix} Av_{r+1} & Av_{r+2} & \cdots & Av_n \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix} = 0_{m\times(n-r)}. \end{aligned} \tag{7.4}$$

Koska V on ortogonaalinen matriisi, niin Määritelmän 1.4 nojalla $V^T V = I_{n\times n}$, missä $I_{n\times n}$ on identtinen matriisi. Lisäksi matriisi V on kääntyvä, joten $VV^{-1} = I_{n\times n}$, jolloin

$$VV^T = (VV^T)(VV^{-1}) = V(V^T V)V^{-1} = VIV^{-1} = VV^{-1} = I_{n\times n}.$$

Lisäksi, koska matriisi V muodostuu lohkomatriiseista V_1 ja V_2, niin lohkomatriisien tulon ja transpoosin laskusääntöjen nojalla

$$I = VV^T = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = V_1 V_1^T + V_2 V_2^T.$$

Tällöin matriisille A saadaan edellisen avulla esitys

$$\begin{aligned} A &= AI = A(V_1 V_1^T + V_2 V_2^T) \\ &= AV_1 V_1^T + AV_2 V_2^T = AV_1 V_1^T, \end{aligned} \tag{7.5}$$

missä $AV_2 V_2^T = 0$, koska yhtälöstä (7.4) nähdään, että $AV_2 = 0$.

Osoitetaan sitten, miten konstruoidaan ortogonaalinen $m \times m$ -matriisi U siten, että $A = U\Sigma V^T$. Olkoon matriisi U_1 siten, että

$$U_1 = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix},$$

missä $u_1, u_2, \dots, u_r \in \mathbb{R}^m$ ovat sarakevektoreita, jotka ovat muotoa

$$u_j = \frac{1}{\sigma_j} A v_j, \text{ kun } j = 1, 2, \dots, r. \tag{7.6}$$

Tällöin matriisitulolle $U_1 \Sigma_1$ pätee kaavan (7.6) ja matriisin V_1 muotoilun nojalla

$$\begin{aligned} U_1 \Sigma_1 &= \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix} \\ &= \begin{bmatrix} u_1 \sigma_1 & u_2 \sigma_2 & \cdots & u_r \sigma_r \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sigma_1} A v_1 \sigma_1 & \frac{1}{\sigma_2} A v_2 \sigma_2 & \cdots & \frac{1}{\sigma_r} A v_r \sigma_r \end{bmatrix} \\ &= \begin{bmatrix} A v_1 & A v_2 & \cdots & A v_r \end{bmatrix} \\ &= A \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix} = A V_1. \end{aligned} \tag{7.7}$$

Osoitetaan seuraavaksi, että matriisin U_1 sarakevektorit muodostavat ortonormaalin joukon. Vektoreille u_i ja u_j pätee kaavan (7.6) mukaisesti

$$\begin{aligned} u_i^T u_j &= \left(\frac{1}{\sigma_i} A v_i \right)^T \left(\frac{1}{\sigma_j} A v_j \right) \\ &= \frac{1}{\sigma_i} \left(v_i^T A^T \right) \frac{1}{\sigma_j} (A v_j) \\ &= \frac{1}{\sigma_i \sigma_j} v_i^T (A^T A v_j), \end{aligned} \tag{7.8}$$

missä $1 \leq i \leq r$ ja $1 \leq j \leq r$. Tässä vektorit v_j ovat matriisin $A^T A$ ominaisvektoreita, jotka vastaavat ominaisarvoja λ_j, joille pätee

$$\lambda_j = \sigma_j^2 > 0, \text{ kun } j = 1, 2, \dots, r.$$

Tällöin ominaisarvojen Määritelmän 6.1 sekä singulaariarvojen Määritelmän 7.2 nojalla

$$A^T A v_j = \lambda_j v_j = \sigma_j^2 v_j.$$

Siispä yhtälö (7.8) saadaan muotoon

$$\begin{aligned} u_i^T u_j &= \frac{1}{\sigma_i \sigma_j} v_i^T (\sigma_j^2 v_j) \\ &= \frac{\sigma_j}{\sigma_i} v_i^T v_j = \begin{cases} 1, & \text{jos } i = j \\ 0, & \text{jos } i \neq j, \end{cases}, \end{aligned}$$

koska joukko $\{v_1, v_2, \dots, v_n\}$ on ortonormaali.

Vektorin u_j määritelmän perusteella jokainen u_j, missä $1 \leq j \leq r$, on matriisin A sarakeavaruudessa. Sarakeavaruuden dimensio on r, joten vektorit $u_1, u_2, \dots, u_r$ muodostavat sarakeavaruuden $C(A)$ ortonormaalin kannan. Lauseen 5.2 perusteella nyt sarakeavaruus $C(A)^\perp = N(A^T)$, jonka dimensio on $m - r$ Lauseen 5.3 perusteella. Olkoon joukko $\{u_{r+1}, u_{r+2}, \dots, u_m\}$ ytimen $N(A^T)$ ortonormaali kanta, ja määritellään, että

$$U_2 = \begin{bmatrix} u_{r+1} & u_{r+2} & \cdots & u_m \end{bmatrix},$$

ja määritellään matriisi U lohkomatriisiksi

$$U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}.$$

Lauseesta 5.3 seuraa, että vektorit $u_1, u_2, \dots, u_m$ muodostavat avaruuden $\mathbb{R}^m$ ortonormaalin kannan. Näin ollen ortogonaalisen matriisin Määritelmän 1.4 nojalla matriisi U on ortogonaalinen. Nyt matriisille $U\Sigma V^T$ pätee matriisien U, Σ ja V lohkomuotojen sekä kaavojen (7.7) ja (7.5) nojalla

$$\begin{aligned} U\Sigma V^T &= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \\ &= \begin{bmatrix} U_1\Sigma_1 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \\ &= U_1\Sigma_1 V_1^T \\ &= AV_1V_1^T = A, \end{aligned}$$

kuten haluttiin. Siispä matriisilla A on singulaariarvohajotelma, joka on muotoa $U\Sigma V^T$.

□

Käydään seuraavaksi läpi esimerkki singulaariarvohajotelman löytämisestä matriisille.

Esimerkki 7.4. Ratkaistaan matriisin A singulaariarvohajotelma, missä

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix},$$

kuten Esimerkissä 6.5. Esimerkissä on ratkaistu matriisin A^TA ominaisarvoiksi $\lambda_1 = 10$ ja $\lambda_2 = 0$, ja näitä vastaaviksi eräiksi ominaisvektoreiksi $x_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ja $x_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. Näin ollen matriisin A singulaariarvot ovat Määritelmän 7.2 mukaisesti

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{10} \qquad \text{ja} \qquad \sigma_2 = \sqrt{\lambda_2} = 0.$$

Ortogonaaliseksi matriisiksi V saatiin Esimerkissä 6.5 seuraava matriisi:

$$V = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix},$$

joten tämän ja kaavan (7.6) avulla löydetään matriisin U ensimmäinen sarakevektori

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}.$$

Matriisin U loput sarakkeet saadaan etsimällä kanta ytimelle $N(A^T)$ ja muuttamalla se Gram-Shmidtin menetelmällä ortonormaaliksi kannaksi ytimelle $N(A^T)$, kuten singulaariarvohajotelman todistuksessa Lauseessa 7.3 on tehty. Tällä tavalla saadaan matriisi

$$U_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix},$$

missä on matriisin U loput kaksi saraketta u_2 ja u_3. Näin ollen matriisi U on

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

ja matriisi Σ on sen määritelmän mukaisesti

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Siispä tuloksi $U\Sigma V^T$ saadaan

$$U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} = A.$$

Näin ollen löydettiin matriisin A singulaariarvohajotelma.

7.1. Pseudokäänteismatriisi

Tarkastellaan tässä alaluvussa singulaariarvohajotelman sovelluksena pseudokäänteismatriisia, esitetään sen määritelmä ja käydään sen toimintaa läpi esimerkin avulla. Pseudokäänteismatriisia hyödynnetään Luvussa 8, ja sen käsittelyssä on käytetty lähteenä Leonin kirjaa [**3**].

Johdatellaan aluksi tarvetta pseudokäänteismatriisille. Jos A on $n \times n$ -neliömatriisi, jonka singulaariarvohajotelma on $A = U\Sigma V^T$, ja jos se on kääntyvä, niin matriisi Σ on myös kääntyvä. Tällöin matriisille A saadaan singulaariarvohajotelman avulla käänteismatriisi

$$A^{-1} = V\Sigma^{-1}U^T.$$

Jos neliömatriisi A ei ole kääntyvä tai A on jokin $m \times n$ -matriisi, jollakin $m \neq n$, jolloin erityisesti se ei ole kääntyvä, niin sille voidaan muodostaa käänteismatriisin sijaan pseudokäänteismatriisi A^+. Määritellään seuraavaksi tämä pseudokäänteismatriisi.

Määritelmä 7.5. Olkoon $m \times n$ -matriisi A, jonka singulaariarvohajotelma on $A = U\Sigma V^T$, ja aste on r. Matriisin A *pseudokäänteismatriisi* on

$$A^+ = V\Sigma^+U^T,$$

missä

$$\Sigma^+ = \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_1^{-1} & 0_{r\times(m-r)} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(m-r)} \end{bmatrix}$$
$$= \begin{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_r} \end{bmatrix} & 0 \\ 0 & 0 \end{bmatrix}.$$

Seuraavassa esimerkissä tarkastellaan 3×2 -matriisin pseudokäänteismatriisia.

Esimerkki 7.6. Olkoon matriisi A kuten Esimerkeissä 7.4 ja 6.5, eli

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

Esimerkissä 7.4 on ratkaistu matriisin A singulaariarvot $\sigma_1 = \sqrt{10}$ ja $\sigma_2 = 0$. Koska singulaariarvo σ_1 on ainoa nollasta poikkeava singulaariarvo, niin pseudokäänteismatriisin Määritelmän 7.5 mukaisesti matriisi Σ^+ on

$$\Sigma^+ = \begin{bmatrix} \frac{1}{\sqrt{10}} & 0_{1\times 2} \\ 0_{1\times 1} & 0_{1\times 2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{10}} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Näin ollen Määritelmän 7.5 nojalla matriisin A pseudokäänteismatriisi on $A^+ = V\Sigma^+U^T$, missä matriisit V ja U ovat kuten singulaariarvohajotelmassa, eli ne on laskettu Esimerkissä 7.4. Siispä matriisin A pseudokäänteismatriisiksi saadaan

$$A^+ = V\Sigma^+U^T = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{10}} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{10} & \frac{1}{10} & 0 \\ \frac{1}{5} & \frac{1}{5} & 0 \end{bmatrix}.$$

LUKU 8

Asteen yksi perturbaatio

Perehdytään matriisiin asteen yksi perturbaatioon, ja siihen liittyvään perturbaatiolauseeseen. Perturbaatiolauseen 8.2 todistuksessa sovelletaan matriisin singulaariarvohajotelmaa, ja lisäksi tarvitaan pseudokäänteismatriisia sekä aiemmin saatuja tuloksia matriisin asteelle. Luvussa on käytetty lähteenä Meyerin artikkelia [**4**], jossa esiintyvää perturbaation todistusta on perusteltu merkittävästi täsmällisemmin tässä tutkielmassa.

Perehdytään tarkemmin asteen yksi perturbaatioon. Tarkastellaan ensin tätä varten tarvittavaa matriisia, jonka aste on yksi, ja joka koostuu kahden vektorin ulkotulosta. Olkoon $m \times n$ -matriisi B, jonka aste on 1, jolloin B on muotoa

$$B = cd^T,$$

missä $c_{m\times 1} \neq 0$ ja $d_{n\times 1} \neq 0$. Nähdään, että matriisin B aste on tosiaan 1, koska matriisille

$$B = cd^T = c\begin{bmatrix} d_1 & d_2 & \cdots & d_n \end{bmatrix} = \begin{bmatrix} d_1c & d_2c & \cdots & d_nc \end{bmatrix}$$

pätee, että sen sarakeavaruuden dimensio on 1, joten asteen Määritelmän 3.5 nojalla matriisin B aste on 1. Siispä $m \times n$ -matriisin A asteen yksi perturbaatiossa lisätään matriisiin A astetta yksi oleva matriisi B, eli tarkastellaan matriisia $A + cd^T$, ja sen astetta. Lauseessa 8.2 osoitetaankin, että $\operatorname{rank}(A + cd^T)$ voi saada arvoja $r-1$, r tai $r+1$ tietyillä ehdoilla.

Seuraavassa Lemmassa 8.1 esitellään erikoistapaus perturbaatiosta, ja sen jälkeen Lauseen 8.2 todistuksessa palautetaan tilanne singulaariarvohajotelman avulla seuraavaan lemmaan. Kyseisessä Lemmassa 8.1 on määritelty matriisi Z, joka on sellaisessa muodossa, että se helpottaa Lauseen 8.2 todistusta, koska matriisilla on vain kolme mahdollista eri astevaihtoehtoa. Selkeyden vuoksi lemmassa on käytetty samoja merkintöjä, kuin mainitussa lauseessa.

Lemma 8.1. *Olkoot $m \times n$ -matriisi A ja vektorit $c = c_{m\times 1} \neq 0, d = d_{n\times 1} \neq 0$. Olkoot myös vektorit $x_2 \in \mathbb{R}^{m-r}$ ja $y_2 \in \mathbb{R}^{n-r}$, ja $(m-r+1) \times (n-r+1)$ -matriisi Z lohkomuodossa siten, että*

$$\begin{aligned} Z &= \begin{bmatrix} 0_{(m-r)\times(n-r)} & x_2 \\ y_2^T & -\alpha \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & \cdots & 0 & x_{2_1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & x_{2_{m-r}} \\ y_{2_1}^T & y_{2_2}^T & \cdots & y_{2_{n-r}}^T & -\alpha \end{bmatrix}, \end{aligned}$$

missä $\alpha = 1 + d^T A^+ c \in \mathbb{R}$, *ja* A^+ *on matriisin* A *pseudokäänteismatriisi. Tällöin matriisin* Z *aste* rank Z *on*

$$\begin{cases} 0, & \textit{jos ja vain jos } x_2 = 0 \textit{ ja } y_2 = 0 \textit{ ja } \alpha = 0, \\ 1, & \textit{jos ja vain jos (joko } x_2 = 0 \textit{ tai } y_2 = 0\textit{) tai (}x_2 = 0 \textit{ ja } y_2 = 0 \textit{ ja } \alpha \neq 0\textit{)}, \\ 2, & \textit{jos ja vain jos } x_2 \neq 0 \textit{ ja } y_2 \neq 0. \end{cases}$$

Todistus. Tarkastellaan ensin tapauksittain implikaation suuntia oikealta vasemmalle.

1. Jos $x_2 = 0$, $y_2 = 0$ ja $\alpha = 0$, niin

$$Z = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

jolloin rank $Z = 0$.

2. Jos $x_2 = 0$, $y_2 = 0$ ja $\alpha \neq 0$, niin

$$Z = \begin{bmatrix} 0 & 0 \\ 0 & -\alpha \end{bmatrix},$$

jolloin rank $Z = 1$.

3. Jos $x_2 = 0$ ja $y_2 \neq 0$ ja α on mikä tahansa reaaliluku, niin

$$Z = \begin{bmatrix} 0 & 0 \\ y_2^T & -\alpha \end{bmatrix}.$$

Tällöin Lauseen 4.7 nojalla rank Z = rank Z^T. Matriisissa Z^T on vain yksi nollasta poikkeava sarake, joten sen asteen täytyy olla rank Z^T = rank $Z = 1$.

4. Jos $x_2 \neq 0$, $y_2 = 0$ ja α on mikä tahansa reaaliluku, niin

$$Z = \begin{bmatrix} 0 & x_2 \\ 0 & -\alpha \end{bmatrix}.$$

Matriisissa Z on vain yksi nollasta poikkeava sarake, jolloin sen aste on rank $Z = 1$.

5. Jos $x_2 \neq 0$, $y_2 \neq 0$ ja α on mikä tahansa reaaliluku, niin

$$Z = \begin{bmatrix} 0 & x_2 \\ y_2^T & -\alpha \end{bmatrix}.$$

Koska Lauseen 4.9 nojalla redusoidun porrasmatriisin johtavien alkioiden määrä vastaa matriisin astetta, niin tehdään matriisille Z alkeisrivioperaatiot siten, että saadaan matriisi redusoituun porrasmatriisimuotoon Lauseen 4.1 mukaisesti. Vaihdetaan vektorin y_2^T sisältämä viimeinen rivi ensimmäiseksi. Sen jälkeen vaihdetaan toiselle riville rivi, missä on jokin vektorin x_2 nollasta poikkeava alkio. Kerrotaan rivi kaksi kyseisen alkion käänteisluvulla, jotta johtavaksi alkioksi saadaan luku 1, ja vähennetään sopivasti kerrottuna toinen rivi alemmista riveistä siten, että niihin jää vain nolla-alkioita. Tällöin saadaan matriisi

$$\begin{bmatrix} y_{2_1}^T & y_{2_2}^T & \cdots & y_{2_{n-r}}^T & -\alpha \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Myös jokin vektorin y_2^T alkioista on nollasta poikkeava, joten ensimmäisen rivin vasemmanpuoleisin nollasta poikkeava alkio on sen johtava alkio. Tälle johtavalle alkiolle saadaan arvoksi yksi kertomalla rivi kyseisen alkion käänteisluvulla. Tällöin johtavat alkiot ovat riveillä 1 ja 2, joten johtavien alkioiden lukumäärä on 2. Tämä muoto on matriisin Z redusoitu porrasmatriisi, joten Lauseen 4.9 nojalla matriisin Z asteelle pätee, että $\operatorname{rank} Z = 2$.

Tarkastellaan lopuksi implikaation käänteisiä suuntia, tässä tapauksessa vasemmalta oikealle. Olkoon $\operatorname{rank} Z = 0$. Tällöin matriisi Z on nollamatriisi, jolloin $x_2 = 0$, $y_2 = 0$ ja $\alpha = 0$. Olkoon sen sijaan $\operatorname{rank} Z = 1$. Tällöin kohtien 1 ja 5 tapaukset vektoreille x_2, y_2 ja luvulle α ovat mahdottomia, koska aste ei tällöin olisi yksi. Jos pätee $x_2 = 0$, niin joko ($y_2 = 0$ ja $\alpha \neq 0$) tai $y_2 \neq 0$. Muussa tapauksessa aste ei voi olla yksi, koska esimerkiksi tapauksessa $x_2 = 0$, $y_2 = 0$ ja $\alpha = 0$ päädytään tapaukseen 1, jolloin aste olisi nolla. Jos sen sijaan pätee $y_2 = 0$, niin joko ($x_2 = 0$ ja $\alpha \neq 0$) tai $x_2 \neq 0$. Muissa tapauksissa aste ei voi olla yksi, kuten edellä. Toisaalta jos pätee $x_2 \neq 0$, niin täytyy olla $y_2 = 0$, koska tapauksessa $x_2 \neq 0$ ja $y_2 \neq 0$ päädytään tapaukseen 5, missä matriiisin Z aste olisi 2. Samoin jos pätee $y_2 \neq 0$, niin samanlaisella päättelyllä täytyy olla, että $x_2 = 0$. Siispä päädytään ainoastaan lemman vaihtoehtoihin (joko $x_2 = 0$ tai $y_2 = 0$) tai ($x_2 = 0$ ja $y_2 = 0$ ja $\alpha \neq 0$). Olkoon sen sijaan $\operatorname{rank} Z = 2$. Tällöin täytyy olla $x_2 \neq 0$ ja $y_2 \neq 0$, koska muuten kohtien 1-4 nojalla asteeksi saataisiin $\operatorname{rank} Z \leq 1$.

Näin ollen matriisin Z asteen vaihtoehdoille saadaan halutut ehdot. □

Todistetaan seuraavaksi perturbaatiolause käyttäen apuna edellistä lemmaa.

LAUSE 8.2. *Olkoon $m \times n$ -matriisi A, jonka aste on* $\operatorname{rank} A = r$. *Olkoot myös vektorit $c = c_{m\times 1} \neq 0, d = d_{n\times 1} \neq 0$. Tällöin aste* $\operatorname{rank}(A + cd^T)$ *on*

$$\begin{cases} r-1, & \textit{joss } c \in C(A) \textit{ ja } d \in C(A^T) \textit{ ja } \alpha = 0, \\ r, & \textit{joss (joko } c \in C(A) \textit{ tai } d \in C(A^T)) \textit{ tai } (c \in C(A) \textit{ ja } d \in C(A)^T \textit{ ja } \alpha \neq 0)\,, \\ r+1, & \textit{joss } c \notin C(A) \textit{ ja } d \notin C(A^T), \end{cases}$$

missä $\alpha = 1 + d^T A^+ c \in \mathbb{R}$, ja A^+ on matriisin A pseudokäänteismatriisi.

TODISTUS. Olkoon matriisin A singulaariarvohajotelma $A = U\Sigma V^T$, kuten Lauseessa 7.3, eli missä matriisi Σ on muotoa

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix} & 0 \\ 0 & 0 \end{bmatrix}.$$

Olkoon myös vektorit $x \in \mathbb{R}^m$ ja $y \in \mathbb{R}^n$ siten, että

$$x = U^T c = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \text{ja} \qquad y = V^T d = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \tag{8.1}$$

missä $x_1, y_1 \in \mathbb{R}^r, x_2 \in \mathbb{R}^{m-r}$ ja $y_2 \in \mathbb{R}^{n-r}$. Nyt kaavan (8.1) avulla saadaan

$$c = Ux \qquad \text{ja} \qquad d = Vy, \tag{8.2}$$

koska matriisit U ja V ovat ortogonaaliset. Siispä matriisi V^T on myös ortogonaalinen ja näin ollen kääntyvä matriisi. Tällöin matriisin $A + cd^T$ asteelle pätee Lemman 4.8 nojalla

$$\begin{aligned} \operatorname{rank}(A + cd^T) &= \operatorname{rank}(U\Sigma V^T + Ux(Vy)^T) \\ &= \operatorname{rank}(U\Sigma V^T + Uxy^T V^T) \\ &= \operatorname{rank}(U(\Sigma + xy^T)V^T) \\ &= \operatorname{rank}(\Sigma + xy^T). \end{aligned} \tag{8.3}$$

Lisätään seuraavaksi matriisiin $\Sigma + xy^T$ yksi muiden sarakkeiden kanssa lineaarisesti riippumaton sarake, jolloin matriisin aste nousee yhdellä:

$$\begin{aligned} \operatorname{rank}(\Sigma + xy^T) + 1 &= \operatorname{rank} \begin{bmatrix} \Sigma + xy^T & x \\ 0 & -1 \end{bmatrix} \\ &= \operatorname{rank} \begin{bmatrix} \Sigma & x \\ y^T & -1 \end{bmatrix} \begin{bmatrix} I_{n\times n} & 0_{n\times 1} \\ (y^T)_{1\times n} & 1 \end{bmatrix}. \end{aligned} \tag{8.4}$$

Tarkastellaan $(n + 1) \times (n + 1)$ -neliömatriisia

$$\begin{bmatrix} I_{n\times n} & 0 \\ y^T & 1 \end{bmatrix},$$

joka on alakolmiomatriisi. Tällöin sen determinantti on diagonaalialkioiden tulo, mikä on tässä tapauksessa 1, koska kaikki diagonaalialkiot ovat ykkösiä. Näin ollen determinantti on nollasta poikkeava, joten tarkasteltava matriisi on kääntyvä. Siispä kaava (8.4) saadaan Lemman 4.8 nojalla muotoon

$$\operatorname{rank}(\Sigma + xy^T) + 1 = \operatorname{rank} \begin{bmatrix} \Sigma & x \\ y^T & -1 \end{bmatrix}. \tag{8.5}$$

Koska edellinen yhtäsuuruus pätee, niin kaavoista (8.3) ja (8.5) seuraa, että

$$\begin{aligned} \operatorname{rank}(A + cd^T) &= \operatorname{rank}(\Sigma + xy^T) \\ &= \operatorname{rank} \begin{bmatrix} \Sigma_{m\times n} & x_{m\times 1} \\ (y^T)_{1\times n} & -1 \end{bmatrix} - 1 \\ &= \operatorname{rank} \begin{bmatrix} \Sigma_1 & 0 & x_1 \\ 0 & 0 & x_2 \\ y_1^T & y_2^T & -1 \end{bmatrix} - 1. \end{aligned} \tag{8.6}$$

Lasketaan sitten matriisitulo $d^T A^+ c$, jolle saadaan kaavan (8.2), pseudokäänteismatriisin Määritelmän 7.5 sekä matriisien U ja V ortogonaalisuuden avulla

$$\begin{aligned} d^T A^+ c = d^T V\Sigma^+ U^T c &= (Vy)^T V\Sigma^+ U^T (Ux) \\ &= y^T V^T V \Sigma^+ U^T U x \\ &= y^T \Sigma^+ x \\ &= \begin{bmatrix} y_1^T & y_2^T \end{bmatrix} \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} \\ &= y_1^T \Sigma_1^{-1} x_1. \end{aligned} \tag{8.7}$$

Muodostetaan kaavan (8.4) matriisille matriisitulo, jossa kerrotaan kyseistä matriisia vasemmalta sopivalla alakolmiomatriisilla ja oikealta sopivalla yläkolmiomatriisilla, jotka ovat kääntyviä matriiseja. Tällöin tulon aste pysyy samana alkuperäisen matriisin kanssa Lauseen 3.7 ja Lemman 4.8 mukaisesti. Tarkastellaan siis tuloa

$$\begin{aligned}
&\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -y_1^T\Sigma_1^{-1} & 0 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 & x_1 \\ 0 & 0 & x_2 \\ y_1^T & y_2^T & -1 \end{bmatrix} \begin{bmatrix} I & 0 & -\Sigma_1^{-1}x_1 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \Sigma_1 & 0 & x_1 \\ 0 & 0 & x_2 \\ -y_1^T\Sigma_1^{-1}\Sigma_1 + y_1^T & y_2^T & -y_1^T\Sigma_1^{-1}x_1 - 1 \end{bmatrix} \begin{bmatrix} I & 0 & -\Sigma_1^{-1}x_1 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \Sigma_1 & 0 & x_1 \\ 0 & 0 & x_2 \\ 0 & y_2^T & -y_1^T\Sigma_1^{-1}x_1 - 1 \end{bmatrix} \begin{bmatrix} I & 0 & -\Sigma_1^{-1}x_1 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \Sigma_1 & 0 & -\Sigma_1\Sigma_1^{-1}x_1 + x_1 \\ 0 & 0 & x_2 \\ 0 & y_2^T & -y_1^T\Sigma_1^{-1}x_1 - 1 \end{bmatrix} \\
&= \begin{bmatrix} \Sigma_1 & 0 & 0 \\ 0 & 0 & x_2 \\ 0 & y_2^T & -\alpha \end{bmatrix} \\
&= \begin{bmatrix} \Sigma & 0 \\ 0 & Z \end{bmatrix},
\end{aligned} \tag{8.8}$$

missä Kaavan (8.7) ja luvun α määritelmän nojalla pätee $-y_1^T\Sigma_1^{-1}x_1 - 1 = -\alpha$, ja matriisi Z on kuten Lemmassa 8.1. Siispä edellisestä nähdään, että kaava (8.6) saadaan Lemman 4.8 avulla muotoon

$$\operatorname{rank}(A + cd^T) = \operatorname{rank} \begin{bmatrix} \Sigma & 0 \\ 0 & Z \end{bmatrix} - 1. \tag{8.9}$$

Osoitetaan seuraavaksi, että

$$\operatorname{rank} \begin{bmatrix} \Sigma & 0 \\ 0 & Z \end{bmatrix} = r + \operatorname{rank} Z. \tag{8.10}$$

Lasketaan kyseisen matriisin aste sen redusoidun porrasmatriisimuodon avulla. Kun sovelletaan alkeisrivioperaatioita lohkomatriisiin (8.10), saadaan se johdettua seuraavaan muotoon

$$\begin{bmatrix} \operatorname{rref}(\begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}) & 0 \\ 0 & \operatorname{rref}(Z) \end{bmatrix}. \tag{8.11}$$

Singulaariarvohajotelman Lauseessa 7.3 ollaan osoitettu, että

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Näin ollen jatketaan alkeisrivioperaatioita kertomalla matriisin (8.11) ensimmäinen rivi alkion σ_1 käänteisluvulla, jotta saadaan johtavaksi alkioksi luku 1. Tehdään samoin muille riveille riviin r asti, jolloin saadaan r kappaletta johtavia alkioita. Matriisin Σ lohkomuodosta nähdään, että sen mahdolliset alimmat rivit ovat nollarivejä. Siispä siirretään seuraavaksi nämä nollarivit, eli rivistä $r+1$ riviin m asti, koko matriisin (8.11) alimmiksi riveiksi. Lauseen 4.9 nojalla matriisin Z redusoidun porrasmatriisin johtavien alkioiden lukumäärä vastaa matriisin Z astetta. Koska matriisin $\mathrm{rref}(Z)$ yläpuolella on nollalohko, niin jokainen johtava alkio on ainoa nollasta poikkeava alkio sarakkeessaan, kuten redusoidussa porrasmatrisiissa halutaan. Tällöin siis on saatu matriisin (8.10) redusoitu parrasmatriisimuoto, josta kyseisen matriisin aste saadaan laskemalla yhteen matriisien Σ ja Z johtavien alkioiden lukumäärä, eli kaava (8.10) on osoitettu. Näin ollen edellisen perustelun nojalla kaava (8.9) saadaan muotoon

$$\operatorname{rank}(A + cd^T) = r + \operatorname{rank} Z - 1. \tag{8.12}$$

Jos matriisit U ja V on ositettu kuten singulaariarvohajotelmassa, eli

$$U = \begin{bmatrix} (U_1)_{m\times r} & U_2 \end{bmatrix} \qquad \text{ja} \qquad V = \begin{bmatrix} (V_1)_{n\times r} & V_2 \end{bmatrix},$$

missä

$$U_1 = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix} \qquad \text{ja} \qquad V_1 = \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix},$$

niin matriisien A, U_1, A^T ja V_1 sarakeavaruuksille pätee yhtäsuuruudet

$$C(A) = C(U_1) \qquad \text{ja} \qquad C(A^T) = C(V_1). \tag{8.13}$$

Osoitetaan nämä matriisien lohkomuotojen avulla. Koska matriisilla A on singulaariarvohajotelma $A = U\Sigma V^T$, niin saadaan

$$\begin{aligned} A = U\Sigma V^T &= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} V^T \\ &= \begin{bmatrix} U_1\Sigma_1 & 0 \end{bmatrix} V^T, \end{aligned} \tag{8.14}$$

missä

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix}.$$

Tässä Lemman 4.8 nojalla kääntyvä matriisi V^T ei vaikuta matriisin asteeseen, joten sarakeavaruuden Määritelmän 3.1 avulla pätee

$$\begin{aligned} C(A) &= C(\begin{bmatrix} U_1\Sigma_1 & 0 \end{bmatrix}) \\ &= C(\begin{bmatrix} \sigma_1 u_1 & \sigma_2 u_2 & \cdots & \sigma_r u_r \end{bmatrix}) \\ &= C(\begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}) \\ &= C(U_1), \end{aligned} \tag{8.15}$$

koska nollasta poikkeavat reaalilukukertoimet, $\sigma_i, i = 1, 2, \ldots, r$, eivät muuta sarakeavaruutta. Sen sijaan matriisin A transpoosille pätee singulaariarvohajotelman nojalla

$$A^T = V\Sigma^T U^T.$$

Tapauksessa $C(A^T) = C(V_1)$ pätee näin ollen sama perustelu matriisin A^T avulla, kuin yllä matriisille A, koska matriisit V ja U^T ovat kääntyviä.

Tarkastellaan seuraavaksi matriisia Z, joka on Lemman 8.1 muotoa

$$Z = \begin{bmatrix} 0 & x_2 \\ y_2^T & -\alpha \end{bmatrix}.$$

Osoitetaan kaavan (8.13) avulla, että tässä vektoreille x_2 ja y_2 pätee

$$x_2 = 0 \qquad \text{jos ja vain jos} \qquad c = Ux \in C(A) \tag{8.16}$$

ja

$$y_2 = 0 \qquad \text{jos ja vain jos} \qquad d = Vy \in C(A^T). \tag{8.17}$$

Jos $x_2 = 0$, niin vektorin c määritelmän mukaisesti ja kaavan (8.13) nojalla

$$c = Ux = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = U_1 x_1 \in C(A).$$

Jos taas $c \in C(A)$, niin

$$c = Ux = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = U_1 x_1 + U_2 x_2.$$

Tästä nähdään, että pätee yhtäsuuruus $c - U_1x_1 - U_2x_2 = 0$. Koska $c \in C(A)$, niin kaavan (8.13) nojalla myös $c \in C(U_1)$, jolloin c on matriisin U_1 sarakkeiden eli sarakavektoreiden $u_1, u_2, \ldots, u_r$ lineaarikombinaatio. Olkoot $\lambda_1, \lambda_2, \ldots, \lambda_r \in \mathbb{R}$ kertoimet matriisin c lineaarikombinaatiossa, jolloin merkitään vektoria c seuraavasti:

$$c = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_r u_r,$$

ja lisäksi voidaan merkitä

$$U_1 x_1 = x_{1_1} u_1 + x_{1_2} u_2 + \cdots + x_{1_r} u_r$$

ja

$$U_2 x_2 = x_{2_1} u_{r+1} + x_{2_2} u_{r+2} + \cdots + x_{2_{m-r}} u_m.$$

Tällöin lauseke $c - U_1x_1 - U_2x_2$ saadaan muotoon

$$\begin{aligned} c - U_1x_1 - U_2x_2 =& \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_r u_r - (x_{1_1} u_1 + x_{1_2} u_2 + \cdots + x_{1_r} u_r) \\ & - (x_{2_1} u_{r+1} + x_{2_2} u_{r+2} + \cdots + x_{2_{m-r}} u_m) \\ =& (\lambda_1 - x_{1_1}) u_1 + (\lambda_2 - x_{1_2}) u_2 + \cdots + (\lambda_r - x_{1_r}) u_r \\ & + (-x_{2_1}) u_{r+1} + (-x_{2_2}) u_{r+2} + \cdots + (-x_{2_{m-r}}) u_m = 0. \end{aligned} \tag{8.18}$$

Koska matriisin U sarakkeet ovat lineaarisesti riippumattomat, niin Määritelmän 1.8 nojalla kaavassa (8.18) kaikkien matriisin U sarakkeiden $u_1, u_1, \ldots, u_m$ kertoimien tulee olla nollia. Erityisesti matriisin U_2 sarakkeiden $u_{r+1}, u_{r+2}, \ldots, u_m$ kertoimien, mitkä vastaavat vektorin x_2 alkioita, tulee olla nollia. Siispä täytyy olla, että vektorille x_2 pätee $x_2 = 0$. Samoin vektorin y_2 tapauksessa samanlaisella päättelyllä saadaan, että $y_2 = 0$ jos ja vain jos $d \in C(A^T)$.

Kaavojen (8.16) ja (8.17) avulla huomataan, että lauseessa esiintyvät ehdot vektoreille c ja d vastaavat Lemman 8.1 ehtoja vektoreille x_2 ja y_2. Näin ollen esimerkiksi väite

$$\begin{aligned} \operatorname{rank}(A + cd^T) = r \qquad & \text{jos ja vain jos (joko } c \in C(A) \text{ tai } d \in C(A^T)) \\ & \text{tai } (c \in C(A) \text{ ja } d \in C(A)^T \text{ ja } \alpha \neq 0) \end{aligned} \tag{8.19}$$

saadaan muotoon

$$\begin{aligned} \operatorname{rank}(A + cd^T) = r \qquad & \text{jos ja vain jos (joko } x_2 = 0 \text{ tai } y_2 = 0) \\ & \text{tai } (x_2 = 0 \text{ ja } y_2 = 0 \text{ ja } \alpha \neq 0). \end{aligned} \tag{8.20}$$

Siispä riittää osoittaa väitelause (8.20) todeksi. Jos matriisin $A + cd^T$ asteelle pätee

$$\operatorname{rank}(A + cd^T) = r,$$

niin kaavan (8.12) nojalla

$$r = \operatorname{rank}(A + cd^T) = r + \operatorname{rank} Z - 1,$$

joten matriisin Z asteen täytyy olla yksi. Tämä vastaa Lemman 8.1 tapausta, missä (joko $x_2 = 0$ tai $y_2 = 0$) tai ($x_2 = 0$ ja $y_2 = 0$ ja $\alpha \neq 0$), joten väite (8.20) pätee. Jos sen sijaan (joko $x_2 = 0$ tai $y_2 = 0$) tai ($x_2 = 0$ ja $y_2 = 0$ ja $\alpha \neq 0$), niin Lemman 8.1 nojalla matriisin Z asteelle pätee, että $\operatorname{rank} Z = 1$. Siispä kaavan (8.12) avulla saadaan

$$\operatorname{rank}(A + cd^T) = r + \operatorname{rank} Z - 1 = r + 1 - 1 = r.$$

Näin ollen väite (8.20) pätee myös toiseen suuntaan, joten lauseen ehdot täyttyvät. Samanlaisella päättelyllä saadaan osoitettua myös muut tapaukset, kun matriisin $A + cd^T$ aste on $r - 1$ tai $r + 1$. Näin ollen on saatu osoitettua lauseen ehdot asteen yksi perturbaatiolle.

□

Tarkastellaan vielä asteen yksi perturbaatiota käytännössä esimerkin avulla.

Esimerkki 8.3. Olkoon matriisi A kuten Esimerkeissä 6.5, 7.4 ja 7.6, eli matriisi A on

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix},$$

jonka aste on $r = 1$. Tällöin matriisilla A on singulaariarvohajotelma $A = U\Sigma V^T$, jonka matriisit U ja V on laskettu Esimerkissä 7.4. Käydään tässä esimerkissä läpi Lauseessa 8.2 esiintyvät tapaukset vektoreille c ja d, ja tarkastellaan matriisin $A+cd^T$ astetta. Koska A on 3×2 -matriisi, niin Lauseen 8.2 nojalla vektorin $c \neq 0$ dimensio on 3×1 ja vektorin $d \neq 0$ dimensio on 2×1. Lisäksi kaavoissa (8.1) ja (8.2) esiintyvät vektorit x ja y ovat muotoa

$$x_{(3\times 1)} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_{2_1} \\ x_{2_2} \end{bmatrix} \qquad \text{ja} \qquad y_{(2\times 1)} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

1. Tapauksessa $c \in C(A)$ ja $d \in C(A^T)$ ja $\alpha = 0$ Lauseen 8.2 ja Lemman 8.1 todistusten nojalla tiedetään, että vektorit x_2 ja y_2 ovat muotoa $x_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ja $y_2 = \begin{bmatrix} 0 \end{bmatrix}$. Lisäksi luvun α määritelmän, joka esiintyy Lauseessa 8.2, nojalla $\alpha = 1 + d^T A^+ c = 0$. Jos merkitään

$$c = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \qquad \text{ja} \qquad d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix},$$

saadaan yhtälö

$$
\begin{aligned}
\alpha &= 1 + \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} \frac{1}{10} & \frac{1}{10} & 0 \\ \frac{1}{5} & \frac{1}{5} & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 1 + \begin{bmatrix} \frac{d_1}{10} + \frac{d_2}{5} & \frac{d_1}{10} + \frac{d_2}{5} & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \\
&= 1 + \frac{c_1 d_1}{10} + \frac{c_1 d_2}{5} + \frac{c_2 d_1}{10} + \frac{c_2 d_2}{5} = 1 + \frac{(c_1 + c_2)(d_1 + 2d_2)}{10} = 0.
\end{aligned} \tag{8.21}
$$

Lisäksi vektoreiden c ja d määritelmän, eli kaavan (8.2), nojalla

$$
c = Ux = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{x_1}{\sqrt{2}} \\ \frac{x_1}{\sqrt{2}} \\ 0 \end{bmatrix},
$$

ja

$$
d = Vy = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} y_1 \\ \frac{2}{\sqrt{5}} y_1 \end{bmatrix},
$$

joten nyt tämän ja kaavan (8.21) nojalla valitaan $x_1 = \sqrt{2}$ ja $y_1 = -\sqrt{5}$, jolloin

$$
c = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{ja} \qquad d = \begin{bmatrix} -1 \\ -2 \end{bmatrix}.
$$

Tällöin matriisiksi $A + cd^T$ tulee

$$
A + cd^T = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},
$$

jolloin $\operatorname{rank}(A + cd^T) = 0 = r - 1$.

2. Tapauksessa $c \in C(A)$ ja $d \notin C(A^T)$ Lauseen 8.2 ja Lemman 8.1 todistusten nojalla tiedetään, että vektorit x_2 ja y_2 ovat muotoa $x_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ja $y_2 \neq 0$. Tällöin vektoreiden c ja d määritelmän, eli kaavan (8.2), nojalla c on kuten kohdassa 1 ja

$$
d = Vy = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} y_1 - \frac{2}{\sqrt{5}} y_2 \\ \frac{2}{\sqrt{5}} y_1 + \frac{1}{\sqrt{5}} y_2 \end{bmatrix},
$$

missä matriisi V on laskettu Esimerkissä 7.4. Kun valitaan $x_1 = \sqrt{2}$ ja $y_1 = y_2 = \sqrt{5}$, niin saadaan

$$
c = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{ja} \qquad d = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.
$$

Tällöin matriisiksi $A + cd^T$ tulee

$$
A + cd^T = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 5 \\ 0 & 5 \\ 0 & 0 \end{bmatrix},
$$

jolloin $\operatorname{rank}(A + cd^T) = 1 = r$.

3. Tapauksessa $c \notin C(A)$ ja $d \notin C(A^T)$ Lauseen 8.2 ja Lemman 8.1 todistusten nojalla tiedetään, että vektoreille x_2 ja y_2 pätee $x_2 \neq 0$ ja $y_2 \neq 0$. Tällöin vektoriksi c saadaan kaavan (8.2) nojalla

$$c = Ux = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_{2_1} \\ x_{2_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_{2_1} \\ \frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_{2_1} \\ x_{2_2} \end{bmatrix},$$

ja d on kuten kohdassa 2. Valitaan $x_1 = x_{2_1} = \sqrt{2}$, $x_{2_2} = 1$ ja $y_1 = y_2 = \sqrt{5}$, jolloin saadaan

$$c = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \qquad \text{ja} \qquad d = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

Tällöin matriisiksi $A + cd^T$ tulee

$$A + cd^T = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 8 \\ -1 & 3 \end{bmatrix},$$

jolloin $\operatorname{rank}(A + cd^T) = 2 = r + 1$.

Kirjallisuutta

[1] Petri Juutinen, *Lineaarinen algebra ja geometria 2*, Jyväskylän yliopisto, 2018.
[2] Tero Kilpeläinen, *Lineaarinen algebra ja geometria 1*, Jyväskylän yliopisto, 2022.
[3] Steven J. Leon, *Linear Algebra with Application*, 9. painos Pearson Education, 2015.
[4] Carl D. Meyer, *Rank My Update, Please*, The American Mathematical Monthly, 2018, no. 1, 61–64.
[5] Mikko Saarimäki, *Reaalisia vektoriavaruuksia ja ominaisarvoja*, Jyväskylän yliopistopaino, 2012.
[6] Gilbert Strang, *Introduction to linear algebra*, 5. painos, Cambridge Press, 2016.
[7] Gilbert Strang ja Cleve Moler, *LU and CR Elimination*, Siam Review, 2022, no. 1, 181–190.