Pseudokäänteismatriisit ja pienimmän neliösumman menetelmä

Mika Backlund

Matematiikan pro gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Kevät 2024

Tiivistelmä: Mika Backlund, *Pseudokäänteismatriisit ja pienimmän neliösumman menetelmä* (engl. *Pseudoinverse and least squares method*), matematiikan pro gradu -tutkielma, 43 s., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, kevät 2024.

Tämän tutkielman tarkoituksena on käsitellä pseudokäänteismatriiseja ja pienimmän neliösumman menetelmää sekä niiden taustalla olevaa lineaarialgebran teoriaa. Pienimmän neliösumman menetelmä on tutkielman huipentuma ja sitä ennen perehdytään laajasti lineaarialgebran teorioihin. Tutkielma alkaa lineaarialgebran perusmääritelmillä luoden samalla pohjaa myöhemmin käsiteltäville aiheille.

Tutkielman ensimmäinen merkittävä teema on matriisin singulaariarvohajotelma, jonka mukaan jokainen äärellisulotteinen matriisi on esitettävissä kolmen matriisin tulona. Tullaan todistamaan väite, että singulaariarvohajotelma on olemassa. Nimi singulaariarvohajotelma on tullut siitä, että yksi hajotelman matriiseista sisältää alkuperäisen matriisin singulaariarvot.

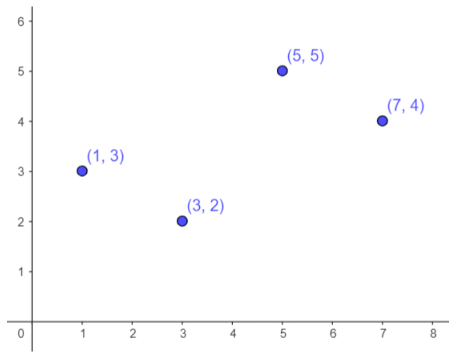
Toinen merkittävä teema on pseudokäänteismatriisi, jonka olemassa olo ja yksikäsitteisyys todistetaan. Tullaan huomaamaan, että pseudokäänteismatriisin ja singulaariarvohajotelman välillä on kytkös; singulaariarvohajotelman matriiseista saadaan muodostettua pseudokäänteismatriisi. Lisäksi käsitellään pseudokäänteismatriisin perusominaisuuksia sekä lohkotun matriisin pseudokäänteismatriisia.

Tutkielman lopussa päästään käsittelemään pienimmän neliösumman menetelmää. Määritellään pienimmän neliösumman ratkaisu ja otetaan käyttöön kaksi laskutapaa, jolla pienimmän neliösumman ratkaisu saadaan. Ensimmäinen laskutapa soveltuu niin sanotuille normaaleille yhtälöryhmille ja jälkimmäinen laskutapa soveltuu yleisesti kaikille yhtälöryhmille. Jälkimmäinen laskutapa kytkeytyy pseudokäänteismatriisiin siten, että sen laskukaavassa esiintyy kyseinen matriisi. Pienimmän neliösumman menetelmää soveltaen saadaan erilaisille pisteaineistoille määritettyä paras mahdollinen kuvaaja.

Sisällys

Johdanto

Tässä matematiikan alan pro gradu -tutkielmassa käsitellään pseudokäänteismatriiseja ja pienimmän neliösumman matemaattista teoriaa, joka on pitkälti lineaarialgebraa. Tutkielman tavoitteena on soveltaa pienimmän neliösumman menetelmän matemaattista teoriaa erilaisille pistejoukoille. Tutkielman pohjalta onkin mahdollista määrittää esimerkiksi suora, joka kuvaa parhaiten Kuvan 0.1 pistejoukkoa avaruudessa $\mathbb{R}^2$. Myös epälineaarisia sovituksia, kuten ympyräsovitus, opitaan tekemään.



Kuva 0.1. Pistejoukko avaruudessa $\mathbb{R}^2$.

Ennenkuin päästään käsittelemään pienimmän neliösumman menetelmää, tulee läpikäydä laajasti lineaarialgebraa. Kaiken lisäksi tutkielmassa halutaan laajentua kompleksilukujen tasolle. Tästä syystä ensimmäisessä luvussa esitellään kompleksilukujen perusominaisuuksia, joita tarvitaan muun muassa uuden kiinnostavan käsitteen, hermiittisen transpoosin määrittelyssä. Allekirjoittaneen tuli palauttaa mieleen joitakin peruskurssien määritelmiä ja siitä syystä jotkin määritelmät saattavat olla lukijalle aikaisemmasta tuttuja.

Toisessa luvussa määritellään äärellisulotteiselle ja kompleksiselle matriisille singulaariarvohajotelma $U\Sigma V^H$. Käsite singulaariarvohajotelma on saanut nimensä hajotelman keskimmäisen matriisin Σ diagonaalialkioista, jotka ovat hajotettavan matriisin singulaariarvoja. On huomionarvoista, että tällainen matriisituloesitys on olemassa jokaiselle matriisille.

Kolmannessa luvussa karakterisoidaan pseudokäänteismatriisi A^+ neljän määrittelyehdon mukaan. Todistetaan myös, että pseudokäänteismatriisi on yksikäsitteinen ja olemassa kaikille matriiseille joukossa $\mathbb{C}^{m\times n}$. Luvun keskeinen havainto on se, että

$A^+ = V\Sigma_k U^H$, missä matriisit U ja V ovat matriisin A singulaariarvohajotelman matriiseja. Tämä tarkoittaa, että pseudokäänteismatriisi on laskettavissa, kun alkuperäisen matriisin singulaariarvohajotelma tunnetaan. Luvussa esitellään joitakin pseudokäänteismatriisin perusominaisuuksia todistuksineen sekä määritellään pseudokäänteismatriisin laskukaava sarakkeittain lohkotulle matriisille. Lohkotuilla matriiseilla on joitakin hyödyllisiä käytännön sovelluksia, joihin ei perehdytä tässä tutkielmassa.

Neljännessä luvussa päästään lopulta käsittelemään pienimmän neliösumman menetelmää. Ongelman voi muotoilla siten, että yritetään löytää pienimpiä ratkaisuja lineaariselle yhtälöryhmälle $A\mathbf{x} = \mathbf{b}$. Olennaista ei kuitenkaan ole se, onko yhtälöryhmällä ratkaisuja olemassa, koska pienimmän neliösumman ratkaisu on aina olemassa. Luvussa annetaan kaksi tapaa määrittää pienimmän neliösumman ratkaisu. Ensimmäinen ratkaisutapa sopii matriiseille $A \in \mathbb{C}^{m\times n}$, joiden aste on n. Jälkimmäinen tapa sopii yleisesti kaikille lineaarisille yhtälöryhmille ja ratkaisussa tarvitaan apuna pseudokäänteismatriisia A^+.

Tutkielmassa sovellettiin useita lähteitä ja niistä on laadittu luettelo tutkielman loppuun. Käytetyin lähde oli Steven J. Leonin sovellettua lineaarialgebraa käsittelevä kirja [**5**], jota sovellettiin tutkielman jokaisessa luvussa. Baratan ja Husseinin teosta [**2**] sovellettiin erityisesti pseudokäänteismatriisin perusominaisuuksien todistamisessa. Lohkomatriisien pseudokäänteismatriiseihin liittyviä tuloksia käsiteltiin Jerzy K. Baksalaryn ja Oskar M. Baksalaryn artikkelin [**1**] pohjalta.

LUKU 1

Lineaarialgebran perusmääritelmiä

Tutkielman ensimmäisessä luvussa lähdetään liikkeelle lineaarialgebran perusmääritelmistä, joista osaa on jo käsitelty peruskursseilla. Näiden määritelmien avulla luodaan pohjaa tuleviin lukuihin. Kaikki määritelmät on luotu kompleksiluvuille toimiviksi ja myös aikaisemmista opinnoista tuntemattomia määritelmiä, lauseita ja lemmoja esitetään. Peruskursseilla esitettyjä tuloksia ei kuitenkaan todisteta enää uudestaan. Tärkeimmät lähteet ovat Leonin kirja [**5**], lineaarialgebran luentomonisteet [**3**] ja [**4**] sekä kompleksilaskennan luentomoniste [**9**].

1.1. Kompleksiluvut

Tutkielmassa halutaan asettautua kompleksilukujen tasolle. Sen vuoksi määritellään kompleksiluku, kompleksikonjugaatti ja kompleksiluvun moduli. Lisäksi esitellään joitain laskusääntöjä. Määrittelyissä tukeudutaan kompleksilaskennan luentomonisteeseen [**9**].

MÄÄRITELMÄ 1.1. Sanotaan, että $z = a + ib$ on kompleksiluku, jos $a, b \in \mathbb{R}$ ja i on imaginääriyksikkö.

MÄÄRITELMÄ 1.2. Kompleksiluvun $z = a+ib$ kompleksikonjugaatti on $\bar{z} = a-ib$.

ESIMERKKI 1.3. Kompleksiluvun $z = 4 + i2$ kompleksikonjugaatti on $\bar{z} = 4 - i2$.

Kompleksikonjugaatille on olemassa seuraavat laskusäännöt:

LAUSE 1.4. *Olkoon $z_1, z_2 \in \mathbb{C}$. Tällöin*

(1) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ *ja*
(2) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

MÄÄRITELMÄ 1.5. Kompleksiluvun $z = a + ib$ moduli on reaaliluku

$$|z| = \sqrt{a^2 + b^2}.$$

ESIMERKKI 1.6. Kompleksiluvun $z = 4 + i2$ moduli on

$$|z| = \sqrt{4^2 + 2^2} = \sqrt{20} = 2\sqrt{5}.$$

1.2. Lohkomatriisit

Lohkotun matriisin ajatellaan koostuvan alimatriiseista eli alkuperäinen matriisi on tavalla tai toisella jaettu pienempiin osiin. Matriisin lohkominen tehdään pysty- ja vaakaviivoja käyttäen. Esimerkiksi matriisi

$$C = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & 0 \\ 0 & -1 & -2 & -3 & -4 \\ -4 & -3 & -2 & -1 & 0 \end{bmatrix}$$

voidaan esittää alimatriisien C_{11}, C_{12}, C_{21} ja C_{22} avulla siten, että

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \left[\begin{array}{ccc|cc} 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & 0 \\ 0 & -1 & -2 & -3 & -4 \\ \hline -4 & -3 & -2 & -1 & 0 \end{array}\right].$$

Matriisin lohkominen ei yleensä ole yksikäsitteistä. Asetetaan lohkomatriisille täsmällinen määritelmä.

Määritelmä 1.7. Olkoon $A \in \mathbb{C}^{m \times n}$. Tällöin

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & & & \\ \vdots & & & \\ A_{r1} & \dots & & A_{rs} \end{bmatrix},$$

missä $A_{ij} \in \mathbb{C}^{m_i \times n_j}$, $m = m_1 + m_2 + \cdots + m_i + \cdots + m_r$ ja
$n = n_1 + n_2 + \cdots + n_j + \cdots + n_s$.

Lineaarialgebran peruskurssilla [**3**] opittiin kertomaan matriiseja keskenään. Seuraavaksi tutustutaan siihen, kuinka lohkotuilla matriiseilla kerrotaan. Tämän aliluvun matriisituloissa oletetaan, että kertolaskut ovat hyvin määriteltyjä eli ensimmäisessä matriisissa on sarakkeita yhtä paljon kuin jälkimmäisessä matriisissa on rivejä. Määritellään siis lohkomatriisien kertolasku.

Määritelmä 1.8. Olkoon

$$A = \begin{bmatrix} A_{11} & \dots & A_{1t} \\ \vdots & & \\ A_{s1} & \dots & A_{st} \end{bmatrix} \text{ ja } B = \begin{bmatrix} B_{11} & \dots & B_{1r} \\ \vdots & & \\ B_{t1} & \dots & B_{tr} \end{bmatrix}.$$

Tällöin

$$AB = \begin{bmatrix} C_{11} & \dots & C_{1r} \\ \vdots & & \\ C_{s1} & \dots & C_{sr} \end{bmatrix}, \text{ missä } C_{ij} = \sum_{k=1}^{t} A_{ik} B_{kj}.$$

Lasketaan vielä esimerkin avulla matriisitulo AB, kun matriisit on lohkottu neljään osaan.

Esimerkki 1.9. Olkoon

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{c|c} 1 & 0 \\ 0 & 1 \\ \hline 0 & 2 \\ 2 & 0 \end{array}\right]$$

ja

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ \hline 2 & 0 & 0 & 2 \end{array}\right].$$

Tällöin

$$\begin{aligned}C_{11} &= A_{11}B_{11} + A_{12}B_{21} \\ &= \begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}2 & 0\end{bmatrix} \\ &= \begin{bmatrix}0 & 1\\2 & 0\end{bmatrix}.\end{aligned}$$

Kun vastaavalla tavalla lasketaan lohkot C_{12}, C_{21} ja C_{22}, niin matriisitulo

$$AB = \left[\begin{array}{cc|cc} 0 & 1 & 1 & 0\\ 2 & 0 & 0 & 2\\ \hline 4 & 0 & 0 & 4\\ 0 & 2 & 2 & 0\end{array}\right] = \begin{bmatrix}0 & 1 & 1 & 0\\2 & 0 & 0 & 2\\4 & 0 & 0 & 4\\0 & 2 & 2 & 0\end{bmatrix}.$$

1.3. Hermiittinen transpoosi

Hermiittinen transpoosi on matriisille suoritettava operaatio, jossa tarvitaan luvun alussa opittua kompleksikonjugaatin käsitettä. Lähteinä on käytetty Leonin kirjaa [**5**] ja lineaarialgebran lähdettä [**3**]. Määritellään seuraavaksi hermiittinen transpoosi.

Määritelmä 1.10. Matriisin $A \in \mathbb{C}^{n\times m}$ hermiittinen transpoosi on $m\times n$-matriisi A^H, jolle

$$A^H = (\overline{A})^T = \overline{A^T}.$$

Hermiittinen transpoosi siis vastaa operaationa tavallista transpoosia mutta on tämän lisäksi kompleksikonjugoitu. Operaation voi suorittaa myös käänteisessä järjestyksessä. Havainnollistetaan hermiittistä transpoosia seuraavaksi.

Esimerkki 1.11. Määritetään matriisin

$$A = \begin{bmatrix}0 & 1\\ 2i & -i\\ 5-i & -4+i\end{bmatrix}$$

hermiittinen transpoosi A^H. Ensin matriisi A transponoidaan eli

$$A^T = \begin{bmatrix}0 & 2i & 5-i\\ 1 & -i & -4+i\end{bmatrix},$$

joka lopuksi kompleksikonjugoidaan, jolloin

$$A^H = \begin{bmatrix}0 & -2i & 5+i\\ 1 & i & -4-i\end{bmatrix}.$$

Samaan tulokseen oltaisiin päästy myös ensin kompleksikonjugoimalla. Lisäksi on huomionarvoista, että reaalikertoimiselle matriisille pätee $A^H = A^T$. Määritellään myös hermiittinen matriisi.

Määritelmä 1.12. Matriisi $A \in \mathbb{C}^{n\times m}$ on hermiittinen, jos $A = A^H$.

Hermiittisen matriisin on oltava neliömatriisi mutta ei kuitenkaan symmetrinen, kuten seuraava esimerkki osoittaa.

Esimerkki 1.13.

$$\begin{bmatrix} 0 & 2i \\ -2i & 1 \end{bmatrix}^H = \begin{bmatrix} 0 & 2i \\ -2i & 1 \end{bmatrix}.$$

Edetään seuraavaksi hermiittisen matriisin peruslaskutoimituksiin, jotka vastaavat transpoosin laskutoimituksia. Seuraavan lauseen todistus vastaa lineaarialgebran kurssilla ollutta todistusta ja sitä esitetä enää uudestaan. Mainittakoon myös, että tulos pätee sillä oletuksella, että matriisien laskutoimitukset ovat hyvin määriteltyjä.

Lause 1.14. *Matriisien hermiittisille transpooseille on voimassa*

(1) $(\alpha A + \beta B)^H = \bar{\alpha} A^H + \bar{\beta} B^H$,
(2) $(A^H)^H = A$ *ja*
(3) $(AB)^H = B^H A^H$ *kaikilla* $\alpha, \beta \in \mathbb{C}$.

1.4. Kompleksiset vektorit

Tässä aliluvussa määritellään kompleksinen vektori, pistetulo ja normi. Lähteenä on käytetty Leonin kirjaa [**5**].

Määritelmä 1.15. Avaruuden $\mathbb{C}^n$ vektori on muotoa

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = (z_1, z_2, \ldots, z_n)^T, \text{ missä } z_i \in \mathbb{C}.$$

Määritelmä 1.16. Avaruuden $\mathbb{C}^n$ vektoreiden $\mathbf{u}$ ja $\mathbf{v}$ pistetulo on

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^H \mathbf{u}.$$

Esimerkki 1.17. Vektoreiden

$$\mathbf{u} = \begin{bmatrix} 4 + i2 \\ i \end{bmatrix} \qquad \text{ja} \qquad \mathbf{v} = \begin{bmatrix} -2 \\ -i \end{bmatrix}$$

välinen pistetulo on

$$\mathbf{v}^H \mathbf{u} = (-2, i) \begin{bmatrix} 4 + i2 \\ i \end{bmatrix} = -2(4 + 2i) + i^2 = -9 - 4i.$$

Määritelmä 1.18. Olkoon $\mathbf{u}$ avaruuden $\mathbb{C}^n$ vektori. Vektorin $\mathbf{u}$ normi on

$$\|\mathbf{u}\| = \sqrt{|u_1|^2 + ... + |u_n|^2}.$$

Lause 1.19. *Olkoon* $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$. *Tällöin*

(1) $\mathbf{u}^H \mathbf{v} = \overline{\mathbf{v}^H \mathbf{u}}$ *ja*
(2) $\mathbf{u}^H \mathbf{u} = \|\mathbf{u}\|^2$.

TODISTUS. Todistetaan kohta (1):

$$\begin{aligned}
\overline{\mathbf{v}^H\mathbf{u}} &= \overline{(\overline{v_1},\dots,\overline{v_n})\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}} = v_1\overline{u_1} + \cdots + v_n\overline{u_n} \\
&= \overline{u_1}v_1 + \cdots + \overline{u_n}v_n = \overline{(u_1,\dots,u_n)}\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\
&= \mathbf{u}^H\mathbf{v}.
\end{aligned}$$

Todistetaan kohta (2):

$$\begin{aligned}
\mathbf{u}^H\mathbf{u} &= (\overline{u_1},\dots,\overline{u_n})\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \overline{u_1}u_1 + \cdots + \overline{u_n}u_n \\
&= \sum_{k=1}^{n} \overline{u_k}u_k = \sum_{k=1}^{n}(a_k - b_k i)(a_i + b_k i) \\
&= \sum_{k=1}^{n}(a_k^2 + b_k^2) = \sum_{k=1}^{n}|u_k|^2 \\
&= \|\mathbf{u}\|^2.
\end{aligned}$$

□

Edellisen lauseen kohta (2) antaa normille vaihtoehtoisen määritelmän

$$\|\mathbf{u}\| = \sqrt{\mathbf{u}^H\mathbf{u}}.$$

Luvun lopuksi esitetään hyödyllinen tulos, jota sovelletaan myöhemmin.

LEMMA 1.20. *Jos vektorit* $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ *ja* $\langle \mathbf{u}, \mathbf{v}\rangle = 0$, *niin* $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

TODISTUS. Kun huomioidaan oletukset ja sovelletaan luvussa opittua, saadaan

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})^H(\mathbf{u}+\mathbf{v}) = \mathbf{u}^H\mathbf{u} + \mathbf{v}^H\mathbf{u} + \mathbf{u}^H\mathbf{v} + \mathbf{v}^H\mathbf{v} \\
&= \|\mathbf{u}\|^2 + \langle \mathbf{u}, \mathbf{v}\rangle + \overline{\langle \mathbf{u}, \mathbf{v}\rangle} + \|\mathbf{v}\|^2 \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.
\end{aligned}$$

□

1.5. Neliömatriisin ominaisuudet

Neliömatriisilla on joitakin hyödyllisiä ominaisuuksia, joita käsiteltiin jo lineaarialgebran peruskurssilla [**3**]. Seuraava tulos kertoo, että jos jokin ehdoista pätee neliömatriisille, niin myös muut ehdot pätevät.

LAUSE 1.21. *Olkoon* $A \in \mathbb{C}^{n\times n}$. *Tällöin seuraavat kohdat ovat ekvivalentteja:*

(1) *A on kääntyvä,*
(2) $\det A \neq 0$,
(3) *matriisin A sarakkeet (tai rivit) ovat lineaarisesti riippumattomia,*

(4) *yhtälöllä* $A\mathbf{x} = \mathbf{b}$ *on yksikäsitteinen ratkaisu jokaiselle* $\mathbf{b} \in \mathbb{C}^n$ *ja*
(5) *yhtälöllä* $A\mathbf{x} = \mathbf{0}$ *on vain triviaalinen ratkaisu* $\mathbf{x} = \mathbf{0}$.

1.6. Matriisin aste

Tässä luvussa määritellään matriisin aste ja lähteenä käytetään Leonin kirjaa [**5**]. Määritelmää seuraavassa esimerkissä määritetään reaalikertoimisen matriisin aste Gauss-Jordan eliminaatiomenetelmällä, joka toimii myös kompleksisella matriisilla.

MÄÄRITELMÄ 1.22. Matriisin $A \in \mathbb{C}^{n\times m}$ aste on matriisin sarakkeiden (tai rivien) virittämän vektoriavaruuden dimensio ja sitä merkitään symbolilla $\operatorname{rank}(A)$.

ESIMERKKI 1.23. Olkoon

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & -5 & 1 \\ 2 & -4 & 6 \end{bmatrix}.$$

Muokkaamalla matriisia A Gauss-Jordanin eliminaatiomenetelmällä, saadaan matriisi

$$U = \begin{bmatrix} 1 & 0 & 13 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}.$$

Koska matriisit A ja U virittävät saman $\mathbb{R}^3$ aliavaruuden ja matriisilla U on selvästi kaksi lineaarisesti riippumatonta riviä, niin $\operatorname{rank}(A) = 2$.

1.7. Matriisien lineaarikuvaukset

Seuraavaksi käsitellään matriisien lineaarikuvauksen ydintä ja kuvajoukkoa kompleksiselle matriisille. Lähteenä on käytetty Leonin kirjaa [**5**]. Lineaarikuvauksen ytimellä tarkoitetaan niiden vektorien joukkoa, jotka lineaarikuvaus kuvaa nollavektoriksi. Kuvajoukko on matriisin lineaarisesti riippumattomien sarakkeiden virittämä aliavaruus. Määritellään nämä käsitteet tarkasti.

MÄÄRITELMÄ 1.24. Olkoon $A \in \mathbb{C}^{m\times n}$ matriisi. Matriisiyhtälön $A\mathbf{x} = \mathbf{0}_m$ ratkaisujoukko on joukko

$$N(A) = \{\mathbf{x} \in \mathbb{C}^n : A\mathbf{x} = \mathbf{0}_m\}.$$

MÄÄRITELMÄ 1.25. Matriisin $A \in \mathbb{C}^{m\times n}$ virittämä $\mathbb{C}^m$ aliavaruus on joukko

$$R(A) = \{\mathbf{b} \in \mathbb{C}^m : \mathbf{b} = A\mathbf{x},\ \text{jollakin } \mathbf{x} \in \mathbb{C}^n\}.$$

Symbolilla $R(A)$ viitataan jatkossa sarakkeiden virittämään aliavaruuteen. Rivien virittämää aliavaruutta merkitään symbolilla $R(A^T)$. Seuraavaksi käsitellään lineaarialgebran peruskurssilla esitettyä dimensiolausetta. Esitetään lähes vastaava tulos, joka poikkeaa peruskurssilta esitetystä siten, että matriisin kuvajoukon dimension tilalla on matriisin aste. Se, että $\operatorname{rank} A = \dim R(A)$ seuraa suoraan matriisin asteen ja kuva-avaruuden määritelmistä.

LAUSE 1.26. *Olkoon* $A \in \mathbb{C}^{m\times n}$. *Tällöin*

$$\dim N(A) + \operatorname{rank}(A) = n.$$

Luvun lopuksi käsitellään matriisiin $A^H A$ liittyviä ominaisuuksia, joita tarvitaan myöhemmin. Seuraava lemma kertoo, että matriiseilla A ja $A^H A$ on sama ydin. Tätä lemmaa sovelletaan heti perään sitä seuraavassa lauseessa, jonka mukaan kyseisillä matriiseilla on sama aste.

LEMMA 1.27. *Kaikille matriiselle $A \in \mathbb{C}^{m\times n}$ on voimassa*

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

TODISTUS. Olkoon $\mathbf{x} \in N(A)$. Jos $A\mathbf{x} = \mathbf{0}$, niin myös $A^H A\mathbf{x} = \mathbf{0}$. Tällöin $N(A) \subseteq N(A^H A)$. Olkoon seuraavaksi $\mathbf{y} \in N(A^H A)$. Tällöin $A^H A\mathbf{y} = \mathbf{0}$, josta seuraa

$$\mathbf{y}^H A^H A\mathbf{y} = \mathbf{0}.$$

Hermiittisen transpoosin laskutoimitusten nojalla

$$(A\mathbf{y})^H A\mathbf{y} = \mathbf{0}.$$

Lauseen 1.19 nojalla

$$\|A\mathbf{y}\|^2 = 0,$$

jolloin

$$A\mathbf{y} = \mathbf{0}.$$

Nyt siis $N(A^H A) \subseteq N(A)$ ja $N(A^H A) = N(A)$. □

LAUSE 1.28. *Kaikille matriiselle $A \in \mathbb{C}^{m\times n}$ on voimassa*

$$\operatorname{rank}(A) = \operatorname{rank}(A^H A).$$

TODISTUS. Lemman 1.27 nojalla $N(A) = N(A^H A)$. Koska matriiseissa A ja $A^H A$ on n kappaletta sarakkeita, niin Lauseesta 1.26 seuraa, että

$$\operatorname{rank}(A) = n - \dim N(A) = n - \dim N(A^H A) = \operatorname{rank}(A^H A).$$

□

1.8. Unitaarinen matriisi

Lineaarialgebran peruskurssilla käsiteltiin ortogonaalisia vektoreita ja matriiseja. Jatketaan määrittelemällä unitaarinen matriisi, jonka reaalikertoimista erikoistapausta sanotaan ortogonaaliseksi matriisiksi. Sen jälkeen esitetään ekvivalenssilause, jonka avulla voidaan helposti tarkistaa, onko matriisi unitaarinen. Lauseen todistus on samanlainen kuin peruskurssilla ja sitä ei todisteta enää uudestaan. Luvussa sovelletaan Leonin kirjaa [**5**].

MÄÄRITELMÄ 1.29. Matriisi $A \in \mathbb{C}^{n\times n}$ on unitaarinen, jos sen sarakevektorit muodostavat ortonormaalin joukon avaruudessa $\mathbb{C}^n$.

LAUSE 1.30. *Matriisi $A \in \mathbb{C}^{n\times n}$ on unitaarinen, jos ja vain jos*

$$A^H A = AA^H = I.$$

Unitaarisen matriisin aste on yhtä suuri kuin sen sarakkeiden määrä. Lisäksi unitaariselle matriisille A pätee

$$A^{-1} = IA^{-1} = A^H AA^{-1} = A^H.$$

Unitaarisella matriisilla on eräs hyödyllinen ominaisuus. Seuraava tulos sanoo, että unitaarisella matriisilla kertominen säilyttää vektorin pituuden.

LAUSE 1.31. *Jos matriisi* $A \in \mathbb{C}^{n\times n}$ *on unitaarinen, niin*

$$\|A\mathbf{x}\| = \|\mathbf{x}\|.$$

TODISTUS. Koska

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^H A\mathbf{x} = \mathbf{x}^H A^H A\mathbf{x} = \mathbf{x}^H\mathbf{x} = \|\mathbf{x}\|^2,$$

niin

$$\|A\mathbf{x}\| = \|\mathbf{x}\|.$$

□

Määritellään seuraavaksi aliavaruuden ortogonaalikomplementti. Kyse on avaruuden $\mathbb{C}^n$ vektoreista, jotka ovat ortogonaalisia jonkin sen aliavaruuden kaikkien vektorien kanssa. Tätä peruskäsitettä tarvitaan välittömästi sitä seuraavassa lauseessa.

MÄÄRITELMÄ 1.32. Olkoon Y avaruuden $\mathbb{C}^n$ aliavaruus. Tällöin joukkoa

$$Y^\perp = \{\mathbf{x} \in \mathbb{C}^n : \mathbf{x}^H\mathbf{y} = 0, \text{ jokaisella } \mathbf{y} \in Y\}$$

sanotaan joukon Y ortogonaalikomplementiksi.

Seuraava tulos kertoo, että lineaarikuvauksen ydin ja kompleksikonjugoidun riviavaruuden kuvajoukon ortogonaalikomplementti ovat samoja joukkoja. Toisaalta tulos kertoo myös sen, että kompleksikonjugoitu riviavaruuden ydin ja sarakeavaruuden kuvajoukon ortogonaalikomplementti ovat myös samoja joukkoja.

LAUSE 1.33. *Jos* $A \in \mathbb{C}^{m\times n}$*, niin* $N(A) = R(A^H)^\perp$ *ja* $N(A^H) = R(A)^\perp$*.*

TODISTUS. Vektorien $\mathbf{x} \in N(A)$ ja $\mathbf{y} = A^H\mathbf{z} \in R(A^H)$ välinen pistetulo

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^H\mathbf{x} = (A^H\mathbf{z})^H\mathbf{x} = \mathbf{z}^H(A^H)^H\mathbf{x} = \mathbf{z}^H A\mathbf{x} = \mathbf{z}^H\mathbf{0} = 0.$$

Tällöin pätee $N(A) \perp R(A^H)$ ja toisaalta myös $N(A) \subset R(A^H)^\perp$. Jos $\mathbf{b} \in R(A^H)^\perp$, niin vektori $\mathbf{b}$ on kohtisuorassa jokaista matriisin A^H saraketta vastaan. Siispä $A\mathbf{b} = \mathbf{0}$ eli $\mathbf{b} \in N(A)$. Tällöin $R(A^H)^\perp \subset N(A)$. Täten

$$N(A) = R(A^H)^\perp.$$

Kun merkitään $B = A^H$, niin saadaan

$$N(A^H) = N(B) = R(B^H)^\perp = R(A)^\perp.$$

□

Lopuksi esitetään kaksi alivaruuksiin liittyvää tulosta todistuksineen. Tulosten välissä määritellään suoran summan käsite. Ensimmäinen tulos kertoo, että aliavaruuden ja sen ortogonaalikomplementin dimensioiden summa vastaa koko avaruuden dimensiota. Toinen tulos kertoo, että mikä tahansa vektoriavaruuden vektori on esitettävissä sen alivaruuden ja aliavaruuden ortogonaalikomplementin vektoreiden suorana summana. Lähteenä on käytetty Leonin kirjaa [**5**].

LAUSE 1.34. *Jos* S *on avaruuden* $\mathbb{C}^m$ *aliavaruus, niin* $\dim S + \dim S^\perp = m$*. Lisäksi jos* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$ *on kanta aliavaruudelle* S *ja* $\{\mathbf{x}_{r+1}, \mathbf{x}_{r+2}, \ldots, \mathbf{x}_m\}$ *on kanta aliavaruudelle* $S^\perp$*, niin* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r, \mathbf{x}_{r+1}, \mathbf{x}_{r+2}, \ldots, \mathbf{x}_m\}$ *on avaruuden* $\mathbb{C}^m$ *kanta.*

Todistus. Jos $S = \{\mathbf{0}\}$, niin $S^\perp = \mathbb{C}^m$. Tällöin

$$\dim S + \dim S^\perp = 0 + m = m.$$

Oletetaan seuraavaksi, että joukossa S on muutakin kuin nollavektori. Toisin sanoen olkoon $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$ joukon S kanta. Olkoon lisäksi $X \in \mathbb{C}^{r\times m}$ siten, että $R(X^T) = S$. Tällöin Lauseen 1.33 nojalla

$$S^\perp = R(X^T)^\perp = N(X)$$

ja Lauseen 1.26 nojalla

$$\dim S^\perp = \dim N(X) = m - r.$$

Siispä $\dim S + \dim S^\perp = m$.

Osoitetaan seuraavaksi, että joukko $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r, \mathbf{x}_{r+1}, \mathbf{x}_{r+2}, \ldots, \mathbf{x}_m\}$ on avaruuden $\mathbb{C}^m$ kanta. Näin on mikäli vektorit $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ ovat lineaarisesti riippumattomia. Olkoon $\mathbf{y} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r$ ja $\mathbf{z} = c_{r+1}\mathbf{x}_{r+1} + \cdots + c_m\mathbf{x}_m$. Joillakin c_i pätee

$$\mathbf{y} + \mathbf{z} = \mathbf{0}$$

eli

$$\mathbf{y} = -\mathbf{z}.$$

Tällöin $\mathbf{y}, \mathbf{z} \in S \cap S^\perp$. Koska aliavaruuksien S ja $S^\perp$ ainoa yhteinen vektori on nollavektori, niin

$$c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r = \mathbf{0} \qquad \text{ja} \qquad c_{r+1}\mathbf{x}_{r+1} + \cdots + c_m\mathbf{x}_m = \mathbf{0}.$$

Oletuksen nojalla vektorit $\mathbf{x}_1, \ldots, \mathbf{x}_r$ ja $\mathbf{x}_{r+1}, \ldots, \mathbf{x}_m$ ovat lineaarisesti riippumattomia, niin

$$c_1 = c_2 = \cdots = c_r = 0 \qquad \text{ja} \qquad c_{r+1} = c_{r+2} = \cdots = c_m = 0.$$

Tällöin vektorit $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ ovat lineaarisesti riippumattomia ja muodostavat kannan avaruudelle $\mathbb{C}^m$. □

Määritelmä 1.35. Olkoon U ja V vektoriavaruuden $W \in \mathbb{C}^n$ aliavaruuksia. Jos jokaiselle $\mathbf{w} \in W$ on olemassa yksikäsitteinen esitys

$$\mathbf{w} = \mathbf{u} + \mathbf{v}, \text{ missä } \mathbf{u} \in U \text{ ja } \mathbf{v} \in V,$$

niin W on aliavaruuksien U ja V suora summa. Lyhyemmin ilmaistuna $W = U \oplus V$.

Lause 1.36. *Jos joukko S on joukon $\mathbb{C}^n$ aliavaruus, niin*

$$\mathbb{C}^n = S \oplus S^\perp.$$

Todistus. Käsitellään ensin triviaalit tapaukset. Jos $S = \{\mathbf{0}\}$, niin välttämättä $S^\perp = \mathbb{C}^n$ ja väite pätee. Väite pätee myös tapauksessa $S = \mathbb{C}^n$. Jatketaan todistusta tapauksella, jossa joukon S dimensio on r, jolle pätee $0 < r < n$. Lauseen 1.34 nojalla jokainen $\mathbf{x} \in \mathbb{C}^n$ on esitettävissä muodossa

$$\mathbf{x} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r + c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n.$$

Edellinen esitys on siis muodostettu joukkojen S ja $S^\perp$ kantavektoreista lisättynä kertoimilla c_i. Merkitään seuraavaksi

$$\mathbf{u} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r \text{ ja } \mathbf{v} = c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n.$$

Tällöin $\mathbf{u} \in S$, $\mathbf{v} \in S^{\perp}$ ja $\mathbf{x} = \mathbf{u} + \mathbf{v}$. Yksikäsitteisyyden osoittamiseksi oletetaan, että vektorille $\mathbf{x}$ on olemassa toinenkin esitys $\mathbf{x} = \mathbf{y} + \mathbf{z}$, missä $\mathbf{y} \in S$, $\mathbf{z} \in S^{\perp}$. Vektorin $\mathbf{x}$ esityksistä saadaan

$$\mathbf{u} + \mathbf{v} = \mathbf{y} + \mathbf{z}$$
$$\mathbf{u} - \mathbf{y} = \mathbf{y} - \mathbf{v}.$$

Koska $\mathbf{u} - \mathbf{y} \in S$ ja $\mathbf{y} - \mathbf{v} \in S^{\perp}$, niin molempien vektoreiden tulee kuulua joukkoon $S \cap S^{\perp}$. Toisaalta tiedetään, että

$$S \cap S^{\perp} = \{\mathbf{0}\}.$$

Tämä tarkoittaa, että

$$\mathbf{u} = \mathbf{y} \text{ ja } \mathbf{v} = \mathbf{z},$$

mikä todistaa yksikäsitteisyyden. □

1.9. Ominaisarvoteoriaa

Lineaarialgebran toisella peruskurssilla [**4**] käsiteltiin ominaisarvoyhtälöitä. Luvussa esitellään kompleksisten matriisien ominaisarvoyhtälö ja todistetaan, että hermiittisen matriisin 1.12 ominaisarvot ovat reaalisia. Lopuksi määritellään matriisin singulaariarvo. Lähteenä on käytetty Leonin kirjaa [**5**].

Määritelmä 1.37. Olkoon $A \in \mathbb{C}^{n \times n}$ matriisi. Skalaari λ on matriisin A ominaisarvo, jos on olemassa $\mathbf{x} \neq \mathbf{0}$ siten, että

$$A\mathbf{x} = \lambda \mathbf{x}.$$

Vektorin $\mathbf{x}$ sanotaan olevan ominaisarvon λ ominaisvektori.

Lause 1.38. *Hermiittisen matriisin ominaisarvot ovat reaalisia.*

Todistus. Olkoon matriisi $A \in \mathbb{C}^{n \times n}$ hermiittinen. Olkoon λ matriisin A ominaisarvo ja $\mathbf{x}$ sitä vastaava ominaisvektori. Jos $\alpha = \mathbf{x}^H A \mathbf{x}$, niin

$$\overline{\alpha} = \alpha^H = (\mathbf{x}^H A \mathbf{x})^H = (A\mathbf{x})^H \mathbf{x} = \mathbf{x}^H A^H \mathbf{x} = \mathbf{x}^H A \mathbf{x} = \alpha.$$

Toisin sanoen $\alpha \in \mathbb{R}$. Lauseen 1.19 nojalla

$$\alpha = \mathbf{x}^H A \mathbf{x} = \mathbf{x}^H \lambda \mathbf{x} = \lambda \|\mathbf{x}\|^2,$$

jolloin

$$\lambda = \frac{\alpha}{\|\mathbf{x}\|^2} \in \mathbb{R}.$$

□

Määritelmä 1.39. Matriisin $A \in \mathbb{C}^{m \times n}$ singulaariarvo σ on symmetrisen matriisin $A^H A$ ominaisarvon λ neliöjuuri. Toisin sanoen

$$\sigma_i = \sqrt{\lambda_i}, \text{ missä } j = 1, \ldots, n.$$

1.10. Diagonalisoituvuus

Matriisin diagonalisoituvuutta käsiteltiin lineaarialgebran toisella peruskurssilla [**4**]. Diagonaalimatriisi on siis neliömatriisi, jonka ei-diagonaalialkiot ovat nollia. Asetetaan seuraavaksi ehto matriisin diagonalisoituvuudelle.

MÄÄRITELMÄ 1.40. Matriisi $A \in \mathbb{C}^{n\times n}$ on diagonalisoituva, jos on olemassa kääntyvä matriisi V ja diagonaalinen matriisi D siten, että

$$A = VDV^{-1} \qquad \text{eli} \qquad V^{-1}AV = D.$$

Tällöin sanotaan, että matriisi V diagonalisoi matriisin A. Jos matriisi V sattuu olemaan unitaarinen, niin lausekkeissa esiintyvä matriisi V^{-1} voidaan korvata matriisilla V^H. Lisäksi matriisin A ominaisarvot ovat diagonaalimatriisin

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

lävistäjäalkiot.

Esitellään lause, jonka mukaan hermiittinen matriisi on diagonalisoitavissa unitaarisen matriisin avulla. Lauseen todistamiseksi oletetaan tiedetyksi, että jokaisella neliömatriisilla A on unitaarinen matriisi V siten, että $V^HAV = T$ on yläkolmiomatriisi. Sen todistukseen ei tässä tutkielmassa perehdytä. Asiasta kiinnostunut lukija voi perehtyä siihen tarkemmin lukemalla lähdettä [**5**].

LAUSE 1.41. *Jos A on hermiittinen matriisi, niin on olemassa unitaarinen matriisi V, joka diagonalisoi matriisin A.*

TODISTUS. On olemassa esitys $V^HAV = T$, missä V on unitaarinen matriisi ja T on yläkolmiomatriisi. Tällöin

$$T^H = (V^HAV)^H = V^HA^HV = V^HAV = T.$$

Siispä matriisi T on hermiittinen ja diagonaalinen matriisi. □

LUKU 2

Singulaariarvohajotelma

Tutkielman toinen pääluku käsittelee matriisin singulaariarvohajotelmaa. Äärellisulotteinen ja kompleksinen matriisi voidaan esittää kolmen matriisin tulona. Nimensä mukaisesti hajotelmaan liittyy esitettävän matriisin singulaariarvot, joihin tutustuttiin lyhyesti ensimmäisessä luvussa. Aluksi asetetaan hajotelmalle tarkka määritelmä. Luvussa seurataan lähdettä [**5**].

Määritelmä 2.1. Olkoon matriisi $A \in \mathbb{C}^{m\times n}$, jolle $m \geq n$. Matriisi A on tällöin esitettävissä singulaariarvohajotelmana siten, että

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

missä $U \in \mathbb{C}^{m\times m}$ on unitaarinen matriisi, $V \in \mathbb{C}^{n\times n}$ on unitaarinen matriisi ja $\Sigma \in \mathbb{R}^{m\times n}$ on matriisi, joka on saatu lisäämällä singulaariarvot sisältävään diagonaalimatriisiin nollarivit $n+1, n+2, \ldots, m$. Lisäksi diagonaalialkiot σ_i ovat ei-negatiivisia reaalilukuja. Toisin sanottuna

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_n \\ \vdots & & & \\ 0 & & & 0 \end{bmatrix}.$$

Seuraavaksi todistetaan, että singulaariarvohajotelma on olemassa kaikille matriiseille joukossa $\mathbb{C}^{m\times n}$. Todistus on pitkähkö ja siihen tarvitaan paljon aikaisemmin opittua teoriaa. Todistuksen yleinen rakenne on sellainen, jossa ensin määritetään hajotelman matriisit U, V ja Σ. Tämän jälkeen näytetään matriisitulon olevan yhtä kuin valittu matriisi.

Lause 2.2. *Jokaisella $m \times n$ matriisilla on olemassa singulaariarvohajotelma.*

Todistus. Olkoon $A \in \mathbb{C}^{m\times n}$ matriisi. Matriisille $A^HA \in \mathbb{C}^{n\times n}$ on voimassa

$$(A^HA)^H \overset{\text{L.1.14}}{=} A^H(A^H)^H \overset{\text{L.1.14}}{=} A^HA.$$

Määritelmän 1.12 nojalla matriisi A^HA on hermiittinen. Lauseen 1.38 nojalla matriisille A^HA pätee, että sen ominaisarvot ovat reaalisia. Lauseen 1.41 nojalla matriisi A^HA on diagonalisoitavissa unitaarisen matriisin $V \in \mathbb{C}^{n\times n}$ avulla. Matriisin A^HA ominaisarvot ovat matriisin

$$V^H(A^HA)V = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

diagonaalialkiot.

Näytetään seuraavaksi toteen, että matriisin $A^H A$ ominaisarvot ovat ei-negatiivisia. Olkoon λ matriisin $A^H A$ ominaisarvo ja $\mathbf{x}$ ominaisvektori. Tällöin

$$\|A\mathbf{x}\|^2 \overset{\text{L.1.19}}{=} (A\mathbf{x})^H A\mathbf{x} \overset{\text{L.1.14}}{=} \mathbf{x}^H A^H A\mathbf{x} \overset{\text{M.1.37}}{=} \mathbf{x}^H \lambda \mathbf{x} = \lambda \mathbf{x}^H \mathbf{x} \overset{\text{L.1.19}}{=} \lambda \|\mathbf{x}\|^2.$$

Edellä mainitusta lausekkeesta saadaan ominaisarvolle johdettua kaava

$$\lambda = \frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2}, \text{ missä } \mathbf{x} \neq \mathbf{0}.$$

Koska kompleksisten vektorien normit ovat ei-negatiivisia, niin ominaisarvot λ ovat ei-negatiivisia. On siis osoitettu, että matriisin $A^H A$ ominaisarvot ovat ei-negatiivisia.

Voidaan olettaa, että matriisin V sarakkeet on aseteltu siten, että ominaisarvot ovat asettuneet matriisin $V^H(A^H A)V$ diagonaalille suuruusjärjestykseen. Toisin sanoen

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

Määritelmän 1.39 nojalla matriisin A singulaariarvot asetetaan siten, että

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

Olkoon $r = \operatorname{rank}(A)$. Lauseen 1.28 nojalla myös $\operatorname{rank}(A^H A) = r$. Tiedetään, että hermiittisen matriisin aste on yhtä suuri kuin sen nollasta poikkeavien ominaisarvojen lukumäärä. Tällöin matriisin $A^H A$ ominaisarvoille pätee

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 \text{ ja } \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0.$$

Sama relaatio pätee myös sen singulaariarvoille:

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

Olkoon

$$V_1 = (\mathbf{v}_1, \ldots \mathbf{v}_r), V_2 = (\mathbf{v}_{r+1}, \ldots \mathbf{v}_n) \text{ ja } \Sigma_1 = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}.$$

Kuvaillaan hiukan määriteltyjä matriiseja. Matriisi $V_1 \in \mathbb{C}^{n \times r}$ sisältää diagonalisoivan matriisin V sarakkeet, joita vastaavat ominaisarvot ovat positiivisia. Matriisi $V_2 \in \mathbb{C}^{n \times (n-r)}$ sisältää diagonalisoivan matriisin V sarakkeet, joita vastaavat ominaisarvot ovat nollia. Matriisin $\Sigma_1 \in \mathbb{R}^{r \times r}$ diagonaalilla on nollasta poikkeavat singulaariarvot $\sigma_1, \sigma_2, \ldots, \sigma_r$ ja muut alkiot ovat nollia.

Konstruoidaan seuraavaksi singulaariarvohajotelman keskimmäinen matriisi Σ. Tämä tapahtuu siten, että laajennetaan matriisia Σ_1 nollariveillä ja nollasarakkeilla niin, että saadaan $m \times n$ matriisi. Toisin sanoen

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & 0 & \dots & 0 \\ & \sigma_2 & & & & & \vdots \\ & & \ddots & & & & \\ & & & \sigma_r & & & \\ 0 & & & & 0 & & \\ \vdots & & & & & \ddots & \\ 0 & & & \dots & & & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix}.$$

Matriisin $V_2 \in \mathbb{C}^{n \times (n-r)}$ sarakkeet $\mathbf{v}_j$ ovat matriisin $A^H A$ ominaisvektoreita, joille

$$A^H A \mathbf{v}_j = 0 \mathbf{v}_j = \mathbf{0}, \text{ missä } j = r+1, \dots, n.$$

Lemman 1.27 nojalla matriisien $A \in \mathbb{C}^{m \times n}$ ja $A^H A$ ytimet ovat samoja. Tällöin matriisitulo $A\mathbf{v}_j$ on myös nollavektori. Matemaattisemmin ilmaistuna

$$AV_2 = (A\mathbf{v}_{r+1}, \dots, A\mathbf{v}_n) = O_{m \times (n-r)}. \tag{2.1}$$

Todettiin aikaisemmin, että matriisi $V \in \mathbb{C}^{n \times n}$ on unitaarinen. Lauseen 1.30 nojalla pätee

$$I = VV^H = V_1 V_1^H + V_2 V_2^H \text{ ja} \tag{2.2}$$

$$A = AI \overset{(2.2)}{=} AV_1 V_1^H + AV_2 V_2^H \overset{(2.1)}{=} AV_1 V_1^H. \tag{2.3}$$

Todistuksessa ollaan päästy vaiheeseen, jossa ollaan konstruoitu singulaariarvohajotelman matriisit $V \in \mathbb{C}^{n \times n}$ ja $\Sigma \in \mathbb{R}^{m \times n}$. Nyt riittää löytää unitaarinen matriisi $U \in \mathbb{C}^{m \times m}$ siten, että

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

Kun edellistä matriisiyhtälöä kerrotaan oikealta matriisilla V, saadaan matriisiyhtälö

$$AV = U\Sigma \tag{2.4}$$

Tarkastelemalla r ensimmäistä saraketta molemmin puolin yhtälöstä (2.4), saadaan

$$AV_r = (A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_r) \text{ ja}$$

$$U\Sigma_r = U \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & & \\ \vdots & & \ddots & \\ & & & \sigma_r \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{bmatrix} = (\mathbf{u}_1 \sigma_1, \mathbf{u}_2 \sigma_2, \dots \mathbf{u}_r \sigma_r).$$

Termejä voidaan vertailla siten, että

$$A\mathbf{v}_j = \sigma_j \mathbf{u}_j, \text{ missä } j = 1, \dots, r.$$

Kerrotaan yhtälöä puolittain luvulla $\frac{1}{\sigma_j}$:

$$\mathbf{u}_j = \frac{1}{\sigma_j} A\mathbf{v}_j, \text{ missä } j = 1, \dots, r. \tag{2.5}$$

Yhtälön (2.5) avulla määritellään matriisi

$$U_1 = (\mathbf{u}_1, \ldots, \mathbf{u}_r),$$

josta seuraa

$$AV_1 = U_1\Sigma_1. \tag{2.6}$$

Olkoon $1 \leq i \leq r$ ja $1 \leq j \leq r$. Matriisin U_1 sarakkeet muodostavat ortonormaalin kannan, koska

$$\begin{aligned}
\mathbf{u}_i^H\mathbf{u}_j &= (\frac{1}{\sigma_i}\mathbf{v}_i^H A^H)(\frac{1}{\sigma_j}A\mathbf{v}_j) \\
&= \frac{1}{\sigma_i\sigma_j}\mathbf{v}_i^H(A^H A\mathbf{v}_j) \\
&= \frac{1}{\sigma_i\sigma_j}\mathbf{v}_i^H\lambda\mathbf{v}_j \\
&= \frac{1}{\sigma_i\sigma_j}\mathbf{v}_i^H\sigma_j^2\mathbf{v}_j \\
&= \frac{\sigma_j}{\sigma_i}\mathbf{v}_i^H\mathbf{v}_j \\
&= \delta_{ij}, \text{ missä } \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.
\end{aligned}$$

Yhtälöstä (2.5) seuraa, että jokainen $\mathbf{u}_j$ kuuluu matriisin A sarakeavaruuteen. Koska joukon $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ dimensio on r, muodostaa se ortonormaalin kannan joukolle $R(A)$. Lauseen 1.26 nojalla joukon $R(A)^\perp \overset{\text{L.1.33}}{=} N(A^H)$ dimensio on $m - r$. Olkoon

$$U_2 = (\mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \ldots, \mathbf{u}_m)$$

joukon $N(A^H)$ ortonormaali kanta. Asetetaan seuraavaksi singulaariarvohajotelman matriisi

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

Lauseen 1.34 nojalla joukko $U_1 \cup U_2$ muodostaa ortonormaalin kannan avaruudelle $\mathbb{C}^m$. Määritelmän 1.29 nojalla matriisi U on unitaarinen.

Nyt ollaan asetettu kaikki kolme singulaariarvohajotelman matriisia. Lopuksi riittää osoittaa, että matriisitulo $U\Sigma V^H$ todella vastaa matriisia A. Siispä

$$\begin{aligned}
U\Sigma V^H &= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix}. \\
&= U_1\Sigma_1 V_1^H \\
&\overset{(2.6)}{=} AV_1V_1^H \\
&\overset{(2.3)}{=} A.
\end{aligned}$$

□

Esimerkki 2.3. Etsitään singulaariarvohajotelman matriisit U, Σ ja V reaaliselle matriisille

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

Täten

$$A^H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad \text{ja} \qquad A^H A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}.$$

Matriisin $A^H A$ ominaisarvot saadaan karakteristisen yhtälön ratkaisuna. Näin ollen

$$\begin{aligned} \det(\lambda I - A^H A) &= \begin{vmatrix} \lambda - 5 & -2 \\ -2 & \lambda - 2 \end{vmatrix} \\ &= (\lambda - 5)(\lambda - 2) - (-2(-2)) \\ &= \lambda^2 - 7\lambda + 6 \\ &= (\lambda - 6)(\lambda - 1). \end{aligned}$$

Matriisin $A^H A$ ominaisarvot ovat siten $\lambda_1 = 6$ ja $\lambda_2 = 1$. Näin ollen matriisin A singulaariarvot ovat

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{6} \qquad \text{ja} \qquad \sigma_2 = \sqrt{\lambda_2} = \sqrt{1} = 1.$$

Nyt voidaan muodostaa matriisi

$$\Sigma_1 = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \end{bmatrix},$$

jonka diagonaalilla on kyseiset singulaariarvot. Kun matriisia laajennetaan nollarivillä, saadaan matriisi

$$\Sigma = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{2.7}$$

Huomaa, että matriisi Σ on nyt samankokoinen kuin matriisi A, kuten pitää. Ominaisarvoa λ_1 vastaava ominaisvektori saadaan yhtälöstä

$$(\lambda_1 I - A^H A)\mathbf{x} = \mathbf{0},$$

josta sijoittamalla saadaan

$$\begin{bmatrix} 6-5 & -2 \\ -2 & 6-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 - 2x_2 \\ -2x_1 + 4x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Yhtälön ratkaisujoukoksi saadaan

$$\begin{cases} x_1 = 2x_2 \\ x_2 \in \mathbb{R} \end{cases}.$$

Ominaisarvoa $\lambda_1 = 6$ vastaava ominaisvektori on siten muotoa $\alpha(2,1)^H$, missä $\alpha \in \mathbb{R}$. Aivan vastaavalla tavalla saadaan määritettyä ominaisarvolle $\lambda_2 = 1$ ominaisvektori $\beta(1,-2)^H$, missä $\beta \in \mathbb{R}$. Valitaan ominaisvektoreiksi $\mathbf{x} = (2,1)^H$ ja $\mathbf{y} = (1,-2)^H$.

Kun kyseiset ominaisvektorit normitetaan, niin saadaan muodostettua singulaariarvohajotelman matriisi V. Toisin sanoen

$$\mathbf{v}_1 = \frac{1}{\|\mathbf{x}\|}\mathbf{x} = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}) \text{ ja}$$
$$\mathbf{v}_2 = \frac{1}{\|\mathbf{y}\|}\mathbf{y} = (\frac{1}{\sqrt{5}}, \frac{-2}{\sqrt{5}}).$$

Normitetut ominaisvektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ ovat matriisin V sarakkeet. Toisin sanoen

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

Seuraavaksi rakennetaan singulaariarvohajotelman matriisi U. Sen ensimmäinen ja toinen sarake osataan laskea suoraan. Siispä

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{30}}{6} \\ \frac{\sqrt{30}}{15} \\ \frac{\sqrt{30}}{30} \end{bmatrix} \text{ ja}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2}A\mathbf{v}_2 = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{bmatrix}.$$

Viimeisen sarakkeen löytämiseksi tulee löytää joukon $N(A^H)$ vektori $\mathbf{x}$. Tämä ratkaistaan yhtälöstä

$$A^H\mathbf{x} = \mathbf{0},$$

josta sijoittamalla saadaan

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Yhtälön ratkaisujoukoksi saadaan

$$\begin{cases} x_1 = -\frac{1}{2}x_2 \\ x_2 = 2x_3 \\ x_3 \in \mathbb{R} \end{cases}.$$

Vektori $(1,-2,-1)^H$ ja sen lineaarikombinaatiot ovat joukon $N(A^H)$ vektoreita. Valitaan vektori $(1,-2,-1)^H$, joka on normitettuna $(\frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}, \frac{-1}{\sqrt{6}})^H$. Tämä on haluttu kolmas sarake. Näin ollen

$$U = \begin{bmatrix} \frac{\sqrt{30}}{6} & 0 & \frac{1}{\sqrt{6}} \\ \frac{\sqrt{30}}{15} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{\sqrt{30}}{30} & \frac{-2}{\sqrt{5}} & \frac{-1}{\sqrt{6}} \end{bmatrix} \text{ ja}$$

$$A = U\Sigma V^H = \begin{bmatrix} \frac{\sqrt{30}}{6} & 0 & \frac{1}{\sqrt{6}} \\ \frac{\sqrt{30}}{15} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{\sqrt{30}}{30} & \frac{-2}{\sqrt{5}} & \frac{-1}{\sqrt{6}} \end{bmatrix}\begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \end{bmatrix}.$$

LUKU 3

Pseudokäänteismatriisit

Tässä pääluvussa tutustutaan pseudokäänteismatriisiin, joka on määritettävissä kaikille joukon $\mathbb{C}^{m\times n}$ matriiseille. Pseudokäänteismatriisin määritelmässä sovelletaan lähdettä [**5**], jota on muokattu kompleksisille matriiseille toimivaksi. Aloitetaan siis määrittelemällä pseudokäänteismatriisi.

Määritelmä 3.1. Olkoon $A \in \mathbb{C}^{m\times n}$ ja $A^+ \in \mathbb{C}^{n\times m}$. Matriisi A^+ on matriisin A pseudokäänteismatriisi, jos seuraavat ehdot ovat voimassa:

(1) $AA^+A = A$,
(2) $A^+AA^+ = A^+$,
(3) $(A^+A)^H = A^+A$ ja
(4) $(AA^+)^H = AA^+$.

Lisäksi pseudokäänteismatriisi A^+ on yksikäsitteinen ja olemassa jokaisella matriisilla A. Todistetaan yksikäsitteisyys ja olemassaolo seuraavaksi. Aloitetaan yksikäsitteisyydestä, jonka todistuksessa sovelletaan lähdettä [**6**].

Lause 3.2. *Jos matriisilla A on pseudokäänteismatriisi, niin se on yksikäsitteinen.*

Todistus. Olkoon $A \in \mathbb{C}^{n\times m}$. Oletetaan, että B ja C ovat matriisin A kaksi $m \times n$ -pseudokäänteismatriisia. Tällöin

$$\begin{aligned}
AB &\overset{\text{M.3.1}}{=} ACAB \overset{\text{M.3.1}}{=} (AC)^H(AB)^H \\
&\overset{\text{L.1.14}}{=} C^HA^HB^HA^H \overset{\text{L.1.14}}{=} C^H(ABA)^H \\
&\overset{\text{M.3.1}}{=} C^HA^H \overset{\text{L.1.14}}{=} (AC)^H \\
&\overset{\text{M.3.1}}{=} AC.
\end{aligned}$$

Lausekkeessa sovellettiin pseudokäänteismatriisin määritelmän kohtia (1) ja (4) sekä hermiittisen transpoosin laskusääntöä (3). On siis osoitettu, että $AB = AC$. Vastaavalla tavalla voidaan osoittaa, että $BA = CA$. Edellä mainituista tuloksista ja pseudokäänteismatriisin ominaisuuksista seuraa, että

$$B = BAB = BAC = CAC = C.$$

Toisin sanoen matriisit B ja C ovat samoja. Siispä jokaisella matriisilla A on korkeintaan yksi pseudokäänteismatriisi. $\square$

Seuraava lause osoittaa, että pseudokäänteismatriisi on olemassa kaikille matriiseille. Lausetta seuraa todistus, jossa sovelletaan aikaisemmin esitettyä singulaariarvohajotelmaa. Todistuksen lähteenä on [**5**].

Lause 3.3. *Jokaisella $m\times n$ matriisilla on olemassa $n\times m$ pseudokäänteismatriisi.*

TODISTUS. Olkoon $A \in \mathbb{C}^{m\times n}$. Matriisin singulaariarvohajotelman nojalla

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

missä $U \in \mathbb{C}^{m\times m}$ on unitaarinen matriisi, $\Sigma \in \mathbb{R}^{m\times n}$ on singulaariarvot sisältävä matriisi ja $V \in \mathbb{C}^{n\times n}$ on unitaarinen matriisi. Määritellään seuraavaksi $n\times m$ -kertoiminen matriisi A^+, jolle osoitetaan Määritelman 3.1 ehdot (1)-(4). Olkoon matriisi

$$A^+ = V\Sigma_k U^H, \text{ missä}$$

$$\Sigma_k = \left[\begin{array}{ccc|c} \frac{1}{\sigma_1} & & & \\ & \ddots & & O \\ & & \frac{1}{\sigma_r} & \\ \hline & O & & O \end{array}\right] \in \mathbb{R}^{n\times m}.$$

Toisin sanoen matriisi Σ_k saadaan matriisista Σ muuttamalla nollasta poikkeavat singulaariarvot käänteisluvuiksi ja lisäksi matriisi on vielä transponoitu. Todistuksen jatkon kannalta on syytä huomata, että

$$\Sigma\Sigma_k\Sigma = \Sigma \qquad \text{ja} \qquad \Sigma_k\Sigma\Sigma_k = \Sigma_k \tag{3.1}$$

sekä

$$(\Sigma_k\Sigma)^H = \Sigma_k\Sigma \qquad \text{ja} \qquad (\Sigma\Sigma_k)^H = \Sigma\Sigma_k. \tag{3.2}$$

Valitulle matriisille A^+ on voimassa

$$\begin{aligned} AA^+A &= (U\Sigma V^H)(V\Sigma_k U^H)(U\Sigma V^H) \\ &= U\Sigma(V^H V)\Sigma_k(U^H U)\Sigma V^H \\ &\overset{\text{L.1.30}}{=} U(\Sigma\Sigma_k\Sigma)V^H \overset{(3.1)}{=} U\Sigma V^H = A \end{aligned}$$

ja

$$\begin{aligned} A^+AA^+ &= (V\Sigma_k U^H)(U\Sigma V^H)(V\Sigma_k U^H) \\ &= V\Sigma_k(U^H U)\Sigma(V^H V)\Sigma_k U^H \\ &\overset{\text{L.1.30}}{=} V(\Sigma_k\Sigma\Sigma_k)U^H \overset{(3.1)}{=} V\Sigma_k U^H = A^+. \end{aligned}$$

Pseudokäänteismatriisin ehdot (1) ja (2) ovat siis voimassa. Lisäksi pätee

$$\begin{aligned} A^+A &= (V\Sigma_k U^H)(U\Sigma V^H) \\ &= V\Sigma_k(U^H U)\Sigma V^H \\ &\overset{\text{L.1.30}}{=} V(\Sigma_k\Sigma)V^H \end{aligned}$$

ja

$$\begin{aligned} AA^+ &= (U\Sigma V^H)(V\Sigma_k U^H) \\ &= U\Sigma(V^H V)\Sigma_k U^H \\ &\overset{\text{L.1.30}}{=} U(\Sigma\Sigma_k)U^H. \end{aligned}$$

Koska myös

$$(A^+A)^H = (V(\Sigma_k\Sigma)V^H)^H = V\Sigma^H\Sigma_k^H V^H = V(\Sigma_k\Sigma)^H V^H \overset{(3.2)}{=} V(\Sigma_k\Sigma)V^H$$

ja

$$(AA^+)^H = (U(\Sigma\Sigma_k)U^H)^H = U\Sigma_k^H\Sigma^H U^H = U(\Sigma\Sigma_k)^H U^H \overset{(3.2)}{=} U(\Sigma\Sigma_k)U^H,$$

niin pseudokäänteismatriisin ehdot (3) ja (4) tulevat osoitetuksi. On siis osoitettu, että matriisi A^+ on matriisin A pseudokäänteismatriisi. □

Tähän mennessä on määritelty pseudokäänteismatriisi sekä todistettu sen yksikäsitteisyys ja olemassaolo. Seuraavaksi käsitellään pseudokäänteismatriin A^+ laskemista esimerkin avulla valitulle matriisille A, joka on jo tuttu Esimerkistä 2.3. Laskussa sovelletaan edellisen luvun singulaariarvohajotelmaa.

Esimerkki 3.4. Matriisin

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

singulaariarvohajotema on

$$U\Sigma V^H = \begin{bmatrix} \frac{\sqrt{30}}{6} & 0 & \frac{1}{\sqrt{6}} \\ \frac{\sqrt{30}}{15} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{\sqrt{30}}{30} & \frac{-2}{\sqrt{5}} & \frac{-1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \end{bmatrix}.$$

Tällöin

$$A^+ = V\Sigma_k U^H = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{6}^{-1} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{30}}{6} & \frac{\sqrt{30}}{15} & \frac{\sqrt{30}}{30} \\ 0 & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{-1}{3} \\ \frac{1}{6} & \frac{-1}{3} & \frac{5}{6} \end{bmatrix}.$$

3.1. Pseudokäänteismatriisien perusominaisuudet

Pseudokäänteismatriiseille tunnetaan useita erilaisia perusominaisuuksia, joita esitellään seuraavaksi. Niiden todistuksissa seurataan lähdettä [**2**].

Lause 3.5. *Jokaisella $A \in \mathbb{C}^{m\times n}$ on voimassa:*

(1) $(A^+)^+ = A$,
(2) $(A^+)^T = (A^T)^+$
(3) $\overline{A^+} = (\overline{A})^+$,
(4) $(A^+)^H = (A^H)^+$ *ja*
(5) $(\alpha A)^+ = \alpha^{-1}A^+$ *kaikilla* $\alpha \neq 0$.

Todistus. Todistetaan kohdat (1)-(5) järjestyksessä. Kohdassa (1) sovelletaan singulaariarvohajotelmaa, jonka perusteella $A = U\Sigma V^H$. Toisaalta Lauseen 3.3 perusteella $A^+ = V\Sigma_k U^H$. Kun matriisin Σ_k diagonaalialkiot muutetaan käänteisalkioikseen ja vielä transponoidaan saadaan alkuperäinen matriisi Σ. Siispä

$$(A^+)^+ = U\Sigma V^H = A.$$

Kohta (2):

$$\begin{aligned} (A^T)^+ &= ((U\Sigma V^H)^T)^+ = (\overline{V}\Sigma^T U^T)^+ \\ &= \overline{U}\Sigma_k^T\overline{V}^H = \overline{U}\Sigma_k^T V^T \\ &= (V\Sigma_k U^H)^T = (A^+)^T. \end{aligned}$$

Kohta (3):

$$\begin{aligned}\overline{A^+} &= \overline{V\Sigma_k U^H} = \overline{V}\Sigma_k \overline{U^H} \\ &= \overline{V}\Sigma_k \overline{U}^H = (\overline{U\Sigma V^H})^+ \\ &= (\overline{A})^+.\end{aligned}$$

Kohta (4):

$$\begin{aligned}(A^+)^H &= (V\Sigma_k U^H)^H = U(V\Sigma_k)^H \\ &= U\Sigma_k^H V^H = (V\Sigma^H U^H)^+ \\ &= (V(U\Sigma)^H)^+ = ((U\Sigma V^H)^H)^+ \\ &= (A^H)^+.\end{aligned}$$

Kohta (5):

$$\begin{aligned}(\alpha A)^+ &= (\alpha U\Sigma V^H)^+ = (U\Sigma_\alpha V^H)^+ \\ &= V\Sigma_{\alpha,k}U^H = \alpha^{-1}V\Sigma_k U^H \\ &= \alpha^{-1}A^+.\end{aligned}$$

□

Seuraavassa lauseessa ja sen todistuksessa seurataan lähdettä [**2**].

LAUSE 3.6. *Pseudokäänteismatriisille on voimassa seuraavat yhtälöt:*

(1) $A^+ = A^+(A^+)^H A^H$,
(2) $A^+ = A^H(A^+)^H A^+$,
(3) $A = AA^H(A^+)^H$,
(4) $A = (A^+)^H A^H A$,
(5) $A^H = A^H A A^+$ *ja*
(6) $A^H = A^+ A A^H$.

TODISTUS. Kohdat (1)-(4) todistetaan käyttämällä Määritelmää 3.1 ja hermiittisen matriisin laskusääntöjä. Toisin sanoen

$$A^+ = A^+AA^+ = A^+(AA^+)^H = A^+(A^+)^H A^H,$$

$$A^+ = A^+AA^+ = (A^+A)^H A^+ = A^H(A^+)^H A^+,$$

$$A = AA^+A = A(A^+A)^H = AA^H(A^+)^H$$

ja

$$A = AA^+A = (AA^+)^H A = (A^+)^H A^H A.$$

Kohdat (5) ja (6) todistetaan ottamalla kohtien (3) ja (4) yhtälöistä hermiittinen transpoosi. Toisin sanoen

$$A^H = ((A^+)^H A^H A)^H = ((AA^+)^H A)^H = A^H((AA^+)^H)^H = A^H AA^+$$

ja

$$A^H = (AA^H(A^+)^H)^H = (A(A^+A)^H)^H = ((A^+A)^H)^H A^H = A^+AA^H.$$

□

Seuraava tulos antaa vaihtoehtoisen tavan laskea pseudokäänteismatriisi A^+ matriisille A. Tulosta sovelletaan tutkielmassa myöhemmin. Lähteenä on [**2**].

LAUSE 3.7. *Jokaiselle matriisille* $A \in \mathbb{C}^{m\times n}$ *pätee* $(AA^H)^+ = (A^H)^+A^+$, *josta seuraa*

$$A^+ = A^H(AA^H)^+ = (A^HA)^+A^H.$$

TODISTUS. Olkoon $B = (A^H)^+A^+$. Tällöin

$$AA^H \overset{\text{L.3.6}}{=} AA^H(A^+)^H A^H = AA^H(A^+)^H A^+AA^H \overset{\text{L.3.5}}{=} (AA^H)B(AA^H).$$

Lisäksi pätee

$$B = (A^H)^+A^+ = (A^+)^H A^+AA^+ = (A^+)^H A^+AA^H(A^+)^H A^+ = BAA^HB.$$

Huomataan myös, että

$$(AA^H)B = (AA^H(A^H)^+)A^+ \overset{\text{L.3.6}}{=} AA^+$$

ja

$$B(AA^H) = (A^+)^H(A^+AA^H) = (A^+)^H A^H.$$

Saadut tulokset osoittavat, että matriisi B on matriisin AA^H pseudokäänteismatriisi. On siis osoitettu, että $(AA^H)^+ = (A^H)^+A^+$. Kun esitetyn lauseen oletuksessa vaihdetaan matriisin A paikalle A^H ja matriisin A^H paikalle A, saadaan

$$(A^HA)^+ = A^+(A^H)^+. \tag{3.3}$$

Seuraavaksi rakennetaan lauseen jälkimmäinen väite. Toisin sanoen

$$A^H(AA^H)^+ = A^H(A^H)^+A^+ \overset{\text{L.3.6(2)}}{=} A^+$$

ja

$$(A^HA)^+A^H \overset{(3.3)}{=} A^+(A^H)^+A^H \overset{\text{L.3.6(1)}}{=} A^+.$$

□

3.2. Lohkomatriisien pseudokäänteismatriisi

Tutkielmassa halutaan sivuta lohkomatriisin pseudokäänteismatriisia. Tätä varten määritellään lyhyesti ortogonaalisen projektiomatriisin käsite ja esitetään joitakin hyödyllisiä tuloksia jatkon kannalta. Todistetaan esimerkiksi, että matriisit AA^+ ja $I - AA^+$ ovat ortogonaalisia projektiomatriiseja.

MÄÄRITELMÄ 3.8. Matriisi $P \in \mathbb{C}^{n\times n}$ on ortogonaalinen projektiomatriisi, jos

$$P^2 = P = P^H.$$

ESIMERKKI 3.9. Matriisi $P = AA^+$ on ortogonaalinen projektiomatriisi, koska

$$(AA^+)^2 = AA^+AA^+ \overset{\text{M.3.1}}{=} AA^+ \overset{\text{M.3.1}}{=} (AA^+)^H.$$

Lisäksi matriisi $I - P$ on ortogonaalinen projektiomatriisi, koska

$$\begin{aligned}(I-P)^2 &= (I-P)(I-P)\\ &= I - 2P + P^2\\ &= I - P\end{aligned}$$

ja

$$(I-P)^H = I^H + (-P)^H = I - P^H = I - (AA^+)^H = I - AA^+ = I - P.$$

Vastaavalla tavalla voidaan osoittaa, että A^+A ja $I - A^+A$ ovat ortogonaalisia projektiomatriiseja. Seuraava lemma osoittaa, että ortogonaalisella projektiomatriisilla on eräs kiinnostava ominaisuus.

LEMMA 3.10. *Jos A on ortogonaalinen projektiomatriisi ja $\mathbf{z} = A\mathbf{x} \in R(A)$, niin $A\mathbf{z} = \mathbf{z}$.*

TODISTUS.

$$A\mathbf{z} = A(A\mathbf{x}) = A^2\mathbf{x} = A\mathbf{x} = \mathbf{z}.$$

□

Mainittakoon, että matriisi AA^+ on ortogonaaliprojektio joukolle $R(A)$, koska $AA^+(A\mathbf{x}) = A\mathbf{x}$. Vastaavalla tavalla matriisi A^+A on ortogonaaliprojektio joukolle $R(A^H)$, koska $A^+A(A^H\mathbf{x}) = A^H\mathbf{x}$. Edellisessä päättelyssä sovellettiin pseudokäänteismatriisin perusominaisuuksia. Todetaan myös, että $I - AA^+$ on ortogonaaliprojektio joukolle $N(A^H)$ ja $I - A^+A$ on ortogonaaliprojektio joukolle $N(A)$. Tämän osoittaa myös seuraava esimerkki.

ESIMERKKI 3.11. Huomataan, että

$$\begin{aligned}
A(I - A^+A) &= AI - AA^+A = A - A = O,\\
(I - AA^+)A &= IA - AA^+A = A - A = O,\\
A^H(I - AA^+) = A^HI - A^H(AA^+)^H &= A^H - (AA^+A)^H = A^H - A^H = O \text{ ja}\\
(I - A^+A)A^H = IA^H - (A^+A)^HA^H &= A^H - (AA^+A)^H = A^H - A^H = O.
\end{aligned}$$

Vielä lopuksi esitetään eräs hyödyllinen tulos, jonka todistuksessa on sovellettu lähdettä [**7**].

LAUSE 3.12. *Jos $A \in \mathbb{C}^{n\times n}$ on ortogonaalinen projektiomatriisi, niin kaikille matriiseille $B \in \mathbb{C}^{m\times n}$ pätee*

$$A(BA)^+ = (BA)^+ \qquad \textit{ja} \qquad (AB)^+A = (AB)^+.$$

TODISTUS. Molempien yhtälöiden todistus noudattaa samaa kaavaa. Todistetaan vain ensimmäinen yhtälö. Olkoon $C = BA$ ja $D = A(BA)^+$. Väite pätee, jos matriisi D on matriisin C peudokäänteismatriisi. Riittää siis tutkia, pätevätkö Määritelmän 3.1 ehdot matriisille D. Koska

$$\begin{aligned}
CDC &= BAA(BA)^+BA\\
&= BA(BA)^+BA\\
&= BA\\
&= C,
\end{aligned}$$

$$\begin{aligned}
DCD &= A(BA)^+BAA(BA)^+\\
&= A(BA)^+BA(BA)^+\\
&= A(BA)^+\\
&= D,
\end{aligned}$$

$$\begin{aligned}(DC)^H &= (A(BA)^+BA)^H = ((BA)^+BA)^H A^H \\ &= (BA)^+BAA^H = (BA)^+BAA \\ &= (BA)^+BA = ((BA)^+BA)^H \\ &= A^H B^H ((BA)^+)^H = AB^H((BA)^+)^H \\ &= AAB^H((BA)^+)^H = AA^H B^H ((BA)^+)^H \\ &= A((BA)^+BA)^H = A(BA)^+BA \\ &= DC\end{aligned}$$

ja

$$\begin{aligned}(CD)^H &= (BAA(BA)^+)^H \\ &= (BA(BA)^+)^H \\ &= BA(BA)^+ \\ &= BAA(BA)^+ \\ &= CD,\end{aligned}$$

niin väite pätee. □

Vielä ennen luvun päätulosta, asetetaan kaksi hyödyllistä lemmaa, jotka on esitetty lähteessä [**1**]. Näistä ensimmäinen kertoo, että pseudokäänteismatriisilla ja ortogonaaliprojektiolla on yhteys. Jälkimmäinen lemma liittyy lohkotun matriisin kuvajoukkoihin.

LEMMA 3.13. *Matriisi $G \in \mathbb{C}^{n\times m}$ on matriisin $A \in \mathbb{C}^{m\times n}$ pseudokäänteismatriisi jos ja vain jos $R(G^H) \subseteq R(A)$ ja GA on ortogonaaliprojektio joukolle $R(A^H)$.*

TODISTUS. Lemman sanoma voidaan kirjoittaa auki seuraavalla tavalla:

$$G = A^+ \iff G = LA^H, \text{ jollakin } \mathrm{L} \in \mathbb{C}^{n\times n} \text{ ja } GA = A^+A.$$

Aloitetaan todistamalla oikea suunta. Olkoon $L = A^+(A^+)^H$, jolloin

$$LA^H = A^+(A^+)^H A^H = A^+(AA^+)^H = A^+AA^+ = A^+ = G.$$

Todistetaan seuraavaksi vasen suunta. Oletusten nojalla pätee

$$LA^HA = A^+A.$$

Kerrotaan puolittain oikealta matriisilla A^+, jolloin

$$LA^HAA^+ = A^+AA^+.$$

Koska $AA^+ = (AA^+)^H = (A^+)^HA^H = (A^H)^+A^H$, niin

$$LA^H(A^H)^+A^H = A^+AA^+.$$

Pseudokäänteismatriisin määrittelyehtojen nojalla

$$LA^H = A^+.$$

Siispä

$$G = A^+.$$

□

LEMMA 3.14. *Olkoon*

$$A = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \in \mathbb{C}^{m\times n},$$

eli matriisi A on lohkottu sarakkeittain. Olkoon $P_i \in \mathbb{C}^{m\times m}$ ja $Q_i \in \mathbb{C}^{m\times m}$ ortogonaaliprojektioita siten, että

$$P_i = A_i A_i^+ \text{ ja } Q_i = I - P_i,\ i = 1, 2.$$

Tällöin seuraavat väitteet ovat ekvivalenttejä:

(1) $R(A_1) \cap R(A_2) = \{\mathbf{0}\}$,
(2) $R(A_1^H) = R(A_1^H Q_2)$ *ja*
(3) $R(A_2^H) = R(A_2^H Q_1)$.

TODISTUS. Osoitetetaan ensin, että kohdasta (2) seuraa kohta (1). Oletuksen nojalla

$$\begin{aligned} \operatorname{rank}(Q_2 A_1) &= \operatorname{rank}((A_1^H Q_2^H)^H) \\ &= \operatorname{rank}((A_1^H Q_2)^H) \\ &= \dim R((A_1^H Q_2)^H) \\ &= \dim R(A_1^H Q_2) \\ &= \dim R(A_1^H) \\ &= \operatorname{rank}(A_1). \end{aligned}$$

Yhdistämällä totena pidetyt yhtälöt

$$\operatorname{rank}(\begin{bmatrix} A_1 & A_2 \end{bmatrix}) = \operatorname{rank}(A_2) + \operatorname{rank}(Q_2 A_1)$$

ja

$$\operatorname{rank}(\begin{bmatrix} A_1 & A_2 \end{bmatrix}) = \operatorname{rank}(A_1) + \operatorname{rank}(A_2) - \dim\,[R(A_1) \cap R(A_2)],$$

niin saadun tuloksen $\operatorname{rank}(A_1) = \operatorname{rank}(Q_2 A_1)$ nojalla

$$\dim\,[R(A_1) \cap R(A_2)] = 0.$$

Siispä kohta (1) pätee. Aivan vastaavalla tavalla voidaan näyttää, että kohdasta (3) seuraa kohta (1).

Osoitetaan seuraavaksi käänteinen suunta, eli kohdasta (1) seuraa kohta (2). Kohdan (1) nojalla

$$\dim\,[R(A_1) \cap R(A_2)] = 0,$$

jolloin

$$\operatorname{rank}(\begin{bmatrix} A_1 & A_2 \end{bmatrix}) = \operatorname{rank}(A_1) + \operatorname{rank}(A_2)$$

ja

$$\operatorname{rank}(\begin{bmatrix} A_1 & A_2 \end{bmatrix}) = \operatorname{rank}(A_2) + \operatorname{rank}(Q_2 A_1).$$

Voidaan tehdä päätelmä, että

$$\operatorname{rank}(A_1) = \operatorname{rank}(Q_2 A_1),$$

josta saadaan

$$\dim R(A_1^H) = \dim R(A_1^H Q_2).$$

Lisäksi koska yleisesti pätee $R(A_1^H Q_2) \subseteq R(A_1^H)$, niin välttämättä

$$R(A_1^H) = R(A_1^H Q_2).$$

Siis kohta (2) pätee. Vastaavalla tavalla osoitetaan, että kohdasta (1) seuraa kohta (3). □

Nyt ollaan valmiita esittämään luvun päätulos, joka antaa pseudokäänteismatriisin laskukaavan sarakkeittain lohkotulle matriisille. Myös riveittäin lohkotulle matriisille on olemassa kaava, jota ei esitellä tässä tutkielmassa. Esitettävän lauseen väitteet ovat ekvivalentteja keskenään ja lausetta voi käyttää pseudokäänteismatriisin laskemisessa, kun jokin väitteistä pätee. Lausetta seuraa todistus, jonka lähteenä on [**1**].

LAUSE 3.15. *Olkoon*

$$A = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \in \mathbb{C}^{m\times n},$$

eli matriisi A on lohkottu sarakkeittain. Olkoon lisäksi $Q_1, Q_2 \in \mathbb{C}^{m\times m}$ ortogonaaliprojektioita, kuten Lemmassa 3.14. Olkoon myös matriisi

$$G = \begin{bmatrix} (Q_2A_1)^+ \\ (Q_1A_2)^+ \end{bmatrix} \in \mathbb{C}^{n\times m}.$$

Tällöin seuraavat väitteet ovat ekvivalenttejä:

(1) $G = A^+$,
(2) $G \in A\{1\} = \{G \in \mathbb{C}^{n\times m} | AGA = A\}$ *ja*
(3) $R(A_1) \cap R(A_2) = \{\mathbf{0}\}$.

TODISTUS. Jos väite (1) pätee, niin pseudokäänteismatriisin Määritelmän 3.1 nojalla $G \in A\{1\}$, eli väite (2) pätee.

Oletetaan, että väite (2) pätee. Tällöin

$$\begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} (Q_2A_1)^+ \\ (Q_1A_2)^+ \end{bmatrix} \begin{bmatrix} A_1 & A_2 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \end{bmatrix},$$

josta voi päätellä, että

$$A_1(Q_2A_1)^+A_1 = A_1 \qquad \text{ja} \qquad A_2(Q_1A_2)^+A_2 = A_2. \tag{3.4}$$

Lauseen 3.12 nojalla

$$(Q_2A_1)^+ = (Q_2A_1)^+Q_2 \qquad \text{ja} \qquad (Q_1A_2)^+ = (Q_1A_2)^+Q_1,$$

jotka sijoittamalla yhtälöihin (3.4), saadaan

$$A_1(Q_2A_1)^+Q_2A_1 = A_1 \qquad \text{ja} \qquad A_2(Q_1A_2)^+Q_1A_2 = A_2.$$

Otetaan hermiittinen transpoosi, jolloin saadaan

$$A_1^H = A_1^HQ_2(A_1^HQ_2)^+A_1^H \qquad \text{ja} \qquad A_2^H = A_2^HQ_1(A_2^HQ_1)^+A_2^H.$$

Kun merkitään, että $B = (A_1^HQ_2)^+A_1^H$ ja $C = (A_2^HQ_1)^+A_2^H$, niin

$$A_1^H = A_1^HQ_2B \qquad \text{ja} \qquad A_2^H = A_2^HQ_1C.$$

Kuva-avaruuksille pätee

$$R(A_1^H) = R(A_1^HQ_2B) \subseteq R(A_1^HQ_2) \qquad \text{ja} \qquad R(A_2^H) = R(A_2^HQ_1C) \subseteq R(A_2^HQ_1).$$

Koska triviaalisti pätee $R(A_1^HQ_2) \subseteq R(A_1^H)$, niin välttämättä $R(A_1^H) = R(A_1^HQ_2)$. Vastaavasti saadaan, että $R(A_2^H) = R(A_2^HQ_1)$. Tulos vastaa Lemman 3.14 väitteitä (2) ja (3). Tällöin $R(A_1) \cap R(A_2) = \{\mathbf{0}\}$ eli väitteestä (2) seuraa väite (3).

Oletetaan seuraavaksi, että väite (3) pätee. Väitteen (1) todistamisessa tullaan soveltamaan Lemmaa 3.13. Näytetään, että $LA^H = G$, jollakin $L \in \mathbb{C}^{n\times n}$ ja $GA = A^+A$. Huomataan ensin, että

$$\begin{aligned} GA &= \begin{bmatrix}(Q_2A_1)^+\\(Q_1A_2)^+\end{bmatrix}\begin{bmatrix}A_1 & A_2\end{bmatrix}\\ &= \begin{bmatrix}(Q_2A_1)^+Q_2\\(Q_1A_2)^+Q_1\end{bmatrix}\begin{bmatrix}A_1 & A_2\end{bmatrix}\\ &= \begin{bmatrix}(Q_2A_1)^+Q_2A_1 & (Q_2A_1)^+Q_2A_2\\(Q_1A_2)^+Q_1A_1 & (Q_1A_2)^+Q_1A_2\end{bmatrix}\\ &= \begin{bmatrix}(Q_2A_1)^+Q_2A_1 & O\\ O & (Q_1A_2)^+Q_1A_2\end{bmatrix}. \end{aligned}$$

Matriisille GA on voimassa

$$(GA)^2 = GA = (GA)^H,$$

koska

$$((Q_2A_1)^+Q_2A_1)^2 = (Q_2A_1)^+Q_2A_1 \qquad \text{ja} \qquad ((Q_1A_2)^+Q_1A_2)^2 = (Q_1A_2)^+Q_1A_2$$

sekä

$$((Q_2A_1)^+Q_2A_1)^H = (Q_2A_1)^+Q_2A_1 \qquad \text{ja} \qquad ((Q_1A_2)^+Q_1A_2)^H = (Q_1A_2)^+Q_1A_2.$$

Tämä siis osoittaa, että GA on ortogonaalinen projektiomatriisi. Todetaan vielä, että

$$\begin{aligned} GAA^H &= \begin{bmatrix}(Q_2A_1)^+Q_2A_1A_1^H\\(Q_1A_2)^+Q_1A_2A_2^H\end{bmatrix} = \begin{bmatrix}(Q_2A_1)^+(Q_2A_1)A_1^HQ_2\\(Q_1A_2)^+(Q_1A_2)A_2^HQ_1\end{bmatrix}\\ &= \begin{bmatrix}(Q_2A_1)^+(Q_2A_1)(Q_2A_1)^H\\(Q_1A_2)^+(Q_1A_2)(Q_1A_2)^H\end{bmatrix} \overset{\text{L.3.6}}{=} \begin{bmatrix}(Q_2A_1)^H\\(Q_1A_2)^H\end{bmatrix}\\ &= \begin{bmatrix}A_1^HQ_2\\A_2^HQ_1\end{bmatrix} = \begin{bmatrix}A_1^H\\A_2^H\end{bmatrix}\\ &= A^H. \end{aligned}$$

Tulos $GAA^H = A^H$ tarkoittaa, että matriisi GA on ortogonaalinen projektiomatriisi, joka projisoi joukolle $R(A^H)$. Koska ortogonaalinen projektiomatriisi on yksikäsitteinen, niin välttämättä $GA = A^+A$. Yritetään seuraavaksi muodostaa matriisi $L \in \mathbb{C}^{n\times n}$. Olkoon

$$L = \begin{bmatrix}L_{11} & L_{12}\\L_{21} & L_{22}\end{bmatrix},$$

missä

$$\begin{gathered} L_{11} = (A_1^HQ_2A_1)^+, \qquad L_{12} = -(A_1^HQ_2A_1)^+A_1^H(A_2^+)^H,\\ L_{21} = -(A_2^HQ_1A_2)^+A_2^H(A_1^+)^H \qquad \text{ja} \qquad L_{22} = (A_2^HQ_1A_2)^+. \end{gathered}$$

Kun P_i ja Q_i ovat kuten Lemmassa 3.14, niin

$$\begin{aligned}
LA^H &= \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} A_1^H \\ A_2^H \end{bmatrix} \\
&= \begin{bmatrix} L_{11}A_1^H + L_{12}A_2^H \\ L_{21}A_1^H + L_{22}A_2^H \end{bmatrix} \\
&= \begin{bmatrix} (A_1^H Q_2 A_1)^+ A_1^H - (A_1^H Q_2 A_1)^+ A_1^H (A_2^+)^H A_2^H \\ -(A_2^H Q_1 A_2)^+ A_2^H (A_1^+)^H A_1^H + (A_2^H Q_1 A_2)^+ A_2^H \end{bmatrix} \\
&= \begin{bmatrix} (A_1^H Q_2 A_1)^+ A_1^H - (A_1^H Q_2 A_1)^+ A_1^H P_2 \\ -(A_2^H Q_1 A_2)^+ A_2^H P_1 + (A_2^H Q_1 A_2)^+ A_2^H \end{bmatrix} \\
&= \begin{bmatrix} (A_1^H Q_2 A_1)^+ A_1^H (I - P_2) \\ (A_2^H Q_1 A_2)^+ A_2^H (I - P_1) \end{bmatrix} \\
&= \begin{bmatrix} (A_1^H Q_2 A_1)^+ A_1^H Q_2 \\ (A_2^H Q_1 A_2)^+ A_2^H Q_1 \end{bmatrix} \\
&= \begin{bmatrix} (A_1^H Q_2 Q_2 A_1)^+ A_1^H Q_2 \\ (A_2^H Q_1 Q_1 A_2)^+ A_2^H Q_1 \end{bmatrix} \\
&= \begin{bmatrix} ((Q_2 A_1)^H Q_2 A_1)^+ (Q_2 A_1)^H \\ ((Q_1 A_2)^H Q_1 A_2)^+ (Q_1 A_2)^H \end{bmatrix} \\
&\overset{\text{L.3.7}}{=} \begin{bmatrix} (Q_2 A_1)^+ \\ (Q_1 A_2)^+ \end{bmatrix} \\
&= G.
\end{aligned}$$

Väite (1) seuraa Lemmasta 3.13. On siis osoitettu, että lauseen kohdat (1), (2) ja (3) ovat yhtäpitäviä. □

Esimerkki 3.16. Lasketaan valitulle lohkomatriisille pseudokäänteismatriisi. Olkoon

$$A_1 = \begin{bmatrix} 2+i \\ 1-i \\ 1 \end{bmatrix} \text{ ja } A_2 = \begin{bmatrix} 2i & 1 \\ i & 0 \\ -1 & 2 \end{bmatrix},$$

jolloin

$$A = \begin{bmatrix} A_1 & A_2 \end{bmatrix} = \left[\begin{array}{c|cc} 2+i & 2i & 1 \\ 1-i & i & 0 \\ 1 & -1 & 2 \end{array}\right].$$

Koska matriisin $\begin{bmatrix} A_1 & A_2 \end{bmatrix}$ determinantti on -7, niin matriisin sarakkeet ovat lineaarisesti riippumattoimia. Tällöin $R(A_1) \cap R(A_2) = \{\mathbf{0}\}$, jolloin voidaan soveltaa Lausetta 3.15. Lasketaan tarvittavia matriiseja:

$$A_2^+ = \begin{bmatrix} \frac{1}{11} - \frac{4i}{11} & \frac{-5i}{22} & \frac{-1}{22} + \frac{2i}{11} \\ \frac{1}{11} - \frac{2i}{11} & \frac{i(-1+i)}{11} & \frac{5}{11} + \frac{i}{11} \end{bmatrix},$$

$$A_1^+ = \begin{bmatrix} \frac{1}{8} - \frac{i}{4} & \frac{1}{8} + \frac{i}{8} & \frac{1}{8} \end{bmatrix},$$

$$Q_2 = \begin{bmatrix} 0,182 & -0,367 + 0,091i & -0,091 \\ -0,367 - 0,091i & 0,777 & 0,182 + 0,045i \\ -0,091 & 0,182 - 0,045i & 0,045 \end{bmatrix} \text{ ja}$$

$$Q_1 = \begin{bmatrix} 0,375 & -0,125-0,375i & -0,25-0,125i \\ -0,125+0,375i & 0,75 & -0,125+0,125i \\ -0,25+0,125i & -0,125-0,125i & 0,875 \end{bmatrix}.$$

Matriisien

$$Q_2A_1 = \begin{bmatrix} 0,636i \\ 0,318-1,273i \\ -0,318i \end{bmatrix} \text{ ja } Q_1A_2 = \begin{bmatrix} 0,625+0,75i & -0,125-0,25i \\ -0,625+0,375i & -0,375+0,625i \\ -1-0,625i & 1,5+0,125i \end{bmatrix}$$

pseudokänteismatriisit $(Q_2A_1)^+$ ja $(Q_1A_2)^+$ saadaan singulaariarvohajotelman avulla. Siispä

$$(Q_1A_2)^+ = \begin{bmatrix} 0,286-0,286i & -0,429-0,286i & -0,143+0,143i \\ 0,143 & -0,286-0,429i & 0,429 \end{bmatrix} \text{ ja}$$
$$(Q_2A_1)^+ = \begin{bmatrix} -0,286i & 0,143+0,572i & -0,143i \end{bmatrix}.$$

Pseudokäänteismatriisi on siten

$$\begin{bmatrix} A_1 & A_2 \end{bmatrix}^+ = \left[\begin{array}{ccc} -0,286i & 0,143+0,572i & -0,143i \\ \hline 0,286-0,286i & -0,429-0,286i & -0,143+0,143i \\ 0,143 & -0,286-0,429i & 0,429 \end{array}\right].$$

LUKU 4

Pienimmän neliösumman menetelmä

Luvussa on tavoitteena määritellä pienimmän neliösumman ratkaisu $\mathbf{u} \in \mathbb{C}^n$ lineaariselle yhtälöryhmälle

$$A\mathbf{x} = \mathbf{b}, \text{ missä } A \in \mathbb{C}^{m \times n} \text{ ja } \mathbf{b} \in \mathbb{C}^m.$$

Pienimmän neliösumman menetelmässä on tarkoitus löytää vektori $\mathbf{u}$, joka minimoi lausekkeen

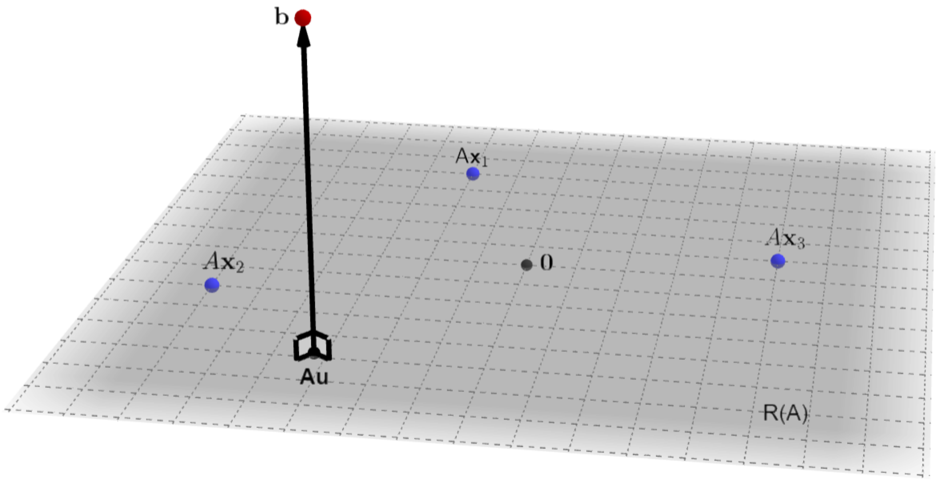
$$\|A\mathbf{u} - \mathbf{b}\|$$

eli ekvivalentisti minimoi lausekkeen

$$\|A\mathbf{u} - \mathbf{b}\|^2.$$

Tutkielmassa käytetään ilmaisua pienimmän neliösumman ratkaisu lineaariselle yhtälöryhmälle vaikka yleisella tasolla ei välttämättä ole kyse yhtälöryhmän ratkaisuista.

Lineaarialgebran peruskurssilta [**3**] tiedetään, että lineaarisella yhtälöryhmällä voi olla nolla, yksi tai äärettömän monta ratkaisua. Jos ratkaisuja on yksi tai ääretön määrä, niin ne kaikki ovat pienimmän neliösumman ratkaisuja. Tällöin edellä mainitut lausekkeet ovat nollia. Jos yhtälöryhmällä ei ole ratkaisuja, niin on kuitenkin olemassa pienimmän neliösumman ratkaisu. Viimeisimmän tilanteen geometriaa on havainnollistettu kuvassa 4.1. Kyseinen ratkaisu on siis vektori $\mathbf{u}$, jolle $A\mathbf{u} \in R(A)$ on lyhin etäisyys vektorista $\mathbf{b}$.



KUVA 4.1. Vektorin $\mathbf{b}$ lyhin etäisyys kuva-avaruuteen R(A).

Edetään kohti pienimpään neliösummaan liittyvää karakterisointilausetta. Jos vektori $\mathbf{u}$ on pienimmän neliosumman ratkaisu yhtälölle $A\mathbf{x} = \mathbf{b}$ ja merkitään $\mathbf{p} = A\mathbf{u}$, niin matriisin A sarakeavaruudessa vektori $\mathbf{p}$ on lähimpänä vektoria $\mathbf{b}$. Esitetään seuraavaksi lause, jonka mukaan tällainen vektori $\mathbf{p}$ on olemassa yksikäsitteisesti. Lauseen muotoilu ja todistus ovat lähteestä [**5**].

LAUSE 4.1. *Olkoon S avaruuden $\mathbb{C}^m$ aliavaruus. Jokaisella $\mathbf{b} \in \mathbb{C}^m$ on olemassa yksikäsitteinen vektori $\mathbf{p} \in S$, joka on lähimpänä vektoria $\mathbf{b}$ eli*

$$\|\mathbf{b} - \mathbf{y}\| > \|\mathbf{b} - \mathbf{p}\| \text{ kaikilla } \mathbf{y} \neq \mathbf{p} \in S.$$

Lisäksi vektori $\mathbf{p}$ on on lähimpänä vektoria $\mathbf{b} \in \mathbb{C}^m$ jos ja vain jos $\mathbf{b} - \mathbf{p} \in S^\perp$

TODISTUS. Lauseen 1.36 nojalla $\mathbb{C}^m = S \oplus S^\perp$, jolloin vektori $\mathbf{b} \in \mathbb{C}^m$ voidaan esittää muodossa

$$\mathbf{b} = \mathbf{p} + \mathbf{z}, \text{ missä } \mathbf{p} \in S \text{ ja } \mathbf{z} \in S^\perp.$$

Valitaan vektori $\mathbf{y} \neq \mathbf{p} \in S$ ja huomataan, että

$$\begin{aligned}\|\mathbf{b} - \mathbf{y}\|^2 &= \|(\mathbf{b} - \mathbf{p}) + (\mathbf{p} - \mathbf{y})\|^2 \\ &\overset{\text{L.1.20}}{=} \|\mathbf{b} - \mathbf{p}\|^2 + \|\mathbf{p} - \mathbf{y}\|^2\end{aligned}$$

Vektorit $\mathbf{b} - \mathbf{p} \in S^\perp$ ja $\mathbf{p} - \mathbf{y} \in S$ ovat ortogonaalisia, joten Lemmaa 1.20 saattoi soveltaa. Koska $\|\mathbf{p} - \mathbf{y}\|^2 > 0$ ja $\|\mathbf{b} - \mathbf{p}\|^2 > 0$, niin välttämättä

$$\|\mathbf{b} - \mathbf{y}\| > \|\mathbf{b} - \mathbf{p}\|.$$

Osoitetaan seuraavaksi Lauseen ekvivalenttiosa. Jos $\mathbf{b} - \mathbf{p} \in S^\perp$ ja $\mathbf{p} \in S$, niin aikaisempi todistaa vektorin $\mathbf{p}$ olevan lähinnä vektoria $\mathbf{b}$. Todistetaan seuraavaksi käänteinen osa. Jos $\mathbf{p} \in S$ on lähinnä vektoria $\mathbf{b}$, niin välttämättä $\mathbf{b} - \mathbf{p}$ on kohtisuorassa aliavaruutta S vasten. Siispä $\mathbf{b} - \mathbf{p} \in S^\perp$. □

Ollaan näytetty toteen, että vektori $\mathbf{p}$ on olemassa yksikäsitteisesti jokaiselle $\mathbf{b} \in \mathbb{C}^m$. Seuraavaksi on järkevää pohtia, kuinka pienimmän neliösumman ratkaisu $\mathbf{u}$ oikein löydetään. Merkitään ensin, että $r(\mathbf{u}) = \mathbf{b} - \mathbf{p} \in R(A)^\perp$. Lauseen 1.33 mukaan

$$N(A^H) = R(A)^\perp,$$

jolloin myös pätee

$$r(\mathbf{u}) \in N(A^H).$$

Matriisin ytimen määritelmän nojalla

$$\mathbf{0} = A^H r(\mathbf{u}) = A^H(\mathbf{b} - \mathbf{p}) = A^H(\mathbf{b} - A\mathbf{u}),$$

josta seuraa

$$A^H A\mathbf{x} = A^H\mathbf{b}. \tag{4.1}$$

Selvittääkseen pienimmän nelösumman ratkaisu lineaarisesta yhtälöryhmästä $A\mathbf{x} = \mathbf{b}$, tuleekin ratkaista lineaarinen yhtälöryhmä (4.1). Tämän yhtälöryhmän yhtälöitä kutsutaan normaaleiksi yhtälöiksi, koska yhtälöryhmässä on sarakkeita yhtä monta kuin rivejä ($n \times n$). Seuraava tulos karakterisoi pienimmän neliösumman ratkaisun normaaleille yhtälöille. Lauseen muotoilu ja todistus ovat lähteestä [**5**].

LAUSE 4.2. *Jos matriisin $A \in \mathbb{C}^{m\times n}$ aste on n, niin normaaleilla yhtälöillä*

$$A^H A\mathbf{x} = A^H\mathbf{b}$$

on yksikäsitteinen ratkaisu

$$\mathbf{u} = (A^H A)^{-1}A^H\mathbf{b},$$

missä $\mathbf{u}$ on pienimmän neliösumman ratkaisu lineaariselle yhtälöryhmälle $A\mathbf{x} = \mathbf{b}$.

TODISTUS. Todistetaan ensin, että matriisi $A^H A \in \mathbb{C}^{n\times n}$ on kääntyvä. Lauseen 1.21 nojalla matriisi $A^H A$ on kääntyvä, jos yhtälörymällä

$$A^H A\mathbf{x} = \mathbf{0}$$

on vain triviaaliratkaisu $\mathbf{z} = \mathbf{0}$. Olkoon vektori $A\mathbf{z} \in N(A^H)$ jokin ratkaisu edellä mainitulle yhtälöryhmälle. Toisaalta pätee $A\mathbf{z} \in R(A) = N(A^H)^\perp$. Koska joukkojen $N(A^H)$ ja $N(A^H)^\perp$ leikkaukseen kuuluu vain nollavektori, niin

$$A\mathbf{z} = \mathbf{0}.$$

Toisaalta tiedetään, että matriisin $A \in \mathbb{C}^{m\times n}$ aste on n. Tällöin sen sarakkeet ovat lineaarisesti riippumattomat, jolloin Lauseen 1.21 nojalla yhtälörymällä

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

on vain triviaaliratkaisu $\mathbf{x} = \mathbf{0}$. Tällöin myös $\mathbf{z} = \mathbf{0}$, ja matriisin $A^H A$ kääntyvyys on osoitettu. Lauseen 1.21 nojalla

$$\mathbf{u} = (A^H A)^{-1}A^H\mathbf{b},$$

on yksikäsitteinen ratkaisu normaaleille yhtälöryhmille ja samalla pienimmän neliösumman ratkaisu yhtälörymälle

$$A\mathbf{x} = \mathbf{b}.$$

□

4.1. Pienimmän neliösumman sovitteita

Pienimmän neliösumman menetelmällä voidaan asettaa pistejoukolle sovite, joka kuvaa aineistoa parhaiten. Tässä aliluvussa esitellään suoran, eksponenttifunktion ja ympyrän sovitteet eri aineistoille.

4.1.1. Lineaarinen sovite. Käsitellyn teorian innoittamana halutaan löytää suoran yhtälö

$$y = c_0 + c_1 x,$$

joka pienimmän neliösumman menetelmällä kuvaa seuraavaa aineistoa:

$$\begin{array}{c|c|c|c|c} x & x_1 & x_2 & \dots & x_n \\ \hline y & y_1 & y_2 & \dots & y_n \end{array}.$$

Havaintoaineiston perusteella saadaan muodostettua yhtälö

$$y_i = c_0 + c_1 x_i,$$

joka kuvaa yksittäiseen pisteeseen sovitettua suoraa. Näistä yhtälöistä muodostuu lineaarinen yhtälöryhmä, jonka matriisimuoto on seuraava:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Yhtälöryhmä voidaan esittää myös lyhyesti

$$A\mathbf{c} = \mathbf{y}.$$

Esimerkki 4.3. Sovitetaan havaintoaineistoon

x	1	3	5	7
y	3	2	5	4

lineaarinen kuvaaja pienimmän neliösumman menetelmän mukaan. Havaintoaineisto voidaan esittää muodossa

$$A\mathbf{c} = \mathbf{y},$$

missä

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 1 & 7 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} \text{ ja } \mathbf{y} = \begin{bmatrix} 3 \\ 2 \\ 5 \\ 4 \end{bmatrix}.$$

Yhtälöryhmän ratkaisut vastaavat normaalien yhtälöiden

$$A^H A\mathbf{x} = A^H\mathbf{b}$$

ratkaisuja. Lauseen 4.2 nojalla

$$\mathbf{c} = (A^H A)^{-1} A^H \mathbf{y}.$$

Matriisi $A^H A$ on kääntyvä Lauseen 1.21 nojalla. Käänteismatriisi $(A^H A)^{-1}$ voidaan laskea Gauss-Jordan eliminaatiomenetelmän avulla. Näin toimimalla saadaan

$$(A^H A)^{-1} = \begin{bmatrix} \frac{21}{20} & \frac{-1}{5} \\ \frac{-1}{5} & \frac{1}{20} \end{bmatrix}$$

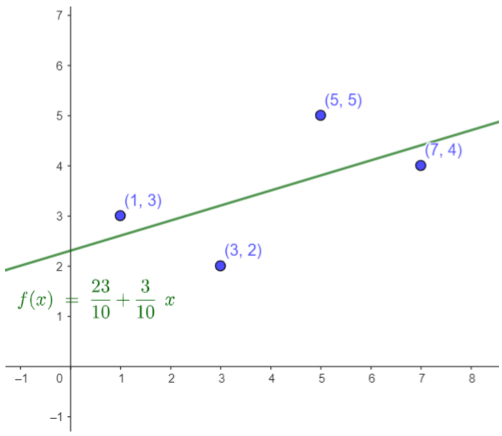
ja

$$\mathbf{c} = \begin{bmatrix} \frac{21}{20} & \frac{-1}{5} \\ \frac{-1}{5} & \frac{1}{20} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 5 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{23}{10} \\ \frac{3}{10} \end{bmatrix}.$$

Kuvaavin lineaarinen sovite aineistolle on siten

$$y = \frac{23}{10} + \frac{3}{10}x,$$

jota on havainnollistettu Kuvassa 4.2.



KUVA 4.2. Pistejoukkoa parhaiten kuvaava suora.

Joskus voi olla järkevää tehdä korkeamman asteen sovituksia aineistolle. Jos oletetaan, että aineistoa kuvaa parhaiten jokin asteen m yhtälö, niin lineaarista yhtälöryhmää tulee laajentaa siten, että

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^m \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

4.1.2. Eksponenttifunktio sovitteena. Etsitään aineistolle

x	1	2	3	4
y	$\frac{1}{2}$	$\frac{3}{2}$	2	$\frac{7}{2}$

paras mahdollinen eksponentiaalista kasvua kuvaava funktio. Tällöin funktio on muotoa

$$y = c_1 + c_0 e^x.$$

Ratkaisu saadaan yhtälöryhmän $A\mathbf{c} = \mathbf{y}$ ratkaisuna eli yhtälöryhmän

$$\begin{bmatrix} e^1 & 1 \\ e^2 & 1 \\ e^3 & 1 \\ e^4 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 2 \\ \frac{7}{2} \end{bmatrix}$$

ratkaisuna. Koska

$$(A^H A)^{-1} = \begin{bmatrix} \frac{21}{20} & \frac{-1}{5} \\ \frac{-1}{5} & \frac{1}{20} \end{bmatrix},$$

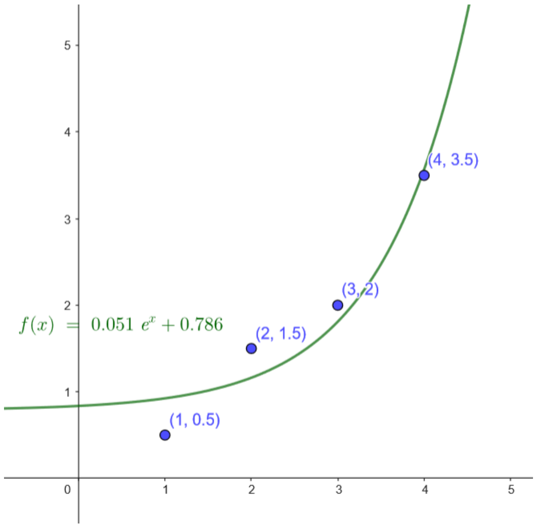
niin

$$\mathbf{c} = \begin{bmatrix} 0,0006 & -0,012854 \\ -0,012854 & 0,522483 \end{bmatrix} \begin{bmatrix} e^1 & 1 \\ e^2 & 1 \\ e^3 & 1 \\ e^4 & 1 \end{bmatrix}^H \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 2 \\ \frac{7}{2} \end{bmatrix} \approx \begin{bmatrix} 0,051 \\ 0,786 \end{bmatrix}.$$

Aineista kuvaa parhaiten funktio

$$y = 0,786 + 0,051e^x,$$

jota on havainnollistettu Kuvassa 4.3.



KUVA 4.3. Pistejoukkoa parhaiten kuvaava eksponenttifunktio.

4.1.3. Ympyrä sovitteena. Seuraavaksi yritetään löytää pistejoukolle ympyränmuotoinen sovite.

ESIMERKKI 4.4. Sovitetaan pistejoukolle

x	-1	1	2	4
y	-1	-2	2	1

ympyrä, jonka yhtälö on muotoa

$$(x-c_1)^2+(y-c_2)^2=r^2.$$

Sovitteen löytämiseksi tulee selvittää keskipiste (c_1, c_2) ja säde r. Ympyrän yhtälö saadaan muotoon

$$2xc_1+2yc_2+(r^2-c_1^2-c_2^2)=x^2+y^2.$$

Kun merkitään $c_3=r^2-c_1^2-c_2^2$, niin yhtälö on muotoa

$$2xc_1+2yc_2+c_3=x^2+y^2.$$

Tämä voidaan esittää lineaarisena yhtälöryhmänä $A\mathbf{c}=\mathbf{y}$ siten, että

$$\begin{bmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ 2x_3 & 2y_3 & 1 \\ 2x_4 & 2y_4 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} x_1^2+y_1^2 \\ x_2^2+y_2^2 \\ x_3^2+y_3^2 \\ x_4^2+y_4^2 \end{bmatrix}.$$

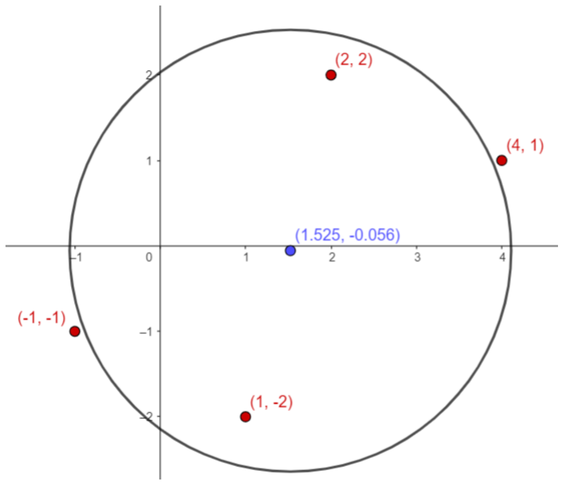
Sijoittamalla pisteparit matriiseihin, saadaan määritettyä matriisit $(A^H A)^{-1}$, A^H ja $\mathbf{y}$. Kuten edellisessä esimerkissä, ratkaisu saadaan normaalien yhtälöiden ratkaisuna

$$\begin{aligned}\mathbf{c} &= (A^H A)^{-1} A^H \mathbf{y} \\ &= \begin{bmatrix} \frac{147}{5528} & \frac{-23}{1382} & \frac{-58}{691} \\ \frac{-23}{1328} & \frac{26}{691} & \frac{46}{691} \\ \frac{-58}{691} & \frac{41}{691} & \frac{357}{691} \end{bmatrix} \begin{bmatrix} -2 & 2 & 4 & 8 \\ -2 & -4 & 4 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 8 \\ 17 \end{bmatrix} \\ &= \begin{bmatrix} \frac{4215}{2764} \\ \frac{-39}{691} \\ \frac{2357}{691} \end{bmatrix} \approx \begin{bmatrix} 1,525 \\ -0,056 \\ 3,411 \end{bmatrix}\end{aligned}$$

Sopivin ympyräsovite aineistolle olisi

$$(x - 1,525)^2 + (y + 0,056)^2 = 2,396^2,$$

jonka kuvaaja on esitetty Kuvassa 4.4.



KUVA 4.4. Pistejoukkoa parhaiten kuvaava ympyrä.

4.2. Ratkaisu pseudokäänteismatriisin avulla

Aikaisemmin esitetty tapa määrittää pienimmän neliösumman ratkaisu perustui siihen, että lineaarisen matriisiyhtälön $A\mathbf{x} = \mathbf{b}$ matriisin $A \in \mathbb{C}^{m \times n}$ aste on n. Jos matriisin A aste $r < n$, niin edellä mainitut tulokset eivät päde. Pseudokäänteimatriisin avulla voidaan kuitenkin määrittää ratkaisu yleisessä tapauksessa. Tämä ratkaisu ei ole yksikäsitteinen mutta sen normi on pienempi kuin minkään muun ratkaisun normi.

Tässä vaiheessa lukijalla tulisi olla hallussa singulaariarvohajotelman ja pseudokäänteismatriisin perusteet. Muussa tapauksessa kannattaa palata aikaisempiin lukuihin ja kerrata tärkeimmät määritelmät ja tulokset. Muistin virkistykseksi seuraavassa lauseessa esiintyvä matriisi Σ_k on määritelty Lauseen 3.3 todistuksessa. Seuraavan lauseen ja sen todistuksen muotoilu on Leonin kirjasta [**5**].

LAUSE 4.5. *Olkoon* $\mathbf{b} \in \mathbb{C}^m$. *Jos matriisin* $A \in \mathbb{C}^{m\times n}$ *aste on* $r < n$, *niin vektori*

$$\mathbf{x} = A^{+}\mathbf{b} = V\Sigma_k U^H \mathbf{b}$$

minimoi lausekkeen $\|\mathbf{b} - A\mathbf{x}\|^2$. *Lisäksi jos* $\mathbf{z}$ *on mikä tahansa muu vektori, joka minimoi lausekkeen* $\|\mathbf{b} - A\mathbf{x}\|^2$, *niin* $\|\mathbf{z}\| > \|\mathbf{x}\|$.

TODISTUS. Määritellään matriisit

$$\mathbf{c} = U^H\mathbf{b} = \begin{bmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{bmatrix} \text{ ja } \mathbf{y} = V^H\mathbf{x} = \begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix},$$

missä $\mathbf{x} \in \mathbb{C}^n$ ja $\mathbf{c}_1, \mathbf{y}_1 \in \mathbb{C}^r$. Tällöin

$$\begin{aligned}
\|\mathbf{b} - A\mathbf{x}\|^2 &\overset{1.31}{=} \|U^H(\mathbf{b} - A\mathbf{x})\|^2\\
&= \|U^H\mathbf{b} - U^H(U\Sigma V^H)\mathbf{x}\|^2\\
&\overset{1.30}{=} \|U^H\mathbf{b} - \Sigma(V^H\mathbf{x})\|^2\\
&= \|\mathbf{c} - \Sigma\mathbf{y}\|^2\\
&= \left\| \begin{bmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{bmatrix} - \begin{bmatrix}\Sigma_1 & O\\ O & O\end{bmatrix}\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} \right\|^2\\
&= \left\| \begin{bmatrix}\mathbf{c}_1 - \Sigma_1\mathbf{y}_1\\ \mathbf{c}_2\end{bmatrix} \right\|^2\\
&= \|\mathbf{c}_1 - \Sigma_1\mathbf{y}_1\|^2 + \|\mathbf{c}_2\|^2.
\end{aligned}$$

Vektori $\mathbf{c}_2$ on riippumaton vektorista $\mathbf{x}$, joten normi $\|\mathbf{b} - A\mathbf{x}\|^2$ on pienin, jos

$$\|\mathbf{c}_1 - \Sigma_1\mathbf{y}_1\| = 0.$$

Edellinen on yhtäpitävä seuraavalle yhtälölle:

$$\mathbf{c}_1 = \Sigma_1\mathbf{y}_1, \text{ missä}$$

$$\Sigma_1 = \begin{bmatrix}\sigma_1 & 0 & \dots & 0\\ 0 & \sigma_2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_r\end{bmatrix} \in \mathbb{R}^{r\times r}.$$

Kerrotaan yhtälöä puolittain matriisilla Σ_1^{-1}, jolloin saadaan

$$\mathbf{y}_1 = \Sigma_1^{-1}\mathbf{c}_1,$$

missä

$$\Sigma_1^{-1} = \begin{bmatrix}\sigma_1^{-1} & 0 & \dots & 0\\ 0 & \sigma_2^{-1} & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_r^{-1}\end{bmatrix} \in \mathbb{R}^{r\times r}.$$

Siispä vektori $\mathbf{x}$ on pienimmän neliösumman ratkaisu, jos

$$\mathbf{x} = V\mathbf{y} = V\begin{bmatrix}\Sigma_1^{-1}\mathbf{c}_1\\ \mathbf{y}_2\end{bmatrix}.$$

Voidaan valita, että vektori $\mathbf{y}_2 = \mathbf{0}$. Tällöin

$$\begin{aligned}\mathbf{x} &= V \begin{bmatrix} \Sigma_1^{-1}\mathbf{c}_1 \\ \mathbf{0} \end{bmatrix} \\ &= V \begin{bmatrix} \Sigma_1^{-1} & O \\ O & O \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} \\ &= V\Sigma_k U^H \mathbf{b} \\ &= A^+\mathbf{b}.\end{aligned}$$

Saatiin todistettua, että $\mathbf{x} = A^+\mathbf{b}$ on minimoi lausekkeen $\|\mathbf{b} - A\mathbf{x}\|^2$. Pitää vielä todistaa, että muille minimoiville vektoreille $\mathbf{z}$ pätee $\|\mathbf{z}\| > \|\mathbf{x}\|$. Vektori $\mathbf{z}$ määritellään kuten $\mathbf{x}$, eli

$$\mathbf{z} = V\mathbf{y} = V \begin{bmatrix} \Sigma_1^{-1}\mathbf{c}_1 \\ \mathbf{y}_2 \end{bmatrix},$$

missä $\mathbf{y}_2 \neq \mathbf{0}$. Tällöin

$$\|\mathbf{z}\|^2 \overset{1.31}{=} \|\mathbf{y}\|^2 = \|\Sigma_1^{-1}\mathbf{c}_1\|^2 + \|\mathbf{y}_2\|^2 > \|\Sigma_1^{-1}\mathbf{c}_1\|^2 = \|\mathbf{x}\|^2,$$

jolloin $\|\mathbf{z}\| > \|\mathbf{x}\|$. □

Esitellään seuraavaksi kaksi havainnollistavaa esimerkkiä pseudokäänteismatriisin soveltamisesta pienimmän neliösumman ratkaisun löytämisessä. Ensimmäisessä esimerkissä on kompleksikertoiminen matriisi $A \in \mathbb{C}^{n\times n}$ jonka aste $r < n$. Jälkimmäisessä esimerkissä on kompleksikertoiminen lohkomatriisi.

Esimerkki 4.6. Etsitään pienimmän neliösumman ratkaisu lineaariselle yhtälöryhmälle $A\mathbf{x} = \mathbf{b}$, missä

$$A = \begin{bmatrix} i & i \\ i & i \end{bmatrix} \text{ ja } \mathbf{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Huomataan, että $\operatorname{rank} A = 1$, joka on pienempi kuin matriisin sarakkeiden määrä. Tällöin ratkaisussa tulee soveltaa Lausetta 4.5. Pseudokäänteismatriisiksi saadaan

$$A^+ = V\Sigma_k U^H = \begin{bmatrix} \frac{-\sqrt{2}i}{2} & \frac{-\sqrt{2}}{2} \\ \frac{-\sqrt{2}i}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}^H = \begin{bmatrix} -\frac{i}{4} & -\frac{i}{4} \\ -\frac{i}{4} & -\frac{i}{4} \end{bmatrix}.$$

Tällöin

$$\mathbf{x} = A^+\mathbf{b} = \begin{bmatrix} -\frac{i}{4} \\ -\frac{i}{4} \end{bmatrix}.$$

Lauseen 4.5 nojalla normi $\|\mathbf{x}\| = \sqrt{(-\frac{1}{4})^2 + (-\frac{1}{4})^2} = \frac{1}{2\sqrt{2}}$ on pienempi kuin muiden ratkaisujen normit.

Esimerkki 4.7. Olkoon

$$\begin{bmatrix} A & B \end{bmatrix} = \left[\begin{array}{c|c} i & -i \\ i & -i \\ -i & 0 \end{array}\right] \text{ ja } \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Koska $\operatorname{rank} \begin{bmatrix} A & B \end{bmatrix} = 2$, niin tehtävässä voitaisiin soveltaa Lausetta 4.2. Sovelletaan kuitenkin Lausetta 4.5, koska se on yleispätevä. Ensin tarvitsee siis määrittää matriisit

$Q_1 = I - AA^+$, $Q_2 = I - BB^+$ ja niitä varten lasketaan matriisi A^+ ja matriisi B^+. Saadaan

$$Q_1 = I - AA^+ = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

ja

$$Q_2 = I - BB^+ = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Lopuksi määritetään matriisien Q_1B ja Q_2A pseudokäänteismatriisit singulaariarvohajotelman avulla. Saadaan

$$(Q_2A)^+ = \begin{bmatrix} 0 & 0 & i \end{bmatrix}$$

ja

$$(Q_1B)^+ = \begin{bmatrix} \frac{i}{2} & \frac{i}{2} & i \end{bmatrix}.$$

Näin ollen

$$\begin{bmatrix} A & B \end{bmatrix}^+ = \begin{bmatrix} (Q_2A)^+ \\ (Q_1B)^+ \end{bmatrix} = \begin{bmatrix} 0 & 0 & i \\ \frac{i}{2} & \frac{i}{2} & i \end{bmatrix}.$$

Tällöin

$$\mathbf{x} = \begin{bmatrix} A & B \end{bmatrix}^+ \mathbf{b} = \begin{bmatrix} i \\ i \end{bmatrix}.$$

Tämä ratkaisu on yksikäsitteinen ja $\|\mathbf{x}\| = \sqrt{1+1} = \sqrt{2}$.

Kirjallisuutta

[1] Jerzy K. Baksalary ja Oskar Maria Baksalary: *Particular formulae for the Moore-Penrose inverse of a columnwise partitioned matrix.* Linear Algebra and its applications 421 (2007) 16-23.

[2] Joao Carlos Alves Barata ja Mahir S. Hussein: *The Moore–Penrose Pseudoinverse: A Tutorial Review of the Theory.* Braz J Phys 42, 146–165 (2012).

[3] Petri Juutinen: *Lineaarinen algebra ja geometria 1.* Jyväskylän yliopisto. 2018.

[4] Petri Juutinen: *Lineaarinen algebra ja geometria 2.* Jyväskylän yliopisto. 2020.

[5] Steven J. Leon: *Linear algebra with applications.* University of Massachusetts, 8. painos, 2010.

[6] Ross MacAusland: *The Moore-Penrose Inverse and Least Squares.* University of Puget Sound, 2014.

[7] Anthony A. Maciejewski ja Charles A. Klein: *Obstacle Avoidance for Kinematically Redundant Manipulators in Dynamically Varying Environments.* The International Journal of Robotics Research 1985; 4; 109.

[8] Dan Margalit ja Joseph Rabino: *Interactive linear algebra.* Georgia Institute of Technology, 2019.

[9] Antti Vähäkangas: *Kompleksilaskenta.* Jyväskylän yliopisto, 2018.