Tampereen yliopisto

Vesa Rantanen

MATRIISIYHTÄLÖN $XAX = B$ RATKAISEMISESTA

Yhtälön ratkaisukaavan johtaminen reaalisten definiittien $n \times n$-matriisien tapauksessa

Kandidaatintyö
Tekniikan ja luonnontieteiden tiedekunta
Tarkastaja: Dos. Heikki Orelma
Marraskuu 2022

TIIVISTELMÄ

Vesa Rantanen: Matriisiyhtälön $XAX = B$ ratkaisemisesta
Kandidaatintyö
Tampereen yliopisto
Tekniikan ja luonnontieteiden kandidaattiohjelma
Marraskuu 2022

Tässä työssä johdetaan ratkaisukaava matriisiyhtälölle $XAX = B$ tapauksessa, jossa reaaliset $n \times n$-matriisit A ja B ovat definiittejä. Työ on suunnattu lukijoille, joilla on pohjatietoja matriisilaskennasta. Työn alussa käydään läpi teoriaa, jota tarvitaan yhtälön ratkaisemisessa. Tarvittavat perustiedot matriiseista esitellään vain notaatioiden tasolla.

Koska yhtälön ratkaisu on työssä pääasiallisesti kohdennettu definiiteille matriiseille, työssä määritellään matriisien eri definiittisyydet ja esitetään tarpeellisia lauseita todistuksineen. Näiden lauseiden avulla definiittisyydet yhdistetään ominaisarvoihin sekä matriisin determinanttiin ja jälkeen. Työssä keskitytään positiivisesti ja negatiivisesti definiitteihin matriiseihin, mutta myös semidefiniitit ja indefiniitit matriisit määritellään.

Työssä käsitellään kvadraattista matriisiyhtälöä, joten matriisin neliöjuuren käsite on tärkeä yhtälön ratkaisemisen kannalta. Matriisin neliöjuuri määritellään ja positiivisesti definiitin 2×2-matriisin neliöjuurelle johdetaan esityskaava Cayley-Hamilton-lauseen avulla. Lisäksi työssä esitetään matriisin neliöjuuren definiittisyyteen liittyviä lauseita.

Päätuloksena yhtälölle $XAX = B$ saadaan neljä ratkaisukaavaa eri tilanteille, joissa molemmat tunnetut matriisit ovat positiivisesti definiittejä, molemmat tunnetut matriisit ovat negatiivisesti definiittejä, matriisi A on positiivisesti definiitti matriisin B ollessa negatiivisesti definiitti tai päinvastoin. Ratkaisukaavojen käytöstä esitetään konkreettisia esimerkkejä 2×2-tapauksissa. Yhtälöä tutkitaan työssä myös yleisemmin. Yleisessä tapauksessa pystytään päättelemään matriisin X ominaisuuksia annettujen matriisien A ja B kääntyvyyden, determinanttien sekä asteiden avulla.

Avainsanat: Matriisi, Matriisiyhtälö, Definiittisyys, Matriisin neliöjuuri

SISÄLLYSLUETTELO

LYHENTEET JA MERKINNÄT

$\mathbb{R}$	Reaaliluvut
$\mathbb{C}$	Kompleksiluvut
$\mathbb{R}^{m\times n}$	Reaalinen $m \times n$-matriisiavaruus
$\mathbb{C}^{m\times n}$	Kompleksinen $m \times n$-matriisiavaruus
$\|\mathbf{x}\|$	Vektorin **x** normi
A^T	Matriisin A transpoosi
A^*	Matriisin A konjugaattitranspoosi
I	Identiteettimatriisi
$\det(A)$	Matriisin A determinantti
$\mathrm{tr}(A)$	Matriisin A jälki
A^{-1}	Matriisin A käänteismatriisi
λ_i	Ominaisarvot

1. JOHDANTO

Tässä työssä tarkastellaan kvadraattisen matriisiyhtälön ratkaisemista. Lineaarisen matriisiyhtälön $AX = B$ jälkeen luonnollinen yksinkertaisin kvadraattinen matriisiyhtälö on $XAX = B$, jossa A ja B ovat tunnettuja neliömatriiseja. Yhtälö $XAX = B$ on yksinkertainen muoto niin kutsutusta algebrallisesta Riccatin yhtälöstä. Algebrallista Riccatin yhtälöä käytetään esimerkiksi optimoinnissa [1]. Yhtälö saattaa näyttää aluksi yksinkertaiselta, mutta matriisien tapauksessa sen ratkaiseminen on haastavaa. Yhtälön ratkaisu riippuu matriisien A ja B ominaisuuksista, joista tärkein on matriisien definiittisyys. Tämän työn tarkoituksena on esitellä keino matriisiyhtälön $XAX = B$ ratkaisujen löytämiseen, kun A ja B ovat reaalisia definiittejä $n \times n$-matriiseja. Tähän vaaditaan kuitenkin joitakin tietoja matriisien ominaisuuksista ja laskusäännöistä, jotka työssä esitetään.

Työn toisessa luvussa käsitellään tarvittavat esitiedot matriisien perusominaisuuksista, joita tarvitaan seuraavissa luvuissa esitettävien tulosten ymmärtämiseen ja todistamiseen. Luvussa käsitellään esimerkiksi symmetrisyyttä, ominaisarvoja sekä diagonalisointia. Määritelmät ja lauseet esitetään luvussa vain notaatioina ilman todistuksia.

Luvussa 3 määritellään matriisin eri definiittisyydet, jotka ovat oleellisia tunnistaa työssä ratkaistavan yhtälön kannalta. Luvussa esitetään lisäksi lauseita, jotka yhdistävät definiittisyydet ominaisarvoihin, determinantteihin ja matriisin jälkeen, joiden avulla eri definiittisyyksien tunnistaminen helpottuu. Työn kannalta oleellisia ovat positiivisesti ja negatiivisesti definiitit matriisit, joten luvussa ei keskitytä semidefiniitteihin tapauksiin määritelmää pidemmälle. Luvun lopussa esitetään definiittien matriisien laskusääntöjä.

Neljäs luku käsittelee matriisin neliöjuurta. Luvussa määritellään matriisineliöjuuri ja esitetään matriisineliöjuuren definiittisyyteen liittyviä lauseita. Luvussa keskitytään myös positiivisesti definiitin 2×2-matriisin neliöjuuren eksplisiittisen esityskaavan johtamiseen. Esityskaavan johtamiseen käytetään luvussa esiteltyä Cayley-Hamilton-lausetta.

Viidennessä luvussa tutkitaan yhtälön $XAX = B$ ratkaisun ominaisuuksia yleisessä tapauksessa ja johdetaan ratkaisukaava kyseiselle yhtälölle definiittien $n \times n$-matriisien tapauksessa käyttäen työssä aikaisemmin määriteltyjä ja johdettuja matriisien ominaisuuksia. Ratkaisu riippuu yhtälön tunnettujen matriisien definiittisyyksistä, joten saadaan ratkaisukaavat, kun matriisit A ja B ovat positiivisesti definiittejä, A ja B ovat negatiivisesti definiittejä, toinen matriiseista on positiivisesti ja toinen negatiivisesti definiitti sekä päinvastoin. Luvussa esitetään lisäksi esimerkkejä ratkaisukaavasta 2×2-matriisien tapauksissa.

2. TARVITTAVIA ESITIETOJA MATRIISEISTA

Matriisiyhtälöiden, matriisin neliöjuuren ja definiittien matriisien laskusääntöjen ymmärtämiseen tarvitaan joitakin oleellisia esitietoja matriisien ominaisuuksista. Työssä on oletettu, että matriisien peruslaskusäännöt ja jotkin erityistapaukset, kuten neliömatriisi ja identiteettimatriisi, ovat lukijalle tuttuja. Tässä luvussa on esitetty esitietoja vain notaatioiden muodossa ja todistukset on sivuutettu. Ensimmäisessä luvussa lähteinä on käytetty pääosin Johnsonin, Dean Riessin ja Arnoldin teosta Introduction to linear algebra [4] sekä Hornin ja Johnsonin teosta Matrix analysis [2].

2.1 Matriisiavaruudet ja vektorit

Oletetaan, että vektori- ja matriisilaskenta ovat lukijalle entuudestaan tuttuja ja hän tuntee reaalisekä kompleksilukujen joukot. Joukot, joissa on määritelty sekä yhteen- että kertolaskut, ovat matriisiavaruuksia. Reaalista $m \times n$-avaruutta merkitään $\mathbb{R}^{m\times n}$ ja kompleksisten $m \times n$-matriisien avaruutta $\mathbb{C}^{m\times n}$. Matriisiavaruuksien perusteella pystytään määrittelemään myös vektoriavaruudet $\mathbb{R}^{m\times 1}$ ja $\mathbb{C}^{m\times 1}$, joita merkitään lyhyesti $\mathbb{R}^n$ ja $\mathbb{C}^n$.

Reaalisten matriisi- ja vektoriavaruuksien tapauksessa vektorin $\mathbf{x}$ transpoosia eli rivien ja sarakkeiden vaihtoa merkitään $\mathbf{x}^T$. Vastaavasti kompleksisten avaruuksien kohdalla on kyse konjugaattitranspoosista, jossa rivien ja sarakkeiden vaihdon lisäksi kompleksiluvut muuttuvat liittoluvuikseen. Vektorin konjugaattitranspoosia merkitään $\mathbf{x}^*$.

Määritelmä 2.1. Matriisi $A \in \mathbb{R}^{n\times n}$ on *symmetrinen*, jos $A = A^T$.

Vastaavaa ominaisuutta kutsutaan kompleksisessa avaruudessa hermiittisyydeksi.

Lause 2.2. *[2, s. 228] Matriisi A on symmetrinen, jos ja vain jos matriisi $T^T AT$ on symmetrinen kaikilla $T \in \mathbb{R}^{n\times n}$.*

Määritellään vektorin normi, jota kutsutaan myös vektorin pituudeksi, Pythagoraan lauseen avulla.

Määritelmä 2.3. Vektorin

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

normi on

$$\|\mathbf{x}\| = \sqrt{{x_1}^2 + {x_2}^2 + \cdots + {x_n}^2}.$$

Vektorin normi voidaan määrittää myös matriisitulon avulla.

Lause 2.4. *[4, s. 68] Olkoon* $\mathbf{x} \in \mathbb{C}^n$*. Vektorin normi on tällöin* $\|\mathbf{x}\| = \sqrt{\mathbf{x}^*\mathbf{x}}$.

2.2 Determinantti ja jälki

Neliömatriisien tärkeitä ominaisuuksia ovat determinantti ja jälki. Määritellään aluksi matriisin determinantti ja esitetään hyödyllinen tulos tulomatriisin determinantille.

Määritelmä 2.5. Olkoon $A \in \mathbb{C}^{n \times n}$ matriisi, jonka alkiot ovat muotoa a_{ij}. Tällöin sen *determinantti* määritellään vaakarivin suhteen rekursiivisesti

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij}),$$

missä $\det(M_{ij})$ ovat alideterminantteja. Matriisi M_{ij} saadaan matriisista A, kun on poistetaan i:nnes rivi ja j:nnes sarake. Determinantti voidaan määritellä myös sarakkeen suhteen.

Tulomatriisin tapauksessa determinantin voi ratkaista laskemalla ensin tulomatriisin arvon. Tulomatriisin determinantti voidaan kuitenkin myös selvittää tulontekijöiden determinanttien avulla.

Lause 2.6. *[4, s. 466] Olkoot* $A, B \in \mathbb{C}^{n \times n}$*. Tällöin*

$$\det(AB) = \det(A)\det(B).$$

Luvuissa 3 ja 4 tarvitaan determinantin lisäksi matriisin jäljen määritelmän tuntemista. Matriisin jälki on matriisin diagonaalialkioiden summa.

Määritelmä 2.7. Matriisin $A \in \mathbb{C}^{n \times n}$ *jälki* on

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

kun $A = [a_{ij}]$.

2.3 Matriisin kääntyvyys

Usein halutaan ratkaista muotoa $A\mathbf{x} = \mathbf{b}$ oleva matriisiyhtälö. Jos A on neliömatriisi, tällainen yhtälö voidaan ratkaista käänteismatriisin A^{-1} avulla $\mathbf{x} = A^{-1}\mathbf{b}$, jos käänteismatriisi on olemassa. Huomioitavaa on, että kaikki matriisit eivät ole kuitenkaan kääntyviä. Matriiseja, jotka eivät ole kääntyviä, kutsutaan singulaarisiksi matriiseiksi.

Määritelmä 2.8. Neliömatriisi A on *kääntyvä*, jos ja vain jos on olemassa neliömatriisi A^{-1}, jolle $AA^{-1} = I$ ja $A^{-1}A = I$. Matriisia A^{-1} kutsutaan tällöin matriisin A *käänteismatriisiksi*.

Karakterisoidaan seuraavaksi matriisin kääntyvyyttä.

Lause 2.9. *[4, s. 101][2, s. 14] Olkoon $A \in \mathbb{C}^{n\times n}$. Tällöin seuraavat väitteet ovat toistensa kanssa ekvivalentteja:*

(a) A on kääntyvä.

(b) Matriisin A sarakevektorit ovat lineaarisesti riippumattomia.

(c) Yhtälöllä $A\mathbf{x} = \mathbf{b}$ on yksiselitteinen ratkaisu.

(d) A on riviekvivalentti identiteettimatriisin I kanssa.

(e) $\det(A) \neq 0$.

Lause 2.10. *[4, s. 99] Jos matriisit A ja B ovat kääntyviä, on niillä seuraavat ominaisuudet:*

(a) $(A^{-1})^{-1} = A$

(b) $(AB)^{-1} = B^{-1}A^{-1}$ eli matriisi AB on kääntyvä

(c) $(A^T)^{-1} = (A^{-1})^T$.

Käänteismatriisin etsiminen perustuu Gaussin eliminointimenetelmään, mutta esimerkiksi 2×2-matriisin käänteismatriisille on olemassa ratkaisukaava.

Lause 2.11. *[4, s. 98] Olkoon*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

kääntyvä. Tällöin sen käänteismatriisi on

$$A^{-1} = \frac{1}{\det(\mathrm{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Kääntyvän neliömatriisin erikoistapaus on ortogonaalinen matriisi.

Määritelmä 2.12. Kääntyvä neliömatriisi A on *ortogonaalinen*, jos $A^{-1} = A^T$.

2.4 Ominaisarvot ja diagonalisointi

Tässä alaluvussa määritetään ominaisarvoihin liittyviä käsitteitä ja niiden ominaisuuksia. Matriisin ominaisarvojen ja -vektorien käsitettä tarvitaan esimerkiksi matriisien potenssien laskemisessa, sekä työssä myöhemmin esiteltävien definiittisyyden sekä neliöjuuren tutkimisessa. Lisäksi alaluvussa määritellään diagonalisoituvuus.

Määritelmä 2.13. Olkoon $A \in \mathbb{R}^{n\times n}$ matriisi. Jos on olemassa nollasta poikkeava vektori $\mathbf{x} \in \mathbb{C}^n$ ja skalaari $\lambda \in \mathbb{C}$, jotka toteuttavat yhtälön

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

on λ matriisin A *ominaisarvo* ja vektori $\mathbf{x}$ sitä vastaava *ominaisvektori*.

Määritelmä 2.14. Neliömatriisin A *karakteristiseksi polynomiksi* kutsutaan polynomia

$$p_A(\lambda) = \det(A - \lambda I).$$

Matriisin ominaisarvot voidaan selvittää karakteristisen polynomin nollakohdista. Neliömatriisin ominaisarvojen määrä riippuu neliömatriisin koosta.

Lause 2.15. *[4, s. 300] Matriisilla $A \in \mathbb{C}^{n\times n}$ on monikerrat mukaan lukien n ominaisarvoa $\lambda_i \in \mathbb{C}$.*

Reaalisen matriisin ominaisarvot voivat olla kompleksisia. Symmetrisen matriisin ominaisarvot ovat kuitenkin aina reaalisia.

Lause 2.16. *[4, s. 319] Jos matriisi $A \in \mathbb{R}^{n\times n}$ on symmetrinen, niin sen ominaisarvot $\lambda_i \in \mathbb{R}$.*

Ominaisarvojen ja -vektoreiden avulla tietynlaiset matriisit voidaan jakaa osiin eli diagonalisoida, jolloin matriisin käsitteleminen ja esimerkiksi matriisien potenssien laskemisen suorittaminen helpottuu.

Määritelmä 2.17. Jos matriisi voidaan kirjoittaa muodossa $A = SDS^{-1}$, jossa

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{2.1}$$

ja matriisi S koostuu ominaisarvoja vastaavista ominaisvektoreista $\mathbf{x}_i$, sanotaan sen olevan *diagonalisoituva*.

Lause 2.18. *[4, s. 334–335] Jos A on symmetrinen matriisi, se voidaan diagonalisoida ortogonaalisen matriisin Q avulla siten, että*

$$A = QDQ^T,$$

missä D on ominaisarvoista koostuva diagonaalimatriisi.

2.5 Matriisien kongruenssi

Kongruentit matriisit voidaan esittää toistensa avulla.

Määritelmä 2.19. Neliömatriisit $A, B \in \mathbb{R}^{n\times n}$ ovat kongruentteja, jos on olemassa matriisi T, joka toteuttaa yhtälön

$$A = TBT^T.$$

Kompleksisessa matriisiavaruudessa tätä kutsutaan $*$-kongruenssiksi.

Lause 2.20. *[2, s.282] Symmetriset matriisit A ja B ovat kongruentteja, jos ja vain jos niillä on sama määrä positiivisia ja negatiivisia ominaisarvoja. Vastaavasti hermiittiset matriisit A ja B ovat* $*$*-kongruentteja, jos ja vain jos niillä on sama määrä positiivisia ja negatiivisia ominaisarvoja.*

Lauseen perusteella voidaan todeta, että symmetrinen matriisi on kongruentti aina ominaisarvomatriisinsa kanssa.

2.6 Matriisin aste

Oletetaan, että lukija tietää, että matriisi voidaan esittää vaaka- tai pystyvektoreiden avulla. Vektoreille keskeinen käsite on lineaarinen riippumattomuus. Kun vektorit ovat keskenään lineaarisesti riippumattomia, niitä ei voida esittää toistensa avulla eli ne eivät ole toistensa monikertojen summia. Lineaarisen riippumattomuuden avulla voidaan määritellä matriisin aste.

Määritelmä 2.21. *Matriisin aste* $\mathrm{rank}(A)$ on lineaarisesti riippumattomien rivien (tai sarakkeiden) lukumäärä matriisissa.

Lause 2.22. *[2, s. 14] Olkoon $A \in \mathbb{C}^{n\times n}$ kääntyvä matriisi. Tällöin* $\mathrm{rank}(A) = n$. *Vastaavasti ei-kääntyvän eli singulaarisen matriisin $B \in \mathbb{C}^{n\times n}$ tapauksessa,* $\mathrm{rank}(A) < n$.

Lause 2.23. *[2, s. 13] Olkoon $A \in \mathbb{C}^{k\times m}$ ja $B \in \mathbb{C}^{m\times n}$. Tällöin*

$$\mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}.$$

Matriisin asteen määritelmää ja edellä olevia lauseita voidaan käyttää matriisiyhtälöiden ratkaisemisessa, kuten myöhemmin työssä huomataan.

3. MATRIISIEN DEFINIITTISYYS

Definiittisyys on tärkeä symmetrisen matriisin ominaisuus, jonka perusteella pystytään päättelemään sen muita ominaisuuksia. Symmetrisen matriisin definiittisyys määritellään kvadraattisen muodon $\mathbf{x}^T A\mathbf{x}$ avulla. Kvadraattisen muodon tulontekijöiden dimensioiden perusteella havaitaan, että sen arvo on skalaari. Luokitellaan aluksi eri definiittisyydet.

Määritelmä 3.1. Symmetrinen matriisi $A \in \mathbb{R}^{n\times n}$ on

(a) *positiivisesti definiitti*, kun $\mathbf{x}^T A\mathbf{x} > 0, \forall\, \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$

(b) *positiivisesti semidefiniitti*, kun $\mathbf{x}^T A\mathbf{x} \geq 0, \forall\, \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$

(c) *negatiivisesti definiitti*, kun $\mathbf{x}^T A\mathbf{x} < 0, \forall\, \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$

(d) *negatiivisesti semidefiniitti*, kun $\mathbf{x}^T A\mathbf{x} \leq 0, \forall\, \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$

(e) *indefiniitti*, jos se ei ole mikään edellisistä.

Matriisin ollessa positiivisesti tai negatiivisesti definiitti, sen on oltava silloin myös positiivisesti tai negatiivisesti semidefiniitti [2, s. 430]. Matriisin definiittisyyttä voidaan tutkia kvadraattisen muodon lisäksi myös edellisessä luvussa käytyjen ominaisarvojen avulla seuraavan lauseen mukaisesti.

Lause 3.2. *[4, s. 487] Olkoon $A \in \mathbb{R}^{n\times n}$ symmetrinen matriisi, jonka ominaisarvot ovat λ_i, kun $i = 1, 2, \ldots, n$. Tällöin A on*

(a) positiivisesti definiitti, jos ja vain jos $\lambda_i > 0$

(b) positiivisesti semidefiniitti, jos ja vain jos $\lambda_i \geq 0$

(c) negatiivisesti definiitti, jos ja vain jos $\lambda_i < 0$

(d) negatiivisesti semidefiniitti, jos ja vain jos $\lambda_i \leq 0$

(e) indefiniitti, jos ja vain jos λ_i saa sekä positiivisia, että negatiivisia arvoja.

Todistus. Todistus seuraa lähteen [2, s. 430] todistusta ja täydentää sitä lähteen [3, s. 114–115] avulla.

Todistetaan aluksi kohdat $(a) - (d)$. Oletetaan aluksi, että matriisi A on positiivisesti definiitti. Olkoon $\mathbf{x}$ matriisin A ominaisvektori ja λ sitä vastaava ominaisarvo eli

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

Tämä voidaan kertoa puolittain vasemmalta ominaisvektorin transpoosilla $\mathbf{x}^T$, jolloin saadaan

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

Vektorien laskusääntöjen nojalla voidaan kirjoittaa $\mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$, jossa $\|\mathbf{x}\|$ on matriisin normi, ja edellisestä yhtälöstä saadaan

$$\mathbf{x}^T A\mathbf{x} = \lambda\|\mathbf{x}\|^2.$$

Määritelmän 3.1 perusteella matriisin A ollessa positiivisesti definiitti $\mathbf{x}^T A\mathbf{x} > 0$, ja koska vektorin $\mathbf{x}$ normi on myös aina positiivinen skalaari, on todistettu, että $\lambda_i > 0$ matriisin A ollessa positiivisesti definiitti.

Oletetaan nyt, että $\lambda_i > 0$. Lauseen 2.18 mukaan

$$A = QDQ^T,$$

missä Q on ortogonaalinen matriisi, joka sisältää matriisin A normeeratut ominaisvektorit ja diagonaalimatriisi D sisältää matriisin A ominaisarvot diagonaalillaan kuten yhtälössä (2.1). Yhtälö voidaan kertoa puolittain vektorilla $\mathbf{x}^T$ vasemmalta ja vektorilla $\mathbf{x}$ oikealta, jolloin saadaan

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T QDQ^T\mathbf{x}.$$

Kirjoitetaan $\mathbf{y} = \mathbf{x}^T Q$ eli yhtälö on nyt muotoa

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}D\mathbf{y}^T,$$

jossa $\mathbf{y} = \left(y_1, y_2, \ldots, y_n\right)^T$. Matriisitulot auki laskettaessa saadaan

$$\mathbf{x}^T A\mathbf{x} = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \cdots + \lambda_n {y_n}^2 > 0.$$

Näin ollen A on positiivisesti definiitti, kun sen ominaisarvot ovat positiivisia.

Todistus on vastaava negatiivisesti definiitissä ja semidefiniiteissä tapauksissa, joten kyseisten tapausten todistuksia ei työssä erikseen esitetä.

Todistetaan vielä kohta (e). Oletetaan, että matriisi A on indefiniitti, jonka ominaisvektori on $\mathbf{x}$ ja sitä vastaava ominaisarvo λ. Edellisten kohtien todistuksen perusteella saadaan

$$\mathbf{x}^T A\mathbf{x} = \lambda\|\mathbf{x}\|^2.$$

Koska matriisi A on indefiniitti, yhtälön vasen puoli voi olla positiivinen tai negatiivinen, jolloin ominaisarvoiksi on myös tultava positiivisia ja negatiivisia arvoja vektorinormin ollessa positiivinen.

Oletetaan sitten, että λ_i saa sekä positiivisia ja negatiivisia arvoja. Kuten aiemmassa todistuksessa johdettiin,

$$\mathbf{x}^T A\mathbf{x} = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \cdots + \lambda_n {y_n}^2.$$

Koska ominaisarvot voivat olla sekä positiivisia, että negatiivisia, voi $\mathbf{x}^T A\mathbf{x}$ saada positiivisia ja negatiivisia arvoja eli matriisi A on indefiniitti. □

Lauseen 2.13 avulla huomataan helposti, jos matriisilla A on ominaisarvot $\lambda_i > 0$, niin matriisilla $-A$ on ominaisarvot $-\lambda_i < 0$. Näin ollen jos matriisi A on positiivisesti (semi)definiitti, matriisi $-A$ on negatiivisesti (semi)definiitti. [2, s. 430]

Definiittisyyttä voidaan tutkia muutenkin kuin ominaisarvojen näkökulmasta.

Lause 3.3. *Jos symmetrisen matriisin $A \in \mathbb{R}^{2\times 2}$ determinantti* $\det(A)$ *sekä jälki* $\operatorname{tr}(A)$ *ovat positiivisia, matriisi A on positiivisesti definiitti.*

Todistus. Olkoon

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

jossa $a, b, c \in \mathbb{R}$ ja $a, c \neq 0$. Oletetaan, että $\det(A) = ac - b^2 > 0$ ja jälki $\operatorname{tr}(A) = a + c > 0$. Ratkaistaan Määritelmän 2.14 mukaisen karakteristisen polynomin nollakohdat eli

$$\det(A - \lambda I) = 0.$$

Kun lasketaan tämä determinantti, saadaan

$$\lambda^2 - \operatorname{tr}(A)\lambda + \det(A) = 0,$$

mistä ratkaisemalla edelleen

$$\lambda = \frac{\operatorname{tr}(A) \pm \sqrt{\operatorname{tr}(A)^2 - 4\ \det(A)}}{2}.$$

Yhtälöstä nähdään, että λ on aina positiivinen, sillä

$$\operatorname{tr}(A) > \sqrt{\operatorname{tr}(A)^2 - 4\det(A)},$$

kun $\det(A) > 0$ ja $\operatorname{tr}(A) > 0$. Tämän lisäksi diskriminanttia tutkimalla huomataan, että yhtälöllä on aina ratkaisu, sillä

$$\begin{aligned}
D &= \operatorname{tr}(A)^2 - 4\det(A) \\
&= (a + c)^2 - 4(ac - b^2) \\
&= a^2 + 2ac + c^2 - 4ac + 4b^2 \\
&= a^2 - 2ac + c^2 + 4b^2 \\
&= (a - c)^2 + 4b^2 \geq 0.
\end{aligned}$$

Kun $\mathrm{tr}(A)$ ja $\det(A)$ ovat positiivisia, ovat myös matriisin A ominaisarvot aina positiiviset, jolloin Lauseen 3.2 nojalla matriisi A on positiivisesti definiitti. □

Usein tulee selvittää myös tulomatriisin definiittisyys. Erikseen tämän tiedon hakeminen esimerkiksi ominaisarvojen avulla ei ole aina helpoin tai nopein tapa, vaan tulomatriisin definiittisyys voidaan päätellä myös tulontekijöiden definiittisyyden avulla. Seuraava lause ja sen todistus on kirjoitettu lähteen [2, s. 485–486] perusteella.

Lause 3.4. *Jos matriisit A ja B ovat positiivisesti definiittejä, tällöin niiden tulo on positiivisesti definiitti.*

Todistus. Olkoot matriisit A ja B positiivisesti definiittejä. Lauseen 2.20 perusteella on olemassa kääntyvä matriisi $T \in \mathbb{R}^{n\times n}$, joka toteuttaa yhtälön

$$T^{-1}AT^{-T} = I,$$

sillä A ja I ovat kongruentteja molempien matriisien kaikkien ominaisarvojen ollessa positiivisia. Matriisi $T^T BT$ on puolestaan symmetrinen (Lause 2.2), joten voidaan kirjoittaa $Q^T(T^T BT)Q = D$, jossa Q on ortogonaalinen ja D on diagonaalimatriisi, jonka alkiot ovat matriisin B ominaisarvoja. Merkitään seuraavaksi $S = TQ$, jolloin

$$S^{-1}AS^{-T} = Q^T T^{-1}AT^{-T}Q = Q^T IQ = I$$

ja

$$S^T BS = Q^T T^T BTQ = D.$$

Näin ollen matriisit A ja B saadaan muotoon $A = SS^T$ ja $B = S^{-T}DS^{-1}$. Täten tulomatriisiksi saadaan

$$AB = SS^T S^{-T}DS^{-1} = SDS^{-1}.$$

Matriisi SDS^{-1} on lauseen 3.2 mukaan positiivisesti definiitti eli matriisi AB on myös. □

Lause 3.5. *Olkoon matriisi A positiivisesti definiitti ja matriisi B negatiivisesti definiitti. Tällöin niiden tulo on negatiivisesti definiitti.*

Todistus. Lauseen todistus on vastaava kuin Lauseelle 3.4, joten sitä ei ole erikseen työssä avattu. □

4. MATRIISIN NELIÖJUURI

Kvadraattisia yhtälöitä ratkaistaessa on neliöjuuren käsite usein tärkeä. Kuten reaaliluvulla, myös kääntyvällä matriisilla on neliöjuuri. Matriisin neliöjuuri voidaan määrittää eri tavoin, esimerkiksi alaluvussa 2.4 esiteltyä diagonalisointia avuksi käyttäen. Kuitenkin matriisin neliöjuurelle on olemassa 2×2-tapauksessa esityskaava, joka johdetaan tässä luvussa positiivisesti definiiteille matriiseille. Luvussa esitellään myös matriisin neliöjuuren liittyviä ominaisuuksia.

4.1 Määritelmä ja ominaisuuksia

Määritelmä 4.1. Matriisi B on neliömatriisin A *neliöjuuri*, jos ja vain jos $A = BB$, jolloin merkitään $A^{1/2} = B$.

Lause 4.2. *[5, s. 154] Positiivisesti definiitillä matriisilla on yksikäsitteinen positiivisesti definiitti neliöjuuri.*

Todistus. Todistus seuraa lähdettä [5, s. 154].

Olkoon A positiivisesti definiitti diagonalisoituva matriisi, jonka ominaisarvot ovat $\lambda_1, \lambda_2, \dots, \lambda_n$. Tällöin se voidaan esittää Lauseen 2.17 mukaan muodossa $A = SDS^{-1}$, missä D on matriisin A ominaisarvoista koostuva diagonaalimatriisi ja S koostuu näitä vastaavista ominaisvektoreista. Kun otetaan matriisista A neliöjuuri, saadaan

$$A^{\frac{1}{2}} = SD^{\frac{1}{2}}S^{-1} = S \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}^{\frac{1}{2}} S^{-1}.$$

Laskemalla tulo

$$A^{\frac{1}{2}}A^{\frac{1}{2}} = SD^{\frac{1}{2}}S^{-1}SD^{\frac{1}{2}}S^{-1} = SDS^{-1} = A,$$

huomataan, että matriisi $A^{\frac{1}{2}}$ on matriisin A neliöjuuri. Diagonaalimatriisin D neliöjuuri on

$$D^{\frac{1}{2}} = \begin{pmatrix} \pm\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \pm\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm\sqrt{\lambda_n} \end{pmatrix}.$$

Koska matriisi A on positiivisesti definiitti matriisi, Lauseen 3.2 mukaan sen kaikki ominaisarvot λ_i ovat positiivisia ja näin ollen saadaan matriisin A neliöjuureksi reaalinen matriisi

$$A^{\frac{1}{2}} = S \begin{pmatrix} \pm\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \pm\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm\sqrt{\lambda_n} \end{pmatrix} S^{-1}.$$

Huomataan, että $A^{\frac{1}{2}}$ on positiivisesti definiitti vain silloin, kun kaikki matriisin D diagonaalialkiot ovat positiivisia eli positiivisesti definiitillä matriisilla on olemassa täsmälleen yksi positiivisesti definiitti neliöjuuri. □

Kuten reaaliluvun neliöjuuri, myös reaalisen matriisin neliöjuuri voi olla kompleksinen.

Lause 4.3. *Jos A on positiivisesti definiitti matriisi, jonka neliöjuuri on $A^{\frac{1}{2}}$, on negatiivisesti definiitin matriisin $-A$ neliöjuuri $iA^{\frac{1}{2}}$.*

Todistus. Jos matriisi A on positiivisesti definiitti,

$$A = A^{\frac{1}{2}}A^{\frac{1}{2}}.$$

Näin ollen negatiivisesti definiitti matriisi $-A$ voidaan esittää muodossa

$$-A = i^2A = (iA^{\frac{1}{2}})(iA^{\frac{1}{2}}),$$

mistä huomataan, että $(-A)^{\frac{1}{2}} = iA^{\frac{1}{2}}$. □

4.2 Cayley-Hamilton-lause

Merkitään k:nnen asteen polynomia

$$q(t) = b_kt^k + b_{k-1}t^{k-1} + \cdots + b_1t + b_0.$$

Yhtälöön on mahdollista sijoittaa neliömatriisi muuttujan t paikalle. Matriiseille lasketaan tällöin potenssit vastaavasti kuin matriisien laskusääntöjen mukaan. Seuraavaa lausetta tarvitaan luvun lopussa esiteltävän esityskaavan johtamisessa.

Lause 4.4. *[4, s. 540] Olkoon $q(t)$ $k-$asteinen polynomi ja olkoon $A \in \mathbb{R}^{n\times n}$ matriisi, joka toteuttaa yhtälön $A\mathbf{x} = \lambda\mathbf{x}$, kun $\mathbf{x}$ on nollasta poikkeava vektori. Tällöin*

$$q(A)\mathbf{x} = (b_k A^k + b_{k-1}A^{k-1} + \cdots + b_1 A + b_0 I)\mathbf{x} = q(\lambda)\mathbf{x}$$

eli $q(\lambda)$ on matriisin $q(A)$ ominaisarvo.

Todistus. Todistus seuraa lähteen [4, s. 541] todistusta.

Sijoittamalla polynomiin $q(t)$ muuttujan t paikalle matriisi $A \in \mathbb{R}^{n\times n}$ ja kertomalla polynomi vektorilla $\mathbf{x} \neq \mathbf{0}$, saadaan

$$\begin{aligned} q(A)\mathbf{x} &= (b_k A^k + b_{k-1}A^{k-1} + \cdots + b_1 A + b_0)\mathbf{x} \\ &= b_k A^k\mathbf{x} + b_{k-1}A^{k-1}\mathbf{x} + \cdots + b_1 A\mathbf{x} + b_0\mathbf{x}. \end{aligned}$$

Koska $A\mathbf{x} = \lambda\mathbf{x}$ ja tiedetään, että $A^i\mathbf{x} = \lambda^i\mathbf{x}$, kun $i = 1, 2, \ldots$, polynomi saadaan muotoon

$$\begin{aligned} q(A)\mathbf{x} &= b_k\lambda^k\mathbf{x} + b_{k-1}\lambda^{k-1}\mathbf{x} + \cdots + b_1\lambda\mathbf{x} + b_0\mathbf{x} \\ &= (b_k\lambda^k + b_{k-1}\lambda^{k-1} + \cdots + b_1\lambda + b_0)\mathbf{x} \\ &= q(\lambda)\mathbf{x}. \end{aligned}$$

Näin ollen $q(A)\mathbf{x} = q(\lambda)\mathbf{x}$. □

4.3 Positiivisesti definiitin 2×2-matriisin neliöjuuri

Kaikkien matriisien neliöjuuren selvittämiseen ei tarvita diagonalisointia. Hyödyntäen edellä esiteltyä Cayley-Hamilton-lausetta, voidaan positiivisesti definiitin 2×2-matriisin neliöjuurelle johtaa esityskaava lähteen [6] mukaisesti.

Olkoon $A \in \mathbb{R}^{2\times 2}$ positiivisesti definiitti matriisi, joka voidaan kirjoittaa yleisesti muotoon

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

ja olkoon matriisi $C \in \mathbb{R}^{2\times 2}$ matriisin A neliöjuuri eli

$$A = C^2. \tag{4.1}$$

Määritelmän 2.14 mukaan tämän neliöjuurimatriisin karakteristinen polynomi on

$$p_C(\lambda) = \lambda^2 - \mathrm{tr}(C)\lambda + \det(C),$$

jossa λ on matriisin C ominaisarvo. Cayley-Hamilton-lauseen mukaan polynomiin voidaan sijoittaa matriisi C, jolloin

$$p_C(C) = C^2 - \mathrm{tr}(C)C + \det(C)I.$$

Kun ratkaistaan tämän funktion nollakohdat, saadaan yhtälö

$$C^2 - \mathrm{tr}(C)C + \det(C)I = 0,$$

josta edelleen

$$A - \mathrm{tr}(C)C + \det(C)I = 0.$$

Lauseen 2.6 perusteella tiedetään, että $\det(C)^2 = \det(A)$, jolloin $\det(C) = \pm\sqrt{\det(A)}$. Yhtälö saadaan tällöin muotoon

$$\mathrm{tr}(C)C = A \pm \sqrt{\det(A)}I.$$

Merkitään $\tau = \mathrm{tr}(C)$. Koska matriisi A on positiivisesti definiitti, $\tau \neq 0$. Tällöin jokainen neliöjuuri on muotoa

$$C = \frac{1}{\tau}(A \pm \sqrt{\det(A)}I). \tag{4.2}$$

Sijoitetaan tämä yhtälöön (4.1) ja saadaan

$$\begin{aligned} & A^2 \pm 2\sqrt{\det(A)}A + (\det(A))I = A\tau^2 \\ \Leftrightarrow\ \ & A^2 \pm (2\sqrt{\det(A)} - \tau^2)A + (\det(A))I = 0. \end{aligned}$$

Matriisin A karakteristisen polynomin $p_A(A)$ nollakohdan perusteella $A^2 = \mathrm{tr}(A)A - \det(A)I$, joten

$$\begin{aligned} (\mathrm{tr}(A)A - \det(A)I) \pm (2\sqrt{\det(A)} - \tau^2)A + (\det(A))I = 0 \\ \Leftrightarrow\ \ (\mathrm{tr}(A) \pm 2\sqrt{\det(A)} - \tau^2)A = 0. \end{aligned}$$

Koska A ei ole nollamatriisi, voidaan todeta, että

$$\tau = \mathrm{tr}(C) = \pm\sqrt{\mathrm{tr}(A) \pm 2\sqrt{\det(A)}}.$$

Sijoittamalla puolestaan tämä yhtälöön (4.2), saadaan neliöjuurimatriisiksi

$$C = \pm\frac{A \pm \sqrt{\det(A)}I}{\sqrt{\mathrm{tr}(A) \pm 2\sqrt{\det(A)}}}.$$

Kirjoitetaan selkeyden vuoksi, että $s = \pm\sqrt{\det(A)}$, jolloin edellinen yhtälö voidaan yksinkertaistaa muotoon

$$C = \frac{1}{\tau}(A + sI). \tag{4.3}$$

Esityskaavasta huomataan, että neliöjuurelle voi olla kaksi tai neljä eri ratkaisua positiivisesti definiitin 2×2-matriisin tapauksessa, joista yksi on positiivisesti definiitti Lauseen 4.2 mukaan.

Esimerkki 4.5. Määritetään matriisin

$$A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \tag{4.4}$$

neliöjuuret käyttäen edeltävää menetelmää, kun $a \in \mathbb{R}$. Aloitetaan laskemalla matriisin jälki ja determinantti, joiden arvoiksi saadaan $\mathrm{tr}(A) = 2$ ja $\det(A) = 1$. Tällöin saataisiin $s = \pm 1$ ja $\tau = \pm\sqrt{2 \pm 2}$, mutta huomataan, jos $s = -1$, niin $\tau = 0$, jolloin kaavaa 4.3 ei ole määritelty. Täten valitaan $s = 1$, jolloin $\tau = \pm 2$. Matriisin A neliöjuureksi saadaan

$$C = \frac{1}{\pm 2} \begin{pmatrix} 1+1 & a \\ 0 & 1+1 \end{pmatrix} = \pm \begin{pmatrix} 1 & \frac{a}{2} \\ 0 & 1 \end{pmatrix},$$

joista

$$C = \begin{pmatrix} 1 & \frac{a}{2} \\ 0 & 1 \end{pmatrix}$$

on positiivisesti definiitti.

5. MATRIISIYHTÄLÖ $XAX = B$

Reaalilukujen tapauksessa yhtälö

$$xax = b$$

on helppo ratkaista, mutta parametrien a ja b vaikutusta ratkaisuun ei aina tule ajateltua. Vaikka yhtälö on triviaali, se on mielenkiintoinen ja sen ratkaiseminen voidaan jakaa seuraaviin tapauksiin:

(a) Kun $a = 0$ ja $b \neq 0$, yhtälöllä ei ole ratkaisua.

(b) Kun $a = 0$ ja $b = 0$, jokainen $x \in \mathbb{R}$ toteuttaa yhtälön.

(c) Kun $a \neq 0$ ja $b = 0$, yhtälöllä on yksikäsitteinen ratkaisu $x = 0$.

(d) Kun $a \neq 0$ ja $b \neq 0$, ratkaisut riippuvat vakioiden merkeistä seuraavasti:

  (i) Kun a ja b ovat samanmerkkiset, ratkaisu on $x = \pm\sqrt{\frac{b}{a}}$.

  (ii) Kun a ja b ovat erimerkkiset, ratkaisu on $x = \pm i\sqrt{\frac{b}{a}}$.

Matriisiyhtälön ratkaisun ominaisuuksia voidaan päätellä samaan tapaan, mutta nollasta poikkeavien matriisien A ja B tapauksessa ratkaisun määrittäminen on monimutkaisempaa.

5.1 Ratkaisun ominaisuuksia yleisessä tapauksessa

Perehdytään tässä alaluvussa, mitä matriisiyhtälöstä

$$XAX = B \tag{5.1}$$

ja sen ratkaisujen ominaisuuksista tiedetään yleisesti, kun $A, B \in \mathbb{C}^{n\times n}$. Tutkitaan aluksi determinanttien arvoja soveltamalla Lausetta 2.6 yhtälöön ja saadaan

$$\det(XAX) = \det(X)\det(A)\det(X),$$

mistä edelleen

$$\det(X)^2\det(A) = \det(B). \tag{5.2}$$

Tämän perusteella voidaan päätellä, mihin matriisiavaruuteen matriisi X kuuluu.

Lause 5.1. *Olkoon $XAX = B$, jossa $A, B \in \mathbb{C}^{n\times n}$.*

(a) Jos $\det(A)$ *ja* $\det(B)$ *ovat samanmerkkiset,* $\det(X) \in \mathbb{R}$. *Tämä ei kuitenkaan välttämättä tarkoita, että $X \in \mathbb{R}^{n\times n}$.*

(b) Jos $\det(A)$ *ja $det(B)$ ovat erimerkkiset, $det(X) \in \mathbb{C}$, jolloin $X \in \mathbb{C}^{n\times n}$.*

Todistus. Todistuksen triviaalisuuden vuoksi sitä ei työssä erikseen esitetä. □

Lauseen 2.9 avulla yhtälöstä (5.2) voidaan päätellä myös matriisin X kääntyvyys eri tilanteissa.

Lause 5.2. *Olkoon $XAX = B$, jossa $A, B \in \mathbb{C}^{n\times n}$.*

(a) Jos A ja B ovat kääntyviä, on myös matriisin X oltava kääntyvä.

(b) Jos A on kääntyvä ja B ei ole, X ei ole kääntyvä.

(c) Jos A ei ole kääntyvä ja B on kääntyvä, yhtälölle ei ole ratkaisua.

(d) Jos A eikä B ole kääntyviä, matriisin X kääntyvyyttä ei voida päätellä.

Todistus. Todistuksen triviaalisuuden vuoksi sitä ei työssä erikseen esitetä. □

Matriisin X kääntyvyyttä voidaan tutkia myös matriisien asteiden avulla Lausetta 2.23 käyttäen.

Esimerkki 5.3. Osoitetaan Lauseen 5.2 kohta (c) tutkimalla yhtälössä esiintyvien matriisien asteita. Olkoon matriisi $A \in \mathbb{C}^{n\times n}$ singulaarinen ja $B \in \mathbb{C}^{n\times n}$ kääntyvä. Tehdään vastaoletus, että yhtälöllä on ratkaisu, jolloin Lauseen 2.23 mukaan

$$\operatorname{rank}(B) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(X)\}.$$

Koska matriisi A ei ole kääntyvä, $\operatorname{rank}(A) = k < n$ ja koska B on kääntyvä $\operatorname{rank}(B) = n$. Saadaan

$$\operatorname{rank}(B) \leq \min\{k, \operatorname{rank}(X)\} < n.$$

Saadaan ristiriita, jonka perusteella voidaan todeta, ettei yhtälöllä ole ratkaisua, jos matriisi A ei ole kääntyvä matriisin B ollessa.

Eri matriisiavaruuksien lisäksi yhtälön $XAX = B$ ratkaisut voivat saada eri muotoja. Usein saadut ratkaisut ovat joko täysin reaalisia tai imaginäärisiä kuten skalaarien tapauksessa. Matriisiyhtälölle on kuitenkin mahdollista, että yhtälöllä on olemassa ratkaisu, joka on muotoa

$$X = Y + iZ,$$

kun $Y, Z \in \mathbb{R}^{n\times n}$.

Esimerkki 5.4. Olkoot
$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ ja } B = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}.$$
Tällöin eräs ratkaisu yhtälöön (5.1) on
$$X = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

5.2 Ratkaisu reaalisten definiittien $n \times n$-matriisien A ja B tapauksessa

Tässä alaluvussa yhtälöä $XAX = B$ ratkaistaan reaalisten definiittien matriisien tapauksessa, jolloin tilanne on laskujen kannalta monimutkaisempi, mutta analoginen edellä käydyn erikoistapauksen $n = 1$ kanssa. Aloitetaan yhtälöstä
$$XAX = B, \tag{5.3}$$
kun $A, B \in \mathbb{R}^{n \times n}$. Symmetrisellä matriisilla on aina neliöjuuri, joten voimme sijoittaa $A = CC$ yhtälöön 5.3 ja saadaan
$$XCCX = B.$$
Kertomalla tämä puolittain neliöjuurimatriisilla C sekä vasemmalta että oikealta, saadaan
$$CXCCXC = CBC.$$
Huomataan, että yhtälön vasemmalla puolella on matriisin CXC neliö eli
$$(CXC)^2 = CBC.$$
Kuten skalaarien tapauksessa yhtälön ratkaisu riippui vakioiden merkeistä, riippuu ratkaisu nyt vakiomatriisien A ja B definiittisyyksistä. Oletetaan ensin matriisien A ja B olevan positiivisesti definiittejä. Kun matriisi A positiivisesti definiitti, voidaan aina valita positiivisesti definiitti neliöjuurimatriisi C Lauseen 4.2 perusteella. Lauseen 3.4 avulla huomataan puolestaan, että matriisitulo CBC on positiivisesti definiitti. Näin ollen neliöjuuri voidaan ottaa suoraan tulosta ja saadaan
$$CXC = \pm (CBC)^{\frac{1}{2}}.$$
Tämä puolestaan voidaan kertoa puolittain vasemmalta ja oikealta matriisilla C^{-1}, jolloin
$$X = \pm C^{-1}(CBC)^{\frac{1}{2}}C^{-1},$$
jossa $C = A^{\frac{1}{2}}$.

Tilanne on lähes vastaava, jos vakiomatriisit A ja B ovat negatiivisesti definiittejä. Tällöin kertomalla yhtälö 5.3 puolittain luvulla -1, saadaan yhtälö muotoon

$$X(-A)X = -B,$$

jossa matriisit $-A$ ja $-B$ ovat positiivisesti definiittejä. Näin ollen saadaan lähes sama ratkaisukaava matriisille X kuin positiivisesti definiittien matriisien tapauksessa eli

$$X = \pm C^{-1}(C(-B)C)^{\frac{1}{2}}C^{-1},$$

jossa $C = (-A)^{\frac{1}{2}}$.

Kuten skalaarienkin tapauksessa, yhtälön ratkaisu on kompleksinen, kun matriisit A ja B ovat "erimerkkisiä" eli toinen on positiivisesti ja toinen negatiivisesti definiitti. Oletetaan seuraavaksi, että matriisi A on positiivisesti ja matriisi B negatiivisesti definiitti. Koska matriisi B on negatiivisesti definiitti ja C on valittu positiivisesti definiitiksi, Lauseen 3.5 perusteella myös tulomatriisi CBC on negatiivisesti definiitti. Lauseen 4.3 mukaan sen neliöjuuri on kompleksinen

$$(CBC)^{\frac{1}{2}} = (-C(-B)C)^{\frac{1}{2}} = i(C(-B)C)^{\frac{1}{2}},$$

jossa $C(-B)C$ on positiivisesti definiitti matriisi. Näin ollen ratkaisukaavaksi tässä tapauksessa saadaan

$$X = \pm iC^{-1}(C(-B)C)^{\frac{1}{2}}C^{-1},$$

jossa $C = A^{\frac{1}{2}}$.

Jos matriisi A on negatiivisesti ja matriisi B positiivisesti definiitti, yhtälö (5.1) voidaan kertoa puolittain luvulla -1, jolloin matriisi $-A$ on positiivisesti definiitti. Tällöin matriisi $-B$ on negatiivisesti definiitti, jolloin tulomatriisi $C(-B)C$ on negatiivisesti definiitti. Saadaan

$$(C(-B)C)^{\frac{1}{2}} = (-CBC)^{\frac{1}{2}} = i(CBC)^{\frac{1}{2}}$$

ja ratkaisukaavaksi edelleen

$$X = \pm iC^{-1}(CBC)^{\frac{1}{2}}C^{-1},$$

jossa $C = (-A)^{\frac{1}{2}}$.

Kootaan ratkaisukaavat yhteen lauseeseen.

Lause 5.5. *Olkoot $A, B \in \mathbb{R}^{n\times n}$. Yhtälön $XAX = B$ ratkaisuja ovat*

(a)
$$X = \pm C^{-1}(CBC)^{\frac{1}{2}}C^{-1}, \tag{5.4}$$
jossa $C = A^{\frac{1}{2}}$, kun A ja B ovat positiivisesti definiittejä

(b)
$$X = \pm C^{-1}(C(-B)C)^{\frac{1}{2}}C^{-1}, \tag{5.5}$$
jossa $C = (-A)^{\frac{1}{2}}$, kun A ja B ovat negatiivisesti definiittejä

(c)
$$X = \pm iC^{-1}(C(-B)C)^{\frac{1}{2}}C^{-1}, \tag{5.6}$$
jossa $C = A^{\frac{1}{2}}$, kun A on positiivisesti definiitti ja B on negatiivisesti definiitti

(d)
$$X = \pm iC^{-1}(CBC)^{\frac{1}{2}}C^{-1}, \tag{5.7}$$
jossa $C = (-A)^{\frac{1}{2}}$, jos A on negatiivisesti definiitti ja B on positiivisesti definiitti.

5.3 Esimerkkejä ratkaisukaavan käytöstä 2×2-matriiseilla

Esimerkki 5.6. Ratkaistaan yhtälö $XAX = B$, kun

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \text{ ja } B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

eli molempien vakiomatriisien ollessa positiivisesti definiittejä. Määritetään aluksi matriisin A neliöjuuri. Aloitetaan laskemalla matriisin jälki ja determinantti, joiden arvoiksi saadaan $\mathrm{tr}(A) = 6$ ja $\det(A) = 9$. Tällöin saadaan $s = \pm 3$ ja $\tau = \pm\sqrt{6 \pm 6}$. Käyttäen kaavaa 4.3 saadaan

$$C = \pm \begin{pmatrix} \frac{3\pm 3}{\sqrt{6\pm 6}} & 0 \\ 0 & \frac{3\pm 3}{\sqrt{6\pm 6}} \end{pmatrix}.$$

Nyt kuitenkin huomataan, että vakion s negatiivista arvoa ei voida käyttää, joten saadaan

$$C = \pm \begin{pmatrix} \frac{6}{\sqrt{12}} & 0 \\ 0 & \frac{6}{\sqrt{12}} \end{pmatrix} = \pm \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{3} \end{pmatrix}.$$

Matriisilla A on täten kaksi neliöjuurta, joista voidaan valita positiivisesti definiitti neliöjuuri eli

$$C = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{3} \end{pmatrix}.$$

Tämän neliöjuurimatriisin käänteismatriisiksi saadaan lauseen 2.11 avulla

$$C^{-1} = \frac{1}{3}\begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{3} \end{pmatrix}.$$

Lasketaan sitten positiivisesti definiitin matriisin CBC neliöjuuret Lauseen 4.3 avulla, josta saadaan

$$(CBC)^{\frac{1}{2}} = \pm\frac{1}{\sqrt{15}}\begin{pmatrix} 9 & 3 \\ 3 & 6 \end{pmatrix} \text{ ja } (CBC)^{\frac{1}{2}} = \pm\begin{pmatrix} \sqrt{3} & \sqrt{3} \\ \sqrt{3} & 0 \end{pmatrix}.$$

Nyt sijoittamalla kaikki mahdolliset arvot ratkaisukaavaan 5.4 saadaan ratkaisuiksi

$$X = \pm\frac{1}{\sqrt{15}}\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \text{ ja } X = \pm\frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Yhtälölle on siis neljä eri ratkaisua, jotka kaikki toteuttavat alkuperäisen yhtälön.

Esimerkki 5.7. Ratkaistaan yhtälö nyt, kun

$$A = \begin{pmatrix} -4 & 0 \\ 0 & -1 \end{pmatrix} \text{ ja } B = \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}$$

eli molemmat vakiomatriisit ovat negatiivisesti definiittejä. Saadaan siis yhtälö

$$X\begin{pmatrix} -4 & 0 \\ 0 & -1 \end{pmatrix}X = \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}.$$

Kerrotaan yhtälö puolittain luvulla -1, jolloin matriisit $-A$ ja $-B$ ovat positiivisesti definiittejä. Lasketaan matriisin $-A$ neliöjuuri C ja saadaan positiivisesti definiitiksi ratkaisuksi

$$C = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},$$

jonka käänteismatriisi on

$$C^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Lasketaan seuraavaksi matriisitulon $C(-B)C$ neliöjuuri, josta saadaan neljä ratkaisua

$$(C(-B)C)^{\frac{1}{2}} = \pm \frac{1}{\sqrt{7+4\sqrt{3}}} \begin{pmatrix} 4+2\sqrt{3} & 0 \\ 0 & 3+2\sqrt{3} \end{pmatrix} \text{ ja}$$

$$(C(-B)C)^{\frac{1}{2}} = \pm \frac{1}{\sqrt{7-4\sqrt{3}}} \begin{pmatrix} 4-2\sqrt{3} & 0 \\ 0 & 3-2\sqrt{3} \end{pmatrix}.$$

Lopulta voidaan sijoittaa arvot ratkaisukaavaan (5.5) ja muuttujamatriisille X saadaan mahdolliset arvot

$$X = \pm \frac{1}{4\sqrt{7+4\sqrt{3}}} \begin{pmatrix} 4+2\sqrt{3} & 0 \\ 0 & 12+8\sqrt{3} \end{pmatrix} \text{ ja}$$

$$X = \pm \frac{1}{4\sqrt{7-4\sqrt{3}}} \begin{pmatrix} 4-2\sqrt{3} & 0 \\ 0 & 12-8\sqrt{3} \end{pmatrix}.$$

Nämä sievenevät vielä muotoon

$$X = \pm \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \pm\sqrt{3} \end{pmatrix}.$$

Sijoittamalla nämä alkuperäiseen yhtälöön huomataan näiden toteuttavan sen.

Esimerkki 5.8. Ratkaistaan yhtälö $XAX = B$, kun

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \text{ ja } B = \begin{pmatrix} -2 & -1 \\ -1 & -1 \end{pmatrix}$$

Huomataan, että matriisi A on positiivisesti definiitti ja matriisi B on negatiivisesti definiitti. Edellisen esimerkin perusteella matriisin A neliöjuuri

$$C = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

ja sen käänteismatriisi on puolestaan

$$C^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Nyt voidaan laskea tulon $C(-B)C$ neliöjuuri, jolle saadaan neljä mahdollista arvoa

$$(C(-B)C)^{\frac{1}{2}} = \pm \frac{1}{\sqrt{13}} \begin{pmatrix} 10 & 2 \\ 2 & 3 \end{pmatrix} \text{ ja}$$

$$(C(-B)C)^{\frac{1}{2}} = \pm \frac{1}{\sqrt{5}} \begin{pmatrix} 6 & 2 \\ 2 & -1 \end{pmatrix}.$$

Sijoitetaan saadut arvot ratkaisukaavaan (5.6) ja saadaan ratkaisuiksi

$$X = \pm i \frac{1}{4\sqrt{13}} \begin{pmatrix} 10 & 4 \\ 4 & 12 \end{pmatrix} \text{ ja}$$

$$X = \pm i \frac{1}{4\sqrt{5}} \begin{pmatrix} 6 & 4 \\ 4 & -4 \end{pmatrix}.$$

6. YHTEENVETO

Tässä työssä tutkittiin yhtälön $XAX = B$ ratkaisemista. Työn päätulokset olivat yhtälön $XAX = B$ ratkaisukaavat tilanteessa, jossa matriisit A ja B ovat positiivisesti tai negatiivisesti definiittejä. Ratkaisu riippuu matriisien A ja B definiittisyyksistä, joten ratkaisukaavoja saatiin neljä ja ne esitettiin Lauseessa 5.5. Ratkaisukaavojen käytöstä esitettiin myös konkreettisia esimerkkejä alaluvussa 5.3. Yleisessä tapauksessa yhtälön ratkaisun kannalta tärkeitä tuloksia olivat myös Lauseet 5.2 ja 5.1, joiden avulla voidaan päätellä onko ratkaisua olemassa ja ominaisuuksia yhtälön ratkaisevista matriiseista.

Työssä esitettiin yhtälön ratkaisemiseen tarvittavia tietoja matriisin perusominaisuuksista, kuten käänteismatriiseista ja ominaisarvoista. Oleellista työn kannalta olivat kolmannessa luvussa esitetty matriisien definiittisyyksien määritelmät, joista työn päätuloksen kannalta erityisesti positiivisesti ja negatiivisesti definiittien matriisien määritelmät olivat tärkeitä. Lisäksi alaluvussa 5.3 tarvittiin 2×2-matriisin neliöjuurta, jolle johdettiin esityskaava alaluvussa 4.3.

Työssä saatua päätulosta ei voi yleistää kaikille matriiseille, sillä ratkaisukaavan johtamisessa tutkittiin vain reaalisia positiivisesti ja negatiivisesti definiittejä $n \times n$-matriiseja. Tämä on melko rajattu tarkastelualue, joten työn tarkastelunäkökulmaa voisi seuraavaksi laajentaa koskemaan myös semidefiniittejä tapauksia. Ratkaisukaavassa tutkittavat matriisit oletettiin reaalisiksi, joten yhtälöä voitaisiin tutkia enemmän myös kompleksisessa matriisiavaruudessa.

Suurin osa työn todistuksista ja lauseista on tehty lähteeseen viitaten. Kuitenkin osa työssä laadituista lauseista ja todistuksista ovat työn ohjaajan kanssa itse laadittuja. Näissä lauseissa ja todistuksissa ei täten ole lähdeviittauksia. Osaa lauseista ei työssä myöskään todistettu niiden triviaalisuuden tai työn laajuuden rajaamisen vuoksi.

LÄHTEET

[1] A. Ferrante ja L. Ntogramatzidis. The generalised discrete algebraic Riccati equation in linear-quadratic optimal control. eng. *Automatica (Oxford)* 49.2 (2013), 471–478. ISSN: 0005-1098. URL: `https://doi.org/10.1016/j.sysconle.2014.11.001`.

[2] R. A. Horn ja C. R. Johnson. *Matrix Analysis.* Vol. 2nd ed. Cambridge eText, 2012. ISBN: 9780521839402.

[3] A. Jeffrey. *Matrix Operations for Engineers and Scientists An Essential Guide in Linear Algebra*. eng. 1st ed. 2010. Dordrecht: Springer Netherlands, 2010. ISBN: 1-280-38230-9.

[4] L. W. Johnson, D. R. Riess ja J. T. Arnold. *Introduction to linear algebra*. 5th ed. Pearson Education, Inc, 2002.

[5] M. Pease. *Methods of Matrix Algebra*. New York : Academic Press, 1965.

[6] D. Sullivan. The Square Roots of 2 × 2 Matrices. *Mathematics Magazine* 66.5 (1993), 314–316. ISSN: 0025570X, 19300980. URL: `https://doi.org/10.2307/2690509`.