Tampereen yliopisto

Jaakko Tervonen

HADAMARDIN TULO JA SEN KÄYTTÖ DIGITAALISESSA KUVANKÄSITTELYSSÄ

Diplomityö
Tekniikan ja luonnontieteiden tiedekunta
Tarkastajat: Yliopisto-opettaja Mika Mattila
Yliopistonlehtori Pentti Haukkanen
Joulukuu 2022

TIIVISTELMÄ

Jaakko Tervonen: Hadamardin tulo ja sen käyttö digitaalisessa kuvankäsittelyssä
Diplomityö
Tampereen yliopisto
Teknis-luonnontieteellinen DI-ohjelma
Joulukuu 2022

Perinteisen matriisitulon ohella matriiseille voidaan määritellä myös muita tuloja, kuten Hadamardin tulo, joka eroaa perinteisestä matriisitulosta muun muassa yksinkertaisuutensa ja vaihdannaisuutensa ansiosta. Hadamardin tulo on määritelty kahdelle samankokoiselle matriisille. Hadamardin tulossa tulontekijöiden (matriisien) alkiot kerrotaan keskenään. Hadamardin tulo muistuttaa siis matriisien yhteenlaskua, joka tehdään samaan tapaan alkioittain. Hadamardin tulossa tulokseksi saadaan tulontekijöiden kanssa samankokoinen matriisi.

Tässä diplomityössä tutkittiin Hadamardin tuloa ja sen käyttöä digitaalisessa kuvankäsittelyssä. Työn on tarkoitus olla ymmärrettävissä ilman ulkoisia lähteitä. Työn tärkeimpiä tuloksia ovat Schurin lauseen ja Hadamardin tulon determinanttiepäyhtälön todistus. Schurin lause on keskeinen matriisianalyysissa ja sen avulla voidaan todistaa monia muita ominaisuuksia. Tässä työssä Schurin lausetta käytettiin Hadamardin tulon determinanttiepäyhtälön todistamiseen. Lisäksi todistuksia varten tarvitaan lukuisia erilaisia lauseita ja määritelmiä, jotka esitetään työn alkupuolella.

Digitaaliset kuvat voidaan esittää matriiseina (mustavalkokuvat) tai pinona matriiseja (RGB-kuvat). Matriisilaskennan avulla digitaalisille kuville voidaan tehdä monenlaisia muokkauksia, kuten säätää kuvan kirkkautta tai kontrastia. Tässä työssä on käsitellään erityisesti digitaalisten kuvien konvoluutiota, jonka avulla kuvia voidaan esimerkiksi sumentaa tai terävöittää. Yksi sovellusluvun merkittävimpiä tuloksia on digitaalisen kuvan alueellinen sumentaminen, joka toteutettiin hyödyntämällä digitaalisten kuvien konvoluutiota ja Hadamardin tuloa.

Avainsanat: Hadamardin tulo, Schurin lause, positiivisesti semidefiniitti matriisi, digitaalinen kuvankäsittely, konvoluutio

ABSTRACT

Jaakko Tervonen: The Hadamard product and its use in digital image processing
Master's thesis
Tampere University
Master's Degree Programme in Science and Engineering
December 2022

In addition to the traditional matrix product, other products can also be defined for matrices, such as the Hadamard product, which differs from the usual matrix product due to, among other things, its simplicity and its commutative property. The Hadamard product is defined for two matrices of the same size. In the Hadamard product, the elements of the multiplication factors (matrices) are multiplied together. The Hadamard product is thus similar to the matrix addition, which is done in the same way elementwise. Hadamard product results in a matrix of the same size as the multiplication factors.

This master's thesis studies the Hadamard product and its use in digital image processing. The thesis is meant to be comprehensible without external sources. The most important results of the thesis are the proof of Schur product theorem and the determinant inequality of the Hadamard product. Schur product theorem is fundamental theorem in matrix analysis and it can be used to prove many other properties. In this thesis, Schur product theorem was used to prove the determinant inequality of the Hadamard product. In addition, numerous different theorems and definitions are needed for the proofs, which are presented at the beginning of the thesis.

Digital images can be represented as matrices (grayscale images) or as a stack of matrices (RGB images). Using matrix calculation, digital images can be edited in many ways, such as adjusting the brightness or contrast of the image. This thesis deals in particular with the convolution of digital images, which can be used to blur or sharpen images, for example. One of the most significant results of the application chapter is the regional blurring of the digital image, which was implemented by utilizing the convolution of digital images and the Hadamard product.

Keywords: Hadamard product, Schur product theorem, positive semidefinite matrix, digital image processing, convolution

ALKUSANAT

Pitkä ja uuvuttava aiheenvalintaprosessi loppui syksyllä 2022 saatuani aiheen Mika Mattilalta. Pienen alkukankeuden jälkeen työ lähti edistymään varsin mallikkaasti. Aihe osoittautui juuri sopivaksi minulle ja työn kirjoittaminen oli opettavaista ja oikeastaan ihan mukavaakin puuhaa.

Kiitos Mika Mattilalle mielenkiintoisesta aiheesta sekä selkeästä, jouhevasta ja joutuisasta prosessista, ja kiitos Pentti Haukkaselle työn laatua parantavista kommenteista. Lopuksi haluan vielä kiittää perhettäni, ystäviäni ja opiskelukavereitani, jotka ovat olleet tukena ja seurana opintojeni aikana.

Tampereella, 13. joulukuuta 2022

Jaakko Tervonen

SISÄLLYSLUETTELO

KUVALUETTELO

TAULUKKOLUETTELO

OHJELMA- JA ALGORITMILUETTELO

LYHENTEET JA MERKINNÄT

$\mathbb{R}$	reaaliluvut
$\mathbb{R}^n$	reaalinen n-ulotteinen vektoriavaruus
$\mathbb{C}$	kompeksiluvut
$\mathbb{C}^n$	kompleksinen n-ulotteinen vektoriavaruus
$\mathbf{F}$	kunta
$\mathbf{F}^n$	n-ulotteinen vektoriavaruus yli kunnan $\mathbf{F}$
$M_{m,n}(\mathbf{F})$	$m \times n$-kokoisten matriisien, joiden alkiot ovat kunnan $\mathbf{F}$ skalaareja, joukko
$M_{m,n}$	$m \times n$-kokoisten kompleksisten matriisien joukko
M_n	$n \times n$-kokoisten kompleksisten neliömatriisien joukko
I_n	joukkoon M_n kuuluva identiteettimatriisi; merkitään I, jos matriisin koko selviää kontekstista
$J_{m,n}$	joukkoon $M_{m,n}$ kuuluva ykkösmatriisi; merkitään J, jos matriisin koko selviää kontekstista
$O_{m,n}$	joukkoon $M_{m,n}$ kuuluva nollamatriisi; merkitään O, jos matriisin koko selviää kontekstista
$0_{n,1}$	n-ulotteinen nollavektori; merkitään 0, jos vektorin koko selviää kontekstista
$\overline{A}$	matriisi A kompleksikonjugaatti
A^T	matriisin A transpoosi
A^*	matriisin A konjugaattitranspoosi
A^{-1}	matriisin $A \in M_n$ käänteismatriisi eli inverssi
$\det$	matriisin $A \in M_n$ determinantti
$\operatorname{tr} A$	matriisin $A = [a_{ij}] \in M_{m,n}$ jälki; $\operatorname{tr} A = \sum_i a_{ii}$
λ	matriisin $A \in M_n$ ominaisarvo
μ	matriisin $A \in M_n$ ominaisarvo
$\sigma(A)$	matriisin $A \in M_n$ spektri; matriisin $A \in M_n$ ominaisarvojen muodostama joukko

$A[\alpha, \beta]$	indeksijoukkojen α ja β määräämä matriisin $A \in M_{m,n}$ alimatriisi
$A[\alpha]$	indeksijoukon α määräämä matriisin $A \in M_{m,n}$ pääalimatriisi
$A \circ B$	matriisien $A, B \in M_{m,n}$ Hadamardin tulo
$A \otimes B$	matriisien $A \in M_{m,n}$ ja $B \in M_{p,q}$ Kroneckerin tulo
$\mathcal{N}(A)$	matriisin $A \in M_{m,n}$ nolla-avaruus
$\mathcal{R}(A)$	matriisin $A \in M_{m,n}$ sarakeavaruus
$\\|\cdot\\|$	vektorin normi; yleensä viitataan euklidiseen normiin
$\\|\cdot\\|_2$	euklidinen vektorin normi
$\langle\cdot,\cdot\rangle$	vektorin sisätulo; yleensä euklidinen sisätulo
span S	vektorijoukon S vektorien virittämä aliavaruus
$\oplus$	suora summa
$A * K$	kuvamatriisin $A \in M_{m,n}$ ja ytimen $K \in M_p$ konvoluutio

1. JOHDANTO

Tämän työn tarkoituksena on tutkia Hadamardin tuloa ja sen käyttöä digitaalisessa kuvankäsittelyssä. Digitaalinen kuvankäsittely on looginen sovelluskohde, sillä digitaaliset kuvat voidaan esittää kätevästi ja helposti ymmärrettävällä tavalla joko matriisina (mustavalkokuvat) tai pinona matriiseja (RGB-kuvat).

Perinteinen matriisitulo ei suinkaan ole ainoa tulo, joka voidaan määritellä matriiseille. Matriiseille voidaan määritellä esimerkiksi Hadamardin tulo, jossa samankoisten matriisien vastaavat alkiot kerrotaan keskenään, eli kertolasku suoritetaan alkioittain. Näin tulokseksi saadaan tulontekijöiden kanssa samankokoinen matriisi. Tavallaan Hadamardin tulo muistuttaa siis enemmän matriisien yhteenlaskua, joka suoritetaan samaan tapaan alkioittain, kuin perinteistä matriisituloa. Voidaankin sanoa, että Hadamardin tulo on yksinkeraisempi näistä kahdesta matriisitulosta. Tästä huolimatta Hadamardin tuloa harvoin edes mainitaan lineaarialgebraan liittyvissä teksteissä [8, s.88], mikä tekee sen tutkimisesta jo itsessään mielenkiintoisen. Monet Hadamardin tuloon liittyvistä ominaisuuksista ovat voimassa vain tietyntyyppisille matriiseille, kuten positiivisesti semidefiniiteille matriiseille.

Työn merkittävimpinä aikaansaannoksina voidaan pitää Schurin lauseen todistusta, Hadamardin tulon $A \circ B$ determinanttiin liittyvän epäyhtälön

$$\det A \det B \leq \det(A \circ B)$$

todistamista ja digitaalisen kuvan alueellista sumentamisen toteuttamista konvoluution ja Hadamardin tulon avulla. Todistukset näihin lauseisiin ovat olleet olemassa jo kauan, mutta suomeksi niitä ei välttämättä löydy, ainakaan kovin helposti. Tämä työ ei ole edelläkävijä myöskään digitaalisen kuvan alueellisen sumennuksen toteuttamisessa.

Luvuissa 2 ja 3 annetaan lukijalle melko kattavakin katsaus työn kannalta olennaisiin vektorien ja matriisien perusominaisuuksiin. Näiden lukujen ohittamista suositellaan, mikäli lukijalla on jo jonkin verran kokemusta matriisilaskennasta. Luvussa 4 näitä esitietoja täydennetään perehtymällä normaaleihin matriiseihin ja erityisesti unitaaristen, hermiittisten ja positiviisesti (semi)definiittien matriisien muodostamiin alaluokkiin. Luvussa esitetään monia tärkeitä lauseita, joita tullaan tarvitsemaan myöhemmin. Aiemmin mainitut luvut tarjoavat vankan pohjan Hadamardin tulon tutkimiselle, joka on tehty luvussa 5. Tässä luvussa selviää, mikä yhteys on Kroneckerin tulon ja Hadamardin tulon välillä sekä todistetaan Hadamardin tuloon liittyviä mielenkiintoisia ominaisuuksia, kuten aiemmin mainittu Schurin lause ja determinattiepäyhtälö. Luvussa 6 on esitelty digitaalista kuvankäsittelyä ja Hadamardin tulon merkitystä siinä. Ennen kaikkea on keskitytty digitaalisten kuvien konvoluutioon sumentavilla ytimillä. Liitteissä on työhön epäsuorasti liittyvien aritmeettis-geometrisen epäyhtälön ja algebran peruslauseen lisäksi MATLAB® ohjelmistolla tehdyt digitaaliseen kuvankäsittelyyn liittyvät ohjelmakoodit.

2. VEKTOREISTA

Määritelmä 2.1. Vektoriavaruus V yli kunnan $\mathbf{F}$ on epätyhjä joukko, joka on suljettu yhteenlaskun ja skalaarikertolaskun suhteen. Yhteenlaskun on oltava assosiatiivinen ja kommutatiivinen. Yhteenlaskulla on neutraalialkio (nollavektori, merkitään 0) sekä käänteisalkio joukossa. Lisäksi skalaarikertolaskun ominaisuudet $a(x+y) = ax + ay, (a+b)x = ax + bx, a(bx) = (ab)x$ sekä $ex = x$, missä $e \in \mathbf{F}$ on skalaarikertolaskun neutraalialkio, ovat voimassa kaikilla kunnan alkioilla $a, b \in \mathbf{F}$ ja kaikilla vektoreilla $x, y \in V$.

Vektoriavaruuksista erityisen kiinnostavia ovat kompleksinen vektoriavaruus $\mathbb{C}^n$ ja tähän avaruuteen sisältyvä reaalinen vektoriavaruus $\mathbb{R}^n$. Näiden yhteydessä yhteenlasku ja skalaarikertolasku määritellään alkioittain samoin kuin matriiseille (ks. määritelmä 3.2). Vektoriavaruudesta puhuttaessa viitataan oletusarvoisesti vektoriavaruuteen $\mathbb{C}^n$. Toisinaan vektoriavaruudesta puhutaan lyhemmin avaruutena.

Vektoriavaruuden määritelmästä selviää, että vektori määritellään vektoriavaruuden alkioksi. Vektoriavaruuden $\mathbf{F}^n$ vektoria voidaan kuitenkin ajatella järjestetyksi joukoksi kunnan $\mathbf{F}$ alkioita, jolloin saadaan vektorista konkreettisempi kuva. Vaihtoehtoisesti avaruuden $\mathbf{F}^n$ vektorin havainnollistaminen matriisina, jonka sarakkeiden tai rivien lukumäärä on 1, antaa helposti lähestyttävän näkökulman. Jos sarakkeita on 1, kyseessä on pystyvektori. Jos rivejä on 1, kyseessä on vaakavektori. Vektorit ovat pystyvektoreita ellei toisin mainita.

Aliavaruus on vektoriavaruuden osajoukko, joka on itsessään vektoriavaruus. Vektoriavaruudella ja sen alivaruudella on sama yhteenlasku- ja skalaarikertolaskuoperaatio. Aliavaruuksien leikkaus on aina aliavaruus. Jos S on avaruuden V osajoukko, niin span S on kaikkien joukon S sisältävien aliavaruuksien leikkaus. Jos S on epätyhjä joukko, niin

$$\text{span}\, S = \{a_1v_1 + \cdots + a_kv_k : v_1, \ldots, v_k \in S, a_1, \ldots, a_k \in \mathbf{F} \text{ ja } k = 1, 2, \ldots\}.$$

Jos S on tyhjä joukko, niin se sisältyy jokaiseen vektoriavaruuden V aliavaruuteen. Vektoriavaruuden V kaikkien aliavaruuksien leikkaus on $\{0\}$, joten määritelmä takaa, että span $S = \{0\}$. Vaikka S ei olisi aliavaruus, niin span S on aina aliavaruus. Sanotaankin, että span S on joukon S vektorien virittämä aliavaruus. Jos span $S = V$, joukko S virittää avaruuden V. [9, s.2]

Lineaarikombinaatiolla tarkoitetaan vektoreiden summaa $a_1v_1 + \cdots + a_kv_k$, missä k on positiivinen kokonaisluku, skalaarit $a_1, \ldots, a_k \in \mathbf{F}$ ja $v_1, \ldots, v_k \in V$. Summattavia vektoreita on oltava äärellinen määrä. Täten span $S \neq \emptyset$ muodostuu kaikista joukon S vektorien lineaarikombinaatioista. Vektorijoukon $v_1, \ldots, v_k \in V$ vektorit ovat lineaarisesti riippuvia, jos ja vain jos on olemassa sellaiset skalaarit $a_1, \ldots, a_k \in \mathbf{F}$, että $a_1v_1 + \cdots + a_kv_k = 0$ ja ainakin yksi skalaareista $a_1, \ldots, a_k$ ei ole 0. Vektorit ovat lineaarisesti riippumattomia, jos ne eivät ole lineaarisesti riippuvia. [9, s.2-3]

Vektoriavaruuden V kanta on sellainen lineaarisesti riippumaton joukko avaruuden V vektoreita, että vektorit virittävät avaruuden V. Tällöin jokainen avaruuden V alkio voidaan esittää kannan vektoreiden lineaarikombinaationa täsmälleen yhdellä tavalla. Jos vektorijoukkoon lisätään tai siitä poistetaan vektori tämä ei enää pidä paikkaansa. [9, s.3] Jos vektoriavaruuden V kannassa on n kappaletta vektoreita, niin tällöin jokaisessa tämän vektoriavaruuden kannassa on n kappaletta vektoreita. Tätä kannan muodostavien vektoreiden joukon mahtavuutta (lukumäärää) sanotaan vektoriavaruuden V dimensioksi ja merkitään $\dim V$. Esimerkiksi vektoriavaruuden $\mathbb{R}^n$ dimensio on n, eli $\dim \mathbb{R}^n = n$, ja vektoriavaruuden $\mathbb{C}^n$ dimensio on n yli kunnan $\mathbb{C}$, mutta dimensio on $2n$ yli kunnan $\mathbb{R}$. Vektoriavaruuden $\mathbf{F}^n$ kantaa $e_1, \ldots, e_n$, missä vektorin e_i i. alkio on 1 ja muut alkiot ovat nollia, sanotaan standardikannaksi. [9, s.4]

Vektorin normi ja vektorien sisätulo ovat funktiota, jotka määritellään aksioomien avulla. Määritelmien aksioomien avulla voidaan päätellä monia muitakin ominaisuuksia. Esimerkiksi normin määritelmän kohdista (a) ja (b) seuraa, että mikä tahansa nollasta eroava vektori x voidaan normeerata yksikkövektoriksi, eli vektoriksi, jonka normi on 1. Tämä tapahtuu asettamalla $u = \frac{x}{\|x\|}$, sillä $\|u\| = \|\frac{x}{\|x\|}\| = \frac{\|x\|}{\|x\|} = 1$. [9, s.314]

Määritelmä 2.2. Olkoon V vektoriavaruus yli kunnan $\mathbb{C}$. Funktiota $\|\cdot\| : V \to \mathbb{R}$ sanotaan normiksi, jos kaikilla vektoreilla $x, y \in V$ ja skalaareilla $c \in \mathbb{C}$ on voimassa

(a) $\|x\| \geq 0$ ja $\|x\| = 0$, jos ja vain jos $x = 0$,

(b) $\|cx\| = |c|\|x\|$,

(c) $\|x + y\| \leq \|x\| + \|y\|$.

Määritelmä 2.3. Olkoon V vektoriavaruus yli kunnan $\mathbb{C}$. Funktiota $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ sanotaan sisätuloksi, jos kaikilla vektoreilla $x, y, z \in V$ ja skalaareilla $c \in \mathbb{C}$ on voimassa

(a) $\langle x, x \rangle \geq 0$ ja $\langle x, x \rangle = 0$, jos ja vain jos $x = 0$,

(b) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,

(c) $\langle cx, y \rangle = c\langle x, y \rangle$,

(d) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

Sisätulon aksioomista voidaan johtaa hyvin tunnettu Cauhcy-Schwarzin epäyhtälö, joka on voimassa kaikilla sisätuloilla [9, s.315]. Tätä lausetta tarvitaan, kun osoitetaan, että euklidinen normi toteuttaa normin määritelmän aksioomat.

Lause 2.1 (Cauchy-Schwarzin epäyhtälö)**.** *Olkoon $\langle \cdot, \cdot \rangle$ vektoriavaruuden V sisätulo yli kunnan $\mathbb{C}$. Tällöin jokaisella vektorilla $x \in V$ ja $y \in V$ on voimassa*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle.$$

Todistus. (Vrt. [9], Theorem 5.1.4) Olkoon x ja y mitä tahansa vektoriavaruuden V vektoreita. Jos $x = y = 0$, väittämä on tosi. Täten voidaan olettaa, että $y \neq 0$. Olkoon $v = \langle y, y\rangle x - \langle x, y\rangle y$. Käyttämällä sisätulon määritelmän aksioomia (a)-(d) saadaan

$$\begin{aligned}
0 &\overset{\text{(a)}}{\leq} \langle v, v\rangle = \langle \langle y, y\rangle x - \langle x, y\rangle y, \langle y, y\rangle x - \langle x, y\rangle y\rangle \\
&\overset{\text{(b)}}{=} \langle \langle y, y\rangle x, \langle y, y\rangle x - \langle x, y\rangle y\rangle + \langle -\langle x, y\rangle y, \langle y, y\rangle x - \langle x, y\rangle y\rangle \\
&\overset{\text{(d)}}{=} \overline{\langle \langle y, y\rangle x - \langle x, y\rangle y, \langle y, y\rangle x\rangle} - \overline{\langle \langle y, y\rangle x - \langle x, y\rangle y, \langle x, y\rangle y\rangle} \\
&\overset{\text{(b)}}{=} \overline{\langle \langle y, y\rangle x, \langle y, y\rangle x\rangle} - \overline{\langle \langle x, y\rangle y, \langle y, y\rangle x\rangle} - \overline{\langle \langle y, y\rangle x, \langle x, y\rangle y\rangle} + \overline{\langle \langle x, y\rangle y, \langle x, y\rangle y\rangle} \\
&\overset{\text{(c)}}{=} \langle y, y\rangle\overline{\langle y, y\rangle}\langle x, x\rangle - \overline{\langle x, y\rangle}\langle y, y\rangle\overline{\langle y, x\rangle} - \overline{\langle y, y\rangle}\langle x, y\rangle\overline{\langle x, y\rangle} + \overline{\langle x, y\rangle}\langle x, y\rangle\overline{\langle y, y\rangle} \\
&\overset{\text{(d)}}{=} \langle y, y\rangle\langle y, y\rangle\langle x, x\rangle - \overline{\langle x, y\rangle}\langle y, y\rangle\langle x, y\rangle = \langle y, y\rangle^2\langle x, x\rangle - \langle y, y\rangle|\langle x, y\rangle|^2 \\
&= \langle y, y\rangle\left(\langle y, y\rangle\langle x, x\rangle - |\langle x, y\rangle|^2\right).
\end{aligned}$$

Koska oletuksesta $y \neq 0$ seuraa, että $\langle y, y\rangle > 0$, niin epäyhtälö voidaan kirjoittaa muodossa $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$. □

Erityisen tärkeä normi on euklidinen normi $\|\cdot\|_2$, josta käytetään myös nimityksiä 2-normi ja l_2-normi. Euklidinen normi on tunnetuin koon ja etäisyyden mitta avaruuksissa $\mathbb{R}^2$ ja $\mathbb{R}^3$ [9, s.313]. Se määritellään käyttäen euklidista sisätuloa, josta käytetään myös nimityksiä pistetulo ja skalaaritulo. Seuraavassa esimerkissä määritellään euklidinen normi ja euklidinen sisätulo sekä osoitetaan, että nämä toteuttavat määritelmien 2.2 ja 2.3 aksioomat. Euklidisesta normista käytetään jatkossa merkintää $\|\cdot\|$ ellei erikseen haluta korostaa, mikä normi on kyseessä.

Esimerkki 2.1. Olkoot $x, y, z \in \mathbb{C}^n$, ja olkoon $c \in \mathbb{C}$. Määritellään euklidinen sisätulo seuraavasti:

$$\langle x, y\rangle = y^* x = \sum_{k=1}^{n} \overline{y}_k x_k. \tag{2.1}$$

Merkintä y^*x tarkoittaa vektorin y konjugaattitranspoosin ja vektorin x välistä matriisituloa (ks. määritelmät 3.2 ja 3.3). Osoitetaan summien ja kompleksilukujen laskusääntöjä käyttämällä, että määritelmän 2.3 aksioomat toteutuvat.

(a) $\langle x, x\rangle = x^*x = \sum_{k=1}^{n} \overline{x}_k x_k = \sum_{k=1}^{n} |x_k|^2 \geq 0$, sillä $|x_k|^2 \geq 0$ jokaisella indeksillä k.

Jotta $\sum_{k=1}^{n} |x_k|^2 = 0$, on jokaisella indeksillä k oltava $|x_k|^2 = 0$, eli $x_k = 0$, eli $x = 0$.

(b) $\langle x+y, z\rangle = z^*(x+y) = \sum_{k=1}^{n} \overline{z}_k(x_k + y_k) = \sum_{k=1}^{n} \overline{z}_k x_k + \sum_{k=1}^{n} \overline{z}_k y_k$
$= z^*x + z^*y = \langle x, z\rangle + \langle y, z\rangle$.

(c) $\langle cx, y\rangle = y^*(cx) = \sum_{k=1}^{n} \overline{y}_k c x_k = c\sum_{k=1}^{n} \overline{y}_k x_k = c\langle x, y\rangle$.

(d) $\langle x, y\rangle = y^*x = \sum_{k=1}^{n} \overline{y}_k x_k = \sum_{k=1}^{n} \overline{y_k \overline{x}_k} = \overline{\sum_{k=1}^{n} \overline{x}_k y_k} = \overline{x^*y} = \overline{\langle y, x\rangle}$.

Määritellään euklidinen normi euklidisen sisätulon avulla

$$\|x\|_2 = \langle x, x\rangle^{1/2} = (x^*x)^{1/2} = \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2}. \tag{2.2}$$

Osoitetaan, että määritelmän 2.2 aksioomat toteutuvat.

(a) Aksiooma seuraa suoraan sisätulon aksioomasta (a).

(b) Tulos saadaan sisätulon aksioomista (c) ja (d) seuraavasti:

$$\|cx\|_2 = \langle cx, cx\rangle^{1/2} = (c\overline{c}\langle x, x\rangle)^{1/2} = \big(|c|^2\langle x, x\rangle\big)^{1/2} = |c|\langle x, x\rangle^{1/2} = |c|\|x\|_2.$$

(c) Hyödyntämällä Cauchy-Schawarzin epäyhtälöä 2.1 saadaan

$$\begin{aligned}\|x+y\|_2^2 &= \langle x+y, x+y\rangle = \langle x, x+y\rangle + \langle y, x+y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \overline{\langle x, y\rangle} + \langle y, y\rangle = \|x\|_2^2 + 2\mathrm{Re}(\langle x, y\rangle) + \|y\|_2^2 \\ &\leq \|x\|_2^2 + 2|\langle x, y\rangle| + \|y\|_2^2 \leq \|x\|_2^2 + 2(\langle x, x\rangle\langle y, y\rangle)^{1/2} + \|y\|_2^2 \\ &= \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2.\end{aligned}$$

Näin ollen $\|x+y\|_2^2 \leq (\|x\|_2 + \|y\|_2)^2$, joten $\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$.

Kaksi vektoria $x, y, \in \mathbb{C}^n$ ovat ortogonaaliset, jos $\langle x, y\rangle = 0$. Geometrinen tulkinta ortogonaalisuudelle avaruuksissa $\mathbb{R}^2$ ja $\mathbb{R}^3$ on vektorien kohtisuoruus. Vektorijoukko $x_1, \ldots, x_k \in \mathbb{C}^n$ on ortogonaalinen, jos $\langle x_i, x_j\rangle = 0$ aina, kun $i \neq j$, ja $i, j \in \{1, \ldots, k\}$. Ortogonaalinen vektorijoukko nollasta eroavia vektoreita on lineaarisesti riippumaton. Lisäksi jos vektorijoukon jokaisen vektorin euklidinen normi on 1, eli $x_i^*x_i = 1$ jokaisella $i = 1, \ldots, k$, niin sanotaan, että vektorijoukko on ortonormaali. [9, s.15] Ortogonaalisella kannalla tarkoitetaan kantaa, jonka vektorit ovat keskenään ortogonaaliset. Vastaavasti määritellään ortonormaali kanta.

3. PERUSTIETOA MATRIISEISTA

Tässä luvussa käydään läpi työn kannalta olennaisia matriiseihin liittyviä esitietoja suhteellisen kattavasti. Matriisien peruskäsitteet määritellään ja keskeiset ominaisuudet todistetaan. Keskeiset sisällöt on jaettu alaluvuiksi.

Matriisi on eräänlainen matematiikan objekti. Matriiseja usein kuvataan taulukkona, joka sisältää alkioita, kuten lukuja tai lausekkeita. Toinen tapa ajatella matriiseja on vektoriavaruuksien välisinä lineaarikuvauksina. [9, s.5] Ensimmäisenä esitelty tapa on työn kannalta oleellisempi. Matriiseja merkitään usein isoilla kirjaimilla ja matriisin alkioita pienillä kirjaimilla, joiden alaindeksinä on alkion rivin ja sarakkeen indeksi. Esimerkiksi a_{23} on toisen rivin ja kolmannen sarakkeen alkio.

Määritelmä 3.1. Matriisi A on $m \times n$-kokoinen taulukko kunnan $\mathbf{F}$ skalaareja a_{ij}. Tällöin matriisia A merkitään seuraavasti:

$$A_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]_{m\times n} \in M_{m,n}(\mathbf{F}).$$

Merkintä $M_{m,n}(\mathbf{F})$ tarkoittaa kaikkien $m \times n$-kokoisten matriisien, joiden alkiot ovat kunnan $\mathbf{F}$ skalaareja, muodostamaa joukkoa. Matriisin $A = [a_{ij}]_{m\times n} \in M_{m,n}(\mathbb{F})$ päädiagonaali tarkoittaa alkioista $a_{11}, a_{22}, \ldots, a_{qq}$, missä $q = \min\{m, n\}$, muodostettua listaa. Matriisin A p. ylädiagonaali tarkoittaa alkioista $a_{1,p+1}, a_{2,p+2}, \ldots, a_{k,p+k}$, missä $k = \min\{m, n-p\}$ ja $p = 0, 1, 2, \ldots, n-1$, muodostettua listaa. Matriisin A p. aladiagonaali tarkoittaa alkioista $a_{p+1,1}, a_{p+2,2}, \ldots, a_{p+l,l}$, missä $l = \min\{m-p, n\}$ ja $p = 0, 1, 2, \ldots, m-1$, muodostettua listaa. [9, s.5] Jos tekstissä viitataan diagonaaliin ilman erityistä mainintaa, tällä tarkoitetaan päädiagonaalia.

Matriisien merkintätapaa voidaan yksinkertaistaa, jos asiayhteydestä käy selvästi ilmi matriisin koko tai se, minkä kunnan skalaareita matriisi sisältää. Jos matriisin koko on tiedossa, voidaan matriisi kirjoittaa $A = [a_{ij}]$ ilman alaindeksiä, joka kertoo matriisin rivien ja sarakkeiden lukumäärän. Tässä työssä käsitellään oletusarvoisesti kompleksisia matriiseja $M_{m,n}(\mathbb{C})$, vaikka osa matriisien operatioista ja ominaisuuksista olisivatkin voimassa mielivaltaisen kunnan $\mathbb{F}$ skalaareille, minkä vuoksi tätä merkintää voidaan lyhentää muotoon $M_{m,n}$. Lisäksi, jos rajoitutaan tarkastelemaan ehdon $m = n$ toteuttavia neliömatriiseja, käytetään matriisien joukosta merkintää M_n.

3.1 Matriisien laskusääntöjä

Määritelmästä 3.2 on helppo nähdä, että matriisien yhteenlasku on määtitelty vain samankokoisille matriiseille ja että se suoritetaan alkioittain. Skalaarikertolaskussa jokainen alkio kerrotaan kyseisellä skalaarilla. Pistetuloa eli skalaarituloa ei tule sekoittaa skalaarikertolaskuun, sillä se on aivan eri asia. Matriisin transpoosi puolestaan tekee rivivektoreista sarakevektoreita ja sarakevektoreista rivivektoreita. Konjugaattitranspoosi on nimensä mukaan kompleksikonjugaatin ja transpoosin yhdistelmä.

Määritelmä 3.2. Olkoot $A = [a_{ij}], B = [b_{ij}] \in M_{m,n}$ matriiseja ja olkoon $\alpha \in \mathbb{C}$ kompleksiluku. Tällöin matriiseille määritellään yhteenlasku

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] \in M_{m,n},$$

skalaarikertolasku

$$\alpha A = \alpha[a_{ij}] = [\alpha a_{ij}] \in M_{m,n},$$

kompleksikonjugaatti

$$\overline{A} = \overline{[a_{ij}]} = [\overline{a}_{ij}] \in M_{m,n},$$

transpoosi

$$A^T = [a_{ij}]^T = [a_{ji}] \in M_{n,m}$$

ja konjugaattitranspoosi

$$A^* = \overline{A}^T = \overline{A^T} = [\overline{a}_{ji}].$$

Matriisi $-A$ määritellään skalaarikertolaskun avulla seuraavasti: $-A = (-1)A$. Näin ollen matriisien A ja B erotus $A - B$ voidaan määritellä kirjoittamalla erotus matriisien yhteenlaskun ja skalaarikertolaskun avulla seuraavasti: $A - B = A + (-1)B$.

Määritelmä 3.3. Olkoon $A = [A_{ij}] \in M_{m,p}$ ja $B = [b_{ij}] \in M_{p,n}$. Matriisitulo AB määritellään

$$AB = \left[\sum_{k=1}^{p} a_{ik} b_{kj}\right] \in M_{m,n}. \tag{3.1}$$

Matriisitulo on siis määritelty, kun vasemmanpuoleisen tulontekijän sarakkeiden lukumäärä on yhtäsuuri kuin oikeanpuoleisen tulontekijän rivien lukumäärä. Tulona saadun matriisin rivien lukumäärä on sama kuin vasemmanpuoleisen tulontekijän ja sarakkeiden sama kuin oikeanpuoleisen tulontekijän. Alla olevan esimerkin avulla on helppo havaita, että matriisitulo ei välttämättä ole kommutatiivinen. Toisin sanoen matriiseille A ja B ei ole yleisesti voimassa yhtälö $AB = BA$.

Esimerkki 3.1. Olkoot $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ ja $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Silloin

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = BA.$$

Lisäksi neliömatriiseille voidaan määritellä matriisipotenssi hyödyntäen matriisituloa. Eksponentin on oltava positiivinen kokonaisluku. Myöhemmin työssä määritellään matriisin juuri, mutta tällöin matriisin on oltava positiivisesti semidefiniitti.

Määritelmä 3.4. Olkoon $A \in M_n$ ja olkoon k positiivinen kokonaisluku. Tällöin matriisin A potenssi on

$$A^k = \overbrace{AA \cdots A}^{k \text{ kpl}}.$$

Neliömatriisien joukko M_n, matriisien yhteenlasku ja matriisitulo muodostavat renkaan $(M_n, +, \cdot)$. Tämä tarkoittaa, että joukko M_n on Abelin ryhmä yhteenlaskun suhteen, monoidi matriisitulon suhteen ja lisäksi matriisitulo on distributiivinen yhteenlaskun suhteen. Nämä on osoitettu lauseessa 3.1. Tämän renkaan kertolaskun neutraalialkiota kutsutaan identiteettimatriisiksi ja sitä merkitään

$$I = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \in M_n. \tag{3.2}$$

Identiteettimatriisin diagonaalialkiot ovat ykkösiä ja kaikki muut alkiot ovat nollia. Identiteettimatriisilla kertominen ei vaikuta tuloon. Yhtälö $IA = A = AI$ on siis voimassa kaikilla neliömatriiseilla A. Matriisille $B \in M_{m,n}$ on voimassa $I_m B = B = BI_n$. Renkaan yhteenlaskun neutraalialkio on nollamatriisi O, jonka jokainen alkio on 0. Asiayhteydestä riippuen 0 voi viitata reaaliluvun ohella myös nollavektoriin. Jos on olemassa mahdollisuus merkintöjen sekoittamiseen, niin alaindeksillä ilmoitetaan dimensiot. Samoin toimitaan identiteettimatriisin I kanssa.

Lauseen 3.1 ominaisuudet ovat yleisesti voimassa kaikille sopivan kokoisille matriiseille, mutta vain neliömatriisit muodostavat renkaan, sillä matriisitulo ei ole suljettu joukossa $M_{m,n}$. Yhteenlasku on assosiatiivinen, kommutatiivinen, yhteenlaskulla on neutraalialkio ja vasta-alkio (ominaisuudet (a)-(d)). Matriisitulo on assosiatiivinen ja tulolla on neutraalialkio (ominaisuudet (e) ja (f)). Lisäksi tulo on distributiivinen yhteenlaskun suhteen (ominaisuudet (g) ja (h)). Viimeiset kohdat ovat skalaarikertolaskun ja kompleksikonjugaatin ominaisuuksia, eivätkä ne liity renkaisiin.

Lause 3.1. *(Vrt. [6], Theorem 0.3.1 ja alaluku 1.1.2 Theorems, sekä [5] alaluku 1.9) Olkoot $\alpha, \beta \in \mathbb{C}$. Sopivan kokoisilla matriiseille $A = [a_{ij}], B = [b_{ij}]$ ja $C = [c_{ij}]$ on voimassa seuraavat yhteen- ja kertolaskusäännöt:*

(a) $(A + B) + C = A + (B + C)$,

(b) $A + B = B + A$,

(c) $O + A = A + O = A$,

(d) $A + (-A) = (-A) + A = O$,

(e) $(AB)C = A(BC)$,

(f) $IA = A = AI$,

(g) $A(B + C) = AB + AC$,

(h) $(A + B)C = AC + BC$,

(i) $(\alpha + \beta)A = \alpha A + \beta A$,

(j) $\alpha(A + B) = \alpha A + \alpha B$,

(k) $\alpha(\beta A) = (\alpha\beta)A$,

(l) $(\alpha A)B = A(\alpha B) = \alpha AB$,

(m) $\overline{AB} = \overline{A}\,\overline{B}$ *ja* $\overline{\overline{A}} = A$.

Todistus. (a) $(A + B) + C = [a_{ij} + b_{ij}] + [c_{ij}] = [(a_{ij} + b_{ij}) + c_{ij}] = [a_{ij} + b_{ij} + c_{ij}]$
$= [a_{ij} + (b_{ij} + c_{ij})] = [a_{ij}] + [b_{ij} + c_{ij}] = A + (B + C)$.

(b) $A + B = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = B + A$.

(c) $O + A = [0 + a_{ij}] = [a_{ij} + 0] = [a_{ij}] = A$.

(d) $A + (-A) = [a_{ij} + (-a_{ij})] = [(-a_{ij}) + a_{ij}] = (-A) + A = [0] = O$.

(e) $$(AB)C = \left[\sum_{k=1}^{p} a_{ik}b_{kj}\right] C = \left[\sum_{h=1}^{q}\left(\sum_{k=1}^{p} a_{ik}b_{kh}\right)c_{hj}\right] = \left[\sum_{k=1}^{p}\sum_{h=1}^{q} a_{ik}b_{kh}c_{hj}\right]$$
$$= \left[\sum_{k=1}^{p} a_{ik}\left(\sum_{h=1}^{q} b_{kh}c_{hj}\right)\right] = A\left[\sum_{h=1}^{q} b_{ih}c_{hj}\right] = A(BC).$$

(f) Matriisissa I alkio $i_{ij} = 1$, jos $i = j$, ja muulloin $i_{ij} = 0$. Täten
$$IA = \left[\sum_{k=1}^{p} i_{ik}a_{kj}\right] = [a_{ij}] = A = [a_{ij}] = \left[\sum_{k=1}^{p} a_{ik}i_{kj}\right] = AI.$$

(g) $$A(B + C) = A[b_{ij} + c_{ij}] = \left[\sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj})\right] = \left[\sum_{k=1}^{p} a_{ik}b_{kj} + a_{ik}c_{kj}\right]$$
$$= \left[\sum_{k=1}^{p} a_{ik}b_{kj} + \sum_{k=1}^{p} a_{ik}c_{kj}\right] = AB + AC.$$

(h) $$(A + B)C = [a_{ij} + b_{ij}]C = \left[\sum_{k=1}^{p}(a_{ik} + b_{ik})c_{kj})\right] = \left[\sum_{k=1}^{p} a_{ik}c_{kj} + b_{ik}c_{kj}\right]$$
$$= \left[\sum_{k=1}^{p} a_{ik}c_{kj} + \sum_{k=1}^{p} b_{ik}c_{kj}\right] = AC + BC.$$

(i) $(\alpha+\beta)A = (\alpha+\beta)[a_{ij}] = [(\alpha+\beta)a_{ij}] = [\alpha a_{ij}+\beta a_{ij}] = [\alpha a_{ij}]+[\beta a_{ij}] = \alpha A+\beta A.$

(j) $\alpha(A+B) = \alpha[a_{ij}+b_{ij}] = [\alpha(a_{ij}+b_{ij})] = [\alpha a_{ij}+\alpha b_{ij}] = [\alpha a_{ij}]+[\alpha b_{ij}] = \alpha A+\alpha B.$

(k) $\alpha(\beta A) = \alpha[\beta a_{ij}] = \alpha\beta[a_{ij}] = (\alpha\beta)A.$

(l) $(\alpha A)B = \left[\sum_{k=1}^{p}(\alpha a_{ik})b_{kj}\right] = \left[\sum_{k=1}^{p} a_{ik}(\alpha b_{kj})\right] = A(\alpha B) = \left[\alpha \sum_{k=1}^{p} a_{ik}b_{kj}\right] = \alpha AB.$

(m) $\overline{AB} = \left[\sum_{k=1}^{p} \overline{a_{ik}b_{kj}}\right] = \left[\sum_{k=1}^{p} \overline{a}_{ik}\overline{b}_{kj}\right] = \overline{A}\,\overline{B}$ ja $\overline{\overline{A}} = [\overline{\overline{a}}_{ij}] = [a_{ij}] = A.$

□

Lause 3.2. *(Vrt. [13], lause 1.14.15) Sopivan kokoisilla matriiseille $A = [a_{ij}]$ ja $B = [b_{ij}]$ on voimassa seuraavat transpoosin ja konjugaattitranspoosin laskusäännöt:*

(a) $(A^T)^T = A$ *ja* $(A^*)^* = A$.

(b) $(AB)^T = B^T A^T$ *ja* $(AB)^* = B^* A^*$.

(c) $(A + B)^T = A^T + B^T$ *ja* $(A + B)^* = A^* + B^*$.

(d) $(\alpha A)^T = \alpha A^T$ *ja* $(\alpha A)^* = \overline{\alpha} A^*$.

Todistus. (a) $(A^T)^T = [a_{ji}]^T = [a_{ij}] = A$ ja $(A^*)^* = [\overline{a}_{ji}]^* = [\overline{\overline{a}}_{ij}] = [a_{ij}] = A.$

(b) $[AB]_{ij}^T = [AB]_{ji} = \sum_{k=1}^{p} a_{jk}b_{ki}$ ja $[B^T A^T]_{ij} = \sum_{k=1}^{p} b_{ki}a_{jk} = \sum_{k=1}^{p} a_{jk}b_{ki}$. Lisäksi $(AB)^* = \overline{(AB)^T} = \overline{(B^T A^T)} = B^* A^*.$

(c) $(A + B)^T = [a_{ij} + b_{ij}]^T = [a_{ji} + b_{ji}] = A^T + B^T$ ja $(A + B)^* = \overline{(A + B)^T}$ $= \overline{A^T + B^T} = A^* + B^*$.

(d) $(\alpha A)^T = [\alpha a_{ij}]^T = [\alpha a_{ji}] = \alpha A^T$ ja $(\alpha A)^* = [\overline{\alpha a}_{ij}]^T = \overline{\alpha}[\overline{a}_{ji}] = \overline{\alpha} A^*.$

□

3.2 Rivi- ja sarakeoperaatiot sekä redusoitu riviporrasmuoto

On olemassa kolme elementaarista rivi- tai sarakeoperaatiota, joiden avulla matriisia voidaan muokata sellaiseen muotoon, joka helpottaa lineaaristen yhtälöryhmien ratkaisua tai matriisin asteen, inverssin ja determinantin määrittämistä. Jokaista tällaista operaatiota vastaa alkeismatriisi. Rivioperaatioissa muunnettu matriisi saadaan kertomalla operaatiota vastaavalla alkeismatriisilla vasemmalta ja sarakeoperaatiossa oikealta. [9, s.9] Määritellään seuraavaksi nämä kolme rivioperaatiota. Sarakeoperaatiot on mahdollista määritellä vastaavasti.

Ensimmäinen rivioperaatio on rivienvaihto. Jos rivien i ja $j \neq i$ paikkaa vaihdetaan, vastaa se kertomista matriisin kertomista vasemmalta alkeismatriisilla E, joka on identiteettimatriisi, jonka rivien i ja j paikat on vaihdettu. Toinen rivioperaatio on rivin i kertominen skalaarilla $\alpha \in \mathbb{C}\backslash\{0\}$. Tällöin alkeismatriisi E on identiteettimatriisi, jonka rivi i on kerrottu skalaarilla α. Kolmas rivioperaatio on skalaarilla α kerrotun rivin i yhteenlasku johonkin toiseen riviin $j \neq i$. Tällöin alkeismatriisi on identiteettimatriisi, jonka rivin j i:s alkio on nollan sijaan skalaari α.

Alkeismatriisit ovat ei-singulaareja (määritelmä 3.6). Vakuutetaan seuraavaksi lukija tämän tuloksen järkevyydestä. Jos alkeismatriisi E_1 vastaa matriisin rivienvaihtoa, niin voidaan todeta, että $E_1E_1 = I$. Jos alkeismatriisi E_2 vastaa matriisin rivin i kertomista skalaarilla α, tällöin voidaan määritellä alkeismatriisi E_2^{-1}, joka vastaa rivin i kertomista skalaarilla $\frac{1}{\alpha}$. Täten laskemalla voidaan havaita, että $E_2E_2^{-1} = I$. Jos alkeismatriisi E_3 vastaa skalaarilla α kerrotun rivin i lisäämistä toiseen riviin j, niin voidaaan määritellä alkeismatriisi E_3^{-1}, joka vastaa skalaarilla $-\alpha$ kerrotun rivin i lisäämistä riviin j. Laskemalla havaitaan, että $E_3E_3^{-1} = I$.

Määritelmä 3.5 (Redusoitu riviporrasmuoto)**.** Olkoon $R = [r_{ij}] \in M_{m,n}$. Matriisi R on redusoidussa riviporrasmuodossa (RREF-muoto), jos seuraavat ehdot toteutuvat:

(a) Jos matriisilla on nollarivejä, ne ovat alimpana.

(b) Jos rivi ei ole nollarivi, sen johtava alkio (ensimmäinen alkio) on 1.

(c) Jokaisen johtavan alkion sarakkeessa muut alkiot ovat nollia.

(d) Johtavat alkiot muodostavat porrasmaisen kuvion. Jos rivi i ei ole nollarivi ja r_{ik} on sen johtava alkio, tällöin joko $i = m$, rivi $i + 1$ on nollarivi tai rivin $i + 1$ johtava alkio on $r_{i+1,l}$, missä $l > k$.

Jokainen matriisi voidaan redusoida RREF-muotoon hyödyntämällä rivioperaatioita. Jos $R \in M_{m,n}$ on matriisin $A \in M_{m,n}$ RREF-muoto, niin $R = EA$, missä matriisi $E \in M_m$ on tehtyjä rivioperaatioita vastaavien alkeismatriisien tulo. [9, s.11] Matriisi E on ei-singulaarinen, sillä ei-singulaaristen matriisien tulo on ei-singulaarinen lauseen 3.6 nojalla. Matriisi E^{-1} on myös alkeismatriisien tulo. Lisäksi, jos R on neliömatriisi, niin $R = I$ tai matriisi R sisältää ainakin yhden nollarivin.

3.3 Singulaarisuus

Jokaisella nollasta eroavalla kompleksiluvulla α on olemassa kertolaskun käänteisalkio $\alpha^{-1} = \frac{1}{\alpha}$. Kaikilla matriiseilla $A \neq 0$ ei kuitenkaan ole olemassa matriisitulon käänteisalkioita eli inverssiä tai käänteismatriisia. Jos matriisilla on käänteisalkio, sanotaan, että matriisi on ei-singulaarinen tai kääntyvä. Muutoin matriisi on singulaarinen. Ei-singulaarisen matriisin on oltava neliömatriisi. Matriisi $A \in M_{m,n}$, missä $m \neq n$, voi olla vasemman- tai oikeanpuoleinen inverssi, vaikka se ei olisikaan kääntyvä. Toispuoleiset inverssit ovat samat neliömatriiseilla, ja käänteismatriisi on yksikäsitteinen. Täten kun halutaan osoittaa, että neliömatriisin A käänteismatriisi on neliömatriisi B, riittää osoittaa joko $BA = I$

tai $AB = I$. Toisen ollessa tosi on toinenkin välttämättä tosi.

Määritelmä 3.6. Matriisi $A \in M_n$ on ei-singulaarinen (kääntyvä), jos on olemassa sellainen matriisi $B \in M_n$, että

$$BA = AB = I. \tag{3.3}$$

Muulloin matriisi A on singulaarinen.

Lause 3.3. *Jos matriisi $A \in M_n$ on ei-singulaarinen, niin tällöin määritelmän 3.6 matriisi $B \in M_n$ on yksikäsitteinen ja sanotaan, että matriisi B on matriisin A käänteismatriisi tai inverssi.*

Todistus. (Vrt. [14], Theorem 3.6) Oletetaan, että kaksi matriisia $B, C \in M_n$ toteuttavat määritelmän ehdon. Identiteettimatriisin ominaisuuksia ja ei-singulaarisen matriisin määritelmää hyödyntämällä saadaan

$$B = BI = B(AC) = (BA)C = IC = C. \tag{3.4}$$

Näin ollen $B = C$ ja käänteismatriisi on yksikäsitteinen. □

Matriisin A käänteismatriisia merkitään usein A^{-1}. Matriisin singulaarisuuden perusteella voidaan päätellä monenlaisia tärkeitä ominaisuuksia. Onkin hyödyllistä, että singulariteetti voidaan päätellä eri tavoilla. Tällaisia työn kannalta olennaisia, yhtäpitäviä kriteereitä on koottu alla olevaan lauseeseen.

Lause 3.4. *Olkoon $A \in M_n$. Tällöin seuraavat väittämät ovat yhtäpitäviä:*

(a) *Matriisi A on ei-singulaarinen.*

(b) *Inverssi A^{-1} on olemassa.*

(c) *Yhtälöllä $Ax = b$ on yksikäsitteinen ratkaisu jokaisella $b \in \mathbb{C}^n$.*

(d) *Yhtälöllä $Ax = 0$ on vain triviaaliratkaisu $x = 0$.*

(e) *Matriisin A determinantti ei ole nolla, eli $\det A \neq 0$.*

(f) *Luku 0 ei ole matriisin A ominaisarvo.*

Todistus. (Vrt. [14], Theorem 3.12) Kohdat (a) ja (b) ovat selvästi yhtäpitäviä. Todistetaan päättelyketju $(b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (b)$. Jos matriisi A^{-1} on olemassa, niin $x = A^{-1}b$ on eräs yhtälön $Ax = b$ ratkaisu, sillä $A(A^{-1}b) = Ib = b$. Oletetaan, että yhtälöllä on toinenkin ratkaisu y, jolloin $Ay = b \Leftrightarrow y = A^{-1}b = x$. Täten ratkaisu on yksikäsitteinen. Kohta (d) seuraa valitsemalla edellisessä kohdassa $b = 0$. Oletetaan sitten, että yhtälöllä $Ax = 0$ on vain triviaaliratkaisu $x = 0$. Matriisi A voidaan kirjoittaa $A = E^{-1}R$, missä matriisi $E \in M_n$ on alkeismatriisien tulo ja matriisi $R \in M_n$ on RREF-muotoinen matriisi. Täten $Ax = E^{-1}Rx = 0 \Leftrightarrow Rx = 0$. Jos $R = I$, yhtälöllä $Rx = 0$ on vain triviaaliratkaisu. Jos $R \neq I$, niin matriisi R sisältää ainakin yhden nollarivin, mistä seuraa, että yhtälöryhmässä $Rx = 0$ vektorissa x on n tuntematonta skalaaria, mutta yhtälöitä on enintään $n-1$. Tällöin yhtälöryhmällä $Rx = 0$ ei voi olla yksikäsitteistä ratkaisua. On siis oltava $R = I$, joten $A^{-1} = E$ on olemassa.

Kohtien (f) ja (g) yhtäpitävyys on todistettu lauseissa 3.10 ja 3.14. $\square$

Kaikkien sellaisten vektoreiden $b \in \mathbb{C}^m$, jolla yhtälöllä $Ax = b$ on ainakin yksi ratkaisu, muodostamaa joukkoa kutsutaan matriisin $A \in M_{m,n}$ sarakeavaruudeksi. Matriisin $A \in M_{m,n}$ nolla-avaruus on kaikkien yhtälön $Ax = 0$ ratkaisujen $x \in \mathbb{C}^n$ muodostama joukko. Nämä ovat tärkeitä avaruuksien $\mathbb{C}^m$ ja $\mathbb{C}^n$ aliavaruuksia.

Määritelmä 3.7. Olkoon $A \in M_{m,n}$. Tällöin matriisin A sarakeavaruus on $\mathcal{R}(A) = \{b \in \mathbb{C}^m : Ax = b \text{ jollain vektorilla } x \in \mathbb{C}^n\}$ ja nolla-avaruus on $\mathcal{N}(A) = \{x \in \mathbb{C}^n : Ax = 0\}$.

Lause 3.5 (Dimensiolause). *Olkoon matriisi* $A \in M_{m,n}$. *Tällöin* $\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = n$.

Todistus. (Vrt. [14], Theorem 6.19) Oletetaan, että nolla-avaruuden dimensio on positiivinen kokonaisluku r ja $0 < r < n$. Lisäksi oletetaan, että vektoriavaruuden $\mathbb{C}^n$ kannan muodostavat vektorit $v_1, \dots, v_r, v_{r+1}, \dots, v_n$. Tällöin nolla-avaruuden kannan virittää r kappaletta vektoriavaruuden $\mathbb{C}^n$ vektoreita. Rajoittamatta todistuksen yleisyyttä voidaan olettaa, että nämä vektorit ovat $v_1, \dots, v_r$. Toisin sanoen $\mathcal{N}(A) = \text{span}\{v_1, \dots, v_r\}$ ja $Av_1 = \cdots = Av_r = 0$. Olkoon $S = \{Av_{r+1}, \dots, Av_n\}$ vektorijoukko. Osoitetaan, että joukon S vektorit ovat lineaarisesti riippumattomia ja että ne virittävät matriisin A sarakeavaruuden, mistä seuraa, että S on matriisin A sarakeavaruuden kanta. Tutkitaan joukon S vektoreiden lineaarikombinaatiota ja merkitään

$$\alpha_{r+1}(Av_{r+1}) + \cdots + \alpha_n(Av_n) = A(\alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n) = 0.$$

Yhtälö on tosi, jos ja vain jos vektori $\alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n \in \mathcal{N}(A)$. Täten se voidaan kirjoittaa nolla-avaruuden kantavektorien lineaarikombinaationa

$$\begin{aligned}
&\alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n = \alpha_1 v_1 + \cdots + \alpha_r v_r \\
&\alpha_1 v_1 + \cdots + \alpha_r v_r - \alpha_{r+1}v_{r+1} - \cdots - \alpha_n v_n = 0.
\end{aligned}$$

Koska vektorit $v_1, \dots, v_n$ muodostavat kannan, ne ovat lineaarisesti riippumattomia. Tästä seuraa, että edellinen yhtälö toteutuu vain, jos $\alpha_1 = \cdots = \alpha_n = 0$. Näin ollen joukon S vektorit ovat lineaarisesti riippumattomia.

Olkoon $Ax \in \mathcal{R}(A)$ mikä tahansa matriisin A sarakeavaruuden vektori. Koska $x \in \mathbb{C}^n$, se voidaan kirjoittaa vektorien $v_1, \dots, v_n$ lineaarikombinaationa. Täten

$$\begin{aligned} Ax &= A(\alpha_1 v_1 + \cdots + \alpha_r v_r + \alpha_{r+1} v_{r+1} + \cdots + \alpha_n v_n) \\ &= \alpha_1(Av_1) + \cdots + \alpha_r(Av_r) + \alpha_{r+1}(Av_{r+1}) + \cdots + \alpha_n(Av_n) \\ &= 0 + \alpha_{r+1}(Av_{r+1}) + \cdots + \alpha_n(Av_n). \end{aligned}$$

Vektori Ax voidaan siis kirjoittaa vektorien $Av_{r+1}, \dots, Av_n$ lineaarikombinaationa, joten joukko S virittää matriisin A sarakeavaruuden. Tämä ja vektorien lineaarinen riippumattomuus takaavat, että $\mathcal{R}(A) = \text{span}\{Av_{r+1}, \dots, Av_n\}$ ja $\dim \mathcal{R}(A) = n - r$. Nyt

$$\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = (n - r) + r = n.$$

Jos $r = 0$ tai $r = n$, niin vastaavasti kuin edellä voidaan osoittaa, että $\mathcal{R}(A) = \text{span}\, S$ ja $\dim \mathcal{R}(A) = n - r$. Tällöin valitaan $\mathcal{N}(A) = \{0\}$, jolloin $\dim \mathcal{N}(A) = 0$ ja $S = \{Av_1, \dots, Av_n\}$ tai $\mathcal{N}(A) = \{v_1, \dots, v_n\}$, jolloin $\dim \mathcal{N}(A) = n$ ja $S = \{0\}$, tässä järjestyksessä. □

Lause 3.6. *(Vrt. [6], Theorem 1.3.3a, ja [6, s.58], Corollary 4) Olkoot $A, B \in M_n$. Tällöin matriisien tulo AB on ei-singulaarinen ja tällöin $(AB)^{-1} = B^{-1}A^{-1}$, jos ja vain jos matriisit A ja B ovat ei-singulaarisia.*

Todistus. Osoitetaan, että matriisi AB on singulaarinen, jos ja vain jos matriisi A tai B on singulaarinen. Tämä väittämä on yhtäpitävä lauseen väittämän kanssa. Oletetaan aluksi, että matriisi B on singulaarinen. Tällöin lauseen 3.4 nojalla on olemassa jokin sellainen nollasta eroava vektori x, että $Bx = 0$. Täten yhtälöllä $(AB)x = A(Bx) = A0 = 0$ on ei-triviaali ratkaisu riippumatta siitä, onko matriisi A singulaarinen vai ei. Näin ollen lauseesta 3.4 seuraa, että AB on singulaarinen. Jos matriisi A on singulaarinen ja matriisi B on ei-singulaarinen, niin tällöin on olemassa sellaiset vektorit $b, x \neq 0$, että $Ab = 0$ ja $Bx = b$. Näin ollen yhtälöllä $(AB)x = A(Bx) = Ab = 0$ on ei-triviaali ratkaisu ja lauseen 3.4 perusteella AB on singulaarinen. Todistetaan sitten väite toiseen suuntaan. Jos matriisi AB on singulaarinen, niin tällöin on olemassa sellainen nollasta eroava vektori $x \neq 0$, että $ABx = 0$. Jos matriisit A ja B ovat ei-singulaareja, niin tällöin lauseen 3.4 nojalla olisi $ABx \neq 0$, mikä on ristiriidassa oletuksen $ABx = 0$ kanssa. Näin ollen Matriisi A tai B on singulaarinen.

Jos matriisi AB on ei-singulaarinen, niin $(AB)^{-1}$ on yksikäsitteinen. Ei-singulaariselle matriisille on oltava voimassa määritelmän 3.6 yhtälö ja koska

$$(AB)(B^{-1}A^{-1}) = AIA^{-1} = I = B^{-1}IB = (B^{-1}A^{-1})(AB),$$

niin $(AB)^{-1} = B^{-1}A^{-1}$. □

Edellistä lausetta tarvitaan determinantin ominaisuuden $\det AB = \det A \det B$ todistuksessa. Seuraavassa lauseessa on matriisien $A^{-1}, \alpha A, A^T$ ja A^* käänteismatriisit. Kohta (a) osoittaa, että jos matriisin A käänteismatriisi on matriisi B, niin matriisin B käänteismatriisi on A. Kohtien (c) ja (d) perusteella voidaan kirjoittaa $(A^T)^{-1} = (A^{-1})^T = A^{-T}$ ja $(A^*)^{-1} = (A^{-1})^* = A^{-*}$.

Lause 3.7. *(Vrt. [14], Theorem 3.9) Olkoon matriisi $A \in M_n$ ei-singulaarinen ja skalaari $\alpha \in \mathbb{C}$ on nollasta eroava. Tällöin*

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

(b) $(\alpha A)^{-1} = \alpha^{-1} A^{-1}$,

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

(d) $(A^*)^{-1} = (A^{-1})^*$.

Todistus. Osoitetaan, että $AB = I$, josta seuraa $A^{-1} = B$.

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

(b) $(\alpha A)(\alpha^{-1}A^{-1}) = 1 \cdot AA^{-1} = I$.

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

(d) $(A^*)(A^{-1})^* = (A^{-1}A)^* = I^* = I$.

□

3.4 Alimatriisit, ositukset ja lohkomatriisit

Joukon S ositus on sellainen kokoelma joukon S osajoukkoja, että jokainen joukon S alkio on täsmälleen yhden osajoukon alkio. Esimerkiksi joukko $\{1, 2, \ldots, n\}$ on kokoelma osajoukoista, joita sanotaan indeksijoukoiksi, $\alpha_1, \alpha_2, \ldots, \alpha_t$ siten, että jokainen kokonaisluku lukujen 1 ja n välillä kuuluu täsmälleen yhteen indeksijoukkoon. Kyseessä on peräkkäinen ositus, jos joukko S jaetaan sellaisiin indeksijoukkoihin, että ne ovat muotoa $\alpha_1 = \{1, \ldots, i_1\}, \alpha_2 = \{i_1 + 1, \ldots, i_2\}, \ldots, \alpha_t = \{i_{t-1} + 1, \ldots, n\}$. [9, s.16]

Määritelmä 3.8. Olkoon $A \in M_{m,n}$, ja olkoot $\alpha \subseteq \{1, 2, \ldots, m\}$ ja $\beta \subseteq \{1, 2, \ldots, n\}$ joukkoja. Matriisista A saatua matriisia, jossa säilytetään joukon α indeksoimat rivit ja joukon β indeksoimat sarakkeet, sanotaan matriisin A alimatriisiksi ja sitä merkitään

$$A[\alpha, \beta].$$

Jos $\alpha = \beta$, niin kyseessä on pääalimatriisi, jota merkitään

$$A[\alpha] = A[\alpha, \alpha].$$

Jos $A \in M_n$ on neliömatriisi ja $k \in \{1, 2, \dots, n\}$, niin pääalimatriisia

$$A[\{1, 2, \dots, k\}]$$

sanotaan johtavaksi pääalimatriisiksi, ja jos $k \in \{k, \dots, n\}$, niin matriisia

$$A[\{k, \dots, n\}]$$

sanotaaan häntäpääalimatriisiksi.

On hyvin selvää, että kaikki johtavat pääalimatriisit ovat pääalimatriiseja ja kaikki pääalimatriisit ovat alimatriiseja. Joskus alimatriisit on kätevämpää ilmoittaa poistettavien rivien ja sarakkeiden avulla. Tällöin voidaan käyttää joukkojen α ja β komplementteja ilmaisemaan, mitkä rivit ja sarakkeet poistetaan. Perusjoukkona käytetään rivien tai sarakkeiden indeksien muodostamaa joukkoa. Täten esimerkiksi $\alpha^c = \{1, 2, \dots, m\} \setminus \alpha$. Selvennetään vielä määritelmää esimerkin avulla.

Esimerkki 3.2. Olkoon $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ neliömatriisi. Tällöin matriisi

$$B = A[2, \{1, 2, 3\}] = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$$

on matriisin A eräs alimatriisi, matriisi

$$C = A[\{1, 3\}] = \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}$$

on matriisin A eräs pääalimatriisi ja matriisi

$$D = A[\{1, 2\}] = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$$

on matriisin A eräs johtava pääalimatriisi. Vaihtoehtoisesti matriisi B voitaisiin merkitä $B = A[2, \emptyset^c]$ ja päädyttäisiin täysin samaan lopputulokseen.

Matriisin osituksessa matriisi jaetaan sellaisiin alimatriiseihin, että jokainen alkuperäisen matriisin alkio on täsmälleen yhden alimatriisin alkio. Esimerkiksi kätevä tapa esittää matriisit on osittaa matriisi sarakkeiden mukaisesti. Esimerkiksi jos $A \in M_{m,p}$ ja $B \in M_{p,n}$, niin matriisi B voidaan kirjoittaa $B = \begin{bmatrix} b_1 & \dots & b_n \end{bmatrix}$, missä vektori $b_i \in \mathbb{C}^p$ on matriisin B i. sarake. Tällöin matriisitulo AB voidaan esittää muodossa $AB = \begin{bmatrix} Ab_1 & \dots & Ab_n \end{bmatrix}$. [9, s.17]

Jos $\alpha_1, \ldots, \alpha_t$ muodostavat joukon $\{1, \ldots, m\}$ osituksen ja $\beta_1, \ldots, \beta_s$ muodostavat joukon $\{1, \ldots, n\}$ osituksen, niin alimatriisit $A[\alpha_i, \beta_j]$, missä $1 \leq i \leq t$ ja $1 \leq j \leq s$, muodostavat matriisin $A \in M_{m,n}$ osituksen. Jos matriiseille $A \in M_{m,n}$ ja $B \in M_{n,p}$ on muodostettu sellainen ositus, että joukon $\{1, \ldots, n\}$ ositukset ovat samat, niin sanotaan, että matriisit ovat yhteensopivasti ositettu. Tässä tapauksessa voidaan laskea

$$(AB)[\alpha_i, \gamma_j] = \sum_{k=1}^{s} A[\alpha_i, \beta_k]B[\beta_k, \gamma_j], \tag{3.5}$$

missä alimatriisit $A[\alpha_i, \beta_k]$ ja $B[\beta_k, \gamma_j]$ ovat matriisitulon kannalta yhteensopivia. Yhtälön (3.5) vasen puoli on matriisitulon alimatriisi ja oikealla puolella jokainen summan termi on alimatriisien matriisitulo. Näin ollen, jos matriisit A ja B ositetaan yhteensopivasti, niin ositettujen matriisien tulo mukailee tavallista matriistuloa. Vastaavasti, jos matriisit $A, B \in M_{m,n}$ ositetaan yhteensopivasti, niin ositettujen matriisien yhteenlasku on

$$(A+B)[\alpha_i, \beta_j] = A[\alpha_i, \beta_j] + B[\alpha_i, \beta_j] \tag{3.6}$$

ja mukailee tavallista matriisien yhteenlaskua. [9, s.17-18]

Jos matriisi ositetaan rivien ja sarakkeiden peräkkäisen osituksen avulla, niin saadaan ositettu matriisi, jota sanotaan lohkomatriisiksi. Esimerkiksi jos matriisin $A \in M_n$ rivit ja sarakkeet jaetaan samalla peräkkäisellä osituksella osiin $\alpha_1 = \{1, \ldots, k\}$ ja $\alpha_2 = \{k+1, \ldots, n\}$, niin saadaan seuraava lohkomatriisi

$$A = \begin{bmatrix} A[\alpha_1, \alpha_1] & A[\alpha_1, \alpha_2] \\ A[\alpha_2, \alpha_1] & A[\alpha_2, \alpha_2] \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = [A_{ij}], \tag{3.7}$$

missä lohkot ovat $A_{ij} = A[\alpha_i, \alpha_j]$. Kokoa 2×2 olevat lohkomatriisit ovat tärkeimpiä ja käyttökelpoisimpia. [9, s.18] Nollalohkoksi sanotaan lohkoa, jossa on alkioina pelkkiä nollia. Yhtälöt (3.5) ja (3.6) tulevat yhteensopiville lohkomatriiseille $A = [A_{ij}] \in M_{m,p}$ ja $B = [B_{ij}] \in M_{p,n}$ muotoon

$$AB = \left[\sum_{k=1}^{p} A_{ik}B_{kj}\right] \tag{3.8}$$

ja

$$A + B = [A_{ij} + B_{ij}]. \tag{3.9}$$

Lohkomatriisien transpoosi ja konjugaattitranspoosi määritellään seuraavasti: $A^T = \left[A_{ji}^T\right]$ ja $A^* = \left[A_{ji}^*\right]$. Lohkoja käsitellään siis kuten tavallisen matriisin alkioita transpoosissa eli lohkojen rivit ja sarakkeet vaihtuvat keskenään, mutta lisäksi jokaisesta lohkosta otetaan vielä transpoosi tai konjugaatitranspoosi.

3.5 Diagonaalimatriisit ja lohkodiagonaalimatriisit

Matriiseja voidaan luokitella muodon ja ominaisuuksien perusteella. Eri luokille voidaan osoittaa ominaisuuksia, jotka eivät ole voimassa yleisesti kaikille matriiseille. Tällainen luokittelu mahdollistaa matriisien helpomman käsittelyn ja monet ominaisuudet sekä sovellukset toimivat vain tietyn tyyppisille matriiseille. Esimerkiksi diagonaalimatriisien korottaminen potenssiin on hyvinkin triviaalia verrattuna tavallisiin matriisieihin.

Matriisi $D = [d_{ij}] \in M_{m,n}$ on diagonaalimatriisi, jos $d_{ij} = 0$ aina, kun $i \neq j$. Lisäksi jos matriisin kaikki diagonaalialkiot ovat positiivisia tai epänegatiivisia, voidaan matriisista käyttää nimitystä positiivinen tai epänegatiivinen diagonaalimatriisi. Jos matriisi D on neliömatriisi, jonka kaikki diagonaalialkiot ovat yhtäsuuria, niin matriisi kutsutaan skalaarimatriisiksi. Skalaarimatriisi voidaan kirjoittaa identiteettimatriisin avulla $D = \alpha I$, missä $\alpha \in \mathbb{C}$. Matriisitulo skalaarimatriisin kanssa vastaa skalaarikertolaskua. [9, s.30]

Jos $A = [a_{ij}] \in M_{m,n}$ ja $q = \min\{m, n\}$, niin merkintä $\operatorname{diag} A = [a_{11}, \ldots, a_{qq}]^T \in \mathbb{C}^q$ tarkoittaa vektoria, jonka alkioina ovat matriisin A diagonaalialkiot. Ja käänteisesti, jos $x \in \mathbb{C}^n$ ja m sekä n ovat sellaisia positiivisia kokonaislukuja, että $\min\{m, n\} = q$, niin merkintä $\operatorname{diag} x \in M_{m,n}$ tarkoittaa sellaista $m \times n$-kokoista diagonaalimatriisia A, jolle $\operatorname{diag} A = x$ eli jonka diagonaalialkiot ovat vektorista x. Jotta $\operatorname{diag} x$ on hyvin määritelty, kumpikin luku m ja n täytyy olla määritelty. Merkintä $\operatorname{diag}(a_1, \ldots, a_n)$, missä $a_1, \ldots, a_n \in \mathbb{C}$, tarkoittaa aina neliömatriisia $A = [a_{ij}] \in M_n$, jolle $a_{ii} = a_i$ jokaisella $i = 1, \ldots, n$ ja $a_{ij} = 0$, jos $i \neq j$. [9, s.30]

Jos matriisi $A \in M_n$ on jokin mielivaltainen matriisi ja matriisi $D \in M_n$ on diagonaalimatriisi, niin tällöin matriisitulo $DA = [d_{ii}a_{ij}] \in M_n$ kertoo matriisin A rivin i diagonaalimatriisin alkiolla d_{ii}. Vastaavasti matriisitulo $AD = [a_{ij}d_{jj}]$ kertoo matriisin A sarakkeet diagonaalimatriisin alkiolla. Jokaiselle positiiviselle kokonaisluvulle k on voimassa $D^k = \operatorname{diag}(d_{11}^k, \ldots, d_{nn}^k)$. Lisäksi samankokoiset diagonaalimatriisit kommutoivat. Siis jos $D, E \in M_n$ ovat diagonaalimatriiseja, niin $DE = \operatorname{diag}(d_{11}e_{11}, \ldots, d_{nn}e_{nn})$. [9, s.30] Myöhemmin tullaan huomaamaan, että diagonaalimatriisin determinantti on sen diagonaalialkioiden tulo, sen ominaisarvot ovat diagonaalialkiot ja niitä vastaavat ominaisvektorit ovat diagonaalimatriisin sarakkeet.

Määritellään seuraavaksi uusi matriisien operaatio, suora summa. Matriisien A ja B suora summa on lohkomatriisi $A \oplus B$. Jos matriisit A ja B ovat neliömatriiseja, niin suora summa on lohkodiagonaalimatriisi.

Määritelmä 3.9. Olkoon $A \in M_{m,n}$ ja $B \in M_{p,q}$ matriiseja. Tällöin matriisien A ja B suora summa on

$$A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in M_{m+p,n+q}. \tag{3.10}$$

Erityisesti, jos $A \in M_m$ ja $B \in M_n$, niin suora summa $A \oplus B \in M_{m+n}$ on lohkodiagonaalimatriisi.

Esimerkki 3.3. Olkoot $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ ja $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Tällöin suora summa

$$A \oplus B = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 & 0 \\ 7 & 8 & 9 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 & 4 \end{bmatrix}.$$

Lohkodiagonaaliksi kutsutaan matriisia $A \in M_n$, joka on muotoa

$$A = \begin{bmatrix} A_{11} & & 0 \\ & \ddots & \\ 0 & & A_{kk} \end{bmatrix},$$

missä $A_{ii} \in M_{n_i}, i = 1, \ldots, k, \sum_{i=1}^{k} n_i = n$, ja jokainen lohkodiagonaalin ala- ja yläpuolella oleva lohko on nollalohko. Tällainen matriisi on kätevä kirjoittaa suoran summan avulla seuraavasti

$$A = A_{11} \oplus A_{22} \oplus \cdots \oplus A_{kk} = \bigoplus_{i=1}^{k} A_{ii}.$$

Lohkodiagonaalimatriisit yleistävät monet diagonaalimatriisien ominaisuudet, kuten edellä mainitut kommutointia ja determinanttia koskevat ominaisuudet. Esimerkiksi lohkodiagonaalimatriisit $A = \bigoplus\limits_{i=1}^{k} A_{ii}$ ja $B = \bigoplus\limits_{i=1}^{k} B_{ii}$ kommutoivat, jos ja vain jos jokainen pari A_{ii} ja B_{ii} kommutoi. [9, s.30-31]

3.6 Kolmiomatriisit ja lohkokolmiomatriisit

Matriisi $T = [t_{ij}] \in M_{n,m}$ on yläkolmiomatriisi, jos $t_{ij} = 0$ aina, kun $i > j$. Jos $t_{ij} = 0$ aina, kun $i \geq j$, sanotaan, että T on aito yläkolmiomatriisi. Vastaavasti matriisi T on alakolmiomatriisi tai aito alakolmiomatriisi, jos sen transpoosi on yläkolmiomatriisi tai aito yläkolmiomatriisi. Kolmiomatriisi voi olla joko ala- tai yläkolmiomatriisi, ja aito kolmiomatriisi voi olla joko aito ala- tai yläkolmiomatriisi. Yksikkökolmiomatriisi on kolmiomatriisi, jonka kaikki diagonaalialkiot ovat ykkösiä. [9, s.31]

Olkoon $T \in M_{n,m}$ jokin matriisi. Jos T on yläkolmiomatriisi ja $n \leq m$, niin se on muotoa $T = \begin{bmatrix} R & T_2 \end{bmatrix}$. Jos T on yläkolmiomatriisi ja $n \geq m$, niin se on muotoa $T = \left[\begin{smallmatrix} R \\ 0 \end{smallmatrix}\right]$. Tässä $R \in M_{\min\{n,m\}}$ on yläkolmiomatriisi ja T_2 on mielivaltainen (tyhjä, jos $m = n$). Jos T on alakolmiomatriisi ja $n \leq m$, niin se on muotoa $T = \begin{bmatrix} L & 0 \end{bmatrix}$. Jos T on alakolmiomatriisi ja $n \geq m$, niin se on muotoa $T = \left[\begin{smallmatrix} L \\ T_2 \end{smallmatrix}\right]$, jos $n \geq m$. Tässä $L \in M_{\min\{n,m\}}$ on alakolmiomatriisi ja T_2 on mielivaltainen (tyhjä, jos $m = n$). [9, s.31]

Lohkoyläkolmiomatriisiksi kutsutaan matriisia $A \in M_n$, joka on muotoa

$$A = \begin{bmatrix} A_{11} & \bigstar & \bigstar \\ & \ddots & \bigstar \\ O & & A_{kk} \end{bmatrix},$$

missä $A_{ii} \in M_{n_i}, i = 1, \ldots, k, \sum_{i=1}^{k} n_i = n$, ja kaikki lohkodiagonaalin alapuolella olevat lohkot ovat nollalohkoja. Jos lohkodiagonaalit ovat nollalohkoja, niin kyseessä on aito lohkoyläkolmiomatriisi. Vastaavasti matriisi on lohkoalakolmiomatriisi, jos sen transpoosi on lohkoyläkolmiomatriisi, ja aito lohkoalakolmiomatriisi, jos sen transpoosi on aito lohkoyläkolmiomatriisi. Lohkokolmiomatriisi on joko lohkoyläkolmiomatriisi tai lohkoalakolmiomatriisi. Lohkokolmiomatriisi on sekä lohkoala- että lohkoyläkolmiomatriisi, jos ja vain jos se on lohkodiagonaalimatriisi. [9, s.31-32] Merkeillä $\bigstar$ tarkoitetaan mielivaltaisia lohkoja. Jos asiayhteydessä on merkityksettömiä lohkoja, niitä voidaan myös merkitä symbolilla $\bigstar$ sievemmän lopputuloksen toivossa.

3.7 Matriisin jälki ja determinantti

Matriisin jälki ja determinantti ovat tärkeitä skalaareja, jotka kertovat matriisin ominaisuuksista. Näistä matriisin jälki on huomattavasti yksinkertaisempi. Jälki määritellään matriisin diagonaalialkioiden summaksi. Determinantti puolestaan määritellään induktiivisesti alideterminanttien summana tai vaihtoehtoisesti eri permutaatioiden tulona. Jälkeä, determinanttia ja niiden ominaisuuksien tullaan tarvitsemaan eri puolilla työtä, monien lauseiden todistamisessa.

Määritelmä 3.10. Matriisin $A = [a_{ij}] \in M_{m,n}$ jälki on matriisin päädiagonaalialkioiden summa. Tätä merkitään

$$\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{qq} = \sum_{i=1}^{q} a_{ii}, \tag{3.11}$$

missä $q = \min\{m, n\}$.

Lause 3.8. *(Vrt. [13], lause 1.15.3) Olkoon $A \in M_{m,n}$ ja $B \in M_{n,m}$ (kohdassa (b)*

$A, B \in M_{m,n}$*) matriiseja ja* α *jokin kompleksiluku. Tällöin*

(a) $\operatorname{tr} \alpha A = \alpha \operatorname{tr} A$*,*

(b) $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$*,*

(c) $\operatorname{tr} A^T = \operatorname{tr} A$*,*

(d) $\operatorname{tr} A^* = \overline{\operatorname{tr} A}$*,*

(e) $\operatorname{tr} AB = \operatorname{tr} BA$*.*

Todistus. Väitteet seuraavat melko suoraviivaisesti jäljen määritelmästä sekä matriisien ja summien laskusäännöistä.

(a) $\operatorname{tr} \alpha A = \sum_{i=1}^{q} \alpha a_{ii} = \alpha \sum_{i=1}^{q} a_{ii} = \alpha \operatorname{tr} A$.

(b) $\operatorname{tr}(A + B) = \sum_{i=1}^{q} (a_{ii} + b_{ii}) = \sum_{i=1}^{q} a_{ii} + \sum_{i=1}^{q} b_{ii} = \operatorname{tr} A + \operatorname{tr} B$.

(c) Transpoosissa päädiagonaalin alkiot pysyvät samoina, joten $\operatorname{tr} A^T = \operatorname{tr} A$.

(d) $\operatorname{tr} A^* = \operatorname{tr}(\overline{A})^T = \operatorname{tr} \overline{A} = \sum_{i=1}^{q} \overline{a}_{ii} = \overline{\sum_{i=1}^{q} a_{ii}} = \overline{\operatorname{tr} A}$.

(e) Lasketaan aluksi matriisitulot $AB = [\sum_{k=1}^{n} a_{ik} b_{kj}] \in M_m$ ja $BA = [\sum_{h=1}^{m} b_{ih} a_{hj}] \in M_n$. Matriisin jäljen määritelmän mukaan $\operatorname{tr} AB = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} b_{ki}$ ja $\operatorname{tr} BA = \sum_{j=1}^{n} \sum_{h=1}^{m} b_{jh} a_{hj} = \sum_{h=1}^{m} \sum_{j=1}^{n} a_{hj} b_{jh}$, sillä summauksen järjestyksellä ei ole merkitystä. Koska summien indeksien kirjaimet voidaan valita mielivaltaisesti, niin $\operatorname{tr} AB = \operatorname{tr} BA$.

□

Lause 3.9. *(Vrt. [9, s.7], equation 0.2.5.1) Olkoon* $A = [a_{ij}] \in M_{m,n}$ *mielivaltainen matriisi. Tällöin*

$$\operatorname{tr} AA^* = \operatorname{tr} A^* A = \sum_{i,j}^{m,n} |a_{ij}|^2,$$

ja

$$\operatorname{tr} AA^* = 0, \text{ jos ja vain jos } A = 0.$$

Todistus. Matriisitulon määritelmän 3.3 mukaan

$$AA^* = \left[\sum_{k=1}^{n} a_{ik} \overline{a}_{jk} \right].$$

Koska diagonaalialkioiden indekseillä on voimassa $i = j$, niin matriisin AA^* diagonaalialkiot ovat muotoa $\sum_{k=1}^{n} a_{ik} \overline{a}_{ik} = \sum_{k=1}^{n} |a_{ik}|^2$. Matriisitulo $AA^* \in M_m$ on neliömatriisi, joten matriisin jäljen määritelmän mukaan saadaan

$$\operatorname{tr} AA^* = \sum_{i=1}^{m}\sum_{k=1}^{n} |a_{ik}|^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2,$$

missä jälkimmäisessä yhtäsuurudessa vaihdettiin vain indeksikirjain k kirjaimeen j esteettisyyden takia. Edellisen lauseen perusteella $\operatorname{tr} AA^* = \operatorname{tr} A^*A$, joten

$$\operatorname{tr} AA^* = \operatorname{tr} A^*A = \sum_{i,j} |a_{ij}|^2.$$

Tiedetään, että $\operatorname{tr} AA^* = \sum_{i,j} |a_{ij}|^2$ ja $|a_{ij}|^2 \geq 0$ jokaisella $i = 1, \ldots, m$ ja $j = 1, \ldots, n$. Täten jokaisella indeksien i ja j arvolla on voimassa

$$\sum_{i,j} |a_{ij}|^2 = 0 \Leftrightarrow |a_{ij}|^2 = 0 \Leftrightarrow a_{ij} = 0 \Leftrightarrow A = 0,$$

mikä todistaa lauseen jälkimmäisen väitteen. □

Determinantti on määritelty vain neliömatriiseille. Determinantti voidaan määritellä eri tavoilla ja tässä työssä käytetään pääosin Laplacen esitystä. Lausetta 3.16 varten tarvitaan kuitenkin myös vaihtoehtoinen määritelmä, joka esitellään alaluvun lopussa. Alla esitetyssä yhtälössä (3.12) ensimmäinen summa on Laplace laajennus alideterminateilla riviä pitkin ja jälkimmäinen saraketta pitkin [9, s.8].

Määritelmä 3.11. Determinantti määritellään induktiivisesti matriisille $A = [a_{ij}] \in M_n$ seuraavasti. Oletetaan, että determinantti on määritelty joukon M_{n-1} matriiseilla ja merkitään $A_{ij} \in M_{n-1}$ alimatriisia, joka on saatu poistamalla rivi i ja sarake j matriisista A. Tällöin mille tahansa indeksille $i, j \in \{1, \ldots, n\}$ määritellään, että

$$\det A = |A| = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj}. \tag{3.12}$$

Lisäksi asetetaan, että 1×1-kokoisen matriisin determinantti on yksittäisen alkion arvo ja $\det A[\emptyset] = 1$.

Seuraava lause on osana aiempaa lausetta 3.4, mutta se todistetaan vasta nyt. Tämän lauseen avulla on helppo todeta, onko matriisi singulaarinen vai ei. Lause on hyödyllinen myös toiseen suuntaan. Lausetta käytetään esimerkiksi ominaisarvoja määritettäessä.

Lause 3.10. *Olkoon $A \in M_n$. Tällöin matriisi A on ei-singulaarinen, jos ja vain jos $\det A \neq 0$.*

Todistus. (Vrt. [14], Theorem 4.6) Oletetaan, että $\det A \neq 0$. Matriisi A voidaan kirjoittaa RREF-muodossa $R = EA$, missä $R, E \in M_n$ ja E on ei-singulaarinen. Voidaan siis kirjoittaa $A = E^{-1}R$. Lauseen 3.11 kohdan (k) avulla saadaan

$$\det A = \det(E^{-1}R) = \det E^{-1} \det R \neq 0,$$

mikä toteutuu vain, jos $R = I$, koska muulloin $\det R = 0$ kohdan (d) nojalla. Tästä seuraa, että matriisi A on ei-singulaarinen ja $A^{-1} = E$.

Todistetaan väite sitten toiseen suuntaan. Oletetaan, että matriisi A on ei-singulaarinen. Tällöin voidaan kirjoittaa $A = E^{-1}R = E^{-1}I$, joten lauseen 3.11 kohdan (i) ja (k) nojalla

$$\det A = \det(E^{-1}R) = \det(E^{-1}I) = \det E^{-1} \det I = \det E^{-1} \neq 0.$$

□

Alla olevaan lauseeseen on kerätty työn kannalta tärkeitä determinantin ominaisuuksia. Vastaavat ominaisuudet voidaan osoittaa sarakkeille samaan tapaan.

Lause 3.11. *Olkoot $A, B \in M_n$ neliömatriiseja ja α jokin kompleksiluku. Tällöin*

(a) jos matriisin A rivien i ja $j \neq i$ paikkoja vaihdetaan, niin saadun matriisin $E_1 A$ determinantti on $\det(E_1 A) = -\det A$.

(b) jos matriisin A jokin rivi tai sarake kerrotaan skalaarilla α, niin saadun matriisin $E_2 A$ determinantti on $\det(E_2 A) = \alpha \det A$.

(c) jos matriisissa A on kaksi samaa vaakariviä, niin $\det A = 0$.

(d) jos matriisissa A nollarivi, niin $\det A = 0$.

(e) jos matriisin A jokin skalaarilla α kerrottu rivi i lisätään riviin $j \neq i$, niin saadun matriisin $E_3 A$ determinantti on $\det(E_3 A) = \det A$.

(f) $\det A^T = \det A$.

(g) $\det A^* = \overline{\det A}$.

(h) $\det \alpha A = \alpha^n \det A$.

(i) $\det I = 1$.

(j) rivioperaatioita rivienvaihto, skalaarilla kertominen ja rivien yhteenlasku vastaavien alkeismatriisien $E_1, E_2, E_3 \in M_n$ determinantit ovat $\det E_1 = -1, \det E_2 = \alpha$ ja $\det E_3 = 1$.

(k) $\det(EA) = \det E \det A$, missä matriisi $E \in M_n$ on alkeismatriisien tulo. Lisäksi $\det E \neq 0$.

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

(m) jos A on ei-singulaarinen, niin $\det A^{-1} = (\det A)^{-1}$.

Todistus. (Vrt. [14], Theorem 4.3 ja 4.8, sekä [13], lause 1.23.6)

(a) Todistetaan väite induktiolla kaikille matriiseille $A = [a_{ij}] \in M_n$ ja merkitään muunnettua matriisia E_1A, missä E_1 on rivienvaihtoa vastaava alkeismatriisi. Näytetään aluksi, että väite on tosi, kun $n = 2$. Tällöin

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

ja

$$\det(E_1A) = \begin{vmatrix} a_{21} & a_{22} \\ a_{12} & a_{11} \end{vmatrix} = a_{12}a_{21} - a_{11}a_{22} = -\det A.$$

Tehdään induktio-oletus, että väite on tosi, kun $n = p \in \mathbb{Z}$, missä $p \geq 2$. Toisin sanoen oletetaan, että $\det(E_1A) = -\det A$, aina kun $n = p$. Todistetaan tapaus, missä $n = p+1$. Kehitetään determinantti jonkin muun kuin vaihdetun rivin mukaan

$$\begin{aligned} \det(E_1A) &= \sum_{k=1}^{p+1}(-1)^{i+k}a_{ik}\det E_1A_{ik} = \sum_{k=1}^{p+1}(-1)^{i+k}a_{ik}(-\det A_{ik}) \\ &= -\sum_{k=1}^{p+1}(-1)^{i+k}a_{ik}\det A_{ik} = -\det A, \end{aligned}$$

sillä induktio-oletuksen mukaan kaikille alimatriiseille $\det(E_1A_{ik}) = -\det A_{ik}$. Näin ollen induktioperiaatteen nojalla väite on tosi jokaisella $n \geq 2$.

(b) Oletetaan, että rivi i on kerrottu skalaarilla α. Kehitetään muutetun matriisin E_2A determinantti tätä riviä i pitkin. Täten

$$\det(E_2A) = \sum_{k=1}^{n}(-1)^{i+k}\alpha a_{ik}\det A_{ik} = \alpha\sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det A_{ik} = \alpha\det A.$$

(c) Vaihdetaan samanlaisten rivien paikkoja, jolloin matriisi ei muutu. Kuitenkin kohdan (a) nojalla on voimassa $\det A = -\det A$, mikä toteutuu vain kun $\det A = 0$.

(d) Kehitetään determinantti nollarivin suhteen. Tällöin

$$\det A = \sum_{k=1}^{n}(-1)^{i+k}\cdot 0\cdot\det A_{ik} = 0.$$

(e) Oletetaan, että rivi h on kerrottu skalaarilla α ja lisätty riviin i. Kehitetään determinantti muunnetun rivin i suhteen. Täten

$$\begin{aligned}\det(E_3A) &= \sum_{k=1}^{n}(-1)^{i+k}(a_{ik}-\alpha a_{hk})\det A_{ik}\\ &= \sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det A_{ik}-\alpha\sum_{k=1}^{n}(-1)^{i+k}a_{hk}\det A_{ik}\\ &= \det A - 0 = \det A,\end{aligned}$$

sillä jälkimmäinen lauseke on determinantti matriisista, jossa on kaksi samanlaista riviä.

(f) Määritelmän mukaan matriisin transpoosi vaihtaa matriisin rivien ja sarakkeiden paikkoja. Näin ollen jos matriisin $A = [a_{ij}]$ determinantti kehitetään rivin i suhteen saadaan $\det A = \sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det A_{ik}$ ja matriisin $B = [b_{ij}] = A^T = [a_{ji}]$ determinantti kehitetään sarakkeen j suhteen saadaan

$$\begin{aligned}\det B &= \sum_{k=1}^{n}(-1)^{k+j}b_{kj}\det B_{kj} = \sum_{k=1}^{n}(-1)^{k+j}a_{jk}\det A^T_{kj}\\ &= \sum_{k=1}^{n}(-1)^{k+j}a_{jk}\det A_{jk} = \det A.\end{aligned}$$

(g) Koska $\det A^* = \det \overline{A}^T = \det \overline{A}$, riittää osoittaa, että $\det\overline{A} = \overline{\det A}$. Osoitetaan väite determinantin vaihtoehtoisen määritelmän 3.12 avulla. Saadaan

$$\det\overline{A} = \sum_{\sigma}\left(\operatorname{sgn}\sigma\prod_{i=1}^{n}\overline{a}_{i\sigma(i)}\right) = \overline{\sum_{\sigma}\left(\operatorname{sgn}\sigma\prod_{i=1}^{n}a_{i\sigma(i)}\right)} = \overline{\det A}.$$

(h) Soveltamalla kohdan (b) tulosta n kertaa, saadaan haluttu tulos.

(i) Ominaisuus seuraa lauseesta 3.12.

(j) Kohtien (a) ja (i) avulla saadaan, että $\det E_1E_1 = \det I = 1$ ja toisaalta $\det E_1E_1 = -\det E_1$. Täten $\det E_1 = -1$. Kohtien (b) ja (i) avulla saadaan, että $\det E_2^{-1}E_2 = \det I = 1$ ja toisaalta $\det E_2^{-1}E_2 = \frac{1}{\alpha}\det E_2$. Täten $\det E_2 = \alpha$. Kohtien (e) ja (i) avulla saadaan, että $\det E_3^{-1}E_3 = \det I = 1$ ja toisaalta $\det E_3^{-1}E_3 = \det E_3$. Täten $\det E_3 = 1$.

(k) Kohdista (a), (b), (e) ja (j) seuraa, että $\det(E_iA) = \det E_i \det A$, missä E_i on jotain rivioperaatiota vastaava alkeismatriisi. Matriisi E voidaan kirjoittaa alkeismatriisien tulona $E = E_1\cdots E_k$. Kohtaa (j) hyödyntämällä saadaan $\det EA = \det E_1\cdots\det E_k\det A$ ja täten voidaan määritellä

$$\det E = \det E_1\cdots\det E_k \neq 0,$$

sillä $E_i \neq 0$ jokaisella $i \in \{1, \dots, k\}$.

(l) Oletetaan aluksi, että matriisi A on singulaarinen. Tällöin lauseesta 3.6 seuraa, että matriisi AB on singulaarinen ja lauseesta 3.10 että

$$\det(AB) = 0 = 0 \cdot \det B = \det A \det B.$$

Jos matriisi A on ei-singulaarinen, niin se kirjoittaa rivioperaatioita vastaavien alkeismatriisien tulona eli $A = E_A$. Edellistä kohtaa soveltamalla saadaan

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

(m) Koska matriisi A on ei-singulaarinen, niin lauseen 3.10 nojalla $\det A \neq 0$. Kohdista (i) ja (l) seuraa, että $1 = \det I = \det(AA^{-1}) = \det A \det A^{-1}$. Näin ollen $\det A^{-1} = \frac{1}{\det A} = (\det A)^{-1}$.

□

Determinantin laskeminen kolmiomatriiseille on helppoa, sillä determinantti on diagonaalialkioiden tulo. Lause on voimassa myös diagonaalimatriiseille, sillä ne ovat kolmiomatriiseja.

Lause 3.12. *(Vrt. [14], Theorem 4.2) Olkoon $A = [A_{ij}] \in M_n$ kolmiomatriisi. Tällöin $\det A = a_{11} \cdots a_{nn}$.*

Todistus. Oletetaan aluksi, että matriisi A on alakolmiomatriisi. Lähdetään kehittämään matriisin A determinanttia rivien kautta. Ensimmäisen rivin ainoa mahdollisesti nollasta eroava alkio on a_{11} ja alimatriisin $A[\{1\}^c]$ ensimmäisen rivin ainoa mahdollisesti nollasta eroava alkio on a_{22}. Samaan tapaan jatkaen saadaan

$$\begin{aligned}\det(A) &= a_{11} \det A[\{1\}^c] \\ &= a_{11} a_{22} \det A[\{1, 2\}^c] \\ &\vdots \\ &= a_{11} \cdots a_{nn}.\end{aligned}$$

Oletetaan sitten, että A on yläkolmiomatriisi. Tällöin A^T on alakolmiomatriisi ja edellisestä lauseesta seuraa, että $\det A^T = \det A = a_{11} \cdots a_{nn}$. □

Esitellään seuraavaksi determinantin vaihtoehtoinen määritelmä. Määritelmässä permutaatiolla tarkoitetaan bijektiivistä kuvausta $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$. Erilaisia permutaatioita on $n!$ kappaletta. Peruspermutaatiolle on voimassa $\sigma(i) = i$ jokaisella $i = 1, \dots, n$. [9, s.9]

Määritelmä 3.12 (Vaihtoehtoinen määritelmä determinantille)**.** Olkoon $A = [a_{ij}] \in M_n$. Tällöin matriisin A determinantti on

$$\det A = |A| = \sum_{\sigma} \left(\operatorname{sgn} \sigma \prod_{i=1}^{n} a_{i\sigma(i)} \right), \tag{3.13}$$

missä summa otetaan yli kaikkien joukon $\{1, \ldots, n\}$ permutaatioiden ja missä permutaation σ merkkifunktio $\operatorname{sgn} \sigma$ on joko $+1$ tai -1 riippuen siitä, mikä on pienin määrä tehtyjä transpositiota (vaihtoa pareittain), jotta on päästy haluttuun permutaatioon peruspermutaatiosta $\{1, \ldots, n\}$. Jos tehtyjen transpositioiden lukumäärä on parillinen, niin $\operatorname{sgn} \sigma = +1$, ja jos pariton, niin $\operatorname{sgn} \sigma = -1$.

Esimerkki 3.4. Selvennetään vielä edellistä määritelmää laskemalla mielivaltaisen matriisin $A = [a_{ij}] \in M_3$ determinantti. Taulukoidaan summan termit ja niiden laskemiseen tarvittavat tiedot taulukkoon 3.1. Taulukon ensimmäisessä sarakkeessa merkintä $\{1, 2, 3\} \to \{2, 1, 3\}$ tarkoittaa sellaista permutaatiota σ, että $\sigma(1) = 2, \sigma(2) = 1, \sigma(3) = 3$. Ylimmällä rivillä on siis peruspermutaatio.

Taulukko 3.1. *Permutaatiot, niiden merkkifunktion arvot ja summan termit.*

permutaatio σ	$\operatorname{sgn} \sigma$	$\operatorname{sgn} \sigma \prod_{i=1}^{n} a_{i\sigma(i)}$
$\{1, 2, 3\} \to \{1, 2, 3\}$	+1	$+a_{11}a_{22}a_{33}$
$\{1, 2, 3\} \to \{1, 3, 2\}$	-1	$-a_{11}a_{23}a_{32}$
$\{1, 2, 3\} \to \{3, 1, 2\}$	+1	$+a_{13}a_{21}a_{32}$
$\{1, 2, 3\} \to \{3, 2, 1\}$	-1	$-a_{13}a_{22}a_{31}$
$\{1, 2, 3\} \to \{2, 3, 1\}$	+1	$+a_{12}a_{23}a_{31}$
$\{1, 2, 3\} \to \{2, 1, 3\}$	-1	$-a_{12}a_{21}a_{33}$

Näin ollen

$$\begin{aligned}\det A = |A| &= \sum_{\sigma} \left(\operatorname{sgn} \sigma \prod_{i=1}^{n} a_{i\sigma(i)} \right) \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33}.\end{aligned}$$

Esitetään alaluvun lopuksi vielä tärkeä 2×2-kokoisiin lohkomatriiseihin perustuva laskukaava matriisin determinantille. Olkoon $A = [a_{ij}] \in M_n$ ja olkoon $\alpha \subseteq \{1, \ldots, n\}$ sellainen indeksijoukko, että alimatriisi $A[\alpha]$ on ei-singulaarinen. Jos matriisi A ositetaan käyttäen indeksijoukkoja α ja α^c, niin tällöin determinantti voidaan laskea seuraavasti:

$$\det A = \det A[\alpha] \det(A[\alpha^c] - A[\alpha^c, \alpha] A[\alpha]^{-1} A[\alpha, \alpha^c]). \tag{3.14}$$

Tämä on 2×2-kokoisen matriisin determinantin laskukaavan yleistys 2×2-kokoiselle lohkomatriisille. [9, s.24-25]

3.8 Ominaisarvot ja -vektorit

Ominaisarvot ja -vektorit ovat monesti keskeisessä roolissa, kun tutkitaan matriisin ominaisuuksia tai sovelluskohteita. Niiden avulla voidaan esimerkiksi päätellä matriisin singulaarisuus, laskea matriisin jälki tai determinantti. Ominaisarvoja ja -vektoreita tarvitaan myös erilaisissa hajotelmissa, jotka yksinkertaistavat ja helpottavat matriiseilla laskemista. Myöhemmin tullaan hyödyntämään erästä tällaista hajotelmaa, spektraalihajotelmaa, useasti eri lauseita todistaessa. Lisäksi sen avulla voidaan laskea esimerkiksi positiivisesti semidefiniittien matriisien juuria.

Määritelmä 3.13. Olkoon $A \in M_n$ neliömatriisi. Jos luku $\lambda \in \mathbb{C}$ ja nollasta eroava vektori $x \in \mathbb{C}^n$ toteuttavat yhtälön

$$Ax = \lambda x, \tag{3.15}$$

niin tällöin lukua λ sanotaan matriisin A ominaisarvoksi ja vektoria x matriisin A ominaisarvoa λ vastaavaksi ominaisvektoriksi.

Huomionarvoista on, että ominaisarvot ja -vektorit esiintyvät pareittain ja että ominaisvektori ei voi olla nollavektori. Jos nollavektori voisi olla ominaisvektori, niin tällöin jokainen kompleksiluku $\lambda \in \mathbb{C}$ olisi ominaisarvo, sillä $Ax = A0 = 0 = \lambda 0 = \lambda x$. Geometrinen tulkinta ominaisarvoyhtälölle (3.15) on, että matriisin A kertominen ominaisvektorilla x skaalaa ominaisvektoria skalaarilla λ. Jos λ on reaalinen, niin ominaisvektorin pituus muuttuu, mikäli $\lambda \neq 1$ ja vektorin x suunta vaihtuu päinvastaiseksi, jos $\lambda < 0$. Jos ominaisarvo λ on kompleksinen, niin ominaisvektoria kierretään skaalaamisen lisäksi.

Jokaista ominaisarvoa vastaa äärettömän monta ominaisvektoria, sillä jos x on ominaisvektori, niin tällöin αx on myös ominaisvektori jokaisella nollasta eroavalla kompleksiluvulla α. Kaikkien mahdollisten ominaisvektorien muodostamaa joukkoa sanotaan ominaisavaruudeksi. Joskus riittää löytää vain jotkin ominaisavaruuden virittävät vektorit, toisinaan virittäviltä vektoreilta vaaditaan esimerkiksi ortogonaalisuutta. Normeeratut ominaisvektorit ovat käteviä ominaisuutensa $x^*x = 1$ takia.

Ominaisarvojen muodostamaa joukkoa kutsutaan matriisin spektriksi. Esitellään seuraavaksi spektriin liittyviä ominaisuuksia.

Määritelmä 3.14. Matriisin $A \in M_n$ spektri on sen kaikkien ominaisarvojen muodostama joukko. Spektriä merkitään $\sigma(A)$.

Lause 3.13. *(Vrt. [9, s.45], Exercise) Olkoon $A \in M_n$ neliömatriisi, ja olkoon $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$ matriisin A spektri. Tällöin matriisin A kompleksikonjugaatin $\overline{A}$ spektri on*

$$\sigma(\overline{A}) = \overline{\sigma(A)} = \{\overline{\lambda}_1, \dots, \overline{\lambda}_k\}.$$

Todistus. Jos $Ax = \lambda x$, niin tällöin $\overline{A}\overline{x} = \overline{\lambda}\overline{x} \Leftrightarrow \overline{A}y = \overline{\lambda}y$. Näin ollen

$$\lambda \in \sigma(A) \Leftrightarrow Ax = \lambda x \Leftrightarrow \overline{A}y = \overline{\lambda}y \Leftrightarrow \overline{\lambda} \in \sigma(\overline{A}),$$

mistä väittämä seuraa. □

Lause 3.14. *Matriisi $A \in M_n$ on singulaarinen, jos ja vain jos $0 \in \sigma(A)$.*

Todistus. (Vrt. [9], Observation 1.1.7) Matriisi A on singulaarinen, jos ja vain jos $Ax = 0$ jollain nollasta eroavalla vektorilla x. Tämä on voimassa, jos ja vain jos $Ax = 0x$ jollain vektorilla $x \neq 0$ eli jos ja vain jos $\lambda = 0$ on matriisin A ominaisarvo. □

Lause 3.15. *Olkoon $A \in M_n$ ja $\lambda, \mu \in \mathbb{C}$. Tällöin $\lambda \in \sigma(A)$, jos ja vain jos $\lambda + \mu \in \sigma(A + \mu I)$.*

Todistus. (Vrt. [9], Observation 1.1.8) Jos $\lambda \in \sigma(A)$, niin on olemassa sellainen nollasta eroava vektori x, että $Ax = \lambda x$. Näin ollen

$$(A + \mu I)x = Ax + \mu Ix = \lambda x + \mu x = (\lambda + \mu)x,$$

joten $\lambda + \mu \in \sigma(A + \mu I)$. Kääntäen, jos $\lambda + \mu \in \sigma(A + \mu I)$, niin on olemassa sellainen nollasta eroava vektori y, että

$$Ay + \mu y = (A + \mu I)y = (\lambda + \mu)y = \lambda y + \mu y.$$

Edellisestä yhtälöstä saadaan, että $Ay = \lambda y$ ja täten $\lambda \in \sigma(A)$. □

Ominaisarvoyhtälö (3.15) voidaan kirjoittaa

$$(\lambda I - A)x = 0. \tag{3.16}$$

Luku λ on matriisin A ominaisarvo, jos ja vain jos on olemassa jokin sellainen nollasta eroava vektori x, että yhtälö toteutuu. Tämä tarkoittaa, että matriisin $\lambda I - A$ on oltava singulaarinen, mikä toteutuu, jos ja vain jos

$$\det(\lambda I - A) = 0. \tag{3.17}$$

Määritelmä 3.15. Matriisin $A \in M_n$ karakteristinen polynomi on

$$p_A(t) = \det(tI - A).$$

Karakteristisella yhtälöllä tarkoitetaan yhtälöä $p_A(t) = 0$.

Lause 3.16. *Matriisin $A = [a_{ij}] \in M_n$ karakteristinen polynomi on aina astetta n ja on muotoa*

$$p_A(t) = t^n - (\operatorname{tr} A)t^{n-1} + \cdots + (-1)^n \det A. \tag{3.18}$$

Lisäksi $p_A(\lambda) = 0$, jos ja vain jos λ on matriisin A ominaisarvo, joten matriisin spektrissä $\sigma(A)$ on korkeintaan n kompleksilukua.

Todistus. (Vrt. [9], Observation 1.2.4) Määritelmän 3.12 mukaan jokainen matriisin $tI - A$ determinantin summattava termi on tulo, jossa on täsmälleen n matriisin $tI - A$ alkiota. Nämä alkiot ovat kaikki eri riveiltä ja eri sarakkeista, joten jokainen summatava termi on korkeintaan polynomi astetta n. Termin polynomi voi olla astetta n vain, jos jokainen tulontekijä sisältää muuttujan t, mikä tapahtuu vain jos kyseessä on diagonaalialkioiden tulo

$$(t - a_{11}) \cdots (t - a_{nn}) = t^n - (\operatorname{tr} A)t^{n-1} + \cdots . \tag{3.19}$$

Mikä tahansa muu summattava termi sisältää tekijän a_{ij}, missä $i \neq j$, joten samalla rivillä ja samassa sarakkeessa diagonaalialkiot $t - a_{ii}$ ja $t - a_{jj}$ eivät voi olla termissä tulon tekijöinä. Tämän takia muut summattavat termit voivat olla korkeintaan astetta $n-2$. Täten karakteristisessa polynomissa $p_A(t)$ termien t^n ja t^{n-1} kertoimet tulevat pelkästään summan termistä (3.19). Karakteristisen polynomin vakiotermi on $p_A(0) = \det(0I - A) = \det(-A) = (-1)^n \det A$. Väittämä $p_A(\lambda) = 0$ on yhtäpitävä yhtälöiden (3.16) ja (3.17) kanssa. Tiedetään, että astetta $n \geq 1$ olevalla polynomilla on enintään n erisuurta nollakohtaa ja siksi matriisin spektrissä voi olla korkeintaan n kompleksilukua. $\square$

Minkä tahansa matriisin $A \in M_n$, missä $n > 1$, karakteristinen polynomi voidaan kirjoittaa nollakohtien $\alpha_1, \ldots, \alpha_n$ avulla muotoon $p_A(t) = (t - \alpha_1) \cdots (t - \alpha_n)$. Tiedetään, että luvut $\alpha_1, \ldots, \alpha_n$ ovat matriisin A ominaisarvoja riippumatta kertaluvusta. Avaamalla sulut karakteristisesta polynomista saadaan

$$p_A(t) = t^n - (\alpha_1 + \cdots + \alpha_n)t^{n-1} + \cdots + (-1)^n \alpha_1 \ldots \alpha_n. \tag{3.20}$$

Vertaamalla yhtälöitä (3.18) ja (3.20) huomataan, että matriisin A jälki on karakteristisen polynomin nollakohtien summa ja matriisin A determinantti on näiden tulo. Jos jokaisella nollakohdalla on kertaluku 1 eli jos $\alpha_i \neq \alpha_j$ aina kun $i \neq j$, niin tällöin $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, joten $\operatorname{tr} A$ on ominaisarvojen summa ja $\det A$ on ominaisarvojen tulo. Jotta nämä väittämät pysyisivät tosina, on numeroitava ominaisarvot polynomin $p_A(t)$ nollakohtien kertalukujen mukaan. [9, s.50-51]

Määritelmä 3.16. Olkoon $A \in M_n$. Matriisin A ominaisarvon λ (algebrallinen) kertaluku on sitä vastaavan karakteristisen polynomin nollakohdan kertaluku.

Täten, kun huomioidaan ominaisarvojen monikerrat, karakteristinen polynomi voidaan ilmoittaa ominaisarvojen avulla seuraavasti:

$$p_A(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n), \tag{3.21}$$

missä luvut $\lambda_1, \ldots, \lambda_n$ ovat matriisin A ominaisarvot missä tahansa järjestyksessä. Kun viitataan erisuuriin ominaisarvoihin käytetään matriisin spektriä $\sigma(A)$. [9, s.51]

Voidaan todeta, että jokaisella matriisilla $A \in M_n$ on täsmälleen n ominaisarvoa kompleksilukujen joukossa Algebran peruslauseen B nojalla. Lisäksi tärkeät yhteydet matriisin jäljen ja ominaisarvojen sekä determinantin ja ominaisarvojen välillä on koottu alla olevaan lauseeseen.

Lause 3.17. *Olkoon $A \in M_n$. Tällöin $\operatorname{tr} A = \lambda_1 + \cdots + \lambda_n$ ja $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$, missä $\lambda_1, \ldots, \lambda_n$ ovat matriisin A ominaisarvot.*

Todistus. Tulos osoitettu edellä. □

Määritelmä 3.17. Olkoon $A \in M_n$ ja $\lambda \in \sigma(A)$. Kaikkien ominaisarvoyhtälön $Ax = \lambda x$ toteuttavien vektorien $x \in \mathbb{C}^n$ joukkoa sanotaan ominaisarvoa λ vastaavaksi ominaisavaruudeksi. Tämän ominaisavaruuden dimensio on ominaisarvon λ geometrinen kertaluku.

Lause 3.18. *Olkoot $\lambda_1, \ldots, \lambda_k$ matriisin $A \in M_n$ erisuuria ominaisarvoja. Tällöin näitä ominaisarvoja vastaavat ominaisvektorit $x_1, \ldots, x_k$ ovat lineaarisesti riippumattomia.*

Todistus. (Vrt. [9], Lemma 1.3.8) Oletetaan, että on olemassa sellaiset kompleksiluvut $\alpha_1, \ldots, \alpha_k$, että $\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$. Olkoon $B_1 = (A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_k I)$. Tästä matriisien tulosta on tarkoituksella jätetty pois termi $A - \lambda_1 I$. Koska ominaisvektori x_i vastaa ominaisarvoa λ_i, niin saadaan

$$\begin{aligned} B_1 x_i &= (A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_k I) x_i = (A - \lambda_2 I)(A - \lambda_3 I) \cdots (A x_i - \lambda_k x_i) \\ &= (A - \lambda_2 I)(A - \lambda_3 I) \cdots (\lambda_i x_i - \lambda_k x_i) = (A - \lambda_2 I)(A - \lambda_3 I) \cdots (\lambda_i - \lambda_k) x_i \\ &= (\lambda_i - \lambda_2)(\lambda_i - \lambda_3) \cdots (\lambda_i - \lambda_k) x_i. \end{aligned}$$

Vektori $B_1 x_i$ on tulon nollasäännön nojalla nollavektori, jos $2 \leq i \leq k$. Jos $i = 1$, niin vektori $B_1 x_1$ ei ole nollavektori, sillä $\lambda_1 \neq \lambda_j$ jokaisella $j = 2, \ldots, k$ ja $x_1 \neq 0$. Täten

$$\begin{aligned} 0 &= B_1(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k) \\ &= \alpha_1 B_1 x_1 + \alpha_2 B_1 x_2 + \cdots + \alpha_k B_1 x_k \\ &= \alpha_1 B_1 x_1 + 0 + \cdots + 0 = \alpha_1 B_1 x_1, \end{aligned}$$

mikä takaa, että $\alpha_1 = 0$, sillä $B_1 x_1 \neq 0$. Jokaisella $j = 2, \ldots, k$ voidaan määritellä matriisi B_j vastaavasti kuin määriteltiin matriisi B_1, mutta jättämällä matriisien tulosta termi $A - \lambda_j I$ pois termin $A - \lambda_1 I$ sijaan. Näin saadaan, että $\alpha_j = 0$ jokaisella j, joten $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$, eli vektorit $x_1, \ldots, x_k$ ovat lineaarisesti riippumattomia. $\square$

Kolmiomatriisin ominaisarvojen selvittäminen on yksinkertaista, sillä ne ovat kolmiomatriisin diagonaalialkiot. Lisäksi diagonaalimatriiseille ominaisvektorien selvittäminen on triviaalia.

Lause 3.19. *(Vrt. [14], Theorem 4.15) Olkoon $A = [a_{ij}] \in M_n$ kolmiomatriisi. Tällöin matriisin A ominaisarvot ovat matriisin diagonaalialkiot.*

Todistus. Oletetaan, että matriisi A on kolmiomatriisi. Tällöin matriisi $\lambda I - A$ on myös kolmiomatriisi. Lauseesta 3.12 seuraa, että $\det(\lambda I - A) = (\lambda_1 - a_{11}) \cdots (\lambda_n - a_{nn})$. Asettamalla $\det(\lambda I - A) = 0$ on saadaan, että ominaisarvot ovat diagonaalialkiot. $\square$

Lause 3.20. *(Vrt. [9, s.44], Exercise) Diagonaalimatriisin $D \in M_n$ ominaisarvot ovat sen diagonaalialkiot ja ominaisvektorit ovat diagonaalimatriisin sarakkeet.*

Todistus. Koska diagonaalimatriisi on kolmiomatriisi, niin edellisen lauseen perusteella sen ominaisarvot ovat diagonaalialkioita. Kirjoitetaan diagonaalimatriisi pystyvektoreiden avulla

$$D = \begin{bmatrix} d_1 & \ldots & d_n \end{bmatrix} = \begin{bmatrix} \lambda_1 e_1 & \ldots & \lambda_n e_n \end{bmatrix},$$

missä vektorin e_i i:s alkio on 1 ja muut alkiot ovat nollia. Tällöin e_i on ominaisarvoa λ_i vastaava ominaisvektori, sillä $De_i = d_i = \lambda_i e_i$. Myös vektorien e_i monikerrat $d_i = \lambda_i e_i$ ovat ominaisvektoreita. $\square$

3.9 Similaarisuus

Similaarisuuden tutkimisella, englanniksi similarity eli samankaltaisuus, viitataan nimenmukaisesti samankaltaisten matriisien tai lineaarikuvausten ja niiden ominaisuuksien tutkimiseen. Similaarisuus on ekvivalenssirelaatio, joka osittaa neliömatriisien muodostaman joukon M_n, kuten tullaan huomaamaan. Tärkeä huomio on, että similaarisilla matriiseilla on sama karakteristinen polynomi, mistä seuraa muita ominaisuuksia, kuten samat ominaisarvot, jäljet ja determinantit. Seuraavassa luvussa tutustutaan myös unitaariseen similaarisuuteen, joka on tavallista similaarisuutta vielä vahvempi ominaisuus.

Määritelmä 3.18. Olkoot matriisit $A, B \in M_n$. Matriisi B on similaarinen matriisin A kanssa, jos on olemassa sellainen ei-singulaarinen matriisi $S \in M_n$, että

$$B = S^{-1} AS.$$

Relaatio "matriisi B on similaarinen matriisin A kanssa" voidaan lyhentää merkinnällä $B \sim A$. Jos matriisi B on similaarinen diagonaalimatriisin kanssa, sanotaan, että matriisi B on diagonalisoituva.

Esimerkki 3.5. Näytetään, että similaarisuus on ekvivalenssirelaatio joukossa M_n. Toisin sanoen similaarisuus on refleksiivinen, symmetrinen ja transitiivinen relaatio. Olkoon $A, B, C \in M_n$. Similaarisuus on refleksiivinen, sillä $A = I^{-1}AI$, eli $A \sim A$ jokaisella $A \in M_n$. Similaarisuus on symmetrinen, sillä aina kun $B \sim A$, eli $B = S^{-1}AS$, niin $A = SBS^{-1} = T^{-1}BT$, eli $A \sim B$. Lopuksi, similaarisuus on transitiivinen, sillä aina kun $A \sim B$ ja $B \sim C$, niin $A = S^{-1}BS = S^{-1}(T^{-1}CT)S = (TS)^{-1}C(TS)$, eli $A \sim C$.

Kuten mikä tahansa ekvivalenssirelaatio, similaarisuus osittaa joukon M_n erillisiin ekvivalenssiluokkiin. Jokainen ekvivalenssiluokka on joukko kaikista niistä joukon M_n matriiseista, jotka ovat similaarisia ekvivalenssiluokkaa edustavan matriisin kanssa. Jokainen ekvivalenssiluokan matriisi on siis similaarinen muiden luokan matriisien kanssa, mutta ei minkään muun ekvivalenssiluokan matriisin kanssa. Keskeinen havainto on, että saman luokan matriiseilla on monia yhteisiä ominaisuuksia. [9, s.58]

Lause 3.21. *Olkoot $A, B \in M_n$. Jos matriisi B on similaarinen matriisin A kanssa, niin tällöin matriiseilla A ja B on sama karakteristinen polynomi.*

Todistus. (Vrt. [9], Theorem 1.3.3) Lasketaan

$$\begin{aligned} p_B(t) &= \det(tI - B) = \det(tS^{-1}S - S^{-1}AS) = \det(S^{-1}(tI - A)S) \\ &= \det(S^{-1})\det(tI - A)\det S = \frac{1}{\det S}\det(tI - A)\det S \\ &= \det(tI - A) = p_A(t). \end{aligned}$$

□

Lause 3.22. *(Vrt. [9], Corollary 1.3.4) Olkoot $A, B \in M_n$ ja olkoon matriisi A similaarinen matriisin B kanssa. Tällöin*

(a) matriiseilla A ja B on samat ominaisarvot.

(b) jos B on diagonaalimatriisi, niin sen diagonaalialkiot ovat matriisin A ominaisarvot.

(c) matriisien A ja B determinantit ja jäljet ovat yhtäsuuret.

(d) matriisi $B = O$, jos ja vain jos matriisi $A = O$.

(e) matriisi $B = I$, jos ja vain jos matriisi $A = I$.

Todistus. Edellisestä lauseesta seuraa, että matriiseilla A ja B on sama karakteristinen polynomi.

(a) Karakteristinen polynomi voidaan kirjoittaa ominaisarvojen avulla, kuten yhtälössä (3.21). Väite seuraa tästä.

(b) Diagonaalimatriisin ominaisarvot ovat sen diagonaalialkiot lauseen 3.20 nojalla, joten väite seuraa (a)-kohdasta.

(c) Lauseen 3.17 mukaan determinantti ja jälki voidaan kirjoittaa ominaisarvojen avulla. Täten väite seuraa (a)-kohdasta.

(d) Oletuksen perusteella voidaan kirjoittaa $A = S^{-1}BS \Leftrightarrow SAS^{-1} = B$. Jos $B = 0$, niin $A = S^{-1}0S = 0$. Jos $A = 0$, niin $B = S0S^{-1} = 0$.

(e) Jos $B = I$, niin $A = S^{-1}IS = S^{-1}SI = I$. Jos $A = I$, niin $B = SIS^{-1} = ISS^{-1} = I$.

□

Lause 3.23. *(Vrt. [14], Theorem 4.27) Matriisi $A \in M_n$ on diagonalisoituva, jos ja vain jos jokaisen ominaisarvon $\lambda \in \sigma(A)$ algebrallinen ja geometrinen kertaluku ovat yhtäsuuret.*

Todistus. Oletetaan aluksi, että matriisi A on diagonalisoituva. Tällöin voidaan kirjoittaa $S^{-1}AS = \Lambda$, missä $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Ositetaan matriisi $S = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}$ sarakevektoreiden mukaan. Yhtälöstä $AS = S\Lambda$ voidaan päätellä, että jokainen sarake x_j on ominaisarvoa λ_j vastaava ominaisvektori. Koska matriisi S on ei-singulaarinen, niin matriisin sarakevektorit ovat lineaarisesti riippumattomia. Näin ollen jokaista erisuurta ominaisarvoa vastaa täsmälleen ominaisarvon algebrallisen kertaluvun, eli kuinka monta kertaa kyseinen ominaisarvo esiintyy matriisissa Λ, verran lineaarisesti riippumattomia ominaisvektoreita. Geometrinen kertaluku on ominaisavaruuden dimensio, eli ominaisarvoa vastaavien lineaarisesti riippumattomien ominaisvektorien lukumäärä. Toisin sanoen jokaisen ominaisarvon geometrinen kertaluku on yhtäsuuri kuin ominaisarvon algebrallinen kertaluku.

Oletetaan sitten, että jokaisen ominaisarvon λ_i algebrallinen ja geometrinen kertaluku ovat yhtäsuuret. Ominaisarvoyhtälöt $Ax_i = \lambda_i x_i$ voidaan kirjoittaa $AS = S\Lambda$, missä $S = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}$ ja $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Koska oletuksen nojalla jokaista ominaisarvoa vastaa sen algebrallisen kertaluvun verran lineaarisesti riippumattomia ominaisvektoreita ja koska eri ominaisarvoja vastaavat ominaisvektorit ovat lineaarisesti riippumattomia lauseen 3.18 nojalla, niin matriisi S on ei-singulaarinen. Näin ollen $S^{-1}AS = \Lambda$, joten matriisi A on diagonalisoituva. □

4. NORMAALIT MATRIISIT

Normaalien matriisien luokka on joukko, joka muodostuu matriiseista $A \in M_n$, jotka toteuttavat ehdon $AA^* = A^*A$. Tämä luokka on tärkeä joka puolella matriisianalyysia. Se nousee luontaisesti esiin unitaarisen similaarisuuden yhteydessä. Normaalien matriisien luokkaan kuuluvat unitaariset, hermiittiset, vinohermiittiset, reaaliortogonaaliset, reaalisymmetriset ja reaalivinosymmetriset matriisit. [9, s.131] Tässä luvussa tutustutaan erityisesti unitaarisiin, hermiittisiin ja positiivisesti (semi)definiitteihin matriiseihin, jotka ovat hermiittisten matriisien alaluokka. Lisäksi käsitellään unitaarista similaarisuutta ja tullaan muun muassa huomaamaan, että jokainen normaali matriisi on unitaarisesti diagonalisoituva.

Määritelmä 4.1. Neliömatriisi $A \in M_n$ on normaali matriisi, jos se kommutoi oman konjugaattitranspoosinsa kanssa. Toisin sanoen matriisi A on normaali matriisi, jos

$$AA^* = A^*A. \tag{4.1}$$

Normaalin matriisin määritelmän ominaisuudelle (4.1) voidaan antaa myös geometrinen tulkinta, joka voi auttaa lukijaa ymmärtämään ominaisuuden paremmin. Ositetaan matriisit $A \in M_n$ ja $A^T \in M_n$ sarakkeiden mukaan seuraavasti: $A = \begin{bmatrix} c_1 & \dots & c_n \end{bmatrix}$ ja $A^T = \begin{bmatrix} r_1 & \dots & r_n \end{bmatrix}$, missä vektorit c_j ovat matriisin A sarakkeita ja vektorit r_i^T ovat matriisin A rivejä. Näin ollen

$$A^*A = \begin{bmatrix} \overline{c}_1 \\ \vdots \\ \overline{c}_n \end{bmatrix} \begin{bmatrix} c_1 & \dots & c_n \end{bmatrix} = [c_i^* c_j]$$

ja

$$AA^* = \begin{bmatrix} r_1^T \\ \vdots \\ r_n^T \end{bmatrix} \begin{bmatrix} \overline{r}_1 & \dots & \overline{r}_n \end{bmatrix} = [r_i^T \overline{r}_j] = \overline{[r_i^* r_j]}.$$

Tarkastelemalla alkiokohtaisesti matriisiyhtälöä $A^*A = AA^*$ huomataan, että matriisi A on normaali, jos ja vain jos $c_i^* c_j = \overline{r_i^* r_j}$ jokaisella $i, j = 1, \dots, n$. Erityisesti $c_i^* c_i =$

$\|c_i\|_2^2 = \|r_i\|_2^2 = r_i^* r_i$, joten jokaisen rivin k euklidinen normi on yhtäsuuri kuin vastaavan sarakkeen k euklidinen normi. Lisäksi, jos k. rivi on nollarivi, niin tällöin k. sarake on nollasarake. [9, s.132]

Lause 4.1. *(Vrt. [9, s.131], Exercise) Matriisit A ja B ovat normaaleja, jos ja vain jos suora summa $A \oplus B$ on normaali matriisi.*

Todistus. Olkoon $A \in M_n$ ja $B \in M_m$ matriiseja. Nyt

$$(A \oplus B)(A \oplus B)^* = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^* = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} A^* & 0 \\ 0 & B^* \end{bmatrix} \tag{4.2}$$

$$= \begin{bmatrix} AA^* & 0 \\ 0 & BB^* \end{bmatrix} \in M_{n+m} \tag{4.3}$$

ja

$$(A \oplus B)^*(A \oplus B) = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^* \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} A^* & 0 \\ 0 & B^* \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \tag{4.4}$$

$$= \begin{bmatrix} A^*A & 0 \\ 0 & B^*B \end{bmatrix} \in M_{n+m}. \tag{4.5}$$

Jos matriisit A ja B ovat normaaleja, niin $AA^* = A^*A$ ja $BB^* = B^*B$. Tästä seuraa yhtä suuruus yhtälöiden (4.3) ja (4.5) välille, mikä osoittaa, että $A \oplus B$ on normaali. Jos $A \oplus B$ on normaali edellä mainittujen yhtälöiden välillä on yhtäsuuruus, josta seuraa, että on oltava $AA^* = A^*A$ ja $BB^* = B^*B$. Näin ollen matriisit A ja B ovat normaaleja. $\square$

Apulause 4.1. *Olkoon matriisi A ositettu seuraavasti: $A = \left[\begin{smallmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{smallmatrix}\right]$, missä A_{11} ja A_{22} ovat neliömatriiseja. Tällöin matriisi A on normaali, jos ja vain jos A_{11} ja A_{22} ovat normaaleja sekä $A_{12} = 0$. Lohkoyläkolmiomatriisi on normaali, jos ja vain jos jokainen matriisin diagonaalilohko on normaali ja muut kuin diagonaalilohkot ovat nollalohkoja. Erityisesti, yläkolmiomatriisi on normaali, jos ja vain jos se on diagonaalimatriisi.*

Todistus. (Vrt. [9], Lemma 2.5.2) Jos matriisit A_{11} ja A_{22} ovat normaaleja ja $A_{12} = 0$, niin matriisi $A = A_{11} \oplus A_{22}$ on normaalien matriisien suora summa, joka on normaali lauseen 4.1 nojalla. Kääntäen, jos matriisi A on normaali, niin

$$AA^* = \begin{bmatrix} A_{11}A_{11}^* + A_{12}A_{12}^* & \bigstar \\ \bigstar & \bigstar \end{bmatrix} = \begin{bmatrix} A_{11}^*A_{11} & \bigstar \\ \bigstar & \bigstar \end{bmatrix} = A^*A.$$

Täten on oltava $A_{11}^*A_{11} = A_{11}A_{11}^* + A_{12}A_{12}^*$, joten myös näiden matriisien jälkien on oltava yhtä suuret. Hyödyntämällä jäljen ominaisuuksia lauseesta 3.8 saadaan

$$\begin{aligned}\operatorname{tr} A_{11}^*A_{11} &= \operatorname{tr}(A_{11}A_{11}^* + A_{12}A_{12}^*)\\ &= \operatorname{tr} A_{11}A_{11}^* + \operatorname{tr} A_{12}A_{12}^*\\ &= \operatorname{tr} A_{11}^*A_{11} + \operatorname{tr} A_{12}A_{12}^*,\end{aligned}$$

josta saadaan $\operatorname{tr} A_{12}A_{12}^* = 0$. Lauseen 3.9 nojalla yhtälöstä $\operatorname{tr} A_{12}A_{12}^* = 0$ seuraa, että $A_{12} = 0$. Normaali matriisi A on siis matriisien A_{11} ja A_{22} suora summa ja lauseen 4.1 nojalla matriisit A_{11} ja A_{22} ovat normaaleja.

Olkoon $B = [B_{ij}]_{i,j=1}^k \in M_n$ normaali lohkoyläkolmiomatriisi. Toisin sanoen tälle matriisille $B_{ii} \in M_{n_i}$ jokaisella $i = 1, \dots, k$ ja $B_{ij} = 0$, jos $i > j$. Matriisi B voidaan osittaa seuraavasti: $B = \left[\begin{smallmatrix} B_{11} & X \\ 0 & \tilde{B}\end{smallmatrix}\right]$, missä $X = [B_{12} \quad \dots \quad B_{1k}]$ ja $\tilde{B} = [B_{ij}]_{i,j=2}^k$ on lohkoyläkolmiomatriisi. Aiemman perusteella $X = 0$ ja matriisit B_{11} ja $\tilde{B}$ ovat normaaleja. Tämä sama prosessi voidaan toistaa $k-1$ kertaa lohkoyläkolmiomatriisille $\tilde{B}$, mistä voidaan päätellä, että kaikki diagonaalilohkot ovat normaaleja ja muut lohkot nollalohkoja. Päinvastoin, jos lohkoyläkolmiomatriisin $B = [B_{ij}] \in M_n$ diagonaalilohkot ovat normaaleja ja muut lohkot ovat nollalohkoja, niin matriisi voidaan kirjoittaa suoran summan avulla $B = \bigoplus_{i=1}^{k} B_{ii}$. Soveltamalla lausetta 4.1 voidaan osoittaa, että matriisi B on tällöin normaali. Viimeinen väite seuraa tästä, kun asetetaan lohkojen koko olemaan 1×1. □

Lause 4.2. *(Vrt. [9], Problem 2.5.P21 ja 2.5.P54) Olkoon matriisi $A \in M_n$ normaali. Tällöin matriisien A ja A^* nolla- ja sarakeavaruudet ovat samat.*

Todistus. Todistetaan aluksi nolla-avaruutta koskeva väite. Normin määritelmen 2.2 mukaan $Ax = 0 \Leftrightarrow \|Ax\|^2 = 0$. Koska

$$\begin{aligned}\|Ax\|^2 &= \langle Ax, Ax\rangle = (Ax)^*(Ax) = x^*A^*Ax = x^*AA^*x\\ &= (A^*x)^*(A^*x) = \langle A^*x, A^*x\rangle = \|A^*x\|^2,\end{aligned}$$

niin yhtälöillä $A^*x = 0$ ja $Ax = 0$ on samat ratkaisut ja täten sama nolla-avaruus.

Todistetaan sitten, että matriisien A ja A^* sarakeavaruudet ovat samat. Tätä varten tarvitsee aluksi osoittaa tulos, jonka mukaan avaruudet $\mathcal{R}(A) + \mathcal{N}(A)$ ja $\mathbb{C}^n$ ovat samat. Jos $Ax \in \mathcal{R}(A)$ ja $y \in \mathcal{N}(A)$, niin

$$\langle Ax, y\rangle = y^*(Ax) = (A^*y)^*x = \langle x, A^*y\rangle = \langle x, 0\rangle = 0,$$

sillä edellä todettiin, että $Ay = 0 = A^*y$. Näin ollen matriisin A sarake- ja nolla-avaruus ovat keskenään ortogonaaliset. Ortogonaalisuudesta ja dimensiolauseesta 3.5 seuraa, että $\mathcal{R}(A) + \mathcal{N}(A) = \text{span}\,\{\mathcal{R}(A) \cup \mathcal{N}(A)\} = \mathbb{C}^n$. Tämä tulos voidaan osoittaa myös matriisille A^*.

Olkoon $Ax \in \mathcal{R}(A)$. Koska matriisien A ja A^* nolla-avaruudet ovat samat ja $x \in \mathbb{C}^n$, niin aiemman perusteella se voidaan kirjoittaa muodossa $x = y + z$, missä $y \in \mathcal{R}(A^*)$ ja $z \in \mathcal{N}(A)$. On siis olemassa sellainen vektori $c \in \mathbb{C}^n$, että $A^*c = y \Leftrightarrow AA^*c = Ay$. Nyt

$$Ax = A(y+z) = Ay + Az = Ay + 0 = Ay = AA^*c = A^*Ac,$$

joten $Ax \in \mathcal{R}(A^*)$ ja $\mathcal{R}(A) \subseteq \mathcal{R}(A^*)$. Vastaavasti voidaan osoittaa, että $\mathcal{R}(A^*) \subseteq \mathcal{R}(A)$, mistä seuraa, että matriisien A ja A^* sarakeavaruudet ovat samat. $\square$

Apulause 4.2. *(Vrt. [2, s.198], Exercise 4) Olkoon matriisi $A \in M_n$ on normaali. Tällöin jos $Ax = \lambda x$, niin $A^*x = \overline{\lambda}x$, missä λ on matriisin A ominaisarvo ja x sitä vastaava ominaisvektori.*

Todistus. Ominaisarvoyhtälö voidaan kirjoittaa muodossa $(A - \lambda I)x = 0$, missä vektori $x \neq 0$. Laskemalla voidaan todeta, että matriisi $A - \lambda I$ on normaali. Täten lauseen 4.2 nojalla

$$(A - \lambda I)x = 0 \Leftrightarrow (A - \lambda I)^*x = 0 \Leftrightarrow (A^* - \overline{\lambda} I)x = 0 \Leftrightarrow A^*x = \overline{\lambda}x.$$

$\square$

Lause 4.3. *(Vrt. [2, s.198], Exercise 5) Olkoon matriisi $A \in M_n$ normaali. Tällöin matriisin erisuuria ominaisarvoja λ ja μ vastaavat ominaisvektorit $x \in \mathbb{C}^n$ ja $y \in \mathbb{C}^n$ ovat ortogonaaliset.*

Todistus. Oletetaan, että λ ja μ ovat matriisin A erisuuria ominaisarvoja sekä x ja y näitä ominaisarvoja vastaavat ominaisvektorit tässä järjestyksessä. Tällöin apulauseen 4.2 mukaan $Ax = \lambda x \Rightarrow A^*x = \overline{\lambda}x$ ja $Ay = \mu y \Rightarrow A^*y = \overline{\mu}y$. Lasketaan

$$\langle Ax, y\rangle = y^*(Ax) = y^*(\lambda x) = \lambda(y^*x) = \lambda\langle x, y\rangle$$

ja toisaalta

$$\langle Ax, y\rangle = y^*(Ax) = (y^*A)x = (A^*y)^*x = (\overline{\mu}y)^*x = \mu(y^*x) = \mu\langle x, y\rangle.$$

Koska oletuksen perusteella $\mu \neq \lambda$, niin yhtälö $\mu\langle x, y\rangle = \lambda\langle x, y\rangle$ toteutuu vain, jos $\langle x, y\rangle = 0$, eli jos ominaisvektorit x ja y ovat ortogonaaliset. $\square$

4.1 Unitaariset ja unitaarisesti similaariset matriisit

Reaaliortogonaaliset matriisit toteuttavat ehdon $Q^TQ = I$. Unitaarisille matriiseille $U^*U = I$. Unitaarisia matriiseja voidaan siis ajatella ikään kuin reaaliortogonaalisten matriisien yleistyksenä kompleksisten matriisien joukkoon. Unitaariset matriisit ovat normaaleja matriiseja ja niille voidaan antaa vastaavanlainen geometrinen tulkinta. Jos unitaarinen matriisi ositetaan riveihin tai sarakkeisiin u_i, niin $u_i^*u_j = 1$, jos $i = j$, ja $u_i^*u_j = 0$, jos $i \neq j$. Täten jokaisen rivivektorin euklidinen normi on 1, joten yksikään matriisin $U = [u_{ij}]$ alkio ei voi olla itseisarvoltaan suurempi kuin 1. Lisäksi kaksi eri rivivektoria ovat aina keskenään ortogonaaliset. Rivivektorit muodostavat siis ortonormaalin joukon. Vastaava ominaisuus on voimassa myös sarakevektoreille. Tässä alaluvussa käsitellään unitaaristen matriisien lisäksi unitaarista similaarisuutta ja todistetaan tärkeät lauseet 4.8 ja 4.9 koskien Schurin hajotelmaa ja normaalien matriisien spektraalihajotelmaa.

Määritelmä 4.2. Neliömatriisi $U \in M_n$ on unitaarinen, jos $U^*U = I$. Matriisi $U \in M_n(\mathbb{R})$ on ortogonaalinen, jos $U^TU = I$.

Lause 4.4. *Olkoon matriisi $U \in M_n$. Tällöin seuraavat väitteet ovat yhtäpitäviä keskenään:*

(a) Matriisi U on unitaarinen.

(b) Matriisi U on ei-singulaarinen ja $U^* = U^{-1}$.

(c) $UU^* = I$.

(d) Matriisi U^* on unitaarinen.

(e) Matriisin U sarakevektoreiden muodostama joukko on ortonormaali.

(f) Matriisin U rivivektoreiden muodostama joukko on ortonormaali.

Todistus. (Vrt. [9], Theorem 2.1.4) Osoitetaan, että $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d)$. Väittämästä (a) seuraa väittämä (b), sillä jos U^{-1} on olemassa, niin U^{-1} on yksikäsitteinen ja käänteismatriisin määritelmän mukaan U^* on matriisin U käänteismatriisi, sillä matriisi U on unitaarinen, eli $U^*U = I$. Näin ollen U on ei-singulaarinen ja $U^* = U^{-1}$. Käänteismatriisin määritelmä kertoo myös, että jos U on neliömatriisi ja $U^{-1}U = I$, niin tällöin $UU^{-1} = I$, joten väittämästä (b) seuraa väittämä (c). Koska $(U^*)^* = U$, niin väittämä (c) voidaan kirjoittaa $UU^* = (U^*)^*U^* = I$ ja täten matriisi U^* on määritelmän mukaan unitaarinen. Vastaavasti voidaan osoittaa $(d) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a)$, mikä todistaa väittämät (a)-(d) yhtäpitäviksi.

Ositetaan matriisi U sarakkeiden mukaan muotoon $U = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}$. Tällöin $U^*U = I$, mikä tarkoittaa, että $u_i^*u_i = 1$ jokaisella $i = 1, \dots, n$ ja $u_i^*u_j = 0$ aina, kun $i \neq j$. Näin ollen $U^*U = I$ on toinen tapa ilmaista, että matriisin U sarakevektoreiden muodostama joukko on ortonormaali. Täten väitteet (a) ja (e) ovat yhtäpitäviä. Vastaavasti voidaan osoittaa väitteiden (d) ja (f) yhtäpitävyys. □

Lause 4.5. *(Vrt. [9], Observation 2.1.6) Olkoot matriisit $U, V \in M_n$ unitaarisia. Tällöin matriisien tulo UV on unitaarinen matriisi.*

Todistus. Hyödyntämällä oletusta $U^*U = I = V^*V$ saadaan, että matriisien tulo

$$(UV)^*UV = V^*U^*UV = V^*IV = V^*V = I \in M_n$$

on unitaarinen määritelmän mukaan. □

Määritelmä 4.3. Olkoot $A, B \in M_n$ neliömatriiseja. Matriisi A on unitaarisesti similaarinen matriisin B kanssa, jos on olemassa sellainen unitaarinen matriisi U, että $A = UBU^*$. Matriisi A on unitaarisesti diagonalisoituva, jos se on unitaarisesti similaarinen diagonaalimatriisin kanssa.

Lause 4.6. *(Vrt. [9, s.131], Exercise) Olkoot $A \in M_n$ ja $B \in M_n$. Lisäksi olkoon matriisi A normaali ja unitaarisesti similaarinen matriisin B kanssa. Tällöin matriisi B on normaali.*

Todistus. Koska matriisi A on unitaarisesti similaarinen matriisin B kanssa, voidaan kirjoittaa $A = UBU^*$, missä matriisi U on unitaarinen. Matriisi A on normaali, joten

$$AA^* = UBU^*(UBU^*)^* = UBU^*UB^*U^* = UBIB^*U^* = UBB^*U^*$$

ja

$$A^*A = (UBU^*)^*UBU^* = UB^*U^*UBU^* = UB^*IBU^* = UB^*BU^*.$$

Näin ollen

$$\begin{aligned}
& AA^* = A^*A \\
\Leftrightarrow\quad & UBB^*U^* = UB^*BU^* && |U^*\cdot \\
\Leftrightarrow\quad & U^*UBB^*U^* = U^*UB^*BU^* && |\cdot U \\
\Leftrightarrow\quad & IBB^*U^*U = IB^*BU^*U \\
\Leftrightarrow\quad & BB^*I = B^*BI \\
\Leftrightarrow\quad & BB^* = B^*B.
\end{aligned}$$

Täten määritelmän 4.1 mukaan matriisi B on normaali. □

Vastaavasti kuin similaarisuuden kanssa voidaan osoittaa, että unitaarinen similaarisuus on ekvivalenssirelaatio. Unitaarisesta similaarisuudesta seuraa similaarisuus, mutta päinvastainen ei ole voimassa, ja lisäksi unitaarinen similaarisuus jakaa joukon M_n hienompiin ekvivalenssiluokkiin kuin similaarisuus [9, s.95]. Näin ollen similaarisuudesta seuraavat ominaisuudet ovat voimassa myös unitaarisesti similaarisilla matriiseilla. Mainittakoon erityisesti lauseen 3.22 ominaisuudet.

Lause 4.7. *Olkoot matriisit U ja V unitaarisia, olkoot $A = [A_{ij}], B = [b_{ij}] \in M_{n,m}$ ja oletetaan, että $A = UBV$. Tällöin $\sum_{i,j=1}^{n,m} |b_{ij}|^2 = \sum_{i,j=1}^{n,m} |a_{ij}|^2$. Nimenomaan tämä ominaisuus toteutuu, jos $m = n$ ja $V = U^*$ ja täten matriisi A on unitaarisesti similaarinen matriisin B kanssa.*

Todistus. (Vrt. [9], Theorem 2.2.2) Hyödynnetään lausetta 3.9, jolloin riittää osoittaa, että $\mathrm{tr}(A^*A) = \mathrm{tr}(B^*B)$. Nyt

$$\begin{aligned}\mathrm{tr}(A^*A) &= \mathrm{tr}((UBV)^*(UBV)) = \mathrm{tr}(V^*B^*U^*UBV) \overset{1}{=} \mathrm{tr}(V^*B^*BV) = \mathrm{tr}((BV)^*BV) \\ &\overset{2}{=} \mathrm{tr}(BV(BV)^*) = \mathrm{tr}(BVV^*B^*) \overset{1}{=} \mathrm{tr}(BB^*) \overset{2}{=} \mathrm{tr}(B^*B),\end{aligned}$$

mikä todistaa lauseen. Kohdissa 1 hyödynnettiin unitaarisen matriisin määritelmää $U^*U = I = V^*V$ ja kohdissa 2 lauseen 3.8 kohtaa (e). □

Esimerkki 4.1. Olkoon $x \in \mathbb{C}^n$ yksikkövektori. Esimerkin tarkoitus on osoittaa, kuinka on mahdollista muodostaa sellainen unitaarinen matriisi $U = \begin{bmatrix} x & u_2 & \dots & u_n \end{bmatrix}$, että matriisin ensimmäisenä sarakkeena on vektori x.

Kirjoitetaan $x = \begin{bmatrix} x_1 & y^T \end{bmatrix}^T$, missä $x_1 \in \mathbb{C}$ ja $y \in \mathbb{C}^{n-1}$. Valitaan sellainen luku $\theta \in \mathbb{R}$, että $e^{i\theta}x_1 \geq 0$ ja määritellään vektori $z = e^{i\theta}x = \begin{bmatrix} z_1 & \zeta^T \end{bmatrix}^T$, missä luku $z_1 \in \mathbb{R}$ on epänegatiivinen ja $\zeta \in \mathbb{C}^{n-1}$. Tarkastellaan nyt matriisia

$$V_x = \begin{bmatrix} z_1 & \zeta^* \\ \zeta & -I + \frac{1}{1+z_1}\zeta\zeta^* \end{bmatrix} = \begin{bmatrix} V_{x_1} & V_{x_2} \\ V_{x_3} & V_{x_4} \end{bmatrix}. \tag{4.6}$$

On helppo nähdä, että kyseessä on hermiittinen matriisi eli matriisilla on ominaisuus $V_x = V_x^*$. Lasketaan sitten $V_x^*V_X = V_x^2$ hyödyntäen matriisin ositusta. Koska x on yksikkövektori, niin $\|x\|^2 = 1$ ja tästä seuraa, että $\|z\|^2 = |e^{i\theta}|\|x\|^2 = \|x\|^2 = 1$. Lisäksi vektorin z normin neliö voidaan kirjoittaa muodossa

$$\|z\|^2 = \zeta^*\zeta + \|z_1\|^2 = \zeta^*\zeta + z_1^2 = 1 \tag{4.7}$$

ja tätä yhtälöä tullaan hyödyntämään eri muodoissa laskemisessa. Lasketaan aluksi V_x^2

lohkoittain

$$
\begin{aligned}
V_x^2 &= \begin{bmatrix} V_{x_1} & V_{x_2} \\ V_{x_3} & V_{x_4} \end{bmatrix} \begin{bmatrix} V_{x_1} & V_{x_2} \\ V_{x_3} & V_{x_4} \end{bmatrix} = \begin{bmatrix} V_{x_1}V_{x_1} + V_{x_2}V_{x_3} & V_{x_1}V_{x_2} + V_{x_2}V_{x_4} \\ V_{x_3}V_{x_1} + V_{x_4}V_{x_3} & V_{x_3}V_{x_2} + V_{x_4}V_{x_4} \end{bmatrix} \\
&= \begin{bmatrix} z_1^2 + \zeta^*\zeta & z_1\zeta^* + \zeta^*(-I + \frac{1}{1+z_1}\zeta\zeta^*) \\ z_1\zeta + (-I + \frac{1}{1+z_1}\zeta\zeta^*)\zeta & \zeta\zeta^* + (-I + \frac{1}{1+z_1}\zeta\zeta^*)^2 \end{bmatrix}
\end{aligned}
$$

ja selvyyden vuoksi sievennetään jokainen lohko yksitellen aloittaen ensimmäisen rivin ensimmäisestä lohkosta. Hyödyntämällä yhtälöä (4.7) saadaan

$$
z_1^2 + \zeta^*\zeta = 1.
$$

Jatketaan ensimmäisen rivin toisesta sarakkeesta

$$
\begin{aligned}
& z_1\zeta^* + \zeta^*(-I + \frac{1}{1+z_1}\zeta\zeta^*) = z_1\zeta^* - \zeta^* I + \frac{1}{1+z_1}\zeta^*\zeta\zeta^* && |\zeta^*\zeta = 1 - z_1^2 \\
&= z_1\zeta^* - \zeta^* + \frac{1-z_1^2}{1+z_1}\zeta^* = (z_1 - 1 + \frac{1-z_1^2}{1+z_1})\zeta^* && |\text{osittelulaki} \\
&= \frac{z_1^2 - 1 + 1 - z_1^2}{1+z_1}\zeta^* = O_{1,n-1}.
\end{aligned}
$$

Koska toisen rivin ensimmäinen sarake on äskeisen lohkon konjugaattitranspoosi, niin

$$
\begin{aligned}
z_1\zeta + (-I + \frac{1}{1+z_1}\zeta\zeta^*)\zeta &= (z_1\zeta^* + \zeta^*(-I + \frac{1}{1+z_1}\zeta\zeta^*))^* \\
&= (O_{1,n-1})^* = O_{n-1,1}.
\end{aligned}
$$

Sievennetään vielä viimeinen lohko. Kohdassa 1 käytetään yhtälöä $\zeta^*\zeta = 1 - z_1^2$ ja kohdassa 2 osittelulakia. Saadaan

$$
\begin{aligned}
& \zeta\zeta^* + (-I + \frac{1}{1+z_1}\zeta\zeta^*)^2 = \zeta\zeta^* + (I - \frac{2}{1+z_1}\zeta\zeta^* + \big(\frac{1}{1+z_1}\big)^2\zeta\zeta^*\zeta\zeta^*) \\
&\stackrel{1}{=} I + \zeta\zeta^* - \frac{2}{1+z_1}\zeta\zeta^* + \frac{1-z_1^2}{(1+z_1)^2}\zeta\zeta^* \stackrel{2}{=} I + (1 - \frac{2}{1+z_1} + \frac{1-z_1^2}{(1+z_1)^2})\zeta\zeta^* \\
&= I + \frac{(1+z_1)^2 - 2(1+z_1) + 1 - z_1^2}{(1+z_1)^2}\zeta\zeta^* = I + \frac{1 + 2z_1 + z_1^2 - 2 - 2z_1 + 1 - z_1^2}{(1+z_1)^2}\zeta\zeta^* \\
&= I + 0\zeta\zeta^* = I_{n-1,n-1}.
\end{aligned}
$$

Yhdistämällä nämä lohkot saadaan

$$
V_x^2 = \begin{bmatrix} 1 & O_{1,n-1} \\ O_{n-1,1} & I_{n-1,n-1} \end{bmatrix} = I_n.
$$

Tulokseksi saatiin siis identiteettimatriisi. Matriisi V_x on määritelmän 4.2 mukaan unitaarinen, sillä $V_x^* V_x = V_x^2 = I$. Myös matriisi $U = e^{-i\theta} V_x$ on unitaarinen, koska

$$U^* U = (e^{-i\theta} V_x)^* (e^{-i\theta} V_x) = \overline{e^{-i\theta}} V_x^* e^{-i\theta} V_x = e^{i\theta} e^{-i\theta} V_x^* V_x = V_x^* V_x = I.$$

Todetaan vielä, että matriisin U ensimmäinen sarake on muotoa $e^{-i\theta} z = e^{-i\theta} e^{i\theta} x = x$. Näin ollen osoitettiin, kuinka on mahdollista muodostaa sellainen unitaarinen matriisi $U = \begin{bmatrix} x & u_2 & \dots & u_n \end{bmatrix}$, että matriisin ensimmäisenä sarakkeena on yksikkövektori x.

Seuraava lause on ehkä hyödyllisin matriisiteorian perustulos. Tämä on Isaac Schurin ansioita. Lauseen mukaan jokainen neliömatriisi $A \in M_n$ on similaarinen yläkolmiomatriisin kanssa, jonka diagonaalialkioina ovat matriisin A ominaisarvot jossain järjestyksessä. Lauseen todistuksessa tehdään peräkkäisiä redusointeja unitaarisen simillarisuuden avulla. [9, s.101]

Lause 4.8 (Schurin hajotelma)**.** *Olkoon $A \in M_n$ matriisi, jonka ominaisarvot ovat $\lambda_1, \dots, \lambda_n$ missä tahansa järjestyksessä, ja olkoon $x \in \mathbb{C}^n$ sellainen yksikkövektori, että $Ax = \lambda_1 x$. Tällöin on olemassa sellainen unitaarinen matriisi $U = \begin{bmatrix} x & u_2 & \dots & u_n \end{bmatrix} \in M_n$, että $U^* A U = T = [t_{ij}]$ on yläkolmiomatriisi, jonka diagonaalialkioina ovat $t_{ii} = \lambda_i$, missä $i = 1, \dots, n$.*

Todistus. (Vrt. [9], Theorem 2.3.1) Olkoon x matriisin A normeerattu ominaisvektori, joka vastaa ominaisarvoa λ_1. Toisin sanoen $x^* x = 1$ ja $Ax = \lambda_1 x$. Olkoon $U_1 = \begin{bmatrix} x & u_2 & \dots & u_n \end{bmatrix}$ unitaarinen matriisi, jonka ensimmäinen sarake on x. Edellinen esimerkki 4.1 näyttää, kuinka tällainen on mahdollista muodostaa. Tällöin

$$\begin{aligned} U_1^* A U_1 &= U_1^* \begin{bmatrix} Ax & Au_2 & \dots & Au_n \end{bmatrix} = U_1^* \begin{bmatrix} \lambda_1 x & Au_2 & \dots & Au_n \end{bmatrix} \\ &= \begin{bmatrix} x^* \\ u_2^* \\ \vdots \\ u_n^* \end{bmatrix} \begin{bmatrix} \lambda_1 x & Au_2 & \dots & Au_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 x^* x & x^* A u_2 & \dots & x^* A u_n \\ \lambda_1 u_2^* x & & & \\ \vdots & & A_1 & \\ \lambda_1 u_n^* x & & & \end{bmatrix} = \begin{bmatrix} \lambda_1 & \bigstar \\ 0 & A_1 \end{bmatrix}, \end{aligned}$$

sillä unitaarisen matriisin U_1 sarakkeet ovat ortonormaalit eli $u_i^* x = 0$ jokaisella $i = 2, \dots, n$.

Koska similaaristen matriisien ominaisarvot ovat samat, niin

$$\begin{aligned}\det(\lambda I - A) &= \det\left(\lambda I - \begin{bmatrix}\lambda_1 & \bigstar \\ 0 & A_1\end{bmatrix}\right) \\ &= \det\begin{vmatrix}\lambda - \lambda_1 & \bigstar \\ 0 & \lambda I - A_1\end{vmatrix} = (\lambda - \lambda_1)\det(\lambda I - A_1).\end{aligned}$$

Täten alimatriisin $A_1 = [u_i^* A u_j]_{i,j=2}^n \in M_{n-1}$ ominaisarvot ovat $\lambda_2, \ldots, \lambda_n$. Jos $n = 2$, niin ollaan saatu haluttu tulos. Jos näin ei ole, jatketaan prosessia samaan tapaan. Olkoon $\xi \in \mathbb{C}^{n-1}$ matriisin A_1 ominaisarvoa λ_2 vastaava yksikkövektori. Tehdään edellä esitetty redusointi matriisille A_1. Jos $U_2 \in M_{n-1}$ on unitaarinen matriisi, jonka ensimmäinen sarake on vektori ξ, niin saadaan

$$U_2^* A_1 U_2 = \begin{bmatrix}\lambda_2 & \bigstar \\ 0 & A_2\end{bmatrix}.$$

Olkoon $V_2 = [1] \oplus U_2 = \left[\begin{smallmatrix}1 & 0 \\ 0 & U_2\end{smallmatrix}\right]$. Tämä on unitaarinen matriisi, sillä

$$\begin{bmatrix}1 & 0 \\ 0 & U_2^*\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2\end{bmatrix} = \begin{bmatrix}1 & 0 \\ 0 & U_2^* U_2\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & I\end{bmatrix} = I.$$

Nyt

$$\begin{aligned}(U_1V_2)^* A U_1 V_2 = V_2^* U_1^* A U_1 V_2 &= \begin{bmatrix}1 & 0 \\ 0 & U_2^*\end{bmatrix}\begin{bmatrix}\lambda_1 & \bigstar \\ 0 & A_1\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2\end{bmatrix} \\ &= \begin{bmatrix}\lambda_1 & \bigstar \\ 0 & U_2^* A_1\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2\end{bmatrix} = \begin{bmatrix}\lambda_1 & \bigstar \\ 0 & U_2^* A_1 U_2\end{bmatrix} \\ &= \begin{bmatrix}\lambda_1 & \bigstar & \bigstar \\ 0 & \lambda_2 & \bigstar \\ 0 & 0 & A_2\end{bmatrix}.\end{aligned}$$

Jatketaan edelleen redusointia, joka tuottaa unitaariset matriisit $U_i \in M_{n-i+1}$, missä $i = 1, \ldots, n-1$ ja $V_i \in M_n$, missä $i = 2, \ldots, n-2$. Unitaaristen matriisien tulo on unitaarinen lauseen 4.5 nojalla. Täten matriisi $U = U_1 V_2 V_3 \cdots V_{n-2}$ on unitaarinen ja $U^* A U$ on yläkolmiomatriisi. $\square$

Seuraavan lauseen mukaan kohtien (a) ja (b) mukaan matriisi A on normaali, jos ja vain jos matriisi A on unitaarisesti diagonalisoituva. Tästä käytetään nimitystä normaalien matriisien spektraalilause. [9, s.133]

Lause 4.9. *Olkoon $A = [a_{ij}] \in M_n$ matriisi, jonka ominaisarvot ovat $\lambda_1, \ldots, \lambda_n$. Seuraavat väittämät ovat yhtäpitäviä keskenään.*

(a) Matriisi A on normaali.

(b) Matriisi A on unitaarisesti diagonalisoituva.

(c) $\sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2$.

Todistus. (Vrt. [9], Theorem 2.5.3) Osoitetaan päättelyketju $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a)$, mikä todistaa väittämät yhtäpitäviksi.

Lauseen 4.8 avulla voidaan kirjoittaa $A = UTU^*$, missä $U = [u_1 \quad \ldots \quad u_n] \in M_n$ on unitaarinen matriisi ja $T = [t_{ij}] \in M_n$ on yläkolmiomatriisi. Jos matriisi A on normaali, niin tällöin myös matriisi T on normaali lauseen 4.6 nojalla. Apulause 4.1 takaa, että matriisi T on diagonaalimatriisi, joten matriisi A on unitaarisesti similaarinen diagonaalimatriisin kanssa, eli unitaarisesti diagonalisoituva.

Jos on olemassa sellainen unitaarinen matriisi V, että $A = V\Lambda V^*$ ja $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, niin $\operatorname{tr} A^*A = \operatorname{tr} \Lambda^*\Lambda$ lauseen 4.7 nojalla, mikä on kohdan (c) väite.

Matriisin T diagonaalialkiot ovat ominaisarvot $\lambda_1, \ldots, \lambda_n$ jossain järjestyksessä ja täten lauseen 4.7 perusteella $\operatorname{tr} A^*A = \operatorname{tr} T^*T = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i<j} |t_{ij}|^2$. Näin ollen kohdasta (c) seuraa, että $\sum_{i<j} |t_{ij}|^2 = 0$, joten T on diagonaalimatriisi ja matriisi A on unitaarisesti diagonalisoituva.

Yhtälö $A = UTU^*$ voidaan kirjoittaa $U^*AU = T$. Koska T on diagonaalimatriisi, niin se on myös normaali apulauseen 4.1 mukaan. Näin ollen lauseen 4.6 nojalla matriisi A on normaali. $\square$

Normaalin matriisin $A \in M_n$ esitysmuotoa $A = U\Lambda U^*$, missä matriisi U on unitaarinen ja Λ on diagonaalimatriisi, sanotaan matriisin A spektraalihajotelmaksi. Kun matriisin $A \in M_n$ erisuuret ominaisarvot $\lambda_1, \ldots, \lambda_d$ ovat tiedossa, matriisille A voidaan muodostaa spektraalihajotelma noudattamalla seuraavaa ohjeistusta. Määritetään jokaiselle erisuurelle ominaisarvolle λ ominaisavaruuden $\{x \in \mathbb{C}^n : Ax = \lambda x\}$ kanta ja ortonormeerataan (ortogonalisoidaan ja normeerataan) se. Eri ominaisarvoja vastaavat ominaisavaruudet ovat keskenään ortogonaaliset lauseen 4.3 nojalla ja ominaisavaruuksien dimensiot vastaavat ominaisarvojen algebrallista kertalukua lauseen 3.23 nojalla. Täten ominaisavaruuksien yhdiste on avaruuden $\mathbb{C}^n$ ortonormaali kanta. Järjestämällä nämä kantavektorit sarakkeiksi matriisiin $U \in M_n$ saadaan sellainen unitaarinen matriisi, että U^*AU on diagonaalimatriisi. Edellä muodostettu unitaarinen matriisi U ei ole yksikäsitteinen, sillä ominaisavaruuksien kannat voidaan valita monella eri tapaa. [9, s.134]

4.2 Hermiittiset matriisit

Hermiittisten matriisien luokka on monesta näkökulmasta luonnollinen yleistys reaalisymmetrisille matriiseille. Hermiittiinen ja reaalinen matriisi on reaalisymmetrinen matriisi. Reaalisymmetristen matriisien monet tärkeät ominaisuudet eivät ole voimassa kompleksisymmetrisille matriiseille. [9, s.227]. Tässä alaluvussa tutustutaan hermiittisiin matriiseihin ja niiden ominaisuuksiin. Ennen kaikkea on keskitytty tutkimaan hermiittisen matriisin A, sen ominaisarvojen ja luvun x^*Ax välisiä yhteyksiä. Seuraavassa luvussa tullaan käsittelemään positiivisesti (semi)definiittejä matriiseja, joiden yhteys hermiittisiin matriiseihin ja lukuun x^*Ax on erittäin keskeinen.

Määritelmä 4.4. Matriisi $A \in M_n$ on *hermiittinen*, jos se on itsensä konjugaattitranspoosi. Toisin sanoen, matriisi A on hermiittinen, jos yhtälö $A = A^*$ on voimassa. Ehdon $A = -A^*$ täyttävää matriisia kutsutaan *vinohermiittiseksi*.

Esimerkki 4.2. Olkoon $A = [a_{ij}], B = [b_{ij}] \in M_n$ mitä tahansa neliömatriiseja. Tällöin matriisit $A + A^*$, AA^* ja A^*A ovat hermiittisiä. Osoitetaan, että nämä matriisit täyttävät määritelmän 4.4 ehdon. Konjugaattitranspoosin ominaisuuksia hyödyntämällä saadaan, että

$$(A + A^*)^* = A^* + (A^*)^* = A^* + A = A + A^*,$$

$$(AA^*)^* = ((A^*)^*A^*) = AA^*$$

ja

$$(A^*A)^* = (A^*(A^*)^*) = A^*A.$$

Täten väitetyt matriisit ovat hermiittisiä.

Lause 4.10. *(Vrt. [9], Problem 4.1.P4) Olkoot $A = [a_{ij}], B = [b_{ij}] \in M_n$ hermiittisiä matriiseja. Tällöin*

(a) A^k on hermiittinen jokaisella $k = 1, 2, 3, \ldots$.

(b) Jos A on ei-singulaarinen, niin A^{-1} on hermiittinen.

(c) Matriisi $aA + bB$ on hermiittinen kaikilla reaaliluvuilla a ja b.

(d) Matriisin A päädiagonaalin alkiot ovat reaalisia.

(e) Reaalinen ja symmetrinen matriisi on hermiittinen.

Todistus. Hermiittisille matriiseille on voimassa $A = A^*$ ja $B = B^*$.

(a) $(A^k)^* = (\overset{k \text{ kpl}}{AA \cdots A})^* = (\overset{k \text{ kpl}}{A^*A^* \cdots A^*}) = (\overset{k \text{ kpl}}{AA \cdots A}) = A^k$.

(b) $A^{-1} = (A^*)^{-1} = (A^{-1})^*$.

(c) $(aA + bB)^* = \overline{a}A^* + \overline{b}B^* = aA^* + bB^* = aA + bB$.

(d) Määritelmän 4.4 mukaan matriisi A toteuttaa ehdon $A = A^*$, eli $[a_{ij}] = [\overline{a}_{ji}]$. Päädiagonaalin alkioille on siis oltava voimassa ehto $a_{ii} = \overline{a}_{ii}$, mikä toteutuu vain, jos a_{ii} on reaaliluku.

(e) Reaaliselle matriisille $A = \overline{A}$ ja symmetriselle $A = A^T$, joten $A^* = \overline{A}^T = A^T = A$.

□

Lause 4.11. *Olkoon $A = [a_{ij}] \in M_n$. Matriisi A on hermiittinen matriisi, jos ja vain jos vähintään yksi seuraavista ehdoista toteutuu:*

(a) neliömuoto x^*Ax on reaaliluku jokaisella vektorilla $x \in \mathbb{C}^n$,

(b) matriisi A on normaali ja sillä on vain reaalisia ominaisarvoja,

(c) matriisi S^*AS on hermiittinen jokaisella matriisilla $S \in M_n$.

Todistus. (Vrt. [9], Theorem 4.1.3) Oletetaan aluksi, että matriisi A on hermiittinen. Matriisitulon x^*Ax dimensio on 1×1. Kyseessä on siis luku, joten $x^*Ax = (x^*Ax)^T$. Laskemalla saadaan

$$\overline{(x^*Ax)} = \overline{(x^*A^*x)} = \overline{(x^*Ax)^*} = \overline{\overline{(x^*Ax)^T}} = (x^*Ax)^T = x^*Ax.$$

Luku x^*Ax on siis itsensä kompleksikonjugaatti, eli $x^*Ax = \overline{x^*Ax}$. Tästä seuraa, että x^*Ax on oltava reaalinen. Jos $Ax = \lambda x$ ja $x^*x = 1$, niin $\lambda = \lambda x^*x = x^*\lambda x = x^*Ax$, joka on reaalinen äskeisen perusteella. Viimeiseksi matriisi S^*AS on hermiittinen määritelmän 4.4 nojalla, sillä $(S^*AS)^* = S^*A^*(S^*)^* = S^*A^*S = S^*AS$.

Todistetaan sitten väite toiseen suuntaan. Oletetaan, että x^*Ax on reaaliluku jokaisella vektorilla $x \in \mathbb{C}^n$. Täten

$$\begin{aligned}&(x+y)^*A(x+y) = (x^*+y^*)A(x+y) = (x^*A+y^*A)(x+y)\\ &= x^*Ax + x^*Ay + y^*Ax + y^*Ay\end{aligned}$$

on reaaliluku oletuksen nojalla jokaisella $x, y \in \mathbb{C}^n$. Myös x^*Ax ja y^*Ay ovat reaalilukuja, joten voidaan päätellä, että $x^*Ay + y^*Ax$ on oltava reaalinen. Jos valitaan $x = e_k$ ja $y = e_j$, niin saadaan, että $x^*Ay + y^*Ax = a_{kj} + a_{jk}$ on reaaliluku, joten on oltava $\mathrm{Im}\, a_{kj} = -\mathrm{Im}\, a_{jk}$. Jos valitaan $x = ie_k$ ja $y = e_j$, niin saadaan, että $x^*Ay + y^*Ax = -ia_{kj} + ia_{jk}$ on reaaliluku, joten on oltava $\mathrm{Re}\, a_{kj} = \mathrm{Re}\, a_{jk}$. Yhdistämällä nämä reaali- ja imaginaariosat saadaan ominaisuus $a_{kj} = \overline{a}_{jk}$. Koska k ja j ovat mielivaltaisia, niin $A = A^*$ ja määritelmän 4.4 mukaan matriisi A on hermiittinen. Jos matriisi A on normaali, se on unitaarisesti diagonalisoituva lauseen 4.9 nojalla, joten se voidaan kirjoittaa $A = U\Lambda U^*$, missä $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Jos ominaisarvot ovat reaalisia, niin matriisi Λ on myös. Täten $A^* = (U\Lambda U^*)^* = U\overline{\Lambda}U^* = U\Lambda U^* = A$ ja matriisi A on hermiittinen.

Lopuksi oletetaan, että väite (c) on tosi. Valitaan $S = I$, jolloin $S^*AS = A$ on hermiittinen oletuksen perusteella. $\square$

Hermiittiset matriisit ovat normaaleja matriiseja, sillä $AA^* = A^2 = A^*A$. Näin ollen kaikki normaalien matriisien ominaisuudet ovat voimassa myös hermiittisille matriiseille. [9, s.229] Seuraavaa keskeistä lausetta sanotaan hermiittisien matriisien spektraalilauseeksi. Se on normaalien matriisien spektraalilauseen 4.9 erityistapaus. [9, s.135]

Lause 4.12. *Olkoon $A \in M_n$ hermiittinen matriisi, jonka ominaisarvot ovat $\lambda_1, \ldots, \lambda_n$. Olkoon $\Lambda =$ diag$(\lambda_1, \ldots, \lambda_n)$. Tällöin*

(a) ominaisarvot $\lambda_1, \ldots, \lambda_n$ ovat reaalisia,

(b) matriisi A on unitaarisesti diagonalisoituva,

(c) on olemassa sellainen unitaarinen matriisi $U \in M_n$, että $A = U\Lambda U^*$.

Todistus. (Vrt. [9], Theorem 2.5.5) Väite (a) todistettu edellä. Väite (b) seuraa lauseesta 4.9, sillä hermiittinen matriisi on normaali. Väiteessä (b) diagonaalimatriisin alkioiden on oltava ominaisarvoja, sillä unitaarisesti similaareilla matriiseilla on samat ominaisarvot ja diagonaalimatriisin ominaisarvoja ovat sen diagonaalialkiot. Tämän tiedon perusteella väite (c) on vain väite (b) toisin ilmaistuna. $\square$

Kirjoitetaan edellinen lause hieman eri muodossa mukavuussyistä. Tämä on tehty alla.

Lause 4.13. *Matriisi $A \in M_n$ on hermiittinen, jos ja vain jos on olemassa sellainen unitaarinen matriisi $U \in M_n$ ja sellainen reaalinen diagonaalimatriisi $\Lambda \in M_n$, että $A = U\Lambda U^*$.*

Todistus. (Vrt. [9], Theorem 4.1.5) Jos on olemassa sellainen unitaarinen matriisi $U \in M_n$ ja sellainen reaalinen diagonaalimatriisi $\Lambda \in M_n$, että $A = U\Lambda U^*$, niin matriisi A on hermiittinen lauseen 4.11 kohdan (b) nojalla. Todistus toiseen suuntaan seuraa lauseesta 4.12. $\square$

Lause 4.14. *Olkoon matriisi $A \in M_n$. Tällöin x^*Ax on reaalinen ja positiivinen (vastaavasti, epänegatiivinen) jokaisella nollasta eroavalla vektorilla $x \in \mathbb{C}^n$, jos ja vain jos matriisi A on hermiittinen ja kaikki sen ominaisarvot ovat positiivisia (vastaavasti, epänegatiivinen).*

Todistus. (Vrt. [9], Theorem 4.1.8) Oletetaan, että x^*Ax on reaalinen ja positiivinen aina, kun $x \neq 0$. Tällöin lause 4.11 takaa, että matriisi A on hermiittinen. Lisäksi $\lambda = \lambda u^*u = u^*(\lambda u) = u^*Au$, jos yksikkövektori $u \in \mathbb{C}^n$ on matriisin A ominaisarvoa λ vastaava ominaisvektori. Näin ollen oletuksesta seuraa, että $\lambda > 0$.

Oletetaan sitten, että matriisi A on hermiittinen ja että kaikki sen ominaisarvot $\lambda_1, \dots, \lambda_n$ ovat positiivisia. Koska matriisi A on hermiittinen, niin se voidaan kirjoittaa muotoon $A = U\Lambda U^*$, missä $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ ja matriisi $U = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}$ on unitaarinen. Tällöin

$$\begin{aligned} x^*Ax = x^*U\Lambda U^*x &= \begin{bmatrix} x^*u_1 & \dots & x^*u_n \end{bmatrix} \Lambda \begin{bmatrix} u_1^*x & \dots & u_n^*x \end{bmatrix}^T \\ &= \begin{bmatrix} x^*u_1\lambda_1 & \dots & x^*u_n\lambda_n \end{bmatrix} \begin{bmatrix} u_1^*x & \dots & u_n^*x \end{bmatrix}^T \\ &= \sum_{k=1}^n x^*u_k\lambda_k u_k^*x = \sum_{k=1}^n \lambda_k (u_k^*x)^* u_k^*x = \sum_{k=1}^n \lambda_k |u_k^*x|^2. \end{aligned}$$

Oletuksen perusteella on aina $\lambda_k > 0$, joten summan positiivisuus riippuu termeistä u_k^*x. Koska matriisi U^* on unitaarinen, se on ei-singulaarinen ja näin ollen $U^*x = 0$ vain jos $x = 0$. Tästä seuraa, että on olemassa sellainen indeksi $k \in \{1, \dots, n\}$, että $u_k^*x \neq 0$ aina, kun $x \neq 0$. Täten x^*Ax on positiivinen. Todistus väitteelle epänegatiivisuudesta menee samaan tapaan. $\square$

4.3 Positiivisesti definiitit ja semidefiniitit matriisit

Positiivisesti (semi)definiitit matriisit ovat hermiittisiä matriiseja, joilla erityinen positiivisuusominaisuus. Positiivisesti definiittien matriisien voidaan ajatella vastaavan positiivisen luvun yleistystä matriisien joukossa. Näitä hermiittisten matriisien erikoistapauksia hyödynnetään esimerkiksi Hessen matriisin kautta optimoinnissa tutkittaessa lokaaleja ääriarvoja tai funktioiden konveksisuutta. [9, s.425-426] Tämän alaluvun lauseessa 4.20 huomataan, että matriisijuuri on määritelty vain positiivisesti semidefiniiteille matriiseille, mikä vahvistaa ajatusta positiivisen luvun yleistyksestä. Lähes jokaista alaluvun lauseista tullaan tarvitsemaan Schurin lauseen todistamiseen.

Määritelmä 4.5. Hermiittinen matriisi $A \in M_n$ on positiivisesti definiitti, jos

$$x^*Ax > 0 \text{ kaikilla nollasta eroavilla vektoreilla } x \in \mathbb{C}^n. \tag{4.8}$$

Hermiittinen matriisi $A \in M_n$ on positiivisesti semidefiniitti, jos

$$x^*Ax \geq 0 \text{ kaikilla nollasta eroavilla vektoreilla } x \in \mathbb{C}^n. \tag{4.9}$$

Määritelmässä oletetaan, että matriisi on hermiittinen. Tämä on positiivisesti definiitin matriisin määritelmässä tavanomaista, vaikkakin tarpeetonta. [9, s. 430] Jotta määritelmä on mielekäs, on luvun x^*Ax oltava reaalinen, sillä kompleksilukujen muodostama kunta ei ole järjestetty. Luvun x^*Ax reaalisuus puolestaan takaa matriisin A hermiittisyyden lauseen 4.11 nojalla. Huomiona vielä, että positiivisesti semidefiniitissä määritelmässä ei olisi pakko olettaa vektorin x eroavan nollavektorista, sillä jos $x = 0$, niin $x^*Ax = 0$, ja määritelmässä mainittu ehto toteutuu.

Määritelmästä nähdään suoraan, että positiivisesti definiitit matriisit ovat myös positiivisesti semidefiniittejä. Negatiivisesti definiitti matriisi voidaan määritellä vaihtamalla epäyhtälön suuntaa tai vaihtoehtoisesti, jos $-A$ on positiivisesti definiitti, niin A on negatiivisesti definiitti. Edellä esitetyt määritelmät ovat yhtäpitäviä, sillä $x^*Ax < 0 \Leftrightarrow x^*(-A)x > 0$. Vastaavasti voidaan määritellä negatiivisesti semidefiniitti matriisi. Jos matriisi ei ole positiivisesti eikä negatiivisesti definiitti, sanotaan, että matriisi on indefiniitti. [9, s.430]

Lause 4.15. *(Vrt. [9, s.430], Exercise) Jos $A \in M_n$ positiivisesti definiitti, niin matriisit $\overline{A}, A^T, A^*$ ja A^{-1} ovat myös positiivisesti definiittejä.*

Todistus. Olkoon matriisi $A \in M_n$ positiivisesti definiitti ja olkoon $x \in \mathbb{C}^n$ mikä tahansa nollavektorista eroava kompleksinen vektori. Huomataan, että $(x^*Ax)^T = \overline{x^*Ax} = x^*Ax > 0$, koska x^*Ax on reaaliluku. Merkitään $y = \overline{x} \in \mathbb{C}^n \Leftrightarrow y^* = x^T$. Matriisi $\overline{A}$ on positiivisesti definiitti, sillä

$$x^*\overline{A}x = \overline{x^T A\overline{x}} = \overline{y^*Ay} = y^*Ay > 0.$$

Vastaavin argumentein matriisi A^T on positiivisesti definiitti, sillä

$$x^*A^Tx = (x^TA\overline{x})^T = (y^*Ay)^T = y^*Ay > 0.$$

Matriisi A on positiivisesti definiitti, joten matriisi A on hermiittinen, eli $A = A^*$. Näin saadaan

$$x^*A^*x = x^*Ax > 0.$$

Yllä todistetut ominaisuudet voitaisiin osoittaa samaan tapaan myös positiivisesti semidefiniiteille matriiseille. Viimeisen väittämän todistaminen positiivisesti semidefiniiteille matriiseille ei onnistu, sillä nämä saattavat olla singulaarisia. Osoitetaan aluksi, että positiivisesti definiitti matriisi A on ei-singulaarinen. Hyödynnetään epäsuoraa todistusta. Tehdään vastaoletus, että matriisi A on singulaarinen. Tällöin yhtälöllä $Ax = 0$ on jokin ei-triviaali ratkaisu $x \neq 0$. Nyt tälle ratkaisulle on voimassa $x^*Ax = x^*0 = 0 \not> 0$. Päädyttiin ristiriitaan oletuksen kanssa, jonka mukaan A on positiivisesti definiitti. Näin ollen matriisin A on oltava ei-singulaarinen.

Merkitään $Ay = x$. Koska matriisi A on ei-singulaarinen, niin yhtälöllä $Ay = x$ on yksikäsitteinen ratkaisu kaikilla vektoreilla $x \in \mathbb{C}^n$. Matriisi A^{-1} on positiivisesti definiitti, sillä

$$x^*A^{-1}x = (Ay)^*A^{-1}(Ay) = y^*A^*A^{-1}Ay = y^*A^*Iy = y^*A^*y = y^*Ay > 0.$$

□

Lause 4.16. *Olkoon $A \in M_n$ hermiittinen matriisi. Jos A on positiivisesti definiitti, niin tällöin kaikki sen pääalimatriisit ovat positiivisesti definiittejä. Jos A on positiivisesti semidefiniitti, niin tällöin kaikki sen pääalimatriisit ovat positiivisesti semidefiniittejä.*

Todistus. (Vrt. [9], Observation 7.1.3) Todistetaan lauseen ensimäinen väite positiivisesti definiiteille matriiseille. Todistus positiivisesti semidefiniiteille matriiseille etenee samaan tapaan. Olkoon α joukon $\{1, \ldots, n\}$ aito osajoukko. Tarkastellaan alimatriisia $A[\alpha]$. Olkoon $x \in \mathbb{C}^n$ sellainen mielivaltainen vektori, että $x[\alpha] \neq 0$ ja $x[\alpha^c] = 0$. Täten $x \neq 0$ ja

$$x[\alpha]^* A[\alpha] x[\alpha] = x^* A x > 0,$$

sillä $x[\alpha^c] = 0$ ja A on positiivisesti definiitti. Koska vektori $x[\alpha]$ on mielivaltainen, luku $x[\alpha]^* A[\alpha] x[\alpha]$toteuttaa määritelmän 4.5 ehdon (4.8) ja $A[\alpha]$ on näin ollen positiivisesti definiitti. □

Apulause 4.3. *Olkoon $A \in M_n$ positiivisesti semidefiniitti ja olkoon $C \in M_{n,m}$. Tällöin matriisi $C^* A C$ on positiivisesti semidefiniitti.*

Todistus. (Vrt. [9], Observation 7.1.8) Olkoon $x \in \mathbb{C}^m$. Merkitään $y = Cx \in \mathbb{C}^n$. Nyt $x^* C^* A C x = y^* A y \geq 0$, sillä A on positiivisesti semidefiniitti. Näin ollen $C^* A C$ on positiivisesti semidefiniitti. □

Positiivisesti definiittejä ja semidefiniittejä matriiseja voidaan luonnehtia monin, toisinaan yllätävinkin tavoin. [9, s.438] Seuraavassa lauseessa ilmaistaan jo aiemmin todistettu hyödyllinen yhteys hermiitisen matriisin ominaisarvojen ja definiittisyyden välillä.

Lause 4.17. *Hermiittinen matriisi on positiivisesti semidefiniitti, jos ja vain jos sen jokainen ominaisarvo on epänegatiivinen reaaliluku. Vastaavasti, hermiittinen matriisi on positiivisesti semidefiniitti, jos ja vain jos sen jokainen ominaisarvo on positiivinen reaaliluku.*

Todistus. (Vrt. [9], Theorem 7.2.1) Tämä on lause on uudelleen muotoiltu versio lauseesta 4.14. □

Lause 4.18. *Olkoon $A \in M_n$ positiivisesti semidefiniitti matriisi (vastaavasti positiivisesti definiitti). Tällöin $\operatorname{tr} A$, $\det A$ ja matriisin A jokainen päälideterminantti on epänegatiivinen (vastaavasti positiivinen). Lisäksi $\operatorname{tr} A = 0$, jos ja vain jos $A = 0$.*

Todistus. (Vrt. [9], Corollary 7.1.4) Todistetaan lause positiivisesti semidefiniiteille matriiseille. Todistus positiivisesti definiiteille matriiseille etenee samaan tapaan. Matriisin ominaisarvot ovat kaikki epänegatiivisia edellisen lauseen perusteella. Matriisin jälki on epänegatiivinen, sillä se on matriisin ominaisarvojen summa lauseen 3.17 nojalla. Jos

$\operatorname{tr} A = 0$, on matriisin A jokaisen ominaisarvon oltava 0. Koska matriisi A on hermiittisenä diagonalisoituva ja diagonaalimatriisilla on oltava samat ominaisarvot kuin matriisilla A, niin jokaisen diagonaalimatriisin diagonaalialkion on oltava 0. Täten diagonaalimatriisi on nollamatriisi samoin kuin matriisi A lauseen 3.22 nojalla. Matriisin A determinantti epänegatiivinen, sillä se on epänegatiivisten lukujen tulo lauseen 3.17 mukaan. Päälideterminantit ovat pääalimatriisien determinantteja ja lause 4.16 takaa, että pääalimatriisit ovat positiivisesti semidefiniittejä, joten kaikki pääalimatriisin ominaisarvot ovat epänegatiivisia. Tästä seuraa, että päälideterminantit ovat epänegatiivisia. □

Aiemmin todettiin, että positiivisesti definiitin (vastaavasti semidefiniitin) matriisin päälideterminantit ovat positiivisia (vastaavasti epänegatiivisia). Seuraava lause, Sylvesterin kriteeri, antaa vastaavanlaisen mutta käänteisen tuloksen. Sylvesterin kriteerin antaa vaihtoehtoisen tavan osoittaa matriisin positiivinen definiittisyys määritelmän 4.5 ja lauseen 4.17 ohelle. Lauseen (b)-kohdassa on tärkeä huomata, että ensimmäisten $n-1$ johtavan päälideterminantin on oltava nimenomaan positiivisia, eikä pelkkä epänegatiivisuus riitä. Esimerkiksi matriisin $\left[\begin{smallmatrix} 0 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ johtavat päälideterminantit ovat epänegatiivisia, mutta matriisi ei ole positiivisesti semidefiniitti [9, s.439].

Lause 4.19 (Sylvesterin kriteeri). *Olkoon matriisi $A \in M_n$ hermiittinen.*

(a) *Jos jokainen matriisin A johtava (vastaavasti häntäpäälideterminantti) päälideterminantti on positiivinen (mukaan lukien $\det A$), niin matriisi A on positiivisesti definiitti.*

(b) *Jos matriisiin A ensimmäiset $n-1$ johtavaa (vastaavasti häntäpäälideterminanttia) päälideterminanttia ovat positiivisia ja $\det A \geq 0$, niin matriisi A on positiivisesti semidefiniitti.*

Todistus. (Vrt. [9], Theorem 7.2.5)

(a) Todistetaan väittämä induktion avulla. Olkoon $A_k = A[\{1, \ldots, k\}]$, missä $k = 1, \ldots, n$, matriisin A johtava pääalimatriisi. Jos $\det A_1 > 0$, niin matriisin A_1 ainoa ominaisarvo on positiivinen ja näin ollen matriisi A_1 on positiivisesti definiitti lauseen 4.17 nojalla.

Tehdään induktio-oletus. Jos $k \in \{1, \ldots, n-1\}$ ja jokainen johtava päälideterminantti on positiivinen, niin matriisi A_k on positiivisesti definiitti. Tällöin matriisi A_k on ei-singulaarinen, sillä A_k on positiivisesti definiitti, ja A_k^{-1} on positiivisesti definiitti.

Todistetaan sitten tapaus $k+1$. Oletetaan, että matriisin A_{k+1} jokainen johtava päälideterminantti on positiivinen. Koska A_{k+1} on hermiittinen, niin se voidaan kirjoittaa mudossa

$$A_{k+1} = \begin{bmatrix} A_k & y \\ y^* & a \end{bmatrix},$$

missä vektori $y \in \mathbb{C}^k$ ja $a \in \mathbb{C}$. Matriisin A_{k+1} determinantti voidaan laskea lohkomatriisien determinantin kaavalla (3.14). Oletuksen perusteella determinantti on positiivinen. Saadaan

$$\det A_{k+1} = \det A_k \det(a - y^* A_k y) > 0,$$

joten $a - y^* A_k y > 0$. Olkoon $z = \begin{bmatrix} x^T & \alpha \end{bmatrix}^T \in \mathbb{C}^{k+1}$, missä $x \in \mathbb{C}^k$ ja $\alpha \in \mathbb{C}$, mielivaltainen nollasta eroava vektori. Tällöin

$$\begin{aligned}
z^* A_{k+1} z &= x^* A_k x + \overline{\alpha} y^* x + \alpha x^* y + \overline{\alpha} a \alpha \\
&= x^* A_k x + \overline{\alpha} y^* x + \alpha x^* y + (|\alpha|^2 y^* A_k^{-1} y - |\alpha|^2 y^* A_k^{-1} y) + |\alpha|^2 a \\
&= (x^* A_k + \overline{\alpha} y^*) x + (x^* + \overline{\alpha} y^* A_k^{-1}) \alpha y + |\alpha|^2 a - |\alpha|^2 y^* A_k^{-1} y \\
&= (x^* A_k + \overline{\alpha} y^*) x + (x^* A_k A_k^{-1} + \overline{\alpha} y^* A_k^{-1}) \alpha y + |\alpha|^2 a - |\alpha|^2 y^* A_k^{-1} y \\
&= (x^* A_k + \overline{\alpha} y^*) x + (x^* A_k + \overline{\alpha} y^*) \alpha A_k^{-1} y + |\alpha|^2 a - |\alpha|^2 y^* A_k^{-1} y \\
&= (x^* A_k + \overline{\alpha} y^*)(x + \alpha A_k^{-1} y) + |\alpha|^2 a - |\alpha|^2 y^* A_k^{-1} y \\
&= (x^* A_k + \overline{\alpha} y^* A_k^{-1} A_k)(x + \alpha A_k^{-1} y) + |\alpha|^2 a - |\alpha|^2 y^* A_k^{-1} y \\
&= (x^* + \overline{\alpha} y^* A_k^{-1}) A_k (x + \alpha A_k^{-1} y) + |\alpha|^2 a - |\alpha|^2 y^* A_k^{-1} y \\
&= (x + \alpha A_k^{-1} y)^* A_k (x + \alpha A_k^{-1} y) + |\alpha|^2 a - |\alpha|^2 y^* A_k^{-1} y \\
&= c^* A_k c + |\alpha|^2 (a - y^* A_k^{-1} y).
\end{aligned}$$

Tiedetään, että matriisi A_k on positiivisesti definiitti, joten $c^* A_k c > 0$, jos $c \neq 0$, ja $c^* A_k c = 0$, jos $c = 0$. Jos $\alpha = 0$, niin $c = x \neq 0$ ja edelleen $z^* A_k z = c^* A_k c > 0$. Jos $\alpha \neq 0$, niin $z^* A_k z \geq |\alpha|^2 (a - y^* A_k^{-1} y) > 0$. Näin ollen matriisi A_{k+1} on positiivisesti definiitti.

Tarkastellaan sitten väitettä matriisin $A \in M_{2n}$ häntäpääalideterminanteille. Matriisissa A on parillinen määrä rivejä ja sarakkeita. Merkitään indeksien i ja j rivienvaihto $i \leftrightarrow j$. Tehdään matriisille A seuraavat rivienvaihdot $1 \leftrightarrow 2n, 2 \leftrightarrow 2n-1, \ldots, n \leftrightarrow n+1$. Vaihdetaan myös näitä samoja indeksejä vastaavien sarakkeiden paikkoja, jolloin saadaan muunnettu matriisi $\tilde{A}$. Matriisi

$$E = \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} \in M_{2n}$$

on tehtyjä rivi- ja sarakeoperaatioita vastaavien alkeismatriisien tulo. Tällöin muunnettu matriisi voidaan kirjoittaa $\tilde{A} = EAE \Leftrightarrow A = E\tilde{A}E$, sillä $E = E^{-1}$. Matriisit $\tilde{A}$ ja A ovat siis similaariset. Kirjoittamalla matriisit A ja $\tilde{A}$ auki saadaan

$$\tilde{A} = \begin{bmatrix} a_{2n,2n} & a_{2n,2n-1} & a_{2n,2n-2} & \cdots & a_{2n,1} \\ a_{2n-1,2n} & a_{2n-1,2n-1} & a_{2n-1,2n-2} & \cdots & a_{2n-1,1} \\ a_{2n-2,2n} & a_{2n-2,2n-1} & a_{2n-2,2n-2} & \cdots & a_{2n-2,1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1,2n} & a_{1,2n-1} & a_{1,2n-2} & \cdots & a_{11} \end{bmatrix}$$

ja

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1,2n-2} & a_{1,2n-1} & a_{1,2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{2n-2,1} & \cdots & a_{2n-2,2n-2} & a_{2n-2,2n-1} & a_{2n-2,2n} \\ a_{2n-1,1} & \cdots & a_{2n-1,2n-2} & a_{2n-1,2n-1} & a_{2n-1,2n} \\ a_{2n,1} & \cdots & a_{2n,2n-2} & a_{2n,2n-1} & a_{2n,2n} \end{bmatrix}.$$

Huomataan, että matriisin $\tilde{A}$ johtavat alideterminanti vastaavat matriisin A häntäalideterminantteja. Toisin sanoen $\det A[\{k,\dots,2n\}] = \det \tilde{A}[\{1,\dots,2n-k+1\}]$, missä $k = 1,2,\dots,2n$. Jos matriisin A häntäalideterminantit ovat positiivisia, tällöin myös matriisin $\tilde{A}$ johtavat alideterminantit ovat positiivisia. Täten matriisi $\tilde{A}$ on positiivisesti definiitti aiemman todistuksen osan nojalla, eli matriisin $\tilde{A}$ jokainen ominaisarvo on positiivinen lauseen 4.17 nojalla. Similaarisuudesta seuraa, että matriisilla A on samat ominaisarvot lauseen 3.22 nojalla, joten matriisi A on positiivisesti definiitti.

Jos $A \in M_{2n+1}$, suoritetaan rivien ja sarakkeiden vaihdot indekseille $1 \leftrightarrow 2n+1, 2 \leftrightarrow 2n, \dots, n \leftrightarrow n+2$. Vastaavin argumentein kuin edellä voidaan päätellä, että A on positiivisesti definiitti.

(b) Todistus etenee samaan tapaan kuin (a)-kohdassa, mutta tällä kertaa

$$\det A_{k+1} = \det A_k \det(a - y^* A_k y) \geq 0,$$

joten $a - y^* A_k y \geq 0$. Hyödyntämällä tätä saadaan

$$\begin{aligned} z^* A_{k+1} z &= (x + \alpha A_k^{-1} y)^* A_k (x + \alpha A_k^{-1} y) + |\alpha|^2 (a - y^* A_k^{-1} y) \\ &= c^* A_k c + |\alpha|^2 (a - y^* A_k^{-1} y) \geq c^* A_k c \geq 0, \end{aligned}$$

sillä matriisi A_k on positiivisesti definiitti ja vektori $c \in \mathbb{C}^k$ voi mahdollisesti olla nollavektori. Näin ollen matriisi A_{k+1} on positiivisesti semidefiniitti. Väite häntäalideterminanteille seuraa tästä ja similaarisuudesta samaan tapaan kuin edellä.

□

Lause 4.20. *Olkoon $A \in M_n$ hermiittinen ja positiivisesti semidefiniitti matriisi ja olkoon $k \in \{2, 3, \ldots\}$. Tällöin on olemassa sellainen yksikäsitteinen hermiittinen ja positiivisesti semidefiniitti matriisi B, että $B^k = A$.*

Todistus. (Vrt. [9], Theorem 7.2.6) Matriisi A voidaan esittää lauseen 4.12 mukaan muodossa $A = U\Lambda U^*$, missä matriisi $U \in M_n$ on unitaarinen ja $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \in M_n$, missä ominaisarvot $\lambda_1, \ldots, \lambda_n$ ovat epänegatiivisia. Määritellään matriisi $B = U\Lambda^{1/k}U^*$, missä $\Lambda^{1/k} = \text{diag}(+\lambda_1^{1/k}, \ldots, +\lambda_n^{1/k})$ ja jokaisessa tapauksessa otetaan yksikäsitteinen epänegatiivinen k:s juuri. Tällöin matriisi B on hermiittinen lauseen 4.13 nojalla, positiivisesti semidefiniitti lauseen 4.14 nojalla ja $B^k = A$.

Todistetaan vielä yksikäsitteisyys. Oletetaan, että on olemassa toinen hermiittinen ja positiivisesti semidefiniitti matriisi C. Tällöin lauseen 4.12 avulla matriisi C voidaan kirjoittaa $C = V\Omega^{1/k}V^*$, missä V on unitaarinen matriisi ja $\Omega^{1/k} = \text{diag}(+\omega_1^{1/k}, \ldots, +\omega_n^{1/k})$ on diagonaalimatriisi, jossa luvut $\omega_1, \ldots, \omega_n$ ovat epänegatiivisia. Nyt yhtälöstä $B^k = C^k = A$ saadaan

$$\begin{aligned}
U\Lambda U^* &= V\Omega V^* && |\cdot U, V^*\cdot \\
V^*U\Lambda &= \Omega V^*U && |T = V^*U \\
T\Lambda &= \Omega T && |\text{ matriisitulon määritelmä 3.3} \\
[t_{ij}\lambda_j] &= [\omega_i t_{ij}].
\end{aligned}$$

Täten $t_{ij}\lambda_j = \omega_i t_{ij}$ jokaisella $i, j = 1, \ldots, n$, mistä seuraa, että $t_{ij}\lambda_j^{1/k} = \omega_i^{1/k}t_{ij}$. Näin ollen

$$\begin{aligned}
[t_{ij}\lambda_j^{1/k}] &= [\omega_i^{1/k}t_{ij}] \\
T\Lambda^{1/k} &= \Omega^{1/k}T && |T = V^*U \\
V^*U\Lambda^{1/k} &= \Omega^{1/k}V^*U \\
U\Lambda^{1/k}U^* &= V\Omega^{1/k}V^* \\
B &= C,
\end{aligned}$$

mikä todistaa yksikäsitteisyyden. □

Merkintä $A^{1/k}$ on positiivisesti semidefiniitin matriisin A positiivinen yksikäsitteinen k. juuri. Seuraava esimerkki havainnollista, kuinka tällainen juuri voidaan käytännössä laskea.

Esimerkki 4.3. Lasketaan edellisen lauseen todistuksen mukaisesti $\left[\begin{smallmatrix} 5 & 4 \\ 4 & 5 \end{smallmatrix}\right]^{1/2}$. Selvitetään aluksi ominaisarvot ja niitä vastaavat normeeratut ominaisvektorit. Ominaisarvot saadaan yhtälöstä

$$\begin{aligned} &\det(A-\lambda I)=0\\ \Leftrightarrow\quad &\begin{vmatrix} 5-\lambda & 4\\ 4 & 5-\lambda \end{vmatrix}=0\\ \Leftrightarrow\quad &(5-\lambda)^2-16=0\\ \Leftrightarrow\quad &(5-\lambda)^2=16\\ \Leftrightarrow\quad &5-\lambda_1=4\wedge 5-\lambda_2=-4\\ \Leftrightarrow\quad &\lambda_1=1\wedge\lambda_2=9. \end{aligned}$$

Ominaisarvoa $\lambda_1 = 1$ vastaava ominaisvektori x_1 voidaan ratkaista yhtälöstä $(A-\lambda_1 I)x_1 = 0$. Saadaan

$$\begin{bmatrix} 4 & 4 & 0\\ 4 & 4 & 0 \end{bmatrix} \overset{\text{RREF}}{\rightarrow} \begin{bmatrix} 1 & 1 & 0\\ 0 & 0 & 0 \end{bmatrix} \Rightarrow x_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1\\ -1 \end{bmatrix}.$$

Vastaavasti saadaan ominaisarvoa $\lambda_2 = 9$ vastaava ominaisvektori x_2

$$\begin{bmatrix} -4 & 4 & 0\\ 4 & -4 & 0 \end{bmatrix} \overset{\text{RREF}}{\rightarrow} \begin{bmatrix} 1 & -1 & 0\\ 0 & 0 & 0 \end{bmatrix} \Rightarrow x_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1\\ 1 \end{bmatrix}.$$

Näin ollen $U = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$ on unitaarinen lauseen 4.4 nojalla ja $\Lambda^{1/2} = \operatorname{diag}(1,3)$. Nyt voidaan laskea haluttu juuri

$$\begin{aligned} \begin{bmatrix} 5 & 4\\ 4 & 5 \end{bmatrix}^{1/2} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1\\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0\\ 0 & 3 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1\\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 3\\ -1 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1\\ 1 & 1 \end{bmatrix}\\ &= \frac{1}{2}\begin{bmatrix} 4 & 2\\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1\\ 1 & 2 \end{bmatrix}. \end{aligned}$$

5. KRONECKERIN JA HADAMARDIN TULO

Perinteisen matriisitulon ohella matriiseille on mahdollista määritellä erilaisia tuloja. Tässä luvussa tutustutaan kahteen merkittävään tuloon, Kroneckerin tulo ja Hadamardin tulo, ja näiden tulojen ominaisuuksiin. Nämä tulot ovat hyvin erilaisia verrattuna perinteiseen matriisituloon, tavallaan jopa yksinkertaisempia. Matriisitulossa tarvitaan sekä kunnan (tässä työssä kompleksiluvut) skalaarikerto- että yhteenlaskuoperaatiota, mutta Kroneckerin ja Hadamardin tulossa on tarve vain kertolaskulle.

Matriisitulossa 3.3 vasemmanpuoleisen tulontekijän rivien ja oikeanpuoleisen tulontekijän sarakkeiden alkiot kerrotaan keskenään ja summataan yhteen uudeksi alkioksi. Jos määritelmän 3.3 matriisitulo lasketaan, tarvitaan $mn(3p-1)$ kerto- ja yhteenlaskuoperaatiota. Neliömatriisien $A, B \in M_n$ tapauksessa operaatioiden määrä on $3n^3 - n^2$.

Kroneckerin tulo 5.1 on määritelty kaikenkokoisille matriiseille. Siinä vasemmanpuoleisen matriisin jokainen alkio kerrotaan oikeanpuoleisella matriisilla ja saaduista matriiseista muodostetaan lohkomatriisi. Jos määritelmän 5.1 matriiseille lasketaan Kroneckerin tulo ja jokainen lohko laskettaisiin erikseen auki, tarvittaisiin $mnpq$ kertolaskuoperaatiota. Neliömatriisien $A, B \in M_n$ tapauksessa operaatioiden määrä on n^4.

Hadamardin tulo 5.2 on määritelty kahdelle samankokoiselle matriisille. Tulo suoritetaan matriisien yhteenlaskun tapaan alkioittain, mikä tekee siitä yksinkertaisen. Määritelmän tuloon tarvitaan mn kertolaskuoperaatiota. Neliömatriisien $A, B \in M_n$ tapauksessa operaatioiden määrä on n^2.

5.1 Kroneckerin tulo

Jotkin matematiikan historoitsijat ovat kyseenalaistaneet nimityksen Kroneckerin tulo, sillä ei ole näyttöä siitä, että Kronecker olisi ollut edelläkävijä tulon löytämisessä tai käyttämisessä. Tänä päivänä Kroneckerin tulo on kuitenkin vakiinnuttanut asemansa ja on lähes yleisesti käytössä. [10, s.254] Kirjallisuudessa Kroneckerin tulosta käytetään myös nimityksiä suora tulo ja tensoritulo [10, s. 242]. Tässä alaluvussa määritellään Kroneckerin tulo ja esitellään siihen liittyviä perusominaisuuksia ja lauseita. Erityisesti lause 5.6, positiivisesti (semi)definiittien matriisien Kroneckerin tulo on positiivisesti (semi)definiitti, mahdollistaa Schurin lauseen hyvinkin elegantin todistuksen.

Määritelmä 5.1. Matriisien $A \in M_{m,n}$ ja $B \in M_{p,q}$ Kroneckerin tuloa merkitään $A \otimes B$. Tulo määritellään lohkomatriisina seuraavasti

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix} = [a_{ij}B] \in M_{mp,nq}. \tag{5.1}$$

Kroneckerin tulo on määritelty kahdelle mielivaltaisen kokoiselle matriisille. Määritelmästä on helppo nähdä, että matriisien I_m ja I_n Kroneckerin tulo on I_{mn} ja että $A \otimes 0 = 0 \otimes A = 0$. Todetaan vielä esimerkin avulla, että Kroneckerin tulo ei välttämättä ole vaihdannainen.

Esimerkki 5.1. Olkoon $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ ja $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Tällöin

$$A \otimes B = \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix}$$

ja

$$B \otimes A = \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix},$$

joten selkeästi $A \otimes B \neq B \otimes A$.

Ainoastaan matriisijoukko M_1 on suljettu Kroneckerin tulon suhteen. Näin ollen joukko M_n, missä $n > 1$, ei muodosta rengasta yhteenlaskun ja Kroneckerin tulon kanssa, toisin kuin yhteenlasku ja matriisitulo tai yhteenlasku ja Hadamardin tulo, kuten tullaan myöhemmin huomaamaan.

Lause 5.1. *(Vrt. [10], Equations 4.2.3-4.2.8, sekä [10, s.265], Problem 13 ja 18) Olkoon $A = [a_{ij}] \in M_{m,n}$, $B = [b_{ij}] \in M_{p,q}$ ja $C = [c_{ij}] \in M_{r,s}$ mielivaltaisia matriiseja. Olkoon $x = [x_i] \in \mathbb{C}^n$ ja $y^T = [y_j] \in \mathbb{C}^n$ ja olkoon $\alpha \in \mathbb{C}$ jokin kompleksiluku. Poikkeuksena väittämissä (f) ja (g) yhteenlaskettavien matriisien on oltava samankokoisia ja väittämät (h) ja (i) matriisit $A \in M_n$ ja $B \in M_m$ ovat neliömatriiseja. Tällöin*

(a) $xy^T = x \otimes y^T = y^T \otimes x$,

(b) $(\alpha A) \otimes B = \alpha(A \otimes B) = A \otimes (\alpha B)$,

(c) $(A \otimes B)^T = A^T \otimes B^T$,

(d) $(A \otimes B)^* = A^* \otimes B^*$,

(e) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$,

(f) $(A + B) \otimes C = A \otimes C + B \otimes C$,

(g) $A \otimes (B + C) = A \otimes B + A \otimes C$,

(h) $\operatorname{tr}(A \otimes B) = \operatorname{tr} A \operatorname{tr} B$,

(i) $\det(A \otimes B) = (\det A)^m (\det B)^n$.

Todistus. Väittämät seuraavat suoraviivaisesti laskemalla ja hyödyntämällä määritelmiä.

(a) $xy^T = [x_i y_j]$, $x \otimes y^T = [x_i y^T] = [x_i y_j]$ ja $y^T \otimes x = [y_j x] = [y_j x_i] = [x_i y_j]$.

(b) $(\alpha A) \otimes B = [(\alpha a_{ij})B] = \alpha[a_{ij}B] = \alpha(A \otimes B) = [a_{ij}(\alpha B)] = A \otimes (\alpha B)$.

(c) $(A \otimes B)^T = [a_{ij}B]^T = [a_{ji}B^T] = A^T \otimes B^T$.

(d) $(A \otimes B)^* = [a_{ij}B]^* = [\overline{a}_{ji}B^*] = A^* \otimes B^*$.

(e) $(A \otimes B) \otimes C = [a_{ij}B] \otimes C = [a_{ij}b_{hk}C] = [a_{ij}(b_{hk}C)] = [a_{ij}(B \otimes C)] = A \otimes (B \otimes C)$.

(f) $(A + B) \otimes C = [(a_{ij} + b_{ij})C] \otimes C = [a_{ij}C + b_{ij}C] = [a_{ij}C] + [b_{ij}C] = A \otimes C + B \otimes C$.

(g) $A \otimes (B + C) = [a_{ij}(B + C)] = [a_{ij}B + a_{ij}C] = [a_{ij}B] + [a_{ij}C] = A \otimes B + A \otimes C$.

(h) Ominaisuus seuraa lauseista 3.17 ja 5.5 seuravasti:

$$\operatorname{tr}(A \otimes B) = \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \mu_j = \sum_{i=1}^{n} \lambda_i \sum_{j=1}^{m} \mu_j = \operatorname{tr} A \operatorname{tr} B.$$

(i) Ominaisuus seuraa lauseista 3.17 ja 5.5 seuraavasti:

$$\begin{aligned} \det(A \otimes B) &= \prod_{i=1}^{n} \prod_{j=1}^{m} \lambda_i \mu_j = \prod_{i=1}^{n} \lambda_i^m \prod_{j=1}^{m} \mu_j = \prod_{i=1}^{n} \lambda_i^m \det B \\ &= (\det B)^n \prod_{i=1}^{n} \lambda_i^m = (\det A)^m (\det B)^n. \end{aligned}$$

□

Lauseessa 5.1 on esitelty Kroneckerin tulon perusominaisuuksia. Näistä ominaisuuksista seuraa muun muassa, että hermiittisten matriisien Kroneckerin tulo on hermiittinen, sillä $A \otimes B = A^* \otimes B^* = (A \otimes B)^*$. Lisäksi $A \otimes B = 0$, jos ja vain jos $A = 0$ tai $B = 0$. Jos $A = 0$ tai $B = 0$, tulos on triviaali. Jos $A \otimes B = 0$, määritelmän 5.1 nojalla on oltava $a_{ij}B = 0$ jokaisella $i = 1, \ldots, m$ ja $j = 1, \ldots, n$. Tästä seuraa, että $B = 0$ tai $a_{ij} = 0$ jokaisella i, j eli $A = 0$. Eri tuloista voidaan muodostaa yhdistettyjä tuloja. Seuraava lause on esimerkki tällaisesta. Siinä yhdistetään Kroneckerin tuloja ja matriisituloja.

Lause 5.2. *Olkoot* $A \in M_{m,n}$, $B \in M_{p,q}$, $C \in M_{n,k}$ *ja* $D \in M_{q,r}$. *Tällöin*

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

Todistus. (Vrt. [10], Lemma 4.2.10) Merkitään $A = [a_{ik}]$ ja $C = [c_{kj}]$. Tällöin Kroneckerin tulon määritelmän mukaan $A \otimes B = [a_{ik}B]$ ja $C \otimes D = [c_{kj}D]$. Näin ollen

$$\begin{aligned}(A \otimes B)(C \otimes D) = [a_{ik}B][c_{kj}D] &= \left[\sum_{k=1}^{n} a_{ik}Bc_{kj}D\right] = \left[\left(\sum_{k=1}^{n} a_{ik}c_{kj}\right)BD\right] \\ &= \big[[AC]_{ij}BD\big] = AB \otimes CD.\end{aligned}$$

□

Edellisen lauseen avulla on helppo todeta, että ei-singulaarisien matriisien Kroneckerin tulo on myös ei-singulaarinen ja että unitaaristen matriisien Kroneckerin tulo on unitaarinen. Lisäksi lausetta tarvitaan, kun halutaan tutkia Kroneckerin tulon ominaisarvoja ja -vektoreita. Matriisin ominaisarvoista ja -vektoreista voidaan puolestaan päätellä lisää ominaisuuksia.

Lause 5.3. *Jos matriisit $A \in M_m$ ja $B \in M_n$ ovat ei-singulaarisia, niin $A \otimes B$ on myös ja $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.*

Todistus. (Vrt. [10], Corollary 4.2.11) Tarkastellaan tuloa

$$(A \otimes B)(A^{-1} \otimes B^{-1}) \overset{\text{lause } 5.2}{=} AA^{-1} \otimes BB^{-1} = I_m \otimes I_n = I_{mn}.$$

Näin ollen $A \otimes B$ on ei-singulaarinen ja $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ määritelmän 3.6 mukaisesti. □

Lause 5.4. *(Vrt. [15], Theorem 6.17) Jos matriisit $A \in M_m$ ja $B \in M_n$ ovat unitaarisia, niin $A \otimes B$ on myös.*

Todistus. Tarkastellaan tuloa

$$(A \otimes B)^*(A \otimes B) \overset{\text{lause } 5.1}{=} (A^* \otimes B^*)(A \otimes B) \overset{\text{lause } 5.2}{=} A^*A \otimes B^*B = I_m \otimes I_n = I_{mn}.$$

Näin ollen $A \otimes B$ on unitaarinen määritelmän 4.2 mukaisesti. □

Lause 5.5. *Olkoon $A \in M_n$ ja B_m matriiseja, joiden ominaisarvot ovat $\lambda_1, \ldots, \lambda_n$ ja $\mu_1, \ldots, \mu_m$ tässä järjestyksessä. Tällöin matriisin $A \otimes B$ ominaisarvot ovat*

$$\lambda_i \mu_j,$$

missä $i = 1, \ldots, n$ ja $j = 1, \ldots, m$. Erityisesti $\sigma(A \otimes B) = \sigma(B \otimes A)$. Lisäksi, jos ominaisarvoa λ vastaa ominaisvektori $x \in \mathbb{C}^n$ ja ominaisarvoa μ vastaa ominaisvektori $y \in \mathbb{C}^m$, niin ominaisarvoa $\lambda\mu$ vastaa ominaisvektori $x \otimes y \in \mathbb{C}^{nm}$.

Todistus. (Vrt. [10], Theorem 4.2.12) Lauseen 4.8 nojalla on olemassa sellaiset unitaariset matriisit U ja V, että $U^*AU = T_1$ ja $V^*BV = T_2$, missä matriisit T_1 ja T_2 ovat yläkolmiomatriiseja, joiden diagonaalialkiot ovat $\lambda_1, \ldots, \lambda_m$ ja $\mu_i, \ldots, \mu_n$ tässä järjestyksessä. Näin ollen

$$\begin{aligned} T_1 \otimes T_2 = (U^*AU) \otimes (V^*BV) &\overset{\text{lause 5.2}}{=} (U^* \otimes V^*)(AU \otimes BV) \\ &\overset{\text{lause 5.2}}{=} (U^* \otimes V^*)(A \otimes B)(U \otimes V). \end{aligned}$$

□

Lauseen 5.4 mukaan matriisit $(U^* \otimes V^*)$ ja $(U \otimes V)$ ovat unitaarisia. Täten matriisit $A \otimes B$ ja $T_1 \otimes T_2$ ovat unitaarisesti similaarit ja lauseen 3.22 nojalla niillä on samat ominaisarvot. Yläkolmiomatriisin $T_1 \otimes T_2$ ominaisarvot ovat muotoa $\lambda_i \mu_j$, missä $i = 1, \ldots, m$ ja $j = 1, \ldots, n$ lauseen 3.19 mukaan.

Todistetaan sitten jälkimmäinen väite. Tiedetään, että $Ax = \lambda x$ ja $By = \mu y$, joten hyödyntämällä lauseita 5.1 ja 5.2 saadaan

$$(A \otimes B)(x \otimes y) = Ax \otimes By = \lambda x \otimes \mu y = \lambda\mu(x \otimes y).$$

Lause 5.6. *Olkoon matriisit $A \in M_n$ ja $B \in M_m$ positiivisesti (semi)definiittejä. Tällöin Kroneckerin tulo $A \otimes B$ on positiivisesti (semi)definiitti.*

Todistus. (Vrt. [10], Corollary 4.2.13) Matriisit A ja B ovat positiivisesti definiittejä, joten määritelmän 4.5 nojalla ne ovat myös hermiittisiä. Lisäksi lauseesta 4.17 seuraa, että matriisien ominaisarvot ovat positiivisia. Aiemmin todettiin, että Kroneckerin tulo $A \otimes B$ on hermiittinen, jos matriisit A ja B ovat hermiittisiä. Lauseen 5.5 nojalla matriisin $A \otimes B$ ominaisarvot ovat positiivisia. Näin ollen lause 4.17 takaa, että $A \otimes B$ on positiivisesti definiitti. Todistus positiivisesti semidefiniiteille matriiseille etenee vastaavasti. □

5.2 Hadamardin tulo

Kirjallisuudessa Hadamardin tulosta käytetään jossain materiaaleissa intuitiivisesti nimitystä tulo alkioittain tai nimitystä Schurin tulo, koska Issai Schur todisti joitain varhaisia perusominaisuuksia liittyen Hadamardin tuloon [10, s.298]. Moniin ohjelmistoihin on ohjelmoitu Hadamardin tulo valmiiksi. MATLAB® ohjelmistossa matriisitulon ja Hadamardin tulon ero on piste, esimerkiksi matriisitulo saadaa komennolla 'A*B' ja Hadamardin tulo komennolla 'A.*B'. Tässä alaluvussa määritellään Hadamardin tulo ja tutustutaan sen perusominaisuuksiin. Lisäksi tehdään havainto, että Hadamardin tulo on Kroneckerin tulon alimatriisi, ja todistetaan Schurin lause kahdella eri tavalla sekä tutkitaan Hadamardin tulon determinanttiin liittyviä epäyhtälöitä.

Määritelmä 5.2. Olkoon $A = [a_{ij}] \in M_{m,n}$ ja $B = [b_{ij}] \in M_{m,n}$. Tällöin matriisien A ja B välinen Hadamardin tulo on $A \circ B = [a_{ij}b_{ij}] \in M_{m,n}$.

Matriisien $M_{m,n}$ joukko muodostaa kommutoivan renkaan yhteenlaskun ja Hadamardin tulon kanssa. Alaluvussa 3.1 todettiin, että ryhmä $(M_n, +)$ on Abelin ryhmä. Itse asiassa myös $(M_{m,n}, +)$ on Abelin ryhmä. Alla olevasta lauseesta voidaan päätellä, että $(M_{m,n}, \circ)$ on monoidi, eli joukko $M_{m,n}$ on suljettu Hadamardin tulon suhteen, Hadamardin tulo on assosiatiivinen ja tulolla on neutraalialkio. Neutraalialkio on matriisi, jonka jokainen alkio on 1, eli niin sanottu ykkösmatriisi $J \in M_{m,n}$. Lisäksi Hadamardin tulo $\circ$ on kommutatiivinen ja distributiivinen yhteenlaskun suhteen.

Lause 5.7. *(Vrt. [12], Theorem 1.2, 1.3 ja 1.5) Olkoot $A, B, C, J, O \in M_{m,n}$ ja olkoon $\alpha \in \mathbb{C}$. Tällöin seuraavat ominaisuudet ovat voimassa Hadamardin tulolle:*

(a) $A \circ B = B \circ A$,

(b) $A \circ (B \circ C) = (A \circ B) \circ C$,

(c) $A \circ (B + C) = A \circ B + A \circ B$,

(d) $(\alpha A) \circ B = A \circ (\alpha B) = \alpha(A \circ B)$,

(e) $A \circ J = J \circ A = A$,

(f) $A \circ O = O \circ A = O$.

Todistus. Nämä ominaisuudet seuraavat suoraviivaisesti määritelmästä 5.2 ja kunnan $(\mathbb{C}, +, \cdot)$ tunnetuista ominaisuuksista.

(a) $A \circ B = [a_{ij}b_{ij}] = [b_{ij}a_{ij}] = B \circ A$.

(b) $A \circ (B \circ C) = [a_{ij}] \circ ([b_{ij}] \circ [c_{ij}]) = [a_{ij}] \circ [b_{ij}c_{ij}] = [a_{ij}b_{ij}c_{ij}] = [a_{ij}b_{ij}] \circ [c_{ij}] = ([a_{ij}] \circ [b_{ij}]) \circ [c_{ij}] = (A \circ B) \circ C$.

(c) $A \circ (B + C) = [a_{ij}] \circ [b_{ij} + c_{ij}] = [a_{ij}(b_{ij} + c_{ij})] = [a_{ij}b_{ij} + a_{ij}c_{ij}] = [a_{ij}b_{ij}] + [a_{ij}c_{ij}] = A \circ B + A \circ C$.

(d) $(\alpha A) \circ B = [\alpha a_{ij}] \circ [b_{ij}] = [\alpha a_{ij}b_{ij}] = \alpha[a_{ij}b_{ij}] = \alpha(A \circ B) = [a_{ij}(\alpha b_{ij})] = [a_{ij}] \circ [\alpha b_{ij}] = A \circ (\alpha B)$.

(e) $A \circ J = [a_{ij} \cdot 1] = [1 \cdot a_{ij}] = [a_{ij}] = A$.

(f) $A \circ O = [a_{ij} \cdot 0] = [0 \cdot a_{ij}] = [0] = O$.

□

Lause 5.8. *(Vrt. [9, s.478], Exercise) Olkoon matriisit $A, B \in M_n$ hermiittisiä. Tällöin Hadamardin tulo $A \circ B$ on hermiittinen.*

Todistus. Hyödyntämällä määritelmiä 4.4 ja 5.2 saadaan

$$A \circ B = A^* \circ B^* = [\overline{a}_{ji}\overline{b}_{ji}] = [a_{ij}b_{ij}]^* = (A \circ B)^*,$$

mikä todistaa väitteen. □

Seuraavan lauseen perusteella Hadamardin tulo voidaan ilmaista Kroneckerin tulon alimatriisina. Jotkin ominaisuudet, kuten positiivinen definiittisyys, periytyvät matriisilta alimatriisille, mikä tekee lauseesta hyödyllisen.

Lause 5.9. *Jos $A, B \in M_{m,n}$, niin*

$$A \circ B = (A \otimes B)[\alpha, \beta],$$

missä $\alpha = \{1, m+2, 2m+3, \ldots, m^2\}$ ja $\beta = \{1, n+2, 2n+3, \ldots, n^2\}$. Erityisesti, jos $m = n$, niin Hadamardin tulo $A \circ B$ on Kroneckerin tulon $A \otimes B$ pääalimatriisi.

Todistus. (Vrt. [15], Theorem 6.18) Olkoon $e_{mi} \in \mathbb{R}^m$ ja $e_{ni} \in \mathbb{R}^n$ vektoreita, joiden i. alkio on luku 1 ja muut alkiot ovat nollia. Muodostetaan matriisit

$$E_m = \begin{bmatrix} e_{m1} \otimes e_{m1} & \ldots & e_{mm} \otimes e_{mm} \end{bmatrix} \in M_{m^2,m}$$

ja

$$E_n = \begin{bmatrix} e_{n1} \otimes e_{n1} & \ldots & e_{nn} \otimes e_{nn} \end{bmatrix} \in M_{n^2,n}.$$

Matriisien E_m ja E_n sarakkeesta i riviltä $(i-1)(m+1)+1$ tai $(i-1)(n+1)+1$ löytyy luku 1, muut alkiot ovat nollia. Toisin sanoen joukon α tai β indeksoimilta riveiltä löytyy luku 1. Tulo $E_m^T(A \otimes B)E_n$ poimii siis matriisista $A \otimes B$ indeksijoukon α osoittamat rivit ja indeksijoukon β osoittamat sarakkeet ja täten tulo vastaa lauseen alimatriisia $(A \otimes B)[\alpha, \beta]$. Tarkastellaan nyt Hadamardin tulon $A \circ B$ rivin i sarakkeen j alkiota

$$\begin{aligned} a_{ij}b_{ij} &= (e_{mi}^T A e_{nj})(e_{mi}^T B e_{nj}) = (e_{mi}^T A e_{nj}) \otimes (e_{mi}^T B e_{nj}) && \mid \text{lause 5.2} \\ &= (e_{mi}^T \otimes e_{mi}^T)(A e_{nj} \otimes B e_{nj}) && \mid \text{lause 5.2} \\ &= (e_{mi}^T \otimes e_{mi}^T)(A \otimes B)(e_{nj} \otimes e_{nj}) && \mid \text{lause 5.1} \\ &= (e_{mi} \otimes e_{mi})^T (A \otimes B)(e_{nj} \otimes e_{nj}), \end{aligned}$$

mikä on matriisin $E_m^T(A \otimes B)E_n$ rivin i sarakkeen j alkio. Indeksit i ja j ovat mielivaltaisia, joten $A \circ B = E_m^T(A \otimes B)E_n = (A \otimes B)[\alpha, \beta]$. Jos $m = n$, niin $\alpha = \beta$ ja täten Hadamardin tulo $A \circ B$ on Kroneckerin tulon $A \otimes B$ pääalimatriisi määritelmän 3.8 mukaisesti. □

Samaan tapaan kuin lauseessa 5.2 tehtiin, voidaan muodostaa yhdistetty tulo Hadamardin ja Kroneckerin tulolle. Tähän tuloon liittyvä ominaisuus on alla.

Lause 5.10. *(Vrt. [4], Theorem 2.11) Olkoon* $A = [a_{ij}], C = [c_{ij}] \in M_{m,n}$ *ja* $B, D \in M_{p,q}$. *Tällöin yhdistetty tulo*

$$(A \otimes B) \circ (C \otimes D) = (A \circ C) \otimes (B \circ D).$$

Todistus. Laskemalla määritelmien 5.1 ja 5.2 mukaisesti saadaan

$$\begin{aligned}(A \otimes B) \circ (C \otimes D) &= [a_{ij}B] \circ [c_{ij}D] = [(a_{ij}B) \circ (c_{ij}D)] \\ &= [a_{ij}c_{ij}(B \circ D)] = [a_{ij}c_{ij}] \otimes (B \circ D) \\ &= (A \circ C) \otimes (B \circ D),\end{aligned}$$

mikä todistaa väitteen. □

Seuraava apulause antaa kätevän esitysmuodon Hadamardin tulolle tavallisen matriisitulon jälkenä. [9, s.478] Apulausetta hyödynnetään todistettaessa Schurin lausetta.

Apulause 5.1. *Olkoot* $A, B \in M_n$ *ja* $x, y \in \mathbb{C}^n$. *Merkitään* $\operatorname{diag} x \in M_n$ *ja* $\operatorname{diag} y \in M_n$ *diagonaalimatriiseja, joiden diagonaalialkiot ovat vektoreista* x *ja* y. *Tällöin*

$$x^*(A \circ B)y = \operatorname{tr}((\operatorname{diag} \overline{x})A)(\operatorname{diag} y)B^T).$$

Todistus. (Vrt. [9], Lemma 7.5.2) Tehdään aluksi neliömatriisien $C, D \in M_n$ jälkeä koskeva havainto. Määritelmien 3.3 ja 3.10 mukaan matriisin CD^T jälki on

$$\operatorname{tr}(CD^T) = \operatorname{tr}\left(\sum_{k=1}^{n} c_{ik}d_{jk}\right) = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} c_{ik}d_{ik}\right) = \sum_{i,k=1}^{n} c_{ik}d_{ik} = \sum_{i,j=1}^{n} c_{ij}d_{ij}. \tag{5.2}$$

Kirjain k voidaan vaihtaa kirjaimeen j, sillä summan indeksointiin käytettävä kirjain voidaan vapaasti valita.

Olkoot $A = [a_{ij}], B = [b_{ij}], x = [x_i]$ ja $y = [y_i]$. Täten $(\operatorname{diag} \overline{x})A = [\overline{x}_i a_{ij}]$ ja $B \operatorname{diag} y = [b_{ij}y_j]$. Nyt käyttämällä yhtälöä (5.2) ja sen jälkeen matriisitulon määritelmää 3.3 kahdesti saadaan

$$\begin{aligned}\operatorname{tr}((\operatorname{diag} \overline{x})A(\operatorname{diag} y)B^T) &= \operatorname{tr}(((\operatorname{diag} \overline{x})A)(B \operatorname{diag} y)^T) \\ &= \sum_{i,j=1}^{n} (\overline{x}_i a_{ij})(b_{ij}y_j) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} \overline{x}_i a_{ij} b_{ij}\right) y_j \\ &= \left[\sum_{i=1}^{n} \overline{x}_i (a_{ij}b_{ij})\right] y = x^*(A \circ B)y.\end{aligned}$$

□

Lause 5.11 (Schurin lause)**.** *Olkoot* $A \in M_n$ *ja* $B \in M_n$ *positiivisesti (semi)definiittejä* *matriiseja. Tällöin Hadamardin tulo $A \circ B$ on positiivisesti (semi)definiitti.*

Todistus 1. (Vrt. [9], Theorem 7.5.3) Tehdään todistus kolmessa osassa:

(a) $A \circ B$ on positiivisesti semidefiniitti.

(b) jos A on positiivisesti definiitti ja matriisin B jokainen päädiagonaalin alkio on positiivinen, niin $A \circ B$ on positiivisesti definiitti.

(c) jos A ja B ovat positiivisesti definiittejä, niin $A \circ B$ on positiivisesti definiitti.

(a) Olkoot $A = [a_{ij}], B = [b_{ij}] \in M_n$ positiivisesti semidefiniittejä ja olkoon $x = [x_i] \in \mathbb{C}^n$ kompleksinen vektori. Myös matriisi $\overline{B}$ on positiivisesti semidefiniitti lauseen 4.15 perusteella ja siksi on mahdollista laskea $\overline{B}^{1/2}$ lauseen 4.20 mukaisesti. Merkitään $C = (\operatorname{diag} x)\overline{B}^{1/2}$, jolloin $C^* = \overline{B}^{1/2}(\operatorname{diag} \overline{x})$. Hyödyntämällä apulausetta 5.1 saadaan

$$\begin{aligned}
x^*(A \circ B)x &= \operatorname{tr}((\operatorname{diag} \overline{x})A(\operatorname{diag} x)B^T) && \mid \text{hermiittisille matriiseille } B^T = \overline{B} \\
&= \operatorname{tr}((\operatorname{diag} \overline{x})A(\operatorname{diag} x)\overline{B}) && \mid \overline{B} = \overline{B}^{1/2}\overline{B}^{1/2} \\
&= \operatorname{tr}((\operatorname{diag} \overline{x})A(\operatorname{diag} x)\overline{B}^{1/2}\overline{B}^{1/2}) && \mid \text{lause 3.8 kohta (e)} \\
&= \operatorname{tr}(\overline{B}^{1/2}(\operatorname{diag} \overline{x})A(\operatorname{diag} x)\overline{B}^{1/2}) \\
&= \operatorname{tr}(C^*AC).
\end{aligned}$$

Lause 4.3 osoittaa, että matriisi C^*AC on positiivisesti semidefiniitti, joten matriisin ominaisarvot ja jälki ovat epänegatiiviset lauseen 4.18 nojalla. Täten $x^*(A \circ B)x \geq 0$ jokaisella vektorilla $x \in \mathbb{C}^n$, joten $A \circ B$ on positiivisesti semidefiniitti määritelmän mukaan.
(b) Olkoon $\lambda_1 > 0$ matriisin A pienin ominaisarvo ja olkoon $\beta > 0$ matriisin B pienin diagonaalialkio. Olkoon $x = [x_i] \in \mathbb{C}^n$ nollasta eroava vektori. Lauseesta 3.15 seuraa, että

$$\lambda_i \in \sigma(A) \Leftrightarrow \lambda_i - \lambda_1 \in \sigma(A - \lambda_1 I) \text{ jokaisella } i = 1, \ldots, n.$$

Koska $\lambda_i - \lambda_1 \geq 0$, matriisin $A - \lambda_1 I$ ominaisarvot ovat epänegatiivisia. Lisäksi matriisi $A - \lambda_1 I$ on hermiittinen, sillä $(A - \lambda_1 I)^* = A^* - \overline{\lambda_1} I^* = A - \lambda_1 I$, ja näin ollen positiivisesti semidefiniitti lauseen 4.17 nojalla. Täten (a)-kohdasta seuraa, että matriisi $(A - \lambda_1 I) \circ B$ on positiivisesti semidefiniitti. Toisin sanoen

$$0 \leq x^*((A - \lambda_1 I) \circ B)x = x^*(A \circ B - \lambda_1 I \circ B)x = x^*(A \circ B)x - \lambda_1 x^*(I \circ B)x,$$

joten

$$\begin{aligned} x^*(A \circ B)x &\geq \lambda_1 x^*(I \circ B)x = \lambda_1 x^* \operatorname{diag}(b_{11}, \dots, b_{nn})x \\ &= \lambda_1 x^*[b_{ii}x_i] = \lambda_1 \sum_{i=1}^{n} \overline{x}_i b_{ii} x_i = \lambda_1 \sum_{i=1}^{n} b_{ii}|x_i|^2 \\ &\geq \lambda_1 \beta \sum_{i=1}^{n} |x_i|^2 = \lambda_1 \beta \|x\|_2^2 > 0, \end{aligned}$$

mikä todistaa (b)-kohdan väitteen.

(c) Jos matriisi B on positiivisesti definiitti, niin tutkimalla 1×1-kokoisia pääalimatriiseja lauseen 4.16 avulla nähdään, että matriisin B diagonaalialkiot ovat positiivisia. Väite seuraa kohdasta (b). □

Todistus 2. (Vrt. [15], Theorem 6.21) Oletetaan aluksi, että matriisit A ja B ovat positiivisesti definiittejä ja täten Kroneckerin tulo $A \otimes B$ on positiivisesti definiitti lauseen 5.6 nojalla. Lauseen 5.9 nojalla matriisi $A \circ B$ on Kroneckerin tulon $A \otimes B$ pääalimatriisi ja täten lauseen 4.16 mukaan positiivisesti definiitti. □

Schurin lauseesta seuraa runsaasti erilaisia epäyhtälöitä, jotka koskevat esimerkiksi determinantteja, ominaisarvoja tai diagonaalialkioita. Hadamardin epäyhtälö on eräs tällainen epäyhtälö. Se on hyvin keskeinen ja monet muut epäyhtälöt ovat yhtäpitäviä sen kanssa tai sen yleistyksiä. [9, s.505] Hadamardin epäyhtälö (5.3) voidaan kirjoittaa muodossa

$$1 \overset{n \text{ kpl}}{\cdots} 1 \det A \leq \det(I \circ A).$$

Tämä havainto voidaan yleistää muotoon

$$\max\{a_{11} \cdots a_{nn} \det B, b_{11} \cdots b_{nn} \det A\} \leq \det(A \circ B),$$

mikä on tehty lauseessa 5.13. [9, s.509] Seuraavaksi esitellään Hadamardin epäyhtälö ja edellä mainittu yleistys. Näiden avulla todistetaan Hadamardin tuloa koskeva mielenkiintoinen epäyhtälö

$$\det A \det B \leq \det(A \circ B).$$

Lause 5.12 (Hadamardin epäyhtälö)**.** *Olkoon matriisi* $A = [a_{ij}] \in M_n$ *positiivisesti definiitti. Tällöin*

$$\det A \leq a_{11} \cdots a_{nn}. \tag{5.3}$$

Lisäksi epäyhtälössä (5.3) *on yhtäsuuruus, jos ja vain jos* A *on diagonaalimatriisi.*

Todistus. (Vrt. [9], Theorem 7.8.1) Koska matriisi A on positiivisesti definiitti, niin sen diagonaalialkiot ovat positiivisia lauseen 4.16 nojalla. Olkoon $D = \operatorname{diag}(a_{11}^{1/2}, \dots, a_{nn}^{1/2})$.

Määritellään matriisi $C = D^{-1}AD^{-1} = \left[\frac{a_{ij}}{\sqrt{a_{ii}a_{jj}}}\right]$. Matriisi C on positiivisesti definiitti, sillä

$$x^*Cx = x^*D^{-1}AD^{-1}x = ((D^{-1})^*x)^*A(D^{-1}x) = (D^{-1}x)^*A(D^{-1}x) = y^*Ay > 0,$$

missä vektori $y = D^{-1}x \neq 0$. Matriisin C kaikki diagonaalialkiot ovat ykkösiä, joten $\operatorname{tr} C = n$. Olkoon $\lambda_1, \ldots, \lambda_n$ matriisin C ominaisarvot. Ominaisarvot ovat positiivisia, sillä matriisi C on positiivisesti definiitti. Tällöin aritmeettis-geometrinen epäyhtälö A takaa, että

$$\det C = \lambda_1 \cdots \lambda_n \leq \left(\frac{1}{n}(\lambda_1 + \cdots + \lambda_n)\right)^n = \left(\frac{1}{n}\operatorname{tr} C\right)^n = 1, \tag{5.4}$$

missä yhtäsuuruus on voimassa, jos ja vain jos $\lambda_i = 1$ jokaisella $i = 1, \ldots, n$. Koska matriisi C on hermiittinen ja näin ollen diagonalisoituva (similaarinen diagonaalimatriisin kanssa), niin lauseen 3.22 nojalla jokainen $\lambda_i = 1$, jos ja vain jos $C = I$. Täten

$$\det A = \det(DCD) = \det C(\det D)^2 = (\det C)a_{11} \cdots a_{nn} \overset{\text{epäyhtälö (5.4)}}{\leq} a_{11} \cdots a_{nn},$$

missä yhtäsuuruus toteutuu, jos ja vain jos $A = DCD = D^2 = \operatorname{diag}(a_{11}, \ldots, a_{nn})$. □

Apulause 5.2. *Olkoon $A = [a_{ij}] \in M_n$ positiivisesti semidefiniitti. Koska matriisi A on hermiittinen, se voidaan osittaa ja kirjoittaa muodossa $A = \begin{bmatrix} a_{11} & x^* \\ x & A_{22} \end{bmatrix}$, missä matriisi $A_{22} \in M_{n-1}$. Määritellään*

$$\alpha(A) = \begin{cases} \frac{\det A}{\det A_{22}}, & \text{jos } A_{22} \text{ on positiivisesti definiitti,} \\ 0, & \text{muulloin.} \end{cases}$$

Tällöin matriisi $\tilde{A} = \begin{bmatrix} a_{11} - \alpha(A) & x^* \\ x & A_{22} \end{bmatrix}$ on positiivisesti semidefiniitti.

Todistus. (Vrt. [9], Lemma 7.8.15) Jos matriisi A on singulaarinen, ei ole mitään todistettavaa, sillä tällöin $\det A = 0$, mistä seuraa, että $\tilde{A} = A$. Näin ollen voidaan olettaa, että matriisi A on positiivisesti definiitti. Täten jokainen matriisin A pääalimatriisi on myös positiivisesti definiitti lauseen 4.16 nojalla. Sovelletaan Sylvesterin kriteeriä 4.19 häntäpääalimatriiseihin. Häntäalideterminantit $\det(\tilde{A}[\{k, \ldots, n\}]) = \det A[\{k, \ldots, n\}]$ ovat positiivisia jokaisella $k = 2, \ldots, n$ lauseiden 4.16 ja 4.18 nojalla. Lisäksi

$$\det \tilde{A} = \det A - \alpha(A) \det A_{22} = \det A - \det A = 0.$$

Näin ollen matriisi $\tilde{A}$ on positiivisesti semidefiniitti lauseen 4.19 kohdan (b) nojalla. □

Lause 5.13. *Olkoot matriisit $A = [a_{ij}] \in M_n$ ja $B = [b_{ij}] \in M_n$ positiivisesti semidefiniittejä. Tällöin*

$$\max\{a_{11} \cdots a_{nn} \det B, b_{11} \cdots b_{nn} \det A\} \leq \det(A \circ B). \tag{5.5}$$

Todistus. (Vrt. [9], Theorem 7.8.16) Todistetaan väite hyödyntäen induktioperiaatetta. Jos $n = 1$,

$$\max\{a_{11} \det B, b_{11} \det A\} = a_{11}b_{11} = \det(A \circ B),$$

joten väittämä on tosi alkuarvolla. Tehdään sitten induktio-oletus, että väittämä on tosi, aina, kun $n = k \geq 1$, missä k on luonnollinen luku, ja todistetaan tämän avulla tapaus $n = k + 1$. Olkoon matriisi $\tilde{A}$ sama kuin edellisessä apulauseessa 5.2. Koska matriisi $\tilde{A} \in M_{k+1}$ on positiivisesti semidefiniitti, niin Schurin lauseen 5.11 nojalla matriisi $\tilde{A} \circ B \in M_{k+1}$ on positiivisesti semidefiniitti. Täten

$$\begin{aligned} 0 \leq \det(\tilde{A} \circ B) &= \det(A \circ B) - \det \begin{bmatrix} \alpha(A)b_{11} & 0 \\ * & A_{22} \circ B_{22} \end{bmatrix} \\ &= \det(A \circ B) - \alpha(A)b_{11} \det(A_{22} \circ B_{22}). \end{aligned}$$

Käyttämällä induktio-oletusta matriisin $A_{22} \circ B_{22} \in M_k$ determinanttiin saadaan

$$\det(A \circ B) \geq \alpha(A)b_{11}(b_{22} \cdots b_{k+1,k+1} \det A_{22}) = b_{11} \cdots b_{k+1,k+1} \det A.$$

Induktioperiaatteen nojalla matriiseille $A, B \in M_n$ on siis voimassa $b_{11} \cdots b_{nn} \det A \leq \det(A \circ B)$. Koska Hadamardin tulo on vaihdannainen, eli $A \circ B = B \circ A$, niin edeltävä epäyhtälö voidaan kirjoittaa myös muodossa $a_{11} \cdots a_{nn} \det A \leq \det(A \circ B)$. Näistä epäyhtälöistä voidaan muodostaa väittämän epäyhtälö (5.5). □

Lause 5.14. *(Vrt. [9, s.510], Exercise) Olkoon $A, B \in M_n$ positiivisesti semidefiniittejä matriiseja. Tällöin*

$$\det A \det B \leq \det(A \circ B).$$

Todistus. Oletetaan aluksi, että matriisit A ja B ovat molemmat positiivisesti definiittejä. Tällöin käyttämällä lauseita 5.12 ja 5.13 saadaan

$$\det A \det B \leq a_{11} \cdots a_{nn} \det B \leq \det(A \circ B).$$

Oletetaan sitten, että matriisit A ja B ovat positiivisesti semidefiniittejä. Oletetaan lisäksi, että ainakin toinen matriiseista ei ole positiivisesti definiitti. Olkoon tämä matriisi A tässä tapauksessa. Tällöin lauseesta 4.17 seuraa, että luku 0 on matriisin A ominaisarvo ja täten $\det A = 0$. Toisaalta Schurin lauseen 5.11 nojalla matriisi $A \circ B$ on positiivisesti semidefiniitti, joten lauseiden 3.17 ja 4.17 nojalla

$$\det A \det B = 0 \leq \lambda_1 \cdots \lambda_n = \det(A \circ B),$$

missä matriisin $A \circ B$ ominaisarvot $\lambda_1, \ldots, \lambda_n$ ovat epänegatiivisia. $\square$

6. SOVELLUSKOHDE: DIGITAALINEN KUVANKÄSITTELY

Digitaalisen kuvankäsittelyn oleellisimmat sovelluskohteet ovat kuvallisen informaation parantaminen ihmiselle helpommin tulkittavaksi, kuvadatan säilöminen ja siirto sekä kuvien esittäminen koneille hyödyllisellä tavalla [7, s.1]. Tässä luvussa kerrotaan suppeasti, kuinka digitaalinen kuva voidaan esittää matriisien avulla, esitellään tapa digitaalisen kuvan kirkkauden ja kontrastin säätämiseen sekä määritellään digitaalisten kuvien konvoluutio ja havainnollistetaan muutaman eri ytimen käyttöä konvoluutiossa. Kaikki luvussa esitellyt kuvankäsittelyn operaatiot ovat nykyajan kuvankäsittelyohjelmistojen ja älypuhelimien vakio-ominaisuuksia, vaikka niiden toteutus saattaakin erota tämän luvun toteutuksista. Lisäksi kuvankäsittelylle löytyy myös lukuisia sovelluskohteita teollisuuden ja tekniikan aloilla, kuten lääketieteen, tähtitieteen tai konenäön saralla. Liitteestä C löytyvät esitysjärjestyksessä kuvien käsittelyyn liittyvien ohjelmien koodirivit, jotka on kirjoitettu MATLAB® ohjelmiston Live Editorilla. MATLAB® ohjelmistossa kuvia voi näyttää komennolla ’imshow(image)’ ja tallentaa tiedostoiksi komennolla ’imwrite(image, ’filepath.png’). Nämä kuitenkin karsittu koodista pois selkeyden vuoksi. Lisäksi varsinkin jokaisen kuvan näyttäminen hidastaa ohjelmaa huomattavasti. Koodi ei ole missään määrin optimoitua, vaan sen tarkoituksena on vain havainnollistaa, kuinka jokin operaatio olisi mahdollista toteuttaa.

6.1 Digitaalisten kuvien esitysmuodot

Digitaalinen kuva voidaan esittää voidaan esittää äärellisenä 2-ulotteisena diskreettinä funktiona $f(x, y)$, jonka arvo vastaa kuvan kirkkautta tai väriä pisteessä (x, y). Nämä funktion arvot voidaan koota numeerista dataa sisältäväksi taulukoksi. Tämän taulukon alkioita sanotaan pikseleiksi tai kuvan alkioiksi ja yhdessä ne muodostavat digitaalisen kuvan. Jos funktion arvojoukko on 1-ulotteinen, kyseessä on matriisi. Pikselin ympärillä olevia pikseleitä nimitetään pikselin ympäristöksi. Esimerkiksi 3×3-kokoinen pikselin (x, y) ympäristö koostuu 9 pikselistä, jotka muodostavat neliön, jonka keskipisteenä on koordinaattien (x, y) pikseli. Lähes poikkeuksetta ympäristön dimensiot ovat parittomia. [11, s.8-10]

Binäärisessä kuvassa jokainen pikseli on joko musta tai valkoinen, eli 0 tai 1. Mustavalko-

1	1	1	1	1	1	1	1
1	1	0	1	1	0	1	1
1	1	1	1	1	1	1	1
1	0	0	0	0	0	0	1
1	0	1	1	1	1	0	1
1	1	0	1	1	0	1	1
1	1	1	0	0	1	1	1
1	1	1	1	1	1	1	1

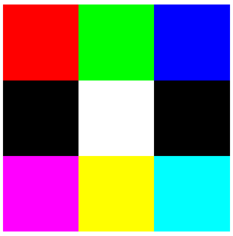
Kuva 6.1. *Binäärinen kuva muodostuu vain luvuista 0 ja 1. Lukua 0 vastaa musta pikseli ja lukua 1 valkoinen.*

kuvassa jokaisen pikselin arvo on kokonaisluku 0 ja 255 väliltä. Erilaisia harmaan sävyjä on siis $256 = 2^8$ kappaletta. Yhden pikselin esittämiseen tarvitaan 8 bittiä muistia, eli täsmälleen 1 tavu. [11, s.12-13] Tällaisesta kuvan datatyypistä käytetään nimitystä uint8 (unsigned 8-bit integer). Mustavalkokuvia voidaan vaihtoehtoisesti esittää muun muassa int8-datatyyppinä (signed 8-bit integer), eli kokonaislukuna väliltä $[-128, 127]$, uint16-datatyyppinä, eli kokonaislukuna väliltä $[0, 65535]$, tai esittää pikselit likulukuina väliltä $[0, 1]$. Useimmissa tapauksissa uint8 on kuitenkin riittävä ja tätä datatyyppiä tullaan pääosin työssä käyttämään.

Kuvassa 6.1 on havainnollistettu 8×8-kokoisen hymiön binääristä kuvaa ja sen matriisia. Matriisissa ei ole värejä eikä pikseleissä näy lukuarvoa. Kuvassa on siis ikään kuin ladottu binäärinen kuva ja matriisi päällekkäin. Toisaalta kyseessä voisi olla myös mustavalkokuva, jonka datatyyppinä käytetään liukulukua väliltä $[0, 1]$.

Värikuville tutuin esitysmuoto lienee RGB-muoto. Tässä muodossa jokaista pikseliä vastaa kolme lukua, jotka yhdessä antavat tietyn värin pikselille. Nämä luvut kertovat, paljonko pikselissä on punaista (R), vihreää (G) ja sinistä (B) väriä. Jos jokainen värikomponentti on datatyypiltään uint8, niin erilaisia värejä on $(2^8)^3 = 2^{24} = 16\ 777\ 216$, eli noin 16 miljoonaa. Tällaisesta kuvasta käytetään myös nimitystä 24-bittinen värikuva. [11, s.13] RGB-kuva voidaan esittää 2-ulotteisena vektoriarvoisena funktiona. Vaihtoehtoisesti RGB-kuvaa voidaan ajatella kolmen matriisin pinona, jossa jokainen matriisi vastaa yhtä värikanavaa [11, s.13].

Kuvassa 6.2 on esitetty 3×3-kokoinen RGB-kuva ja tämän värikanavia vastaavat matriisit. RGB-kuvassa on kaikki sellaiset kombinaatiot, joissa pikselissä väriä on joko mak-



255	0	0
0	255	0
255	255	0

0	255	0
0	255	0
0	255	255

0	0	255
0	255	0
255	0	255

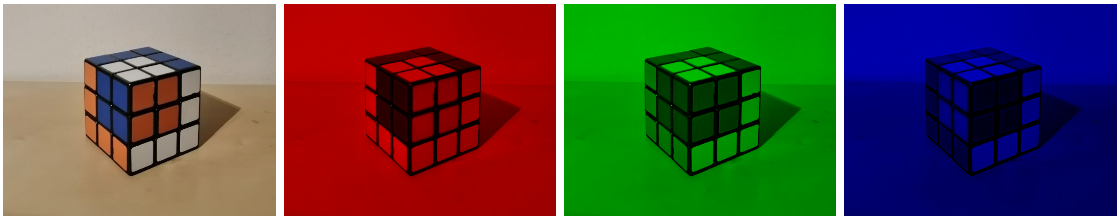
Kuva 6.2. *Ylhäällä 3x3-kokoinen RGB-kuva ja alhaalla sen RGB-värikanavat tässä järjestyksessä. [11, s.14]*

simimäärä 255 tai minimimäärä 0. Näitä värejä on yhteensä 8 kappaletta ja siksi musta esiintyy kahdesti. Kuvasta nähdään, kuinka mikäkin väri muodostuu. Esimerkiksi, että magenta saadaan RGB-arvoilla $[255, 0, 255]$, eli yhdistämällä punaisen ja sinisen värikanavan maksimiarvot sekä vihreän minimiarvo.

6.2 Alkeellisia kuvankäsittelyoperaatioita

Yksinkertaisillakin matriisien laskuoperaatioilla voidaan muokata digitaalisia kuvia. Yhteenlasku ja vastaavasti erotus voidaan suorittaa kahdelle samankokoiselle kuvalle samoin kuin Hadamardin tulo. Hadamardin tulo on erityisen mielenkiintoinen, kun se suoritetaan binäärisen kuvan ja RGB-kuvan eri värikanavien välillä. Tällöin binäärisen kuvan valkoisiin pikseleihin tulee RGB-kuvan pikselin väri. Tällaisen operaation avulla voidaan esimerkiksi rajata jokin muu kuvankäsittelyoperaatio tapahtuvaksi vain jollakin kuvan tietyllä alueella. RGB-kuvia käsiteltäessä jokaista värikanavaa käsitellään erikseen ja lopuksi värikanavat yhdistetään uudeksi RGB-kuvaksi.

Ohjelmassa 6.1 on esitetty, kuinka värikanavien eristys voidaan toteuttaa MATLAB® ohjelmalla. Aluksi ohjelmassa RGB-kuva luetaan muuttujaan 3-ulotteiseksi taulukoksi, jota voidaan ajatella kolmen matriisin pinona. Sen jälkeen alustetaan värikanavat samankokoisiksi taulukoiksi. Lopuksi RGB-kuvan värikanavat kopioidaan värikanaville alustettuihin muuttujiin. Jos värikanavat alustettaisiin matriiseiksi, niin MATLAB® näyttäisi kuvat mustavalkoisina. Saadaan siis näyttävämpi tulos, kun värikanavat esitetään taulukoina, kuten kuvasta 6.3 voi huomata. Tässä kuvassa on esitettynä suurikokoinen RGB-kuva Rubikin kuutiosta ja sen eri värikanavat eristettyinä. Mitä tummempi väri on, sitä vähemmän kyseistä väriä on.



Kuva 6.3. Vasemmalla alkuperäinen kuva ja tämän vieressä kuvasta eristetty punainen, vihreä ja sininen värikanava.

```
% Color Channel Extraction
% read image data to a variable
img=imread('rubik1.jpg');
% initializing variables R, G, B to proper sizes
R=uint8(zeros(size(img)));
G=R;
B=R;
% extracting color channels
R(:,:,1)=img(:,:,1);
G(:,:,2)=img(:,:,2);
B(:,:,3)=img(:,:,3);
```

***Ohjelma 6.1.** RGB-kuvan värikanavien eristys MATLAB® ohjelmistolla*

On olemassa eri tapoja muuntaa RGB-kuva mustavalkokuvaksi. Mustavalkokuvan pikselin arvo voitaisiin laskea esimerkiksi eri värikanavien keskiarvona. Tässä työssä muunnokseen käytetään samaa kaavaa kuin MATLAB® ohjelmiston 'im2gray' käyttää. Tämä kaava on

$$0.2989R + 0.5870G + 0.1140B, \tag{6.1}$$

missä R, G ja B ovat RGB-kuvan eri värikanavia vastaavat matriisit.

Eri ohjelmistoissa, kuten MATLAB® ohjelmistossa, jotkin laskuoperaatiot täytyy suorittaa samaa datatyyppiä oleville tekijöille. Tämä tarkoittaa sitä, että esimerkiksi uint8-datatyyppiä ja double-datatyyppiä oleville matriiseille ei voida suorittaa yhteenlaskua. Laskuoperaatioissa tarvitsee siis hyödyntää tyyppimuunnoksia, eli muunnoksia datatyypistä toiseen. Toisinaan MATLAB® onnistuu tekemään tyyppimuunnoksen automaattisesti. Laskuoperaatiot toimivat samoin uint8-datatyypillä kuin kokonaisluvuilla, mutta jos tulos ei ole kokonaisluku väliltä $[0, 255]$, niin tulos pyöristetään lähimpään tämän välin kokonaisluvuista. Toisin sanoen jos luku a on laskuoperaation tulos ja $a < 0$ tai $a > 255$, niin uint8(a)$= 0$ tai uint8(a)$= 255$ tässä järjestyksessä. Muutoin luku a pyöristyy lähimpään kokonaislukuun ja uint8(0.5)$= 1$.

Matriisien yhteenlaskun ja skalaarikertolaskun avulla voidaan säätää kuvan kirkautta.

Seuraavan yhtälön avulla voidaan säätää kuvan kirkkautta:

$$B = aA + bJ, \tag{6.2}$$

missä A on alkuperäisen kuvan matriisi, B on muokatun kuvan matriisi, J on ykkösmatriisi ja a sekä b ovat skalaareja. Parametreja a ja b muuttamalla voidaan vaikuttaa muokatun kuvan B kirkautteen. Mitä suurempia arvoja parametreille annetaan, sitä kirkkaampi kuva saadaan. Lisäksi yhtälöllä (6.2) voidaan tehdä negatiivikuva, kun valitaan $a = -1$ ja $b = 255$.

Toinen tapa parantaa hämärässä otettuja kuvia on kuvan kontrastin säätäminen. Tämä voidaan tehdä esimerkiksi gammakorjauksella. Gammakorjaus määritellään seuraavasti:

$$B = aA^{\circ\gamma}, \tag{6.3}$$

missä A on alkupeäisen kuvan matriisi, B on muokatun kuvan matriisi, $^\circ$ on potenssi alkioittain (MATLAB® ohjelmistossa .^) ja luvut a ja γ ovat positiivisia. Jos $\gamma > 1$, niin karkeasti kuvailtuna kuvan tummista alueista tulee entistä tummempia, mikä kasvattaa kontrastia. Vastaavasti jos $\gamma < 1$, niin kuvan tummista alueista tulee vaaleita, mikä vähentää kontrastia. Mitä kauempana luku γ on luvusta 1, sitä voimakkaampi on muutos. Gammakorjausta varten pikselien arvot muunnetaan välille $[0, 1]$ välin $[0, 255]$ sijaan.

Ohjelmassa 6.2 hämärässä otettua kuvaa 'lowlight.jpg' käsitellään yhtälöiden (6.2) ja (6.3) avulla eri parametrein. Ohjelmassa on huomioitu tarvittavat tyyppimuunnokset.

```
% Brightness and contrast
% read image
img=imread('lowlight.jpg');
% brighten the image
I1=uint8(1.1*img+60*uint8(ones(size(img))));
% gamma correction
A=double(img)*1/255;
I2=uint8((A.^0.5)*255);
% brighten the gamma corrected image
I3=I2+uint8(25*ones(size(I)));
% small gamma
I4=uint8((A.^0.1)*255);
% big gamma
I5=uint8((A.^5)*255);
```

Ohjelma 6.2. *RGB-kuvan kirkkauden ja kontrastin säätö MATLAB® ohjelmistolla*

Kuvan kirkkauden ja kontranstin säätöä on havainnollistettu kuvassa 6.4. Tässä kuvassa ylimällä rivillä vasemmalla on alkuperäinen kuva, joka on otettu huonossa valaistukses-

Kuva 6.4. *Hämärässä otetun kuvan kirkkauden ja kontrastin säätö. Ylhäällä vasemmalla on alkuperäinen kuva, keskellä kuvan kontrastia ja kirkkautta säädetty ja oikealla pelkästään kirkkautta säädetty. Alhaalla vasemmalla gammakorjaus sopivalla, keskellä liian pienellä ja oikealla liian suurella luvulla γ.*

sa. Keskellä olevalle kuvalle on aluksi tehty gammakorjaus (6.3) parametrein $a = 1$ ja $\gamma = 0.5$ ja tämän jälkeen kirkkautta säädetty yhtälöllä (6.2) parametrein $a = 1$ ja $b = 25$. Toisin kuin alkuperäisestä kuvasta tästä kuvasta erottaa selkeästi, että kyseessä on hautausmaalta otettu kuva. Hautakivet ja vasemmassa reunassa oleva tie tulevat selkeästi näkyviin.

Kuvassa 6.4 ylimmällä rivillä oikealla alkuperäisen kuvan kirkkautta on säädetty yhtälöllä (6.2) parametrein $a = 1.1$ ja $b = 60$ ja alhaalla vasemmalla olevan kuvan kontrastia on muutettu gammakorjauksella (6.3) parametrein $a = 1$ ja $\gamma = 0.5$. Molemmissa kuvissa alkuperäinen kuva selkeytyy huomattavasti. Pelkällä kirkkauden säädöllä saadaan kuitenkin aikaiseksi samea ja usvainen kuva. Kuvan 6.4 alimman rivin kahdessa viimeisessä kuvassa on esitetty, mitä kuvalle tapahtuu, jos valitaan vääränlainen γ. Keskellä on valittu hyvin pieni $\gamma = 0.1$. Kuvasta erottaa yksityiskohtia, mutta kuvan värimaailma on täysin luonnoton. Oikealla $\gamma = 5$, jolloin kuvan kontrastia lisätään. Tällöin kuvassa näkyvät enää vain valot.

6.3 Digitaalisten kuvien konvoluutio

Konvoluutio on tärkeä operaatio digitaalisessa kuvankäsittelyssä. Sitä hyödynnetään esimerkiksi konvoluutioneuroverkoissa, eli nimensä mukaan neuroverkoissa, jotka käyttävät konvoluutiota. Tällaisia neuroverkkoja voidaan puolestaan käyttää vaikkapa esineiden ja asioiden tunnistamiseen ja luokitteluun kuvista. Konvoluutiolla kuvasta voidaan poistaa häiriöitä tai erottaa tiettyjä piirteitä, ominaisuuksia ja malleja kuvista, joiden avulla niiden automaattinen tunnistus ja luokittelu onnistuu.

6.3.1 Määritelmästä

Korrelaatiossa ydin (engl. kernel), eli eräänlainen pieni (verrattuna kuvamatriisiin) matriisi, keskitetään kuvamatriisin päälle siten, että ytimen keskipiste ja kuvamatriisin alkio ovat päällekkäin. Sitten lasketaan Hadamardin tulo alle jäävien kuvamatriisin alkioiden kanssa ja näin saadun matriisin alkiot summataan uudeksi alkioksi korreloituun kuvamatriisiin, josta käytetään myös nimitystä suodatettu kuvamatriisi. Tämän uuden alkion koordinaatit ovat samat kuin ytimen keskipisteen alla olevan kuvamatriisin alkion. Kun uusi alkio on saatu selville ydintä siirretään kuvamatriisin päällä seuraavaan kohtaan, lasketaan Hadamardin tulo ja summataan alkiot uudeksi alkioksi suodatettuun matriisiin. Tätä toistetaan, kunnes kaikki mahdolliset kohdat kuvamatriisista on käyty läpi. Toisin sanoen ydintä ikään kuin liutetaan kuvamatriisin yllä ja tehdään edellä mainitut operaatiot.

Konvoluutio toimii vastaavasti, mutta siinä ydintä kierretään 180° ennen operaatiota. Se on siis korrelaatio, jossa ydintä on kierretty. Yleensä ydin on dimensioiltaan pariton. Periaatteessa ydin voitaisiin määritellä dimensioiltaan myös parilliseksi, mutta dimensioiden ollessa parittomia on ydin huomattavasti yksinkertaisempaa keskittää. [7, s.148-150]

Itse asiassa konvoluution täsmällisessä määritelmässä ydin asetettaisiin kuvamatriisin päälle myös kaikkiin sellaisiin kohtiin, joissa ainakin yksi ytimen ja kuvamatriisin alkio on kohdakkain. Ei siis vain kohtiin, joissa ytimen keskipiste ja kuvamatriisin alkio ovat päällekkäin. Täsmällinen määritelmä tuottaa kuitenkin alkuperäistä kuvamatriisia suuremman tuloksen, joka täytyisi rajata sopivan kokoiseksi. Tämä nähdään esimerkiksi teoksen [7] kuvista figure 3.29 ja figure 3.30. On kuitenkin suoraviivaisempaa toteuttaa konvoluutio siten, että lopputulosta ei tarvitse rajata.

On melko helppo huomata, että kun konvoluutiossa käsitellään reunapikseleitä, niin ytimessä on alkioita, joiden alla ei ole kuvamatriisin alkioita. Näiden reunapikseleiden käsittelyyn on eri vaihtoehtoja. Eräs tapa on tehdä nollareunus (engl. zero padding), eli lisätä reunoille nollarivejä ja -sarakkeita. Muita tapoja on muun muassa monistaa reunapikseleiden arvot nollien tilalle tai jättää tällaiset kohdat pois konvoluutiosta, jolloin suodatetusta kuvasta tulee alkuperäistä pienempi. [7, s.146-147] Alla esitetty työssä käytettävä konvoluution määritelmä, jossa reunat käsitellään nollareunuksen avulla ja konvoluution tulos on alkuperäisen kuvan kanssa samankokoinen.

Määritelmä 6.1 (Digitaalisen kuvan konvoluutio)**.** Olkoon kuvaa tai sen värikanavaa vastaava matriisi $A \in M_{m,n}$ ja $K \in M_{p,q}$ ydin, missä kokonaisluvut p ja q ovat parittomia. Merkitään $d_p = \left\lfloor \frac{p}{2} \right\rfloor$ ja $d_q = \left\lfloor \frac{q}{2} \right\rfloor$. Lisäksi olkoon matriisin A nollareunustettu matriisi

$$N = \begin{bmatrix} O & O & O \\ O & A & O \\ O & O & O \end{bmatrix} \in M_{m+2d_p,n+2d_q}.$$

Ytimen K alle jäävä alimatriisi on muotoa

$$N_{ij} = N[\{i - d_p, i - d_p + 1, \dots, i + d_p\}, \{j - d_q, j - d_q + 1, \dots, j + d_q\}].$$

Tällöin matriisin A ja ytimen K konvoluutiota merkitään $B = A{*}K \in M_{m,n}$. Konvoluution tuloksen, eli suodatetun matriisin B, alkiot ovat muotoa

$$B_{i,j} = (A * K)_{ij} = \sum_{k,h}^{p,q} (N_{i+d_p, j+d_q} \circ EKE)_{kh},$$

missä matriisi $E = \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}$ kiertää ydintä $180°$.

Kuvassa 6.5 on hahmoteltu konvoluution vaiheita hyvin yksinkertaiselle matriisille. Aluksi matriisiin lisätään nollareunus, jonka koon määrää ytimen dimensiot. Tässä tapauksessa riittää lisätä kaksi riviä ja kaksi saraketta. Seuraavaksi konvoluutiossa kierretään ydintä $180°$ ja asetetaan sen keskipiste alkuperäisen kuvamatriisin ensimmäisen alkion kohdalle. Alle jäävän nollareunusmatriisin kanssa lasketaan Hadamardin tulo ja näin saadun matriisin alkiot summataan uudeksi alkioksi suodatetun kuvan matriisiin (konvoluution tulos). Tämä uusi alkio on kehystetty punaisella reunauksella. Sitten ydintä siirretään yhden sarakkeen verran oikealle ja sama operaatio toistetaan. Näin jatketaan, kunnes koko nollareunustettu matriisi on käyty läpi. Korrelaatio antaa saman tuloksen kuin konvoluutio paitsi, että tulosta on kierretty $180°$.

Nollareunus tekee suodatettuun kuvaan kehysmäisen reunan, jos reunan pikselit eivät ole mustia. Ydin, kuva ja näiden koot vaikuttavat siihen, kuinka selkeä kehys muodostuu. Suurilla kuvilla se voi olla huomaamaton, ellei alkuperäisen kuvan reuna ole lähes valkoinen. Kuvaa suurennettaessa tai pienillä kuvilla tämä on kuitenkin havaittavissa, kuten myöhemmin esiteltävistä konvoluutioista voi havaita.

Tässä työssä ytimet ovat neliömatriiseja yksinkertaisuuden vuoksi. Sanotaan, että ydin on symmetrinen, jos sitä kuvaava matriisi on keskiösymmetrinen (engl. centrosymmetric), eli jos ytimelle $K \in M_p$ ja määritelmän 6.1 matriisille $E \in M_p$ on voimassa $KE = EK$. Toisin sanoen ydin on symmetrinen keskipisteensä suhteen. Symmetrisillä ytimillä korrelaatio ja konvoluutio antavat saman tuloksen ja tässä tapauksessa määritelmästä 6.1 voidaan unohtaa matriisi E. Joskus ytimen kierto saadaan tehtyä huomattavasti yksinkertaisemmin, esimerkiksi kertomalla luvulla -1.

Konvoluutio identiteettiytimellä

$$id = \begin{bmatrix} O & O & O \\ O & 1 & O \\ O & O & O \end{bmatrix}$$

0	0	0	0	0
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	0	0

***(a)** Alkuperäistä kuvaa vastaava matriisi.*

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

***(b)** Nollareunustettu matriisi.*

1	2	3
4	5	6
7	8	9

***(c)** Konvoluutiossa käytettävä ydin.*

9	8	7	0	0	0	0
6	5	4	0	0	0	0
3	2	1	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

***(d)** Konvoluution lähtötilanne.*

0	0	0	0	0
0	1	2	3	0
0	4	5	6	0
0	7	8	9	0
0	0	0	0	0

***(e)** Konvoluution tulos.*

0	0	0	0	0
0	9	8	7	0
0	6	5	4	0
0	3	2	1	0
0	0	0	0	0

***(f)** Korrelaation tulos.*

***Kuva 6.5.** Hahmotelma konvoluution vaiheista. Kuvassa näkyy alkuperäinen kuva, nollareunus, ydin, konvoluution lähtötilanne sekä konvoluution ja korrelaation lopputulokset. [7, s.149]* antaa tulokseksi alkuperäisen matriisin. Muuttamalla ytimen kokoa, muotoa ja alkioita voidaan vaikuttaa siihen, millainen lopputulos konvoluutiosta saadaan. Kirjallisuudessa ytimelle voidaan käyttää myös nimityksiä suodatin (engl. filter) tai peite (engl. mask).

Ohjelmassa 6.3 on määritelmän 6.1 mukainen konvoluution suorittava funktio. Funktiossa muodostetaan aluksi määritelmän 6.1 vaatimat muuttujat. Funktio olettaa ytimen olevan neliömatriisi ja tämän lisäksi jo valmiiksi kierretty. Konvoluutio itsessään tapahtuu asettamalla suodatettuun matriisiin 'conv' jokainen alkio yksitellen silmukoiden avulla. Koodissa indeksi i kuvaa riviä ja j saraketta.

```
function newImg = convolution(img, kernel)
    % This function does convolution. As a matter of fact it does
    % correlation, since it assumes the user has already flipped
    % the kernel.
    K=kernel;
    % variables needed for zeropadding
    d=floor(size(K,1)/2);
    dim=size(img);
    % zeropadding
    A=[zeros(d,dim(2)+2*d);
       zeros(dim(1),d) img zeros(dim(1),d);
       zeros(d,dim(2)+2*d)];
    % initialize the conv variable
    conv=zeros(size(img));
    % convolution
    for i=1+d:dim(1)+d-1
        for j=1+d:dim(2)+d-1
            % sum of the hadamard product
            conv(i-d+1,j-d+1)=sum(K.*double(A(i-d:i+d,j-d:j+d)),
                                  'all');
        end
    end
    % return convoluted image
    newImg=uint8(conv);
end
```

Ohjelma 6.3. *Konvoluutiofunktio MATLAB® ohjelmistolla*

6.3.2 Kuvaa sumentavan ytimen konvoluutio

Jos kuvassa esiintyy häiriöitä, voidaan häiriöt yrittää suodattaa pois ytimellä, joka laskee keskiarvon ympäröivistä pikseleistä. Tällaisesta ytimestä käytetään nimitystä keskiarvoydin ja se on muotoa

$$K = \frac{1}{k^2}\begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \in M_k. \tag{6.4}$$

Toinen käyttötarkoitus keskiarvoytimelle on kuvan sumentaminen (engl. blurring), josta käytetään myös termiä pehmentäminen (engl. smoothing). Keskiarvoydin vaikuttaa erityisesti pikseleihin, joiden ympäristössä on suuria eroja pikselien arvoissa. Tällaisia kohtia ovat esimerkiksi kuvassa esiintyvät ääriviivat, kuten myöhemmin kuvan terävöinnin yhteydessä tullaan huomaamaan.

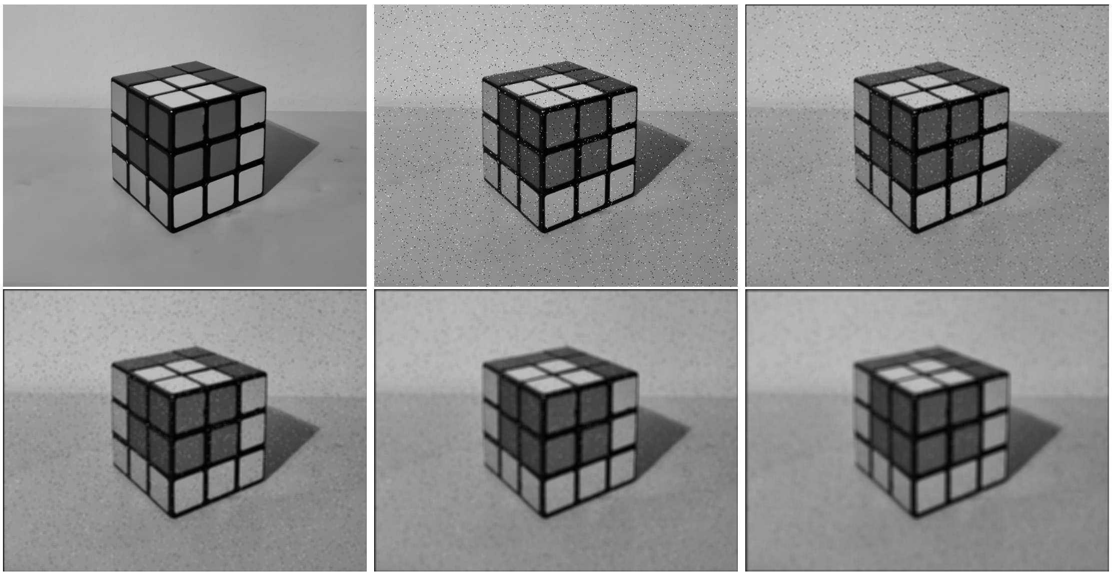
Ohjelmassa 6.4 on aluksi tehty luetusta kuvasta mustavalkokuva kaavalla (6.1). Tämän jälkeen on luotu keinotekoisesti häiriötä kuvaan. Lopuksi on suoritettu konvoluutio eri kokoisilla keskiarvoytimillä silmukassa hyödyntäen funktiota 6.3.

```
% Convolution Blur
img = imread('rubik1.jpg');
% make grayscale image
img=0.2989*img(:,:,1) + 0.5870*img(:,:,2) + 0.1140*img(:,:,3);
% add noise to image
imgNoise=imnoise(img,'salt & pepper',0.02);
% loop over different kernel sizes
for i=[3 5 7 9]
    kernel=1/i^2*ones(i);
    blur=convolution(imgNoise, kernel);
end
```

Ohjelma 6.4. *Kuvan sumentaminen konvoluutiolla MATLAB® ohjelmistolla*

Kuvassa 6.6 vasemmalla ylhäällä on alkuperäinen mustavalkokuva Rubikin kuutiosta ja tämän vieressä kuva, johon on keinotekoisesti lisätty häiriötä, eli tässä tapauksessa mustia ja valkoisia pisteitä. Seuraavat kuvat ovat summenuksen tuloksia yhtälön (6.4) keskiarvoytimellä, missä $k = 3, 5, 7, 9$. Ytimen koot ovat kuvissa suuruusjärjestyksessä. Mitä suurempi ydin, sitä paremmin ydin suodattaa häiriöitä pois, mutta samalla kuva sumenee enemmän. Ytimillä $k = 7$ ja $k = 9$ suodatettuja kuvia tarvitsee suurentaa, jos haluaa havaita häiriöitä.

Lisäksi summenus tai mikä tahansa muu konvoluutio voidaan tehdä vain tietylle alueelle kuvassa esimerkiksi Hadamardin tulon avulla. Aluksi kuvamatriisin $A_{m,n}$ kanssa yhtäsuuri ykkösmatriisi $J_{m,n}$ jaettaisiin kahdeksi sellaiseksi binääriseksi matriisiksi H_A ja H_B, että $H_A + H_B = J_{m,n}$ ja että matriisissa H_B summenettavan alueen alkiot ovat ykkösiä ja muut nollia. Hadamardin tulo binäärisen matriisin kanssa tuottaa alkuperäisen kuvan vain niihin pikseleihin, jotka ovat binäärisessä kuvassa valkoisia, eli ykkösiä. Alueellinen sumennus



Kuva 6.6. *Ylärivin kaksi ensimmäistä kuvaa ovat alkuperäinen kuva ja häiriötä sisältävä kuva. Seuraavat neljä esittävät konvoluution tuloksia häiriötä sisältävälle kuvalle erisuurilla keskiarvoytimillä.*

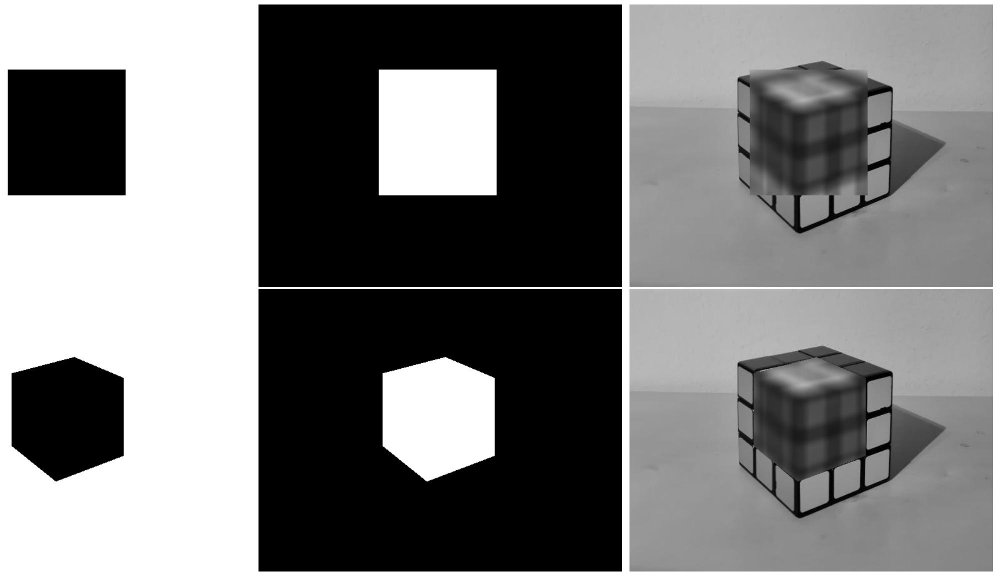
saadaan kaavalla

$$C = H_A \circ A + H_B \circ B, \tag{6.5}$$

missä matriisi A on alkuperäinen kuvamatriisi ja matriisi B on sumennettu kuvamatriisi.

Ohjelmassa 6.5 on toteuttu alueellinen sumennus kaavan (6.5) avulla. Koska luettavat kuvat 'rubik_sq.jpg' ja 'rubik_fs.jpg' ovat .jpg-tiedostomuodossa, ne ovat oletusarvoisesti RGB-kuvia, vaikka jokainen pikseli on joko musta tai valkoinen. Täten jokainen värikanava on sama. Värikanavan jokainen alkio on puolestaan joko 0 tai 255 ja täten kertomalla luvulla $\frac{1}{255}$ mikä tahansa värikanava saadaan binäärinen kuva. Lopuilla riveillä on toteutettuna edellä mainittu kaava (6.5). Tässäkin ohjelmassa käytetään konvoluutiofunktiota 6.3.

```
% Region blur
% read images
img=imread('rubik1.jpg');
H_A1=imread('rubik_sq.jpg');
H_A2=imread('rubik_fs.jpg');
% make grayscale image
img=0.2989*img(:,:,1) + 0.5870*img(:,:,2) + 0.1140*img(:,:,3);
% make binary images of regions to be blurred and their negatives
H_A1=1/255*H_A1(:,:,1);
H_A2=1/255*H_A2(:,:,1);
H_B1=ones(size(H_A1))-double(H_A1);
H_B2=ones(size(H_A2))-double(H_A2);
```



Kuva 6.7. *Ylärivillä esitelty alueelliseen sumentaminen suorakulmion muotoiselle alueelle ja alarivillä vapaamuotoiselle alueelle. Molempien rivien kuvat vasemmalta oikealla ovat* H_A, H_B *ja sumennuksen lopputulos* C.

```
% blur image
B=convolution(img,1/25^2*ones(25));
% make regional blurs
blur_sq=uint8(H_A1).*img+uint8(H_B1).*B;
blur_fs=uint8(H_A2).*img+uint8(H_B2).*B;
```

Ohjelma 6.5. *Tietyn alueen sumentaminen kuvasta konvoluutiolla MATLAB® ohjelmistolla*

Kuvassa 6.7 on esitelty kaavan (6.5) käyttöä jo edellä esiteltyyn Rubikin kuution kuvaan (ks. kuva 6.6). Tarkoituksena on sumentaa Rubikin kuutiossa näkyvän pienemmän kuution alue. Ylärivillä on alueelliseen sumennukseen on käytetty suorakulmion muotoista aluetta ja alarivillä hieman vapaamuotoisempaa, kuusikulmion muotoista aluetta. Rivien ensimmäinen ja toinen kuva ovat binäärisiä ja ne vastaavat matriiseja H_A ja H_B tässä järjestyksessä. Kuva H_A on saatu piirtämällä digitaalisesti alkuperäisen kuvan kokoiselle valkoiselle pohjalle musta kuvio ja kuva H_B on saatu laskemalla $H_B = J_{m,n} - H_A$. Periaatteessa tämä kuvio voisi olla minkä muotoinen tahansa. Rivin viimeinen kuva vastaa kaavan (6.5) matriisia C. Sumennettu matriisi B on saatu 25×25-kokoisella keskiarvoytimellä.

Sumentamiseen voitaisiin käyttää vaihtoehtoisesti Gaussin ydintä, joka muodostetaan Gaussin funktion avulla. Gaussin 2-uloitteinen funktio on $f(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$, missä σ on keskihajonta. Mitä suurempi keskihajonta on, sitä litteämpi funktion kuvaaja on. Gaussin funktion diskreetin approksimaation avulla voidaan muodostaa Gaussin ydin. [11, s.100-

101] Gaussin funktion diskreettiversio on

$$G(x,y) = \frac{f(x,y)}{\sum_x \sum_y f(x,y)}. \tag{6.6}$$

Ytimen keskipisteen ajatellaan olevan origo. Täten 3×3-kokoisessa Gaussin ytimessä diskreettien muuttujien otosavaruudet ovat $\Omega_x = \{-1, 0, 1\}$ ja $\Omega_y = \{-1, 0, 1\}$. Sijoittamalla jokaisen ytimen alkion koordinaatit kaavaan (6.6) saadaan 3×3-kokoinen Gaussin ydin

$$\begin{bmatrix} 0.0751 & 0.1238 & 0.0751 \\ 0.1238 & 0.2042 & 0.1238 \\ 0.0751 & 0.1238 & 0.0751 \end{bmatrix}.$$

Gaussin ydin antaa samanlaisen tuloksen kuin aiemmin esitetty keskiarvoydin, mutta Gaussin ydin painottaa origoa ja sen viereisiä pikseleitä enemmän kuin kulmien pikseleitä.

6.3.3 Kuvan terävöinti sumentavan ytimen avulla

Konvoluution avulla kuvia voidaan myös terävöittää, eli korostaa kuvassa esiintyviä ääriviivoja. Tämä tehdään vähentämällä skaalattu ja sumennettu kuva alkuperäisestä kuvasta. Tällaisen operaation ydin voidaan muodostaa seuraavasti:

$$K = s(id - \frac{1}{k}a), \tag{6.7}$$

missä id on identiteettiydin, a on sumentava ydin ja luvut k ja s ovat skaalauskertoimia. Jotta muokatussa kuvassa olisi sama kirkkaustaso kuin alkuperäisessä täytyy kaikkien matriisiin K alkioiden summa olla 1. Tämä tehdään kertoimella s, joten yhtälön

$$s\left(1 - \frac{1}{k}\right) = 1 \tag{6.8}$$

on oltava voimassa. Sijoittamalla $\frac{s}{k} = t$ edelliseen yhtälöön saadaan, että $s = t + 1$. Täten yhtälö (6.7) voidaan kirjoittaa muodossa

$$K = id + t(id - a). \tag{6.9}$$

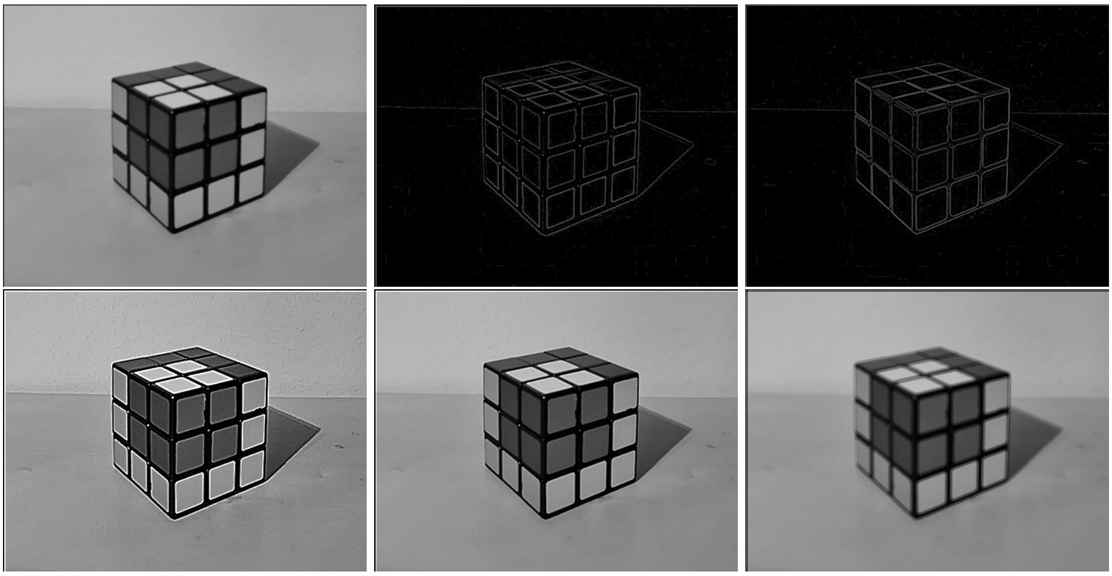
Tämä ydin antaa terävöidyn kuvan, kun $t > 0$. Yhtälössä id vastaa alkuperäistä kuvaa ja $t(id - a)$ skaalattuja ääriviivoja. Mitä suurempi luvun t arvo on yhtälössä (6.9) on, sitä enemmän ääriviivoja korostetaan. [11, s.104-106] Ääriviivat saadaan erotettua kuvasta, vaikka parametri t olisi negatiivinen. Tällöin yhtälön (6.9) ydin antaa kuitenkin tärähtäneeltä ja sumealta näyttävän kuvan. Ero ytimillä $1 \cdot (id - a)$ ja $-1 \cdot (id - a)$ poimituissa ääriviivoissa on hienovarainen, mutta olemassa. Näistä ensimmäinnen ikään kuin poimii ääriviivojen sisäreunat ja jälkimmäinen ulkoreunat. Vähentämällä alkuperäisestä kuvasta ytimellä $-1 \cdot (id - a)$ poimitut ääriviivat, saadaan terävöitetty kuva, jossa ääriviivojen ulkoreunat korostuvat mustalla.

Ohjelmassa 6.6 on toteutettu kuvan terävöinti sumennuksen avulla. Muuttujiin edges1 ja edges2 on tallennettu kuvan ääriviivojen sisä- ja ulkoreunat. Koodissa ääriviivojen kontrastia säädetään kuvaa 6.8 varten, jotta ääriviivat näkyvät selvemmin. Muuttujiin conv1, conv2 ja conv3 tallennetaan konvoluution tulokset yhtälön (6.9) parametrin arvoilla $t = -5, 5, 20$ muodostetuilla ytimillä.

```
% Convolution, Unsharp Masking
% read image and convert it to grayscale
img=imread('rubik1.jpg');
img=0.2989*img(:,:,1) + 0.5870*img(:,:,2) + 0.1140*img(:,:,3);
% create a gaussian 3x3 kernel
a=fspecial("gaussian",[3,3],1);
% create unsharp kernel
id=[ 0 0 0;0 1 0; 0 0 0];
edge_kernel=id-a;
% extract the edges to be sharpened from image
edges1=convolution(img, edge_kernel);
edges2=convolution(img, -edge_kernel);
% modify the contrast of edges
A=double(edges1)*1/255;
edges1=uint8((A.^0.35)*255);
A=double(edges2)*1/255;
edges2=uint8((A.^0.35)*255);
% sharpened image with t=5
conv1=convolution(img, id+5*edge_kernel);
% sharpened image with t=-5
conv2=convolution(img, id-5*edge_kernel);
% sharpened image with t=20
conv3=convolution(img, id+20*edge_kernel);
```

Ohjelma 6.6. *Kuvan terävöinti konvoluutiolla MATLAB® ohjelmistolla*

Kuvassa 6.8 on havainnollistettu digitaalisen kuvan terävöintiä. Se on tehty yhtälöllä (6.9), summenukseen käytetty ydin a on Gaussin 3×3-kokoinen ydin keskihajonnalla $\sigma = 1$. Ylhäällä alkuperäisen kuvan oikealla puolella on kuvat, ääriviivojen sisä- ja ulkoreunoista, jotka on saatu ytimillä $1 \cdot (id - a)$ ja $-1 \cdot (id - a)$. Ääriviivakuvien kontrastia on muutettu, jotta ääriviivat ovat selkeämmin havaittavissa. Alhaalla on yhtälön (6.9) parametrilla $t =$



Kuva 6.8. Kuvan terävöittäminen sumennuksen avulla. Ylhäällä alkuperäinen kuva ja kuvan ääriviivojen sisä- ja ulkoreunat korostettuna. Alhaalla yhtälön (6.9) *tulokset parametrin arvoilla $t = 20, 5, -5$ tässä järjestyksessä.*

$20, 5, -5$ tuotetut kuvat tässä järjestyksessä. Parametrilla $t = 20$ kuvasta tulee paljon terävämpi kuin alkuperäisestä, jopa hieman yliterävöitetty. Osa Rubikin kuution palojen ääriviivoista hehkuu luonnottoman valkoisena. Lisäksi taustalla olevan seinän ja Rubikin kuution alla olevan pöydän yksityiskohdat ja muodot korostuvat. Parametrilla $t = 5$ kuvan tausta säilyy tasaisena. Terävöinnin huomaa ennen kaikkea vertamalla Rubikin kuution varjoa ja palikoiden reunoja alkuperäiseen kuvaan. Parametrilla $t = -5$ kuvasta tulee alkuperäistä sumeampi ja tärähtäneen näköinen.

6.3.4 Konvoluutio RGB-kuville

Konvoluutio voidaan suorittaa myös RGB-kuville. Tällöin kovoluution määritelmää 6.1 käytetään jokaiseen värikanavaan erikseen ja lopuksi konvoluutioiden tulokset yhdistetään uudeksi RGB-kuvaksi.

Ohjelmassa 6.7 on toteutettu konvoluutiofunktio RGB-kuville. Tämä funktio eristää aluksi RGB-kuvan värikanavat ja syöttää ne sitten yksitellen aiemmin esitetylle konvoluutiofunktiolle 6.3. Lopuksi konvoloiduista värikanavista muodostetaan RGB-kuva, jonka funktio palauttaa.

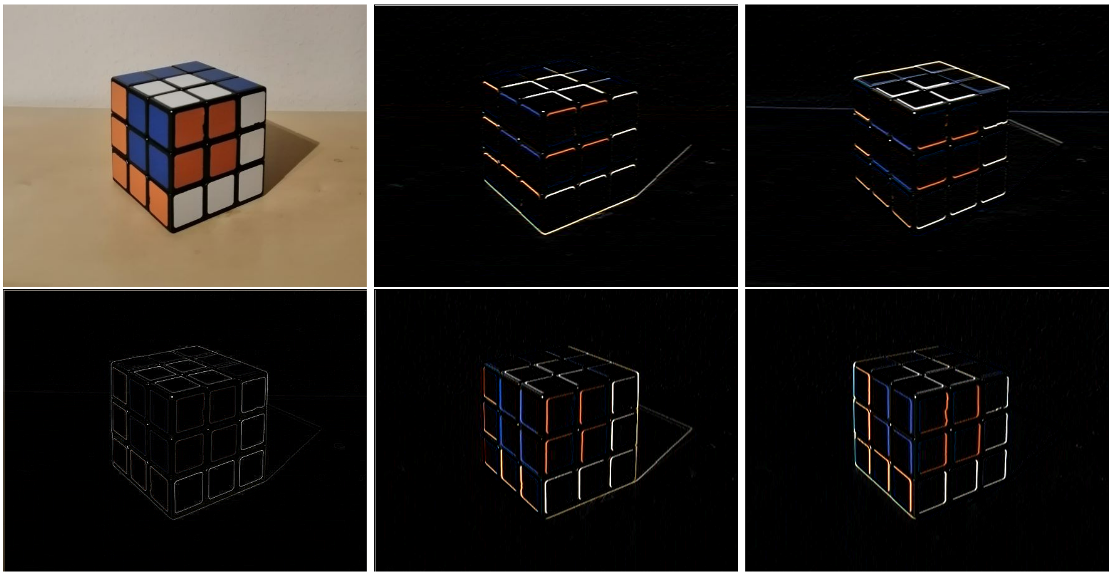
```
function newImg = colorConvolution(img, kernel)
    % This function extract different color channels from
    % RGB-image and convolutes them one by one with
    % convolution-function before combining them and returning
    % the color convoluted image.
```

```
    % convolution of red channel
    dim=size(img);
    newR=zeros(dim(1),dim(2),3);
    newR(:,:,1)=convolution(img(:,:,1), kernel);
    newR=uint8(newR);
    % green channel
    newG=zeros(dim(1),dim(2),3);
    newG(:,:,2)=convolution(img(:,:,2), kernel);
    newG=uint8(newG);
    % blue channel
    newB=zeros(dim(1),dim(2),3);
    newB(:,:,3)=convolution(img(:,:,3), kernel);
    newB=uint8(newB);
    % return image
    newImg=newR+newG+newB;
end
```

Ohjelma 6.7. *Konvoluutiofunktio RGB-kuvalle MATLAB® ohjelmistolla*

Ohjelmassa 6.8 on tehty konvoluutioita RGB-kuvalle. Ytiminä on käytetty erilaisia kuvan ääriviivat poimivia ytimiä. Ohjelmassa luodaan aluksi ytimet. Ytiminä käytetään aiemmin kaavassa (6.9) esitettyä ääriviivat poimivaa ydintä $5(id - a)$, missä a on 3×3-kokoinen Gaussin ydin, ja tämän lisäksi epäsymmetrisen Prewitt-ytimen (6.10) eri kierto-orientaatioita. Kierrot saatu aikaan kertomalla matriisi luvulla -1 tai matriisin transpoosilla. Ytimet K1-K5 ovat ohjelmassa jo valmiiksi kierretyjä ennen kuin ne syötetään konvoluutiofunktiolle. Konvoluutio RGB-kuvalle tapahtuu colorConvolution-funktion 6.7 avulla.

```
% Colorconvolution
% read image
img = imread('rubik1.jpg');
% create a gaussian 3x3 kernel
a=fspecial("gaussian",[3,3],1);
% create unsharp kernel
K1=5*([ 0 0 0;0 1 0; 0 0 0]-a);
% Prewitt kernels
K2=-1*fspecial("prewitt");
K3=-K2;
K4=K2';
K5=-K4;
% create list of kernels
K={K1, K2, K3, K4, K5};
% loop over different kernels
```



Kuva 6.9. *Ylimmällä rivillä vasemmalla alkuperäinen kuva ja tämän alapuolella kuva summentavan ytimen avulla poimituista ääriviivoista. Loput neljä kuvaa ovat konvoluutioita Prewitt-ytimellä. Prewitt-ytimen skannauksen suunta mainittu kuvan yhteydessä.*

```
for i=1:5
    I=colorConvolution(img,K{i});
end
```

Ohjelma 6.8. *RGB-kuvan konvoluutio MATLAB® ohjelmistolla*

Kuvassa 6.9 on esitetty ylimmällä rivillä vasemmalla alkuperäinen kuva ja ytimellä $5(id - a)$ suoritetulla konvoluutiolla korostuvat ääriviivat tämän alapuolella. Loput neljä kuvaa havainnollistavat Prewitt-ytimen konvoluutiota. Toisen rivin ensimmäisessä kuvassa konvoluutio on tehty Prewitt-ytimellä

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}. \tag{6.10}$$

Koska ydin on epäsymmetrinen, sitä täytyy kiertää konvoluutiota varten. Tämä tapahtuu kätevästi kertomalla luvulla -1. Konvoluutiossa keskimmäisen pikselin yläpuolella olevien pikseleiden arvot vaihtuvat negatiivisiksi ja ne summataan alapuolella olevien pikseleiden arvoihin. Jos alapuolen pikseleiden arvot ovat suurempia, eli vaaleampia, niin tällöin konvoluutiossa uuden pikselin arvo on suurempaa kuin 0. Mulloin uusi pikseli saa arvon 0. Näin ollen voidaan ajatella, että Prewitt-ydin käsittelee kuvan ikään kuin ylhäältä alaspäin ja korostaa vaakasuuntaisia ääriviivoja siirryttäessä tummasta vaaleaan sävyyn. Jos luvulla -1 kerrotulla ytimellä (6.10) tehdään konvoluutio, niin saatu ydin korostaa vaakasuuntaisia ääriviivoja siirrettäessä tummasta vaaleaa ja alhaalta ylös. Vaakasuuntaisia äriviivoja korostavien Prewitt-ytimien konvoluutiot näkyvät ensimäisellä rivillä. Kuvassa 6.9 toisella rivillä kaksi viimeistä kuvaa ovat Prewitt-ytimen (6.10) transpoosilla ja luvulla -1 kerrotulla transpoosilla suoritetut konvoluutiot tässä järjestyksessä. Tällöin korostuvat pystysuuntaiset ääriviivat vasemmalta oikealla ja oikealta vasemmalle.

Tärkeää huomata, että konvoluutio suoritetaan kullekin värikanavalle erikseen, mistä johtuen kuviin voisi periaatteessa tulla ääriviivoja joillakin värisävyillä, joita alkuperäisessä kuvassa ei ole havaittavissa. Ytimellä $t(id - a)$ suoritettu konvoluutio poimii kaikensuuntaisia ääriviivoja. Nämä ääriviivat sattuvat olemaan eri harmaansävyjä, mutta kuvasta ja parametrin t arvosta riippuen voi piirtyä muunkin värisiä ääriviivoja. Lisäksi näistä viivoista tulee huomattavasti himmeämpiä kuin Prewitt-ytimellä. Parametria t kasvatettaessa ääriviivat kyllä voimistuvat, mutta samalla ydin alkaa korostaa yksittäisiä pikseleitä varsinkin ääriviivojen läheisyydestä. Toisin sanoen kuvaan alkaa tulla häiriötä.

7. YHTEENVETO

Tässä työssä tutkittiin Hadamardin tuloa, siihen liittyviä ominaisuuksia ja sen käyttöä digitaalisten kuvien käsittelyssä. Työn yhtenä päämääränä oli, että lukija pystyy ymmärtämään tekstin ilman ulkoisia lähteitä. Tästä syystä työ alkaa luvuilla 2 ja 3, joissa esitellään ja kerrataan melko kattavasti vektoreihin ja matriiseihin liittyviä työn kannalta oleellisia perusominaisuuksia. Luku 4 perehdyttää lukijan normaaleihin matriiseihin keskittyen erityisesti unitaaristen, hermiittisten ja positiivisesti semidefiniittien matriisien muodostamiin alaluokkiin. Tämä luku antaa tarvittavat työkalut Hadamardin tulon ominaisuuksien tutkimiselle. Ennen kaikkea hyödyllisyytensä osoittivat hermiittisten matriisien spektraalilause 4.12 sekä lauseet 4.16, 4.17 ja 4.19, joiden avulla voidaan todeta, että jokin matriisi on positiivisesti (semi)definiitti.

Luvussa 5 esiteltiin aluksi Kroneckerin tulo ja sen perusominaisuudet todistuksineen. Tärkeä havainto on, että Hadamardin tulo on Kroneckerin tulon alimatriisi. Tämän takia Kroneckerin tulo on luonnollinen osa lukua. Aivan kuten Kroneckerin tulo, Hadamardinkin tulo eroaa merkittävästi perinteisestä matriisitulosta. Hadamardin tulo erottuu muun muassa yksinkertaisuutensa ja vaihdannaisuutensa ansiosta. Tämän luvun oleellisin tulos on Hadamardin tuloon liittyvä Schurin lause 5.11, joka todistettiin työssä kahdella eri tavalla. Ensiksi hyödyntäen Hadamardin tulon ja matriisin jäljen välistä yhteyttä. Toiseksi hieman siistimmin hyödyntämällä tietoa, että Hadamardin tulo on Kroneckerin tulon alimatriisi. Schurin lauseen mukaan positiivisesti (semi)definiittien matriisien Hadamardin tulo on positiivisesti (semi)definiitti matriisi. Tästä seuraa, että positiivisesti (semi)definiittien matriisien joukko on suljettu Hadamardin tulon suhteen. Tämä lause on muutenkin keskeinen matriisianalyysissa ja sen avulla voidaan osoittaa monia muita ominaisuuksia. Työssä Schurin lausetta hyödynnettiin todistettaessa Hadamardin tulon determinanttiin liittyvä epäyhtälö 5.14, jonka mukaan kahden matriisin determinanttien tulo on pienempää tai yhtäsuurta kuin kyseisten matriisien Hadamardin tulon determinantti.

Luvussa 6 tarkasteltiin digitaalista kuvankäsittelyä ja Hadamardin tulon roolia siinä. Matriisilaskenta soveltuu digitaaliseen kuvankäsittelyyn mainosti, sillä digitaaliset mustavalkokuvat voidaan esittää matriiseina ja RGB-kuvat kolmen matriisin pinona. Mustavalkokuvien käsittely tapahtuu suoraan kuvaa vastaavalle matriisille, kun taas RGB-kuvissa jokaiselle värikanvalle erikseen, minkä jälkeen värikanavat yhdistetään uudeksi RGB-kuvaksi. Jo alkeellisilla matriisioperaatioilla saadaan aikaan huomattavia muutoksia. Tästä esimerkkinä matriisien yhteenlasku ja skalaarikertolasku, joiden avulla voidaan tuottaa muun muassa negatiivikuvia tai säätää kuvan kirkkautta. Lisäksi Hadamardin tulolla voidaan esimerkiksi rajata tiettyjä alueita kuvasta. Hieman edistyneemmällä matematiikalla kuvankäsittelyn mahdollisuudet moninkertaistuvat. Työssä tärkeäksi Hadamardin tulon sovelluskohteeksi valikoitui digitaalisten kuvien konvoluutio, joka on oleellinen osa esimerkiksi konvoluutioneuroverkkoja. Konvoluutiota tutkittiin etenkin sumentavilla ytimillä, joita on erilaisia. Yleisimmät sumentavat ytimet ovat keskiarvoydin ja Gaussin ydin. Aluksi havaittiin, että sumentavalla ytimellä voidaan poistaa häiriöitä kuvasta tai nimenmukaisesti sumentaa kuva. Kuvasta voidaan myös sumentaa vain tietty alue. Tämä toteutettiin työssä hyödyntämällä Hadamardin tuloa. Lisäksi todettiin, kuinka sumentavia ytimiä voi hyödyntää kuvan terävöinnissä, eli ääriviivojen korostamisessa. Lopussa esiteltiin vielä konvoluutiota RGB-kuville vertaamalla erilaisia kuvan ääriviivat erottavia Prewitt-ytimiä terävöinnin yhteydessä esitettyyn ääriviivat erottavaan ytimeen.

LÄHTEET

[1] Edwin. Beckenbach. *An introduction to inequalities*. eng. New mathematical library ; 3. New York: Random House, 1961.

[2] Richard. Bellman. *Introduction to matrix analysis*. eng. New York: McGraw-Hill, 1960.

[3] Lindsay N. Childs. *A Concrete Introduction to Higher Algebra*. eng. 3rd ed. 2009. Undergraduate Texts in Mathematics. New York, NY: Springer New York, 2009. ISBN: 0-387-74725-7.

[4] Pankaj Kumar Das ja Lalit K. Vashisht. ”Traces of Hadamard and Kronecker products of matrices”. eng. *Mathematics for applications* 6.2 (2017), s. 143–150. ISSN: 1805-3610.

[5] Leonard E. Fuller. *Basic matrix theory*. eng. Englewood Cliffs, NJ: Prentice-Hall, 1962.

[6] Jack L. Goldberg. *Matrix theory with applications*. eng. International series in pure and applied mathematics. New York: McGraw-Hill, Inc., 1991. ISBN: 0-07-557200-1.

[7] Rafael C. Gonzalez. *Digital image processing*. eng. 3rd ed. Upper Saddle River (N.J.): Prentice Hall, 2008. ISBN: 0-13-168728-X.

[8] Roger A Horn. ”The Hadamard product”. *Proceedings of Symposia in Applied Mathematics* 40 (1990), s. 87–169. URL: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=tLr2dXsAAAAJ&citation_for_view=tLr2dXsAAAAJ:qjMakFHDy7sC.

[9] Roger A. Horn ja Charles R. Johnson. *Matrix analysis*. eng. 2. ed. Cambridge: Cambridge University Press, 2013. ISBN: 978-0-521-83940-2.

[10] Roger A. Horn ja Charles R. Johnson. *Topics in matrix analysis*. eng. Cambridge: Cambridge University Press, 1991. ISBN: 0-521-30587-X.

[11] Alasdair McAndrew. *A Computational Introduction to Digital Image Processing*. eng. Boca Raton: Chapman ja Hall/CRC, 2016. ISBN: 9781482247329.

[12] Elizabeth Million. *The Hadamard product*. 2007. URL: http://buzzard.ups.edu/courses/2007spring/projects/million-paper.pdf.

[13] Seppo Pohjolainen, Lassi Paunonen ja Mika Mattila. *MATH.APP.410 Matriisilaskenta [luentomoniste]*. Tampereen yliopisto, 2022.

[14] David Poole. *Linear algebra : a modern introduction*. eng. Second edition. Australia: Thomson, 2006. ISBN: 0-534-99845-3.

[15] Fuzhen Zhang. *Matrix theory : basic results and techniques*. eng. Universitext. New York: Springer, 1999. ISBN: 0-387-98696-0.

LIITE A: ARITMEETTIS-GEOMETRINEN EPÄYHTÄLÖ

Lause A.1 (Aritmeettis-geometrinen epäyhtälö). *Olkoot luvut* $\lambda_1, \ldots, \lambda_n$ *epänegatiivisia reaalilukuja. Tällöin*

$$\lambda_1 \cdots \lambda_n \leq \left(\frac{1}{n}(\lambda_1 + \cdots + \lambda_n) \right)^n,$$

missä yhtäsuuruus on voimassa, jos ja vain jos $\lambda_1 = \cdots = \lambda_n$.

Todistus. (Ks. [1], Theorem 4.3). □

LIITE B: ALGEBRAN PERUSLAUSE

Lause B.1 (Algebran peruslause)**.** *Kompleksikertoimisella polynomilla, joka on vähintään astetta* 1*, on ainakin yksi nollakohta kompleksilukujen joukossa.*

Todistus. (Ks. [3], Chapter 15 F. The Fundamental Theorem). □

Lause B.2. *Kompleksikertoimisella polynomilla, joka on astetta* $n \geq 1$*, on täsmälleen* n *nollakohtaa kompleksilukujen joukossa, kun nollakohtien monikerrat on laskettu mukaan.*

Todistus. Todistetaan lause induktiolla. Olkoon $p_n(z)$ astetta $n \geq 1$ oleva kompleksikertoiminen polynomi. Jos $n = 1$, Algebran peruslause B.1 takaa, että polynomilla $p_1(z)$ on nollakohta z_0. Polynomi voidaan kirjoittaa tekijämuodossa $p(z) = \alpha(z - z_0)$, mistä nähdään, että enempää nollakohtia ei voi olla. Oletetaan, että väite on tosi, kun $n = k \in \mathbb{N}$. Jos $n = k + 1$, Algebran peruslauseen B.1 nojalla on ainakin yksi nollakohta z_{k+1}, joten se voidaan kirjoittaa tämän avulla $p_{k+1}(z) = (z - z_{k+1})p_k(z)$. Induktio-oletuksesta seuraa, että polynomilla $p_k(z)$ on k nollakohtaa. Näin ollen polynomilla p_{k+1} on $k + 1$ nollakohtaa. □

LIITE C: DIGITAALISEN KUVANKÄSITTELYN MATLAB® OHJELMISTOLLA

```
% Color Channel Extraction
% read image data to a variable
img=imread('rubik1.jpg');
% initializing variables R, G, B to proper sizes
R=uint8(zeros(size(img)));
G=R;
B=R;
% extracting color channels
R(:,:,1)=img(:,:,1);
G(:,:,2)=img(:,:,2);
B(:,:,3)=img(:,:,3);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Brightness and contrast
% read image
img=imread('lowlight.jpg');
% brighten the image
I1=uint8(1.1*img+60*uint8(ones(size(img))));
% gamma correction
A=double(img)*1/255;
I2=uint8((A.^0.5)*255);
% brighten the gamma corrected image
I3=I2+uint8(25*ones(size(I)));
% small gamma
I4=uint8((A.^0.1)*255);
% big gamma
I5=uint8((A.^5)*255);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
% Convolution Blur
img = imread('rubik1.jpg');
% make grayscale image
img=0.2989*img(:,:,1) + 0.5870*img(:,:,2) + 0.1140*img(:,:,3);
% add noise to image
imgNoise=imnoise(img,'salt␣&␣pepper',0.02);
% loop over different kernel sizes
for i=[3 5 7 9]
    kernel=1/i^2*ones(i);
    blur=convolution(imgNoise, kernel);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Region blur
% read images
img=imread('rubik1.jpg');
H_A1=imread('rubik_sq.jpg');
H_A2=imread('rubik_fs.jpg');
% make grayscale image
img=0.2989*img(:,:,1) + 0.5870*img(:,:,2) + 0.1140*img(:,:,3);
% make binary images of regions to be blurred and their negatives
H_A1=1/255*H_A1(:,:,1);
H_A2=1/255*H_A2(:,:,1);
H_B1=ones(size(H_A1))-double(H_A1);
H_B2=ones(size(H_A2))-double(H_A2);
% blur image
B=convolution(img,1/25^2*ones(25));
% make regional blurs
blur_sq=uint8(H_A1).*img+uint8(H_B1).*B;
blur_fs=uint8(H_A2).*img+uint8(H_B2).*B;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Convolution, Unsharp Masking
% read image and convert it to grayscale
img=imread('rubik1.jpg');
img=0.2989*img(:,:,1) + 0.5870*img(:,:,2) + 0.1140*img(:,:,3);
% create a gaussian 3x3 kernel
```

```
a=fspecial("gaussian",[3,3],1);
% create unsharp kernel
id=[ 0 0 0;0 1 0; 0 0 0];
edge_kernel=id-a;
% extract the edges to be sharpened from image
edges1=convolution(img, edge_kernel);
edges2=convolution(img, -edge_kernel);
% modify the contrast of edges
A=double(edges1)*1/255;
edges1=uint8((A.^0.35)*255);
A=double(edges2)*1/255;
edges2=uint8((A.^0.35)*255);
% sharpened image with t=5
conv1=convolution(img, id+5*edge_kernel);
% sharpened image with t=-5
conv2=convolution(img, id-5*edge_kernel);
% sharpened image with t=20
conv3=convolution(img, id+20*edge_kernel);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Colorconvolution
% read image
img = imread('rubik1.jpg');
% create a gaussian 3x3 kernel
a=fspecial("gaussian",[3,3],1);
% create unsharp kernel
K1=5*([ 0 0 0;0 1 0; 0 0 0]-a);
% Prewitt kernels
K2=-1*fspecial("prewitt");
K3=-K2;
K4=K2';
K5=-K4;
% create list of kernels
K={K1, K2, K3, K4, K5};
% loop over different kernels
for i=1:5
    I=colorConvolution(img,K{i});
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function newImg = convolution(img, kernel)
    % This function does convolution. As a matter of fact it does
    % correlation, since it assumes the user has already flipped
    % the kernel.
    K=kernel;
    % variables needed for zeropadding
    d=floor(size(K,1)/2);
    dim=size(img);
    % zeropadding
    A=[zeros(d,dim(2)+2*d);
        zeros(dim(1),d) img zeros(dim(1),d);
        zeros(d,dim(2)+2*d)];
    % initialize the conv variable
    conv=zeros(size(img));
    % convolution
    for i=1+d:dim(1)+d-1
        for j=1+d:dim(2)+d-1
            % sum of the hadamard product
            conv(i-d+1,j-d+1)=sum(K.*double(A(i-d:i+d,j-d:j+d)),
                                      'all');
        end
    end
    % return convoluted image
    newImg=uint8(conv);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function newImg = colorConvolution(img, kernel)
    % This function extract different color channels from
    % RGB-image and convolutes them one by one with
    % convolution-function before combining them and returning
    % the color convoluted image.

    % convolution of red channel
    dim=size(img);
    newR=zeros(dim(1),dim(2),3);
    newR(:,:,1)=convolution(img(:,:,1), kernel);
```

```
    newR=uint8(newR);
    % green channel
    newG=zeros(dim(1),dim(2),3);
    newG(:,:,2)=convolution(img(:,:,2), kernel);
    newG=uint8(newG);
    % blue channel
    newB=zeros(dim(1),dim(2),3);
    newB(:,:,3)=convolution(img(:,:,3), kernel);
    newB=uint8(newB);
    % return image
    newImg=newR+newG+newB;
end
```

Ohjelma C.1. *Digitaalista kuvankäsittelyä MATLAB® ohjelmistolla*