Tampereen yliopisto

Jussi Harju

UNITAARINEN DIAGONALISOINTI JA SPEKTRAALIHAJOTELMA

Informaatioteknologian ja viestinnän tiedekunta
Pro gradu -tutkielma
Marraskuu 2022

Tiivistelmä

Jussi Harju: Unitaarinen diagonalisointi ja spektraalihajotelma
Pro gradu -tutkielma
Tampereen yliopisto
Matematiikan ja tilastotieteen tutkinto-ohjelma
Marraskuu 2022

Tässä tutkielmassa käsitellään reaalimatriisien ortogonaalista diagonalisointia ja kompleksisten matriisien unitaarista diagonalisointia sekä vertaillaan niiden ominaisuuksia. Ennen unitaarisen diagonalisoinnin käsittelyä tarkastellaan pohjatietona kompleksisia vektoriavaruuksia, matriiseja ja sisätuloavaruuksia. Tutkielmassa osoitetaan, että matriisin on oltava symmetrinen ollakseen ortogonaalisesti diagonalisoituva ja normaali ollakseen unitaarisesti diagonalisoituva. Lisäksi tarkastellaan normaalin matriisin spektraalihajotelmaa.

Tutkielman keskeisimmät tulokset ovat, että reaalimatriisi on ortogonaalisesti diagonalisoituva, jos ja vain jos se on symmetrinen, ja kompleksinen matriisi on unitaarisesti diagonalisoituva, jos ja vain jos se on normaali. Ortogonaaliseen diagonalisoituvuuteen liittyen esitellään oleelliset määritelmät matriisin transpoosi, symmetrinen ja ortogonaalinen matriisi sekä unitaariseen diagonalisoituvuuteen liittyen Hermiten konjugaatti, hermiittinen, unitaarinen ja normaali matriisi.

Lopuksi esitetään normaalin matriisin spektraalihajotelma ja tarkastellaan spektraalihajotelman matriisien ominaisuuksia. Tutkielma on kirjoitettu mukaillen Anthonyn ja Harveyn kirjaa Linear Algebra: Concepts and Methods sekä Antonin ja Rorresin kirjaa Elementary Linear Algebra. Lisäksi lähteinä on käytetty Hornin ja Johnsonin kirjaa Matrix Analysis sekä Prasolovin kirjaa Problems and Theorems in Linear Algebra. Lukijan oletetaan tuntevan matriisien tavanomaisen diagonalisoinnin.

Avainsanat: ortogonaalinen diagonalisointi, kompleksinen sisätuloavaruus, unitaarinen diagonalisointi, normaali matriisi, spektraalihajotelma

Sisällys

1 Johdanto

Tässä tutkielmassa tarkastellaan kompleksisia matriiseja ja erityisesti unitaarista diagonalisointia, mutta ortogonaalinen diagonalisointi esitellään vertailun vuoksi. Lisäksi tarkastellaan normaalin matriisin spektraalihajotelmaa.

Tämän tutkielman luvussa 3 tarkastellaan reaalimatriisien ortogonaalista diagonalisointia. Määritellään matriisin transpoosi ja symmetrinen matriisi, sekä esitetään symmetrisen reaalimatriisin ominaisarvojen reaalisuutta koskeva lause. Ortogonaalinen matriisi määritellään pykälässä 3.2, jossa esitetään myös kolme ortogonaaliseen matriisiin ja sen ominais- ja sarakevektoreihin liittyvää lausetta. Sen jälkeen todetaan, että ortogonaalisen matriisin määräämä lineaarikuvaus säilyttää sisätulon tuloksen reaalisessa vektoriavaruudessa. Pykälässä 3.3 määritellään ortogonaalisesti diagonalisoituva matriisi ja todetaan, että matriisi on sellainen, jos ja vain jos sen ominaisvektorit muodostavat ortonormaalin kannan reaalisessa vektoriavaruudessa. Lisäksi esitetään luvun 3 päälause, spektraalilause, eli matriisi on ortogonaalisesti diagonalisoituva, jos ja vain jos se on symmetrinen.

Luvussa 4 tarkastellaan kompleksisia vektoriavaruuksia, matriiseja ja sisätuloavaruuksia. Määritellään kompleksinen vektoriavaruus, kompleksinen matriisi ja todistetaan reaalimatriisin kompleksisen ominaisarvon konjugaattiin liittyvä lause. Lopuksi tarkastellaan kompleksisia sisätuloavaruuksia ja niiden ominaisuuksia.

Luvussa 5 tarkastellaan kompleksisten matriisien unitaarista diagonalisointia ja verrataan sitä reaalimatriisien ortogonaaliseen diagonalisointiin. Pykälässä 5.1 määritellään Hermiten konjugaatti ja hermiittinen matriisi, kompleksisten matriisien vastineet transpoosille ja symmetriselle reaalimatriisille. Lisäksi osoitetaan, että hermiittisen matriisin ominaisarvot ovat reaalisia ja sen erisuuria ominaisarvoja vastaavat ominaisvektorit ovat ortogonaaliset. Pykälässä 5.2 määritellään unitaarinen matriisi, kompleksinen vastine ortogonaaliselle matriisille, ja todistetaan, että matriisi on unitaarinen, jos ja vain jos sen sarakevektorit muodostavat kompleksisen vektoriavaruuden ortonormaalin kannan ja sen määräämä lineaarikuvaus säilyttää sisätulon tuloksen kompleksisessa vektoriavaruudessa. Pykälässä 5.3 määritellään unitaarisesti diagonalisoituva matriisi ja todistetaan, että matriisi on sellainen, jos ja vain jos sen ominaisvektorit muodostavat ortonormaalin kannan kompleksisessa vektoriavaruudessa. Sitten määritellään normaali matriisi ja todistetaan, että matriisi on unitaarisesti diagonalisoituva, jos ja vain jos se on normaali. Lisäksi osoitetaan, että jos normaalin matriisin kaikki ominaisarvot ovat reaalisia, niin se on hermiittinen.

Luvussa 6 esitetään normaalin matriisin spektraalihajotelma ja tarkastellaan spektraalihajotelman matriisien ominaisuuksia sekä todetaan, että niiden idempotenttisuus tarkoittaa sitä, että ne muodostavat ortogonaaliprojektion.

Luvussa 2 luetellaan aiheiden käsittelyssä tarvittavia määritelmiä ja lauseita. Ensin esitetään vektoriavaruuksiin ja matriisin diagonalisoituvuuteen liittyviä määritelmiä ja lauseita. Sitten esitetään sisätuloavaruuksiin liittyviä määritelmiä ja yksi lause. Lopuksi esitetään kompleksiluvun, kompleksisen konjugaatin ja kompleksiluvun modulin määritelmät.

Lukijalta edellytetään joidenkin lineaarialgebran perusasioiden tuntemista. Edellytetään muun muassa, että lukija tuntee matriisien tavalliset laskutoimitukset, ominaisarvojen ja -vektorien määritelmät, lineaarikuvauksen ja kannan määritelmät sekä matriisin tavanomaisen diagonalisoinnin. Lähdeteoksina käytetään Anthonyn ja Harveyn kirjaa Linear Algebra: Concepts and Methods, Antonin ja Rorresin kirjaa Elementary Linear Algebra, Hornin ja Johnsonin kirjaa Matrix Analysis ja Prasolovin kirjaa Problems and Theorems in Linear Algebra.

2 Esitietoja

2.1 Vektoriavaruuksista ja matriiseista

Tässä pykälässä esitetään vektoriavaruuksiin sekä similaarisiin ja diagonalisoituviin matriiseihin liittyviä oleellisia esitietoja. Esitetään ortogonaalisten vektorien määritelmä, ortonormaalin joukon määritelmä, similaaristen matriisien ominaisarvoja koskeva lause, diagonalisoituvan matriisin määritelmä ja diagonalisoituvan matriisin ominaisvektoreista muodostuvaa kantaa koskeva lause.

Määritelmä 2.1 (Reaalinen vektoriavaruus)**.** Reaalinen *vektoriavaruus* on epätyhjä joukko V, jonka alkioita sanotaan *vektoreiksi*, varustettuna vektorien yhteenlaskulla $\boldsymbol{u} + \boldsymbol{v} \in V$ ja skalaarilla kertomisella $\alpha\boldsymbol{u} \in V$. Laskutoimitukset toteuttavat seuraavat ehdot kaikilla $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ ja kaikilla $\alpha, \beta \in \mathbb{R}$:

1. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ (vaihdannaisuus)
2. $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ (liitännäisyys)
3. On olemassa nollavektori $\boldsymbol{0} \in V$, jolle pätee $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$ kaikilla $\boldsymbol{u} \in V$ (nolla-alkio)
4. kaikilla $\boldsymbol{u} \in V$ on olemassa vektori $-\boldsymbol{u}$, jolle pätee $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$ (vasta-alkio)
5. $\alpha(\boldsymbol{u} + \boldsymbol{v}) = \alpha\boldsymbol{u} + \alpha\boldsymbol{v}$ (osittelulaki)
6. $(\alpha + \beta)\boldsymbol{u} = \alpha\boldsymbol{u} + \beta\boldsymbol{u}$ (osittelulaki)
7. $\alpha(\beta\boldsymbol{u}) = (\alpha\beta)\boldsymbol{u}$ (liitännäisyys)
8. $1\boldsymbol{u} = \boldsymbol{u}$ (neutraalialkio).

Vektori $\boldsymbol{0} \in V$ on *nollavektori*. Vektori $-\boldsymbol{u} \in V$ on vektorin $\boldsymbol{u}$ *vastavektori*. Lukua $\alpha \in \mathbb{R}$ sanotaan *skalaariksi*.

Vektoriavaruuden V vektoria

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_n\boldsymbol{v}_n$$

sanotaan vektorien $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$ *lineaarikombinaatioksi*. Skalaareita $a_1, a_2, \ldots, a_n \in \mathbb{R}$ sanotaan *kertoimiksi*.

Määritelmä 2.2. Olkoot $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$. Tällöin vektorit $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ ovat *lineaarisesti riippumattomia*, jos

$$a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_n\boldsymbol{v}_n = 0 \quad \Rightarrow \quad a_1 = a_2 = \cdots = a_n = 0.$$

Muutoin vektorit ovat *lineaarisesti riippuvia*.

Määritelmä 2.3 (Similaarisuus)**.** Neliömatriisin A sanotaan olevan *similaarinen* neliömatriisin B kanssa, jos on olemassa sellainen kääntyvä matriisi P, että

$$P^{-1}AP = B.$$

Määritelmä 2.4 (Diagonalisoituva matriisi)**.** Neliömatriisia A sanotaan diagonalisoituvaksi, jos se on similaarinen jonkin diagonaalimatriisin D kanssa, eli on olemassa kääntyvä matriisi P siten, että

$$P^{-1}AP = D.$$

Esimerkki 2.1. Olkoon

$$A = \begin{pmatrix} 7 & -15 \\ 2 & -4 \end{pmatrix} \qquad \text{ja} \qquad P = \begin{pmatrix} 5 & 3 \\ 2 & 1 \end{pmatrix}.$$

Tällöin

$$P^{-1} = \begin{pmatrix} -1 & 3 \\ 2 & -5 \end{pmatrix}$$

ja

$$P^{-1}AP = D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Lause 2.1. *Olkoon A $n \times n$ -matriisi. Tällöin A on diagonalisoituva, jos ja vain jos vektoriavaruudella $\mathbb{R}^n$ on olemassa kanta, joka muodostuu matriisin A ominaisvektoreista.*

Todistus. Ks. [1, s. 257–258]. □

Lause 2.2. *Similaarisilla matriiseilla on samat ominaisarvot.*

Todistus. Ks. [1, s. 262–263]. □

2.2 Sisätuloavaruuksista

Tässä pykälässä esitetään sisätuloavaruuksiin liittyviä oleellisia määritelmiä ja lauseita. Reaalisen vektoriavaruuden $\mathbb{R}^n$ kahden vektorin $\boldsymbol{x}$ ja $\boldsymbol{y}$ sisätulo on $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$. Tätä sanotaan *euklidiseksi* tai *standardiksi* sisätuloksi. Sisätuloja on erilaisia kuin edellinen ja missä tahansa vektoriavaruudessa. Määritellään seuraavaksi sisätulo yleisesti.

Määritelmä 2.5 (Sisätulo)**.** Olkoon V vektoriavaruus. Tällöin kuvausta $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$ sanotaan *sisätuloksi*, jos

(i) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$ kaikilla $\boldsymbol{x}, \boldsymbol{y} \in V$,

(ii) $\langle \alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{z} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta\langle \boldsymbol{y}, \boldsymbol{z} \rangle$ kaikilla $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ ja kaikilla $\alpha, \beta \in \mathbb{R}$,

(iii) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ kaikilla $\boldsymbol{x} \in V$, ja $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$, jos ja vain jos $\boldsymbol{x} = \boldsymbol{0}$.

Jos vektoriavaruudessa on määritelty sisätulo, sanotaan sitä *sisätuloavaruudeksi*.

Määritelmä 2.6 (Normi)**.** Oletetaan, että V on sisätuloavaruus ja vektori $\boldsymbol{x} \in V$. Tällöin vektorin $\boldsymbol{x}$ *normi*, tai *pituus*, jota merkitään $\|\boldsymbol{x}\|$, on

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

Sanotaan, että vektori $\boldsymbol{x}$ on *yksikkövektori*, jos sen normi on 1.

Määritelmä 2.7 (Ortogonaaliset vektorit)**.** Olkoon V sisätuloavaruus. Tällöin vektorien $\boldsymbol{x}, \boldsymbol{y} \in V$ sanotaan olevan *ortogonaalisia*, jos $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$. Tällöin merkitään $\boldsymbol{x} \perp \boldsymbol{y}$.

Määritelmä 2.8. Sisätuloavaruuden V joukon S sanotaan olevan *ortonormaali*, jos

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \qquad \text{ja} \qquad \|\boldsymbol{x}\| = 1$$

aina, kun $\boldsymbol{x}, \boldsymbol{y} \in S$ ja $\boldsymbol{x} \neq \boldsymbol{y}$.

Lause 2.3. *Oletetaan, että V on sisätuloavaruus, vektorit $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ ovat pareittain ortogonaalisia ($\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0$, kun $i \neq j$) ja että yksikään ei ole nollavektori. Tällöin vektorijoukko $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ on lineaarisesti riippumaton.*

Todistus. Ks. [1, s. 318]. □

2.3 Kompleksiluvuista

Tässä pykälässä määritellään *kompleksiluku* ja *kompleksinen konjugaatti*, joita käytetään luvuissa 4, 5 ja 6.

Määritelmä 2.9 (Kompleksiluku)**.** Kompleksiluku on muotoa $z = a + ib$, missä a ja b ovat reaalilukuja ja $i^2 = -1$. Kompleksilukujen joukko on

$$\mathbb{C} = \{a + ib \quad | \quad a, b \in \mathbb{R}\}.$$

Määritelmä 2.10 (Kompleksinen konjugaatti)**.** Kompleksiluvun $z = a + ib$ *kompleksinen konjugaatti* on kompleksiluku $\overline{z} = a - ib$.

Tästä määritelmästä seuraa, että jos x on reaaliluku, niin $x = \overline{x}$. Kompleksiluvun $z = a + ib \in \mathbb{C}$ *moduli* on

$$|z| = \sqrt{z\overline{z}} = \sqrt{(a + ib)(a - ib)} = \sqrt{a^2 + b^2},$$

joka on epänegatiivinen reaaliluku.

3 Ortogonaalinen diagonalisointi

Tässä luvussa käsitellään reaalimatriisien ortogonaalista diagonalisointia. Pykälässä 3.1 käsitellään matriisin *transpoosia* ja *symmetrisyyttä*. Pykälässä 3.2 määritellään *ortogonaalinen* matriisi, tutkitaan sarakevektorien ja ominaisvektorien yhteyttä ortogonaalisuuteen sekä osoitetaan, että ortogonaalisen matriisiin määräämä lineaarikuvaus säilyttää sisätulon tuloksen. Näiden tietojen ja ominaisuuksien esittelyn jälkeen pykälässä 3.3 määritellään *ortogonaalisesti diagonalisoituva* matriisi ja tutkitaan, milloin matriisi on sellainen.

3.1 Transpoosi ja symmetriset matriisit

Jos vaihdetaan matriisin rivit ja sarakkeet toistensa paikalle saadaan sen *transpoosi*. Määritellään se seuraavasti (vrt. [1, s. 21]):

Määritelmä 3.1 (Transpoosi)**.** *Transpoosi* $m \times n$ -matriisista

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

on $n \times m$ -matriisi

$$A^{\mathrm{T}} = (a_{ji}) = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

Siis kun matriisi A transponoidaan, niin sen rivistä i tulee matriisin A^{T} sarake i.

Esimerkki 3.1. Matriisin

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

transpoosi on

$$A^{\mathrm{T}} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}.$$

Matriisin transpoosilla on seuraavat ominaisuudet

- $(A^{\mathrm{T}})^{\mathrm{T}} = A$,
- $(cA)^{\mathrm{T}} = cA^{\mathrm{T}}$,
- $(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$,
- $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

Jos matriisi A on kääntyvä, sillä on myös seuraava ominaisuus

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

Seuraavaksi määritellään *symmetrinen* matriisi (vrt. [1, s. 22]).

Määritelmä 3.2 (Symmetrinen matriisi)**.** Matriisin A sanotaan olevan symmetrinen, jos

$$A = A^{\mathrm{T}},$$

eli matriisi on itsensä transpoosi.

Vain neliömatriisit voivat olla symmetrisiä. Jos A on symmetrinen, niin $a_{ij} = a_{ji}$. Matriisi on siis symmetrinen diagonaalinsa suhteen.

Esimerkki 3.2. Matriisi

$$A = \begin{pmatrix} 1 & 4 & 5 \\ 4 & 2 & -7 \\ 5 & -7 & 3 \end{pmatrix}$$

on symmetrinen.

Jos D on diagonaalinen matriisi, niin $d_{ij} = 0 = d_{ji}$ kaikilla $i \neq j$. Täten kaikki diagonaalimatriisit ovat symmetrisiä.

Seuraavaksi esitetään symmetrisiä matriiseja ja niiden ominaisarvoja koskeva lause, joka todistetaan vasta luvussa 5.

Lause 3.1. *Jos reaalimatriisi A on symmetrinen, niin sen ominaisarvot ovat reaalisia.*

Todistus. Ks. lauseen 5.1 seuraus. □

3.2 Ortogonaaliset matriisit

Seuraava lause osoittaa, että symmetrisen matriisin erisuuria ominaisarvoja vastaavat ominaisvektorit ovat ortogonaaliset.

Lause 3.2. *Jos matriisi A on symmetrinen, niin sen erisuuria ominaisarvoja vastaavat ominaisvektorit ovat ortogonaaliset.*

Todistus Ks. [1, s. 332]. □

Esimerkki 3.3. Olkoon

$$A = \begin{pmatrix} 4 & 0 & 4 \\ 0 & 4 & 4 \\ 4 & 4 & 8 \end{pmatrix}$$

symmetrinen matriisi. Sen ominaisarvot ovat $\lambda_1 = 4$, $\lambda_2 = 0$, $\lambda_3 = 12$ ja vastaavat ominaisvektorit esimerkiksi

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \qquad \boldsymbol{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

Nyt

$$\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle = 1 \cdot 1 + (-1) \cdot 1 + 0 \cdot (-1) = 1 + (-1) + 0 = 0,$$

$$\langle \boldsymbol{x}_1, \boldsymbol{x}_3 \rangle = 1 \cdot 1 + (-1) \cdot 1 + 0 \cdot 2 = 1 + (-1) + 0 = 0,$$

ja

$$\langle \boldsymbol{x}_2, \boldsymbol{x}_3 \rangle = 1 \cdot 1 + 1 \cdot 1 + (-1) \cdot 2 = 1 + 1 + (-2) = 0.$$

Matriisilla voi olla erityisen käytännöllinen ominaisuus, joka linkittyy vektorien ortogonaalisuuteen. Tämä määritellään seuraavasti:

Määritelmä 3.3 (Ortogonaalinen matriisi)**.** Neliömatriisin P sanotaan olevan *ortogonaalinen*, jos $P^{\mathrm{T}}P = PP^{\mathrm{T}} = I$, eli matriisin P transpoosi on sen käänteismatriisi.

Esimerkki 3.4. Matriisi

$$P = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{-4}{5} & \frac{3}{5} \end{pmatrix}$$

on ortogonaalinen, koska

$$\begin{aligned} P^{\mathrm{T}}P &= \begin{pmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{-4}{5} & \frac{3}{5} \end{pmatrix} \\ &= \begin{pmatrix} \frac{3}{5} \cdot \frac{3}{5} + \frac{-4}{5} \cdot \frac{-4}{5} & \frac{3}{5} \cdot \frac{4}{5} + \frac{-4}{5} \cdot \frac{3}{5} \\ \frac{4}{5} \cdot \frac{3}{5} + \frac{3}{5} \cdot \frac{-4}{5} & \frac{4}{5} \cdot \frac{4}{5} + \frac{3}{5} \cdot \frac{3}{5} \end{pmatrix} \\ &= \begin{pmatrix} \frac{9}{25} + \frac{16}{25} & \frac{12}{25} + \frac{-12}{25} \\ \frac{12}{25} + \frac{-12}{25} & \frac{16}{25} + \frac{9}{25} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \end{aligned}$$

ja

$$
\begin{aligned}
PP^{\mathrm{T}} &= \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{-4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \\
&= \begin{pmatrix} \frac{3}{5} \cdot \frac{3}{5} + \frac{4}{5} \cdot \frac{4}{5} & \frac{3}{5} \cdot \frac{-4}{5} + \frac{4}{5} \cdot \frac{3}{5} \\ \frac{-4}{5} \cdot \frac{3}{5} + \frac{3}{5} \cdot \frac{4}{5} & \frac{-4}{5} \cdot \frac{-4}{5} + \frac{3}{5} \cdot \frac{3}{5} \end{pmatrix} \\
&= \begin{pmatrix} \frac{9}{25} + \frac{16}{25} & \frac{-12}{25} + \frac{12}{25} \\ \frac{-12}{25} + \frac{12}{25} & \frac{16}{25} + \frac{9}{25} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.
\end{aligned}
$$

Siis $P^{\mathrm{T}}P = PP^{\mathrm{T}} = I$.

Tarkastellaan nyt ortogonaalisen matriisin ja ortogonaalisten vektorien yhteyttä. Seuraava lause osoittaa, että ortogonaalisen matriisin sarakevektorit ovat pareittain ortogonaalisia yksikkövektoreita.

Lause 3.3. *Matriisi P on ortogonaalinen, jos ja vain jos sen sarakevektorit ovat pareittain ortogonaaliset ja jokaisen pituus on* 1.

Todistus Ks. [1, s. 319]. □

Ortogonaalisen matriisin P sarakevektorit muodostavat siis ortonormaalin joukon määritelmän 2.8 mukaan. Seuraava lause osoittaa, että ortogonaalisen matriisin sarakevektorit muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan.

Lause 3.4. *Olkoon P $n \times n$ -matriisi. Tällöin P on ortogonaalinen, jos ja vain jos sen sarakevektorit muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan.*

Todistus Ks. [2, s. 402]. □

Jos matriisi P on ortogonaalinen, niin myös sen transpoosi P^{T} on ortogonaalinen, koska $P = (P^{\mathrm{T}})^{\mathrm{T}}$. Siis $(P^{\mathrm{T}})^{\mathrm{T}}P^{\mathrm{T}} = P^{\mathrm{T}}(P^{\mathrm{T}})^{\mathrm{T}} = I$.

Tästä seuraa, että Lause 3.4 pätee myös *rivivektoreilla*, eli: matriisi P on ortogonaalinen, jos ja vain jos sen rivivektorit muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan.

Ortogonaalisen matriisin P määräämä lineaarikuvaus säilyttää sisätulon tuloksen ja siten minkä tahansa vektorin pituuden, eli $\langle P\boldsymbol{x}, P\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$. Tällaista lineaarikuvausta kutsutaan *isometriaksi*.

Lause 3.5. *Ortogonaalisen matriisin P määräämä lineaarikuvaus säilyttää sisätulon tuloksen vektoriavaruudessa $\mathbb{R}^n$.*

Todistus Ks. [1, s. 355]. □

Kahden vektorin välinen sisätulo on yhtäsuuri kuin niiden ortogonaalisen matriisin P määräämien kuvien sisätulo.

3.3 Ortogonaalinen diagonalisointi

Neliömatriisi A diagonalisoidaan siten, että on olemassa sellainen kääntyvä matriisi P, että $P^{-1}AP = D$, missä D on diagonaalimatriisi. Jos on olemassa ortogonaalinen matriisi P, joka diagonalisoi matriisin A siten, että $P^{-1}AP = P^{\mathrm{T}}AP = D$, niin tätä kutsutaan *ortogonaaliseksi diagonalisoinniksi*. Määritellään nyt ortogonaalisesti diagonalisoituva matriisi (vrt. [1, s. 329]).

Määritelmä 3.4. Neliömatriisin A sanotaan olevan ortogonaalisesti diagonalisoituva, jos on olemassa ortogonaalinen matriisi P siten, että $P^{\mathrm{T}}AP = D$, missä D on diagonaalimatriisi.

Koska P on ortogonaalinen, niin $P^{\mathrm{T}} = P^{-1}$, mistä seuraa, että $P^{\mathrm{T}}AP = P^{-1}AP = D$. Matriisi A on diagonalisoituva, joten lauseen 2.1 mukaan matriisin P sarakkeet ovat matriisin A ominaisvektorit ja muodostavat vektoriavaruuden $\mathbb{R}^n$ kannan. Koska A on ortognaalisesti diagonalisoituva, niin matriisin P sarakkeet, jotka ovat matriisin A ominaisvektorit, muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan. Nämä tiedot yhdistämällä saadaan seuraava lause.

Lause 3.6. *Matriisi A on ortogonaalisesti diagonalisoituva, jos ja vain jos vektoriavaruudelle $\mathbb{R}^n$ on olemassa matriisin A ominaisvektoreista muodostuva ortonormaali kanta.*

Todistus Ks. [2, s. 304, 410]. □

Lauseessa 3.2 todettiin, että jos symmetrisellä $n \times n$ -matriisilla on n erisuurta ominaisarvoa, niin niitä vastaavat ominaisvektorit ovat ortogonaaliset. Tällöin sen ominaisvektorit muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan lauseiden 3.3 ja 3.4 perusteella. Lauseen 3.6 mukaan tällainen matriisi on siis ortogonaalisesti diagonalisoituva. Seuraava lause osoittaa, että mikä tahansa symmetrinen matriisi on ortogonaalisesti diagonalisoituva.

Lause 3.7 (Spektraalilause)**.** *Matriisi A on ortogonaalisesti diagonalisoituva, jos ja vain jos A on symmetrinen.*

Spektraalilauseen todistus Ks. [1, s. 337]. □

Voidaan siis todeta, että matriisi on ortogonaalisesti diagonalisoituva, vain jos se on symmetrinen.

Esimerkki 3.5. Tarkastellaan symmetristä matriisia

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

Sen ominaisarvot ovat

$$\lambda_1 = 8, \quad \lambda_2 = -1 \quad \text{ja} \quad \lambda_3 = -1,$$

ja vastaavista ominaisvektoreista muodostettu matriisi on

$$P = \begin{pmatrix} \frac{2}{3} & \frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{3} & \frac{2}{3\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{3} & \frac{-\sqrt{5}}{3} & 0 \end{pmatrix}.$$

Tällöin $P^{\mathrm{T}}P = I$, joten P on ortogonaalinen matriisi. Nyt

$$P^{\mathrm{T}}AP = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{4}{3\sqrt{5}} & \frac{2}{3\sqrt{5}} & \frac{-\sqrt{5}}{3} \\ \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} & 0 \end{pmatrix} \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & \frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{3} & \frac{2}{3\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{3} & \frac{-\sqrt{5}}{3} & 0 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin A ominaisarvot.

4 Kompleksiset matriisit ja sisätuloavaruudet

4.1 Kompleksiset vektoriavaruudet

Vektoriavaruutta, jossa skalaarikertoimet ovat kompleksilukuja, kutsutaan kompleksiseksi vektoriavaruudeksi. Seuraava määritelmä on muuten sama kuin määritelmä 2.1, paitsi skalaarit ovat kompleksilukuja.

Määritelmä 4.1 (Kompleksinen vektoriavaruus)**.** *Kompleksinen vektoriavaruus* on epätyhjä joukko V, jonka alkioita sanotaan *vektoreiksi*, varustettuna vektorien yhteenlaskulla $\boldsymbol{u} + \boldsymbol{v} \in V$ ja skalaarilla kertomisella $\alpha\boldsymbol{u} \in V$. Laskutoimitukset toteuttavat seuraavat ehdot kaikilla $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ ja kaikilla $\alpha, \beta \in \mathbb{C}$:

1. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ (vaihdannaisuus)
2. $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ (liitännäisyys)
3. On olemassa nollavektori $\boldsymbol{0} \in V$, jolle pätee $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$ kaikilla $\boldsymbol{u} \in V$ (nolla-alkio)
4. kaikilla $\boldsymbol{u} \in V$ on olemassa vektori $-\boldsymbol{u}$, jolle pätee $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$ (vasta-alkio)
5. $\alpha(\boldsymbol{u} + \boldsymbol{v}) = \alpha\boldsymbol{u} + \alpha\boldsymbol{v}$ (osittelulaki)
6. $(\alpha + \beta)\boldsymbol{u} = \alpha\boldsymbol{u} + \beta\boldsymbol{u}$ (osittelulaki)
7. $\alpha(\beta\boldsymbol{u}) = (\alpha\beta)\boldsymbol{u}$ (liitännäisyys)
8. $1\boldsymbol{u} = \boldsymbol{u}$ (neutraalialkio).

Vektori $\boldsymbol{0} \in V$ on *nollavektori*. Vektori $-\boldsymbol{u} \in V$ on vektorin $\boldsymbol{u}$ *vastavektori*. Lukua $\alpha \in \mathbb{C}$ sanotaan *skalaariksi*.

Vektoreiden lineaarikombinaatio on samankaltainen kuin reaalisessa vektoriavaruudessa, paitsi sen kertoimet ovat kompleksilukuja. Täten vektori $\boldsymbol{v} \in V$ on vektoreiden $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$ lineaarikombinaatio, jos

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_n\boldsymbol{v}_n, \quad a_i \in \mathbb{C}.$$

Monet reaalisen vektoriavaruuden $\mathbb{R}^n$ käsitteistä siirtyvät kompleksiseen vektoriavaruuteen $\mathbb{C}^n$, samoin kuin lineaarikombinaation käsite. Lineaarinen riippumattomuus on yksi näistä käsitteistä.

Esimerkki 4.1. Oletetaan, että vektorin $\boldsymbol{e}_i$ jokainen alkio on 0, paitsi i:s alkio, joka on 1, kun $i = 1, 2, \dots, n$. Tällöin vektorit $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$ muodostavat kompleksisen vektoriavaruuden $\mathbb{C}^n$ kannan. Jokaisella $z = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n$ voidaan kirjoittaa

$$z = z_1\boldsymbol{e}_1 + z_2\boldsymbol{e}_2 + \cdots + z_n\boldsymbol{e}_n.$$

Kantaa $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ kutsutaan n-ulotteisen kompleksisen vektoriavaruuden $\mathbb{C}^n$ *luonnolliseksi kannaksi*.

4.2 Kompleksiset matriisit

Matriisia, jonka alkiot ovat kompleksilukuja, sanotaan *kompleksiseksi matriisiksi*, vastaavasti kuten matriisia, jonka alkiot ovat reaalilukuja, sanotaan reaalimatriisiksi. Jos A on kompleksinen $m \times n$ -matriisi, niin $\overline{A}$ on $m \times n$ -matriisi, jonka alkio (i, j) on matriisin A alkion (i, j) kompleksinen konjugaatti. Toisin sanoen, jos $A = (a_{ij})$, niin $\overline{A} = (\overline{a_{ij}})$. Matriisia $\overline{A}$ sanotaan matriisin A *kompleksiseksi konjugaattimatriisiksi*.

Ominaisarvon ja ominaisvektorin käsitteet ovat samat kompleksisilla matriiseilla, kuin reaalimatriiseilla, ja ne etsitään samalla tavalla. Erityisesti vektoriavaruudessa $\mathbb{C}^n$ voidaan diagonalisoida reaalimatriisi kompleksisilla ominaisarvoilla, kuten seuraava esimerkki osoittaa.

Esimerkki 4.2. Olkoon

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Sen ominaisarvot ovat $\lambda_1 = i$ ja $\lambda_2 = -i$ sekä vastaavat ominaisvektorit $\boldsymbol{x}_1 = (-i, 1)^{\mathrm{T}}$ ja $\boldsymbol{x}_2 = (i, 1)^{\mathrm{T}}$. Tarkastellaan nyt, voidaanko matriisi A diagonalisoida samoin kuin reaalimatriisit. Jos P on kääntyvä matriisi sarakevektoreinaan matriisin A ominaisvektorit ja D diagonaalimatriisi diagonaalialkioinaan matriisin A ominaisarvot, niin osoitetaan, että $P^{-1}AP = D$. Olkoon

$$P = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \qquad \text{ja} \qquad D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Nyt

$$P^{-1} = \frac{1}{2i}\begin{pmatrix} -1 & i \\ 1 & i \end{pmatrix}.$$

Tällöin

$$\begin{aligned}
P^{-1}AP &= \frac{1}{2i}\begin{pmatrix}-1 & i\\ 1 & i\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}-i & i\\ 1 & 1\end{pmatrix}\\
&= \frac{1}{2i}\begin{pmatrix}-1 & i\\ 1 & i\end{pmatrix}\begin{pmatrix}1 & 1\\ i & -i\end{pmatrix}\\
&= \frac{1}{2i}\begin{pmatrix}-2 & 0\\ 0 & 2\end{pmatrix}\\
&= \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix} = D.
\end{aligned}$$

Havaitaan tästä esimerkistä, että ominaisarvojen kompleksisia konjugaatteja, $\lambda_1 = i$ ja $\lambda_2 = \overline{\lambda_1} = -1$, vastaavat ominaisvektorit ovat toistensa kompleksisia konjugaattivektoreita, $\boldsymbol{x}_2 = \overline{\boldsymbol{x}}_1$. Tämä on yleisesti tosi reaalimatriiseilla.

Lause 4.1. *Jos A on $n \times n$ -reaalimatriisi, ja jos λ on sen ominaisarvo vastaavalla ominaisvektorilla $\boldsymbol{x}$, niin myös $\overline{\lambda}$ on matriisin A ominaisarvo vastaavalla ominaisvektorilla $\overline{\boldsymbol{x}}$.*

Todistus (vrt. [1, s. 401]). Koska A on reaalimatriisi, niin sen karakteristinen yhtälö on n. asteen polynomi reaalisilla kertoimilla, ja siten mitkä tahansa kompleksiset juuret esiintyvät konjugaattipareina. Tämä tarkoittaa, että jos λ on matriisin A ominaisarvo, niin on myös $\overline{\lambda}$. Jos λ on ominaisarvo vastaavalla ominaisvektorilla $\boldsymbol{x}$, niin $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Otetaan molemmin puolin kompleksinen konjugaatti ja saadaan $\overline{A\boldsymbol{x}} = \overline{\lambda\boldsymbol{x}} \Leftrightarrow A\overline{\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}}$, koska A on reaalimatriisi. Täten $\overline{\boldsymbol{x}}$ on ominaisvektori, joka vastaa ominaisarvoa $\overline{\lambda}$. □

4.3 Kompleksiset sisätuloavaruudet

Kahden reaalisen vektoriavaruuden $\mathbb{R}^n$ vektorin $\boldsymbol{x}$ ja $\boldsymbol{y}$ sisätulo on reaaliluku

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Vektorin $\boldsymbol{x}$ normi on tämän sisätulon määritelmän mukaan $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$. Tämä sisätulon määritelmä ei toimi kompleksisessa vektoriavaruudessa $\mathbb{C}^n$. Esimerkiksi, jos $\boldsymbol{x} = (1, 0, i)^{\mathrm{T}} \in \mathbb{C}^3$, niin selvästi $\boldsymbol{x} \neq \boldsymbol{0}$, mutta

$$\|\boldsymbol{x}\|^2 = x_1^2 + x_2^2 + x_3^2 = 1^2 + 0 + i^2 = 1 - 1 = 0.$$

Tätä määritelmää on muutettava, jotta se toimisi kompleksisessa vektoriavaruudessa. Tarkastellaan pykälässä 2.3 määritettyä kompleksiluvun modulia. Jos $z =$

$a + ib$, niin $|z|^2 = z\overline{z} = a^2 + b^2$ on epänegatiivinen reaaliluku, ja $|z| = 0$, vain jos $z = 0$.

Määritelmä 4.2. Olkoot $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$. Tällöin *standardi kompleksinen sisätulo* on kompleksiluku

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1\overline{y}_1 + x_2\overline{y}_2 + \cdots + x_n\overline{y}_n.$$

Esimerkki 4.3. Jos $\boldsymbol{x} = (1, 4i, 3+i)^{\mathrm{T}}$ ja $\boldsymbol{y} = (i, -3, 1-2i)^{\mathrm{T}}$, niin

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{y} \rangle &= 1(-i) + 4i(-3) + (3+i)(1+2i) \\
&= -i - 12i + (1+7i) \\
&= 1 - 6i.
\end{aligned}$$

Koska

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = x_1\overline{x}_1 + x_2\overline{x}_2 + \cdots + x_n\overline{x}_n = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$$

on vektorin $\boldsymbol{x}$ alkioiden modulien neliösumma, kompleksisen vektorin sisätulo itsensä kanssa on epänegatiivinen reaaliluku. Tällöin vektorin $\boldsymbol{x}$ normi on $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$, ja $\|\boldsymbol{x}\| = 0$, jos ja vain jos $\boldsymbol{x}$ on nollavektori.

Lause 4.2. *Tarkastellaan vektoriavaruutta $\mathbb{C}^n$. Standardi kompleksinen sisätulo*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1\overline{y}_1 + x_2\overline{y}_2 + \cdots + x_n\overline{y}_n$$

toteuttaa seuraavat ehdot kaikilla $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{C}^n$ ja kaikilla $\alpha, \beta \in \mathbb{C}$:

(i) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$

(ii) $\langle \alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{z} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta\langle \boldsymbol{y}, \boldsymbol{z} \rangle$

(iii) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ *ja* $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$*, jos ja vain jos* $\boldsymbol{x} = \boldsymbol{0}$.

Todistus (vrt. [1, s. 402]). Todistetaan kohta (i). Nyt

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{y} \rangle &= x_1\overline{y}_1 + x_2\overline{y}_2 + \cdots + x_n\overline{y}_n \\
&= \overline{y}_1x_1 + \overline{y}_2x_2 + \cdots + \overline{y}_nx_n \\
&= \overline{y_1\overline{x}_1} + \overline{y_2\overline{x}_2} + \cdots + \overline{y_n\overline{x}_n} \\
&= \overline{y_1\overline{x}_1 + y_2\overline{x}_2 + \cdots + y_n\overline{x}_n} \\
&= \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle},
\end{aligned}$$

joten kohta (i) on tosi.

Todistetaan kohta (ii). Nyt

$$\begin{aligned}\langle \alpha \boldsymbol{x} + \beta \boldsymbol{y}, \boldsymbol{z} \rangle &= (\alpha x_1 + \beta y_1)\overline{z}_1 + (\alpha x_2 + \beta y_2)\overline{z}_2 + \cdots + (\alpha x_n + \beta y_n)\overline{z}_n \\ &= \alpha(x_1\overline{z}_1 + x_2\overline{z}_2 + \cdots + x_n\overline{z}_n) + \beta(y_1\overline{z}_1 + y_2\overline{z}_2 + \cdots + y_n\overline{z}_n) \\ &= \alpha\langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta\langle \boldsymbol{y}, \boldsymbol{z} \rangle,\end{aligned}$$

joten kohta (ii) on tosi.

Todistetaan kohta (iii). Huomataan, että jos $x_j = a_j + ib_j$ on vektorin $\boldsymbol{x}$ j:s alkio, niin

$$\begin{aligned}\langle \boldsymbol{x}, \boldsymbol{x} \rangle &= |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 \\ &= a_1^2 + b_1^2 + a_2^2 + b_2^2 + \cdots + a_n^2 + b_n^2\end{aligned}$$

on reaalilukujen neliösumma, joten $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ ja $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$, jos ja vain jos jokainen termi a_j^2 ja b_j^2 on 0, toisin sanoen, jos ja vain jos $\boldsymbol{x} = \boldsymbol{0}$. □

Kuten reaalisilla vektoriavaruuksilla, myös kompleksisten vektoriavaruuksien sisätulolle on yleinen määritelmä.

Määritelmä 4.3 (Kompleksinen sisätulo)**.** Olkoon V kompleksinen vektoriavaruus. Tällöin kuvausta $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{C}$ sanotaan *kompleksiseksi sisätuloksi*, jos

(i) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$ kaikilla $\boldsymbol{x}, \boldsymbol{y} \in V$,

(ii) $\langle \alpha \boldsymbol{x} + \beta \boldsymbol{y}, \boldsymbol{z} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta\langle \boldsymbol{y}, \boldsymbol{z} \rangle$ kaikilla $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ ja kaikilla $\alpha, \beta \in \mathbb{C}$,

(iii) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ kaikilla $\boldsymbol{x} \in V$, ja $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$, jos ja vain jos $\boldsymbol{x} = \boldsymbol{0}$.

Jos vektoriavaruudessa on määritelty kompleksinen sisätulo, sanotaan sitä *kompleksiseksi sisätuloavaruudeksi*. Mistä tahansa kompleksisesta sisätulosta voidaan määritellä normi

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

Määritelmän 4.2 standardi kompleksinen sisätulo on selvästi sisätulo tämän yleisen määritelmän perusteella.

Määritelmästä 4.3 seuraa suoraan kaksi seuraavaa ominaisuutta:

- $\langle \boldsymbol{x}, \alpha \boldsymbol{y} \rangle = \overline{\alpha}\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ kaikilla $\boldsymbol{x}, \boldsymbol{y} \in V$ ja kaikilla $\alpha \in \mathbb{C}$,
- $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$ kaikilla $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$.

Ortogonaalisten vektorien määritelmä kompleksisessa sisätuloavaruudessa on täsmälleen sama kuin reaalisessa sisätuloavaruudessa:

Määritelmä 4.4. Kompleksisen sisätuloavaruuden vektoreiden $\boldsymbol{x}$ ja $\boldsymbol{y}$ sanotaan olevan *ortogonaaliset*, jos

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0.$$

Tällöin merkitään $\boldsymbol{x} \perp \boldsymbol{y}$.

Kompleksisen vektoriavaruuden V vektorijoukko $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ on *ortogonaalinen*, jos $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0$, kun $i \neq j$. Vektorijoukon sanotaan olevan *ortonormaali*, jos jokainen sen vektori on myös yksikkövektori.

Lauseen 2.3 mukaan reaalisen vektoriavaruuden $\mathbb{R}^n$ ortogonaalinen vektorijoukko, missä ei ole nollavektoreita, on lineaarisesti riippumaton. Sama pätee myös kompleksisessa vektoriavaruudessa $\mathbb{C}^n$.

Lause 4.3. *Oletetaan, että V on kompleksinen sisätuloavaruus ja vektorit $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ ovat pareittain ortogonaalisia ($\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0$, kun $i \neq j$) ja yksikään ei ole nollavektori. Tällöin vektorijoukko $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ on lineaarisesti riippumaton.*

Todistus (vrt. [1, s. 405]). Täytyy osoittaa, että jos

$$\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_n \boldsymbol{v}_n = \boldsymbol{0}, \quad \alpha_i \in \mathbb{C},$$

niin $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Olkoon i mikä tahansa kokonaisluku lukujen 1 ja n välissä. Tällöin

$$\langle \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_n \boldsymbol{v}_n, \boldsymbol{v}_i \rangle = \langle \boldsymbol{0}, \boldsymbol{v}_i \rangle = 0,$$

mutta

$$\begin{aligned}
&\langle \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_n \boldsymbol{v}_n, \boldsymbol{v}_i \rangle \\
&\quad = \alpha_1 \langle \boldsymbol{v}_1, \boldsymbol{v}_i \rangle + \cdots + \alpha_{i-1} \langle \boldsymbol{v}_{i-1}, \boldsymbol{v}_i \rangle + \alpha_i \langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle \\
&\qquad + \alpha_{i+1} \langle \boldsymbol{v}_{i+1}, \boldsymbol{v}_i \rangle + \cdots + \alpha_n \langle \boldsymbol{v}_n, \boldsymbol{v}_i \rangle.
\end{aligned}$$

Koska $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0$, kun $i \neq j$, niin

$$\langle \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_n \boldsymbol{v}_n, \boldsymbol{v}_i \rangle = \alpha_i \langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle = \alpha_i \|\boldsymbol{v}_i\|^2.$$

Koska $\boldsymbol{v}_i \neq \boldsymbol{0}$, niin $\|\boldsymbol{v}_i\|^2 \neq 0$, joten $\alpha_i = 0$. Mutta koska i oli mikä tahansa kokonaisluku välillä $[1, k]$, niin tästä seuraa, että

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Vektorijoukko $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ on siis lineaarisesti riippumaton. $\square$

Esimerkki 4.4. Vektorit

$$\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}, \qquad \boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}i\\ 1\end{pmatrix}$$

muodostavat vektoriavaruuden $\mathbb{C}^2$ ortonormaalin kannan. Vektorit ovat ortogonaalisia, koska

$$\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle = \frac{1}{2}\left\langle \begin{pmatrix}1\\ i\end{pmatrix}, \begin{pmatrix}i\\ 1\end{pmatrix}\right\rangle = \frac{1}{2}(1(-i) + i(1)) = 0,$$

ja kunkin vektorin normi on 1, koska

$$\langle \boldsymbol{v}_1, \boldsymbol{v}_1\rangle = \frac{1}{2}\left\langle \begin{pmatrix}1\\ i\end{pmatrix}, \begin{pmatrix}1\\ i\end{pmatrix}\right\rangle = \frac{1}{2}(1(1) + i(-i)) = 1$$

ja

$$\langle \boldsymbol{v}_2, \boldsymbol{v}_2\rangle = \frac{1}{2}\left\langle \begin{pmatrix}i\\ 1\end{pmatrix}, \begin{pmatrix}i\\ 1\end{pmatrix}\right\rangle = \frac{1}{2}(i(-i) + 1(1)) = 1.$$

Vektorit $\boldsymbol{v}_1$ ja $\boldsymbol{v}_2$ muodostavat siis vektoriavaruuden $\mathbb{C}^2$ kannan, koska ne ovat lineaarisesti riippumattomat ja vektoriavaruuden $\mathbb{C}^2$ dimensio on 2.

5 Unitaarinen diagonalisointi

Tässä luvussa tarkastellaan kompleksisia matriiseja ja niiden unitaarista diagonalisointia ja verrataan niitä reaalimatriisien ortogonaaliseen diagonalisointiin. Pykälässä 5.1 tarkastellaan Hermiten konjugaattia (*konjugaattitranspoosi*) ja hermiittisiä matriiseja, pykälässä 5.2 unitaarisia matriiseja, ja pykälässä 5.3 unitaarista diagonalisointia ja normaaleja matriiseja.

5.1 Hermiten konjugaatti ja hermiittiset matriisit

Jos A on kompleksinen matriisi, niin sen *Hermiten konjugaatti* on matriisi $\overline{A}^{\mathrm{T}}$, jota merkitään notaatiolla A^*. Huomataan, että reaalimatriisin transpoosilla on vastaavat ominaisuudet kuin Hermiten konjugaatilla. Lisäksi, kuten reaalimatriisin transpoosilla on tärkeä rooli ortogonaalisessa diagonalisoinnissa, niin Hermiten konjugaatilla on vastaavasti kompleksimatriisin unitaarisessa diagonalisoinnissa.

Määritelmä 5.1 (Hermiten konjugaatti)**.** Olkoon A $m \times n$ -matriisi kompleksialkioilla. Matriisin A *Hermiten konjugaatti* on

$$A^* = \overline{A}^{\mathrm{T}},$$

missä $A = (a_{ij})$, $\overline{A} = (\overline{a_{ij}})$ ja $A^* = \overline{A}^{\mathrm{T}} = (\overline{a_{ji}})$.

Esimerkki 5.1. Kompleksimatriisin

$$A = \begin{pmatrix} i & 5-3i & 2+i \\ 3 & 1+2i & 4-9i \end{pmatrix}$$

Hermiten konjugaatti on

$$A^* = \begin{pmatrix} -i & 3 \\ 5+3i & 1-2i \\ 2-i & 4+9i \end{pmatrix}.$$

Hermiten konjugaatilla on seuraavat ominaisuudet

- $(A^*)^* = A$,
- $(cA)^* = \overline{c}A^*$,
- $(A+B)^* = A^* + B^*$,

- $(AB)^* = B^* A^*$.

Verrataan näitä ominaisuuksia pykälässä 3.1 esitettyihin reaalimatriisin transpoosin ominaisuuksiin ja huomataan, että ne ovat vastaavat. Huomataan myös, että jos A on reaalimatriisi, niin $A^* = \overline{A}^{\mathrm{T}} = A^{\mathrm{T}}$.

Palautetaan mieleen, että reaalimatriisi A on symmetrinen, jos $A = A^{\mathrm{T}}$, kuten määriteltiin pykälässä 3.1. Symmetrisen matriisin kompleksinen vastine on *hermiittinen* matriisi.

Määritelmä 5.2 (Hermiittinen matriisi)**.** Olkoon A kompleksinen $n \times n$ -matriisi. Matriisia A kutsutaan *hermiittiseksi*, jos

$$A = A^*.$$

Reaalialkioinen hermiittinen matriisi on symmetrinen matriisi, koska $A = A^* = A^{\mathrm{T}}$. Miltä hermiittinen matriisi sitten näyttää? Jos $A = (a_{ij}) = A^* = (\overline{a_{ij}})$, niin diagonaalialkioiden tulee olla reaalilukuja, koska silloin $a_{ii} = \overline{a_{ii}}$. Lisäksi alkioiden a_{ij} ja a_{ji} on oltava toistensa kompleksiset konjugaatit.

Esimerkki 5.2. Matriisi

$$A = \begin{pmatrix} 1 & 1+2i & 4-i \\ 1-2i & -3 & i \\ 4+i & -i & 2 \end{pmatrix}$$

on hermiittinen matriisi.

Symmetristen matriisien ortogonaalisessa diagonalisoinnissa todettiin lauseessa 3.1, että symmetrisen matriisin ominaisarvot ovat reaalilukuja. Nyt tämä voidaan todistaa. Tulos on seuraavan lauseen seuraus.

Lause 5.1. *Jos matriisi A on hermiittinen matriisi, niin sen ominaisarvot ovat reaalisia.*

Todistus (vrt. [1, s. 409]). Oletetaan, että λ on matriisin A ominaisarvo, jota vastaava ominaisvektori on $\boldsymbol{x}$. Tällöin

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \quad \text{ja} \quad \boldsymbol{x} \neq \boldsymbol{0}.$$

Kerrotaan tämä yhtälö vasemmalta puolelta vektorin $\boldsymbol{x}$ Hermiten konjugaatilla. Nyt

$$\boldsymbol{x}^* A\boldsymbol{x} = \boldsymbol{x}^* \lambda\boldsymbol{x} = \lambda\boldsymbol{x}^*\boldsymbol{x} = \lambda\|\boldsymbol{x}\|^2,$$

missä vektorin $\boldsymbol{x}$ normi on positiivinen reaaliluku. Kun otetaan kompleksinen konjugaattitranspoosi yhtälön $A\boldsymbol{x} = \lambda\boldsymbol{x}$ molemmilta puolilta saadaan

$$(A\boldsymbol{x})^* = (\lambda\boldsymbol{x})^*,$$

josta tulee

$$\boldsymbol{x}^*A^* = \overline{\lambda}\boldsymbol{x}^*.$$

Kerrotaan sitten yhtälö oikealta vektorilla $\boldsymbol{x}$. Nyt

$$\boldsymbol{x}^*A^*\boldsymbol{x} = \overline{\lambda}\boldsymbol{x}^*\boldsymbol{x} = \overline{\lambda}\|\boldsymbol{x}\|^2.$$

Koska A on hermiittinen, niin $\boldsymbol{x}^*A\boldsymbol{x} = \boldsymbol{x}^*A^*\boldsymbol{x}$, mistä seuraa, että

$$\lambda\|\boldsymbol{x}\|^2 = \overline{\lambda}\|\boldsymbol{x}\|^2.$$

Koska $\|\boldsymbol{x}\|^2 \neq 0$, niin $\lambda = \overline{\lambda}$, joten λ on reaalinen. □

Seuraus 5.1. *Jos reaalimatriisi A on symmetrinen, niin sen ominaisarvot ovat reaalisia.*

Kuten symmetrisillä reaalimatriiseilla todettiin lauseessa 3.2, myös hermiittisen matriisin erisuuria ominaisarvoja vastaavat ominaisvektorit ovat ortogonaaliset.

Lause 5.2. *Jos matriisi A on hermiittinen, niin sen erisuuria ominaisarvoja vastaavat ominaisvektorit ovat ortogonaaliset.*

Todistus (vrt. [1, s. 332]). Oletetaan, että λ ja μ ovat matriisin A mitkä tahansa kaksi erisuurta ominaisarvoa, ja oletetaan, että $\boldsymbol{x}$ ja $\boldsymbol{y}$ ovat vastaavat ominaisvektorit. Tällöin $A\boldsymbol{x} = \lambda\boldsymbol{x}$ ja $A\boldsymbol{y} = \mu\boldsymbol{y}$. Nyt

$$\boldsymbol{y}^*A\boldsymbol{x} = \boldsymbol{y}^*\lambda\boldsymbol{y} = \langle\lambda\boldsymbol{x}, \boldsymbol{y}\rangle = \lambda\langle\boldsymbol{x}, \boldsymbol{y}\rangle.$$

Toisaalta, koska $A = A^*$, niin

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

Lauseen 5.1 mukaan hermiittisen matriisin ominaisarvot ovat reaalisia, joten $\overline{\mu} = \mu$, ja siten näistä kahdesta muodosta saadaan

$$\lambda\langle\boldsymbol{x}, \boldsymbol{y}\rangle = \mu\langle\boldsymbol{x}, \boldsymbol{y}\rangle \quad \Leftrightarrow \quad (\lambda - \mu)\langle\boldsymbol{x}, \boldsymbol{y}\rangle.$$

Koska μ ja λ ovat erisuuria, niin $\mu - \lambda \neq 0$, joten $\langle\boldsymbol{x}, \boldsymbol{y}\rangle = 0$. Täten $\boldsymbol{x}$ ja $\boldsymbol{y}$ ovat ortogonaaliset. □

5.2 Unitaariset matriisit

Ortogonaalinen matriisi määriteltiin (Määritelmä 3.3) pykälässä 3.2. Sen kompleksinen vastine on *unitaarinen* matriisi.

Määritelmä 5.3 (Unitaarinen matriisi)**.** Kompleksinen $n \times n$ -matriisi P on *unitaarinen*, jos $PP^* = P^*P = I$. Toisin sanoen, matriisi P on unitaarinen, jos sillä on käänteismatriisi P^*.

Tästä määritelmästä seuraa, että jos P on unitaarinen matriisi, niin samoin on P^*.

Esimerkki 5.3. Olkoon

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

matriisi. Tällöin

$$\begin{aligned} PP^* &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}} \cdot \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \cdot \frac{-i}{\sqrt{2}} + \frac{i}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{-i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \cdot \frac{-i}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} + \frac{-i^2}{2} & \frac{-i}{2} + \frac{i}{2} \\ \frac{i}{2} + \frac{-i}{2} & \frac{-i^2}{2} + \frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \end{aligned}$$

ja

$$\begin{aligned} P^*P &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{-i}{\sqrt{2}} \cdot \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \cdot \frac{i}{\sqrt{2}} + \frac{-i}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{i}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \cdot \frac{i}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} + \frac{-i^2}{2} & \frac{i}{2} + \frac{-i}{2} \\ \frac{-i}{2} + \frac{i}{2} & \frac{-i^2}{2} + \frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I. \end{aligned}$$

Siis $PP^* = P^*P = I$, joten P on unitaarinen matriisi.

Unitaarinen matriisi P reaalialkioilla on ortogonaalinen matriisi, koska $P^* = P^{\mathrm{T}}$. Lauseessa 3.4 osoitettiin, että matriisi P on ortogonaalinen, jos ja vain jos sen sarakevektorit muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan. Todistetaan vastaava tulos unitaarisille matriiseille.

Lause 5.3. *Olkoon P $n \times n$ -matriisi. Tällöin P on unitaarinen, jos ja vain jos sen sarakevektorit muodostavat vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan.*

Todistus (vrt. [1, s. 410]). Oletetaan, että matriisi P on unitaarinen. Olkoot $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ matriisin P sarakevektorit. Tällöin matriisin P^* rivivektorit ovat näiden sarakevektorien kompleksiset konjugaattitranspoosit.

Koska $I = P^*P$, niin

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1^* \\ \boldsymbol{x}_2^* \\ \vdots \\ \boldsymbol{x}_n^* \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \end{pmatrix}$$

Muodostamalla yhtälöt näiden matriisin alkioista saadaan $\boldsymbol{x}_i^*\boldsymbol{x}_j = \langle \boldsymbol{x}_j, \boldsymbol{x}_i \rangle = 0$, kun $i \neq j$, ja $\boldsymbol{x}_i^*\boldsymbol{x}_i = \langle \boldsymbol{x}_i, \boldsymbol{x}_i \rangle = 1$, kun $i = j$. Tämä tarkoittaa täsmälleen sitä, että sarakevektorit $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ muodostavat ortonormaalin vektorijoukon. Täten ne ovat lineaarisesti riippumattomia, ja koska niitä on n kappaletta, muodostavat ne vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan.

Oletetaan sitten, että matriisin P sarakevektorit muodostavat vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan. Tällöin matriisitulon P^*P tulee olla identiteettimatriisi, kuten yllä osoitettiin, joten $P^*P = I$. Nyt $P^* = P^{-1}$, joten myös $PP^* = I$. □

Koska P^* on unitaarinen, tämä tulos pätee myös matriisin P rivivektoreihin.

Kuten ortogonaalisilla matriiseilla todettiin lauseessa 3.5, myös unitaarisen matriisin P määräämä lineaarikuvaus on *isometria*, eli se säilyttää sisätulon tuloksen, ja siten minkä tahansa vektorin pituuden. Itse asiassa tämä ominaisuus määrittää unitaarisen matriisin, kuten seuraava lause osoittaa.

Lause 5.4. *Matriisi P on unitaarinen, jos ja vain jos sen määräämä lineaarikuvaus säilyttää kompleksisen sisätulon tuloksen vektoriavaruudessa $\mathbb{C}^n$, eli $\langle P\boldsymbol{x}, P\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ kaikilla $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$.*

Todistus (vrt. [1, s. 411]). Oletetaan ensin, että P on unitaarinen matriisi. Tällöin

$$\langle P\boldsymbol{x}, P\boldsymbol{y} \rangle = (P\boldsymbol{y})^*(P\boldsymbol{x}) = \boldsymbol{y}^*P^*P\boldsymbol{x} = \boldsymbol{y}^*I\boldsymbol{x} = \boldsymbol{y}^*\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle,$$

joten P säilyttää sisätulon tuloksen.

Oletetaan sitten, että on olemassa matriisi P, jolle

$$\langle P\boldsymbol{x}, P\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$$

kaikilla $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$. Olkoon $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ vektoriavaruuden $\mathbb{C}^n$ luonnollinen kanta. Tällöin $P\boldsymbol{e}_i = \boldsymbol{v}_i$, missä $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ ovat matriisin P sarakevektorit. Nyt

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \langle P\boldsymbol{e}_i, P\boldsymbol{e}_j\rangle = \langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle,$$

mistä voidaan päätellä, että matriisin P sarakevektorit muodostavat vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan, ja P on täten unitaarinen matriisi lauseen 5.3 mukaan. □

5.3 Unitaarinen diagonalisointi ja normaalit matriisit

Pykälässä 3.3 määriteltiin (Määritelmä 3.4) ortogonaalisesti diagonalisoituva matriisi. *Unitaarisesti* diagonalisoituva matriisi määritellään vastaavasti.

Määritelmä 5.4. Neliömatriisin A sanotaan olevan *unitaarisesti diagonalisoituva*, jos on olemassa sellainen unitaarinen matriisi P, että $P^*AP = D$, missä D on diagonaalimatriisi.

Lauseessa 3.6 osoitettiin, että reaalimatriisi on ortogonaalisesti diagonalisoituva, jos ja vain jos vektoriavaruudelle $\mathbb{R}^n$ on olemassa sen ominaisvektoreista muodostuva ortonormaali kanta. Vastaava tulos pätee myös kompleksisille matriiseille.

Lause 5.5. *Matriisi A on unitaarisesti diagonalisoituva, jos ja vain jos vektoriavaruudelle $\mathbb{C}^n$ on olemassa matriisin A ominaisvektoreista muodostuva ortonormaali kanta.*

Todistus (vrt. [1, s. 412]). Oletetaan ensin, että matriisi A on unitaarisesti diagonalisoituva, eli $P^*AP = D$. Koska matriisi P diagonalisoi matriisin A, niin matriisin P sarakevektorit ovat matriisin A ominaisvektorit, ja ne muodostavat vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan.

Oletetaan sitten, että matriisin A ominaisvektorit muodostavat vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan. Tällöin matriisi P, jonka sarakevektorit ovat nämä kantavektorit, on unitaarinen lauseen 5.3 mukaan. Koska matriisin P sarakevektorit ovat matriisin A ominaisvektorit saadaan yhtälö $AP = PD$, missä D on diagonaalimatriisi diagonaalialkioinaan matriisin A ominaisarvot. Koska P on unitaarinen matriisi, niin

$$AP = PD \quad \Leftrightarrow \quad P^*AP = D,$$

joten A on unitaarisesti diagonalisoituva matriisi. □

Spektraalilauseessa 3.7 todettiin, että vain symmetrinen matriisi on ortogonaalisesti diagonalisoituva. On kiinnostavaa, mikä on vastaava tulos kompleksisille matriiseille. Hermiittiset matriisit ovat unitaarisesti diagonalisoituvia, mutta ne eivät ole ainoita sellaisia matriiseja. On olemassa laajempi joukko kompleksisia matriiseja, jotka ovat unitaarisesti diagonalisoituvia. Näitä kutsutaan *normaaleiksi matriiseiksi*.

Määritelmä 5.5 (Normaali matriisi)**.** Kompleksisen $n \times n$ -matriisin A sanotaan olevan *normaali*, jos

$$AA^* = A^*A.$$

Jokainen hermiittinen matriisi on normaali, koska $AA^* = AA = A^*A$. Myös jokainen unitaarinen matriisi on normaali, koska $AA^* = I = A^*A$, mutta kaikki hermiittiset tai unitaariset matriisit eivät ole normaaleita.

Esimerkki 5.4. Olkoon

$$A = \begin{pmatrix} 1 & 1+i & 1 \\ -1+i & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$$

matriisi. Tällöin

$$AA^* = \begin{pmatrix} 1 & 1+i & 1 \\ -1+i & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1-i & -1 \\ 1-i & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 1 & -1-i \\ 1 & 4 & 1-i \\ -1+i & 1+i & 3 \end{pmatrix}$$

ja

$$A^*A = \begin{pmatrix} 1 & -1-i & -1 \\ 1-i & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1+i & 1 \\ -1+i & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 1 & -1-i \\ 1 & 4 & 1-i \\ -1+i & 1+i & 3 \end{pmatrix}.$$

Nyt $AA^* = A^*A$, joten matriisi on normaali. Huomataan, että $A \neq A^*$ ja $AA^* = A^*A \neq I$, joten matriisi A ei ole hermiittinen eikä unitaarinen.

Lisäksi jokainen diagonaalimatriisi on normaali. Olkoon
$D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$, missä d_i on alkio paikassa (i, i) ja muut alkiot ovat nollia. Tällöin D^* on diagonaalimatriisi diagonaalialkioin $\overline{d}_i$. Nyt

$$DD^* = \operatorname{diag}(|d_1|^2, |d_2|^2, \ldots, |d_n|^2) = D^*D.$$

Täten diagonaalimatriisi D on normaali ja diagonaalimatriisin DD^* alkiot ovat reaalisia.

Huomautus. Diagonaalimatriisit eivät välttämättä ole hermiittisiä eikä unitaarisia.

Todistetaan seuraava tärkeä lause, joka on kompleksisten matriisien vastine reaalimatriisien spektraalilauseelle (Lause 3.7).

Lause 5.6. *Matriisi A on unitaarisesti diagonalisoituva, jos ja vain jos A on normaali.*

Määritellään nyt unitaarisesti similaariset matriisit ja todistetaan kolme apulausetta, joita käytetään lauseen 5.6 todistuksessa.

Määritelmä 5.6. Olkoot A ja B neliömatriiseja. Matriisin A sanotaan olevan *unitaarisesti similaarinen* matriisin B kanssa, jos on olemassa sellainen unitaarinen matriisi P, että $A = PBP^*$.

Apulause 5.1 (Schurin hajotelma)**.** Olkoot $\lambda_1, \lambda_2, \ldots, \lambda_n$ kompleksisen $n \times n$ -matriisin A ominaisarvot. Tällöin on olemassa unitaarinen matriisi P siten, että

$$P^*AP = T = [t_{ij}],$$

missä T on kompleksinen yläkolmiomatriisi, jonka diagonaalialkiot $t_{ii} = \lambda_i$, kun $i = 1, 2, \ldots, n$.

Kompleksinen matriisi on siis unitaarisesti similaarinen sellaisen yläkolmiomatriisin kanssa, jonka diagonaalialkiot ovat sen ominaisarvot.

Todistus (vrt. [3, s. 101]). Olkoot A kompleksinen $n \times n$ -matriisi ja $\boldsymbol{x}$ sen ominaisarvoa λ_1 vastaava normeerattu ominaisvektori. Nyt siis $\boldsymbol{x}^*\boldsymbol{x} = 1$ ja $A\boldsymbol{x} = \lambda_1\boldsymbol{x}$. Olkoon $P_1 = (\boldsymbol{x} \quad \boldsymbol{p}_2 \quad \cdots \quad \boldsymbol{p}_n)$ unitaarinen matriisi, jonka ensimmäinen sarakevektori on $\boldsymbol{x}$. Tällöin

$$\begin{aligned}
P_1^*AP_1 &= P_1^*\begin{pmatrix} A\boldsymbol{x} & A\boldsymbol{p}_2 & \cdots & A\boldsymbol{p}_n \end{pmatrix} \\
&= P_1^*\begin{pmatrix} \lambda_1\boldsymbol{x} & A\boldsymbol{p}_2 & \cdots & A\boldsymbol{p}_n \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{x}^* \\ \boldsymbol{p}_2^* \\ \vdots \\ \boldsymbol{p}_n^* \end{pmatrix}\begin{pmatrix} \lambda_1\boldsymbol{x} & A\boldsymbol{p}_2 & \cdots & A\boldsymbol{p}_n \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1\boldsymbol{x}^*\boldsymbol{x} & \boldsymbol{x}^*A\boldsymbol{p}_2 & \cdots & \boldsymbol{x}^*A\boldsymbol{p}_n \\ \lambda_1\boldsymbol{p}_2^*\boldsymbol{x} & & & \\ \vdots & & A_1 & \\ \lambda_1\boldsymbol{p}_n^*\boldsymbol{x} & & & \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & \star \\ 0 & A_1 \end{pmatrix},
\end{aligned}$$

koska matriisin P_1 sarakevektorit ovat ortonormaalit. Alimatriisin A_1 ominaisarvot ovat $\lambda_2, \lambda_3, \ldots, \lambda_n$.

Jos $n = 2$, niin lause on todistettu. Jos $n > 2$, niin olkoon $\boldsymbol{y} \in \mathbb{C}^{n-1}$ alimatriisin A_1 ominaisarvoa λ_2 vastaava normeerattu ominaisvektori.

Suoritetaan edellinen reduktio alimatriisille A_1. Jos P_2 on unitaarinen $(n-1) \times (n-1)$ -matriisi, jonka ensimmäinen sarakevektori on $\boldsymbol{y}$, niin

$$P_2^* A_1 P_2 = \begin{pmatrix} \lambda_2 & \star \\ 0 & A_2 \end{pmatrix}.$$

Olkoon $Q_2 = [1] \oplus P_2$, joka on matriisien $[1]$ ja P_2 suora summa (ks. [3, s. 30]), ja

$$(P_1 Q_2)^* A P_1 Q_2 = Q_2^* P_1^* A P_1 Q_2 = \begin{pmatrix} \lambda_1 & \star & \star \\ 0 & \lambda_2 & \star \\ 0 & 0 & A_2 \end{pmatrix}.$$

Jatketaan tätä reduktiota, joka tuottaa unitaarisia $(n-i+1) \times (n-i+1)$ -matriiseja P_i, missä $i = 1, 2, \ldots, n-1$, ja unitaarisia $n \times n$ -matriiseja Q_j, missä $j = 2, 3, \ldots, n-2$.

Matriisi

$$P = P_1 Q_2 Q_3 \cdots V_{n-2}$$

on unitaarinen ja $P^* A P$ on kompleksinen yläkolmiomatriisi. □

Huomautus. Jos kompleksisen matriisin A kaikki ominaisarvot ovat reaalisia, matriisiksi P voidaan valita myös ortogonaalinen reaalimatriisi.

Apulause 5.2. Olkoot A ja B unitaarisesti similaarisia matriiseja. Jos A on normaali matriisi, niin myös matriisi B on normaali.

Todistus. Olkoot A normaali $n \times n$ -matriisi ja B $n \times n$ -matriisi. Koska A ja B ovat unitaarisesti similaarisia, niin $B = P^* A P$, missä P on unitaarinen matriisi. Tällöin

$$\begin{aligned} B^* B &= (P^* A P)^* (P^* A P) \\ &= P^* A^* P P^* A P \\ &= P^* A^* A P \\ &= P^* A A^* P \\ &= P^* A P P^* A^* P \\ &= (P^* A P)(P^* A^* P) \\ &= (P^* A P)(P^* A P)^* \\ &= B B^*, \end{aligned}$$

joten B on normaali matriisi. □

Apulause 5.3. Olkoon A normaali matriisi. Jos A on yläkolmiomatriisi, niin se on diagonaalimatriisi.

Todistus (vrt. [4, s. 82]). Olkoon A normaali yläkolmiomatriisi. Tällöin $a_{ij} = 0$, kun $i \geq j$, joten matriisien A^*A ja AA^* diagonaalialkiot ovat

$$(A^*A)_{ii} = \sum_{k=1}^{i} |a_{ki}|^2 \qquad \text{ja} \qquad (AA^*)_{ii} = \sum_{k=i}^{n} |a_{ik}|^2.$$

Kun $i = 1$, niin

$$|a_{11}|^2 = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2,$$

mistä seuraa, että

$$a_{12} = a_{13} = \cdots = a_{1n} = 0,$$

eli matriisin A ensimmäisellä rivillä vain alkio a_{11} voi olla nollasta poikkeava.

Kun $i = 2$, niin

$$|a_{12}|^2 + |a_{22}|^2 = |a_{22}|^2 + |a_{23}|^2 + \cdots + |a_{2n}|^2.$$

Koska $a_{12} = 0$, niin matriisin A toisella rivillä vain alkio a_{22} voi olla nollasta poikkeava. Kun tätä toistetaan jokaiselle riville, huomataan, että A on diagonaalimatriisi. Täten mikä tahansa normaali yläkolmiomatriisi on diagonaalimatriisi. □

Nyt voidaan todistaa Lause 5.6.

Todistus. Oletetaan ensin, että A on unitaarisesti diagonalisoituva matriisi. Tällöin on olemassa unitaarinen matriisi P ja diagonaalimatriisi D siten, että $P^*AP = D$. Kirjoitetaan tämä yhtälö muodossa $A = PDP^*$. Nyt

$$\begin{aligned} AA^* &= (PDP^*)(PDP^*)^* \\ &= (PDP^*)(PD^*P^*) \\ &= PD(P^*P)D^*P^* \\ &= P(DD^*)P^* \end{aligned}$$

Samoin

$$\begin{aligned} A^*A &= (PDP^*)^*(PDP^*) \\ &= (PD^*P^*)(PDP^*) \\ &= PD^*(P^*P)DP^* \\ &= P(D^*D)P^* \end{aligned}$$

Koska D on diagonaalimatriisi, niin se on normaali, joten $P(DD^*)P^* = P(D^*D)P^*$, eli $AA^* = A^*A$. Matriisi A on siis normaali.

Oletetaan sitten, että A on normaali matriisi ja T on yläkolmiomatriisi siten, että se on unitaarisesti similaarinen matriisin A kanssa, eli $P^*AP = T$, missä P on unitaarinen matriisi. Koska matriisit A ja T ovat unitaarisesti similaarisia Schurin hajotelman (Apulause 5.1) mukaan, niin apulauseen 5.2 mukaan myös T on normaali matriisi. Koska T on yläkolmiomatriisi ja normaali, niin apulauseen 5.3 mukaan se on diagonaalimatriisi. Normaali matriisi A on täten unitaarisesti diagonalisoituva. $\square$

Normaali matriisi A diagonalisoidaan unitaarisesti siten, että ensin löydetään sen ominaisarvot. Jokaista ominaisarvoa vastaavalle ominaisavaruudelle löydetään sitten ortonormaali kanta. Tällöin kaikkien tällaisten ominaisvektorien joukko muodostaa vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan. Lauseen 5.6 sisältö mahdollistaa tämän aina. Muodostetaan matriisi P, jonka sarakevektorit ovat matriisin A ominaisvektorit. Tällöin P on unitaarinen ja $P^*AP = D$, missä D on diagonaalimatriisi diagonaalialkioinaan matriisin A vastaavat ominaisarvot.

Ortogonaalinen diagonalisointi on unitaarisen diagonalisoinnin erikoistapaus, missä A on symmetrinen reaalimatriisi. Annetaan nyt esimerkki kompleksisen matriisin unitaarisesta diagonalisoinnista.

Esimerkki 5.5. Tarkastellaan erästä normaalia matriisia. Olkoon

$$A = \begin{pmatrix} 1 & 2+i \\ 2-i & 5 \end{pmatrix}.$$

Sen ominaisarvot ovat

$$\lambda_1 = 0 \qquad \text{ja} \qquad \lambda_2 = 6,$$

ja vastaavista ominaisvektoreista muodostettu matriisi on

$$P = \begin{pmatrix} \frac{2+i}{\sqrt{6}} & \frac{2+i}{\sqrt{30}} \\ \frac{-1}{\sqrt{6}} & \frac{5}{\sqrt{30}} \end{pmatrix}.$$

Tällöin $P^*P = I$, joten P on unitaarinen matriisi. Nyt

$$P^*AP = \begin{pmatrix} \frac{2-i}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \\ \frac{2-i}{\sqrt{30}} & \frac{5}{\sqrt{30}} \end{pmatrix} \begin{pmatrix} 1 & 2+i \\ 2-i & 5 \end{pmatrix} \begin{pmatrix} \frac{2+i}{\sqrt{6}} & \frac{2+i}{\sqrt{30}} \\ \frac{-1}{\sqrt{6}} & \frac{5}{\sqrt{30}} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 6 \end{pmatrix}$$

on diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin A ominaisarvot.

Tiedetään, että kaikki hermiittiset matriisit ovat normaaleja ja kuten lauseessa 5.1 todistettiin, niin hermiittisellä matriisilla on reaaliset ominaisarvot. Todistetaan nyt, että jos normaalin matriisin kaikki ominaisarvot ovat reaalisia, niin se on hermiittinen.

Lause 5.7. *Olkoon A normaali matriisi. Jos matriisin A kaikki ominaisarvot ovat reaalisia, niin se on hermiittinen.*

Todistus (vrt. [1, s. 415]). Koska matriisi A on normaali, on se unitaarisesti diagonalisoituva. Olkoon P sellainen unitaarinen matriisi, että $P^*AP = D$, missä D on diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin A ominaisarvot. Koska oletuksen mukaan matriisin A ominaisarvot ovat reaalisia, niin $D^* = D$. Tällöin $A = PDP^*$ ja

$$A^* = (PDP^*)^* = PD^*P^* = PDP^* = A,$$

joten matriisi A on hermiittinen. □

6 Spektraalihajotelma

Tarkastellaan nyt matriisin A unitaarista diagonalisointia eri tavalla. Olkoon P unitaarinen matriisi siten, että $P^*AP = D$, missä D on diagonaalimatriisi diagonaalialkioinaan matriisin A ominaisarvot ja matriisin P sarakevektorit ovat vastaavat ominaisvektorit $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$. Koska matriisi P on unitaarinen, niin $P^*P = PP^* = I$. Yhtälöä $P^*P = I$ käytettiin lauseen 5.3 todistuksessa osoittamaan, että matriisin P sarakevektorit muodostavat vektoriavaruuden $\mathbb{C}^n$ ortonormaalin kannan. Siis

$$I = P^*P = \begin{pmatrix} \boldsymbol{x}_1^* \\ \boldsymbol{x}_2^* \\ \vdots \\ \boldsymbol{x}_n^* \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1^*\boldsymbol{x}_1 & \boldsymbol{x}_1^*\boldsymbol{x}_2 & \ldots & \boldsymbol{x}_1^*\boldsymbol{x}_n \\ \boldsymbol{x}_2^*\boldsymbol{x}_1 & \boldsymbol{x}_2^*\boldsymbol{x}_2 & \ldots & \boldsymbol{x}_2^*\boldsymbol{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x}_n^*\boldsymbol{x}_1 & \boldsymbol{x}_n^*\boldsymbol{x}_2 & \ldots & \boldsymbol{x}_n^*\boldsymbol{x}_n \end{pmatrix},$$

missä alkio $\boldsymbol{x}_i^*\boldsymbol{x}_j = \langle \boldsymbol{x}_j, \boldsymbol{x}_i \rangle$ on kompleksiluku (tässä tapauksessa 1 tai 0). Mutta mitä tietoa voidaan johtaa toisesta matriisitulosta $PP^* = I$ sarakevektorien suhteen. Matriisien kertolaskulla saadaan

$$I = PP^* = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1^* \\ \boldsymbol{x}_2^* \\ \vdots \\ \boldsymbol{x}_n^* \end{pmatrix} = \boldsymbol{x}_1\boldsymbol{x}_1^* + \boldsymbol{x}_2\boldsymbol{x}_2^* + \cdots + \boldsymbol{x}_n\boldsymbol{x}_n^*,$$

missä $\boldsymbol{x}_i\boldsymbol{x}_i^* = E_i$ on $n \times n$ -matriisi. Se on $n \times 1$ -sarakevektorin $\boldsymbol{x}_i$ ja $1 \times n$ -rivivektorin $\boldsymbol{x}_i^*$ matriisitulo. Edellinen yhtälö voidaan siis kirjoittaa muodossa

$$I = E_1 + E_2 + \cdots + E_n.$$

Tämä pätee kaikilla unitaarisilla matriiseilla P. Kiinnostavinta on erityisesti tapaus, jossa matriisin P sarakevektorit ovat matriisin A ominaisvektorit.

Lause 6.1 (Spektraalihajotelma)**.** *Olkoot A normaali matriisi ja $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ matriisin A ominaisvektoreista muodostuva ortonormaali joukko vastaavilla ominaisarvoilla $\lambda_1, \lambda_2, \ldots, \lambda_n$. Tällöin*

$$A = \lambda_1\boldsymbol{x}_1\boldsymbol{x}_1^* + \lambda_2\boldsymbol{x}_2\boldsymbol{x}_2^* + \cdots + \lambda_n\boldsymbol{x}_n\boldsymbol{x}_n^*.$$

Todistus (vrt. [1, s. 415]). Olkoon P unitaarinen matriisi, jonka sarakevektorit ovat normaalin matriisin A ominaisvektorit. Tällöin

$$I = \boldsymbol{x}_1\boldsymbol{x}_1^* + \boldsymbol{x}_2\boldsymbol{x}_2^* + \cdots + \boldsymbol{x}_n\boldsymbol{x}_n^*.$$

Kerrotaan yhtälön molemmat puolet vasemmalta matriisilla A. Nyt

$$\begin{aligned} A = AI &= A(\boldsymbol{x}_1\boldsymbol{x}_1^* + \boldsymbol{x}_2\boldsymbol{x}_2^* + \cdots + \boldsymbol{x}_n\boldsymbol{x}_n^*) \\ &= A\boldsymbol{x}_1\boldsymbol{x}_1^* + A\boldsymbol{x}_2\boldsymbol{x}_2^* + \cdots + A\boldsymbol{x}_n\boldsymbol{x}_n^* \\ &= \lambda_1\boldsymbol{x}_1\boldsymbol{x}_1^* + \lambda_2\boldsymbol{x}_2\boldsymbol{x}_2^* + \cdots + \lambda_n\boldsymbol{x}_n\boldsymbol{x}_n^*. \end{aligned}$$

□

Normaalin matriisin A spektraalihajotelma on:

$$A = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_n E_2,$$

missä $E_i = \boldsymbol{x}_i\boldsymbol{x}_i^*$.

Esimerkki 6.1. Olkoon

$$A = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}.$$

Sen ominaisarvot ovat $\lambda_1 = 1$ ja $\lambda_2 = 4$ ja vastaavat ominaisvektorit esimerkiksi

$$\boldsymbol{x}_1 = \begin{pmatrix} \frac{-1-i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \qquad \text{ja} \qquad \boldsymbol{x}_2 = \begin{pmatrix} \frac{1+i}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} \end{pmatrix}.$$

Tällöin

$$E_1 = \boldsymbol{x}_1\boldsymbol{x}_1^* = \begin{pmatrix} \frac{-1-i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \frac{-1+i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{-1-i}{3} \\ \frac{-1+i}{3} & \frac{1}{3} \end{pmatrix}$$

ja

$$E_2 = \boldsymbol{x}_2\boldsymbol{x}_2^* = \begin{pmatrix} \frac{1+i}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} \end{pmatrix} \begin{pmatrix} \frac{1-i}{\sqrt{6}} & \sqrt{\frac{2}{3}} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1+i}{3} \\ \frac{1-i}{3} & \frac{2}{3} \end{pmatrix}.$$

Täten matriisin A spektraalihajotelma on

$$\begin{aligned} A = 1E_1 + 4E_2 &= 1\begin{pmatrix} \frac{2}{3} & \frac{-1-i}{3} \\ \frac{-1+i}{3} & \frac{1}{3} \end{pmatrix} + 4\begin{pmatrix} \frac{1}{3} & \frac{1+i}{3} \\ \frac{1-i}{3} & \frac{2}{3} \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{3} & \frac{-1-i}{3} \\ \frac{-1+i}{3} & \frac{1}{3} \end{pmatrix} + \begin{pmatrix} \frac{4}{3} & \frac{4+4i}{3} \\ \frac{4-4i}{3} & \frac{8}{3} \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{3} + \frac{4}{3} & \frac{-1-i}{3} + \frac{4+4i}{3} \\ \frac{-1+i}{3} + \frac{4-4i}{3} & \frac{1}{3} + \frac{8}{3} \end{pmatrix} \\ &= \begin{pmatrix} \frac{6}{3} & \frac{3+3i}{3} \\ \frac{3-3i}{3} & \frac{9}{3} \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}. \end{aligned}$$

Tarkastellaan lähemmin matriiseja E_i. Kuten seuraava lause osoittaa, ne ovat hermiittisiä, idempotentteja ja ne toteuttavat yhtälön $E_iE_j = \mathbf{0}$, kun $i \neq j$, missä $\mathbf{0}$ on nollamatriisi.

Lause 6.2. *Jos* $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ *on vektoriavaruuden* $\mathbb{C}^n$ *ortonormaali kanta, niin matriisilla* $E_i = \boldsymbol{x}_i\boldsymbol{x}_i^*$ *on seuraavat ominaisuudet:*

(i) $E_i^* = E_i$*,*

(ii) $E_i^2 = E_i$*,*

(iii) $E_iE_j = 0$*, kun* $i \neq j$*.*

Todistus (vrt. [1, s. 418]). Todistetaan kohta (i). Matriisi E_i on *hermiittinen*, koska

$$E_i^* = (\boldsymbol{x}_i\boldsymbol{x}_i^*)^* = (\boldsymbol{x}_i^*)^*\boldsymbol{x}_i^* = \boldsymbol{x}_i\boldsymbol{x}_i^* = E_i.$$

Todistetaan kohta (ii). Matriisi E_i on *idempotentti*, koska

$$\begin{aligned} E_i^2 = E_iE_i &= (\boldsymbol{x}_i\boldsymbol{x}_i^*)(\boldsymbol{x}_i\boldsymbol{x}_i^*) = \boldsymbol{x}_i(\boldsymbol{x}_i^*\boldsymbol{x}_i)\boldsymbol{x}_i^* \\ &= \boldsymbol{x}_i\langle \boldsymbol{x}_i, \boldsymbol{x}_i\rangle \boldsymbol{x}_i^* = \boldsymbol{x}_i \cdot 1 \cdot \boldsymbol{x}_i^* \\ &= E_i. \end{aligned}$$

Todistetaan kohta (iii). Kun $i \neq j$, niin

$$\begin{aligned} E_iE_j &= (\boldsymbol{x}_i\boldsymbol{x}_i^*)(\boldsymbol{x}_j\boldsymbol{x}_j^*) = \boldsymbol{x}_i(\boldsymbol{x}_i^*\boldsymbol{x}_j)\boldsymbol{x}_j^* \\ &= \boldsymbol{x}_i\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle \boldsymbol{x}_j^* = \boldsymbol{x}_i \cdot 0 \cdot \boldsymbol{x}_j^* \\ &= \mathbf{0}. \end{aligned}$$

□

Koska jokainen matriisi E_i on idempotentti, se indusoi projektion. Kun ortonormaalin kannan kantavektori kerrotaan matriisilla E_i, saadaan

$$E_i\boldsymbol{x}_i = (\boldsymbol{x}_i\boldsymbol{x}_i^*)\boldsymbol{x}_i = \boldsymbol{x}_i(\boldsymbol{x}_i^*\boldsymbol{x}_i) = \boldsymbol{x}_i\langle \boldsymbol{x}_i, \boldsymbol{x}_i\rangle = \boldsymbol{x}_i \cdot 1 = \boldsymbol{x}_i$$

ja

$$E_i\boldsymbol{x}_j = (\boldsymbol{x}_i\boldsymbol{x}_i^*)\boldsymbol{x}_j = \boldsymbol{x}_i(\boldsymbol{x}_i^*\boldsymbol{x}_j) = \boldsymbol{x}_i\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle = \boldsymbol{x}_i \cdot 0 = \mathbf{0}.$$

Olkoon $\boldsymbol{v} \in \mathbb{C}^n$. Vektori $\boldsymbol{v}$ voidaan nyt kirjoittaa yksikäsitteisenä lineaarikombinaationa muodossa $\boldsymbol{v} = a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_n\boldsymbol{x}_n$. Tällöin

$$\begin{aligned} E_i\boldsymbol{v} &= E_i(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_n\boldsymbol{x}_n) \\ &= a_1E_i\boldsymbol{x}_1 + \cdots + a_{i-1}E_i\boldsymbol{x}_{i-1} + a_iE_i\boldsymbol{x}_i + a_{i+1}E_i\boldsymbol{x}_{i+1} + \cdots + a_n\boldsymbol{x}_n \\ &= a_i\boldsymbol{x}_i. \end{aligned}$$

Siis E_i on ortogonaaliprojektio vektoriavaruudelta $\mathbb{C}^n$ vektorin $\boldsymbol{x}_i$ virittämälle vektorialiavaruudelle.

Lähteet

[1] Anthony, Martin and Harvey, Michele. *Linear Algebra: Concepts and Methods*. Cambridge: Cambridge University Press, 2012.

[2] Anton, Howard and Chris, Rorres. *Elementary Linear Algebra: Applications Version*. Eleventh edition. New York: Wiley, 2015.

[3] Horn, Roger and Johnson, Charles. *Matrix Analysis*. Second edition. Cambridge: Cambridge University Press, 1994.

[4] Prasolov, Viktor. *Problems and Theorems in Linear Algebra*. American Mathematical Society, 1994.