Tampereen yliopisto

Anni Sipilä

KOMPLEKSIMATRIISIEN SPEKTRAALIHAJOTELMA

Informaatioteknologian ja viestinnän tiedekunta
Kandidaattitutkielma
Toukokuu 2020

Tiivistelmä

Anni Sipilä: Kompleksimatriisien spektraalihajotelma
Kandidaattitutkielma
Tampereen yliopisto
Matematiikan ja tilastotieteen tutkinto-ohjelma
Toukokuu 2020

Tämän tutkielman tarkoituksena on esitellä kompleksimatriisien spektraalihajotelma ja sen olemassaoloon tarvittavia keskeisiä ehtoja.

Tutkielman kannalta tärkeää teoriaa ovat kompleksimatriisien ominaisarvot, -vektorit ja -avaruudet, sekä kompleksilukujen ja -matriisien kompleksikonjugaatit. Välttämätöntä on myös ymmärtää kompleksisen sisätuloavaruuden ominaisuudet ja muun muassa vektoreiden ortogonaalisuus ja ortonormaalius.

Eräs työn keskeisiä tuloksia on, että kompleksimatriisi on unitaarisesti diagonalisoituva, jos ja vain jos matriisi on normaali matriisi tai jos kompleksiselle vektoriavaruudelle on olemassa kyseisen matriisin ominaisvektoreista koostuva ortonormaali kanta. Keskeistä on täten esitellä hermiittiset, unitaariset ja diagonaalimatriisit ja todeta, että ne ovat normaaleja matriiseja. Läpi käydään myös hermiittisen konjugaatin määritelmä ja ominaisuuksia. Lopuksi todetaan, että normaalille matriisille on mahdollista muodostaa spektraalihajotelma. Käydään myös läpi spektraalihajotelman $A = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_n E_n$ matriisin E_i ominaisuuksia ja todetaan, että sen idempotettisuus tarkoittaa, että se muodostaa projektion.

Tutkielma on kirjoitettu Anthonyn ja Harveyn kirjaa Linear Algebra Concepts and Methods, Hornin ja Johnsonin kirjaa Matrix Analysis ja Haukkasen luentomonistetta Lineaarialgebra 1B mukaillen.

Avainsanat: Spektraalihajotelma, kompleksinen, normaali matriisi, sisätuloavaruus, unitaarinen diagonalisoituvuus

Sisältö

1 Johdanto

Tässä tutkielmassa tarkastellaan kompleksimatriisien spektraalihajotelmaa ja todistetaan siihen liittyviä keskeisiä ehtoja. Aluksi käydään läpi kompleksilukujen ja erityisesti kompleksimatriisien ominaisuuksia: niiden kompleksikonjugaatit, ominaisarvot, -vektorit ja -avaruudet.

Luvussa 3 määritellään ja todistetaan standardi kompleksinen sisätulo ja todetaan, että se todella on sisätulo avaruudessa $\mathbb{C}^n$ sisätulon yleisen määritelmän mukaan. Todetaan myös, että vektoreiden ortogonaalisuuden ja ortonormaaliuden määritelmät säilyvät sellaisenaan siirryttäessä reaalisista sisätuloavaruuksista kompleksisiin. Tämän jälkeen käydään normaalit matriisit läpi esimerkein ja esitellään niiden ominaisuuksia. Erityisesti huomataan, että hermiittinen matriisi on kompleksinen vastine reaaliselle symmetriselle matriisille ja että sen ominaisarvot ovat reaalisia. Täten saadaan myös todistettua, että hermiittisen matriisin eri ominaisarvoja vastaavat ominaisvektorit ovat ortogonaalisia. Samassa luvussa todetaan vielä, että unitaarinen matriisi on kompleksinen vastine reaaliselle ortogonaaliselle matriisille ja käydään läpi sen sarakkeiden ja rivienominaisuuksia kompleksisessa vektoriavaruudessa. Luvun lopuksi todetaan, että sekä hermiittiset, unitaariset että diagonaalimatriisit ovat normaalimatriiseja.

Luvun 4 alussa tarkastellaan ensin yleisemmin matriisien diagonalisoituvuutta, josta sitten siirrytään matriisien unitaariseen diagonalisoituvuuteen. Todistetaan matriisin olevan unitaarisesti diagonalisoituvissa, jos kompleksiselle vektoriavaaruudelle on olemassa kyseisen matriisin ominaisvektoreista koostuva ortonormaali kanta tai jos matriisi on normaali matriisi. Jälkimmäinen todistetaan Schurin lauseen avulla. Lopuksi annetaan esimerkkejä ja todetaan vielä reaalilukuihin rinnastaen, että normaalimatriisi, jonka kaikki ominaisarvot ovat reaalisia, on hermiittinen matriisi.

Luvussa 5 todistetaan kompleksimatriisien spektraalihajotelman esittelevä lause ja annetaan hajotelmasta esimerkki. Huomataan vielä, että hajotelman matriisilla E_i on olemassa ominaisuuksia kuten hermiittisyys ja idempotenttisuus. Matriisin E_i idempotenttisuuden vuoksi huomataan myös, että se on ortogonaali projektio kompleksiselta vektoriavaruudelta vektorin $\mathbf{x}_i$ virittämälle aliavaruudelle.

Lukijalta oletetaan perustietämys kompleksiluvuista ja algebran sekä lineaarialgebran alkeet. Erityisesti lukijan tulisi hallita vektoriavaruudet, aliavaruudet, matriisien algebraa ja ominaisuuksia, sekä vektoriavaruuden kanta.

2 Kompleksimatriisien ominaisuuksista

Tutkielmassa käsitellään kompleksimatriiseja eli matriiseja, joiden alkiot kuuluvat kompleksilukuihin. Ominaisarvojen ja ominaisvektoreiden ominaisuudet ovat samat kompleksimatriiseille kuin reaalisillekin ja ne myös löydetään samoin keinoin. Tämän luvun alalukujen 2.1 ja 2.2 määritelmät ovat mukailtu Anthonyn ja Harveyn teoksen $[1, s.390, 399 - 401]$ mukaisesti ja alalukujen 2.3 ja 2.4 määritelmät ovat mukailtu Pentti Haukkasen luentomonisteesta $[2, s.51 - 57]$.

2.1 Kompleksiluvuista

Määritelmä 2.1. *Kompleksiluvuksi* kutsutaan lukua z, joka esitetään muodossa $z = a + ib$, kun $a, b \in \mathbb{R}$ ja $i^2 = -1$. Tällaisten lukujen joukkoa merkitään

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

Määritelmä 2.2. Kompleksiluvun $z = a + ib$ *kompleksikonjugaatti* saadaan, kun vaihdetaan sen imaginääriosan etumerkki. Luvun z kompleksikonjugaatiksi kutsutaan siis lukua $\overline{z} = a - ib$.

2.2 Kompleksimatriiseista

Määritelmä 2.3. Olkoon $A \in \mathbb{C}^{m \times n}$. Matriisin A *kompleksikonjugaatiksi* $\overline{A} \in \mathbb{C}^{n \times n}$ kutsutaan matriisia, jonka (i, j)-alkio on matriisin A (i, j)-alkion kompleksikonjugaatti. Kun siis $A = (a_{ij})$, niin $\overline{A} = (\overline{a_{ij}})$.

Lause 2.4. *Olkoon $A \in \mathbb{R}^{n \times n}$. Jos $\lambda \in \mathbb{C}$ on sen kompleksinen ominaisarvo ja sitä vastaava ominaisvektori on $\mathbf{v}$, niin tällöin myös $\overline{\lambda}$ on matriisin A ominaisarvo, ja sitä vastaava ominaisvektori on $\overline{\mathbf{v}}$.*

Todistus. Koska A on reaalinen matriisi, niin sen karakteristinen yhtälö on n. asteen polynomi reaalisilla kertoimilla. Täten kaikki kompleksiset juuret ilmenevät konjugaattipareina. Tällöin siis, jos $\lambda \in \mathbb{C}$ on matriisin A ominaisarvo, samoin on myös $\overline{\lambda}$. Jos ominaisarvoa λ vastaava ominaisvektori on $\mathbf{v}$, niin tällöin $A\mathbf{v} = \lambda\mathbf{v}$. Kun yhtälöstä otetaan puolittain kompleksikonjugaatit, saadaan $\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}}$, ja koska A on reaalinen, niin edelleen saadaan $A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$. Yhtälöstä nähdään nyt, että $\overline{v}$ on ominaisarvoa $\overline{\lambda}$ vastaava ominaisvektori. $\square$

2.3 Ominaisarvot ja ominaisvektorit

Määritelmä 2.5. Olkoon A $n \times n$*-matriisi*. Tällöin luvun $\lambda \in \mathbb{C}$ sanotaan olevan matriisin A *ominaisarvo*, jos on olemassa $\mathbf{x} \neq \mathbf{0}$ siten, että

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

Tällöin vektorin $\mathbf{x}$ sanotaan olevan ominaisarvoa λ vastaava *ominaisvektori*.

Määritelmä 2.6. Matriisin A ominaisarvojen joukkoa

$$\sigma(A) = \{ \lambda_1, \lambda_2, \ldots, \lambda_n \}$$

kutsutaan matriisin A *spektriksi*.

Määritelmä 2.7. Olkoon $A \in \mathbb{C}^{n\times n}$. Tällöin

$$\det(\lambda I - A)$$

kutsutaan matriisin A *karakteristiseksi polynomiksi*.

Lause 2.8. *Olkoon $A \in \mathbb{C}^{n\times n}$. Tällöin λ on matriisin A ominaisarvo, jos ja vain jos se toteuttaa yhtälön*

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

Tätä yhtälöä kutsutaan matriisin A karakteristiseksi yhtälöksi.

2.4 Ominaisavaruudet

Määritelmä 2.9. Olkoon $A \in \mathbb{C}^{n\times n}$ ja olkoon λ matriisin A ominaisarvo. Tällöin ominaisarvon λ *ominaisavaruus* on aliavaruus

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

Esimerkki 2.10. Tarkastellaan matriisia

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

Tällöin

$$\begin{aligned} \det(\lambda I - A) = 0 \iff & \begin{vmatrix} \lambda - 5 & -2 \\ -4 & \lambda + 1 \end{vmatrix} = 0 \\ \iff & (\lambda - 5)(\lambda + 1) - 8 = 0 \\ \iff & \lambda - 3\lambda + 2 = 0, \end{aligned}$$

josta matriisin A ominaisarvoiksi saadaan $\lambda = 2 - \sqrt{17}$ tai $\lambda = 2 + \sqrt{17}$.

Esimerkki 2.11. Tarkastellaan jälleen matriisia

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

Ratkaistaan nyt sen ominaisarvoja $2 - \sqrt{17}$ ja $2 + \sqrt{17}$ vastaavat ominaisvektorit etsimällä epätriviaali ratkaisu yhtälölle $(\lambda I - A)\mathbf{x} = \mathbf{0}$. Ominaisarvoa $2 - \sqrt{17}$ vastaava ominaisvektori saadaan, kun ratkaistaan yhtälö $((2 - \sqrt{17})I - A)\mathbf{x} = \mathbf{0}$. Tästä saadaan yhtälöryhmä

$$\begin{cases} (-3 - \sqrt{17})x - 2y = 0 \\ -4x + (3 - \sqrt{17})y = 0. \end{cases}$$

Kun merkitään $y = t$, $t \in \mathbb{R}$, saadaan yhtälöryhmästä ratkaistua, että ominaisarvoa $\lambda = 2 - \sqrt{17}$ vastaavat ominaisvektorit ovat muotoa

$$\mathbf{x} = t \begin{bmatrix} (3 - \sqrt{17})/4 \\ 1 \end{bmatrix}.$$

Vastaavaasti saadaan ratkaistua, että ominaisarvoa $\lambda = 2 + \sqrt{17}$ vastaavat ominaisvektorit ovat muotoa

$$\mathbf{x} = t \begin{bmatrix} (3 + \sqrt{17})/4 \\ 1 \end{bmatrix}.$$

Nyt ominaisarvoa $\lambda = 2 - \sqrt{17}$ vastaavan ominaisavaruuden kannaksi saadaan

$$\left\{ \begin{bmatrix} (3 - \sqrt{17})/4 \\ 1 \end{bmatrix} \right\}$$

ja ominaisarvoa $\lambda = 2 + \sqrt{17}$ vastaavan ominaisavaruuden kannaksi saadaan

$$\left\{ \begin{bmatrix} (3 + \sqrt{17})/4 \\ 1 \end{bmatrix} \right\}.$$

3 Kompleksisesta vektoriavaruudesta ja normaaleista matriiseista

3.1 Kompleksinen vektoriavaruus ja kompleksinen sisätuloavaruus

Kompleksinen vektoriavaruus on kuin reaalinen vektoriavaruus, mutta sen alkiot ovat kompleksilukuja, jolloin vektoriavaruuden aksioomat eroavat siten, että skalaarit kuuluvat kompleksilukuihin. Tämä luku on tehty mukaillen Anthonyn ja Harveyn $[1, s.401 - 415]$ esitystapaa.

Määritelmä 3.1. Olkoot $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Tällöin *standardi kompleksinen sisätulo* on

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

Lause 3.2. *Standardi kompleksinen sisätulo*

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

toteuttaa seuraavat ominaisuudet kaikille $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ *ja kaikille* $\alpha, \beta \in \mathbb{C}$*:*

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

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

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

Todistus. Todistetaan ensin kohta (i). Nyt

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

Todistetaan sitten kohta (ii). Nyt

$$\begin{aligned}
\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle &= (\alpha x_1 + \beta y_1)\overline{z_1} + (\alpha x_2 + \beta y_2)\overline{z_2} + \cdots + (\alpha x_n + \beta y_n)\overline{z_n} \\
&= \alpha(x_1\overline{z_1} + x_2\overline{z_2} + \cdots + x_n\overline{z_n}) + \beta(y_1\overline{z_1} + y_2\overline{z_2} + \cdots + y_n\overline{z_n}) \\
&= \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle .
\end{aligned}$$

Kohdassa (*iii*) kun $x_j = a_j + ib_j$ on vektorin $\mathbf{x}$ j:nnes komponentti saadaan, että

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

on reaalilukujen neliöiden summa, joten $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$. Tällöin siis $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, jos ja vain jos jokainen termi a_j^2 ja b_j^2 on nolla eli kun $\mathbf{x}$ on nollavektori $\mathbf{x} = \mathbf{0}$. □

Määritelmä 3.3. Olkoon V kompleksinen vektoriavaruus. *Yleiseksi kompleksiseksi sisätuloksi* vektoriavaruudessa V kutsutaan kuvausta vektoriparilta $\mathbf{x}, \mathbf{y}$ kompleksiluvuille, josta käytetään merkintää $\langle \mathbf{x}, \mathbf{y} \rangle$ ja joka toteuttaa seuraavat ominaisuudet:

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

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

(iii) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ on reaaliluku kaikilla $\mathbf{x} \in V$ ja $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, jos ja vain jos $\mathbf{x} = \mathbf{0}$ on vektoriavaruuden V nollavektori.

Huomautus. Vektoriavaruutta V kompleksisella sisätulolla varustettuna kutsutaan *kompleksiseksi sisätuloavaruudeksi*. Nyt vektorin $\mathbf{x} \in V$ pituus määritellään seuraavasti:

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

Täten nähdään, että aiemmin määritelty ja myöhemmin tutkielmassa käytettävä sisätulo avaruudessa $\mathbb{C}^n$ on sisätulo sisätulon yleisen määritelmän mukaan.
Määritelmästä seuraa suoraan vielä kaksi ominaisuutta:

(a) $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$ kaikille $\mathbf{x}, \mathbf{y} \in V$ ja kaikille $\alpha \in \mathbb{C}$.

(b) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ kaikille $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$.

3.2 Ortogonaalisista vektoreista

Ortogonaalisten vektoreiden määritelmä säilyy sellaisenaan siirryttäessä reaalisista sisätuloavaruuksista kompleksisiin:

Määritelmä 3.4. Vektoreiden $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ sanotaan olevan *ortogonaaliset* kompleksisessa sisätuloavaruudessa, jos

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

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

Huomautus. Vektorijoukon $\{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \}$ sanotaan olevan *ortogonaalinen* kompleksisessa sisätuloavaruudessa V, jos jokainen siihen kuuluva vektoripari on keskenään ortogonaalinen eli $\left\langle \mathbf{v}_i, \mathbf{v}_j \right\rangle = 0$, kun $i \neq j$.

Huomautus. Vektorijoukkoa kutsutaan *ortonormaaliksi*, jos lisäksi jokainen siihen kuuluva vektori on yksikkövektori $||\mathbf{v}_i|| = 1$ eli jos $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$.

Lause 3.5. *Olkoon V kompleksinen sisätuloavaruus ja olkoot vektorit $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ pareittain ortogonaalisia eli $\left\langle \mathbf{v}_i, \mathbf{v}_j \right\rangle = 0$, kun $i \neq j$. Oletetaan myös, että yksikään vektoreista ei ole nollavektori. Tällöin vektorijoukko $\{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \}$ on lineaarisesti riippumaton joukko.*

Todistus. Olkoon $\alpha_i \in \mathbb{C}$. Osoitetaan, että jos vektorijoukon $\{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \}$ lineaarikombinaatio on nollavektori eli

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0},$$

niin skalaarien on oltava lukuja 0 eli on oltava $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$.
Olkoon nyt $i = 1, 2, \ldots, k$. Tällöin

$$\langle \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k, \mathbf{v}_i \rangle = \langle \mathbf{0}, \mathbf{v}_i \rangle = 0$$

eli

$$\begin{aligned}
&\langle \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k, \mathbf{v}_i \rangle \\
&= \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_i \rangle + \cdots + \alpha_{i-1} \langle \mathbf{v}_{1-1}, \mathbf{v}_i \rangle + \alpha_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle + \alpha_{i+1} \langle \mathbf{v}_{i+1}, \mathbf{v}_i \rangle + \cdots + \alpha_k \langle \mathbf{v}_k, \mathbf{v}_k \rangle .
\end{aligned}$$

Nyt, koska vektorit ovat pareittain ortogonaalisia, niin $\left\langle \mathbf{v}_i, \mathbf{v}_j \right\rangle = 0$, kun $i \neq j$. Täten siis myös $\alpha_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle = 0$ eli $\alpha_i ||\mathbf{v}_i||^2 = 0$. Ja koska vektorijoukon vektoreista yksikään ei ole nollavektori eli $\mathbf{v_i} \neq \mathbf{0}$, niin on myös oltava $||\mathbf{v}_i||^2 \neq 0$, mistä seuraa, että $\alpha_i = 0$. Täten, koska $i = 1, 2, \ldots, k$, niin saadaan, että

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

□

3.3 Hermiittinen konjugaatti ja hermiittinen matriisi

Kompleksimatriisin hermiittinen konjugaatti saadaan, kun matriisi transponoidaan ja sen jokainen alkio muutetaan omaksi kompleksikonjugaatikseen. Koska matriisin transponointi ja matriisin kompleksikonjugaatin muodostaminen kommutoivat keskenään, ei hermiittistä konjugaattia muodostaessa ole merkitystä, missä järjestyksessä operaatiot suoritetaan. Hermiittinen matriisi taas on kompleksinen vastine reaaliselle symmetriselle matriisille.

Määritelmä 3.6. Olkoon matriisi $A \in \mathbb{C}^{m\times n}$. Tällöin matriisin A *hermiittiseksi konjugaatiksi* kutsutaan matriisia

$$A^* = \overline{A}^T.$$

Tällöin, jos $A = (a_{ij})$, niin $\overline{A} = (\overline{a_{ij}})$ ja $A^* = \overline{A^T} = (\overline{a_{ji}})$.

Huomautus. Olkoot $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Tällöin kompleksista sisätuloa voidaan merkitä matriisien tulon avulla seuraavasti:

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

Ja edelleen

$$\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^*\mathbf{x} = ||\mathbf{x}||^2.$$

Huomautus. Matriisin hermiittinen konjugaatti toteuttaa seuraavat ominaisuudet:

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

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

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

Esimerkki 3.7. Matriisin

$$A = \begin{bmatrix} 2+8i & 11 & i \\ 7-i & 8+4i & 16-6i \end{bmatrix}$$

hermiittinen konjugaatti on

$$A^* = \begin{bmatrix} 2-8i & 7+i \\ 11 & 8-4i \\ -i & 16+6i \end{bmatrix}.$$

Määritelmä 3.8. Matriisin $A \in \mathbb{C}^{n\times n}$ sanotaan olevan *hermiittinen*, jos

$$A = A^*.$$

Huomautus. Hermiittinen matriisi reaalisilla alkoilla on symmetrinen matriisi, sillä tällöin $A = A^* = A^T$.

Esimerkki 3.9. Hermiittisen matriisin diagonaalialkioiden tulee olla reaalilukuja, ja diagonaalin molemmin puolin toisiaan vastaavien alkioiden tulee olla toistensa kompleksikonjugaatteja. Matriisi

$$A = \begin{bmatrix} 11 & 20 & 20-i \\ 20 & 26 & 9+8i \\ 20+i & 9-8i & 3 \end{bmatrix}$$

on hermiittinen matriisi.

Lause 3.10. *Jos matriisi A on hermiittinen matriisi, niin sen ominaisarvot ovat reaalisia.*

Todistus. Oletetaan, että λ on matriisin A ominaisarvo ja että $\mathbf{v}$ on sitä vastaava ominaisvektori.Tällöin $A\mathbf{v} = \lambda\mathbf{v}$ ja $\mathbf{v} \neq \mathbf{0}$. Kerrotaan tämä yhtälö vasemmalta vektorin $\mathbf{v}$ hermiittisellä konjugaatilla, jolloin saadaan

$$\mathbf{v}^* A\mathbf{v} = \mathbf{v}^* \lambda\mathbf{v} = \lambda\mathbf{v}^*\mathbf{v} = \lambda||\mathbf{v}||^2,$$

missä vektorin $\mathbf{v}$ normi on positiivinen reaaliluku. Toisaalta, kun otetaan yhtälön $A\mathbf{v} = \lambda\mathbf{v}$ molemmista puolista niiden kompleksikonjugaattien transpoosi, saadaan

$$(A\mathbf{v})^* = (\lambda\mathbf{v})^*,$$

josta saadaan

$$\mathbf{v}^* A^* = \overline{\lambda}\mathbf{v}^*.$$

Kun viimeisin yhtälö kerrotaan puolittain oikealta vektorilla $\mathbf{v}$ saadaan

$$\mathbf{v}^* A^*\mathbf{v} = \overline{\lambda}\mathbf{v}^*\mathbf{v} = \overline{\lambda}||\mathbf{v}||^2.$$

Nyt, koska A on hermiittinen, jolloin $\mathbf{v}^* A\mathbf{v} = \mathbf{v}^* A^*\mathbf{v}$, niin siitä seuraa, että

$$\lambda||\mathbf{v}||^2 = \overline{\lambda}||\mathbf{v}||^2.$$

Koska $||\mathbf{v}||^2 \neq 0$, saadaan, että $\lambda = \overline{\lambda}$. Täten λ on reaalinen. □

Lause 3.11. *Jos matriisi $A \in \mathbb{C}^{n\times n}$ on hermiittinen matriisi, niin sen eri ominaisarvoja vastaavat ominaisvektorit ovat ortogonaalisia.*

Todistus. Olkoot λ ja μ matriisin A ominaisarvoja siten, että $\lambda \neq \mu$. Tällöin

$$A\mathbf{x} = \lambda\mathbf{x}, \quad A\mathbf{y} = \mu\mathbf{y},$$

missä $\mathbf{x}$, $\mathbf{y}$ ovat ominaisarvoja λ, μ vastaavat ominaisvektorit. Tällöin

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

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

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

Nyt, koska hermiittisen matriisin ominaisarvot ovat reaalisia, niin $\overline{\mu} = \mu$. Tällöin kahdesta aiemmasta yhtälöstä saadaan yhtäsuuruus $\lambda \langle \mathbf{x}, \mathbf{y} \rangle = \mu \langle \mathbf{x}, \mathbf{y} \rangle$ eli

$$(\lambda - \mu) \langle \mathbf{x}, \mathbf{y} \rangle = 0.$$

Koska $\lambda \neq \mu$, niin saadaan, että $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Täten ominaisvektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat ortogonaalisia. □

3.4 Unitaariset matriisit ja normaalit matriisit

Unitaarinen matriisi on kompleksinen vastine reaaliselle ortogonaaliselle matriisille. Todetaan, että sekä hermiittiset, unitaariset että diagonaalimatriisit ovat normaaleja matriiseja, ja että normaalit matriisit ovat unitaarisesti diagonalisoitavissa.

Määritelmä 3.12. Kompleksimatriisin $P \in \mathbb{C}^{n \times n}$ sanotaan olevan *unitaarinen*, jos

$$PP^* = P^* P = I$$

eli kun matriisilla P on käänteismatriisi P^*.

Huomautus. Tästä seuraa suoraan, että jos matriisi P on unitaarinen matriisi, niin myös P^* on unitaarinen matriisi.

Esimerkki 3.13. Matriisi

$$A = \begin{bmatrix} \frac{1+i}{2} & \frac{1-i}{2} \\ \frac{1-i}{2} & \frac{1+i}{2} \end{bmatrix}$$

on unitaarinen matriisi.

Lause 3.14. *Matriisi $P \in \mathbb{C}^{n \times n}$ on unitaarinen, jos ja vain jos sen sarakkeet ovat kompleksisen vektoriavaruuden $\mathbb{C}^n$ ortonormaali kanta.*

Todistus. Merkitään matriisin P sarakkeita $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Täten siis matriisin P^* rivit ovat näiden sarakkeiden kompleksikonjugaattien transpooseja.

Oletetaan aluksi, että matriisi P on unitaarinen. Tällöin koska määritelmän mukaan $I = P^*P$, niin saadaan

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^* \\ \mathbf{x}_2^* \\ \vdots \\ \mathbf{x}_n^* \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^*\mathbf{x}_1 & \mathbf{x}_1^*\mathbf{x}_2 & \cdots & \mathbf{x}_1^*\mathbf{x}_n \\ \mathbf{x}_2^*\mathbf{x}_1 & \mathbf{x}_2^*\mathbf{x}_2 & \cdots & \mathbf{x}_2^*\mathbf{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_n^*\mathbf{x}_1 & \mathbf{x}_n^*\mathbf{x}_2 & \cdots & \mathbf{x}_n^*\mathbf{x}_n \end{bmatrix}.$$

Nyt, koska $\mathbf{x}_i^*\mathbf{x}_j = \left\langle \mathbf{x}_j, \mathbf{x}_i \right\rangle = 0$, kun $i \neq j$ ja $\mathbf{x}_i^*\mathbf{x}_i = \langle \mathbf{x}_i, \mathbf{x}_i \rangle = 1$, kun $i = j$, niin sarakkeiden muodostama vektorijoukko $\{ \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \}$ on ortonormaalinen. Täten ne ovat lineaarisesti riippumattomia, ja koska niitä on n kappaletta, niin ne muodostavat kompleksisen vektoriavaruuden $\mathbb{C}^n$ kannan.
Oletetaan sitten, että matriisin P sarakkeet ovat kompleksisen vektoriavaruuden $\mathbb{C}^n$ kanta. Tällöin matriisitulon P^*P ratkaisun tulee olla identiteettimatriisi $P^*P = I$. Siis $P^* = P^{-1}$, jolloin myös $PP^* = I$. □

Huomautus. Koska nyt myös matriisi P^* on unitaarinen, niin Lause 3.14. pätee vastaavasti myös matriisin rivien ollessa kompleksisen vektoriavaruuden $\mathbb{C}^n$ kanta.

Lause 3.15. *Matriisi $P \in \mathbb{C}^{n\times n}$ on unitaarinen, jos ja vain jos sen määrittelemä lineaarikuvaus säilyttää kompleksisen sisätulon eli kun*

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

kaikille $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$.

Todistus. Oletetaan ensin, että $P \in \mathbb{C}^{n\times n}$ unitaarinen matriisi. Tällöin

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

joten matriisin P määrittelemä lineaarikuvaus säilyttää kompleksisen sisätulon. Oletetaan sitten matriisin P säilyttävän sisätulo, eli

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

kaikille $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Merkitään vektoriavaruuden $\mathbb{C}^n$ luonnollista kantaa $\{ \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \}$. Tällöin $P\mathbf{e}_i = \mathbf{v}_i$, missä $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ ovat matriisin P sarakkeita. Täten

$$\left\langle \mathbf{v}_i, \mathbf{v}_j \right\rangle = \left\langle P\mathbf{e}_i, P\mathbf{e}_j \right\rangle = \left\langle \mathbf{e}_i, \mathbf{e}_j \right\rangle,$$

mistä voidaan päätellä, että matriisin P sarakkeet ovat vektoriavaruuden $\mathbb{C}^n$ ortonormaali kanta, ja siis P on unitaarinen matriisi. □

Määritelmä 3.16. Matriisin $A \in \mathbb{C}^{n \times n}$ sanotaan olevan *normaali matriisi*, jos

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

Huomautus. Nyt siis jokainen hermiittinen matriisi on normaali matriisi, sillä $AA^* = AA = A^*A$.

Huomautus. Myös jokainen unitaarinen matriisi on normaali matriisi, sillä $AA^* = I = A^*A$.

Huomautus. Myös jokainen diagonaalimatriisi on normaali matriisi. Merkitään $D = diag(d_1, \ldots, d_n)$. Merkitään matriisin D kompleksikonjugaattia $D^* = diag(\overline{d_1}, \ldots, \overline{d_n})$. Tällöin, koska

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

niin D on normaali matriisi, ja lisäksi matriisin DD^* diagonaalialkiot ovat reaalisia.

4 Unitaarinen diagonalisoituvuus

Tämä luku on tehty mukaillen Anthonyn ja Harveyn [1, s.232, 256, 412 − 415] esitystapaa, poislukien lause 4.5, ks. [3, s.79], ja lause 4.6, ks. [3, s.101].

4.1 Matriisin diagonalisoituvuus

Määritelmä 4.1. Olkoot $A, B \in \mathbb{C}^{n\times n}$. Matriisien A ja B sanotaan olevan *similaariset*, jos on olemassa ei-singulaarinen matriisi $S \in \mathbb{C}^{n\times n}$ siten, että

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

Määritelmä 4.2. Olkoon $A \in \mathbb{C}^{n\times n}$. Matriisin A sanotaan olevan *diagonalisoituva*, jos se on similaarinen diagonaalimatriisin kanssa eli on olemassa diagonaalimatriisi D ja ei-singulaarinen matriisi P siten, että $P^{-1}AP = D$.

4.2 Unitaarinen diagonalisoituvuus

Määritelmä 4.3. Matriisin $A \in \mathbb{C}^{n\times n}$ sanotaan olevan *unitaarisesti diagonalisoituva*, jos on olemassa sellainen unitaarinen matriisi $P \in \mathbb{C}^{n\times n}$, että $P^*AP = D$, missä D on diagonaalimatriisi.

Lause 4.4. *Matriisi $A \in \mathbb{C}^{n\times n}$ voidaan unitaarisesti diagonalisoida, jos ja vain jos vektoriavaruudelle $\mathbb{C}^n$ on olemassa ortonormaali kanta, joka koostuu matriisin A ominaisvektoreista.*

Todistus. Oletetaan, että matriisi A on unitaarisesti diagonalisoituva eli $P^*AP = D$. Nyt, koska matriisi P diagonalisoi matriisin A, niin matriisin P sarakkeet ovat vektoriavaruuden $\mathbb{C}^n$ kanta, joka koostuu matriisin A ominaisvektoreista. Koska P on unitaarinen, niin matriisin P sarakkeet ovat vektoriavaruuden $\mathbb{C}^n$ ortonormaali kanta, koostuen matriisin A ominaisvektoreista.

Kääntäen, jos matriisin A ominaisvektorit ovat vektoriavaruuden $\mathbb{C}^n$ ortonormaali kanta, niin tällöin se matriisi P, jonka sarakkeet ovat tämän kannan virittävät vektorit, on unitaarinen. Koska nämä vektorit ovat matriisin A ominaisvektoreita, niin saadaan, että $AP = PD$, missä D on diagonaalimatriisi ominaisvektoreita vastaavilla ominaisarvoilla. Koska $P^{-1} = P^*$, niin saadaan $P^*AP = D$, jolloin A on unitaarisesti diagonalisoitu. □

Lause 4.5. *Schurin lause. Olkoot $\lambda_1, \ldots, \lambda_n$ matriisin $A \in \mathbb{C}^{n\times n}$ ominaisarvot. Tällöin on olemassa unitaarinen matriisi $M \in \mathbb{C}^{n\times n}$ siten, että*

$$U^* AU = T = [t_{ij}],$$

missä $T \in \mathbb{C}^{n\times n}$ on yläkolmiomatriisi, jonka diagonaalialkiot $t_{ii} = \lambda_i$, kun $i = 1, \ldots, n$. Tällöin siis jokainen matriisi A on unitaarisesti ekvivalentti sellaisen kolmiomatriisin kanssa, jonka diagonaalialkiot ovat matriisin A ominaisarvot määrätyssä järjestyksessä. Lisäksi, jos $A \in \mathbb{C}^{n\times n}$ ja kaikki matriisin A ominaisarvot ovat reaalisia, matriisiksi U voidaan valita myös reaalinen, ortogonaalinen matriisi.

Todistus. Ks. [3, s.79] □

Lause 4.6. *Matriisi $A \in \mathbb{C}^{n\times n}$ on unitaarisesti diagonalisoituva, jos ja vain jos A on normaali matriisi.*

Todistus. Todistetaan ensin, että vain normaalit matriisit voidaan unitaarisesti diagonalisoida. Oletetaan, että A on unitaarisesti diagonalisoituva. Tällöin on olemassa unitaarinen matriisi P ja diagonaalinen matriisi D siten, että $P^* AP = D$. Tästä saadaan, että $A = PDP^*$. Tällöin

$$AA^* = (PDP^*)(PDP^*)^* = (PDP^*)(PD^* P^*) = PD(P^* P)D^* P^* = P(DD^*)P^*.$$

Samoin saadaan

$$A^* A = (PDP^*)^*(PDP^*) = (PD^* P^*)(PDP^*) = PD^*(P^* P)DP^* = P(D^* D)P^*.$$

Koska D on diagonaalinen, niin se on myös normaali, joten $P(DD^*)P^* = P(D^* D)P^*$. Täten voidaan myös päätellä, että A on normaali.

Todistetaan sitten, että jos A on normaali matriisi, niin tällöin se on myös unitaarisesti diagonalisoituva. Oletetaan nyt, että $T \in \mathbb{C}^{n\times n}$ on yläkolmiomatriisi, joka on Schurin lauseen 4.5 mukaisesti unitaarisesti ekvivalentti matriisin A kanssa eli $T = U^* AU$, missä $U \in \mathbb{C}^{n\times n}$. Nyt, koska matriisi T on unitaarisesti ekvivalentti matriisin A kanssa, niin matriisin A normaalius on ekvivalenttia matriisin T normaaliuden kanssa.

Nyt siis jos matriisi A on normaali, niin myös matriisi T on normaali. Kuitenkin, jotta kolmiomatriisi voi olla normaali, täytyy sen olla diagonaalimatriisi. Matriisin T diagonaalisuus voidaan nähdä siitä, että matriisitulojen $T^* T$ ja TT^* diagonaalialkiot ovat samat.

Matriisitulon $T^* T$ $(1, 1)$-alkio on sama, kuin matriisitulon TT^* $(1, 1)$-alkio ja siis

$$\bar{t}_{11} t_{11} = t_{11} \bar{t}_{11} + \sum_{j=2}^{n} t_{1j} \bar{t}_{1j} = |t_{11}|^2 + \sum_{j=2}^{n} |t_{1j}|^2.$$

Tästä huomataan, että $\sum_{j=2}^{n} |t_{1j}|^2 = 0$ on epänegatiivisten termien summa, jolloin termeistä jokaisen tulee olla 0. Täten saadaan, että

$$t_{1j} = 0, \quad j = 2, \ldots, n.$$

Matriisitulojen T^*T ja TT^* $(2,2)$-alkiot ovat samat ja siis

$$\bar{t}_{22} t_{22} = t_{22}\bar{t}_{22} + \sum_{j=3}^{n} t_{2j}\bar{t}_{2j} = |t_{22}|^2 + \sum_{j=3}^{n} |t_{2j}|^2.$$

Tällöin samasta syystä, kuin aiemmin saadaan, että

$$t_{2j} = 0, \quad j = 3, \ldots, n.$$

Samaan tapaan, kun oletetaan tunnetuksi

$$t_{ij} = 0, \ j > i \text{ ja } i = 1, \ldots, k-1$$

voidaan todeta, että

$$t_{ij} = 0, \ j > i \text{ kun } i = k.$$

Kun vastaava osoitetaan peräkkäin jokaiselle diagonaalialkiolle, voidaan lopuksi todeta, että

$$t_{ij} = 0, \ j > i \text{ kun } i = 1, \ldots, n.$$

Nyt, koska T on yläkolmiomatriisi, niin

$$t_{ij} = 0, \ j < i \text{ kun } i = 1, \ldots, n$$

ja täten voidaan todeta, että T on diagonaalimatriisi ja täten unitaarisesti diagonalisoituva. □

Esimerkki 4.7. Matriisi unitaarisesti diagonalisoidaan ratkaisemalla aluksi sen karakteristinen yhtälö, jonka avulla ratkaistaan sen ominaisarvot. Jokaiselle ominaisarvolle löydetään ortonormaali kanta sitä vastaavalle ominaisavaruudelle. Täten kaikkien sellaisten ominaisvektoreiden joukko on sen vektoriavaruuden $\mathbb{C}^n$ ortonormaali kanta. Sitten muodostetaan matriisi P asettamalla nämä ominaisvektorit sarakkeiksi. Täten P on unitaarinen ja $P^*AP = D$, missä D on diagonaalimatriisi vastaavilla ominaisarvoilla.

Esimerkki 4.8. Matriisi

$$A = \begin{bmatrix} 4 & 4+i \\ 4-i & 20 \end{bmatrix}$$

on hermiittinen matriisi ja siten normaali matriisi, jolloin se on unitaarisesti diagonalisoitavissa. Ratkaisemalla matriisin A karakteristinen yhtälö, sen ominaisarvoiksi saadaan $\lambda_1 = 21$ ja $\lambda_2 = 3$. Näitä ominaisarvoja vastaaviksi ominaisvektoreiksi saadaan ratkaistua

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -i+4 \end{bmatrix} \text{ ja } \mathbf{x}_2 = \begin{bmatrix} 1 \\ \frac{1}{17}i - \frac{4}{17} \end{bmatrix}.$$

Nyt, koska vektorin $\mathbf{x}_1$ pituus on $||\mathbf{x}_1|| = \sqrt{18}$ ja vektorin $\mathbf{x}_2$ pituus on $||\mathbf{x}_2|| = \sqrt{\frac{18}{17}}$, niin ominaisvektorit saadaan normalisoitua muotoon

$$\mathbf{x}_1 = \begin{bmatrix} \frac{1}{\sqrt{18}} \\ \frac{-i+4}{\sqrt{18}} \end{bmatrix} \text{ ja } \mathbf{x}_2 = \begin{bmatrix} \frac{1}{\sqrt{\frac{18}{17}}} \\ \frac{\frac{1}{17}i - \frac{4}{17}}{\sqrt{\frac{18}{17}}} \end{bmatrix}.$$

Tällöin asettamalla

$$P = \begin{bmatrix} \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{\frac{18}{17}}} \\ \frac{-i+4}{\sqrt{18}} & \frac{\frac{1}{17}i - \frac{4}{17}}{\sqrt{\frac{18}{17}}} \end{bmatrix} \text{ ja } D = \begin{bmatrix} 21 & 0 \\ 0 & 3 \end{bmatrix}$$

saadaan, että $P^*AP = P^{-1}AP = D$.

Lause 4.9. *Olkoon A normaali matriisi. Jos matriisin A kaikki ominaisarvot ovat reaalisia, matriisi A on hermiittinen matriisi.*

Todistus. Koska matriisi A on normaali matriisi, se voidaan unitaarisesti diagonalisoida. Olkoon P unitaarinen matriisi siten, että $P^*AD = D$, missä D on diagonaalimatriisi matriisin A ominaisarvoista. Nyt, koska matriisilla A on reaaliset ominaisarvot, niin $D^* = D$. Tällöin $A = PDP^*$ ja

$$A^* = (PDP^*)^* = PDP^* = A$$

eli matriisi A on hermiittinen matriisi. □

5 Spektraalihajotelma

Tämä on tehty mukaillen Anthonyn ja Harveyn [1, s.415 – 418] esitystapaa.

Lause 5.1. *Olkoon* $A \in \mathbb{C}^{n\times n}$ *normaali matriisi ja olkoot* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *ortonormaali joukko matriisin* A *ominaisvektoreita ja niitä vastaavat ominaisarvot* $\lambda_1, \lambda_2, \ldots, \lambda_n$*. Tällöin*

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

Matriisin A *spektraalihajotelma noudattaa siis kaavaa*

$$A = \lambda_1E_1 + \lambda_2E_2 + \cdots + \lambda_nE_n,$$

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

Todistus. Koska matriisi $P \in \mathbb{C}^{n\times n}$ on unitaarinen matriisi, jonka sarakkeet ovat matriisin $A \in \mathbb{C}^{n\times n}$ ominaisvektorit, niin

$$I = PP^* = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n\end{bmatrix}\begin{bmatrix}\mathbf{x}_1^* \\ \mathbf{x}_2^* \\ \cdots \\ \mathbf{x}_n^*\end{bmatrix} = \mathbf{x}_1\mathbf{x}_1^* + \mathbf{x}_2\mathbf{x}_2^* + \cdots + \mathbf{x}_n\mathbf{x}_n^*.$$

Nyt kertomalla yhtälö puolittain vasemmalta matriisilla A, saadaan

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

□

Esimerkki 5.2. Esimerkissä 4.7 unitaarisesti diagonalisoitiin matriisi

$$A = \begin{bmatrix}4 & 4+i \\ 4-i & 20\end{bmatrix}$$

ja löydettiin matriisit

$$P = \begin{bmatrix}\frac{1}{\sqrt{18}} & \frac{1}{\sqrt{\frac{18}{17}}} \\ \frac{-i+4}{\sqrt{18}} & \frac{\frac{1}{17}i - \frac{4}{17}}{\sqrt{\frac{18}{17}}}\end{bmatrix} \text{ ja } D = \begin{bmatrix}21 & 0 \\ 0 & 3\end{bmatrix},$$

joille pätee $P^*AP = P^{-1}AP = D$. Tällöin matriisin A spektraalihajotelma on siis $A = 21E_1 + 3E_2$. Nyt matriisiksi E_1 saadaan

$$E_1 = \mathbf{x}_1\mathbf{x}_1^* = \begin{bmatrix} \frac{1}{\sqrt{18}} \\ \frac{-i+4}{\sqrt{18}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{18}} & \frac{-i+4}{\sqrt{18}} \end{bmatrix} = \begin{bmatrix} \frac{1}{18} & \frac{1}{18}i + \frac{2}{9} \\ \frac{1}{18}i + \frac{2}{9} & \frac{4}{9}i + \frac{5}{6} \end{bmatrix}$$

ja matriisiksi E_2 saadaan

$$E_2 = \mathbf{x}_2\mathbf{x}_2^* = \begin{bmatrix} \frac{1}{\sqrt{\frac{18}{17}}} \\ \frac{\frac{1}{17}i - \frac{4}{17}}{\sqrt{\frac{18}{17}}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{\frac{18}{17}}} & \frac{\frac{1}{17}i - \frac{4}{17}}{\sqrt{\frac{18}{17}}} \end{bmatrix} = \begin{bmatrix} \frac{17}{18} & \frac{1}{18}i - \frac{2}{9} \\ \frac{1}{18}i - \frac{2}{9} & -\frac{4}{153}i + \frac{5}{102} \end{bmatrix}.$$

Täten matriisin A spektraalihajotelma on

$$A = \begin{bmatrix} \frac{1}{18} & \frac{1}{18}i + \frac{2}{9} \\ \frac{1}{18}i + \frac{2}{9} & \frac{4}{9}i + \frac{5}{6} \end{bmatrix} + \begin{bmatrix} \frac{17}{18} & \frac{1}{18}i - \frac{2}{9} \\ \frac{1}{18}i - \frac{2}{9} & -\frac{4}{153}i + \frac{5}{102} \end{bmatrix}.$$

Lause 5.3. *Olkoon* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *vektoriavaruuden* C^n *ortonormaali kanta. Tällöin matriisilla* $E_i = \mathbf{x}_i\mathbf{x}_i^*$ *on seuraavat ominaisuudet:*

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

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

(iii.) $E_iE_j = \mathbf{0}$, *kun* $i \neq j$, *ja missä* $\mathbf{0}$ *on nollamatriisi.*

Huomautus. Täten kohtien (i) ja (ii) mukaan matriisi E_i on hermiittinen ja idempotentti.

Todistus. Todistetaan kohta (i.). Matriisit ovat nyt hermiittisiä, sillä

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

Todistetaan kohta (ii.). Matriisit ovat nyt idempotenttisia, sillä

$$E_i^2 = E_iE_i = (\mathbf{x}_i\mathbf{x}_i^*)(\mathbf{x}_i\mathbf{x}_i^*) = \mathbf{x}_i(\mathbf{x}_i^*\mathbf{x}_i)\mathbf{x}_i^* = \mathbf{x}_i \langle \mathbf{x}_i, \mathbf{x}_i \rangle \mathbf{x}_i^* = \mathbf{x}_i 1 \mathbf{x}_i = E_i.$$

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

$$E_iE_j = (\mathbf{x}_i\mathbf{x}_i^*)(\mathbf{x}_j\mathbf{x}_j^*) = \mathbf{x}_i(\mathbf{x}_i^*\mathbf{x}_j)\mathbf{x}_j^* = \mathbf{x}_i \left\langle \mathbf{x}_j, \mathbf{x}_i \right\rangle \mathbf{x}_j^* = \mathbf{x}_i \cdot 0 \cdot \mathbf{x}_j^* = 0.$$

□

Huomautus. Se, että jokainen matriisi E_i on idempotentti tarkoittaa sitä, että se kuvaa projektion. Nyt, kun matriisia E_i kerrotaan ortonormaalin kannan vektoreille, saadaan

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

ja

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

Olkoon $\mathbf{v} \in \mathbb{C}^n$. Vektori $\mathbf{v}$ voidaan nyt esittää lineaarikombinaationa

$$\mathbf{v} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n.$$

Tällöin saadaan

$$\begin{aligned} E_i\mathbf{v} =& E_i(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n) \\ =& a_1E_i\mathbf{x}_1 + a_2E_i\mathbf{x}_2 + \cdots + a_{i-1}E_i\mathbf{x}_{i-1} + a_iE_i\mathbf{x}_i + a_{i+1}E_i\mathbf{x}_{i+1} + \cdots + a_nE_i\mathbf{x}_n = a_i\mathbf{x}_i. \end{aligned}$$

Tällöin E_i on siis ortogonaaliprojektio avaruudelta $\mathbb{C}^n$ vektorin $\mathbf{x}_i$ virittämälle aliavaruudelle.

Lähteet

[1] Anthony, M ja Harvey, M. *Linear Algebra Concepts and Methods*. Cambridge University Press, 2012.

[2] Haukkanen, P. *Lineaarialgebra 1B*. Tampereen Yliopisto, 2020.

[3] Horn, R ja Johnson, C. *Matrix Analysis*. Cambridge University Press, 1985.