Tampereen yliopisto

Sanni Miina

HERMIITTISET JA UNITAARISET MATRIISIT

Informaatioteknologian ja viestinnän tiedekunta
Kandidaattitutkielma
Maaliskuu 2021

Tiivistelmä

Sanni Miina: Hermiittiset ja unitaariset matriisit

Kandidaattitutkielma

Tampereen yliopisto

Matematiikan ja tilastotieteen tutkinto-ohjelma

Maaliskuu 2021

Matriisit ovat lineaarialgebran yksi keskeisimmistä aiheista. Tässä tutkielmassa keskitytään tarkastelemaan muutamaa matriisiluokkaa ja matriisien diagonalisointia. Tutkielman pääpaino on kompleksisen avaruuden hermiittisissä ja unitaarisissa matriiseissa. Lisäksi tarkastellaan reaalisen avaruuden symmetrisiä ja ortogonaalisia matriiseja. Tutkielman keskeisessä roolissa on matriisien diagonalisoituvuus ja ensimmäisissä luvuissa esitetään sen ymmärtämiseen vaadittavia matriisien ominaisuuksia.

Tutkielman toisessa luvussa päästään tutustumaan tutkielman aiheeseen, kun siinä esitellään myöhemmissä vaiheissa hyödynnettäviä määritelmiä. Tärkeitä tietoja ymmärtää ovat myöhemmin useasti esiintyvät ominaisarvot ja ominaisvektorit. Luvussa määritellään myös ortogonaalisuus sekä ominaisarvojen etsimisen kannalta tärkeä karakteristinen yhtälö.

Kolmannessa luvussa päästään tutkielman aiheeseen, kun siirrytään käsittelemään ensin reaalisista matriiseista symmetrisiä ja ortogonaalisia, sekä kompleksisista hermiittisiä ja unitaarisia matriiseja. Matriisien määrittelemisen lisäksi tarkastellaan, ovatko ominaisarvot reaalisia sekä ovatko niitä vastaavat ominaisvektorit ortogonaalisia. Reaalisilla symmetrisillä ja kompleksisilla hermiittisillä matriiseilla on monia yhtäläisiä ominaisuuksia. Tällaisia ovat muun muassa ominaisarvojen reaalisuus ja ominaisvektorien ortogonaalisuus. Myös reaalisilla ortogonaalisilla ja kompleksisilla unitaarisilla matriiseilla on yhtäläisyyksiä. Nämä perustuvat siihen, että unitaarinen matriisi on ortogonaalinen, jos sen alkiot ovat reaalisia.

Viimeisessä luvussa siirrytään matriisien diagonalisoituvuuden käsittelemiseen. Ensin määritellään reaalisen matriisin ortogonaalinen diagonalisoituvuus ja toisena kompleksisen matriisin unitaarinen diagonalisoituvuus. Diagonalisoituvuus on yksi hyödyllisimmistä tekniikoista matriiseja sovellettaessa ja siksi tärkeä käsiteltävä aihe.

Siinä hyödynnetään aiemmissa luvuissa esitettyjä määritelmiä ja lauseita, erityisesti ominaisarvot ja ominaisvektorit ovat tärkeä osa diagonalisointia. Luvun lopussa esitetään vielä hermiittisen matriisin diagonalisoinnin neljä keskeistä vaihetta.

Avainsanat: matriisi, hermiittinen, unitaarinen, diagonalisoituvuus

Sisällys

1 Johdanto

Matriisit ovat lineaarialgebran yksi keskeisimmistä aiheista. Tässä tutkielmassa tarkastellaan muutamia reaalisia ja kompleksisia matriiseja sekä niiden diagonalisointia. Tutkielman alussa esitellään loppuosan kannalta tarvittavia tietoja. Tällaisia ovat sisätuloon, ominaisarvoihin ja ortogonaalisuuteen liittyvät määritelmät. Pääasiassa määritelmät esitetään vain kompleksisille avaruuksille $\mathbb{C}^n$.

Kolmannessa luvussa käsitellään ensin reaalisten matriisien symmetrisiä ja ortogonaalisia muotoja. Tämän jälkeen siirrytään kompleksisten hermiittisten ja unitaaristen matriisien käsittelyyn. Hermiittisillä ja unitarisilla matriiseilla on samanlaisia ominaisuuksia kuin symmetrisillä ja ortogonaalisilla matriiseilla. Yksi huomio tästä on se, että sekä symmetristen että hermiittisten matriisien ominaisarvot ovat aina reaalisia ja niiden ominaisvektorit ortogonaalisia.

Viimeisessä luvussa käsitellään matriisin diagonalisoituvuutta. Diagonalisoituvuudessa hyödynnetään aiemmissa luvuissa esiteltyjä tietoja. Pykälässä 4.1 määritellään reaalisen matriisin ortogonaalinen diagonalisoituvuus ja pykälässä 4.2 kompleksisen matriisin unitaarinen diagonalisoituvuus.

Lukijalta edellytetään joidenkin matriisilaskentaan kuuluvien asioiden tuntemista. Muun muassa perustietämys matriisilaskennasta, esimerkiksi transpoosi, kompleksikonjugaatti, laskusäännöt sekä matriisin tavallinen diagonalisointi ovat oleellisia esitietoja. Päälähteenä on käytetty Martin Anthonyn ja Michele Harveyn teosta Linear Algebra Concepts and Methods [1]. Lisäksi apuna on käytetty David C. Mellon kirjaa Invitation to Linear Algebra [2].

2 Esitietoja

Luvussa 2 käsitellään lyhyesti luvuissa 3 ja 4 tarvittavia tietoja matriiseihin ja diagonalisoituvuuteen liittyen (vrt. [1, s. 248,401,404], [2, s. 215,245]).

Määritelmä 2.1 (Kompleksinen sisätulo)**.** *Standardi kompleksinen sisätulo* vektoreille $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ on kompleksinen luku $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, joka määritellään kaavalla

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

Määritelmä 2.2 (Normi)**.** Olkoon V kompleksinen sisätuloavaruus ja vektori $\boldsymbol{x} \in V$. Vektorin $\boldsymbol{x}$ *normi* on

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

missä $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ on standardi kompleksinen sisätulo.

Määritelmä 2.3 (Ominaisarvo ja ominaisvektori)**.** Olkoon A $n \times n$-matriisi. Matriisin A *ominaisarvo* on $\lambda \in \mathbb{C}$, jos on olemassa vektori $\boldsymbol{x} \neq 0$ siten, että

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

Tällöin vektori $\boldsymbol{x}$ on ominaisarvoa λ vastaava *ominaisvektori*.

Esimerkki 2.1. Olkoon 3×3-matriisi

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

Sen ominaisarvot ovat $\lambda_1 = 2$, $\lambda_2 = 2 - \sqrt{2}$ ja $\lambda_3 = 2 + \sqrt{2}$. Ominaisarvoa λ_1 vastaa ominaisvektori on $\boldsymbol{x}_1 = [1, 0, -1]^T$ ja tällöin

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

Huomataan, että ominaisvektori ei voi olla koskaan nollavektori, sillä jos $\boldsymbol{x} = \mathbf{0}$, niin $A(\mathbf{0}) = \lambda(\mathbf{0})$. Tällöin vastaus olisi triviaali ja mikä tahansa ominaisarvo toteuttaisi yhtälön.

Määritelmä 2.4 (Karakteristinen polynomi ja yhtälö)**.** Olkoon $A \in \mathbb{C}^{n \times n}$. Matriisin *A karakteristinen polynomi* on

$$|A - \lambda I|$$

ja yhtälöä

$$|A - \lambda I| = 0$$

sanotaan matriisin *A karakteristiseksi yhtälöksi.*

Määritelmä 2.5 (Ortogonaaliset vektorit)**.** Oletetaan, että V on kompleksinen sisätuloavaruus. Tällöin vektorit $\boldsymbol{x}, \boldsymbol{y} \in V$ ovat *ortogonaaliset*, jos

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

Vektoreiden ortogonaalisuutta merkitään myös $\boldsymbol{x} \perp \boldsymbol{y}$.

Määritelmä 2.6 (Ortonormaali joukko)**.** Kompleksisen sisätuloavaruuden V vektorijoukkoa $\{ \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \}$ sanotaan *ortonormaaliksi joukoksi*, jos

$$\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle = 0, \text{ kun } i \neq j, \quad \text{ja} \quad \langle \boldsymbol{x}_i, \boldsymbol{x}_i \rangle = \|\boldsymbol{x}_i\|^2 = 1.$$

Tällöin vektorit ovat ortogonaalisia ja jokainen vektori on yksikkövektori.

3 Matriisit

Luvussa 3 esitellään kolme matriisiluokkaa ja niiden ominaisuuksia. Ensimmäisenä symmetrinen ja ortogonaalinen matriisi, seuraavaksi hermiittinen matriisi ja lopuksi unitaarinen matriisi (vrt. [1, s. 332,337,319-320,407-412], [2, s. 233,251-254]).

3.1 Symmetrinen ja ortogonaalinen matriisi

Tässä pykälässä käsitellään reaalisista matriiseista symmetristä ja ortogonaalista matriisia.

Määritelmä 3.1 (Symmetrinen matriisi)**.** Matriisi $A \in \mathbb{R}^{n\times n}$ on *symmetrinen*, jos

$$A = A^T,$$

missä A^T on matriisin A transpoosi.

Esimerkki 3.1. Matriisi

$$A = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 0 & 8 \\ 2 & 8 & 3 \end{bmatrix}$$

on symmetrinen, sillä sen diagonaalin suhteen vastakkaiset alkiot ovat yhtäsuuria eli $a_{ij} = a_{ji}$ kaikilla $i \neq j$.

Lause 3.1. *Jos A on symmetrinen matriisi, sen kaikki ominaisarvot ovat reaalisia.*

Todistus. Oletetaan, että λ on matriisin A ominaisarvo ja $\boldsymbol{x}$ sitä vastaava ominaisvektori. Tällöin

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

Tarkastellaan yhtälön eri puolten transpoosia. Koska A on symmetrinen matriisi, niin $A = A^T$ ja tällöin

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

Matriisin A kaikki alkiot ovat reaalisia ja kun otataan kompleksikonjugaatti yhtälön molemmin puolin, saadaan

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

Kerrotaan nyt alkuperäinen yhtälö $A\boldsymbol{x} = \lambda\boldsymbol{x}$ vasemmalta puolittain vektorilla $\overline{\boldsymbol{x}}^T$ ja yhtälö $\overline{\boldsymbol{x}}^T A = \overline{\lambda}\overline{\boldsymbol{x}}^T$ oikealta puolittain vektorilla $\boldsymbol{x}$. Saadaan

$$\begin{aligned}\overline{\boldsymbol{x}}^T A\boldsymbol{x} &= \lambda\overline{\boldsymbol{x}}^T\boldsymbol{x}\\ \overline{\boldsymbol{x}}^T A\boldsymbol{x} &= \overline{\lambda}\overline{\boldsymbol{x}}^T\boldsymbol{x}.\end{aligned}$$

Yhdistämällä yllä olevat yhtälöt saadaan

$$\lambda\overline{\boldsymbol{x}}^T\boldsymbol{x} - \overline{\lambda}\overline{\boldsymbol{x}}^T\boldsymbol{x} = 0$$

ja edelleen

$$(\lambda - \overline{\lambda})\overline{\boldsymbol{x}}^T\boldsymbol{x} = 0.$$

Tutkitaan saatua yhtälöä. Olkoon

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Tällöin

$$\overline{\boldsymbol{x}}^T\boldsymbol{x} = \begin{bmatrix} \overline{x}_1 & \overline{x}_2 & \cdots & \overline{x}_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2.$$

Koska $\boldsymbol{x}$ on ominaisvektori, se ei voi olla nollavektori eli $\boldsymbol{x} \neq \boldsymbol{0}$. Tällöin $\overline{\boldsymbol{x}}^T\boldsymbol{x} \neq 0$. Nyt siis $(\lambda - \overline{\lambda})\overline{\boldsymbol{x}}^T\boldsymbol{x} = 0$ pätee vain, kun $\lambda = \overline{\lambda}$, jolloin $\lambda \in \mathbb{R}$. □

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

Todistus. Olkoot λ ja μ matriisin A erisuuret ominaisarvot ja olkoot vektorit $\boldsymbol{x}$ ja $\boldsymbol{y}$ ominaisarvoja vastaavat ominaisvektorit. Tällöin $A\boldsymbol{x} = \lambda\boldsymbol{x}$ ja $A\boldsymbol{y} = \mu\boldsymbol{y}$. Etsitään kaksi erilaista esitysmuotoa matriisitulolle $\boldsymbol{x}^T A\boldsymbol{y}$.

Tarkastellaan tilannetta ensin yhtälön $A\boldsymbol{y} = \mu\boldsymbol{y}$ avulla. Tällöin saadaan

$$\boldsymbol{x}^T A\boldsymbol{y} = \boldsymbol{x}^T(A\boldsymbol{y}) = \boldsymbol{x}^T(\mu\boldsymbol{y}) = \mu\boldsymbol{x}^T\boldsymbol{y}.$$

Muistetaan, että matriisi A on symmetrinen eli $A = A^T$. Tällöin yhtälöä $A\boldsymbol{x} = \lambda\boldsymbol{x}$ hyödyntämällä saadaan

$$\boldsymbol{x}^T A\boldsymbol{y} = \boldsymbol{x}^T A^T\boldsymbol{y} = (\boldsymbol{x}^T A^T)\boldsymbol{y} = (A\boldsymbol{x})^T\boldsymbol{y} = (\lambda\boldsymbol{x})^T\boldsymbol{y} = \lambda\boldsymbol{x}^T\boldsymbol{y}.$$

Yhdistämällä saadut yhtälöt $\boldsymbol{x}^T A\boldsymbol{y} = \mu\boldsymbol{x}^T\boldsymbol{y}$ ja $\boldsymbol{x}^T A\boldsymbol{y} = \lambda\boldsymbol{x}^T\boldsymbol{y}$ saadaan

$$\mu\boldsymbol{x}^T\boldsymbol{y} = \lambda\boldsymbol{x}^T\boldsymbol{y}$$

ja edelleen

$$(\mu - \lambda)\boldsymbol{x}^T\boldsymbol{y} = 0.$$

Kuitenkin oletuksen mukaan $\mu \neq \lambda$, joten $\mu - \lambda \neq 0$. Tästä seuraa, että $\boldsymbol{x}^T\boldsymbol{y} = \langle\boldsymbol{x},\boldsymbol{y}\rangle = 0$. Tämä tarkoittaa, että vektorit $\boldsymbol{x}$ ja $\boldsymbol{y}$ ovat ortogonaaliset. □

Määritelmä 3.2 (Ortogonaalinen matriisi)**.** Matriisi $P \in \mathbb{R}^{n\times n}$ on *ortogonaalinen*, jos

$$P^T P = I = PP^T.$$

Huomautus. Määritelmästä 3.2 seuraa, että ortogonaalinen martriisi P on ei-singulaarinen ja $P^{-1} = P^T$.

Lause 3.3. *$n \times n$-matriisi P on ortogonaalinen, jos ja vain jos matriisin P sarakevektorit ovat ortogonaaliset ja jokaisen vektorin pituus on 1.*

Todistus. Olkoon matriisi $P = (\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n)$. Silloin P^T on matriisi, jonka rivit ovat $\boldsymbol{x}_1^T,\boldsymbol{x}_2^T,\ldots,\boldsymbol{x}_n^T$. Tällöin $P^T P = I$ voidaan esittää muodossa

$$P^T P = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1^T\boldsymbol{x}_1 & \boldsymbol{x}_1^T\boldsymbol{x}_2 & \cdots & \boldsymbol{x}_1^T\boldsymbol{x}_n \\ \boldsymbol{x}_2^T\boldsymbol{x}_1 & \boldsymbol{x}_2^T\boldsymbol{x}_2 & \cdots & \boldsymbol{x}_2^T\boldsymbol{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x}_n^T\boldsymbol{x}_1 & \boldsymbol{x}_n^T\boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n^T\boldsymbol{x}_n \end{bmatrix} = I.$$

Todistetaan ensin lauseen toinen suunta. Oletetaan, että $P^T P = I$. Yllä olevien matriisien yhtäsuuruudesta saadaan

$$\boldsymbol{x}_i^T\boldsymbol{x}_i = \langle\boldsymbol{x}_i,\boldsymbol{x}_i\rangle = \|\boldsymbol{x}_i\|^2 = 1 \quad \text{ja} \quad \boldsymbol{x}_i^T\boldsymbol{x}_j = \langle\boldsymbol{x}_i,\boldsymbol{x}_j\rangle = 0, \text{jos } i \neq j.$$

Tämän perusteella vektorit ovat yksikkövektoreita ja sarakevektorit $\boldsymbol{x}_i$ ja $\boldsymbol{x}_j$ ovat ortogonaaliset.

Toiseen suuntaan oletetaan, että sarakkeet ovat pareittain ortogonaaliset yksikkövektorit. Tällöin

$$\|\boldsymbol{x}_i\|^2 = \langle\boldsymbol{x}_i,\boldsymbol{x}_i\rangle = \boldsymbol{x}_i^T\boldsymbol{x}_i = 1 \quad \text{ja} \quad \langle\boldsymbol{x}_i,\boldsymbol{x}_j\rangle = \boldsymbol{x}_i^T\boldsymbol{x}_j = 0, \text{kun } i \neq j,$$

joten matriisi P on ortogonaalinen ja $P^T P = I$. □

3.2 Hermiittinen matriisi

Reaalista symmetristä matriisia vastaavaa kompleksista matriisia sanotaan hermiittiseksi matriisiksi. Jos hermiittisen matriisin alkiot ovat reaalisia, niin $A = A^* = A^T$. Hermiittisellä matriisilla onkin vastaavia ominaisuuksia kuin symmetrisellä reaalisella matriisilla. Tässä pykälässä määritellään hermiittinen matriisi ja käsitellään sen ominaisuuksia.

Määritelmä 3.3 (Hermiittinen konjugaatti)**.** Olkoon A $m \times n$-matriisi, missä $(a_{ij}) \in \mathbb{C}$. Tällöin matriisin A *hermiittinen konjugaatti* eli *konjugaattitranspoosi* A^* on

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

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

Huomautus. Matriisin konjugaattitranspoosi toteuttaa seuraavan yhtälön:

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

Määritelmä 3.4 (Hermiittinen matriisi)**.** Matriisi $A \in \mathbb{C}^{n \times n}$ on *hermiittinen*, jos

$$A = A^*.$$

Huomautus. Hermiittisen matriisin määritelmän perusteella tulee seuraavien yhtäsuuruuksien toteutua $A = (a_{ij}) = (\overline{a}_{ji}) = A^*$ ja $a_{ii} = \overline{a}_{ii}$. Siis diagonaalialkioiden tulee olla reaalisia ja vastakkaisten alkioiden tulee olla toistensa kompleksiset konjugaatit.

Esimerkki 3.2. Matriisi

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

on hermiittinen, sillä

$$A^* = \overline{A}^T = \begin{bmatrix} \overline{1} & \overline{1+i} \\ \overline{1-i} & \overline{-1} \end{bmatrix} = \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix} = A.$$

Lause 3.4. *Hermiittisen matriisin A kaikki ominaisarvot ovat reaalisia.*

Todistus. Olkoon λ matriisin A ominaisarvo ja $\boldsymbol{x}$ sitä vastaava ominaisvektori. Tällöin

$$A\boldsymbol{x} = \lambda \boldsymbol{x}. \tag{3.1}$$

Sen kompleksinen konjugaattitranspoosi on

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

ja tiedetään, että $A = A^*$. Näin ollen määritelmän 3.3 jälkeisen huomautuksen perusteella

$$\boldsymbol{x}^*A = \overline{\lambda}\boldsymbol{x}^*. \tag{3.2}$$

Kerrotaan yhtälön (3.1) molemmat puolet vasemmalta vektorilla $\boldsymbol{x}^*$ ja yhtälön (3.2) molemmat puolet oikealta vektorilla $\boldsymbol{x}$. Saadaan

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

Yhdistämällä nämä kaksi yhtälöä saadaan

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

ja edelleen

$$(\lambda - \overline{\lambda})\boldsymbol{x}^*\boldsymbol{x} = (\lambda - \overline{\lambda})\|\boldsymbol{x}\|^2 = 0.$$

Koska $\boldsymbol{x}$ on ominaisvektori, niin $\|\boldsymbol{x}\| \neq 0$. Näin ollen täytyy $\lambda = \overline{\lambda}$ ja tällöin $\lambda \in \mathbb{R}$. □

Lause 3.5. *Jos matriisi A on hermiittinen, sen erisuuria ominaisarvoja vastaavat ominaisvektorit ovat ortogonaalisia euklidisen sisätulon suhteen avaruudessa $\mathbb{C}^n$.*

Todistus. Olkoot λ_1 ja λ_2 matriisin A erisuuria ominaisarvoja. Olkoot $\boldsymbol{x}_1$ ja $\boldsymbol{x}_2$ niitä vastaavat ominaisvektorit, jolloin

$$A\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1 \tag{3.3}$$
$$A\boldsymbol{x}_2 = \lambda_2\boldsymbol{x}_2. \tag{3.4}$$

Lauseen todistamiseksi tulee osoittaa, että

$$\langle\boldsymbol{x}_1, \boldsymbol{x}_2\rangle = \boldsymbol{x}_1^*\boldsymbol{x}_2 = 0,$$

missä $\langle\boldsymbol{x}_1, \boldsymbol{x}_2\rangle$ on euklidinen sisätulo. Kun otetaan yhtälöstä (3.3) puolittain konjugaattitranspoosi, saadaan

$$\boldsymbol{x}_1^*A = \lambda_1\boldsymbol{x}_1^*. \tag{3.5}$$

Kerrotaan yhtälö (3.4) vasemmalta puolelta vektorilla $\boldsymbol{x}_1^*$ ja yhtälö (3.5) oikealta vektorilla $\boldsymbol{x}_2$. Saadaan

$$\boldsymbol{x}_1^* A \boldsymbol{x}_2^* = \lambda_2 \boldsymbol{x}_1^* \boldsymbol{x}_2$$
$$\boldsymbol{x}_1^* A \boldsymbol{x}_2^* = \lambda_1 \boldsymbol{x}_1^* \boldsymbol{x}_2.$$

Lopuksi yhdistämällä kaksi viimeisintä yhtälöä saadaan

$$(\lambda_2 - \lambda_1)\boldsymbol{x}_1^* \boldsymbol{x}_2 = (\lambda_2 - \lambda_1)\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle = 0.$$

Oletuksen mukaan $(\lambda_2 - \lambda_1) \neq 0$, joten täytyy $\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle = 0$. Näin ollen vektorit $\boldsymbol{x}_1$ ja $\boldsymbol{x}_2$ ovat ortogonaaliset. □

Esimerkki 3.3. Osoitetaan, että hermiittisen matriisin

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

ominaisarvot ovat reaalisia ja ominaisvektorit ortogonaaliset.

Ratkaistaan ominaisarvot karakteristisen yhtälön $|A - \lambda I| = 0$ avulla. Saadaan

$$\begin{vmatrix} 1-\lambda & 1-i \\ 1+i & -1-\lambda \end{vmatrix} = (1-\lambda)((-1)-\lambda) - (1-i)(1+i) = \lambda^2 - 3 = 0,$$

jonka ratkaisut ovat $\lambda_1 = -\sqrt{3}$ ja $\lambda_2 = \sqrt{3}$. Ominaisarvot ovat reaaliset.

Ratkaistaan matriisin A molempia ominaisarvoja vastaavat ominaisvektorit lausekkeen $(A - \lambda I)\boldsymbol{x} = \boldsymbol{0}$ avulla.

Ominaisarvolle $\lambda_1 = \sqrt{3}$ saadaan

$$(A - \sqrt{3}I)\boldsymbol{x}_1 = \begin{bmatrix} 1-\sqrt{3} & 1-i \\ 1+i & -1-\sqrt{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ja tästä ominaisvektori

$$\boldsymbol{x}_1 = \begin{bmatrix} \frac{1+\sqrt{3}}{2} + i\frac{-1-\sqrt{3}}{2} \\ 1 \end{bmatrix}.$$

Ominaisarvolle $\lambda_2 = -\sqrt{3}$ saadaan

$$(A - (-\sqrt{3})I)\boldsymbol{x}_2 = \begin{bmatrix} 1+\sqrt{3} & 1-i \\ 1+i & -1+\sqrt{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ja ominaisvektori

$$\boldsymbol{x}_2 = \begin{bmatrix} \frac{1-\sqrt{3}}{2} + i\frac{-1+\sqrt{3}}{2} \\ 1 \end{bmatrix}.$$

Osoitetaan vielä, että ominaisvektorit ovat ortogonaaliset. Nyt

$$\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle = \boldsymbol{x}_1^* \boldsymbol{x}_2 = \begin{bmatrix} \frac{1+\sqrt{3}}{2} - i\frac{1+\sqrt{3}}{2} & 1 \end{bmatrix} \begin{bmatrix} \frac{1-\sqrt{3}}{2} + i\frac{-1+\sqrt{3}}{2} \\ 1 \end{bmatrix} = 0.$$

3.3 Unitaarinen matriisi

Unitaariset matriisit ovat kompleksisia vastineita reaalisille ortogonaalisille matriiseille. Unitaarinen matriisi P, jossa on vain reaalisia alkioita, on ortogonaalinen, kun $P^* = P^T$. Tässä pykälässä määritellään unitaarinen matriisi ja käsitellään sen ominaisuuksia.

Määritelmä 3.5 (Unitaarinen matriisi)**.** Matriisi $P \in \mathbb{C}^{n\times n}$ on *unitaarinen*, jos

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

Huomautus. Määritelmästä 3.5 seuraa, että unitaarinen matriisi P on aina ei-singulaarinen ja $P^{-1} = P^*$.

Esimerkki 3.4. Matriisi

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

on unitaarinen, sillä

$$PP^* = \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1-i & 1+i \\ 1+i & 1-i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = P^*P.$$

Lause 3.6. *$n \times n$-neliömatriisi P on unitaarinen, jos ja vain jos sarakevektorit muodostavat avaruuden $\mathbb{C}^n$ ortonormaalin kannan.*

Todistus. Olkoot $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ matriisin P sarakevektorit. Kompleksisella konjugaattitranspoosilla saadaan rivivektorit matriisille P^*.

Oletetaan ensin, että P on unitaarinen ja $I = P^*P$. Tällöin

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

Yhtälöstä nähdään, että $\boldsymbol{x}_i^*\boldsymbol{x}_j = \langle \boldsymbol{x}_j, \boldsymbol{x}_i \rangle = 0$, jos $i \neq j$, ja $\boldsymbol{x}_i^*\boldsymbol{x}_i = \langle \boldsymbol{x}_i, \boldsymbol{x}_i \rangle = 1$, jos $i = j$. Näin ollen sarakkeet $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n\}$ muodostavat ortonormaalin joukon ja täten ovat lineaarisesti riippumattomia. Joukossa on n vektoria, joten ne muodostavat avaruuden $\mathbb{C}^n$ kannan.

Seuraavaksi oletetaan, että matriisin P sarakevektorit muodostavat avaruuden $\mathbb{C}^n$ ortonormaalin kannan. Tällöin $P^*P = I$ kuten yllä osoitettiin. Näin ollen $P^* = P^{-1}$ ja $PP^* = I$. □

Lause 3.7. *Matriisi P on unitaarinen, jos ja vain jos matriisin P määräämä lineaaritransformaatio säilyttää standardisen kompleksisen sisätulon, toisin sanoen* $\langle P\boldsymbol{x}, P\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$ *kaikilla* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$.

Todistus. Oletetaan, 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 kompleksisen sisätulon.

Oletetaan sitten, että matriisille P pätee $\langle P\boldsymbol{x}, P\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$ kaikilla $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$. Merkitään avaruuden $\mathbb{C}^n$ luonnollista kantaa $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$. Silloin $P\boldsymbol{e}_i = \boldsymbol{x}_i$, missä $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ ovat matriisin P sarakevektorit. Tällöin

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

mistä voidaan päätellä, että matriisin P sarakevektorit ovat avaruuden $\mathbb{C}^n$ ortonormaali kanta ja näin ollen P on unitaarinen matriisi. □

4 Matriisin diagonalisoituvuus

Luvussa 4 käsitellään matriisin diagonalisoituvuutta (vrt. [1, s. 329-339, 412-413], [2, s. 255-257]). Diagonalisointi on yksi hyödyllisimmistä tekniikoista matriiseja sovellettaessa. Siinä hyödynnetään matriisin ominaisarvoja sekä ominaisvektoreita.

4.1 Ortogonaalinen diagonalisoituvuus

Tässä pykälässä määritellään reaalisen matriisin ortogonaalinen diagonalisoituvuus sekä esitetään muutama lause siihen liittyen.

Määritelmä 4.1 (Ortogonaalinen diagonalisoituvuus)**.** Matriisi $A \in \mathbb{R}^{n\times n}$ on *ortogonaalisesti diagonalisoituva*, jos on olemassa sellainen ortogonaalinen matriisi $P \in \mathbb{R}^{n\times n}$, että

$$P^T AP = D,$$

missä D on diagonaalimatriisi.

Lause 4.1. *Matriisi A on ortogonaalisesti diagonalisoituva, jos ja vain jos on olemassa matriisin A ominaisvektoreista koostuva avaruuden $\mathbb{R}^n$ ortonormaali kanta.*

Todistus. Oletetaan, että matriisi A on ortogonaalisesti diagonalisoituva. Matriisin A diagonalisoituvuus tarkoittaa, että matriisin P sarakkeet koostuvat matriisin A ominaisvektoreista ja ovat avaruuden $\mathbb{R}^n$ kanta (vrt. [1, s. 259]). Se, että matriisi A on ortogonaalisesti diagonalisoituva tarkoittaa, että matriisin P sarakkeet koostuvat matriisin A ominaisvektoreiden ortonormaalista joukosta ja ovat avaruuden $\mathbb{R}^n$ ortonormaali kanta (vrt. [1, s. 321]). Näin ollen tämä suunta pätee.

Seuraavaksi oletetaan, että matriisin A ominaisvektorit muodostavat avaruuden $\mathbb{R}^n$ ortonormaalin kannan. Tällöin näistä ominaisvektoreista muodostuva matriisi P on ortogonaalinen (vrt. [1, s. 321]). Matriisin P ollessa ortogonaalinen $P^T = P^{-1}$ ja edelleen $P^T AP = P^{-1}AP = D$, missä D on diagonaalimatriisi. Näin ollen matriisi A on ortogonaalisesti diagonalisoituva. □

Lause 4.2. *Matriisi A on ortogonaalisesti diagonalisoituva, jos ja vain jos A on symmetrinen.*

Todistus. Todistetaan ensin suunta, jos A on ortogonaalisesti diagonalisoituva, se on symmetrinen. Nyt jos A on ortogonaalisesti diagonalisoituva, on olemassa ortogonaalinen matriisi P ja diagonaalimatriisi D siten, että $P^T AP = P^{-1}AP = D$. Ratkaistaan tästä matriisi A:

$$A = PDP^{-1} = PDP^T.$$

Otetaan transpoosi yhtälöstä puolittain. Koska D on diagonaalimatriisi, tiedetään, että $D^T = D$. Tällöin

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A,$$

mistä nähdään, että A on symmetrinen. Siis vain symmetrinen matriisi voi olla ortogonaalisesti diagonalisoituva.

Todistetaan toinen suunta eli kaikille symmetrisille matriiseille A on olemassa ortogonaalinen matriisi P ja diagonaalimatriisi D siten, että $P^T AP = D$.

Lause pätee kaikille 1×1 -matriiseille, sillä ne ovat valmiiksi diagonaalisia. Tällöin voidaan valita $P = I$, jolloin P on ortogonaalinen matriisi.

Tarkastellaan seuraavaksi yleistä tilannetta $n \geq 2$ ja oletetaan, että lause pätee kaikille $(n-1)\times(n-1)$-kokoisille symmetrisille matriiseille. Olkoon A jokin $n \times n$ -kokoinen symmetrinen matriisi. Tiedetään, että matriisilla A on reaaliset ominaisarvot. Olkoon λ_1 jokin matriisin A ominaisarvo ja $\boldsymbol{x}_1$ sitä vastaava ominaisvektori, jolle pätee $\|\boldsymbol{x}_1\| = 1$. Vektorin $\boldsymbol{x}_1$ virittämän vektoriavaruuden $Lin(\boldsymbol{x}_1)$ kanta on $\boldsymbol{x}_1$. Teoksesta Linear Algebra löytyvän lauseen [1, s. 190] perusteella kanta voidaan laajentaa avaruuden $\mathbb{R}^n$ kannaksi muotoon $\{\boldsymbol{x}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n\}$. Muutetaan laajennettu kanta avaruuden $\mathbb{R}^n$ ortonormaaliksi kannaksi $B = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ Gram-Schmidtin prosessilla [1, s. 321].

Olkoon P sellainen matriisi, jonka sarakkeet koostuvat joukon B vektoreista alkaen vektorista $\boldsymbol{x}_1$. Lauseen 3.3 mukaan matriisi P on tällöin ortogonaalinen. Teoksen Linear Algebra lauseen [1, s. 231] mukaan $P^T AP = P^{-1}AP$ on lineaarikuvauksen $T : \boldsymbol{x} \longrightarrow A\boldsymbol{x}$ matriisi kannassa B. Kuitenkin lauseen [1, s. 230] mukaan tiedetään, että matriisin $P^T AP$ ensimmäinen sarake on lineaarikuvauksen $T(x_1)$ koordinaattivektori kannan B suhteen. Nyt $T(\boldsymbol{x}_1) = A\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1$, joten koordinaattivektori on

$$\begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Tästä seuraa, että on olemassa sellaiset luvut $d_1, d_2, \ldots, d_{n-1}$ ja $c_{i,j}$, missä $i, j = 1, \ldots, n-1$, että matriisi $P^T AP$ on muotoa

$$P^T AP = \begin{bmatrix} \lambda_1 & d_1 & \cdots & d_{n-1} \\ 0 & c_{1,1} & \cdots & c_{1,n-1} \\ 0 & c_{2,1} & \cdots & c_{2.n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & c_{n-1,1} & \cdots & c_{n-1,n-1} \end{bmatrix}.$$

Mutta matriisi A on symmetrinen, joten myös matriisi $P^T AP$ on symmetrinen ja tällöin

$$(P^T AP)^T = P^T A^T P = P^T AP.$$

Matriisin $P^T AP$ symmetrisyydellä on kaksi välitöntä seurausta:

- $d_1 = d_2 = \cdots = d_{n-1} = 0$
- $(n-1) \times (n-1)$-matriisi $C = (c_{i,j})$ on symmetrinen.

Nyt voimme kirjoittaa

$$P^T AP = \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & C \end{bmatrix},$$

missä $\mathbf{0}$ on $(n-1)$-nollavektori ja C on $(n-1) \times (n-1)$-kokoinen symmetrinen matriisi.

Alussa oletettiin, että lause pätee kaikille $(n-1) \times (n-1)$-symmetrisille matriiseille, joten se pätee matriisille C. Tällöin on olemassa $(n-1) \times (n-1)$-kokoinen ortogonaalinen matriisi R siten, että $R^T CR = D$, missä D on diagonaalimatriisi. Tarkastellaan $n \times n$-matriisia

$$Q = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R \end{bmatrix}.$$

Nyt Q on ortogonaalinen matriisi, koska R on ortogonaalinen ja tällöin matriisin Q sarakkeet $2, 3, \ldots, n$ ovat keskenään ortogonaaliset ja pituudeltaan 1 ja koska lisäksi ensimmäinen sarake on ortogonaalinen kaikkien muiden sarakkeiden kanssa ja pituudeltaan 1. Olkoon $S = PQ$. Tällöin S on ortogonaalinen, koska P ja Q ovat ortogonaalisia ja näin ollen

$$S^{-1} = (PQ)^{-1} = Q^{-1}P^{-1} = Q^T P^T = (PQ)^T = S^T.$$

Tarkastellaan matriisia $S^T AS$. Saadaan

$$\begin{aligned}
S^T AS &= (PQ)^T A(PQ) \\
&= Q^T P^T APQ \\
&= Q^T (P^T AP)Q \\
&= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R \end{bmatrix}^T \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & C \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R \end{bmatrix} \\
&= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R^T \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & C \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & R^T CR \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & D \end{bmatrix},
\end{aligned}$$

joka on diagonaalimatriisi, koska D on diagonaalinen. Näin on osoitettu todeksi, että on olemassa ortogonaalinen matriisi S siten, että $S^T AS$ on diagonaalinen. □

4.2 Unitaarinen diagonalisoituvuus

Tässä pykälässä määritellään kompleksisen matriisin unitaarinen diagonalisoituvuus, käsitellään siihen liittyviä lauseita ja esitellään hermiittisen matriisin diagonalisoinnin vaiheet.

Määritelmä 4.2 (Unitaarinen diagonalisoituvuus)**.** Matriisi $A \in \mathbb{C}^{n\times n}$ on *unitaarisesti diagonalisoituva*, jos on olemassa unitaarinen matriisi $P \in \mathbb{C}^{n\times n}$ siten, että

$$P^* AP = D,$$

missä D on diagonaalimatriisi.

Lause 4.3. *Matriisi A voi olla unitaarisesti diagonalisoituva, jos ja vain jos on olemassa matriisin A ominaisvektoreista koostuva avaruuden $\mathbb{C}^n$ ortonormaali kanta.*

Todistus. Oletetaan, että matriisi A on unitaarisesti diagonalisoituva siten, että $P^* AP = D$. Matriisin P diagonalisoidessa matriisin A, matriisin P sarakkeet ovat avaruuden $\mathbb{C}^n$ kanta ja koostuvat matriisin A ominaisvektoreista. Matriisin P ollessa unitaarinen, sen vektorit ovat avaruuden $\mathbb{C}^n$ ortonormaali kanta. Toisin sanoen, jos matriisi P

unitaarisesti diagonalisoi matriisin A, niin matriisin P sarakkeet koostuvat matriisin A ominaisvektoreista ja muodostavat avaruuden $\mathbb{C}^n$ ortonormaalin kannan.

Toiseen suuntaan oletetaan, että matriisin A ominaisvektorit ovat avaruuden $\mathbb{C}^n$ ortonormaali kanta. Tällöin matriisi P, jonka sarakkeet muodostuvat näistä kantavektoreista, on unitaarinen. Vektoreiden ollessa matriisin A ominaisvektorit, saadaan $AP = PD$. Tässä D on diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin A ominaisarvot. Koska $P^* = P^{-1}$, niin $P^*AP = D$ ja tällöin matriisi A on unitaarisesti diagonalisoituva. □

Määritelmä 4.3 (Normaali matriisi)**.** Matriisi $A \in \mathbb{C}^{n\times n}$ on *normaali*, jos

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

Lause 4.4. *Matriisi A on unitaarisesti diagonalisoituva, jos ja vain jos matriisi A on normaali matriisi.*

Todistus. Todistetaan lauseesta vain suuntaan, jos matriisi A on unitaarisesti diagonalisoituva, se on normaali. Toisin sanoen vain normaali matriisi voidaan unitaarisesti diagonalisoida. Oletetaan, että matriisi A on unitaarisesti diagonalisoituva. Tällöin on olemassa unitaarinen matriisi P ja diagonaalimatriisi D siten, että $P^*AP = D$. Ratkaisemalla A saadaan $A = PDP^*$. Silloin

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

Samalla tavalla

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

Kun matriisi D on diagonaalimatriisi, se on normaali ja $P(DD^*)P^* = P(D^*D)P^*$. Tästä voidaan päätellä, että matriisi A on normaali. □

Lause 4.5. *Jokainen hermiittinen $n\times n$-matriisi A on unitaarisesti diagonalisoituva.*

Todistus. Jokainen hermiittinen matriisi on normaali, kun $AA^* = AA = A^*A$. Määritelmän 3.4 mukaan matriisi on hermiitinen, jos $A = A^*$. Tällöin $AA^* = AA = A^*A$ pätee kaikille hermiittisille matriiseille ja hermiittinen matriisi on näin ollen aina normaali. Nyt todistuksen loppu mukailee lauseen 4.4 todistusta, jossa osoitetaan, että vain normaalit matriisit ovat unitaarisesti diagonalisoituvia. □

Hermiittisen matriisin diagonalisointi

1. Etsitään matriisin A ominaisarvot ja huomioidaan niiden kertaluvut.

2. Jos λ on yksinkertainen ominaisarvo ja $\boldsymbol{x}$ on sitä vastaava ominaisvektori, normalisoidaan tämä ominaisvektori

$$\boldsymbol{u}_1 = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}.$$

   Toistetaan tämä kohta jokaiselle matriisin A yksinkertaiselle ominaisarvolle.

3. Jos λ on moninkertainen ominaisarvo ja sitä vastaa ominaisvektorit $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\ldots$, $\boldsymbol{x}_k$, sovelletaan Gram-Schmidtin menetelmää ominaisvektorien joukkoon. Saadaan ortonormaali joukko ominaisvektoreita $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\ldots$, $\boldsymbol{u}_k$. Toistetaan tämä kohta jokaiselle moninkertaiselle ominaisarvolle.

4. Muodostetaan matriisi P. Sen sarakkeet ovat kohdissa 2 ja 3 saadut normalisoidut ominaisvektorit. Silloin P on unitaarinen matriisi, joka diagonalisoi matriisin A. Lopulta P^*AP on diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin A ominaisarvot.

Esimerkki 4.1. Diagonalisoidaan hermiittinen matriisi

$$A = \begin{bmatrix} 2 & -i \\ i & 2 \end{bmatrix}.$$

Ensimmäisenä selvitetään matriisin A ominaisarvot. Ratkaistaan ne yhtälöstä

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & -i \\ i & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3 = 0.$$

Ratkaisuksi saadaan yksinkertaiset ominaisarvot $\lambda_1 = 3$ ja $\lambda_2 = 1$. Seuraavaksi selvitetään ominaisarvoja vastaavat ominaisvektorit yhtälön $(A - \lambda I)\boldsymbol{x} = \boldsymbol{0}$ avulla, jolloin saadaan

$$(A - 3I)\boldsymbol{x}_1 = \begin{bmatrix} -1 & -i \\ i & -1 \end{bmatrix} \boldsymbol{x}_1 = \boldsymbol{0}, \quad \boldsymbol{x}_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

$$(A - 1I)\boldsymbol{x}_2 = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \boldsymbol{x}_2 = \boldsymbol{0}, \quad \boldsymbol{x}_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Koska $\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle = 0$, niin vektorit ovat ortogonaaliset ja matriisi A on ortogonaalisesti diagonalisoituva.

Normalisoidaan vielä ominaisvektorit, jolloin saadaan

$$\boldsymbol{u}_1 = \frac{\boldsymbol{x}_1}{\|\boldsymbol{x}_1\|} = \begin{bmatrix} \frac{-i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \boldsymbol{u}_2 = \frac{\boldsymbol{x}_2}{\|\boldsymbol{x}_2\|} = \begin{bmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Lopuksi muodostetaan unitaarinen matriisi P yllä saaduista vektoreista $\boldsymbol{u}_1$ ja $\boldsymbol{u}_2$:

$$P = \begin{bmatrix} \frac{-i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Nyt matriisin P avulla saadaan diagonalisoitua matriisi A:

$$P^T A P = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = D.$$

Lähteet

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

[2] Mello, D. C. *Invitation to Linear Algebra*. Boca Raton, FL: Chapman and Hall/CRC, 2017.