Jussi Harju

# ORTOGONAALINEN DIAGONALISOINTI JA NELIÖMUODOT

Informaatioteknologian ja viestinnän tiedekunta
Kandidaattitutkielma
Huhtikuu 2021

# Tiivistelmä

Jussi Harju: Ortogonaalinen diagonalisointi ja neliömuodot
Kandidaattitutkielma
Tampereen yliopisto
Matematiikan ja tilastotieteen tutkinto-ohjelma
Huhtikuu 2021

---

Tässä tutkielmassa käsitellään matriisien ortogonaalista diagonalisointia, sekä matriiseihin liittyviä neliömuotoja. Tutkielmassa osoitetaan, että matriisin on oltava symmetrinen ollakseen ortogonaalisesti diagonalisoituva. Aihe on rajattu käsittelemään vain reaalimatriiseja, eli kompleksiset matriisit ovat jätetty tutkielman ulkopuolelle.

Luvussa 2 luetellaan aiheen käsittelyssä tarvittavia määritelmiä ja lauseita. Näistä oleellisimmat määritelmät ovat diagonalisoituvan matriisin ja ortogonaalisen matriisin määritelmät. Tämän luvun lauseet ovat merkittäviä myöhempiä todistuksia varten.

Luvussa 3 käsitellään matriisin ortogonaalista diagonalisointia. Aluksi esitellään ortogonaalisesti diagonalisoituvan matriisin määritelmä. Seuraavaksi todistetaan, että matriisi on ortogonaalisesti diagonalisoituva, vain jos sen ominaisvektoreista muodostuu reaalisen vektoriavaruuden ortonormaali kanta. Tämän jälkeen osoitetaan, että symmetrinen matriisi, jonka kaikki ominaisarvot ovat erisuuria, on ortogonaalisesti diagonalisoituva. Ennen tätä esitellään tärkeä apulause, että symmetrisen matriisin kaikki ominaisarvot ovat reaalilukuja. Sen todistus jätetään kuitenkin tämän tutkielman ulkopuolelle. Seuraavaksi todistetaan yleinen tapaus, tämän luvun päälause, eli spektraalilause. Lauseessa osoitetaan, että matriisi on ortogonaalisesti diagonalisoituva, jos se on symmetrinen, ja vain symmetriset matriisit ovat ortogonaalisesti diagonalisoituvia. Ennen spektraalilauseen todistusta esitetään muutama todistukseen tarvittava apulause.

Luvussa 4 tarkastellaan matriisien neliömuotoja. Aluksi esitellään neliömuodon määritelmä ja neliömuodon definiittisyyden määritelmä. Seuraavaksi esitellään lause, jonka mukaan neliömuodon definiittisyyden muoto on määritettävissä matriisin ominaisarvojen positiivisuuden tai negatiivisuuden avulla. Lauseesta todistetaan positiivisesti definiittiin matriisiin liittyvä kohta. Sitten todistetaan matriisin johtavien pääminorien etumerkkien ja matriisin positiivisen ja negatiivisen definiittisyyden yhteys. Tätä ennen esitellään johtavan pääminorin määritelmä. Viimeiseksi todistetaan, että ortogonaalisen matriisin määräämä lineaarikuvaus säilyttää sisätulon tuloksen reaalisessa vektoriavaruudessa. Tutkielma on kirjoitettu mukaillen Anthonyn ja Harveyn kirjaa Linear Algebra: Concepts and Methods sekä Antonin ja Rorresin kirjaa Elementary Linear Algebra.

Avainsanat: symmetriset matriisit, ortogonaalinen diagonalisointi, neliömuodot

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

# Sisällys

# 1 Johdanto

Tämän tutkielman aihe on rajattu siten, että kompleksiset matriisit eivät kuulu siihen, eli tässä tutkielmassa käsitellään vain reaalimatriiseja.

Tämän tutkielman luvussa 3 tarkastellaan matriisien ortogonaalista diagonalisointia. Ensimmäiseksi todistetaan, että matriisi on ortogonaalisesti diagonalisoituva, jos ja vain jos sen ominaisvektorit muodostavat ortonormaalin kannan reaalisessa vektoriavaruudessa. Se todistetaan pykälässä 3.1. Toiseksi todistetaan, että jos matriisin ominaisarvot ovat erisuuret, niin se on ortogonaalisesti diagonalisoituva. Tämä tulos esitetään pykälässä 3.2. Kolmanneksi todistetaan yleinen tapaus, eli milloin matriisi on ortogonaalisesti diagonalisoituva, vaikka sillä olisikin yhtäsuuria ominaisarvoja. Se todistetaan pykälässä 3.3.

Tämän tutkielman luvussa 4 tarkastellaan neliömuotoja ja joitakin niiden ominaisuuksia. Ensimmäiseksi esitellään ja määritellään neliömuoto. Tämä tehdään pykälässä 4.1. Seuraavaksi käsitellään neliömuodon definiittisyytä, ja todistetaan esimerkiksi, että neliömuoto on positiivisesti definiitti, jos ja vain jos siihen liittyvän matriisin kaikki ominaisarvot ovat positiivisia. Se todistetaan pykälässä 4.2. Sitten todistetaan, että ortogonaalisen matriisin määräämä lineaarikuvaus säilyttää sisätulon tuloksen reaalisessa vektoriavaruudessa. Tämä tulos esitetään pykälässä 4.3.

Valmisteluna luvussa 3 käsiteltyä ortogonaalista diagonalisointia varten luvussa 2 esitetään luettelonomaisesti muutamia tarvittavia määritelmiä ja lauseita.

Lukijalta edellytetään joidenkin lineaarialgebran perusasioiden tuntemista. Edellytetään muun muassa, että lukija tuntee matriisien tavalliset laskutoimitukset, ominaisarvojen ja -vektorien määritelmät, sekä lineaarikuvauksen ja kannan määritelmät. Lähdeteoksina käytetään Anthonyn ja Harveyn kirjaa Linear Algebra: Concepts and Methods sekä Antonin ja Rorresin kirjaa Elementary Linear Algebra.

# 2 Esitietoja

## 2.1 Tarvittavien käsitteiden määritelmiä

Luvussa 2 esitetään lyhyesti muutamia aiheen käsittelyssä tarvittavia apuneuvoja. Tässä pykälässä esitetään neljä määritelmää: diagonalisoituvan matriisin, ortogonaalisten vektorien, ortogonaalisen matriisin ja ortonormaalin joukon määritelmät (vrt. [1, s. 256, 317, 319–320]).

**Määritelmä 2.1.** Neliömatriisia $A$ sanotaan diagonalisoituvaksi, jos se on similaarinen jonkin diagonaalimatriisin $D$ kanssa, eli on olemassa kääntyvä matriisi $P$ siten, että

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

**Määritelmä 2.2** (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 2.3** (Ortogonaalinen matriisi)**.** Neliömatriisin $P$ sanotaan olevan ortogonaalinen, jos $P^{\mathrm{T}}P = PP^{\mathrm{T}} = I$, eli matriisin $P$ transpoosi on sen käänteismatriisi.

**Määritelmä 2.4** (Ortonormaali joukko)**.** Sisätuloavaruuden $V$ joukon $S$ sanotaan olevan ortonormaali joukko, jos

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

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

## 2.2 Tarvittavien lauseiden esittely

Tässä pykälässä esitetään similaaristen matriisien ominaisarvoja koskeva lause, matriisin diagonalisoituvuutta koskeva lause, sekä kaksi ortogonaalisia matriiseja koskevaa lausetta.

**Lause 2.1.** *Similaarisilla matriiseilla on samat ominaisarvot.*

*Todistus* (ks. [1, s. 262–263]). □

**Lause 2.2.** *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 2.3.** *Matriisi $P$ on ortogonaalinen, jos ja vain jos sen sarakkeet vektoreina ovat pareittain ortogonaaliset ja jokaisen pituus on* 1.

*Todistus* (ks. [1, s. 319]). □

**Lause 2.4.** *Olkoon $P$ $n \times n$ -matriisi. Tällöin $P$ on ortogonaalinen, jos ja vain jos sen sarakkeet muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan.*

*Todistus* (ks. [2, s. 402]). □

**Lause 2.5.** *Olkoon $A$ $n \times n$ -matriisi. Matriisin $A$ determinantti on yhtäsuuri kuin sen ominaisarvojen tulo.*

*Todistus* (ks. [1, s. 253]). □

**Lause 2.6.** *Olkoon $A$ $n \times n$ -matriisi. Matriisin $A$ jälki,* $\mathrm{tr}(A)$*, on yhtäsuuri kuin sen ominaisarvojen summa.*

*Todistus* (ks. [1, s. 254–256]). □

# 3 Ortogonaalinen diagonalisointi

## 3.1 Ortogonaalinen diagonalisointi

Ensimmäiseksi määritellään ortogonaalisesti diagonalisoituva matriisi (vrt. [1, s. 329]).

**Määritelmä 3.1.** 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 2.2 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 3.1.** *Matriisi $A$ on ortogonaalisesti diagonalisoituva, jos ja vain jos vektoriavaruudelle $\mathbb{R}^n$ on olemassa matriisin $A$ ominaisvektoreista muodostuva ortonormaali kanta.*

*Todistus* (vrt. [2, s. 304, 410]). Oletetaan ensin, että matriisi $A$ on ortogonaalisesti diagonalisoituva, eli $P^{\mathrm{T}}AP = D$. Tällöin $A$ ja $D$ ovat similaarisia matriiseja. Lauseen 2.1 mukaan similaarisilla matriiseilla on samat ominaisarvot, ja koska $D$ on diagonaalimatriisi, niin sen ominaisarvot ovat sen diagonaalialkiot. Merkitään $P = (\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \dots \quad \boldsymbol{v}_n)$. Tällöin

$$
\begin{aligned}
P^{\mathrm{T}}AP = D &\Leftrightarrow AP = PD \\
&\Leftrightarrow A(\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \dots \quad \boldsymbol{v}_n) = (\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \dots \quad \boldsymbol{v}_n)\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}. \\
&\Leftrightarrow (A\boldsymbol{v}_1 \quad A\boldsymbol{v}_2 \quad \dots \quad A\boldsymbol{v}_n) = (\lambda_1\boldsymbol{v}_1 \quad \lambda_2\boldsymbol{v}_2 \quad \dots \quad \lambda_n\boldsymbol{v}_n),
\end{aligned}
$$

mistä seuraa, että $A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1$, $A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2, \dots, A\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n$, joten vektorit $\boldsymbol{v}_1$, $\boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ ovat matriisin $A$ ominaisvektoreita. Koska matriisi $P$ on ortogonaalinen, niin lauseen 2.4 mukaan matriisin $A$ ominaisvektorit $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_n$ muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan.

Oletetaan sitten, että matriisin $A$ ominaisvektorit muodostavat vektoriavaruuden $\mathbb{R}^n$ ortonormaalin kannan. Muodostetaan matriisi $P$, jossa sarakkeina ovat nämä ominaisvektorit, eli $P = (\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \ldots \quad \boldsymbol{v}_n)$. Tällöin $P$ on lauseen 2.4 mukaan ortogonaalinen matriisi. Koska $A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1$, $A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n$, niin

$$(A\boldsymbol{v}_1 \quad A\boldsymbol{v}_2 \quad \ldots \quad A\boldsymbol{v}_n) = (\lambda_1\boldsymbol{v}_1 \quad \lambda_2\boldsymbol{v}_2 \quad \ldots \quad \lambda_n\boldsymbol{v}_n), \tag{3.1}$$

kun $\boldsymbol{v}_i$ on ominaisarvoa $\lambda_i$ vastaava ominaisvektori ja $i = 1, 2, \ldots, n$. Kirjoitetaan yhtälö (3.1) muodossa $AP = PD$, missä $D$ on diagonaalimatriisi, jonka diagonaalialkiot ovat matriisin $A$ ominaisarvot. Koska $P$ on ortogonaalinen matriisi, niin

$$AP = PD \quad \Leftrightarrow \quad P^{\mathrm{T}}AP = D,$$

joten $A$ on ortogonaalisesti diagonalisoituva matriisi. □

## 3.2 Erisuuret ominaisarvot

Aluksi esitellään tärkeä apulause, jonka todistus jätetään tämän tutkielman ulkopuolelle, koska siinä olisi luonnollista käyttää kompleksisia matriiseja (ks. [1, s. 409]).

**Apulause 3.1.** Jos $A$ on symmetrinen matriisi, niin kaikki sen ominaisarvot ovat reaalilukuja.

Seuraavaksi tässä pykälässä osoitetaan, että jos symmetrisellä $n \times n$ -matriisilla on $n$ erisuurta ominaisarvoa, niin se on ortogonaalisesti diagonalisoituva. Siihen tarvitaan seuravaa lausetta.

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

*Todistus* (vrt. [1, s. 332]). Oletetaan, että $\lambda$ ja $\mu$ 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 $A\boldsymbol{y} = \mu\boldsymbol{y}$.

Koska $A\boldsymbol{y} = \mu\boldsymbol{y}$, niin

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

Muistetaan myös, että $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Koska $A$ on symmetrinen, niin $A = A^{\mathrm{T}}$. Sijoittamalla ja käyttämällä matriisin transpoosin ominaisuuksia saadaan

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

Näistä kahdesta muodosta saadaan

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

Koska $\mu$ ja $\lambda$ ovat erisuuria, niin $\mu - \lambda \neq 0$, joten $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = \langle\boldsymbol{x}, \boldsymbol{y}\rangle = 0$. Tällöin $\boldsymbol{x}$ ja $\boldsymbol{y}$ ovat ortogonaaliset. □

**Esimerkki 3.1.** Olkoon $A = \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix}$. Sen ominaisarvot ovat

$$\lambda_1 = 10 \quad \text{ja} \quad \lambda_2 = 0,$$

ja vastaavat ominaisvektorit ovat

$$\boldsymbol{x} = \begin{pmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{pmatrix} \quad \text{ja} \quad \boldsymbol{y} = \begin{pmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}.$$

Näiden ominaisvektorien sisätulo on

$$\langle\boldsymbol{x}, \boldsymbol{y}\rangle = \frac{1}{\sqrt{10}} \cdot \frac{-3}{\sqrt{10}} + \frac{3}{\sqrt{10}} \cdot \frac{1}{\sqrt{10}} = \frac{-3}{10} + \frac{3}{10} = 0,$$

joten ne ovat ortogonaaliset.

Lause 3.2 osoittaa, että jos symmetrisellä $n \times n$ -matriisilla on $n$ erisuurta ominaisarvoa ja jos otetaan $n$ vastaavaa ominaisvektoria, niin mitkä tahansa kaksi näistä ominaisvektoreista ovat keskenään ortogonaalisia. Normeeraamalla ominaisvektorit saadaan niistä yksikkövektorit. Nyt nämä ominaisvektorit muodostavat ortonormaalin joukon, joka on vektoriavaruuden $\mathbb{R}^n$ kanta. Siis lauseen 3.1 mukaan matriisi on ortogonaalisesti diagonalisoituva. Jos $P$ on matriisi sarakkeinaan nämä ominaisvektorit, niin $P^{-1}AP = D$ on diagonaalimatriisi alkioinaan matriisin $A$ ominaisarvot. Koska matriisin $P$ sarakkeet muodostavat ortonormaalin joukon, niin lauseen 2.3 mukaan $P$ on ortogonaalinen matriisi, joten $P^{-1} = P^{\mathrm{T}}$ eli $P^{\mathrm{T}}AP = D$. Toisin sanoen, saadaan seuraava lause.

**Lause 3.3.** *Oletetaan, että matriisi $A$ on symmetrinen ja sillä on $n$ erisuurta ominaisarvoa. Otetaan $n$ vastaavaa normeerattua ominaisvektoria, kukin pituudeltaan* 1. *Muodostetaan matriisi $P$, jossa on sarakkeina nämä normeeratut ominaisvektorit. Tällöin $P^{-1} = P^{\mathrm{T}}$ eli $P$ on ortogonaalinen matriisi ja $P^{\mathrm{T}}AP = D$, diagonaalimatriisi, jonka alkiot ovat matriisin $A$ ominaisarvot.*

*Todistus.* Koska matriisilla $A$ on $n$ erisuurta ominaisarvoa, niin sen kaikki ominaisvektorit $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ ovat lauseen 3.2 mukaan ortogonaalisia. Koska nämä ominaisvektorit ovat normeerattuja yksikkövektoreita, niin $\|\boldsymbol{v}_i\| = 1$, $i = 1, 2, \ldots, n$. Merkitään $P = (\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \ldots \quad \boldsymbol{v}_n)$. Tällöin

$$\begin{aligned}
P^{\mathrm{T}}P &= \begin{pmatrix} \boldsymbol{v}_1^{\mathrm{T}} \\ \boldsymbol{v}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{v}_n^{\mathrm{T}} \end{pmatrix} (\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \ldots \quad \boldsymbol{v}_n) = \begin{pmatrix} \|\boldsymbol{v}_1\| & \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_2 & \ldots & \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_n \\ \boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{v}_1 & \|\boldsymbol{v}_2\| & \ldots & \boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_1 & \boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{v}_2 & \ldots & \|\boldsymbol{v}_n\| \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = I,
\end{aligned}$$

joten $P$ on ortogonaalinen. Nyt

$$\begin{aligned}
AP &= (A\boldsymbol{v}_1 \quad A\boldsymbol{v}_2 \quad \ldots \quad A\boldsymbol{v}_n) = (\lambda_1\boldsymbol{v}_1 \quad \lambda_2\boldsymbol{v}_2 \quad \ldots \quad \lambda_n\boldsymbol{v}_n) \\
&= (\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \ldots \quad \boldsymbol{v}_n) \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.
\end{aligned}$$

Kirjoitetaan yllä oleva matriisitulo muodossa $PD$. Tällöin

$$AP = PD \quad \Leftrightarrow \quad P^{\mathrm{T}}AP = D,$$

missä diagonaalimatriisin $D$ alkiot ovat matriisin $A$ ominaisarvot. □

## 3.3 Yleinen tapaus

Edellisessä pykälässä todistettiin, että symmetrinen matriisi on ortogonaalisesti diagonalisoituva, jos sen ominaisarvot ovat erisuuret. Seuraavaksi todistetaan, että mikä tahansa symmetrinen matriisi on ortogonaalisesti diagonalisoituva.

**Lause 3.4** (Spektraalilause)**.** *Matriisi $A$ on ortogonaalisesti diagonalisoituva, jos ja vain jos $A$ on symmetrinen.*

Lauseen 3.4 todistamiseen tarvitaan muutama apulause, jotka esitellään alla.

**Apulause 3.2.** Olkoon $V$ äärellisulotteinen vektoriavaruus ja $W$ sen aliavaruus. Tällöin $\dim(W) \leq \dim(V)$. Jos $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_r\}$ on aliavaruuden W kanta, niin on olemassa $s = \dim(V) - \dim(W)$ vektoria $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_s \in V$ siten, että

$$\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_r, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_s\}$$

on vektoriavaruuden $V$ kanta.

*Todistus.* Ks. [1, s. 190]. □

**Apulause 3.3.** Oletetaan, että $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ ja $B' = \{\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots, \boldsymbol{v}'_m\}$ ovat vektoriavaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^m$ kantoja. Olkoot $T\colon \mathbb{R}^n \to \mathbb{R}^m$ lineaarikuvaus ja $M = A_{[B,B']}$ $m \times n$ -matriisi, jonka $i$:s sarake on yhtäsuuri kuin $[T(\boldsymbol{v}_i)]_{B'}$, lineaarikuvauksen $T(\boldsymbol{v}_i)$ koordinaattivektori kannan $B'$ suhteen. Tällöin

$$[T(\boldsymbol{x})]_{B'} = M[\boldsymbol{x}]_B \qquad \forall \boldsymbol{x} \in \mathbb{R}^n.$$

*Todistus.* Ks. [1, s. 230]. □

**Seuraus 3.1.** *Oletetaan, että $T\colon \mathbb{R}^n \to \mathbb{R}^n$ on lineaarikuvaus ja $B = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ vektoriavaruuden $\mathbb{R}^n$ kanta. Olkoot $A$ $n \times n$ -matriisi siten, että $T(\boldsymbol{x}) = A\boldsymbol{x}$, ja $P = (\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \ldots \quad \boldsymbol{x}_n)$ matriisi, jonka sarakkeet ovat kannan $B$ vektorit. Tällöin*

$$[T(\boldsymbol{x})]_B = P^{-1}AP[\boldsymbol{x}]_B \qquad \forall \boldsymbol{x} \in \mathbb{R}^n.$$

Nyt tarvittavat apulauseet on esitelty, joten lause 3.4 voidaan todistaa.

*Spektraalilauseen todistus* (vrt. [1, s. 337]). Osoitetaan ensin, että jos $A$ on ortogonaalisesti diagonalisoituva, niin $A$ on symmetrinen. Tällöin $A = P^{\mathrm{T}}DP$, ja

$$A^{\mathrm{T}} = (P^{\mathrm{T}}DP)^{\mathrm{T}} = P^{\mathrm{T}}(P^{\mathrm{T}}D)^{\mathrm{T}} = P^{\mathrm{T}}(D^{\mathrm{T}}P) = P^{\mathrm{T}}DP = A,$$

joten $A$ on symmetrinen.

Osoitetaan sitten, että jos $A$ on symmetrinen, niin $A$ on ortogonaalisesti diagonalisoituva. Sovelletaan induktiota matriisin koon suhteen.

Jokainen $1 \times 1$ -matriisi on diagonaalimatriisi, joten voidaan valita $P = I$, joka on ortogonaalinen matriisi. Lause on siis tosi, kun $n = 1$.

Oletetaan sitten, että lause pätee kaikille symmetrisille $(n-1) \times (n-1)$ -matriiseille.

Olkoon $A$ mikä tahansa symmetrinen $n \times n$ -matriisi. Olkoon $\lambda_1$ matriisin $A$ jokin ominaisarvo ja olkoon $\boldsymbol{v}_1$ vastaava ominaisvektori, jolle pätee $\|\boldsymbol{v}_1\| = 1$. Apulauseen 3.2 mukaan vektoriavaruuden $\mathrm{Lin}\{\boldsymbol{v}_1\}$ kanta $\{\boldsymbol{v}_1\}$ voidaan laajentaa vektoriavaruuden $\mathbb{R}^n$ kannaksi $\{\boldsymbol{v}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_n\}$. Käytetään sitten Gram-Schmidtin prosessia muuntamaan tämä kanta vektoriavaruuden $\mathbb{R}^n$ ortonormaaliksi kannaksi $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$.

Olkoon $P$ matriisi, jonka sarakkeet ovat kannan $B$ vektorit, vektorin $\boldsymbol{v}_1$ ollessa ensimmäinen sarake. Tällöin lauseen 2.4 mukaan $P$ on ortogonaalinen, ja seurauksen 3.1 mukaan $P^{\mathrm{T}}AP = P^{-1}AP$ on lineaarikuvauksen $T(\boldsymbol{x}) = A\boldsymbol{x}$ matriisi kannan $B$ suhteen. Apulauseen 3.3 mukaan matriisin $P^{\mathrm{T}}AP$ ensimmäinen sarake on lineaarikuvauksen $T(\boldsymbol{v}_1)$ koordinaattivektori kannan $B$ suhteen. Nyt $T(\boldsymbol{v}_1) = A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1$, joten tämä koordinaattivektori on

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

Tästä seuraa, että joillakin luvuilla $d_1, d_2, \ldots, d_{n-1}$ ja $c_{(i,j)}$, kun $i, j = 1, 2, \ldots, n-1$, $P^{\mathrm{T}}AP$ voidaan esittää muodossa

$$P^{\mathrm{T}}AP = \begin{pmatrix} \lambda_1 & d_1 & \ldots & d_{n-1} \\ 0 & c_{(1,1)} & \ldots & c_{(1,n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & c_{(n-1,1)} & \ldots & c_{(n-1,n-1)} \end{pmatrix}.$$

Koska $A$ on symmetrinen, niin myös $P^{\mathrm{T}}AP$ on, koska

$$(P^{\mathrm{T}}AP)^{\mathrm{T}} = P^{\mathrm{T}}A^{\mathrm{T}}P = P^{\mathrm{T}}AP.$$

Koska tämä matriisi on symmetrinen, niin tästä seuraa, että $d_1 = d_2 = \cdots = d_{n-1} = 0$, ja $(n-1) \times (n-1)$ -matriisi $C = \begin{pmatrix} c_{(i,j)} \end{pmatrix}$ on symmetrinen. Voidaan siis merkitä

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

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

Koska aiemmin oletettiin, että lause pätee kaikille symmetrisille $(n-1) \times (n-1)$ -matriiseille, niin se pätee myös matriisille $C$. Tämä tarkoittaa, että on olemassa ortogonaalinen $(n-1) \times (n-1)$ -matriisi $R$ siten, että $R^{\mathrm{T}}CR = D$, missä $D$ on diagonaalimatriisi. Tarkastellaan $n \times n$ -matriisia

$$Q = \begin{pmatrix} \lambda_1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & R \end{pmatrix}.$$

Koska matriisi $R$ on ortogonaalinen, niin matriisin $Q$ sarakkeet $2, 3, \dots, n$ ovat keskenään ortogonaalisia ja niiden pituus on 1. Lisäksi ensimmäinen sarake on ortogonaalinen muiden sarakkeiden kanssa ja sen pituus on 1. Matriisi $Q$ on siis ortogonaalinen. Olkoon $S = PQ$. Tällöin

$$S^{-1} = (PQ)^{-1} = Q^{-1}P^{-1} = Q^{\mathrm{T}}P^{\mathrm{T}} = (PQ)^{\mathrm{T}} = S^{\mathrm{T}},$$

joten $S$ on ortogonaalinen. Tarkastellaan sitten matriisia $S^{\mathrm{T}}AS$. Nyt

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

joka on diagonaalimatriisi, koska $D$ on diagonaalinen. Symmetriselle matriisille $A$ on siis olemassa ortogonaalinen matriisi $S$ siten, että $S^{\mathrm{T}}AS$ on diagonaalinen. Täten lause pätee $n \times n$ -matriiseille. □

**Esimerkki 3.2.** Tarkastellaan matriisia, jonka kaikki ominaisarvot eivät ole erisuuria.

Olkoon

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

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 Neliömuodot

## 4.1 Neliömuoto

Ortogonaalisen diagonalisoinnin yksi käytännöllinen sovellus on neliömuotojen analyysi.

Kahden muuttujan $x$ ja $y$ neliömuoto voidaan esittää muodossa

$$q(x,y) = ax^2 + 2cxy + by^2.$$

Tämä voidaan esittää myös muodossa

$$q = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x},$$

missä $A$ on symmetrinen matriisi $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ ja

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \boldsymbol{x}^{\mathrm{T}} = (x \quad y).$$

**Määritelmä 4.1** (Neliömuoto)**.** Olkoot $A$ symmetrinen $n \times n$ -neliömatriisi ja $\boldsymbol{x} \in \mathbb{R}^n$. Tällöin

$$q = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$$

on matriisiin $A$ liittyvä *neliömuoto*.

**Esimerkki 4.1.** Olkoon

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} \qquad \text{ja} \qquad \boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Tällöin matriisiin $A$ liittyvä neliömuoto on

$$q = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} = (x \quad y) \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 4x^2 + 4xy + 3y^2.$$

## 4.2 Neliömuotojen definiittisyys

Esimerkissä 4.1 $q(x,y) \geq 0$ kaikilla $x, y \in \mathbb{R}$. Lisäksi $q(x,y) = 0$ vain, kun $x = y = 0$. Tärkeä yleinen kysymys onkin, että onko neliömuoto aina positiivinen, kun $\boldsymbol{x}$ ei ole nollavektori. Sen selvittämiseen voidaan käyttää ominaisarvoja ja ortogonaalista diagonalisointia. Ensiksi esitellään neliömuodon definiittisyyden määritelmä.

**Määritelmä 4.2.** Olkoon $q(\boldsymbol{x})$ neliömuoto. Tällöin

- $q(\boldsymbol{x})$ on positiivisesti definiitti, jos $q(\boldsymbol{x}) > 0$ kaikilla $\boldsymbol{x} \in \mathbb{R}^n$ ja $q(\boldsymbol{x}) = 0$, vain kun $\boldsymbol{x} = \mathbf{0}$
- $q(\boldsymbol{x})$ on positiivisesti semidefiniitti, jos $q(\boldsymbol{x}) \geq 0$ kaikilla $\boldsymbol{x} \in \mathbb{R}^n$
- $q(\boldsymbol{x})$ on negatiivisesti definiitti, jos $q(\boldsymbol{x}) < 0$ kaikilla $\boldsymbol{x} \in \mathbb{R}^n$ ja $q(\boldsymbol{x}) = 0$, vain kun $\boldsymbol{x} = \mathbf{0}$
- $q(\boldsymbol{x})$ on negatiivisesti semidefiniitti, jos $q(\boldsymbol{x}) \leq 0$ kaikilla $\boldsymbol{x} \in \mathbb{R}^n$
- $q(\boldsymbol{x})$ on indefiniitti, jos on olemassa sellaiset $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$, että $q(\boldsymbol{x}_1) < 0$ ja $q(\boldsymbol{x}_2) > 0$.

**Lause 4.1.** *Oletetaan, että neliömuodolle $q(\boldsymbol{x})$ on olemassa matriisiesitys $q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} A\boldsymbol{x}$. Tällöin*

- *$q$ on positiivisesti definiitti, jos ja vain jos matriisin $A$ kaikki ominaisarvot ovat positiivisia*
- *$q$ on positiivisesti semidefiniitti, jos ja vain jos matriisin $A$ kaikki ominaisarvot ovat ei-negatiivisia*
- *$q$ on negatiivisesti definiitti, jos ja vain jos matriisin $A$ kaikki ominaisarvot ovat negatiivisia*
- *$q$ on negatiivisesti semidefiniitti, jos ja vain jos matriisin $A$ kaikki ominaisarvot ovat ei-positiivisia*
- *$q$ on indefiniitti, jos ja vain jos matriisilla $A$ on olemassa sellaiset ominaisarvot $\lambda_1$ ja $\lambda_2$, että $\lambda_1 < 0$ ja $\lambda_2 > 0$.*

*Todistus* (vrt. [1, s. 342]). Todistetaan lauseen ensimmäinen kohta. Tarkastellaan neliömuotoa $q = \boldsymbol{x}^{\mathrm{T}} A\boldsymbol{x}$, missä $A$ on symmetrinen ja ortogonaalisesti diagonalisoituva siten, että $P^{\mathrm{T}}AP = D$, missä $D$ on diagonaalimatriisi. Olkoon $\boldsymbol{x} = P\boldsymbol{z}$ eli $\boldsymbol{z} = P^{\mathrm{T}}\boldsymbol{x}$. Tällöin

$$q = \boldsymbol{x}^{\mathrm{T}} A\boldsymbol{x} = P\boldsymbol{z}^{\mathrm{T}} AP\boldsymbol{z} = \boldsymbol{z}^{\mathrm{T}}(P^{\mathrm{T}}AP)\boldsymbol{z} = \boldsymbol{z}^{\mathrm{T}} D\boldsymbol{z}.$$

Diagonaalimatriisin $D$ alkiot ovat matriisin $A$ ominaisarvot, koska ne ovat similaarisia matriiseja. Olkoon $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$ koordinaattivektori matriisin $A$ ominaisvektorien muodostaman ortonormaalin kannan suhteen. Tällöin

$$q = \boldsymbol{z}^{\mathrm{T}} D\boldsymbol{z} = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2,$$

joka on neliöiden lineaarikombinaatio.

Oletetaan nyt, että kaikki ominaisarvot ovat positiivisia. Tällöin voidaan sanoa, että $q \geq 0$ kaikilla $z$, ja $q = 0$ vain, kun $z$ on nollavektori. Koska $\boldsymbol{x} = Pz$ ja $z = P^{\mathrm{T}}\boldsymbol{x}$, niin $\boldsymbol{x} = \boldsymbol{0}$, jos ja vain jos $z = \boldsymbol{0}$. Tällöin määritelmän 4.2 mukaan neliömuoto on positiivisesti definiitti, jos matriisin $A$ kaikki ominaisarvot ovat positiivisia.

Oletetaan sitten, että neliömuoto on positiivisesti definiitti siten, että $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} > 0$ kaikilla $\boldsymbol{x} \neq \boldsymbol{0}$. Jos $\boldsymbol{u}_i$ on ominaisarvoa $\lambda_i$ vastaava yksikkövektoriksi normeerattu ominaisvektori, niin

$$\boldsymbol{u}_i^{\mathrm{T}}A\boldsymbol{u}_i = \boldsymbol{u}_i^{\mathrm{T}}\lambda_i\boldsymbol{u}_i = \lambda_i\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{u}_i = \lambda_1\|\boldsymbol{u}_i\|^2 = \lambda_i > 0,$$

joten matriisin $A$ ominaisarvot ovat positiivisia, jos neliömuoto on positiivisesti definiitti. $\square$

**Esimerkki 4.2.** Määritetään erään symmetrisen $2 \times 2$ -matriisin neliömuodon definiittisyys. Olkoon $A = \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix}$. Tällöin sen ominaisarvot ovat

$$\lambda_1 = 5 \quad \text{ja} \quad \lambda_2 = -3,$$

joten lauseen 4.1 nojalla matriisin $A$ neliömuoto on indefiniitti. Etsitään vielä sellaiset $\boldsymbol{x}_1$ ja $\boldsymbol{x}_2$, että määritelmän 4.2 ehdot täyttyvät. Olkoot $\boldsymbol{x}_1 = (-1, 1)$ ja $\boldsymbol{x}_2 = (1, 1)$. Tällöin

$$q(\boldsymbol{x}_1) = \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -6 < 0$$

ja

$$q(\boldsymbol{x}_2) = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 10 > 0,$$

joten on olemassa sellaiset $\boldsymbol{x}_1$ ja $\boldsymbol{x}_2$, että $q(\boldsymbol{x}_1) < 0$ ja $q(\boldsymbol{x}_2) > 0$.

Symmetrisen matriisin $A$ sanotaan olevan *positiivisesti definiitti*, jos sen neliömuoto $q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ on positiivisesti definiitti. Vastaavasti sanotaan myös negatiivisesti definiiteistä, positiivisesti ja negatiivisesti semidefiniiteistä sekä indefiniiteistä matriiseista.

Lauseen 4.1 seurauksena, matriisin $A$ positiivisen ja negatiivisen definiittisyyden määrittämiseksi, sen ominaisarvoista tarvitsee tietää vain niiden etumerkit, ei niiden arvoja. Näiden ominaisarvojen etumerkit voidaan selvittää suoraan matriisista $A$.

Tarkastellaan tapausta, jossa $A$ on symmetrinen $2 \times 2$ -matriisi,

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

Olkoon matriisin $A$ ominaisarvot $\lambda_1$ ja $\lambda_2$. Nyt lauseiden 2.5 ja 2.6 mukaan

$$\det(A) = ab - c^2 = \lambda_1\lambda_2 \qquad \text{ja} \qquad \operatorname{tr}(A) = a + b = \lambda_1 + \lambda_2.$$

Jos $\det(A) = ab - c^2 > 0$, niin molemmilla ominaisarvoilla $\lambda_1$ ja $\lambda_2$ on samat etumerkit. Alkioilla $a$ ja $b$ on myös oltava samat etumerkit, koska $ab > c^2 \geq 0$, joten ominaisarvojen etumerkit voidaan päätellä alkion $a$ etumerkistä.

**Esimerkki 4.3.** Olkoon

$$A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}.$$

Tällöin matriisilla $A$ on kaksi reaalista ominaisarvoa $\lambda_1$ ja $\lambda_2$, joiden tulee toteuttaa seuraavat yhtälöt:

$$\det(A) = \lambda_1\lambda_2 = 50 \qquad \text{ja} \qquad \operatorname{tr}(A) = \lambda_1 + \lambda_2 = 15.$$

Koska $\lambda_1\lambda_2 > 0$, molemmilla ominaisarvoilla on sama etumerkki eikä kumpikaan ole 0. Koska $\lambda_1 + \lambda_2 > 0$, molemmat ominaisarvot ovat positiivisia. Siten matriisi $A$ on positiivisesti definiitti.

Samoin, jos $\det(A) = ab - c^2 < 0$, niin ominaisarvoilla on eri etumerkit, ja tällöin matriisi $A$ on indefiniitti. Joten voidaan päätellä seuraavaa.

- Jos $\det(A) > 0$ ja $a > 0$, niin $\lambda_1 > 0$, $\lambda_2 > 0$ ja $A$ on positiivisesti definiitti.
- Jos $\det(A) > 0$ ja $a < 0$, niin $\lambda_1 < 0$, $\lambda_2 < 0$ ja $A$ on negatiivisesti definiitti.
- Jos $\det(A) < 0$, niin ominaisarvot $\lambda_1$ ja $\lambda_2$ ovat erimerkkisiä ja $A$ on indefiniitti.

Jos $\det(A) = 0$, niin voidaan päätellä, että yksi ominaisarvoista on 0.

Tämä tulos voidaan yleistää kaikkiin symmetrisiin $n \times n$ -matriiseihin, ja siihen tarvitaan johtavan pääminorin määritelmää.

**Määritelmä 4.3.** Jos $A$ on $n \times n$ -matriisi, niin sen johtavat pääminorit ovat $n$ determinanttia muodostettuna matriisin $A$ ensimmäisistä $r$:stä rivistä ja $r$:stä sarakkeesta, kun $r = 1, 2, \ldots, n$, eli

$$a_{11}, \quad \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \ldots \quad , \quad \det(A).$$

**Esimerkki 4.4.** Olkoon

$$A = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 10 & -3 \\ 1 & -3 & 2 \end{pmatrix}.$$

Tällöin sen johtavat pääminorit ovat

$$\det(5) = 5, \quad \det \begin{pmatrix} 5 & 2 \\ 2 & 10 \end{pmatrix} = 46 \quad \text{ja} \quad \det(A) = 25.$$

Seuraavaksi esitellään kaksi lausetta, joiden mukaan symmetrisen matriisin definiittisyys voidaan määrittää sen johtavien pääminorien avulla. Lauseiden todistukset kuitenkin sivuutetaan niiden laajuudesta johtuen.

**Lause 4.2.** *Oletetaan, että $A$ on symmetrinen $n \times n$ -matriisi. Tällöin $A$ on positiivisesti definiitti, jos ja vain jos sen kaikki johtavat pääminorit ovat positiivisia.*

*Todistus* (Ks. [1, s. 347–350]). □

**Lause 4.3.** *Oletetaan, että $A$ on symmetrinen $n \times n$ -matriisi. Tällöin $A$ on negatiivisesti definiitti, jos ja vain jos sen parilliset johtavat pääminorit ovat positiivisia ja parittomat negatiivisia.*

*Todistus.* Sivuutetaan. □

**Esimerkki 4.5.** Olkoon

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

Tällöin sen johtavat pääminorit ovat

$$\det(-3) = -3, \quad \det \begin{pmatrix} -3 & 2 \\ 2 & -5 \end{pmatrix} = 11 \quad \text{ja} \quad \det(A) = -6,$$

joten matriisi $A$ on negatiivisesti definiitti.

## 4.3 Ortogonaalisen matriisin määräämä lineaarikuvaus

**Lause 4.4.** *Ortogonaalisen matriisin $P$ määräämä lineaarikuvaus säilyttää sisätulon tuloksen vektoriavaruudessa $\mathbb{R}^n$.*

*Todistus* (vrt. [1, s. 355]). Olkoon $T\colon \mathbb{R}^n \to \mathbb{R}^n\colon T(\boldsymbol{x}) = P\boldsymbol{x}$. Olkoot $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$. Tällöin

$$\langle P\boldsymbol{v}, P\boldsymbol{w}\rangle = (P\boldsymbol{v})^{\mathrm{T}}(P\boldsymbol{w}) = \boldsymbol{v}^{\mathrm{T}}P^{\mathrm{T}}P\boldsymbol{w} = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = \langle \boldsymbol{v}, \boldsymbol{w}\rangle.$$

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

**Esimerkki 4.6.** Olkoon

$$\boldsymbol{v} = (1, 2, 1) \qquad \text{ja} \qquad \boldsymbol{w} = (3, 0, 1).$$

Tällöin

$$\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 4.$$

Olkoon sitten

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

ortogonaalinen matriisi. Tällöin

$$P\boldsymbol{v} = \begin{pmatrix} \frac{2}{3} & \frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{3} & \frac{2}{3\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{3} & \frac{-\sqrt{5}}{3} & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{11\sqrt{5}}{15} + \frac{2}{3} \\ \frac{1}{3} - \frac{2\sqrt{5}}{15} \\ \frac{2}{3} - \frac{2\sqrt{5}}{3} \end{pmatrix} \quad \text{ja} \quad P\boldsymbol{w} = \begin{pmatrix} \frac{2}{3} & \frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{3} & \frac{2}{3\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{3} & \frac{-\sqrt{5}}{3} & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} + 2 \\ 1 - \frac{2\sqrt{5}}{5} \\ 2 \end{pmatrix}.$$

Nyt

$$\langle P\boldsymbol{v}, P\boldsymbol{w}\rangle = \begin{pmatrix} \frac{11\sqrt{5}}{15} + \frac{2}{3} \\ \frac{1}{3} - \frac{2\sqrt{5}}{15} \\ \frac{2}{3} - \frac{2\sqrt{5}}{3} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \frac{1}{\sqrt{5}} + 2 \\ 1 - \frac{2\sqrt{5}}{5} \\ 2 \end{pmatrix} = \frac{8\sqrt{5}}{5} + \frac{31}{15} + \frac{3}{5} - \frac{4\sqrt{5}}{15} + \frac{4 - 4\sqrt{5}}{3} = 4,$$

eli $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 4 = \langle P\boldsymbol{v}, P\boldsymbol{w}\rangle$.

# 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.