TURUN YLIOPISTO

Matriisihajotelmista

Maiju Virta

Pro gradu -tutkielma
Toukokuu 2020

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

TURUN YLIOPISTO
Matematiikan ja tilastotieteen laitos

Virta, Maiju: Matriisihajotelmista
Pro gradu -tutkielma, 31 s.
Matematiikka
Toukokuu 2020

Tämä matematiikan opettajalinjan pro gradu -tutkielma käsittelee matriisihajotelmia. Matriisihajotelmassa matriisi esitetään kahden tai useamman yksinkertaista muotoa olevan matriisin tulona. Matriisihajotelmia käytetään esimerkiksi lineaaristen yhtälöryhmien ratkaisemiseen. Hajotelmilla on monia eri käyttötarkoituksia ja jokaiseen tilanteeseen sopii tietynlainen hajotelma, jonka vuoksi hajotelmia on monenlaisia.

Tutkielma on suunnattu oppimateriaaliksi aiheesta kiinnostuneille, joilla on kuitenkin perustietämys lineaarialgebrasta; esimerkiksi henkilöille, jotka ovat käyneet lineaarialgebran kurssin. Tutkielman alussa kerrataan jonkin verran lineaarialgebrasta tuttuja asioita.

Tutkielma jakaantuu kahteen osaan. Ensimmäinen osa koostuu luvusta 1, jossa esitellään matriisihajotelmissa tarvittavaa pohjatietoa. Tämän luvun on tarkoitus toimia lineaarialgebran kertauksena. Toinen osa koostuu luvusta 2, joka käsittelee matriisihajotelmia. Tutkielmaan on valittu QR-hajotelma, LU-hajotelma, spektraalihajotelma ja -esitys, singulaariarvohajotelma ja polaarihajotelma. Jokaisesta hajotelmasta annetaan myös esimerkki havainnollistamaan sen muodostamista.

Asiasanat: matriisihajotelma, lineaarialgebra, QR-hajotelma, LU-hajotelma, spektraalihajotelma, spektraaliesitys, singulaariarvohajotelma, polaarihajotelma

Sisältö

1 Johdanto

Matematiikan kehittyminen on peräisin ihmisten halusta ratkaista käytännön ongelmia. Näin on syntynyt myös lineaariset yhtälöt ja yhtälöryhmät. Matriisit ja determinantit ovat kehittyneet, kun yhtälöryhmien ratkaisulle on etsitty sääntöjä ja ratkaisumenetelmiä. Alunperin matriisiteoria on korostanut enemmän determinantteja kuin matriiseja. Matriisilaskennan kehittymisen katsotaan alkaneen 1600-luvulla, vaikka matriisien ja determinanttien historia alkaa jo 300-100 -luvuilta eKr. [3]

Matriisiyhtälön ratkaiseminen on usein haastavaa, joten siksi matriisilaskennassa pyritään usein esittämään matriisi kahden tai useamman yksinkertaista muotoa olevan matriisin tulona eli matriisihajotelmana. Matriisihajotelma siis helpottaa matriisiyhtälön ratkaisemista. Hajotelmia on monia erilaisia, sillä eri tilanteissa tarvitaan eri hajotelmia. Tässä tutkielmassa esitellään muutamia yleisiä matriisihajotelmia.

Työn on tarkoitus toimia matriisihajotelmia käsittelevänä itseopiskelumateriaalina aiheesta kiinnostuneelle henkilölle. Pohjatietona henkilöllä kuitenkin oletetaan olevan perustiedot lineaarialgebrasta, esimerkiksi lineaarialgebran kurssilta.

Tutkielman alussa luvussa 2 käydään läpi pohjatietoa. Tarkoituksena on, että luku toimii kertausmateriaalina lineaarialgebran aiheille, joita tarvitaan matriisihajotelmia käsiteltäessä. Luku käsittelee mm. vektoriavaruutta, matriiseja ja matriisien ominaisuuksia. Luvussa 3 päästään itse matriisihajotelmiin, johon on valittu QR-hajotelma, LU-hajotelma, spektraalihajotelma ja -esitys, singulaariarvohajotelma ja polaarihajotelma. Hajotelmat esitellään yksi kerrallaan ja jokaisesta hajotelmasta annetaan esimerkki.

Työ pohjautuu vahvasti kirjallisuuteen. Pääasiallisina lähteinä ovat toimineet Strangin kirja *Introduction to Linear Algebra* [6] ja Koppisen luentomoniste *Matriisilaskenta* [2].

2 Pohjatietoa

Tässä työssä matriiseja tarkastellaan selvyyden ja yksinkertaisuuden vuoksi pääosin yli reaalilukujen. Työn lopulla luvussa 3.5 syvennytään lyhyesti myös kompleksilukuihin.

2.1 Vektoriavaruus

Vektoriavaruus on joukko, jonka alkiot ovat *vektoreita* ja jolle on määritelty laskutoimitukset summa ja skalaarilla kertominen. Esimerkiksi

$$\mathbb{R}^n = \{\bar{x} = (x_1, ..., x_n) \mid x_i \in \mathbb{R}\},$$

on vektoriavaruus, johon kuuluu laskutoimitukset

$$\begin{aligned} \bar{x} + \bar{y} &= (x_1 + y_1, ..., x_n + y_n) \quad \text{ja} \\ a\bar{x} &= (ax_1, ..., ax_n), \end{aligned}$$

missä $\bar{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, $\bar{y} = (y_1, ..., y_n) \in \mathbb{R}^n$ ja $a \in \mathbb{R}$.

Määritelmä 2.1. Jos $(U, +, \cdot)$ ja $(V, +, \cdot)$ ovat vektoriavaruuksia, joille on määritelty summa ja skalaarilla kertominen, sekä $U \subseteq V$, vektoriavaruus U on vektoriavaruuden V *aliavaruus.*

Aliavaruuskriteerin mukaan vektoriavaruuden V osajoukko $U \neq \emptyset$ on aliavaruus, jos ja vain jos

1. $X + Y \in U$, kun $X, Y \in U$ ja
2. $aX \in U$, kun $a \in \mathbb{R}$ ja $X \in U$.

Määritelmä 2.2. Jokainen vektoriavaruuden V osajoukko S *virittää* aliavaruuden

$$L(S) = \{a_1\bar{x}_1 + ... + a_k\bar{x}_k \mid a_i \in \mathbb{R},\ \bar{x}_i \in S\}.$$

Aliavaruuskriteerin perusteella $L(S)$ on todella vektoriavaruuden V aliavaruus.

Vektoriavaruuden V osajoukko $S = \{x_1, ..., x_k\}$ on *lineaarisesti riippuva*, jos on olemassa sellaiset $a_1, ..., a_k \in \mathbb{R}$, että jokin $a_i \neq 0$ ja

$$a_1x_1 + ... + a_kx_k = \bar{0}.$$

Jos osajoukko S ei ole lineaarisesti riippuva, se on *lineaarisesti riippumaton.*

Määritelmä 2.3. Osajoukko $S = \{x_1, ..., x_k\}$ on vektoriavaruuden V *kanta*, jos osajoukko on lineaarisesti riippumaton ja virittää vektoriavaruuden V. Jos vektoriavaruudella V on tällainen kanta, vektoriavaruutta V kutsutaan *äärellisulotteiseksi* ja sen *dimensio* $\dim(V) = k$.

2.2 Matriisit

Matriisi muistuttaa suorakulmaista taulukkoa, joka on jaettu riveihin ja sarakkeisiin. Matriisi sisältää alkioita a_{ij} ja sen rakenne on seuraavanlainen:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{m \times n}, \quad a_{ij} \in \mathbb{R},$$

on $m \times n$ *-matriisi*, jossa siis m on vaakarivien määrä ja n pystyrivien eli sarakkeiden määrä. Kaikkien $m \times n$ -matriisien joukkoa merkitään $\mathcal{M}_{m \times n}(\mathbb{R})$, missä $\mathbb{R}$ tarkoittaa, että matriisin alkiot ovat reaalilukuja. *Neliömatriiseille* $m = n$ ja $\mathcal{M}_{m \times n}(\mathbb{R}) = \mathcal{M}_n(\mathbb{R})$ eli vaaka- ja pystyrivejä on yhtä paljon. Neliömatriiseissa diagonaalin alkioilla eli *diagonaalialkioilla* tarkoitetaan alkioita a_{ij}, kun $i = j$, eli alkioita $a_{11}, a_{22}, ..., a_{nn}$.

Matriiseja voidaan myös laskea yhteen ja kertoa keskenään. Seuraavaksi määritellään, miten nämä laskutoimitukset suoritetaan matriiseilla.

Määritelmä 2.4. Jos $A, B \in \mathcal{M}_{m \times n}(\mathbb{R})$, niin summa $A + B = (c_{ij}) \in \mathcal{M}_{m \times n}(\mathbb{R})$, missä $c_{ij} = a_{ij} + b_{ij}$. Jos $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ ja $B \in \mathcal{M}_{n \times r}(\mathbb{R})$, niin tulo $AB = (c_{ij}) \in \mathcal{M}_{m \times r}(\mathbb{R})$, missä $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$.

Matriisien summa siis lasketaan seuraavasti:

$$
\begin{aligned}
A + B &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \\
&= \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} = C.
\end{aligned}
$$

Matriisien kertolasku taas lasketaan seuraavasti:

$$
\begin{aligned}
AB &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{m1} & a_{11}b_{12} + \cdots + a_{1n}b_{m2} & \cdots & a_{11}b_{1n} + \cdots + a_{1n}b_{mn} \\ a_{21}b_{11} + \cdots + a_{2n}b_{m1} & a_{21}b_{12} + \cdots + a_{2n}b_{m2} & \cdots & a_{21}b_{1n} + \cdots + a_{2n}b_{mn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{m1} & a_{m1}b_{12} + \cdots + a_{mn}b_{m2} & \cdots & a_{m1}b_{1n} + \cdots + a_{mn}b_{mn} \end{pmatrix} \\
&= \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} = C.
\end{aligned}
$$

Matriisien muokkaaminen ns. vaakarivioperaatioilla helpottaa usein matriisin erilaisten ominaisuuksien tarkastelussa. Määritellään seuraavaksi matriisin vaakarivioperaatiot.

1. Matriisin rivien i ja j vaihtaminen.
2. Matriisin rivin i alkioiden kertominen vakiolla c.
3. Matriisin rivin j lisääminen riviin i kerrottuna vakiolla d.

Jos matriisi B saadaan matriisista A edellisillä operaatioilla, matriisit ovat *riviekvivalentit*.

Matriisilaskuissa matriisin muuttaminen ns. *redusoiduksi porrasmatriisiksi* edellä mainituilla operaatioilla on usein hyödyllistä. Määritellään ensin käsite *porrasmatriisi*.

Määritelmä 2.5. Matriisi on *porrasmatriisi*, jos seuraavat ehdot toteutuvat:

1. mahdolliset nollarivit ovat alimpana ja
2. kullakin rivillä ensimmäinen nollasta poikkeava alkio on ylemmän rivin ensimmäisen nollasta poikkeavan alkion oikealla puolella.

Usein jo porrasmuoto on riittävä, mutta kutakin matriisia vastaa useampi kuin yksi matriisin kanssa riviekvivalentti porrasmatriisi, kun taas redusoitu porrasmatriisi on kullekin matriisille yksikäsitteinen. Seuraavaksi määritellään käsite redusoitu porrasmatriisi.

Määritelmä 2.6. Matriisi on redusoitu porrasmatriisi, jos seuraavat ehdot toteutuvat:

1. matriisi on porrasmatriisi,
2. jokaisen rivin ensimmäinen nollasta poikkeava alkio on yksi ja
3. ensimmäinen nollasta poikkeava alkio on sarakkeensa ainoa nollasta poikkeava alkio.

Seuraavassa esimerkissä havainnollistetaan matriisin saattamista vaakarivioperaatioilla (redusoituun) porrasmuotoon ns. Gaussin eliminaatiomenetelmällä.

Esimerkki 2.7. Muuta matriisi

$$\begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 0 & 2 & 1 \\ -1 & -2 & 0 & 3 \end{pmatrix}$$

redusoiduksi porrasmatriisiksi Gaussin eliminaatioalgoritmilla.

Aloitetaan vaihtamalla ensimmäisen ja toisen rivin paikat, jotta saadaan ensimmäisen rivin ensimmäiseksi nollasta poikkeavaksi alkioksi luku 1:

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 2 \\ -1 & -2 & 0 & 3 \end{pmatrix}$$

Muutetaan luvun 1 alla olevat alkiot nolliksi lisäämällä ensimmäinen rivi toiseen riviin luvulla -2 kerrottuna ja kolmanteen riviin luvulla 1 kerrottuna:

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -2 & 2 & 4 \end{pmatrix}.$$

Kerrotaan toinen rivi luvulla -1, jonka jälkeen lisätään toinen rivi kolmanteen riviin luvulla 2 kerrottuna, jotta saadaan myös toisen rivin ensimmäiseksi nollasta poikkeavaksi alkioksi luku 1 ja sen alla olevat alkiot nolliksi. Saadaan matriisi

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 4 & 4 \end{pmatrix},$$

joka on iseasiassa porrasmatriisi. Jatketaan kertomalla kolmas rivi luvulla $\frac{1}{4}$ ja lisäämällä sen jälkeen kolmas rivi toiseen riviin kerrottuna luvulla -1 ja ensimmäiseen riviin kerrottuna luvulla -2:

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

jolloin saadaan redusoitu porrasmatriisi. Kaikki yllä esitetyt matriisit ovat siis keskenään riviekvivalentteja.

Usein matriisilaskuissa esiintyy *identiteettimatriisi* I, joka on neliömatriisi, jonka diagonaalialkiot ovat ykkösiä ja muut alkiot ovat nollia. Esimerkiksi 3×3 -identiteettimatriisi on

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Matriisin $A = (a_{ij})$ *transpoosi* $A^T = (a_{ji})$ saadaan, kun matriisi *transponoidaan* eli sen sarakkeet vaihdetaan riveiksi tai vastaavasti rivit muutetaan sarakkeiksi. Esimerkiksi 3×3 -matriisin

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

transpoosi

$$A^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

Mikäli matriisi ja sen transpoosi ovat samat eli $A = A^T$, matriisi A on *symmetrinen.* Lisäksi $(AB)^T = B^T A^T$.

Määritelmä 2.8. Neliömatriisi $A \in \mathcal{M}_n(\mathbb{R})$ on *säännöllinen*, jos on olemassa sellainen $B \in \mathcal{M}_n(\mathbb{R})$, että $AB = I = BA$. Matriisi B on matriisin A *käänteismatriisi*, eli $B = A^{-1}$, jos $AB = I$ tai $BA = I$. Matriisia, joka ei ole säännöllinen, sanotaan *singulaariseksi.*

Käänteismatriisi A^{-1} neliömatriisille voidaan laskea esimerkiksi Gaussin ja Jordanin menetelmällä. Menetelmässä matriisin A rinnalle asetetaan identiteettimatriisi, jonka jälkeen Gaussin eliminaatioalgoritmilla muokataan matriisista A identiteettimatriisi, jolloin alkuperäisestä identiteettimatriisista tulee käänteismatriisi A^{-1}.

Esimerkki 2.9. Etsi käänteismatriisi P^{-1} matriisille

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

Asetetaan matriisi P ja identiteettimatriisi rinnakkain

$$(P \mid I_3) = \begin{pmatrix} 1 & 1 & 1 & | & 1 & 0 & 0 \\ -1 & 0 & 2 & | & 0 & 1 & 0 \\ 1 & -1 & 1 & | & 0 & 0 & 1 \end{pmatrix}$$

ja muokataan Gaussin eliminaatioalgoritmilla matriisista P identiteettimatriisi.

Lisätään ensimmäinen rivi toiseen riviin kerrottuna luvulla 1 ja kolmanteen riviin kerrottuna luvulla -1

$$\begin{pmatrix} 1 & 1 & 1 & | & 1 & 0 & 0 \\ 0 & 1 & 3 & | & 1 & 1 & 0 \\ 0 & -2 & 0 & | -1 & 0 & 1 \end{pmatrix},$$

jonka jälkeen lisätään kolmas rivi ensimmäiseen riviin kerrottuna luvulla $\frac{1}{2}$ ja toinen rivi kolmanteen riviin kerrottuna luvulla 2

$$\begin{pmatrix} 1 & 0 & 1 & | & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 3 & | & 1 & 1 & 0 \\ 0 & 0 & 6 & | & 1 & 2 & 1 \end{pmatrix}.$$

Kerrotaan kolmas rivi luvulla $\frac{1}{6}$, jonka jälkeen lisätään se toiseen riviin kerrottuna luvulla -3 ja ensimmäiseen riviin kerrottuna luvulla -1, jolloin saadaan

$$\begin{pmatrix} 1 & 0 & 0 & | & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & | & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & | & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

Näin ollen käänteismatriisi

$$P^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

2.3 Determinantti

Neliömatriiseille voidaan laskea *determinantti*, joka kertoo esimerkiksi sen, onko matriisi säännöllinen vai ei. Determinantti voidaan määritellä usealla eri tavalla, joista yksi esitellään seuraavaksi.

Määritelmä 2.10. Neliömatriisin $A = (a_{ij})_{n \times n}$ determinantti on

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}), \quad \text{kun} \quad i = 1, ..., n,$$

missä matriisi A_{ij} saadaan poistamalla matriisista A i:nnes rivi ja j:nnes sarake. Matriisin determinantti voidaan siis edellisen kaavan perusteella laskea ns. *kehittämällä* se i:nen vaakarivin mukaan. Kaavaa toistetaan kunnes matriisi A_{ij} on 2×2-matriisi, jolloin matriisin A determinantti

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Esimerkki 2.11. Laske determinantti matriisille

$$A = \begin{pmatrix} -2 & 3 & 2 \\ 0 & 1 & -1 \\ 1 & 2 & 4 \end{pmatrix}.$$

Matriisin A determinantti voidaa nyt laskea kehittämällä se ensimmäisen rivin mukaan:

$$\begin{aligned} \det(A) &= \begin{vmatrix} -2 & 3 & 2 \\ 0 & 1 & -1 \\ 1 & 2 & 4 \end{vmatrix} \\ &= -2 \cdot \begin{vmatrix} 1 & -1 \\ 2 & 4 \end{vmatrix} - 3 \cdot \begin{vmatrix} 0 & -1 \\ 1 & 4 \end{vmatrix} + 2 \cdot \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} \\ &= -2(1 \cdot 4 - (-1) \cdot 2) - 3(0 \cdot 4 - (-1) \cdot 1) + 2(0 \cdot 2 - 1 \cdot 1) \\ &= -2 \cdot 6 - 3 \cdot 1 + 2 \cdot (-1) = -17. \end{aligned}$$

Kerrataan seuraavaksi joitakin determinantin perusominaisuuksia, joiden todistukset löytyvät esimerkiksi kirjasta [6]:

1. Matriisin transpoosin determinantti $\det(A^T) = \det(A)$.
2. Jos matriisin A kaksi vaaka- tai pystyriviä vaihdetaan keskenään, determinantin $\det(A)$ etumerkki vaihtuu.
3. Determinantti $\det(A) = 0$, jos matriisissa A on kaksi samaa vaaka- tai pystyriviä.
4. Matriisin tulon determinantti $\det(AB) = \det(A)\det(B)$.

Jos determinantti $\det(A) = 0$, neliömatriisi A on singulaarinen. Jos determinantti $\det(A) \neq 0$, matriisi A on ei-singulaarinen eli säännöllinen.

Esimerkiksi *yläkolmiomatriisin* A, jossa siis päälävistäjän alapuolella olevat alkiot ovat nollia, determinantti lasketaan seuraavasti määritelmän 2.10 mukaisesti:

$$\begin{aligned} \det(A) &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} \\ &= \cdots = a_{11}a_{22}\cdots a_{nn}. \end{aligned}$$

Vastaavasti matriisia, jonka päälävistäjän yläpuolella olevat alkiot ovat nollia, kutsutaan *alakolmiomatriisiksi*. Alakolmiomatriisi on siis seuraavanlainen:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Myös alakolmiomatriisin determinantti on diagonaalin alkioiden tulo, kuten yläkolmiomatriisillakin.

Esimerkki 2.12. Lasketaan edellisen esimerkin matriisille

$$A = \begin{pmatrix} -2 & 3 & 2 \\ 0 & 1 & -1 \\ 1 & 2 & 4 \end{pmatrix}$$

determinantti muokkaamalla matriisi ensin vaakarivimuunnoksilla yläkolmiomatriisiksi. Aloitetaan lisäämällä ensimmäinen rivi kolmanteen riviin kerrottuna luvulla $\frac{1}{2}$, jolloin

$$\begin{pmatrix} -2 & 3 & 2 \\ 0 & 1 & -1 \\ 0 & \frac{7}{2} & 5 \end{pmatrix}.$$

Lisätään vielä toinen rivi kolmanteen riviin kerrottuna luvulla $\frac{-7}{2}$, jolloin saadaan yläkolmiomatriisi

$$\begin{pmatrix} -2 & 3 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & \frac{17}{2} \end{pmatrix}.$$

Lasketaan determinantti soveltaen yläkolmiomatriisin determinantin kaavaa, jonka mukaan determinantti

$$\det(A) = \begin{vmatrix} -2 & 3 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & \frac{17}{2} \end{vmatrix} = -2 \cdot 1 \cdot \frac{17}{2} = -17.$$

Tulos todellakin pitää paikkansa, sillä se on sama kuin edellisessä esimerkissä matriisille A laskettu determinantti.

Seuraavan lauseen mukaisesti determinantin avulla voidaan myös helposti tutkia, onko tarkasteltu matriisi säännöllinen.

Lause 2.13. *Matriisi A on säännöllinen, jos ja vain jos sen determinantti on erisuuri kuin nolla eli* $\det(A) \neq 0$.

2.4 Sisätulo ja ortogonaalisuus

Kokoa $m \times 1$ tai $1 \times n$ olevat matriisit ovat vektoreita. Matriisin sarakkeet eli *pystyrivit* $\bar{a}_j^T = (a_{1j}, ..., a_{mj}) \in \mathbb{R}^m$ muodostavat *pystyriviavaruuden* $P(A) = L(\bar{a}_1, ..., \bar{a}_n)$. *Vaakarivit* $\bar{a}^i = (a_{i1}, ..., a_{in}) \in \mathbb{R}^n$ vastaavasti muodostavat *vaakariviavaruuden* $V(A) = L(\bar{a}^1, ..., \bar{a}^m)$. Seuraavaksi esitellään muutamia vektorien ominaisuuksia.

Määritelmä 2.14. Vektoreiden $\bar{x} = \{x_1, x_2, ..., x_n\}$ ja $\bar{y} = \{y_1, y_2, ..., y_n\}$ *sisätuloa* merkitään $\langle \bar{x}, \bar{y} \rangle$ ja se määritellään

$$\langle \bar{x}, \bar{y} \rangle = x_1 y_1 + x_2 y_2 + ... + x_n y_n.$$

Määritelmä 2.15. Vektorin $\bar{x} = \{x_1, x_2, ..., x_n\}$ *pituus* on

$$|\bar{x}| = \sqrt{\langle \bar{x}, \bar{x} \rangle} = \sqrt{|x_1|^2 + |x_2|^2 + ... + |x_n|^2}.$$

Esimerkki 2.16. Laske vektoreiden $\bar{x} = (1,0,1)$ ja $\bar{y} = (-1,1,0)$ sisätulo, sekä vektorin $\bar{x}$ pituus. Vektroreiden $\bar{x}$ ja $\bar{y}$ sisätulo

$$\begin{aligned}\langle \bar{x}, \bar{y} \rangle &= \langle (1,0,1), (-1,1,0) \rangle \\ &= 1 \cdot (-1) + 0 \cdot 1 + 1 \cdot 0 = -1\end{aligned}$$

ja vektorin $\bar{x}$ pituus

$$\begin{aligned}|\bar{x}| &= \sqrt{\langle \bar{x}, \bar{x} \rangle} = \sqrt{\langle (1,0,1), (1,0,1) \rangle} \\ &= \sqrt{|1|^2 + |0|^2 + |1|^2} = \sqrt{2}.\end{aligned}$$

Määritelmä 2.17. Vektorit $\bar{x}$ ja $\bar{y}$ ovat *kohtisuoria* eli *ortogonaalisia*, jos sisätulo $\langle \bar{x}, \bar{y} \rangle = 0$. Vektorijoukko $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_k\}$ on ortogonaalinen, jos sisätulo $\langle \bar{x}_i, \bar{x}_j \rangle = 0$, kun $i \neq j$. Lisäksi joukko on *ortonormaali*, jos $|\bar{x}_i| = 1$ kaikilla $i \in \{1, ..., k\}$.

Seuraavaksi kerrataan vektorin kantaesitys, jonka todistus löytyy esimerkiksi kirjasta [6].

Lause 2.18. *Jos* $\{\bar{u}_1, ..., \bar{u}_m\}$ *on aliavaruuden* U ortonormaalikanta*, niin vektorin* $\bar{x} \in U$ kantaesitys *on*

$$\bar{x} = \sum_{i=1}^{m} \langle \bar{x}, \bar{u}_i \rangle \bar{u}_i.$$

Määritelmä 2.19. Matriisi $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ on ortogonaalinen, jos $A^T A = I$.

Huomautus 2.20. Jos matriisi A on neliömatriisi ja ortogonaalinen, eli $A^T A = I$, niin matriisin A transpoosin on oltava matriisin A käänteismatriisi eli $A^T = A^{-1}$.

2.5 Matriisin aste

Luvut n ja m kertovat matriisin koon, mutta eivät välttämättä lineaarisen systeemin todellista kokoa. Matriisin todellisen koon määrää tietyssä mielessä matriisin A *aste*, joka saadaan pysty- ja vaakariviavaruuksien dimensioista.

Määritelmä 2.21. Matriisin $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ aste $r(A) = \dim V(A) = \dim P(A)$ eli vaaka- ja pystyriviavaruuksien dimensio.

Matriisin vaakariviavaruuden kanta ja aste voidaan määrittää seuraavan esimerkin mukaisesti saattamalla matriisi ensin vaakarivioperaatioilla (redusoituun) porrasmuotoon.

Esimerkki 2.22. Laske aste matriisille

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & 7 & 12 \end{pmatrix}.$$

Muokataan matriisista A ensin redusoitu porrasmatriisi. Aloitetaan lisäämällä ensimmäinen rivi toiseen riviin kerrottuna luvulla -4 ja kolmanteen riviin kerrottuna luvulla -2

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 3 & 6 \end{pmatrix}.$$

Lisätään toinen rivi kolmanteen riviin, jonka jälkeen kerrotaan toinen rivi luvulla $\frac{-1}{3}$, jolloin saadaan porrasmatriisi

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Lisätään vielä toinen rivi ensimmäiseen riviin kerrottuna luvulla -2, jolloin saadaan redusoitu porrasmatriisi

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Vaakariviavaruuden kanta on $\{(1,0,1),\ (0,1,2)\}$. Matriisin A aste $r(A) = 2$, sillä vaakarivien dimensio on kaksi. Matriisin asteessa ei siis oteta huomioon nollariviä, joka vaakarivimuunnoksilla muodostui.

Seuraava lemma, jonka todistus löytyy luentomonisteesta [2], antaa ylärajan matriisien tulon asteelle.

Lemma 2.23. *Jos $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ ja $B \in \mathcal{M}_{n\times k}(\mathbb{R})$, niin matriisin AB aste*

$$r(AB) \leqslant \min\{r(A), r(B)\}.$$

Seuraavassa lauseessa saadaan ensimmäinen matriisihajotelma.

Lause 2.24. *Jos matriisi $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ ja $r(A) = r$, on olemassa sellaiset matriisit $B \in \mathcal{M}_{m\times r}(\mathbb{R})$ ja $C \in \mathcal{M}_{r\times n}(\mathbb{R})$, että $A = BC$ ja $r(B) = r(C) = r$.*

Todistus. Valitaan matriisin A pystyriviavaruudelle $P(A)$ kanta $\{\bar{b}_1, ..., \bar{b}_r\}$. On olemassa sellaiset kertoimet $c_{ij} \in \mathbb{R}$, että matriisin A pystyrivit

$$\bar{a}_j = \sum_{k=1}^{r} c_{kj}\bar{b}_k \quad (j = 1, ..., n).$$

Jos merkitään $\bar{b}_k = (b_{1k}, ..., b_{mk})^T$, niin

$$a_{ij} = \sum_{k=1}^{r} c_{kj}b_{ik} = \sum_{k+1}^{r} b_{ik}c_{kj} \quad (i = 1, ..., m,\ j = 1, ..., n).$$

Näin ollen saadaan matriisit $B = (b_{ij})_{m\times r}$ ja $C = (c_{ij})_{r\times n}$, ja $A = BC$. Koska matriisissa B on r pystyriviä, matriisin B aste $r(B) \leqslant r$. Myös matriisin C aste $r(C) \leqslant r$, koska matriisissa C on r vaakariviä.

Lemman 2.23 mukaan matriisin A aste on $r(A) = r(BC) \leqslant \min\{r(B), r(C)\}$, mutta jos $r(B) < r$ tai $r(C) < r$, niin myös $r(A) < r$, mikä ei pidä paikkansa. Siispä $r(B) = r(C) = r$. □

Tätä matriisihajotelmaa tarkastellaan lähemmin esimerkin kera luvussa 3.

2.6 Ominaisarvot ja -vektorit

Määritelmä 2.25. Matriisin $A \in \mathcal{M}_n(\mathbb{R})$ *karakteristinen polynomi* eli *ominaisarvopolynomi* on

$$c_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}.$$

Määritelmä 2.26. Olkoon $A \in \mathcal{M}_n(\mathbb{R})$. Luku λ on matriisin A *ominaisarvo* ja $\bar{x} \in \mathbb{R}^n$, $\bar{x} \neq \bar{0}$, on siihen kuuluva *ominaisvektori*, jos $A\bar{x} = \lambda\bar{x}$. *Ominaisavaruus* $V_\lambda = \{\bar{x} \in \mathbb{R}^n | A\bar{x} = \lambda\bar{x}\}$ sisältää ominaisarvoon λ kuuluvat ominaisvektorit sekä nollavektorin $\bar{0}$.

Huomautus 2.27. Luku λ on matriisin A ominaisarvo, jos ja vain jos λ on yhtälön $\det(A - \lambda I) = 0$ ratkaisu. Karakteristista polynomia voidaan siis käyttää ominaisarvojen etsimiseen ratkaisemalla yhtälö $\det(A - \lambda I) = 0$.

Geometrisesti ominaisvektorilla tarkoitetaan suuntaa ja ominaisarvolla sen skaalausta. Jos ominaisarvo $\lambda_i \geqslant 1$, vektorin suunta säilyy ja pituus kasvaa. Jos $0 < \lambda_i < 1$, suunta säilyy, mutta pituus pienenee. Lisäksi ominaisarvon ollessa negatiivinen, suunta muuttuu vastakkaissuuntaiseksi.

Määritelmä 2.28. Matriisin $A \in \mathcal{M}_n(\mathbb{R})$ ominaisarvon $\lambda_i \in \mathbb{R}$, *algebrallinen kertaluku* on k_i, kun λ_i on karakteristisen polynomin k_i-kertainen nollakohta. Ominaisavaruuden V_{λ_i} dimensio on ominaisarvon λ_i *geometrinen kertaluku.*

Esimerkki 2.29. Etsi ominaisarvot ja ominaisvektorit matriisille

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{pmatrix}.$$

Etsitään matriisin A ominaisarvot ratkaisemalla yhtälö $\det(A - \lambda I) = 0$:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 0 & 4 \\ 0 & 6-\lambda & 0 \\ 4 & 0 & 2-\lambda \end{vmatrix} \\ &= (2-\lambda)(6-\lambda)(2-\lambda) + 4(-4(6-\lambda)) \\ &= (6-\lambda)((2-\lambda)^2 - 16) \\ &= (6-\lambda)^2(-2-\lambda) = 0, \end{aligned}$$

jolloin matriisin A ominaisarvot ovat $\lambda_2 = 6$ ja $\lambda_1 = -2$. Ominaisarvon 6 algebrallinen kertaluku on 2, sillä 6 on 2-kertainen karakteristisen polynomin nollakohta. Vastaavasti ominaisarvon -2 algebrallinen kertaluku on 1.

Etsitään ominaisvektorit yhtälöllä $(A - \lambda I)\bar{x} = 0$ sijoittamalla vuorollaan kukin ominaisarvo. Sijoittamalla ominaisarvo $\lambda_1 = -2$ yllä olevaan yhtälöön voidaan määrittää siihen kuuluvat ominaisvektorit:

$$\begin{pmatrix} 2-(-2) & 0 & 4 \\ 0 & 6-(-2) & 0 \\ 4 & 0 & 2-(-2) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

josta saadaan yhtälöryhmä

$$\begin{cases} 4x + 4z = 0 \\ 8y = 0 \\ 4x + 4z = 0, \end{cases}$$

jolloin ominaisavaruus $V_{\lambda_1} = \{(x, y, z) \in \mathbb{R} \mid x = -z,\ y = 0\}$ ja eräs ominaisvektori $\bar{x}_1 = (1, 0, -1)$. Ominaisarvon -2 geometrinen kertaluku on myös 1, sillä ominaisavaruuden dimensio $\dim(V_{\lambda_1}) = 1$.
Sijoittamalla ominaisarvo $\lambda_2 = 6$ yllä olevaan yhtälöön saadaan ominaivektori:

$$\begin{pmatrix} 2-6 & 0 & 4 \\ 0 & 6-6 & 0 \\ 4 & 0 & 2-6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

josta saadaan yhtälöryhmä

$$\begin{cases} -4x + 4z = 0 \\ 0 = 0 \\ 4x - 4z = 0, \end{cases}$$

jolloin ominaisavaruus $V_{\lambda_2} = \{(x, y, z) \in \mathbb{R} \mid x = z\}$ ja eräs ominaisvektori $\bar{x}_2 = (1, -1, 1)$. Myös esimerkiksi vektori $\bar{x} = (1, 0, 1)$ kuuluu ominaisavaruuteen V_{λ_2} ja on siis ominaisarvoon $\lambda_2 = 6$ kuuluva ominaisvektori. Vektorit $(1, -1, 1)$ ja $(1, 0, 1)$ ovat lineaarisesti riippumattomia toisistaan. Näin ollen $\dim(V_{\lambda_2}) = 2$, joten ominaisarvon 6 geometrinen kertaluku on 2.

2.7 Matriisien ominaisuuksista

Seuraavaksi esitellään muutamia matriisin ominaisuuksia, joita tarvitaan matriisihajotelmissa.

Määritelmä 2.30. Matriisit $A, B \in \mathcal{M}_n(K)$ ovat *similaariset*, jos on olemassa sellainen säännöllinen matriisi P, että $A = P^{-1}BP$.

Lause 2.31. *Similaarisilla matriiseilla on sama ominaisarvopolynomi.*

Todistus. Olkoon matriisi $A \in \mathcal{M}_n(\mathbb{R})$ similaarinen matriisin $B \in \mathcal{M}_n(\mathbb{R})$ kanssa eli $A = P^{-1}BP$, missä P on säännöllinen matriisi. Matriisin A ominaisarvopolymomi $\det(A - \lambda I)$ ja matriisin B ominaisarvopolynomi $\det(B - \lambda I)$ ovat seuraavan perusteella yhtäsuuret:

$$\begin{aligned} \det(A - \lambda I) &= \det(P^{-1}BP - \lambda P^{-1}IP) = \det(P^{-1}(B - \lambda I)P) \\ &= \det(P^{-1}) \det(B - \lambda I) \det(P) = \det(B - \lambda I) \end{aligned}$$

□

Määritelmä 2.32. Matriisi A on on *diagonalisoituva*, jos se on similaarinen diagonaalimatriisin kanssa eli on olemassa säännöllinen matriisi $P \in \mathcal{M}_n(K)$ ja sellaiset $\lambda_1, ..., \lambda_n \in K$, että

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n).$$

Lauseen 2.31 mukaan diagonalisoituvan matriisin A ominaisarvopolynomi $\det(A-\lambda I) = \det(P^{-1}AP-\lambda I) = \det(\mathrm{diag}(\lambda_1-\lambda, ..., \lambda_n-\lambda)) = (\lambda_1-\lambda)\cdots(\lambda_n-\lambda)$, joten $\lambda_1, ..., \lambda_n$ ovat matriisin A ominaisarvot.

Kerrataan seuraavaksi joitakin matriisin diagonalisoituvuuten liittyviä ominaisuuksia, joiden todistukset löytyvät esimerkiksi kirjasta [6].

Lause 2.33. *Oletetaan, että matriisin $A \in \mathcal{M}_n(\mathbb{R})$ erisuuret ominaisarvot ovat $\lambda_1, .., \lambda_s$, ja että lisäksi ominaisarvon λ_i algebrallinen kertaluku on k_i ja geometrinen kertaluku n_i. Silloin seuraavat ehdot ovat ekvivalentit:*

1. *Matriisi A on diagonalisoituva.*

2. *Ominaisarvon λ_i algebrallinen kertaluku on yhtä suuri kuin geometrinen kertaluku (eli $k_i = n_i$) kaikilla $i \in \{1, ..., s\}$.*

3. *Jokaiseen ominaisarvoon λ_i kuuluu k_i lineaarisesti riippumatonta ominaisvektoria.*

Esimerkki 2.34. Tutki, onko edellisen esimerkin matriisi

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{pmatrix}$$

diagonalisoituva. Lauseen 2.33 kohdan 2 perusteella voidaan sanoa, että matriisi A on diagonalisoituva, sillä ominaisarvon $\lambda_1 = -2$ algebralliset ja geometriset kertaluvut ovat samat eli $k_1 = n_1 = 1$ sekä vastaavasti ominaisarvon $\lambda_2 = 6$ kertaluvut ovat samat eli $k_2 = n_2 = 2$.

Tutkitaan vielä määritelmän 2.32 mukaan, onko matriisin A diagonalisoituva. Matriisin A ominaisvektorit ovat $\bar{x}_1 = (1, 0, -1)$, $\bar{x}_2 = (1, -1, 1)$ ja $\bar{x}_3 = (1, 0, 1)$. Muodostetaan matriisi P, missä pystyvektoreina ovat matriisin A ominaisvektorit:

$$P = (\bar{x}_1^T \mid \bar{x}_2^T \mid \bar{x}_3^T) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}.$$

Matriisin P käänteismatriisi

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

Seuraavaksi lasketaan

$$
\begin{aligned}
P^{-1}AP &= \begin{pmatrix} \frac{1}{2} & 0 & \frac{-1}{2} \\ 0 & -1 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \operatorname{diag}(-2, 6, 6).
\end{aligned}
$$

Määritelmän 2.32 mukaan $P^{-1}AP = \operatorname{diag}(\lambda_1, ..., \lambda_n)$. Esimerkissä 2.24 saatiin matriisin A ominaisarvoiksi $\lambda_1 = -2$ ja $\lambda_2 = 6$, joten matriisi A on diagonalisoituva.

Määritelmä 2.35. Neliömatriisi A on *idempotentti*, jos $A^2 = A$

Määritelmä 2.36. Matriisi $A \in \mathcal{M}_n(\mathbb{R})$ on *positiivisesti definiitti*, jos matriisi A on symmetrinen ja sen ominaisarvot ovat positiivisia ($\lambda_i > 0$). Jos symmetrisen matriisin A omainaisarvot ovat epänegatiivisia ($\lambda_i \geqslant 0$), matriisi A on *positiivisesti semidefiniitti.*

Vastaavasti määritellään *negatiivisesti definiitti* ($\lambda_i < 0$) ja *negatiivisesti semidefiniitti* ($\lambda_i \leqslant 0$) matriisi.

3 Matriisihajotelmista

Usein matriisiyhtälön ratkaiseminen on hidasta ja epäkäytännöllistä, joten tämän helpottamiseksi matriisit pyritään usein esittämään kahden tai useamman yksinkertaista muotoa olevan matriisin tulona. Tälläistä esitystä kutsutaan *matriisihajotelmaksi.*

Hajotelma helpottaa matriisiyhtälön ratkaisemisessa ja saattaa myös antaa hyödyllistä tietoa matriisista. Matriisihajotelmille on monia eri käyttötarkoituksia ja jokaiseen tilanteeseen sopii tietynlainen hajotelma, jonka vuoksi hajotelmia on monia erilaisia.

3.1 Matriisihajotelma asteen avulla

Aloitetaan perehtymällä lauseen 2.24 mukaiseen matriisihajotelmaan $A = BC$. Matriisi B muodostetaan matriisin A pystyriviavaruuden kannasta, joka saadaan matriisin A transpoosin A^T redusoidun porrasmuodon vaakariveistä. Eli jos vektorit $\bar{x}_1, ..., \bar{x}_k$ ovat matriisin A transpoosin A^T redusoidun porrasmuodon nollasta poikkeavat vaakarivit, matriisin B pystyvektorit ovat $\bar{b}_1 = \bar{x}_1^T$,...,$\bar{b}_k = \bar{x}_k^T$. Vektorit $\bar{b}_1, ..., \bar{b}_k$ muodostavat $P(A)$:n kannan. Matriisi C saadaan, kun lausutaan matriisin A pystyrivit tässä kannassa:

$$
\begin{aligned}
a_1^T &= c_1 b_1 + ... + c_n b_k \\
&\vdots \\
a_k^T &= c_1 b_1 + ... + c_n b_k.
\end{aligned}
$$

Esimerkki 3.1. Muodostetaan lauseen 2.24 mukainen matriisihajotelma $A = BC$ matriisille

$$A = \begin{pmatrix} 1 & 2 & 2 & 4 & 6 \\ 1 & 2 & 3 & 6 & 9 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix}.$$

Transponoidaan matriisi A, jolloin saadaan

$$A^T = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \\ 2 & 3 & 1 \\ 4 & 6 & 2 \\ 6 & 9 & 3 \end{pmatrix}.$$

Muutetaan matriisi A^T redusoiduksi porrasmatriisiksi. Lisätään ensimmäinen rivi toiseen riviin kerrottuna luvulla -2, kolmanteen riviin kerrottuna luvulla -2, neljänteen riviin kerrottuna luvulla -4 ja viidenteen riviin kerrottuna luvulla -6:

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 3 & 3 \end{pmatrix}.$$

Muokataan vaakarivien järjestystä niin, että nollarivi on alimpana, jolloin saadaan matriisi

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Lisätään toinen rivi kolmanteen riviin kerrottuna luvulla -2, neljänteen riviin kerrottuna luvulla -3 ja ensimmäiseen riviin kerrottuna luvulla -1, jolloin saadaan redusoitu porrasmatriisi

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

Matriisin A aste $r(A) = 2$, joten B on 3×2-matriisi ja C on 2×5-matriisi. Myös $r(B) = r(C) = 2$.

Muodostetaan matriisi B, kun redusoidun porrasmatriisin nollavektorista poikkeavat vaakarivit ovat matriisin B pystyrivit eli $\bar{b}_1 = (1, 0, -1)^T$ ja $\bar{b}_2 = (0, 1, 1)^T$, jolloin

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

Lausutaan matriisin A pystyrivit tässä kannassa:

$$\begin{aligned}(1,1,0)^T &= \bar{b}_1 + \bar{b}_2\\ (2,2,0)^T &= 2\bar{b}_1 + 2\bar{b}_2\\ (2,3,1)^T &= 2\bar{b}_1 + 3\bar{b}_2\\ (4,6,2)^T &= 4\bar{b}_1 + 6\bar{b}_2\\ (6,9,3)^T &= 6\bar{b}_1 + 9\bar{b}_2,\end{aligned}$$

ja muodostetaan kertoimista matriisi C:

$$C = \begin{pmatrix}1 & 2 & 2 & 4 & 6\\ 1 & 2 & 3 & 6 & 9\end{pmatrix}.$$

Tarkistamalla huomataan, että todella

$$A = \begin{pmatrix}1 & 2 & 2 & 4 & 6\\ 1 & 2 & 3 & 6 & 9\\ 0 & 0 & 1 & 2 & 3\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 2 & 4 & 6\\ 1 & 2 & 3 & 6 & 9\end{pmatrix} = BC$$

3.2 QR-hajotelma

Lineaarialgebrassa *QR-hajotelma* on matriisihajotelma, jossa matriisi A voidaan esittää ortogonaalisen matriisin Q ja yläkolmiomatriisin R tulona. QR-hajotelmaa käytetään mm. lineaaristen yhtälöryhmien ratkaisemiseen, matriisin ominaisarvojen etsimiseen, determinantin itseisarvon määrittämiseen sekä käänteismatriisin etsimiseen.

QR-hajotelman muodostamiseen on useita menetelmiä, joista yksi hyödyntää Gramin-Schmidtin ortogonalisointimenetelmää. Esitellään seuraavaksi Gramin-Schmidtin ortogonalisointimenetelmä: Olkoon $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_k\} \subseteq \mathbb{R}^n$. Konstruoidaan vektorit $\bar{y}_i$ rekursiivisesti:

$$\begin{aligned}\bar{y}_1 &= \bar{x}_1\\ \bar{y}_j &= \bar{x}_j - \sum_{i=1}^{j-1} a_{ji}\bar{y}_i, \text{ kun j} = 2, ..., \text{k},\end{aligned}$$

missä

$$a_{ji} = \begin{cases} 0, & \text{jos } \bar{\text{y}}_\text{i} = \bar{0},\\ \dfrac{\langle \bar{x}_j, \bar{y}_i\rangle}{|y_i|^2}, & \text{jos } \bar{\text{y}}_\text{i} \neq \bar{0}.\end{cases}$$

Vektorit $\bar{y}_1, \bar{y}_2, ..., \bar{y}_k$ ovat ortogonaaliset ja kaikille $j = 1, ..., k$ on voimassa $L(\bar{x}_1, ..., \bar{x}_j) = L(\bar{y}_1, ..., \bar{y}_j)$. Ortogonaalisesta joukosta saadaan ortonormaali poistamalla mahdolliset nollavektorit ja jakamalla ortogonaaliset vektorit pituudellaan.

Lause 3.2. *Olkoon $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ ($m \geqslant n$) sellainen matriisi, että matriisin A aste $r(A) = n$. Silloin on olemassa ortogonaalinen matriisi $Q \in \mathcal{M}_{m\times n}(\mathbb{R})$ ja yläkolmiomatriisi $R \in \mathcal{M}_n(\mathbb{R})$, että $A = QR$.*

Todistus. Olkoon matriisi $A = (\bar{a}_1| \ ... \ | \ \bar{a}_n)$, missä pystyrivit $\bar{a}_j^T \in \mathbb{R}^m$. Koska matriisin A aste $r(A) = n$, niin pystyrivit $\bar{a}_j^T$ ovat lineaarisesti riippumattomia. Gramin-Schmidtin ortogonalisointimenetelmällä saadaan ortonormaalit vektorit $\bar{y}_1, ..., \bar{y}_n$, missä $\bar{y}_i \in \mathbb{R}^m$, ja jolle $L(\bar{a}_1, ..., \bar{a}_j) = L(\bar{y}_1, ..., \bar{y}_j) \quad (j = 1, ..., n)$, missä $\{\bar{y}_1, ..., \bar{y}_j\}$ on ortonormaalikanta. Lauseen 2.18 mukaan vektorin $\bar{a}_j$ kantaesitys on

$$\bar{a}_j = \sum_{k=1}^{j} \langle \bar{a}_j, \bar{y}_k \rangle \bar{y}_k.$$

Olkoon matriisi $R = (r_{ij}) \in \mathcal{M}_n(\mathbb{R})$, missä $r_{ij} = \langle \bar{a}_j, \bar{y}_i \rangle$, kun $i \leqslant j$ ja $r_{ij} = 0$, kun $i > j$. Vektorin $\bar{a}_j$ kantaesityksen mukaan $A = QR = Q(\bar{r}_1|...|\bar{r}_n) = (Q\bar{r}_1|...|Q\bar{r}_n)$, missä matriisi $Q = (\bar{y}_1|...| \ \bar{y}_n)$. $\square$

Huomautus 3.3. Jos matriisi A on säännöllinen, niin hajotelma on yksikäsitteinen, jos matriisi R on positiivinen yläkolmiomatriisi. Lisäksi jos A on neliömatriisi, matriisi Q on ortogonaalinen eli $Q^T Q = QQ^T = I$.

QR-hajotelman matriisin Q pystyvektorit $\bar{q}_j^T = (q_{1j}, ..., q_{mj})$ saadaan ortonormalisoimalla matriisin A pystyvektorit $\bar{a}_j^T = (a_{1j}, ..., a_{mj})$. Matriisin A pystyvektorit ortogonalisoidaan Gramin-Schmidtin menetelmällä seuraavasti:

$$\begin{aligned}
\bar{q}_1' &= \bar{a}_1, \\
\bar{q}_2' &= \bar{a}_2 - \frac{\langle \bar{a}_2, \bar{q}_1 \rangle}{|\bar{q}_1|^2} \bar{q}_1, \\
&\vdots \\
\bar{q}_n' &= \bar{a}_n - \frac{\langle \bar{a}_n, \bar{q}_{n-1} \rangle}{|\bar{q}_{n-1}|^2} \bar{q}_{n-1} - ... - \frac{\langle \bar{a}_n, \bar{q}_1 \rangle}{|\bar{q}_1|^2} \bar{q}_1.
\end{aligned}$$

Nämä ortogonaaliset vektorit $\bar{q}_1', ..., \bar{q}_n'$ ortonormalisoidaan jakamalla vektorit pituudellaan, jolloin

$$\begin{aligned}
\bar{q}_1 &= \frac{\bar{q}_1'}{|\bar{q}_1'|}, \\
&\vdots \\
\bar{q}_n &= \frac{\bar{q}_n'}{|\bar{q}_n'|},
\end{aligned}$$

jotka ovat matriisin Q pystyvektorit. Matriisi $Q = (\bar{q}_1^T \mid \cdots \mid \bar{q}_n^T)$.

Yläkolmiomatriisi R muodostuu matriisin A pystyvektoreiden $\bar{a}_j^T = (a_{1j}, ..., a_{mj})$ ja matriisin Q pystyvektoreiden $\bar{q}_j^T = (q_{1j}, ..., q_{mj})$ sisätuloista seuraavasti:

$$R = \begin{pmatrix} \langle \bar{a}_1, \bar{q}_1 \rangle & \langle \bar{a}_2, \bar{q}_1 \rangle & \cdots & \langle \bar{a}_n, \bar{q}_1 \rangle \\ 0 & \langle \bar{a}_2, \bar{q}_2 \rangle & \cdots & \langle \bar{a}_n, \bar{q}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle \bar{a}_n, \bar{q}_n \rangle \end{pmatrix}.$$

Esimerkki 3.4. Muodosta QR-hajotelma matriisille

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

Ortogonalisoidaan matriisin A pystyvektoreiden joukko $\{\bar{x}_1, \bar{x}_2, \bar{x}_3\}$, missä $\bar{x}_1 = (1,1,0)$, $\bar{x}_2 = (1,0,1)$ ja $\bar{x}_3 = (0,1,1)$ Gramin-Schmidtin menetelmällä:

$$\begin{aligned}
\bar{y}_1 &= \bar{x}_1 = (1,1,0) \\
\bar{y}_2 &= \bar{x}_2 - \frac{\langle \bar{x}_2, \bar{y}_1 \rangle}{|\bar{y}_1|^2}\bar{y}_1 = \bar{x}_2 - \frac{\langle (1,0,1),(1,1,0)\rangle}{|(1,1,0)|^2}\bar{y}_1 = \bar{x}_2 - \frac{1}{2}\bar{y}_1 \\
&= (1,0,1) - \frac{1}{2}(1,1,0) = (\frac{1}{2}, -\frac{1}{2}, 1) \\
\bar{y}_3 &= \bar{x}_3 - \frac{\langle \bar{x}_3, \bar{y}_1 \rangle}{|\bar{y}_1|^2}\bar{y}_1 - \frac{\langle \bar{x}_3, \bar{y}_2 \rangle}{|\bar{y}_2|^2}\bar{y}_2 = \bar{x}_3 - \frac{1}{2}\bar{y}_1 - \frac{1}{3}\bar{y}_2 \\
&= (0,1,1) - \frac{1}{2}(1,1,0) - \frac{1}{3}(\frac{1}{2}, -\frac{1}{2}, 1) = (-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}).
\end{aligned}$$

Ortonormalisoidaan ortogonaaliset vektorit $\bar{y}_1$, $\bar{y}_2$ ja $\bar{y}_3$ jakamalla vektorit pituudellaan:

$$\begin{aligned}
\bar{y}_1' &= \frac{1}{\sqrt{2}}(1,1,0) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0) \\
\bar{y}_2' &= \frac{1}{\sqrt{3/2}}(\frac{1}{2}, -\frac{1}{2}, 1) = (\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}) \\
\bar{y}_3' &= \frac{\sqrt{3}}{2}(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}) = (-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}).
\end{aligned}$$

Näin ollen ortogonaalinen matriisi

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Yläkolmiomatriisi

$$R = \begin{pmatrix} \langle \bar{x}_1, \bar{y}_1' \rangle & \langle \bar{x}_2, \bar{y}_1' \rangle & \langle \bar{x}_3, \bar{y}_1' \rangle \\ 0 & \langle \bar{x}_2, \bar{y}_2' \rangle & \langle \bar{x}_3, \bar{y}_2' \rangle \\ 0 & 0 & \langle \bar{x}_3, \bar{y}_3' \rangle \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix}.$$

Tarkistamalla huomataan, että todella

$$QR = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = A$$

QR-hajotelmaa käytetään esimerkiksi lineaaristen yhtälöryhmien ratkaisemisessa. Lineaarinen yhtälöryhmä

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

voidaan esittää matriiseilla muodossa $A\bar{x} = \bar{b}$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

QR-hajotelmaa voidaan hyödyntää ratkaisussa seuraavasti: Matriisi Q on ortogonaalinen, joten $Q^TQ = I$.

$$\begin{aligned} A\bar{x} &= b \\ \Leftrightarrow QR\bar{x} &= b \\ \Leftrightarrow Q^TQR\bar{x} &= Q^Tb \\ \Leftrightarrow R\bar{x} &= Q^Tb, \end{aligned}$$

joka on usein helpompi ratkaista kuin matriisin A käänteismatriisi.

3.3 LU-hajotelma

LU-hajotelmaa käytetään numeerisessa analyysissa ja lineaarialgebrassa lineaaristen yhtälöryhmien ratkaisemiseen sekä käänteismatriisien muodostamiseen ja determinantin laskemiseen. Tietokoneet useimmiten käyttävät LU-hajotelmaa lineaariyhtälöiden ratkaisemisessa. LU-hajotelma muodostetaan hajottamalla säännöllinen matriisi A oikeanlaisilla pysty- ja/tai vaakarivien muunnoksilla kahteen tekijään: alakolmiomatriisiin L ja yläkolmiomatriisiin U.

Lause 3.5. *Matriisi $A \in \mathcal{M}_n(\mathbb{R})$ on säännöllinen, jos ja vain jos*

$$A = LU,$$

missä $L \in \mathcal{M}_n(\mathbb{R})$ alakolmiomatriisi, jonka diagonaalilla on pelkkiä ykkösiä, ja $U \in \mathcal{M}_n(\mathbb{R})$ yläkolmiomatriisi.

Katso lauseen todistus viitteestä [3].

3×3-matriisin A LU-hajotelma on seuraavanlainen:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

LU-hajotelman muodostaminen aloitetaan yläkolmiomatriisista U. Tämä matriisi muodostetaan matriisista A Gaussin eliminaatioalgoritmilla eli matriisin A pysty- ja/tai vaakarivejä lisätään muihin pysty- ja/tai vaakariveihin kerrottuna jollakin luvulla kunnes päädytään yläkolmiomatriisiin. Matriisin $L = (l_{ij})$ $(i > j)$ alkiot tulevat matriisin U muodostamisessa käytetyistä kertoimista niin, että l_{ij} on sen kertoimen vastaluku, jolla rivi j on kerrottu lisättäessä riviin i.

Esimerkki 3.6. Muodostetaan LU-hajotelma matriisille

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

Aloitetaan lisäämällä matriisin A ensimmäinen rivi toiseen riviin, jolloin saadaan matriisi

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

Seuraavaksi lisätään ensimmäinen rivi kerrottuna luvulla -1 kolmanteen riviin, jolloin saadaan yläkolmiomatriisi

$$U = \begin{pmatrix} -1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Muodostetaan alaolmiomatriisi edellisistä kertoimista, jolloin

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

Todella matriisin U saamiseksi ensimmäinen rivi $(j = 1)$ lisättiin toiseen riviin $(i = 2)$, joten matriisin L alkio $l_{21} = -1$. Samoin ensimmäinen rivi $(j = 1)$ lisättiin kolmanteen riviin $(i = 3)$ kerrottuna luvulla -1, joten alkio $l_{31} = 1$. Koska toista riviä ei lisätty kolmanteen riviin, alkio $l_{32} = 0$.

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} = A$$

Lause 3.7. *Jos $A = LU$ on neliömatriisi, niin*

$$\det(A) = \det(LU) = \det(L)\det(U) = \det(U) = u_{11}u_{22}\cdots u_{nn}.$$

Todistus. Determinantin perusominaisuuksien (kohta 4) mukaan hajotelman determinantti on sama kuin tekijöiden determinanttien tulo eli $\det(LU) = \det(L)\det(U)$. Määritelmän 2.10 mukaan matriisin L determinantti on diagonaalin alkioiden tulo, joka siis on 1. Näin ollen matriisin A determinantti on sama kuin matriisin U determinantti, joka on myös diagonaalin alkioiden tulo. Siispä $\det(A) = u_{11}u_{22}\cdots u_{nn}$. □

3.4 Spektraalihajotelma ja -esitys

Spektraalihajotelmassa matriisi hajotetaan kanoniseen muotoon, missä matriisi esitetään ominaisarvojen ja ominaisvektoreiden avulla. Tästä hajotelmasta käytetään myös nimitystä *ominaisarvohajotelma*. Vain diagonalisoituvilla matriiseilla on spektraalihajotelma.

Määritelmä 3.8. Matriiseja $E_1, E_2, ..., E_m$ sanotaan *ortogonaalisesti idempotenteiksi*, jos matriisit $E_1, E_2, ..., E_m$ ovat *idempotentteja* ja $E_iE_j = 0$, kun $i \neq j$ $(i, j = 1, ..., m)$. Lisäksi jos $E_1 + ... + E_m = I$, niin sanotaan, että $\{E_1, E_2, ..., E_m\}$ on *täysi joukko ortogonaalisia idempotentteja*.

Esimerkki 3.9. Olkoot matriisit $E_1, ..., E_m$ ortogonaalisia idempotentteja.

1. Osoitetaan, että jos $\{E_1, ..., E_m\}$ on täysi joukko ortogonaalisia idempotentteja, niin silloin myös $\{PE_1P^{-1}, ..., PE_mP^{-1}\}$ on täysi joukko ortogonaalisia idempotentteja, kun P on säännöllinen matriisi.

   Matriisit $PE_1P^{-1}, ..., PE_mP^{-1}$ ovat idempotentteja, sillä

   $$(PE_iP^{-1})^2 = (PE_iP^{-1})(PE_iP^{-1}) = PE_iE_iP^{-1} = PE_iP^{-1}.$$

   Edelleen matriisit ovat ortogonaalisesti idempotentteja, sillä

   $$(PE_iP^{-1})(PE_jP^{-1}) = PE_iE_jP^{-1} = 0.$$

   Lisäksi $\{PE_1P^{-1}, ..., PE_mP^{-1}\}$ on täysi joukko ortogonaalisia idempotentteja, sillä

   $$PE_1P^{-1} + ... + PE_nP^{-1} = P(E_1 + ... + E_n)P^{-1} = PP^{-1} = I.$$

2. Samoin matriisit $E_{11}, ..., E_{nn}$ muodostavat täyden joukko ortogonaalisia idempotentteja, missä matriisi $E_{kk} = (e_{ij}) \in \mathcal{M}_n(\mathbb{R})$ on se matriisi, jonka alkio e_{kk} on yksi ja muut alkiot ovat nollia. Eli esimerkiksi $E_{11} = diag(1, 0, ..., 0)$, $E_{22} = diag(0, 1, 0, ..., 0)$ ja $E_{nn} = diag(0, ..., 0, 1)$. Osoitetaan, että $\{E_{11}, E_{22}, \ldots, E_{nn}\}$ on täysi joukko ortogonaalisia idempotetteja. Matriisit $E_{11}, ..., E_{nn}$ ovat ortogonaalisia idempotentteja, sillä

   $$E_{ii}^2 = E_{ii} \text{ ja } E_{ii}E_{jj} = 0.$$

   Lisäksi $\{E_{11}, ..., E_{nn}\}$ on täysi joukko ortogonaalisia idempotentteja, sillä

   $$E_{11} + ... + E_{nn} = I.$$

Määritelmä 3.10. (i) Matriisin A *spektraalihajotelma* on $A = \sum_{i=1}^{s} \lambda_i E_i$, missä $\lambda_1, ..., \lambda_s$ ovat matriisin A erisuuret ominaisarvot ja $\{E_1, ..., E_n\}$ on täysi joukko ortogonaalisia idempotentteja.

(ii) Matriisin A *spektraaliesitys* on $A = \sum_{i=1}^{n} \lambda_i E_i$, missä $\lambda_1, ..., \lambda_n$ ovat matriisin A ominaisarvot ja $\{E_1, ..., E_s\}$ on täysi joukko ortogonaalisia idempotentteja.

Huomautus 3.11. Spektraaliesityksestä saadaan spektraalihajotelma, kun yhdistetään yhtä suuret ominaisarvot.

Lause 3.12. *Jos $A \in \mathcal{M}_n(\mathbb{R})$, seuraavat ehdot ovat ekvivalentit:*

(i) matriisi A on diagonalisoituva,

(ii) matriisilla A on spektraaliesitys ja

(iii) matriisilla A on spektraalihajotelma.

Todistus. Osoitetaan, että diagonalisoituvalla matriisilla on spektraaliesitys. Koska matriisi A on diagonalisoituva, se on similaarinen diagonaalimatriisin kanssa eli $A = PDP^{-1}$, missä P on matriisin A ominaisvektoreista muodostuva säännöllinen matriisi. Matriisi $D = \text{diag}(\lambda_1, ..., \lambda_n) = \lambda_1 E_{11} + ... + \lambda_n E_{nn}$, missä λ_i ovat matriisin A ominaisarvot ja matriisit $\{E_{11}, ..., E_{nn}\}$ ovat esimerkin 3.9 täysi joukko ortogonaalisia idempotentteja.

Näin ollen

$$\begin{aligned} A &= P(\lambda_1 E_{11} + ... + \lambda_n E_{nn})P^{-1} \\ &= \lambda_1 P E_{11} P^{-1} + ... + \lambda_n E_{nn} P^{-1} \\ &= \lambda_1 E_1 + ... + \lambda_n E_n, \end{aligned}$$

kun $E_i = PE_{ii}P^{-1}$. Koska matriisit E_{ii} muodostavat täyden joukon ortogonaalisia idempotentteja, niin samoin muodostavast matriisit E_i. Näin saatiin matriisin A spektraaliesitys.

Huomautuksen 3.11 mukaan spektraaliesityksestä saadaan spektraalihajotelma yhdistämällä yhtä suuret ominaisarvot, jolloin

$$A = \lambda_1 E_1 + ... + \lambda_s E_s,$$

missä $\lambda_1, ..., \lambda_s$ ovat matriisin A erisuuret ominaisarvot.

Vielä olisi näytettävä, että jos matriisilla A on spektraalihajotelma, niin silloin se on diagonalisoituva. Tämä ei kuitenkaan jatkon kannalta ole erityisen mielenkiintoinen. Sen vuoksi todistus sivuutetaan tässä esityksessä, mutta se on luettavissa esimerkiksi luentomonisteesta [2]. □

Seuraavaksi käydään läpi esimerkki spektraalihajotelman ja -esityksen muodostamisesta.

Esimerkki 3.13. Muodosta spektraalihajotelma matriisille

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

Lasketaan matriisin A määritelmän 2.26 mukaiset ominaisarvot:

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

Näin ollen matriisilla A on kolme ominaisarvoa: $\lambda_1 = 0$, $\lambda_2 = 1$ ja $\lambda_3 = 3$.

Etsitään ominaisarvolle $\lambda_1 = 0$ ominaisvektori $\bar{x}_1$ ratkaisemalla yhtälö $(A - \lambda_1 I)\bar{x}_1 = \bar{0}$:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

josta saadaan yhtälöryhmä

$$\begin{cases} x + y = 0 \\ x + 2y + z = 0 \\ y + z = 0, \end{cases}$$

jolloin $x = z = -y$ ja eräs ominaisvektori $\bar{x}_1 = (1, -1, 1)$. Samoin etsitään ominaisvektorit ominaisarvoille $\lambda_2 = 1$ ja $\lambda_3 = 3$: $\bar{x}_2 = (1, 0, -1)$ ja $\bar{x}_3 = (1, 2, 1)$.

Matriisin A diagonalisointia varten muodostetaan ominaisvektoreista matriisi

$$P = (\bar{x}_1^T \mid \bar{x}_2^T \mid \bar{x}_3^T) = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix}.$$

Käytetään esimerkissä 2.9 matriisille P etsittyä käänteismatriisia

$$P^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

Koska

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix} = PDP^{-1},$$

niin matriisi A on similaarinen diagonaalimatriisin kanssa ja matriisi A on diagonalisoituva.

Spektraalihajotelma

$$\begin{aligned} A &= P(\lambda_1 E_{11} + \lambda_2 E_{22} + \lambda_3 E_{33})P^{-1} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} \left(0 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix} \\ &= 1 \cdot \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + 3 \cdot \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix} = \lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3 \end{aligned}$$

Tässä esimerkissä spektraalihajotelma ja -esitys ovat samat, sillä matriisin A kaikki ominaisarvot ovat erisuuria.

Spektraaliesitykset helpottavat matriiseilla laskemista. Etenkin matriisin potenssin laskeminen on tehokkaampaa. Nimittäin

$$A^k = (\sum_{i=1}^{n} \lambda_i E_i)^k = \sum_{i=1}^{n} \lambda_i^k E_i = \lambda_1^k E_1 + ... + \lambda_n^k E_n,$$

koska $\{E_1, \ldots, E_n\}$ on ortogonaalinen joukko idempotentteja matriiseja sekä $E_i^2 = I$ ja $E_i E_j = 0$ kaikilla $i \neq j$.

Esimerkki 3.14. Laske edellisen esimerkin matriisille A^{100}. Edellisessä esimerkissä muodostimme matriisille A spektraalihajotelman

$$A = 1 \cdot \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + 3 \cdot \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

Nyt

$$\begin{aligned} A^{100} &= 1 \cdot \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + 3 \cdot \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix} \\ &= 1^{100} \cdot \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + 3^{100} \cdot \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{pmatrix} \\ &= \begin{pmatrix} \frac{3+3^{100}}{6} & 3^{99} & \frac{-3+3^{100}}{6} \\ 3^{99} & 2 \cdot 3^{99} & 3^{99} \\ \frac{-3+3^{100}}{6} & 3^{99} & \frac{3+3^{100}}{6} \end{pmatrix}. \end{aligned}$$

3.5 Singulaariarvohajotelma

Tässä luvussa perehdytään jälleen yhteen matriisihajotelmaan, joka antaa arvokasta tietoa matriisista. Yksinkertaisuuden vuoksi pysymme edelleen matriiseissa, joiden alkiot ovat reaalilukuja. Kappaleen 3.2 QR-hajotelmassa näimme, että kerrottaessa matriisi toiselta puolelta ortogonaalisella matriisilla, saadaan yläkolmiomatriisi. Jos toiselta puolelta kertomalla saadaan tuotettua yläkolmiomatriisi, miten yksinkertainen matriisi saadaan, jos kerrotaan matriisi myös toiselta puolelta mahdollisesti erilaisella ortogonaalisella matriisilla? Vastaus kysymykseen on, että näin muodostuu samaan aikaan sekä yläkolmiomatriisi että alakolmiomatriisi, eli diagonaalimatriisi. Kyseinen hajotelma on *singulaariarvohajotelma*, josta käytetään myös lyhennettä SVD (singular value decomposition).

Määritelmä 3.15. Matriisin $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ *singulaariarvoja* ovat

$$\sigma_i = \sqrt{\lambda_i} \quad (i = 1, ..., k),$$

missä $k = r(A)$ ja $\lambda_1, ..., \lambda_k$ ovat matriisin $A^T A$ positiiviset ominaisarvot.

Matriisien $A^T A$ ja AA^T ominaisarvot kertalukuineen ovat samat ja niitä on siis $r(A) = k$ kappaletta. Lisäksi matriisit $A^T A$ ja AA^T ovat diagonalisoituvia. Näiden tulosten todistukset löytyvät esimerkiksi lähteestä [2].

Lause 3.16. *Matriisin AA^T ominaisarvoon λ kuuluva ominaisvektori on*

$$\bar{u}_i = \frac{1}{\sqrt{\lambda}} A\bar{v}_i \quad (i = 1, ..., s),$$

missä vektori $\bar{v}_i$ on matriisin A^TA ominaisarvoon λ kuuluva ominaisvektori.

Todistus. Olkoon matriisin A^TA positiivinen ominaisarvo λ ja siihen kuuluvien ominaisvektorien ortogonaalinen joukko $\{\bar{v}_1, ..., \bar{v}_s\}$ ($\subseteq \mathbb{R}^m$). Merkitään

$$\bar{u}_i = \frac{1}{\sqrt{\lambda}} A\bar{v}_i \quad (\in \mathbb{R}^n) \quad (i = 1, ..., s).$$

Vektorit $\bar{u}_i$ ovat matriisin AA^T ominaisarvoon λ kuuluvia ominaisvektoreita, sillä

$$AA^T\bar{u}_i = \frac{1}{\sqrt{\lambda}} AA^TA\bar{v}_i = \frac{1}{\sqrt{\lambda}} A(\lambda\bar{v}_i) = \frac{\lambda}{\sqrt{\lambda}} A\bar{v}_i = \sqrt{\lambda}A\bar{v}_i = \lambda\bar{u}_i.$$

□

Lause 3.17. *Matriisin $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ singulaariarvohajotelma on*

$$A = U\Sigma V^T,$$

missä $U \in \mathcal{M}_m(\mathbb{R})$ ja $V \in \mathcal{M}_n(\mathbb{R})$ ovat ortogonaalisia matriiseja, sekä $D = \mathrm{diag}(\sigma_1, \cdots, \sigma_k)$ ja matriisi

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_{m\times n}(\mathbb{R}),$$

kun matriisin A singulaariarvot ovat $\sigma_1 \geqslant \cdots \geqslant \sigma_k$.

Huomaa, että matriisin Σ nollamatriisit valitaan niin, että matriisi on kokoa $m \times n$, kun matriisi D on aina neliömatriisi. Myös matriisit U ja V ovat siis neliömatriiseja. Katso lauseen todistus viitteestä [2].

Singulaariarvohajotelma muodostetaan seuraavin vaihein:

1. Muodosta matriisit A^TA ja AA^T.

2. Laske matriisin A^TA ominaisarvot $\lambda_1, ..., \lambda_n$ ja numeroi ne seuraavasti

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k > 0 \quad \text{ja} \quad \lambda_{k+1} = \cdots = \lambda_n = 0.$$

   Lisäksi laske matriisin A singulaariarvot $\sigma_1, ..., \sigma_k$, missä $\sigma_i = \sqrt{\lambda_i}$.

3. Etsi matriisin A^TA ominaisvektorit $\bar{v}_i$ kullekin ominaisarvolle λ_i $(i = 1, ..., n)$. Muodosta vektoreista ortonormaalikanta $\{\bar{v}_1, ..., \bar{v}_n\}$.

4. Laske matriisin AA^T ominaisarvot ja etsi matriisin AA^T ominaisvektorit $\bar{u}_i$ kullekin ominaisarvolle λ_i $(i = 1, ..., k)$, kun

$$\bar{u}_i = \frac{1}{\sqrt{\lambda_i}} A\bar{v}_i = \frac{1}{\sigma_i} A\bar{v}_i \quad (i = 1, ..., k).$$

   Muodosta vektoreista ortonormaalikanta $\{\bar{u}_1, ..., \bar{u}_k\}$.

5. Täydennä ortonormaalikanta $\{\bar{u}_1, ..., \bar{u}_k\}$ koko avaruuden $\mathbb{R}^n$ ortonormaaliksi kannaksi $\{\bar{u}_1, ..., \bar{u}_n\}$ etsimällä ominaisvektorit $\bar{u}_i$, kun $i = k+1, ..., n$.
6. Muodosta matriisit $U = \big(\bar{u}_1 | \cdots | \bar{u}_m\big) \in \mathcal{M}_m(\mathbb{R})$, $V = \big(\bar{v}_1 | \cdots | \bar{v}_n\big) \in \mathcal{M}_n(\mathbb{R})$ ja $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_{m\times n}(\mathbb{R})$, missä $D = \mathrm{diag}(\sigma_1, ..., \sigma_k)$. Lisäksi muodosta matriisista V matriisi V^T.

Esimerkki 3.18. Muodosta singulaariarvohajotelma matriisille

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

Muodostetaan aluksi matriisit A^TA ja AA^T, kun

$$A^T = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}:$$

$$A^TA = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \text{ ja}$$
$$AA^T = \begin{pmatrix} -2 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Lasketaan matriisin A^TA ominaisarvot ratkaisemalla karakteristisen polynomin nollakohdat:

$$\det(A^TA - \lambda I) = \lambda^2 - 6\lambda + 8 = 0$$

Yhtälön ratkaisuna saadaan ominaisarvot $\lambda_1 = 4$ ja $\lambda_2 = 2$, jolloin singulaariarvot ovat $\sigma_1 = \sqrt{4} = 2$ ja $\sigma_2 = \sqrt{2}$.

Kun lasketaan matriisin A^TA ominaisarvoille $\lambda_1 = 4$ ja $\lambda_2 = 2$ ominaisvektorit, kuten esimerkissä 2.29, saadaan

$$\bar{v}_1 = (1, 0)^T \text{ ja}$$
$$\bar{v}_2 = (0, -1)^T.$$

Huomaa, että ominaisvektoreiden tulee olla ortonormaaleja, eli niiden pituuksien tulee olla yksi ja ne ovat keskenään ortogonaalisia.

Lasketaan matriisin AA^T ominaisarvot ratkaisemalla karakteristisen polynomin nollakohdat:

$$\det(AA^T - \lambda I) = -\lambda^3 + 6\lambda^2 - 8\lambda = 0$$

Yhtälön ratkaisuna saadaan ominaisarvot $\lambda_1' =$, $\lambda_2' = 2$ ja $\lambda_3' = 0$.

Muodostetaan vektorit $\bar{u}_i = \frac{1}{\sigma_i} A\bar{v}_i$

$$\bar{u}_1 = \frac{1}{2}\begin{pmatrix} -2 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

$$\bar{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -2 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Täydennetään vektoreiden $\bar{u}_i$ kanta etsimällä ominaisarvolle $\lambda_3' = 0$ ortonormaali ominaisvektori, joka on $\bar{u}_3 = \frac{1}{\sqrt{2}}(0, 1, 1)^T$.

Näin ollen saadaan matriisit

$$U = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \text{ ja}$$

$$V^T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Lisäksi $D = \operatorname{diag}(2, \sqrt{2})$, joten

$$\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}.$$

Tarkistamalla huomataan, että todella

$$U\Sigma V^T = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} = A.$$

3.5.1 Yleistetty käänteismatriisi

Singulaarisella matriisilla ei ole käänteismatriisia, mutta käänteismatriisin käsite voidaan yleistää *yleistetyksi käänteismatriisiksi*, joka voidaan määrittää kaikille matriiseille.

Määritelmä 3.19. Matriisin $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ yleistetty käänteismatriisi $A^+ \in \mathcal{M}_{n\times m}(\mathbb{R})$ on se yksikäsitteinen matriisi $X \in \mathcal{M}_{n\times m}(\mathbb{R})$, joka toteuttaa seuraavat ehdot:

1. $(AX)^T = AX$,
2. $(XA)^T = XA$,
3. $AXA = A$,
4. $XAX = X$.

Näiden tulosten todistukset löytyvät esimerkiksi lähteestä [2].

Lause 3.20. *Jos matriisin A singulaariarvohajotelma on $A = U\Sigma V^T$, missä*

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

niin yleistetty käänteismatriisi $A^+ = V\Sigma^+U^T$*, missä*

$$\Sigma^+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Todistus. Osoitetaan, että määritelmän 3.19 ehdot toteutuvat, kun matriisin $A = U\Sigma V^T$ yleistetty käänteismatriisi $A^+ = V\Sigma^+U^T$.

Koska matriisit U ja V ovat ortogonaalisia, matriisi $U^TU = I = UU^T$ ja $V^TV = I = VV^T$. Koska määritelmän 3.19 ehdot toteutuvat eli $(\Sigma\Sigma^+)^T = \Sigma\Sigma^+$, $(\Sigma^+\Sigma)^T = \Sigma^+\Sigma$, $\Sigma\Sigma^+\Sigma = \Sigma$ ja $\Sigma^+\Sigma\Sigma^+ = \Sigma^+$, niin Σ^+ on matriisin Σ yleistetty käänteismatriisi.

Näytetään seuraavaksi, että määritelmän 3.19 ehdot ovat voimassa matriisille $A^+ = V\Sigma^+U^T$:

1. $(AA^+)^T = ((U\Sigma V^T)(V\Sigma^+U^T))^T = (U\Sigma\Sigma^+U^T)^T = U\Sigma\Sigma^+U^T$
   $= (U\Sigma V^T)(V\Sigma^+U^T) = AA^+$
2. $(A^+A)^T = ((V\Sigma^+U^T)(U\Sigma V^T))^T = (V\Sigma^+\Sigma V^T)^T = V\Sigma^+\Sigma V^T$
   $= (V\Sigma^+U^T)(U\Sigma V^T) = A^+A$
3. $AA^+A = (U\Sigma V^T)(V\Sigma^+U^T)(U\Sigma V^T) = U\Sigma\Sigma^+\Sigma V^T = U\Sigma V^T = A$
4. $A^+AA^+ = (V\Sigma^+U^T)(U\Sigma V^T)(V\Sigma^+U^T) = V\Sigma^+\Sigma\Sigma^+U^T = V\Sigma^+U^T = A^+$

Ehdot toteutuvat, joten matriisi $A^+ = V\Sigma^+U^T$ on matriisin $A = U\Sigma V^T$ yleistetty käänteismatriisi. □

Huomaa, että A^+ ja Σ^+ ovat kooltaan samat kuin matriisi A^T ja että $D^{-1} = \mathrm{diag}(\sigma_1^{-1}, ..., \sigma_k^{-1})$.

Esimerkki 3.21. Muodostetaan lauseen 3.20 mukainen yleistetty käänteismatriisi A^+ edellisen esimerkin matriisille

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

$$\begin{aligned} A^+ &= V\Sigma^+U^T \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \end{aligned}$$

3.6 Polaarihajotelma

Neliömatriisi A voidaan hajottaa matriisituloksi HW, missä W on unitaarinen matriisi ja matriisi H on positiivisesti semidefiniitti. Molemmat matriisit ovat neliömatriiseja ja saman kokoisia keskenään. Tällaista hajotelmaa kutsutaan *polaarihajotelmaksi.*

Tähän asti matriisihajotelmat on rajattu vain reaalisiin matriiseihin, mutta polaarihajotelmassa kompleksiluvuilla on oleellinen rooli, joten perehdytään kompleksilukuihin seuraavaksi.

3.6.1 Kompleksiluvut

Tämä luku pohjautuu Weissteinin julkaisuun *Complex Number* [7] ja Kivelän julkaisuun *Kompleksiluvut* [1]. Lukualueen laajenemisen luonnollisista luvuista kokonaislukujen ja rationaalilukujen kautta reaalilukuihin oletetaan johtuvan tarpeesta löytää ratkaisu yhä uusille yhtälötyypeille. Myös kompleksiluvut ovat syntyneet halusta löytää yhtälölle $x^2 + 1 = 0$ ratkaisu, jota ei löydy reaalilukujen joukosta.

Suoraviivainen ratkaisu ongelmaan on ottaa käyttöön uusi ns. luku i, jolla on ominaisuus $i^2 = -1$, jolloin yhtälölle $x^2 + 1 = 0$ saadaan kaksi ratkaisua, i ja $-i$. Tämä ratkaisu jättää kuitenkin avoimeksi, mikä i oikeastaan on.

Kompleksilukujen määrittelyssä otetaan lähtökohdaksi xy-taso

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\},$$

jota aletaan kutsua *kompleksitasoksi* tai *kompleksilukujoukoksi* ja sille käytetään symbolia $\mathbb{C}$. Kutsutaan joukon alkioita, $z = (x, y)$, kompleksiluvuiksi.

Määritelmä 3.22. Joukon $\mathbb{C}$ alkioille määritellään yhteenlasku ja kertolasku seuraavasti:

$$\begin{aligned} z_1 + z_2 &= (x_1, y_1) + (x_2, y_2) = (x_1 + x_2,\ y_1 + y_2) \text{ ja} \\ z_1 z_2 &= (x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2,\ x_1y_2 + x_2y_1). \end{aligned}$$

Alkiot (x_1, y_1) ja (x_2, y_2) ovat yhtäsuuria joukossa $\mathbb{R}^2$, jos ja vain jos $x_1 = x_2$ ja $y_1 = y_2$. Näin ollen kompleksiluvuilla voidaan laskea samoilla laskusäännöillä kuin reaaliluvuilla. On olemassa kompleksiluvut *nolla* (0,0) ja *ykkönen* (1,0), jotka käyttäytyvät kuten reaaliluvut 0 ja 1. Erityisasemassa on kompleksiluku (0,1), joka nimetään *imaginaariyksiköksi* ja merkitään sitä symbolilla i. Kertolaskusäännön mukaan imaginaariyksikön tulo itsensä kanssa on (-1,0), jolloin $i^2 = -1$.

Yhdistetään reaalilukujoukko $\mathbb{R}$ ja kompleksitaso $\mathbb{C}$ samaistamalla $x \in \mathbb{R}$ ja $(x, 0) \in \mathbb{C}$, jolloin $x = (x, 0)$. Yhteenlasku- ja kertolaskusääntöjen mukaan

$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1)(y, 0) = x + yi.$$

Määritelmä 3.23. Kompleksiluku z on muotoa

$$z = x + yi,$$

jossa on *reaaliosa* Re $z = x$ ja *imaginaariosa* Im $z = y$.

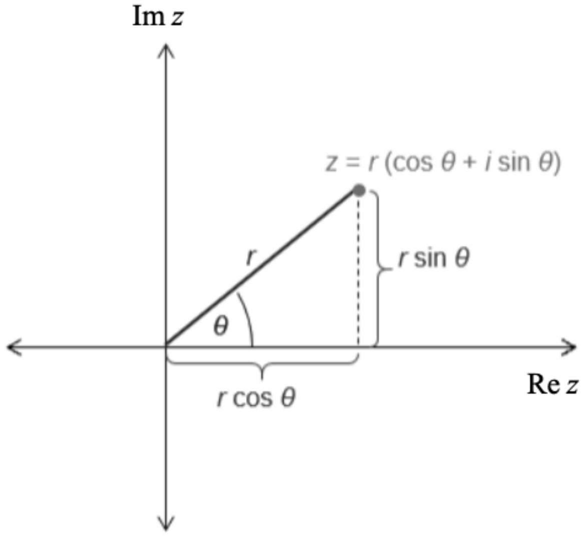
Määritelmä 3.24. Kompleksiluvun z *kompleksikonjugaatti* on

$$\bar{z} = x - yi.$$

Määritelmä 3.25. Jokaisella kompleksiluvulla $z = x + yi$ on polaarimuoto

$$z = r\cos\theta + ir\sin\theta = r(\cos\theta + i\sin\theta) = re^{i\theta},$$

missä $r = |z|$ ja θ on napakulma (Kuva 1).



Kuva 1: Kompleksiluvun z polaariesitys.

Jos sovelletaan kompleksilukujen polaarimuotoa 1×1-matriisiin, $e^{i\theta}$ olisi ortonormaali matriisi ja $r \geqslant 0$ olisi positiivisesti semidefiniitti matriisi. Kun laajennetaan tätä ajatusta $n \times n$-matriisiin, saadaan unitaarinen matriisi W ja positiivisesti semidefiniitti matriisi H, jolloin polaarihajotelma $A = HW$.

Kompleksisilla matriiseilla on joitakin ominaisuuksia, joita reaalisilla matriiseilla ei ole. Seuraavaksi käydään läpi muutamia näistä.

Määritelmä 3.26. Matrisiin $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{C})$ *adjungoitu matriisi* on $A^* = \bar{A}^T = (\bar{a}_{ji}) \in \mathcal{M}_{n\times m}(\mathbb{C})$. Lisäksi jos $A^* = A$, matriisi A on *itseadjungoitunut.*

Lisäksi $(AB)^* = B^*A^*$.

Huomautus 3.27. Reaalisille matriiseille adjungoitu matriisi on sama kuin matriisin transpoosi eli $A^* = A^T$.

Määritelmä 3.28. Matriisi $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ on *unitaarinen*, jos sen pystyrivit muodostavat ortonormaalin joukon eli $A^*A = I$. Tällöin myös vaakarivit muodostavat ortonormaalin joukon, sillä ehdot $A^*A = I$ ja $AA^* = I$ ovat yhtäpitävät.

Huomautus 3.29. Reaalisille matriiseille unitaarisuus ja ortogonaalisuus ovat ekvivalentit.

Matriisin polaarihajotelma saadaan singulaariarvohajotelmasta seuraavasti.

Lause 3.30. *Jos matriisin singulaariarvohajotelma $A = U\Sigma V^*$, matriisin polaarihajotelma $A = HW$, missä $H = U\Sigma U^*$ ja $W = UV^*$.*

Todistus. Osoitetaan, että $H = U\Sigma U^*$ on positiivisesti semidefiniitti. Koska $H^* = (U\Sigma U^*)^* = U\Sigma^* U^* = U\Sigma U^* = H$, niin matriisi H on itseadjungoitu, ja lisäksi matriisien H ja Σ ominaisarvot ovat samat ($\geqslant 0$) lauseen 2.26 mukaan, joten matriisi H on positiivisesti semidefiniitti. Osoitetaan vielä, että matriisi $W = UV^*$ on unitaarinen. Koska $WW^* = (UV^*)(UV^*)^* = UV^*VU^* = I$, niin matriisi W on unitaarinen. $\square$

Polaarihajotelma voi myös olla muotoa $A = U\Sigma V^* = (UV^*)(U\Sigma U^*)$, jolloin $UV^* = W'$ ja $U\Sigma U^* = H'$, eli $A = W'H'$, missä matriisi W' on unitaarinen ja matriisi H' on positiivisesti semidefiniitti. Tällöin $W \neq W'$ ja $H \neq H'$, vaikka matriisi A on sama.

Esimerkki 3.31. Muodostetaan matriisille

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

polaarihajotelma, kun matriisin singulaariarvohajotelma on

$$A = U\Sigma V^* = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Muodostetaan ensin matriisit H ja W:

$$H = U\Sigma U^* = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \sqrt{\frac{2}{5}} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$
$$W = UV^* = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix}.$$

Näin ollen

$$HW = \sqrt{\frac{2}{5}} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \frac{1}{\sqrt{10}} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = A.$$

Lähteet

[1] Kivelä, Simo K.: *Kompleksiluvut.*
http://matta.hut.fi/matta/kompleksiluvut/cluvut.pdf (5.9.2018).

[2] Koppinen, Markku: *Matriisilaskenta*, luentomoniste, Turun yliopisto, 2012.

[3] Miller, G. A.: *On the History of Determinants.*
http://www.jstor.org/stable/2299112 (27.4.2020)

[4] Olver, Peter J., Shakiban, Chehrzad: *Applied Linear Albegra, Second Edition*, Springer, 2018.

[5] Shores, Thomas S.: *Applied Linear Algebra and Matrix Analysis, Second Edition*, Springer, 2018.

[6] Strang, Gilbert: *Introduction to Linear Algebra, Fifth Edition*, Cambridge, 2016.

[7] Weisstein, Eric W.: *Complex Number.* WolframMathWorld.
http://mathworld.wolfram.com/ComplexNumber.html (5.9.2018).

[8] Wikipedia: *Matrix decomposition*
https://en.wikipedia.org/wiki/Matrix_decomposition (23.2.2020)