TAMPEREEN YLIOPISTO
Pro gradu -tutkielma

Simo Jaakkola

Ortogonaalisuudesta

Informaatiotieteiden yksikkö
Matematiikka
Joulukuu 2015

Tampereen yliopisto
Informaatiotieteiden yksikkö
JAAKKOLA, SIMO: Ortogonaalisuudesta
Pro gradu -tutkielma, 30 s.
Matematiikka
Marraskuu 2015

Tiivistelmä

Tässä tutkielmassa tarkastelemme ortogonaalisuutta ja sen sovelluksia. Yleisesti ortogonaalisuudella tarkoitetaan kohtisuoruutta. Pyrimme havainnollistamaan ortogonaalisuutta ja sen sovelluksia kuvien ja esimerkkien avulla.

Aloitamme tarkastelun ortogonaalisuuden määritelmästä avaruudessa $\mathbb{R}^n$. Tutustumme tutkielman alussa myös ortogonaalisiin joukkoihin, kantoihin ja projektioihin. Projektioita havainnollistamme kuvien avulla. Esittelemme myös ortogonaalisen lemman ja Gram-Schmidtin ortogonalisointi algoritmin.

Ortogonaalisuuden määritelmästä siirrymme sen sovelluksiin. Tässä tutkielmassa esittelemme ortogonaalisen diagonalisoinnin ja määrittelemme positiivisesti definiitin matriisin. Näihin perehdymme useiden esimerkkien avulla. Positiiviesti definiittien matriisien yhteydessä esittelemme Choleskyn hajotelman ja sen algoritmin. Lopuksi tutkimme neliömuotojen sovelluksia ja havainnollistamme niitä kuvien avulla.

Lukijalta vaadimme lineaarialgebran perusteiden tuntemista. Lähdeteoksina käytämme Keith Nicholsonin kirjaa Linear Algebra with Applications ja David C. Layn kirjaa Linear Algebra and Its Applications.

Sisältö

1 Johdanto

Tämän tutkielman luvussa 2 tarkastelemme ortogonaalisuutta avaruudessa $\mathbb{R}^n$. Aluksi kertaamme kuitenkin pistetulon määritelmän. Tästä etenemme ortogonaalisuuden määritelmään ja tutkimme ortogonaalisia joukkoja ja kantoja. Esittelemme tässä luvussa myös Pythagoran lauseen ja etsimme esimerkin avulla matriisin riviavaruuden ortogonaalisen kannan. Lopuksi tarkastelemme projektioita aliluvussa 2.3.

Seuraavassa luvussa tarkastelemme ortogonaalista diagonalisointia. Tätä varten kertaamme diagonaalimatriisin määritelmän ja määrittelemme ortogonaalimatriisin. Näiden avulla diagonalisoimme symmetrisiä matriiseja esimerkkien avulla.

Neljännessä luvussa esittelemme positiivisesti definiittien matriisien määritelmän ja ominaisuuksia. Tässä luvussa tutustumme myös Choleskyn hajotelmaan esimerkkien avulla.

Tutkielman viimeisessä luvussa tarkastelemme neliömuotojen sovelluksia. Aluksi määrittelemme neliömuodon. Luvussa esittelemme muuttujien vaihdon ja lisäksi tarkastelemme neliömuotoja käyttämällä ortogonaalista diagonalisointia.

Lukijalta edellytämme lineaarialgebran perusteiden tuntemista ja tulkintaa. Ensisijaisena lähdeteoksena käytämme Keith Nicholsonin kirjaa Linear Algebra with Applications. Toisena lähteenä käytämme David C. Layn kirjaa Linear Algebra and Its Applications.

2 Ortogonaalisuus avaruudessa $\mathbb{R}^n$

Luvussa 2 tarkastelemme ortogonaalisuutta avaruudessa $\mathbb{R}^n$. Aluksi kertaamme pistetulon määritelmän, jonka avulla tarkastelemme ortogonaalisia joukkoja ja kantoja. Seuraavaksi muodostamme ortogonaalisen kannan Gram-Schmidtin algoritmilla. Lopuksi käymme läpi projektioita havainnollistavien esirnerkkien avulla.

2.1 Pistetulon määritelmä

Olkoon $X = [x_1\ x_2 \dots x_n]^T \in \mathbb{R}^n$ ja $Y = [y_1\ y_2 \dots y_n]^T \in \mathbb{R}^n$. Nyt *pistetulon* $X \cdot Y$ määritelmä on

$$X \cdot Y = X^T Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Vektorin X *pituus* $\|X\|$ määritellään seuraavasti:

$$\|X\| = \sqrt{X \cdot X} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Tämä pätee esimerkiksi avaruuksissa $\mathbb{R}^2$ ja $\mathbb{R}^3$. Myös pistetulo vastaa matriisien kertolaskua. Mikäli X,$Y \in \mathbb{R}^n$ kirjoitetaankin sarakkeiden sijaan riveiksi, niin pistetulo $X \cdot Y = XY^T$.

Esimerkki 2.1. Jos $X = [2\ 0\ \ -1\ 1] \in \mathbb{R}^4$ ja $Y = [0\ \ -1\ 3\ \ -3] \in \mathbb{R}^4$, niin

$$X \cdot Y = 0 + 0 - 3 - 3 = -6 \text{ ja } \|X\|^2 = 4 + 0 + 1 + 1 = 6.$$

Lause 2.1. *Kaikille $X, Y, Z \in \mathbb{R}^n$ pätee*

1. $X \cdot Y = Y \cdot X$
2. $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$
3. $(aX) \cdot Y = X \cdot (aY) = a(X \cdot Y)$ *kaikilla* $a \in \mathbb{R}$
4. $\|X\| \geq 0$, *ja* $\|X\| = 0$, *jos ja vain jos* $X = 0$.

Todistus sivuutetaan.

2.2 Ortogonaaliset joukot ja kannat

Määritelmä 2.1. Kaksi vektoria avaruudessa $\mathbb{R}^3$ ovat *ortogonaaliset*, jos niiden pistetulo on nolla. Toisin sanoen avaruuteen $\mathbb{R}^n$ kuuluvia vektoreita X ja Y sanotaan ortogonaalisiksi, jos $X \cdot Y = 0$. Joukkoa $\{X_1, X_2, \dots, X_n\} \in \mathbb{R}^n$ sanotaan *otrogonaaliseksi joukoksi*, jos kaikilla i:n arvoilla $X_i \neq 0$ ja $X_i \cdot X_j = 0$, kun $i \neq j$. Mikäli lisäksi kaikilla i:n arvoilla $\|X_i\| = 1$, sanotaan joukkoa $\{X_1, X_2, \dots, X_n\}$ *ortonormaaliksi joukoksi.*

Selvästi 0 on ortogonaalinen kaikille vektoreille, ja 0 on myös ainoa vektori avaruudessa $\mathbb{R}^n$, joka on ortogonaalinen itsensä kanssa. Jätämme nollavektorin pois ortogonaalisista joukoista, koska olemme ensisijaisesti kiinnostuneita ortogonaalisista kannoista.

Esimerkki 2.2. Avaruuden $\mathbb{R}^n$ luonnollinen kanta on ortonormaali.

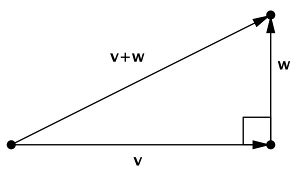
Huomautus: Osoitetaan, että $\frac{1}{\|X\|}X$ on ainoa vektorin X positiivinen monikerta joka on yksikkövektori (eli pituus $= 1$), kun $X \in \mathbb{R}^n$ ja $X \neq 0$.

Nyt jos $t > 0$, niin $\|tX\|^2 = tX \cdot tX = t^2\|X\|^2$. Koska $t > 0$, niin $t^2\|X\|^2 = 1$ jos ja vain jos $t = \frac{1}{\|X\|}$.

Näin ollen, esimerkiksi vektorin $[1\ -1\ 2\ 0] \in \mathbb{R}^4$ kanssa samansuuntainen yksikkövektori on $\frac{1}{\sqrt{6}}[1\ \ -1\ 2\ 0]$.

Jos $\{X_1, X_2, \ldots, X_m\}$ on ortogonaalinen, niin myös $\{a_1X_1, a_2X_2, \ldots, a_mX_m\}$ on ortogonaalinen kaikilla $a_i \neq 0$. Erityisesti $\{\frac{1}{\|X_1\|}X_1, \ldots, \frac{1}{\|X_m\|}X_m\}$ on ortonormaali, ja tätä joukkoa sanomme ortogonaalisen joukon $\{X_1, X_2, \ldots, X_m\}$ normalisoinnin tulokseksi.

Esimerkki 2.3. Olkoon $E_1 = [1\ 1\ 1\ {-1}]^T, E_2 = [1\ 0\ 1\ 2]^T, E_3 = [-1\ 0\ 1\ 0]^T, E_4 = [-1\ 3\ \ -1\ 1]^T$. Nyt $\{E_1, E_2, E_3, E_4\}$ on ortogonaalinen avaruudessa $\mathbb{R}^4$. Normalisoinnin jälkeen, vastaava ortonormaali joukko on $\{\frac{1}{2}E_1, \frac{1}{\sqrt{6}}E_2, \frac{1}{\sqrt{2}}E_3, \frac{1}{2\sqrt{3}}E_4\}$.



Kuva 1: Keskenään kohtisuorat vektorit $\mathbf{v}$ ja $\mathbf{w}$.

Olkoot $\mathbf{v}$ ja $\mathbf{w}$ keskenään kohtisuorat vektorit, kuvan 1 osoittamalla tavalla. Täten hypotenuusaa vastaava vektori on suunnikassäännön mukaisesti $\mathbf{v} + \mathbf{w}$. Nyt Pythagoraan lause voidaan esittää muodossa $\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$. Tämä siksi, että $\{\mathbf{v}, \mathbf{w}\}$ on ortogonaalinen joukko. Tästä pääsemme seuraavaksi esiteltävään Pythagoraan lauseen yleiseen muotoon.

Lause 2.2. *(Pythagoraan lause) Jos X ja Y ovat ortogonaaliset avaruudessa $\mathbb{R}^n$, niin $\|X + Y\|^2 = \|X\|^2 + \|Y\|^2$.*

Todistus. (vrt. [2, s.271]). Pituuden määritelmän, lauseen 2.1 ja ortogonaalisuuden määritelmän perusteella

$$\begin{aligned}\|X+Y\|^2 &= (X+Y)\cdot(X+Y)\\ &= X\cdot X + X\cdot Y + Y\cdot X + Y\cdot Y\\ &= \|X\|^2 + 0 + 0 + \|Y\|^2 = \|X\|^2 + \|Y\|^2. \qquad \square\end{aligned}$$

Lause 2.3. *Kaikki avaruuden $\mathbb{R}^n$ ortogonaaliset vektorijoukot ovat lineaarisesti riippumattomia.*

Todistus. (vrt. [2, s.271]). Jos $\{X_1, X_2, \ldots, X_m\}$ on ortogonaalinen ja $r_1X_1 + r_2X_2 + \cdots + r_mX_m = 0$, niin X_1:n pistetuloksi saamme:

$$\begin{aligned}0 = X_1\cdot 0 &= X_1\cdot(r_1X_1 + r_2X_2 + \cdots + r_mX_m)\\ &= r_1(X_1\cdot X_1) + r_2(X_1\cdot X_2) + \cdots + r_m(X_1\cdot X_m)\\ &= r_1\|X_1\|^2 + 0 + \cdots + 0.\end{aligned}$$

Nyt $r_1 = 0$, koska $X_1 \neq 0$. Vastaavasti $r_2 = r_3 = \cdots = r_m = 0$. $\square$

Ortogonaaliset kannat ovat käteviä, koska esittäessämme vektorin sen kantavektorien lineaarikombinaationa kertoimille on olemassa täsmälliset kaavat.

Lause 2.4. *Jos $\{E_1, E_2, \ldots, E_n\}$ on avaruuden $\mathbb{R}^n$ ortogonaalinen kanta, niin*

$$X = \frac{X\cdot E_1}{\|E_1\|^2}E_1 + \frac{X\cdot E_2}{\|E_2\|^2}E_2 + \cdots + \frac{X\cdot E_n}{\|E_n\|^2}E_n$$

kaikilla $X \in \mathbb{R}^n$.

Todistus. (vrt. [2, s.271]). Jos $X = r_1E_1 + \cdots + r_iE_i + \cdots + r_nE_n$, niin

$$X\cdot E_i = r(E_1\cdot E_i) + \cdots + r_i\|E_i\|^2 + \cdots + r_n(E_n\cdot E_i) = r_i\|E_i\|^2.$$

Koska $\|E_i\|^2 \neq 0$, saamme lauseen 2.4. $\square$

Esimerkki 2.4. Ortogonaalinen joukko $\{E_1, E_2, E_3, E_4\}$ esimerkissä 2.3 on lauseen 2.3 nojalla avaruuden $\mathbb{R}^4$ kanta. Jos $X = [a\ b\ c\ d]^T$, niin lauseella 2.3 saamme $X = r_1E_1 + r_2E_2 + r_3E_3 + r_4E_4$, missä

$$\begin{aligned}r_1 &= \frac{X\cdot E_1}{\|E_1\|^2}E_1 = \frac{1}{4}(a+b+c-d)\\ r_2 &= \frac{X\cdot E_2}{\|E_2\|^2}E_2 = \frac{1}{6}(a+c+2d)\\ r_3 &= \frac{X\cdot E_3}{\|E_3\|^2}E_3 = \frac{1}{2}(-a+c)\\ r_4 &= \frac{X\cdot E_4}{\|E_4\|^2}E_4 = \frac{1}{12}(-a+3b-c+d).\end{aligned}$$

Jos $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ on lineaarisest riiippumaton vekotriavaruudessa ja jos $\mathbf{v}_{m+1}$ ei kuulu joukkoon span$\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, niin $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}\}$ on lineaarisesti riippumaton [2, Theorem 1 §6.4]. Ortogonaalisille joukoille avaruudessa $\mathbb{R}^n$ ilmoitamme vastaavan seuraavasti.

Lause 2.5. *(Ortogonaalinen lemma) Olkoon $\{E_1, E_2, \ldots, E_m\}$ ortogonaalinen joukko avaruudessa $\mathbb{R}^n$ ja $X \in \mathbb{R}^n$. Olkoon*

$$E_{m+1} = X - \frac{X \cdot E_1}{\|E_1\|^2} E_1 - \frac{X \cdot E_2}{\|E_2\|^2} E_2 - \cdots - \frac{X \cdot E_m}{\|E_m\|^2} E_m.$$

Silloin:

1. *E_{m+1} on ortogonaalinen kaikille $E_1, E_2, \ldots, E_m$,*

2. *jos $X \notin span\{E_1, \ldots, E_m\}$, niin $E_{m+1} \neq 0$ ja $\{E_1, \ldots, E_m, E_{m+1}\}$ on ortogonaalinen joukko.*

Todistus. (vrt. [2, s.272]). Merkitään $t_i = (X \cdot E_i)/\|E_i\|^2$ jokaiselle $i = 1, 2, \ldots, m$. Nyt kun $1 \leq k \leq m$,

$$\begin{aligned} E_{m+1} \cdot E_k &= (X - t_1 E_1 - \cdots - t_k E_k - \cdots - t_m E_m) \cdot E_k \\ &= X \cdot E_k - t_1(E_1 \cdot E_k) - \cdots - t_k(E_k \cdot E_k) - \cdots - t_m(E_m \cdot E_k) \\ &= X \cdot E_k - t_k \|E_k\|^2 \\ &= 0. \end{aligned}$$

Tämä todistaa kohdan 1. Kohta 2 seuraa, koska $E_{m+1} \neq 0$, jos $X \notin \text{span}\{E_1, \ldots, E_m\}$. □

Ortogonaalisella lemmalla on kolme tärkeää seurausta avaruudelle $\mathbb{R}^n$. Ensimmäinen näistä on laajennus ortonormaaleille joukoille, missä jokainen riippumaton joukko, on osa kantaa[2, Theorem 2 §6.4].

Lause 2.6. *Olkoon U avaruuden $\mathbb{R}^n$ aliavaruus.*

1. *Jokainen ortogonaalinen osajoukko $\{E_1, \ldots, E_m\} \subseteq U$ on osa U:n jotakin ortogonaalista kantaa.*

2. *Jos $U \neq 0$, sillä on ortogonaalinen kanta.*

Todistus. Ks. [2, s.272] □

Voimme kehittää lauseen 2.6 kohtaa 2. Ortogonaalisen lemman toinen seuraus on toimenpide, jolla mikä tahansa aliavaruuden U kanta $\{X_1, \ldots, X_m\} \in \mathbb{R}^n$ voidaan systemaattisesti muuttaa aliavaruuden U ortogonaaliseksi kannaksi $\{E_1, \ldots, E_m\}$. Vektorit E_i konstruoidaan yksi kerrallaan vektorien X_i

avulla. Aluksi, $E_1 = X_1$. Seuraavaksi $X_2 \notin \text{span}\{E_1\}$, koska $\{X_1, X_2\}$ on lineaarisesti riippumaton. Valitaan

$$E_2 = X_2 - \frac{X_2 \cdot E_1}{\|E_1\|^2} E_1.$$

Siis $\{E_1, E_2\}$ on ortogonaalisen lemman mukaisesti ortogonaalinen. Lisäksi $\text{span}\{E_1, E_2\} = \text{span}\{X_1, X_2\}$, joten $X_3 \notin \text{span}\{E_1, E_2\}$.Täten $\{E_1, E_2, E_3\}$ on ortogonaalinen, missä

$$E_3 = X_3 - \frac{X_3 \cdot E_1}{\|E_1\|^2} E_1 - \frac{X_3 \cdot E_2}{\|E_2\|^2} E_2.$$

Jälleen $\text{span}\{E_1, E_2, E_3\} = \text{span}\{X_1, X_2, X_3\}$, joten $X_4 \notin \text{span}\{E_1, E_2, E_3\}$. Prosessi jatkuu aina m:nteen iterointiin asti, jossa konstruoimme ortogonaalisen joukon $\{E_1, \ldots, , E_m\}$ siten, että $\text{span}\{E_1, E_2, \ldots, E_m\} = \text{span}\{X_1, X_2, \ldots X_m\} = U$. Seurauksena $\{E_1, E_2, \ldots, E_m\}$ on haluttu U:n ortogonaalinen kanta. Tämän prosessin esitämme seuraavaksi tiivistettynä.

Lause 2.7. *(Gram-Schmidtin ortogonalisointialgoritmi)*
Jos $\{X_1, X_2, \ldots, X_m\}$ on mikä tahansa aliavaruuden $U \subseteq \mathbb{R}^n$ kanta, konstruoimme vektorit $E_1, E_2, \ldots, E_m \in U$ seuraavasti:

$$\begin{aligned}
&E_1 = X_1 \\
&E_2 = X_2 - \frac{X_2 \cdot E_1}{\|E_1\|^2} E_1 \\
&E_3 = X_3 - \frac{X_3 \cdot E_1}{\|E_1\|^2} E_1 - \frac{X_3 \cdot E_2}{\|E_2\|^2} E_2 \\
&\ldots \\
&E_k = X_k - \frac{X_k \cdot E_1}{\|E_1\|^2} E_1 - \frac{X_k \cdot E_2}{\|E_2\|^2} E_2 - \cdots - \frac{X_k \cdot E_{k-1}}{\|E_{k-1}\|^2} E_{k-1}
\end{aligned}$$

kaikille $k = 2, 3, \ldots, m$. Silloin

1. *$\{E_1, E_2, \ldots, E_m\}$ on U:n ortogonaalinen kanta,*
2. *$\text{span}\{E_1, E_2, \ldots, E_m\} = \text{span}\{X_1, X_2, \ldots, X_m\}$ kaikilla $k = 2, 3, \ldots, m$.*

Esimerkki 2.5. Etsi ortogonaalinen kanta matriisin

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

riviavaruudelle.
Ratkaisu:
Merkitään symboleilla X_1, X_2, X_3 matriisin A rivejä. Nyt $\{X_1, X_2, X_3\}$ on

lineaarisesti riippumaton. Merkitään $E_1 = X_1$. Edellisen algoritmin avulla saamme

$$E_2 = X_2 - \frac{X_2 \cdot E_1}{\|E_1\|^2} E_1 = [2\ \ -1\ 4\ 0] - \frac{[2\ \ -1\ 4\ 0] \cdot [2\ 1\ 1\ \ -1]}{\|[2\ 1\ 1\ \ -1]\|^2} [2\ 1\ 1\ \ -1]$$
$$= [2\ \ -1\ 4\ 0] - \frac{7}{7}[2\ 1\ 1\ \ -1] = [0\ \ -2\ 3\ \ -1]$$

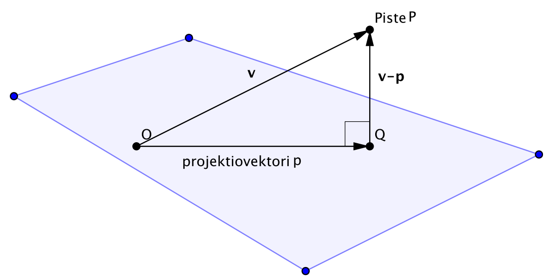
$$E_3 = X_3 - \frac{X_3 \cdot E_1}{\|E_1\|^2} E_1 - \frac{X_3 \cdot E_2}{\|E_2\|^2} E_2 = X_3 - \frac{0}{7} E_1 - \frac{-3}{25} E_2 = \frac{1}{25}[0\ 19\ 9\ 22].$$

Täten $\{[2\ 1\ 1\ -1], [0\ -2\ 3\ -1], \frac{1}{25}[0\ 19\ 9\ 22]\}$ on muodostettu ortogonaalinen kanta.

Huomataan, että vektori $\frac{X \cdot E_i}{\|E_i\|^2} E_i$ säilyy muuttumattomana jos E_i:n paikalla käytetään E_i:n skalaarimonikertaa ($\neq 0$). Nyt jos vasta konstruoitu E_i kerrotaan jossain vaiheessa Gram-Schmidtin algoritmia skalaarilla ($\neq 0$), tätä seuraavat E:t säilyvät muuttumattomina. Tämä on käytännöllistä varsinaisessa laskemisessa.

2.3 Projektiot

Tässä luvussa tarkastelemme projektion määritelmää esimerkkien avulla.



Kuva 2: Projektion havainnollistaminen.

Tutkitaan pistettä P ja origon O kautta kulkevaa tasoa. Haluamme löytää tasolta pisteen Q, joka on lähimpänä pistettä P. Geometristen havaintojemme pohjalta tiedämme, että kyseinen piste Q on olemassa. Merkitään P:n paikkavektoria, eli vektoria O:sta P:hen, merkillä $\mathbf{v}$. Nyt meidän pitää löytää paikkavektori $\mathbf{p}$ kuvan 2 osoittamalla tavalla. Jälleen tiedämme, että jos $\mathbf{p}$ valitaan siten, että vektori $\mathbf{v} - \mathbf{p}$ on kohtisuorassa valitun tason kanssa, on $\mathbf{p}$ haluamamme vektori.

Nyt voimme tehdä kaksi havaintoa. Ensimmäiseksi huomaamme, että tasolla olevien pisteiden paikkavektorien joukko U on avaruuden $\mathbb{R}^3$ aliavaruus. Toiseksi, koska vektori $\mathbf{v}-\mathbf{p}$ on kohtisuorassa valitun tason kanssa, tarkoittaa se myös, että $\mathbf{v}-\mathbf{p}$ on ortogonaalinen kaikille joukon U vektoreille. Ortogonaalisen lemman avulla voimme löytää $\mathbf{p}$:n yleisemmällä tavalla.

Määritelmä 2.2. Jos U on avaruuden $\mathbb{R}^n$ aliavaruus, määritellään U:n *ortogonaalinen komplementti* $U^\perp$ seuraavasti:

$$U^\perp = \{X \in \mathbb{R}^n | X \cdot Y = 0, Y \in U\}.$$

Esimerkki 2.6. Etsi $U^\perp$, jos $U = \text{span}\{[1\ -1\ 2\ 0], [1\ 0\ -2\ 3]\}$, kun $U \in \mathbb{R}^4$

Ratkaisu: $X = [x\ y\ z\ w] \in U^\perp$, jos ja vain jos se on ortogonaalinen sekä vektorin $[1\ -1\ 2\ 0]$ että vektorin $[1\ 0\ -2\ 3]$ kanssa. Näin ollen

$$\begin{aligned} x - y + 2z &= 0 \\ x - 2z + 3w &= 0. \end{aligned}$$

Gaussin eliminoinnin avulla saamme $U^\perp = \text{span}\{[2\ 4\ 1\ 0], [3\ 3\ 0\ -1]\}$.

Olkoon $\{E_1, E_2, \ldots, E_m\}$ ortogonaalinen kanta $\mathbb{R}^n$:n aliavaruudelle U. Olkoon $X \in \mathbb{R}^n$ ja tarkastellaan vektoria

$$P = \frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 + \cdots + \frac{X \cdot E_m}{\|E_m\|^2} E_m.$$

Tällöin $P \in U$ ja $X - P \in U^\perp$. Lisäksi P on riippumaton U:n ortogonaalisen kannan valinnasta. Tämän todistamiseksi oletetaan, että $\{E_1', E_2', \ldots, E_m'\}$ on toinen U:n ortogonaalinen kanta, ja kirjoitetaan

$$P' = \frac{X \cdot E_1'}{\|E_1'\|^2} E_1' + \frac{X \cdot E_2'}{\|E_2'\|^2} E_2' + \cdots + \frac{X \cdot E_m'}{\|E_m'\|^2} E_m'.$$

Vastaavasti kuten edellä, saamme $P' \in U$ ja $X - P' \in U^\perp$. Kirjoitetaan vektori $P - P'$ muotoon

$$P - P' = (X - P') - (X - P).$$

Huomataan, että $P - P' \in U$ ja $P - P' \in U^\perp$. Täten $P - P'$ täytyy olla nolla, eli $P = P'$ kuten väitimmekin.

Vektori P riippuu vain vektorista X ja aliavaruudesta U, eikä U:n ortogonaalisen kannan $\{E_1, E_2, \ldots, E_m\}$ valinnasta. Täten voimme tehdä seuraavan määritelmän.

Määritelmä 2.3. Olkoon U avaruuden $\mathbb{R}^n$ aliavaruus, jolla on ortogonaalinen kanta $\{E_1, E_2, \ldots, E_m\}$. Jos $X \in \mathbb{R}^n$, niin vektoria

$$\text{proj}_U(X) = \frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 + \cdots + \frac{X \cdot E_m}{\|E_m\|^2} E_m$$

sanotaan X:n *ortogonaaliseksi projektioksi* avaruudelle U. Määrittelemme, että $\text{proj}_0(X) = 0$. Edeltävä todistaa seuraavan lauseen ensimmäisen osan.

Lause 2.8. *(Projektiolause) Jos U on avaruuden $\mathbb{R}^n$ aliavaruus ja $X \in \mathbb{R}^n$, niin kirjoitamme $P = \text{proj}_U(X)$. Tällöin*

1. *$P \in U$ ja $X - P \in U^\perp$,*
2. *P on lähimpänä X:ää oleva vektori, joka kuuluu U:hun siinä mielessä, että*
   $$\|X - P\| < \|X - Y\| \text{ kaikilla } Y \in U, Y \neq P,$$
3. *$\dim U + \dim U^\perp = n$.*

Todistus. (vrt. [2, s.275]).

1. Tämä on todistettu edellä.
2. Kirjoitetaan $X - Y = (X - P) + (P - Y)$. Nyt $P - Y \in U$ ja on siten ortogonaalinen vektorille $X - P$. Pythagoraan lauseen nojalla saamme
   $$\|X - Y\|^2 = \|X - P\|^2 + \|P - Y\|^2 > \|X - P\|^2,$$
   koska $P - Y \neq 0$.
3. Jos $U = 0$, niin $U^\perp = \mathbb{R}^n$ ja kohta (3) pätee. Jos $U^\perp = 0$, niin kohdan (1) mukaisesti $U = \mathbb{R}^n$ ja jälleen kohta (3) pätee. Täten oletetaan, että $U \neq 0$ ja $U^\perp \neq 0$. Olkoon $\{E_1, \ldots, E_m\}$ U:n ortogonaalinen kanta, ja vastaavasti $\{F_1, \ldots, F_k\}$ $U^\perp$:n ortogonaalinen kanta. Täten $\{E_1, \ldots, E_m, F_1, \ldots, F_k\}$ on joukon $U^\perp$ määritelmän nojalla ortogonaalinen. Riittää, kun osoitamme, että joukko virittää avaruuden $\mathbb{R}^n$, ja tämä seuraa kohdasta (1). $\square$

Esimerkki 2.7. Olkoon $U = \text{span}\{X_1, X_2\} \subseteq \mathbb{R}^4$, missä $X_1 = [1\ 1\ 0\ 1]$ ja $X_2 = [0\ 1\ 1\ 2]$. Jos $X = [3\ -1\ 0\ 2]$, niin etsi X:ää lähimpänä oleva vektori, joka kuuluu U:hun. Esitä X avaruudelle U ortogonaalisen ja avaruuteen U kuuluvan vektorin summana.

Ratkaisu: $\{X_1, X_2\}$ on lineaarisesti riippumaton, muttei ortogonaalinen. Gram-Schmidtin prosessilla saamme U:lle ortogonaalisen kannan $\{E_1, E_2\}$, missä $E_1 = X_1 = [1\ 1\ 0\ 1]$ ja

$$E_2 = X_2 - \frac{X_2 \cdot E_1}{\|E_1\|^2} E_1 = X_2 - \frac{3}{3} E_1 = [-1\ 0\ 1\ 1]$$

Nyt voimme määrittää projektion käyttämällä $\{E_1, E_2\}$:

$$P = \operatorname{proj}_U(X) = \frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 = \frac{4}{3} E_1 + \frac{-1}{3} E_2 = \frac{1}{3} [5\ 4\ \ -1\ 3].$$

Täten P on X:ää lähimpänä oleva vektori, joka kuuluu U:hun, ja $X - P = \frac{1}{3}[4\ -7\ 1\ 3]$ on ortogonaalinen kaikille U:n vektoreille. Haluttu X:n hajotelma on siten

$$X = P + (X - P) = \frac{1}{3}[5\ 4\ \ -1\ 3] + \frac{1}{3}[4\ \ -7\ 1\ 3]$$

Esimerkki 2.8. Etsi tason $2x + y - z = 0$ piste, joka on lähimpänä pistettä $P_0 = (2, -1, -3)$.
Ratkaisu: Taso on aliavaruus U, jonka pisteet (x, y, z) toteuttavat ehdon $z = 2x + y$. Siis

$$U = \{(s, t, 2s + t) | s, r \in \mathbb{R}\} = \operatorname{span}\{(0, 1, 1), (1, 0, 2)\}$$

Gram-Schmidtin prosessilla saadaan U:lle ortogonaalinen kanta $\{\mathbf{e_1}, \mathbf{e_2}\}$, missä $\mathbf{e_1} = (0, 1, 1)$ ja $\mathbf{e_2} = (1, -1, 1)$. Näin ollen paikkavektoria $\mathbf{v} = (2, -1, -3)$ lähimpänä oleva vektori on

$$\operatorname{proj}_U(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{e_1}}{\|\mathbf{e_1}\|^2} \mathbf{e_1} + \frac{\mathbf{v} \cdot \mathbf{e_2}}{\|\mathbf{e_2}\|^2} \mathbf{e_2} = -2\mathbf{e_1} + 0\mathbf{e_2} = (0, -2, -2).$$

Täten pistettä $P_0 = (2, -1, -3)$ lähimpänä oleva U:n piste on $Q(0, -2, -2)$.

3 Ortogonaalinen diagonalisointi

Luvussa 3 kertaamme aluksi diagonalisoituvan matriisin määritelmän ja määrittelemme ortogonaalimatriisin. Luvun tärkeimpänä tuloksena huomaamme, että symmetrisillä matriiseilla on ortogonaalinen kanta. Lopuksi diagonalisoimme symmetrisiä matriiseja esimerkkien avulla.

3.1 Ortogonaalimatriisit

Olkoon A $n \times n$-matriisi. Se on diagonalisoituva, jos ja vain jos sillä on n lineaarisesti riippumatonta ominaisvektoria. Nyt matriisi P, jonka sarakkeina ovat nämä ominaisvektorit, on A:n *diagonalisoiva* matriisi. Tällöin matriisi $P^{-1}AP$ on diagonaalimatriisi.

Kuten luvussa 2 huomasimme, avaruuden $\mathbb{R}^n$ mielekkäimmät kannat ovat yleensä ortogonaaliset kannat. Nyt haluamme luonnollisesti tutkia millä $n \times n$-matriiseilla on ortogonaalinen kanta ominaisvektoreina. Huomaamme myöhemmin, että nämä matriisit ovat symmetrisiä.

Jos matriisilla A on n kpl ortogonaalisia ominaisvektoreita, niistä saadaan normalisoinnin avulla ortonormaaleja. Vastaavalla diagonalisoivalla matriisilla P on ortonormaalit sarakkeet. Tällaiset matriisit ovat helposti käännettävissä.

Lause 3.1. *Seuraavat ehdot ovat yhtäpitäviä $n \times n$-matriisille P.*

1. *P on kääntyvä ja $P^{-1} = P^T$.*
2. *P:n rivit ovat ortonormaaleja.*
3. *P:n sarakkeet ovat ortonormaaleja.*

Todistus. (vrt. [2, s.279]). Tiedetään, että ensimmäinen ehto on yhtäpitävä ehdon $PP^T = I$ kanssa [2, Theorem 5 §2.3]. Merkitään symboleilla $X_1, X_2, \ldots, X_n$ matriisin P rivejä. Täten X_j^T on matriisin P^T sarake j. Nyt matriisin PP^T (i,j) alkio vastaa siis pistetuloa $X_i \cdot X_j$. Tällöin $PP^T = I$ tarkoittaa, että pistetulo $X_i \cdot X_j = 0$, kun $i \neq j$ ja $X_i \cdot X_j = 1$, kun $i = j$. Täten ehto (1) on yhtäpitävä ehdon (2) kanssa. Ehtojen (1) ja (3) yhtäpitävyys todistetaan vastaavasti. □

Nyt $n \times n$-matriisia P sanotaan *ortogonaalimatriisiksi*, jos se toteuttaa yhden (ja täten kaikki) lauseen 3.1 kohdista.

Esimerkki 3.1. Matriisi $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ on ortogonaalinen kaikille kulman θ arvoille.

Ollakseen ortogonaalimatriisi ei kuitenkaan riitä, että matriisin A rivit ovat ortogonaaliset.

Esimerkki 3.2. (vrt. [2, s.279]). Matriisilla

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

on ortogonaaliset rivit, mutta se ei ole ortogonaalimatriisi. Kuitenkin, jos rivit normalisoidaan saadaan ortogonaalimatriisi

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

Nyt myös sarakkeet ovat ortonormaalit.

Esimerkki 3.3. Jos P ja Q ovat ortogonaalimatriiseja, niin PQ on ortogonaalinen, ja tällöin myös $P^{-1} = P^T$ on ortogonaalinen.

Ratkaisu: P ja Q ovat kääntyviä matriiseja, joten myös PQ on kääntyvä ja $(PQ)^{-1} = Q^{-1}P^{-1} = Q^T P^T = (PQ)^T$. Näin ollen PQ on ortogonaalinen. Vastaavasti $(P^{-1})^{-1} = P = (P^T)^T = (P^{-1})^T$ ja täten P^{-1} on ortogonaalinen.

Neliömatriisia A sanotaan *ortogonaalisesti diagonalisoituvaksi*, kun on olemassa ortogonaalinen matriisi P siten, että $P^{-1}AP = P^T AP$ on diagonaalinen. Tällä ehdolla karakterisoimme symmetriset matriisit.

3.2 Symmetriset matriisit

Lause 3.2. *(Pääakselilause) Seuraavat ehdot ovat yhtäpitäviä $n \times n$-matriisille A.*

1. *A:lla on ominaisvektoreiden muodostama ortonormaali joukko.*
2. *A on ortogonaalisesti diagonalisoituva.*
3. *A on symmetrinen.*

Todistus. Ks. [2, s.280] □

Symmetrisen matriisin A ominaisvektorien ortonormaalia joukkoa sanotaan A:n *pääakselien* joukoksi. Lausetta 3.2 sanotaan myös *spektrilauseeksi* tai *spektraalilauseeksi* ja ominaisarvojen joukkoa sanotaan matriisin *spektriksi*.

Esimerkki 3.4. Etsi ortogonaalinen matriisi P siten, että $P^{-1}AP$ on diagonaalinen, kun

$$A = \begin{bmatrix} 3 & 0 & 7 \\ 0 & 5 & 0 \\ 7 & 0 & 3 \end{bmatrix}.$$

Ratkaisu: A:n karakteristinen polynomi on

$$\begin{aligned}
c_A(x) &= \det \begin{bmatrix} x-3 & 0 & 7 \\ 0 & x-5 & 0 \\ 7 & 0 & x-3 \end{bmatrix} \\
&= (x-3)\begin{bmatrix} x-5 & 0 \\ 0 & x-3 \end{bmatrix} - 0 \begin{bmatrix} 0 & 0 \\ 7 & x-3 \end{bmatrix} + 7 \begin{bmatrix} 0 & x-5 \\ 7 & 0 \end{bmatrix} \\
&= (x-3)((x-5)(x-3) - 0 \cdot 0) + 7(0 \cdot 0 - 7(x-5)) \\
&= x^3 - 11x^2 - 10x + 200 \\
&= (x-10)(x-5)(x+4).
\end{aligned}$$

Näin ollen ominaisarvoiksi saamme $\lambda_1 = 10, \lambda_2 = 5$ ja $\lambda_3 = -4$. Ominaisarvoja vastaavat ominaisvektorit saadaan seuraavasti:

$$\begin{aligned}
\lambda_1 = 10 : \begin{bmatrix} 7 & 0 & -7 \\ 0 & 5 & 0 \\ -7 & 0 & 7 \end{bmatrix} &\to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ ja nyt } X_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \\
\lambda_2 = 5 : \begin{bmatrix} 2 & 0 & -7 \\ 0 & 0 & 0 \\ -7 & 0 & -2 \end{bmatrix} &\to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \text{ ja nyt } X_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \\
\lambda_3 = -4 : \begin{bmatrix} -7 & 0 & -7 \\ 0 & -9 & 0 \\ -7 & 0 & -7 \end{bmatrix} &\to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ ja nyt } X_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.
\end{aligned}$$

Hämmästyttävästi ominaisvektorit ovat keskenään ortogonaaliset. Nyt $\|X_1\|^2 = 2, \|X_2\|^2 = 1$ ja $\|X_3\|^2 = 2$ ja tällöin

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}}X_1 & X_2 & \frac{1}{\sqrt{2}}X_3 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

on ortogonaalinen (tällöin $P^{-1} = P^T$) ja

$$P^T A P = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

diagonalisointialgoritmin perusteella.

Esimerkin 3.4 ominaisvektorien ortogonaalisuus ei ole sattumaa. Lauseen [2, Theorem 4 §5.3] perusteella, ne ovat lineaarisesti riippumattomia. Matriisin symmetrisyydestä seuraa, että ominaisvektorit ovat ortogonaalisia. Tämän todistamiseen vaaditaan seuraava lause symmetrisistä matriiseista.

Lause 3.3. *Jos A on symmetrinen $n \times n$ matriisi, niin*

$$(AX) \cdot Y = X \cdot (AY)$$

kaikille sarakkeille X ja $Y \in \mathbb{R}^n$.

Todistus. (vrt. [2, s.281]). Tiedetään, että $X \cdot Y = X^T Y$ kaikille sarakkeille X ja Y. Koska $A^T = A$, saamme

$$(AX) \cdot Y = (AX)^T Y = X^T A^T Y = X^T AY = X \cdot (AY). \qquad \square$$

Lause 3.4. *Symmetrisen matriisin A ominaisarvoja vastaavat ominaisvektorit ovat ortogonaalisia.*

Todistus. (vrt. [2, s.282]). Olkoon $AX = \lambda X$ ja $AY = \mu Y$, missä $\lambda \neq \mu$. Lauseen 3.3 avulla saamme

$$\lambda(X \cdot Y) = (\lambda X) \cdot Y = (AX) \cdot Y = X \cdot (AY) = X \cdot (\mu Y) = \mu(X \cdot Y)$$

Näin ollen $(\lambda - \mu)(X \cdot Y) = 0$ ja tällöin $X \cdot Y = 0$, koska $\lambda \neq \mu$. $\square$

Nyt symmetrisen $n \times n$ matriisin diagonalisointi on selkeää. Etsitään erisuuret ominaisarvot ja jokaisen ominaisavaruuden ortogonaalinen kanta. Tähän saatetaan tarvita Gram-Schmidtin algoritmia. Lauseen 3.4 mukaisesti kaikkien näiden kantavektoreiden joukko on ortonormaali ja sisältää n vektoria.

Lause 3.5. *(Kolmiointilause) Olkoon A on $n \times n$-matriisi, jolla on reaaliset ominaisarvot, on olemassa matriisi P siten, että $P^T AP$ on yläkolmiomatriisi. Myös alakolmiomatriisi on olemassa.*

Todistus. (vrt. [2, s.283]). Olkoon $AX_1 = \lambda_1 X_1$, missä $\|X_1\| = 1$. Olkoon $\{X_1, X_2, \ldots, X_n\}$ avaruuden $\mathbb{R}^n$ ortonormaali kanta ja $P_1 = [X_1 \ldots X_n]$. Tällöin P_1 on ortogonaalinen ja $P_1^T AP_1 = \begin{bmatrix} \lambda_1 & X \\ 0 & A_1 \end{bmatrix}$ lohkomuodossa. Induktiolla olkoon $Q^T A_1 Q = T_1$ yläkolmiomatriisi, missä Q ortogonaalinen ja kokoa $(n-1) \times (n-1)$. Tällöin $P_2 = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}$ on ortogonaalinen. Täten myös $P = P_1 P_2$ on ortogonaalinen ja $P^T AP = \begin{bmatrix} \lambda_1 & XQ \\ 0 & T_1 \end{bmatrix}$ on yläkolmiomatriisi. $\square$

Lauseen 3.5 todistus ei anna keinoa matriisin P konstruoimiseen. Tähän vaadittava algoritmi esitellään lauseen 3.5 kehitetyssä versiossa, ks. [2, luku 8.7].

Kun kyseessä on diagonaalimatriisi, yläkolmiomatriisin ominaisarvot ovat diagonaalilla. Koska lauseessa 3.5, matriiseilla A ja $P^T AP$ on sama determinantti ja jälki, voimme tehdä seuraavan päätelmän:

Seuraus 3.6. Jos A on $n\times n$ matriisi, jolla on reaaliset ominaisarvot $\lambda_1, \lambda_2, \dots,$ λ_n, niin $\det A = \lambda_1\lambda_2\cdots\lambda_n$ ja $\mathrm{tr}A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Tämä seuraus pätee, vaikka kaikki ominaisarvot eivät olisikaan reaalisia. Itse asiassa, lause 3.5 pätee myös kaikille $n \times n$ kompleksimatriiseille.

4 Positiivisesti definiitit matriisit

Symmetrisen matriisin ominaisarvot ovat reaalisia. Tässä luvussa tutkimme matriiseja, joilla on positiiviset ominaisarvot. Näillä matriiseilla on useita käyttökohteita tieteessä ja tekniikan aloilla. Luvussa 4 esittelemme positiivisesti definiittien matriisien määritelmän ja tutustumme Choleskyn hajotelmaan.

4.1 Positiivisesti definiitit matriisit ja pääalimatriisit

Määritelmä 4.1. Neliömatriisi on positiivisesti definiitti, jos se on symmetrinen ja ominaisarvot ovat positiivisia.

Symmetrisyyden vuoksi, pääakselilause on olennainen osa teoriaa.

Lause 4.1. *Jos A on positiivisesti definiitti, niin se on kääntyvä ja* $\det A > 0$.

Todistus. (vrt. [2, s.286]). Jos A on kokoa $n \times n$ ja sen ominaisarvot ovat $\lambda_1, \lambda_2, \ldots, \lambda_n$, niin seuraukse 3.6 nojalla $\det A = \lambda_1 \lambda_2 \ldots \lambda_n > 0$. □

Jos X on sarake avaruudessa $\mathbb{R}^n$ ja A mikä tahansa reaalinen $n \times n$ matriisi, tarkastelemme 1×1 matriisia $X^T AX$ reaalilukuna. Tällä sopimuksella kuvaamme positiivisesti definiittejä matriiseja seuraavasti:

Lause 4.2. *Symmetrinen matriisi A on positiivisesti definiitti, jos ja vain jos $X^T AX > 0$, kun sarake $X \neq 0$ ja $X \in \mathbb{R}^n$.*

Todistus. (vrt. [2, s.286]). Pääakselilauseen nojalla olkoon $P^T AP = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, missä $P^{-1} = P^T$ ja λ_i:t ovat A:n ominaisarvot. Sarake $X \in \mathbb{R}^n$ kirjoitetaan muotoon $Y = P^T X = [y_1\ y_2\ \ldots\ y_n]^T$. Tällöin

$$X^T AX = Y^T DY = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

Jos A on positiivisesti definiitti ja $X \neq 0$, niin $X^T AX > 0$, koska jokin $y_j \neq 0$. Oletetaan käänteisesti, että $X^T AX > 0$ aina, kun $X \neq 0$. Olkoon $X = PE_j \neq 0$, missä E_j on matriisin I_n sarake j. Tällöin $Y = E_j$, ja siten $\lambda_j = X^T AX > 0$. □

Esimerkki 4.1. Olkoon U mikä tahansa kääntyvä neliömatriisi. Osoita, että $A = U^T U$ on positiivisesti definiitti.

Ratkaisu: Kirjoitetaan $q(X) = X^T AX$, missä $X \in \mathbb{R}^n$. Jos $X \neq 0$, niin

$$q(X) = X^T AX = X^T (U^T U)X = (UX)^T (UX) = \|UX\|^2 > 0,$$

koska $UX \neq 0$ (U on kääntyvä). Tämä seuraa lauseesta 4.2.

On merkittävää, että esimerkin 4.1 käänteislause pätee myös. Itse asiassa jokainen positiivisesti definiitti matriisi A voidaan ilmoittaa muodossa $A = U^T U$, missä U on yläkolmiomatriisi, jolla on diagonaalilla positiiviset alkiot. Ennen kuin osoitamme tämän, esittelemme toisen keskeisessä roolissa olevan käsitteen positiivisesti definiiteille matriiseille.

Määritelmä 4.2. Jos A on mikä tahansa neliömatriisi, niin merkitään symbolilla ${}^{(r)}A$ avulla A:n vasemman yläkulman $r \times r$ alimatriisia. Tarkentaen ${}^{(r)}A$ muodostetaan matriisista A poistamalla viimeiset $n-r$ rivit ja sarakkeet. Matriiseja ${}^{(1)}A, {}^{(2)}A, \ldots {}^{(n)}A = A$ kutsutaan matriisin A *pääalimatriiseiksi*.

Esimerkki 4.2. Olkoon $A = \begin{bmatrix} 8 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 5 \end{bmatrix}$. Nyt ${}^{(1)}A = \begin{bmatrix} 8 \end{bmatrix}$, ${}^{(2)}A = \begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$ ja ${}^{(3)}A = A$.

Lemma 4.3. Jos A on positiivisesti definiitti matriisi, niin sen jokainen pääalimatriisi ${}^{(r)}A$ on myös positiivisesti definiitti, kun $r = 1, 2, \ldots, n$.

Todistus. (vrt. [2, s.287]). Olkoon $A = \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix}$ lohkomuodossa. Jos $Y \neq 0$ ja $Y \in \mathbb{R}^n$, merkitään $X = \begin{bmatrix} Y \\ 0 \end{bmatrix} \in \mathbb{R}^n$. Tällöin $X \neq 0$ ja koska A on positiivisesti definiitti, niin

$$0 < X^T A X = [Y^T \ 0] \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix} \begin{bmatrix} Y \\ 0 \end{bmatrix} = Y^T ({}^{(r)}A) Y.$$

Tästä huomaamme, että lauseen 4.2 mukaisesti ${}^{(r)}A$ on positiivisesti definiitti. □

4.2 Choleskyn hajotelma

Jos matriisi A on positiivisesti definiitti, lemmasta 4.3 ja lauseesta 4.1 huomaamme, että $\det{}^{(r)}A > 0$ kaikilla $r = 1, 2, \ldots, n$. Tämä todistaa osan seuraavaa lausetta, joka sisältää käänteisen tapauksen esimerkille 4.1 ja karakterisoi positiivisesti definiittejä matriiseja symmetristen matriisien joukossa.

Lause 4.4. *Seuraavat ehdot ovat yhtäpitävät symmetriselle neliömatriisille A.*

1. *A on positiivisesti definiitti.*
2. *$\det({}^{(r)}A) > 0$ kaikilla $r = 1, 2, \ldots, n$.*
3. *$A = U^T U$, missä U on yläkolmiomatriisi, jolla on diagonaalilla positiiviset alkiot.*

Lisäksi kohdan (3) hajotelma on yksikäsitteinen ja sitä sanotaan A:n Choleskyn hajotelmaksi.

Todistus. (vrt. [2, s.287-288]). Esimerkin 4.1 nojalla (3) $\implies$ (1) , ja lemman 4.3 ja lauseen 4.1 nojalla (1) $\implies$ (2). Oletataan (2) todeksi, ja jatketaan todistusta induktion avulla. Jos $n = 1$, niin $A = [a]$, missä $a > 0$ kohdan (2) nojalla. Valitaan $U = [\sqrt{a}]$. Jos $n > 1$, niin merkitään $B =^{(n-1)} A$. Täten B on symmetrinen ja toteuttaa kohdan (2). Nyt induktiolla saamme $B = U^T U$ kuten kohdassa (3), missä U on $(n-1) \times (n-1)$ matriisi. Koska A on symmetrinen, sillä on lohkomuoto $A = \begin{bmatrix} B & P \\ P^T & b \end{bmatrix}$, missä P on sarake avaruudessa $\mathbb{R}^{n-1}$ ja $b \in \mathbb{R}$. Jos merkitsemme $X = (U^T)^{-1}P$ ja $c = b - X^T X$, lohkojen kertolaskulla saamme

$$A = \begin{bmatrix} U^T U & P \\ P^T & b \end{bmatrix} = \begin{bmatrix} U^T & 0 \\ X^T & 1 \end{bmatrix} \begin{bmatrix} U & X \\ 0 & c \end{bmatrix}.$$

Ottamalla determinantit ja käyttämällä lausetta [2, Theorem 5 §3.1], saamme $\det A = \det(U^T)\ \det U \cdot c = c(\det\ U)^2$. Seurauksena $c > 0$, koska kohdan (2) nojalla $\det A > 0$. Edellä mainitun kehitelmän nojalla, voimme kirjoittaa $A = \begin{bmatrix} U^T & 0 \\ X^T & \sqrt{c} \end{bmatrix} \begin{bmatrix} U & X \\ 0 & \sqrt{c} \end{bmatrix}$. Koska matriisin U diagonaalialkiot ovat positiivisia, kohta (3) pätee.

Osoitamme vielä hajotelman yksikäsitteisyyden. Oletetaan, että $A = U^T U = U_1^T U_1$ ovat kaksi Choleskyn hajotelmaa. Merkitään $D = U U_1^{-1} = (U^T)^{-1} U_1^T$. Tällöin D on yläkolmiomatriisi, koska $D = U U_1^{-1}$, ja myös alakolmiomatriisi, koska $D = (U^T)^{-1} U_1^T$. Toisin sanoen, D on diagonaalimatriisi. Täten $U = DU_1$ ja $U_1 = DU$, joten riittää, kun osoitamme, että $D = I$. Eliminoimalla U_1 saamme $U = D^2 U$, joten $D^2 = I$, koska U on kääntyvä. Matriisin D diagonaalialkioiden positiivisuudesta seuraa, että $D = I$. □

Hämmästyttävää on, että Choleskyn hajotelmassa käytetty matriisi U on helppo muodostaa matriisista A rivioperaatioiden avulla. Tämän kannalta olennainen on seuraavaksi esitettävän algoritmin ensimmäinen vaihe, joka on mahdollinen jokaiselle positiivisesti definiitille matriisille A.

Choleskyn hajotelman algoritmi 4.3. Jos A on positiivisesti definiitti matriisi, Choleskyn hajotelma $A = U^T U$ voidaan muodostaa seuraavilla vaiheilla.

1. Viedään matriisi A positiivisilla diagonaalialkioilla varustettuun yläkolmiomatriisiin U_1 käyttämällä rivioperaatioita, jolloin jokainen operaatio lisää rivin vakiolla kerrottuna alemmalle riville.

2. Muodostetaan U matriisista U_1 jakamalla jokainen matriisin U_1 rivi, kyseisen rivin diagonaalialkion neliöjuurella.

Todistus. Ks. [2, s.289] $\square$

Esimerkki 4.4. Muodosta Choleskyn hajotelma matriisille $A = \begin{bmatrix} 20 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 5 \end{bmatrix}$.

Ratkaisu: Matriisi A on lauseen 4.4 perusteella positiivisesti definiitti. Täten Choleskyn hajotelman algoritmin mukaisesti

$$A = \begin{bmatrix} 20 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 5 \end{bmatrix} \to \begin{bmatrix} 20 & 4 & 5 \\ 0 & \frac{6}{5} & 2 \\ 0 & 2 & \frac{15}{4} \end{bmatrix} \to \begin{bmatrix} 20 & 4 & 5 \\ 0 & \frac{6}{5} & 2 \\ 0 & 0 & \frac{5}{12} \end{bmatrix} = U_1.$$

Täten algoritmin toisen vaiheen mukaisesti

$$U = \begin{bmatrix} 2\sqrt{5} & \frac{2}{\sqrt{5}} & \frac{5}{2\sqrt{5}} \\ 0 & \frac{6}{\sqrt{30}} & \frac{10}{\sqrt{30}} \\ 0 & 0 & \frac{5}{2\sqrt{15}} \end{bmatrix} = \frac{1}{30} \begin{bmatrix} 60\sqrt{5} & 12\sqrt{5} & 15\sqrt{5} \\ 0 & 6\sqrt{30} & 10\sqrt{30} \\ 0 & 0 & 5\sqrt{15} \end{bmatrix}.$$

Nyt huomataan, että $U^T U = A$.

5 Sovelluksia neliömuotoihin

Polynomia $x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 + x_2x_3$ sanotaan muuttujien x_1, x_2 ja x_3 neliömuodoksi. Tässä luvussa osoitamme, että muuttujien vaihto voidaan aina tehdä niin, ettei uuden muuttujan (y_1, y_2 ja y_3) avulla muodostetussa neliömuodossa ole ristitermejä y_1y_2, y_1y_3 tai y_2y_3. Lisäksi tarkastelemme tätä kaikilla äärellisisllä luvuilla käyttämällä ortogonaalista diagonalisointia. Tällä on useita sovelluksia esimerkiksi tilastotieteissä, fysiikassa, lukuteoriassa ja geometriassa.

Määritelmä 5.1. Muuttujien $x_1, x_2, \ldots, x_n$ *neliömuoto* q on termien $x_1^2, x_2^2, \ldots, x_n^2$ ja ristitermien $x_1x_2, x_1x_3, x_2x_3, \ldots$ lineaarikombinaatio. Jos $n = 3$, niin neliömuoto

$$q = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + a_{13}x_1x_3 + a_{31}x_3x_1 + a_{23}x_2x_3 + a_{32}x_3x_2.$$

Yleisessä muodossa

$$q = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \ldots.$$

Tämä summa voidaan esittää kätevästi myös matriisien tulona

$$q = q(X) = X^TAX,$$

missä $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ ja $A = [a_{ij}]$ on reaalinen neliömatriisi. Huomattava, jos $i \neq j$, niin erilliset termit $a_{ij}x_ix_j$ ja $a_{ji}x_jx_i$ on lueteltu ja ne sisältävät termin x_ix_j. Ne voidaan korvata muodoilla

$$\frac{1}{2}(a_{ij} + a_{ji})x_ix_j \text{ ja } \frac{1}{2}(a_{ij} + a_{ji})x_jx_i$$

vastaavasti muuttamatta summaa. Voimme siis rajoituksetta olettaa, että termien x_ix_j ja x_jx_i kertoimet ovat samat.Toisin sanoen, voimme olettaa matriisin A olevan symmetrinen.

Esimerkki 5.1. (vrt. [1, s.402]). Kirjoita $q = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$ muodossa $q(X) = X^TAX$, missä A on symmetrinen 3×3 matriisi.

Ratkaisu: Termien x_1^2, x_2^2 ja x_3^2 kertoimet ovat matriisin A diagonaalilla. Jotta A on symmetrinen termin x_ix_j, kun $i \neq j$, kerroin täytyy jakautua tasan alkioiden (i, j) ja (j, i) kesken. Termin x_1x_3 kerroin on 0. Nyt

$$q(X) = [x_1 \ x_2 \ x_3] = \begin{bmatrix} 5 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Esimerkki 5.2. (vrt. [1, s.402]). Olkoon $q(X) = x_1^2 - 8x_1x_2 - 5x_2^2$. Määritä $q(X)$, kun $X = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ ja $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

Ratkaisu: Selvästi

$$\begin{aligned} q(-3,1) &= (-3)^2 - 8(-3)(1) - 5(1)^2 = 28 \\ q(2,-2) &= (2)^2 - 8(2)(-2) - 5(-2)^2 = 16 \\ q(1,-3) &= (1)^2 - 8(1)(-3) - 5(-3)^2 = -20. \end{aligned}$$

Joissain tapauksissa neliömuotoja on helpompi käyttää, kun niissä ei ole ristitulotermejä. Eli kun neliömuodon matriisi on diagonaalimatriisi. Ristitulotermi voidaan eliminoida tekemällä sopiva muuttujan vaihto.

Oletamme jatkossa, että kaikki neliömuodot voidaan esittää muodossa

$$q(X) = X^T AX,$$

missä A on symmetrinen matriisi. Tässä muodossa ongelmana on löytää uudet muuttujat $y_1, y_2, \ldots, y_n$, jotka eivät muuttujien $x_1, x_2, \ldots, x_n$ tapaan neliömuotoa q luodessa muodosta ristitermejä. Merkitään

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Tästä voimme tutkia onko $q = Y^T DY$, missä D on diagonaalinen. Tämä voidaan toteuttaa aina, eikä yllätyksenä tule, että matriisi D saadaan diagonalisoimalla symmetrinen matriisi A. Itse asiassa, pääakselilauseen nojalla, matriisille A ortogonaalinen ja sen diagonalisoiva matriisi P voidaan muodostaa

$$P^T AP = D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Diagonaalialkiot $\lambda_1, \lambda_2, \ldots, \lambda_n$ (eivät välttämättä erisuuria) ovat matriisin A ominaisarvot toistettuna kertalukunsa mukaisesti, ja matriisin P sarakkeet ovat matriisin A vastaavat ominaisvektorit.

Nyt määritetään uudet muuttujat matriisille Y seuraavien yhtälöiden avulla:

$$X = PY \text{ ja yhtäpitävästi } Y = P^T X.$$

Sijoittamalla $q(X) = X^T AX$ saamme

$$q = (PY)^T A(PY) = Y^T (P^T AP)Y = Y^T DY = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

Näin ollen tällä muuttujan vaihdolla saamme halutun yksinkertaistuksen neliömuodolle q.

Lause 5.1. *(Diagonalisointilause) Olkoon* $q = X^T AX$ *muuttujien* $x_1, x_2, \ldots, x_n$ *neliömuoto, missä* $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ *ja* A *on symmetrinen neliömatriisi. Olkoon* P *sellainen ortogonaalinen matriisi, että* $P^T AP$ *on diagonaalimatriisi, ja määritetään uudet muutujat* $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$*, missä*

$$X = PY \text{ ja yhtäpitävästi } Y = P^T X.$$

Jos q *ilmoitetaan uusien muuttujien* $y_1, y_2, \ldots, y_n$ *avulla, saadaan*

$$q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

missä $\lambda_1, \lambda_2, \ldots, \lambda_n$ *ovat matriisin* A *ominaisarvot monikerran mukaisesti kerrottuna.*

Olkoon $q = X^T AX$ neliömuoto, missä A on symmetrinen matriisi, ja olkoot $\lambda_1, \lambda_2, \ldots, \lambda_n$ sen reaaliset ominaisarvot monikerran mukaisesti kerrottuna. Matriisin A ortonormaaleja ominaisvektoreita vastaavaa joukkoa $\{E_1, \ldots, E_n\}$ kutsutaan neliömuodon q *pääakselien* joukoksi. Nimi selitetään myöhemmin. Lauseessa 5.1 esiintyvä ortogonaalinen matriisi P, esitetään nyt muodossa $P = [E_1 \ \ldots \ E_n]$, joten muuttujat X ja Y liittyvät toisiinsa seuraavasti:

$$X = PY = [E_1 \ E_2 \ \ldots \ E_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = y_1 E_1 + y_2 E_2 + \cdots + y_n E_n.$$

Täten uudet muuttujat y_i ovat kertoimet, kun X esitetään avaruuden $\mathbb{R}^n$ ortonormaalin kannan $\{E_1, \ldots E_n\}$ suhteen. Erityisesti kertoimet y_i saadaan muodossa $y_i = X \cdot E_i$ lauseen 2.4 mukaisesti. Nyt voimme määrittää neliömuodon q helposti ominaisarvojen λ_i ja pääakselien E_i avulla:

$$q = q(X) = \lambda_1 (X \cdot E_1)^2 + \cdots + \lambda_n (X \cdot E_n)^2$$

Esimerkki 5.3. Etsi uudet muuttujat y_1, y_2, ja y_3 siten, että

$$q = 7x_1^2 + x_2^2 + x_3^2 + 8x_1x_2 + 8x_1x_3 - 16x_2x_3$$

on diagonaalimuoto, ja etsi vastaavat pääakselit.

Ratkaisu:

Neliömuoto saadaan siis, kun $q = X^T AX$, missä

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ A = \begin{bmatrix} 7 & 4 & 4 \\ 4 & 1 & -8 \\ 4 & -8 & 1 \end{bmatrix}.$$

Normaaleilla laskutoimituksilla saamme

$$c_A(x) = \det(xI - A) = (x-9)^2(x+9),$$

joten vastaavat ominaisarvot ovat $\lambda_1 = \lambda_2 = 9$ ja $\lambda_3 = -9$. Vastaavat ortonormaalit ominaisvektorit ovat pääakselit:

$$F_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \ F_2 = \frac{1}{3}\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \ F_3 = \frac{1}{3}\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

Matriisi

$$P = [F_1 \ F_2 \ F_3] = \frac{1}{3}\begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$$

on näin ollen ortogonaalinen, ja $P^{-1}AP = P^T AP$ on diagonaalinen. Tästä johtuen uusille muuttujille Y ja vanhoille muuttujille X on yhteys, koska $Y = P^T X$ ja $X = PY$. Nyt kysytyt uudet muuttujat ovat

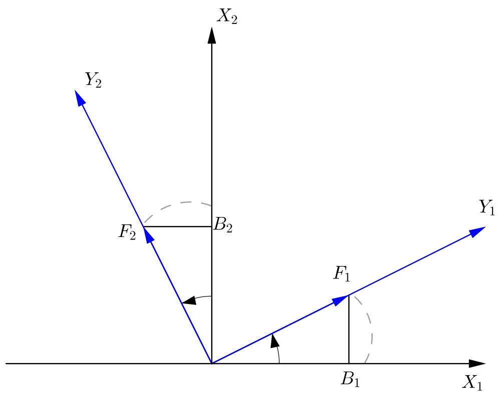
$$\begin{aligned} y_1 &= \frac{1}{3}(2x_1 + 2x_2 - x_3) \\ y_2 &= \frac{1}{3}(2x_1 - x_2 + 2x_3) \\ y_3 &= \frac{1}{3}(-x_1 + 2x_2 + 2x_3). \end{aligned}$$

Jos nämä muuttujat x_i vaihdetaan alkuperäiseen neliömuotoon q, saadaan

$$q = 9y_1^2 + 9y_2^2 - 9y_3^2,$$

joka on siis haluttu diagonaalimuoto.

On opettavaista tarkastella neliömuotoja kahden muuttujan x_1 ja x_2 avulla. Tällöin pääakselit voidaan aina löytää kiertämällä vastapäivään akseleita X_1 ja X_2 kulman θ suuruisesti origoon nähden. Näin saadaan siis uudet akselit Y_1 ja Y_2, kuten kuvasta 3 huomaamme.

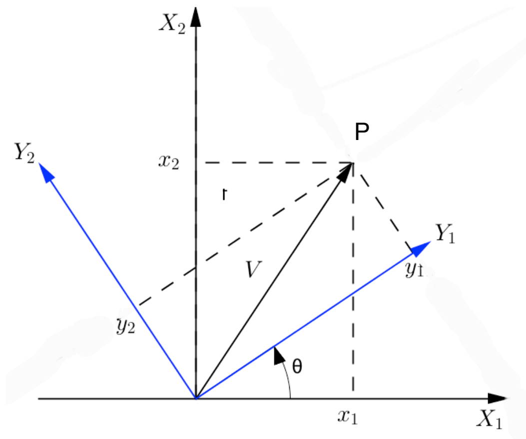


Kuva 3: Akselien kääntäminen.

Jotta noudatamme diagonalisointilausetta 5.1 kirjoitamme järjestetyt parit sarakkeiksi. Täten akseita kiertämällä, akselien X_1 ja X_2 koordinaattivektorit $B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ja $B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ siirtyvät akselien Y_1 ja Y_2 koordinaattivektoreille F_1 ja F_2. Yksinkertaisen trigonometrian avulla saamme

$$F_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad \text{ja} \quad F_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}.$$

Näiden avulla voidaan määrittää akselien Y_1 ja Y_2 koordinaatit y_1 ja y_2 pisteessä P, akselien X_1 ja X_2 pisteen P kordinaattien x_1 ja x_2 avulla. Jos V on pisteen P paikkavektori, kuten kuvassa 4, niin alkuperäisten akselien X_1 ja X_2 tapauksessa $V = x_1B_1 + x_2B_2$ ja uusien akselien Y_1 ja Y_2 kohdalla $V = y_1F_1 + y_2F_2$.



Kuva 4: Uudet koordinaatit.

Näin ollen

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = V = y_1F_1 + y_2F_2 = y_1 \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + y_2 \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

on kaava muuttujan vaihdolle, kun akseleita käännetään. Huomataan, että matriisi $P = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ on ortogonaalinen, joten edellä esitetty kaava muuttujan vaihdolle saa muodon $X = PY$, kuten diagonalisointilauseessa 5.1.

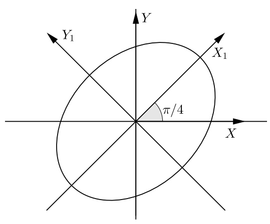
Lause 5.2. *Olkoon neliömuoto* $q = ax^2 + bxy + cy^2$*, missä* a, b *ja* c *eivät kaikki ole* 0.

1. *Koordinaattiakseleita voidaan pyörittää origoon nähden vastapäivään siten, että uudessa koordinaatistossa neliömuodolla* q *ei ole ristitermiä.*
2. *Yhtälön* $q = ax^2 + bxy + cy^2 = 1$ *kuvaaja on ellipsi, jos* $b^2 - 4ac < 0$ *ja hyperbeli, jos* $b^2 - 4ac > 0$.

Todistus. Ks. [2, s.310] □

Esimerkki 5.4. Olkoon $x^2 + xy + y^2 = 1$. Etsi akselien kierto siten, ettei yhtälöllä ole ristitermejä.

Ratkaisu: Tässä $a = b = c = 1$ lauseen 5.2 mukaisesti. Nyt $\cos\theta = \frac{1}{\sqrt{2}} = \sin\theta$. Täten $\theta = \frac{\pi}{4}$ kelpaa. Jos uudet muuttujat ovat x_1 ja y_1, niin $x = \frac{1}{\sqrt{2}}(x_1 - y_1)$ ja $y = \frac{1}{\sqrt{2}}(x_1 + y_1)$. Nyt yhtälö saadaan muotoon $3x_1^2 + y_1^2 = 2$. Yhtälön kuvaaja on ellipsi ja kulma θ on valittu siten, että uudet akselit X_1 ja Y_1 ovat ellipsin symmetria-akselit kuvan 5 mukaisesti. Ominaisvektorit $F_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ja $F_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ osoittavat näiden symmetria-akseleiden paikan. Tämä on myös syy pääakselin nimelle.



Kuva 5: Ellipsi symmetria-akseleineen.

Kaikkien ortogonaalisten matriisien P determinantti on joko 1 tai -1 (koska $PP^T = I$). Ortogonaalisen matriisin $\begin{bmatrix} \cos\,\theta & -\sin\,\theta \\ \sin\,\theta & \cos\,\theta \end{bmatrix}$ akselien pyörityksestä saatujen matriisien determinantti on 1. Yleisemmin, kaikille neliömuodoille $q = X^TAX$, ortogonaalinen matriisi P, jolle P^TAP on diagonaalinen, voidaan aina valita siten, että $\det P = 1$ vaihtamalla keskenään kaksi ominaisarvoa. Kirjan [2] luvussa 9 osoitetaan, että ortogonaaliset 2×2 matriisit joiden determinantti on 1, vastaavat akselien kääntöjä. Vastaava toteutuu myös 3×3, kun akseleita käännetään origon lävistäjän mukaisesti. Tämä laajentaa lausetta 5.2: Jokainen kahden tai kolmen muuttujan neliömuoto voidaan diagonalisoida kääntämällä koordinaatistoa.

Viitteet

[1] David C. Lay *Linear Algebra and Its Applications*, 4th ed. update, Addison-Wesley, 2012

[2] Nicholson W. Keith. *Linear Algebra with Applications*, 4th ed.,The McGraw-Hill Ryerson limited, a subsdiary of the McGraw-Hill Companies, 2003.