4.4 Hermiittiset muodot ja polaarinen hajotelma

Olkoon $\langle , \rangle$ sisätulo äärellisulotteisessa K-vektoriavaruudessa V. Olemme aikaisemmin näyttäneet, että jokainen lineaarinen muoto $L\colon V \to \mathbb{R}$ voidaan esittää sisätulon avulla muodosa L_w, missä $L(v) = \langle v, w\rangle$. Tämä aliluku aloitetaan tutkimalla samantyylinen ongelma Hermittisille muodoille.

Palautetaan mieleen, että Hermiittinen muoto $H\colon V \times V \to K$ on 3/2-lineaarinen kuvaus, joka toteuttaa lisäksi ehdon $H(x,y) = \overline{H(y,x)}$. Kuten tiedämme jo, tästä ehdosta erityisesti seuraa, että $H(x,x) \in \mathbb{R}$ kaikilla $x \in V$. Hermittinen muoto on *positiivisesti semidefiniitti* jos $H(x,x) \geq 0$ kaikilla $x \in V$. Se on *positiivisesti definiitti* jos $H(x,x) > 0$ kaikilla $x \neq 0$.

Olkoon $F\colon V \times V \to K$ 3/2-lineaarinen muoto ja olkoon $\mathbf{e} = (e_1, \dots, e_n)$ V:n ortonormaali kanta. Määritellään F:n matriisi $A = [F]_{\mathbf{e}} = (a_{ij}) \in M(n \times n; K)$ tässä kannassa ehdoilla

$$a_{ij} = F(e_i, e_j).$$

Lemma 4.41. *Olkoot $F, \mathbf{e}$ ja $A = [F]_{\mathbf{e}}$ kuten yllä. Tällöin*
a) kaikilla $x, y \in V$ pätee

$$F(x,y) = \langle Ax, y\rangle = \langle x, A^* y\rangle.$$

b) F on Hermittiinen jos ja vain jos A on itseadjungoitu matriisi.
c) F on positiivisesti semidefiniitti jos ja vain jos A on itseadjungoitu matriisi jonka ominaisarvot ovat kaikki ei-negatiivisia.
d) F on positiivisesti definiitti (eli on sisätulo V:ssä) jos ja vain jos A on itseadjungoitu matriisi jonka ominaisarvot ovat kaikki aidosti positiivisia.

Todistus. Esitetään x ja y kannassa $\mathbf{e}$,

$$x = \sum x_i e_i, \; y = \sum y_i e_i.$$

Tällöin, koska F on 3/2-lineaarinen, pätee

$$F(x,y) = \sum_{i,j} x_i \overline{y_j} F(e_i, e_j) = \sum_{i,j} x_i \overline{y_j} a_{ij}.$$

Toisaalta

$$Ax = \sum_{j=1}^{n} c_j e_j,$$

missä

$$c_j = \sum_{i=1}^{n} a_{ij} x_i.$$

Nyt

$$\langle Ax, y \rangle = \sum_{j=1}^{n} c_j \overline{y_j} = \sum_{i,j} a_{ij} x_i \overline{y_j}.$$

Näin ollen

$$F(x, y) = \langle Ax, y \rangle.$$

Toinen yhtälö $F(x, y) = \langle x, A^* y \rangle$ saadaan tästä suoraan adjungaatin määritelmän avulla.
b) Jos F on Hermittinen, niin

$$a_{ji} = F(e_j, e_i) = \overline{F(e_i, e_j)} = \overline{a_{ij}},$$

joten A on itseadjungoitu.
Kääntäen jos A on itseadjungoitu ja x, y ovat V:n alkiot, niin a)-kohdan nojalla

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

c-d) Proposition 4.37 nojalla jokainen itseadjungoitu matriisi A on diagonalisoituva ja lisäksi kaikki sen ominaisarvot ovat reaalisia. Olkoon $\mathbf{e} = (e_1, \ldots, e_n)$ V:n ortonormaali kanta, joka koostuu A:n ominaisarvoista, tällöin $A(e_i) = r_i e_i$, missä $r \in \mathbb{R}$. Tässä siis ajatelemme matriisi lineaarisena kuvauksena $V \to V$. Olkoon $x \in V$ ja esitetään se tässä kannassa muodossa $x = \sum x_i e_i$. Tällöin

$$Ax = \sum_{i=1}^{n} r_i x_i,$$

joten a)-kohdan nojalla

$$F(x, x) = \langle Ax, x \rangle = \sum_{i=1}^{n} r_i x_i \overline{x_i} = \sum_{i=1}^{n} r_i \left| x_i \right|^2.$$

Nyt jos F on positiivisesti semidefiniitti Hermiittinen muoto, niin erityisesti

$$r_i = F(e_i, e_i) \geq 0$$

kaikilla $i = 1, \ldots, n$ ja yhtälö on aito jos F on positiivisesti definiitti.
Kääntäen jos A:n ominaisarvot ovat ei-negatiivisia, jokaisella $x \in V$ pätee

$$F(x, x) = \sum_{i=1}^{n} r_i \left| x_i \right|^2 \geq 0.$$

Näin ollen F on tällöin positiivisesti semidefiniitti. Jos taas kaikki ominaisarvot ovat aidosti positiivisesti ja $x \in V$ yllä pätee $F(x,x) > 0$ ja F on positiivisesti definiitti. $\square$

Sanomme itseadjungoitu matriisi $A \in M(n \times n; K)$ *positiivisesti semidefiniitiksi* jos sen ominaisarvot ovat ei-negatiivisia ja vastaavasti positiivisesti definiitiksi jos sen ominaisarvot ovat kaikki ei-negatiivisia. Puhutaan myös negatiivisesti definiitista (kaikki ominaisarvot negatiivisia) ja negatiivisesti semi-definiitista (kaikki ominaisarvot ei-positiivisia) itseadjunoiduista matriiseista.
Edellisesta Lemmasta seuraa, että Hermiitisen muodon $F\colon V \times V \to K$ matriisi $A = [F]_{\mathbf{e}}$ jonkun ortonormaalin kannan $\mathbf{e}$ suhteen on positiivisesti semidefiniitti jos ja vain jos kaikilla $v \in V$

$$F(v,v) \geq 0.$$

Erityisesti F on sisätulo jos ja vain jos A on positiivisesti definiitti.
Saamme siis seuraavan tuloksen.

Lemma 4.42. *Olkoon $(V; \langle , \rangle)$ äärellisulotteinen sisätuloavaruus, jossa siis kiinitämme eräs sisätulo $\langle , \rangle$. Tällöin mikä tahansa muu sisätulo $(,)$ samassa avaruudessa voidaan määritellä kaavalla*

$$(x,y) = \langle Ax, y \rangle$$

jollakin positiivisesti definiitilla itseadjungoidulla matriisilla A. Erityisesti K^n:ssä voidaan mielivaltainen sisätulo $\langle , \rangle$ määritellä pistetulon avulla,

$$\langle x, y \rangle = (Ax) \cdot y,$$

missä A positiivisesti definitti matriisi.

Itdeadjunoidu matriisi A on ei-definiitti, jos lauseke $\langle Ax, x \rangle$ saa sekä aidosti positiivisia, että aidosti negatiivisia arvoja eli ei ole positiivisesti semidefiniitti ja negatiivisesti semidefiniitti.

Esimerkki 4.43. *Olkoon $f\colon \mathbb{R} \to \mathbb{R}$ kahdesti derivoituva funktio. Tunnetusti sen* ääriarvot *sijaitsevat derivaatan f' nollakohdissa eli* kriittisissä pisteissä*. Toisaalta kääntäinen ei päde eli kriittisessä pisteessä ei vältämättä ole ääriarvoja ja vaikka olisikin, pitää vielä selvittää onko se lokaali maksimi vai minimi. Tämän selvittämiseksi on kehitetty kaikenlaisia menetelmiä, joista yksi perustuu toisen derivatan f'' tutkimiseen. Jos kriittisessä pisteessä a pätee $f''(a) > 0$ niin siellä funktiolla on (aito) lokaali minimi, jos taas $f''(a) < 0$ - aito lokali maksimi.*

Tätä lähestymistapa voidaan yleistää funktioihin $f\colon \mathbb{R}^n \to \mathbb{R}$*, jolloin toinen derivaatta korvataan niin sanotulla* Hessen matriisilla*, jonka kertoimet ovat* f*:n toista kertalukua olevat osittaisderivaatat (annetussa pisteessä* x_0*). Tämä on siis matriisi*

$$H = \begin{bmatrix} D_{11}(x_0) & D_{12}(x_0) & \dots & D_{1n}(x_0) \\ D_{21}(x_0) & D_{22}(x_0) & \dots & D_{2n}(x_0) \\ \dots & \dots & \dots & \dots \\ D_{n1}(x_0) & D_{n2}(x_0) & \dots & D_{nn}(x_0) \end{bmatrix}.$$

Oletetaan, että funktio f *on vähintään* C^2*-luokkaa, eli sillä on kaikki toisen kertalun jatkuvat osittaisderivaatat. Analyysin kurssilla todistetaan, että tällöin matriisi on symmetrinen (peräkkäisessä osittaisderivoinnissa lopputulos ei riippu siitä, missä järjetyksessä otetaan osittaisderivaata).*
C^2*-funktiolle* $f\colon \mathbb{R}^n \to \mathbb{R}$ *on voimassa* toisen kertaluvun Taylorin kaava

$$f(x) = f(a) + \sum_{i=1}^{n} D_i f(x_0)(x_i - a_i) + \frac{1}{2}(\sum_{i,j} D_{ij}(x_0)(x_i - a_i)(x_j - a_j)) + h(x),$$

missä h *menee nollaan kun* $x \to x_0$*.*
Vältämättön (mutta ei riittävä) ehto sille, että f*:llä on pisteessä* x_0 *ääriarvo on se, että* x_0*:ssä on krittinen piste, eli kaikki sen ensimmäisen kertaluvun osittaisderivat häviää siinä pisteessä. Toisin sanoen täytyy olla* $D_i(x_0) = 0$ *kaikilla* $i = 1, \dots, n$*. Näin ollen krittisessä pisteessä Taylorin kaava antaa*

$$f(x) - f(a) = \frac{1}{2}(H(x-a) \cdot (x-a)) + h(x).$$

Näin ollen, jos H *on positiivisesi definittii, tarpeeksi lähellä* a*:ta* $f(x) - f(a) > 0$ *kun* $x \neq a$*, joten* a*:ssä on lokaali minimi. Jos taas* H *on negatiivisesti definiitti, pisteessä* a *funktiolla on lokaali maksimi. Jos* H *on ei-definiitti,* a*:ssä ei ole ääriarvoa.*
Matriisin H *definiittisyyttä voimme tutkia lineaarialgebrallisin keinoin. Voidaan osoittaa esimerkiksi, että* $\mathbb{R}$*-kertoiminen symmetrinen matriisi* A *on positiivisesti definiitti jos ja vain jos* $\det A_1 = a_{11} > 0, \det A_2 > 0, \dots, \det A_{n-1} > 0, \det A_n = \det A > 0$*. Tässä* A_i *on* $(i \times i)$*-matriisi, jonka sarakkeet ovat* A*:n sarakkeiden* $A^1, \dots, A^i$ *rajoittumat riviin* i *asti.*
Tämä tulos on käytännössä erittäin hyödyllinen, sillä determinaanttien laskiminen on helpompi kuin esim. ominaisarvojen laskeminen.

Lemma 4.44. *Olkoon* $A \in M(n \times n; K)$*. Tällöin* A *on positiivisesti semidefiniitti jos ja vain jos on olemassa matriisi* $B \in M(n \times n; K)$ *jolle* $A = BB^*$*. Lisäksi, jos* A *on positiivisesti semidefiniitti on olemassa jopa yksikäsitteinen positiivisesti semidefiniitti matriisi* B *jolle* $A = B^*B = B^2$*.*

Todistus. Vältämättömyys on helppo osoittaa - jos $A = BB^*$, niin $A^* = (BB^*)^* = (B^*)^*B^* = BB^* = A$. Lisäksi jos $BB^*(x) = rx$ jollakin negatiivisella $r \in \mathbb{R}$, saadaan

$$0 \leq \langle B^*x, B^*x \rangle = \langle BB^*x, x \rangle = r\langle x, x \rangle,$$

mikä on mahdollista ainoastaan kun $x = 0$. Näin ollen B^*B:n ominaisarvot ei voi olla negatiivisia.
Oletetaan kääntäen, että A on positiivesesti semidefiniitti. Näytetään, että on olemassa yksikäsitteinen positiivisesti semidefiniitti matriisi B jolle $B^2 = BB^* = A$.
Olkoon B sellainen positiivisesti semidefiniitti matriisi, että $B^2 = A$. Olkoot $r_1, \dots, r_m$ kaikki sen erilaiset ominaisarvot. Tällöin kaikki ne ovat reaalisia ja ei-negatiivia ja lisäksi

$$V = \oplus_{i=1}^m V_{r_i}.$$

Toisaalta, koska A on positiivesesti semidefiniitti, pätee yhtähyvin

$$V = \oplus_{j=1}^p W_{s_j},$$

missä $s_1, \dots, s_p$ kaikki A:n ominaisarvot, myös reaaliset ja ei-negatiiviset.
Koska $A = B^2$, aliavaruudessa V_{r_i} pätee

$$A(x) = B^2(x) = r_i^2 x.$$

Näin ollen jokaisella $i = 1, \dots, m$, r_i^2 on jokin s_j. Koska kuvaus $x \mapsto x^2$ on positiivisten reaalilukujen joukossa bijektiivinen, $m = p$ ja voimme olettaa, että $r_i^2 = s^i$ jokaisella $i = 1, \dots, m$. Lisäksi olemme näyttäneet, että tällöin $V_{r_i} \subset W_{s_i}$, joten pakko olla $V_{r_i} = W_{s_i}$.
Kuvaus B on siis yksikäsitteinen ja määritelty jokaisessa A:n ominaisarvoaliavaruudessa W_s kaavalla $B(x) = \sqrt{s}x$. Koska voidaan myös määritellä lineaarinen kuvaus, joka on tällä tavalla määritelty, B on myös olemassa. $\square$

Propositio 4.45. Polaarinen hajotelma.
Jokainen neliömatriisi $A \in M(n \times n; K)$ voidaan esittää muodossa US, missä $S \in M(n \times n; K)$ itseadjungoitu positiivisesti semidefiniitti ja $U \in M(n \times n; K)$ unitaarinen. Lisäksi S on yksikäsitteinen. U on yksikäsitteinen jos A on kääntyvä. Tällöin S on jopa positiivisesti definiitti.

Todistus. Jos $A = US$, missä S positiivisesti semidefiniitti ja U unitaarinen, niin

$$A^*A = S^*U^*US = S^2.$$

Edellisen proposition mukaan A^*A on positiivisesti semidefiniitti ja on olemassa yksikäsitteinen positiivisesti semidefiniitti S siten, että $S^2 = A^*A$.

Se osoittaa sen, että S on yksikäsitteisesti määritelty. Jos A on kääntyvä, lopputodistus on helppo - ensinnäkin tällöin

$$(\det S)^2 = \det(S^2) = \det(A^*A) = \det A^* \det A = \overline{\det A} \det A = |\det A|^2 > 0,$$

mikä osoittaa sen, että myös S on kääntyvä. Siitä seuraa, että sen ominaisarvot ovat aidosti positiiviset, joten S positiivisesti definiitti ja lisäksi pakko olla $U = AS^{-1}$. Jäljellä vain sen osoittaminen, että tällä tavalla määritelty U on unitaarinen. Se on helppo lasku -

$$U^*U = (S^{-1})^*A^*AS = S^{-1}S^2S^{-1} = \mathrm{id}\,.$$

Tässä olemme käyttäneet tietoa siitä, että kääntyvän itseadjungoidun matriisin käänteismatriisi on myös itseadjungoitu (helppo tarkistaa).

Jos A ei ole kääntyvä, ei ole myös S:kään joten näin ei voi menetellä. Idea on seuraava - jos tällainen U on olemassa, niin, koska U säilyttää normit, täytyisi olla

$$|Ax| = |U(Sx)| = |Sx|\,.$$

Todistus etenee käänteisessä järjestyksessä - aloitetaan näyttämällä, että edellinen yhtälö on voimassa. Ajatelemme matriisit lineaarisina kuvauksina $K^n \to K^n$ (missä avaruus varustetaan tavallisella pistetulolla). Olkoon $x \in K^n$. Koska $S^2 = A^*A$ ja S on itseadjunoitu, saadaan

$$|Ax|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle = \langle S^2x, x\rangle =$$
$$= \langle Sx, Sx\rangle = |Sx|^2\,.$$

Näin ollen $|Ax| = |Sx|$ kaikilla $x \in K^n$. Koska haluamme unitarisen kuvauksen (matriisin) U jolle $US = A$, S:n kuvajoukossa $\operatorname{Im} S$ U:n täytyy toteuttaa ehdon $Uy = Ax$, missä $y = Sx$. Käytämme tätäkin havaintoa konstruktiossa hyväksi. Nimittäin, määritellään U ensin aliavaruudessa $W = \operatorname{Im} S$ kaavalla $Uw = Sx$, jossa x on sellainen, että $w = Sx$. Koska S ei ole vältämättä kääntyvä, tällainen x ei vältämättä ole yksikäsitteinen, joten meidän osoittaa ensin, että määritelmä ei riipu x:n valinnasta. Meidän on siis näytettävä, että jos $Sx = Sy$, niin $Ax = Ay$. Koska molemmat S ja A ovat lineaarisia, riittää näyttää, että jos $Su = 0$, myös $Au = 0$ (eli $\operatorname{Ker} S \subset \operatorname{Ker} A$). Tarvitsemamme väite sitten seuraa tästä, kun asetetaan $u = x - y$.
Mutta jos $Su = 0$, niin

$$|Au| = |Su| = 0,$$

mistä seuraa, että myös $Au = 0$.
Kuvaus $U\colon W \to K^n$ on siis hyvin määritelty ja toteuttaa ehdon $US = A$. Lisäksi se säilyttää normit, sillä jos $y = Sx \in W = \operatorname{Im} S$, $x \in K^n$, niin

$$|Uy|^2 = |USx|^2 = |Ax|^2 = |Sx|^2 = |y|^2\,.$$

Kuvaus U on siis unitarinen (Lemma 4.22), erityisesti injektio (sama Lemma 4.22). Näin ollen $\operatorname{Im} U = W'$ on sama ulottuvuutta kuin W. Todistus on valmis, kunhan osoitetaan vielä, että U voidaan laajentaa koko avaruuden unitaariseksi kuvaukseksi.
Esitetään K^n suorina summina $K^n = W \subset W^\perp = W' \subset W'\perp$. Koska $\dim W = \dim W'$, myös $\dim W^\perp = n - \dim W = n - \dim W' = \dim W'\perp$. Valitaan molemmille alivaruudelle $W^\perp$ ja $W'^\perp$ ortonormaalit kannat $(e_1, \ldots, e_k)$ ja $(f_1, \ldots, f_k)$ ja määritellään lineaarinen $U' \colon W^\perp \to W'\perp$ ehdolla $U'(e_i) = f_i$, $i = 1, \ldots, k$. Lemman 4.22 nojalla U' on unitaarinen. Laitetaan näin määritellyt kuvaukset $U \colon W \to W'$ ja $U' \colon W^\perp \to'^\perp$ "yhteen" määritelemällä kuvaus $U \colon K^n \to K^n$ kaavalla $U(x + y) = U(x) + U'(y)$, $x \in W$, $y \in W^\perp$. Tämä on hyvin määritelty ja yksikäsitteinen, koska $W \oplus W^\perp = K^n$ (Propositio 4.15). Helposti nähdään, että näin määritelty kuvaus on unitaarinen. Propositio todistettu.
Kun A ei ole kääntyvä, ei ole myöskään S (sillä $\det S = |\det A|$), joten $W^\perp$ on aito aliavaruus. Tällöin on olemassa ainakin kaksi tapaa jatkaa annettu unitaarinen kuvaus $W \colon W'$ koko avaruuden unitaariseksi kuvaukseksi, joten kun A ei ole kääntyvä polaarinen hajotelma ei ole koskaan yksikäsitteinen. □

Kun edellinen tulos sovelletaan matriisin A konjugaattiin, saadaan esitys $A^* = US$, missä U unitaarinen ja S positiivisesti semidefiniitti. Ottamalla tästä konjugaatin, saadaan A:lle esitys $A = SU'$, missä U' on unitaarinen. Tätäkin hajotelma sanotaan polaariseksi.
On olemassa hyödyllinen analogiaa matriisien ja kompleksilukujen välillä. Voidaan ajatella, että itseadjungoidut matriisit vastaavat reaalilukuja, positiivisesti definiitit matriisit posiitivisia reaalilukuja, unitaariset matriisit yksikköympyrän alkiota $z \in S^1, |z| = 1$. Tällöin edellä todistettu polaarinen hajotelma $A = SU$ vastaa tuttua kompleksilukujen polaarista esitystä $z = rt$, missä $r = \sqrt{|z|} \geq 0$ ja $t \in S^1$. Tästä nimitys "polaarihajotelma" tuleekin.

Sovellus integraalilaskennassa.
Yksiulotteisen (Riemann)-integraalin "muuttujavaihtokaava" on muotoa

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(x)dx.$$

Kun tätä yleistetään moniulotteiseen avaruuteen eli funktiolle $f \colon \mathbb{R}^n \to \mathbb{R}$, $g \colon \mathbb{R}^n \to \mathbb{R}^n$ derivaatta $g'(x)$ pitää jäljeen kerran korvata lineaarikuvauksen $g'(x) \colon \mathbb{R}^n \to \mathbb{R}^n$ determinantilla ja kaava näyttää seuraavalta -

$$\int_{g(A)} f(x)dx = \int_E (f \circ g)(x) \left|\det g'(x)\right| dx.$$

Tässä tietenkin pitää olettaa, että integroituva alue A sekä funktiot f, g ovat tarpeeksi "siistiä" ja säännöllisiä.
Muuttujavaihtokaava todistetaan seuraavaksi - ensin käytetään hyväksi sitä, että $g'(x)$ on g:n "lineaarinen approksimaatio", joten lokaalisti g voidaan korvata derivaatallaan ja riittää osoittaa, että kaava pätee kun $g = L$ on lineaarinen kuvaus. Tämä taas palautuu yksinkertaisempaan väitteeseen, jonka mukaan

$$m(L(E)) = |\det L|\, m(A), \tag{4.46}$$

missä m on "mitta" eli geometrisen joukon "tilavuus" (pinta-ala jne.) $\mathbb{R}^n$:ssä, A tarpeeksi siisti integroituva alue ja $L\colon \mathbb{R}^n \to \mathbb{R}^n$ lineaarinen kuvaus. Näiden välivaiheiden täsmällinen perustelu ja läpikäynti on tietysti pitkä prosessi täynnä yksityiskohtia, joihin emme voi tässä mennä.
Todistuksen ydin pilee kuitenkin kaavan 4.46 todistamisessa. Se taas voidaan tehdä polaarisen hajotelman avulla. Nimittäin voimme kirjoittaa L:n matriisi (standardin kannan suhteen) muodossa $A = OS = OO'DO'^{-1}$, missä D diagonaalinen ja O, O' ortogonaalisia. Tässä siis käytämme hyväksi myös sitä, että symmetrinen matriis S on diagonalisoituva ortonormaalissa kannassa.
Koska $\det L = \det A = \det O \det O' \det D \det O'^{-1}$, väite riittää todistaa kun A on joko diagonalinen matriisi tai ortogonaalinen.
Molemmat ovat helppoja - diagonaalinen matriisi vain venyttää koordinaatisuorat ja voimme laskea suoraan miten kappaleen mitta muuttuu tällaisessa muunnoksessa. Ortogonaalinen kuvaus taas säilyttää etäisyydet, joten myös säilyttää pallot ja niiden mitat. Koska jokaisen joukon mitta voidaan approksimoida mielivaltaisen tarkasti pallojen avulla (tämä osoitetaan esim. Mitta ja integraali-kurssilla), tästä seuraa, että $m(O(A)) = m(A) = |\det O| m(A)$, sillä ortogonaalisen kuvauksen determinaantin itseisarvo on aina 1. Siis kaava 4.46 pätee sekä ortogonaalisille että diagonalisille matriiseille, polaarihajotelman avulla nähdään, että se pätee kaikille L.