

7 LINEAARIAVARUDET

Ch:LinSpaces

7.1 Rakenteiden hierarkiaa.

Sec:AlgStruct

Matematiikan historian suurimmat saavutuksen on saatu, kun on tutkittu rakenteita. Seuraavassa luetellaan ne rakenteet, jotka ovat parhaiten kestäneet ajan kulutusta. Seuraavan parin sivun sisältämä luettelo kannattaa ensin lukea kursoorisesti (pysähtymättä yksityiskohtiin). Annettuja määritelmiä ei kaikkia ole nyt numeroitu, ja numeroimattomia ei kysytä tentissä. Jotkut luettelossa olevat termit esiintyvät usein myöhemmillä kursseilla. Nämä määritelmät on numeroitu, ja ne pitää muistaa.

Joukossa A määritelty laskutoimitus $\hat{+}$ on kuvaus $\hat{+} : A \times A \rightarrow A, (a, b) \mapsto a \hat{+} b$. Joskus joukon A alkio kerrotaan joukon K alkiolla. Tälle kertolaskulle käytämme laskutoimituksen kaltaista merkintää $\circ : K \times A \rightarrow A, (k, a) \mapsto k \circ a$.

ath1040c4s1es1

Esimerkki 123 Olkoon $A = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ ja määritellään laskutoimitukset

$$a \hat{+} b = \text{jakojäännös, kun } a + b \text{ jaetaan } 5\text{:llä, ja} \quad (7.1)$$

$$a \hat{*} b = \text{jakojäännös, kun } a \cdot b \text{ jaetaan } 5\text{:llä.} \quad (7.2)$$

Silloin laskutoimitusten laskutaulut ovat seuraavat:

$\hat{+}$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\hat{*}$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Näillä laskutoimituksilla on monia hyvin tuttuja ominaisuuksia. Esim:

$$a \hat{+} (b \hat{+} c) = (a \hat{+} b) \hat{+} c \quad (7.3)$$

$$a \hat{*} (b \hat{+} c) = (a \hat{*} b) \hat{+} (a \hat{*} c) \quad (7.4)$$

Jos tämä oli ensimmäinen kerta, kun luit tästä joukosta ja näistä laskutoimituksista, niin on luonnollista, jos asia tuntuu keinotekoiselta. Myöhemmin tulet tapaamaan tämän tai vastaavan rakenteen uudelleen.

On syntynyt tapa sanoa, että kun oliojoukossa A on määritelty laskutoimitus $\hat{+}$ ja pari $(A, \hat{+})$ toteuttaa joukon aksioomia, niin parilla $(A, \hat{+})$ on tähän aksioomajoukkoon liittyvä strukturi.

Parhaiten ajan koetusta ovat kestäneet seuraavat struktuurit:

Ryhmä. Pari $(A, \hat{+})$ on ryhmä (group), jos se toteuttaa seuraavat aksioomat:
 (1) $a\hat{+}(b\hat{+}c) = (a\hat{+}b)\hat{+}c$, kaikille $a, b, c \in A$, (eli laskutoimitus on liitännäinen)
 (2) On olemassa neutraalialkio $o \in A$ siten, että $o\hat{+}a = a\hat{+}o = a$ kaikille $a \in A$
 (3) jokaiselle $a \in A$ on olemassa vasta-alkio $a' \in A$ siten, että $a'\hat{+}a = a\hat{+}a' = o$.

Abelin ryhmä. Ryhmä $(A, \hat{+})$ on Abelin ryhmä (abelian group), jos se edellisen lisäksi toteuttaa aksiooman:

(4) $a\hat{+}b = b\hat{+}a$, kaikilla $a, b \in A$, (eli laskutoimitus on vaihdannainen).

Rengas. Kolmikko $(A, \hat{+}, \hat{*})$ on rengas (ring), jos

- (1) $(A, \hat{+})$ on Abelin ryhmä,
 (2) $a\hat{*}(b\hat{*}c) = (a\hat{*}b)\hat{*}c$, kaikilla $a, b, c \in A$, (eli kertolasku on liitännäinen),
 (3) on voimassa osittelulait:

$$a\hat{*}(b\hat{+}c) = (a\hat{*}b)\hat{+}(a\hat{*}c) \quad (7.5)$$

$$(b\hat{+}c)\hat{*}a = (b\hat{*}a)\hat{+}(c\hat{*}a) \quad (7.6)$$

kaikilla $a, b, c \in A$.

Vaihdannainen (kommutatiivinen) rengas. Rengas $(A, \hat{+}, \hat{*})$ on vaihdannainen rengas (commutative ring), jos edellisen lisäksi

(4) $a\hat{*}b = b\hat{*}a$, kaikilla $a, b \in A$, (eli kertolasku on vaihdannainen).

Ykkösellinen rengas. Rengas $(A, \hat{+}, \hat{*})$ on ykkösellinen rengas (ring with unit), jos

(4) on olemassa yksikkö-alkio e , jolle $e\hat{*}a = a\hat{*}e = a$ kaikille $a \in A$.

Jakorengas. Ykkösellinen rengas $(A, \hat{+}, \hat{*})$ on jakorengas (division ring), jos

(5) $e \neq o$,

(6) jokaiselle $a \in A, a \neq o$ on olemassa kertolaskun käänteisalkio a^{-1} siten, että $(a^{-1})\hat{*}a = a\hat{*}(a^{-1}) = e$.

Kunta. Kommutatiivinen jakorengas $(A, \hat{+}, \hat{*})$ on kunta (field).

Kunnan määritelmä on tässä kumulatiivisessa määritelmäketjussa hämäävän yksinkertainen. On syytä koota yhteenveto kunnan laskutoimitusten ominaisuuksista.

(1) Yhteenlasku on vaihdannainen ja liitännäinen, sillä on neutraali-alkio 0 ja jokaisella alkiolla a on yhteenlaskun vasta-alkio $-a$.

(2) Kertolasku on vaihdannainen ja liitännäinen, sillä on ykkösalkio e ja jokaisella yhteenlaskun neutraali-alkiosta poikkeavalla alkiolla a on kertolaskun käänteisalkio a^{-1} . Tästä seuraa, että jakolasku on mahdollista.

(3) Yhteen- ja kertolasku noudattavat osittelulakeja, joten lausekkeiden sieventämisessä voidaan käyttää tavanomaisia numerolaskennasta tuttuja keinoja.

Siis kaikki mahdolliset numerolaskennasta tutut laskulait ovat voimassa.

Esimerkki 124 Reaalilukujen joukko varustettuna tavallisella yhteen- ja kertolaskulla, $(\mathbb{R}, +, \cdot)$ on kunta. $(\mathbb{Z}, +, \cdot)$ ei ole kunta. Esimerkissä (I23) esiintynyt struktuuri $(\mathbb{Z}_5, \hat{+}, \hat{*})$ on kunta.

Olkoon $(A, \hat{+})$ Abelin ryhmä ja $(K, \bar{+}, \bar{*})$ ykkösellinen rengas. Olkoon lisäksi määritelty skalaarilla kertominen $\circ : K \times A \rightarrow A$.

Moduli. Nelikko $(A, (K, \bar{+}, \bar{*}), \hat{+}, \circ)$ on K-moduli (K -module), jos

- (1) $(A, \hat{+})$ on Abelin ryhmä,
- (2) $(K, \bar{+}, \bar{*})$ on ykkösellinen rengas (ns. kerroinrengas),
- (3) skalaarilla kertominen toteuttaa aksioomat:

$$\lambda \circ (\mu \circ a) = (\lambda \bar{*} \mu) \circ a, \quad \text{kaikilla } \lambda, \mu \in K, a \in A$$

$$(\lambda \bar{+} \mu) \circ a = \lambda \circ a \hat{+} \mu \circ a, \quad \text{kaikilla } \lambda, \mu \in K, a \in A$$

$$\lambda \circ (a \hat{+} b) = \lambda \circ a \hat{+} \lambda \circ b, \quad \text{kaikilla } \lambda \in K, a, b \in A$$

$$e \circ a = a, \quad \text{kaikilla } a \in A. \quad (\text{Tässä } e \text{ on kerroinrenkaan } K \text{ yksikköalkio.})$$

Lineaariavaruus. K -moduli $(A, K, \hat{+}, \circ)$ on K-kertoiminen lineaariavaruus (linear space over K), jos kerroinrengas $(K, \bar{+}, \bar{*})$ on kunta;

Algebra. Viisikko $(A, K, \hat{+}, \hat{*}, \circ)$ on K -algebra, jos

- (1) $(A, K, \hat{+}, \circ)$ on K -kertoiminen lineaariavaruus,
- (2) $(A, \hat{+}, \hat{*})$ on rengas ja
- (3) skalaarilla kertominen ja kertolasku toteuttavat aksiooman:

$$\lambda \circ (a \hat{*} b) = (\lambda \circ a) \hat{*} b = a \hat{*} (\lambda \circ b), \quad \text{kaikille } \lambda \in K, a, b \in A$$

Jakoalgebra. K -algebra $(A, K, \hat{+}, \hat{*}, \circ)$ on K -jakoalgebra, jos

- (2) rengas $(A, \hat{+}, \hat{*})$ on jakorengas.

Edellä esiintyvät aksioomat ovat niitä laskulakeja, joita numeroilla laskemaan tottunut ihminen pitää itsestäänselvyyksinä. Matriisien tapauksessa havaitsimme, että eräs numerolaskennan itsestäänselvyys - kertolaskun vaihdantalaki ole voimassa. Tämä on tärkeä havainto, kun nyt ryhdymme laskemaan numeroiden sijasta yleisemmillä olioilla (vektoreilla, matriiseilla, funktioilla jne.)

Kun tarkastelemme jotakin oliojoukkoa ja siinä määriteltyjä laskutoimituksia, niin mitä vahvempi struktuuri sillä on asteikolla "ryhmä < rengas < jakorengas < kunta" tai asteikolla " K -moduli < K -kertoiminen lineaariavaruus < K -algebra < K -jakoalgebra", sitä helpompaa on olioilla laskeminen. Mitä useampi aksiooma on voimassa, sitä useampi numerolaskennan itsestäänselvyys on voimassa.

Kahdessa edellisessä kappaleessa perehdyimme matriiseihin. Neliömatriisit muodostavat renkaan, mutta eivät kommutoivaa rengasta eivätkä jakorengasta. Kun neliömatriisien joukko varustetaan kerroinkunnalla R , ne muodostavat R -algebran, mutta eivät R -jakoalgebraa.

Jatkossa keskitymme alkeellisempaan struktuuriin, nimittäin lineaariavaruuteen, jonka kerroinkunta on reaalilukujen kunta. Tämä merkitsee siis sitä, että olioiden tuloa ei ole määritelty, mutta niiden summa ja olioiden kertominen reaaliluvulla noudattavat kaikkia vanhastaan tuttuja laskulakeja.

7.2 $\mathbb{R}$ - ja $\mathbb{C}$ -kertoiminen lineaariavaruus

Sec:RCLinSpace

Tässä esityksessä tärkein ja eniten käsitelty lineaariavaruus on Euklidinen vektoriavaruus $(\mathbb{R}^n, \mathbb{R}, +, \circ) = \mathbb{R}^n = E^n$ (engl. Euclidean Vector Space), jonka kerroinkunta on $\mathbb{R}$.

Merkintä $\mathbb{R}^n$ on loogisempi, mutta merkintää E^n käytetään paljon. Merkintä korostaa, sitä, että vektoriavaruus ei ole pelkkä karteesinen tulo-joukko, vaan sillä on vektoriavaruuden rakenne, eli yhteenlasku ja kerroinkunta. Kumpikin merkintä esiintyy jatkossa.

Kompleksinen lineaariavaruus $(\mathbb{C}^n, \mathbb{C}, +, \circ) = \mathbb{C}^n$ (engl. Complex Vector Space) on tärkeä, mutta se jätetään nyt hetkeksi aikaa taka-alalle. Asiaan palataan kappaleessa ^{$\mathbb{C}^n$} (??) sivulla ??.

Usein virheellisesti lineaariavaruus käsitetään samaksi, kuin E^n , mutta tämä ei ole oikein. Monet funktiojoukot voidaan helposti varustaa lineaariavaruuden struktuurilla. Tällaiset lineaariset funktioavaruudet ovat erittäin hyödyllisiä olioita ja monessa suhteessa E^n :stä poikkeavia. Seuraavassa yleiset lineaariavaruuksiin liittyvät määritelmät ja tulokset muotoillaan koskemaan reaalikertoimista lineaariavaruutta $L = (L, \mathbb{R}, +, \circ)$. Tällainen tulos on sellaisenaan voimassa myös, kun $L = E^n$. Jos väite esitetään E^n :lle, niin se ei välttämättä ole voimassa kaikille lineaariavaruuksille.

Puramme ensin edellisen kappaleen hierarkiseen määritelmäketjuun piiloutuneet ominaisuudet seuraavaan määritelmään. Huomaa, että tässä oikeastaan toistetaan jo todettua.

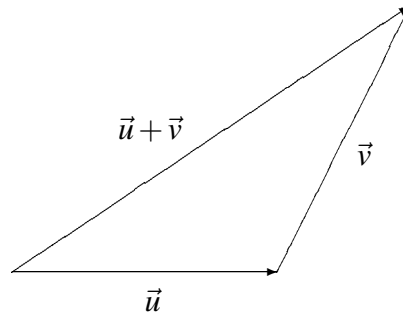
Määritelmä 125 Nelikko $(L, \mathbb{R}, +, \circ)$ on reaalinen lineaariavaruus (eli reaalinen vektoriavaruus), jos yhteenlasku $L \times L \rightarrow L, (\vec{u}, \vec{v}) \mapsto \vec{a} + \vec{b}$ ja reaaliluvulla kertominen

$\mathbb{R} \times L \rightarrow L, (\lambda, \vec{u}) \mapsto \lambda \circ \vec{u}$ toteuttavat seuraavat ehdot:

- (A1) $\vec{u} + \vec{v} = \vec{v} + \vec{u},$ kaikilla $\vec{u}, \vec{v} \in L,$
- (A2) $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ kaikilla $\vec{u}, \vec{v}, \vec{w} \in L,$
- (A3) on olemassa (yksikäsitteinen) nolla-alkio $\vec{0} \in L$ siten, että $\vec{u} + \vec{0} = \vec{u}$ kaikilla $\vec{u} \in L,$
- (A4) jokaista $\vec{u} \in L$ vastaa (yksikäsitteinen) vasta-alkio $-\vec{u}$ siten, että $\vec{u} + (-\vec{u}) = \vec{0},$ (merkitään $\vec{v} - \vec{u} = \vec{v} + (-\vec{u})$)
- (A5) $\lambda(\mu \circ \vec{u}) = (\lambda \cdot \mu) \circ \vec{u},$ kaikilla $\lambda, \mu \in \mathbb{R}, \vec{u} \in L,$
- (A6) $(\lambda + \mu) \circ \vec{u} = \lambda \circ \vec{u} + \mu \circ \vec{u},$ kaikilla $\lambda, \mu \in \mathbb{R}, \vec{u} \in L,$
- (A7) $\lambda \circ (\vec{u} + \vec{v}) = \lambda \circ \vec{u} + \lambda \circ \vec{v},$ kaikilla $\lambda \in \mathbb{R}, \vec{u}, \vec{v} \in L,$
- (A8) $1 \circ \vec{u} = \vec{u}$ kaikilla $\vec{u} \in L,$

Avaruuden L alkioita sanomme vektoreiksi. Jos laskutoimitukset ovat asiayhteydestä selvät sanomme, että ” L on reaalinen lineaariavaruus”.

Esimerkki 126 Tason vektoreiden joukko on reaalinen lineaariavaruus



Tason vektori voidaan esittää lukuparina $\vec{u} = (x, y)$, jolloin $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ ja $\lambda \circ (x, y) = (\lambda \cdot x, \lambda \cdot y)$. Suora lasku osoittaa, että nolla-alkio on $(0, 0)$ ja aksioomat (A1 - A8) ovat voimassa.

Esimerkki 127 Joukko $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n)^T \mid x_i \in \mathbb{R}, \text{ kaikille } 1 \leq i \leq n\}$ varustettuna pystyvektorien yhteenlaskulla ja reaalityluvulla kertomisella

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad \lambda \circ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot a_1 \\ \lambda \cdot a_2 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix}$$

on reaalinen lineaariavaruus $E^n = (\mathbb{R}^n, \mathbb{R}, +, \circ)$.

Esimerkki 128 $m \times n$ -matriisien joukko on reaalinen lineaariavaruus.

Esimerkki 129 Olkoon $A \neq \emptyset$ ja olkoon $\mathcal{F}(A, \mathbb{R})$ kaikkien kuvausten $f: A \rightarrow \mathbb{R}$ joukko. Joukko $\mathcal{F}(A, \mathbb{R})$ on reaalinen lineaariavaruus, kun määrittelemme

$$\begin{aligned}(f+g)(x) &= f(x) + g(x), \quad \forall x \in A \\ (\lambda \circ f)(x) &= \lambda \cdot f(x), \quad \forall x \in A\end{aligned}$$

Lause 130 Olkoon L reaalinen lineaariavaruus. Silloin

$$\begin{aligned}(1) \quad 0 \circ \vec{u} &= \vec{0}, \quad \forall \vec{u} \in L, \\ (2) \quad \lambda \circ \vec{0} &= \vec{0}, \quad \forall \lambda \in \mathbb{R}, \\ (3) \quad [\lambda \circ \vec{u} = \vec{0}] &\Leftrightarrow [\lambda = 0 \text{ tai } \vec{u} = \vec{0}], \\ (4) \quad (-1) \circ \vec{u} &= -\vec{u}, \quad \forall \vec{u} \in L.\end{aligned}$$

Todistus.

$$\begin{aligned}(1) \quad 0 \circ \vec{u} &\stackrel{(A3)}{=} 0 \circ \vec{u} + \vec{0} \stackrel{(A4)}{=} 0 \circ \vec{u} + (0 \circ \vec{u} + (- (0 \circ \vec{u}))) \stackrel{(A3)}{=} (0 \circ \vec{u} + 0 \circ \vec{u}) - (0 \circ \vec{u}) \\ &\stackrel{(A6)}{=} (0 + 0) \circ \vec{u} - (0 \circ \vec{u}) = \vec{0}\end{aligned}$$

$$(2) \quad \lambda \circ \vec{0} \stackrel{(A3)}{=} \lambda \circ \vec{0} + \vec{0} \stackrel{(A4)}{=} \lambda \circ \vec{0} + \lambda \circ \vec{0} + (- (\lambda \circ \vec{0})) \stackrel{(A7)}{=} \lambda \circ (\vec{0} + \vec{0}) - (\lambda \circ \vec{0}) = \vec{0}$$

$$(3) \quad \boxed{\Leftarrow} \text{ seuraa kohdista (1) ja (2).}$$

$$\boxed{\Rightarrow} \text{ Olkoon } \lambda \circ \vec{u} = \vec{0}. \text{ Jos } \lambda \neq 0, \text{ niin } \lambda^{-1} \text{ on olemassa ja}$$

$$\vec{u} \stackrel{(A8)}{=} 1 \circ \vec{u} = (\lambda^{-1} \lambda) \circ \vec{u} \stackrel{(A5)}{=} \lambda^{-1} (\lambda \circ \vec{u}) = \lambda^{-1} \circ \vec{0} \stackrel{(1)}{=} \vec{0}$$

$$(4) \quad \vec{u} + (-1) \circ \vec{u} \stackrel{(A8)}{=} 1 \circ \vec{u} - 1 \circ \vec{u} \stackrel{(A6)}{=} (1 - 1) \circ \vec{u} = 0 \circ \vec{u} \stackrel{(1)}{=} \vec{0}$$

□

Määritelmä 131 Lineaariavaruuden L osajoukko M on lineaarinen aliavaruus, jos

$$\begin{aligned}(1) \quad \vec{0} &\in M \\ (2) \quad \vec{u}, \vec{v} \in M &\Rightarrow \vec{u} + \vec{v} \in M \\ (3) \quad \lambda \in \mathbb{R}, \vec{u} \in M &\Rightarrow \lambda \circ \vec{u} \in M\end{aligned}$$

(Engl. subspace)

Lineaarinen aliavaruus siis sisältää nollavektorin ja se on yhteenlaskun ja skalaarilla kertomisen suhteen suljettu osajoukko.

Lause 132 Olkoon M_1 ja M_2 kaksi lineaariavaruuden L aliavaruutta. Silloin:

(1) Aliavaruuksien leikkaus

$$M_1 \cap M_2 = \{\vec{u} \in L \mid \vec{u} \in M_1 \text{ ja } \vec{u} \in M_2\}$$

on L :n lineaarinen aliavaruus.

(1) Aliavaruuksien summa

$$M_1 + M_2 = \{\vec{u} + \vec{v} \in L \mid \vec{u} \in M_1 \text{ ja } \vec{v} \in M_2\}$$

on L :n lineaarinen aliavaruus.

Todistus. HT

□

Tyhjä joukko ei ole aliavaruus, mutta pelkän nollavektorin sisältävä joukko $\{\vec{0}\}$ on aliavaruus (triviaali aliavaruus). Koska jokainen aliavaruus sisältää nolavektorin, sisältää kahden aliavaruuden leikkaus aina ainakin nollavektorin.

Määritelmä 133 Jos M_1 ja M_2 ovat kaksi lineaariavaruuden L aliavaruutta, joiden leikkaus on triviaali ($M_1 \cap M_2 = \{\vec{0}\}$), niin niiden summaa sanotaan suoraksi summaksi ja sitä merkitään

$$M_1 \oplus M_2.$$

c4s2th1

Lause 134 Jos lineaariavaruus L on kahden aliavaruutensa M_1 ja M_2 suora summa, eli $L = M_1 \oplus M_2$, niin jokaista vektoria $\vec{u} \in L$ vastaa yhdet ja vain yhdet komponenttivektorit $\vec{u}_1 \in M_1$ ja $\vec{u}_2 \in M_2$ siten, että

$$\vec{u} = \vec{u}_1 + \vec{u}_2.$$

Todistus. Koska $L = M_1 \oplus M_2$ on olemassa ainakin yhdet komponenttivektorit $\vec{u}_1 \in M_1$ ja $\vec{u}_2 \in M_2$, joille $\vec{u} = \vec{u}_1 + \vec{u}_2$. On siis osoitettava, että muita ei löydy.

Olkoon $\vec{w}_1 \in M_1$ ja $\vec{w}_2 \in M_2$ toiset komponentit siten, että $\vec{u} = \vec{w}_1 + \vec{w}_2$. Nyt

$$\begin{aligned} \vec{u}_1 + \vec{u}_2 &= \vec{w}_1 + \vec{w}_2 \\ \Leftrightarrow \underbrace{\vec{u}_1 - \vec{w}_1}_{\in M_1} &= \underbrace{\vec{w}_2 - \vec{u}_2}_{\in M_2} \end{aligned}$$

Koska $M_1 \cap M_2 = \{\vec{0}\}$, niin viimeinen yhtälö voi olla tosi vain silloin, kun LHS ja RHS ovat nollavektoreita. Siis $\vec{w}_1 = \vec{u}_1$ ja $\vec{w}_2 = \vec{u}_2$. $\square$

Seuraavassa määritelty joukko on hyvin tärkeä ja usein esiintyvä aliavaruus.

Kerrataan ensin jo luvussa 2 esiintynyt lineaarikombinaation käsite.

Määritelmä 135 Olkoot $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n \in L$ lineaariavaruuden L vektoreita, ja olkoot $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ reaalityyppisiä lukuja. Silloin vektori $\lambda_1 \circ \vec{u}_1 + \lambda_2 \circ \vec{u}_2 + \dots + \lambda_n \circ \vec{u}_n \in L$ on lineaarikombinaatio vektoreista $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ kertoimin $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. (Engl. linear combination.)

Määritelmä 136 Olkoon L reaalinen lineaariavaruus ja olkoon M sen osajoukko. Joukon M virittämä aliavaruus, $\text{span}(M)$, on kaikkien M :n alkioiden äärellisten lineaarikombinaatioiden joukko. Sovimme edellisen lisäksi, että tyhjä joukko virittää triviaalin aliavaruuden, $\text{span}(\emptyset) = \{\vec{0}\}$. (Engl. subspace spanned by M .)

Lause 137 Olkoon L reaalinen lineaariavaruus ja olkoon M sen osajoukko. Silloin $\text{span}(M)$ on L :n lineaarinen aliavaruus.

Todistus. Jos $M = \emptyset$, niin $\text{span}(M) = \{\vec{0}\}$ on aliavaruus. Jos $M \neq \emptyset$, niin on olemassa $\vec{u} \in M$ joten $\vec{0} = 0 \circ \vec{u} \in \text{span}(M)$. Siis aliavaruuden määritelmän ehto (1) on voimassa.

(2) Jos $\vec{v}, \vec{w} \in \text{span}(M)$, niin ne voidaan esittää äärellisinä lineaarikombinaatioina

$$\begin{aligned}\vec{v} &= \lambda_1 \circ \vec{u}_1 + \dots + \lambda_m \circ \vec{u}_m, \text{ ja} \\ \vec{w} &= \lambda_{m+1} \circ \vec{u}_{m+1} + \dots + \lambda_n \circ \vec{u}_n, \text{ joten} \\ \vec{v} + \vec{w} &= \lambda_1 \circ \vec{u}_1 + \dots + \lambda_n \circ \vec{u}_n \in \text{span}(M)\end{aligned}$$

(3) Jos $\lambda \in \mathbb{R}$ ja $\vec{v} = \lambda_1 \circ \vec{u}_1 + \dots + \lambda_m \circ \vec{u}_m$, niin

$$\lambda \circ \vec{v} = (\lambda \cdot \lambda_1) \circ \vec{u}_1 + \dots + (\lambda \cdot \lambda_m) \circ \vec{u}_m \in \text{span}(M)$$

$\square$

Esimerkki 138 Olkoon L tasovektoreiden muodostama vektoriavaruus. Olkoon $\vec{i}$ yksikön mittainen vektori x -akselin suuntaan, ja olkoon $\vec{j}$ yksikön mittainen vektori y -akselin suuntaan. Silloin $L = \text{span}(\vec{i}, \vec{j})$.

7.3 Sisätulo ja normi

Sec:InnerPNorm

Etenemme asteittain niin, että ensin esittelemme E^n :ssä sisätulon ja kolme erilaista normia. Näiden käsitteiden avulla määrittelemme kohtisuoruuden ja pituuden käsitteet. Lopuksi vielä annamme määritelmät yleisen lineaariavaruuden tapauksessa.

7.3.1 Sisätulo E^n :ssä

Määritelmä 139 Kahden vektorin $\vec{a} = (a_1, a_2, \dots, a_n)^T \in \mathbb{R}^n$ ja $\vec{b} = (b_1, b_2, \dots, b_n)^T \in \mathbb{R}^n$ sisätulo $\langle \vec{a} | \vec{b} \rangle$ määritellään kaavalla

$$\langle \vec{a} | \vec{b} \rangle = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \vec{a} \cdot \vec{b}$$

(Kyseessä on siis jo ennemmin määritelty pistetulo, mutta nyt otamme käyttöön kulmasulkumerkinnän.)

Määritelmästä seuraa välittömästi, että sisätulo on vaihdannainen ja kummankin tekijänsä suhteen lineaarinen. Lisäksi nollavektori on ainoa vektori, jonka sisätulo itsensä kanssa on nolla. Kokoamme nämä selvät ominaisuudet lauseeksi, jotta ne myöhemmin löytyvät helposti.

c4s3th1 **Lause 140** Olkoot $\vec{u}, \vec{v} \in \mathbb{R}^n$ ja $\lambda \in \mathbb{R}$. Silloin

- (S1) $\langle \vec{u} | \vec{v} \rangle = \langle \vec{v} | \vec{u} \rangle$,
- (S2) $\langle \vec{u} | \vec{v} + \vec{w} \rangle = \langle \vec{u} | \vec{v} \rangle + \langle \vec{u} | \vec{w} \rangle$,
- (S3) $\lambda \cdot \langle \vec{u} | \vec{v} \rangle = \langle \lambda \vec{u} | \vec{v} \rangle = \langle \vec{u} | \lambda \vec{v} \rangle$,
- (S4) $\langle \vec{u} | \vec{u} \rangle \geq 0$ ja $\langle \vec{u} | \vec{u} \rangle = 0$ jos ja vain jos $\vec{u} = \vec{0}$

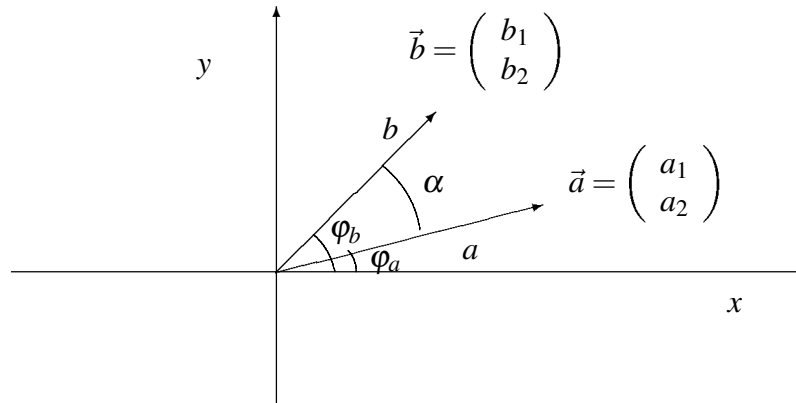
Todistus. HT.

Kaksi- ja kolmiulotteisissa Euklididissa vektoriavaruuksissa E^2 ja E^3 sisätulolla on seuraava koulukurssista tuttu geometrinen tulkinta.

Lause 141 Tarkastellaan E^2 :n vektoreita $\vec{a} = (a_1, a_2)^T$ ja $\vec{b} = (b_1, b_2)^T$. Vektoreiden pituudet ovat $a = \sqrt{a_1^2 + a_2^2}$ ja $b = \sqrt{b_1^2 + b_2^2}$. Jos vektoreiden välinen kulma on α , niin

$$\langle \vec{a} | \vec{b} \rangle = a \cdot b \cdot \cos \alpha$$

c4s3fig1



Kuva 27. vektoreiden $\vec{a}$ ja $\vec{b}$ sisätulo.

Todistus. Olkoon vektorin $\vec{a}$ ja vaaka-akselin välinen kulma φ_a ja vastaavasti vektorin $\vec{b}$ ja vaaka-akselin välinen kulma φ_b . Silloin (Ks. kuva (c4s3fig1))

$$\begin{cases} a_1 = a \cos \varphi_a \\ a_2 = a \sin \varphi_a \end{cases} \quad \text{ja} \quad \begin{cases} b_1 = b \cos \varphi_b \\ b_2 = b \sin \varphi_b \end{cases}.$$

Trigonometristen funktioiden kaavojen mukaan $\cos(x-y) = \cos x \cos y + \sin x \sin y$. Siis

$$\begin{aligned} \langle \vec{a} | \vec{b} \rangle &= a_1 b_1 + a_2 b_2 = a \cos(\varphi_a) \cdot b \cos(\varphi_b) + a \sin(\varphi_a) \cdot b \sin(\varphi_b) \\ &= ab \cos(\varphi_b - \varphi_a) = ab \cos \alpha \end{aligned}$$

□

Edeltävä tulos pitää paikkansa myös E^3 :ssa. Emme nyt todista tätä, vaan myöhemmin, palaamme asiaan. Joka tapauksessa teemme nyt seuraavat yleistykset.

Määritelmä 142 Vektorit $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$ ovat

(1) kohtisuorassa toisiaan vastaan ($\vec{a} \perp \vec{b}$) eli ortogonaaliset (engl. *orthogonal*), jos

$$\langle \vec{a} | \vec{b} \rangle = 0.$$

(2) yhdensuuntaiset ($\vec{a} \parallel \vec{b}$), jos on olemassa reaaliluku $\lambda \in \mathbb{R}$ siten, että

$$\vec{a} = \lambda \vec{b}$$

(3) samansuuntaiset ($\vec{a} \uparrow \vec{b}$), jos on olemassa positiivinen reaaliluku $\lambda \in \mathbb{R}$ siten, että

$$\vec{a} = \lambda \vec{b}, \lambda > 0$$

(4) vastakkaissuuntaiset ($\vec{a} \updownarrow \vec{b}$), jos on olemassa negatiivinen reaaliluku $\lambda \in \mathbb{R}$ siten, että

$$\vec{a} = \lambda \vec{b}, \lambda < 0$$

Olkon $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$ mitkä tahansa kaksi vektoria. Vektori $\vec{a}$ voidaan aina lausua kahden vektorin summana $\vec{a} = \vec{a}_b + \vec{a}_\perp$ niin, että ensimmäinen vektoreista on $\vec{b}$:n suuntainen ja toinen on $\vec{b}$:tä vastaan kohtisuorassa. Jos nimittäin asetamme

$$\vec{a}_b = \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} \quad (7.7)$$

$$\vec{a}_\perp = \vec{a} - \vec{a}_b = \vec{a} - \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} \quad (7.8)$$

niin selvästi vektoreiden summa on $\vec{a}$, ensimmäinen vektori on yhdensuuntainen $\vec{b}$:n kanssa ja lineaarisuuden perusteella (ks. (S2) ja (S3) lauseessa (I40))

$$\langle \vec{a}_\perp | \vec{b} \rangle = \left\langle \vec{a} - \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} \middle| \vec{b} \right\rangle \quad (7.9)$$

$$= \langle \vec{a} | \vec{b} \rangle - \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \langle \vec{b} | \vec{b} \rangle = 0 \quad (7.10)$$

c4s3esim1

Esimerkki 143 Tarkastellaan esimerkkiä, jonka vektorit voimme piirtää (ks. kuva (I33)). Olkoon $\vec{a} = (2, 6)^T \in \mathbb{R}^2$ ja $\vec{b} = (8, 4)^T \in \mathbb{R}^2$. Silloin

$$\vec{a}_b = \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} = \left(\frac{2 \cdot 8 + 6 \cdot 4}{8 \cdot 8 + 4 \cdot 4} \right) \vec{b} = 0.5 \begin{pmatrix} 8 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad (7.11)$$

$$\vec{a}_\perp = \vec{a} - \vec{a}_b = \begin{pmatrix} 2 \\ 6 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix} \quad (7.12)$$

Määritelmä 144 Olkoon $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$. Vektori

$$\vec{a}_b = \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b}$$

on vektorin $\vec{a}$ projektio vektorin $\vec{b}$ suuntaan (vektoriprojektio).

Määritelmä 145 Olkoon $M \subseteq \mathbb{R}^n$ lineaarinen aliavaruus.

(1) Olkoon $\vec{u} \in \mathbb{R}^n$ vektori. Jos

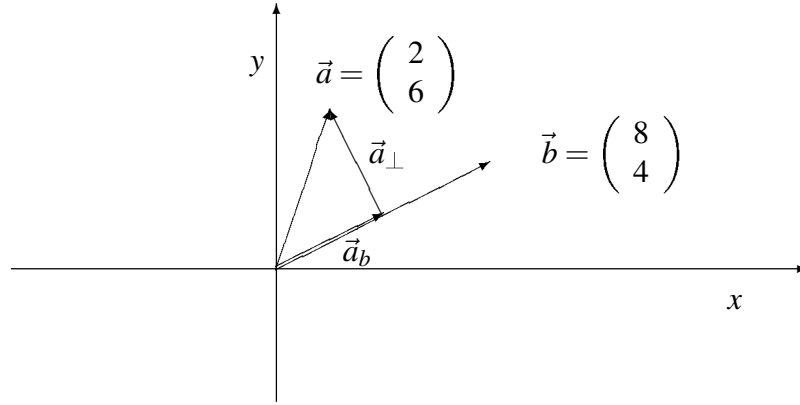
$$\vec{u} \perp \vec{v}, \quad \forall \vec{v} \in M,$$

niin sanomme, että $\vec{u}$ on kohtisuorassa aliavaruutta M vastaan ja merkitsemme

$$\vec{u} \perp M.$$

(2) Aliavaruuden M ortogonaalikomplementti on

$$M^\perp = \{ \vec{u} \in \mathbb{R}^n \mid \vec{u} \perp M \}.$$



c4s3fig2

Kuva 28. Vektorin $\vec{a}$ projektio vektorin $\vec{b}$ suuntaan.

Lause 146 Olkoon $M \subseteq \mathbb{R}^n$ lineaarinen aliavaruus. Silloin myös $M^\perp$ on lineaarinen aliavaruus.

Todistus. On siis osoitettava, että kaikilla $\vec{a}, \vec{b} \in M^\perp$ ja $\lambda \in \mathbb{R}$

$$\begin{aligned}\vec{a} + \vec{b} &\in M^\perp \text{ ja} \\ \lambda \vec{a} &\in M^\perp\end{aligned}$$

Olkoon nyt $\vec{a}, \vec{b} \in M^\perp$ ja $\lambda \in \mathbb{R}$ kuten edellä. Silloin

$$\begin{aligned}\forall \vec{c} \in M: \quad \langle \vec{a} + \vec{b} | \vec{c} \rangle &= \langle \vec{a} | \vec{c} \rangle + \langle \vec{b} | \vec{c} \rangle = 0 \Rightarrow \vec{a} + \vec{b} \in M^\perp \\ \forall \vec{c} \in M: \quad \langle \lambda \vec{a} | \vec{c} \rangle &= \lambda \langle \vec{a} | \vec{c} \rangle = 0 \Rightarrow \lambda \vec{a} \in M^\perp\end{aligned}$$

□

Käytännössä suurin osa $\mathbb{R}^n$:n aliavaruuksista on viritettyjä, joten seuraava lause on lähes aina käytettävissä, kun tutkimme kohtisuoruutta.

c4s3th2

Lause 147 Olkoon $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q) \subseteq \mathbb{R}^n$ vektoreiden $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ virittämä aliavaruus ja olkoon $\mathbf{V}$ matriisi, jonka k :s sarake on $\vec{v}_k$. Olkoon $\vec{u} \in \mathbb{R}^n$. Silloin

$$\vec{u} \perp M. \Leftrightarrow \vec{u} \perp \vec{v}_k, \forall k = 1, 2, \dots, q \quad (7.13)$$

$$\Leftrightarrow \mathbf{V}^T \vec{u} = \vec{0} \quad (7.14)$$

Todistus. Helppo HT.

□

Jos siis vektori on kohtisuorassa jokaista virittäjää vastaan, se on kohtisuorassa koko aliavaruutta vastaan.

Seuraava lause on ”tekninen temppu”, mutta se tekee monet tarkastelut helpoiksi.

Lause 148 Olkoon $\vec{b} \in \mathbb{R}^n$, $\vec{c} \in \mathbb{R}^m$ ja olkoon $\mathbf{A}$ $m \times n$ -matriisi. Silloin

$$\langle \mathbf{A}\vec{b} | \vec{c} \rangle = \langle \vec{b} | \mathbf{A}^T \vec{c} \rangle$$

Todistus.

$$\begin{aligned} \langle \mathbf{A}\vec{b} | \vec{c} \rangle &= \sum_{k=1}^m (\mathbf{A}\vec{b})_k c_k = \sum_{k=1}^m \left(\sum_{j=1}^n a_{kj} b_j \right) c_k \\ &= (a_{11}b_1 + a_{12}b_2 + \cdots + a_{1n}b_n)c_1 \\ &\quad + (a_{21}b_1 + a_{22}b_2 + \cdots + a_{2n}b_n)c_2 \\ &\quad \cdots + (a_{m1}b_1 + a_{m2}b_2 + \cdots + a_{mn}b_n)c_m \\ &= b_1(a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m) \\ &\quad + b_2(a_{12}c_1 + a_{22}c_2 + \cdots + a_{m2}c_m) \\ &\quad \cdots + b_n(a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m) \\ &= \sum_{j=1}^n b_j \left(\sum_{k=1}^m a_{kj} c_k \right) = \sum_{j=1}^n b_j (\mathbf{A}^T \vec{c})_j = \langle \vec{b} | \mathbf{A}^T \vec{c} \rangle \end{aligned}$$

□

Edellisen lauseen yhteydessä on tärkeätä muistaa transpoosin peruskaavat

$$(\mathbf{A}^T)^T = \mathbf{A} \quad (7.15)$$

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T \quad (7.16)$$

$$(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} \quad (7.17)$$

Edellä esimerkissä ^{|c4s3esim1|}(143) jaοimme vektorin $\vec{a} \in \mathbb{R}^n$ kahteen osaan $\vec{a} = \vec{a}_b + \vec{a}^\perp$ siten, että ensimmäinen komponentti on vektorin $\vec{b}$ suuntainen ja toinen on vektoria $\vec{b}$ vastaan kohtisuora. Lineaarisen aliavaruuden $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q)$ tapauksessa voimme erään ehdon vallitessa jakaa vektorin $\vec{a} \in \mathbb{R}^n$ kahteen komponenttiin $\vec{a} = \vec{a}_M + \vec{a}_{M^\perp}$ siten että $\vec{a}_M \in M$ ja $\vec{a}_{M^\perp} \in M^\perp$. Komponentit osoittautuvat yksikäsitteisiksi.

Ehto, joka nyt joudutaan olettamaan, voidaan myöhemmin kokonaan poistaa pienellä lauseen uudelleenmuotoilulla. Voisimme lykätä tarkastelun myöhemmäksi, mutta otamme tämän nyt esimerkkinä sisätulon ominaisuuksista.

c4s3th6

Lause 149 Olkoon $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q) \subseteq \mathbb{R}^n$ vektoreiden $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ virittämä lineaarinen aliavaruus ja olkoon $\mathbf{V}$ matriisi, jonka k :s sarake on $\vec{v}_k$.

Teemme nyt erikoisen oletuksen olettamalla, että matriisi $\mathbf{V}^T \mathbf{V}$ on säännöllinen (eli on olemassa $(\mathbf{V}^T \mathbf{V})^{-1}$).

Silloin

$$\mathbb{R}^n = M \oplus M^\perp.$$

Todistus. Ensin osoitamme, että jokaisella $\vec{a} \in \mathbb{R}^n$ on olemassa vektorit $\vec{a}_M \in M$ ja $\vec{a}_{M^\perp} \in M^\perp$ siten, että $\vec{a} = \vec{a}_M + \vec{a}_{M^\perp}$.

Tekemämme oletuksen perusteella voimme asettaa

$$\vec{a}_M = \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a}, \quad \text{ja} \quad \vec{a}_{M^\perp} = \vec{a} - \vec{a}_M$$

(1) Jos merkitsemme $\vec{c} = (c_1, c_2, \dots, c_q)^T = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a}$, niin

$$\vec{a}_M = \mathbf{V} \vec{c} = c_1 \vec{v}_1 + \dots + c_q \vec{v}_q \in M.$$

Siis ensimmäinen komponentti on aliavaruuden vektori.

(2) Toiseksi tutkimme toisen komponentin kohtisuoruutta lauseen (I47) mukaisesti

$$\mathbf{V}^T \vec{a}_{M^\perp} = \mathbf{V}^T \vec{a} - \mathbf{V}^T \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} = \vec{0}$$

Siis toinen komponentti on kohtisuorassa aliavaruutta M vastaan.

(3) Kolmanneksi toteamme summan suoraksi osoittamalla, että $M \cap M^\perp = \{\vec{0}\}$. Olkoon siis $\vec{u} \in M \cap M^\perp$. Silloin erityisesti

$$\begin{aligned} \vec{u} \perp \vec{u} &\Leftrightarrow \langle \vec{u} | \vec{u} \rangle = u_1^2 + u_2^2 + \dots + u_n^2 = 0 \\ &\Leftrightarrow \vec{u} = \vec{0} \end{aligned}$$

□

Lause sanoo sen, että mikä tahansa vektori voidaan lausua kahden vektorin summan niin, että ensimmäinen yhteenlaskettava kuuluu aliavaruuteen M ja toinen yhteenlaskettava on kohtisuorassa aliavaruutta M vastaan. Tämä on jatkossa erittäin hyödyllinen tulos. Lauseessa oleva lisäehto $\det(\mathbf{V}^T \mathbf{V}) \neq 0$ tulee jatkossa hoidettua pois.

Esimerkki 150 Edellisen lauseen todistus on hämäävän helppo. Testaamme vielä fundamentaalilla tavalla, ovatko vektorit $\vec{a}_{M^\perp}$ ja $\vec{v}_k$, $1 \leq k \leq q$ kohtisuorassa:

$$\langle \vec{a}_{M^\perp} | \vec{v}_k \rangle = \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} | \vec{v}_k \rangle \quad (7.18)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} | \mathbf{V} \vec{e}_k \rangle \quad (7.19)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} | \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.20)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | ((\mathbf{V}^T \mathbf{V})^{-1})^T \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.21)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | ((\mathbf{V}^T \mathbf{V})^T)^{-1} \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.22)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.23)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | \vec{e}_k \rangle \quad (7.24)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \vec{a} | \mathbf{V} \vec{e}_k \rangle \quad (7.25)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \vec{a} | \vec{v}_k \rangle = 0 \quad (7.26)$$

7.3.2 $\mathbb{R}^n$:n normeja

Vektoreiden $\vec{a} = (a_1, a_2)^T \in \mathbb{R}^2$ ja $\vec{b} = (b_1, b_2, b_3)^T \in \mathbb{R}^3$ pituudet on koulukurssissa laskettu kaavoilla

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2}, \text{ ja}$$

$$|\vec{b}| = \sqrt{b_1^2 + b_2^2 + b_3^2}.$$

Normi yleistää pituuden idean $\mathbb{R}^n$:ään. Perusidea on se, että pienellä vektorilla on pieni normi ja suurella vektorilla on suuri normi. Kaksi vektoria ovat lähellä toisiaan, jos niiden erotuksen normi on pieni.

Määrittelemme kolme erilaista normia.

c4s3defnorm

Määritelmä 151 Vektorin $\vec{a} = (a_1, a_2, \dots, a_n)^T \in \mathbb{R}^n$ 1-normi $\|\vec{a}\|_1$, 2-normi $\|\vec{a}\|_2$ ja ääretön-normi $\|\vec{a}\|_\infty$ määritellään kaavoilla

$$(1) \quad \|\vec{a}\|_1 = |a_1| + |a_2| + \dots + |a_n| = \sum_{k=1}^n |a_k|$$

$$(2) \quad \|\vec{a}\|_2 = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2} = \sqrt{\langle \vec{a} | \vec{a} \rangle}$$

$$(3) \quad \|\vec{a}\|_\infty = \max\{|a_1|, |a_2|, \dots, |a_n|\} = \max_{1 \leq k \leq n} |a_k|$$

CauchySchwarz

Lause 152 (Cauchy-Schwarz) Kaikille $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$ on voimassa epäyhtälö

$$|\langle \vec{a} | \vec{b} \rangle| \leq \|\vec{a}\|_2 \|\vec{b}\|_2$$

Todistus. Seuraavassa todistuksessa normi $\|\cdot\|$ tarkoittaa 2-normia. Sisätulon ominaisuuksista seuraa, että kaikilla $\lambda \in \mathbb{R}$ on voimassa

$$0 \leq \langle \vec{a} + \lambda \vec{b} | \vec{a} + \lambda \vec{b} \rangle = \|\vec{a}\|^2 + \lambda^2 \|\vec{b}\|^2 + 2\lambda \langle \vec{a} | \vec{b} \rangle$$

Kun tähän sijoitetaan

$$\lambda = -\frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle},$$

niin saadaan

$$0 \leq \|\vec{a}\|^2 - \frac{\langle \vec{a} | \vec{b} \rangle^2}{\|\vec{b}\|^2},$$

mistä seuraa lauseen väite. □

Kaikki kolme normia toteuttavat kolme yksinkertaista ominaisuutta.

Lause 153 Kaikilla $\vec{u}, \vec{v} \in \mathbb{R}^n$ ja $\lambda \in \mathbb{R}$ edellä määritellyt 1-normi, 2-normi ja ääretön-normi toteuttavat seuraavat ehdot

- (N1) $\|\vec{u}\| \geq 0$ ja $\|\vec{u}\| = 0$ jos ja vain jos $\vec{u} = \vec{0}$
- (N2) $\|\lambda \vec{u}\| = |\lambda| \cdot \|\vec{u}\|$,
- (N3) $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$, (kolmioepäyhtälö)

Todistus. Kohdat (N1) ja (N2) seuraavat suoraan määritelmistä, mutta kolmioepäyhtälö kannattaa käydä tarkemmin läpi: Merkitään todistuksen ajan $\vec{u} + \vec{v} = \vec{w}$

1-normi: Olkoon x ja y reaalilukuja. Tarvittaessa nimet vaihtaen voimme olettaa, että $|x| \geq |y|$. Nyt

$$|x+y| = \begin{cases} |x| + |y|, & \text{kun } x \text{ ja } y \text{ ovat saman merkkiset, ja} \\ |x| - |y| \leq |x| + |y|, & \text{kun } x \text{ ja } y \text{ ovat erimerkkiset.} \end{cases}$$

Aina on siis kahdelle reaaliluvulle totta, että $|x+y| \leq |x| + |y|$. Nyt

$$\begin{aligned} \|\vec{w}\|_1 &= |u_1 + v_1| + |u_2 + v_2| + \cdots + |u_n + v_n| \\ &\leq |u_1| + |v_1| + |u_2| + |v_2| + \cdots + |u_n| + |v_n| = \|\vec{u}\|_1 + \|\vec{v}\|_1 \end{aligned}$$

2-normi: Cauchy-Schwarzin epäyhtälöä käyttäen saamme

$$\begin{aligned}\|\vec{w}\|_2^2 &= \langle \vec{u} + \vec{v} | \vec{u} + \vec{v} \rangle = \|\vec{u}\|_2^2 + 2\langle \vec{u} | \vec{v} \rangle + \|\vec{v}\|_2^2 \\ &\leq \|\vec{u}\|_2^2 + 2|\langle \vec{u} | \vec{v} \rangle| + \|\vec{v}\|_2^2 \\ &\leq \|\vec{u}\|_2^2 + 2\|\vec{u}\|_2\|\vec{v}\|_2 + \|\vec{v}\|_2^2 = (\|\vec{u}\|_2 + \|\vec{v}\|_2)^2\end{aligned}$$

∞ -normi: Olkoon $\|\vec{w}\|_\infty = |w_i|$, $\|\vec{u}\|_\infty = |u_j|$ ja $\|\vec{v}\|_\infty = |v_k|$. Silloin

$$\|\vec{w}\|_\infty = |w_i| = |u_i + v_i| \leq |u_i| + |v_i| \leq |u_j| + |v_k| = \|\vec{u}\|_\infty + \|\vec{v}\|_\infty$$

□

Kolmen erilaisen normin käyttö ei ole tässä vaiheessa kovin motivoivaa. Jos laadimme numeerisia algoritmeja, niin 1-normin ja ∞ -normin laskeminen on helpompaa kuin 2-normin laskeminen. Ongelma voidaan useimmiten helposti kiertää.

Kun myöhemmin määrittelemme vektorinormiin liittyvän vastaavan matriisinormin, alkavat syntyvät kolme vaihtoehtoa erota oleellisesti toisistaan. Palaamme tähän myöhemmin.

Normia tarvitaan yleensä, kun halutaan tutkia, ovatko kaksi vektoria ”melkein samat”, eli ”lähellä toisiaan”. Jos esimerkiksi $\|\vec{a} - \vec{b}\|_\infty = 10^{-15}$, niin vektorit ovat käytännössä melkein identtiset, sillä kaikilla k on $|a_k - b_k| \leq 10^{-15}$. ”Lähellä oleminen” ei nyt riipu käytetystä normista, sillä helposti nähdään, että

$$\frac{1}{\sqrt{n}}\|\vec{u}\|_2 \leq \|\vec{u}\|_1 \leq \|\vec{u}\|_2 \quad (7.27)$$

$$\frac{1}{n}\|\vec{u}\|_\infty \leq \|\vec{u}\|_1 \leq \|\vec{u}\|_\infty \quad (7.28)$$

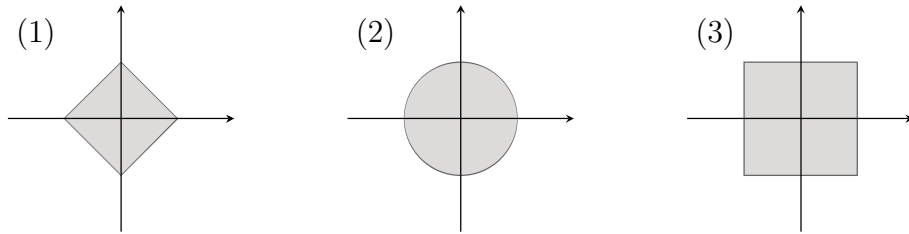
$$\|\vec{u}\|_1 \leq \|\vec{u}\|_2 \leq \sqrt{n}\|\vec{u}\|_1 \quad (7.29)$$

$$\frac{1}{\sqrt{n}}\|\vec{u}\|_\infty \leq \|\vec{u}\|_2 \leq \|\vec{u}\|_\infty \quad (7.30)$$

$$\|\vec{u}\|_1 \leq \|\vec{u}\|_\infty \leq n\|\vec{u}\|_1 \quad (7.31)$$

$$\|\vec{u}\|_2 \leq \|\vec{u}\|_\infty \leq \sqrt{n}\|\vec{u}\|_2 \quad (7.32)$$

Erityisesti, jos rajankäynnissä normi lähestyy nollaa yhden normin mielessä, niin myös muut kaksi normia lähestyvät nollaa. Kuvaan  (29) on piirretty kaikki pisteet $\vec{z} = (x, y)^T$, joille on totta, että (1) $\|\vec{z}\|_1 \leq 1$, (2) $\|\vec{z}\|_2 \leq 1$, (3) $\|\vec{z}\|_\infty \leq 1$.



Kuva 29. Pisteet, jotka ovat enintään etäisyydellä 1 origosta: (1) $\|\vec{z}\|_1 \leq 1$, (2) $\|\vec{z}\|_2 \leq 1$, (3) $\|\vec{z}\|_\infty \leq 1$

fig056

7.3.3 Yleinen tarkastelu

Määritelmä 154 Olkoon L reaallinen lineaariavaruus. Jos on olemassa kuvaus

$$L \times L \rightarrow \mathbb{R}, (\mathbf{a}, \mathbf{b}) \mapsto \langle \mathbf{a} | \mathbf{b} \rangle,$$

joka toteuttaa seuraavat aksioomat

- (S1) $\langle \mathbf{u} | \mathbf{v} \rangle = \langle \mathbf{v} | \mathbf{u} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in L$
- (S2) $\langle \mathbf{u} | \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} | \mathbf{w} \rangle, \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in L$
- (S3) $\lambda \cdot \langle \mathbf{u} | \mathbf{v} \rangle = \langle \lambda \mathbf{u} | \mathbf{v} \rangle = \langle \mathbf{u} | \lambda \mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in L, \forall \lambda \in \mathbb{R}$
- (S4) $\langle \mathbf{u} | \mathbf{u} \rangle \geq 0$ ja $\langle \mathbf{u} | \mathbf{u} \rangle = 0$ jos ja vain jos $\mathbf{u} = \mathbf{0}$

niin kuvausta sanotaan sisätuloksi ja lukua $\langle \mathbf{u} | \mathbf{v} \rangle$ sanotaan vektoreiden $\mathbf{u}$ ja $\mathbf{v}$ sisätuloksi (engl. inner product).

Määritelmä 155 Olkoon L reaallinen lineaariavaruus. Jos on olemassa kuvaus

$$L \rightarrow \mathbb{R}, \mathbf{u} \mapsto \|\mathbf{u}\|,$$

joka toteuttaa seuraavat aksioomat

- (N1) $\|\mathbf{u}\| \geq 0$ ja $\|\mathbf{u}\| = 0$ jos ja vain jos $\mathbf{u} = \vec{0}$
- (N2) $\|\lambda \mathbf{u}\| = |\lambda| \cdot \|\mathbf{u}\|, \quad \forall \mathbf{u} \in L, \forall \lambda \in \mathbb{R}$
- (N3) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|, \quad \forall \mathbf{u}, \mathbf{v} \in L,$

niin kuvausta sanotaan normiksi ja lukua $\|\mathbf{u}\|$ sanotaan vektorin $\mathbf{u}$ normiksi (engl. norm).

Yleisessä tapauksessa sisätulo ja normi voidaan määritellä monella tavalla. (Huomaa miten tämä näkyy edellä esitettyjen määritelmien muotoilussa. Määritelmä ei sano mikä normi on, vaan se kertoo millä ehdolla kuvauksesta saa sanoa, että se on normi.)

Jokaiseen sisätuloon $\langle \mathbf{u} | \mathbf{v} \rangle$ liittyy sen avulla määritelty normi $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u} | \mathbf{u} \rangle}$. (Usein sanotaan, että ”sisätulo indusoi normin”). Toisaalta normiin ei tarvitse liittyä mitään sisätuloa.

$\mathbb{R}^n$:ssä sisätulo indusoi 2-normin, mutta 1-normiin ja ∞ -normiin ei liity sisätuloa, eikä siis myöskään kohtisuoruutta.

Esimerkki 156 Olkoon $L = C([-1, 1]) = \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ on jatkuva kuvaus}\}$. L on reaalinen lineaariavaruus, ja voimme määritellä sisätulon ja normin kaavoilla

$$\langle f | g \rangle = \int_{-1}^1 f(x)g(x) \, dx \quad (7.33)$$

$$\|f\| = \sqrt{\langle f | f \rangle} = \left(\int_{-1}^1 (f(x))^2 \, dx \right)^{1/2} \quad (7.34)$$

Lasketaan seuraavien funktioiden normit ja niiden väliset sisätulot:

$$e(x) = 1, \quad f(x) = x, \quad g(x) = x^2.$$

$$\|e\| = \left(\int_{-1}^1 (1)^2 \, dx \right)^{1/2} = \sqrt{2}$$

$$\|f\| = \left(\int_{-1}^1 x^2 \, dx \right)^{1/2} = \left(\left[\frac{1}{3}x^3 \right]_{-1}^1 \right)^{1/2} = \sqrt{2/3}$$

$$\|g\| = \left(\int_{-1}^1 (x^2)^2 \, dx \right)^{1/2} = \left(\left[\frac{1}{5}x^5 \right]_{-1}^1 \right)^{1/2} = \sqrt{2/5}$$

$$\langle e | f \rangle = \int_{-1}^1 1 \cdot x \, dx = 0$$

$$\langle e | g \rangle = \int_{-1}^1 1 \cdot x^2 \, dx = \frac{2}{3}$$

$$\langle f | g \rangle = \int_{-1}^1 x \cdot x^2 \, dx = 0$$

Siis $e \perp f$ ja $f \perp g$, mutta $e \not\perp g$. (Kohtisuoruus ei siis ole transitiivinen relaatio.)

Cauchy-Schwarz -epäyhtälön todistus (152) sivulla 137 perustuu sisätulon ja siihen liittyvän normin ominaisuuksiin. Epäyhtälö on siis tässäkin tapauksessa voimassa.

$$\begin{aligned} |\langle f | g \rangle| &\leq \|f\| \cdot \|g\| \\ \Leftrightarrow \left| \int_{-1}^1 f(x)g(x) \, dx \right| &\leq \left(\int_{-1}^1 (f(x))^2 \, dx \right)^{1/2} \left(\int_{-1}^1 (g(x))^2 \, dx \right)^{1/2} \end{aligned}$$

7.4 Lineaarinen kuvaus

Sec:LinearMap

Tämä on lyhyt kappale, jossa on vain muutama määritelmä. Jokainen määritelmä ja lause on tärkeä. Sanaa ”lineaarinen” käytetään paljon eri yhteyksissä. Matematiikan ulkopuolella sana merkitsee yleensä jotakin suoraa tai mutkatonta.

Kerran kirjoittaja joutui kansainvälisessä konferenssissa kuuntelemaan esitelmää, jossa tutkija määritteli amerikkalaisen tavan käydä liikeneuvottelua ”lineaariseksi” ja kiinalaisen tavan käydä liikeneuvottelua ”epälineaariseksi”. Esitelmässä vertailtiin näitä kahta neuvottelukulttuuria. Koska esitelmän otsikossa oli sana ”lineaarinen”, sai tutkija esittää esitelmänsä matemaattisessa sessiossa. Esitelmä oli hauska, mutta täysin väärässä paikassa.

Matematiikassa sana ”lineaarinen” liitetään aina kahden lineaariavaruuden väliseen kuvaukseen.

Määritelmä 157 Olkoon $(L, +, \circ)$ ja $(M, \hat{+}, \hat{\circ})$ reaalisia lineaariavaruuksia, ja olkoon $f : L \rightarrow M, \vec{u} \mapsto f(\vec{u})$ kuvaus. f on lineaarinen eli lineaarikuvaus, jos

$$\begin{aligned} f(\vec{a} + \vec{b}) &= f(\vec{a}) \hat{+} f(\vec{b}), \quad \forall \vec{a}, \vec{b} \in L, \text{ ja} \\ f(\lambda \circ \vec{a}) &= \lambda \hat{\circ} f(\vec{a}), \quad \forall \lambda \in \mathbb{R}, \forall \vec{a} \in L. \end{aligned}$$

Lineaarikuvaus on siis kuvaus, joka kuvaa vektorin vektorille.

kuvaus_c4s3es1

Esimerkki 158 Jos f on määritelty seuraavasti: $f : E^3 \rightarrow E^3, \vec{u} \mapsto f(\vec{u}) = 2\vec{u}$, niin esimerkiksi

$$f\left(\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}\right) = 2 \circ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 6 \end{pmatrix}.$$

Esimerkki 159 Jokainen $m \times n$ -matriisi $\mathbf{A}$ määrittelee kuvauksen $f_{\mathbf{A}} : E^n \rightarrow E^m, \vec{u} \mapsto \mathbf{A}\vec{u}$. Tämä kuvaus on lineaarinen, sillä

$$\begin{aligned} f_{\mathbf{A}}(\vec{a} + \vec{b}) &= \mathbf{A}(\vec{a} + \vec{b}) = \mathbf{A}\vec{a} + \mathbf{A}\vec{b} = f_{\mathbf{A}}(\vec{a}) + f_{\mathbf{A}}(\vec{b}), \text{ ja} \\ f_{\mathbf{A}}(\lambda \circ \vec{a}) &= \mathbf{A}(\lambda \vec{a}) = \lambda(\mathbf{A}\vec{a}) = \lambda \circ f_{\mathbf{A}}(\vec{a}). \end{aligned}$$

kuvmat_c4s3th1

Lause 160 Olkoon $f : E^n \rightarrow E^m, \vec{u} \mapsto f(\vec{u})$ lineaarinen kuvaus. Silloin on olemassa $m \times n$ -matriisi $\mathbf{A}$, jolle

$$f(\vec{v}) = \mathbf{A}\vec{v}, \quad \forall \vec{v} \in E^n.$$

Todistus. Käytetään seuraavassa konstruktiossa vanhoja tuttuja vektoreita (ks. sivu) $\vec{e}_k = (\delta_{1k}, \delta_{2k}, \dots, \delta_{nk})^T$. Jos siis esimerkiksi $n = 4$, niin

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \text{ ja } \vec{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Muodostamme matriisin $\mathbf{A}$ siten, että sen k :s sarake on $f(\vec{e}_k)$. Olkoon nyt $\vec{v} = (v_1, v_2, \dots, v_n)^T \in E^n$ mikä tahansa vektori. Suora lasku osoittaa, että

$$\begin{aligned} f(\vec{v}) &= f(v_1\vec{e}_1 + v_2\vec{e}_2 + \dots + v_n\vec{e}_n) \\ &= v_1f(\vec{e}_1) + v_2f(\vec{e}_2) + \dots + v_nf(\vec{e}_n) \\ &= v_1\vec{a}_{\bullet 1} + v_2\vec{a}_{\bullet 2} + \dots + v_n\vec{a}_{\bullet n} \\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \mathbf{A}\vec{v} \end{aligned}$$

□

Esimerkki 161 Esimerkin [linkuvaus_c4s3es1](#) (158) lineaarikuvauksen matriisi on

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

Huomaa, että lause [linkumat_c4s3th1](#) (160) koskee vain Euklidisia vektoriavaruuksia E^n . Syy on tietenkin se, ettei esimerkiksi funktioavaruuksien tapauksessa funktion kertominen matriisilla ole hyvin määritelty. Sen jälkeen, kun kanta ja dimensio on määritelty, voidaan väite yleistää äärellisulotteisille lineaariavaruuksille. Mutta senkin jälkeen on paljon tärkeitä lineaariavaruuksia, joiden välisiä lineaarikuvauksia ei voi esittää matriisien avulla. Asiaan palataan tämän luvun viimeisessä kappaleessa.

Lause 162 Kahdesta lineaarikuvauksesta yhdistämällä saatu yhdistetty kuvaus on myös lineaarikuvaus.

Todistus. Olkoon $g : L \rightarrow M$, ja $f : M \rightarrow N$ lineaarikuvauksia. Silloin

$$\begin{aligned}(f \circ g)(\vec{a} + \vec{b}) &= f(g(\vec{a} + \vec{b})) = f(g(\vec{a}) + g(\vec{b})) = f(g(\vec{a})) + f(g(\vec{b})) \\ &= (f \circ g)(\vec{a}) + (f \circ g)(\vec{b}), \text{ ja} \\ (f \circ g)(\lambda \circ \vec{a}) &= f(g(\lambda \vec{a})) = f(\lambda g(\vec{a})) = \lambda f(g(\vec{a})) \\ &= \lambda (f \circ g)(\vec{a}).\end{aligned}$$

□

Lause 163 Jos $f : L \rightarrow M$ on lineaarikuvaus, niin $f(\vec{0}) = \vec{0}$.

Todistus. $f(\vec{0}) = f(0 \circ \vec{0}) = 0 \circ f(\vec{0}) = \vec{0}$.

□

7.5 (x, y) -Tason siirrot, venytykset, peilaukset ja kierrot

Sec:R2Transf

Tässä kappaleessa tutkimme yksinkertaisia kuvauksia $E^2 \rightarrow E^2$. Otamme käyttöön koulusta tutun tavan piirtää (x, y) -koordinaatisto, jonka piste kohdassa x korkeudella y vastaa vektoria $(x, y)^T$.

Käytämme seuraavassa kuvapisteele $f((x, y)^T)$ merkintää $(u, v)^T$. Piirrämme samaan kuvaan (x, y) -koordinaston päälle (u, v) -koordinaatiston. Näin voimme havainnollisesti esittää kuvapisteen ja alkukuvapisteen samassa kuvassa. (Jos haluamme tiivistää esityksemme äärimmilleen, kirjoitamme: $\vec{z} = (x, y)^T$, $\vec{w} = (u, v)^T$ ja $f(\vec{z}) = \vec{w}$.)

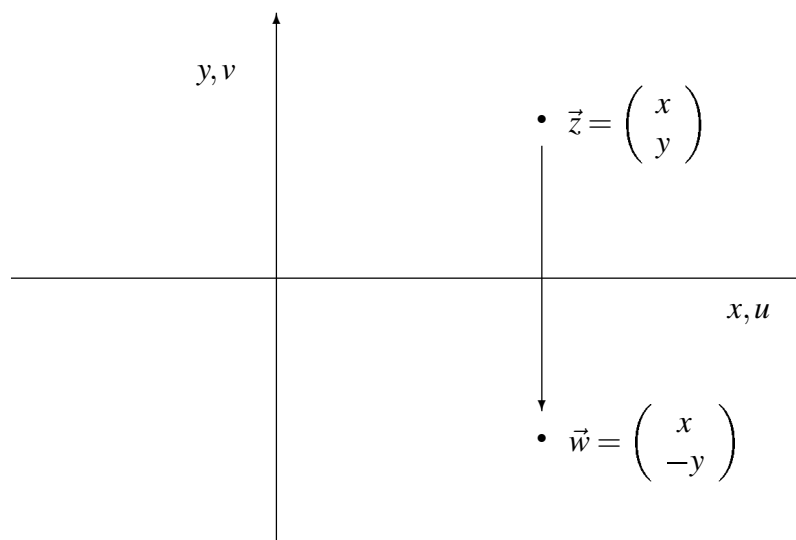
Peilaus x -akselin suhteen. Kuvaus $p_x : E^2 \rightarrow E^2, (x, y)^T \mapsto (x, -y)^T$ peilaa pisteen x -akselin toiselle puolelle. (Ks. kuva (7.7)).

Kuvaus p_x voidaan esittää myös matriisin avulla

$$p_x \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

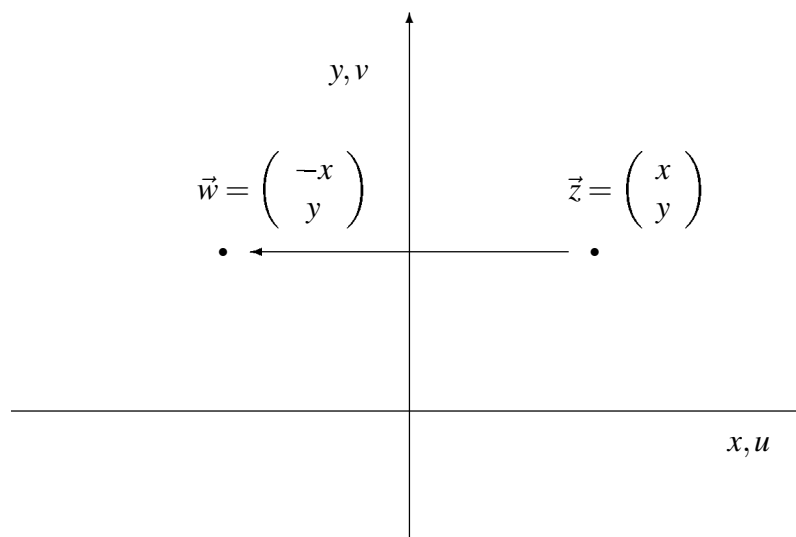
Koska kuvaus voidaan määritellä näin matriisin avulla, se on lineaarikuvaus.

c4s4fig1



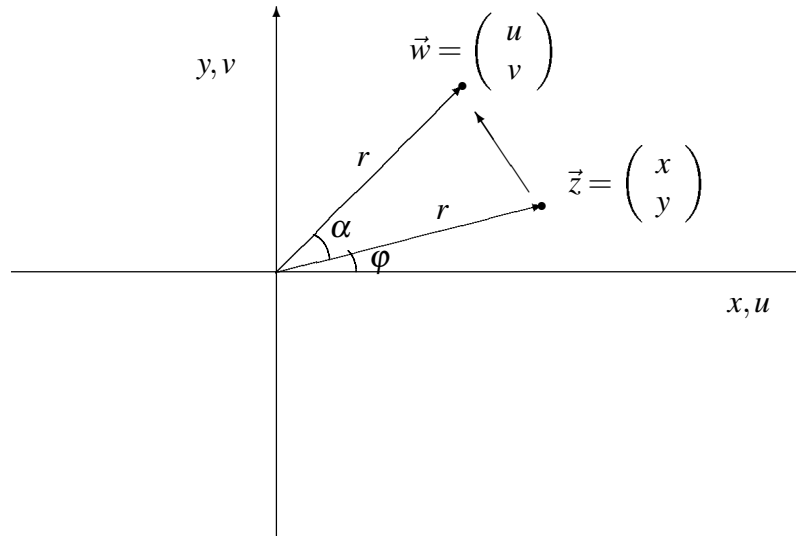
Kuva 30. Peilaus x -akselin suhteen.

c4s4fig2



Kuva 31. Peilaus y -akselin suhteen.

c4s4fig3



Kuva 32. Kierto kulman α verran origon ympäri.

Peilaus y-akselin suhteen. Kuvaus $p_x: E^2 \rightarrow E^2, (x, y)^T \mapsto (-x, y)^T$ peilaa pisteen x -akselin toiselle puolelle. (Ks. kuva (??)).

Kuvaus p_y voidaan esittää myös matriisin avulla

$$p_y \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Koska kuvaus voidaan määritellä näin matriisin avulla, se on lineaarikuvaus.

Kierto kulman α verran origon ympäri. Määritellään, että piste $\vec{w} = r_\alpha(\vec{z})$ on yhtä kaukana origosta kuin $\vec{z}$, mutta sen paikkavektorin ja x -akselin välinen kulma on α :n verran suurempi kuin vastaava pisteeseen $\vec{z}$ liittyvä kulma. (Ks. kuva (??)).

Jos alkukuva on piste

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

niin trigonometristen kaavojen avulla saamme välittömästi

$$\begin{aligned} \begin{pmatrix} u \\ v \end{pmatrix} &= \begin{pmatrix} r \cos(\alpha + \varphi) \\ r \sin(\alpha + \varphi) \end{pmatrix} = \begin{pmatrix} r \cos \alpha \cos \varphi - r \sin \alpha \sin \varphi \\ r \sin \alpha \cos \varphi + r \cos \alpha \sin \varphi \end{pmatrix} \\ &= \begin{pmatrix} x \cos \alpha - y \sin \alpha \\ x \sin \alpha + y \cos \alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Kierto origon ympäri on siis lineaarinen kuvaus!

Siirto matkan $\vec{d}$ verran. Siirto $s_{\vec{d}}: E^2 \rightarrow E^2, \vec{z} \mapsto \vec{z} + \vec{d}$ ei ole lineaarikuvaus, sillä $s_{\vec{d}}(\vec{0}) \neq \vec{0}$.

Siis: kierto origon ympäri on lineaarinen kuvaus, mutta siirto ei ole lineaarinen kuvaus.

Siirto on käytännössä tärkeä kuvaus ja sitä paljon tarvitaan. On muistettava, että kun tässä kappaleessa mainittuja kuvauksia yhdistetään, niin jos yksikin osakuvaus on siirto, niin yhdistetty kuvaus ei ole lineaarinen, eikä sitä voida esittää yhden matriisin avulla.

Peilaus suoran $y = x$ suhteen. Peilaus suoran $y = x$ suhteen saadaan yhdistämällä kolme edellä kuvattua kuvausta: (1) kierto kulman $-\pi/4$ verran (2) peilaus x -akselin suhteen (3) kierto kulman $\pi/4$ verran. Kuvaukseen liittyvä matriisi on siis

$$\begin{aligned} A &= \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(-\pi/4) & -\sin(-\pi/4) \\ \sin(-\pi/4) & \cos(-\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(\pi/4) & \sin(\pi/4) \\ -\sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \begin{pmatrix} \cos(\pi/4) & \sin(\pi/4) \\ \sin(\pi/4) & -\cos(\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

7.6 Lineaarinen aliavaruus ja span

7.7 Ydin ja kuva

Sec:MapKer

Teemme ensin yleiset tarkastelut ja sitten keskitymme matriiseihin.

Määritelmä 164 Olkoon $f: L_1 \rightarrow L_2$ lineaarikuvaus. Sanomme, että L_1 on kuvauksen lähtöavaruus (engl. domain) ja L_2 on kuvauksen maaliavaruus (engl. codomain).

Lineaarikuvauksen ydin (engl. kernel) on joukko

$$\ker(f) = \{\vec{u} \in L_1 \mid f(\vec{u}) = \vec{0}\} \subseteq L_1.$$

Lineaarikuvauksen kuva (engl. range) on joukko

$$\text{ran}(f) = \{\vec{w} \in L_2 \mid f(\vec{u}) = \vec{w} \text{ jollakin } \vec{u} \in L_1\} \subseteq L_2.$$

Ydin on siis eräänlainen ”nollakohtien” joukko ja kuva on täsmälleen se mitä peruskurssilla jo on määritelty kuvauksen kuvaksi.

c4s6th2

Lause 165 Olkoon $f : L_1 \rightarrow L_2$ lineaarikuvaus. Lineaarikuvauksen ydin $\ker(f)$ on L_1 :n lineaarinen aliavaruus, ja kuva $\text{ran}(f)$ on L_2 :n lineaarinen aliavaruus.

Todistus. HT. □

Lause 166 Olkoon $f : L_1 \rightarrow L_2$ lineaarikuvaus.

- (1) $\vec{0}_1 \in \ker(f)$ ja $\vec{0}_2 \in \text{ran}(f)$.
- (2) Lineaarikuvaus f on injektio, jos ja vain jos $\ker(f) = \{\vec{0}_1\}$.
- (3) Lineaarikuvaus on surjektio, jos ja vain jos $\text{ran}(f) = L_2$.

Todistus. (1) $f(\vec{0}_1) = f(0 \circ \vec{0}_1) = 0 \circ f(\vec{0}_1) = \vec{0}_2$.

(2) $\Rightarrow$ Olkoon f injektio. Silloin $\vec{0}_2$:n alkukuvia voi olla vain yksi. Siis $f^{-1}(\{\vec{0}_2\}) = \ker(f) = \{\vec{0}_1\}$.

$\Leftarrow$ Olkoon $\ker(f) = \{\vec{0}_1\}$ ja olkoon $\vec{u} \in L_1$ ja $\vec{v} \in L_1$ sellaiset, että $f(\vec{u}) = f(\vec{v})$. Kuvauksen f lineaarisuuden perusteella

$$f(\vec{u} - \vec{v}) = f(\vec{u}) - f(\vec{v}) = \vec{0}_2$$

Siis $\vec{u} - \vec{v} \in \ker(f)$. Oletuksen mukaan ydin on triviaali, joten

$$\begin{aligned}\vec{u} - \vec{v} &= \vec{0} \\ \Leftrightarrow \vec{u} &= \vec{v}\end{aligned}$$

(3) Seuraa suoraan surjektion ja kuvan määritelmästä. □

Samoista asioista käytetään hieman erilaisia merkintöjä, kun lineaarikuvaus on matriisin määräämä Euklidisten lineaariavaruuksien välinen kuvaus.

Määritelmä 167 Olkoon $f : E^n \rightarrow E^m, f(\vec{u}) = \mathbf{A}\vec{u}$ lineaarikuvaus. Tämän kuvauksen ydin on matriisin $\mathbf{A}$ ydin (tai nolla-avaruus) (engl. null space)

$$\text{null}(\mathbf{A}) = \{\vec{u} \in E^n \mid \mathbf{A}\vec{u} = \vec{0}\}$$

Kuvauksen kuva on matriisin $\mathbf{A}$ kuva (tai sarake-avaruus) (engl. column space)

$$\text{col}(\mathbf{A}) = \{\vec{v} \in E^m \mid \mathbf{A}\vec{u} = \vec{v}, \text{ jollakin } \vec{u} \in E^n\}$$

Matriisin sarake-avaruus on helppo luonnehtia sanomalla, että:

$$\begin{aligned}
 \text{col}(\mathbf{A}) &= \{\mathbf{A}\vec{u} \mid \vec{u} \in E^n\} \\
 &= \{u_1\vec{a}_{\bullet 1} + u_2\vec{a}_{\bullet 2} + \cdots + u_n\vec{a}_{\bullet n} \mid u_1, u_2, \dots, u_n \in \mathbb{R}\} \\
 &= \text{span}\{\vec{a}_{\bullet 1}, \vec{a}_{\bullet 2}, \dots, \vec{a}_{\bullet n}\} \\
 &= \text{matriisin } \mathbf{A} \text{ sarakkeiden virittämä } E^n\text{:n aliavaruus}
 \end{aligned}$$

Matriisin ytimen määrittäminen on hieman mutkikas tehtävä, mutta se perustuu tuttuun tekniikkaan, joten esitämme esimerkin.

c4s6es1 **Esimerkki 168** Olkoon tutkittavana 4×6 -matriisi

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & 3 & 0 & 3 & 0 \\ 0 & 1 & 1 & 2 & 3 & 0 \\ 1 & 4 & 5 & -1 & 4 & 1 \\ 2 & 0 & 2 & -3 & -1 & 2 \end{pmatrix}$$

Määritämme matriisin ytimen. On siis löydettävä kaikki yhtälön $\mathbf{M}\vec{x} = \vec{0}$ ratkaisut. Sitä varten alamme ratkaista tätä yhtälöryhmää. Käytämme rivioperaatioita algortmiversion II mukaisesti, ja aina kun pivotointi epäonistuu, jätämme sarakkeen paikalleen ja jatkamme pivotointia seuraavasta sarakkeesta.

$$\begin{aligned}
 &\left(\begin{array}{cccccc|c} [1] & 2 & 3 & 0 & 3 & 0 & 0 \\ 0 & 1 & 1 & 2 & 3 & 0 & 0 \\ 1 & 4 & 5 & 2 & 7 & 1 & 0 \\ 2 & 0 & 2 & -4 & -2 & 2 & 0 \end{array} \right) \left| \begin{array}{c} -1 \\ \leftarrow \\ \leftarrow \\ \leftarrow \end{array} \right. \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 2 & 3 & 0 & 3 & 0 & 0 \\ 0 & [1] & 1 & 2 & 3 & 0 & 0 \\ 0 & 2 & 2 & 2 & 4 & 1 & 0 \\ 0 & -4 & -4 & -4 & -8 & 2 & 0 \end{array} \right) \left| \begin{array}{c} \leftarrow \\ -2 \\ \leftarrow \\ \leftarrow \end{array} \right. \left| \begin{array}{c} -2 \\ +4 \\ \leftarrow \end{array} \right. \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 0 & 1 & -4 & -3 & 0 & 0 \\ 0 & (1) & 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & [-2] & -2 & 1 & 0 \\ 0 & 0 & 0 & 4 & 4 & 2 & 0 \end{array} \right) \left| \begin{array}{c} \leftarrow \\ \leftarrow \\ -2 \\ \leftarrow \end{array} \right. \left| \begin{array}{c} \leftarrow \\ +1 \\ +2 \\ \leftarrow \end{array} \right. \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 0 & 1 & 0 & 1 & -2 & 0 \\ 0 & (1) & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & (-2) & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & [4] & 0 \end{array} \right) \left| \begin{array}{c} \leftarrow \\ \leftarrow \\ +1/2 \\ \leftarrow \end{array} \right. \left| \begin{array}{c} \leftarrow \\ \leftarrow \\ -1/4 \\ -1/4 \end{array} \right. \left| \begin{array}{c} \leftarrow \\ \leftarrow \\ \leftarrow \\ :4 \end{array} \right. : (-2) \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & (1) & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & (1) & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1) & 0 \end{array} \right)
 \end{aligned}$$

Siirretään kolmas ja viides sarake *RHS*:iin ja kirjoitetaan yhtälöt yhtälöiksi. Lisätään yhtälöihin kaksi triviaalia yhtälöä $x_3 = x_3$ ja $x_5 = x_5$.

$$\begin{cases} x_1 = -x_3 - x_5 \\ x_2 = -x_3 - x_5 \\ x_3 = x_3 \\ x_4 = -x_5 \\ x_5 = x_5 \\ x_6 = 0 \end{cases}$$

Jos valitsemme arvot $x_3 = a \in \mathbb{R}$ ja $x_5 = b \in \mathbb{R}$, niin ytimeen kuuluvalla vektorille $\vec{x}$ saadaan esitys

$$\vec{x} = a \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} = a\vec{u}_1 + b\vec{u}_2$$

Koska a ja b voidaan valita vapaasti, kuuluvat kaikki mahdolliset vektoreiden $\vec{u}_1$ ja $\vec{u}_2$ lineaarikombinaatiot matriisin $\mathbf{M}$ ytimeen. Siis

$$\text{null}(\mathbf{M}) = \text{span}\{\vec{u}_1, \vec{u}_2\}.$$

c4s6th3

Lause 169 Olkoon $\mathbf{A}$ $m \times n$ -matriisi ja olkoon $f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \vec{x} \mapsto \mathbf{A}\vec{x}$ matriisin generoima lineaarikuvaus. Silloin

(1) jos f on injektio, niin $m \geq n$.

(2) jos f on surjektio, niin $m \leq n$.

(3) jos f on bijektio, niin $m = n$.

Todistus. (1) Jos f on injektio, niin lauseen ^(c4s6th2)(I65) mukaan $\ker(\mathbf{A}) = \{\vec{0}\}$. Silloin tulee esimerkin ^(c4s6es1)(I68) mukaisessa konstruktiossa pivotointien onnistua jokaisessa sarakkeessa, joten on mahdotonta, että rivejä olisi vähemmän kuin sarakkeita.

(2) Osoitamme, että jos $m > n$, niin f ei voi olla surjektio. Olkoon siis $m > n$. Ratkaisemme Rivioperaatioiden avulla yhtälöryhmän $\mathbf{A}\vec{x} = \vec{0}$. Koska kerroinmatriisissa rivejä on enemmän kuin sarakkeita, on viimeisen kaavion viimeinen rivi muotoa $0 = 0$. Jos ratkaisun aikana tehtyjen rivioperaatioiden jono vastaa matriiseilla $\mathbf{E}_p \cdot \mathbf{E}_2 \mathbf{E}_1$ kertomista, niin asetamme $\vec{b} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdot \mathbf{E}_p^{-1} \vec{e}_n$. Jos nyt ratkaisemme yhtälöryhmän $\mathbf{A}\vec{x} = \vec{b}$ rivioperaatioilla, niin viimeisen kaavion viimeinen rivi on tyyppiä $0 = 1$. Siis vektorilla $\vec{b} \in \mathbb{R}^n$ ei ole alkukuvaa kuvauksessa f .

(3) Seuraa kohdista (1) ja (2). □

c4s6th1

Lause 170 $n \times n$ -neliömatriisin $\mathbf{A}$ ydin on triviaali, jos ja vain jos matriisin determinantti eroaa nolasta, $\det(\mathbf{A}) \neq 0$.

Todistus. Edellisen esimerkin mukainen konstruktio ei löydä yhtään ytimeen kuuluvaa vektoria, jos ja vain jos pivointi onnistuu joka sarakkeessa ratkaistaessa yhtälöryhmää $\mathbf{A}\vec{x} = \vec{0}$ Gaussin algoritmilla. Tämä on yhtäpitävää sen kanssa, että $\det(\mathbf{A}) \neq 0$. $\square$

8 KANNAT JA ORTOGONAALISUUS

:KantaOrthogon

8.1 Lineaarinen riippumattomuus

Sec:LinIndep

Lineaarinen riippumattomuus on oikeastaan jo määritelty, mutta kirjoitamme määritelmät uudestaan tähän kappaleeseen, jotta esitys olisi helpommin luettavissa. Lisäksi otamme käyttöön uuden tavan erotella 'joukot', 'jonot' ja 'systeemit'.

Vektorijoukolle, jossa on q vektoria käytetään merkintää $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$. Alkioiden järjestyksellä ei nyt ole väliä, vaikka käytännössä ne joudutaankin aina kirjoittamaan johonkin järjestykseen. Sama alkio ei toistu siistityssä vektori-luettelossa.

Vektorijonossa $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ vektoreiden järjestys on listan mukainen ja sama vektori voi esiintyä useamman kuin yhden kerran. Matematiikassa tyypillinen tapa merkitä jono kaarisulkeita käyttäen voisi nyt aiheuttaa väärinkäsityksiä. Siksi käytämme vektorijonolle samaa merkintää kuin vektorijoukolle!

Vektorisysteemissä $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ voi sama vektori esiintyä useamman kuin yhden kerran, mutta vektoreiden järjestyksellä ei ole väliä. (Voidaan sanoa, että kaksi vektorisysteemiä ovat samat, jos niiden luetteloista saadaan permutoimalla samat.)

Esimerkki 171 (1) Olkoon $L = \{\mathbf{a}, \mathbf{b}\}$, $M = \{\mathbf{a}, \mathbf{a}, \mathbf{b}\}$ ja $N = \{\mathbf{a}, \mathbf{b}, \mathbf{a}\}$. Silloin vektorijoukkoina $L = M = N$; vektorijonoina $M \neq N$; mutta vektorisysteeminä $M = N$.

(2) a) Jos viisipäiväisen loman aikana tiskarijoukko on $\{\text{Matti, Heikki, Ville}\}$, niin nämä kolme henkilöä siis tiskaavat.

b) Jos loman aikana tiskarijono on $\{\text{Matti, Ville, Heikki, Ville, Ville}\}$, niin luettelo kertoo täsmällisesti tiskivuorojen järjestyksen.

c) Jos tiskarisysteemi on $\{\text{Matti, Heikki, Ville, Ville, Ville}\}$, niin kunkin osalta on tiedossa miten monena päivänä joutuu tiskaamaan, mutta järjestys on jätetty auki.

Seuraavassa lähes aina, kun puhumme vektorisysteemistä, joudumme käytännössä käsittelemään vektorijonoa. Kun sanomme, että vektorisysteemillä on tietty ominaisuus,

niin tarkoitamme, että kyseinen ominaisuus on vastaavan vektorijonon kaikilla permutaatioilla. (Siis ”järjestyksellä ei ole väliä”.)

Määritelmät annetaan taas yleisessä muodossa, mutta tärkeimmät sovellukset koskevat $\mathbb{R}^n$:ää.

Määritelmä 172 L :n vektorisysteemi $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippuva, eli sidottu (engl. linearly dependent), jos on olemassa reaalityyppiset $\lambda_1, \lambda_2, \dots, \lambda_q \in \mathbb{R}$ siten, että

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \dots + \lambda_q \mathbf{u}_q = \mathbf{0}$$

ja ainakin yksi kertoimista eroaa nolasta ($\lambda_k \neq 0$), jollakin $k = 1, 2, \dots, q$.

Vektorisysteemi $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippumaton, eli vapaa (engl. linearly independent), jos se ei ole sidottu.

Määritelmästä seuraa välittömästi seuraavat tulokset:

Lause 173 Tarkastellaan vektorisysteemiä $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$.

- (1) Jos jokin systeemin vektoreista on nollavektori, niin systeemi on sidottu.
- (1b) Jos systeemi on vapaa, niin mikään vektoreista ei ole nollavektori.
- (2) Jos sama vektori esiintyy luettelossa kaksi kertaa, niin systeemi on sidottu.
- (2b) Jos systeemi on vapaa, niin $\mathbf{u}_i \neq \mathbf{u}_j$ aina, kun $i \neq j$.
- (3) Jos jokin tarkasteltavan systeemin osasysteemi on sidottu, niin systeemi on sidottu.
- (3b) Jos systeemi on vapaa, niin jokainen sen osasysteemi on vapaa.

Todistus.HT.

□

Määritelmä 174 Olkoon $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ vektorijono ja olkoot $\lambda_1, \lambda_2, \dots, \lambda_q \in \mathbb{R}^n$ reaalityyppiset. Lauseketta

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \dots + \lambda_q \mathbf{u}_q$$

sanomme linearikombinaatioksi vektoreista $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ ja lukuja $\lambda_1, \lambda_2, \dots, \lambda_q$ sanomme linearikombinaation kertoimiksi. Jos kaikki kertoimet ovat nollia, $\lambda_1 = \lambda_2 = \dots = \lambda_q = 0$, niin sanomme, että lineaarikombinaatio on triviaali. Jos jokin kertoimista eroaa nolasta, $\lambda_k \neq 0$, niin sanomme, että lineaarikombinaatio on ei-triviaali.

Jos vektori $\mathbf{a}$ voidaan lausua summana

$$\mathbf{a} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_q \mathbf{u}_q,$$

niin sanomme, että $\mathbf{a}$ voidaan lausua lineaarikombinaationa vektoreista $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$. Jos edellä ainakin yksi kerroin eroaa nollasta, niin sanomme, että vektori $\mathbf{a}$ voidaan lausua ei-triviaalina lineaarikombinaationa vektoreista $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$.

Lause 175 Jos vektorisysteemi $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippuva, niin jokin systeemin vektoreista voidaan lausua ei-triviaalina lineaarikombinaationa muista systeemin vektoreista.

Todistus. HT

□

$\mathbb{R}^n$:ssä liitämme vektorijonoon (-systeemiin) $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ matriisiin $\mathbf{V}$, jonka k :s sarakke on jonon k :s vektori. Siis

$$(\mathbf{V})_{\bullet k} = \vec{v}_k.$$

Jos $\vec{a} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_q \vec{v}_q$, niin

$$\mathbf{V}\vec{\lambda} = \vec{a}, \text{ missä } \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_q \end{pmatrix}.$$

Näillä merkinnöillä vektorisysteemin lineaarinen riippumattomuus voidaan luonnehtia seuraavasti:

$$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\} \text{ on vapaa} \quad (8.1)$$

$$\Leftrightarrow \mathbf{V}\vec{\lambda} = \vec{0}, \text{ jos ja vain jos } \vec{\lambda} = \vec{0} \quad (8.2)$$

$$\Leftrightarrow \ker(\mathbf{V}) = \{\vec{0}\} \quad (8.3)$$

Seuraavassa lauseessa tätä ideaa kehitetään edelleen.

c4s7th1

Lause 176 Tarkastellaan $\mathbb{R}^n$:n vektorisysteemiä $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ ja siihen liittyvää matriisiä $\mathbf{V}$. Silloin seuraavat ehdot ovat yhtäpitävät

$$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\} \text{ on vapaa} \quad (8.4)$$

$$\Leftrightarrow \ker(\mathbf{V}) = \{\vec{0}\} \quad (8.5)$$

$$\Leftrightarrow \ker(\mathbf{V}^T \mathbf{V}) = \{\vec{0}\} \quad (8.6)$$

$$\Leftrightarrow \det(\mathbf{V}^T \mathbf{V}) \neq 0 \quad (8.7)$$

$$\Leftrightarrow \mathbf{V}^T \mathbf{V} \text{ on säännöllinen} \quad (8.8)$$

Todistus. $(8.4) \Leftrightarrow (8.5)$ seuraa suoraan määritelmistä.

$(8.5) \Leftrightarrow (8.6)$ Jos $\vec{u} \in \ker(V)$, niin $\mathbf{V}^T \mathbf{V} \vec{u} = \mathbf{V}^T \vec{0} = \vec{0}$. Siis $\ker(V) \subseteq \ker(\mathbf{V}^T \mathbf{V})$. Toisaalta olkoon $\vec{u} \in \ker(\mathbf{V}^T \mathbf{V})$. Silloin

$$0 = \langle \mathbf{V}^T \mathbf{V} \vec{u} | \vec{u} \rangle = \langle \mathbf{V} \vec{u} | \mathbf{V} \vec{u} \rangle \Rightarrow \mathbf{V} \vec{u} = \vec{0}$$

Siis $\ker(\mathbf{V}^T \mathbf{V}) \subseteq \ker(V)$.

$(8.6) \Leftrightarrow (8.7)$ Seuraa siitä, että $\mathbf{V}^T \mathbf{V}$ on neliömatriisi ja lauseesta (I70) sivulla 150.

$(8.7) \Leftrightarrow (8.8)$ Seuraa lauseesta (I04) sivulla 102. □

Esimerkki 177 Tutkitaan onko vektorijono $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ lineaarisesti riippumaton, kun

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -3 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} 4 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Tutkitaan asia (1) ”kynällä ja paperilla” ja (2) ”tietokonetta käyttäen”.

(1) Etsimme ei-triviaaleja λ -kertoimia, joilla yhtälö $\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0}$ on tosi. Jos löytyy ei-triviaalit kertoimet, jono on sidottu, ja jos ei löydy, niin jono on vapaa. Ryhdyimme työhön:

$$\begin{aligned} \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0} &\Leftrightarrow \begin{cases} \lambda_1 + 2\lambda_2 + 4\lambda_3 = 0 \\ \lambda_2 - \lambda_3 = 0 \\ 2\lambda_1 - \lambda_2 = 0 \\ -3\lambda_1 + \lambda_3 = 0 \end{cases} \\ &\sim \left(\begin{array}{ccc|c} [1] & 2 & 4 & 0 \\ 0 & 1 & -1 & 0 \\ 2 & -1 & 0 & 0 \\ -3 & 0 & 1 & 0 \end{array} \right) \begin{array}{l} -2 \\ \leftarrow \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & [1] & -1 & 0 \\ 0 & -5 & -8 & 0 \\ 0 & 6 & 12 & 0 \end{array} \right) \begin{array}{l} +5 \\ \leftarrow \\ \leftarrow \end{array} \\ &\sim \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & (1) & -1 & 0 \\ 0 & 0 & [-13] & 0 \\ 0 & 0 & 18 & 0 \end{array} \right) \begin{array}{l} +18/13 \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & (1) & -1 & 0 \\ 0 & 0 & (-13) & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \\ &\Leftrightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0 \end{aligned}$$

Siis vektorisysteemi on vapaa, eli lineaarisesti riippumaton.

(2) Matlabilla saman asian voi todeta lauseen (I76) perusteella:

```
>> U = [1 0 2 -3 ; 2 1 -1 0 ; 4 -1 0 1]'
```

```
U =
```

```

     1     2     4
     0     1    -1
     2    -1     0
    -3     0     1
```

```
>> det (U' * U)
```

```
ans =
```

```
820
```

```
>>
```

Koska siis $\det(\mathbf{U}^T \mathbf{U}) \neq 0$, niin vektorisysteemi on vapaa.

Esimerkki 178 Otetaan vielä esimerkki, jossa johtopäätös on erilainen. Tutkitaan onko vektorijono $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ lineaarisesti riippumaton, kun

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -3 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} -1 \\ -2 \\ 8 \\ -9 \end{pmatrix}.$$

(1) Etsimme taas ei-triviaaleja λ -kertoimia, joilla yhtälö $\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0}$ on tosi. Jos löytyy ei-triviaalit kertoimet, jono on sidottu, ja jos ei löydy, niin jono on vapaa. Ryhdymme työhön:

$$\begin{aligned} \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0} &\Leftrightarrow \begin{cases} \lambda_1 + 2\lambda_2 - \lambda_3 = 0 \\ \lambda_2 - 2\lambda_3 = 0 \\ 2\lambda_1 - \lambda_2 + 8\lambda_3 = 0 \\ -3\lambda_1 - 9\lambda_3 = 0 \end{cases} \\ &\sim \left(\begin{array}{ccc|c} [1] & 2 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 2 & -1 & 8 & 0 \\ -3 & 0 & -9 & 0 \end{array} \right) \begin{array}{l} -2 \\ +3 \\ \leftarrow \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|c} (1) & 2 & 1 & 0 \\ 0 & [1] & -2 & 0 \\ 0 & -5 & 10 & 0 \\ 0 & 6 & -12 & 0 \end{array} \right) \begin{array}{l} \\ +5 \\ \leftarrow \\ \leftarrow \end{array} \\ &\sim \left(\begin{array}{ccc|c} (1) & 2 & 1 & 0 \\ 0 & (1) & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \Leftrightarrow \begin{cases} \lambda_1 = 3a \\ \lambda_2 = 2a \\ \lambda_3 = a \end{cases}, \quad a \in \mathbb{R} \end{aligned}$$

Koska löytyy ei-triviaalit kertoimet, vektorisysteemi on sidottu, eli lineaarisesti riippuva.

(2) Matlabilla saman asian voi todeta lauseen ^{lc4s7th1}(176) perusteella:

```
>> U = [1 0 2 -3 ; 2 1 -1 0 ; -1 -2 8 -9]'
```

```
U =
```

```

      1      2     -1
      0      1     -2
      2     -1      8
     -3      0     -9
```

```
>> det (U' * U)
```

```
ans =
```

```
0
```

```
>>
```

Koska siis $\det(\mathbf{U}^T \mathbf{U}) = 0$, niin vektorisysteemi on sidottu.

Lause 179 Olkoon $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ lineaarisesti riippumaton $\mathbb{R}^n$:n vektorisysteemi. Jos $\vec{a} \in \mathbb{R}^n$, $\vec{a} \neq \vec{0}$ ja $\vec{a} \perp \vec{v}_k$, $\forall k = 1, \dots, q$, niin vektorisysteemi

$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{a}\}$ on vapaa.

Todistus. Olkoon $\mathbf{V}$ se $n \times q$ -matriisi, jonka sarakkeina ovat vektorit $\vec{v}_k, k = 1, \dots, q$ ja olkoon $\mathbf{W} = (\mathbf{V} | \vec{a})$ vastaava $n \times (q+1)$ -matriisi, jonka viimeinen sarake on vektori $\vec{a}$. Oletuksesta $\vec{a} \neq \vec{0}$ seuraa, että $\|\vec{a}\| > 0$ ja oletuksesta $\vec{a} \perp \vec{v}_k, \forall k$ seuraa, että $\mathbf{V}^T \vec{a} = \vec{0}$. Nyt voimme laskea suoraan

$$\begin{aligned} \mathbf{W}^T \mathbf{W} &= (\mathbf{V} | \vec{a})^T (\mathbf{V} | \vec{a}) = \begin{pmatrix} \mathbf{V}^T \\ \vec{a}^T \end{pmatrix} (\mathbf{V} | \vec{a}) \\ &= \left(\begin{array}{c|c} \mathbf{V}^T \mathbf{V} & \mathbf{V}^T \vec{a} \\ \hline \vec{a}^T \mathbf{V} & \vec{a}^T \vec{a} \end{array} \right) = \left(\begin{array}{c|c} \mathbf{V}^T \mathbf{V} & \vec{0} \\ \hline \vec{0}^T & \|\vec{a}\|^2 \end{array} \right) \end{aligned}$$

$$\Rightarrow \det(\mathbf{W}^T \mathbf{W}) = \|\vec{a}\|^2 \det(\mathbf{V}^T \mathbf{V}) \neq 0$$

Siis lauseen [\(I 76\)](#) mukaan vektorijono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{a}\}$ on vapaa. □

Yhdistämällä lause [\(I 76\)](#) lauseeseen [\(I 49\)](#) saadaan seuraava tulos, jota myöhemmin pystymme vielä parantamaan:

Lause 180 Jos $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_q\}$ on lineaarisesti riippumaton $\mathbb{R}^n$:n vektorisysteemi ja $M = \text{span}(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_q)$, niin

$$\mathbb{R}^n = M \oplus M^\perp$$

Seuraa suoraan lauseesta [\(I 76\)](#) sivulla [I 53](#) ja lauseesta [\(I 49\)](#) sivulla [I 35](#). □

8.2 Kanta ja dimensio

Sec:BaseDim

c4s8def1

Määritelmä 181 Lineaariavaruuden L vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaariavaruuden L kanta, jos

1. Jokainen L :n vektori $\mathbf{a} \in L$ voidaan lausua lineaarikombinaationa jonon F vektoreista

$$\mathbf{a} = \mu_1 \mathbf{u}_1 + \mu_2 \mathbf{u}_2 + \dots + \mu_q \mathbf{u}_q.$$

2. Jono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on vapaa.

Jos $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on kanta ja $\mathbf{a} = \mu_1 \mathbf{u}_1 + \mu_2 \mathbf{u}_2 + \dots + \mu_q \mathbf{u}_q$, niin kertoimet $\mu_1, \mu_2, \dots, \mu_q$ ovat vektorin $\mathbf{a}$ koordinaatit kannassa F .

Koordinaateista muodostettua vektoria $\vec{\mu} = (\mu_1, \mu_2, \dots, \mu_q)^T \in \mathbb{R}^q$ sanomme vektorin $\mathbf{a}$ koordinaattivektoriksi (kannassa F).

Esimerkki 182 (1) Vektorijono $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ eli vektorit

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

on $\mathbb{R}^n$:n kanoninen kanta. Kantaa sanotaan kanoniseksi, koska sillä on monta yksinkertaista ominaisuutta: jokaisen kantavektorin normi on 1, kantavektorit ovat keskenään kohtisuorat.

(2) Jos lineaariavaruuden L vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippumaton, niin se on virittämänsä aliavaruuden $M = \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q)$ kanta.

Jos vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaariavaruuden L kanta ja vektori $\mathbf{a} \in L$ on mikä tahansa L :n vektori, niin vektorin $\mathbf{a}$ koordinaatit kannassa F ovat yksikäsitteiset. Siis on olemassa yhdet ja vain yhdet koordinaatit. Jos nimittäin olisi olemassa kahdet erilaiset koordinaatit

$$\mathbf{a} = a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_q \mathbf{u}_q,$$

$$\mathbf{a} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_q \mathbf{u}_q, \quad \text{ja } a_k \neq \alpha_k \text{ jollakin } k,$$

niin silloin olisi

$$(a_1 - \alpha_1) \mathbf{u}_1 + (a_2 - \alpha_2) \mathbf{u}_2 + \dots + (a_q - \alpha_q) \mathbf{u}_q = \mathbf{0}, \quad (a_k - \alpha_k) \neq 0$$

mikä on mahdotonta, koska kanta on vapaa.

Lause 183 Vektorin koordinaatit kannassa ovat yksikäsitteiset.

Todistus. Perustelu edellä. □

Lause 184 Olkoon L lineaariavaruus ja olkoon vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ sen kanta. Silloin identtinen kuvaus

$$\text{id} : L \rightarrow L, \mathbf{a} \mapsto \mathbf{a}$$

ja kuvaus, joka kuvaa vektorin sen koordinaattivektorille kannassa F

$$v_F : L \rightarrow \mathbb{R}^q, \mu_1 \mathbf{u}_1 + \mu_2 \mathbf{u}_2 + \dots + \mu_q \mathbf{u}_q \mapsto (\mu_1, \mu_2, \dots, \mu_q)^T$$

ovat bijektiivisiä lineaarikuvauksia.

Todistus. Olkoon $\mathbf{a} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_q \mathbf{u}_q$, $\mathbf{b} = \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \dots + \beta_q \mathbf{u}_q$ ja $\lambda \in \mathbb{R}$. Silloin

$$\begin{aligned} \text{id}(\mathbf{a} + \mathbf{b}) &= \mathbf{a} + \mathbf{b} = \text{id}(\mathbf{a}) + \text{id}(\mathbf{b}) \\ \text{id}(\lambda \mathbf{a}) &= \lambda \mathbf{a} = \lambda \text{id}(\mathbf{a}) \\ v_F(\mathbf{a} + \mathbf{b}) &= \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_q + \beta_q \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_q \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_q \end{pmatrix} = v_F(\mathbf{a}) + v_F(\mathbf{b}) \\ v_F(\lambda \mathbf{a}) &= \begin{pmatrix} \lambda \alpha_1 \\ \lambda \alpha_2 \\ \vdots \\ \lambda \alpha_q \end{pmatrix} = \lambda v_F(\mathbf{a}) \end{aligned}$$

Siis kumpikin kuvaus on lineaarinen. Kumpikin kuvaus on selvästi bijektio. □

Lause 185 Olkoon vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ lineaariavaruuden L kanta. Jos $G = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ on toinen lineaariavaruuden L kanta, niin

$$k = q.$$

Toisin sanoen: Kaikissa mahdollisissa L :n kannoissa on sama määrä kantavektoreita.

Todistus. Seuraavassa kaaviossa kaikki kuvaukset ovat bijektiivisiä lineaarikuvauksia.

$$\begin{array}{ccc} & \text{id} & \\ L & \longrightarrow & L \\ v_F \downarrow & & \downarrow v_G \\ \mathbb{R}^q & & \mathbb{R}^k \end{array}$$

Yhdistämällä nämä kuvaukset saamme bijektiivisen lineaarikuvauksen

$$v_G \circ \text{id} \circ v_F^{-1} : \mathbb{R}^q \rightarrow \mathbb{R}^k$$

Tämä on lauseen (I69) ^{c4s6th3} mukaan mahdollista vain, jos $k = q$. □

Koska lineaariavaruuden kannan kantavektoreiden lukumäärä on sama kaikilla kannoilla, voimme määritellä seuraavasti.

Määritelmä 186 Jos lineaariavaruudella L on kanta, jossa on k kantavektoria, niin sanomme, että lineaariavaruuden L dimensio, $\dim(L)$, on k

Esimerkki 187 (1)

$$\dim(\mathbb{R}^n) = n$$

(2) Jos $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippumaton vektorijono, niin

$$\dim(\text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q)) = q$$

Siis vapaan vektorijonon virittämän aliavaruuden dimensio on sama kuin virittäjiensä lukumäärä.

Kaikki tason vektorit voidaan lausua kahden $\{\vec{i}, \vec{j}\}$ vektorin lineaarikombinaatioina, joten tason dimensio on 2. Toisaalta kolme tasovektoria muodostaa aina sidotun systeemin. Yleisemmin: jos vektorijonossa on liikaa vektoreita, se ei voi olla vapaa.

c4s8th5

Lause 188 Olkoon L lineaariavaruus, ja $\dim(L) = n$. Tarkastellaan L :n vektorijonoa $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$. Jos $q > n = \dim(L)$, niin vektorijono F on sidottu.

Todistus. (1) Todistetaan ensin se erikoistapaus, jossa $L = \mathbb{R}^n$ ja $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$. Olkoon $\mathbf{V}$ se $n \times q$ -matriisi, jonka sarakkeina ovat jonon F vektorit. Lähdemme nyt etsimään ei-triviaaleja kertoimia λ_i niin, että

$$\begin{aligned} \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_q \vec{v}_q &= \vec{0} \\ \Leftrightarrow \mathbf{V} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_q \end{pmatrix} &= \vec{0} \end{aligned}$$

Tämä on homogeeninen yhtälöryhmä, jonka kerroinmatriisissa $\mathbf{V}$ on n riviä ja q saraketta. Koska sarakkeita on enemmän kuin rivejä, jää Gaussin algoritmin mukaisessa ratkaisussa aina osa sarakkeista pivotoimatta, joten yhtälöryhmälle löytyy ei-triviaaleja ratkaisuja. Siis jono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ on sidottu.

(2) Nyt palaamme alkuperäiseen väitteeseen. Koska L :n dimensio on n , on olemassa kanta $G = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$, jossa on n kantavektoria. Olkoon nyt $\vec{v}_k = v_G(\mathbf{v}_k)$ vektorin $\mathbf{v}_k$ koordinaattivektori kannassa G . Todistuksen alkuosan mukaan nyt on olemassa ei-triviaalit kertoimet λ_i siten, että

$$\mathbf{V}\vec{\lambda} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \dots + \lambda_q\vec{v}_q = \vec{0}.$$

Silloin

$$\begin{aligned} \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \dots + \lambda_q\mathbf{v}_q &= \sum_{i=1}^q \lambda_i \left(\sum_{k=1}^n v_{ki} \mathbf{u}_k \right) \\ &= \lambda_1(v_{11}\mathbf{u}_1 + v_{21}\mathbf{u}_2 + \dots + v_{n1}\mathbf{u}_n) \\ &\quad + \lambda_2(v_{12}\mathbf{u}_1 + v_{22}\mathbf{u}_2 + \dots + v_{n2}\mathbf{u}_n) \\ &\quad + \dots \\ &\quad + \lambda_q(v_{1q}\mathbf{u}_1 + v_{2q}\mathbf{u}_2 + \dots + v_{nq}\mathbf{u}_n) \\ &= \mathbf{u}_1(v_{11}\lambda_1 + v_{12}\lambda_2 + \dots + v_{1q}\lambda_q) \\ &\quad + \mathbf{u}_2(v_{21}\lambda_1 + v_{22}\lambda_2 + \dots + v_{2q}\lambda_q) \\ &\quad + \dots \\ &\quad + \mathbf{u}_n(v_{n1}\lambda_1 + v_{n2}\lambda_2 + \dots + v_{nq}\lambda_q) \\ &= \sum_{k=1}^n \mathbf{u}_k (\mathbf{V}\vec{\lambda})_k = \vec{0} \end{aligned}$$

□

c4s8th6

Lause 189 *Olkoon L lineaariavaruus, ja $\dim(L) = n$. Jos L :n vektorijono $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ on vapaa, niin se on L :n kanta.*

Todistus. Kannan määritelmän mukaan nyt riittää osoittaa, että vektorijono virittää L :n. Olkoon sitä varten $\mathbf{a}$ mikä tahansa L :n vektori. Silloin jonossa $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{a}\}$ on ”liian monta vektoria ollakseen vapaa”. Edellisen lauseen mukaan $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{a}\}$ on sidottu. Siis joku vektoreista voidaan lausua muiden lineaarikombinaationa. Siinä yhtälössä $\mathbf{a}$ on mukana ja voidaan ratkaista F :n vektoreiden lineaarikombinaationa. □

c4s8th7

Lause 190 Olkoon L lineaariavaruus, ja $\dim(L) = n$. Tarkastellaan L :n vektorijonoa $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$. Jos $q < n$, ja vektorijono F on vapaa, niin on olemassa $n - q$ vektoria $\mathbf{a}_j \in L, j = 1, \dots, n - q$ siten, että jono

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q, \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-q}\}$$

on L :n kanta. (Sanomme, että vapaa jono $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ on täydennetty kannaksi vektoreilla $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-q}\}$.)

Todistus. Koska $\dim(L) = n$, on lineaariavaruudella kanta $G = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$. Jokin kannan G kantavektoreista $\mathbf{u}_k$ ei ole lausuttavissa jonon $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ vektoreiden lineaarikombinaationa, muuten jono $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ olisi L :n kanta, mikä on mahdonta, koska siinä ei ole tarpeeksi vektoreita. Lisätään kantavektori $\mathbf{u}_k$ vektorijonoon viimeiseksi. Uusi täydennetty vektorijono $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q, \mathbf{u}_k\}$ on vapaa ja siinä on $q + 1$ vektoria.

Toistamalla äskeistä menettelyä saamme lopulta täydennetyn vektorijonon, joka on vapaa ja jossa on n vektoria. Se on kanta. $\square$

c4s8th7b

Lause 191 Olkoon $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ vapaa jono lineaariavaruuden $\mathbb{R}^n$ vektoreita. Jos $\vec{b} \in L$ on kohtisuorassa kaikkia jonon vektoreita vastaan, niin täydennetty jono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{b}\}$ on vapaa.

Todistus. Kysymys on nyt siitä voidaanko $\vec{b}$ lausua ei-triviaalina lineaarikombinaationa jonon F vektoreista. Olkoon nyt

$$\vec{b} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_q \vec{v}_q = \mathbf{V} \vec{\lambda}$$

Koska $\vec{b}$ on kohtisuorassa jonon vektoreita vastaan, on

$$\vec{0} = \mathbf{V}^T \vec{b} = \mathbf{V}^T \mathbf{V} \vec{\lambda}$$

Lauseen (II 76) mukaan tämä on mahdollista vain, kun $\lambda_1 = \lambda_2 = \dots = \lambda_q = 0$. Siis täydennetty jono on vapaa. $\square$

c4s8th8

Lause 192 Tarkastellaan n -ulotteisen lineaariavaruuden $\mathbb{R}^n$ vapaata vektorijonoa $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ ja sen virittämää aliavaruutta $M = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$. Jos $q < n$, niin on olemassa $n - q$ vektoria $\vec{b}_j \in M^\perp, j = 1, \dots, n - q$ siten, että jono

$$G = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{b}_1, \vec{b}_2, \dots, \vec{b}_{n-q}\}$$

on L :n kanta ja jono

$$H = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_{n-q}\}$$

on aliavaruuden ortogonaalikomplementin $M^\perp$ kanta.

Erityisesti

$$\dim(M) + \dim(M^\perp) = n$$

Todistus. Lauseen (I90) todistuksen konstruktio tapahtui vaiheittain siten, että joka askeleessa jonoon lisätään yksi vektori $\vec{a}_k$. Koska $\mathbb{R}^n = M \oplus M^\perp$ on suora summa, joten lauseen (I34) mukaan vektori $\vec{a}_k$ voidaan yksikäsitteisellä tavalla jakaa kahden vektorin summaksi $\vec{a}_k = \vec{a}_k^M + \vec{a}_k^\perp$ siten, että $\vec{a}_k^M \in M$ ja $\vec{a}_k^\perp \in M^\perp$. Asetetaan $\vec{b}_k = \vec{a}_k^\perp$.

Lauseen (I91) mukaan täydennetty jono on vapaa, ja toistamalla menettelyä saamme täydennetyn jonon

$$G = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{a}_1^\perp, \vec{a}_2^\perp, \dots, \vec{a}_{n-q}^\perp\},$$

joka on vapaa ja siis $\mathbb{R}^n$:n kanta.

Jono $H = \{\vec{a}_1^\perp, \vec{a}_2^\perp, \dots, \vec{a}_{n-q}^\perp\}$ on vapaan jonon osajonona myös vapaa. Jäljellä on enää sen osoittaminen, että H virittää $M^\perp$:n.

Olkoon $\vec{u} \in M^\perp$ mikä tahansa ortogonaalikomplementin vektori. Myös vektori $\vec{u}$ voidaan jakaa kahden komponentin summaksi

$$\vec{u} = \vec{u}^M + \vec{u}^\perp = \vec{0} + \vec{u}$$

Jos siis vektori $\vec{u}$ kirjoitetaan kannassa G , saadaan

$$\vec{u} = \underbrace{\mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \dots + \mu_q \vec{v}_q}_{=\vec{u}^M = \vec{0}} + \underbrace{\mu_{q+1} \vec{a}_1^\perp + \mu_{q+2} \vec{a}_2^\perp + \dots + \mu_n \vec{a}_{n-q}^\perp}_{=\vec{u}^\perp = \vec{u}}$$

Koska jono $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ on vapaa, on ensimmäisten koordinaattien oltava nollia, $\mu_1 = \mu_2 = \dots = \mu_q = 0$. Siis

$$\vec{u} \in \text{span}(\vec{a}_1^\perp, \vec{a}_2^\perp, \dots, \vec{a}_{n-q}^\perp)$$

□

Kootaan joitakin aikasemmin todettuja tuloksia yhteen seuraavasti:

Lause 193 Tarkastellaan $\mathbb{R}^n$:n vektorisysteemiä $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ ja siihen liittyvää matriisia $\mathbf{V}$.

(1) Jos $q < n$, niin systeemi on vapaa, jos ja vain jos $\det(\mathbf{V}^T \mathbf{V}) \neq 0$.

(2) Jos $q = n$, niin systeemi on vapaa, jos ja vain jos $\det(\mathbf{V}) \neq 0$.

(3) Jos $q > n$, niin systeemi on sidottu.

Todistus. (1) Todistettiin lauseessa (I76 sivu 153). (2) Kun $q = n$, matriisi $\mathbf{V}$ on neliö-matriisi, ja $\det(\mathbf{V}^T \mathbf{V}) = (\det(\mathbf{V}))^2$. (3) Todettiin lauseessa (I88). $\square$

Lopuksi tarkastelemme tärkeitä kysymystä siitä, miten vektorijonon virittämälle aliavaruudelle saadaan kanta, jos virittäjäjono ei ole vapaa.

Olkoon lähtökohtanamme $\mathbb{R}^n$:n vektorijono

$$F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$$

ja uusi vektorijono, jossa alkutilanteessa on vain yksi vektori

$$G = \{\vec{v}_1\}.$$

Käytään alkuperäisen vektorijonon F vektorit läpi alkaen toisesta ja päättyen viimeiseen siten, että jokaisen vektorin $\vec{v}_k$ kohdalla tehdään seuraava konstruktio:

Jos tutkittavaa vektoria $\vec{v}_k$ ei voi lausua sen edellä alkuperäisessä vektorijonossa olevien vektorien lineaarikombinaationa ($\vec{v}_k \notin \text{span}(\vec{v}_1, \dots, \vec{v}_{k-1})$), niin lisätään vektori $\vec{v}_k$ uuteen jonoon G .

Kun kaikki alkuperäisen jonon vektorit on näin käyty läpi, on uudessa jonossa G alkuperäisen jonon osajono. Tästä uudesta vektorijonosta sanomme, että se on saatu karsimalla alkuperäinen jono vapaaksi.

Helpoin tapa tehdä karsiminen, on muodostaa vektorijonoa $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ vastaava matriisi $\mathbf{V}$, ja tehdä matriisille Gaussin algoritmi mukaiset rivioperaatiot. Aina, kun pivotointi jossakin sarakkeessa epäonnistuu, se johtuu siitä, että kyseinen sarake on lausuttavissa edellisten sarakkeiden lineaarikombinaationa. Teemme siis siten, että Vektorijonosta poistetaan kaikki vektorit, joita vastaavassa $\mathbf{V}$:n sarakkeessa pivotointi epäonnistui tai jäi tekemättä.

Otamme ensin esimerkin, jotta menettely tulee varmasti ymmärrettyä oikein.

Esimerkki 194 Karsimme vapaaksi vektorijonon

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 0 \\ 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -4 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ -1 \end{pmatrix} \right\}$$

Muodostamme asiaankuuluvan matriisin ja ryhdymme pivotoimaan

$$\begin{pmatrix} [1] & 2 & 1 & 1 & 0 \\ 2 & 5 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 5 & -4 & 2 \\ 0 & 2 & 4 & -2 & -1 \end{pmatrix} \begin{array}{l} -2 \\ \leftarrow \\ \\ \leftarrow \\ \leftarrow \end{array} \sim \begin{pmatrix} (1) & 2 & 1 & 1 & 0 \\ 0 & [1] & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 3 & 6 & -3 & 2 \\ 0 & 2 & 4 & -2 & -1 \end{pmatrix} \begin{array}{l} \\ -3 \\ \leftarrow \\ \leftarrow \\ \leftarrow \end{array} -2 \\ \sim \begin{pmatrix} (1) & 2 & 1 & 1 & 0 \\ 0 & (1) & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & [1] \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{array}{l} \\ \\ -2 \\ \leftarrow \\ \leftarrow \end{array} +1 \sim \begin{pmatrix} (1) & 2 & 1 & 1 & 0 \\ 0 & (1) & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & (1) \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Siis kolmas ja neljäs vektori karsitaan pois; jäljelle jäävät ensimmäinen, toinen ja viides vektori. Vapaaksi karsittu jono on

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ -1 \end{pmatrix} \right\}.$$

Lause 195 Olkoon

$$F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$$

$\mathbb{R}^n$:n vektorijono ja olkoon

$$G = \{\vec{v}_{k_1}, \vec{v}_{k_2}, \dots, \vec{v}_{k_r}\}$$

jonosta F vapaaksi karsittu jono. Silloin G on vapaa ja virittää saman aliavaruuden kuin F . Toisin sanoen G on aliavaruuden $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q)$ kanta.

Todistus. Jos G olisi sidottu, niin jokin G :n vektoreista voitaisiin lausua sitä edeltävien vektoreiden lineaarikombinaationa. Tämä on mahdotonta, koska tällaiset vektorit on

karsittu pois. Siis G on vapaa.

Olkoon toiseksi

$$\vec{a} = \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \cdots + \mu_q \vec{v}_q$$

mikä tahansa aliavaruuden M vektori. Käymme termit läpi alkaen viimeisestä. Jos $\vec{v}_k$ karsittiin jonosta pois, niin se voidaan korvata lineaarikombinaationa edeltävistä vektoreista. Sijoitetaan tämä lineaarikombinaatio vektorin $\vec{v}_k$ paikalle ja sievennetään lauseke. Jatketaan seuraavasta termistä. Näin jatkaen löytyy lineaarikombinaatio, jossa on vain karsitun jonon vektoreita. $\square$

Esimerkki 196 Edellisessä esimerkissä

$$\vec{v}_3 = -3\vec{v}_1 + 2\vec{v}_2, \text{ ja}$$

$$\vec{v}_4 = \vec{v}_2 - \vec{v}_3$$

Siis

$$\begin{aligned} & \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \mu_3 \vec{v}_3 + \mu_4 \vec{v}_4 + \mu_5 \vec{v}_5 \\ = & \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \mu_3 \vec{v}_3 + \mu_4 (\vec{v}_2 - \vec{v}_3) + \mu_5 \vec{v}_5 \\ = & \mu_1 \vec{v}_1 + (\mu_2 + \mu_4) \vec{v}_2 + (\mu_3 - \mu_4) \vec{v}_3 + \mu_5 \vec{v}_5 \\ = & \mu_1 \vec{v}_1 + (\mu_2 + \mu_4) \vec{v}_2 + (\mu_3 - \mu_4) (-3\vec{v}_1 + 2\vec{v}_2) + \mu_5 \vec{v}_5 \\ = & (\mu_1 - 3\mu_3 + 3\mu_4) \vec{v}_1 + (\mu_2 + 2\mu_3 - \mu_4) \vec{v}_2 + \mu_5 \vec{v}_5 \in \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_5\} \end{aligned}$$

8.3 Matriisin säännöllisyysaste

Sec:MatRank

Tässä kappaleessa tarkastelemme $m \times n$ -matriisia $\mathbf{A}$. Merkitään matriisin k :ttä saraketta vektorina $\vec{a}_k$. Kappaleessa ^{lydin}(??) määrittelimme matriisin sarakeavaruuden

$$\text{coll}(\mathbf{A}) = \text{span}\{\vec{a}_1, \dots, \vec{a}_n\},$$

joka on siis matriisin sarakkeiden virittämä $\mathbb{R}^m$:n aliavaruus, ja ytimen

$$\text{null}(\mathbf{A}) = \{\vec{x} \in \mathbb{R}^n \mid \mathbf{A}\vec{x} = \vec{0}\},$$

joka on $\mathbb{R}^n$:n aliavaruus.

Määritelmä 197 $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste, eli rangi, $\text{rank}(\mathbf{A})$, on matriisin sarakeavaruuden dimensio

$$\text{rank}(\mathbf{A}) = \dim(\text{coll}(\mathbf{A}))$$

ja matriisin nulliteetti on sen ytimen dimensio

$$\dim(\text{null}(\mathbf{A})).$$

Nulliteetille emme ota käyttöön mitään erityistä merkintää. Kun käsitettä tarvitaan, se on luonnollista merkitä ytimen avulla. Rangia käytetään paljon, sillä aiemmasta tiedämme, että

$$\begin{aligned} n \times n - \text{neliömatriisi } \mathbf{A} \text{ on säännöllinen} \\ \Leftrightarrow \text{rank}(\mathbf{A}) = n \end{aligned}$$

Tässä tapauksessa englanninkielinen kirjallisuus ei yleensä sano, että matriisi olisi ”regular”, vaan se sanoo, että matriisi on ”a matrix of full rank”.

Aiemmasta tiedämme, että rangi ja nulliteetti saadaan määritettyä tekemällä matriisille Gaussin algoritmin mukaiset rivioperaatiot. Onnistuneiden pivotointien lukumäärä on rangi ja epäonnistuneiden tai tekemättä jääneiden pivotointien lukumäärä on nulliteetti (vrt. esimerkki (I68) sivulla I48). Siis

Lause 198 $m \times n$ -matriisille $\mathbf{A}$ on voimassa

$$\text{rank}(\mathbf{A}) + \dim(\text{null}(\mathbf{A})) = n$$

Todistus. Jos ydin on triviaali, niin lauseen (I76) mukaan matriisin sarakkeet muodostavat vapaan systeemin ja lause on tosi.

Jos matriisin ydin ei ole triviaali, niin

$$\begin{aligned} \text{col}(\mathbf{A}^T)^\perp &= \{\vec{y} \in \mathbb{R}^m \mid \langle \vec{y} | \vec{z} \rangle = 0, \forall \vec{z} \in \text{col}(\mathbf{A}^T)\} \\ &= \{\vec{y} \in \mathbb{R}^m \mid \langle \vec{y} | \mathbf{A}^T \vec{x} \rangle = 0, \forall \vec{x} \in \mathbb{R}^n\} \\ &= \{\vec{y} \in \mathbb{R}^m \mid \langle \mathbf{A} \vec{y} | \vec{x} \rangle = 0, \forall \vec{x} \in \mathbb{R}^n\} \\ &= \text{null}(\mathbf{A}) \end{aligned} \tag{8.9}$$

Kaava tarkoittaa nyt sitä, että ytimen kanta voidaan aina konstruoida. Transpoosin sarakkeavaruuden kanta saadaan karsimalla sarakkeista ja ortogonaalikomplementin (eli ytimen) kanta saadaan lauseen (I92) konstruktiolla.

Olkoon nyt vektorijono $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_q\}$ ytimen kanta. Täydennetään tämä jono $\mathbb{R}^n$:n kannaksi $\{\vec{b}_1, \dots, \vec{b}_q, \vec{c}_1, \dots, \vec{c}_{n-q}\}$. Käytetään tässäkin lauseen (I92) konstruktiota.

Seuraavaksi osoitamme, että vektorijono $\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\}$ on aliavaruuden $\text{col}(\mathbf{A})$ kanta. Jos näin on, niin $\dim(\text{null}(\mathbf{A})) = q$ ja $\text{rank}(\mathbf{A}) = \dim(\text{col}(\mathbf{A})) = n - q$, joten

lauseen väite on tosi.

Ensin osoitamme, että $\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\}$ virittää aliavaruuden $\text{col}(\mathbf{A})$. Olkoon sitä varten $\vec{y} \in \mathbb{R}^m$ mikä tahansa $\text{col}(\mathbf{A})$:n vektori. Silloin on olemassa $\vec{x} \in \mathbb{R}^n$ siten, että

$$\vec{y} = x_1 \vec{a}_1 + \dots + x_n \vec{a}_n = \mathbf{A}\vec{x}.$$

Vektorilla $\vec{x}$ on yksikäsitteinen esitys $\mathbb{R}^n$:n kannassa $\{\vec{b}_1, \dots, \vec{b}_q, \vec{c}_1, \dots, \vec{c}_{n-q}\}$

$$\vec{x} = \mu_1 \vec{b}_1 + \dots + \mu_q \vec{b}_q + \mu_{q+1} \vec{c}_1 + \dots + \mu_n \vec{c}_{n-q}.$$

Silloin koska $\mathbf{A}\vec{b}_k = \vec{0}, \forall k$, saamme

$$\begin{aligned} \vec{y} &= \mathbf{A}(\mu_1 \vec{b}_1 + \dots + \mu_q \vec{b}_q + \mu_{q+1} \vec{c}_1 + \dots + \mu_n \vec{c}_{n-q}) \\ &= \mu_{q+1} \mathbf{A}\vec{c}_1 + \dots + \mu_n \mathbf{A}\vec{c}_{n-q} \\ &\in \text{span}\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\} \end{aligned}$$

Seuraavaksi osoitamme, että $\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\}$ on vapaa.

$$\begin{aligned} \lambda_1 \mathbf{A}\vec{c}_1 + \lambda_2 \mathbf{A}\vec{c}_2 + \dots + \lambda_{n-q} \mathbf{A}\vec{c}_{n-q} &= \vec{0} \\ \Leftrightarrow \mathbf{A}(\lambda_1 \vec{c}_1 + \lambda_2 \vec{c}_2 + \dots + \lambda_{n-q} \vec{c}_{n-q}) &= \vec{0} \\ \Leftrightarrow \lambda_1 \vec{c}_1 + \lambda_2 \vec{c}_2 + \dots + \lambda_{n-q} \vec{c}_{n-q} &\in \text{null}(\mathbf{A}) \end{aligned}$$

Nyt löytyy sellaiset kertoimet $\mu_1, \dots, \mu_q$, että

$$\begin{aligned} \lambda_1 \vec{c}_1 + \lambda_2 \vec{c}_2 + \dots + \lambda_{n-q} \vec{c}_{n-q} - \mu_1 \vec{b}_1 - \dots - \mu_q \vec{b}_q &= \vec{0} \\ \Leftrightarrow \lambda_1 = \lambda_2 = \dots = \lambda_{n-q} = \mu_1 = \dots = \mu_q &= 0 \end{aligned}$$

(Viimeinen ekvivalenssi johtuu siitä, että jono $\{\vec{b}_1, \dots, \vec{b}_q, \vec{c}_1, \dots, \vec{c}_{n-q}\}$ on kantana vapaa.) $\square$

Lause 199 $m \times n$ -matriisille $\mathbf{A}$ on voimassa

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T)$$

Todistus. Kaavan [\(8.9\)](#) mukaan $\text{col}(\mathbf{A}^T)^\perp = \text{null}(\mathbf{A})$, joten lauseen [\(192\)](#) mukaan

$$\text{rank}(\mathbf{A}^T) + \dim(\text{null}(\mathbf{A})) = n$$

Edellisen lauseen mukaan

$$\text{rank}(\mathbf{A}) + \dim(\text{null}(\mathbf{A})) = n.$$

Ratkaisemalla rangit kummastakin yhtälöstä saadaan

$$\text{rank}(\mathbf{A}) = n - \dim(\text{null}(\mathbf{A})) = \text{rank}(\mathbf{A}^T)$$

□

Siis $m \times n$ -matriisin säännöllisyysaste on korkeintaan pienempi luvuista m ja n . Rangin laskeminen ei ole aina ongelmatonta. Ongelmat ovat numeerisia, eikä niihin puututa tässä kappaleessa. Singulaariarvo-hajoituksen avulla saamme numeerisesti hyvin käyttökelpoisen tavan laskea säännöllisyysaste. Toistaiseksi oletamme, että matriisin säännöllisyysaste tunnetaan tarvittaessa tarkasti.

8.4 Pseudoinverssi ja ortogonaaliprojektio

Sec:MatPInv

Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n ja $m > n$. ($\text{rank}(\mathbf{A}) = n$) Pohdimme (aina vain) yhtälöryhmän

$$\mathbf{A}\vec{x} = \vec{b}$$

ratkaisua. Ratkaisu on olemassa, jos

$$\begin{aligned} \vec{b} &\in \text{col}(\mathbf{A}) \\ \Leftrightarrow \quad &\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n, \vec{b}\} \text{ on sidottu} \\ \Leftrightarrow \quad &\det((\mathbf{A} | \vec{b})^T (\mathbf{A} | \vec{b})) = 0 \end{aligned}$$

Ratkaisun olemassaolo on siis helppo tarkistaa.

Jos kerroinmatriisin QR-hajotelma on tiedossa $\mathbf{A} = \mathbf{Q}\mathbf{R}$, niin yhtälöryhmällä $\mathbf{A}\vec{x} = \vec{b}$ on ratkaisu, jos ja vain jos yhtälöryhmällä

$$\mathbf{R}\vec{x} = \mathbf{Q}^T \vec{b} = \vec{c} \tag{8.10}$$

on ratkaisu. Jos RHS:n $m - n$ viimeistä koordinaattia ovat nolliä, $c_{n+1} = \dots = c_m = 0$, niin yhtälöryhmällä on yksikäsitteinen ratkaisu, joka saadaan yksinkertaisten palautuvien sijoitusten avulla.

Jos $\mathbf{A}$ olisi neliömatriisi, niin saisimme ratkaisun laskettua käänteismatriisin avulla $\vec{x} = \mathbf{A}^{-1}\vec{b}$. Nyt kuitenkin käänteismatriisia ei ole olemassa.

Lause 200 Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n ja $m > n$. Jos $\vec{b} \in \text{col}(\mathbf{A})$, niin yhtälöryhmän

$$\mathbf{A}\vec{x} = \vec{b}$$

ratkaisu saadaan kertomalla RSH kerroinmatriisin pseudoinverssillä $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$. Siis

$$\vec{x} = \mathbf{A}^\dagger \vec{b}.$$

Todistus. Oletuksen mukaan on olemassa $\vec{x}$ siten, että $\mathbf{A}\vec{x} = \vec{b}$. Nyt

$$\mathbf{A}^\dagger \vec{b} = [(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T] \mathbf{A} \vec{x} = \vec{x}$$

□

Pseudoinverssi QR-hajotelman avulla.

Jos $m \times n$ -matriisin $\mathbf{A}$, jolle $\text{rank}(\mathbf{A}) = n$, QR-hajotelma on

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

niin

$$\begin{aligned} \mathbf{A}^\dagger &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \\ &= ((\mathbf{Q}\mathbf{R})^T \mathbf{Q}\mathbf{R})^{-1} (\mathbf{Q}\mathbf{R})^T \\ &= ((\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{Q}^T \\ &= \mathbf{R}^\dagger \mathbf{Q}^T \end{aligned}$$

Pseudoinverssi singulaariarvohajotelman avulla.

Jos $m \times n$ -matriisin $\mathbf{A}$, jolle $\text{rank}(\mathbf{A}) = n$, singulaariarvohajotelma on

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T,$$

niin

$$\begin{aligned} \mathbf{A}^\dagger &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \\ &= \left[\left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right]^{-1} \left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \\ &= \left[\mathbf{V} \begin{pmatrix} \mathbf{\Sigma}^T & \mathbf{0} \end{pmatrix} \mathbf{U}^T \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right]^{-1} \left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \\ &= [\mathbf{V} \mathbf{\Sigma}^2 \mathbf{V}^T]^{-1} \left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \\ &= [\mathbf{V} \mathbf{\Sigma}^{-2} \mathbf{V}^T] (\mathbf{V} \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \end{pmatrix} \mathbf{U}^T) \\ &= \mathbf{V} \begin{pmatrix} \mathbf{\Sigma}^{-1} & \mathbf{0} \end{pmatrix} \mathbf{U}^T \end{aligned}$$

8.4.1 Ortogonaaliprojektio

Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n . Oletamme, että vektori $\vec{y} = (y_1, y_2, \dots, y_n)^T$ ei kuulu $\mathbf{A}$:n sarakeavaruuteen

$$\vec{y} \notin \text{col}(\mathbf{A}) = \text{span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}.$$

Asetamme tehtäväksi nyt etsiä se sarakeavaruuden vektori $\vec{x} \in \text{col}(\mathbf{A})$, jolla

$$\|\vec{x} - \vec{y}\|_2$$

saa pienimmän mahdollisen arvon. Muutaman välivaiheen kautta osoitamme, että tehtävän ratkaisu on

$$\vec{x} = \mathbf{A}\mathbf{A}^\dagger \vec{y}.$$

Lause 201 Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n . Jos $\vec{x} = \mathbf{A}\mathbf{A}^\dagger \vec{y}$, niin

$$(\vec{x} - \vec{y}) \perp \text{col}(\mathbf{A}).$$

Todistus.

$$\mathbf{A}^T(\vec{x} - \vec{y}) = \mathbf{A}^T \mathbf{A} \mathbf{A}^\dagger \vec{y} - \mathbf{A}^T \vec{y} = 0 = \mathbf{A}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \vec{y} - \mathbf{A}^T \vec{y} = 0$$

□

c4s21thortproj

Lause 202 Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n . Jos $\vec{x} = \mathbf{A}\mathbf{A}^\dagger \vec{y}$, niin $\vec{x}$ on lähinnä vektoria $\vec{y}$ oleva sarakeavaruuden $\text{col}(\mathbf{A})$ vektori. Siis

$$\|\vec{x} - \vec{y}\|_2 = \min_{\vec{w} \in \text{col}(\mathbf{A})} \|\vec{w} - \vec{y}\|_2.$$

Todistus. Olkoon $\vec{w} \in \text{col}(\mathbf{A})$ mikä tahansa sarakeavaruuden vektori. Silloin

$$\vec{w} - \vec{y} = (\vec{w} - \vec{x}) + (\vec{x} - \vec{y})$$

ja $(\vec{w} - \vec{x}) \perp (\vec{x} - \vec{y})$. Silloin Pythagoran lauseen nojalla

$$\begin{aligned} \|\vec{w} - \vec{y}\|_2^2 &= \|\vec{w} - \vec{x}\|_2^2 + \|\vec{x} - \vec{y}\|_2^2 \\ \Rightarrow \|\vec{w} - \vec{y}\|_2 &\geq \|\vec{x} - \vec{y}\|_2 \end{aligned}$$

□

Sec:BaseOrt

8.5 Ortonormeerattu kanta, Gram-Schmidt

Tässä vaiheessa selvä perusongelma on vektorin koordinaattien laskeminen annetussa kannassa. Olkoon $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ lineaariavaruuden $L \subseteq \mathbb{R}^m$ kanta ja $\mathbf{V}$ $m \times n$ -matriisi, jonka sarakkeina ovat kantavektorit. Oletamme seuraavassa, että vektori $\vec{a}$ kuuluu kannan virittämään aliavaruuteen. Vektorin $\vec{a}$ koordinaatit $\alpha_1, \dots, \alpha_m$ ja niistä muodostettu koordinaattivektori $\vec{\alpha} = (\alpha_1 \dots \alpha_m)^T$ toteuttavat yhtälöt

$$\begin{aligned} \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_m \vec{v}_m &= \vec{a} \\ \Leftrightarrow \mathbf{V} \vec{\alpha} &= \vec{a}. \end{aligned} \quad (8.11)$$

Koska kanta on vapaa, on $\text{rank}(\mathbf{V}) = n$, ja koordinaattivektori voidaan kätevästi ratkaista kaavalla

$$\vec{\alpha} = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a}. \quad (8.12) \quad \text{kooreq}$$

Huomaa, että tämä kaava antaa väärän vastauksen, jos $\vec{a} \notin \text{col}(\mathbf{V})$. Silloin ei nimittäin koordinaatteja ole olemassa, mutta kaavasta sellaiset saadaan laskettua. Ne ovat $\text{coll}(\mathbf{V})$:n lähimpänä $\vec{a}$:ta olevan vektorin koordinaatit.

Tarkastellaan ongelmaa hieman eri näkökulmasta. Lasketaan vektorin $\vec{a}$ sisätulot kantavektoreiden kanssa.

$$\begin{aligned} \langle \vec{a} | \vec{v}_i \rangle &= \langle \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_m \vec{v}_m | \vec{v}_i \rangle \\ &= \alpha_1 \langle \vec{v}_1 | \vec{v}_i \rangle + \alpha_2 \langle \vec{v}_2 | \vec{v}_i \rangle + \dots + \alpha_m \langle \vec{v}_m | \vec{v}_i \rangle \end{aligned}$$

Jos nyt $\langle \vec{v}_i | \vec{v}_j \rangle = \delta_{ij}$ eli

$$\vec{v}_i \perp \vec{v}_j, \quad \forall i \neq j \quad \text{ja} \quad (8.13)$$

$$\|\vec{v}_i\| = 1, \quad \forall i, \quad (8.14)$$

niin

$$\alpha_i = \langle \vec{a} | \vec{v}_i \rangle$$

eli

$$\vec{a} = \langle \vec{a} | \vec{v}_1 \rangle \vec{v}_1 + \langle \vec{a} | \vec{v}_2 \rangle \vec{v}_2 + \dots + \langle \vec{a} | \vec{v}_m \rangle \vec{v}_m \quad (8.15) \quad \text{kooreq2}$$

Esitys (8.15) on yksinkertainen ja helppo muistaa. Esitys on käytettävissä vain jos käytetty kanta toteuttaa ehdot (8.13, 8.14). Annamme ehdot toteuttaville kannoille nimen, jotta voimme jatkossa puhua niistä sujuvasti

Määritelmä 203 Lineaariavaruuden L vektorisysteemi $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ on ortonormaalinen, jos systeemin vektorit ovat pareittain kohtisuorassa keskenään,

$$i \neq j \Rightarrow \mathbf{v}_i \perp \mathbf{v}_j.$$

Jos lisäksi

$$\|\mathbf{v}_i\| = \sqrt{\langle \mathbf{v}_i | \mathbf{v}_i \rangle} = 1, \quad \forall i = 1, \dots, n,$$

niin systeemi on ortonormaali (ortonormeerattu).

Laksut usein yksinkertaistuvat, kun käytetty kanta on ortonormitettu. Koulukurssista tutut vektorit $\vec{i} = (1, 0, 0)^T$, $\vec{j} = (0, 1, 0)^T$ ja $\vec{k} = (0, 0, 1)^T$ muodostavat $\mathbb{R}^3$:n kannan $F = \{\vec{i}, \vec{j}, \vec{k}\}$. Toisaalta myös jono $G = \{2\vec{i}, 2\vec{j}, 2\vec{k}\}$ on $\mathbb{R}^3$:n kanta. Jos vektorin $\vec{a}$ koordinaatit kannassa F ovat $(1 \ 1 \ 1)^T$, niin sen koordinaatit kannassa G ovat $(0.5 \ 0.5 \ 0.5)^T$. Nyt on selvää, ettei vektorin pituutta voi laskea suoraan koordinaateista ilman jonkinlaista tietoa koordinaatteihin liittyvästä kannasta. Ortonormaalisuus riittää.

Lause 204 (Pythagoraan lauseen yleistys) Olkoon L reaalinen lineaariavaruus, jossa on määritelty sisätulo, ja olkoon vektorijono $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ sen ortonormaali kanta. Jos

$$\begin{aligned} \mathbf{a} &= a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_q \mathbf{v}_q, \quad \text{ja} \\ \mathbf{b} &= b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \dots + b_q \mathbf{v}_q, \end{aligned}$$

niin

$$\begin{aligned} \langle \mathbf{a} | \mathbf{b} \rangle &= a_1 b_1 + a_2 b_2 + \dots + a_q b_q \\ \|\mathbf{a}\|^2 &= \langle \mathbf{a} | \mathbf{a} \rangle = a_1^2 + a_2^2 + \dots + a_q^2. \end{aligned}$$

Todistus. Suora lasku. □

Siis vektorin pituus ja kahden vektorin kohtisuoruus voidaan tutkia suoraan koordinaattien avulla, JOS KANTA ON ORTONORMITETTU.

Seuraavaksi kuvailemme Gram-Schmidt -algoritmin, jonka avulla saamme aina muutettua $\mathbb{R}^n$:n aliavaruuden kannan ortonormeeratuksi. Oletamme, että vektorijono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ on vapaa eli aliavaruuden $M = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ kanta. Muodostamme M :lle uuden kannan $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_q\}$ seuraavasti.

Gram-Schmidt:

(1) Ensimmäiseksi asetamme

$$\vec{b}_1 = \frac{1}{\|\vec{v}_1\|} \vec{v}_1. \quad (8.16) \quad \boxed{\text{GrSm1}}$$

(2) Toiseksi asetamme

$$\vec{a}_2 = \vec{v}_2 - \langle \vec{v}_2 | \vec{b}_1 \rangle \vec{b}_1 \quad (8.17)$$

$$\vec{b}_2 = \frac{1}{\|\vec{a}_2\|} \vec{a}_2. \quad (8.18)$$

(3) Kolmanneksi asetamme

$$\vec{a}_3 = \vec{v}_3 - \langle \vec{v}_3 | \vec{b}_1 \rangle \vec{b}_1 - \langle \vec{v}_3 | \vec{b}_2 \rangle \vec{b}_2 \quad (8.19)$$

$$\vec{b}_3 = \frac{1}{\|\vec{a}_3\|} \vec{a}_3. \quad (8.20)$$

(4...q) k:s kantavektori saadaan asettamalla

$$\vec{a}_k = \vec{v}_k - \sum_{j=1}^{k-1} \langle \vec{v}_k | \vec{b}_j \rangle \vec{b}_j \quad (8.21)$$

$$\vec{b}_k = \frac{1}{\|\vec{a}_k\|} \vec{a}_k. \quad (8.22)$$

Lause 205 Olkoon $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ vapaa jono $\mathbb{R}^n$:n vektoreita, eli virittämässään aliavaruuden $M = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ kanta. Gram-Schmidt -menetelmällä (8.16-8.22) muodostettu vektorijono $G = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_q\}$ on M :n ortonormaali kanta. (GrSm1)

Todistus. Vektorit $\vec{v}_1, \dots, \vec{v}_q$ ovat selvästi yksikkövektoreita, kunhan mikään $\vec{a}_k$ ei ole nollavektori. Riittää siis osoittaa, että jono $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_q\}$ on vapaa ja ortogonaalinen.

Tehdään induktiotarkastelu:

Induktion lähtökohta: jono $\{\vec{a}_1\}$ on vapaa ja ortogonaalinen.

Induktio-oletus: jono $\{\vec{a}_1, \dots, \vec{a}_{k-1}\}$ on vapaa ja ortogonaalinen.

Induktioväite: jono $\{\vec{a}_1, \dots, \vec{a}_k\}$ on vapaa ja ortogonaalinen.

Induktiotodistus: Jos pystymme osoittamaan, että $\vec{a}_k \perp \vec{a}_s$ kaikilla $s = 1, \dots, k-1$, niin jono $\{\vec{a}_1, \dots, \vec{a}_k\}$ on ortogonaalinen ja lauseen (I91) mukaan vapaa. (c4s8th7b)

Yhtälöt ^{|GrSm6}(8.21) ja ^{|GrSm7}(8.22) voidaan tiivistää yhdeksi yhtälöksi

$$\vec{a}_k = \vec{v}_k - \sum_{j=1}^{k-1} \rho_{kj} \vec{a}_j, \quad \rho_{kj} = \frac{\langle \vec{v}_k | \vec{a}_j \rangle}{\langle \vec{a}_j | \vec{a}_j \rangle}$$

Induktio-oletuksen mukaan kaikilla $s < k$ on

$$\begin{aligned} \langle \vec{a}_k | \vec{a}_s \rangle &= \left\langle \vec{v}_k - \sum_{j=1}^{k-1} \rho_{kj} \vec{a}_j \middle| \vec{a}_s \right\rangle \\ &= \langle \vec{v}_k | \vec{a}_s \rangle - \sum_{j=1}^{k-1} \rho_{kj} \underbrace{\langle \vec{a}_j | \vec{a}_s \rangle}_{=0, \text{ kun } j \neq s} \\ &= \langle \vec{v}_k | \vec{a}_s \rangle - \rho_{ks} \langle \vec{a}_s | \vec{a}_s \rangle \\ &= 0 \end{aligned}$$

□

Gram-Schmidt -menettely on käsin laskettaessa raskas. Se onkin enemmän teoreettinen asia, kuin käytännöllinen työkalu.

Lause 206 Jos $\mathbb{R}^n$:n aliavaruudessa M on nollavektorin lisäksi muitakin vektoreita, niin aliavaruudella on ortonormeerattu kanta.

Todistus. Olkoon $\vec{a} \in \mathbb{R}^n$ ja $\vec{a} \neq \vec{0}$. Silloin jono $\{\vec{a}\}$ on vapaa jono, jonka voimme täydentää $\mathbb{R}^n$:n kannaksi. Tämä kanta voidaan edelleen ortonormeerata Gram-Schmidt -menetelmällä. □

Esimerkki 207 Tarkastellaan $\mathbb{R}^5$:n vektoreita $\vec{v}_1 = (1, 0, 2, -1, 3)^T$, $\vec{v}_2 = (3, 1, -1, 2, 0)^T$ ja $\vec{v}_3 = (0, 1, 3, 5, -1)^T$. Muodostetaan näiden vektoreiden virittämän aliavaruuden ortonormeerattu kanta Matlabilla seuraavasti.

```
>> v1 = [1 0 2 -1 3]';
>> v2 = [3 1 -1 2 0]';
>> v3 = [0 1 3 5 -1]';
```

```
>> b1 = v1/norm(v1)
b1 =
    0.2582
         0
```

```
0.5164
-0.2582
0.7746

>> a2 = v2 - (v2'*b1)*b1;
>> b2 = a2/norm(a2)
b2 =
    0.7936
    0.2588
   -0.2243
    0.5003
    0.0518

>> a3 = v3 - (v3'*b1)*b1 - (v3'*b2)*b2;
>> b3 = a3/norm(a3)
b3 =
   -0.2637
    0.0842
    0.6624
    0.6847
   -0.1255

>> %Tarkistuksia
>> b1'*b2
ans =
   -2.7756e-017

>> b1'*b3
ans =
   -8.3267e-017

>> b2'*b3
ans =
    1.1102e-016

>> norm(b1)
ans =
    1

>> norm(b2)
ans =
```

1

```
>> norm(b3)
ans =
     1
>>
```

8.6 Ortogonaaliset matriisit, QR-hajoitelma

Sec:MatOrt

Määritelmä 208 Neliömatriisi $\mathbf{A}$ on ortogonaalinen, jos

$$\mathbf{A}^{-1} = \mathbf{A}^T.$$

Määritelmä on hämäävän yksinkertainen. Seuraava lause kertoo paremmin mistä oikein on kysymys.

Lause 209 Olkoon $\mathbf{A}$ neliömatriisi. Silloin seuraavat kolme ominaisuutta ovat yhtäpitävät:

- (1) $\mathbf{A}$ on ortogonaalinen.
- (2) $\mathbf{A}$:n sarakkeet muodostavat ortonormeeratun systeemin.
- (3) $\mathbf{A}$:n rivit muodostavat ortonormeeratun systeemin.

Todistus. Ensinnä osoitetaan, että (1):stä seuraa (2) ja (3).

$$\begin{aligned} (1) \quad &\Leftrightarrow \mathbf{A}^{-1} = \mathbf{A}^T \\ &\Leftrightarrow \begin{cases} \mathbf{A}^T \mathbf{A} = \mathbf{I} & \Rightarrow (2) \\ \mathbf{A} \mathbf{A}^T = \mathbf{I} & \Rightarrow (3) \end{cases} \end{aligned}$$

Toiseksi osoitamme, että (2):sta seuraa (1). Oletamme siis, että $\mathbf{A}$ on neliömatriisi, jonka sarakkeet muodostavat ortonormeeratun systeemin. Silloin $\mathbf{A}^T \mathbf{A} = \mathbf{I}$, joten $\det(\mathbf{A}) \neq 0$. Siis $\mathbf{A}$ on säännöllinen eli sillä on käänteismatriisi. Nyt

$$\mathbf{A} \mathbf{A}^T = \underbrace{\mathbf{A} \mathbf{A}^T \mathbf{A}}_{=\mathbf{I}} \mathbf{A}^{-1} = \mathbf{I}$$

Kolmanneksi osoitamme, että (3):sta seuraa (1)

$$\begin{aligned}
 (3) \mathbf{A}: \text{lle} &\Leftrightarrow (2) \mathbf{A}^T: \text{lle} \\
 &\Rightarrow (1) \mathbf{A}^T: \text{lle} \\
 &\Leftrightarrow (\mathbf{A}^T)^{-1} = (\mathbf{A}^T)^T \\
 &\Leftrightarrow (\mathbf{A}^{-1})^T = \mathbf{A} \\
 &\Leftrightarrow \mathbf{A}^{-1} = \mathbf{A}^T \\
 &\Leftrightarrow (1) \mathbf{A}: \text{lle}
 \end{aligned}$$

□

Lause 210 Olkoon $n \times n$ -matriisi $\mathbf{A}$ ortogonaalinen ja $\vec{b}, \vec{c} \in \mathbb{R}^n$.

(1) Jos $\vec{b} \perp \vec{c}$, niin $\mathbf{A}\vec{b} \perp \mathbf{A}\vec{c}$.

(2) $\|\mathbf{A}\vec{b}\|_2 = \|\vec{b}\|_2$

(3) Jos $\mathbb{R}^n$:n vektorijono $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ on $\mathbb{R}^n$:n ortonormitettu kanta, niin vektorijono $\{\mathbf{A}\vec{u}_1, \mathbf{A}\vec{u}_2, \dots, \mathbf{A}\vec{u}_n\}$ on myös ortonormitettu kanta.

(Verbaalisti: Ortonormaalilla matriisilla kertominen säilyttää kohtisuoruuden ja 2-normin.)

Todistus. HT.

□

Lause 211 $n \times n$ -matriisi $\mathbf{A}$, jonka sarakkeet ovat lineaarisesti riippumattomat (siis $\text{rank}(\mathbf{A}) = n$), voidaan aina kirjoittaa muotoon

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

missä $\mathbf{Q}$ on ortogonaalinen $n \times n$ -matriisi ja $\mathbf{R}$ on $n \times n$ -yläkolmiomatriisi ($r_{ij} = 0$, kun $i > j$), jonka lävistäjäalkiot ovat nollasta eroavat.

Todistus. Matriisin $\mathbf{A}$ sarakevektorit $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ muodostavat vapaan jonon, joten voimme ortonormeerata sen Gram-Schmidt -menetelmällä jonoksi $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$. Kirjoitamme uudelleen Gram-Schmidt -menetelmän yhtälöt (8.21) ja (8.22) tähän tilanteeseen kuuluvien merkinnöin.

$$\vec{c}_k = \vec{a}_k - \sum_{j=1}^{k-1} \langle \vec{a}_k | \vec{q}_j \rangle \vec{q}_j \quad (8.23)$$

$$\vec{q}_k = \frac{1}{\|\vec{c}_k\|} \vec{c}_k. \quad (8.24)$$

Yhtälöt voidaan yhdistää seuraavasti

$$\vec{a}_k = \sum_{j=1}^k r_{jk} \vec{q}_j \quad (8.25)$$

ja $r_{kk} \neq 0, \forall k$. Voimme lisäksi asettaa $r_{jk} = 0$, kun $j > k$. Edellinen yhtälö pitää kirjoittaa jokaiselle k :n arvolla. Siis n yhtälöä. Nämä kaikki yhtälöt saadaan koottua siistiin pakettiin kirjoittamalla seuraava matriisiyhtälö

$$\begin{pmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{pmatrix} = \begin{pmatrix} \vec{q}_1 & \dots & \vec{q}_n \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix} \quad (8.26)$$

$$\Leftrightarrow \mathbf{A} = \mathbf{QR}$$

□

Todistus on valmis, mutta koska on olemassa vaara, että yhtälö (8.26) jää ymmärtämättä, kirjoitamme sen auki poikkeuksellisen perusteellisesti

$$\begin{pmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{pmatrix} = \begin{pmatrix} \vec{q}_1 & \dots & \vec{q}_n \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \dots & q_{nn} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix}$$

Laskemalla matriisikertolasku RHS:ssa saadaan

$$\vec{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = \begin{pmatrix} r_{11}q_{11} \\ r_{11}q_{21} \\ \vdots \\ r_{11}q_{n1} \end{pmatrix} = r_{11}\vec{q}_1$$

$$\vec{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} = \begin{pmatrix} r_{12}q_{11} + r_{22}q_{12} \\ r_{12}q_{21} + r_{22}q_{22} \\ \vdots \\ r_{12}q_{n1} + r_{22}q_{n2} \end{pmatrix} = r_{12}\vec{q}_1 + r_{22}\vec{q}_2$$

ja niin edelleen yhtälön (8.25) mukaisesti.

QR-hajoitelman olemassaolo perusteltiin edellä Gram-Schmidt algoritmin avulla. QR-hajoitelmaa ei käytännössä kannata laskea Gram-Schmidt -algoritmillä, vaan siihen tarkoitukseen on olemassa numeerisesti luotettavampia menetelmiä, joita emme nyt käsittele. (Ns. Housholderin transformaatiot. ks Kivelä: 'Matriisilasku ja Lineaarialgebra' ss. 123–127.)

Esimerkki 212 Jos käytettävissä on Matlab-ohjelma, niin voimme pyytää QR-hajoitelman yksinkertaisella komennolla. Matriisiin

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

QR-hajoitelma saadaan seuraavasti:

```
>> A = [1 2 0 1 ; 3 2 1 0 ; 0 1 2 1 ; 4 -1 3 1]
```

```
A =
```

```

1      2      0      1
3      2      1      0
0      1      2      1
4     -1      3      1
```

```
>> [Q,R]=qr(A)
```

```
Q =
```

```

-0.1961    -0.6026     0.1407    -0.7606
-0.5883    -0.5022     0.2272     0.5916
         0    -0.3264    -0.9414     0.0845
-0.7845     0.5273    -0.2056    -0.2535
```

```
R =
```

```

-5.0990    -0.7845    -2.9417    -0.9806
         0    -3.0634     0.4269    -0.4018
         0         0    -2.2724    -1.0064
         0         0         0    -0.9297
```

```
>>
```

QR-hajoitelma voidaan muodostaa myös $m \times n$ -matriisille, jos $m > n$ ja $\text{rank}(\mathbf{A}) = n$. Sama konstruktio kuin edellä, johtaa $m \times n$ matriisiin $\mathbf{Q}_1$ ja $n \times n$ -yläkolmiomatriisiin

$\mathbf{R}_1$ siten, että matriisin $\mathbf{Q}_1$ sarakkeet muodostavat ortonormeeratun systeemin, $\mathbf{R}_1$:n diagonaali-alkiot eroavat nolasta ja

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R}_1.$$

Koska matriisin $\mathbf{Q}_1$ sarakevektorit $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$ muodostavat vapaan jonon, se voidaan täydentää $\mathbb{R}^m$:n kannaksi ja sen jälkeen ortonormeerata. Näin saadaan uusi vektorijono $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n, \vec{w}_1, \dots, \vec{w}_{m-n}\}$, jossa ensimmäiset vektorit ovat $\mathbf{Q}_1$:n sarakkeet ja loput vektorit ovat uusia. Näistä koottu matriisi

$$\mathbf{Q} = (\mathbf{Q}_1 | \mathbf{W})$$

on ortogonaalinen $m \times m$ -matriisi. Täydennetään edellä saatua $\mathbf{R}_1$ matriisia lisäämällä siihen $m - n$ nollariviä alimmiksi riveiksi. Saatua matriisi

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{pmatrix}$$

on $m \times n$ matriisi, jossa diagonaalien alapuolella on vain nollia. Nyt

$$\mathbf{QR} = (\mathbf{Q}_1 | \mathbf{W}) \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{pmatrix} = \mathbf{Q}_1 \mathbf{R}_1 = \mathbf{A}$$

Esimerkki 213 Matriisin

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

QR-hajoitelma saadaan seuraavasti:

```
>> A = [1 2 ; 3 2 ; 0 1 ; 4 -1]
```

```
A =
```

```
1      2
3      2
0      1
4     -1
```

```
>> [Q,R] = qr(A)
```

```
Q =
```

```
-0.1961    -0.6026    -0.0039    -0.7735
-0.5883    -0.5022    -0.3283     0.5421
0          -0.3264     0.9116     0.2498
-0.7845     0.5273     0.2472    -0.2132
```

```

R =
    -5.0990    -0.7845
         0    -3.0634
         0         0
         0         0
>>

```

Matlabin komento on myös mahdollista antaa muodossa, jossa lisäykset jää pois

```

>> [Q,R] = qr(A,0)
Q =
    -0.1961    -0.6026
    -0.5883    -0.5022
         0    -0.3264
    -0.7845     0.5273

R =
    -5.0990    -0.7845
         0    -3.0634
>>

```

Nyt siis saimme $\mathbf{Q}_1:n$ ja $\mathbf{R}_1:n$

8.7 Lineaarikuvauksen matriisi

Sec:LatLinMap

Tässä kappaleessa oletamme, että L ja M ovat äärellisulotteisia lineaariavaruuksia. Olkoon $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ lineaariavaruuden L kanta ja olkoon $G = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ lineaariavaruuden M kanta. (Toisin sanoen $\dim(L) = n$ ja $\dim(M) = m$.)

Jokaisella L :n vektorilla $\mathbf{x}$ on yksikäsitteiset koordinaatit $(x_1, x_2, \dots, x_n)$ kannassa F siten, että

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n$$

Vastaavasti jokaisella M :n vektorilla $\mathbf{y}$ on yksikäsitteiset koordinaatit $(y_1, y_2, \dots, y_m)$ kannassa G siten, että

$$\mathbf{y} = y_1 \mathbf{w}_1 + y_2 \mathbf{w}_2 + \dots + y_m \mathbf{w}_m$$

Olkoon $f : L \rightarrow M$ on lineaarikuvaus. Jos tunnemme kantavektoreiden $\mathbf{v}_i$ kuvat, niin koko kuvaus tunnetaan täydellisesti. Olkoon nimittäin

$$f(\mathbf{v}_i) = a_{1i}\mathbf{w}_1 + a_{2i}\mathbf{w}_2 + \cdots + a_{mi}\mathbf{w}_m, \quad \text{kaikilla } i = 1, 2, \dots, n$$

Jos nyt $\mathbf{y} = f(\mathbf{x})$, niin

$$\begin{aligned} \mathbf{y} &= y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_n\mathbf{w}_m \\ &= f(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n) \\ &= x_1f(\mathbf{v}_1) + x_2f(\mathbf{v}_2) + \cdots + x_nf(\mathbf{v}_n) \\ &= x_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m) \\ &\quad + x_2(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m) \\ &\quad + \cdots + \\ &\quad x_n(a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m) \end{aligned}$$

Vertaamalla ensimmäisen ja viimeisen lausekkeen kertoimia keskenään, näemme, että

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\Leftrightarrow \quad \vec{y}_G = \mathbf{A}\vec{x}_F$$

Muotoilemme edellisen lauseeksi:

Lause 214 Olkoon $f : L \rightarrow M$ lineaarikuvaus. Jos $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ on lineaariavaruuden L kanta, ja $G = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ on lineaariavaruuden M kanta, niin on olemassa kantoihin F ja G ja lineaarikuvaukseen f liittyvä $m \times n$ -matriisi $\mathbf{A}$ siten, että seuraava pitää paikkansa:

Jos $\mathbf{x} \in L$ ja $f(\mathbf{x}) = \mathbf{y} \in M$, niin

$$\vec{y}_G = \mathbf{A}\vec{x}_F,$$

missä $\vec{y}_G$ on $\mathbf{y}$:n koordinaattivektori kannassa G ja $\vec{x}_F$ on $\mathbf{x}$:n koordinaattivektori kannassa F . (Matriisin $\mathbf{A}$ k :s sarake on vektorin $f(\mathbf{v}_k)$ koordinaattivektori kannassa G .) $\square$

Esimerkki 215 Olkoon P_4 korkeintaan astetta 4 olevien polynomifunktioiden lineaariavaruus. Kantana voimme käyttää viisikkoa

$$\{1, x, x^2, x^3, x^4\}$$

Derivaattakuvauksen

$$D : P_4 \rightarrow P_4, p(x) \mapsto p'(x)$$

matriisi on nyt

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

Esimerkki 216 Tarkastellaan toiseksi matriisin

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

määrittää lineaarikuvausta $f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f_A(\vec{x}) = \mathbf{A}\vec{x}$. Jos varustamme lähtö- ja maaliavaruudet luonnollisilla kannoilla

$$F = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad \text{ja} \quad G = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

niin kuvauksen f_A matriisi on $\mathbf{A}$ itse.

Toisaalta on hyvin luonnollista rajoittaa kuvauksen maaliavaruus aliavaruuteen $M = \text{col}(\mathbf{A}) = \text{span}(\vec{a}_1, \vec{a}_2)$. Nyt kuvauksen $f_A : \mathbb{R}^2 \rightarrow M$, $f_A(\vec{x}) = \mathbf{A}\vec{x}$ maaliavaruuden kannaksi on luonnollista valita vapaa vektorijono

$$H = \left\{ \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} \right\}.$$

Tällä kanta-valinnalla kuvauksen f_A matriisi on

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

Kolmanneksi siirrymme ortonormitettuihin koordinaatteihin. F on ortonormitettu, mutta H ei ole. Ortonormitamme vektorijonon H . Muodostetaan vastaava matriisi $\mathbf{H}$ ja sen QR-hajoitelma

```
>> H=[1 2 ; 2 -1 ; 2 2]
```

```
H =
```

```
    1    2
    2   -1
    2    2
```

```
>> [Q1,R1] = qr(H,0)
```

```
Q1 =
```

```
-0.3333    0.5788
-0.6667   -0.7029
-0.6667    0.4134
```

```
R1 =
```

```
-3.0000   -1.3333
         0    2.6874
```

```
>>
```

Siiis uusi ortonormitettu maaliavaruuden kanta on

$$H_2 = \left\{ \begin{pmatrix} -0.3333 \\ -0.6667 \\ -0.6667 \end{pmatrix}, \begin{pmatrix} 0.5788 \\ -0.7029 \\ 0.4134 \end{pmatrix} \right\}$$

Määritämme nyt kuvauksen matriisin tällä kanta valinnalla.

$$f(\vec{e}_1) = A\vec{e}_1 = Q_1 R_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = Q_1 \vec{r}_1$$

$$f(\vec{e}_2) = A\vec{e}_2 = Q_1 R_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = Q_1 \vec{r}_2$$

ja kuvauksen matriisi on tällä kanta-valinnalla

$$A_f = R_1 = \begin{pmatrix} -3 & -1.3333 \\ 0 & 2.6874 \end{pmatrix}.$$

Tärkeintä on nyt muistaa, että kuvauksen matriisi riippuu valituista lähtö ja maaliavaruuden kannoista!

8.8 Kannan vaihto

Sec:BaseChange

Tarkastellaan kahta n -ulotteisen lineaariavaruuden L kantaa $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ ja $\hat{E} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_n\}$. Kannan E kantavektorit voidaan lausua yksikäsitteisellä tavalla kannassa $\hat{E}$

$$\begin{cases} \mathbf{e}_1 = a_{11}\hat{\mathbf{e}}_1 + a_{21}\hat{\mathbf{e}}_2 + \dots + a_{n1}\hat{\mathbf{e}}_n \\ \mathbf{e}_2 = a_{12}\hat{\mathbf{e}}_1 + a_{22}\hat{\mathbf{e}}_2 + \dots + a_{n2}\hat{\mathbf{e}}_n \\ \vdots \\ \mathbf{e}_n = a_{1n}\hat{\mathbf{e}}_1 + a_{2n}\hat{\mathbf{e}}_2 + \dots + a_{nn}\hat{\mathbf{e}}_n \end{cases}$$

Huomaa, että indeksointi on nyt tehty erikoisesti. Edelläoleva on helpointa muistaa seuraavasti

$$E^T = \mathbf{A}^T \hat{E}^T \Leftrightarrow E = \hat{E} \mathbf{A}.$$

Olkoon nyt $\mathbf{u}$ lineaariavaruuden L vektori, jolla on eri kannoissa esitykset

$$\begin{aligned} \mathbf{u} &= x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n, \quad \text{ja} \\ \mathbf{u} &= \hat{x}_1 \hat{\mathbf{e}}_1 + \hat{x}_2 \hat{\mathbf{e}}_2 + \dots + \hat{x}_n \hat{\mathbf{e}}_n. \end{aligned}$$

Silloin

$$\begin{aligned} \mathbf{u} &= \sum_{j=1}^n x_j \mathbf{e}_j = \sum_{j=1}^n x_j \left(\sum_{k=1}^n a_{kj} \hat{\mathbf{e}}_k \right) \\ &= x_1 (a_{11} \hat{\mathbf{e}}_1 + a_{21} \hat{\mathbf{e}}_2 + \dots + a_{n1} \hat{\mathbf{e}}_n) \\ &\quad + x_2 (a_{12} \hat{\mathbf{e}}_1 + a_{22} \hat{\mathbf{e}}_2 + \dots + a_{n2} \hat{\mathbf{e}}_n) \\ &\quad + \dots + \\ &\quad + x_n (a_{1n} \hat{\mathbf{e}}_1 + a_{2n} \hat{\mathbf{e}}_2 + \dots + a_{nn} \hat{\mathbf{e}}_n) \\ &= (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) \hat{\mathbf{e}}_1 \\ &\quad (a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n) \hat{\mathbf{e}}_2 \\ &\quad + \dots \\ &\quad (a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n) \hat{\mathbf{e}}_n \end{aligned}$$

Viimeisestä muodosta voimme lukea koordinaatit hatullisessa kannassa. Siis

$$\begin{cases} \hat{x}_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \hat{x}_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ \hat{x}_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{cases}$$

Merkitsemme seuraavasti.

Kannat:

$$E = (\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_n), \quad \text{ja} \quad \hat{E} = (\hat{\mathbf{e}}_1 \ \hat{\mathbf{e}}_2 \ \dots \ \hat{\mathbf{e}}_n)$$

Vektorin $\mathbf{u}$ koordinaattivektorit kannoissa:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \text{ missä } \mathbf{u} = E\vec{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$$

$$\vec{\hat{x}} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_n \end{pmatrix}, \text{ missä } \mathbf{u} = \hat{E}\vec{\hat{x}} = \hat{x}_1\hat{\mathbf{e}}_1 + \cdots + \hat{x}_n\hat{\mathbf{e}}_n$$

Näillä merkinnöillä saamme tuloksen.

Lause 217 Jos $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ ja $\hat{E} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_n\}$ ovat kaksi lineaariavaruuden L kantaa ja jos $E^T = \mathbf{A}^T \hat{E}^T$ (eli $E = \hat{E}\mathbf{A}$), niin vektorin koordinaatit toteuttavat yhtälöt

$$\vec{\hat{x}} = \mathbf{A}\vec{x}$$

$$\vec{x} = \mathbf{A}^{-1}\vec{\hat{x}}$$

□

Esimerkki 218 Olkoon $E = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ ja $\hat{E} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ kaksi lineaariavaruuden kantaa siten, että

$$\begin{cases} \mathbf{e}_1 = \hat{\mathbf{e}}_1 + 2\hat{\mathbf{e}}_2 - 3\hat{\mathbf{e}}_3 \\ \mathbf{e}_2 = \hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_2 \\ \mathbf{e}_3 = \hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3 \end{cases} \quad \text{eli}$$

$$\Leftrightarrow \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -3 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \\ \hat{\mathbf{e}}_3 \end{pmatrix}$$

a) Laskemme vektorin $\mathbf{u} = 2\hat{\mathbf{e}}_1 + 3\hat{\mathbf{e}}_2 - \hat{\mathbf{e}}_3$ koordinaatit hatuttomassa kannassa.

b) Laskemme vektorin $\mathbf{v} = 4\mathbf{e}_1 - 2\mathbf{e}_2 + \mathbf{e}_3$ koordinaatit hatullisessa kannassa.

```
>> A=[1 2 -3 ; 1 1 0 ; 0 1 1]'
```

```
A =
```

```
    1    1    0
    2    1    1
   -3    0    1
```

```
>> uhat = [2; 3; -1]
```

```
uhat =
```

```
2
```

```
3
```

```
-1
```

```
>> v = [4; -2; 1]
```

```
v =
```

```
4
```

```
-2
```

```
1
```

```
>> u = inv(A)*uhat
```

```
u =
```

```
0.5000
```

```
1.5000
```

```
0.5000
```

```
>> vhat = A*v
```

```
vhat =
```

```
2
```

```
7
```

```
-11
```

Siiis

$$\mathbf{u} = 2\hat{\mathbf{e}}_1 + 3\hat{\mathbf{e}}_2 - \hat{\mathbf{e}}_3 = 0.5\mathbf{e}_1 + 1.5\mathbf{e}_2 + 0.5\mathbf{e}_3$$

$$\mathbf{v} = 4\mathbf{e}_1 - 2\mathbf{e}_2 + \mathbf{e}_3 = 2\hat{\mathbf{e}}_1 + 7\hat{\mathbf{e}}_2 - 11\hat{\mathbf{e}}_3$$

Lause 219 Kahden ortonormitetun kannan välinen kannanvaihto-matriisi on ortogonaalinen.

Todistus. Jos

$$\mathbf{e}_j = a_{1j}\hat{\mathbf{e}}_1 + a_{2j}\hat{\mathbf{e}}_2 + \cdots + a_{nj}\hat{\mathbf{e}}_n, \quad \forall j,$$

niin ortonormitusten perusteella

$$\delta_{ij} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = \vec{a}_i \cdot \vec{a}_j$$

□

Seuraavaksi mietimme, miten kannan vaihto vaikuttaa lineaarikuvauksen matriisiin. Olkoon $f : L \rightarrow L$ lineaarikuvaus, jossa lähtö- ja maaliavaruuksien ollessa sama lineaariavaruus L . Olkoon kannassa E kuvauksen matriisi $\mathbf{M}$. Jos $f(\mathbf{x}) = \mathbf{y}$, niin

$$\vec{y} = \mathbf{M}\vec{x}.$$

Jos $\hat{E}$ on toinen kanta siten, että kantojen välinen kannanvaihto-matriisi on $\mathbf{A}^T$, niin $E = \hat{E}\mathbf{A}$ ja

$$\vec{y} = \mathbf{A}\mathbf{M}\mathbf{A}^{-1}\vec{\hat{x}}$$

9 OMINAISARVOTEHTÄVÄ

Ch:EigSystem

9.1 Kompleksinen lineaariavaruus

Sec:CMatrix

9.1.1 Kompleksiluvut

Pian tulemme tarvitsemaan kompleksisen lineaariavaruuden alkeita. Tätä varten käymme läpi kompleksilukujen alkeet. Asiasta on oma kurssi, joten nyt toteamme asioita vain sen verran, että saamme tässä kurssissa tarvittavat asiat esitettyä.

Kompleksilukujen lukupariesitys. Kompleksiluvut voidaan esittää lukupareina

$$\mathbb{C} = \{(x, y) \mid x, y \in \mathbb{R}\}$$

joiden joukossa on määritelty yhteen- ja kertolasku seuraavasti

$$(x_1, y_1) \hat{+} (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad (9.1)$$

$$(x_1, y_1) \hat{*} (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) \quad (9.2)$$

Kolmikko $(\mathbb{C}, \hat{+}, \hat{*})$ on kunta (ks. sivu 102). Kunnassa laskeminen on helppoa, kunhan kompleksilukujen esittäminen saadaan helpoksi seuraavalla merkinnällä.

Kompleksilukujen esitys imaginääriyksikön avulla. Sovimme seuraavan lyhennysmerkinnän

$$\begin{aligned} (0, 1) &=_{def} i = \underline{\text{imaginääriyksikkö}} \\ (1, 0) &=_{def} 1 \\ \longrightarrow (x, y) &= x + iy \end{aligned}$$

Imaginääriyksikön tärkein ominaisuus, on

$$i^2 = -1$$

Tämän jälkeen kompleksiluvuilla laskeminen sujuu automaattisesti oikein. Laskusääntö ^{plus} ^{kerto} (9.1) ja (9.2) saavat nyt esityksen

$$\begin{aligned} (x_1 + iy_1) + (x_2 + iy_2) &= x_1 + x_2 + i(y_1 + y_2) \\ (x_1 + iy_1)(x_2 + iy_2) &= x_1 x_2 + ix_1 y_2 + ix_2 y_1 + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1) \end{aligned}$$

Määritelmä 220 Kompleksiluvun $z = x + iy \in \mathbb{C}$ reaaliosa on

$$\operatorname{Re} z = x.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ imaginaariosa on

$$\operatorname{Im} z = y.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ liittoluku on

$$z^* = x - iy.$$

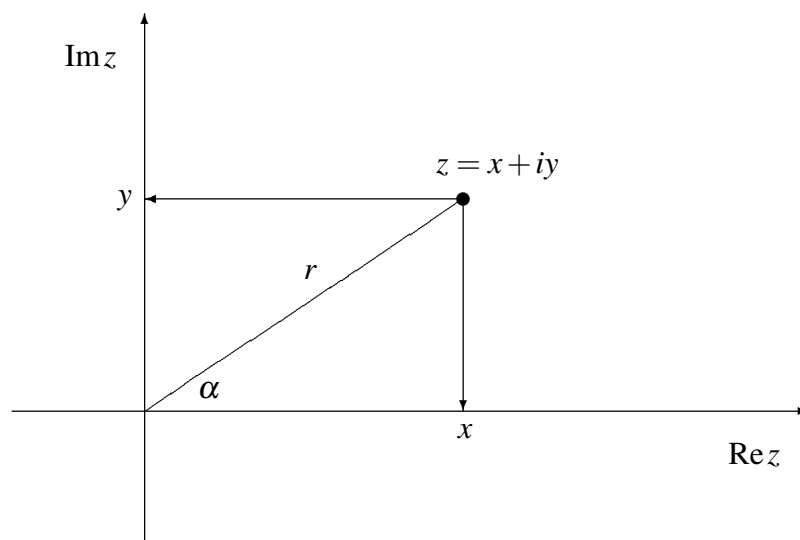
Kompleksiluvun $z = x + iy \in \mathbb{C}$ moduli on

$$|z| = \sqrt{x^2 + y^2} = \sqrt{zz^*}.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ argumentti on

$$\arg z = \arctan\left(\frac{y}{x}\right).$$

Kompleksitaso. Kompleksiluku $z = x + iy$ ”sijoitetaan” xy -tasoon pisteeseen (x, y) . Jos nyt $r = |z|$ ja $\alpha = \arg z$, niin suoraan kuvasta nähdään, että r on pisteen etäisyys origosta ja α on pisteen paikkavektorin ja x -akselin välinen kulma. Silloin



$$x = r \cos \alpha \quad \text{ja} \quad y = r \sin \alpha$$

9.1.2 Kompleksiset vektorit ja matriisit

Kompleksiset vektorit $\mathbb{C}^n$ ja matriisit määritellään nyt kaavioiksi, joihin on sijoitettu kompleksilukuja. Suora lasku osoittaa, että viisikko $(\mathbb{C}^n, \mathbb{C}, +, *, \circ)$ on kompleksiker-toiminen lineaariavaruus.

Muutama asia pitää määritellä uudelleen.

Määritelmä 221 (1) Matriisin $\mathbf{A} = (a_{ij})$ kompleksikonjugaatti on matriisi

$$\mathbf{A}^* = (a_{ij}^*).$$

(Jokainen kaavion luku siis konjugoidaan.)

(2) $m \times n$ -matriisin $\mathbf{A}$ hermitoitu matriisi on

$$\mathbf{A}^H = (\mathbf{A}^*)^T$$

(Matriisi siis konjugoidaan ja transponoidaan.)

(3) Kahden kompleksisen vektorin $\vec{z} \in \mathbb{C}^n$ ja $\vec{w} \in \mathbb{C}^n$ sisätulo määritellään kaavalla

$$\langle \vec{z} | \vec{w} \rangle = z_1^* w_1 + z_2^* w_2 + \cdots + z_n^* w_n = \vec{z}^H \vec{w}$$

Tehdyt määritelmät eivät muuta aiemmin tehtyjä reaalisten vektoreiden ja matriisien vastaavia määritelmiä. Jos nimittäin kaaviossa on vain reaalilukuja, niin herminointi palautuu transponoinniksi.

Kompleksisten vektorien ja matriisien tapauksessa erään aiemmat kaavat saavat uuden asun. Seuraavat kaavat ovat suoria seuraksia tehdyistä määritelmistä, joten emme esitä niille erillisiä todistuksia.

$$(\mathbf{A}^H)^H = \mathbf{A} \quad (9.3)$$

$$(\mathbf{A} + \mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H \quad (9.4)$$

$$(\lambda \mathbf{A})^H = \lambda^* \mathbf{A}^H \quad (9.5)$$

$$(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H \quad (9.6)$$

$$\langle \vec{z} | \vec{w} \rangle = \langle \vec{w} | \vec{z} \rangle^* \quad (9.7)$$

$$\langle \lambda \vec{z} | \vec{w} \rangle = \lambda^* \langle \vec{z} | \vec{w} \rangle \quad (9.8)$$

$$\langle \vec{z} | \lambda \vec{w} \rangle = \lambda \langle \vec{z} | \vec{w} \rangle \quad (9.9)$$

$$\langle \lambda \vec{z} | \vec{w} \rangle = \langle \vec{z} | \lambda^* \vec{w} \rangle \quad (9.10)$$

$$\langle \mathbf{A} \vec{z} | \vec{w} \rangle = \langle \vec{z} | \mathbf{A}^H \vec{w} \rangle \quad (9.11)$$

$$\|\vec{z}\| = \sqrt{\langle \vec{z} | \vec{z} \rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2} \in \mathbb{R} \quad (9.12)$$

Määritelmä 222 Neliömatriisi $\mathbf{A}$ on

$$\begin{aligned} \text{symmetrinen, jos} \quad & \mathbf{A} = \mathbf{A}^T, \\ \text{hermiittinen, jos} \quad & \mathbf{A} = \mathbf{A}^H, \\ \text{ortogonaalinen, jos} \quad & \mathbf{A}^{-1} = \mathbf{A}^T, \\ \text{unitaarinen, jos} \quad & \mathbf{A}^{-1} = \mathbf{A}^H. \end{aligned}$$

9.2 Matriisin ominaisarvot

Sec:MatEig

9.2.1 Perusmääritelmät

Tässä kappaleessa oletamme skalaarit, vektorit ja matriisit kompleksisiksi. Vaikka lähtökohta olisikin reaalin matriisi, niin laskujen aikana saattaa tulla esiin kompleksilukuja ja me nyt hyväksymme sen. Tämä helpottaa asian tarkastelua suuresti.

Seuraavaksi tarkastelemme teknistä ongelmaa, joka on yksinkertaisissa tapauksissaan mahdollista ratkaista suoraan laskemalla. Ensin opimme ratkaisemaan ominaisarvotävän helpossa tapauksessa. Myöhemmin opimme lisää ratkaisun ominaisuuksista, ja niistä ehdoista joilla ratkaisu on olemassa.

Määritelmä 223 Olkoot $\mathbf{A}$ $n \times n$ -matriisi, $\vec{x} \in \mathbb{R}^n, \vec{x} \neq \vec{0}$ vektori ja $\lambda \in \mathbb{R}$ reaalityyppinen luku. Jos

$$\mathbf{A}\vec{x} = \lambda\vec{x}, \tag{9.13} \quad \text{eigveceq}$$

niin sanomme, että λ on matriisin $\mathbf{A}$ ominaisarvo (engl. *eigenvalue*) ja vektori $\vec{x}$ on ominaisarvoon λ liittyvä matriisin $\mathbf{A}$ ominaisvektori (engl. *eigenvector*).

Jos λ on matriisin $\mathbf{A}$ ominaisarvo, niin homogeenisella yhtälöryhmällä

$$(\mathbf{A} - \lambda \mathbf{I})\vec{x} = \vec{0}$$

on ei-triviaali ratkaisu. Tämä on mahdollista vain jos

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \tag{9.14}$$

$$\Leftrightarrow \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{9.15}$$

Ominaisarvotehtävän ratkaisu tehdään $(n + 1)$:ssä vaiheessa:

(1) Ensin ratkaistaan ominaisarvot karakteristisesta yhtälöstä

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0. \quad (9.16)$$

Karakteristinen yhtälö on λ :n suhteen astetta n oleva yhtälö, joten sillä on korkeintaan n juurta (ominaisarvoa).

(2) Jokaista karakteristisen yhtälön juurta (ominaisarvoa) λ kohti ratkaistaan vastaava ominaisvektori homogeenisesta yhtälöryhmästä

$$\mathbf{A}\vec{x} = \lambda\vec{x}. \quad (9.17)$$

c4s15es1

Esimerkki 224 Ratkaise matriisin

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

ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

$$\Leftrightarrow \begin{vmatrix} 2-\lambda & 2 & 0 \\ -1 & 3-\lambda & 2 \\ 0 & 2 & 2-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow (2-\lambda) \begin{vmatrix} 3-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} - 2 \begin{vmatrix} -1 & 2 \\ 0 & 2-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow (2-\lambda)[(3-\lambda)(2-\lambda) - 2 \cdot 2] - 2[-1 \cdot (2-\lambda) - 0 \cdot 2] = 0$$

$$\Leftrightarrow (2-\lambda)[\lambda^2 - 5\lambda + 2] + 2(2-\lambda) = 0$$

$$\Leftrightarrow (2-\lambda)(\lambda^2 - 5\lambda + 4) = 0$$

$$\Leftrightarrow (2-\lambda) = 0 \text{ tai } (\lambda^2 - 5\lambda + 4) = 0$$

$$\Leftrightarrow \lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 4$$

(2) Ominaisarvoa $\lambda_1 = 1$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned}
 & \mathbf{A}\vec{x} = \lambda_1 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 1 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + 2x_2 = x_1 \\ -x_1 + 3x_2 + 2x_3 = x_2 \\ 2x_2 + 2x_3 = x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} [x_1] + 2x_2 = 0 \\ -x_1 + 2x_2 + 2x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \\ \end{array} \\
 \Leftrightarrow & \begin{cases} (x_1) + 2x_2 = 0 \\ + [4x_2] + 2x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \begin{array}{l} \\ -1/2 \\ \leftarrow \end{array} \\
 \Leftrightarrow & \begin{cases} (x_1) + 2x_2 = 0 \\ + (4x_2) + 2x_3 = 0 \\ 0 = 0 \end{cases} \begin{array}{l} \rightarrow x_1 = -2x_2 \\ \rightarrow x_3 = -2x_2 \\ \end{array}
 \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} -2a \\ a \\ -2a \end{pmatrix}, a \in \mathbb{R}.$$

Yleensä vektori valitaan niin, että sen pituus on yksi. Siis $(-2a)^2 + a^2 + (-2a)^2 = 1$, josta ratkeaa $a = 1/3$. Siis lopullinen ominaisarvoon $\lambda_1 = 1$ liittyvä ominaisvektori on

$$\vec{x}_1 = \begin{pmatrix} -2/3 \\ 1/3 \\ -2/3 \end{pmatrix}.$$

(3) Ominaisarvoa $\lambda_2 = 2$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned}
 & \mathbf{A}\vec{x} = \lambda_2 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + 2x_2 = 2x_1 \\ -x_1 + 3x_2 + 2x_3 = 2x_2 \\ 2x_2 + 2x_3 = 2x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} 2x_2 = 0 \\ -x_1 + x_2 + 2x_3 = 0 \\ 2x_2 = 0 \end{cases} \begin{array}{l} \rightarrow x_2 = 0 \\ \rightarrow x_1 = 2x_3 \\ \end{array}
 \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} 2b \\ 0 \\ b \end{pmatrix}, b \in \mathbb{R}.$$

Vektori valitaan taas niin, että sen pituus on yksi. Siis $(2b)^2 + 0^2 + (b)^2 = 1$, josta ratkeaa $b = 1/\sqrt{5}$. Siis lopullinen ominaisarvoon $\lambda_2 = 2$ liittyvä ominaisvektori on

$$\vec{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}.$$

(4) Ominaisarvoa $\lambda_3 = 4$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned} & \mathbf{A}\vec{x} = \lambda_3\vec{x} \\ \Leftrightarrow & \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ \Leftrightarrow & \begin{cases} 2x_1 + 2x_2 = 4x_1 \\ -x_1 + 3x_2 + 2x_3 = 4x_2 \\ 2x_2 + 2x_3 = 4x_3 \end{cases} \\ \Leftrightarrow & \left\{ \begin{array}{l} [-2x_1] + 2x_2 = 0 \\ -x_1 - x_2 + 2x_3 = 0 \\ 2x_2 - 2x_3 = 0 \end{array} \right. \begin{array}{l} -1/2 \\ \leftarrow \\ \end{array} \\ \Leftrightarrow & \left\{ \begin{array}{l} (-2x_1) + 2x_2 = 0 \\ \quad [-2x_2] + 2x_3 = 0 \\ \quad 2x_2 - 2x_3 = 0 \end{array} \right. \begin{array}{l} \\ +1 \\ \leftarrow \end{array} \\ \Leftrightarrow & \left\{ \begin{array}{l} (-2x_1) + 2x_2 = 0 \\ \quad (-2x_2) + 2x_3 = 0 \\ \quad 0 = 0 \end{array} \right. \begin{array}{l} \rightarrow x_1 = x_2 \\ \rightarrow x_3 = x_2 \\ \end{array} \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} c \\ c \\ c \end{pmatrix}, c \in \mathbb{R}.$$

Vektori valitaan taas niin, että sen pituus on yksi. Siis $c^2 + c^2 + c^2 = 1$, josta ratkeaa $c = 1/\sqrt{3}$. Siis lopullinen ominaisarvoon $\lambda_3 = 4$ liittyvä ominaisvektori on

$$\vec{x}_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

Vastauksena ongelmaan voimme siis luetella ominaisarvot ja niihin liittyvät ominaisvektorit

$$\lambda_1 = 1, \vec{x}_1 = \begin{pmatrix} -2/3 \\ 1/3 \\ -2/3 \end{pmatrix}, \text{ ja } \lambda_2 = 2, \vec{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}, \text{ ja } \lambda_3 = 4, \vec{x}_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$$

Saman matriisin ominaisarvot ja ominaisvektorit saadaan Matlabin komennolla

$[V,D] = \text{eigs}(A)$ seuraavasti:

```
>> A = [2 2 0 ; -1 3 2 ; 0 2 2]
```

```
A =
```

```
    2    2    0
   -1    3    2
    0    2    2
```

```
>> [V,D] = eigs(A)
```

```
V =
```

```
    0.5774   -0.8944    0.6667
    0.5774   -0.0000   -0.3333
    0.5774   -0.4472    0.6667
```

```
D =
```

```
    4.0000         0         0
         0    2.0000         0
         0         0    1.0000
```

```
>>
```

Siis matriisin V sarakkeet ovat ominaisvektoreita ja diagonaalimatriisin D diagonaalialkiot ovat vastaavat ominaisarvot.

On melko tavallista, että karakteristisella yhtälöllä on kompleksisia juuria. Nämä juuret ovat myös ominaisarvoja ja niitä vastaavat ominaisvektorit ovat yleensä myös kompleksisia. Kompleksisen esimerkin laskeminen käsin on melko työläs tehtävä, joten tyydymme nyt tietokoneella laskettuun esimerkkiin.

c4s15es2 **Esimerkki 225** >> A = [2 2 0 ; -1 3 2 ; 0 2 -1]

```
A =
```

```
    2    2    0
   -1    3    2
    0    2   -1
```

```
>> [V,D] = eigs(A)
```

```
V =
```

```
    0.8165         0.8165         0.1851
    0.3585 - 0.3711i    0.3585 + 0.3711i   -0.3476
    0.2178 - 0.1403i    0.2178 + 0.1403i    0.9192
```

```

D =
    2.8782 - 0.9089i    0    0
         0    2.8782 + 0.9089i    0
         0         0    -1.7563
>>

```

Edellisen esimerkin matriisi ei näyttänyt hankalalta, mutta silti ominaisarvot ja ominaisvektorit ovat kompleksiset. Tässä vaiheessa helposti tulee epätoivoinen tunne siitä, että asia on liian mutkikas.

Jotta epätoivo ei ahdistaisi toteamme kaksi tulosta, jotka perustelemme pian:

- (1) Symmetrisen matriisin ($\mathbf{A}^T = \mathbf{A}$) ominaisarvot ovat reaaliset.
- (2) Ns. normaalin matriisin ($\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T$) ominaisvektoreista voidaan muodostaa $\mathbb{R}^n$:n kanta.

Useissa sovelluksissa tarkasteltava matriisi on symmetrinen, ja tiedämme siis, että ominaisarvot eivät ole kompleksisia.

Kun karakteristisen yhtälön, $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$, kaikki termit viedään vasemmalle puolelle yhtälöä, niin sinne syntyy karakteristinen polynomi $P(\lambda)$. Karakteristinen polynomi voidaan aina myös saattaa tulomuotoon

$$\begin{aligned}
 P(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) \\
 &= (-1)^n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \cdots + \alpha_0
 \end{aligned} \tag{9.18}$$

$$= (-1)^n (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_q)^{k_q} \tag{9.19}$$

Jos matriisin karakteristinen polynomi on muotoa ^{karpoltulo} (9.19), niin sanomme, että ominaisarvon λ_j algebraalinen kertaluku on k_j . Kompleksianalyysin kurssilla opimme, että kertalukujen summa on $k_1 + k_2 + \cdots + k_q = n$.

$\mathbb{C}^n$:n lineaarinen aliavaruus $\ker(\mathbf{A} - \lambda_j \mathbf{I})$ sisältää kaikki ominaisarvoon λ_j liittyvät ominaisvektorit ja nollavektorin (joka ei ole ominaisvektori). Tämän aliavaruuden dimensiota

$$\rho(\lambda_j) = \dim \ker(\mathbf{A} - \lambda_j \mathbf{I}) \geq 1$$

sanomme ominaisarvon λ_j geometriseksi kertaluvuksi. Geometrinen kertaluku kertoo, miten monta lineaarisesti riippumatonta ominaisvektoria voidaan liittää ominaisarvoon λ_j .

Normaalisti matriisin ominaisarvojen algebralliset- ja geometriset kertaluvut yhtyvät. Jos näin ei tapahdu, niin sanomme, että matriisi on defektiivinen. Defektiivisten matriisien tutkiminen ylittää tämän peruskurssin tavoitteet. Termi mainitaan tässä, jotta opiskelija osaa alkaa etsiä lisätietoa kirjallisuudesta, jos laskentaohjelma antaa termin osana virheilmoitusta.

c4s15es3

Esimerkki 226 Tutkitaan matriisin

$$\mathbf{M} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

Ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö saa muodon

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

Siis ominaisarvon $\lambda_1 = 2$ algebrallinen kertaluku on 2, ja ominaisarvon $\lambda_2 = 4$ algebrallinen kertaluku on 1.

(2) Seuraavaksi määritämme ominaisarvoon $\lambda_1 = 2$ liittyvät ominaisvektorit.

$$\begin{aligned}
 & \mathbf{M}\vec{x} = \lambda_1 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 & = & 2x_1 \\ & 3x_2 + x_3 & = & 2x_2 \\ & x_2 + 3x_3 & = & 2x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} & 0 & = & 0 & \rightarrow x_1 = \text{vapaa} \\ x_2 + x_3 & = & 0 & \rightarrow x_2 = -x_3 \\ x_2 + x_3 & = & 0 \end{cases} \\
 \Leftrightarrow & \begin{cases} x_1 & = & x_1 \\ x_2 & = & -x_3 \\ x_3 & = & x_3 \end{cases} \\
 \Leftrightarrow & \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}
 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ geometrinen kertaluku on 2 ja ominaisvektorien virittäjiksi voidaan valita

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{ja} \quad \vec{x}_2 = \begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

(3) Seuraavaksi määritämme ominaisarvoon $\lambda_2 = 4$ liittyvät ominaisvektorit.

$$\begin{aligned}
 & \mathbf{M}\vec{x} = \lambda_2 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 & = & 4x_1 \\ & 3x_2 + x_3 & = & 4x_2 \\ & x_2 + 3x_3 & = & 4x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} -2x_1 & = & 0 & \rightarrow x_1 = 0 \\ & -x_2 + x_3 & = & 0 & \rightarrow x_2 = x_3 \\ & x_2 + -x_3 & = & 0 \end{cases}
 \end{aligned}$$

Ominaisvektoriksi voidaan valita

$$\vec{x}_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Sama matlabilla:

```
>> M = [2 0 0 ; 0 3 1 ; 0 1 3]
```

```
M =
```

```
      2      0      0
      0      3      1
      0      1      3
```

```
>> [V,D] = eigs(M)
```

```
V =
```

```
      0      0      1.0000
  0.7071  -0.7071      0
  0.7071   0.7071      0
```

```
D =
```

```
      4      0      0
      0      2      0
      0      0      2
```

```
>>
```

c4s15es4

Esimerkki 227 Tutkitaan matriisin

$$\mathbf{Q} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

Ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö saa muodon

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

Siis ominaisarvon $\lambda_1 = 2$ algebrallinen kertaluku on 2, ja ominaisarvon $\lambda_2 = 3$ geometrinen kertaluku on 1.

(2) Seuraavaksi määritämme ominaisarvoon $\lambda_1 = 2$ liittyvät ominaisvektorit.

$$\begin{aligned}
 Q\vec{x} &= \lambda_1 \vec{x} \\
 \Leftrightarrow \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow \begin{cases} 2x_1 + x_2 &= 2x_1 \\ x_2 - x_3 &= 2x_2 \\ 2x_2 + 4x_3 &= 2x_3 \end{cases} \\
 \Leftrightarrow \begin{cases} x_2 &= 0 \\ -x_2 - x_3 &= 0 \\ 2x_2 + 2x_3 &= 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \end{array} \\
 \Leftrightarrow \begin{cases} x_2 &= 0 \\ -x_3 &= 0 \\ 2x_2 + 2x_3 &= 0 \end{cases} \begin{array}{l} \rightarrow x_1 = \text{vapaa} \\ \rightarrow x_2 = x_3 = 0 \end{array} \\
 \Leftrightarrow \vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ geometrinen kertaluku on 1.

(3) Seuraavaksi määritämme ominaisarvoon $\lambda_2 = 3$ liittyvät ominaisvektorit.

$$\begin{aligned}
 Q\vec{x} &= \lambda_2 \vec{x} \\
 \Leftrightarrow \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 3 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow \begin{cases} 2x_1 + x_2 &= 3x_1 \\ x_2 - x_3 &= 3x_2 \\ 2x_2 + 4x_3 &= 3x_3 \end{cases} \\
 \Leftrightarrow \begin{cases} -x_1 + x_2 &= 0 \\ -2x_2 - x_3 &= 0 \\ 2x_2 + x_3 &= 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \end{array} \\
 \Leftrightarrow \begin{cases} -x_1 + x_2 &= 0 \\ -2x_2 - x_3 &= 0 \\ 0 &= 0 \end{cases} \begin{array}{l} \rightarrow x_1 = x_2 \\ \rightarrow x_3 = -2x_2 \end{array}
 \end{aligned}$$

Ominaisvektoriksi voidaan valita

$$\vec{x}_2 = \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}$$

Koska ensimmäisen ominaisarvon algebrallinen kertaluku on 2 ja geometrinen kertaluku on 1, on matriisi defektiivinen.

Sama matlabilla:

```
>> Q = [2 1 0 ; 0 1 -1 ; 0 2 4]
Q =
     2     1     0
     0     1    -1
     0     2     4

>> [V,D] = eigs(Q)
V =
    0.4082    1.0000    1.0000
    0.4082   -0.0000         0
   -0.8165    0.0000         0

D =
     3     0     0
     0     2     0
     0     0     2

>>
```

Defektiivisyys paljastuu nyt siitä, että tuloksena saadut ominisvektorit eivät muodosta vapaata systeemiä, sillä

```
>> det(V'*V)
ans =
    4.6259e-017

>>
```

Ominaisarvoihin ja ominaisvektoreihin palaamme takaisin, mutta ensin toteamme joi-takin uusia asioita.

9.2.2 Geometrinen tulkinta helpossa tapauksessa

Edellä on annettu vain tekninen määritelmä, joka saattaa joskus johtaa kompleksisiin ominaisarvoihin ja ominaisvektoreihin. Jotta asiasta saisi nyt konkreettisen hyödyllisen mielikuvan, oletamme tässä alakappaleessa, että $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa $\lambda_1, \lambda_2, \dots, \lambda_n$. Tämä yksinkertaistaa kovasti tarkastelua ja antaa mahdollisuuden nähdä, mitä merkitystä ominaisarvoilla on.

eigbase

Lause 228 Olkoon $\mathbf{A}$ $n \times n$ -matriisi, jonka ominaisarvot ovat reaaliset. Matriisin $\mathbf{A}$ ominaisvektorit muodostavat $\mathbb{R}^n$:n kannan, jos ja vain jos $\mathbf{A}$ ei ole defektiivinen.

Todistus. $\Rightarrow$ Jos $\mathbf{A}$ on defektiivinen, niin lineaarisesti riippumattomien ominaisvektorien lukumäärä on pienempi kuin n , joten ominaisvektorit eivät voi virittää $\mathbb{R}^n$:ää.

$\Leftarrow$ Oletetaan seuraavaksi, että $\mathbf{A}$ ei ole defektiivinen. Luetellaan matriisin ominaisarvot suuruusjärjestyksessä $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ siten, että jokainen ominaisarvo toistuu luettelossa kertalukunsa mukaisesti. Olkoot $\{\vec{x}_1, \dots, \vec{x}_n\}$, vastaavat ominaisvektorit. Koska $\mathbf{A}$ ei ole defektiivinen, voidaan samaan ominaisarvoon liittyvät ominaisvektorit valita vapaiksi.

Ominaisarvoon λ_1 liittyvä ominaisvektori $\{\vec{x}_1\}$ muodostaa vapaan systeemin. Oletamme seuraavaksi, että ominaisarvoihin $(\lambda_1, \lambda_2, \dots, \lambda_{p-1})$ liittyvät ominaisvektorit $\{\vec{x}_1, \dots, \vec{x}_{p-1}\}$ muodostavat vapaan systeemin. Osoitamme seuraavaksi, että ominaisarvoihin $(\lambda_1, \lambda_2, \dots, \lambda_p)$ liittyvät ominaisvektorit $\{\vec{x}_1, \dots, \vec{x}_{p-1}, \vec{x}_p\}$ muodostavat vapaan systeemin. oletamme sitä varten, että

$$\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_p \vec{x}_p = \vec{0}. \quad (9.20) \quad \text{eigbase0}$$

Oletamme, että k viimeistä ominaisvektoria liittyvät samaan ominaisarvoon ($\lambda_{p-k+1} = \lambda_{p-k+2} = \dots = \lambda_p$). Nyt

$$\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_{p-k} \vec{x}_{p-k} = -\mu_{p-k+1} \vec{x}_{p-k+1} - \dots - \mu_p \vec{x}_p = \vec{a}. \quad (9.21) \quad \text{eigbase1}$$

missä vektori $\vec{a}$ on ominaisarvoon λ_p liittyvä ominaisvektori. Kertomalla yhtälö (9.21) puolittain vasemmalta $\mathbf{A}$:lla saadaan

$$\mu_1 \lambda_1 \vec{x}_1 + \mu_2 \lambda_2 \vec{x}_2 + \dots + \mu_{p-k} \lambda_{p-k} \vec{x}_{p-k} = \lambda_p \vec{a}, \quad (9.22) \quad \text{eigbase2}$$

Yhdistämällä (9.21) ja (9.22) saadaan

$$\mu_1 \left(1 - \frac{\lambda_1}{\lambda_p}\right) \vec{x}_1 + \mu_2 \left(1 - \frac{\lambda_2}{\lambda_p}\right) \vec{x}_2 + \dots + \mu_{p-k} \left(1 - \frac{\lambda_{p-k}}{\lambda_p}\right) \vec{x}_{p-k} = \vec{0}.$$

Tämä on mahdollista vain, jos $\mu_1 = \mu_2 = \dots = \mu_{p-k} = 0$. Siis $\vec{a} = \vec{0}$, mikä edelleen on mahdollista vain, jos $\mu_{p-k+1} = \mu_{p-k+2} = \dots = \mu_p = 0$. Olemme siis osoittaneet, että

$$(\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_p \vec{x}_p = \vec{0}) \Rightarrow (\mu_1 = \mu_2 = \dots = \mu_p = 0).$$

Induktioaskel on valmis, ja lauseen todistus saadaan induktiolla p :n suhteen. $\square$

Lause 229 Jos $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa $\lambda_1, \lambda_2, \dots, \lambda_n$, niin vastaavista ominaisvektoreista muodostettu jono $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ on vapaa ja siis $\mathbb{R}^n$:n kanta.

Todistus. Suora seuraus edellisestä. □

Tarkastellaan kuvausta $f: \mathbb{R}^n \rightarrow \mathbb{R}^n, \vec{u} \mapsto \mathbf{A}\vec{u}$. Kun käytämme $\mathbb{R}^n$:n kantana ominaisvektoreista muodostettua kantaa, niin jokainen vektori voidaan lausua kannassa

$$\vec{u} = u_1 \vec{x}_1 + u_2 \vec{x}_2 + \dots + u_n \vec{x}_n$$

ja vektorin kuva on vastaavasti

$$\begin{aligned} f(\vec{u}) &= \mathbf{A}\vec{u} = \mathbf{A}(u_1 \vec{x}_1 + u_2 \vec{x}_2 + \dots + u_n \vec{x}_n) \\ &= u_1 \lambda_1 \vec{x}_1 + u_2 \lambda_2 \vec{x}_2 + \dots + u_n \lambda_n \vec{x}_n \end{aligned}$$

joten kuvauksen matriisi on

$$\mathbf{M}_f = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Geometrinen tulkinta on nyt siis se, että kun vektori lausutaan kantavektoreiden lineaarikombinaationa, se jaetaan kantavektoreiden suuntaisiin komponentteihin. Kerrottaessa matriisilla $\mathbf{A}$ jokainen komponentti tulee kerrottua vastaavalla ominaisarvolla. Jos $|\lambda_j| > 1$, niin komponentti venyy, jos $|\lambda_j| < 1$, niin komponentti kutistuu.

9.2.3 Ominaisarvohajoitelma

Merkitsemme $\mathbf{X}$:llä matriisia, jonka sarakkeina ovat matriisin $\mathbf{A}$ ominaisvektorit ja $\mathbf{D}$:llä disagonaalimatriisia, jonka diagonaalilla ovat matriisin $\mathbf{A}$ ominaisarvot. Jos nyt matriisin $\mathbf{A}$ ominaisvektorit muodostavat ortonormitetun systeemin, niin

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^T. \tag{9.23}$$

eigsys haj

Kaavan ^{leigsyshaj}(9.23) perusteluksi riittää osoittaa, että $\mathbf{XDX}^T \vec{x}_j = \lambda_j \vec{x}_j = \mathbf{A} \vec{x}_j$ kaikilla $j = 1, 2, \dots, n$. Tämä taas on suora lasku, sillä

$$\begin{aligned} \mathbf{X}^T \vec{x}_j &= \vec{e}_j \\ \Rightarrow \mathbf{DX}^T \vec{x}_j &= \mathbf{D} \vec{e}_j = \lambda_j \vec{e}_j \\ \Rightarrow \mathbf{XDX}^T \vec{x}_j &= \lambda_j \mathbf{X} \vec{e}_j = \lambda_j \vec{x}_j \end{aligned}$$

Ominaisarvohajoittelman avulla on helppo perustella joitakin tuloksia.

Lause 230 Jos neliömatriiseilla $\mathbf{A}$ ja $\mathbf{B}$ ovat samat ominaisvektorit ja ominaisvektorit muodostavat ortonormitetun jonon, niin

$$\mathbf{AB} = \mathbf{BA}$$

eli matriisit kommutoivat.

Todistus. Merkitään $\mathbf{A}$:n ominaisarvoista muodostettua diagonaalimatriisia $\mathbf{D}_A$:lla ja vastaavasti $\mathbf{B}$:n ominaisarvoista muodostettua diagonaalimatriisia $\mathbf{D}_B$:lla. Diagonaalimatriisit kommutoivat, joten

$$\begin{aligned} \mathbf{AB} &= \mathbf{XD}_A \mathbf{X}^T \mathbf{XD}_B \mathbf{X}^T = \mathbf{XD}_A \mathbf{D}_B \mathbf{X}^T = \mathbf{XD}_B \mathbf{D}_A \mathbf{X}^T \\ &= \mathbf{XD}_B \mathbf{X}^T \mathbf{XD}_A \mathbf{X}^T = \mathbf{BA} \end{aligned}$$

□

Matriisin potenssit on nyt helppo laskea käyttäen hajoitelmaa ^{leigsyshaj}(9.23):

$$\begin{aligned} \mathbf{A}^2 &= \mathbf{AA} = \mathbf{XDX}^T \mathbf{XDX}^T \\ &= \mathbf{XD}^2 \mathbf{X}^T = \mathbf{X} \begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix} \mathbf{X}^T \\ \mathbf{A}^n &= \mathbf{XD}^n \mathbf{X}^T = \mathbf{X} \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^n \end{pmatrix} \mathbf{X}^T \end{aligned}$$

Lause 231 (Cayley-Hamilton) Olkoon neliömatriisilla $\mathbf{A}$ ominaisarvohajoitelma $\mathbf{A} = \mathbf{XDX}^T$, missä ominaisvektoreista muodostettu matriisi on ortogonaalinen. Matriisi $\mathbf{A}$ toteuttaa oman karakteristisen yhtälönsä.

Todistus.

$$\begin{aligned} P(\mathbf{A}) &= P(\mathbf{XDX}^T) = \mathbf{XP(D)X}^T \\ &= \mathbf{X} \begin{pmatrix} P(\lambda_1) & 0 & \cdots & 0 \\ 0 & P(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P(\lambda_n) \end{pmatrix} \mathbf{X}^T = \mathbf{0} \end{aligned}$$

□

9.3 Similaarisuus ja matriisin diagonalisointi

Sec:MatSim

Määritelmä 232 Neliömatriisit $\mathbf{A}$ ja $\mathbf{B}$ ovat similaariset, jos on olemassa säännöllinen matriisi $\mathbf{T}$ siten, että

$$\mathbf{A} = \mathbf{T}^{-1}\mathbf{BT}.$$

Matriisia $\mathbf{T}$ sanomme similariteettimuunnosmatriisiksi ja muunnosta $\mathbf{B} \mapsto \mathbf{T}^{-1}\mathbf{BT}$ sanomme similariteettimuunnokseksi.

Jos similariteettimuunnosmatriisi on unitaarinen, sanomme että matriisit ovat unitaarisimilaariset, eli

$$\mathbf{A} = \mathbf{T}^H\mathbf{BT}$$

Lause 233 Jos matriisit $\mathbf{A}$ ja $\mathbf{B}$ ovat similaariset, $\mathbf{A} = \mathbf{T}^{-1}\mathbf{BT}$, niin

- (1) matriiseilla on sama karakteristinen polynomi,
- (2) matriiseilla on samat ominaisarvot
- (3) jos $\vec{x}$ on $\mathbf{A}$:n ominaisvektori, niin $\mathbf{T}\vec{x}$ on vastaava $\mathbf{B}$:n ominaisvektori.

Todistus. (1)

$$\begin{aligned} P_A(\lambda) &= \det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{T}^{-1}(\mathbf{B} - \lambda\mathbf{I})\mathbf{T}) \\ &= \det(\mathbf{T}^{-1})\det(\mathbf{B} - \lambda\mathbf{I})\det(\mathbf{T}) \\ &= \det(\mathbf{B} - \lambda\mathbf{I}) = P_B(\lambda) \end{aligned}$$

(2) Seuraa suoraan edellisestä.

(3)

$$\mathbf{B}(\mathbf{T}\vec{x}) = \mathbf{T}(\mathbf{T}^{-1}\mathbf{BT})\vec{x} = \mathbf{TA}\vec{x} = \mathbf{T}\lambda\vec{x} = \lambda(\mathbf{T}\vec{x})$$

□

Määritelmä 234 $n \times n$ -matriisin $\mathbf{A}$ jälki (trace) on sen diagonaalialkioiden summa

$$\operatorname{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}$$

Lause 235

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$$

Todistus. Suora lasku osoittaa, että

$$\begin{aligned} \operatorname{tr}(\mathbf{AB}) &= (\mathbf{AB})_{11} + (\mathbf{AB})_{22} + \cdots + (\mathbf{AB})_{nn} \\ &= (a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}) \\ &\quad + (a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2}) \\ &\quad + \cdots + (a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn}) \\ &= (b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1n}a_{n1}) \\ &\quad + (b_{21}a_{12} + b_{22}a_{22} + \cdots + b_{2n}a_{n2}) \\ &\quad + \cdots + (b_{n1}a_{1n} + b_{n2}a_{2n} + \cdots + b_{nn}a_{nn}) \\ &= (\mathbf{BA})_{11} + (\mathbf{BA})_{22} + \cdots + (\mathbf{BA})_{nn} = \operatorname{tr}(\mathbf{BA}) \end{aligned}$$

□

Lause 236 Similaarisilla matriiseilla on samat jäljet.

Todistus. Olkoon $\mathbf{A} = \mathbf{T}^{-1}\mathbf{BT}$. Silloin

$$\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{T}^{-1}\mathbf{BT}) = \operatorname{tr}(\mathbf{BTT}^{-1}) = \operatorname{tr}(\mathbf{B})$$

□

Jos herää kysymys siitä, ovatko kaksi annettua matriisia similaariset keskenään, niin ensin kannattaa laskea matriisien jäljet ja verrata niitä. Jos jäljet eivät ole samat, niin matriisit eivät voi olla similaariset.

Merkitään jatkossa diagonaalista matriisia merkinnällä

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

Määritelmä 237 Neliömatriisi $\mathbf{A}$ on diagonalisoituva, jos se on similaarinen jonkin diagonaalimatriisiin $\mathbf{D}$ kanssa, eli

$$\mathbf{D} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T}.$$

Jos kuvauksen $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ matriisi on diagonalisoituva, niin sopivalla kannan vaihdolla saadaan kuvauksen matriisi diagonaaliseksi. Tämä saattaa ratkaisevalla tavalla helpottaa kuvauksen tarkastelua. Seuraavassa esimerkki.

Esimerkki 238 Tarkastellaan differentiaaliyhtälöryhmää

$$\begin{cases} x_1'(t) = 3x_1(t) - 0.5x_2(t) - 0.5x_3(t) \\ x_2'(t) = 0.5x_2(t) + 1.5x_3(t) \\ x_3'(t) = 1.5x_2(t) + 0.5x_3(t) \end{cases}$$

Jätetään jatkossa monessa kohdassa ”(t)” pois. Tämä yksinkertaistaa lausekkeita ja saa ne helpommin luettaviksi. Nyt siis

$$\begin{pmatrix} dx_1/dt \\ dx_2/dt \\ dx_3/dt \end{pmatrix} = \begin{pmatrix} 3 & -0.5 & -0.5 \\ 0 & 0.5 & 1.5 \\ 0 & 1.5 & 0.5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (9.24) \quad \boxed{\text{dyryhma}}$$

Merkitsemme nyt

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \frac{d\vec{x}}{dt} = \begin{pmatrix} dx_1/dt \\ dx_2/dt \\ dx_3/dt \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 3 & -0.5 & -0.5 \\ 0 & 0.5 & 1.5 \\ 0 & 1.5 & 0.5 \end{pmatrix}$$

Voimme siis nyt kirjoittaa yhtälön ^{dyryhma}(9.24) muotoon

$$\frac{d\vec{x}}{dt} = \mathbf{A}\vec{x}$$

Kerroinmatriisi on diagonalisoituva, sillä

$$\mathbf{D} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T},$$

kun

$$\mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \mathbf{T}^{-1} = \begin{pmatrix} 1 & -0.5 & -0.5 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & -0.5 \end{pmatrix}$$

Nyt muodostamme uudet funktiot, jotka ovat lineaarikombinaatioita alkuperäisistä:

$$\begin{cases} w_1(t) = x_1(t) - 0.5x_2(t) - 0.5x_3(t) \\ w_2(t) = 0.5x_2(t) + 0.5x_3(t) \\ w_3(t) = 0.5x_2(t) - 0.5x_3(t) \end{cases}$$

$$\Leftrightarrow \vec{w} = \mathbf{T}^{-1} \vec{x}$$

Nyt suora lasku osoittaa, että w -funktiot toteuttavat hyvin yksinkertaiset differentiaaliyhtälöt

$$\begin{aligned} & \begin{cases} dw_1/dt = dx_1/dt - 0.5 dx_2/dt - 0.5 dx_3/dt \\ dw_2/dt = 0.5 dx_2/dt + 0.5 dx_3/dt \\ dw_3/dt = 0.5 dx_2/dt - 0.5 dx_3/dt \end{cases} \\ \Leftrightarrow & \begin{pmatrix} dw_1/dt \\ dw_2/dt \\ dw_3/dt \end{pmatrix} = \mathbf{T}^{-1} \frac{d\vec{x}}{dt} = \mathbf{T}^{-1} \mathbf{A} \vec{x} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T} \vec{w} = \mathbf{D} \vec{w} \\ \Leftrightarrow & \begin{pmatrix} dw_1/dt \\ dw_2/dt \\ dw_3/dt \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \\ \Leftrightarrow & \begin{cases} dw_1/dt = 3w_1(t) \\ dw_2/dt = 2w_2(t) \\ dw_3/dt = -w_3(t) \end{cases} \end{aligned}$$

Tämän ryhmän ratkaisufunktiot voidaan kirjoittaa suoraan

$$\vec{w} = \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{pmatrix} = \begin{pmatrix} C_1 e^{3t} \\ C_2 e^{2t} \\ C_3 e^{-t} \end{pmatrix},$$

ja lopulta

$$\vec{x} = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \mathbf{T} \vec{w} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} C_1 e^{3t} \\ C_2 e^{2t} \\ C_3 e^{-t} \end{pmatrix} = \begin{pmatrix} C_1 e^{3t} + C_2 e^{2t} \\ C_2 e^{2t} + C_3 e^{-t} \\ C_2 e^{2t} - C_3 e^{-t} \end{pmatrix}.$$

Lause 239 Olkoon $\mathbf{A}$ $n \times n$ -matriisi, jonka ominaisarvot ovat reaaliset. Matriisi $\mathbf{A}$ on diagonalisoituva, jos ja vain jos $\mathbf{A}$ ei ole defektiivinen.

Todistus. $\Rightarrow$ Olkoon

$$\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n) = \mathbf{T}^{-1} \mathbf{A} \mathbf{T}.$$

Nyt

$$\mathbf{T} \mathbf{D} = \mathbf{A} \mathbf{T}.$$

Edellisen yhtälön LHS:n k :s sarake on $\lambda_k \vec{t}_{\bullet,k}$ ja RHS:n k :s sarake on $\mathbf{A} \vec{t}_{\bullet,k}$. Siis $\mathbf{T}$:n sarakkeet ovat matriisin $\mathbf{A}$ ominaisvektorit. Koska $\mathbf{T}$ on säännöllinen sen rangi on n , joten $\mathbf{A}$:n ominaisvektorit muodostavat vapaan systeemin. Siis $\mathbf{A}$ ei ole defektiivinen.

$\Leftarrow$ Olkoon $\mathbf{A}$ ei-defektiivinen. Silloin lauseen (228) ^{|c4s15theigbase} mukaan matriisilla $\mathbf{A}$ on n ominaisvektoria, jotka muodostavat vapaan systeemin. Tehdään ominaisvektoreista matriisi $\mathbf{T}$ ja ominaisarvoista diagonaalimatriisi $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Silloin vastaavasti kuin edellä

$$\begin{aligned} \mathbf{T}\mathbf{D} &= \mathbf{A}\mathbf{T} \\ \Rightarrow \mathbf{D} &= \mathbf{T}^{-1}\mathbf{A}\mathbf{T}. \end{aligned}$$

□

Lause 240 Jos $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa, niin matriisi on diagonalisoituva

Todistus. Lauseen (229) ^{|c4s15threalbase} mukaan matriisi ei ole defektiivinen jolloin lauseen (239) ^{|c4s16thdiag1} mukaan se on diagonalisoituva. □

9.4 Schurin muoto ja normaalit matriisit

Sec:MatSchur

Tässä kappaleessa esitämme todistukset vain tapauksessa, jossa matriisin ominaisarvot ovat reaaliset. Todistukset on mahdollista yleistää kompleksisten ominaisarvojen tapaukseen (Ks. Kivelä: ”Matriisilasku ja Lineaarialgebra” sivut 176–178).

Lause 241 Olkoot $\lambda_1, \lambda_2, \dots, \lambda_n$ reaalisen $n \times n$ -matriisin $\mathbf{A}$ ominaisarvot. On olemassa unitaarimatriisi $\mathbf{M}$ siten, että

$$\mathbf{M}^H \mathbf{A} \mathbf{M} = \begin{pmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & \mathbf{C} \end{pmatrix},$$

^{thschur1} missä matriisin $\mathbf{C}$ ominaisarvot ovat $\lambda_2, \dots, \lambda_n$.

Todistus. Olkoon $\vec{x}$ ominaisarvoon λ_1 liittyvä normitettu ominaisvektori (ts. $\vec{x}^T \vec{x} = 1$). Oletamme nyt, että λ_1 on reaalinen. (Kompleksinen tapaus ks. Kivelä: ”Matriisilasku ja Lineaarialgebra” sivut 176–178.) Täydennetään yhden vektorin systeemi $\{\vec{x}\}$ ortonormeeratuksi kannaksi $\{\vec{x}, \vec{v}_1, \vec{v}_2, \dots, \vec{v}_{n-1}\}$. Muodostetaan näistä vektoreista (sarakkeisiin sijoittamalla) matriisi

$$\mathbf{M} = \begin{pmatrix} \vec{x} & \mathbf{V} \end{pmatrix},$$

joka on sarakkeiden ortonormeerauksen perusteella unitaarinen. Edelleen ortonormeerauksen takia $\mathbf{V}^T \vec{x} = \vec{0}$. Siis

$$\begin{aligned} \mathbf{M}^H \mathbf{A} \mathbf{M} &= \begin{pmatrix} \vec{x}^T \\ \mathbf{V}^T \end{pmatrix} \mathbf{A} \begin{pmatrix} \vec{x} & \mathbf{V} \end{pmatrix} = \begin{pmatrix} \vec{x}^T \mathbf{A} \vec{x} & \vec{x}^T \mathbf{A} \mathbf{V} \\ \mathbf{V}^T \mathbf{A} \vec{x} & \mathbf{V}^T \mathbf{A} \mathbf{V} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & \vec{x}^T \mathbf{A} \mathbf{V} \\ \vec{0} & \mathbf{C} \end{pmatrix} = \mathbf{G} \end{aligned}$$

Koska matriisit $\mathbf{A}$ ja $\mathbf{G}$ ovat similaariset, niillä on samat karakteristiset polynomit. Siis

$$P_A(\lambda) = P_G(\lambda) = (\lambda_1 - \lambda)P_C(\lambda)$$

ja $\mathbf{C}$:n ominaisarvot ovat $\lambda_2, \dots, \lambda_n$. □

Lause 242 $n \times n$ -matriisi $\mathbf{A}$ on unitaarisimilaarinen yläkolmiomatriisin kanssa.

$$\mathbf{M}^H \mathbf{A} \mathbf{M} = \begin{pmatrix} \lambda_1 & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & \lambda_2 & u_{23} & \cdots & u_{2n} \\ 0 & 0 & \lambda_3 & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Yläkolmiomatriisin lävistäjäalkiot ovat $\mathbf{A}$:n ominaisarvot.

Todistus. Tehdään induktio n :n suhteen.

Kun $n = 1$, väite on tosi, koska 1×1 -matriisi on valmiiksi yläkolmiomatriisi.

Oletamme seuraavaksi, että väite on tosi, kun $n = k - 1$ ja osoitamme sen silloin olevan tosi myös, kun $n = n$. Lauseen (241) mukaan on olemassa unitaarinen $k \times k$ -matriisi $\mathbf{M}_1$ siten, että

$$\mathbf{M}_1^H \mathbf{A} \mathbf{M}_1 = \begin{pmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & \mathbf{C}_1 \end{pmatrix},$$

missä λ_1 on matriisin $\mathbf{A}$ ominaisarvo ja matriisin $\mathbf{C}_1$ ominaisarvot ovat $\lambda_2, \lambda_3, \dots, \lambda_k$.

Induktio-oletuksen mukaan on olemassa unitaarinen $(k-1) \times (k-1)$ -matriisi $\mathbf{M}_2$ siten, että

$$\mathbf{C} = \mathbf{M}_2^H \mathbf{C}_1 \mathbf{M}_2$$

on yläkolmiomatriisi, jonka lävistäjäalkiot ovat $\lambda_2, \lambda_3, \dots, \lambda_k$. Nyt matriisi

$$\mathbf{M} = \mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix}$$

on unitaarinen ja

$$\begin{aligned}
 \mathbf{M}^H \mathbf{A} \mathbf{M} &= \left(\mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \right)^H \mathbf{A} \mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix}^H \mathbf{M}_1^H \mathbf{A} \mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix}^H \begin{pmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & \mathbf{C}_1 \end{pmatrix} \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2^H \end{pmatrix} \begin{pmatrix} \lambda_1 & \vec{b}^T \mathbf{M}_2 \\ \vec{0} & \mathbf{C}_1 \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} \lambda_1 & \vec{b}^T \mathbf{M}_2 \\ \vec{0} & \mathbf{C} \end{pmatrix}
 \end{aligned}$$

□

Esimerkki 243 Esimerkkien (224 sivu 193), (225 sivu 196), (226 sivu 198) ja (227 sivu 200) matriisien Schurin muodot ovat

```
>> help schur
```

```
SCHUR Schur decomposition.
```

```
[U,T] = SCHUR(X) produces a quasitriangular Schur matrix T and
a unitary matrix U so that X = U*T*U' and U'*U = EYE(SIZE(U)).
X must be square.
```

```
T = SCHUR(X) returns just the Schur matrix T.
```

```
If X is complex, the complex Schur form is returned in matrix T.
The complex Schur form is upper triangular with the eigenvalues
of X on the diagonal.
```

```
If X is real, two different decompositions are available.
```

```
SCHUR(X,'real') has the real eigenvalues on the diagonal and the
complex eigenvalues in 2-by-2 blocks on the diagonal.
```

```
SCHUR(X,'complex') is triangular and is complex if X has complex
eigenvalues. SCHUR(X,'real') is the default.
```

```
See RSF2CSF to convert from Real to Complex Schur form.
```

Esimerkki (224 sivu 193):

```
>> A=[2 2 0 ; -1 3 2 ; 0 2 2]
A =
     2     2     0
    -1     3     2
     0     2     2

>> [U,T] = schur(A)
U =
    0.5774    -0.7071   -0.4082
    0.5774     0.7071   -0.4082
    0.5774     0.0000    0.8165

T =
    4.0000     2.4495   -0.0000
         0     2.0000    1.7321
         0         0     1.0000

>>
```

Esimerkki (225 sivu 196):

```
>> A = [2 2 0 ; -1 3 2 ; 0 2 -1]
A =
     2     2     0
    -1     3     2
     0     2    -1

>> [U,T] = schur(A)
U =
    0.8490   -0.5199   -0.0938
    0.4602    0.8150   -0.3522
    0.2595    0.2559    0.9312

T =
    2.8782     2.4575   -0.7676
   -0.3362     2.8782    0.7785
         0         0   -1.7563

>> [U,T] = schur(A,'complex')
U =
    0.8165 + 0.0046i    0.4821 + 0.3035i   -0.0938
    0.3564 + 0.3731i   -0.2006 - 0.7547i   -0.3522
```

```

0.2170 + 0.1415i  -0.0273 - 0.2548i  0.9312

T =
2.8782 + 0.9089i  -1.3569 - 1.6306i  -0.7623 - 0.1001i
0                2.8782 - 0.9089i  -0.5658 + 0.5329i
0                0                -1.7563

>>
\end{verbatim}

```

Esimerkki (\ref{c4s15es3} sivu \pageref{c4s15es3}):

```

\begin{verbatim}
>> M = [2 0 0 ; 0 3 1 ; 0 1 3]
M =
2      0      0
0      3      1
0      1      3

```

```

>> [U,T] = schur(M)
U =
1.0000      0      0
0    -0.7071    0.7071
0     0.7071    0.7071

```

```

T =
2      0      0
0      2      0
0      0      4

```

```

>>

```

Esimerkki (\ref{c4s15es4} sivu \pageref{c4s15es4}):

```

>> Q = [2 1 0 ; 0 1 -1 ; 0 2 4]
Q =
2      1      0
0      1     -1
0      2      4

```

```

>> [U,T] = schur(Q)
U =
1.0000      0      0

```

$$\begin{pmatrix} 0 & -0.7071 & -0.7071 \\ 0 & 0.7071 & -0.7071 \end{pmatrix}$$

$$T = \begin{pmatrix} 2.0000 & -0.7071 & -0.7071 \\ 0 & 2.0000 & -3.0000 \\ 0 & 0 & 3.0000 \end{pmatrix}$$

>>

Määritelmä 244 Neliömatriisi $\mathbf{A}$ on normaali, jos

$$\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H.$$

Reaalisen matriisin tapauksessa normaalisuus merkitsee siis, että

$$\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T.$$

Erityisesti hermiittinen ($\mathbf{A}^H = \mathbf{A}$) ja unitaarinen ($\mathbf{A}^H = \mathbf{A}^{-1}$) sekä reaalisten matriisien tapauksessa symmetrinen ($\mathbf{A}^T = \mathbf{A}$) ja ortogonaalinen ($\mathbf{A}^T = \mathbf{A}^{-1}$) ovat normaalin matriisin erikoistapauksia.

Lause 245 Olkoon neliömatriisi $\mathbf{A}$ unitaarisimilaarinen matriisin $\mathbf{B}$ kanssa ($\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{B}$). Silloin $\mathbf{A}$ on normaali, jos ja vain jos $\mathbf{B}$ on normaali.

Todistus. $\Rightarrow$ Oletamme siis, että $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H$. Nyt

$$\begin{aligned} \mathbf{B}^H \mathbf{B} &= (\mathbf{U}^H \mathbf{A} \mathbf{U})^H (\mathbf{U}^H \mathbf{A} \mathbf{U}) = (\mathbf{U}^H \mathbf{A}^H \mathbf{U}) (\mathbf{U}^H \mathbf{A} \mathbf{U}) = \mathbf{U}^H \mathbf{A}^H \mathbf{A} \mathbf{U} \\ &= \mathbf{U}^H \mathbf{A} \mathbf{A}^H \mathbf{U} = (\mathbf{U}^H \mathbf{A} \mathbf{U}) (\mathbf{U}^H \mathbf{A}^H \mathbf{U}) = \mathbf{B} \mathbf{B}^H \end{aligned}$$

$\Leftarrow$ Toisinpäin $\mathbf{A} = \mathbf{U} \mathbf{B} \mathbf{U}^H$ ja oletamme, että $\mathbf{B}^H \mathbf{B} = \mathbf{B} \mathbf{B}^H$. Nyt

$$\begin{aligned} \mathbf{A}^H \mathbf{A} &= (\mathbf{U} \mathbf{B} \mathbf{U}^H)^H (\mathbf{U} \mathbf{B} \mathbf{U}^H) = (\mathbf{U} \mathbf{B}^H \mathbf{U}^H) (\mathbf{U} \mathbf{B} \mathbf{U}^H) = \mathbf{U} \mathbf{B}^H \mathbf{B} \mathbf{U}^H \\ &= \mathbf{U} \mathbf{B} \mathbf{B}^H \mathbf{U}^H = (\mathbf{U} \mathbf{B} \mathbf{U}^H) (\mathbf{U} \mathbf{B}^H \mathbf{U}^H) = \mathbf{A} \mathbf{A}^H \end{aligned}$$

□

Lause 246 Matriisi $\mathbf{A}$ on unitaarisimilaarinen diagonaalimatriisin kanssa, jos ja vain jos se on normaali.

Todistus. $\Rightarrow$ Olkoon $U^H A U = D$, missä U on unitaarinen ja D on diagonaalinen. Koska diagonaalimatriisi on normaali, on edellisen lauseen perusteella myös A normaali.

$\Leftarrow$ Seuraavaksi oletamme, että A on normaali. Koska A :lla on Schurin muoto, se on unitaarisimilaarinen yläkolmiomatriisin kanssa. Koska A on normaali, myös yläkolmiomatriisi on normaali. On helppo nähdä, että normaali yläkolmiomatriisi on välttämättä diagonaalinen. $\square$

Lause 247 Normaalien $n \times n$ -matriisien ominaisvektoreista voidaan muodostaa ortonormeerattu jono.

Todistus. Edellisen lauseen mukaan normaali matriisi A on unitaarisimilaarinen diagonaalimatriisin kanssa. Siis

$$U^H A U = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

missä U on unitaarinen ja D on diagonaalinen. Silloin

$$A U = U D \\ \Leftrightarrow A \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \vec{u}_1 & \lambda_2 \vec{u}_2 & \cdots & \lambda_n \vec{u}_n \end{pmatrix}.$$

Siis U :n sarakkeet ovat ominaisvektoreita, ja U :n unitaarisuudesta seuraa, että vektori-systeemi $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ on vapaa ja ortonormeerattu. $\square$

Lause 248 Hermiittisen matriisin ($A = A^H$) ja reaalisen symmetrisen matriisin ($A = A^T$) Schurin muoto on diagonaalinen ja ominaisarvot ovat reaaliset.

Todistus. Lauseen ^{lause 247} (246) mukaan matriisin Schurin muoto on diagonaalinen. Olkoon $A = A^H$ ja $U^H A U = D$, missä U on unitaarinen ja D on diagonaalinen Schurin muoto. Silloin

$$D^H = (U^H A U)^H = U^H A^H U = U^H A U = D.$$

Tästä seuraa, että D :n lävistäjäalkiot, eli A :n ominaisarvot ovat reaaliset. $\square$

YHTEENVETO: Symmetrisen $n \times n$ -matriisin ominaisarvot ovat reaaliset, ei johda ongelmia, vaan ominaisarvot ovat reaaliset ja ominaisvektoreista voidaan muodostaa $\mathbb{R}^n$:n ortonormitettu kanta.

Lause 249 Olkoon $n \times n$ -neliömatriisin $\mathbf{A}$ ominaisarvot $\lambda_1, \lambda_2, \dots, \lambda_n$. Silloin

$$\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$$

ja

$$\operatorname{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Todistus. Matriisi on unitaarisimilaarinen Schurin muotonsa kanssa

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{S},$$

missä $\mathbf{U}$ on unitaarinen ja $\mathbf{S}$ on yläkolmiomatriisi, jonka diagonaalialkiot ovat $\lambda_1, \lambda_2, \dots, \lambda_n$. Silloin koska similaarisilla matriiseilla on samat determinantit ja jäljet, niin

$$\begin{aligned} \det(\mathbf{A}) &= \det(\mathbf{S}) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n, & \text{ja} \\ \operatorname{tr}(\mathbf{A}) &= \operatorname{tr}(\mathbf{S}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n \end{aligned}$$

□

9.5 Matriisihajoitelmia

Sec:MatDecomp

Aiemmissä kappaleissa on jo esiintynyt kolme matriisi-hajoitelmaa. Seuraavassa on tiivistelmä näistä kolmesta.

QR-hajoitelma.

(1) Jos $\mathbf{A}$ on $m \times n$ -matriisi, jolle $\operatorname{rank}(\mathbf{A}) = n$ (eli $\mathbf{A}$ on ”a matrix of full rank”), niin matriisilla on QR-hajoitelma

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \tag{9.25}$$

missä $\mathbf{Q}$ on ortogonaalinen $m \times m$ matriisi, ja $\mathbf{R}$ on $m \times n$ yläkolmiomatriisi, jonka diagonaalialkiot eroavat nolasta.

(2) Jos $\mathbf{A}$ on $m \times n$ -matriisi, jolle $\operatorname{rank}(\mathbf{A}) < n$, niin matriisilla on QR-hajoitelma

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \tag{9.26}$$

missä $\mathbf{Q}$ on ortogonaalinen $m \times m$ matriisi, ja $\mathbf{R}$ on $m \times n$ yläkolmiomatriisi, jonka diagonaalilla on nollia.

(3) Jos $\mathbf{A}$ on $m \times n$ -matriisi, jolle $\text{rank}(\mathbf{A}) = n$, niin matriisilla on QR-hajoitelma

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R}_1, \quad (9.27)$$

missä $\mathbf{Q}_1$ on $m \times n$ matriisi jonka sarakevektorit muodostavat ortonormitetun jonon (siis $\mathbf{Q}_1^T \mathbf{Q}_1 = \mathbf{I}$, mutta $\mathbf{Q}_1 \mathbf{Q}_1^T \neq \mathbf{I}$), ja $\mathbf{R}_1$ on $n \times n$ yläkolmiomatriisi, jonka diagonaalialkiot eroavat nolasta.

c4s18qresim1

Esimerkki 250

```
>> A = [1 2 3 ; 0 1 1 ; 1 1 1; 3 -1 2]
A =
     1     2     3
     0     1     1
     1     1     1
     3    -1     2

>> [Q,R] = qr(A)
Q =
   -0.3015   -0.7559    0.4303   -0.3906
         0   -0.3780    0.1632    0.9113
   -0.3015   -0.3780   -0.8753   -0.0000
   -0.9045    0.3780    0.1484    0.1302

R =
   -3.3166         0   -3.0151
         0   -2.6458   -2.2678
         0         0    0.8753
         0         0         0

>> [Q1,R1] = qr(A,0)
Q1 =
   -0.3015   -0.7559    0.4303
         0   -0.3780    0.1632
   -0.3015   -0.3780   -0.8753
   -0.9045    0.3780    0.1484

R1 =
   -3.3166         0   -3.0151
         0   -2.6458   -2.2678
         0         0    0.8753

>> B = [1 2 3 ; 0 1 1 ; 1 1 2; 3 -1 2]
B =
     1     2     3
     0     1     1
     1     1     2
```

```

      3      -1      2
>> [Q2,R2] = qr(B)
Q2 =
-0.3015    -0.7559    -0.4128    -0.4090
      0    -0.3780    -0.2028    0.9033
-0.3015    -0.3780     0.8745     0.0382
-0.9045     0.3780    -0.1539     0.1236

R2 =
-3.3166      0    -3.3166
      0    -2.6458    -2.6458
      0      0     0.0000
      0      0      0
>>

```

Ominaisarvohajoitelma.

Jos reaalinen $n \times n$ -neliömatriisi $\mathbf{A}$ on symmetrinen ($\mathbf{A}^T = \mathbf{A}$), niin ominaisarvot ovat reaaliset ja ominaisvektoreista voidaan muodostaa ortonormitettu kanta. Lisäksi matriisilla on ominaisarvohajoitelma

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^T, \quad (9.28)$$

missä diagonaalimatriisin $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ lävistäjällä ovat matriisin ominaisarvot ja matriisin $\mathbf{X}$ sarakkeissa ovat vastaavat normitetut ominaisvektorit.

Matriisin diagonalisointi.

Jos reaalinen $n \times n$ -neliömatriisi $\mathbf{A}$ ei ole defektiivinen (eli se on normaali $\mathbf{A}^T \mathbf{A} = \mathbf{A}\mathbf{A}^T$, esimerkiksi symmetrinen tai ortogonaalinen), niin se on similaarinen diagonaalimatriisin kanssa

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}, \quad (9.29)$$

missä similariteetti-muunnosmatriisi $\mathbf{X}$ voidaan muodostaa lineaarisesti riippumattomista ominaisvektoreista ja matriisin $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ lävistäjällä ovat vastaavat matriisin ominaisarvot. (Huomaa, että nyt $\mathbf{X}$:n ei tarvitse olla ortonormeerattu.)

Schurin-muoto.

Jokainen reaalinen $n \times n$ -matriisi $\mathbf{A}$ on unitaarisimilaarinen Schurin-muotonsa $\mathbf{S}$ kanssa

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{U}^H, \quad (9.30)$$

missä $\mathbf{U}$ on unitaarinen ($\mathbf{U}^H = \mathbf{U}^{-1}$) ja $\mathbf{S}$ on yläkolmiomatriisi, jonka diagonaalilla on matriisin $\mathbf{A}$ ominaisarvot.

9.5.1 Jordanin muoto

Jordanin lohkoksi sanotaan $r \times r$ matriisia

$$\mathbf{J}(\lambda, r) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

$n \times n$ Jordanin matriisi on lohkomatriisi, jonka lavistäjä-lohkot ovat Jordanin lohkoja, ja muut lohkot ovat nolla-lohkoja.

$$\left(\begin{array}{c|c|c|c} \mathbf{J}(\lambda_1, r_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{J}(\lambda_2, r_2) & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}(\lambda_q, r_q) \end{array} \right)$$

Esimerkki 251

$$\begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & 0 & & & \\ & & & 3 & 0 & & \\ & & & & 3 & 0 & \\ & & & & & 5 & 1 \\ & & & & & & 5 \end{pmatrix} = \begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & & & & \\ \hline & & & 3 & & & \\ \hline & & & & 3 & & \\ \hline & & & & & 5 & 1 \\ & & & & & & 5 \end{pmatrix}$$

Matriisin Jordanin muoto.

Jokainen reaalinen $n \times n$ -neliömatriisi $\mathbf{A}$ on similaarinen Jordanin-matriisin kanssa

$$\mathbf{A} = \mathbf{X}\mathbf{J}\mathbf{X}^{-1}, \quad (9.31) \quad \boxed{\text{Jordan}}$$

missä similariteetti-muunnosmatriisin $\mathbf{X}$ sarakkeita sanotaan yleistetyiksi ominaisvektoreiksi ja $\mathbf{J} = \text{diag}(\mathbf{J}(\lambda_1, r_1), \mathbf{J}(\lambda_2, r_2), \dots, \mathbf{J}(\lambda_q, r_q))$, $r_1 + r_2 + \cdots + r_q = n$.

Perustelu Jordanin muodon olemassaololle on melko mutkikas ja sivuutamme sen tässä. (Ks. Simo Kivel: ”Matriisilasku ja Lineaarialgebra”, sivut 183–190.)

Yhtälö (9.31) voidaan myös kirjoittaa muotoon

$$\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{J},$$

mistä seuraa, että yleistetty ominaisvektori toteuttaa toisen seuraavista yhtälöistä

$$(1) \quad \mathbf{A}\vec{x}_k = \lambda_k \vec{x}_k, \quad \text{tai}$$

$$(2) \quad \mathbf{A}\vec{x}_k = \lambda_k \vec{x}_k + \vec{x}_{k-1}.$$

9.6 Matriisin singulaariarvohajoitelma

Sec:MatSVD

Kun halutaan saada mielikuva lineaarikuvauksesta

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^n, \mathbf{x} \mapsto \mathbf{A}\mathbf{x}, \quad (9.32)$$

niin eräs hyvä tapa on laskea kuvavektorin pituus, kun $\|\mathbf{x}\| = 1$. Jos λ on matriisin ominaisarvo ja $\mathbf{x}$ on sitä vastaava normeerattu ominaisvektori, niin

$$\|\mathbf{A}\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda| \quad (9.33)$$

Jos ominaisvektorit virittävät $\mathbb{R}^n$:n, niin saamme varsin hyvän käsityksen kuvauksesta. Valitettavasti kaikkien matriisien ominaisvektorit eivät ole reaalisia. Onneksi tähän löytyy seuraava ratkaisu.

Olkoon $\mathbf{A}$ mikä tahansa (!) $(m \times n)$ -matriisi. Tutkitaan kuvavektorin normin sijasta normin neliötä

$$\|\mathbf{A}\mathbf{x}\|^2 = (\mathbf{A}\mathbf{x})^T \mathbf{A}\mathbf{x} = \mathbf{x}^T \mathbf{A}^T \mathbf{A}\mathbf{x} \quad (9.34)$$

Matriisi $\mathbf{A}^T \mathbf{A}$ on symmetrinen, sen ominaisarvot ovat reaaliset. Jos λ_j on matriisin $\mathbf{A}^T \mathbf{A}$ ominaisarvo ja $\mathbf{v}_j$ on sitä vastaava normeerattu ominaisvektori, niin

$$\|\mathbf{A}\mathbf{v}_j\|^2 = \mathbf{v}_j^T \mathbf{A}^T \mathbf{A}\mathbf{v}_j = \mathbf{v}_j^T \lambda_j \mathbf{v}_j = \lambda_j \quad (9.35)$$

Matriisin $\mathbf{A}^T \mathbf{A}$ ominaisarvot ovat siis ei-negatiiviset ja ne yleensä järjestetään suuruusjärjestykseen $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0$. Voimme määritellä

Määritelmä 252 $(m \times n)$ -matriisin $\mathbf{A}$ singulaariarvot (*singular values*) ovat

$$\sigma_j = \sqrt{\lambda_j}, j = 1, \dots, n$$

missä $\lambda_1, \dots, \lambda_n$ ovat matriisin $\mathbf{A}^T \mathbf{A}$ ominaisarvot.

Siis $\|\mathbf{A}\mathbf{v}_j\| = \sigma_j$.

Lause 253 Jos $(m \times n)$ -matriisin $\mathbf{A}$ säännöllisyysaste on $\text{Rank}(\mathbf{A}) = r$, niin on olemassa hajoitelma

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T,$$

missä $\mathbf{U}$ on ortogonaalinen $(m \times m)$ -matriisi ja $\mathbf{V}$ on ortogonaalinen $(n \times n)$ -matriisi ja

$$\mathbf{D} = \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$$

Todistus. Matriisi $\mathbf{A}^H \mathbf{A}$ on hermiittinen $n \times n$ -matriisi ja sen ominaisarvot ovat ≥ 0 . Luetellaan ominaisarvot $\lambda_j = \sigma_j^2$ suuruusjärjestyksessä

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0.$$

(Huomaa, että r ensimmäistä ominaisarvoa ovat positiivisia ja loput nollia.) Vastaavat ortonormitetut ominaisvektorit ovat $\vec{v}_1, \dots, \vec{v}_n$. Merkitään

$$\mathbf{V} = (\mathbf{V}_1 \mid \mathbf{V}_2) = (\vec{v}_1 \quad \dots \quad \vec{v}_r \mid \vec{v}_{r+1} \quad \dots \quad \vec{v}_n)$$

Matriisin $\mathbf{A}^H \mathbf{A}$ ominaisarvohajoitelma on

$$\begin{aligned} \mathbf{A}^H \mathbf{A} &= \mathbf{V} \begin{pmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^H, \quad \mathbf{\Sigma}^2 = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2) \\ \Rightarrow \begin{pmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} &= \begin{pmatrix} \mathbf{V}_1^H \\ \mathbf{V}_2^H \end{pmatrix} \mathbf{A}^H \mathbf{A} (\mathbf{V}_1 \mid \mathbf{V}_2) \\ &\Rightarrow \begin{cases} \mathbf{V}_1^H \mathbf{A}^H \mathbf{A} \mathbf{V}_1 = \mathbf{\Sigma}^2 \\ \mathbf{A} \mathbf{V}_2 = \mathbf{0} \end{cases} \end{aligned}$$

Asetetaan

$$\mathbf{U}_1 = \mathbf{A} \mathbf{V}_1 \mathbf{\Sigma}^{-1},$$

jolloin edeltävän mukaan $\mathbf{U}_1^H \mathbf{U}_1 = \mathbf{I}$, ts. $\mathbf{U}_1$:n sarakkeet muodostavat ortonormeeratun jonon. $\mathbf{U}_1$ voidaan täydentää unitaariseksi matriisiksi $\mathbf{U} = (\mathbf{U}_1 \mid \mathbf{U}_2)$. Unitarisuudesta seuraa, että $\mathbf{U}_2^H \mathbf{U}_1 = \mathbf{0}$. Nyt

$$\begin{aligned} \mathbf{U}^H \mathbf{A} \mathbf{V} &= \begin{pmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{pmatrix} \mathbf{A} (\mathbf{V}_1 \mid \mathbf{V}_2) = \begin{pmatrix} \mathbf{U}_1^H \mathbf{A} \mathbf{V}_1 & \mathbf{U}_1^H \mathbf{A} \mathbf{V}_2 \\ \mathbf{U}_2^H \mathbf{A} \mathbf{V}_1 & \mathbf{U}_2^H \mathbf{A} \mathbf{V}_2 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{\Sigma}^{-1} \mathbf{V}_1^H \mathbf{A}^H \mathbf{A} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{U}_2^H \mathbf{U}_1 \mathbf{\Sigma} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \end{aligned}$$

□

Singulaariarvohajoitelman suuri etu on siinä, että **reaalisen $m \times n$ -matriisin $\mathbf{A}$ singulaariarvohajoitelma on aina olemassa ja se on reaalinen!** Singulaariarvohajoitelman laskemiseen on olemassa hyvin luotettavia ja hyvin toimivia algoritmeja, joihin emme nyt syvenny.

Lause 254 Jos $n \times n$ -matriisilla $\mathbf{A}$ on n positiivista singulaariarvoa

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad \mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n),$$

niin $\mathbf{A}$ on säännöllinen ja sen käänteismatriisi on

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T, \quad \mathbf{\Sigma}^{-1} = \text{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}).$$

Todistus. HT

□

Esimerkki 255 ^{|c4s18qresiml} Esimerkin (250) matriisien singulaariarvohajotelmat saadaan seuraavasti.

```
>> help svd
```

```
SVD      Singular value decomposition.
```

```
[U,S,V] = SVD(X) produces a diagonal matrix S, of the same
dimension as X and with nonnegative diagonal elements in
decreasing order, and unitary matrices U and V so that
X = U*S*V' .
```

```
S = SVD(X) returns a vector containing the singular values.
```

```
[U,S,V] = SVD(X,0) produces the "economy size"
decomposition. If X is m-by-n with m > n, then only the
first n columns of U are computed and S is n-by-n.
```

```
See also SVDS, GSVD.
```

```
Overloaded methods
```

```
help sym/svd.m
```

```
>> A = [1 2 3 ; 0 1 1 ; 1 1 1 ; 3 -1 2]
```

```
A =
```

```
    1    2    3
    0    1    1
    1    1    1
    3   -1    2
```

```
>> [U,S,V] = svd(A)
```

```
U =
```

```
-0.6930    0.5189    0.3129   -0.3906
-0.2102    0.3333    0.1192    0.9113
-0.3263    0.1290   -0.9364   -0.0000
-0.6075   -0.7765    0.1048    0.1302
```

```

S =
    4.9476         0         0
         0    2.8685         0
         0         0    0.5412
         0         0         0

V =
   -0.5744   -0.5862   -0.5713
   -0.2658    0.7937   -0.5472
   -0.7742    0.1625    0.6117

>> B = [1 2 3 ; 0 1 1 ; 1 1 2 ; 3 -1 2]
B =
     1     2     3
     0     1     1
     1     1     2
     3    -1     2

>> [U2,S2,V2] = svd(B)
U2 =
   -0.6625    0.4727   -0.4128   -0.4090
   -0.2045    0.3179   -0.2028    0.9033
   -0.4580    0.1548    0.8745    0.0382
   -0.5563   -0.8072   -0.1539    0.1236

S2 =
    5.2577         0         0
         0    2.8907         0
         0         0    0.0000
         0         0         0

V2 =
   -0.5306   -0.6206   -0.5774
   -0.2722    0.7698   -0.5774
   -0.8028    0.1492    0.5774

>>

```

Esimerkin (227) defektiivisen matriisin singulaariarvohajoitelma on

```

>> Q = [2 1 0 ; 0 1 -1 ; 0 2 4]
Q =
     2     1     0
     0     1    -1
     0     2     4

>> [U,S,V] = svd(Q)

```



```

U =
    0.1213    -0.9359   -0.3307
   -0.1005   -0.3430    0.9339
    0.9875     0.0801    0.1357

S =
    4.5221         0         0
         0     2.2793         0
         0         0     1.1642

V =
    0.0537   -0.8212   -0.5681
    0.4414   -0.4908    0.7512
    0.8957    0.2911   -0.3361
>>

```

Ei siis lainkaan ongelmia defektiivisyydestä.

9.7 Matriisinormit

Sec:MatNorms

Lineaarinen yhtälöryhmä $A\vec{x} = \vec{b}$ on hyväkuntoinen, jos pienet muutokset kerroinmatriisissa A tai RHS-vektorissa $\vec{b}$ aiheuttavat vain pieniä muutoksia ratkaisuvektorissa $\vec{x}$. Muussa tapauksessa ryhmä on huonokuntoinen.

c4s19es1

Esimerkki 256 Ryhmän

$$\begin{cases} 0.9999x_1 - 1.0001x_2 = 1 \\ x_1 - x_2 = 1 \end{cases}$$

yksikäsitteinen ratkaisu on $x_1 = 0.5$ ja $x_2 = -0.5$. Jos b_2 :ta hiukan muutetaan, saadaan

$$\begin{cases} 0.9999x_1 - 1.0001x_2 = 1 \\ x_1 - x_2 = 1 + \varepsilon \end{cases} \Leftrightarrow \begin{cases} x_1 = 0.5 + 5000.5\varepsilon \\ x_2 = -0.5 + 4999.5\varepsilon \end{cases}$$

Ratkaisu on hyvin herkkä lähtötietojen muutoksille, joten yhtälöryhmä on huonokuntoinen.

Tämän ilmiön ymmärtämiseksi tutustumme seuraavaksi matriisin normiin. Vektorinormit $\|\cdot\|_1$, $\|\cdot\|_2$ ja $\|\cdot\|_\infty$ esiteltiin määritelmässä (151) sivulla 136.

$$(1) \quad \|\vec{a}\|_1 = |a_1| + |a_2| + \cdots + |a_n| = \sum_{k=1}^n |a_k|$$

$$(2) \quad \|\vec{a}\|_2 = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} = \sqrt{\langle \vec{a} | \vec{a} \rangle}$$

$$(3) \quad \|\vec{a}\|_\infty = \max\{|a_1|, |a_2|, \dots, |a_n|\} = \max_{1 \leq k \leq n} |a_k|$$

Lineaariavaruuden L normi on funktio $N : L \rightarrow \mathbb{R}$ siten, että

$$(N1) \quad N(x) \geq 0, \forall x \in L \text{ ja } N(x) = 0 \Leftrightarrow x = \theta \in L$$

$$(N2) \quad N(\alpha x) = |\alpha|N(x), \forall x \in L, \alpha \in \mathbb{R}$$

$$(N3) \quad N(x+y) \leq N(x) + N(y), \forall x, y \in L$$

Normia on tapana merkitä $N(x) = \|\vec{x}\|$. Ja kahden vektorin $\vec{x}$ ja $\vec{y}$ väliseksi etäisyydeksi sanomme niiden erotuksen normia $\|\vec{x} - \vec{y}\|$.

Valittu vektorinormi indusoi matriiseille vastaavan operaattorinormin.

$$\|\mathbf{A}\|_p = \max_{\vec{x} \neq \vec{0}} \frac{\|\mathbf{A}\vec{x}\|_p}{\|\vec{x}\|_p} = \max_{\|\vec{x}\|_p=1} \|\mathbf{A}\vec{x}\|_p \quad (9.36) \quad \boxed{\text{matnormi}}$$

Erityisesti, kun $p = 1$, $p = 2$ tai $p \rightarrow \infty$, saadaan $m \times n$ -matriisin $\mathbf{A}$ normeille kaavat

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^m |a_{ij}| \quad (= \text{itseisarvojen sarakesummien maksimi})$$

$$\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}| \quad (= \text{itseisarvojen rivisummien maksimi})$$

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A})} = \sigma_{\max}(\mathbf{A})$$

missä $\lambda_{\max}(\mathbf{A}^T \mathbf{A})$ on matriisin $\mathbf{A}^T \mathbf{A}$ itseisarvoltaan suurin ominaisarvo eli matriisin $\mathbf{A}$ suurin singulaariarvo .

Unitaarimatriisin kaikki singulaariarvot ovat ykkösiä ($\mathbf{U} = \mathbf{U}\mathbf{I}\mathbf{I}^T$), joten unitaarimatriisin 2-normi on $\|\mathbf{U}\|_2 = 1$.

Edellä määriteltyjen operaattorinormien lisäksi $m \times n$ -matriisille $\mathbf{A}$ käytetään varsin paljon myös Frobenius normia

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (9.37)$$

Määritelmän seuraa, että matriisin operaattorinormeilla on ominaisuudet

$$\|\mathbf{A}\vec{x}\|_p \leq \|\mathbf{A}\|_p \|\vec{x}\|_p \quad (9.38)$$

$$\|\mathbf{A}\mathbf{B}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p \quad (9.39)$$

$$\|\mathbf{A}^n\|_p \leq \|\mathbf{A}\|_p^n \quad (9.40)$$

9.8 Matriisin kuntoluku

Tarkastellaan nyt yhtälöryhmää, jossa kerroinmatriisi on $n \times n$ -matriisi

$$\mathbf{A}\vec{x} = \vec{b}.$$

Jos RHS muuttuu arvoon $\vec{b} + \Delta\vec{b}$, niin vastaava uusi ratkaisu on $\vec{x} + \Delta\vec{x}$. Arvioimme seuraavassa suhteellista virhettä normin $\|\cdot\|_p$ mielessä.

$$\begin{aligned} & \begin{cases} \mathbf{A}(\vec{x} + \Delta\vec{x}) = \vec{b} + \Delta\vec{b} \\ \mathbf{A}\vec{x} = \vec{b} \end{cases} \\ \Rightarrow & \Delta\vec{x} = \mathbf{A}^{-1}\Delta\vec{b} \\ \Rightarrow & \|\Delta\vec{x}\|_p \leq \|\mathbf{A}^{-1}\|_p \|\Delta\vec{b}\|_p \end{aligned} \quad (9.41)$$

Toisaalta vastaavasti

$$\begin{aligned} \|\vec{b}\|_p &= \|\mathbf{A}\vec{x}\|_p \leq \|\mathbf{A}\|_p \|\vec{x}\|_p \\ \Rightarrow & \frac{1}{\|\vec{x}\|_p} \leq \frac{\|\mathbf{A}\|_p}{\|\vec{b}\|_p} \end{aligned} \quad (9.42)$$

Yhdistämällä epäyhtälöt ^{mev1}(9.41) ja ^{mev2}(9.42) saadaan

$$\frac{\|\Delta\vec{x}\|_p}{\|\vec{x}\|_p} \leq \underbrace{\|\mathbf{A}\|_p \|\mathbf{A}^{-1}\|_p}_{=\text{Cond}_p(\mathbf{A})} \frac{\|\Delta\vec{b}\|_p}{\|\vec{b}\|_p} \quad (9.43) \quad \boxed{\text{cond1}}$$

Kaavassa ^{cond1}(9.43) olevaa kerrointa $\text{Cond}_p(\mathbf{A}) = \|\mathbf{A}\|_p \|\mathbf{A}^{-1}\|_p$ sanotaan matriisin kuntoluvuksi. Hyväkuntoisen yhtälöryhmän kerroinmatriisin kuntoluku on pieni (kuitenkin aina ≥ 1).

Tarkastellaan edelleen yhtälöryhmää

$$\mathbf{A}\vec{x} = \vec{b}$$

Jos kerroinmatriisi muuttuu arvoon $\mathbf{A} + \Delta\mathbf{A}$, niin vastaava uusi ratkaisu on $\vec{x} + \Delta\vec{x}$. Arvioimme seuraavassa suhteellista virhettä normin $\|\cdot\|_p$ mielessä.

$$\begin{aligned} & (\mathbf{A} + \Delta\mathbf{A})(\vec{x} + \Delta\vec{x}) = \vec{b} \\ \Leftrightarrow & \mathbf{A}\Delta\vec{x} + \Delta\mathbf{A}(\vec{x} + \Delta\vec{x}) = \vec{0} \\ \Leftrightarrow & \Delta\vec{x} = -\mathbf{A}^{-1}\Delta\mathbf{A}(\vec{x} + \Delta\vec{x}) \\ \Rightarrow & \|\Delta\vec{x}\|_p \leq \|\mathbf{A}^{-1}\|_p \|\Delta\mathbf{A}\|_p \|\vec{x} + \Delta\vec{x}\|_p \\ & = \text{Cond}(\mathbf{A}) \cdot \frac{\|\Delta\mathbf{A}\|_p}{\|\mathbf{A}\|_p} \cdot \|\vec{x} + \Delta\vec{x}\|_p \end{aligned}$$

Voimme siis arvioida

$$\frac{\|\Delta \vec{x}\|_p}{\|\vec{x}\|_p} \approx \frac{\|\Delta \vec{x}\|_p}{\|\vec{x} + \Delta \vec{x}\|_p} \leq \text{Cond}(\mathbf{A}) \cdot \frac{\|\Delta \mathbf{A}\|_p}{\|\mathbf{A}\|_p} \quad (9.44)$$

Jos matriisin kuntoluku on suuri, on odotettavissa ongelmia laskettaessa matriisilla.
Pieni kuntoluku on hyvä kuntoluku.

Kahden matriisin tulon kuntolukua voimme arvioida seuraavasti

$$\begin{aligned} \text{Cond}_p(\mathbf{AB}) &= \|\mathbf{AB}\|_p \|(\mathbf{AB})^{-1}\|_p = \|\mathbf{AB}\|_p \|\mathbf{B}^{-1} \mathbf{A}^{-1}\|_p \\ &\leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p \|\mathbf{B}^{-1}\|_p \|\mathbf{A}^{-1}\|_p = \text{Cond}(\mathbf{A}) \text{Cond}(\mathbf{B}) \end{aligned}$$

Koska unitaarinen matriisi säilyttää 2-normin, niin unitaarille matriisille $\mathbf{U}$ pätee

$$\begin{aligned} \|\mathbf{U}\|_2 &= 1, \quad \text{ja} \\ \text{Cond}_2(\mathbf{U}) &= 1. \end{aligned}$$

Vaikka unitaarinen (ortogonaalinen) matriisi onkin käsin laskettaessa työläs olio, niin laskettaessa tietokoneella numeerisesti, se siisti ja ongelmaton olio. Unitaarisella matriisilla kertominen ei kasvata lausekkeiden kuntolukuja.

Esimerkki 257 Pyydämme vielä Matlabilta arvion kuntoluvusta esimerkissä ^{|c4s19es1}(256) esiintyneelle kerroinmatriisille

$$\begin{pmatrix} 0.9999 & -1.0001 \\ 1 & -1 \end{pmatrix}.$$

Kun katsoo esimerkkiä ^{|c4s19es1}(256), niin ensimmäinen ajatus on, että kuntoluku olisi noin 5000, mutta asia täytyy tutkia tarkemmin:

Esimerkissä suhteellinen virhe RHS:ssä oli

$$\frac{\|\Delta \vec{b}\|_1}{\|\vec{b}\|_1} = \frac{\varepsilon}{2}$$

ja ratkaisun suhteellinen virhe oli

$$\frac{\|\Delta \vec{x}\|_1}{\|\vec{x}\|_1} = \frac{10000\varepsilon}{1}$$

ja näiden suhde on siis 20000.

```
>> help cond
COND    Condition number with respect to inversion.
COND(X) returns the 2-norm condition number (the ratio of the
largest singular value of X to the smallest). Large condition
numbers indicate a nearly singular matrix.

COND(X,P) returns the condition number of X in P-norm:

      NORM(X,P) * NORM(INV(X),P).

where P = 1, 2, inf, or 'fro'.

Class support for input X:
      float: double, single

See also rcond, condest, condeig, norm, normest.

>> A = [0.9999 -1.0001 ; 1 -1]
A =
    0.9999    -1.0001
    1.0000    -1.0000

>> cond(A,1)
ans =
    2.0001e+004

>> cond(A,2)
ans =
    2.0000e+004

>> cond(A,inf)
ans =
    2.0001e+004

>> cond(A,'fro')
ans =
    2.0000e+004

>>
```

9.9 Neliömuodot ja definiittisyys

Sec:FormDef

Seuraavassa määriteltävä neliömuoto ei ole lineaarikuvaus, vaikka sen ominaisuuksia tutkitaan matriisien avulla. Tässä kurssissa käsiteltävät sovellukset liittyvät optimoin-

tiin.

Määritelmä 258 (Epälineaarinen) kuvaus $\mathbb{R}^n \rightarrow \mathbb{R}$

$$Q(\mathbf{x}) = \sum q_{ij}x_i x_j$$

on $\mathbb{R}^n$:n neliömuoto.

Neliömuoto voidaan aina esittää symmetrisen matriisin avulla.

Esimerkki 259

$$\begin{aligned} & x_1^2 + x_2^2 + 4x_1x_2 - 2x_2x_3 \\ = & x_1x_1 + x_2x_2 + 2x_1x_2 + 2x_2x_1 - x_2x_3 - x_3x_2 \\ = & x_1(x_1 + 2x_2) + x_2(x_2 + 2x_1 - x_3) - x_3(x_2) \\ = & \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 - x_3 \\ -x_2 \end{pmatrix} \\ = & \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned} \quad (9.45)$$

$$(9.46)$$

□

Voimme siis aina olettaa, että neliömuoto voidaan esittää symmetrisen matriisin $\mathbf{A}$ avulla

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \quad (9.47)$$

Määritelmä 260 Symmetrinen $(n \times n)$ -neliömatriisi $\mathbf{A}$ on positiivisesti definiitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} > 0, \text{ kaikilla } \mathbf{x} \neq \mathbf{0}.$$

negatiivisesti definiitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} < 0, \text{ kaikilla } \mathbf{x} \neq \mathbf{0},$$

positiivisesti semidefiniitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0, \text{ kaikilla } \mathbf{x} \in \mathbb{R}^n,$$

negatiivisesti semidefiniitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0, \text{ kaikilla } \mathbf{x} \in \mathbb{R}^n,$$

indefiniitti, jos se ei ole mitään edeltävistä.

Lause 261 Olkoon $\mathbf{A}$ symmetrinen $(n \times n)$ -neliömatriisi, jonka reaaliset ominaisarvot ovat $\lambda_1, \lambda_2, \dots, \lambda_n$.

(1) Jos kaikki ominaisarvot ovat positiivisia ($\lambda_j > 0$), niin $\mathbf{A}$ on positiivisesti definiitti.

(2) Jos kaikki ominaisarvot ovat negatiivisia ($\lambda_j < 0$), niin $\mathbf{A}$ on negatiivisesti definiitti.

(3) Jos kaikki ominaisarvot ovat ei-negatiivisia ($\lambda_j \geq 0$), niin $\mathbf{A}$ on positiivisesti semi-definiitti.

(4) Jos kaikki ominaisarvot ovat ei-positiivisia ($\lambda_j \leq 0$), niin $\mathbf{A}$ on negatiivisesti semi-definiitti.

(5) Jos on olemassa sekä positiivinen ominaisarvo, että negatiivinen ominaisarvo, niin $\mathbf{A}$ on indefiniitti.

Esimerkki 262 Tutkitaan definiittisyys matriisille

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad (9.48)$$

Lasketaan ensin ominaisarvot karakteristisesta yhtälöstä

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \quad (9.49)$$

$$\Leftrightarrow \begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 1-\lambda & -1 \\ 0 & -1 & -\lambda \end{vmatrix} = 0 \quad (9.50)$$

$$\Leftrightarrow (1-\lambda)[(1-\lambda)(-\lambda)-1]-2 \cdot 2 \cdot (-\lambda) = 0$$

$$\Leftrightarrow (1-\lambda)(\lambda^2 - \lambda - 1) + 4\lambda = 0$$

$$\Leftrightarrow -\lambda^3 + 2\lambda^2 + 2\lambda - 1 = 0$$

$$\Leftrightarrow (\lambda + 1)(-\lambda^2 + 3\lambda - 1) = 0$$

$$\Leftrightarrow -(\lambda + 1)\left(\lambda - \frac{3+\sqrt{5}}{2}\right)\left(\lambda - \frac{3-\sqrt{5}}{2}\right) = 0$$

Ominaisarvot ovat siis

$$\lambda_1 = \frac{3+\sqrt{5}}{2} \approx 2.618 \quad \lambda_2 = \frac{3-\sqrt{5}}{2} \approx 0.382 \quad \lambda_3 = -1$$

Siis matriisi on indefiniitti.

Pienten yksinkertaisten matriisien tapauksessa on usein helpompaa soveltaa seuraavaa lausetta

Lause 263 Olkoon $\mathbf{A}$ ($n \times n$) neliömuoto. Lasketaan apudeterminantit (vasemmasta ylänurkasta) $|D_1|, |D_2|, \dots, |D_n|$ asettamalla

$$|D_1| = a_{11} \quad \text{ja} \quad |D_k| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}$$

Jos kaikki apudeterminantit ovat positiivisia, $|D_k| > 0$, niin matriisi on positiivisesti definiitti.

Jos parittomat apudeterminantit ovat negatiivisia, $|D_k| < 0$, kun $k = 1, 3, \dots$, ja parilliset ovat positiivisia, $|D_k| > 0$, kun $k = 2, 4, \dots$, niin matriisi on negatiivisesti definiitti.

Todistus. Siuvuutetaan. □

Esimerkki 264 Tutkitaan vielä kerran definiittisyys matriisille

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad (9.51)$$

Apudeterminantit ovat

$$|D_1| = 1 > 0 \quad (9.52)$$

$$|D_2| = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -1 < 0 \quad (9.53)$$

$$|D_3| = \begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{vmatrix} = -1 < 0 \quad (9.54)$$

Siis matriisi on indefiniitti.

Hakemisto

antisymmetrinen matriisi, 79

arccos, 19

arccot, 19

arcsin, 19

arctan, 19

arcus-funktiot, 18

aste, 10

cos, 16

Gauss-Jordan algoritmi, 54

Gaussin eliminointikeino, 54

gooni, 10

grad, 10

homogeeninen yhtälöryhmä, 65

hypotenuusa, 12

kateetti, 12

kosinifunktio, 16

kosinilause, 15

matriisi, 69

neliömatrissi, 74

nollamatriisi, 74

piiru, 11

pistetulo, 32, 33

pystyvektori, 70

Pythagoran lause, 13

rad, 10

radiaani, 10

ristitulo, 32, 35

sin, 16

sinifunktio, 16

sinilause, 15

skalaariprojektio, 38, 40

suorakulmainen kolmio, 12

symmetrinen matriisi, 79

terävä kulma, 11

transpoosi, 78

tylppä kulma, 11

vaakavektori, 70

vektoriprojektio, 38, 39

voima, 29

voiman momentti, 41, 43

yksikkömatrissi, 74