Lineaarialgebra ja matriisilaskenta I-III

Pekka Pankka

17. kesäkuuta 2024

Lukijalle

Lineaarialgebra ja matriisilaskenta I–III luentomonisteet on tarkoitettu kurssimateriaaliksi Helsingin yliopiston Matemaattisten tieteiden kandiohjelman vastaavan nimisille kursseille. Näillä kursseilla käsitellään matematiikan kandiopinnoissa ja matematiikan sovellusaloilla tarvittavan lineaarialgebran perusteet kolmen periodin aikana.

Tämän kurssimateriaalin tavoitteena on yhdistää matematiikan perus- ja aineopintokurssit Lineaarialgebra ja matriisilaskenta I ja II tilastotieteen valinnaisten aineopintojen kurssiin Lineaarialgebra ja matriisilaskenta III.

Kurssi Lineaarialgebra ja matriisilaskenta I sijoittuu matematiikan perusopintohin ja kurssi Lineaarialgebra ja matriisilaskenta II matematiikan aineopintoihin. Kurssi Lineaarialgebra ja matriisilaskenta III on puolestaan aiemmin sijoittunut tilastotieteen aineopintohin. Yhdessä nämä kurssit muodostavat kokonaisuuden, joka alkaa lineaaristen yhtälöryhmien ratkaisemisesta ja päättyy matriisien normaalimuotoihin kattaen sekä konkreettisemman matriisien teorian että abstraktimman vektoriavaruuksien teorian.

Tämä kolme periodia kestävä kurssikokonaisuus on pituudeltaan poikkeuksellinen. Yleensä lineaarialgebran peruskurssi on yhden lukukauden mittainen ja peruskurssin materiaali rajataan ainoastaan kurssien Lineaarialgebra ja matriisilaskenta I ja II aiheisiin ja matriisin normaalimuotoja koskevia tuloksia käsitellään rajoitetusti. Kolmen periodin mittainen kokonaisuus mahdollistaa kuitenkin aihepiirin kattavan käsittelyn. Tämä on tärkeää erityisesti teorian soveltamista ajatellen, sillä teorian tärkeät sovellukset perustuvat neliömatriisien normaalimuodoista kumpuaviin matriisien syvällisiin ominaisuuksiin.

Ensimmäinen kurssi *Lineaarialgebra ja matriisilaskenta I* keskittyy korostetusti matriisien teoriaan. Konkreettisimmin tämä näkyy sarakeavaruuden $\mathbb{R}^{n\times 1}$ korotustumisena tavallisemman euklidisen avaruuden $\mathbb{R}^n$ sijaan. Tästä on hyötyä sekä kurssin materiaalin konkreettisessa soveltamisessa että teoreettisesti. Konkreettisimmillaan tämä näkyy yhdenmukaisina merkintöinä matriiseille ja vektoreille sekä matriisitulon luonnollisuutena. Teoreettisemmin tästä valinnasta on hyötyä aliavaruuksien teorian soveltamisessa matriisien nolla- ja sarakeavaruuksiin sekä lineaarisen geometrian ja matriisiyhtälöiden välisen yhteyden korostumisena. Huomautuksena sanottakoon, että aliavaruuksien yleinen dimensioteoria on tässä esityksessä siirretty seuraavalle kurssille *Lineaarialgebra ja matriisilaskenta II.*

Determinanttien teoriaa on käsitelty tavallista laajemmin kurssin *Lineaarialgebra ja matriisilaskenta I* materiaalissa käyttäen tilavuusmuodon käsitettä. Tämän laajemman käsittelyn ajatuksena on koota kaikki tarvittava yhteen lähteeseen myöhempää käyttöä varten. Tilavuusmuodon esittely antaa mahdollisuuden johtaa tarkasti determinanttien teorian yleisimmät tulokset determinantin intuitiivisesta tulkinnasta vektoreiden virittämän suunnikkaan tilavuutena. Teoreettisempi käsittely voidaan kuitenkin ohittaa luennoilla ja harjoitustehtävissä keskittyen determinantin laskemiseen liittyviin kysymyksiin. Kurssi päättyy ominaisarvojen teorian esittelyyn. Luvun tavoitteena on antaa perustaidot ominaisarvojen ja ominaisvektoreiden löytämiseen sekä neliömatriisin diagonalisoimiseen tilanteessa, jossa matriisin ominaisvektoreista voidaan muodostaa kanta. Ominaisarvojen olemassaoloteoriaan palataan matriisihajotelmien teorian yhteydessä kurssilla *Lineaarialgebra ja matriisilaskenta III*.

Toinen kurssi *Lineaarialgebra ja matriisilaskenta II* käsittelee äärellisulotteisten vektoriavaruuksien yleistä teoriaa. Kurssin alkuosa käsitelee äärellisesti viritettyjen avaruuksien dimensioteoriaa ja näiden välisten lineaarikuvausten ominaisuuksia. Paljastuu, että yleisen teorian puitteissa voidaan ratkaista helposti ensimmäisellä kurssilla avoimiksi jääneet kysymykset matriisien nolla- ja sarakeavaruuksista. Lisäksi osoitetaan, että äärellisulotteisten vektoriavaruuksien ja niiden välisten lineaarikuvausten yleinen teoria voidaan palauttaa kannan käsitteen avulla laskennallisempaan matriisien teoriaan. Kurssin jälkipuoli käsittelee vektoriavaruuksien geometriaa sisätulon käsitteen avulla ja vektoriavaruuksien lineaarisia itseiskuvauksia eli operaattoreita. Ensin käsitellään sisätuloavaruuksien yleistä teoriaa ja kohtisuoruuteen liittyviä käsitteitä. Tämän jälkeen siirrytään tarkastelemaan lineaaristen operaattoreiden perusteoriaa neliömatriisien teorian pohjalta. Viimeisessä luvussa käsitellään ortogonaaliprojektiota ja transpoosin vastinetta yleisissä sisätuloavaruuksissa eli adjungaatteja.

Kolmas ja viimeinen kurssi keskittyy lineaaristen operaattorien, tai matriisien kielellä neliömatriisien, normaalimuotojen teoriaan. Tämä normaalimuotojen teoria vastaa kysymykseen millaisessa muodossa matriisi voidaan esittää kannanvaihdon avulla. Tärkein tulos on symmetristen matriisien spektraalilause, jonka mukaan symmetrinen matriisi on aina diagonalisoituva. Spektraalilauseesta yhdessä semidefiniittien matriisien teorian kanssa, johdetaan polaarihajotelma ja singulaariarvohajotelma yleisille matriiseille. Kurssi päättyy yleistettyihin ominaisavaruuksiin ja kompleksiseen Jordanin hajotelmaan. Näiden tulosten käsittelemistä varten tarvitaan kaikkea kursseilla I–III aiemmin käsiteltyä materiaalia.

Muu materiaali

Nämä luentomuistiinpanot on tarkoitettu kattavaksi materiaaliksi näille kursseille. Materiaali sisältää myös lukuja ja liitteitä, jotka voidaan sivuuttaa yleiskuvan siitä kärsimättä. Yhteen opukseen ei ole kuitenkaan mahdollista sisällyttää kaikkia asioita tai näkökulmia esityksen siitä kärsimättä. Lukijalle suositellaankin tutustumista myös muihin lineaarialgebraa käsittelevien kirjoihin ja luentomuistiinpanoihin.

Kiitokset

Kiitokset Anne-Maria Ernvall-Hytöselle, Tuomo Kuuselle ja Samuli Siltaselle keskusteluista ja kommenteista lineaarialgebran opettamiseen liittyen. Kiitokset myös Leena

Kalliovirralle ja Petteri Piiroiselle kannustuksesta kirjoittaa muistiinpanot kurssille Lineaarialgebra ja matriisilaskenta III, joka antoi sysäyksen kirjoittaa nämä muistiinpanot.

Suuret kiitokset kurssien Lineaarialgebra ja matriisilaskenta I-III ohjaajille vuonna 2020-2021 Paavo Aholalle, Susanna Heikkilälle, Linda Obsilovalle ja Teemu Lundströmille kommenteista ja korjauksista muistiinpanojen ensimmäiseen versioon. Erityiset kiitokset Esko Kivelälle, Sami Laineelle, Jiahao Lille, Sonja Madetojalle ja Henrik Mannerströmille tekstin tarkasta lukemisesta ja monien painovirheiden löytämisestä.

Helsingissä 17. kesäkuuta 2024
Pekka Pankka

Kirjallisuutta

[1] S. Axler. *Linear algebra done right.* Undergraduate Texts in Mathematics. Springer, Cham, third edition, 2015.

[2] J. Häsä, L. Oinonen ja J. Rämö. *Johdatus lineaarialgebraan – osa I.* 2017.

[3] D. C. Lay. *Linear Algebra and its applications.* Addison-Wesley, third edition, 2006.

[4] J. Liesen ja V. Mehrman. *Linear Algebra.* Springer, 2015.

[5] J. Möttönen. *Lineaarialgebra ja matriisilaskenta III.* 2019.

[6] L. Oinonen ja J. Rämö. *Johdatus lineaarialgebraan – osa II.* 2016.

Sisällys

Osa I

Lineaarialgebra ja matriisilaskenta I

Lukijalle

Kurssin sisältö koostuu kuuden eri otsikolle alle kootuista aiheista:

- lineaarisista yhtälöryhmistä,
- matriiseista,
- aliavaruuksista,
- determinantista ja
- ominaisarvoista.

Todellisuudessa tämä kurssi käsittelee kuitenkin matriiseja, sillä

- lineaaristen yhtälöryhmien ratkaisumetodi johdattaa matriiseihin,
- aliavaruudet ovat tapa ymmärtää matriisien ja yhtälöryhmien ratkaisujen välistä yhteyttä,
- determinantti on tapa ymmärtää matriisin kääntyvyys kertoimien avulla ja
- ominaisarvojen ja -vektoreiden yksi sovellus on ymmärtää yhtälöryhmien ratkaiseminen matriisien sisäisen rakenteen avulla.

Kurssin painotukset olisi voitu valita toisinkin. Olisi myös mahdollista aloittaa geometrisemmalla yleisellä teorialla. Näin ei kuitenkaan ole tehty lähinnä aliavaruus luvun (luku 3) esityksessä tehtyjen valintojen johdosta. Luvussa 3 käsitellään lineaarialgebran keskeisiä vapauden ja virittämisen käsitteitä. Matriiseihin liittyvät sarake- ja nolla-avaruudet antavat näille käsitteille konkreettisen sovelluskohteen, jonka avulla voidaan helposti havaita yleisemmän teorian tarve ja merkitys. Yleisempi teoria tehdään – ja vastaukset syntyneisiin kysymyksiin saadaan – heti kurssin Lineaarialgebra ja matriisilaskenta II alussa.

Luku 1

Lineaariset yhtälöryhmät

Luvun tavoitteet

- Yhtälöryhmän ratkaiseminen rivioperaatiolla.
- Yhtälöryhmän ratkaisujoukon tulkitseminen.
- Yhtälöryhmän supistettu porrasmuoto.
- Yhtälöryhmän ratkaisujoukon muuttumattomuus rivioperaatioissa.

1.1 Motivointi

Aloitetaan tutulla suoran yhtälöllä

$$y = ax + b, \tag{1.1}$$

koska se on yksinkertaisin ja sovelletuin lineaarinen yhtälö. Tämä yhtälö esiintyy esimerkiksi, kun tarkastellaan vaikkapa matkan määräytymistä nopeudesta, nopeuden kiihtyvyydestä, tuotannon määräytymistä ajasta ja varastosta, varallisuuden määräymistä palkasta, hinnan määräytyminen kappalehinnasta jne.

Suorat tasossa

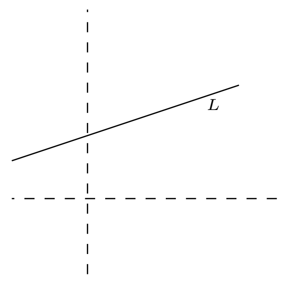
Kun puhekielessä puhutaan yhtälöstä $y = ax + b$, niin silloin tarkoitetaan niitä tason $\mathbb{R}^2$ pisteitä (x, y), jotka toteuttavat yhtälön $y = ax + b$ eli joukkoa

$$L = \{(x, y) \in \mathbb{R}^2 \colon x \in \mathbb{R},\ y = ax + b\},$$

joka voidaan kirjoittaa myös muodossa

$$L = \{(x, ax + b) \in \mathbb{R}^2 \colon x \in \mathbb{R}\};$$

katso kuva 1.1



Kuva 1.1: Yhtälön $y = ax + b$ määräämä suora L tasossa $\mathbb{R}^2$. Luku b kertoo suoran L ja y-akselin leikkauspisteen ja luku a suoran L kulmakertoimen.

Tässä terminologiassa suoralla siis tarkoitetaan yhtälön $y = ax + b$ ratkaisuja (x, y). Yhtälössä $y = ax + b$ luvut a ja b ovat kiinteitä reaalilukuja (nk. kertoimia), jotka määrittelevät yhtälön ja siten joukon L. Vastaavasti x ja y ovat toisistaan riippuvia lukuja, joita tarkastellaan lukuparina (x, y). Lukua x kutsutaan usein vapaaksi muuttujaksi ja lukua y, jonka ajatellaan riippuvan luvusta x, sidotuksi muuttujaksi. Muuttujien x ja y välistä riippuvuutta kutsutaan lineaariseksi riippuvuudeksi.

Leikkaavat suorat

Sovelluksissa ollaan usein kiinnostuneita tilanteesta, jossa halutaan tietää, että milloin kahden eri suoran yhtälön avulla määriteltyä asiaa kohtaavat. Tälläisiä tilanteita ovat esimerkiksi kysymykset, että missä tai milloin eri nopeudella toisiaan vastaan kulkevat kappaleet kohtaavat, millaisilla tuotantomäärillä tuotantotapaa kannattaa vaihtaa jne.

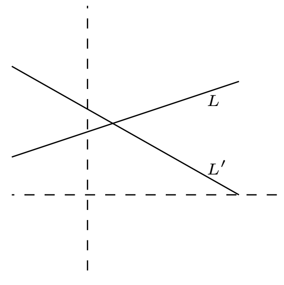
Tälläisissä tilanteissa tarkastellaan esimerkiksi kahta suoraa

$$L = \{(x, y) \in \mathbb{R}^2 \colon y = ax + b\}$$

ja

$$L' = \{(x, y) \in \mathbb{R}^2 \colon y = a'x + b'\}$$

missä $a, b, a', b' \in \mathbb{R}$ ja näiden leikkauspistettä; katso kuva 1.2.



Kuva 1.2: Leikkaavat suorat L ja L' tasossa $\mathbb{R}^2$.

Suorien L ja L' leikkauspisteiksi kutsutaan sellaisia tason $\mathbb{R}^2$ pisteitä, jotka kuuluvat molempiin suoriin L ja L' eli joukon

$$L \cap L' = \{(x,y) \in \mathbb{R}^2 \colon y = ax + b\} \cap \{(x,y) \in \mathbb{R}^2 \colon y = a'x + b'\}$$

pisteitä. Leikkausjoukko voidaan tulkita kahdella tavalla. Toisaalta leikkausjoukko voidaan kirjoittaa muodossa

$$\begin{aligned} L \cap L' &= \{(x,y) \in \mathbb{R}^2 \colon y = ax + b\} \cap \{(x,y) \in \mathbb{R}^2 \colon y = a'x + b'\} \\ &= \{(x,y) \in \mathbb{R}^2 \colon y = ax + b,\ y = a'x + b'\}. \end{aligned}$$

Toisaalta yhtälöt $y = ax + b$ ja $y = a'x + b'$ voidaan kirjoittaa yhtälöpariksi

$$\begin{cases} y &= ax + b \\ y &= a'x + b' \end{cases} \tag{1.2}$$

Näin ollen suorien L ja L' leikkauspisteet ovat täsmälleen yhtälöparin (1.2) ratkaisuja.

Esimerkki 1.1.1. *Ratkaistaan konkreettinen yhtälöpari*

$$\begin{cases} y &= 2x + 3 \\ y &= -(1/2)x + 5 \end{cases}$$

Oletetaan, että (x,y) *on tämän yhtälöparin ratkaisu. Tällöin sekä* $y = 2x + 3$ *että* $y = -(1/2)x + 5$*. Näin ollen luku* x *toteuttaa yhtälön*

$$2x + 3 = -(1/2)x + 5,$$

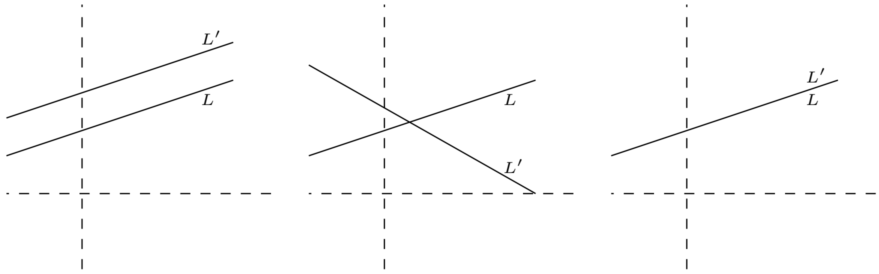
eli yhtälön $(5/2)x = 2$*. Tästä saadaan, että* $x = 4/5$ *ja* $y = 2x + 3 = 2(4/5) + 3 = 8/5 + 3 = 23/5$*.*

Koska emme a priori *tienneet, että yhtälöparilla on ratkaisuja, niin tarkistetaan, että* $(x,y) = (4/5, 23/5)$ *todellakin on ratkaisu. Koska* $2x + 3 = 2(4/5) + 3 = 23/5 = y$ *ja* $-(1/2)x + 5 = -(1/2)(4/5) + 5 = -2/5 + 5 = 23/5 = y$*, niin* $(4/5, 23/5)$ *toteuttaa molemmat yhtlöt. Näin ollen* $(4/5, 23/5)$ *on yhtälöparin ratkaisu.*

Kuva 1.2 ei ole sikäli kattava, että kaksi suoraa voivat kohdata itseasiassa kolmella tavalla: Annetut suorat L ja L' joko leikkaavat, eivät leikkaa, tai ovat samat suorat; katso kuva 1.3.

Kuvasta 1.3 on helppo päätellä, että nämä kolme tapausta liittyvät suorien $y = ax+b$ ja $y = a'x + b'$ kertoimiin a, b ja a', b' seuraavasti:

- suorat L ja L' eivät leikkaa, jos $a = a'$ ja $b \neq b'$,
- suorat L ja L' leikkaavat täsmälleen yhdessä pisteessä, jos $a \neq a'$, ja
- suorat L ja L' ovat sama suora, jos $a = a'$ ja $b = b'$.



Kuva 1.3: Kolme eri vaihtoehtoa suorille L ja L'.

Suoran yhtälön yleinen muoto

Suoran yhtälö $y = ax + b$ ei ole yleisin mahdollinen. Sen avulla ei voi esittää pystysuoria, eli joukkoja

$$L = \{(x, y) \in \mathbb{R}^2 \colon x = b\}$$

missä $b \in \mathbb{R}$.[1]

Nämä molemmat tapaukset saadaan kuitenkin esitettyä yleisemmällä yhtälöllä

$$a_1 x + a_2 y = b, \tag{1.3}$$

missä $a_1, a_2, b \in \mathbb{R}$ ovat yhtälön kiinteitä kertoimia, joista vähintään toinen on nollasta poikkeava. Jälleen tarkkaan ottaen suoralla tarkoitetaan tässä yhtälön (1.3) ratkaisujoukkoa

$$L = \{(x, y) \in \mathbb{R}^2 \colon a_1 x + a_2 y = b\}.$$

Tarkistetaan vielä, että pystysuorat ja yhtälön (1.1) määräämät suorat ovat suoria yhtälön (1.3) mielessä, eli että niiden yhtälöt voidaan kirjoittaa tässä yleisemmässä muodossa.

Pystysuoran $x = b$ tilanteessa riittää valita $a_1 = 1$ ja $a_2 = 0$, joten keskitytään tarkastelemaan suoraa $y = ax + b$. Siirtämällä termi ax yhtälön toiselle puolelle saadaan

$$-ax + y = b.$$

Näin ollen kertoimiksi a_1 ja a_2 voidaan valita $a_1 = -a$ ja $a_2 = 1$.

Tässä yhteydessä on tärkeää huomata, että yhtälöä $y = ax + b$ muokattiin edellä sellaisella tavalla, että yhtälön ratkaisut eivät muutu muokkauksessa. Tämä vastaa yleistä ns. rivioperaation periaatetta, jota tarkastellaan luvussa 1.2.

On myös helppo havaita, että yhtälön avulla ei voi kuvata muita suoria kuin pystysuoria ja sellaisia, joita voidaan esittää yhtälön (1.1) avulla.

Suoran yhtälön yleisellä muodolla (1.3) on kuitenkin tärkeä ominaisuus, jota ei yhtälöllä (1.1) ole: sama suora voidaan esittää useilla eri yleisessä muodossa olevilla

[1] Huomaa, että suoran L määritelmä tulisi kirjoittaa täydellisemmin muodossa $L = \{(x, y) \in \mathbb{R}^2 \colon x = b, y \in \mathbb{R}\}$. Näin ei kuitenkaan kirjallisuudessa yleensä tehdä ja jatkossa noudatetaan tätä tapaa.

yhtälöillä. Tarkemmin tämä havainto on seuraava. Olkoon

$$L = \{(x_1, x_2) \in \mathbb{R}^2 \colon a_1 x + a_2 y = b\}$$

tason $\mathbb{R}^2$ suora. Tällöin jokaisella $\lambda \in \mathbb{R} \setminus \{0\}$ suora

$$L' = \{(x_1, x_2) \in \mathbb{R}^2 \colon \lambda a_1 x + \lambda a_2 y = \lambda b\}$$

on sama suora kuin L, eli $L' = L$. Tämä huomio tulee korostumaan seuraavissa luvuissa.

Avaruuden $\mathbb{R}^3$ tasot

Siirrytään nyt tason suorista tarkastelemaan avaruuden $\mathbb{R}^3$ tasoja. Sovelluksissa tälläinen tilanne syntyy esimerkiksi silloin, kun tarkastellaan ilmiötä, joka riippuu kahdesta parametrista.

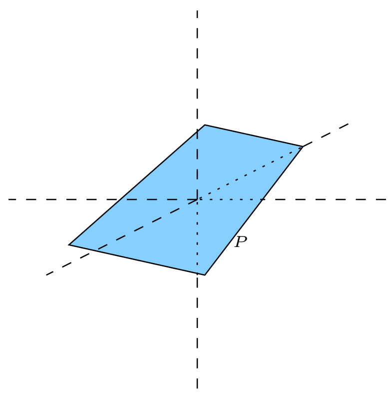
Avaruudessa $\mathbb{R}^3$ jokaisella pisteellä on kolme koordinaattia, joita tässä luvussa merkitään (x, y, z). Tarkastellaan nyt yhtälöä

$$z = a_1 x + a_2 y + b$$

eli joukkoa

$$P = \{(x, y, z) \in \mathbb{R}^3 \colon z = a_1 x + a_2 y + b\},$$

missä $a_1, a_2, b \in \mathbb{R}$; katso kuva 1.4.



Kuva 1.4: Eräs taso P avaruudessa $\mathbb{R}^3$.

Kuten suorien tapauksessa havaitaan, että yhtälöllä $z = a_1 x + a_2 y + b$ ei voida kuvata kaikkia avaruuden $\mathbb{R}^3$ tasoja; esimerkiksi yhtälön $x + y = 1$ kuvaamaa avaruuden $\mathbb{R}^3$ tasoa $P = \{(x, y, z) \in \mathbb{R}^3 \colon x + y = 1\}$ ei voi esittää yhtälön avulla $z = a_1 x + a_2 y + b$, koska joukon P pisteiden z-koordinaatti ei riipu x- ja y-koordinaattien arvoista.

Kuten suorien tapauksessa onkin luonnollista yhtälön $z = a_1 x + a_2 y + b$ sijaan tarkastella yhtälöä

$$a_1 x + a_2 y + a_3 z = b, \tag{1.4}$$

missä $a_1, a_2, a_3, b \in \mathbb{R}$ ja jokin luvuista a_1, a_2, a_3 ei ole nolla. Näin ollen taso on siis joukko

$$P = \{(x, y, z) \in \mathbb{R}^3 \colon a_1x + a_2y + a_3z = b\},$$

missä jokin luvuista $a_1, a_2, a_3 \in \mathbb{R}$ ei ole nolla.

Huomaa, että tässä muotoilussa yksikään koordinaatti ei ole erikoisasemassa ja että tämän yleisen muotoilun avulla voidaan käsitellä kaikki eri tapaukset mukaanlukien koordinaattiakselien suuntaiset tasot.

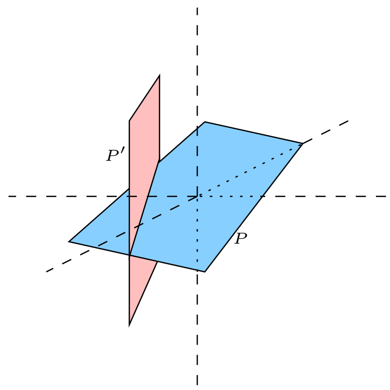
Tarkastellaan nyt kahden avaruuden $\mathbb{R}^3$ tason

$$P = \{(x, y, z) \in \mathbb{R}^3 \colon a_1x + a_2y + a_3z = b\}$$

ja

$$P' = \{(x, y, z) \in \mathbb{R}^3 \colon a_1'x + a_2'y + a_3'z = b'\}$$

leikkausta, eli yhteisiä pisteitä; katso kuva 1.5.



Kuva 1.5: Kaksi tasoa P ja P' avaruudessa $\mathbb{R}^3$.

Kuten leikkaavien suorien tapauksessa, jälleen leikkausjoukko $P \cap P'$ vastaa täsmälleen yhtälöparin

$$\begin{cases} a_1x + a_2y + a_3z &= b \\ a_1'x + a_2'y + a_3'z &= b' \end{cases} \tag{1.5}$$

ratkaisuja $(x, y, z) \in \mathbb{R}^3$, eli sellaisia pisteitä, jotka toteuttavat molemmat yhtälöparin yhtälöt.

Esimerkki 1.1.2. *Ratkaistaan konkreettinen yhtälöpari*

$$\begin{cases} x + 2y + z &= 1 \\ 2x + y - z &= 0 \end{cases} \tag{1.6}$$

käyttämällä samaa ideaa kuin yhtälöparin (1.1.1) *tapauksessa.*

Havaitaan ensin, että siirtämällä termejä yhtälöpari voidaan kirjoittaa muodossa

$$\begin{cases} z &= -x - 2y + 1 \\ z &= 2x + y \end{cases}$$

Oletetaan nyt, että (x, y, z) *on yhtälöparin ratkaisu. Tällöin* $-x - 2y + 1 = 2x + y$*. Siirtämällä jälleen termejä saadaan tämä yhtälö muotoon* $-3x + 1 = 3y$ *eli saadaan yhtälö* $y = -x + 1/3$*. Näin ollen* $z = 2x + y = 2x + (-x + 1/3) = x + 1/3$*, eli saadaan, että koordinaatit* y *ja* z *voidaan kirjoittaa koordinaatin* x *avulla, mutta koordinaatille* x *ei ole ehtoja. Näin ollen saadaan, että* (x, y, z) *on piste* $(x, -x + 1/3, x + 1/3)$*, missä* $x \in \mathbb{R}$*. Yhtälöparin* (1.6) *ratkaisu ei siis ole yksittäinen piste vaan joukko*

$$\{(x, -x + 1/3, x + 1/3) \in \mathbb{R}^3 \colon x \in \mathbb{R}\}.$$

Jälleen tulisi tarkistaa, että pisteet $(x, -x + 1/3, x + 1/3)$ *ovat yhtälöparin* (1.6) *ratkaisuja, mutta sivuutetaan tämä yksityiskohta tällä kertaa.*

Kuten tason suorien tapauksessa, tasojen P ja P' leikkaukselle $P \cap P'$ on kolme vaihtoehtoa: leikkaus on tyhjä, jos tasot ovat saman suuntaisia mutta $P \neq P'$, leikkaus on suora (kuten kuvassa 1.5), jos tasot ovat eri suuntaisia, tai leikkaus on taso, jos $P = P'$. On ehkä hiukan yllättävää, että tasojen P ja P' leikkaus ei voi olla yksittäinen piste. Tämä on helppo perustella myöhemmin käyttäen kehitettyä teoriaa. Toisaalta on helppo löytää esimerkki, että kolmen tason leikkaus voi olla piste.

Huomautus 1.1.3. *Toinen mielenkiintoinen havainto on, että avaruuden* $\mathbb{R}^3$ *suora on itseasiassa aina kahden tason leikkaus. Tämä tietysti herättää kysymyksen, että miten määritellään avaruuden* $\mathbb{R}^3$ *suora.*

Vaikka on intuiviisesti selvää, mikä suoran tulisi olla, määritellään se kuitenkin tarkasti käyttäen seuraavaa määritelmää. Joukko $L \subset \mathbb{R}^3$ *on* suora*, jos on olemassa sellainen nollasta poikkeava vektori* $v \in \mathbb{R}^3$ *ja sellainen vektori* $c \in \mathbb{R}^3$*, että*

$$L = \{c + tv \colon t \in \mathbb{R}\}.$$

Tässä määritelmässä suora on määritelty parametrin t *avulla ja suoraa* L *kutsutaan sen vuoksi joskus parametrisoiduksi suoraksi.*

Tästä määritelmästä seuraa luonnollinen jatkokysymys, että voidaanko avaruuden $\mathbb{R}^3$ *taso määritellä myös parametrisoidusti, eli kahden parametrin avulla. Annettu tason määritelmä on yhtäpitävää seuraavan määritelmän kanssa: joukko* $P \subset \mathbb{R}^3$ *on taso, jos on olemassa sellaiset nollasta poikkeavat vektorit* $v, w \in \mathbb{R}^3$ *ja sellainen vektori* $c \in \mathbb{R}^3$*, että* $v \neq rw$ *kaikilla* $r \in \mathbb{R}$ *ja*

$$P = \{c + tv + sw \colon t, s \in \mathbb{R}\}.$$

Tämä kysymys johdattaa meidät itseasiassa dimension käsitteeseen ja siihen huomioon, että suorat ovat 1*-ulotteisia ja että tasot ovat* 2*-ulotteisia. Näitä asoita käsitellään luvussa 3.*

Palataan nyt yhtälöparin ratkaisemisen problematiikkaan. Tason suoriin liittyvät yhtälöparit voidaan ratkaista käsin ilman teoriaa kuten esimerkissä 1.1.1. Esimerkin 1.1.2 yhtälöpari (1.5) kuitenkin osoittaa, että yhtälöiden ja muuttujien lisääntyessä yleisempien yhtälöryhmien tehokkaaseen ratkaisemiseen tarvitaan teoriaa.

1.2 Yhtälöryhmän määritelmä

Aloitetaan alusta määritelemällä yhtälöryhmän ratkaisut[2].

Määritelmä 1.2.1. *Olkoot $m, n \in \mathbb{Z}_+$ ja $a_{ji} \in \mathbb{R}$ kaikilla $j \in \{1, \dots, m\}$ ja $i \in \{1, \dots, n\}$ sekä $b_j \in \mathbb{R}$ kaikilla $j \in \{1, \dots, m\}$. Vektoria $(x_1, \dots, x_n) \in \mathbb{R}^n$, joka toteuttaa yhtälöt*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases} \tag{1.7}$$

kutsutaan yhtälöryhmän (1.7) *ratkaisuksi. Lukuja a_{ji} kutsutaan* yhtälöryhmän (1.7) kertoimiksi *ja lukuja b_j* yhtälöryhmän (1.7) vakioiksi.

Esimerkki 1.2.2. *Vektori $x = (1, 0, 3) \in \mathbb{R}^3$ on yhtälöryhmän*

$$\begin{cases} x_1 & +2x_2 & -x_3 & = -2 \\ & -x_2 & +x_3 & = 3 \end{cases}$$

eli yhtälöryhmän

$$\begin{cases} x_1 & +2x_2 & -x_3 & = -2 \\ 0 \cdot x_1 & -x_2 & +x_3 & = 3 \end{cases}$$

ratkaisu, sillä

$$\begin{cases} 1 & +2 \cdot 0 & -3 & = -2 \\ & -0 & +3 & = 3 \end{cases}$$

Myös vektori $y = (3, -2, 1)$ on saman yhtälöryhmän ratkaisu, sillä

$$\begin{cases} 3 & +2 \cdot (-2) & -1 & = -2 \\ & -(-2) & +1 & = 3 \end{cases}$$

Huomautus 1.2.3. *Esimerkissä 1.2.2 käytettiin yleistä merkintätapaa, että termiä $a_{ji}x_i$ ei merkitä lainkaan, jos kerroin a_{ji} on nolla.*

Koska yhtälöryhmällä voi olla useita ratkaisuja (tai ei lainkaan ratkaisuja), niin on hyödyllistä määritellä yhtälöryhmän ratkaisujoukon käsite tarkasti.

Määritelmä 1.2.4. *Yhtälöryhmän* (1.7) *ratkaisujoukko $R \subset \mathbb{R}^n$ on yhtälöryhmän* (1.7) *kaikkien ratkaisujen joukko, eli*

$$R = \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon (x_1, \dots, x_n) \text{ on yhtälöryhmän (1.7) ratkaisu}\}.$$

Huomautus 1.2.5. *Yhtälöryhmän ratkaisujoukon määritelmä on hieman tautologisen oloinen. Tämän vuoksi on luontevaa huomata, että vektori $(x_1, \dots, x_n)$ on yhtälöryhmän*

[2] Huomaa, että yhtälöryhmää pidetään käsitteenä, jota ei tarkemmin formalisoida.

(1.7) *ratkaisu, jos ja vain jos* $(x_1, \dots, x_n)$ *on jokaisen yhtälöryhmän* (1.7) *yhtälön ratkaisu. Näin ollen ratkaisujoukko* R *on siis itseasiassa leikkausjoukko*

$$\begin{aligned} R = &\{(x_1, \dots, x_n) \in \mathbb{R}^n \colon a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1\} \\ &\cap \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2\} \\ &\vdots \\ &\cap \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m\}. \end{aligned}$$

Muutama kommentti määritelmistä 1.2.1 ja 1.2.4 on paikallaan.

Huomautus 1.2.6. *Luvut* a_{ji} *yhdessä lukujen* b_j *kanssa määräävät yhtälöryhmän* (1.7)*. Lukuja* $x_1, \dots, x_n$ *kutsutaan yleensä* yhtälöryhmän (1.7) muuttujiksi*. Muuttujan käsitettä ei jatkossa määritellä tarkemmin formaalisti, vaan jatkossa puhutaan vektoreista* $(x_1, \dots, x_n)$ *ja sen kertoimista (tai koordinaateista)* $x_1, \dots, x_n$*. Yhtälöitä* $a_{j1}x_1 + \dots + a_{jn}x_n = b_j$ *kutsutaan yhtälöryhmän* (1.7) *riveiksi.*

Toinen – ja laajempi – kommentti liittyy termeihin *vektori* ja *avaruus* $\mathbb{R}^n$.

Avaruus $\mathbb{R}^n$

Avaruus $\mathbb{R}^n$ määritellään jokaisella $n \in \mathbb{Z}_+$ seuraavasti. Joukkona $\mathbb{R}^n$ koostuu alkiosta $(x_1, \dots, x_n)$, missä jokainen $x_1, \dots, x_n$ on reaaliluku, eli joukon $\mathbb{R}^n$ alkiot ovat sellaisia reaalilukujen jonoja, joissa on n alkiota. Joukko $\mathbb{R}^n$ voidaan siis kirjoittaa muodossa

$$\mathbb{R}^n = \{(x_1, \dots, x_n) \colon x_1, \dots, x_n \in \mathbb{R}\}.$$

Lukuja $x_1, \dots, x_n \in \mathbb{R}$ kutsutaan vektorin $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ koordinaateiksi.

Huomautus 1.2.7. *Lukijaa voi jäädä mietitettymään jonon määritelmä, eli mikä on n:nän reaaliluvun jono. Formaalisti jono* $(x_1, \dots, x_n)$ *voidaan määritellä kuvauksena* $x \colon \{1, \dots, n\} \to \mathbb{R}$*, josta käytetään merkintää* $x = (x_1, \dots, x_n)$*, missä* $x_i = x(i)$ *kaikilla* $i \in \{1, \dots n\}$.

Joukon $\mathbb{R}^n$ alkiota kutsutaan vektoreiksi ja joukkoa $\mathbb{R}^n$ kutsutaan tässä yhteydessä avaruudeksi (tai tarkemmin vektoriavaruudeksi), koska siinä on määritelty kaksi laskutoimitusta: yhteenlasku $+ \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ ja skalaarikertolasku $\cdot \colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ kaavoilla

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

ja

$$a(x_1, \dots, x_n) = (ax_1, \dots, ax_n),$$

missä $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$ ja $a \in \mathbb{R}$. Huomaa, että avaruuden $\mathbb{R}^n$ vektoreiden $(x_1, \dots, x_n)$ ja $(y_1, \dots, y_n)$ tuloa ei ole määritelty.

Näille laskutoimituksille pätee seuraavat ominaisuudet, jotka otetaan annettuina. Olkoot $x, y, z \in \mathbb{R}^n$ ja $a, b \in \mathbb{R}$. Tällöin

- $(x+y)+z = x+(y+z)$,
- $x+y = y+x$,
- $a(x+y) = ax+ay$,
- $(ab)x = a(bx)$ ja
- $(a+b)x = ax+bx$.

Vektorilla $0 = (0,\dots,0) \in \mathbb{R}^n$ ja luvulla $1 \in \mathbb{R}$ on ominaisuudet, että $x+0 = x$ ja $1x = x$ kaikilla $x \in \mathbb{R}^n$. Lisäksi jokaisella $x \in \mathbb{R}^n$ vektoria $(-1)x$ merkitään lyhyesti $-x$. Näillä merkinnöillä saadaan jokaisella $x \in \mathbb{R}^n$, että $x - x = x + (-1)x = (1 + (-1))x = 0$. Vektoria $-x$ kutsutaan vektorin x *vastavektoriksi.*

Huomautus 1.2.8. *Nämä laskutoimitusten + ja · ominaisuudet tekevät joukosta* $\mathbb{R}^n$ *näillä laskutoimituksilla varustettuna ns. vektoriavaruuden. Tähän palataan kurssilla Lineaarialgebra ja matriisilaskenta II.*

Yhtälöryhmän ratkaisun antaminen parametrimuodossa

Avaruuden $\mathbb{R}^n$ laskutoimituksia voidaan käyttää yhtälöryhmän ratkaisujoukon esittämiseen geometrisesti. Tarkastellaan ensin esimerkin avulla.

Esimerkki 1.2.9. *Tarkastellaan uudelleen esimerkin 1.2.2 yhtälöryhmän*

$$\begin{cases} x_1 & +2x_2 & -x_3 & = -2 \\ & -x_2 & +x_3 & = 3 \end{cases} \tag{1.8}$$

ratkaisua $(x_1, x_2, x_3) \in \mathbb{R}^3$*. Jälkimmäisen yhtälön nojalla*

$$x_2 = x_3 - 3.$$

Sijoittamalla tämä ensimmäiseen yhtälöön saadaan

$$x_1 = -2x_2 + x_3 - 2 = -2(x_3 - 3) + x_3 - 2 = -2x_3 + 6 + x_3 - 2 = -x_3 + 4.$$

Näin ollen ratkaisuvektorin (x_1, x_2, x_3) *kertoimet* x_1 *ja* x_2 *voidaan esittää kertoimen* x_3 *avulla. Käyttäen avaruuden* $\mathbb{R}^3$ *vektorilaskutoimituksia havaitaan, että*

$$(x_1, x_2, x_3) = (-x_3+4, x_3-3, x_3) = (4,-3,0) + (-x_3, x_3, x_3) = (4,-3,0) + x_3(-1,1,1).$$

Näin on havaittu, että yhtälöryhmän (1.8) *ratkaisut sisältyvät joukkoon*

$$\{(4,-3,0) + x_3(-1,1,1) \in \mathbb{R}^3 \colon x_3 \in \mathbb{R}\}, \tag{1.9}$$

joka geometrisesti tulkittuna on avaruuden $\mathbb{R}^3$ *suora. Nyt on helppo tarkistaa, että jokainen vektori* $(4,-3,0) + x_3(-1,1,1)$ *on yhtälöryhmän* (1.8) *ratkaisu jokaisella* $x_3 \in \mathbb{R}$*. Näin ollen joukko* (1.9) *on yhtälöryhmän* (1.8) *ratkaisujoukko.*

1.2.1 Yhtälöt $0 = 0$ ja $0 = b$ yhtälöryhmän riveinä

Palataan nyt tarkastelmaan yhtälöryhmää (1.7). Määritelmässä 1.2.1 sallitaan, että jollain yhtälöryhmän rivillä kaikki kertoimet ovat nollia, eli että jokin yhtälöryhmän (1.7) yhtälöistä

$$a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$$

onkin muotoa

$$0x_1 + \cdots + 0x_n = b_j. \tag{1.10}$$

Tämä vaikuttaa tarpeettomalta yleisyydeltä, mutta kuten seuraavassa luvussa huomataan, tällä on tärkeä rooli yhtälöryhmien ratkaisemisessa.

Koska luku $b_j \in \mathbb{R}$ on määrätty yhtälöryhmää (1.7) määriteltäessä, yhtälöllä (1.10) kaksi mahdollista seurausta.

Tapaus $b_j = 0$

Jos $b_j = 0$, eli yhtälö on

$$0x_1 + \cdots + 0x_n = 0, \tag{1.11}$$

niin tällöin yhtälön ratkaisuksi kelpaavat kaikki joukon $\mathbb{R}^n$ vektorit, eli

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon 0x_1 + \cdots + 0x_n = 0\} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon 0 = 0\} = \mathbb{R}^n.$$

Tässä tapauksessa yhtälöryhmän (1.7) ratkaisujoukko on sama kuin yhtälöryhmällä, josta yhtälö $0x_1 + \cdots + 0x_n = 0$ on jätetty pois, eli alkuperäisen yhtälöryhmän ratkaisut ovat samat kuin yhtälöryhmällä

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ \vdots & \vdots \\ a_{(j-1)1}x_1 + a_{(j-1)2}x_2 + \cdots + a_{(j-1)n}x_n & = b_{j-1} \\ a_{(j+1)1}x_1 + a_{(j+1)2}x_2 + \cdots + a_{(j+1)n}x_n & = b_{(j+1)} \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases} \tag{1.12}$$

Perustellaan tämä väite lyhyesti. Merkitään

$$R_k = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k\}$$

jokaisella $k \in \{1, \ldots, m\}$. Tällöin yhtälöryhmän (1.7) ratkaisujoukko on $R_1 \cap \cdots \cap R_m$. Toisaalta yhtälöryhmän (1.12) ratkaisujoukko on $R_1 \cap \cdots R_{j-1} \cap R_{j+1} \cap \cdots R_m$. Koska

$$\begin{aligned} R_1 \cap \cdots \cap R_m &= R_1 \cap \cdots \cap R_{j-1} \cap R_j \cap R_{j+1} \cap \cdots \cap R_m \\ &= R_1 \cap \cdots \cap R_{j-1} \cap \mathbb{R}^n \cap R_{j+1} \cap \cdots \cap R_m \\ &= R_1 \cap \cdots \cap R_{j-1} \cap R_{j+1} \cap \cdots \cap R_m, \end{aligned}$$

niin yhtälöryhmillä (1.7) ja (1.12) on samat ratkaisujoukot.

Tapaus $b_j \neq 0$

Jos $b_j \neq 0$, mutta kertoimet $a_1, \ldots, a_n$ ovat nollia, niin tällöin yhtälöllä (1.10) ei ole ratkaisuja, eli

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon 0x_1 + \cdots + 0x_n = b_j\} = \emptyset.$$

Tällöin ei myöskään yhtälöryhmällä (1.7) ole ratkaisuja. Perustellaan myös tämä havainto lyhyesti. Edellisen kohdan merkintöjä käyttäen havaitaan, että

$$\begin{aligned} R_1 \cap \cdots \cap R_m &= R_1 \cap \cdots \cap R_{j-1} \cap R_j \cap R_{j+1} \cap \cdots \cap R_m \\ &= R_1 \cap \cdots \cap R_{j-1} \cap \emptyset \cap R_{j+1} \cap \cdots \cap R_m = \emptyset, \end{aligned}$$

eli yhtälöryhmällä (1.7) ei ole ratkaisuja.

1.3 Rivioperaatiot

Tarkastellaan nyt yhtälöryhmän konkreettista ratkaisemista ns. rivioperaatioiden avulla. Rivioperaatioiden ideana on sieventää yhtälöryhmän yhtälöitä sellaiseen muotoon, että ratkaisut voidaan helposti lukea yhtälöryhmästä. Näissä operaatiossa yhtälöryhmän ratkaisut eivät tietenkään saa muuttua.

Vakiolla kertominen

Ensimmäinen rivioperaatio on yksittäisen yhtälöryhmän rivin kertominen nollasta poikkeavalla vakiolla. Tarkastellaan merkintöjen yksinkertaistamiseksi tätä varten yhtä yksittäistä yhtälöä

$$a_1x_1 + \cdots + a_nx_n = b \tag{1.13}$$

missä $a_1, \ldots, a_n \in \mathbb{R}$ ja $b \in \mathbb{R}$. Selvästi, jos yhtälön (1.13) kertoo molemmilta puolilta vakiolla $\lambda \neq 0$, niin saadaan yhtälö

$$\lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b \tag{1.14}$$

jolla on samat ratkaisut, koska yhtälö (1.14) voidaan palauttaa takaisin alkuperäiseen kertomalla vakiolla $1/\lambda$. Kirjataan tämä nyt tarkasti.

Lemma 1.3.1. *Olkoot $a_1, \ldots, a_n \in \mathbb{R}$ ja $b \in \mathbb{R}$. Tällöin jokaisella $\lambda \neq 0$ yhtälöillä*

$$a_1x_1 + \cdots + a_nx_n = b$$

ja

$$\lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b$$

on samat ratkaisujoukot eli

$$\begin{aligned} &\{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon a_1x_1 + \cdots + a_nx_n = b\} \\ &= \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon \lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b\}. \end{aligned}$$

Todistus. Osoitetaan ratkaisujoukot samoiksi. Merkitään selvyyden vuoksi

$$R = \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon a_1x_1 + \cdots + a_nx_n = b\}$$

ja

$$R_\lambda = \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon \lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b\}.$$

Osoitetaan, että $R \subset R_\lambda$ ja että $R_\lambda \subset R$. Näistä seuraa suoraan joukkojen R ja R_λ yhtäsuuruus, eli $R = R_\lambda$, ja väite on tämän jälkeen todistettu.

Olkoon $x \in R$. Tällöin vektorille $x = (x_1, \dots, x_n)$ pätee

$$a_1x_1 + \cdots + a_nx_n = b.$$

Näin ollen

$$\lambda b = \lambda \left(a_1x_1 + \cdots + a_nx_n\right) = \lambda a_1x_1 + \cdots \lambda a_nx_n.$$

Koska vektori x toteuttaa joukon R_λ määrittelevän ehdon, niin $x \in R_\lambda$. Näin on osoitettu, että $R \subset R_\lambda$.

Osoitetaan nyt, että $R_\lambda \subset R$. Olkoon $x = (x_1, \dots, x_n) \in R_\lambda$. Tällöin

$$\lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b.$$

Koska $\lambda \neq 0$, niin yhtälön molemmat puolet voidaan kertoa luvulla $1/\lambda$, jolloin saadaan

$$\begin{aligned} b &= (1/\lambda)\lambda b = (1/\lambda)\left(\lambda a_1x_1 + \cdots + \lambda a_nx_n\right) \\ &= (1/\lambda)\lambda a_1x_1 + \cdots + (1/\lambda)\lambda a_nx_n \\ &= a_1x_1 + \cdots + a_nx_n. \end{aligned}$$

Näin ollen x toteuttaa joukon R määrittelevän ehdon, eli $x \in R$. Näin on osoitettu $R_\lambda \subset R$. □

Rivien yhteenlasku

Toista rivioperaatiota varten tarkastellaan yhtälöparia, eli kahden yhtälön ryhmää

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \end{cases} \tag{1.15}$$

Nyt yhtälöparin ensimmäinen rivi voidaan laskea puolittain yhteen yhtälöparin toisen rivin kanssa yhtälöryhmän ratkaisujen siitä muuttumatta, eli yhtälöparilla

$$\begin{cases} a_{11}x_1 + & \cdots + & a_{1n}x_n &= b_1 \\ (a_{21} + a_{11})x_1 + & \cdots + & (a_{2n} + a_{1n})x_n &= b_2 + b_1 \end{cases} \tag{1.16}$$

on samat ratkaisut kuin yhtälöparilla (1.15). Heuristisesti syy tähän on se, että ensimmäisen rivin lisääminen toiseen riviin voidaan kumota vähentämällä ensimmäinen rivi yhtälöryhmän toisesta rivistä. Kirjataan tämä havainto nyt lemmaksi. Koska lemman todistus on hyvin samankaltainen kuin lemman 1.3.1 todistus, se jätetään lukijalle harjoitustehtäväksi.

Lemma 1.3.2. *Olkoon $n \in \mathbb{Z}_+$, $a_{ji} \in \mathbb{R}$ kaikilla $j \in \{1,2\}$ ja $i \in \{1,\dots,n\}$, sekä $b_1, b_2 \in \mathbb{R}$. Tällöin yhtälöryhmillä*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \end{cases}$$

ja

$$\begin{cases} a_{11}x_1 + & \cdots + & a_{1n}x_n &= b_1 \\ (a_{21}+a_{11})x_1 + & \cdots + & (a_{2n}+a_{1n})x_n &= b_2 + b_1 \end{cases}$$

on samat ratkaisujoukot.

Rivien järjestyksen vaihtaminen

Viimeinen rivioperaatio on yhtälöryhmän rivien järjestyksen vaihtaminen. Tarkastellaan nyt yleistä yhtälöryhmää

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases} \tag{1.17}$$

kuten määritelmässä 1.2.1, missä $a_{ji} \in \mathbb{R}$ ja $b_j \in \mathbb{R}$ kaikilla $j \in \{1,\dots,m\}$ ja $i \in \{1,\dots,n\}$.

Merkitään jälleen joukolla R_j tämän yhtälöryhmän j:nnen rivin yhtälön ratkaisuja, eli

$$R_j = \{(x_1,\dots,x_n) \in \mathbb{R}^n \in a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j\}$$

jokaisella $j \in \{1,\dots,n\}$.

Yhtälöryhmän (1.17) ratkaisujoukko on näillä merkinnöillä

$$R_1 \cap R_2 \cap \cdots \cap R_m. \tag{1.18}$$

Jos yhtälöryhmän rivien paikkaa vaihdetaan, niin joukot $R_1,\dots,R_m$ eivät muutu, ainoastaan niiden järjestys vaihtuu. Koska joukkojen järjestys ei vaikuta niiden leikkaukseen, niin joukko (1.18) ei muutu rivien järjestystä muutettaessa. Näin ollen yhtälöryhmän ratkaisut eivät riipu yhtälöryhmän rivien järjestyksestä.

Yllä kuvattu todistus voidaan kirjoittaa myös tarkemmin. Tällöin tulee kuitenkin ensin kuvata, mitä tarkoittaa rivien järjestyksen vaihtaminen ja osoittaa, että leikkausjoukko ei muutu joukkojen $R_1,\dots,R_m$ järjestystä vaihdettaessa.

Tämän intuitiivisen havainnon muotoilu tarkasti on hieman työlästä antamatta varsinaisesti uutta ymmärrystä asiasta, joten sen voi sivuuttaa ensimmäisillä lukukerroilla. Kirjataan tulos kuitenkin täydellisyyden vuoksi lemmaksi. Se on helpointa muotoilla permutaation käsitettä käyttäen seuraavasti.

Joukon $\{1,\dots,m\}$ bijektiota $\sigma\colon \{1,\dots,m\} \to \{1,\dots,m\}$ kutsutaan *permutaatioksi.*

Lemma 1.3.3. *Olkoot $m, n \in \mathbb{Z}_+$ sekä $a_{ji} \in \mathbb{R}$ ja $b_j \in \mathbb{R}$ kaikilla $j \in \{1, \dots, m\}$ ja $i \in \{1, \dots, n\}$. Tällöin jokaisella permutaatiolla $\sigma\colon \{1, \dots, m\} \to \{1, \dots, m\}$ yhtälöryhmillä*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases} \tag{1.19}$$

ja

$$\begin{cases} a_{\sigma(1)1}x_1 + a_{\sigma(1)2}x_2 + \cdots + a_{\sigma(1)n}x_n & = b_{\sigma(1)} \\ a_{\sigma(2)1}x_1 + a_{\sigma(2)2}x_2 + \cdots + a_{\sigma(2)n}x_n & = b_{\sigma(2)} \\ \vdots & \vdots \\ a_{\sigma(m)1}x_1 + a_{\sigma(m)2}x_2 + \cdots + a_{\sigma(m)n}x_n & = b_{\sigma(m)} \end{cases} \tag{1.20}$$

on samat ratkaisujoukot.

Todistus. Olkoon jälleen

$$R_j = \{(x_1, \dots, x_n) \in \mathbb{R}^n \in a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j\}$$

jokaisella $j \in \{1, \dots, n\}$. Tällöin yhtälöryhmän (1.19) ratkaisujoukko on

$$R = R_1 \cap R_2 \cap \cdots \cap R_m.$$

Toisaalta yhtälöryhmän (1.20) ratkaisujoukko on

$$R_\sigma = R_{\sigma(1)} \cap R_{\sigma(2)} \cap \cdots \cap R_{\sigma(m)}.$$

Osoitetaan, että $R = R_\sigma$.

Oletetaan, että $x \in R$. Osoitetaan ensin, että $R_{\sigma(j)}$ jokaisella $j \in \{1, \dots, m\}$. Olkoon $j \in \{1, \dots, m\}$. Tällöin $\sigma(j) \in \{1, \dots, m\}$. Koska $x \in R$, niin leikkausjoukon määritelmän nojalla $x \in R_k$ jokaisella $k \in \{1, \dots, m\}$. Erityisesti $x \in R_{\sigma(j)}$. Näin ollen $x \in R_{\sigma(j)}$ jokaisella $j \in \{1, \dots, m\}$. Näin ollen $x \in R_{\sigma(1)} \cap \cdots \cap R_{\sigma(m)} = R_\sigma$.

Oletetaan nyt, että $x \in R_\sigma$. Osoitetaan, että $\tilde{x} \in R_j$ jokaisella $j \in \{1, \dots, m\}$. Olkoon $j \in \{1, \dots, m\}$. Koska σ on bijektio, niin on olemassa sellainen $k \in \{1, \dots, m\}$, että $\sigma(k) = j$. Koska $x \in R_\sigma$, niin $x \in R_{\sigma(k)}$. Näin ollen $x \in R_j$. Koska $x \in R_j$ jokaisella $j \in \{1, \dots, m\}$, niin $x \in R$. □

Huomautus 1.3.4. *Tarkkaavainen lukija jo luultavasti huomasikin, että edellisellä todistuksessa todistettiin itseasiassa, että leikkausjoukko ei riipu leikkauksen jäsenien järjestyksestä.*

1.4 Yhtälöryhmän ratkaiseminen rivioperaatioilla

Kootaan nyt lemmojen 1.3.1, 1.3.2 ja 1.3.3 tulokset yhteen yhdeksi ratkaisumetodiksi:

- Yhtälöryhmän rivejä voi kertoa nollasta eroavalla vakiolla ratkaisujen muuttumatta.
- Yhtälöryhmän rivejä voi laskea yhteen ratkaisujen muuttumatta.
- Yhtälöryhmän rivien paikkaa voi vaihtaa ratkaisujen muuttumatta.

Ensimmäisestä ja toisesta operaatiosta saadaan suoraan:

- Yhtälöryhmän rivin voi kertoa vakiolla ja lisätä toiseen riviin yhtälöryhmän ratkaisujen muuttumatta.

Yhtälöryhmiä, joilla on samat ratkaisut kutsutaan *ekvivalenteiksi*. Yhtälöryhmän ratkaisemisella tarkoitetaan jatkossa sellaisen ekvivalentin yhtälöryhmän löytämistä, jonka ratkaisut voidaan helposti lukea. Tarkastellaan tätä ensin konkreettisen esimerkin avulla.

Esimerkki 1.4.1. *Olkoon*

$$\begin{cases} x_1 & +4x_2 & -2x_3 & = 5 \\ 2x_1 & -x_2 & +x_3 & = 1 \\ -x_1 & -2x_2 & & = 0 \end{cases}$$

Ensimmäinen rivi voidaan kertoa luvulla -2 ja sitten lisätä toiseen riviin. Samalla voidaan ensimmäinen rivi lisätä kolmanteen riviin. Tällöin saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} x_1 & +4x_2 & -2x_3 & = 5 \\ (2-2)x_1 & (-1-8)x_2 & +(1+4)x_3 & = 1-10 \\ (-1+1)x_1 & +(-2+4)x_2 & -2x_3 & = 0+5 \end{cases}$$

eli yhtälöryhmä

$$\begin{cases} x_1 & +4x_2 & -2x_3 & = 5 \\ & -9x_2 & +5x_3 & = -9 \\ & +2x_2 & -2x_3 & = 5 \end{cases}$$

Näiden operaatioiden seurauksena havaitaan, että muuttuja x_1 riippuu siis muuttujista x_2 ja x_3. Tämä kuvailee ratkaisumetodin tavoitteen: tavoitteena on pyrkiä tilanteeseen, jossa mahdollisimman moni muuttuja esiintyy vain yhdellä rivillä.

Jatketaan yhtälöryhmän sieventämistä jakamalla ensin yhtälöryhmän viimeinen rivi vakiolla 2, jolloin saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} x_1 & +4x_2 & -2x_3 & = 5 \\ & -9x_2 & +5x_3 & = -9 \\ & x_2 & -x_3 & = 5/2 \end{cases}$$

Kerrotaan nyt viimeinen rivi vakiolla -4 ja lisätään ensimmäiseen sekä samanaikaisesti kerrotaan se vakiolla 9 ja lisätään toiseen riviin. Tällöin saadaan yhtälöryhmä

$$\begin{cases} x_1 & & +2x_3 & = -5 \\ & & -4x_3 & = 27/2 \\ & x_2 & -x_3 & = 5/2 \end{cases}$$

Näin ollen toisesta yhtälöstä voidaan lukea, että $x_3 = -27/8$. Nyt voidaan edetä kahta reittiä. Toisaalta voidaan suoraan ratkaista muuttujien x_2 ja x_1 arvot yhtälöistä 1 ja 3: arvot ovat $x_2 = 5/2 + x_3 = 5/2 + (-27/8) = -7/8$ ja $x_1 = -5 - 2x_3 = -5 - 2(-27/8) = -5 + 27/4 = 7/4$.

Sama päättely voidaan tehdä systemaattisemmin hyödyntämällä rivioperaatioita. Jaetaan ensin toinen yhtälö luvulla -4 ja vaihdetaan rivien 2 ja 3 paikkaa. Tällöin saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} x_1 & & +2x_3 & = & -5 \\ & x_2 & -x_3 & = & 5/2 \\ & & x_3 & = & -27/8 \end{cases}$$

Lisäämällä viimeinen rivi riviin 2 ja vähentämällä se rivistä 1 kahdesti saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} x_1 & & & = & 7/4 \\ & x_2 & & = & -7/8 \\ & & x_3 & = & -27/8 \end{cases}$$

Tämä päättää yhtälöryhmän ratkaisemisen.

Esimerkissä 1.4.1 yhtälöryhmällä oli yksikäsitteinen ratkaisu. Tarkastellaan nyt esimerkkejä, joissa yhtälöryhmällä ei ole ratkaisua tai äärettömän monta ratkaisua. Aloitetaan jälkimmäisestä.

Esimerkki 1.4.2. *Olkoon*

$$\begin{cases} 2x_1 & +4x_2 & -2x_3 & = 4 \\ 2x_1 & -x_2 & +x_3 & = 1 \\ -x_1 & -2x_2 & +x_3 & = -2 \end{cases}$$

yhtälöryhmä. Lisäämällä viimeinen rivi kahdesti ensimmäiseen saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} 0x_1 & +0x_2 & +0x_3 & = & 0 \\ 2x_1 & -x_2 & +x_3 & = & 1 \\ -x_1 & -2x_2 & +x_3 & = & -2 \end{cases}$$

eli yhtälöryhmä

$$\begin{cases} & & 0 & = & 0 \\ 2x_1 & -x_2 & +x_3 & = & 1 \\ -x_1 & -2x_2 & +x_3 & = & -2 \end{cases}$$

Yhtälöryhmän ensimmäinen rivi voidaan siis jättää pois ratkaisujen siitä muuttumatta, joten voidaan siirtyä tarkastelemaan ekvivalenttia yhtälöryhmää

$$\begin{cases} 2x_1 & -x_2 & +x_3 & = & 1 \\ -x_1 & -2x_2 & +x_3 & = & -2 \end{cases}$$

Lisäämäällä jälkimmäinen rivi ensimmäiseen kahdesti, kertomalla toinen rivi vakiolla (-1) *ja vaihtamalla rivien järjestys saadaan yhtälöryhmä*

$$\begin{cases} x_1 & +2x_2 & -x_3 & = & 2 \\ & -5x_2 & +3x_3 & = & -3 \end{cases}$$

joka voidaan sieventää muotoon

$$\begin{cases} x_1 & +2x_2 & -x_3 & = & 2 \\ & x_2 & -(3/5)x_3 & = & 3/5 \end{cases}$$

Vähentämällä jälkimmäinen rivi ensimmäisestä kahdesti saadaan

$$\begin{cases} x_1 & & +(1/5)x_3 & = & 4/5 \\ & x_2 & -(3/5)x_3 & = & 3/5 \end{cases}$$

Tässä tilanteessa huomataan, että muuttujat x_1 *ja* x_2 *riippuvat muuttujasta* x_3 *yhtälöillä* $x_1 = 4/5 - (1/5)x_3$ *ja* $x_2 = 3/5 + (3/5)x_3$*, mutta muuttujalle* x_3 *ei ole ehtoja, eli* $x_3 \in \mathbb{R}$*. Näin ollen yhtälöryhmän ratkaisujoukko on*

$$\begin{aligned} &\{(x_1, x_2, x_3) \in \mathbb{R}^3 \colon x_1 = 4/5 - (1/5)x_3,\ x_2 = 3/5 + (3/5)x_3,\ x_3 \in \mathbb{R}\} \\ &= \{(4/5 - x_3/5, 3/5 + 3x_3/5, x_3) \in \mathbb{R}^3 \colon x_3 \in \mathbb{R}\}. \end{aligned}$$

Ratkaisun havainnollistamiseksi on kuitenkin mielekkäänpää kirjoittaa ratkaisut, ns. parametrimuodossa eli käyttäen jälleen avaruuden $\mathbb{R}^3$ *laskutoimituksia. Olkoon* $x_3 \in \mathbb{R}$*. Tällöin*

$$\begin{aligned} \left(\frac{4}{5} - \frac{x_3}{5}, \frac{3}{5} + \frac{3}{5}x_3, x_3\right) &= \left(\frac{4}{5}, \frac{3}{5}, 0\right) + \left(-\frac{x_3}{5}, \frac{3}{5}x_3, x_3\right) \\ &= \left(\frac{4}{5}, \frac{3}{5}, 0\right) + x_3\left(-\frac{1}{5}, \frac{3}{5}, 1\right). \end{aligned}$$

Näin ollen yhtälöryhmän ratkaisujoukko on

$$\{(4/5, 3/5, 0) + x_3(-1/5, 3/5, 1) \in \mathbb{R}^3 \colon x_3 \in \mathbb{R}\}.$$

Koska vapaa parametri x_3 *ei enää riippu ratkaisujen kolmanteen koordinaattiin, niin on tyypillistä vaihtaa se toiseen symboliin, esimerkiksi* t*. Ratkaisujoukko on siis*

$$\{(4/5, 3/5, 0) + t(-1/5, 3/5, 1) \in \mathbb{R}^3 \colon t \in \mathbb{R}\}.$$

Tässä esityksessä on se etu, että siitä nähdään suoraan, että ratkaisujoukko on itseasiassa vektorin $(-1/5, 3/5, 1)$ *suuntainen suora, joka kulkee pisteen* $(4/5, 3/5, 0)$ *kautta.*

Tarkastellaan nyt esimerkkiä, jossa yhtälöryhmällä ei ole ratkaisuja.

Esimerkki 1.4.3. *Olkoon*

$$\begin{cases} x_1 & & +x_3 & = 5 \\ x_1 & +x_2 & +2x_3 & = 1 \\ & x_2 & +x_3 & = 0 \end{cases}$$

yhtälöryhmä. Vähentämällä viimeinen rivi toisesta rivistä saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} x_1 & & +x_3 & = 5 \\ x_1 & & +x_3 & = 1 \\ & x_2 & +x_3 & = 0 \end{cases}$$

Selvästi kahdella ensimmäisellä yhtälöllä ei ole yhteisiä ratkaisuja. Formaalisti tämä voidaan todeta vähentämällä toinen yhtälö ensimmäisestä, jolloin saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} & & 0 & = 4 \\ x_1 & & +x_3 & = 1 \\ & x_2 & +x_3 & = 0 \end{cases}$$

jolla ei ole ratkaisuja.

Rivioperaatioiden merkitseminen laskuihin

Käytännön laskuissa on hyvä merkitä näkyviin käytetyt rivioperaatiot. Laskuissa voi käyttää esimerkiksi symbolia $\sim$ merkitsemään ekvivalentteja yhtälöryhmiä. Seuraavassa esimerkissä käydään läpi esimerkin 1.4.1 lasku lyhennetyin merkinnöin.

Esimerkki 1.4.4 (Esimerkin 1.4.1 toisinto)**.**

$$\begin{cases} x_1 & +4x_2 & -2x_3 & = 5 & | \\ 2x_1 & -x_2 & +x_3 & = 1 & |-2R_1 \\ -x_1 & -2x_2 & & = 0 & |+R_1 \end{cases} \sim \begin{cases} x_1 & +4x_2 & -2x_3 & = 5 & \\ & -9x_2 & +5x_3 & = -9 & \\ & +2x_2 & -2x_3 & = 5 & |\cdot(1/2) \end{cases}$$

$$\sim \begin{cases} x_1 & +4x_2 & -2x_3 & = 5 & |-4R_3 \\ & -9x_2 & +5x_3 & = -9 & |+9R_3 \\ & x_2 & -x_3 & = 5/2 & \end{cases}$$

$$\sim \begin{cases} x_1 & & +2x_3 & = -5 & \\ & & -4x_3 & = 27/2 & |\cdot(-1/4) \\ & x_2 & -x_3 & = 5/2 & \end{cases}$$

$$\underset{R_2\leftrightarrow R_3}{\sim} \begin{cases} x_1 & & +2x_3 & = & -5 & |-2R_3 \\ & x_2 & -x_3 & = & 5/2 & |+R_3 \\ & & x_3 & = & -27/8 & \end{cases}$$

$$\sim \begin{cases} x_1 & & & = & 7/4 \\ & x_2 & & = & -7/8 \\ & & x_3 & = & -27/8 \end{cases}$$

Tarkastellaan vielä yhtä uutta esimerkkiä.

Esimerkki 1.4.5. *Etsitään yhtälöryhmän*

$$\begin{cases} x_1 & -x_2 & +2x_3 & -x_4 & = & 3 \\ x_1 & +x_2 & & +x_4 & = & 3 \end{cases}$$

ratkaisut $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$.

Käyttäen samoja merkintöjä kuin esimerkissä 1.4.4 saadaan, että

$$\begin{aligned}
\begin{cases} x_1 & -x_2 & +2x_3 & -x_4 & = & 3 \\ x_1 & +x_2 & & +x_4 & = & 3 \quad | - R_1 \end{cases}
&\sim \begin{cases} x_1 & -x_2 & +2x_3 & -x_4 & = & 3 \\ & +2x_2 & -2x_3 & +2x_4 & = & 0 \quad | \cdot (1/2) \end{cases} \\
&\sim \begin{cases} x_1 & -x_2 & +2x_3 & -x_4 & = & 3 \quad | + R_2 \\ & x_2 & -x_3 & +x_4 & = & 0 \end{cases} \\
&\sim \begin{cases} x_1 & & +x_3 & & = & 3 \\ & x_2 & -x_3 & +x_4 & = & 0 \end{cases}
\end{aligned}$$

Tällä kertaa havaitaan, että muuttujat x_1 *ja* x_3 *riippuvat toisistaan ja että muuttujat* x_2*,* x_3 *ja* x_4 *riippuvat toisistaan. Tarkemmin sanottuna, yhtälöryhmän ratkaisun* (x_1, x_2, x_3, x_4) *koordinaatit toteuttavat yhtälöt* $x_1 = -x_3 + 3$ *ja* $x_2 = x_3 - x_4$*. Näin ollen ratkaisujoukko on*

$$\begin{aligned}
&\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \colon x_1 = -x_3 + 3,\ x_2 = x_3 - x_4,\ x_3 \in \mathbb{R},\ x_4 \in \mathbb{R}\} \\
&= \{(-x_3 + 3, x_3 - x_4, x_3, x_4) \in \mathbb{R}^4 \colon x_1 = -x_3 + 3,\ x_2 = x_3 - x_4,\ x_3 \in \mathbb{R},\ x_4 \in \mathbb{R}\}.
\end{aligned}$$

Ratkaisujoukko voidaan kirjoittaa parametrimuodossa seuraavasti. Olkoot $x_3, x_4 \in \mathbb{R}$*. Tällöin*

$$\begin{aligned}
(-x_3 + 3, x_3 - x_4, x_3, x_4) &= (3, 0, 0, 0) + (-x_3, x_3 - x_4, x_3, x_4) \\
&= (3, 0, 0, 0) + (-x_3, x_3, x_3, 0) + (0, -x_4, 0, x_4) \\
&= (3, 0, 0, 0) + x_3(-1, 1, 1, 0) + x_4(0, -1, 0, 1).
\end{aligned}$$

Näin ollen ratkaisujoukko on itseasiassa joukko

$$\{(3, 0, 0, 0) + x_3(-1, 1, 1, 0) + x_4(0, -1, 0, 1) \in \mathbb{R}^4 \colon x_3, x_4 \in \mathbb{R}\}.$$

jossa parametrien x_3 *ja* x_4 *nimet voidaan vaihtaa symboleihin* t *ja* s*, eli ratkaisujoukko on*

$$\{(3, 0, 0, 0) + t(-1, 1, 1, 0) + s(0, -1, 0, 1) \in \mathbb{R}^4 \colon t, s \in \mathbb{R}\}.$$

Geometrisesti ratkaisujoukko on siis taso, joka sisältää vektoreiden $(-1, 1, 1, 0)$ *ja* $(0, -1, 0, 1)$ *suuntaiset suorat ja joka kulkee pisteen* $(3, 0, 0, 0)$ *kautta.*

1.4.1 Yhtälöryhmän porrasmuoto; Gaussin ja Jordanin menetelmä

Edellä kuvatuissa esimerkeissä herää kysymys, että milloin yhtälöryhmän muokkaaminen kannattaa lopettaa, eli milloin rivioperaatiot eivät enää sievennä yhtälöryhmää lisää. Emme vastaa tähän kysymykseen formaalisti, vaan kiertäen esittelemällä yhtälöryhmän supistetun porrasmuodon, joka on yhtälöryhmän muokkaamisen varsinainen tavoite.

Määritelmä 1.4.6. *Yhtälöryhmä*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{1n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n & = b_m \end{cases}$$

on supistetussa porrasmuodossa, *jos*

1. *rivillä, joka ei ole nollarivi ensimmäinen nollasta poikkeava kerroin on* 1 *ja muut saman muuttujan kertoimet ovat nollia.*

2. *nollarivin alla ei ole nollasta poikkeavia rivejä ja*

3. *allekkaisilla riveillä, jotka eivät ole molemmat nollarivejä, ensimmäinen nollasta poikkeava kerroin tulee jälkimmäisellä rivillä myöhemmin,*

eli toisin sanoen kaikilla $j \in \{1, \dots, m\}$ ja $i \in \{1, \dots, n\}$ pätee

1. *jos $a_{ji} \neq 0$ ja $a_{j\ell} = 0$ kaikilla $\ell < i$, niin $a_{ji} = 1$ ja $a_{ri} = 0$ kaikilla $r \neq j$,*

2. *jos $a_{jp} = 0$ kaikilla $p = 1, \dots, n$, niin $a_{rp} = 0$ kaikilla $r > j$ ja $p \in \{1, \dots, n\}$ ja*

3. *jos $a_{ji} = 1$ ja $a_{j\ell} = 0$ kaikilla $1 \leq \ell < i$, niin $a_{(j+1)\ell} = 0$ kaikilla $1 \leq \ell \leq i$.*

Supistetun porrasmuodon yhteydessä käytetään yleensä seuraavia termejä.

Määritelmä 1.4.7. *Ehdossa* (3) *kerrointa a_{ji} kutsutaan* rivin j johtavaksi kertoimeksi *ja vastavaa muuttujaa x_i kutsutaan* sidotuksi muuttujaksi. *Niitä muuttujia x_ℓ, joiden kertoimet $a_{j\ell}$ eivät ole minkään rivin johtavia kertoimia kutsutaan* vapaiksi muuttujiksi.

Sidottujen ja vapaiden muuttujien terminologiassa sidottu muuttuja riippuu siis suuremmalla indeksillä varustetuista vapaista muuttujista.

Esimerkki 1.4.8. *Yhtälöryhmä*

$$\begin{cases} x_1 & +3x_2 & & +2x_4 & = 5 \\ & & x_3 & -3x_4 & = 2 \end{cases}$$

on supistetussa porrasmuodossa ja valitun terminologian mukaisesti muuttujat x_1 ja x_3 ovat sidottuja sekä muuttujat x_2 ja x_4 vapaita. Tässä tilanteessa, mikäli (x_1, x_2, x_3, x_4) on tämän yhtälöryhmän ratkaisu, niin tällöin lukujen x_1, x_2 ja x_4 arvot riippuvat toisistaan ja vastaavasti lukujen x_3 ja x_4 riippuvat toisistaan. Toisaalta luvut x_1 ja x_3 määräytyvät yksikäsitteisesti, kun luvut x_2 ja x_4 on valittu, eli

$$\begin{cases} x_1 & = -3t - 2s + 5 \\ x_2 & = t \\ x_3 & = +3s + 2 \\ x_4 & = s \end{cases},$$

missä $t, s \in \mathbb{R}$.

Huomautus 1.4.9. *Edellisessä esimerkissä olisi voitu järjestää, että luvut x_2 ja x_4 olisi ilmaistu lukujen x_1 ja x_3 avulla. Tämä ei kuitenkaan ole yleisen sidottujen ja vapaiden muuttujien terminogian mukaista.*

Supistetussa porrasmuodossa olevalla yhtälöryhmällä ei tarvitse olla ratkaisuja.

Esimerkki 1.4.10. *Yhtälöryhmä*

$$\begin{cases} x_1 & +x_2 & & & = 5 \\ & & x_3 & & = 2 \\ & & & x_4 & = 0 \\ & & & 0 & = 8 \\ & & & 0 & = 0 \end{cases}$$

on supistetussa porrasmuodossa, mutta yhtälön $0 = 8$ *vuoksi sillä ei ole ratkaisuja.*

On helppo löytää esimerkkejä yhtälöryhmistä, jotka eivät ole supistetussa porramuodossa, mutta voidaan helposti siihen saattaa.

Esimerkki 1.4.11. *Mikään yhtälöryhmistä*

$$\begin{cases} 3x_1 & & & & = 5 \\ & x_2 & & +4x_4 & = 0 \\ & & x_3 & & = 2 \end{cases}, \begin{cases} x_1 & & & & = 5 \\ & & & 0 & = 8 \\ & x_2 & & +4x_4 & = 0 \end{cases},$$

$$\begin{cases} & x_2 & & = 0 \\ x_1 & & -x_3 & = 5 \end{cases} \quad ja \quad \begin{cases} x_1 & & +3x_3 & & = 5 \\ & x_2 & & +4x_4 & = 0 \\ & & x_3 & & = 2 \end{cases}$$

ei ole supistetussa porrasmuodossa. Jokainen näistä voidaan kuitenkin saattaa supistettuun porrasmuotoon yhdellä rivioperaatioilla.

Tärkein tulos rivioperaatioihin liityen on, että jokainen yhtälöryhmä voidaan rivioperaatioiden avulla saattaa supistettuun porrasmuotoon. Tätä kutsutaan Gaussin ja Jordanin lauseeksi.

Lause 1.4.12 (Gauss–Jordan)**.** *Jokaisella yhtälöryhmällä on supistettu porrasmuoto.*

Lauseen väite on intuitiivisesti selvä. Todistus on kuitenkin merkinnällisesti haastava. Tämän vuoksi lukija voi sivuuttaa todistuksen tässä vaiheessa.

Huomautus 1.4.13. *Itseasiassa supistettu porrasmuoto on yksikäsitteinen, kun yksitteisyys ymmärretään oikein. Koska nollarivien järjestystä ei ole kiinnitetty, tällä tarkoitetaan sitä, että annettua yhtälöryhmää*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n & = b_m \end{cases}$$

vastaavassa supistetussa porrasmuodossa olevassa yhtälöryhmässä

$$\begin{cases} \tilde{a}_{11}x_1 + \tilde{a}_{12}x_2 + \tilde{a}_{13}x_3 + \cdots + \tilde{a}_{1n}x_n & = \tilde{b}_1 \\ \tilde{a}_{21}x_1 + \tilde{a}_{22}x_2 + \tilde{a}_{23}x_3 + \cdots + \tilde{a}_{2n}x_n & = \tilde{b}_2 \\ \vdots & \vdots \\ \tilde{a}_{m1}x_1 + \tilde{a}_{m2}x_2 + \tilde{a}_{m3}x_3 + \cdots + \tilde{a}_{mn}x_n & = \tilde{b}_m \end{cases}$$

kertoimet $\tilde{a}_{ji}$ on yksikäsitteisesti määrätty. Tämä tulos todistetaan liitteessä A luvun 3 tuloksia hyödyntäen.

Lauseen 1.4.12 todistus. Olkoon

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n & = b_m \end{cases} \tag{1.21}$$

yhtälöryhmä. Osoitetaan, että tämä yhtälöryhmä voidaan viedä supistettuun porrasmuotoon rivioperaatioilla.

Jos yhtälöryhmä koostuu vain nollariveistä, on se supistetussa porrasmuodossa ja todistus voidaan lopettaa. Näin ollen voidaan olettaa, että yhtälöryhmä ei koostu pelkistä nollariveistä. Olkoon $\ell_1 \in \{1, \ldots, n\}$ pienin sellainen indeksi, että kertoimien $a_{1\ell_1}, \ldots, a_{m\ell_1}$ joukossa on nollasta poikkeava kerroin.

Jos näitä nollasta poikkeavia kertoimia on vain yksi, jaetaan rivi tällä kertoimella ja siirretään tämä rivi yhtälöryhmän ensimmäiseksi riviksi. Jos kertoimien $a_{1\ell_1}, \ldots, a_{m\ell_1}$ joukossa on vähintaan kaksi nollasta poikkeavaa kerrointa, niin olkoon $j_1 \in \{1, \ldots, m\}$ pienin sellainen indeksi $j \in \{1, \ldots, m\}$, että $a_{j1} \neq 0$. Oletetaan nyt, että $j > j_i$ on sellainen indeksi, jolla $a_{j\ell_1} \neq 0$. Kerrotaan nyt yhtälöryhmän rivi j_1 vakiolla $a_{j\ell_1}/a_{j_1\ell_1}$ ja vähennetään rivistä j. Mikäli näin saatu rivi j on nollarivi, siirretään se yhtälöryhmän alimmaiseksi. Jaetaan näiden rivioperaatioden jälkeen rivi j_1 luvulla $a_{j_1\ell_1}$ ja siirretään se yhtälöryhmän ensimmäiseksi riviksi. Suorittamalla nämä rivioperaatiot saadaan yhtälöryhmän 1.21 kanssa ekvivalentti yhtälöryhmä

$$\begin{cases} x_{\ell_1} & +a^1_{1(\ell_1+1)}x_{\ell_1+1} & +a^1_{1(\ell_1+2)}x_{\ell_1+2} & +\cdots & +a^1_{1n}x_n & = & b^1_1 \\ & a^1_{2(\ell_1+1)}x_{\ell_1+1} & +a^1_{2(\ell_1+2)}x_{\ell_1+2} & +\cdots & +a^1_{2n}x_n & = & b^1_2 \\ & \vdots & \vdots & & \vdots & & \vdots \\ & a^1_{m(\ell_1+1)}x_{\ell_1+1} & +a^1_{m(\ell_1+2)}x_{\ell_1+2} & +\cdots & +a^1_{mn}x_n & = & b^1_m \end{cases}$$

missä luvut a^1_{ji} ja b^1_j on saatu edellisen yhtälöryhmän kertoimista edellä kuvatulla tavalla. Huomaa, että tässä yhtälöryhmässä ei ole kirjoitettu näkyviin muuttujia $x_1, \ldots, x_{\ell_1-1}$, koska niiden kaikki kertoimet ovat nollia. Tämä ei kuitenkaan tarkoita, että ratkaisujoukkoa olisi muutettu vaan, että molempien yhtälöryhmien ratkaisujoukko on avaruuden $\mathbb{R}^n$ osajoukko.

Toistetaan nyt edellä kuvattu prosessi yhtälöryhmälle

$$\begin{cases} a^1_{2(\ell_1+1)}x_{\ell_1+1} & +a^1_{2(\ell_1+2)}x_{\ell_1+2} & +\cdots & +a^1_{2n}x_n & = & b^1_2 \\ \vdots & \vdots & & \vdots & & \vdots \\ a^1_{m(\ell_1+1)}x_{\ell_1+1} & +a^1_{m(\ell_1+2)}x_{\ell_1+2} & +\cdots & +a^1_{mn}x_n & = & b^1_m \end{cases}$$

Jos tämä yhtäryhmä on koostuu nollariveistä, niin on se on jo supistetussa porramuodossa. Jos näin ei ole, niin toistamalla yllä kuvatut askeleet, saadaan tämä yhtälöryhmä ekvivalenttiin muotoon

$$\begin{cases} x_{\ell_2} & +a^1_{1(\ell_2+1)}x_{\ell_2+1} & +a^1_{1(\ell_2+2)}x_{\ell_2+2} & +\cdots & +a^2_{1n}x_n & = & b^2_1 \\ & a^2_{2(\ell_2+1)}x_{\ell_2+1} & +a^2_{2(\ell_1+2)}x_{\ell_1+2} & +\cdots & +a^1_{2n}x_n & = & b^2_2 \\ & \vdots & \vdots & & \vdots & & \vdots \\ & a^2_{m(\ell_2+1)}x_{\ell_2+1} & +a^2_{m(\ell_2+2)}x_{\ell_2+2} & +\cdots & +a^2_{mn}x_n & = & b^2_m \end{cases}$$

missä $\ell_1 < \ell_2 \leq n$.

Jatkamalla näin löydetään sellaiset indeksit $1 \leq \ell_1 < \ell_2 < \cdots < \ell_k \leq n$, että yhtälöryhmä (1.21) on ekvivalentti sellaisen yhtälöryhmän

$$\begin{cases} a^k_{11}x_1 + a^k_{12}x_2 + a^k_{13}x_3 + \cdots + a^k_{1n}x_n & = b^k_1 \\ a^k_{21}x_1 + a^k_{22}x_2 + a^k_{23}x_3 + \cdots + a^k_{1n}x_n & = b^k_2 \\ \vdots & \vdots \\ a^k_{m1}x_1 + a^k_{m2}x_2 + a^k_{m3}x_3 + \cdots + a^k_{mn}x_n & = b^k_m \end{cases}$$

kanssa, että jokaisella $p \in \{1, \ldots, k\}$ pätee $a^k_{p\ell_p} = 1$ ja $a^k_{j\ell_p} = 0$ kaikilla $j > p$.

Mikäli $a^k_{j\ell_p} \neq 0$ jollain $j < p$, niin vähennetään yhtälöryhmän rivi p vakiolla $a^k_{j\ell_p}$ kerrottuna rivistä j. Huomaa, että rivistä j ei tällöin tule nollariviä, koska $a^k_{j\ell_j} \neq 0$ ja $a^k_{pj} = 0$. Näin saadaan ekvivalentti yhtälöryhmä

$$\begin{cases} \tilde{a}_{11}x_1 + \tilde{a}_{12}x_2 + \tilde{a}_{13}x_3 + \cdots + \tilde{a}_{1n}x_n & = & \tilde{b}_1 \\ \tilde{a}_{21}x_1 + \tilde{a}_{22}x_2 + \tilde{a}_{23}x_3 + \cdots + \tilde{a}_{1n}x_n & = & \tilde{b}_2 \\ \vdots & & \vdots \\ \tilde{a}_{m1}x_1 + \tilde{a}_{m2}x_2 + \tilde{a}_{m3}x_3 + \cdots + \tilde{a}_{mn}x_n & = & \tilde{b}_m, \end{cases}$$

joka on supistetussa porrasmuodossa. Tämä päättää todistuksen. □

Syy porrasmuodon hyödyllisyyteen on se, että sen avulla voidaan suoraan selvittää yhtälöryhmän ratkaisut nk. takaisin sijoittamisen tekniikalla kuten esimerkeissa (1.4.1)–(1.4.5) on tehty. Lisäksi yhtälöryhmän ratkaisujen lukumäärä voidaan suoraan selvittää supistetusta porrasmuodosta. Kirjataan tästä yleinen tulos sellaisen yhtälöryhmän tapauksessa, jossa ei ole nollarivejä.

Lause 1.4.14. *Olkoon*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n & = b_m \end{cases}$$

supistetussa porrasmuodossa oleva yhtälöryhmä, jolla on ratkaisu.
Tällöin

1. *yhtälöryhmällä on yksikäsitteinen ratkaisu, jos kaikki sen muuttujat ovat sidottuja, ja*

2. *yhtälöryhmällä on äärettömän monta ratkaisua, jos sillä on yksikin vapaa muuttuja.*

Huomautus 1.4.15. *Koska yhtälöryhmällä on ratkaisu, niin sen nollarivit ovat muotoa* $0 = 0$*. Huomaa myös, että lauseen 1.4.14 ensimmäinen kohta on voimassa ainoastaan tilanteessa*

$$\begin{cases} x_1 & & & & = b_1 \\ & x_2 & & & = b_2 \\ & & \ddots & & \vdots \\ & & & x_n & = b_n \\ & & & 0 & = 0 \\ & & & & \vdots \\ & & & 0 & = 0 \end{cases}$$

Myös tämän lauseen todistuksen sivuuttaa tässä vaiheessa. Lukija voi palata tähän todistukseen esimerkiksi luvun 3 jälkeen. Konkreettisissa tilanteissa lauseen antamat tulokset on helppo lukea annetusta yhtälöryhmästä.

Esimerkki 1.4.16. *Yhtälöryhmä*

$$\begin{cases} x_1 & +x_2 & & = & 0 \\ & & x_3 & = & 5 \end{cases}$$

on supistetussa porrasmuodossa. Sen sidotut muuttujat ovat x_1 *ja* x_3 *ja muuttuja* x_2 *on sen ainoa vapaa muuttuja. Yhtälöryhmän ratkaisu* (x_1, x_2, x_3) *on muotoa*

$$\begin{cases} x_1 & = -t \\ x_2 & = t \\ x_3 & = 5, \end{cases}$$

missä $t \in \mathbb{R}$*, eli*

$$(x_1, x_2, x_3) = (-t, t, 5) = (0, 0, 5) + t(-1, 1, 0).$$

Lauseen 1.4.14 muotoiluun liittyen on hyvä huomata, että muuttujia ei koskaan tiputeta pois yhtälöryhmästä, koska tämä muuttaa yhtälöryhmää ja ratkaisujoukkoa. Otetaan tästä esimerkki.

Esimerkki 1.4.17. *Kahden yhtälön ja kolmen muuttujan yhtälöryhmällä*

$$\begin{cases} x_1 & & +0x_3 & = 0 \\ & x_2 & +0x_3 & = 0 \end{cases}$$

on vapaa muuttuja x_3. Näin ollen sillä on äärettömän monta ratkaisua. Ratkaisut ovat muotoa $t(0,0,1) \in \mathbb{R}^3$, missä $t \in \mathbb{R}$.

Jos muuttujaa x_3 ei lainkaan huomioitaisi, tutkittaisiin yhtälöryhmää kahden muuttujan ja kahden yhtälön yhtälöryhmää

$$\begin{cases} x_1 & & = 0 \\ & x_2 & = 0 \end{cases}$$

joka on täysin eri yhtälöryhmä ja jonka ratkaisut ovat avaruudessa $\mathbb{R}^2$ avaruuden $\mathbb{R}^3$ sijaan. Itseasiassa tämän yhtälöryhmän ainoa ratkaisu on $(0,0)$.

Lauseen 1.4.14 todistus. Koska yhtälöryhmällä on ratkaisu, ei supistettu porrasmuoto sisällä rivejä $0 = b_j$, missä $b_j \neq 0$. Koska nollarivit $0 = 0$ eivät vaikuta ratkaisuihin, niin ne voidaan yhtälöryhmästä pois, eli olettaa, että yhtälöryhmässä ei ole lainkaan nollarivejä.

Oletetaan ensn, että yhtälöryhmän kaikki muuttujat ovat sidottuja. Tällöin sidotun muuttujan määritelmän nojalla jokaisella rivillä j on täsmälleen yksi nollasta poikkeava kerroin $a_{jj} = 1$. Tällöin vektori $(b_1, \ldots, b_n)$ on yhtälöryhmän ainoa ratkaisu.

Oletetaan nyt, että yhtälöryhmällä on vapaita muuttujia ja osoitetaan, että tällöin yhtälöryhmällä on äärettömästi ratkaisuja. Käsitellään yhtälöryhmää riveittäin kahdessa eri tapauksessa.

Olkoon $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$ yhtälöryhmän sellainen rivi, jolla vain yksi kertoimista a_{ji} on nollasta poikkeva, eli yhtälö on $a_{ji_j}x_{i_j} = b_j$, jollain $i_j \in \{1, \ldots, n\}$. Supistetun porrasmuodon määritelmän mukaan tällöin $a_{ji_j} = 1$. Valitaan $x_{i_j} = b_j$.

Tarkastellaan nyt yhtälöryhmän yhtälöä $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$, jossa vähintään kaksi kertoimista on nollasta poikkeavia. Huomaa, että jos $a_{ji} \neq 0$, niin tällöin $a_{ri} = 0$ kaikilla $r \neq j$, eli että vastaava kerroin ei ole nollasta poikkeava millään muulla rivillä. Olkoon $i_j \in \{1, \ldots, n\}$ pienin niistä indekseistä, joilla $a_{ji} \neq 0$. Valitaan nyt $x_{i_j} = b_j$ ja valitaan $x_i = 0$ niillä indekseillä $i_j < i \leq n$, joilla $a_{ji} \neq 0$. Tässä ei ole ristiriitaa aiempien valintojen kanssa, koska muuttuja x_{i_j} on sidottu muuttuja, jonka arvoa ei määritellä minkään muun rivin yhteydessä, ja muut muuttujat x_i, joille määriteltiin arvo tässä yhteydessä, eivät ole sidottuja muuttujia, joten ne saavat joka kerta saman arvon 0. Näillä valinnoilla yhtälö $a_{j1}x_1 + \cdots + a_{jn}x_m = b_j$ on voimassa.

Koska yhtälöryhmässä ei ole (enää) nollarivejä, niin nyt luku x_i on määrätty jokaisella indeksillä $i \in \{1, \ldots, n\}$. Koska näillä lukujen $x_1, \ldots, x_n$ valinnoilla jokainen yhtälöryhmän yhtälö on voimassa, niin $(x_1, \ldots, x_n)$ on yhtälöryhmän ratkaisu.

Olkoon nyt $t \in \mathbb{R}$. Koska on oletettu, että yhtälöryhmässä on vapaa muuttuja, niin yhtälöryhmässä on sellainen rivi j, jolla on vähintään kaksi nollasta poikeavaa kerrointa a_{ji_j} ja $a_{j\ell}$. Valitaan nyt muuttujan x_ℓ arvoksi t ja asettaa $x_{i_j} = b_j - a_{j\ell}t$. Tällöin myös vektori $(x_1, \ldots, x_n)$ on yhtälöryhmän ratkaisu. Koska $t \in \mathbb{R}$ oli mielivaltainen, niin havaitaan, että yhtälöryhmällä on tässä tilanteessa äärettömän monta ratkaisua. $\square$

Luku 2

Matriisit

Luvun tavoitteet

- Matriisien laskutoimitukset, erityisesti matriisien kertolasku.
- Käänteismatriisin määritelmä ja etsiminen.
- Rivioperaatioiden ja alkeismatriisien välinen yhteys.

2.1 Motivointi: Yhtälöryhmän matriisimerkintä

Yhtälöryhmiä ratkottaessa herää kysymys, että onko pakko raahata muuttujia $x_1, \ldots, x_n$ mukana laskuissa. Vastaus tähän käytännönläheiseen kysymykseen on tietysti, että ei ole pakko.

Yhtälöryhmän

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

ratkaisut riippuvat ainoastaan kertoimista $a_{ji} \in \mathbb{R}$ ja vakioista $b_1, \ldots, b_m \in \mathbb{R}$. Toisaalta tiedetään, että rivioperaatiot ainoastaan muuttavat kertoimia a_{ji} ja lukuja $b_1, \ldots, b_m$. Näin ollen on selvää, että mitään informaatiota yhtälöryhmän ratkaisuista ei kadoteta, jos yhtälöryhmä kirjoitetaan tiivistetysti muodossa

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]. \tag{2.1}$$

Huomautus 2.1.1. *Huomaa, että tässä merkinnässä kertoimet a_{ji} ja vakiot b_j on erotettu toisistaan viivalla. Tätäkään ei aina tehdä. Erottava viiva kuitenkin helpottaa merkinnän ymmärtämistä käytännön tilanteissa. Vertaa esimerkiksi merkintöjä*

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 0 & -1 \end{bmatrix} \text{ ja } \left[\begin{array}{cc|c} 1 & 1 & 3 \\ 2 & 0 & -1 \end{array}\right].$$

Ratkaistaan esimerkin 1.4.5 yhtälöryhmä käyttäen tätä merkintää

Esimerkki 2.1.2. *Kirjoitetaan yhtälöryhmä*

$$\begin{cases} x_1 & -x_2 & +2x_3 & -x_4 & = & 3 \\ x_1 & +x_2 & & +x_4 & = & 3 \end{cases}$$

muodossa

$$\left[\begin{array}{rrrr|r} 1 & -1 & 2 & -1 & 3 \\ 1 & 1 & 0 & 1 & 3 \end{array}\right]$$

ja sovelletaan esimerkissä 1.4.4 käytettyjä merkintöjä ratkaisussa. Tällöin saadaan

$$\begin{aligned} \left[\begin{array}{rrrr|r} 1 & -1 & 2 & -1 & 3 \\ 1 & 1 & 0 & 1 & 3 \end{array}\right] \begin{matrix} \\ -R_1 \end{matrix} & \sim \left[\begin{array}{rrrr|r} 1 & -1 & 2 & -1 & 3 \\ 0 & 2 & -2 & 2 & 0 \end{array}\right] \begin{matrix} \\ \cdot\frac{1}{2} \end{matrix} \\ & \sim \left[\begin{array}{rrrr|r} 1 & -1 & 2 & -1 & 3 \\ 0 & 1 & -1 & 1 & 0 \end{array}\right] \begin{matrix} +R_2 \\ \\ \end{matrix} \\ & \sim \left[\begin{array}{rrrr|r} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & -1 & 1 & 0 \end{array}\right] . \end{aligned}$$

Alkuperäinen yhtälöryhmä on siis ekvivalentti supistetussa porrasmuodossa olevan yhtälöryhmän

$$\begin{cases} x_1 & & +x_3 & & = & 3 \\ & x_2 & -x_3 & +x_4 & = & 0 \end{cases}$$

kanssa. Näin ollen

$$\begin{cases} x_1 & = & -t+3 \\ x_2 & = & +t-s \\ x_3 & = & t \\ x_4 & = & s \end{cases}$$

missä $t, s \in \mathbb{R}$.

Yhtälöryhmän matriisimerkintä herättää ajatuksen, että onko näitä merkintöjä käyttäen mahdollista tehokkaasti ratkaista useita yhtälöryhmiä samoilla kertoimilla a_{ji} mutta eri vakioilla $b_1, \ldots, b_m$. Käsitellään tätä esimerkin avulla.

Esimerkki 2.1.3. *Tarkastellaan samaa yhtälöryhmää kuin esimerkissä 2.1.2*

$$\begin{cases} x_1 & -x_2 & +2x_3 & -x_4 & = & 3 \\ x_1 & +x_2 & & +x_4 & = & 3 \end{cases} \tag{2.2}$$

ja samoilla kertoimilla määriteltyä yhtälöryhmää

$$\left\{ \begin{array}{rrrrrr} y_1 & -y_2 & +2y_3 & -y_4 & = & 0 \\ y_1 & +y_2 & & +y_4 & = & -2 \end{array} \right. \tag{2.3}$$

Yhtälöryhmät (2.2) *ja* (2.3) *voidaan ratkaista samanaikaisesti kirjoittamalla ne muotoon*

$$\left[\begin{array}{rrrr|rr} 1 & -1 & 2 & -1 & 3 & 0 \\ 1 & 1 & 0 & 1 & 3 & -2 \end{array}\right]$$

ja käyttämällä samoja rivioperaatioita kuin esimerkissä 2.1.2. Tällöin

$$\begin{aligned} \left[\begin{array}{rrrr|rr} 1 & -1 & 2 & -1 & 3 & 0 \\ 1 & 1 & 0 & 1 & 3 & -2 \end{array}\right] \begin{array}{l} \\ -R_1 \end{array} & \sim \left[\begin{array}{rrrr|rr} 1 & -1 & 2 & -1 & 3 & 0 \\ 0 & 2 & -2 & 2 & 0 & -2 \end{array}\right] \begin{array}{l} \\ \cdot\frac{1}{2} \end{array} \\ & \sim \left[\begin{array}{rrrr|rr} 1 & -1 & 2 & -1 & 3 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \end{array}\right] \begin{array}{l} +R_2 \\ \\ \end{array} \\ & \sim \left[\begin{array}{rrrr|rr} 1 & 0 & 1 & 0 & 3 & -1 \\ 0 & 1 & -1 & 1 & 0 & -1 \end{array}\right] \end{aligned}$$

Näin ollen yhtälöryhmän (2.2) *ratkaisuiksi saadaan vektorit* $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$, *joille pätee*

$$\begin{cases} x_1 = -t + 3 \\ x_2 = t - s \\ x_3 = t \\ x_4 = s, \end{cases}$$

missä $t, s \in \mathbb{R}$. *Yhtälöryhmän* (2.3) *ratkaisuiksi saadaan vektorit* $(y_1, y_2, y_3, y_4) \in \mathbb{R}^4$, *joille pätee*

$$\begin{cases} y_1 = -t - 1 \\ y_2 = t - s - 1 \\ y_3 = t \\ y_4 = s, \end{cases}$$

missä $t, s \in \mathbb{R}$.

Esimerkistä 2.1.3 saa vaikutelman, että merkinnässä (2.1) lukuja pystysuoran erotusviivan molemmilla puolilla käsitellään samalla tavalla. Tämä ei ole sattumaa vaan havainto johdattaa matriisin ja matriisiyhtälön käsitteisiin, jossa merkinnässä (2.1) viivalla erotetut puolet tulkitaan omina matriiseinaan ja yhtälöryhmän kertoimet $x_1, \ldots, x_n$ merkitään omaksi matriisikseen.

2.2 Matriisit ja sarakkeet

Kuten avaruuden $\mathbb{R}^n$ vektorit voidaan määritellä jonoina, myös matriiseille on formaalisti korrekti määritelmä. Aloitetaan kuitenkin hieman selkeämmällä luonnollisella määritelmällä.

Määritelmä 2.2.1. *Olkoot $m, n \in \mathbb{Z}_+$. Lukujen $a_{ji} \in \mathbb{R}$, missä $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, n\}$, taulukkoa*

$$A = [a_{ji}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

*kutsutaan $m \times n$-*matriisiksi*. Lukuja a_{ji} kutsutaan matriisin A* kertoimiksi*. Kaikkien $m \times n$-matriisien joukkoa merkitään $\mathbb{R}^{m\times n}$.*

Aloitetaan huomautuksilla merkinnöistä ja konventioista.

Huomautus 2.2.2. *Jatkossa 1×1-matriisi $[a] \in \mathbb{R}^{1\times 1}$ samastetaan luvun $a \in \mathbb{R}$ kanssa, eli avaruus $\mathbb{R}^{1\times 1}$ samastetaan reaalilukujen $\mathbb{R}$ kanssa.*

Huomautus 2.2.3. *Mikäli matriisin $A \in \mathbb{R}^{m\times n}$ kertoimille ei ole valittu merkintää, niitä merkitään symboleilla A_{ji}, eli luvut $A_{ji} \in \mathbb{R}$ ovat sellaisia lukuja, että*

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}.$$

Tätä konventioita käytetään jatkossa muutamia kertoja. Usein on kuitenkin mielekäänpää merkitä matriisin A kertoimia symboleilla a_{ji} ja kirjoittaa $A = [a_{ji}]$.

Matriisien yhteydessä puhutaan matriisin sarakkeista ja riveistä. Tällä tarkoitetaan seuraavia alkuperäiseen matriisiin liittyviä matriiseja.

Määritelmä 2.2.4. *Olkoon $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ matriisi. Matriisia*

$$A_{\cdot i} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} \in \mathbb{R}^{m\times 1}$$

*kutsutaan matriisin A i:*nneksi sarakkeeksi *ja matriisia*

$$A_{j\cdot} = \begin{bmatrix} a_{j1} & a_{j2} & \cdots & a_{jn} \end{bmatrix} \in \mathbb{R}^{1\times n}$$

*j:*nneksi riviksi*.*

Yleisemmin puhutaan sarakevektoreista ja rivivektoreista.

Määritelmä 2.2.5. *Joukon $\mathbb{R}^{n\times 1}$ matriiseja kutsutaan* sarakevektoreiksi *ja joukon $\mathbb{R}^{1\times n}$ matriiseja* rivivektoreiksi*.*

Syy tälle terminologialle paljastuu tämän luvun aikana.

Esimerkki 2.2.6. *Määritelmän 2.2.4 matriisissa* $A \in \mathbb{R}^{m\times n}$ *on* m *riviä ja* n *saraketta.*

Esimerkki 2.2.7. *Matriisit*

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 5 & -1 \end{bmatrix}, \; B = \begin{bmatrix} 1 & -1 \\ 2 & 5 \\ 0 & -1 \end{bmatrix},$$

ovat luvuista $1, -1, 0, 2, 5, -1$ *muodostettuja erilaisia matriiseja. Matriisissa* A *on* 2 *riviä ja* 3 *saraketta. Matriissa* B *puolestaan on* 3 *riviä ja* 2 *saraketta. Matriisin* A *toinen rivi on matriisi* $\begin{bmatrix} 2 & 5 & -1 \end{bmatrix}$ *ja vastaavasti matriisin* B *toinen rivi on matriisi* $\begin{bmatrix} 2 & 5 \end{bmatrix}$.

Matriisin sarakemerkintä

Matriisia $A \in \mathbb{R}^{m\times n}$ on usein mielekästä merkitä sarakeittain, eli muodossa

$$A = \begin{bmatrix} A_{\cdot 1} & \cdots & A_{\cdot n} \end{bmatrix}.$$

Toisaalta monissa tilanteissa on hyödyllistä määritellä matriisi annettujen sarakevektoreiden avulla. Annetaan tätä varten tarkka määritelmä.

Määritelmä 2.2.8. *Olkoot* $v_1, \ldots, v_n \in \mathbb{R}^{m\times 1}$ *sarakevektoreita*

$$v_i = \begin{bmatrix} v_{1i} \\ \vdots \\ v_{mi} \end{bmatrix}$$

jokaisella $i \in \{1, \ldots, n\}$*. Tällöin matriisilla* $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ *tarkoitetaan matriisia*

$$A = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{m1} & v_{m2} & & v_{mn} \end{bmatrix}.$$

Esimerkki 2.2.9. *Olkoot*

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^{3\times 1} \text{ ja } v_2 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \in \mathbb{R}^{3\times 1}.$$

Tällöin

$$A = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

on matriisi

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

Lisäksi $A_{\cdot 1} = v_1$ *ja* $A_{\cdot 2} = v_2$.

Termi taulukko matriisin määritelmässä

Palataan nyt matriisin määritelmään 2.2.1, jossa käytetään määrittelemätöntä käsitettä *taulukko*. Tämä voidaan ratkaista seuraavasti.

Palautetaan mieleen, että jonolla $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ tarkoitetaan itseasiassa funktiota $x\colon \{1, \dots, n\} \to \mathbb{R}$, jolle pätee $x(i) = x_i$ kaikilla $i \in \{1, \dots, n\}$. Vastaavasti voidaan sanoa, että $m \times n$-matriisi A on itseasiassa funktio $A\colon \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{R}$. Lukujen a_{ji} määräämä $m \times n$-matriisi $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ on siis funktio $A\colon \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{R}$, jolle pätee $A(j, i) = a_{ji}$ kaikilla $j \in \{1, \dots, m\}$ ja $i \in \{1, \dots, n\}$. Tämä antaa määritelmän taulukon käsitteelle, jota matriisin määritelmässä käytettiin intuitiivisessa mielessä. Tätä formaalisti tarkkaa matriisin määritelmää on kuitenkin hankala hyödyntää käytännössä. Matriisi merkinnän tarkoituksena onkin korvata tämä formaalisti oikea määritelmä käytännöllisemmällä merkinnällä. Huomautettakoon, että tätä formaalia määritelmää ei jatkossa tarvita, vaan määritelmä 2.2.1 riittää mainiosti.

Huomautettakoon kuitenkin, että käyttäen tätä formaalia määritelmää, matriisien samuutta ei tarvitse erikseen käsitellä, koska matriisit A ja B ovat sama matriisi, jos ja vain jos kuvaukset $A\colon \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{R}$ ja $B\colon \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{R}$ ovat sama kuvaus. Lisäksi havaitaan, että $\mathbb{R}^{m\times n}$ on itseasiassa kaikkien funktioiden $\{1, \dots, m\} \times \{1, \dots n\} \to \mathbb{R}$ joukko.

2.2.1 Matriisien yhteenlasku ja skalaarikertolasku

Matriiseille määritellään yhteenlasku seuraavasti.

Määritelmä 2.2.10. *Olkoot $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ ja $B = [b_{ji}] \in \mathbb{R}^{m\times n}$ matriiseja. Tällöin $A + B \in \mathbb{R}^{m\times n}$ on matriisi, jolle pätee*

$$(A + B)_{ji} = a_{ji} + b_{ji}.$$

jokaisella $j \in \{1, \dots, m\}$ ja $i \in \{1, \dots, n\}$. Matriisia $A + B$ kutsutaan matriisien A ja B summaksi.

Matriisit lasketaan siis yhteen kertoimittain, eli

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

Näin ollen matriisien A ja B pitää olla saman kokoisia, jotta ne voidaan laskea yhteen.

Esimerkki 2.2.11. *Matriisien*

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 5 & 0 \end{bmatrix} \text{ ja } B = \begin{bmatrix} 1 & -1 & 4 \\ 0 & 3 & -2 \end{bmatrix}$$

summa $A + B$ on

$$A + B = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 5 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -1 & 4 \\ 0 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 1+1 & 0-1 & -1+4 \\ 2+0 & 5+3 & 0-2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 3 \\ 2 & 8 & -2 \end{bmatrix}.$$

Määritellään nyt matriisien skalaarikertolasku, eli matriisin $A \in \mathbb{R}^{m\times n}$ ja luvun $\lambda \in \mathbb{R}$ tulo.

Määritelmä 2.2.12. *Olkoot $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ ja $\lambda \in \mathbb{R}$. Tällöin $\lambda A \in \mathbb{R}^{m\times n}$ on matriisi, jolle pätee*

$$(\lambda A)_{ji} = \lambda a_{ji}$$

jokaisella $j \in \{1, \dots, m\}$ ja $i \in \{1, \dots, n\}$. Matriisia λA matriisin A ja luvun λ tuloksi.

Myös skalaarikertolasku suoritetaan kertoimittain, eli

$$\lambda \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}$$

Huomautus 2.2.13. *Huomaa, että määritelmissä 2.2.10 ja 2.2.12 matriisit $A + B$ ja λA ovat itseasiassa matriisien nimiä. Ne ovat sellaisia matriiseja, jotka määräytyvät matriiseista A ja B ja luvusta λ annettujen sääntöjen mukaisesti.*

Formaalisti matriisien $A+B$ ja λA määritelmien avulla määritellään $m\times n$-matriisien yhteenlasku $+\colon \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$, $(A, B) \mapsto A + B$, ja skalaarikertolasku $\cdot\colon \mathbb{R} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$, $(\lambda, A) \mapsto \lambda A$.

Joukon $\mathbb{R}^{m\times n}$ matriisien yhteenlaskulle ja skalaarikertolaskulle pätevät samat ominaisuudet kuin avaruuden $\mathbb{R}^n$ vektoreille. Kirjataan tämä havainto lemmaksi, mutta jätetään todistus, joka perustuu reaalilukujen vastaaviin ominaisuuksiin, kiinnostuneelle lukijalle.

Lemma 2.2.14 (Matriisien yhteenlasku- ja skalaarikertolaskulait)**.** *Olkoot $A, B, C \in \mathbb{R}^{m\times n}$ ja $a, b \in \mathbb{R}$. Tällöin*

1. $(A + B) + C = A + (B + C)$,
2. $A + B = B + A$,
3. $a(A + B) = aA + aB$,
4. $(ab)A = a(bA)$ *ja*
5. $(a + b)A = aA + bA$.

Huomautus 2.2.15 (Kiinnostuneelle lukijalle)**.** *Kuten yllä todettiin, nämä matriiseihin liittyvät laskulait voi todistaa suoraan matriisien yhteenlaskun ja skalaarikertolaskun määritelmistä käyttäen reaalilukujen vastaavia ominaisuuksia. Toinen tapa on havaita, että – samoin kuin vektoreiden laskutoimitukset – $m\times n$-matriisien laskutoimitukset ovat itseasiassa funktioiden $\{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{R}$ laskutoimituksia.*

2.3 Matriisitulo

Määritellään nyt matriisien tulo. Toisin kuin matriisien yhteenlasku ja skalaarikertolasku, matriisien tuloa ei määritellä kertoimittain. Matriisien tulo voidaan motivoida kahdella eri tavalla: yhtälöryhmillä tai kuvausten yhdistämisellä. Aloitetaan kuitenkin motivoinnin sijaan suoraan matriisitulon brutaalilla määritelmällä.

Määritelmä 2.3.1. *Matriisien $A = [a_{ji}] \in \mathbb{R}^{m \times n}$ ja $B = [b_{ji}] \in \mathbb{R}^{n \times k}$* tulo *$AB$ on matriisi $AB \in \mathbb{R}^{m \times k}$, jonka alkiot on määritelty kaavalla*

$$(AB)_{ji} = \sum_{\ell=1}^{n} a_{j\ell} b_{\ell i}$$

kaikilla $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, k\}$.

Määritelmän mukaan tulomatriisin AB kerroin $(AB)_{ji}$ saadaan kertomalla matriisin A j:nnen rivin kertoimet matriisin B i:nnen sarakkeen vastaavien kertoimien kanssa ja laskemalla tulot yhteen. Tämä tarkoittaa, että matriisien A ja B tulo AB on määritelty ainoastaan sellaisille matriiseille A ja B, joilla matriisin A sarakkeiden lukumäärä on sama kuin matriisin B rivien lukumäärä. Lisäksi matriisien A ja B järjestyksellä on väliä kuten seuraavat esimerkit osoittavat.

Esimerkki 2.3.2. *Olkoot*

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 3}$$

ja

$$B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ -2 & 2 \end{bmatrix} \in \mathbb{R}^{3 \times 2}.$$

Lasketaan $AB \in \mathbb{R}^{2 \times 2}$ ja $BA \in \mathbb{R}^{3 \times 3}$. Aloitetaan matriisista AB.

Lasketaan ensin kerroin $(AB)_{11}$. Koska $(AB)_{11} = A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31}$ niin havaitaan, että

$$(AB)_{11} = 1 \cdot 1 + 2 \cdot 0 + (-1) \cdot (-2) = 1 + 2 = 3.$$

Vastaavasti voidaan laskea muutkin kertoimet. Näin saadaan, että

$$AB = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ 0 & -1 \end{bmatrix}.$$

Huomautus 2.3.3. *Matriisituloja laskettaessa tulee helposti virheitä. Käytännön laskuissa, kuten matriistulon AB kohdalla edellisessä esimerkissä, voi virheiden välttämiseksi käyttää merkintätapaa*

$$\begin{array}{cc} & \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ -2 & 2 \end{bmatrix}, \\ \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 3 & -4 \\ 0 & -1 \end{bmatrix} \end{array}$$

josta pystyy helposti lukemaan tulomatriisin AB kertoimeen $(AB)_{ji}$ vaadittavan matriisin A rivin j ja matriisin B sarakkeen i.

Esimerkki 2.3.4. *Olkoot matriisit A ja B kuten esimerkissä 2.3.14. Lasketaan nyt tulo BA käyttäen edellisen huomautuksen tapaa.*

Näillä merkinnöillä matriisiksi BA saadaan

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

eli

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

Matriisitulon yhteys yhtälöryhmiin

Matriistulon ja yhtälöryhmien välinen yhteys perustuu huomioon rivi ja sarakevektoreiden tulosta. Käsitellään tätä ensin esimerkin muodossa.

Esimerkki 2.3.5. *Rivivektorin*

$$a = \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \in \mathbb{R}^{1\times n}$$

ja sarakevektorin

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

tulo ax on 1×1-matriisi (eli reaaliluku)

$$ax = \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \in \mathbb{R}. \tag{2.4}$$

Edellisessä esimerkissä symbolien a ja x valinta perustuu yhtälöryhmistä tuttuihin merkintöihin. Näillä merkinnöillä tulo ax vastaa itseasiassa täsmälleen yleisen yhtälöryhmän ensimmäisen rivin vasenta puolta. Sovelletaan nyt tätä kaavaa matriisin $A \in \mathbb{R}^{m\times n}$ tapauksessa.

Esimerkki 2.3.6. *Olkoon $A \in \mathbb{R}^{m\times n}$ kertoimien a_{ji} matriisi, eli*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

ja olkoon

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

sarakevektori. Tällöin laskemalla matriisitulo riveittäin saadaan

$$Ax = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix},$$

eli vektorin Ax j:s kerroin $(Ax)_j$ on rivivektorin $\begin{bmatrix} a_{j1} & \cdots & a_{jn} \end{bmatrix}$ ja sarakevektorin x tulo.

Edellinen esimerkki jo oleellisesti paljasti matriisitulon ja yhtälöryhmän välinen yhteyden. Seuraavassa esimerkissä tätä käsitelään tätä kuitenkin vielä tarkasti.

Esimerkki 2.3.7. *Olkoon $(x_1, \ldots, x_n) \in \mathbb{R}^n$ yhtälöryhmän*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \qquad \vdots & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases} \tag{2.5}$$

ratkaisu.

Merkitään nyt

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}, \ x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1} \text{ ja } b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^{m \times 1}.$$

Tällöin

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax.$$

Näin ollen yhtälöryhmän (2.5) ratkaiseminen vastaa sarakevektorin x ratkaisemista yhtälöstä

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{2.6}$$

eli sarakevektorin x ratkaisemista matriisiyhtälöstä

$$Ax = b.$$

Huomaa, että käyttäen luvun alussa käyttöön otettua merkintää tämä yhtälö vastaa yhtälöä

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

joka voidaan nyt kirjoittaa lyhyesti muodossa

$$\left[A \mid b\right].$$

Matriisitulo ja sarakkeet

Käsitellään vielä yleistä matriisituloa tilanteessa, jossa matriisi on ilmoitettu käyttäen sarakemerkintää. Vaikka konkreettisilla matriiseilla käsinlaskettaessa seuraava tulos ei anna erityistä etua, on sillä konseptuaalisesti erittäin suuri merkitys matriisitulon hahmottamisen kannalta.

Lemma 2.3.8. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi ja olkoon $B \in \mathbb{R}^{n\times k}$ sellainen matriisi, jonka sarakkeet ovat $b_1, \ldots, b_k \in \mathbb{R}^{n\times 1}$, eli että $B = \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix}$. Tällöin*

$$AB = \begin{bmatrix} Ab_1 & \cdots & Ab_k \end{bmatrix}.$$

Lemma siis sanoo, että tulomatriisin AB sarakkeet saadaan kertomalla matriisin B sarakkeet matriisilla A.

Esimerkki 2.3.9. *Olkoot*

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \text{ ja } B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Tällöin matriisin B sarakeet ovat

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ ja } \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Koska

$$\begin{array}{cc} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 2 \\ 2 \end{bmatrix} \end{array}, \quad \begin{array}{cc} & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{array} \quad \text{ja} \quad \begin{array}{cc} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 2 \\ 1 \end{bmatrix} \end{array},$$

niin

$$AB = \begin{bmatrix} 2 & 0 & 2 \\ 2 & 1 & 1 \end{bmatrix}.$$

Lemman 2.3.8 todistus. Pitää osoittaa, että matriisi $AB \in \mathbb{R}^{m\times k}$ on matriisi $C = \begin{bmatrix} Ab_1 & \cdots & Ab_k \end{bmatrix}$, eli että niille on samat kertoimet, eli että $(AB)_{ji} = C_{ji}$ kaikilla $j \in \{1\ldots, m\}$ ja $i \in \{1, \ldots, k\}$. Riittää kuitenkin osoittaa, että niillä on samat sarakevektorit, koska tällöin niillä on myös samat kertoimet.

Olkoot

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

ja

$$b_i = \begin{bmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{bmatrix}$$

jokaisella $i \in \{1, \ldots, k\}$.

Olkoot $j \in \{1 \ldots, m\}$ ja $i \in \{1, \ldots, k\}$. Tällöin matriisitulon määritelmän nojalla

$$(AB)_{ji} = a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni}.$$

Matriisin AB i:s sarake on siis sarakevektori

$$(AB)_{\cdot i} = \begin{bmatrix} a_{11}b_{1i} + a_{12}b_{2i} + \cdots + a_{1n}b_{ni} \\ \vdots \\ a_{m1}b_{1i} + a_{m2}b_{2i} + \cdots + a_{mn}b_{ni} \end{bmatrix}.$$

Toisaalta

$$Ab_i = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{1i} \\ \vdots \\ b_{ni} \end{bmatrix} = \begin{bmatrix} a_{11}b_{1i} + a_{12}b_{2i} + \cdots + a_{1n}b_{ni} \\ \vdots \\ a_{m1}b_{1i} + a_{m2}b_{2i} + \cdots + a_{mn}b_{ni} \end{bmatrix}$$

eli

$$(AB)_{\cdot i} = Ab_i$$

jokaisella $i \in \{1, \ldots, k\}$. Näin ollen

$$AB = \begin{bmatrix} Ab_1 & \cdots & Ab_k \end{bmatrix}.$$

□

Edellinen lause motivoi seuraavan havainnon, jota tullaan jatkossa käyttämään useasti. Havaintoa varten tarvitaan merkintä.

Merkintä 2.3.10. *Merkitään*

$$e_i = \begin{bmatrix} e_{1i} \\ \vdots \\ e_{ni} \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

sarakevektoria, jonka kertoimille pätee

$$e_{ji} = \begin{cases} 1, & j = i \\ 0, & j \neq i. \end{cases}$$

Huomautus 2.3.11. *Huomaa, että merkinnästä e_i ei voi päätellä mihin avaruuteen $\mathbb{R}^{n\times 1}$ vektori e_i kuuluu. Tämä seuraa kuitenkin yleensä kontekstista tai voidaan ilmaista kirjoittamalla $e_i \in \mathbb{R}^{n\times 1}$. Vektoreille $e_1, \ldots, e_n$ annetaan luvussa 3 nimi. Ne ovat avaruuden $\mathbb{R}^{n\times 1}$ ns. standardikannan alkioita.*

Esimerkki 2.3.12. *Jos $n = 2$, niin*

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ ja } e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Jos $n = 3$, niin

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ ja } e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Lemmasta 2.3.8 seuraa nyt, että matriisin $A \in \mathbb{R}^{m\times n}$ sarakkeet voidaan poimia kertomalla matriisia sarakevektoreilla $e_1, \ldots, e_n$.

Korollaari 2.3.13. *Olkoon $A \in \mathbb{R}^{m\times n}$. Tällöin*

$$A = \begin{bmatrix} Ae_1 & \cdots & Ae_n \end{bmatrix}$$

Todistus. Riittää osoittaa, että Ae_i on matriisin sarake $A_{\cdot i}$. Olkoon

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

ja $i \in \{1, \ldots, n\}$. Tällöin

$$Ae_i = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} e_{1i} \\ \vdots \\ e_{ni} \end{bmatrix} = \begin{bmatrix} a_{11}e_{1i} + a_{12}e_{2i} + \cdots + a_{1n}e_{ni} \\ a_{21}e_{1i} + a_{22}e_{2i} + \cdots + a_{2n}e_{ni} \\ \vdots \\ a_{m1}e_{1i} + a_{m2}e_{2i} + \cdots + a_{mn}e_{ni} \end{bmatrix} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} = A_{\cdot i}.$$

□

Matriisi kuvauksena

Matriisitulon avulla matriisille $A \in \mathbb{R}^{m\times n}$ voidaan antaa tulkinta kuvauksena. Tämä perustuu havaintoon, että jokaisella sarakevektorilla $x \in \mathbb{R}^{n\times 1}$, tulo Ax on sarakeavaruuden $\mathbb{R}^{m\times 1}$ alkio. Näin ollen

$$f_A \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}, \quad x \mapsto Ax,$$

on hyvin määritelty kuvaus sarakeavaruudesta $\mathbb{R}^{n\times 1}$ sarakeavaruuteen $\mathbb{R}^{m\times 1}$.

Esimerkki 2.3.14. *Olkoon*

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

kuten esimerkissä ja olkoon $f_A\colon \mathbb{R}^{3\times 1} \to \mathbb{R}^{2\times 1}$ *kuvaus* $x \mapsto Ax$. *Etsitään kuvaukselle* f_A *kaava sarakevektorin alkioiden avulla.*

Olkoon

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^{3\times 1}.$$

Tällöin

$$f_A(x) = Ax = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_2 \end{bmatrix}.$$

Kuvaus f_A on tärkeä esimerkkitapaus niin sanotuista lineaarikuvauksista. Kirjataan lemmaksi kuvauksen f_A tärkeimmät ominaisuudet. Ne seuraavat välittömästi sarakevektoreiden laskusäännöistä.

Lemma 2.3.15. *Olkoon* $A \in \mathbb{R}^{m\times n}$ *matriisi ja* $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$, $x \mapsto Ax$, *kuvaus. Tällöin kaikilla* $x, y \in \mathbb{R}^{n\times 1}$ *ja* $a \in \mathbb{R}$ *pätee sekä* $f_A(x + y) = f_A(x) + f_A(y)$ *että* $f_A(ax) = af_A(x)$.

Todistus. Olkoot $x, y \in \mathbb{R}^{n\times 1}$. Kuvauksen f_A määritelmän ja matriisitulon laskusääntöjen nojalla pätee

$$f_A(x + y) = A(x + y) = Ax + Ay = f_A(x) + f_A(y).$$

Lisäksi

$$f_A(ax) = A(ax) = aAx = af_A(x).$$

□

Matriisitulo ja rivivektorit

Matriisitulon yhteys sarakkeisiin herättää kysymyksen matriisitulon ja rivivektoreiden vastaavasta yhteydessä. Tälläinen yhteys on olemassa. Otetaan käyttöön seuraava merkintä.

Merkintä 2.3.16. *Rivivektori* $e_j^* \in \mathbb{R}^{1\times n}$ *on vektori* $e_j^* = \begin{bmatrix} e_{j1}^* & \cdots & e_{in}^* \end{bmatrix}$, *missä*

$$e_{ji}^* = \begin{cases} 1, & j = i \\ 0, & j \neq i. \end{cases}$$

Esimerkki 2.3.17. *Olkoon* $n = 3$. *Tällöin*

$$e_1^* = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},\ e_2^* = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \text{ ja } e_3^* = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$$

Rivivektori $e_j^* \in \mathbb{R}^{1\times n}$ muistuttaa sarakevektoria $e_j \in \mathbb{R}^{n\times 1}$; niillä on samat kertoimet samassa järjestyksessä. Ainoa ero on, että toinen on sarakevektori ja toinen on rivivekori. Tämä on kuitenkin suuri ero matriisitulon suhteen. Tämän havainnon motivoimana otetaan seuraavia laskuja varten käyttöön merkintä, jonka avulla matriisi voidaan ilmaista rivien avulla. Vertaa seuraavaa merkintää matriisin ilmaisemiseen sarakkeiden avulla.

Merkintä 2.3.18. *Matriisia $A \in \mathbb{R}^{m\times n}$ merkitään*

$$A = \begin{bmatrix} - & A_1 & - \\ & \vdots & \\ - & A_m & - \end{bmatrix},$$

missä $A_j = A_{j\cdot}$ on matriisin A j:s rivi.

Syy näille merkinnöille on seuraava havainto, jota voi verrata lemmaan 2.3.8. Todistus on samanlainen kuin lemmassa 2.3.8 ja korollaarissa 2.3.13, joten se jätetään harjoitustehtäväksi.

Lemma 2.3.19. *Olkoot $A \in \mathbb{R}^{m\times n}$ ja $B \in \mathbb{R}^{n\times k}$. Tällöin*

$$AB = \begin{bmatrix} - & A_1B & - \\ & \vdots & \\ - & A_mB & - \end{bmatrix}.$$

missä $A_j \in \mathbb{R}^{1\times n}$ on matriisin A rivi $A_{j\cdot}$. Erityisesti

$$B = \begin{bmatrix} - & e_1^*B & - \\ & \vdots & \\ - & e_m^*B & - \end{bmatrix}.$$

2.3.1 Matriisitulon perusominaisuudet

Matriisitulolla on seuraavat perusominaisuudet. Laskut ovat suoraviivaisia ja ne ensimmäisestä lukuunottamatta jätetään kiinnostuneelle lukijalle harjoitustehtäväksi.

Lause 2.3.20. *Olkoot $A, \tilde{A} \in \mathbb{R}^{m\times n}$, $B, \tilde{B} \in \mathbb{R}^{n\times k}$, $C \in \mathbb{R}^{k\times p}$ ja $\lambda \in \mathbb{R}$. Tällöin*

1. $(AB)C = A(BC)$,
2. $(A+\tilde{A})B = AB + \tilde{A}B$,
3. $A(B+\tilde{B}) = AB + A\tilde{B}$ *ja*
4. $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

Todistus. Todistetaan väite $(AB)C = A(BC)$. Huomaa, että $AB \in \mathbb{R}^{m\times k}$, $BC \in \mathbb{R}^{n\times p}$ ja $(AB)C, A(BC) \in \mathbb{R}^{m\times p}$.

Olkoot $j \in \{1,\dots,m\}$ ja $i \in \{1,\dots p\}$. Tällöin

$$\begin{aligned}((AB)C)_{ji} &= \sum_{r=1}^{k}(AB)_{jr}C_{ri} = \sum_{r=1}^{k}\left(\sum_{\ell=1}^{n}A_{j\ell}B_{\ell r}\right)C_{ri}\\ &= \sum_{r=1}^{k}\sum_{\ell=1}^{n}\sum_{r=1}^{k}A_{j\ell}B_{\ell r}C_{ri} = \sum_{\ell=1}^{n}A_{j\ell}\left(\sum_{r=1}^{k}B_{\ell r}C_{ri}\right)\\ &= \sum_{\ell=1}^{n}A_{j\ell}(BC)_{\ell i} = (A(BC))_{ji}\,.\end{aligned}$$

□

Huomautus 2.3.21. *Vaikka matriisitulon perusominaisuuksia ei todistettukaan, on niillä tärkeä merkitys. Ne antavat laskulait, joilla tuloja lasketaan. Erityisen huomattavaa on, että, että ensimmäinen väite $(AB)C = A(BC)$, joka on ns. matrisiitulon assoatiivisuus, sanoo, että matriisituloissa ei tarvitse kirjoittaa sulkuja ilmoittamaan järjestystä tulon laskemisessa. Näin ollen voidaan kirjoittaa ABC, koska sekä tulkinta $(AB)C$ että tulkinta $A(BC)$ antavat saman tuloksen. Tämä on erityisen käytännöllistä, kun matriisitulossa on useampia matriiseja.*

Huomautus 2.3.22. *Matriisitulon assosiatiivisuuden, eli ominaisuuden $(AB)C = A(BC)$ käsittelylle oli erityinen syy. Tämä ominaisuus nimittäin antaa uuden näkökulman matriisitulon määritelmälle kuvausten yhdistämisen kautta.*

Määritellään jokaisella matriisilla $M \in \mathbb{R}^{m\times n}$ kuvaus $f_M\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ kaavalla $x \mapsto Mx$.

Olkoot nyt $A \in \mathbb{R}^{m\times n}$ ja $B \in \mathbb{R}^{n\times k}$ matriiseja. Tällöin kuvaukset $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ ja $f_B\colon \mathbb{R}^{k\times 1} \to \mathbb{R}^{n\times 1}$ voidaan yhdistää kuvaukseksi $f_A \circ f_B\colon \mathbb{R}^{k\times 1} \to \mathbb{R}^{m\times 1}$, jolla on kaava

$$(f_A \circ f_B)(x) = f_A(f_B(x)) = f_A(Bx) = A(Bx)$$

jokaisella $x \in \mathbb{R}^{k\times 1}$.

Koska $A(Bx) = (AB)x$, niin

$$(f_A \circ f_B)(x) = A(Bx) = (AB)x = f_{AB}(x).$$

jokaisella $x \in \mathbb{R}^{k\times 1}$.

Matriisitulo liittyy siis näin kuvausten yhdistämiseen.

2.4 Neliömatriisien käänteismatriisit

2.4.1 Johdanto

Tavallisilla reaaliluvuilla yhtälö

$$ax = b$$

ratkaistaan jakamalla luvulla a, jos $a \neq 0$. Herää kysymys onko voidaanko matriisiyhtälöstä

$$Ax = b,$$

ratkaista sarakevektori x samalla tavalla, eli löytämällä sellainen matriisi A^{-1}, että

$$x = A^{-1}b.$$

Kuten reaalilukujenkin tapauksessa ei voida jakaa nollalla, ei tässäkään tapauksessa voida olettaa, että tälläistä matriisia A^{-1} on aina olemassa. Yksi tämän kurssin pääasioita onkin selvittää milloin matriisi A^{-1} voidaan löytää.

Myöhemmin tullaan osoittamaan, että ainoastaan sellaisella matriisilla, jolla on sama määrä rivejä kuin sarakeita, voi olla käänteismatriisi. Koska näillä matriiseilla tulee olemaan tärkeä rooli myös muissa lineaarialgeberan ilmiöissä, annetaan näille matriiseille jo tässä vaiheessa nimi.

Määritelmä 2.4.1. *Matriisi $A \in \mathbb{R}^{m \times n}$ on neliömatriisi, jos sillä on sama määrä rivejä ja sarakeita, eli $m = n$.*

2.4.2 Identiteettimatriisi

Reaalilukujen tapaukssa luvun $\lambda \in \mathbb{R} \setminus \{0\}$ käänteisluvulla λ^{-1} tarkoitetaan sellaista lukua, jolle pätee $\lambda\lambda^{-1} = 1$, missä luvulla 1 on tuttu ominaisuus $x \cdot 1 = x = 1x$ jokaisella $x \in \mathbb{R}$. Määritellään tätä varten yksikkö- eli identitettimatriisin käsite.

Merkintä 2.4.2. *Matriisia $I \in \mathbb{R}^{n \times n}$, jolle pätee*

$$I_{ji} = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}$$

sanotaan identiteettimatriisiksi.

Esimerkki 2.4.3. *Jos $n = 2$, niin identiteettimatriisi on*

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Jos $n = 3$, niin identiteettimatriisi on

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Huomautus 2.4.4. *Identiteettimatriisi $I \in \mathbb{R}^{n \times n}$ on siis jokaisella $n \in \mathbb{Z}_+$ matriisi*

$$I = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix}.$$

Matriisia I kutsutaan identiteettimatriisiksi seuraavasta syystä.

Lemma 2.4.5. *Jokaisella $x \in \mathbb{R}^{n\times 1}$ pätee $Ix = x$.*

Todistus. Olkoon

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Koska

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} e_{11} & \cdots & e_{1n} \\ \vdots & & \vdots \\ e_{n1} & \cdots & e_{nn} \end{bmatrix},$$

niin

$$Ix = \begin{bmatrix} e_{11} & \cdots & e_{1n} \\ e_{21} & \cdots & e_{2n} \\ \vdots & & \vdots \\ e_{n1} & \cdots & e_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} e_{11}x_1 + e_{12}x_2 + \cdots + e_{1n}x_n \\ e_{21}x_1 + e_{22}x_2 + \cdots + e_{2n}x_n \\ \vdots \\ e_{n1}x_1 + e_{n2}x_2 + \cdots + e_{nn}x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x.$$

□

Identiteettimatriisilla on sama ominaisuus kuin luvulla 1, eli $AI = A$ ja $IB = B$ aina, kun matriisitulo on määritelty. Kirjataan tämä lemmaksi.

Lemma 2.4.6. *Olkoon $I \in \mathbb{R}^{n\times n}$ identiteettimatriisi. Tällöin jokaisella $A \in \mathbb{R}^{m\times n}$ ja $B \in \mathbb{R}^{n\times k}$ pätee $AI = A$ ja $IB = B$.*

Todistus. Merkitään $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ ja $B = \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix}$, missä $a_1, \ldots, a_n \in \mathbb{R}^{m\times 1}$ ja $b_1, \ldots, b_k \in \mathbb{R}^{n\times 1}$.

Ensimmäinen väite seuraa matriisitulon sarakemuodosta ja vektoreiden e_i määritelmästä (eli tarkemmin lemmoista 2.3.8 ja 2.3.13) , sillä

$$AI = A \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} = \begin{bmatrix} Ae_1 & \cdots & Ae_n \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} = A.$$

Osoitetaan nyt $IB = B$. Lemman 2.3.8 2.4.5 perusteella pätee

$$IB = I \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix} = \begin{bmatrix} Ib_1 & \cdots & Ib_k \end{bmatrix} = \begin{bmatrix} b_1 & \cdots & Ib_k \end{bmatrix} = B.$$

Tämä päättää todistuksen. □

2.4.3 Käänteismatriisin määritelmä

Nyt voidaan määritellä matriisin käänteismatriisin käsite. Huomaa, että määritelmä ei sano, että jokaisella neliömatriisilla olisi käänteismatriisi.

Määritelmä 2.4.7. *Matriisi $B \in \mathbb{R}^{n\times n}$ on matriisin $A \in \mathbb{R}^{n\times n}$* käänteismatriisi, *jos $AB = I$ ja $BA = I$. Tällöin merkitään $B = A^{-1}$. Matriisia, jolla on käänteismatriisi kutsutaan* kääntyväksi.

Esimerkki 2.4.8. *Matriisin*
$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$
käänteismatriisi on
$$B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix},$$
sillä
$$AB = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1+0 & 1-1 \\ 0+0 & 0+1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$
ja vastaavasti
$$BA = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Muutama huomio määritelmästä on paikallaan.

Huomautus 2.4.9. *Käänteismatriisi on määritelty ainoastaan sellaiselle matriisille A jolla on sama määrä rivejä ja sarakeita. Syy tähän on seuraava. Jos $m > n$ ja $A \in \mathbb{R}^{m\times n}$, niin tällöin millään matriisilla $B \in \mathbb{R}^{n\times m}$ tulomatriisi AB ei ole identtinen matriisi $I \in \mathbb{R}^{m\times m}$. Tämä syvällinen tulos perustuu luvun 3 teoriaan.*

Huomautus 2.4.10. *Määritelmässä vaaditaan, että $AB = I$ ja $BA = I$. Itseasiassa riittäisi vaatia vain toinen ehdoista, vaikka yleisesti matriisitulossa ei tulontekijöiden paikkaa voida vaihtaa. Tämä syvällinen tulos seuraa lauseesta 3.10.1.*

Huomautus 2.4.11. *Mikäli matriisi on kääntyvä, niin sen käänteismatriisi on yksikäsitteinen. Tämä havaitaan seuraavasti. Oletetaan, että $A \in \mathbb{R}^{n\times n}$ neliömatriisi ja että matriisit $B \in \mathbb{R}^{n\times n}$ ja $C \in \mathbb{R}^{n\times n}$ ovat sellaisia, että $AB = BA = I$ ja $AC = CA = I$. Tällöin*
$$B = BI = B(AC) = (BA)C = IC = C.$$

2.4.4 Käänteismatriisin selvittäminen yhtälöryhmän avulla

Herää kysymys, mistä käänteismatriisi keksittiin esimerkissä 2.4.8. Se ratkaistaan vastaavasta yhtälöryhmästä, kuten seuraava esimerkki osoittaa.

Esimerkki 2.4.12. *Olkoon*
$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$
sama matriisi kuin edellisessä esimerkissä. Halutaan löytää sellainen matriisi $B \in \mathbb{R}^{2\times 2}$, että $AB = I$ ja $BA = I$.

Ratkaistaan ensin yhtälö $AB = I$ yhtälöryhmän avulla seuraavasti. Koska $I = \begin{bmatrix} e_1 & e_2 \end{bmatrix}$, niin riittää löytää sellaiset sarakevektorit $b_1, b_2 \in \mathbb{R}^{2\times 1}$, joille pätee $Ab_1 = e_1$ ja $Ab_2 = e_2$. Tämä siksi, että tällöin matriisille $B = \begin{bmatrix} b_1 & b_2 \end{bmatrix}$ pätee
$$AB = A\begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 \end{bmatrix} = I.$$

Yhtälöt $Ab_1 = e_1$ ja $Ab_2 = e_2$ voidaan ratkaista samanaikaisesti käyttämällä esimerkin 2.1.3 tapaa, eli ratkaisemalla yhtälöryhmä

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right].$$

Vähentämällä toinen rivi ensimmäisestä rivistä saadaan ekvivalentti yhtälöryhmä

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 \end{array}\right], \tag{2.7}$$

joka on supistetussa porrasmuodossa. Näin ollen saadaan kuten esimerkissä 2.1.3, että

$$b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ ja } b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

eli

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

Ratkaisu on siis itseasiassa juurikin se matriisi, joka on muodostunut yhtälöryhmän (2.7) *oikealla puolelle.*

Koska tässä vaiheessa on ratkaistu vasta yhtälö $AB = I$, niin pitää vielä tarkistaa, että matriisi B toteuttaa myös yhtälön $BA = I$. Tämä tarkistus voidaan tehdä suoralla laskulla.

Huomautus 2.4.13. *Edellisen esimerkin viimeinen huomio pätee yleisesti. Olkoon $A \in \mathbb{R}^{n\times n}$. Jos matriisiyhtälöä $AB = I$ vastaavan yhtälöryhmän*

$$\left[A \mid I\right]$$

supistettu porrasmuoto on

$$\left[I \mid B\right],$$

niin tällöin matriisi $B \in \mathbb{R}^{n\times n}$ on yhtälöryhmän $AB = I$ ratkaisu. Perustelu tälle huomiolle saadaan kirjoittamalla yhtälöryhmä $\left[I \mid B\right]$ kuten esimerkissä 2.4.12

Esimerkki 2.4.14. *Matriisillla*

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

ei ole käänteismatriisia. Tämä voidaan osoittaa, joko osoittamalla, että yhtälöllä $AB = I$ ei ole ratkaisua tai käyttämällä tässä luvussa kehitettävää teoriaa. Tässä esimerkissä tämä kysymys ratkaistaan kirjoittamalla yhtälö $AB = I$ yhtälöryhmäksi.

Merkitään $B = \begin{bmatrix} b_1 & b_2 \end{bmatrix}$, missä $b_1, b_2 \in \mathbb{R}^{2\times 1}$. Koska

$$AB = A\begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 \end{bmatrix},$$

niin saadaan, että yhtälö $AB = I$ vastaa yhtälöitä $Ab_1 = e_1$ ja $Ab_2 = e_2$. Toisaalta nämä yhtälöt voidaan kirjoittaa yhteen yhtälöryhmään kuten esimerkissä 2.1.3, jolloin saadaan yhtälöryhmä

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right]$$

Vähentämällä ensimmäinen rivi jälkimmäisestä saadaan yhtälöryhmä

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{array}\right]$$

jolla ei selvästi ole ratkaisuja. Näin ollen alkuperäisellä yhtälöllä $AB = I$ ei ole ratkaisua.

2.4.5 Käänteismatriisin perusominaisuudet

Käänteismatriisilla on seuraavat perusominaisuudet.

Lause 2.4.15. *Olkoot $A, B \in \mathbb{R}^{n\times n}$ kääntyviä matriiseja. Tällöin*

1. *tulomatriisi AB on kääntyvä ja $(AB)^{-1} = B^{-1}A^{-1}$ sekä*
2. *käänteismatriisi A^{-1} on kääntyvä ja $(A^{-1})^{-1} = A$.*

Todistus. Koska matriisit A ja B ovat kääntyviä, niin voidaan laskemalla selvittää onko matriisi $B^{-1}A^{-1}$ matriisin käänteismatriisi. Koska

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = A(BB^{-1})A^{-1} = AIA^{-1} = (AI)A^{-1} = AA^{-1} = I.$$

ja

$$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}B = I,$$

niin matriisi $B^{-1}A^{-1}$ on matriisin AB käänteismatriisi. Matriisi AB on siis kääntyvä.

Osoitetaan nyt, että A on matriisin A^{-1} käänteismatriisi $(A^{-1})^{-1}$. Tämä on lähes tautologia. Koska A^{-1} on matriisin käänteismatriisi, niin $(A^{-1})A = I$ ja $A(A^{-1}) = I$. Näin ollen A on matriisin A^{-1} käänteismatriisi, eli $(A^{-1})^{-1} = A$. Erityisesti matriisi A^{-1} on siis kääntyvä. □

2.5 Kääntyvyyslause

Matriisien kääntyvyyslause (lause 2.5.1) antaa kolme yhtäpitävää ehtoa neliömatriisin $A \in \mathbb{R}^{n\times n}$ kääntyvyydelle. Nämä ehdot liittyvät yhtälön $Ax = 0$ ratkaisuihin, yhtälöryhmän $\left[A \mid I\right]$ supistettuun porrasmuotoon ja matriisin A esittämiseen ns. alkeismatriisien avulla. Yhtälöryhmien rivioperaatioilla on keskeinen rooli lauseen todistuksessa, eli näiden ehtojen osoittamisessa yhtäpitäviksi.

Lause 2.5.1 (Neliömatriisien kääntyvyyslause)**.** *Olkoon $A \in \mathbb{R}^{n\times n}$ neliömatriisi. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *A on kääntyvä,*

2. *yhtälön $Ax = 0$ ainoa ratkaisu on $x = 0$.*

3. *yhtälöryhmän $\left[A \mid 0\right]$ supistettu porrasmuoto on $\left[I \mid 0\right]$ ja*

4. *matriisi A voidaan kirjoittaa alkeismatriisien tulona.*

Lauseen todistus jaetaan kahteen osaan. Ensimmäisessä vaiheessa todistetaan implikaatiot $1 \Rightarrow 2 \Rightarrow 3$, joita varten ei tarvita aleismatriisien määritelmää. Tämän jälkeen määritellään alkeismatriisit ja käsitellä niiden yhteyttä rivioperaatioihin. Tämän jälkeen todistetaan loput implikaatiot $3 \Rightarrow 4 \Rightarrow 1$. Yhdistämällä todistukset saadaan, että väitteet ovat yhtäpitäviä.

Lauseen 2.5.1 todistuksen alkuosa. Osoitetaan, että $(1) \Rightarrow (2)$, eli oletetaan, että A on kääntyvä matriisi ja osoitetaan, että yhtälön $Ax = 0$ ainoa ratkaisu on $x = 0$. Olkoon $x \in \mathbb{R}^{n\times 1}$ yhtälön $Ax = 0$ ratkaisu. Koska A on kääntyvä, niin $x = A^{-1}Ax = A^{-1}0 = 0$. Näin ollen väite seuraa.

Osoitetaan nyt, että $(2) \Rightarrow (3)$, eli oletetaan, että yhtälön $Ax = 0$ ainoa ratkaisu on $x = 0$. Olkoon $\left[\tilde{A} \mid 0\right]$ yhtälöryhmän $\left[A \mid 0\right]$ supistettu porrasmuoto. Koska $x = 0$ on yhtälön $Ax = 0$ ainoa ratkaisu, niin lauseen 1.4.14 perusteella jokaisella yhtälöryhmän $\left[\tilde{A} \mid 0\right]$ rivillä on täsmälleen yksi nollasta poikkeava kerroin. Tällöin supistetun porrasmuodon vaatimusten mukaan supistettu porrasmuoto on $\left[I \mid 0\right]$. □

2.5.1 Alkeismatriisit

Alkeismatriisit ovat kääntyviä matriiseja, joilla kertominen vastaa yhtälöryhmän rivioperaatioita. Kuten rivioperaatioita niin myös alkeismatriiseja on siis periaatteessa kolmea tyyppiä: niihin, jotka vaihtavat rivien järjestystä, niihin, jotka lisäävät rivin toiseen riviin, ja niihin, jotka kertovat riviä vakiolla. Näistä kaksi viimeistä voidaan kuitenkin myös yhdistää.

Rivioperaatioiden avulla annettu yhtälöryhmä voitiin saattaa supistettuun porrasmuotoon. Vastaava tulos matriisien kielellä on, että annettu matriisi voidaan alkeismatriiseilla kertomalla viedä supistettuun porrasmuotoon.

Matriisin rivin kertominen vakiolla

Aloitetaan helpoimmasta rivioperaatiosta, eli rivin kertomisesta vakiolla. Otetaan käyttöön seuraava merkintä.

Merkintä 2.5.2. *Olkoot $1 \le p \le n$ ja $\lambda \in \mathbb{R}$. Matriisi $D_{p,\lambda}$ on se matriisi, jonka p:s* *rivi on rivivektori λe_p^* ja jokaisella $j \neq p$, rivi j on rivivektori e_j^*, eli*

$$D_{p,\lambda} = \begin{bmatrix} - & e_1^* & - \\ & \vdots & - \\ - & \lambda e_p^* & - \\ - & \vdots & - \\ - & e_n^* & - \end{bmatrix}$$

Esimerkki 2.5.3. *Olkoon $n = 3$. Tällöin*

$$D_{1,\lambda} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ D_{2,\lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ ja } D_{3,\lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Huomautus 2.5.4. *Matriisi $D_{p,\lambda}$ voidaan määritellä myös kertoimittain kaavalla*

$$(D_{p,\lambda})_{ji} = \begin{cases} 1, & j = i, \ i \neq p \\ \lambda, & j = i = p \\ 0, & j \neq i \end{cases}$$

kaikilla $j, i \in \{1, \ldots, n\}$ tai sarakkeittain kaavalla

$$D_{p,\lambda} = \begin{bmatrix} e_1 & \cdots & e_{p-1} & \lambda e_p & e_{p+1} & \cdots & e_n \end{bmatrix}.$$

Huomautus 2.5.5. *Selvästi matriisi $D_{p,\lambda}$ on kääntyvä, jos ja vain jos $\lambda \neq 0$. Lisäksi suora lasku osoittaa, että $D_{p,\lambda}^{-1} = D_{p,1/\lambda}$ kaikilla $1 \leq p \leq n$ ja $\lambda \neq 0$.*

Osoitetaan nyt, että matriiseilla $D_{p,\lambda}$ on halutut ominaisuudet, eli että matriisitulo $D_{p,\lambda}A$ vastaa matriisin A rivin p kertomista luvulla λ. Tarkastellaan ensin erikoistapausta.

Esimerkki 2.5.6. *Olkoon $n = 3$ ja*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Tällöin tulomatriiisit $D_{p,\lambda}A$ saadaan laskuista

$$\begin{array}{cc} & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \end{array}, \quad \begin{array}{cc} & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \end{array}$$

ja

$$
\begin{array}{cc}
 & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix} & \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{bmatrix}
\end{array}.
$$

Näin ollen matriisi $D_{p,\lambda}A$ saadaan matriisista A kertomalla A:n p:nnen rivin alkiot vakiolla λ.

Todistetaan nyt yleinen tulos, että matriisilla $D_{p,\lambda}$ kertominen vastaa p:nnen rivin kertomista vakiolla λ.

Lemma 2.5.7. *Olkoot $1 \leq p \leq n$ ja $\lambda \in \mathbb{R}$. Tällöin jokaisella*

$$
A = \begin{bmatrix} - & a_1^* & - \\ & \vdots & \\ - & a_n^* & - \end{bmatrix} \in \mathbb{R}^{n \times n},
$$

pätee

$$
D_{p,\lambda}A = \begin{bmatrix} - & a_1^* & - \\ & \vdots & \\ - & \lambda a_p^* & - \\ & \vdots & \\ - & a_n^* & - \end{bmatrix}.
$$

Todistus. Koska

$$
e_j^* A = a_j^*
$$

jokaisella $j \in \{1, \dots, n\}$, niin lemman 2.3.19 perusteella

$$
D_{p,\lambda}A = \begin{bmatrix} - & e_1^* A & - \\ & \vdots & \\ - & \lambda e_p^* A & - \\ & \vdots & \\ - & e_n^* A & - \end{bmatrix} = \begin{bmatrix} - & a_1^* & - \\ & \vdots & \\ - & \lambda a_p^* & - \\ & \vdots & \\ - & a_n^* & - \end{bmatrix}.
$$

$\square$

Matriisin rivien järjestyksen vaihtaminen matriisitulon avulla

Aloitetaan rivioperaatiosta, joka vaihtaa yhtälöryhmän rivien järjestyksen. Paljastuu, että tämä palautuu identtisen matriisin rivien järjestyksen vaihtamiseen.

Merkintä 2.5.8. *Olkoot $1 \leq p < r \leq n$. Matriisi $S^{pr} \in \mathbb{R}^{n\times n}$ on se matriisi, joka saadaan matriista I vaihtamalla rivien p ja r paikkaa, eli*

$$S^{pr} = \begin{bmatrix} - & e_1^* & - \\ & \vdots & \\ - & e_r^* & - \\ & \vdots & \\ - & e_p^* & - \\ & \vdots & \\ - & e_n^* & - \end{bmatrix}$$

Huomautus 2.5.9. *Matriisi S^{pr} voidaan määritellä kertoimittain kaavalla*

$$(S^{pr})_{ji} = \begin{cases} 1, & (j,i) = (p,r) \text{ tai } (j,i) = (r,p) \\ 1, & i = j \text{ ja } j \neq p, r \\ 0, & \text{muutoin.} \end{cases}$$

Esimerkki 2.5.10. *Olkoon $n = 3$. Tällöin*

$$S^{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ S^{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ S^{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ ja } S^{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Seuraava esimerkki näytää, että tapauksessa $n = 3$ matriisit S^{pr} vaihtavat matriisin A rivit p ja r päittäin.

Esimerkki 2.5.11. *Olkoon*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Tällöin

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

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{bmatrix} \text{ ja}$$

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

Todistetaan nyt yleinen tulos, että matriisin A rivien p ja r paikka voidaan vaihtaa kertomalla matriisilla S^{pr}.

Lemma 2.5.12. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi*

$$A = \begin{bmatrix} - & A_1 & - \\ & \vdots & \\ - & A_m & - \end{bmatrix},$$

missä $A_j \in \mathbb{R}^{1\times n}$ on matriisin A rivivektori $A_{j\cdot}$. Tällöin kaikilla $1 \le p < r \le m$ pätee

$$\begin{bmatrix} - & A_1 & - \\ & \vdots & \\ - & A_r & - \\ & \vdots & \\ - & A_p & - \\ & \vdots & \\ - & A_m & - \end{bmatrix} = S^{pr}A.$$

Todistus. Olkoot $1 \le p < r \le m$. Tällöin lemman 2.3.19 nojalla

$$S^{pr}A = \begin{bmatrix} - & (S^{pr})_{1\cdot}A & - \\ & \vdots & \\ - & (S^{pr})_{m\cdot}A & - \end{bmatrix} = \begin{bmatrix} - & e_1^*A & - \\ & \vdots & \\ - & e_r^*A & - \\ & \vdots & \\ - & e_p^*A & - \\ & \vdots & \\ - & e_m^*A & - \end{bmatrix} = \begin{bmatrix} - & A_1 & - \\ & \vdots & \\ - & A_r & - \\ & \vdots & \\ - & A_p & - \\ & \vdots & \\ - & A_m & - \end{bmatrix}.$$

□

Koska matriisin $S^{pr}A$ rivit p ja r voidaan vaihtaa takaisin alkuperäisille paikoilleen kertomalla uudelleen matriisilla S^{pr} saadaan, että $S^{pr}S^{pr}A = A$ kaikilla $1 \le p < r \le m$ ja kaikilla matriisilla $A \in \mathbb{R}^{m\times n}$. Näin ollen erityisesti

$$S^{pr}S^{pr} = S^{pr}S^{pr}I = I$$

eli $(S^{pr})^{-1} = S^{pr}$. Matriisi S^{pr} on siis kääntyvä ja itsensä käänteismatriisi.

Matriisin rivin lisääminen toiseen matriisitulon avulla

Määritellään ensin apumatriisit $E^{pr} \in \mathbb{R}^{n\times n}$ kaikilla $p, r \in \{1, \dots, n\}$, joilla $p \neq r$.

Merkintä 2.5.13. *Olkoot $p, r \in \{1, \dots n\}$, $p \neq r$. Matriisi $E^{pr} \in \mathbb{R}^{n\times n}$ on se matriisi, jolle pätee*

$$(E^{pr})_{ji} = \begin{cases} 1, & (j,i) = (p,r) \\ 0, & (j,i) \neq (p,r). \end{cases}$$

Huomautus 2.5.14. *Matriisiin E^{pr} liittyvät ehdot voidaan kirjoittaa ns. Kroneckerin deltan avulla. Kroneckerin deltalla tarkoitetaan lukua $\delta_{ji} \in \{0, 1\}$, joka määräytyy kaavasta*

$$\delta_{ji} = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}$$

Tällä merkinnällä matriisille E^{pr} pätee $(E^{pr})_{ji} = \delta_{pj}\delta_{ri}$ kaikilla $j, i \in \{1, \ldots, n\}$.[1]

Esimerkki 2.5.15. *Olkoon $n = 2$. Tällöin*

$$E^{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ ja } E^{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Seuraava esimerkki selittää matriisien E^{pr} tarkoituksen.

Esimerkki 2.5.16. *Olkoon*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Tällöin

$$E^{12}A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} \\ 0 & 0 \end{bmatrix}$$

ja

$$E^{21}A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ a_{11} & a_{12} \end{bmatrix}.$$

Edellisestä esimerkistä voi päätellä, että heuristisesti tulomatriisi $E^{pr}A$ poimii matriisin A r:nnen rivin ja sijoittaa sen p:nnelle riville. Näin ollen matriisin A kertominen matriisilla $I + E^{pr}$ vastaa rivioperaatiota, jossa matriisin A riviin p lisätään rivi r, sillä

$$(I + E^{pr})A = IA + E^{pr}A = A + E^{pr}A.$$

Tämä rivioperaatio kumoutuu, kun rivistä p vähennetään rivi r, eli kun matriisia A kerrotaan matriisilla $I - E^{pr}$. Matriisien kielellä tämä tarkoittaa, että matriisi $I - E^{pr}$ on matriisin $I + E^{pr}$ käänteismatriisi. Kirjataan tämä tulos lemmaksi.

Lemma 2.5.17. *Olkoot $p, r \in \{1, \ldots, n\}$, $p \neq r$. Tällöin*

$$(I + E^{pr})^{-1} = I - E^{pr}.$$

Todistus. Havaitaan aluksi, että

$$(I + E^{pr})(I - E^{pr}) = I(I - E^{pr}) + E^{pr}(I - E^{pr}) = I - E^{pr} + E^{pr} - E^{pr}E^{pr} = I - E^{pr}E^{pr}$$

ja että vastavasti

$$(I - E^{pr})(I + E^{pr}) = I(I + E^{pr}) - E^{pr}(I + E^{pr}) = I + E^{pr} - E^{pr} - E^{pr}E^{pr} = I - E^{pr}E^{pr}.$$

[1]Kroneckerin deltan voi määritellä myös funktiona. Kysy luennoitsijalta, mitä tällä tarkoitetaan.

Näin ollen riittää osoittaa, että $E^{pr}E^{pr} = 0$.

Olkoot $j, i \in \{1, \ldots, n\}$. Jos $j \neq p$, niin $(E^{pr})_{jk} = 0$ kaikilla $k \in \{1, \ldots, n\}$, joten

$$(E^{pr}E^{pr})_{ji} = \sum_{k=1}^{n} (E^{pr})_{jk}(E^{pr})_{ki} = 0.$$

Toisaalta, jos $j = p$, niin tällöin $(E^{pr})_{jk} \neq 0$ ainoastaan tapauksessa $k = r$. Koska $p \neq r$, niin $(E^{pr})_{ri} = 0$ kaikilla $i \in \{1, \ldots, n\}$. Näin ollen

$$(E^{pr}E^{pr})_{ji} = \sum_{k=1}^{n} (E^{pr})_{jk}(E^{pr})_{ki} = (E^{pr})_{pr}(E^{pr})_{ri} = 0.$$

Tulomatriisi $E^{pr}E^{pr}$ on siis nollamatriisi. □

Huomautus 2.5.18. *Kuten luvussa 1 havaittiin, usein rivioperaatioita yhdistellään. Tyypillinen esimerkki on, että matriisin A riviin p lisätään rivi r kerrottuna luvulla λ. Matriisioperaationa tämä vastaa matriisilla $I + \lambda E^{pr}$ kertomista.*

Tarkasti ottaen tämä yhdistetty rivioperaatio voidaan kuitenkin kolmivaiheisena: Kerrotaan matriisin A r:s rivi vakiolla λ, lisätään rivi r riviin p, jaetaan rivi r vakiolla λ. Tätä rivioperaatioiden yhdistettä vastaa tulomatriisi $D_{r,1/\lambda}(I + E^{pr})D_{r,\lambda}$.

Perustellaan, että nämä kaksi matriisia ovat samat. Havaitaan aluksi, että

$$\begin{aligned} D_{r,1/\lambda}(I + E^{pr})D_{r,\lambda} &= \left(D_{r,1/\lambda}I + D_{r,1/\lambda}E^{pr}\right) D_{r,\lambda} \\ &= \left(D_{r,1/\lambda} + D_{r,1/\lambda}E^{pr}\right) D_{r,\lambda} \\ &= D_{r,1/\lambda}D_{r,\lambda} + D_{r,1/\lambda}E^{pr}D_{r,\lambda} \\ &= I + D_{r,1/\lambda}E^{pr}D_{r,\lambda}. \end{aligned}$$

Näin ollen riittää osoittaa, että

$$D_{r,1/\lambda}E^{pr}D_{r,\lambda} = \lambda E^{pr}.$$

Tämän voi osoittaa kahdessa vaiheessa seuraavasti.

Osoitetaan ensin, että $D_{r,1/\lambda}E^{pr} = E^{pr}$. Olkoot $j, i \in \{1, \ldots, n\}$. Jos $j \neq r$, niin

$$(D_{r,1/\lambda}E^{pr})_{ji} = \sum_{k=1}^{n} (D_{r,1/\lambda})_{jk}(E^{pr})_{ki} = (D_{r,1/\lambda})_{jj}(E^{pr})_{ji} = (E^{pr})_{ji}.$$

Toisaalta, jos $j = r$, niin tällöin $(E^{pr})_{ji} = 0$ kaikilla $i \in \{1, \ldots, n\}$, koska $p \neq r$. Tällöin

$$(D_{r,1/\lambda}E^{pr})_{ji} = \sum_{k=1}^{n} (D_{r,1/\lambda})_{jk}(E^{pr})_{ki} = 0(E^{pr})_{ji} = 0 = (E^{pr})_{ji}.$$

Näin ollen $D_{r,1/\lambda}E^{pr} = E^{pr}$.

Osoitetaan nyt, että $E^{pr}D_{r,\lambda} = \lambda E^{pr}$. Olkoot $j, i \in \{1, \ldots, n\}$. Jos $i = r$, niin

$$(E^{pr}D_{r,\lambda})_{ji} = \sum_{k=1}^{n}(E^{pr})_{jk}(D_{r,\lambda})_{ki} = (E^{pr})_{ji}\lambda = \lambda(E^{pr})_{ji}.$$

Toisaalta, jos $i \neq r$, niin $(E^{pr})_{ji} = 0$. Tällöin

$$(E^{pr}D_{r,\lambda})_{ji} = \sum_{k=1}^{n}(E^{pr})_{jk}(D_{r,\lambda})_{ki} = 0 = \lambda(E^{pr})_{ji}.$$

Näin ollen $E^{pr}D_{r,\lambda} = \lambda E^{pr}$

Yhdistämällä tehdyt havainnot saadaan

$$I + D_{r,1/\lambda}E^{pr}D_{r,\lambda} = I + (D_{r,1/\lambda}E^{pr})D_{r,\lambda} = I + E^{pr}D_{r,\lambda} = I + \lambda E^{pr}.$$

2.5.2 Kääntyvyyslauseen todistus

Matriiseja S^{pr}, $D_{r,\lambda}$ ja $I + E^{pr}$ kutsutaan alkeismatriiseiksi. Edellä tehtyjen havaintojen perusteella ne ovat kääntyviä. Lauseen 2.4.15 perusteella myös niiden tulot ovat kääntyviä. Kääntyvyyslause sanoo, että itseasiassa kaikki kääntyvät matriisit ovat tuloja alkeismatriiseista. Kääntyvyyslause siis sitoo yhteen kaikki tähän mennessä käsitellyt asiat.

Neliömatriisien kääntyvyyslauseen (lause 2.5.1) todistuksen loppuosa. Osoitetaan implikaatiot (3) ⇒ (4) ⇒ (1). Yhdessä jo todistettujen implikaatioiden kanssa tämä todistaa oletusten yhtäpitävyyden.

Osoitetaan nyt, että (3) ⇒ (4). Koska yhtälöryhmä $\left[A \mid 0\right]$ vastaa yhtälöä

$$Ax = 0$$

ja yhtälöryhmä $\left[A \mid 0\right]$ saadaan rivioperaatioilla supistettuun porrasmuotoon $\left[I \mid 0\right]$, niin on olemassa näitä rivioperaatioita vastaavat alkeismatriisit $B_1, \ldots, B_k$, missä jokaisella $\ell \in \{1, \ldots, k\}$ joko $B_\ell = S^{pr}$, $B_\ell = D_{r,\lambda}$ tai $B_\ell = I + E^{rp}$ jollain $r, p \in \{1, \ldots, n\}$, $p \neq r$, ja $\lambda \in \mathbb{R}$, eli

$$\left[A \mid 0\right] \sim \left[B_1 A \mid 0\right] \sim \left[B_2 B_1 A \mid 0\right] \sim \cdots \sim \left[B_k \cdots B_2 E_1 A \mid 0\right],$$

missä

$$B_k B_{k-1} \cdots B_2 B_1 A = I.$$

Näin ollen

$$A = (B_k B_{k-1} \cdots B_2 B_1)^{-1} = B_1^{-1} \cdots B_k^{-1}.$$

Koska alkeismatriisien käänteismatriisit ovat alkeismatriiseja, niin A on alkeismatriisien tulo.

Osoitetaan nyt viimeinen väite eli (4) ⇒ (1). Oletetaan, että A on alkeismatriisien tulo, eli $A = B_1 \cdots B_k$, joillain alkeismatriiseilla $B_1, \ldots, B_k \in \mathbb{R}^{n \times n}$. Osoitetaan, että matriisi $B = B_k^{-1} \cdots B_1^{-1}$ on matriisin A käänteismatriisi.

Koska

$$\begin{aligned} AB &= (B_1 \cdots B_k)(B_k^{-1} \cdots B_1^{-1}) = B_1 \cdots B_{k-1} B_k B_k^{-1} B_{k-1}^{-1} \cdots B_1^{-1} \\ &= B_1 \cdots B_{k-1} B_{k-1}^{-1} \cdots B_1^{-1} = \cdots = B_1 B_1^{-1} = I \end{aligned}$$

ja

$$\begin{aligned} BA &= (B_k^{-1} \cdots B_1^{-1})(B_1 \cdots B_k) = B_k^{-1} \cdots B_2^{-1} B_1^{-1} B_1 B_2 \cdots B_k \\ &= B_k^{-1} \cdots B_2^{-1} B_2 \cdots B_k = \cdots = B_k^{-1} B_k = I, \end{aligned}$$

niin B on matriisin A käänteismatriisi. Erityisesti A on kääntyvä. $\square$

Luku 3

Aliavaruudet

Tavoitteet

- Käsitteet: lineaarikombinaatio, vapaus, virittäminen ja kanta.
- Matriisin sarakeavaruuden yhteys yhtälön $Ax = b$ ratkaisemiseen ja nolla-avaruuden yhteys ratkaisun yksikäsitteisyyteen.
- Matriisin sarake- ja nolla-avaruuksien kantojen yhteys yhtälöryhmän sidottuihin ja vapaisiin muuttujiin.

3.1 Motivointi: Homogeeniset yhtälöryhmät

Tarkastellaan matriisiyhtälöä

$$Ax = b$$

missä $A \in \mathbb{R}^{m\times n}$ ja $b \in \mathbb{R}^{n\times 1}$.

Luvuissa 1 ja 2 ei erityisesti korostettu, mutta niiden perusteella voidaan kuitenkin havaita, että yhtälöiden $Ax = b$ ja $Ax = 0$ ratkaisuilla on jotakin yhteistä. Tarkemmin sanottuna kaikki yhtälön $Ax = b$ ratkaisut saadaan muodostamalla yhdestä ainoasta yhtälön $Ax = b$ ratkaisusta lisäämällä siihen yhtälön $Ax = 0$ ratkaisuja. Tämän voi muotoilla esimerkiksi seuraavasti.

Lemma 3.1.1. *Olkoot $A \in \mathbb{R}^{m\times n}$ ja $b \in \mathbb{R}^{n\times 1}$. Oletetaan, että yhtälöllä $Ax = b$ on jokin ratkaisu $y \in \mathbb{R}^{m\times 1}$. Tällöin vektorille $x \in \mathbb{R}^{m\times 1}$ pätee $Ax = b$ ratkaisu, jos ja vain jos $x = y + z$, missä $z \in \mathbb{R}^{m\times 1}$ on sellainen vektori, että $Az = 0$.*

Todistus. Olkoot $y \in \mathbb{R}^{m\times 1}$ ja $z \in \mathbb{R}^{m\times 1}$ sellaisia vektoreita, että $Ay = b$ ja $Az = 0$. Tällöin

$$A(y + z) = Ay + Az = b + 0 = b,$$

eli vektori $y+z$ on yhtälön $Ax = b$ ratkaisu. Toisaalta, jos $y \in \mathbb{R}^{m\times 1}$ on sellainen vektori, että $Ay = b$, ja $y' \in \mathbb{R}^{m\times 1}$ toinen sellainen vektori, että $Ay' = b$, niin tällöin vektorille $z = y' - y$ pätee

$$Az = A(y' - y) = Ay' - Ay = b - b = 0$$

eli z on yhtälön $Ax = 0$ ratkaisu ja $y' = y + z$. □

Yhtälöä $Ax = 0$ kutsutaan *yhtälöä $Ax = b$ vastaavaksi homogeeniseksi yhtälöksi* ja sen ratkaisuilla on mielenkiintoinen ominaisuus, että ratkaisujen summat ja skaalaukset ovat myös ratkaisuja. Tämän voi puolestaan kirjoittaa seuraavasti.

Lemma 3.1.2. *Olkoot $A \in \mathbb{R}^{m\times n}$ ja $b \in \mathbb{R}^{m\times 1}$. Jos $y, y' \in \mathbb{R}^{n\times 1}$ ovat sellaisia, että $Ay = 0$ ja $Ay' = 0$, niin tällöin kaikilla $a, b \in \mathbb{R}$ vektori $x = ay + by' \in \mathbb{R}^{n\times 1}$ on yhtälön $Ax = 0$ ratkaisu.*

Todistus. Tämä on suora lasku:

$$A(ay + by') = A(ay) + A(by') = aAy + bAy' = 0 + 0 = 0.$$

□

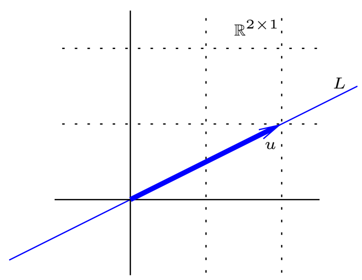
Homogeenisen yhtälön $Ax = 0$ ratkaisujoukkoa kutsutaan näiden ominaisuuksien vuoksi aliavaruudeksi. Aliavaruudelle voidaan määritellä kanta – ja lopulta dimensio – ja siten tarkastella yhtälön $Ax = b$ ratkaisujoukon suuruutta tarkemmin.

Aliavaruudet geometrisesti

Aliavaruuden määritelmää voidaan lähestyä myös geometrisesti, joka johtaa aliavaruuden virittämisen käsitteeseen. Kuten jo tiedetään, avaruuden $\mathbb{R}^{n\times 1}$ sarakevektoreita voidaan laskea yhteen ja kertoa reaaliluvuilla, eli skalaareilla. Sarakeavaruuden $\mathbb{R}^{n\times 1}$ origon kautta kulkevilla suorilla ja tasoilla on samankaltainen ominaisuus kuin yhtälön $Ax = 0$ ratkaisuilla eli vektoreiden summat ja venytykset kuuluvat samalla suoralle tai tasolle.

Esimerkki 3.1.3. *Olkoon $u = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^{2\times 1}$ ja määritellään $L = \{tu \in \mathbb{R}^{2\times 1} : t \in \mathbb{R}\}$.*

Tällöin L on origon kautta kulkeva suora, sillä $0 = 0u \in L$. On huomattavaa, että kaikki suoran L vektorit ovat annetun vektorin u venytyksiä tu. Myöhemmin tässä luvussa määritellään, että suora L on vektorin u virittämä. Suora L ja vektori u on havainnollistettu kuvassa 3.1.



Kuva 3.1: Suora L ja vektori u sarakeavaruudessa $\mathbb{R}^{2\times 1}$.

Olkoot $v, w \in L$. Tällöin $v = tu$ ja $w = su$ jollain $t, s \in \mathbb{R}$. Näin ollen $v + w = tu + su = (t + s)u$. Koska $t + s \in \mathbb{R}$, niin $(t + s)u \in L$ ja $v + w \in L$. Toisaalta, jos $a \in \mathbb{R}$, niin $av = a(tu) = (at)u \in L$. Näin ollen L sisältää kaikki vektoreidensa summat ja venytykset.

Muokataan edellistä esimerkkiä hieman.

Esimerkki 3.1.4. *Olkoot*

$$u = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \ \tilde{u} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \in \mathbb{R}^{3\times 1}$$

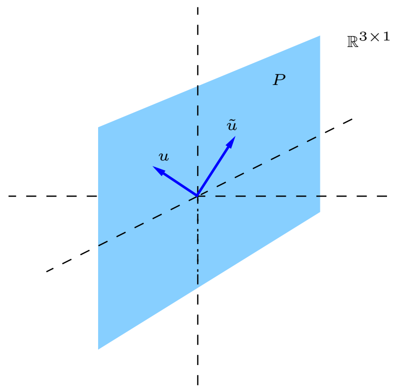
ja määritellään

$$P = \{tu + s\tilde{u} \in \mathbb{R}^{3\times 1} \colon t, s \in \mathbb{R}\}.$$

Nyt P on origon kautta kulkeva taso, sillä $0 = 0u + 0\tilde{s} \in P$. Koska kaikki tason P vektorit saadaan vektoreiden u ja $\tilde{u}$ venytysten summina sanotaan, että tason P vektorit ovat vektoreiden u ja $\tilde{u}$ lineaarikombinaatioita ja että vektorit u ja $\tilde{u}$ virttävät aliavaruuden P. Taso P, joka on itseasiassa joukko

$$P = \{ \begin{bmatrix} 0 \\ y \\ z \end{bmatrix} \in \mathbb{R}^{3\times 1} \colon y, z \in \mathbb{R}\}$$

on havainnollistettu kuvassa 3.2.



Kuva 3.2: Taso P ja vektorit u ja $\tilde{u}$ sarakeavaruudessa $\mathbb{R}^{3\times 1}$.

Aivan samalla argumentilla kuin edellisessä esimerkissä osoitetaan, että P sisältää vektoreidensa summat ja venytykset. Olkoot $v_1, v_2 \in P$. Tällöin $v_1 = t_1 u + s_1 \tilde{u}$ ja $v_2 = t_2 u + s_2 \tilde{u}$ jollain $t_1, t_2, s_1, s_2 \in \mathbb{R}$. Näin ollen

$$v + w = t_1 u + s_1 \tilde{u} + t_2 u + s_2 \tilde{u} = (t_1 + t_2)u + (s_1 + s_2)\tilde{u}.$$

Koska $t_1 + t_2 \in \mathbb{R}$ ja $s_1 + s_2 \in \mathbb{R}$, niin $v + w \in P$ Toisaalta, jos $a \in \mathbb{R}$, niin

$$av = a(t_1u + s_1\tilde{u}) = (at_1)u + (as_1)\tilde{u} \in P.$$

Näin ollen P on aliavaruus.

3.2 Aliavaruuden määritelmä

Avaruuden $\mathbb{R}^{n\times 1}$ epätyhjää osajoukkoa, joka sisältää kaikki vektoreidensa summat ja skaalaukset, kutsutaan aliavaruudeksi. Yleensä oletus, että joukko on epätyhjä korvataan oletuksella, että joukko sisältää avaruuden $\mathbb{R}^{n\times 1}$ origon.

Määritelmä 3.2.1. *Osajoukko $V \subset \mathbb{R}^{n\times 1}$ on* avaruuden $\mathbb{R}^{n\times 1}$ aliavaruus, *jos $0 \in V$ ja kaikilla $v, w \in V$ ja $a \in \mathbb{R}$ pätee sekä*

- *$v + w \in V$ että*
- *$av \in V$.*

Huomautus 3.2.2. *Useimmiten aliavaruuden määritelmä annetaan avaruuden $\mathbb{R}^n$ tapauksessa tai yleisemmin vektoriavaruuksien osajoukoille. Kuten on jo huomattu, avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^{n\times 1}$ välillä ei kuitenkaan ole suurta eroa: molemmat koostuvat alkioista, jotka määräytyvät n:stä reaaliluvusta, toisen alkiot kirjoitetaan jonona ja toisen sarakkeena. Näin ollen merkintöjä $\mathbb{R}^n$ ja $\mathbb{R}^{n\times 1}$ voitaisiin käyttää ristiin tässä luvussa ilman suurempia ongelmia. Avaruus $\mathbb{R}^{n\times 1}$ on valittu tässä määritelmän lähtökohdaksi, koska luvun tärkeimmät sovelluskohteet liittyvät matriiseihin ja sarakevektoreihin.*

Huomautus 3.2.3. *Aliavaruuden määritelmän ensimmäinen ehto yleistyy vektoreiden äärellisille summille, eli että $v_1 + \cdots + v_k \in V$ kaikilla $v_1, \ldots, v_k \in V$. Tämä havaitaan seuraavasti. Olkoon $V \subset \mathbb{R}^{n\times 1}$ aliavaruus ja olkoot $v_1, \ldots, v_k \in V$. Tällöin $v_1 + v_2 \in V$ aliavaruuden määritelmän perusteella. Jos $v_1 + \cdots + v_{j-1} \in V$ jollain $j \in \{1, \ldots, k\}$, niin $v_1 + \cdots + v_{j-1} + v_j = (v_1 + \cdots + v_{j-1}) + v_j \in V$ aliavaruuden määritelmän perusteella. Näin ollen induktiolla saadaan, että $v_1 + \cdots + v_k \in V$.*

Jo käsiteltyjen esimerkkien lisäksi helpoimmat esimerkit aliavaruuksista ovat $\{0\}$ ja koko avaruus $\mathbb{R}^{n\times 1}$.

Esimerkki 3.2.4. *Osajoukko, joka sisältää ainoastaan nolla-vektorin 0, eli osajoukko $\{0\} \subset \mathbb{R}^{n\times 1}$ on aliavaruus. Formaalisti tämä havaitaan seuraavasti. Olkoot $v, w \in \{0\}$ ja $a \in \mathbb{R}$. Tällöin $v = 0$ ja $w = 0$. Näin ollen $v + w = 0$. Lisäksi $av = a \cdot 0 = 0$.*

Esimerkki 3.2.5. *Koko avaruus $\mathbb{R}^{n\times 1}$ on itsensä osajoukko. Selvästi $0 \in \mathbb{R}^{n\times 1}$. Yhteenlaskun ja skalaarilla kertomisen määritelmien nojalla $v + w \in \mathbb{R}^{n\times 1}$ ja $av \in \mathbb{R}^{n\times 1}$ kaikilla $v, w \in \mathbb{R}^{n\times 1}$ ja $a \in \mathbb{R}$. Näin ollen $\mathbb{R}^{n\times 1}$ on (itsensä) aliavaruus.*

On helppo keksiä osajoukkoja, joiden vektoreita laskemalla yhteen tai skaalaarilla kertomalla ei saada saman osajoukon alkiota. Huomaa, että edes kaikki suorat ja tasot eivät ole aliavaruuksia kuten seuraavat esimerkit osoittavat.

Esimerkki 3.2.6. *Olkoon*

$$L = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^{2\times 1} \colon x_1 = 1,\ x_2 \in \mathbb{R} \right\} \subset \mathbb{R}^{2\times 1}.$$

Nyt $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *ja* $w = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ *ovat suoran* L *pisteitä, mutta piste*

$$v + w = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

ei kuulu suoraan L*, koska vektorin* $v+w$ *ensimmäinen koordinaatti ei ole* 1*. Vastaavasti myöskään vektori* $4v$ *ei kuulu suoraan* L*. Itseasiassa voidaan havaita, että kaikilla* $v, w \in L$ *pätee* $v + w \notin L$*, ja että* $av \in L$ *ainoastaan, jos* $a = 1$*.*

Edellisestä esimerkistä voi muokata esimerkin tasosta, joka ei ole aliavaruus.

Esimerkki 3.2.7. *Olkoon*

$$P = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \colon x_1 = 1,\ x_2 \in \mathbb{R},\ x_3 \in \mathbb{R} \right\} \subset \mathbb{R}^{3\times 1}.$$

Kuten edellisessä esimerkissä jälleen kaikilla vektoreilla $v, w \in P$ *pätee, että* $v + w \notin P$*.*

Aliavaruuden määritelmässä olevat ehdot voidaan myös yhdistää yhdeksi ehdoksi ns. *aliavaruuskriteerioksi.* Tämä nopeuttaa usein konkreettisen osajoukon osoittamista aliavaruudeksi.

Lemma 3.2.8. *Epätyhjä osajoukko* $V \subset \mathbb{R}^n$ *on aliavaruus, jos ja vain jos* $av + w \in V$ *kaikilla* $v, w \in V$ *ja* $a \in \mathbb{R}$*.*

Todistus. Oletetaan, että V on aliavaruus. Olkoot $v, w \in V$ ja $a \in \mathbb{R}$. Koska V on aliavaruus, niin $av \in V$. Näin ollen $(av) + w \in V$, koska V on aliavaruus.

Oletetaan nyt, että $V \subset \mathbb{R}^n$ on sellainen epätyhjä osajoukko, että $av + w \in V$ kaikilla $v, w \in V$ ja $a \in \mathbb{R}$. Osoitetaan, että V on aliavaruus. Olkoot $v, w \in V$ ja $a \in \mathbb{R}$. Koska $v = 1 \cdot v$, niin $v + w = (1 \cdot v) + w \in V$. Toisaalta oletuksen nojalla $(-1 \cdot w) + w \in V$. Näin ollen $0 = (-w) + w = (-1 \cdot w) + w \in V$ ja $av = (av) + 0 \in V$. Näin ollen V on aliavaruus. □

3.3 Esimerkki: Matriisin nolla-avaruus aliavaruutena

Matriisiin $A \in \mathbb{R}^{n\times k}$ voidaan aina liittää kaksi luonnollista aliavaruutta: matriisin sarakeavaruus $\mathrm{Col}(A)$, joka on avaruuden $\mathbb{R}^{n\times 1}$ aliavaruus ja matriisin nolla-avaruus $\mathrm{Null}(A)$, joka on avaruuden $\mathbb{R}^{k\times 1}$ aliavaruus.

Käsitellään näistä tärkeistä esimerkeistä ensin matriisin nolla-avaruutta, eli yhtälön $Ax = 0$ ratkaisujen joukkoa. Kerrataan vielä ennen formaalin määritelmän antamista, että tämä joukko todellakin on aliavaruus.

Lemma 3.3.1. *Olkoon $A \in \mathbb{R}^{n\times k}$. Tällöin joukko*

$$\mathrm{Null}(A) = \{x \in \mathbb{R}^{k\times 1} \colon Ax = 0\} \subset \mathbb{R}^{k\times 1}$$

on avaruuden $\mathbb{R}^{k\times 1}$ aliavaruus.

Todistus. Koska $A0 = 0$, niin $0 \in \mathrm{Null}(A)$. Näin ollen $\mathrm{Null}(A)$ on epätyhjä ja väite voidaan todistaa aliavaruuskriteerion (lemma 3.2.8) avulla. Olkoot $x, y \in \mathrm{Null}(A)$ ja $a \in \mathbb{R}$. Koska $Ax = 0$ ja $Ay = 0$, niin matriisitulon ominaisuuksien nojalla saadaan

$$A(ax + y) = A(ax) + Ay = A(ax) = a(Ax) = 0.$$

Näin ollen $ax + y \in \mathrm{Null}(A)$. Osajoukko $\mathrm{Null}(A) \subset \mathbb{R}^{k\times 1}$ on siis aliavaruus. □

Määritelmä 3.3.2. *Matriisin $A \in \mathbb{R}^{n\times k}$* nolla-avaruus *on aliavaruus*

$$\mathrm{Null}(A) = \{x \in \mathbb{R}^{k\times 1} \colon Ax = 0\}.$$

Huomautus 3.3.3. *Määrittelemällä kuvaus $f_A \colon \mathbb{R}^{k\times 1} \to \mathbb{R}^{n\times 1}$, $x \mapsto Ax$, voidaan tulkita, että* $\mathrm{Null}(A)$ *on itseasiassa origon alkukuva tässä kuvauksessa eli* $\mathrm{Null}(A) = f_A^{-1}(0)$.

Lauseen 1.4.14 tulosta voidaan tarkentaa matriisin A nolla-avaruuden käsitteen avulla.

Lause 3.3.4. *Olkoot $A \in \mathbb{R}^{m\times n}$ ja $b \in \mathbb{R}^{m\times 1}$. Jos yhtälöryhmällä $\left[A \mid b\right]$ on ratkaisu $x \in \mathbb{R}^{n\times 1}$, niin yhtälöryhmän $\left[A \mid b\right]$ koko ratkaisujoukko on*

$$x + \mathrm{Null}(A) = \{x + y \in \mathbb{R}^{n\times 1} \colon y \in \mathrm{Null}(A)\}.$$

Huomautus 3.3.5. *Huomaa, että edellisessä määritelmässä $x + \mathrm{Null}(A)$ on joukon $\{x + y \in \mathbb{R}^{n\times 1} \colon y \in \mathrm{Null}(A)\}$ nimi, eli summa merkintää ei ajatella laskutoimituksena. Lauseeseen sisältyy lisäksi havainto, että jos $z \in \mathbb{R}^{n\times 1}$ on jokin toinen yhtälöryhmän $\left[A \mid b\right]$ ratkaisu, niin tällöin $z + \mathrm{Null}(A) = x + \mathrm{Null}(A)$.*[1]

Tarkastellaan ennen lauseen 3.3.4 todistusta konkreettista esimerkkiä.

Esimerkki 3.3.6. *Olkoot*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \text{ ja } b = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Yhtälöryhmän $\left[A \mid b\right]$ supistettu porrasmuoto on

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & 0 \end{array}\right].$$

Näin ollen sen ratkaisut (x_1, x_2) ovat muotoa

$$x_1 + 2x_2 = 3$$

[1] Algebrassa joukkoa $x + \mathrm{Null}(A)$ kutsutaan aliavaruuden $\mathrm{Null}(A)$ sivuluokaksi ja liittyy tekijäavaruuden käsitteeseen.

eli

$$\begin{cases} x_1 &= -2t+3 \\ x_2 &= t, \end{cases}$$

missä $t \in \mathbb{R}$. *Alkuperäisen yhtälön* $Ax = b$ *ratkaisut ovat siis muotoa*

$$x = \begin{bmatrix} -2t+3 \\ t \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Huomaa, että vektori $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ *on siis yksi yhtälöryhmän* $\begin{bmatrix} A \mid b \end{bmatrix}$ *ratkaisuista.*

Tarkastellaan nyt homogeenista yhtälöryhmää $\begin{bmatrix} A \mid 0 \end{bmatrix}$. *Tämän yhtälöryhmän supistettu porrasmuoto on*

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Yhtälöryhmän $\begin{bmatrix} A \mid 0 \end{bmatrix}$ *ratkaisut* (x_1, x_2) *ovat muotoa siis*

$$\begin{cases} x_1 &= -2t \\ x_2 &= t \end{cases},$$

eli

$$x = t \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

missä $t \in \mathbb{R}$. *Näin ollen*

$$\mathrm{Null}(A) = \left\{ t \begin{bmatrix} -2 \\ 1 \end{bmatrix} \in \mathbb{R}^{2\times 1} \colon t \in \mathbb{R} \right\}.$$

Yhdistämällä tämä jo saatuun tulokseen havaitaan, että yhtälöryhmän $\begin{bmatrix} A \mid b \end{bmatrix}$ *ratkaisujoukko on*

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} + \left\{ t \begin{bmatrix} -2 \\ 1 \end{bmatrix} \in \mathbb{R}^{2\times 1} \colon t \in \mathbb{R} \right\};$$

katso kuva 3.3

Lauseen 3.3.4 todistus. Olkoon $y \in \mathrm{Null}(A)$. Tällöin

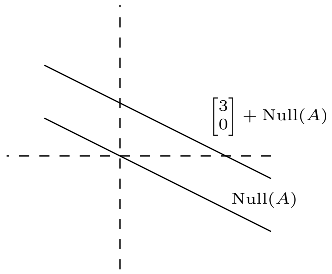
$$A(x+y) = Ax + Ay = Ax + 0 = Ax = b.$$

Näin ollen $x + y$ on yhtälöryhmän $\begin{bmatrix} A \mid b \end{bmatrix}$ ratkaisu.

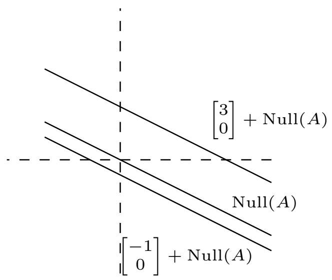
Oletetaan nyt, että $z \in \mathbb{R}^{n\times 1}$ on yhtälöryhmän $\begin{bmatrix} A \mid b \end{bmatrix}$ ratkaisu. Olkoon $y = z - x$. Tällöin $x + y = z$ ja

$$Ay = A(z-x) = Az - Ax = b - b = 0,$$

eli $y \in \mathrm{Null}(A)$. □



Kuva 3.3: Esimerkin 3.3.6 matriisin A nolla-avaruus Null(A) ja suora $\begin{bmatrix}3\\0\end{bmatrix} + \mathrm{Null}(A)$ sarakeavaruudessa $\mathbb{R}^{2\times 1}$.



Kuva 3.4: Esimerkin 3.3.7 matriisin A nolla-avaruus Null(A) sekä suorat $\begin{bmatrix}3\\0\end{bmatrix} + \mathrm{Null}(A)$ ja $\begin{bmatrix}-1\\0\end{bmatrix} + \mathrm{Null}(A)$ sarakeavaruudessa $\mathbb{R}^{2\times 1}$.

Lauseen 3.3.4 varsinainen merkitys on siinä, että se paljastaa yhtälöryhmien $\begin{bmatrix}A \mid b\end{bmatrix}$ ja $\begin{bmatrix}A \mid b'\end{bmatrix}$ ratkaisujoukot olevan oleellisesti samanlaisia: jos yhtälöryhmillä on ratkaisuja, niin tällöin molemmat ratkaisujoukot saadaan siirtämällä matriisin A nolla-avaruutta Null(A) jollain yhtälöryhmän ratkaisulla. Tarkastellaan tätä vielä esimerkin avulla.

Esimerkki 3.3.7. *Olkoot*

$$A = \begin{bmatrix}1 & 2\\2 & 4\end{bmatrix} \text{ ja } b = \begin{bmatrix}3\\6\end{bmatrix}.$$

kuten esimerkissä 3.3.6 ja olkoon

$$b' = \begin{bmatrix}-1\\-2\end{bmatrix}.$$

Esimerkissä 3.3.6 osoitettiin, että yhtälöryhmän $\begin{bmatrix}A \mid b\end{bmatrix}$ *ratkaisujoukko on*

$$\begin{bmatrix}3\\0\end{bmatrix} + \mathrm{Null}(A),$$

missä

$$\mathrm{Null}(A) = \left\{ t \begin{bmatrix} -2 \\ 1 \end{bmatrix} \in \mathbb{R}^{2\times 1} \colon t \in \mathbb{R} \right\}.$$

Toisaalta vastaavasti voidaan päätellä, että yhtälöryhmän $\left[A \mid b'\right]$ *ratkaisujoukko on*

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} + \mathrm{Null}(A).$$

Molempien yhtälöryhmien raktkaisujoukot on havainnollistettu kuvassa 3.4.

3.4 Lineaarikombinaatiot ja virittäminen

Esimerkit 3.1.3 ja 3.1.4 antoivat vihjeen, kuinka aliavaruuksia voi helposti muodostaa. Sanotaankin, että esimerkissä 3.1.3 aliavaruus L on yhden vektorin virittämä ja että 3.1.4 aliavaruus P on kahden vektorin virittämä. Formalisoidaan nyt nämä ajatukset aloittamalla lineaarikombinaation määritelmästä.

Määritelmä 3.4.1. *Vektoria* $v \in \mathbb{R}^{n\times 1}$ *sanotaan* vektoreiden $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ lineaarikombinaatioksi*, jos on olemassa sellaiset luvut* $a_1, \dots, a_k \in \mathbb{R}$*, että*

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k.$$

Vektoreiden $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ kaikkien lineaarikombinaatioiden joukkoa merkitään

$$\mathrm{Sp}(v_1, \dots, v_k) = \{a_1 v_1 + \cdots + a_k v_k \in \mathbb{R}^{n\times 1} \colon a_1, \dots, a_k \in \mathbb{R}\}.$$

Esimerkki 3.4.2. *Esimerkissä 3.1.3 suora* L *on itseasiassa* $\mathrm{Sp}(u)$ *eli*

$$L = \mathrm{Sp}\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} \right).$$

Esimerkissä 3.1.4 taso P *on itseasiassa* $\mathrm{Sp}(u, \tilde{u})$ *eli*

$$L = \mathrm{Sp}\left(\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right).$$

Ei ole sattumaa, että edeltävissä esimerkeissä vektoreiden lineaarikombinaatioiden joukko on aina aliavaruus.

Lause 3.4.3. *Olkoot* $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ *vektoreita. Tällöin* $\mathrm{Sp}(v_1, \dots, v_k)$ *on avaruuden* $\mathbb{R}^{n\times 1}$ *aliavaruus.*

Todistus. Koska $0 = 0v_1 + \cdots + 0v_k \in \mathrm{Sp}(v_1, \dots, v_k)$, niin joukko $\mathrm{Sp}(v_1, \dots, v_k)$ on epätyhjä. Näin ollen voidaan käyttää aliavaruuskriteeriota (lemma 3.2.8).

Olkoot $v, w \in \mathrm{Sp}(v_1, \dots, v_k)$ ja $a \in \mathbb{R}$. Lineaarikombinaatioiden joukon määritelmän perusteella, on olemassa sellaiset luvut $a_1, \dots, a_k \in \mathbb{R}$ ja $b_1, \dots, b_k \in \mathbb{R}$, että $v = a_1 v_1 +$

$\cdots + a_k v_k$ ja $w = b_1 v_1 + \cdots + b_k v_k$. Järjestelemällä vektoreihin $v_1, \ldots, v_k$ liittyviä termejä saadaan

$$\begin{aligned} av + w &= a\,(a_1 v_1 + \cdots + a_k v_k) + (b_1 v_1 + \cdots + b_k v_k) \\ &= (aa_1)v_1 + \cdots + (aa_k)v_k + b_1 v_1 + \cdots + b_k v_k \\ &= (aa_1 + b_1)v_1 + \cdots + (aa_k + b_k)v_k. \end{aligned}$$

Merkitään nyt $c_j = aa_j + b_j \in \mathbb{R}$ jokaisella $j = 1, \ldots, k$. Tällöin

$$av + w = c_1 v_1 + \cdots + c_k v_k \in \mathrm{Sp}(v_1, \ldots, v_k).$$

Osajoukko $\mathrm{Sp}(v_1, \ldots, v_k)$ on siis aliavaruus lemman 3.2.8 perusteella. □

Määritelmä 3.4.4. *Aliavaruutta $V \subset \mathbb{R}^{n\times 1}$ sanotaan* vektoreiden $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ jonon $(v_1, \ldots, v_k)$ virittämäksi aliavaruudeksi, *jos*

$$V = \mathrm{Sp}(v_1, \ldots, v_k).$$

Jonoa $(v_1, \ldots, v_k)$ sanotaan aliavaruuden V virittäväksi jonoksi.

Huomautus 3.4.5. *Huomaa, että mikäli jono on ns. tyhjä jono (), eli siinä ei ole vektoreita, niin tällöin määritellään $\mathrm{Sp}() = \{0\}$. Tähän (hieman erikoiseen) määritelmään palataan hieman myöhemmin.*

Aliavaruuden vektorit voidaan yleensä kirjoittaa usealla eri tavalla virittäjien lineaarikombinaationa kuten seuraava esimerkki osoittaa.

Esimerkki 3.4.6. *Olkoot*

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \; v_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \;\textit{ja}\; v_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

avaruuden $\mathbb{R}^{3\times 1}$ vektoreita. Huomaa, että $v_3 = v_1 + v_2$. Näin ollen vektori

$$v = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}$$

voidaan ilmasta vektoreiden v_1, v_2, v_3 lineaarikombinaationa sekä muodossa

$$v = 2v_3$$

että muodossa

$$v = 2v_1 + 2v_2.$$

Itseasiassa v voidaan ilmaista vektoreiden v_1, v_2, v_3 lineaarikombinaationa äärettömän monella tavalla, sillä jokaisella $t \in \mathbb{R}$ pätee

$$v = tv_1 + tv_2 + (2 - t)v_3.$$

3.4.1 Viritetyn aliavaruuden ominaisuudet

Palataan nyt virittämisen määritelmään ja kommentoidaan tyhjän jonon tapausta. Syy tähän määritelmän laajennukseen on seuraava. Vektoreiden $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ lineaarikombinaatioiden joukko $\mathrm{Sp}(v_1, \dots, v_k)$ voidaan määritellä myös pienimpänä sellaisena avaruuden $\mathbb{R}^{n\times 1}$ aliavaruutena, joka sisältää kaikki vektoreiden $v_1, \dots, v_n$ lineaarikombinaatiot. Koska $\{0\}$ on pienin avaruuden $\mathbb{R}^{n\times 1}$ aliavaruus ja vektoreita $v_1, \dots, v_k$ ei ole annettu, niin $\{0\}$ toteuttaa nämä vaatimukset.

Todistetaan vielä täydellisyyden vuoksi, että määritelmän 3.4.4 mukaan määritelty aliavaruus $\mathrm{Sp}(v_1, \dots, v_k)$ on todellakin pienin vektorit $v_1, \dots, v_k$ sisältävä aliavaruus.

Lause 3.4.7. *Olkoot $v_1, \dots, v_k \in \mathbb{R}^n$ ja olkoon $V \subset \mathbb{R}^n$ sellainen aliavaruus, että $v_1, \dots, v_k \in V$. Tällöin $\mathrm{Sp}(v_1, \dots, v_k) \subset V$.*

Huomaa, että lauseen muotoilussa ei puhuta pienimmästä aliavaruudesta, sen sijaan todetaan, että aliavaruus V, joka sisältää vektorit $v_1, \dots, v_k$ sisältää koko aliavaruuden $\mathrm{Sp}(v_1, \dots, v_k)$, eli että aliavaruus $\mathrm{Sp}(v_1, \dots, v_k)$ on pienempi kuin V. Näin ollen inkluusion määräämän järjestyksen suhteen $\mathrm{Sp}(v_1, \dots, v_k)$ on pienempi (tai yhtäsuuri) kuin mikään muu vektorit $v_1, \dots, v_k$ sisältävä aliavaruus.

Lauseen 3.4.7 todistus. Osoitetaan, että jokaisella $v \in \mathrm{Sp}(v_1, \dots, v_k)$ pätee $v \in V$.

Olkoon $v \in \mathrm{Sp}(v_1, \dots, v_k)$. Tällöin on olemassa sellaiset luvut $a_1, \dots, a_k \in \mathbb{R}$, että $v = a_1v_1 + \cdots + a_kv_k$. Koska $v_j \in V$ jokaisella $j = 1, \dots, k$ ja V on aliavaruus, niin $a_jv_j \in V$ jokaisella $j = 1, \dots, k$. Näin ollen, koska V on aliavaruus, niin $a_1v_1 + \cdots + a_kv_k \in V$. Näin ollen $v \in V$. □

Lauseesta 3.4.7 seuraa erityisesti, että *vektoreiden $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ lineaarikombinaatioiden $w_1 \dots, w_\ell \in \mathbb{R}^{n\times 1}$ lineaarikombinaatiot ovat vektoreiden $v_1, \dots, v_k$ lineaarikombinaatioita.* Kirjataan tämä korollaariksi.

Korollaari 3.4.8. *Olkoot $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ ja olkoot $w_1, \dots, w_\ell \in \mathrm{Sp}(v_1, \dots, v_k)$. Tällöin*

$$\mathrm{Sp}(w_1, \dots, w_\ell) \subset \mathrm{Sp}(v_1, \dots, v_k).$$

3.5 Esimerkki: Matriisin sarakeavaruus aliavaruutena

Kun tarkastellaan vektoreiden $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ virittämää aliavaruutta $\mathrm{Sp}(v_1, \dots, v_n)$, tarkastellaan samalla näiden vektoreiden jonoa. Tämän jonon $(v_1, \dots, v_n)$ tarkasteleminen puolestaan oleellisesti tarkoittaa, että tarkastellaan matriisia $\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. Aliavaruutta $\mathrm{Sp}(v_1, \dots, v_n)$ kutsutaankin matriisin sarakeavaruukseksi.

Määritelmä 3.5.1. *Matriisin $A = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \in \mathbb{R}^{n\times k}$* sarakeavaruus *on aliavaruus*

$$\mathrm{Col}(A) = \mathrm{Sp}(v_1, \dots, v_k) \subset \mathbb{R}^{n\times 1}.$$

Tämän määritelmän etuna on se, että se kertoo suoraan aliavaruuden $\mathrm{Col}(A)$ virittäjät ja sen, että annettu joukko on aliavaruus. Määritelmä voitaisiin toki antaa toisinkin, kuten seuraava lemma osoittaa.

Lemma 3.5.2. *Olkoon $A \in \mathbb{R}^{n\times k}$. Tällöin*

$$\mathrm{Col}(A) = \{Ax \in \mathbb{R}^{n\times 1} \colon x \in \mathbb{R}^{k\times 1}\}.$$

Todistus. Olkoon $v \in \mathrm{Col}(A)$. Sarakeavaruuden määritelmän mukaan v on sarakevektoreiden $v_1, \dots, v_k$ lineaarikombinaatio, eli on olemassa sellaiset luvut $a_1, \dots, a_k \in \mathbb{R}$, että

$$v = a_1 v_1 + \cdots + a_k v_k.$$

Merkitään nyt

$$x = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} \in \mathbb{R}^{k\times 1}.$$

Tällöin

$$v = a_1 v_1 + \cdots + a_k v_k = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} = Ax.$$

Näin ollen $v \in \{Ax \in \mathbb{R}^{n\times 1} \colon x \in \mathbb{R}^{k\times 1}\}$, eli

$$\mathrm{Col}(A) \subset \{Ax \in \mathbb{R}^{n\times 1} \colon x \in \mathbb{R}^{k\times 1}\}.$$

Toisaalta, selvästi kaikilla

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{R}^{k\times 1}$$

pätee

$$Ax = x_1 v_1 + \cdots + x_k v_k \in \mathrm{Col}(A),$$

joten

$$\{Ax \in \mathbb{R}^{n\times 1} \colon x \in \mathbb{R}^{k\times 1}\} \subset \mathrm{Col}(A).$$

Näin ollen

$$\mathrm{Col}(A) = \{Ax \in \mathbb{R}^{n\times 1} \colon x \in \mathbb{R}^{k\times 1}\},$$

joten väite on todistettu. □

Huomautus 3.5.3. *Edellinen lemma osoittaa, että* $\mathrm{Col}(A)$ *on itseasiassa kuvauksen $f_A \colon \mathbb{R}^{k\times 1} \to \mathbb{R}^{n\times 1}$, $x \mapsto Ax$, kuvajoukko.*

Kirjataan vielä ylös edellisen seuraus, että matriisin A sarakeavaruus $\mathrm{Col}(A)$ sisältää täsmälleen sellaiset vektorit b, että yhtälöryhmällä $\begin{bmatrix} A \mid b \end{bmatrix}$ on ratkaisu.

Korollaari 3.5.4. *Olkoon $A = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \in \mathbb{R}^{n\times k}$. Tällöin $b \in \mathrm{Col}(A)$, jos ja vain jos yhtälöryhmällä $\begin{bmatrix} A \mid b \end{bmatrix}$ on ratkaisu.*

Todistus. Olkoon $b \in \mathrm{Col}(A)$. Tällöin lemman 3.5.2 perusteella on olemassa sellainen $x \in \mathbb{R}^{k\times 1}$, että $b = Ax$. Näin ollen yhtälöryhmällä $\left[A \mid b\right]$ on ratkaisu.

Olkoon nyt $b \in \mathbb{R}^{n\times 1}$ sellainen vektori, että yhtälöryhmällä $\left[A \mid b\right]$ on ratkaisu, eli on olemassa sellainen

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{R}^{k\times 1},$$

jolle pätee $Ax = b$. Tällöin

$$b = Ax = x_1v_1 + \cdots + x_kv_k \in \mathrm{Sp}(v_1, \ldots, v_k) = \mathrm{Col}(A).$$

Näin ollen $b \in \mathrm{Col}(A)$. $\square$

3.5.1 Kommentti

Tässä kohtaa herää luonnollinen kysymys:

Ongelma. *Ovatko avaruuden $\mathbb{R}^{n\times 1}$ kaikki aliavaruudet matriisien sarakeavaruuksia, eli jos $V \subset \mathbb{R}^{n\times 1}$ on aliavaruus, niin onko olemassa sellaiset vektorit $v_1, \ldots, v_k$, että $V = \mathrm{Sp}(v_1, \ldots, v_k)$?*

Kysymys on erittäin syvällinen sillä se kysyy oikeasti seuraavaa: Jos $V \subset \mathbb{R}^{n\times 1}$ on aliavaruus, niin voidaanko valita äärellinen määrä vektoreita $v_1, \ldots, v_k \in V$, että kaikki muut aliavaruuden V vektorit voidaan kirjoittaa näiden vektoreiden lineaarikombinaationa.

Vastaus tähän kysymykseen on myönteinen. Tämä syvällinen fakta todistetaan luvun lopussa ja siihen palataan kurssilla Lineaarialgebra ja matriisilaskenta II yleisemmässä äärellisulotteisten vektoriavaruuksien kontekstissa. Itseasiassa tulemme lopulta todistamaan paljon enemmän: avaruuden $\mathbb{R}^{n\times 1}$ aliavaruuden V virittämiseen vektoreita tarvitaan aina korkeintaan n kappaletta ja että nämä vektorit voidaan valita siten, että jokainen aliavaruuden V vektori voidaan ilmaista yksikäsitteisenä lineaarikombinaationa annetuista vektoreista.

Tämä projekti on jaettu kahteen osaan ja tässä luvussa esitellään lineaarisen riippumattomuuden ja kannan käsitteet sarakeavaruuden $\mathbb{R}^{n\times 1}$ tapauksessa, jotka antavat jälkimmäiseen osan vastauksesta.

3.6 Lineaarinen riippumattomuus

Määritelmä 3.6.1. *Vektorit $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ ovat* lineaarisesti riippumattomia, *jos yhtälön*

$$a_1v_1 + \cdots + a_kv_k = 0,$$

missä $a_1, \ldots, a_k \in \mathbb{R}$, ainoa ratkaisu on $a_1 = a_2 = \cdots = a_n = 0$.

Lineaarinen riippumattomuus on lineaarialgebrassa oleellinen käsite, koska lineaarisesti riippumattomuudesta seuraa, että tarkasteltavien vektoreiden lineaarikombinaatiot ovat yksikäsitteisiä. Todistetaan tämä tärkeä huomio nyt tarkasti.

Lause 3.6.2. *Olkoot $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ lineaarisesti riippumattomia vektoreita ja $v \in \mathrm{Sp}(v_1, \dots, v_k)$. Tällöin on olemassa sellaiset yksikäsitteiset luvut $x_1, \dots, x_k \in \mathbb{R}$, että*

$$v = x_1 v_1 + \ldots + x_k v_k.$$

Todistus. Olkoon $v \in \mathrm{Sp}(v_1, \dots, v_k)$. Aliavaruuden $\mathrm{Sp}(v_1, \dots, v_k)$ määritelmän nojalla on olemassa sellaiset luvut $x_1, \dots, x_k \in \mathbb{R}$, että

$$v = x_1 v_1 + \cdots + x_k v_k.$$

Oletetaan nyt, että $y_1, \dots, y_k \in \mathbb{R}$ ovat sellaisia lukuja, että

$$v = y_1 v_1 + \cdots + y_k v_k.$$

Tällöin uudelleen järjestelemällä summat saadaan

$$(y_1 - x_1)v_1 + \cdots + (y_k - x_k)v_k = (y_1 v_1 + \cdots + y_k v_k) - (x_1 v_1 + \cdots + x_k v_k) = v - v = 0.$$

Koska vektorit $v_1, \dots, v_k$ ovat lineaarisesti riippumattomia, niin $y_1 - x_1 = \cdots = y_k - x_k = 0$, eli $y_i = x_i$ kaikilla $i = 1, \dots, k$. □

3.6.1 Lineaarisen riippumattomuuden selvittäminen

Luonnollisin tapa selvittää vektoreiden $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ lineaarinen riippumattomuus on ratkaista ongelma matriisiyhtälöllä $Ax = 0$, missä $A = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \in \mathbb{R}^{n\times k}$. Vektoreiden $v_1, \dots, v_k$ linaarinen riippumattomuus on itseasiassa yhtäpitävää sen kanssa, että matriisin A nolla-avaruus on triviaali, eli pätee $\mathrm{Null}(A) = \{0\}$. Tämä havainto kytkee yhteen lineaarisen riippumattomuuden, matriisiyhtälön ja nolla-avaruuden, joten tehdään se nyt tarkasti.

Lemma 3.6.3. *Olkoon $A = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \in \mathbb{R}^{n\times k}$. Tällöin vektorit $v_1, \dots, v_k$ ovat lineaarisesti riippumattomia, jos ja vain jos yhtälön $Ax = 0$ ainoa ratkaisu on $x = 0$, eli* $\mathrm{Null}(A) = \{0\}$.

Todistus. Oletetaan ensin, että vektorit $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$ ovat linearisesti riippumattomia. Osoitetaan, että yhtälön $Ax = 0$ ainoa ratkaisu on $x = 0$. Olkoon

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{R}^{k\times 1}$$

sellainen vektori, että $Ax = 0$. Tällöin

$$0 = Ax = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} = x_1 v_1 + \cdots x_k v_k.$$

Koska vektorit $v_1, \ldots, v_k$ ovat lineaarisesti riippumattomia, niin $x_1 = \cdots = x_k = 0$. Näin ollen $x = 0$.

Oletetaan nyt, että $\mathrm{Null}(A) = \{0\}$, eli että $x = 0$ on yhtälön $Ax = 0$ ainoa ratkaisu. Osoitetaan, että vektorit $v_1, \ldots, v_k$ ovat lineaarisesti riippumattomia. Olkoot $a_1, \ldots, a_k \in \mathbb{R}$ sellaisia lukuja, että

$$a_1 v_1 + \cdots + a_k v_k = 0. \tag{3.1}$$

Muodostetaan luvuista $a_1, \ldots, a_k$ vektori

$$x = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} \in \mathbb{R}^{k \times 1}.$$

Yhtälön (3.1) perusteella pätee

$$Ax = 0.$$

Oletuksen nojalla $x = 0$, joten $a_1 = \cdots = a_k = 0$. Vektorit $v_1, \ldots, v_k$ ovat siis lineaarisesti riippumattomia. □

3.6.2 Sidotut ja vapaat jonot

Vektoreita $v_1, \ldots, v_k \in \mathbb{R}^{n \times 1}$, jotka eivät ole lineaarisesti riippumattomia toisistaan, sanotaan *lineaarisesti riippuviksi*. Sen sijaan, että puhutaan vektoreiden kokoelmasta (eli joukosta), on luonnollista määritellä nämä käsitteet käyttäen vektoreiden jonoa.

Määritelmä 3.6.4. *Vektoreiden $v_1, \ldots, v_k \in \mathbb{R}^{n \times 1}$ jono $(v_1, \ldots, v_k)$ on* vapaa*, jos vektorit $v_1, \ldots, v_k$ ovat lineaarisesti riippumattomia. Jonoa $(v_1, \ldots, v_k)$ sanotaan* sidotuksi*, jos se ei ole vapaa.*

Sidotun jonon idea paljastuu parhaiten seuraavasta lemmasta, joka sanoo, että jos jono on sidottu, niin silloin joku vektoreista voidaan kirjoittaa muiden lineaarikombinaationa.

Lemma 3.6.5. *Olkoot $v_1, \ldots, v_k \in \mathbb{R}^{n \times 1}$ vektoreita. Jono $(v_1, \ldots, v_k)$ on sidottu, jos ja vain jos on olemassa sellainen $j \in \{1, \ldots, k\}$ ja luvut $b_1, \ldots, b_{j-1} \in \mathbb{R}$, että*

$$v_j = b_1 v_1 + \cdots + b_{j-1} v_{j-1},$$

eli $v_j \in \mathrm{Sp}(v_1, \ldots, v_{j-1})$.

Todistus. Oletetaan ensin, että jono $(v_1, \ldots, v_k)$ on sidottu, eli että vektorit $v_1, \ldots, v_k$ eivät ole lineaarisesti riippumattomia. Tällöin on olemassa sellaiset luvut $a_1, \ldots, a_k \in \mathbb{R}$, että

$$a_1 v_1 + \cdots + a_k v_k = 0$$

ja että $a_i \neq 0$ jollain $i \in \{1, \ldots, k\}$. Näin ollen on olemassa suurin sellainen indeksi $j \in \{1, \ldots, k\}$, että $a_j \neq 0$, eli

$$a_1 v_1 + \cdots + a_{j-1} v_{j-1} + a_j v_j = 0.$$

Vähentämällä vektori a_jv_j yhtälön molemmilta puolilta saadaan, että

$$a_1v_1 + \cdots + a_{j-1}v_{j-1} = -a_jv_j.$$

Koska $a_j \neq 0$, niin saadaan, että

$$v_j = -\frac{a_1}{a_j}v_1 - \cdots - \frac{a_{j-1}}{a_j}v_{j-1}.$$

Voidaan siis valita luvut $b_i = -a_i/a_j$ jokaisella $i \in \{1, \dots, j-1\}$. Tämä päättää ensimmäisen suunnan todistuksen.

Oletetaan nyt, että on olemassa sellainen indeksi $j \in \{1, \dots, k\}$ ja sellaiset luvut $b_1, \dots, b_{j-1} \in \mathbb{R}$, että

$$v_j = b_1v_1 + \cdots + b_{j-1}v_{j-1}.$$

Valitaan nyt $a_i = -b_i$ jokaisella $i \in \{1, \dots, j-1\}$, $a_j = 1$ ja $a_i = 0$ jokaisella $a_i \in \{j+1, \dots, k\}$. Tällöin uudelleen ryhmittelemällä termit saadaan

$$\begin{aligned} a_1v_1 + \cdots + a_kv_k &= a_1v_1 + \cdots + a_{j-1}v_{j-1} + a_jv_j + a_{j+1}v_{j+1} + \cdots + a_kv_k \\ &= (-b_1)v_1 + \cdots + (-b_{j-1})v_{j-1} + (b_1v_1 + \cdots + b_{j-1}v_{j-1}) \\ &= (b_1 - b_1)v_1 + \cdots + (b_{j-1} - b_{j-1})v_{j-1} = 0. \end{aligned}$$

Jono $(v_1, \dots, v_k)$ on siis sidottu. Tämä päättää todistuksen. □

Yleinen todistuksissa esiintyvä havainto on, että vapaan jonon osajono on myös vapaa jono. Kirjataan tämä havainto lemmaksi lukua 3.8 varten.

Lemma 3.6.6. *Olkoon* $(v_1, \dots, v_k) \in \mathbb{R}^{n\times 1}$ *vapaa jono. Tällöin jono* $(v_{i_1}, \dots, v_{i_d})$*, missä* $1 \leq i_1 < \cdots < i_d \leq k$*, on vapaa.*

Todistus. Olkoot $a_1, \dots, a_d \in \mathbb{R}$ sellaisia, että

$$a_1v_{i_1} + a_2v_{i_2} + \cdots + a_dv_{i_d} = 0.$$

Olkoot nyt $b_1, \dots, b_k \in \mathbb{R}$ sellaisia lukuja, että $b_{i_\ell} = a_\ell$ jokaisella $\ell \in \{1, \dots, d\}$ ja $b_i = 0$ muutoin. Tällöin

$$b_1v_1 + b_2v_2 + \cdots + b_kv_k = a_1v_{i_1} + a_2v_{i_2} + \cdots + a_dv_{i_d} = 0.$$

Koska $(v_1, \dots, v_k)$ on vapaa jono, niin $b_1 = b_2 = \cdots = b_k = 0$. Näin ollen $a_1 = a_2 = \cdots = a_d = 0$. Jono $(v_{i_1}, \dots, v_{i_d})$ on siis vapaa. □

3.6.3 Vapaat jonot ja matriisit

Kuten luvussa 3.6.1 havaittiin jonon osoittaminen vapaaksi vastaa jonon vektoreista muodostetun matriisin nolla-avaruuden tutkimista. Tehdään nyt toinen vapauteen ja matriiseihin liittyvä havainto, joka voidaan muotoilla myös seuraavasti: jos annettujen vektoreiden lineaarikombinaatiot muodostavat vapaan jonon, niin tällöin myös alkuperäiset vektorit muodostavat vapaan jonon. Kirjatataan tämä hieman heurstinen havainto seuraavasti.

Lause 3.6.7. *Olkoon* $A \in \mathbb{R}^{m\times n}$ *ja olkoot* $v_1, \dots, v_k \in \mathbb{R}^{n\times 1}$*. Jos jono* $(Av_1, \dots, Av_k)$ *on avaruuden* $\mathbb{R}^{m\times 1}$ *vapaa jono, niin tällöin* $(v_1, \dots, v_k)$ *on avaruuden* $\mathbb{R}^{n\times 1}$ *vapaa jono.*

Huomautus 3.6.8. *Tämä tulos on kaikkein luonnollisinta muotoilla käyttäen kuvausta* $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$, $x \mapsto Ax$*. Tällöin vaite voidaan muotoilla seuraavasti: jos vektoreiden* $v_1, \dots, v_k$ *kuvavektoreiden jono* $(f_A(v_1), \dots, f_A(v_k))$ *on vapaa, niin tällöin alkuperäinen jono* $(v_1, \dots, v_k)$ *on myös vapaa.*

Huomaa, että väite ei päde toisinpäin: *vapaalle jonolle* $(v_1, \dots, v_k)$ *ei päde, että jono* $(Av_1, \dots, Av_n)$ *olisi välttämättä vapaa.* Esimerkki tälläisesti tilanteesta saadaan valitsemalla matriisi A nollamatriisiksi. Yhtäpitävyys on kuitenkin totta, jos matriisi A on kääntyvä.

Korollaari 3.6.9. *Olkoon* $A \in \mathbb{R}^{n\times n}$ *kääntyvä matriisi. Tällöin avaruuden* $\mathbb{R}^{n\times 1}$ *jono* $(v_1, \dots, v_k)$ *on vapaa, jos ja vain jos jono* $(Av_1, \dots, Av_k)$ *on vapaa.*

Todistus. Sovelletaan lausetta 3.6.7 matriiseihin A ja A^{-1}. □

Lauseen 3.6.7. Oletetaan, että jono $(Av_1, \dots, Av_k)$ on vapaa ja osoitetaan, että jono $(v_1, \dots, v_k)$ on vapaa.

Olkoot $a_1, \dots, a_k \in \mathbb{R}$ sellaisia lukuja, että

$$a_1v_1 + \dots + a_kv_k = 0.$$

Tällöin

$$a_1Av_1 + \dots + a_kAv_k = A(a_1v_1 + \dots + a_kv_k) = A0 = 0.$$

Koska jono $(Av_1, \dots, Av_k)$ on vapaa, niin $a_1 = \dots = a_k = 0$. Näin ollen jono $(v_1, \dots, v_k)$ on osoitettu vapaaksi. □

3.7 Aliavaruuden kanta

Virittämisen ja vapauden yhdistävä käsite on kanta.

Määritelmä 3.7.1. *Jono* $(v_1, \dots, v_k)$ *on aliavaruuden* $V \subset \mathbb{R}^n$ *on* kanta*, jos*

1. *jono* $(v_1, \dots, v_k)$ *virittää aliavaruuden* V*, eli* $V = \mathrm{Sp}(v_1, \dots, v_k)$*, ja*
2. *jono* $(v_1, \dots, v_k)$ *on vapaa.*

Huomautus 3.7.2. *Koska tyhjä jono* () *on määritelmän mukaan sekä vapaa jono että virittää aliavaruuden* $\{0\}$*, niin aliavaruudella* $\{0\}$ *on kanta* ()*, ns. tyhjä kanta.*

3.7.1 Vektorin kertoimet kannassa

Kannan määritelmä sanoo siis, että jokainen alivaruuden V vektori $v \in V$ voidaan kirjoittaa vektoreiden $v_1, \ldots, v_k$ lineaarikombinaationa (kohta 1)

$$v = a_1 v_1 + \cdots + a_k v_k$$

ja että tämä lineaarikombinaatio on yksikäsitteinen (kohta 2), eli että luvut $a_1, \ldots, a_k \in \mathbb{R}$ ovat yksikäsitteisiä. Lineaarikombinaation yksikäsitteisyys seuraa lauseesta 3.6.2. Toisaalta, jos jokainen vektori $v \in V$ voidaan kirjoittaa yksikäsitteisenä lineaarikombinaationa vektoreista $v_1, \ldots, v_k$, niin tällöin jono $(v_1, \ldots, v_k)$ on aliavaruuden V kanta. Kirjataan tämä havainto lemmaksi.

Lemma 3.7.3. *Olkoot $v_1, \ldots, v_k \in V$. Tällöin jono $(v_1, \ldots, v_k)$ on aliavaruuden V kanta, jos ja vain jos jokaisella $v \in V$ on olemassa yksikäsitteiset sellaiset luvut $a_1, \ldots, a_k \in \mathbb{R}$, että $v = a_1 v_1 + \cdots + a_k v_k$.*

Todistus. Oletetaan, että $(v_1, \ldots, v_k)$ on aliavaruuden V kanta. Olkoon $v \in V$. Koska $V = \mathrm{Sp}(v_1, \ldots, v_k)$, niin on olemassa sellaiset luvut $x_1, \ldots, x_k \in \mathbb{R}$, että $v = x_1 v_1 + \cdots + x_k v_k$. Koska jono $(v_1, \ldots v_k)$ on vapaa, niin lauseen 3.6.2 perusteella luvut $x_1, \ldots, x_k$ ovat yksikäsitteisiä.

Oletetaan nyt, että jokaisella $v \in V$ on olemassa yksikäsitteiset sellaiset luvut $a_1, \ldots, a_k \in \mathbb{R}$, että $v = a_1 v_1 + \cdots + a_k v_k$ ja osoitetaan, että $(v_1, \ldots, v_k)$ on aliavaruuden V kanta. Koska $V = \mathrm{Sp}(v_1, \ldots, v_k)$, riittää osoittaa, että jono $(v_1, \ldots, v_k)$ on vapaa. Olkoot $a_1, \ldots, a_k \in \mathbb{R}$ sellaisia lukuja, että $a_1 v_1 + \cdots + a_k v_k = 0$. Koska $0 v_1 + \cdots + 0 v_k = 0$, niin lukujen $a_1, \ldots, a_k$ yksikäsitteisyyden nojalla $a_1 = \cdots = a_k = 0$. Jono $(v_1, \ldots, v_k)$ on näin ollen vapaa. $\square$

Määritelmä 3.7.4. *Olkoon $V \subset \mathbb{R}^{n \times 1}$ aliavaruus ja $(v_1, \ldots, v_k)$ aliavaruuden V kanta. Lukujen $a_1, \ldots, a_k \in \mathbb{R}$ jonoa $(a_1, \ldots, a_k)$ kutsutaan vektorin $v \in V$* kertoimiksi *tai* koordinaateiksi kannassa $(v_1, \ldots, v_k)$*, jos*

$$v = a_1 v_1 + \cdots + a_k v_k.$$

Esimerkki 3.7.5. *Avaruuden $\mathbb{R}^{n \times 1}$ standardikanta $(e_1, \ldots, e_n)$ muodostuu vektoreista luvussa 2 esitellyistä vektoreista*

$$e_i = \begin{bmatrix} e_{1i} \\ \vdots \\ e_{ni} \end{bmatrix} \in \mathbb{R}^{n \times 1},$$

missä

$$e_{ji} = \delta_{ji} = \begin{cases} 1, & j = i \\ 0, & j \neq i. \end{cases}$$

Selvästi jokaisella

$$v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

pätee

$$v = x_1 e_1 + \cdots + x_n e_n,$$

missä luvut $x_1, \ldots, x_n$ *ovat selvästi yksikäsitteisiä. Standardikanta* $(e_1, \ldots, e_n)$ *on siis kanta.*

On tärkeää huomata, että vektorin koordinaatit riippuvat annetusta kannasta. Tarkastellaan tätä esimerkin avulla.

Esimerkki 3.7.6. *Vektorin*

$$v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^{2\times 1}$$

koordinaatit avaruuden $\mathbb{R}^{2\times 1}$ *standarikannassa* (e_1, e_2) *ovat (tietenkin)* x_1 *ja* x_2*, eli koordinaatit muodostavat jonon* (x_1, x_2)*.*

Muokataan nyt standardikantaa (e_1, e_2) *ja tarkastellaan vektoreita* $v_1 = e_1 + e_2$ *ja* $v_2 = e_1 - e_2$*. Osoitetaan, että* (v_1, v_2) *on avaruuden* $\mathbb{R}^{2\times 1}$ *kanta. Lemman 3.7.3 nojalla riittää osoittaa, että jokainen vektori*

$$v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 e_1 + x_2 e_2$$

voidaan kirjoittaa yksikäsitteisesti vektoreiden v_1 *ja* v_2 *lineaarikombinaationa. Tällöin tulemme samalla selvittäneeksi vektorin* $v = x_1 e_1 + x_2 e_2$ *kertoimet kannassa* (v_1, v_2)*.*

Ratkaistaan tätä varten luvut y_1 *ja* y_2 *yhtälöstä*

$$x_1 e_1 + x_2 e_2 = y_1 v_1 + y_2 v_2. \tag{3.2}$$

Koska

$$\begin{aligned} y_1 v_1 + y_2 v_2 &= y_1(e_1 + e_2) + y_2(e_1 - e_2) \\ &= y_1 e_1 + y_1 e_2 + y_2 e_1 - y_2 e_2 \\ &= (y_1 + y_2)e_1 + (y_1 - y_2)e_2, \end{aligned}$$

niin yhtälö (3.2) *on yhtäpitävä yhtälön*

$$x_1 e_1 + x_2 e_2 = (y_1 + y_2)e_1 + (y_1 - y_2)e_2$$

kanssa. Koska (e_1, e_2) *on avaruuden* $\mathbb{R}^2$ *kanta, niin kertoimien yksikäsitteisyyden nojalla saadaan, että yhtälö* (3.2) *on yhtäpitävä yhtälöryhmän*

$$\begin{cases} y_1 + y_2 &= x_1 \\ y_1 - y_2 &= x_2 \end{cases}$$

kanssa. Ratkaisemalla yhtälöryhmä saadaan ratkaisu

$$\begin{cases} y_1 &= \frac{1}{2}(x_1 + x_2) \\ y_2 &= \frac{1}{2}(x_1 - x_2), \end{cases}$$

eli

$$v = \frac{1}{2}(x_1 + x_2)v_1 + \frac{1}{2}(x_1 - x_2)v_2.$$

Koska lineaarikombinaatio on selvästi yksikäsitteinen, on (v_1, v_2) *avaruuden* $\mathbb{R}^2$ *kanta. Lisäksi havaittiin, että vektorin* $v = x_1e_1 + x_2e_2$ *koodinaatiti tässä kannassa ovat* $((x_1 + x_2)/2, (x_1 - x_2)/2)$*.*

Erityisesti siis vektoreiden

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

koordinaatit kannassa (v_1, v_2) *ovat* $(1/2, 1/2)$ *ja* $(1/2, -1/2)$*.*

Huomautus 3.7.7. *Edellisen esimerkin metodi perustuu vektorin* v *kirjoittamiseen kahdella eri tavalla kannassa* (e_1, e_2)*, joka antaa yhtälön vektorin koordinaateille. Käytännössä perustelu on kuitenkin usein mielekkäämpää selvittää kertoimet kirjoittamalla yhtälö* (3.2) *muodossa*

$$\begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

eli tässä tapauksessa yhtälöryhmäksi

$$\left[\begin{array}{cc|c} 1 & 1 & x_1 \\ 1 & -1 & x_2 \end{array}\right].$$

Tämän jälkeen kertoimet y_1 *ja* y_2 *voidaan ratkaista tästä yhtälöryhmästä viemällä se supistettuun porrasmuotoon.*

3.8 Esimerkki: Matriisin sarakeavaruuden kanta

Yleisen matriisin $A \in \mathbb{R}^{m \times n}$ sarakeavaruuden $\mathrm{Col}(A)$ kanta selvitetään kahdessa vaiheessa: ensin etsitään matriisin supistettu porrasmuoto ja tämän jälkeen luetaan haluttu kanta tästä porrasmuodosta. Aloitetaan tapauksessa, jossa matriisi on jo supistetussa porrasmuodossa.

3.8.1 Supistetussa porrasmuodossa oleva matriisi

Esitellään ensin tarvittava terminologia.

Määritelmä 3.8.1. *Matriisi* $B = [b_{ji}] \in \mathbb{R}^{m \times n}$ *on* supistetussa porrasmuodossa*, jos vastaava homogeeninen yhtälöryhmä* $\left[\begin{array}{c|c} B & 0 \end{array}\right]$ *on supistetussa porrasmuodossa.*

Esimerkki 3.8.2. *Matriisia*

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

vastaava homogeeninen yhtälöryhmä $\left[B \mid 0\right]$ *on*

$$\begin{cases} x_1 & -4x_2 & & 2x_4 & = & 0 \\ & & x_3 & -x_4 & = & 0 \\ & & & 0 & = & 0 \end{cases}$$

joka on supistetussa porrasmuodossa. Näin ollen B on supistetussa porrasmuodossa.

Tarkastellaan nyt esimerkin 3.8.2 tapauksessa kuinka sarakeavaruudelle löydetään kanta.

Esimerkki 3.8.3. *Olkoon*

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

ja olkoot v_1, v_2, v_3, v_4 matriisin B sarakkeet, eli

$$B = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}.$$

Huomataan jo tässä vaiheessa, että sarakeet v_1 ja v_3 ovat standarikannan alkiota eli

$$v_1 = e_1 \text{ ja } v_3 = e_2.$$

Samoin huomataan, että

$$v_2 = -4v_1 \text{ ja } v_4 = 2v_1 + (-1)v_3.$$

Näin ollen kaikki matriisin B sarakkeet voidaan kirjoittaa sarakkeiden v_1 ja v_3 lineaarikombinaationa, eli $v_1, v_2, v_3, v_4 \in \mathrm{Sp}(v_1, v_3)$. Koska lineaarikombinaatioiden lineaarikombinaatiot ovat alkuperäisten vektoreiden lineaarikombainaatioita (korollaari 3.4.8), niin

$$\mathrm{Col}(B) = \mathrm{Sp}(v_1, v_2, v_3, v_4) \subset \mathrm{Sp}(v_1, v_3).$$

Koska $\mathrm{Sp}(v_1, v_3) \subset \mathrm{Col}(B)$*, niin saadaan, että*

$$\mathrm{Col}(B) = \mathrm{Sp}(v_1, v_3).$$

Toisaalta, koska v_1 ja v_3 ovat standardikannan vektoreita, niin (v_1, v_3) on vapaan jonon (e_1, e_2, e_3, e_4) osajono (e_1, e_3). Näin ollen (v_1, v_3) on vapaa jono lemman 3.6.6 perusteella. Koska (v_1, v_3) on aliavaruuden $\mathrm{Col}(B)$ *vapaa virittäjäjono, niin se on aliavaruuden* $\mathrm{Col}(B)$ *kanta.*

Analysoidaan nyt esimerkin 3.8.3 päättelyä. Edellisessä esimerkissä matriisiin B sarakkeet v_1 ja v_3 vastaavat täsmälleen yhtälöryhmän $\left[B \mid 0\right]$ sidottujia muuttujia. Kaikille supistetussa porrasmuodossa oleville matriisille pätee, että sidottuja muuttujia vastaavat sarakkeet ovat myös aina standardikannan alkiota ja näin ollen, että sidottuja muuttujia vastaavien sarakkeiden jono on aina vapaa.

Esimerkissä tehtiin näiden lisäksi havainto, että vapaita muuttujia vastaavat sarekkeet voidaan kirjoittaa sidottuja muuttujia vastaavien sarekkeiden lineaarikombinaationa. Tästä seuraa, että kaikki sarakeavaruuden vektorit voidaan kirjoittaa sidottuja muuttujia vastaavien sarakkeiden (eli tässä sarekkeiden v_1 ja v_3) lineaarikombinaatioina. Tästä seuraa, että yhtälöryhmän $\left[B \mid 0\right]$ sidottuja muuttujia vastaavat sarekkeet muodostavat aliavaruuden $\mathrm{Col}(B)$ kannan.

Kirjataan nämä havainnot nyt yleiseksi lauseeksi, vaikka edellinen esimerkin analyysi onkin jo oleellisesti väitteen todistus.

Lause 3.8.4. *Olkoon $B = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ supistetussa porrasmuodossa oleva matriisi. Olkoot lisäksi $1 \leq i_1 < \cdots < i_d \leq n$ yhtälöryhmän $\left[B \mid 0\right]$ sidottuja muuttujia vastaavat indeksit. Tällöin $(v_{i_1}, \ldots v_{i_d})$ on sarakeavaruuden $\mathrm{Col}(B)$ kanta.*

Todistus. Olkoon

$$v_i = \begin{bmatrix} v_{1i} \\ \vdots \\ v_{mi} \end{bmatrix}$$

jokaisella $i \in \{1, \ldots, n\}$. Tällöin

$$B = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{m1} & \cdots & v_{mn} \end{bmatrix}.$$

Osoitetaan ensin, että $(v_{i_1}, \ldots, v_{i_d})$ on vapaa jono. Tehdään ensin havainto, että sidotun muuttujan määritelmän nojalla, vektorilla v_{i_ℓ} on täsmälleen yksi nollasta poikkeava kerroin $v_{i_\ell \ell}$, joka on 1. Näin ollen jokaisella $\ell \in \{1, \ldots, d\}$ pätee $v_{i_\ell} = e_\ell$. Jono $(v_{i_1}, \ldots, v_{i_d})$ on näin ollen jono $(e_1, \ldots, e_d)$, joka on lemman 3.6.6 perusteella vapaa.

Osoitetaan nyt, että jono $(v_{i_1}, \ldots, v_{i_d})$ virittää aliavaruuden $\mathrm{Col}(B)$. Korollaarin 3.4.8 perusteella riittää osoittaa, että $v_i \in \mathrm{Sp}(v_{i_1}, \ldots, v_{i_d})$ jokaisella $i \in \{1, \ldots, n\}$.

Olkoon $i \in \{1, \ldots, n\}$. Supistetun porrasmuodon määritelmän nojalla, vektorin v_i rivin j kerroin on nollasta poikkeava ainoastaan tapauksessa, jossa v_i vastaa sidottua muuttujaa tai riviä j vastaavan sidotun muuttujan indeksille i_j pätee $i_j < i$. Molemmissa tapauksissa vektori v_i voidaan siis kirjoittaa lineaarikombinaationa vektoreista $e_1, \ldots, e_j$ eli vektoreista $v_{i_1}, \ldots, v_{i_j}$, missä indeksi j on suurin sellainen indeksi, että $i_j \leq i$. Näin ollen $v \in \mathrm{Sp}(v_{i_1}, \ldots, v_{i_j})$ ja erityisesti $v \in \mathrm{Sp}(v_{i_1}, \ldots, v_{i_d})$. Tämä päättää todistuksen. □

3.8.2 Yleinen tapaus

Tarkastellaan nyt yleista matriisia $A \in \mathbb{R}^{m\times n}$ ja sen sarakeavaruutta $\mathrm{Col}(A)$. Kuten luvussa 2 osoitettiin, rivioperaatiot, jotka saattavat matriisin A supistettuun porrasmuotoon B, voidaan kirjoittaa alkeismatriisien avulla. Lauseessa 2.5.1 tätä käytettiin neliömatriisien kääntyvyyden tarkasteluun, mutta yleisemmin samalla päättelyllä havaitaan, että mikäli matriisia $B \in \mathbb{R}^{m\times n}$ vastaava yhtälöryhmä $\left[B \mid \tilde{b}\right]$ saadaan rivioperaatioilla yhtälöryhmästä $\left[A \mid b\right]$, niin tällöin on olemassa sellaiset näitä rivioperaatioita vastaavat alkeismatriisit $E_1, \dots, E_k \in \mathbb{R}^{m\times m}$, että

$$A = E_k \cdots E_1 B = PB,$$

missä matriisi $P = E_k \cdots E_1 \in \mathbb{R}^{m\times m}$ on kääntyvä.

Näin ollen matriisitulon määritelmän nojalla saadaan, että

$$A = PB = P\begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} Pb_1 & \cdots & Pb_n \end{bmatrix},$$

missä $b_1, \dots, b_n$ ovat matriisin B sarakkeet. Sovelletaan tätä havaintoa tilanteessa, jossa B on supistetussa porrasmuodossa. Paljastuu, että sarakeavaruuden $\mathrm{Col}(A)$ kanta voidaan löytää laskematta lainkaan matriisia P. Sarakeavaruuden $\mathrm{Col}(A)$ kannaksi voidaan valita ne matriisin A sarekkeet, jotka vastavaavat supistetun normaalimuodon sidottujen muuttujien indeksejä.

Lause 3.8.5. *Olkoot $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ matriisi, $B \in \mathbb{R}^{m\times n}$ matriisin A supistettu porrasmuoto ja $1 \le i_1 < \cdots < i_d \le n$ yhtälöryhmän $\left[B \mid 0\right]$ sidottujen muuttujien indeksien jono. Tällöin $(v_{i_1}, \dots, v_{i_d})$ on sarakeavaruuden $\mathrm{Col}(A)$ kanta.*

Todistus. Olkoot $b_1, \dots, b_n \in \mathbb{R}^{m\times 1}$ matriisin B sarakkeet, eli olkoon $B = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}$. Lauseen 3.8.4 nojalla $(b_{i_1}, \dots, b_{i_d})$ on aliavaruuden $\mathrm{Col}(B)$ kanta.

Koska yhtälöryhmä $\left[B \mid \tilde{b}\right]$ saadaan yhtälöryhmästä $\left[A \mid b\right]$ rivioperaatiolla, niin on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{m\times m}$, että

$$A = PB.$$

Osoitetaan, että jono $(Pb_{i_1}, \dots, Pb_{i_d})$ eli jono $(v_{i_1}, \dots, v_{i_d})$ on aliavaruuden $\mathrm{Col}(A)$ kanta.

Osoitetaan ensin, että jono $(Pb_{i_1}, \dots, Pb_{i_d})$ virittää aliavaruuden $\mathrm{Col}(A)$. Havaitaan ensin, että $v_i = Pb_i$ jokaisella $i \in \{1, \dots, n\}$, joten $\mathrm{Col}(A) = \mathrm{Sp}(v_1, \dots, v_n) = \mathrm{Sp}(Pb_1, \dots, Pb_n)$. Koska $\mathrm{Sp}(Pb_{i_1}, \dots, Pb_{i_d}) \subset \mathrm{Col}(A)$, niin riittää osoittaa, että $\mathrm{Col}(A) \subset \mathrm{Sp}(Pb_{i_1}, \dots, Pb_{i_d})$. Olkoon $i \in \{1, \dots, n\}$. Koska $b_i \in \mathrm{Sp}(b_{i_1}, \dots, b_{i_d})$, niin on olemassa sellaiset $a_1, \dots, a_d \in \mathbb{R}$, että $b_i = a_1 b_{i_1} + \cdots + a_d b_{i_d}$. Tällöin $Pb_i = P(a_1 b_{i_1} + \cdots + a_d b_{i_d}) = a_1 Pb_{i_1} + \cdots + a_d Pb_{i_d} \in \mathrm{Sp}(Pb_{i_1}, \dots, Pb_{i_d})$. Näin ollen korollaarin 3.4.8 perusteella $\mathrm{Sp}(Pb_1, \dots, Pb_n) \subset \mathrm{Sp}(Pb_{i_1}, \dots, Pb_{i_d})$.

Osoitetaan nyt, että jono $(Pb_{i_1}, \dots, Pb_{i_d})$ on vapaa. Olkoot $a_1, \dots, a_d \in \mathbb{R}$ sellaisia lukuja, että

$$a_1 Pb_{i_1} + \cdots + a_d Pb_{i_d} = 0.$$

Tällöin

$$P(a_1 b_{i_1} + \cdots + a_d b_{i_d}) = 0.$$

Koska P on kääntyvä, niin $a_1 b_{i_1} + \cdots + a_d b_{i_d} = 0$. Koska $(b_{i_1}, \dots, b_{i_d})$ on vapaa jono, niin $a_1 = \cdots = a_d = 0$. Näin ollen jono $(Pb_1, \dots, Pb_{i_d})$ on vapaa.

Jono $(Pb_1, \dots, Pb_n)$ on siis aliavaruuden $\mathrm{Col}(A)$ vapaa virittäjä jono, eli kanta. □

Tarkastellaan nyt konkreettista esimerkkiä tästä tilanteesta.

Esimerkki 3.8.6. *Olkoon*

$$A = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix} = \begin{bmatrix} 1 & -4 & 0 & 2 \\ -1 & 4 & 1 & -3 \\ 2 & -8 & 0 & 4 \end{bmatrix}.$$

Tällöin lisäämällä ensimmäinen rivi toiseen riviin ja vähentämällä ensimmäinen rivi luvulla 2 kerrottuna saadaan esimerkin 3.8.2 matriisi

$$B = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \end{bmatrix} = \begin{bmatrix} 1 & -4 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Koska matriisin B sidotut muuttujat ovat x_1 ja x_3, niin tiedetään, että $\mathrm{Col}(B) = \mathrm{Sp}(b_1, b_3)$. *Näin ollen lauseen 3.8.5 perusteella* $\mathrm{Col}(A) = \mathrm{Sp}(v_1, v_3)$ *eli*

$$\mathrm{Col}(A) = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right).$$

Tuloksen voi tarkistaa varmistamalla, että vektorit v_1 ja v_3 ovat lineaarisesti riippumattomia ja tarkistamalla, että v_2 ja v_4 ovat vektoreiden lineaarikombinaatioita. Tätä ei kuitenkaan tässä tehdä.

Huomautus 3.8.7. *Edellisessä esimerkissä* $\mathrm{Col}(B) = \mathrm{Sp}(b_1, b_3) = \mathrm{Sp}(e_1, e_2)$, *mutta* $\mathrm{Col}(A) = \mathrm{Sp}(v_1, v_3) \neq \mathrm{Sp}(e_1, e_2)$. *Tuleekin siis muistaa, että vaikka sarakeavaruuksien* $\mathrm{Col}(A)$ *ja* $\mathrm{Col}(B)$ *kantavektorit sijaitsevat matriisien samoissa sarakkeissa ne ovat kuitenkin eri sarakevektoreita, eli sarakeavaruudet* $\mathrm{Col}(A)$ *ja* $\mathrm{Col}(B)$ *ovat eri aliavaruuksia.*

Huomautus 3.8.8. *Vaikka edellisessä huomautuksessa painotettiikin sarakeavaruuksien* $\mathrm{Col}(A)$ *ja* $\mathrm{Col}(B)$ *eroa, liittyvät ne kuitenkin toisiinsa kääntyvän matriisin P välityksellä. Jos P on sellainen kääntyvä matriisi, että $A = PB$, niin tällöin*

$$\begin{aligned} \mathrm{Col}(A) &= \{Ax \in \mathbb{R}^{m\times 1} \colon x \in \mathbb{R}^{n\times 1}\} \\ &= \{PBx \in \mathbb{R}^{m\times 1} \colon x \in \mathbb{R}^{n\times 1}\} \\ &= \{Py \in \mathbb{R}^{m\times 1} \colon y = Bx,\ x \in \mathbb{R}^{n\times 1}\} \\ &= \{Py \in \mathbb{R}^{m\times 1} \colon y \in \mathrm{Col}(B)\}. \end{aligned}$$

Näin ollen $\mathrm{Col}(A)$ *on aliavaruuden* $\mathrm{Col}(B)$ *kuva kuvauksessa* $f_P \colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{m\times 1}$. *Vastaavaa havaintoa on käytetty lauseen 3.8.5 todistuksessa päättelyssä, että* $\mathrm{Col}(A) = \mathrm{Sp}(Pb_{i_1}, \ldots, Pb_{i_d})$.

3.9 Esimerkki: Matriisin nolla-avaruuden kanta

Etsitään nyt matriisin $A \in \mathbb{R}^{m\times n}$ nolla-avaruudelle $\mathrm{Null}(A)$ kanta. Tehdään tämä kahdella tavalla. Tarkastellaan ensin esimerkin avulla, kuinka yhtälöryhmän $\left[A \mid 0\right]$ parametrimuotoinen ratkaisu antaa tavan löytää avaruuden $\mathrm{Null}(A)$ kanta. Esimerkki paljastaa, että nolla-avaruuden kantavektorit vastaavat matriisia vastaavan yhtälöryhmän vapaita muuttujia. Tämän faktan todistus on käsitelty tarkasti liitteessä B.

Esimerkki 3.9.1. *Tarkastellaan matriisia*

$$A = \begin{bmatrix} 1 & -1 & 2 & -1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 4}.$$

Tällöin nolla-avaruuden $\mathrm{Null}(A)$ *määrittäminen vastaa yhtälön* $Ax = 0$ *eli yhtälöryhmän*

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right]$$

ratkaisemista. Lisäämällä toinen rivi ensimmäiseen saadaan tämä yhtälöryhmä supistettuun porrasmuotoon

$$\left[\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right]$$

jonka ratkaisujoukko koostuu täsmälleen vektoreista $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$*, joille pätee*

$$\begin{cases} x_1 & = & -2t \\ x_2 & = & -s \\ x_3 & = & t \\ x_4 & = & s, \end{cases}$$

missä $t, s \in \mathbb{R}$.

Näin ollen yhtälön $Ax = 0$ *ratkaisun ovat*

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2t \\ -s \\ t \\ s \end{bmatrix} = t \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

missä $t, s \in \mathbb{R}$.

Merkitään nyt

$$v_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ ja } v_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}.$$

Tällöin

$$\mathrm{Null}(A) = \mathrm{Sp}(v_1, v_2).$$

Osoitetaan nyt, että (v_1, v_2) *on aliavaruuden* $\mathrm{Null}(A)$ *kanta. Koska vektorit* v_1 *ja* v_2 *virittävät aliavaruuden* $\mathrm{Null}(A)$*, niin riittää osoittaa, että* (v_1, v_2) *on vapaa.*

Vektoreista v_1 *ja* v_2 *voidaan suoraan tehdä seuraava havainto: Vektorissa* v_1 *on nollasta poikkeavia kertoimia ainoastaan sellaisilla riveillä, joilla vastaava kerroin vektorissa* v_2 *on kerroin nolla. Sama havainto pätee myös toisin päin.*[2] *Tästä havainnosta*

[2]Huomaa, että on mahdollista tehdä sellainenkin esimerkki, että molemmissa vektoreissa on jollain rivillä kerroin nolla. Tälläisen esimerkin saa vaikkapa vaihtamalla matriisin A ensimmäisellä rivillä kertoimen 2 luvuksi 0.

seuraa suoraan, että vektorit v_1 ja v_2 ovat lineaarisesti riippumattomia. Tehdään tämä nyt tarkasti. Olkoot $a_1, a_2 \in \mathbb{R}$ sellaisia, että

$$a_1 v_1 + a_2 v_2 = 0.$$

Koska

$$a_1 v_1 + a_2 v_2 = a_1 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2a_1 \\ -a_2 \\ a_1 \\ a_2 \end{bmatrix},$$

niin yhtälöstä $a_1 v_1 + a_2 v_2 = 0$ saadaan yhtälöt

$$\begin{cases} -2a_1 &= 0 \\ -a_2 &= 0 \\ a_1 &= 0 \\ -a_2 &= 0 \end{cases}$$

kertoimille. Näin ollen $a_1 = a_2 = 0$, eli vektorit v_1 ja v_2 ovat lineaarisesti riippumattomia.

Näin on osoitettu, että (v_1, v_2) on avaruuden $\operatorname{Null}(A)$ kanta.

Edellisen esimerkin metodia voi soveltaa kaikkissa tilanteissa, joissa matriisin A nolla-avaruus $\operatorname{Null}(A)$ halutaan selvittää käsin. Suurempia matriiseja varten tarvitaan kuitenkin argumentti, joka perustellee vektoreista v_1 ja v_2 tehdyn havainnon. Lisäksi on hyödyllistä havaita, että nolla-avaruudella $\operatorname{Null}(A)$ on sellainen kanta, jonka alkioiden määrä on täsmälleen yhtälöryhmän $\begin{bmatrix} A \mid 0 \end{bmatrix}$ supistetun porrasmuodon $\begin{bmatrix} B \mid 0 \end{bmatrix}$ vapaiden muuttujien määrä.

3.9.1 Kommentteja

Yhdistämällä luvun 3.8 sarakeavaruuden kantaa koskeva tulos ja liitteen B nolla-avaruuden kantaa koskeva tulos saadaan, että yhtäköryhmän $\begin{bmatrix} A & 0 \end{bmatrix}$ supistetun porrasmuodon sidotut muuttujat antavat sarakeavaruuden $\operatorname{Col}(A)$ kanta-alkioiden määrän ja vapaat muuttujat antavat nolla-avaruuden $\operatorname{Null}(A)$ kanta-alkioiden määrän. Koska muuttuja on joko vapaa tai sidottu ja $m \times n$-matriisin A yhtälöryhmässä on n muuttujaa, niin havaitaan, että sarakeavaruuden kanta-alkioiden ja nolla-avaruuden kanta-alkioden lukumäärien summa on n.

Herää kuitenkin kysymys, että onko mahdollista löytää aliavaruuksille $\operatorname{Col}(A)$ ja $\operatorname{Null}(A)$ jollain toisella menetelmällä jotkin toiset kannat, joissa on eri määrä alkiota. Tästä erikoistapauksena herää kysymys voidaanko matriisille A löytää kaksi erilaista supistettua porrasmuotoa, jotka antaisivat eri määrät kanta-alkiota. Vastaus molempiin kysymyksiin on negatiivinen.

Ensimmäinen kysymys johdattaa dimension käsitteeseen. Tätä käsitellään lyhyesti tämän luvun lopussa ja tarkemmin kurssin Lineaarialgebra ja matriisilaskenta II alussa. Toinen kysymys on ratkaistu liitteessä A.

3.10 Sovellus: Matriisin kääntyvyyden karakterisaatioita

Sovelletaan nyt sarake- ja nolla-avaruuksista tehtyjä havaintoja neliömatriisin kääntyvyyteen. Seuraava lause ei anna konkreettista menetelmää matriisin kääntämiseen, mutta se antaa yhtäpitäviä ehtoja matriisin kääntyvyydelle.

Lause 3.10.1. *Olkoon $A \in \mathbb{R}^{n\times n}$ neliömatriisi. Tällöin seuraavat ehdot ovat yhtäpitäviä*

1. *A on kääntyvä,*
2. *matriisin A sarakkeet ovat lineaarisesti riippumattomia,*
3. *matriisin A sarakkeet virittävät avaruuden $\mathbb{R}^{n\times 1}$,*
4. *matriisin A sarakkeet muodostavat avaruuden $\mathbb{R}^{n\times 1}$ kannan,*
5. *yhtälöllä $Ax = b$ on täsmälleen yksi ratkaisu jokaisella $b \in \mathbb{R}^{n\times 1}$,*

Lauseella 3.10.1 on kaksi sovellusta. Kirjataan molemmat niistä korollaareiksi.

Korollaari 3.10.2. *Matriisi $A \in \mathbb{R}^{n\times n}$ on kääntyvä, jos ja vain jos on olemassa sellainen matriisi $B \in \mathbb{R}^{n\times n}$, että $AB = I$ tai $BA = I$.*

Todistus. Mikäli matriisi A on kääntyvä, niin matriisiksi B voidaan valita käänteismatriisi A^{-1} ja tällöin sekä $AB = I$ että $BA = I$.

Olkoon nyt $A \in \mathbb{R}^{n\times n}$ neliömatriisi. Oletetaan, että on olemassa sellainen matriisi $B \in \mathbb{R}^{n\times n}$, jolla $AB = I$. Hyödynnetään nyt tuloksia, että $\mathrm{Col}(AB) \subset \mathrm{Col}(A)$ ja että $\mathrm{Null}(A) \subset \mathrm{Null}(BA)$, jotka jätetään harjoitustehtäväksi.

Koska $\mathbb{R}^{n\times 1} = \mathrm{Col}(AB) \subset \mathrm{Col}(A)$, niin matriisin A sarakeet virittävät avaruuden $\mathbb{R}^{n\times 1}$. Näin ollen lauseen 3.10.1 nojalla matriisi A on kääntyvä. Jos puolestaan on olemassa sellainen matriisi $B \in \mathbb{R}^{n\times n}$, että $BA = I$, niin tällöin $\mathrm{Null}(A) \subset \mathrm{Null}(BA) = \mathrm{Null}(I) = \{0\}$. Näin ollen matriisin A sarakkeet ovat lineaarisesti riippumattomia lemman 3.6.3 perusteella. Näin ollen matriisi A on kääntyvä myös tässä tapauksessa. □

Korollaari 3.10.3. *Olkoon $(v_1, \dots, v_n)$ jono avaruuden $\mathbb{R}^{n\times 1}$ vektoreita. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *$(v_1, \dots, v_n)$ on vapaa,*
2. *$(v_1, \dots, v_n)$ virittää avaruuden $\mathbb{R}^{n\times 1}$ ja*
3. *$(v_1, \dots, v_n)$ on avaruuden $\mathbb{R}^{n\times 1}$ kanta.*

Todistus. Olkoon $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. Lauseen 3.10.1 perusteella jokainen ehdoista (1)–(3) on yhtäpitävää sen kanssa onko matriisi A kääntyvä. Ehdot ovat siis yhtäpitäviä. □

Kirjataan vielä huomio yhtälöryhmien ratkaisemisesesta.

Huomautus 3.10.4. *Lemman 3.6.3 perusteella ehto* (2) *on yhtäpitävää sen kanssa, että yhtälön* $Ax = 0$ *ainoa ratkaisu on* $x = 0$*, että lemman 3.5.2 perusteella ehto* (3) *on yhtäpitävää sen kanssa, että yhtälöllä* $Ax = b$ *on aina ratkaisu. Lauseen 3.10.1 tärkein seuraus onkin, että nämä kaksi ominaisuutta ovat neliömatriiseille aina samanaikaisesti voimassa.*

Lauseen todistus on pitkähkö ja sen voi ohittaa ensimmäisellä lukukerralla.

Lauseen 3.10.1 todistus. Merkitään $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$ matriisin A sarakkeita, eli $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$.

Osoitetaan implikaatio (1) $\Rightarrow$ (2). Oletetaan, että matriisi A on kääntyvä. Osoitetaan, että sarakevektorit $v_1, \ldots, v_n$ ovat lineaarisesti riippumattomia. Olkoot $a_1, \ldots, a_n \in \mathbb{R}$ sellaisia, että

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Merkitään

$$x = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

Tällöin

$$Ax = 0.$$

Koska oletuksen mukaan A on kääntyvä, niin

$$x = A^{-1}0 = 0.$$

Näin ollen $a_1 = \cdots = a_n = 0$. Vektorit $v_1, \ldots, v_n$ ovat siis lineaarisesti riippumattomia. Tämä päätää implikaation (1) $\Rightarrow$ (2) todistuksen.

Osoitetaan nyt implikaatio (2) $\Rightarrow$ (3). Oletetaan, että sarake vektorit $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$ ovat lineaarisesti riippumattomia. Olkoon $B \in \mathbb{R}^{n\times n}$ matriisin A supistettu porrasmuoto ja $P \in \mathbb{R}^{n\times n}$ sellainen kääntyvä matriisi, että $A = PB$. Koska P on kääntyvä ja matriisin A sarakkeet ovat lineaarisesti riippumattomia, niin tällöin myös matriisin B sarakkeet ovat lineaarisesti riippumattomia lauseen 3.6.7 nojalla. Koska B sen sarakkeet ovat lineaarisesti riippumattomia, niin $\mathrm{Null}(B) = \emptyset$ lemman 3.6.3 perusteella. Koska B on supistetussa porrasmuodossa, niin näin ollen yhtälöryhmällä $\left[B \mid 0\right]$ ei ole vapaita muuttujia lauseen 1.4.14 perusteella. Näin ollen yhtälöryhmän $\left[B \mid 0\right]$ kaikki muuttujat ovat sidottuja. Koska matriisilla B on sama määrä rivejä ja sarakkeita, niin näin ollen $B = I$. Näin ollen $A = PB = PI = P$. Osoitetetaan tämän avulla, että $\mathrm{Col}(A) = \mathbb{R}^{n\times 1}$. Olkoon $y \in \mathbb{R}^{n\times 1}$ ja $x = P^{-1}(y) \in \mathbb{R}^{n\times 1}$. Tällöin $y = P(P^{-1}y) = Px = Ax$. Näin ollen $\mathbb{R}^{n\times 1} \subset \mathrm{Col}(A)$, eli $\mathrm{Col}(A) = \mathbb{R}^{n\times 1}$. Tämä päättää implikaation (2) $\Rightarrow$ (3) todistuksen.

Osoitetaan nyt implikaatio (3) $\Rightarrow$ (4). Oletetaan, että $\mathrm{Col}(A) = \mathbb{R}^{n\times 1}$. Olkoon jälleen B matriisin A supistettu porrasmuoto ja $P \in \mathbb{R}^{n\times n}$ sellainen kääntyvä matriisi, että $A = PB$. Osoitetaan, että $\mathrm{Col}(B) = \mathbb{R}^{n\times 1}$.

Olkoon $y \in \mathbb{R}^{n\times 1}$ ja $x = P^{-1}y$. Koska $\mathrm{Col}(A) = \mathbb{R}^{n\times 1}$, niin $x \in \mathrm{Col}(A)$. Näin ollen on olemassa sellainen $z \in \mathbb{R}^{n\times 1}$, että $x = Az$. Koska $B = P^{-1}A$, niin $Bz = P^{-1}Az = P^{-1}x = y$. Näin ollen $y \in \mathrm{Col}(B)$. Tämä osoittaa, että $\mathrm{Col}(B) = \mathbb{R}^{n\times 1}$.

Koska B on supistetussa porrasmuodossa, niin matriisin B sarakkeet ovat standardikannan vektoreita. Koska $\mathrm{Col}(B) = \mathbb{R}^{n\times 1}$, niin saadaan, että $e_i \in \mathrm{Col}(B)$ jokaisella $i \in \{1, \ldots, n\}$. Näin ollen $B = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} = I$. Näin ollen $A = P$. Koska matriisi P on kääntyvä, niin sen sarakkeet ovat lineaarisesti riippumattomia. Näin ollen matriisin A sarakkeet ovat lineaarisesti riippumattomia. Koska $\mathrm{Col}(A) = \mathbb{R}^{n\times 1}$, niin matriisin A sarakkeet muodostavat avaruuden $\mathbb{R}^{n\times 1}$ kannan. Tämä päättää implikaation (3) $\Rightarrow$ (4) todistuksen.

Osoitetaan nyt implikaatio (4) $\Rightarrow$ (5). Olkoon $b \in \mathbb{R}^{n\times 1}$. Koska $(v_1, \ldots, v_n)$ on avaruuden $\mathbb{R}^{n\times 1}$ kanta, niin vektorilla b on yksikäsitteiset kertoimet $x_1, \ldots, x_n \in \mathbb{R}$ kannassa $(v_1, \ldots, v_n)$, eli

$$b = x_1 v_1 + \cdots + x_n v_n = Ax,$$

missä

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Näin ollen yhtälöllä $Ax = b$ on ratkaisu. Jos sarakevektorille

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$$

pätee $Ay = b$, niin tällöin

$$b = Ay = y_1 v_1 + \cdots + y_n v_n,$$

eli $y_i = x_i$ jokaisella $i \in \{1, \ldots, n\}$ kertoimien yksikäsitteisyyden nojalla. Näin ollen yhtälöllä $Ax = b$ on yksikäsitteinen ratkaisu.

Osoitetaan nyt viimeinen implikaatio (5) $\Rightarrow$ (1). Olkoon $(e_1, \ldots, e_n)$ avaruuden $\mathbb{R}^{n\times 1}$ standardikanta. Koska jokaisella yhtälöllä $Ax = b$ on yksikäsitteinen ratkaisu, niin jokaisella $i \in \{1, \ldots, n\}$ on olemassa sellainen vektori $u_i \in \mathbb{R}^{n\times 1}$, että

$$Au_i = e_i.$$

Olkoon nyt

$$B = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Tällöin

$$AB = A \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} Au_1 & \cdots & Au_n \end{bmatrix} = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} = I.$$

Osoitetaan nyt, että myös $BA = I$. Olkoot $w_1, \ldots, w_n \in \mathbb{R}^{n\times 1}$ matriisin BA sarakevektorit, eli $BA = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix}$, ja osoitetaan, että $w_i = e_i$ jokaisella $i \in \{1, \ldots, n\}$.

Koska $(BA)e_i = w_i$, niin

$$Aw_i = A\left((BA)e_i\right) = A\left(B(Ae_i)\right) = AB(Ae_i) = I(Ae_i) = Ae_i.$$

Näin ollen

$$A(w_i - e_i) = 0.$$

Koska kaikilla yhtälöillä $Ax = b$ on yksikäsitteinen ratktaisu ja $A0 = 0$, niin $w_i = e_i$. Näin ollen

$$BA = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} e_1 & \cdots e_n \end{bmatrix} = I.$$

Tämä päättää implikaation (5) $\Rightarrow$ (1) todistukset ja siten koko lauseen todistuksen. $\square$

3.11 Aliavaruuden dimensio

Matriisin sarakeavaruuden kantaa varten kehitetyllä teorialla voidaan osoittaa kaksi tärkeää aliavaruuksiin liittyvää tulosta.

Lause 3.11.1. *Jokainen aliavaruus $V \subset \mathbb{R}^{n\times 1}$ on jonkin $n \times k$-matriisin, missä $1 \le k \le n$, sarakeavaruus. Erityisesti sillä on kanta, jossa on korkeintaan n alkiota.*

Lause 3.11.2. *Jokaisessa aliavaruuden $V \subset \mathbb{R}^{n\times 1}$ kannassa on sama määrä alkiota.*

Yhdessä nämä lauseet sanovat, että avaruuden $\mathbb{R}^{n\times 1}$ aliavaruudella on dimensio.

Määritelmä 3.11.3. *Aliavaruuden $V \subset \mathbb{R}^{n\times 1}$* dimensio $\dim V$ *on aliavaruuden V kannan alkioiden lukumäärä.*

Matriisien kielellä tätä lukua kutsutaan matriisin asteeksi.

Määritelmä 3.11.4. *Matriisin $A \in \mathbb{R}^{n\times k}$* aste $\mathrm{rank}(A)$ *on sarakeavaruuden* $\mathrm{Col}(A)$ *dimensio eli* $\mathrm{rank}(A) = \dim \mathrm{Col}(A)$.

Huomautus 3.11.5. *Koska matriisin A sarakeavaruuden* $\mathrm{Col}(A)$ *dimensio on matriisin A lineaarisesti riippumattomien sarakkeiden lukumäärä, niin tämä otetaan joskus matrsiisin asteen määritelmäksi.*

Näihin tuloksiin palataan laajemmin kurssilla Lineearialgebra ja matriisilaskenta II yleisten vektoriavaruuksien kontektissa. Annetaan lauseille 3.11.1 ja 3.11.2 nyt kuitenkin lyhyet todistukset.

Lauseen 3.11.1 todistus. Käsitellään kaksi eri tapausta. Jos V on nolla-avaruus $\{0\}$, niin tällöin voidaan valita matriisiksi A, mikä tahansa matriisi, jonka kertoimet ovat nollia, esimerkiksi yhden sarakeen matriisi $A = [0] \in \mathbb{R}^{n\times 1}$. Avaruuden V kanta on ().

Olkoon nyt $V \subset \mathbb{R}^{n\times 1}$ aliavaruus, joka ei ole ole nolla-avaruus. Valitaan nyt avaruudesta V vektoreita seuraavalla menetelmällä. Olkoon $v_1 \in V$ nollasta poikkeava vektori.

Oletetaan nyt, että $1 \le k \le n$ on sellainen luku, että on valittu vektorit $v_1, \ldots, v_k \in V$, joille pätee $v_i \notin \mathrm{Sp}(v_1, \ldots, v_{i-1})$ kaikilla $1 \le i \le k$. Jos $\mathrm{Sp}(v_1, \ldots, v_k) = V$ tai $k = n$, niin lopetetetaan vektoreiden valitseminen. Jos $k < n$ ja $\mathrm{Sp}(v_1, \ldots, v_k) \ne V$, niin tällöin $V \setminus \mathrm{Sp}(v_1, \ldots, v_k) \ne \emptyset$ ja voidaan valita vektori $v_{k+1} \in V \setminus \mathrm{Sp}(v_1, \ldots, v_k)$.

Olkoon nyt $A = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \in \mathbb{R}^{n\times k}$ ja osoitetaan, että $\mathrm{Col}(A) = \mathrm{Sp}(v_1, \ldots, v_k) = V$. Jos edellä vektoreiden valinta lopetettiin, koska $\mathrm{Sp}(v_1, \ldots, v_k) = V$, niin väite on automaattisesti tosi. Voidaan siis olettaa, että $k = n$. Huomaa, että vektoreiden $v_1, \ldots, v_n$ valinnan perusteella $\mathrm{Sp}(v_1, \ldots, v_n) \subset V$.

Tällöin $A \in \mathbb{R}^{n\times n}$ on neliömatriisi. Osoitetaan, että matriisin A sarakkeet ovat lineaarisesti riippumattomia. Olkoot $a_1, \ldots, a_n \in \mathbb{R}$ sellaisia lukuja, että

$$a_1v_1 + \cdots + a_nv_n = 0.$$

Olkoon $i \in \{1, \ldots, n\}$ suurin sellainen indeksi, että $a_i \neq 0$. Tällöin

$$v_i = -\frac{a_1}{a_i}v_1 + \cdots + (-\frac{a_{i-1}}{a_i})v_{i-1}.$$

Tämä on ristiriita vektoreiden $v_1, \ldots, v_n$ valinnan perusteella. Näin ollen matriisin A sarakkeet ovat lineaarisesti riippumattomia.

Lauseen 3.10.1 perusteella matriisin A sarakkeet muodostavat avaruuden $\mathbb{R}^{n\times 1}$ kannan. Näin ollen $\mathbb{R}^{n\times 1} = \mathrm{Sp}(v_1, \ldots, v_n) \subset V$, eli $V = \mathrm{Sp}(v_1, \ldots, v_n) = \mathbb{R}^{n\times 1}$. Tämä päättää todistuksen. □

Lauseen 3.11.2 todistus. Koska nolla-avaruudella $V = \{0\}$ on ainoastaan tyhjä kanta, niin voidaan olettaa, että $V \neq \{0\}$. Olkoon $(v_1, \ldots, v_k)$ aliavaruuden V kanta ja olkoon $(w_1, \ldots, w_m)$ jokin toinen aliavaruuden V kanta.

Osoitetaan ensin, että $m \leq k$. Tehdään vastaoletus, että $m > k$. Merkitään $A = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \in \mathbb{R}^{n\times k}$. Koska jono $(v_1, \ldots, v_k)$ on aliavaruuden V kanta, niin vektorit $w_1, \ldots, w_m$ voidaan kirjoittaa (yksikäsitteisesti) vektoreiden $v_1, \ldots, v_k$ lineaarikombinaatioina, eli jokaisella $i \in \{1, \ldots, m\}$ on yksikäsitteinen vektori $u_i \in \mathbb{R}^{k\times 1}$, jolle pätee $w_i = Au_i$. Tarkastellaan nyt matriisia $B = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix} \in \mathbb{R}^{k\times m}$.

Olkoon $C \in \mathbb{R}^{k\times m}$ matriisin B supistettu porrasmuoto. Koska matriisissa C on enemmän sarakkeita kuin rivejä, yhtälöryhmässä $\left[C \mid 0\right]$ on vapaa muuttuja ja siten sillä on äärettömän monta ratkaisua lauseen 1.4.14 perusteella. Näin ollen myös yhtälöryhmällä $\left[B \mid 0\right]$ on äärettömän monta ratkaisua, eli $\mathrm{Null}(B) \neq \{0\}$. Näin ollen matriisin B sarakkeet eivät ole lineaarisesti riippumattomia lemman 3.6.3 perusteella, eli on olemassa sellaiset luvut $a_1, \ldots, a_m \in \mathbb{R}$, jotka eivät kaikki ole nollia, että

$$a_1u_1 + \cdots + a_mu_m = 0.$$

Tällöin

$$a_1w_1 + \cdots + a_mw_m = a_1Au_1 + \cdots + a_mAu_m = A(a_1u_1 + \cdots + a_mu_m) = 0.$$

Näin ollen vektorit $w_1, \ldots, w_m$ eivät ole lineaarisesti riippumattomia. Tämä on ristiriita. Näin ollen $m \leq k$.

Vaihtamalla kantojen $(v_1, \ldots, v_k)$ ja $(w_1, \ldots, w_m)$ roolit havaitaan, että $k \leq m$. Näin ollen $k = m$. □

Luku 4

Lineaarinen geometria

4.1 Motivointi: Pythagoraasta pistetuloon

Tässä luvussa tarkastellaan lineaarialgebran ja geometrian välistä yhteyttä. Tarkastellaan motivoivana esimerkkinä seuraavaa ilmiötä, joka liittyy vektorin koordinaatteihin kannassa.

Pythagoraan lause sanoo, että tason $\mathbb{R}^{2\times 1}$ pisteen $v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ etäisyys origosta on

$$\sqrt{x_1^2 + x_2^2}.$$

Tätä vektorin v ns. *euklidista pituutta* merkitään jatkossa itseisarvomerkinnällä

$$|v| = \left| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right| = \sqrt{x_1^2 + x_2^2}. \tag{4.1}$$

Euklidista pituutta $|v|$ kutsutaan jatkossa lyhyesti *pituudeksi.*

Koska vektorin v koordinaatit standardikannassa ovat (x_1, x_2), niin herää seuraava kysymys:

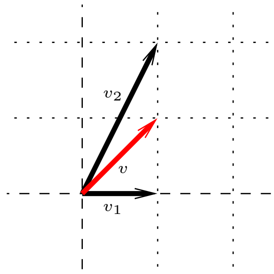
Ongelma. *Voiko vektorin pituuden laskea aina sen kertoimista annetussa kannassa kaavalla* (4.1)*?*

Kuten kysymyksen asettelusta voi arvata, vastaus on, että näin ei voida tehdä. Tarkastellaan nyt konkreettista esimerkkiä.

Esimerkki 4.1.1. *Olkoon* (v_1, v_2) *se avaruuden* $\mathbb{R}^{2\times 1}$ *kanta, missä* $v_1 = e_1$ *ja* $v_2 = e_1 + 2e_2$*. Tarkastellaan avaruuden* $\mathbb{R}^{2\times 1}$ *vektoria* $v = e_1 + e_2$ *tässä kannassa; katso kuva 4.1.*

Pythagoraan lause sanoo, että pisteen $(1, 1)$ *etäisyys origosta on* $\sqrt{2}$*. Toisaalta, jos vektori* v *esitetään kannassa* (v_1, v_2)*, niin saadaan, että*

$$v = \frac{1}{2}v_1 + \frac{1}{2}v_2.$$



Kuva 4.1: Kanta (v_1, v_2) ja vektori v.

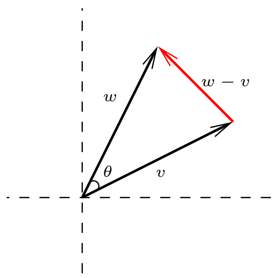
Tästä havaitaan, että pisteen $(1,1)$ *etäisyyttä origosta ei siis voi laskea vektorin* v *kertoimista kannassa* (v_1, v_2)*, sillä*

$$\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{1}{4}} = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}} \neq \sqrt{2}.$$

Parempi vastaus yllä asetettuun ongelmaan on, että *vektorin pituus voidaan laskea kertoimista, jos kanta on ortonormaali*, eli *jos kannan alkiot ovat kohtisuorassa toisiaan vastaan ja pituudeltaan ykkösen mittaisia.*

Paljastuu, että sekä kahden vektorin välinen kohtisuoruus että vektoreiden pituus voidaan määritellä yhden käsitteen – *pistetulon* – avulla. Näiden käsitteiden avulla on puolestaan suoraviivaista määritellä sekä ortonormaalin kannan että ortogonaalimatriisin käsitteet. Pistetulon käsitte puolestaan voidaan motivoida kulman käsitteen avulla.

Esimerkki 4.1.2. *Olkoot* $v = (x_1, x_2)$ *ja* $w = (y_1, y_2)$ *vektoreita vektoriavaruudessa* $\mathbb{R}^2$*. Oletetaan, että kumpikaan vektoreista ei ole nollavektori* 0*; katso kuva 4.2. Tarkastellaan kolmiota jonka kärkipisteet ovat* v*,* w *ja origo sekä vektoreiden* v *ja* w *välistä kulmaa* $\theta \in [0, 2\pi[$ *origossa.*



Kuva 4.2: Vektorit v ja w sekä $w - v$.

Merkitään vektoreiden v*,* w *ja* $w - v$ *pituuksia symboleilla* a*,* b *ja* c*. Tasogeometrian kosinilauseen perusteella*

$$c^2 = a^2 + b^2 - 2ab\cos\theta.$$

(Huomaa, että jos v ja w ovat kohtisuorassa, niin $\theta = \pi/2$ ja $\cos\theta = 0$. Tällöin kosinilause palautuu Pythagoraan lauseeksi.)

Toisaalta Pythagoraan lauseen perusteella tiedetään, että vektoreiden v, w ja $w - v$ pituudet ovat

$$a = \sqrt{x_1^2 + x_2^2}, \quad b = \sqrt{y_1^2 + y_2^2}, \quad \text{ja} \quad c = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}.$$

Näin ollen tiedetään, että

$$(y_1 - x_1)^2 + (y_2 - x_2)^2 = x_1^2 + y_1^2 + x_2^2 + y_2^2 - 2ab\cos\theta.$$

Sieventämisen jälkeen saadaan

$$-2x_1x_2 - 2y_1y_2 = -2ab\cos\theta$$

eli

$$\cos\theta = \frac{x_1x_2 + y_1y_2}{ab} = \frac{x_1x_2 + y_1y_2}{\sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2}}.$$

Vektoreiden v ja w välinen kulma θ riippuu siis vektoreiden v ja w pituuksien lisäksi funktion $\cdot\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$((x_1, y_1), (x_2, y_2)) \mapsto x_1x_2 + y_1y_2,$$

antamasta arvosta.

4.2 Pistetulo ja pituus

Otetaan edellä olevan esimerkin funktio $\cdot\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ lähtökohdaksi ja määritellään vektoreiden välinen pistetulo seuraavasti.

Määritelmä 4.2.1. *Avaruuden $\mathbb{R}^{n\times 1}$* pistetulo *on funktio $\cdot\colon \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1} \to \mathbb{R}$,*

$$\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right) \mapsto x_1y_1 + \cdots + x_ny_n.$$

Avaruuden $\mathbb{R}^{n\times 1}$ vektoreiden v ja w pistetuloa merkitään $v \cdot w$.

Huomautus 4.2.2. *Huomaa, että pistetulon argumentit ovat vektoreita ja pistetulon arvo on reaaliluku. Kysymyksessä ei siis kahden alkion tulo samassa mielessä kuin reaaliluvuilla. Huomaa myös, että pistemerkintää $\cdot$ käytetään jatkossa tarkoittaen sekä vektoreiden pistetuloa että kahden reaaliluvun tuloa. Yleensä tässä ei ole sekaantumisen vaaraa.*

Esimerkki 4.2.3. *Olkoot*

$$v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ ja } w = \begin{bmatrix} 4 \\ -5 \\ 6 \end{bmatrix}$$

avaruuden $\mathbb{R}^{3\times 1}$ *vektoreita. Tällöin*

$$v \cdot w = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ -5 \\ 6 \end{bmatrix} = 1 \cdot 4 + 2 \cdot (-5) + 3 \cdot 6 = 4 - 10 + 18 = 12.$$

Seuraavassa esimerkissä lasketaan avaruuden $\mathbb{R}^{n\times 1}$ standardikannan alkioiden väliset pistetulot.

Esimerkki 4.2.4. *Olkoot* e_i *ja* e_j *avaruuden* $\mathbb{R}^n$ *standardikannan vektoreita. Merkitään*

$$e_i = \begin{bmatrix} e_{1i} \\ \vdots \\ e_{ni} \end{bmatrix} \text{ ja } e_j = \begin{bmatrix} e_{1j} \\ \vdots \\ e_{nj} \end{bmatrix}.$$

Nyt

$$e_i \cdot e_j = e_{1i}e_{1j} + \cdots + e_{ni}e_{nj}.$$

Tarkastellaan ensin tapausta $i \neq j$. *Koska* $e_{ki} = 0$ *kaikilla* $k \neq i$ *ja* $e_{kj} = 0$ *kaikilla* $k \neq j$, *niin* $e_{ki}e_{kj} = 0$ *kaikilla* $k \in \{1, \ldots n\}$. *Näin ollen* $e_i \cdot e_j = 0$. *Toisaalta, jos* $i = j$, *niin* $e_{ii} = e_{ij} = 1$ *ja* $e_{ki} = e_{kj} = 0$ *kaikilla* $k \neq i$. *Näin ollen* $e_i \cdot e_j = e_i \cdot e_i = 1$.

Luvun alussa Pythagoraan lause yhdistettiin tason $\mathbb{R}^{2\times 1}$ vektorien pituuteen. Nyt Pythagoraan lause otetaan pituuden määritelmäksi kaikissa avaruuksissa $\mathbb{R}^{n\times 1}$.

Määritelmä 4.2.5. *Avaruuden* $\mathbb{R}^{n\times 1}$ *vektorin* $v = x_1e_1 + \cdots + x_ne_n$ (euklidinen) pituus *on*

$$|v| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Vektori $v \in \mathbb{R}^{n\times 1}$ *on* yksikkövektori, *jos* $|v| = 1$.

Vektorin pituus on luonnollisella tavalla yhteydessä vektorin sisätuloon itsensä kanssa. Kirjataan tämä havainto lemmaksi.

Lemma 4.2.6. *Olkoon* $v \in \mathbb{R}^{n\times 1}$. *Tällöin*

$$|v| = \sqrt{v \cdot v}.$$

Todistus. Olkoon $v = x_1e_1 + \cdots + x_ne_n \in \mathbb{R}^{n\times 1}$. Tällöin

$$v \cdot v = x_1x_1 + \cdots x_nx_n = x_1^2 + \cdots + x_n^2 = |v|^2.$$

□

Huomautus 4.2.7. *Huomaa, että ainoastaan nolla-vektorin pituus on nolla, eli että vektorille* $v \in \mathbb{R}^{n\times 1}$ *pätee* $|v| = 0$, *jos ja vain jos* $v = 0$.

4.2.1 Pistetulon perusominaisuudet

Pistetulolla on kolme perusominaisuutta.

Lause 4.2.8. *Olkoot $u, v, w \in \mathbb{R}^{n\times 1}$ ja $a \in \mathbb{R}$. Tällöin*

$$(u+v)\cdot w = u\cdot w + v\cdot w \tag{4.2}$$

ja

$$(av)\cdot w = a(v\cdot w). \tag{4.3}$$

Lisäksi

$$v\cdot w = w\cdot v. \tag{4.4}$$

Tulkitaan vielä lauseen sisältö ennen sen todistamista.

Huomautus 4.2.9. *Viimeinen ehto (4.4) sanoo, että pistetulo $\cdot$ on symmetrinen argumenttiensa suhteen, eli että ei ole merkitystä, missä järjestyksessä vektoreita tarkastellaan.*

Huomautus 4.2.10. *Ensimmäiset kaksi omainaisuutta voidaan puolestaan yhdistää sanomalla, että pistetulo on lineaarinen kummankin argumenttinsa suhteen. Funktiota $f\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$ sanotaa* lineaariseksi, *jos kaikilla $v, v' \in \mathbb{R}^{n\times 1}$ ja $a \in \mathbb{R}$ pätee sekä $f(v+w) = f(v) + f(w)$ että $f(av) = af(v)$.*

Pistetulon määritelmän ehtojen (4.2) ja (4.3) perusteella jokaisella $w \in \mathbb{R}^{n\times 1}$ funktio $f_w\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$, $v \mapsto v\cdot w$, on lineaarinen. Symmetrisyyden nojalla funktio f_w voitaisiin antaa myös kaavalla $v \mapsto w\cdot v$. Näin ollen pistetulo on lineaarinen kummankin argumenttinsa suhteen eli ns. bilineaarinen.

Todistetaan nyt lause 4.2.8.

Lauseen 4.2.8 todistus. Olkoot

$$u = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},\ v = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \text{ ja } w = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

Tällöin

$$u+v = \begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

ja

$$\begin{aligned}(u+v)\cdot w &= \begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix}\cdot\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \\ &= (x_1+y_1)z_1 + \cdots (x_n+y_n)z_n \\ &= (x_1z_1 + y_1z_1) + \cdots + (x_nz_n + y_nz_n).\end{aligned}$$

Uudelleen ryhmittelemällä saadaan, että

$$\begin{aligned}(u+v)\cdot w &= (x_1z_1+y_1z_1)+\cdots+(x_nz_n+y_nz_n)\\ &= x_1z_1+\cdots+x_nz_n+y_1z_1+\cdots+y_nz_n\\ &= u\cdot w+v\cdot w.\end{aligned}$$

Tämä todistaa ensimmäisen väitteen eli kaavan (4.2).

Osoitetaan nyt (4.3). Koska

$$av = \begin{bmatrix} ay_1 \\ \vdots \\ ay_n \end{bmatrix},$$

niin uudelleen ryhmittelemällä saadaan

$$\begin{aligned}(av)\cdot w = (ay_1)z_1+\cdots+(ay_n)z_n &= a(y_1z_1)+\cdots+a(y_nz_n)\\ = a\left(y_1z_1+\cdots+y_nz_n\right) &= a(v\cdot w).\end{aligned}$$

Tämä todistaa toisen väitteen.

Symmetrisyyden osoittamiseksi riittää havaita, että

$$v\cdot w = y_1z_1+\cdots y_nz_n = z_1y_1+\cdots+z_ny_n = w\cdot v.$$

Väite on näin todistettu. □

Pistetulon bilineaarisuutta käytetään konkreettisissa laskuissa jatkuvasti ilman erillistä mainintaa. Tämä pätee erityisesti, kun vektorit on ilmoitettu standardikannassa, kuten seuraavassa esimerkissä.

Esimerkki 4.2.11. *Olkoot $v_1 = (e_1+e_2)/\sqrt{2}$ ja $v_2 = (e_1-e_2)/\sqrt{2}$ tason $\mathbb{R}^{2\times 1}$ vektoreita. Tällöin*

$$\begin{aligned}v_1\cdot v_2 &= \Big((e_1+e_2)/\sqrt{2}\Big)\cdot\Big((e_1-e_2)/\sqrt{2}\Big)\\ &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\left((e_1+e_2)\cdot(e_1-e_2)\right)\\ &= \frac{1}{2}\left(e_1\cdot(e_1-e_2)+e_2\cdot(e_1-e_2)\right)\\ &= \frac{1}{2}\left(e_1\cdot e_1+e_1\cdot(-e_2)+e_2\cdot e_1+e_2\cdot(-e_2)\right)\\ &= \frac{1}{2}\left(e_1\cdot e_1-e_1\cdot e_2+e_1\cdot e_2-e_2\cdot e_2\right)\\ &= \frac{1}{2}\left(e_1\cdot e_1-e_2\cdot e_2\right) = \frac{1}{2}(1-1) = 0.\end{aligned}$$

Hieman abstraktimpi esimerkki saadaan pituuden neliöstä.

Esimerkki 4.2.12. *Olkoot $v, w \in \mathbb{R}^{n\times 1}$ vektoreita. Tällöin*

$$\begin{aligned}
|v+w|^2 &= (v+w)\cdot(v+w) \\
&= v\cdot(v+w) + w\cdot(v+w) \\
&= v\cdot v + v\cdot w + w\cdot v + w\cdot w \\
&= |v|^2 + 2v\cdot w + |w|^2.
\end{aligned}$$

4.2.2 Kohtisuoruus

Esimerkissä 4.1.2 havaittiin, että tason $\mathbb{R}^2$ vektorit v ja w ovat kohtisuorassa toisiaan vastaan, jos ja vain jos $v \cdot w = 0$. Tämän motivoimana määritellään avaruuden $\mathbb{R}^{n\times 1}$ vektoreiden kohtisuoruus seuraavasti.

Määritelmä 4.2.13. *Avaruuden $\mathbb{R}^{n\times 1}$ vektorit v ja w ovat* kohtisuorassa toisiaan vastaan*, jos $v \cdot w = 0$.*

Esimerkki 4.2.14. *Avaruuden $\mathbb{R}^{3\times 1}$ vektorit $e_1 - e_2$ ja $2e_1 + 2e_2 + 2e_3$ ovat kohtisuorassa toisiaan vastaan, sillä*

$$(e_1 - e_2)\cdot(2e_1 + 2e_2 + 2e_3) = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = 1\cdot 2 + (-1)\cdot 2 + 0 \cdot 2 = 2 - 2 = 0.$$

Geometrinen intuitiomme sanoo, että kohtisuorat vektorit ovat lineaarisesti riippumattomia. Osoitetaan tämä nyt tarkasti lähtien kohtisuoruuden määritelmästä.

Lause 4.2.15. *Olkoot $v_1, \dots, v_k \in \mathbb{R}^n$ nollasta poikkeavia vektoreita, jotka ovat kohtisuorassa toisiaan vastaan eli $v_i \cdot v_j = 0$ kaikilla $i \neq j$. Tällöin jono $(v_1, \dots, v_k)$ on vapaa.*

Todistus. Olkoot $a_1, \dots, a_k \in \mathbb{R}$ sellaisia reaalilukuja, että

$$a_1 v_1 + \cdots + a_k v_k = 0.$$

Osoitetaan, että $a_j = 0$ jokaisella $j \in \{1, \dots, k\}$.

Olkoon $j \in \{1, \dots, k\}$. Koska $v_j \cdot v_i = 0$ kaikilla $i \neq j$, niin pistetulon bilineaarisuuden nojalla saadaan

$$\begin{aligned}
v_j \cdot (a_1 v_1 + \cdots + a_k v_k) &= a_1(v_j \cdot v_1) + \cdots + a_k(v_j \cdot v_k) \\
&= a_j v_j \cdot v_j = a_j |v_j|^2.
\end{aligned}$$

Näin ollen

$$a_j |v_j|^2 = v_j \cdot (a_1 v_1 + \cdots + a_k v_k) = 0.$$

Koska $|v_j|^2 > 0$ jokaisella $j \in \{1, \dots, k\}$, niin saadaan, että $a_j = 0$ jokaisella $j \in \{1, \dots, k\}$, joka haluttiin todistaa. $\square$

4.3 Ortonormaali kanta

Määritellään nyt ortonormaalin kannan käsite ja osoitetaan, että ortonormaaleilla kannoilla on ominaisuus, että vektorin pituus voidaan laskea vektorin koordinaateista kannan suhteen.

Määritelmä 4.3.1. *Avaruuden $\mathbb{R}^{n\times 1}$ aliavaruuden $V \subset \mathbb{R}^{n\times 1}$ kanta $(v_1, \ldots, v_n)$ on* ortonormaali, *jos*

1. $v_i \cdot v_j = 0$ *kaikilla* $i \neq j$ *ja*

2. $|v_i| = 1$ *jokaisella* $i \in \{1, \ldots n\}$.

Huomautus 4.3.2. *Kroneckerin symboleita käyttäen ortonormaalin kannan ehdot voitaisiin kirjoittaa myös lyhyesti muodossa*

$$v_i \cdot v_j = \delta_{ij}$$

kaikilla $i, j \in \{1, \ldots, n\}$.

Esimerkki 4.3.3. *Esimerkin 4.2.4 perusteella avaruuden* $\mathbb{R}^{n\times 1}$ *standardikannan* $(e_1, \ldots, e_n)$ *vektorit ovat kohtisuorassa toisiaan vastaan. Lisäksi* $|e_i| = \sqrt{e_i \cdot e_i} = 1$ *kaikilla* $i \in \{1, \ldots, n\}$. *Näin ollen* $(e_1, \ldots, e_n)$ *on ortonormaali kanta.*

Esimerkki 4.3.4. *Olkoot* $v_1 = (e_1 + e_2)/\sqrt{2}$ *ja* $v_2 = (e_1 - e_2)/\sqrt{2}$ *tason* $\mathbb{R}^{2\times 1}$ *vektoreita. Esimerkin 4.2.11 perusteella* $v_1 \cdot v_2 = 0$. *Näin ollen jono* (v_1, v_2) *on vapaa ja siten avaruuden* $\mathbb{R}^{2\times 1}$ *kanta korollaarin 3.10.3 perusteella.*

Koska

$$|v_1| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1$$

ja

$$|v_2| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1,$$

niin (v_1, v_2) *on tason* $\mathbb{R}^{2\times 1}$ *ortonormaali kanta.*

Esimerkki 4.3.5. *Olkoot*

$$u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix} \in \mathbb{R}^{3\times 1} \text{ ja } u_2 = \frac{1}{\sqrt{3}}\begin{bmatrix}-1\\1\\1\end{bmatrix} \in \mathbb{R}^{3\times 1}$$

sekä olkoon

$$P = \mathrm{Sp}(u_1, u_2) \subset \mathbb{R}^{3\times 1}.$$

Koska $u_1 \cdot u_2 = 0$, *niin jono* (u_1, u_2) *on vapaa. Näin ollen* (u_1, u_2) *on aliavaruuden* P *kanta. Koska lisäksi* $u_1 \cdot u_1 = 1$ *ja* $u_2 \cdot u_2 = 1$, *niin* (u_1, u_2) *on aliavaruuden* P *ortonormaali kanta.*

4.3.1 Pythagoraan lause ortonormaalille kannalle

Todistetaan nyt ortonormaalin kannan perusominaisuus, että ortonormaalille kannalle pätee *Pythagoraan lause*, eli että vektorin pituus voidaan laskea sen koordinaateista ortonormaalissa kannassa.

Lause 4.3.6 (Pythagoraan lause ortonormaalille kannalle)**.** *Olkoon $V \subset \mathbb{R}^n$ avaruuden $\mathbb{R}^n$ aliavaruus ja olkoon $(u_1, \ldots, u_k)$ aliavaruuden V ortonormaali kanta. Olkoon $v = a_1u_1 + \cdots + a_ku_k \in V$. Tällöin*

$$|v| = \sqrt{a_1^2 + \cdots + a_k^2}.$$

Todistus. Todistus on suora lasku, jossa käytetään pistetulon bilineaarisuutta ja tietoa, että $u_j \cdot u_i = \delta_{ji}$ kaikilla $j, i \in \{1, \ldots, k\}$.

Olkoon

$$v = a_1u_1 + \cdots + a_ku_k \in V.$$

Tällöin

$$\begin{aligned}
|v|^2 = v \cdot v &= (a_1u_1 + \cdots + a_ku_k) \cdot (a_1u_1 + \cdots + a_ku_k) \\
&= \sum_{j=1}^{k} a_ju_j \cdot (a_1u_1 + \cdots + a_ku_k) \\
&= \sum_{j=1}^{k} \sum_{i=1}^{k} a_ju_j \cdot (a_iu_i) \\
&= \sum_{j=1}^{k} \sum_{i=1}^{k} a_ja_i \left(u_j \cdot u_i\right).
\end{aligned}$$

Koska $u_j \cdot u_i = 0$ kaikilla $j \neq i$ ja $u_j \cdot u_j = 1$ kaikilla j, niin saadaan

$$|v|^2 = \sum_{j=1}^{k} a_ja_j = \sum_{j=1}^{k} a_j^2 = a_1^2 + \cdots + a_k^2.$$

Väite seuraa ottamalla neliöjuuri. □

4.3.2 Sarakevektorin koordinaatit ortonormaalissa kannassa

Todistetaan seuraavaksi, että vektorin koordinaatit ortonormaalissa kannassa voidaan laskea pistetulon avulla. Tämä tarkoittaa sitä, että vektorin koordinaatit voidaan laskea suoraan eikä niiden selvittämiseen tarvita yhtälöryhmän ratkaisemista.

Lause 4.3.7. *Olkoon $V \subset \mathbb{R}^n$ avaruuden $\mathbb{R}^n$ aliavaruus ja olkoon $(u_1, \ldots, u_k)$ aliavaruuden V ortonormaali kanta. Olkoon $v \in V$. Tällöin*

$$v = (u_1 \cdot v)u_1 + \cdots + (u_k \cdot v)u_k.$$

Lauseen todistus perustuu seuraavaan huomioon.

Lemma 4.3.8. *Olkoon $V \subset \mathbb{R}^n$ avaruuden $\mathbb{R}^n$ aliavaruus ja olkoon $(u_1, \dots, u_k)$ aliavaruuden V ortonormaali kanta. Jos vektori $v \in V$ on kohtisuorassa kaikkia kantavektoreita $u_1, \dots, u_k$ vastaan, niin $v = 0$.*

Todistus. Oletetaan, että $v \in V$ on sellainen vektori, että $u_i \cdot v = 0$ kaikilla $i \in \{1, \dots, k\}$. Koska $(u_1, \dots, u_k)$ on aliavaruuden V kanta, niin on olemassa sellaiset yksikäsitteliset luvut $a_1, \dots, a_k \in \mathbb{R}$, että

$$v = a_1 u_1 + \cdots + a_k u_k.$$

Olkoon nyt $i \in \{1, \dots, k\}$. Oletuksen nojalla $v \cdot u_i = 0$. Koska $(u_1, \dots, u_k)$ on ortonormaali kanta, niin

$$\begin{aligned} 0 = v \cdot u_i &= (a_1 u_1 + \cdots + a_k u_k) \cdot u_i \\ &= a_1(u_1 \cdot u_i) + \cdots + a_k(u_k \cdot u_i) \\ &= a_i. \end{aligned}$$

Näin ollen $a_i = 0$ jokaisella $i \in \{1, \dots, k\}$, eli $v = 0$. □

Lauseen 4.3.7 todistus. Olkoon

$$w = (u_1 \cdot v)u_1 + \cdots + (u_k \cdot v)u_k \in V.$$

Osoitetaan, että $v - w = 0$. Lemma 4.3.8 nojalla riittää osoittaa, että $(v - w) \cdot u_i = 0$ jokaisella $i \in \{1, \dots, n\}$.

Olkoon $i \in \{1, \dots, k\}$. Tällöin

$$\begin{aligned} (v - w) \cdot u_i &= v \cdot u_i - w \cdot u_i \\ &= v \cdot u_i - ((u_1 \cdot v)u_1 + \cdots + (u_k \cdot v)u_k) \cdot u_i \\ &= v \cdot u_i - u_i \cdot v = 0. \end{aligned}$$

Näin ollen $v - w = 0$ eli $v = w$. Tämä todistaa väitteen. □

4.3.3 Kannan ortonormeeraus: Gram–Schmidt

Yksi koko lineaarialgebran tärkeimmistä tuloksista on, että jokaisesta kannasta voidaan siirtyä ortonormaaliin kantaan niin sanotulla Gram–Schimidt ortonormeerausprosessilla. Todistetaan tästä tuloksesta versio, jossa osa kannan vektoreista ovat kohtisuorassa toisiaan vastaan olevia yksikkövektoreita.

Lause 4.3.9 (Gram–Schmidt). *Olkoon $V \subset \mathbb{R}^{n \times 1}$ aliavaruus ja $(v_1, \dots, v_k)$ aliavaruuden V sellainen kanta, että vektorit $v_1, \dots, v_m$ ovat kohtisuorassa toisiaan vastaan olevia yksikkövektoreita jollain $0 \leq m \leq k$. Tällöin on olemassa sellainen aliavaruuden V ortonormaalikanta $(u_1, \dots, u_k)$, että $u_i = v_i$ jokaisella $1 \leq i \leq m$ ja että jokaisella $m + 1 \leq \ell \leq k$ pätee*

$$\mathrm{Sp}(u_1, \dots, u_\ell) = \mathrm{Sp}(v_1, \dots, v_\ell). \tag{4.5}$$

Lauseen 4.3.9 tulkinta on seuraava. Jos on annettu aliavaruus V ja sille sellainen kanta $(v_1, \ldots, v_k)$, jossa vektorit $v_1, \ldots, v_m$ ovat toisiaan vastaan kohtisuorassa olevia yksikkövektoreita, niin loput vektoreista $v_{m+1}, \ldots, v_k$ voidaan vaihtaa sellaisiin yksikkövektoreihin $u_{m+1}, \ldots, u_k$, jotka ovat kohtisuorassa sekä toisiaan että vektoreita $v_1, \ldots, v_m$ vastaan. Lisäksi yhtälö (4.5) ilmaisee, että tässä vaihdossa vektoreiden järjestys ei muutu siinä mielessä, että uusien vektoreiden virittämät aliavaruudet ovat samat kuin vanhojen.

Huomautus 4.3.10. *Huomaa, että yhden vektorin (v_1) jono toteuttaa automaattisesti kohtisuoruusehdon, eli tapauksessa $m = 1$, ainoa oletus on, että vektori v_1 on yksikkövektori. Tapaus $m = 0$ tulkitaan siten, että vektoreille $v_1, \ldots, v_m$ ei ole asetettu ehtoja.*

Huomautus 4.3.11. *Usein ortonormaaliin kantaan liityvä Gram–Schmidt ortonormeeraus tulos kirjoitetaan muodossa, että* jokaisella alivaruudella $V \subset \mathbb{R}^n$ on ortonormaali kanta. *Tämä tulos saadaan yhdistämällä Gram–Schmidt ortonormeerausprossiin aiempi tulos, että jokaisella aliavaruudella on kanta.*

Todistetaan nyt lause 4.3.9. Todistus perustuu seuraavaan lemmaan, joka kertoo, kuinka ortonormaalia jonoa voidaan laajentaa uudella vektorilla. Lemma voidaan ajatella lauseen 4.3.9 todistuksen induktioaskeleeksi.

Lemma 4.3.12. *Olkoot $V \subset \mathbb{R}^n$ aliavaruus, $(u_1, \ldots, u_m)$ jono aliavaruuden V toisiaan vastaan kohtisuorassa olevia yksikkövektoreita ja $v \in V$ sellainen vektori, että $v \notin \mathrm{Sp}(u_1, \ldots, u_m)$. Tällöin on olemassa sellainen yksikkövektori $u_{m+1} \in V$, että jonon $(u_1, \ldots, u_{m+1})$ vektorit ovat toisiaan vastaan kohtisuorassa olevia yksikkövektoreita ja $\mathrm{Sp}(u_1, \ldots, u_m, v) = \mathrm{Sp}(u_1, \ldots, u_{m+1})$.*

Todistus. Olkoon

$$w = (v \cdot u_1)u_1 + \cdots + (v \cdot u_m)u_m \in V.$$

Tällöin $w \in \mathrm{Sp}(u_1, \ldots, u_m)$.

Koska $v \notin \mathrm{Sp}(u_1, \ldots, u_m)$, niin $v \neq w$ ja $|v - w| > 0$. Näin ollen voidaan määritellä

$$u_{m+1} = \frac{v - w}{|v - w|} \in V.$$

Selvästi u_{m+1} on yksikkövektori, eli $|u_{m+1}| = 1$.

Koska

$$u_{m+1} = \frac{1}{|v - w|}v - \frac{1}{|v - w|}w,$$

niin $u_{m+1} \in \mathrm{Sp}(u_1, \ldots, u_m, v)$.

Osoitetaan nyt, että

$$\mathrm{Sp}(u_1, \ldots, u_{m+1}) = \mathrm{Sp}(u_1, \ldots, u_m, v).$$

Koska $u_{m+1} \in \mathrm{Sp}(u_1, \ldots, u_m, v)$, niin tämän todistamiseksi riittää osoittaa, että $v \in \mathrm{Sp}(u_1, \ldots, u_{m+1})$. Koska $u_{m+1}, w \in \mathrm{Sp}(u_1, \ldots, u_{m+1})$ ja

$$|v-w|u_{m+1} + w = |v-w|\frac{v-w}{|v-w|} + w = v - w + w = v,$$

niin $v \in \mathrm{Sp}(u_1, \ldots, u_{m+1})$.

Jäljellä on osoittaa, että vektori u_{m+1} on kohtisuorassa kaikkia vektoreita $u_1, \ldots, u_m$ vastaan. Olkoon $1 \leq j \leq m$. Koska jonon $(u_1, \ldots, u_m)$ vektorit ovat keskenään kohtisuorassa olevia yksikkövektoreita, niin

$$\begin{aligned} w \cdot u_j &= ((v \cdot u_1)u_1 + \cdots + (v \cdot u_m)u_m) \cdot u_j \\ &= (v \cdot u_1)u_1 \cdot u_j + \cdots + (v \cdot u_m)u_m \cdot u_j \\ &= (v \cdot u_j)u_j \cdot u_j = v \cdot u_j. \end{aligned}$$

Näin ollen

$$\begin{aligned} u_{m+1} \cdot u_j &= \frac{v-w}{|v-w|} \cdot u_j \\ &= \frac{1}{|v-w|}(v_i - w_i) \cdot u_j \\ &= \frac{1}{|v-w|}(v \cdot u_j - w \cdot u_j) \\ &= \frac{1}{|v-w|}(v \cdot u_j - v \cdot u_j) = 0. \end{aligned}$$

Tämä päättää todistuksen. □

Todistetaan nyt lause 4.3.9.

Lauseen 4.3.9 todistus. Olkoon $(v_1, \ldots, v_k)$ aliavaruuden V kanta ja olkoon $0 \leq m \leq k$ sellainen luku, että jonon $(v_1, \ldots v_m)$ vektorit ovat toisiaan vastaan kohtisuorassa olevia yksikkövektoreita.

Jos $m \neq 0$, niin asetetaan $u_i = v_i$ jokaisella $1 \leq i \leq m$. Jos $m = k$, niin tällöin $(u_1, \ldots, u_k)$ on aliavaruuden V ortonormaali kanta. Voidaan siis olettaa, että $m < k$.

Muodostetaan ortonormaali kanta $(u_1, \ldots, u_k)$ induktiolla seuraavasti.

Olkoon $i = m+1$. Tällöin soveltamalla lemmaa 4.3.12 vektoriin $v = v_{m+1}$ havaitaan, että on olemassa sellainen yksikkövektori $u_{m+1} \in V$, että jonon $(u_1, \ldots, u_{m+1})$ vektorit ovat toisiaan vastaan kohtisuorassa olevia yksikkövektoreita, joille pätee

$$\mathrm{Sp}(u_1, \ldots, u_{m+1}) = \mathrm{Sp}(v_1, \ldots, v_{m+1}).$$

Oletetaan nyt, että $m \leq i < k$ on sellainen indeksi, että on olemassa jono $(u_1, \ldots, u_i)$ toisiaan vastaan kohtisuorassa olevia yksikkövektoreita, joille pätee

$$\mathrm{Sp}(u_1, \ldots, u_\ell) = \mathrm{Sp}(v_1, \ldots, v_\ell)$$

jokaisella $1 \leq \ell \leq i$. Soveltamalla lemmaa 4.3.12 vektoriin $v = v_{i+1}$ havaitaan, että on olemassa sellainen yksikkövektori $u_{i+1} \in V$, että jonon $(u_1, \ldots, u_{i+1})$ vektorit ovat toisiaan vastaan kohtisuorassa olevia yksikkövektoreita, joille pätee

$$\mathrm{Sp}(u_1, \ldots, u_{i+1}) = \mathrm{Sp}(v_1, \ldots, v_{i+1}).$$

Tämä päättää todistuksen. □

Huomautus 4.3.13. *Lauseen 4.3.9 todistus paljastaa, että ortonormaalikanta löydetään selkeän algoritmin avulla. Tämän vuoksi lausetta 4.3.9 kutsutaan* kannan Gram–Schmidt ortonormeerausprosessiksi. *Edellisessä todistuksessa induktioaskel vastaa tätä algoritmia. Algoritmin toimivuus taas puolestaan on sisällytetty lemmaan 4.3.12.*

Tilanteessa, jossa ei vielä tiedetä kannan $(v_1, \ldots, v_k)$ sisältävän ortogonaalisia yksikkövektoreita (eli tilanteessa $m = 0$), algoritmi on seuraava:

1. *Valitaan*
$$u_1 = \frac{v_1}{|v_1|}.$$

2. *Oletetaan, että on löydetty vektorit $u_1, \ldots, u_m$.*

3. *Valitaan vektori u_{m+1} kaavalla*
$$u_{m+1} = \frac{v_{m+1} - w_{m+1}}{|v_{m+1} - w_{m+1}|},$$
missä $w_{m+1} = (v_{m+1} \cdot u_1)u_1 + \cdots + (v_{m+1} \cdot u_m)u_m$.

4. *Toistetaan askeleita 2 ja 3 kunnes $m + 1 = k$.*

Sovelletaan edellisen huomautuksen algoritmia konkreettisessa tilanteessa.

Esimerkki 4.3.14. *Olkoot*

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{ja} \quad v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Koska $v_2 - v_1 = e_1$ ja $v_1 - e_1 = e_2$, niin $\mathrm{Sp}(v_1, v_2) = \mathbb{R}^{2\times 1}$. Näin ollen (v_1, v_2) on avaruuden $\mathbb{R}^{2\times 1}$ kanta korollaarin 3.10.3 perusteella.

Etsitään ensin vektori u_1. Koska $|v_1| = \sqrt{2}$, niin

$$u_1 = \frac{v_1}{|v_1|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Etsitään nyt vektori u_2. Koska

$$v_2 \cdot u_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \frac{1}{\sqrt{2}} (2 \cdot 1 + 1 \cdot 1) = \frac{3}{\sqrt{2}},$$

niin

$$w_2 = (v_2 \cdot u_1)u_1 = \frac{3}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}\right) = \frac{3}{2}\begin{bmatrix}1\\1\end{bmatrix}.$$

Näin ollen

$$v_2 - w_2 = \begin{bmatrix}2\\1\end{bmatrix} - \frac{3}{2}\begin{bmatrix}1\\1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix}.$$

ja

$$|v_2 - w_2| = \frac{1}{2}\sqrt{2} = \frac{1}{\sqrt{2}}.$$

Näin saadaan, että

$$u_2 = \frac{v_2 - w_2}{|v_2 - w_2|} = \frac{1}{|v_2 - w_2|}(v_2 - w_2) = \sqrt{2}\frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}.$$

Tarkistetaan vielä, että vektorit u_1 ja u_2 ovat kohtisuorassa toisiaan vastaan. Koska

$$u_1 \cdot u_2 = \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}\right) \cdot \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}\right) = \left(\frac{1}{\sqrt{2}}\right)^2 (1 \cdot 1 + 1 \cdot (-1)) = 0,$$

niin näin pätee.

4.3.4 Gram–Schmidt ortonormeerauksen matriisitulkinta

Yksi syy lauseen 4.3.9 tarkemmalle muotoilulle on seuraava Gram–Schmidt prosessin yhteys yläkolmiomatriiseihin. Annetaan tätä varten yleinen yläkolmiomatriisin määritelmä.

Määritelmä 4.3.15. *Matriisi $A = [a_{ji}] \in \mathbb{R}^{m \times n}$ on* yläkolmiomatriisi, *jos $a_{ji} = 0$ kaikilla $j > i$.*

Esimerkki 4.3.16. *Matriisit*

$$\begin{bmatrix}2 & 0 & 1\\0 & -1 & 2\end{bmatrix} \text{ ja } \begin{bmatrix}3 & 1 & 1\\0 & 1 & 2\\0 & 0 & 3\end{bmatrix}$$

ovat yläkolmiomatriiseja.

Lause 4.3.17. *Olkoon $(v_1, \dots, v_n)$ avaruuden $\mathbb{R}^{n \times 1}$ kanta ja $P = \begin{bmatrix}v_1 & \cdots & v_n\end{bmatrix} \in \mathbb{R}^{n \times n}$. Tällöin on olemassa sellainen avaruuden $\mathbb{R}^n$ ortonormaalikanta $(u_1, \dots, u_n)$, että*

$$P = UR,$$

missä $U = \begin{bmatrix}u_1 & \cdots & u_n\end{bmatrix} \in \mathbb{R}^{n \times n}$ on ortogonaalimatriisi ja matriisi $R \in \mathbb{R}^{n \times n}$ on yläkolmiomatriisi.

Todistus. Olkoon $(u_1, \dots, u_n)$ avaruuden $\mathbb{R}^n$ ortonormaalikanta kuten lauseessa 4.3.9, eli että

$$\mathrm{Sp}(u_1, \dots, u_\ell) = \mathrm{Sp}(v_1, \dots, v_\ell) \tag{4.6}$$

jokaisella $1 \le \ell \le n$.

Tarkastellaan ortonormaalin kannan $(u_1, \dots, u_n)$ kantavektoria u_i. Ehdon (4.6) perusteella tiedetään, että $v_i \in \mathrm{Sp}(u_1, \dots, u_i)$, eli että vektori v_i on kantavektoreiden $u_1, \dots, u_i$ lineaarikombinaatio

$$v_i = a_{1i}u_1 + \cdots + a_{ii}u_i.$$

Vektorin v_i esitys kannassa $(u_1, \dots, u_n)$ on siis

$$v_i = a_{1i}u_1 + \cdots + a_{ii}u_i + 0u_{i+1} + \cdots + 0u_n.$$

Olkoon $a_i \in \mathbb{R}^n$ sarakevektori

$$a_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ii} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

jokaisella $i = 1, \dots, n$. Tällöin

$$v_i = Ua_i$$

jokaisella $i = 1, \dots, n$, eli

$$P = UR,$$

missä

$$R = [a_{ji}] = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Sarakevektoreiden a_i määritelmän perusteella, $a_{ji} = 0$ kaikilla $j > i$. Näin ollen matriisi R on yläkolmiomatriisi. □

4.4 Pistetulon matriisiesitys: Matriisin transpoosi

Avaruuden $\mathbb{R}^{n \times 1}$ vektoreiden

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ ja } y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

pistetulo voidaan kirjoittaa matriisitulona seuraavasti

$$x \cdot y = x_1y_1 + \cdots + x_ny_n = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Pistetulo voidaan siis tulkita matriisitulona, jossa sarakevektoria y kerrotaan vasemmalta rivivektorilla, jossa on sama kertoimet kuin sarakevektorissa x. On kuitenkin huomattavaa, että rivivektori $\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ ei kuitenkaan ole vektori x. Formaalisti tämä ero havaitaan siitä, että siinä, missä $x \in \mathbb{R}^{n\times 1}$, niin $\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \in \mathbb{R}^{1\times n}$. On kuitenkin luontevaa ajatella, että rivivektori $\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ vastaa sarakevektoria x. Tätä vastaavuutta kutsutaan transpoosiksi ja se voidaan määritellä kaikille matriiseille.

Määritelmä 4.4.1. *Matriisin* $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ transpoosi *on matriisi* $A^t = [b_{ij}] \in \mathbb{R}^{n\times m}$*, jonka rivit ovat matriisin* A *sarakkeet, eli kaikilla* $i \in \{1, \dots, n\}$ *ja* $j \in \{1, \dots, m\}$ *pätee* $b_{ij} = a_{ji}$.

Huomautus 4.4.2. *Transpoosille on useita merkintöjä:* $t(A)$, $\tau(A)$, A^T, ${}^tA$ *tai* A'. *Näissä luentomuistiinpanoissa käytetään ainoastaan merkintää* A^t.

Matriisin

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}$$

transpoosi on siis matriisi

$$A^t = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & \ddots & & a_{m2} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & & a_{mn} \end{bmatrix}.$$

Esimerkki 4.4.3.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad \text{ja} \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}^t = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}.$$

Suoraan transpoosin määritelmästä on helppo johtaa laskusääntöjä matriisin transpoosille. Kirjataan niistä ylös muutamia.

Lemma 4.4.4. *Olkoot* $A, B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times k}$ *ja* $a \in \mathbb{R}$. *Tällöin*

1. $(A^t)^t = A$,
2. $(A + B)^t = A^t + B^t$,
3. $(cA)^t = cA^t$ *ja*
4. $(AC)^t = C^tA^t$.

Erityisesti, jos matriisi $E \in \mathbb{R}^{n\times n}$ *on kääntyvä, niin* $(E^{-1})^t = (E^t)^{-1}$.

Huomaa, että kohdassa (4) transpoosi vaihtaa matriisien A ja C järjestyksen.

Todistus. Todistetaan kohdan (4) väite. Muut väitteet jätetään harjoitustehtäviksi.

Olkoot $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ ja $C = [c_{ji}] \in \mathbb{R}^{n\times k}$. Tällöin jokaisella $j \in \{1,\dots,m\}$ ja $i \in \{1,\dots,k\}$ pätee

$$(AC)^t_{ji} = (AC)_{ij} = \sum_{\ell=1}^{n} a_{i\ell}c_{\ell j} = \sum_{\ell=1}^{n}(A^t)_{\ell i}(C^t)_{j\ell} = \sum_{\ell=1}^{n}(C^t)_{j\ell}(A^t)_{\ell i} = (C^tA^t)_{ji}.$$

Näin ollen

$$(AC)^t = C^tA^t.$$

□

Tärkein syy kiinnostuksestamme transpoosiin liittyy siis huomioon, että pistetulo voidaan transpoosin avulla kirjoittaa matriisitulona:

$$v\cdot w = v^tw. \tag{4.7}$$

Tähän havaintoon liittyy toinen helposti todistettava ja hieman yleisempi havainto matriisin transpoosin ja pistetulon välisestä yhteydestä.

Lemma 4.4.5. *Olkoot $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^{n\times 1}$ ja $y \in \mathbb{R}^{m\times 1}$. Tällöin*

$$(Ax)\cdot y = x\cdot A^ty$$

Todistus. Yhtälön (4.7) ja lemman 4.4.4 kohdan (4) perusteella

$$(Ax)\cdot y = (Ax)^ty = x^tA^ty = x^t(A^ty) = x\cdot(A^ty).$$

□

Pistetulon ja transpoosin välisellä yhteydellä on suora sovellus, joka kirjataan esimerkiksi.

Esimerkki 4.4.6. *Halutaan etsiä kaikki sellaiset vektorit $v \in \mathbb{R}^{n\times 1}$, jotka ovat kohtisuorassa annettuja vektoreita $v_1,\dots,v_m \in \mathbb{R}^{n\times 1}$ vastaan.*

Koska $v_i\cdot v = 0$ jokaisella $i \in \{1,\dots,m\}$, niin $v_i^tv = 0$ jokaisella $i \in \{1,\dots,m\}$. Olkoon nyt $A \in \mathbb{R}^{m\times n}$ sellainen matriisi, jonka riveinä ovat vektorit v_i^t, eli

$$A = \begin{bmatrix} -- & v_1^t & --- \\ & \vdots & \\ -- & v_m^t & -- \end{bmatrix}.$$

Nyt matriisitulon ominaisuuksien perusteella

$$Av = \begin{bmatrix} v_1^tv \\ \vdots \\ v_m^tv \end{bmatrix} = 0.$$

Näin ollen v on kohtisuorassa vektoreita $v_1,\dots,v_m$ vastaan, jos ja vain jos $Av = 0$, eli $v \in \mathrm{Null}(A)$.

Tehdään sama vielä konkreettisesti.

Esimerkki 4.4.7. *Olkoot*

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{3\times 1} \text{ ja } v_2 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} \in \mathbb{R}^{3\times 1}.$$

Etsitään kaikki vektorit $v \in \mathbb{R}^{3\times 1}$, jotka ovat kohtisuorassa vektoreita v_1 ja v_2 vastaan.

Olkoon

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 2 \end{bmatrix}^t = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 3}.$$

Vektori v on kohtisuorassa vektoreita v_1 ja v_2 vastaan, jos ja vain jos $Av = 0$.

Matriisin A supistettu porrasmuoto on

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

Koska $\mathrm{Null}(A) = \mathrm{Null}(B) = \mathrm{Sp}(e_1 - e_3)$, *niin vektori v on kohtisuorassa vektoreita v_1 ja v_2 vastaan, jos ja vain jos $v \in \mathrm{Sp}(e_1 - e_3)$ eli $v = t(e_1 - e_3)$ joillain $t \in \mathbb{R}$.*

4.5 Ortogonaaliset matriisit

Transpoosin avulla voidaan ortonormaalin kannan vektoreista muodostetulle matriisille antaa luonnollinen tulkinta.

Olkoon $(u_1, \dots, u_n)$ avaruuden $\mathbb{R}^{n\times 1}$ ortonormaalikanta ja olkoon $U = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}$. Koska $u_j \cdot u_i = \delta_{ji}$ kaikilla $j, i \in \{1, \dots, n\}$, niin havaitaan, että

$$u_j^t u_i = \delta_{ji}$$

kaikilla $j, i \in \{1, \dots, n\}$, eli $U^t U = I$.

Tälläisiä matriiseja kutsutaan ortogonaalisiksi. Yleisemmin määritellään seuraavasti.

Määritelmä 4.5.1. *Matriisi $U \in \mathbb{R}^{m\times n}$ on* ortogonaalinen, *jos $U^t U = I \in \mathbb{R}^{n\times n}$.*

Esimerkki 4.5.2. *Gram–Schmidt prosessissa saatu matriisi U on ortogonaalinen.*

Ei ole sattuma, että ortonormaalit kannat antavat esimerkkejä ortogonaalisista matriiseista: *jokainen ortogonaalinen neliömatriisi on kääntyvä ja sen sarakkeet muodostaa ortonormaalin kannan.*

Lause 4.5.3. *Olkoon $U = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \in \mathbb{R}^{n\times n}$ ortogonaalinen matriisi. Tällöin U on kääntyvä ja $(u_1, \dots, u_n)$ on ortonormaali kanta.*

Todistus. Osoitetaan ensin, että U on kääntyvä. Lauseen 3.10.1 perusteella riittää osoittaa, että $\mathrm{Null}(U) = \{0\}$. Olkoon $x \in \mathbb{R}^{n\times 1}$ sellainen, että $Ux = 0$. Tällöin

$$0 = (Ux)^t Ux = x^t U^t Ux = x^t Ix = x^t x = x \cdot x = |x|^2.$$

Näin ollen $x = 0$, eli $\mathrm{Null}(U) = \{0\}$. Matriisi U on siis kääntyvä.

Osoitetaan nyt, että $(u_1, \ldots, u_n)$ on ortonormaalikanta. Koska U on kääntyvä, niin $(u_1, \ldots, u_n)$ on kanta. Lisäksi

$$u_j \cdot u_i = u_j^t u_i = (U^t U)_{ji} = \delta_{ji}$$

kaikilla $j, i \in \{1, \ldots, n\}$. Näin ollen $(u_1, \ldots, u_n)$ on ortonormaali kanta. □

Ortogonaalisen neliömatriisin kääntyvyyden seurauksena saadaan mielenkiintoinen havainto, että myös ortogonaalisen matriisin transpoosi on sen käänteismatriisi.

Korollaari 4.5.4. *Olkoon $U \in \mathbb{R}^{n\times n}$ ortogonaalinen. Tällöin $U^{-1} = U^t$ ja $UU^t = I$. Erityisesti U^{-1} on ortogonaalinen.*

Todistus. Koska $U^t U = I$, niin $U^t = U^{-1}$. Näin ollen $UU^t = UU^{-1} = I$. Lisäksi $(U^{-1})^t U^{-1} = (U^{-1})^t U^t = (UU^{-1})^t = I$. □

Huomautus 4.5.5. *Huomaa, että yleisesti ortogonaalisille $m \times n$-matriisille U ei päde $UU^t = I$, kun $m \neq n$. Esimerkki tälläisestä matriisista on*

$$U = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \in \mathbb{R}^{2\times 1}.$$

Tällöin $U^t U = 1$, eli U on ortogonaalinen, mutta

$$UU^t = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \neq I.$$

4.6 Sovellus: Matriisin QR-hajotelma

Ortonormaaleilla matriiseilla on tärkeä rooli matriisien esittämisessä. Tähän palataan jatkossa toistuvasti, mutta tehdään tässä vaiheessa niistä jo ensimmäinen, eli matriisin QR-hajotelma, joka sanoo, että matriisi voidaan aina ilmaista ortogonaalimatriisin ja yläkolmiomatriisin tulona.

Lause 4.6.1 (QR-hajotelma)**.** *Olkoon $A \in \mathbb{R}^{m\times n}$. Tällöin on olemassa sellainen ortogonaalimatriisi $Q \in \mathbb{R}^{m\times m}$ ja yläkolmiomatriisi $R \in \mathbb{R}^{m\times n}$, että*

$$A = QR.$$

Lausetta varten tarvitaan pieni aputulos, joka sanoo, että yläkolmiomatriisien tulo on yläkolmiomatriisi. Tämä on hyvin luonnollinen tulos, kuten seuraava esimerkki osoittaa.

Esimerkki 4.6.2. *Olkoot*

$$R = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \text{ ja } T = \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

yläkolmiomatriiseja. Tällöin TR on matriisi

$$\begin{array}{cc} & \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 3 & -5 \\ 0 & 4 \\ 0 & 0 \end{bmatrix} \end{array}.$$

Kirjataan tämä yleisessä muodossa lemmaksi, joka jätetään lukijalle harjoitustehtäväksi.

Lemma 4.6.3. *Olkoot $R \in \mathbb{R}^{m \times n}$ ja $T \in \mathbb{R}^{n \times k}$ yläkolmiomatriiseja. Tällöin RT on yläkolmiomatriisi.*

Lauseen 4.6.1 todistus. Olkoon $A \in \mathbb{R}^{m \times n}$ matriisi ja olkoon $B \in \mathbb{R}^{m \times n}$ matriisin A supistettu porrasmuoto. Tällöin on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{m \times m}$, että $A = PB$.

Merkitään $P = \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$. Koska P on kääntyvä, niin lauseen 3.10.1 perusteella sen sarakkeet $v_1, \ldots, v_m$ muodostavat avaruuden $\mathbb{R}^{m \times 1}$ kannan. Gram-Schmidt ortonormeeraus prosessin seurauksena (lause 4.3.17) on olemassa sellainen ortonormaalikanta $(u_1, \ldots, u_m)$ ja sellainen yläkolmiomatriisi $T \in \mathbb{R}^{m \times m}$, että

$$P = QT,$$

missä $Q = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}$.

Koska Q on ortogonaalimatriisi ja matriisi $R = TB$ on yläkolmiomatriisi lemman 4.6.3 perusteella, niin

$$A = PB = QTB = QR$$

on haluttu tulos. □

Luku 5

Determinantti

Deteriminantti on neliömatriisin kertoimista laskettava luku, joka kertoo onko annettu matriisi kääntyvä: *matriisi on kääntyvä, jos ja vain jos sen determinantti ei ole nolla.*

Determinantti voidaan määritellä rekursiivisesti määrittelemällä ensin, että 2×2-matriisin

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

determinantti on

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

Tämän jälkeen $n \times n$-matriisin $A \in \mathbb{R}^{n\times n}$ determinantti $\det A$ voidaan määritellä rekursiivisesti, niin sanottujen kehityskaavojen avulla, matriisin A $(n-1)\times(n-1)$-alimatriisien determinantien avulla. Vaikka tämä lähestymistapa antaa selkeän tavan laskea matriisin determinantti, siitä on vaikea päätellä, miksi yllä annettu kääntyvyystulos on totta. Determinantin kehityskaavoihin ja determinantin laskemiseen palataan myöhemmin tässä luvussa, mutta determinantin määritteleminen aloitetaan toisesta näkökulmasta, joka selittää, miksi determinantti mittaa matriisin kääntyvyyttä.

5.1 Motivointi

Determinantin määritelmän motivoiva peruskysymys on seuraava:

Ongelma. *Voidaanko neliömatriisin $A \in \mathbb{R}^{n\times n}$ kääntyvyys selvittää ilman, että ratkaistaan käänteismatriisi yhtälöryhmästä $\left[A \mid I\right]$ tai että selvitetään muodostavatko matriisin sarakkeet avaruuden $\mathbb{R}^{n\times 1}$ kannan?*[1]

Syy kysyä tälläistä kysymystä on, että yhtälöä $Ax = b$ ratkaistaessa tai sarakevektoreiden ominaisuuksia selvitettäessä joudutaan tekemään valintoja tai päättelyjä. Mikäli tavoite on ainoastaan ainoastaan selvittää matriisin A kääntyvyys olisi mielekkäänpää selvittää kääntyvyys suoralla laskulla. Koska neliömatriisin A kääntyvyys

[1]Luvun 3 perusteella tiedetään, että nämä kysymykset ovat oikeasti yksi ja sama kysymys.

on yhtäpitävää sarakkeiden lineaarisen riippumattomuuden kanssa, edellinen ongelma voidaan siis muotoilla uudestaan seuraavasti:

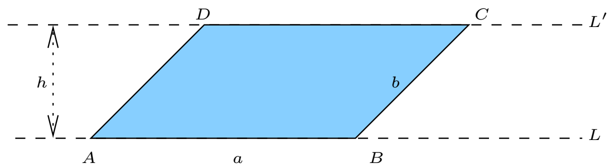
Ongelma. *Onko mahdollista selvittää neliömatriisin $A \in \mathbb{R}^{n \times n}$ sarakkeiden lineaarinen riippumattomuus suoralla laskulla?*

Vastaus tähän kysymykseen on tietysti myönteinen ja, kuten luvussa 6 paljastuu, tällä ominaisuudella on tärkeä (teoreettinen) merkitys selvitettäessä matriisien ominaisuuksia.

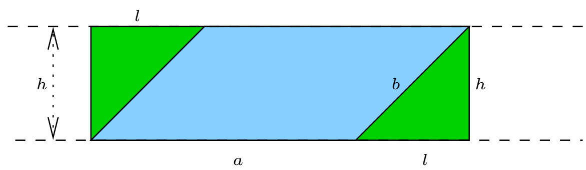
Kysymyksen ratkaisussa tarvittavien käsitteiden motivoimiseksi tarkastellaan aluksi suunnikkaan pinta-alan ja vektoreiden lineaarisen riippumattomuuden välistä yhteyttä. Pinta-alaa käsitellään tässä yhteydessä epäformaalisti lukio-opinnoista tutulla intuitiivisella tavalla. Sen avulla voidaan kuitenkin motivoida tarvittava alternoivan multilineaarikuvauksen käsite.

5.1.1 Suunnikkaan pinta-ala

Kuten tunnettua suorakaide, jonka sivujen pituudet ovat a ja b, pinta-ala on ab. Samoin suorakulmaisen kolmion, jonka kateettien pituudet ovat a ja b pinta-ala on $\frac{ab}{2}$. Tarkastellaan nyt suunnikasta $ABCD$ kuten kuvassa 5.1, jonka sivujen pituudet ovat a ja b. Tämän suorakaiteen pinta-ala on ah, missä h on suorien L ja L' etäisyys. Perustelu tälle faktalle on kuvan 5.2 kuvatekstissä.



Kuva 5.1: Suunnikas, jonka sivujen pituudet ovat a ja b, sekä samansuuntaiset suorat L ja L'.

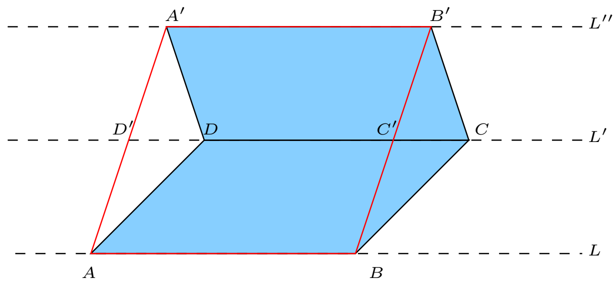


Kuva 5.2: Sinisen suunnikkaan alle on piiretty vihreä suorakulmio, jonka pinta-ala on $(a+l)h$, missä $l = \sqrt{b^2 - h^2}$. Koska vihreiden kolmioiden yhteenlaskettu pinta-ala on $2 \cdot (lh/2) = lh$, niin sinisen suunnikkaan pinta-ala on $(a+l)h - lh = ah$.

Lasku osoittaa, että mikäli suunnikkaan $ABCD$ pinta-ala halutaan ilmoittaa sivunpituuden a avulla, niin tällöin toinen tulontekijä pinta-alassa ei ole toinen sivunpituus b vaan samansuuntaisten suorien L ja L' välisenen etäisyys. Erityisesti havaittaan, että mikäli suorien L ja L' välinen etäisyys on nolla, niin suunnikkaan pinta-ala on nolla.

5.1.2 Kaksi erillistä suunnikasta

Sovelletaan edellistä havaintoa tilanteessa, jossa kahdella erillisellä suunnikkaalla $ABCD$ ja $A'B'CD$ on yhteinen sivu CD, katso kuva 5.3. Tässä tilanteesa suunnikkaiden $ABCD$ ja $A'B'CD$ yhteenlaskettu pinta-ala on suunnikkaan $ABB'A'$ pinta-ala. Tämän faktan voi perustella seuraavasti: Koska janoilla AB, CD ja $C'D'$ on sama pituus, niin suunnikkaan $ABCD$ pinta-ala on sama kuin suunnikkaan $ABC'D'$ ja vastaavasti suunnikkaan $A'B'CD$ pinta-ala on sama kuin suunnikkaan $A'B'C'D'$. Näin ollen suorakulmioiden $ABCD$ ja $A'B'CD$ yhteenlaskettu pinta-ala on suorakulmion $ABB'A'$ pinta-ala.



Kuva 5.3: Suunnikkaan $ABB'A'$ pinta-ala on suunnikkaiden $ABCD$ ja $A'B'CD$ yhteenslaskettu pinta-ala.

On tärkeää huomata, että edellisessä päättelyssä käytettiin ilman erillistä mainintaa tietoa, että suorakulmiot $ABCD$ ja $A'B'CD$ eivät ole päällekkäisiä.

5.1.3 Pinta-alasta vektoreihin

Palataan nyt käyttämään tällä kurssilla käytettyjä merkintöjä. Olkoot $v, w \in \mathbb{R}^2$. Tason $\mathbb{R}^2$ osajoukkoa

$$S(v,w) = \{tv + sw \in \mathbb{R}^2 \colon t, s \in [0,1]\} \subset \mathbb{R}^2$$

kutsutaan *vektoreiden v ja w virittämäksi suunnikkaaksi*; katso kuva 5.4. Huomaa, että suunnikkaan $S(v,w)$ sivun pituudet ovat $|v|$ ja $|w|$.

Jotta voitaisiin hyödyntää edellisen luvun havaintoja, otetaan käyttöön merkintä $u + S(v,w)$, jolla tarkoitetaan suunnikkaan $S(v,w)$ siirtoa vektorilla $u \in \mathbb{R}^2$, eli tarkemmin

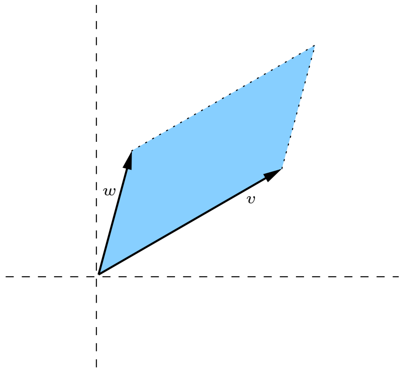
$$u + S(v,w) = \{u + u' \in \mathbb{R}^2 \colon u' \in S(v,w)\} = \{u + tv + sw \in \mathbb{R}^2 \colon t, s \in [0,1]\} \subset \mathbb{R}^2.$$

Huomaa, että geometrisesti myös joukko $u + S(v,w)$ on tason $\mathbb{R}^2$ suunnikas, jonka sivujen pituudet ovat $|v|$ ja $|w|$. Lisäksi on selvää, että suunnikkaan pinta-ala ei muutu siirretäessä, eli että suunnikkailla $u + S(v,w)$ ja $S(v,w)$ on sama pinta-ala.

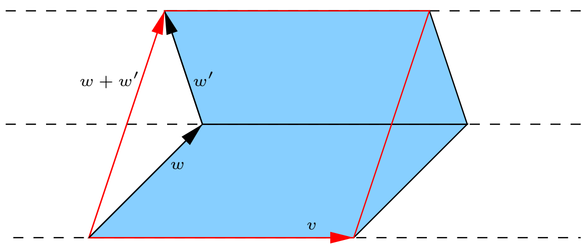
Käyttäen näitä merkintöjä voidaan kuvan 5.3 suunnikkaat $ABCD$ ja $A'B'CD$ ilmaista vektoreilla $v, w, v' \in \mathbb{R}^2$ kuten kuvassa 5.5.

Merkitään nyt $\mathrm{Ala}(v,w)$ suunnikkaan $S(v,w)$ pinta-alaa kaikilla $v, w \in \mathbb{R}^2$. Tällöin kuvassa 5.5 olevien suunnikkaiden pinta-aloille pätee

$$\mathrm{Ala}(v, w + w') = \mathrm{Ala}(v,w) + \mathrm{Ala}(v,w'). \tag{5.1}$$



Kuva 5.4: Vektoreiden v ja w virittämä suunnikas $S(v, w)$.



Kuva 5.5: Sunnikas $ABB'A'$ on ilmaistu vektoreiden v ja w avulla ja suunnikas $ABCD$ vektoreiden v ja w' avulla.

Koska pinta-ala ei riipu vektoreiden järjestyksestä, niin vastaavasti pätee

$$\mathrm{Ala}(w + w', v) = \mathrm{Ala}(w, v) + \mathrm{Ala}(w', v).$$

Intuitiomme pinta-alasta kertoo myös, miten pinta-ala skaalautuu, jos vektoreita skaalataan. Näillä merkinnöillä tämä tarkoittaa, että

$$\mathrm{Ala}(\lambda v, w) = |\lambda|\mathrm{Ala}(v, w) = \mathrm{Ala}(v, \lambda w).$$

missä $\lambda \in \mathbb{R}$. Itseisarvot keskimmäisessä termissä seuraavat siitä, että pinta-ala on aina ei-negatiivinen.

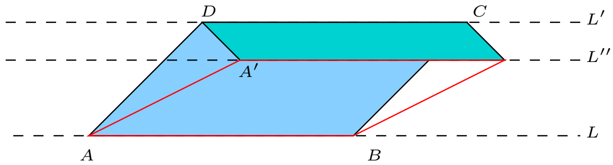
Miksi olemme kiinnostuneita pinta-alasta lineaarialgebran yhteydessä? Tärkein huomio on, että tason vektorit v ja w ovat lineaarisesti riippuvia, jos ja vain jos ne ovat samalla suoralla. Näin ollen vektorit v ja w ovat lineaarisesti riippuvia, jos ja vain jos suunnikkaan pinta-ala $S(v, w)$ on nolla!

Tämä havainto yleistyy kaikkiin avaruuksiin $\mathbb{R}^{n\times 1}$ sellaisessa muodossa, että vektorit $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$ ovat lineaarisesti riippumattomia, jos ja vain jos niiden virittämän suunnikkaan

$$S(v_1, \ldots, s_n) = \{t_1 v_1 + \cdots + t_n v_n \in \mathbb{R}^{n\times 1} \colon t_1, \ldots, t_n \in [0, 1]\}$$

n-ulotteinen tilavuus on nolla. Emme tarvitse jatkossa tilavuuden käsitettä ja siihen liittyvä teoria viedään päätökseen kurssilla *Mitta ja integraali*. Tässä luvussa keskitytään taustalla olevaan lineaarialgebraan.

Palataan vielä hetkiseksi pinta-alan ominaisuuksiin. Tehtyjen havaintojen perusteella pinta-ala vaikuttaisi käyttyvän lähes kuin lineaarinen funktio molempien argumenttiensa suhteen. Palautetaan mieleen, että funktio $f\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$ on lineaarinen, jos $f(v+v') = f(v) + f(v')$ ja $f(\lambda v) = \lambda f(v)$ kaikilla $v, v' \in \mathbb{R}^{n\times 1}$ ja $\lambda \in \mathbb{R}$. Havainnot kuitenkin osoittavat, että näin ei täsmälleen ottaen ole, sillä vektorin skaalauksissa tulee ottaa huomioon skaalauksen itseisarvo ja yhtälö (5.1) ei ole voimassa, jos suunnikkaat ovat päällekkäisiä; tästä esimerkkinä kuvan 5.6 tapaus.



Kuva 5.6: Suunnikkaiden päällekkäisyyden vuoksi suunnikkaan $ABB'A'$ pinta-ala ei ole suunnikkaiden $ABCD$ ja $A'B'CD$ pinta-alojen summa.

Paljastuu kuitenkin, että pinta-ala ei ole primäärisin objekti, jota voimme tutkia, vaan olemassa sellainen funktio $\mathrm{vol}_{\mathbb{R}^2}\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, joka on lineaarinen molempien argumenttiensa suhteen ja jolle pätee $\mathrm{Ala}(v,w) = |\mathrm{vol}_{\mathbb{R}^2}(v,w)|$. Mielenkiintoisesti tämä funktio $\mathrm{vol}_{\mathbb{R}^2}$, jota kutsutaan tason $\mathbb{R}^2$ tilavuusmuodoksi, ei ole symmetrinen vaan antisymmetrinen eli $\mathrm{vol}_{\mathbb{R}^2}(v,w) = -\mathrm{vol}_{\mathbb{R}^2}(w,v)$ kaikilla $v, w \in \mathbb{R}^2$. Tämä antisymmetrisyys on avain vektoreiden lineaariseen riippumattomuuteen ja determinantin määritelmään, kuten kohta huomataan.

5.2 2×2-matriisin determinantti

Tämän luvun tarkoituksena on selittää, miksi 2×2-matriisin

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

determinantille

$$\det A = \det \begin{bmatrix} a & c \\ b & d \end{bmatrix} = ad - bc$$

pätee seuraava tulos:

Lause 5.2.1. *Neliömatriisi $A \in \mathbb{R}^{2\times 2}$ on kääntyvä, jos ja vain jos $\det A \neq 0$.*

Kuten edellisessä luvussa käsiteltiin, tämä lause siis vastaa tulosta, että matriisin $A \in \mathbb{R}^{2\times 2}$ sarakkeet ovat lineaarisesti riippumattomia, jos ja vain jos $\det A \neq 0$.

Lause todistetaan käyttämällä apuna tilavuusmuotojen, eli yleisemmin alternoivien bilineaarikuvausten, ominaisuuksia. Syy tähän on se, että tällä pohjustetaan yleisien $n \times n$-matriisien determinanttien ominaisuuksien osoittamiseen tarvittavia menetelmiä.

Lauseelle 5.2.1 on mahdollista kirjoittaa suora todistus muokkaamalla lauseen 5.2.7 todistusta. Tämä jätetään lukijalle harjoitustehtäväksi.

5.2.1 Tason $\mathbb{R}^{2\times 1}$ alternoivat bilineaarimuodot

Lukijalle

Tässä luvussa käsittely poikkeaa hieman totutusta. Luku alkaa alternoivan bilineaarimuodon määritelmällä, josta on vaikea saada otetta. Tämän jälkeen todistetaan alternoivien bilineaarimuotojen määritelmästä suoraan seuraavia perusominaisuuksia. Vasta tämän jälkeen annetaan ensimmäinen esimerkki. Ei siis kannata pelästyä, vaikka alternoivan bilineaarimuodon määritelmä vaikuttaa abstraktilta. Luvun loppuun mennessä paljastuu, mistä on kysymys.

Määritelmä 5.2.2. *Funktio* $\omega\colon \mathbb{R}^{2\times 1}\times\mathbb{R}^{2\times 1}\to\mathbb{R}$ *on* alternoiva bilineaarimuoto*, jos* ω *on*

- bilineaarinen*, eli jokaisella* $w\in\mathbb{R}^{2\times 1}$ *funktiot* $\mathbb{R}^{2\times 1}\to\mathbb{R}$*,* $v\mapsto\omega(v,w)$*, ja* $\mathbb{R}^{2\times 1}\to\mathbb{R}$*,* $v\mapsto\omega(w,v)$*, ovat lineaarisia, ja*
- alternoiva*, eli* $\omega(v,v)=0$ *kaikilla* $v\in\mathbb{R}^{2\times 1}$*.*

Bilineaarisuudesta ja alternoivuudesta yhdessä seuraa, että alternoiva bilineaarimuoto on antisymmetrinen, eli argumenttien paikan vaihtaminen vaihtaa funktion arvon merkin.

Lemma 5.2.3. *Olkoon* $\omega\colon\mathbb{R}^{2\times 1}\times\mathbb{R}^{2\times 1}\to\mathbb{R}$ *alternoiva bilineaarimuoto. Tällöin kaikilla* $v,w\in\mathbb{R}^{2\times 1}$ *pätee*

$$\omega(v,w)=-\omega(w,v).$$

Todistus. Olkoot $v,w\in\mathbb{R}^{2\times 1}$. Tällöin alternoivuuden perusteella pätee

$$\omega(v+w,v+w)=0.$$

Toisaalta bilineaarisuuden nojalla pätee

$$\omega(v+w,v+w)=\omega(v,v+w)+\omega(w,v+w)=\omega(v,v)+\omega(v,w)+\omega(w,v)+\omega(w,w).$$

Jälleen alternoivuuden perusteella pätee $\omega(v,v)=0$ ja $\omega(w,w)=0$. Näin ollen

$$0=\omega(v+w,v+w)=\omega(v,w)+\omega(w,v),$$

joten $\omega(v,w)=-\omega(w,v)$. □

Alternoivuudesta seuraa myös, että alternoiva bilineaarimuoto $\omega\colon\mathbb{R}^{2\times 1}\times\mathbb{R}^{2\times 1}\to\mathbb{R}$ on täysin määrätty, kun sen arvo $\omega(e_1,e_2)$ standardikannan alkioilla e_1 ja e_2 on kiinnitetty. Kirjataan tämä lemmaksi.

Lemma 5.2.4. *Olkoon* $\omega\colon \mathbb{R}^{2\times 1}\times\mathbb{R}^{2\times 1}\to\mathbb{R}$ *alternoiva bilineaarimuoto. Tällöin kaikilla tason* $\mathbb{R}^{2\times 1}$ *vektoreilla* $v = ae_1 + be_2$ *ja* $w = ce_1 + de_2$ *pätee*

$$\omega(v, w) = \omega(ae_1 + be_2, ce_1 + de_2) = (ad - bc)\omega(e_1, e_2). \tag{5.2}$$

Todistus. Bilineaarisuuden nojalla pätee

$$\begin{aligned}\omega(ae_1 + be_2, ce_1 + de_2) &= \omega(ae_1, ce_1 + de_2) + \omega(be_2, ce_1 + de_2)\\ &= \omega(ae_1, ce_1) + \omega(ae_1, de_2) + \omega(be_2, ce_1) + \omega(be_2, de_2)\\ &= ac\omega(e_1, e_1) + ad\omega(e_1, e_2) + bc\omega(e_2, e_1) + bd\omega(e_2, e_2).\end{aligned}$$

Koska $\omega(e_1, e_1) = \omega(e_2, e_2) = 0$ ja $\omega(e_2, e_1) = -\omega(e_1, e_2)$, niin

$$\omega(ae_1 + be_2, ce_1 + de_2) = ad\omega(e_1, e_2) - bc\omega(e_1, e_2) = (ad - bc)\,\omega(e_1, e_2).$$

□

Tason $\mathbb{R}^{2\times 1}$ tilavuusmuoto $\mathrm{vol}_{\mathbb{R}^{2\times 1}}$

Lemman 5.2.4 kaava (5.2) antaa myös vihjeen kuinka määritellä tason $\mathbb{R}^{2\times 1}$ alternoivia bilineaarimuotoja. Suoralla laskulla havaitaan, että itseasiassa funktio $\mathrm{vol}_{\mathbb{R}^{2\times 1}}\colon \mathbb{R}^{2\times 1}\times \mathbb{R}^{2\times 1}\to\mathbb{R}$, joka on määritelty kaavalla

$$\mathrm{vol}_{\mathbb{R}^{2\times 1}}\left(\begin{bmatrix}a\\b\end{bmatrix}, \begin{bmatrix}c\\d\end{bmatrix}\right) = ad - bc \tag{5.3}$$

kaikilla $\begin{bmatrix}a\\b\end{bmatrix}, \begin{bmatrix}c\\d\end{bmatrix} \in \mathbb{R}^{2\times 1}$, on sellainen alternoiva bilineaarimuoto, että $\mathrm{vol}_{\mathbb{R}^{2\times 1}}(e_1, e_2) = 1$. Tämä jätetään lukijalle harjoitustehtäväksi.

Määritelmä 5.2.5. *Alternoivaa bilineaarimuotoa* $\mathrm{vol}_{\mathbb{R}^2}\colon \mathbb{R}^{2\times 1}\times\mathbb{R}^{2\times 1}\to\mathbb{R}$*, jolle pätee* $\mathrm{vol}_{\mathbb{R}^2}(e_1, e_2) = 1$ *standardikannan alkoilla* (e_1, e_2)*, kutsutaan* tason $\mathbb{R}^2$ tilavuusmuodoksi.

Huomautus 5.2.6. *Lemman 5.2.4 tulos voidaan myös ilmaista muodossa, että kaikki tason alternoivat bilineaarimuodot ovat tilavuusmuodon skaalauksia. Tämä väite jätetään kiinostuneelle lukijalle.*

Lemmasta 5.2.4 seuraa myös, että tilavuusmuoto tunnistaa vektoreiden lineaarisen riippumattomuuden.

Lause 5.2.7. *Tason* $\mathbb{R}^{2\times 1}$ *vektorit* v *ja* w *ovat lineaarisesti riippumattomia, jos ja vain jos* $\mathrm{vol}_{\mathbb{R}^{2\times 1}}(v, w) \neq 0$*.*

Todistus. Oletetaan ensin, että vektorit v ja w ovat lineaarisesti riippuvia toisistaan. Tällöin on olemassa sellaiset vakiot $a, b \in \mathbb{R}$, jotka eivät molemmat ole nollia ja joille pätee $av + bw = 0$. Voidaan olettaa, että $a \neq 0$. Tällöin lineaarisuuden nojalla

$$0 = \mathrm{vol}_{\mathbb{R}^{2\times 1}}(0, w) = \mathrm{vol}_{\mathbb{R}^{2\times 1}}(av{+}bw, w) = a\mathrm{vol}_{\mathbb{R}^{2\times 1}}(v, w){+}b\mathrm{vol}_{\mathbb{R}^{2\times 1}}(w, w) = a\mathrm{vol}_{\mathbb{R}^{2\times 1}}(v, w).$$

Näin ollen $\mathrm{vol}_{\mathbb{R}^2}(v, w) = 0$.

Oletetaan nyt, että $\mathrm{vol}_{\mathbb{R}^2}(v, w) = 0$ ja osoitetaan, että vektorit v ja w ovat lineaarisesti riippuvia toisistaan. Voidaan olettaa, että $v \neq 0$, koska tilanteessa $v = w = 0$ väite on tosi. Olkoot $v = ae_1 + be_2$ ja $w = ce_1 + de_2$. Tällöin

$$0 = \mathrm{vol}_{\mathbb{R}^{2\times 1}}(v, w) = (ad - bc)\mathrm{vol}_{\mathbb{R}^{2\times 1}}(e_1, e_2) = ad - bc$$

eli $ad = bc$. Koska $v \neq 0$, niin joko $a \neq 0$ tai $b \neq 0$. Voidaan olettaa, että $a \neq 0$. Näin ollen $d = bc/a$ ja

$$w = ce_1 + de_2 = ce_1 + (bc/a)e_2 = \frac{c}{a}(ae_1 + be_2) = \frac{c}{a}v.$$

Vektorit v ja w ovat siis lineaarisesti riippuvia toisistaan. □

5.2.2 2×2-matriisin determinantti ja lineaarinen riippumattomuus

Todistetaan nyt lause 5.2.1. Havaitaan ensin, että tilavuusmuodon $\mathrm{vol}_{\mathbb{R}^{2\times 1}}$ määritelmän nojalla matriisin

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \in \mathbb{R}^2$$

determinantille $\det A = ad - bc$ pätee

$$\det A = \mathrm{vol}_{\mathbb{R}^2}(Ae_1, Ae_2), \tag{5.4}$$

sillä $Ae_1 = ae_1 + be_2$ ja $Ae_2 = ce_1 + de_2$, joten

$$\mathrm{vol}_{\mathbb{R}^{2\times 1}}(Ae_1, Ae_2) = \mathrm{vol}_{\mathbb{R}^{2\times 1}}(ae_1 + be_2, ce_1 + de_2) = ad - bc = \det A.$$

Lauseen 5.2.1 todistus. Kaavan 5.4 perusteella

$$\det A = \mathrm{vol}_{\mathbb{R}^2}(Ae_1, Ae_2).$$

Lauseen 5.2.7 perustella vektorit Ae_1 ja Ae_2 ovat lineaarisesti riippumattomia, jos ja vain jos $\mathrm{vol}_{\mathbb{R}^{2\times 1}}(Ae_1, Ae_2) \neq 0$. Näin ollen matriisin A sarakkeet ovat lineaarisesti riippumattomia, jos ja vain jos $\det A \neq 0$. Näin ollen matriisi A on kääntyvä, jos ja vain jos $\det A \neq 0$. Väite on näin todistettu. □

5.3 Extra: Johdatus $n \times n$-neliömatriisin determinanttiin

Lukijalle

Tämän luvun voi alilukua 5.3.2 lukuunottamatta sivuuttaa ensimmäisellä (ja toisellakin lukukerralla) ja siirtyä suoraan lukuun 5.4. Tässä luvussa perustellaan, miksi yleisen $n \times n$-matriisin determinantti määritellään kaavalla 5.6. Lukija voi kuitenkin ottaa kaavan myös annettuna. Determinantin ominaisuudet perustellaan käyttäen tässä luvussa esiteltyjä tuloksia. Näidenkin tulosten todistukset voi sivuuttaa ensimmäisellä lukukerralla.

5.3.1 Alternoivat multilineaarikuvaukset ja lineaarinen riippumattomuus

Määritelmä 5.3.1. *Funktio* $\omega\colon \mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1}\to\mathbb{R}$, $(v_1,\dots,v_k)\mapsto\omega(v_1,\dots,v_n)$, *on* multilineaarikuvaus, *jos jokaisella* $i\in\{1,\dots,n\}$ *ja kaikilla vektoreilla* $v,w,v_1,\dots,v_n\in\mathbb{R}^{n\times 1}$ *ja vakioilla* $a\in\mathbb{R}$ *pätee sekä*

$$\begin{aligned}\omega(v_1,\dots,v_{i-1},v+w,v_{i+1},\dots,v_n) &= \omega(v_1,\dots,v_{i-1},v,v_{i+1},\dots,v_n)\\ &\quad+\omega(v_1,\dots,v_{i-1},w,v_{i+1},\dots,v_n)\end{aligned}$$

että

$$\omega(v_1,\dots,v_{i-1},av_i,v_{i+1},\dots,v_n)=a\omega(v_1,\dots,v_{i-1},v_i,v_{i+1},\dots,v_n).$$

Muutama huomautus on paikallaan.

Huomautus 5.3.2. *Kuvausta* $(v_1,\dots,v_n)\mapsto\omega(v_1,\dots,v_n)$ *sanotaan multilineaariseksi, jos se on siis lineaarinen erikseen jokaisen muuttujan* v_i *suhteen. Huomaa, että edellisessä määritelmässä ehdot voidaan kirjoittaa lyhyesti muodossa, että jokaisella* $i\in\{1,\dots,n\}$ *kuvaus* $v\mapsto\omega(v_1,\dots,v_{i-1},v,v_{i+1},\dots,v_n)$ *on lineaarikuvaus. Tapauksessa* $n=2$ *tämä on täsmälleen bilineaarisuuden määritelmä. Multilineaarikuvaukset ovat siis bilineaarikuvausten yleistys.*

Huomautus 5.3.3. *Multilineaarikuvauksen määritelmässä funktion* ω *argumenttina on* n *kappaletta avaruuden* $\mathbb{R}^{n\times 1}$ *vektoreita. Määritelmän tasolla näin ei tarvitsisi olla, vaan voitaisiin tarkastella yleisempiä multilineaarikuvauksia, joiden argumenttina on jonkin muu määrä avaruuden* $\mathbb{R}^{n\times 1}$ *vektoreita. Tämä johtaa tensorin käsitteeseen. Emme kuitenkaan tarvitse tätä yleistystä. Aiheesta kiinnostuneet lukijat ohjaataan juttelemaan luennotsijan kanssa tai tutustumaan multilineaarialgebraan.*

Määritelmä 5.3.4. *Multilineaarikuvaus* $\omega\colon\mathbb{R}^n\times\cdots\times\mathbb{R}^n\to\mathbb{R}$ *on* alternoiva, *jos kaikilla indekseillä* $1\le i<j\le n$ *ja vektoreilla* $v,v_1,\dots,v_n\in\mathbb{R}^n$ *pätee*

$$\omega(v_1,\dots,v_{i-1},v,v_{i+1},\dots,v_{j-1},v,v_{j+1},\dots,v_n)=0.$$

Huomautus 5.3.5. *Yllä oleva määritelmä voidaan tulkita sanomalla, että alternoiva multilineaarikuvaus saa arvon nolla, jos jokin vektori esiintyy kahdesti argumenttina olevassa jonossa* $(v_1,\dots,v_n)$*. Jos* $n=2$*, niin tämä ehto on täsmälleen edellisen luvun alternoivuuden määritelmä.*

Kuten bilineaarikuvausten kohdalla, alternoivuudesta seuraa jälleen kahden vektorin paikan vaihtaminen argumentissa muuttaa funktion arvon merkin. Kirjataan tämä ylös lemmaksi. Koska todistus on sama kuin bilineaarisessa tilanteessa, jätetään se harjoitustehtäväksi.

Lemma 5.3.6. *Olkoon* $\omega\colon\mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1}\to\mathbb{R}$ *alternoiva multilineaarikuvaus. Tällöin kaikilla* $v_1,\dots,v_n\in\mathbb{R}^{n\times 1}$ *ja* $1\le i<j\le n$ *pätee*

$$\omega(v_1,\dots,v_{i-1},v_j,v_{i+1},\dots,v_{j-i},v_i,v_{j+1},\dots v_n)=-\omega(v_1,\dots,v_n).$$

Perustavanlaatuinen huomio multilinaarikuvauksista on, että multilineaarisuudesta ja alternoivuudesta yhdessä seuraa, että alternoivat multilineaarikuvaukset tunnistavat vektoreiden lineaarisen riippuvuuden. Kirjataan tämä tärkeä havainto lauseeksi.

Lause 5.3.7. *Olkoon* $\omega\colon \mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1}\to\mathbb{R}$ *alternoiva multilineaarikuvaus. Jos vektorit* $v_1,\ldots,v_n\in\mathbb{R}^n$ *ovat lineaarisesti riippuvia, niin*

$$\omega(v_1,\ldots,v_n)=0.$$

Todistus. Koska vektorit $v_1,\ldots,v_n$ ovat lineaarisesti riippuvia, niin on olemassa sellaiset vakiot $a_1,\ldots,a_n\in\mathbb{R}$, jotka eivät kaikki ole nollia, että

$$a_1v_1+\cdots+a_nv_n=0.$$

Olkoon nyt $i\in\{1,\ldots,n\}$ sellainen indeksi, että $a_i\neq 0$. Tällöin

$$v_i=-\frac{1}{a_i}\left(a_1v_1+\cdots+a_{i-1}v_{i-1}+a_{i+1}v_{i+1}+\cdots+a_nv_n\right)=-\frac{1}{a_i}\sum_{\substack{j=1\\ j\neq i}}^{n}a_jv_j.$$

Koska ω on multilineaarinen, niin

$$\begin{aligned}\omega(v_1,\ldots,v_n)&=\omega(v_1,\ldots,v_{i-1},v_i,v_{i+1},\ldots,v_n)\\&=\omega\left(v_1,\ldots,v_{i-1},-\frac{1}{a_i}\sum_{\substack{j=1\\ j\neq i}}^{n}a_jv_j,v_{i+1},\ldots,v_n\right)\\&=-\frac{1}{a_i}\sum_{\substack{j=1\\ j\neq i}}^{n}a_j\omega(v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_n).\end{aligned}$$

Koska summassa indeksi j kuuluu joukkoon $\{1,\ldots,i-1,i+1,\ldots,n\}$, niin huomataan, että jonossa $(v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_n)$ vektori v_j toistuu aina kahdesti. Näin ollen alternoivuuden nojalla $\omega(v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_n)=0$ jokaisella summan indeksillä j. Näin ollen

$$\omega(v_1,\ldots,v_n)=-\frac{1}{a_i}\sum_{\substack{j=1\\ j\neq i}}^{n}a_j\omega(v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_n)=0.$$

$\square$

Huomautus 5.3.8. *On syvällinen tulos, että lauseen 5.3.7 tulos pätee myös toiseen suuntaan, eli että nollakuvauksesta poikkeavalle alternoivalle multilineaarikuvaukselle* $\omega\colon \mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1}\to\mathbb{R}$ *pätee* $\omega(v_1,\ldots,v_n)=0$*, jos ja vain jos vektorit* $v_1,\ldots,v_n$ *ovat toisistaan lineaarisesti riippuvia. Tämä tulos on toinen muotoilu tavoiteltavalle tulokselle, että determinantti kertoo matriisin kääntyvyyden. Tämä tulos riittää todistaa deteriminanttien tapaukssa ja tehdään luvussa 5.8.*

5.3.2 Permutaatiot

Permutaatio on toinen nimi bijektiolle joukolta itselleen.

Määritelmä 5.3.9. *Olkoon* $n \in \mathbb{N}$. *Joukon* $\{1, \dots, n\}$ *bijektiota* $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ *kutsutaan* permutaatioksi. *Joukon* $\{1, \dots, n\}$ *kaikkien permutaatioiden joukkoa merkitään*

$$\mathrm{Sym}(n) = \{\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\} \mid \sigma \text{ on permutaatio}\}.$$

Huomautus 5.3.10. *Joukolle* $\mathrm{Sym}(n)$ *on kirjallisuudessa useita merkintöjä, esimerkiksi* S_n *ja* $\mathrm{Perm}(n)$. *Merkintä* $\mathrm{Sym}(n)$ *on valittu siksi, että se permutaatioita kutsutaan myös* symmetrioiksi *ja ne muodostavat niin sanotun ryhmän, jossa laskutoimituksena on kuvausten yhdistäminen. Merkintä* $\mathrm{Sym}(n)$ *kuvastaa tätä yhteyttä. Ryhmiä käsitellään kursseilla* Algebralliset rakenteet I & II.

Esimerkki 5.3.11. *Joukko* $\mathrm{Sym}(2)$ *koostuu täsmälleen kahdesta permutaatiosta. Tämä voidaan päätellä seuraavasti. Olkoon* $\tau\colon \{1, 2\} \to \{1, 2\}$ *permutaatio. Jos* $\tau(1) = 1$, *niin* $\tau(2) = 2$ *ja* τ *on identtinen kuvaus* id. *Jos* $\tau(1) \neq 1$, *niin* $\tau(1) = 2$ *ja* $\tau(2) = 1$, *koska* τ *on bijektio. Näin ollen joukossa* $\mathrm{Sym}(2)$ *on täsmälleen kaksi permutaatiota: identtinen permaatio* id *ja permutaatio* τ, *joka vaihtaa alkioden* 1 *ja* 2 *paikan, eli jolla pätee* $\tau(1) = 2$. *Näille permutaatioille pätee* $\mathrm{sign}(\mathrm{id}) = 1$ *ja* $\mathrm{sign}(\tau) = -1$.

Joukko $\mathrm{Sym}(3)$ *puolestaan koostuu kuudesta permutaatiosta ja yleisemmin* $\mathrm{Sym}(n)$ *koostuu* $n!$ *permutaatiosta. Nämä lukumäärät eivät kuitenkaan ole tärkeitä käsittelyn kannalta.*

Permutaatioon $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ liittyvä peruskäsite on permutaation merkki $\mathrm{sign}(\sigma)$, jota jatkossa tullaan käyttämään laskemaan alternoivan multilineaarikuvauksen arvoa tilanteessa, jossa permutaatio muuttaa argumenttien järjestystä.

Permutaation merkin määritelmän taustalla on havainto, että jokaiselle permutaatiolle $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ voidaan laskea kuinka monta kertaa σ vaihtaa luvut $i < j$ päinvastaiseen suuruusjärjestykseen, eli kuinka monella parilla (i, j), missä $1 \leq i < j \leq n$, pätee $\sigma(i) > \sigma(j)$. Paljastuu, että tämä lukumäärä ei sinällään ole tärkeä, mutta on hyödyllistä tietää onko se parillinen vai pariton. Koska symboleita $+$ ja $-$ on vaikea käyttää kaavoissa, kirjoitetaan nämä määritelmät tarkemmin seuraavasti.

Olkoon $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ permutaatio. Merkitään

$$N(\sigma) = \#\{(i, j)\colon 1 \leq i < j \leq n,\ \sigma(i) > \sigma(j)\} \in \mathbb{N},$$

missä symboli risuaita $\#$ on merkintä joukon alkioiden lukumäärälle.

Määritelmä 5.3.12. *Permutaation* $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ merkki $\mathrm{sign}(\sigma)$ *on luku*

$$\mathrm{sign}(\sigma) = (-1)^{N(\sigma)} \in \{1, -1\}.$$

Esimerkki 5.3.13. *Olkoon* $\sigma\colon \{1, 2, 3\} \to \{1, 2, 3\}$ *permutaatio, joka on määritelty kaavoilla* $\sigma(1) = 2$, $\sigma(2) = 3$ *ja* $\sigma(3) = 1$. *Tässä tapauksessa lukupareja* (i, j), *jossa* $1 \leq i < j \leq 3$, *on täsmälleen kolme kappaletta, eli lukuparit* $(1, 2)$, $(1, 3)$ *ja* $(2, 3)$.

Koska $(\sigma(1),\sigma(2)) = (2,3)$, $(\sigma(1),\sigma(3)) = (2,1)$ *ja* $(\sigma(2),\sigma(3)) = (3,1)$, *niin saadaan, että*

$$N(\sigma) = \#\{(i,j)\colon 1 \le i < j \le 3,\ \sigma(i) > \sigma(j)\} = \#\{(1,3),(2,3)\} = 2.$$

Näin ollen

$$\mathrm{sign}(\sigma) = (-1)^{N(\sigma)} = (-1)^2 = 1.$$

Esimerkki 5.3.14. *Tarkastellaan permutaatiota* $\sigma\colon \{1,2,3,4\} \to \{1,2,3,4\}$, *joka on määritelty kaavoilla* $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 2$ *ja* $\sigma(4) = 4$, *eli* σ *vaihtaa lukujen* 2 *ja* 3 *paikkaa. Tässä tapauksessa ainoa lukupari* (i,j), *jossa* $1 \le i < j \le 4$ *ja jolle pätee* $\sigma(i) > \sigma(j)$, *on pari* $(2,3)$. *Näin ollen* $N(\sigma) = 1$ *ja* $\mathrm{sign}(\sigma) = (-1)^{N(\sigma)} = -1$.

Esimerkki 5.3.15. *Identtisen permutaation* $\mathrm{id}\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$, *eli permutaation, jolle pätee* $\mathrm{id}(i) = i$ *kaikilla* $i \in \{1,\ldots,n\}$, *merkki on aina* 1, *sillä* $N(\mathrm{id}) = 0$.

Permutaation merkki voidaan myös määritellä toisella tavalla, että permutaation merkki on positiivinen, jos annettu permutaatio voidaan kirjoittaa yhdisteenä parillisesta määrästä permutaatioita, jotka vaihtavat täsmälleen kahden alkion paikkaa. Tälläisiä permutaatioita kutsutaan transpositioiksi ja esimerkki tälläisestä permutaatioista on annettu esimerkissä 5.3.14. Nämä tulokset on käsitelty liitteessä C.1.

5.3.3 Yleisen alternoivan multilineaarikuvauksen kaava

Permutaation käsite antaa meille oikean työkalun kirjoittaa yleinen kaava alternoivan multilineaarikuvauksen ω arvolle $\omega(v_1,\ldots,v_n)$ vektoreiden $v_j = v_{j1}e_1 + \cdots + v_{jn}e_n$ kertoimien avulla.

Lause 5.3.16. *Olkoon* $\omega\colon \mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1} \to \mathbb{R}$ *alternoiva multilineaarikuvaus. Tällöin kaikilla* $v_1,\ldots,v_n \in \mathbb{R}^n$, *missä* $v_i = v_{1i}e_1 + \cdots + v_{ni}e_n$, *pätee*

$$\omega(v_1,\ldots,v_n) = \left(\sum_{\sigma\in\mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(n)n}\right)\omega(e_1,\ldots,e_n). \tag{5.5}$$

Lauseen 5.3.16 todistuksen perustana on havainto, että alternoivan multilineaarikuvaksen syötteenä olevia vektoreiden järjestystä muutettaessa, eli *permutoitaessa*, alternoivan multilineaarikuvauksen arvo muuttuu tämän permutaation merkillä. Koska tätäkin tulosta tarvitaan myöhemmin, kirjataan tulos lauseeksi.

Lause 5.3.17. *Olkoon* $\omega\colon \mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1} \to \mathbb{R}$ *alternoiva multilineaarikuvaus, olkoot* $v_1,\ldots,v_n \in \mathbb{R}^n$ *vektoreita ja* $\sigma\colon \{1,\ldots,n\} \to \{1,\ldots n\}$ *permutaatio. Tällöin*

$$\omega(v_{\sigma(1)},\ldots,v_{\sigma(n)}) = \mathrm{sign}(\sigma)\omega(v_1,\ldots,v_n).$$

Nämä kaksi tulosta (eli lauseet 5.3.16 ja 5.3.17) ovat neliömatriisin determinantin yleisen määritelmän ja ominaisuuksien perusta. Lauseiden todistukset eivät kuitenkaan ole olennaisia determinantin määritelmän käyttämisen kannalta, joten todistukset on siirretty lukuun C.2.

Ennen determinantin määritelmään siirtymistä, tehdään kaksi huomiota. Ensimmäinen huomio on, että kaava (5.5) on kaavan (5.2) yleistys. Tämä havainto tehdään tarkasti seuraavassa esimerkissä.

Esimerkki 5.3.18. *Olkoon* $n = 2$ *ja* $\omega\colon \mathbb{R}^{2\times 1} \times \mathbb{R}^{2\times 1} \to \mathbb{R}$ *alternoiva bilineaarikuvaus. Olkoot myos* $v_1 = v_{11}e_1 + v_{12}e_2 = ae_1 + be_2 \in \mathbb{R}^2$ *ja* $v_2 = v_{21}e_1 + v_{22}e_2 = ce_1 + de_2 \in \mathbb{R}^2$.

Näin ollen tapauksessa $n = 2$ *saadaan, että*

$$\begin{aligned}\left(\sum_{\sigma\in\mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(n)n}\right) &= \mathrm{sign}(\mathrm{id}) v_{\mathrm{id}(1)1} v_{\mathrm{id}(2)2} + \mathrm{sign}(\tau) v_{\tau(1)1} v_{\tau(2)2} \\ &= v_{11}v_{22} - v_{21}v_{12} = ad - bc.\end{aligned}$$

Toinen havainto lauseesta 5.3.16 on, että avaruuden $\mathbb{R}^{n\times 1}$ yhteydessä voidaan määritellä tilavuusmuoto $\mathrm{vol}_{\mathbb{R}^n}$ kuten tason $\mathbb{R}^2$ tapauksessa.

Määritelmä 5.3.19. *Funktiota* $\mathrm{vol}_{\mathbb{R}^n}\colon \mathbb{R}^{n\times 1} \times \cdots \times \mathbb{R}^{n\times 1} \to \mathbb{R}$, *joka on määritelty kaavalla*

$$\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \ldots, v_n) = \sum_{\sigma\in\mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(n)n},$$

missä $v_i = v_{1i}e_1 + \cdots + v_{ni}e_n \in \mathbb{R}^{n\times 1}$ *kaikilla* $i \in \{1, \ldots, n\}$, *kutsutaan avaruuden* $\mathbb{R}^n$ tilavuusmuodoksi.

Tilavuusmuoto $\mathrm{vol}_{\mathbb{R}^n}$ on alternoiva multilineaarikuvaus.

Lause 5.3.20. *Funktio* $\mathrm{vol}_{\mathbb{R}^n}\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ *on alternoiva multilineaarikuvaus, jolle pätee* $\mathrm{vol}_{\mathbb{R}^n}(e_1, \ldots, e_n) = 1$.

Tämänkin lauseen todistuksessa hyödynnetään aputuloksia, joten todistus käsitellään litteessä C.1.

Huomautus 5.3.21. *Tapauksessa* $n \geq 3$ *tilavuusmuodon* $\mathrm{vol}_{\mathbb{R}^n}$ *kaavan soveltaminen konkreettisiin vektoreihin tuottaa pitkiä laskuja, koska joukossa* $\mathrm{Sym}(n)$ *on* $n!$ *alkiota. Palautetaan mieleen, että kertoma* $n!$ *on määritelty kaavalla* $n! = 1 \cdot 2 \cdots n$, *eli* $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, *jne. Näitä laskuja ei ole (koskaan) mielekästä laskea käsin. Konkreettiseen laskemiseen palataan luvussa 5.9.*

5.4 Determinantin määritelmä

Nyt olemme valmiita antamaan neliömatriisin determinantin määritelmän yleisessä tapauksessa. Määritelmää varten tarvitaan luvussa 5.3.2 määritelty permutaation käsite.

Määritelmä 5.4.1. *Neliömatriisin* $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ determinantti *on luku*

$$\det A = \sum_{\sigma\in\mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}. \tag{5.6}$$

Aloitetaan avaamalla kaavan sisältöä.

Huomautus 5.4.2. *Olkoon* $n = 3$ *ja* $A = [a_{ji}] \in \mathbb{R}^{3\times 3}$*. Tarkastellaan aluksi yhtä permutaatiota* $\sigma \in \mathrm{Sym}(3)$*. Tulo*

$$a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}$$

muodostuu niin, että permutaatio σ *valitsee jokaisesta sarakkeesta yhden alkion. Jos esimerkiksi* $\sigma(1) = 3$*,* $\sigma(2) = 1$ *ja* $\sigma(3) = 2$*, niin tulo on* $a_{31}a_{12}a_{23}$ *eli tällöin on valittu alkiot*

$$\begin{bmatrix} a_{11} & \mathbf{a_{12}} & a_{13} \\ a_{21} & a_{22} & \mathbf{a_{23}} \\ \mathbf{a_{31}} & a_{32} & a_{33} \end{bmatrix}$$

Determinantin kaavassa, kun summataan yli permutaatioiden, käydään siis läpi kaikki mahdolliset tavat valita jokaisesta sarakkeesta alkio sellaisella tavalla, että samalta riviltä ei koskaan valita kahta alkiota. Tässä tilanteessa kaavan (5.6) *summa käy siis läpi seuraavat tapaukset:*

$$\begin{bmatrix} \mathbf{a_{11}} & a_{12} & a_{13} \\ a_{21} & \mathbf{a_{22}} & a_{23} \\ a_{31} & a_{32} & \mathbf{a_{33}} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{a_{11}} & a_{12} & a_{13} \\ a_{21} & a_{22} & \mathbf{a_{23}} \\ a_{31} & \mathbf{a_{32}} & a_{33} \end{bmatrix},$$
$$\begin{bmatrix} a_{11} & \mathbf{a_{12}} & a_{13} \\ \mathbf{a_{21}} & a_{22} & a_{23} \\ a_{31} & a_{32} & \mathbf{a_{33}} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & \mathbf{a_{13}} \\ \mathbf{a_{21}} & a_{22} & a_{23} \\ a_{31} & \mathbf{a_{32}} & a_{33} \end{bmatrix},$$
$$\begin{bmatrix} a_{11} & \mathbf{a_{12}} & a_{13} \\ a_{21} & a_{22} & \mathbf{a_{23}} \\ \mathbf{a_{31}} & a_{32} & a_{33} \end{bmatrix} \quad \textit{ja} \quad \begin{bmatrix} a_{11} & a_{12} & \mathbf{a_{13}} \\ a_{21} & \mathbf{a_{22}} & a_{23} \\ \mathbf{a_{31}} & a_{32} & a_{33} \end{bmatrix}.$$

Lopuksi determinantin kaava ottaa huomioon käytetyn permutaation σ *merkin ennen tulojen laskemista yhteen.*

Lasketaan nyt identiteettimatriisin I determinantti suoraan määritelmästä.

Esimerkki 5.4.3. *Olkoon* $I = [e_{ji}] \in \mathbb{R}^{n\times n}$ *identiteettimatriisi. Tällöin* $e_{ji} = 1$*, jos* $j = i$*, ja* $e_{ji} = 0$ *muulloin.*

Olkoon nyt $\sigma \in \mathrm{Sym}(n)$*. Tällöin*

$$e_{\sigma(1)1}\cdots e_{\sigma(n)n} = \begin{cases} 1, & \sigma(i) = i \textit{ jokaisella } i \in \{1,\ldots,n\} \\ 0, & \textit{muutoin} \end{cases}$$

Koska ensimmäinen ehto tarkoittaa, että $\sigma = \mathrm{id}$*, niin saadaan, että*

$$\det I = \sum_{\sigma\in\mathrm{Sym}(n)} \mathrm{sign}(\sigma)e_{\sigma(1)1}e_{\sigma(2)2}\cdots e_{\sigma(n)n} = \mathrm{sign}(\mathrm{id})e_{11}e_{22}\cdots e_{nn} = 1.$$

Yllä oleva huomio ja identiteettimatriisin determinantin laskeminen osoittavat, että determinantin kaavan merkitystä on vaikea hahmottaa matriisin kertoimien tasolla. Luonnollisempi tulkinta määritelmälle saadaankin tilavuusmuodon avulla.[2]

Lemma 5.4.4. *Olkoon $A \in \mathbb{R}^{n\times n}$ neliömatriisi. Tällöin*

$$\det A = \mathrm{vol}_{\mathbb{R}^{n\times 1}}(Ae_1, Ae_2, \ldots, Ae_n),$$

missä $\mathrm{vol}_{\mathbb{R}^{n\times 1}}$ on avaruuden $\mathbb{R}^{n\times 1}$ tilavuusmuoto.

Todistus. Olkoon $A = [a_{ji}] \in \mathbb{R}^{n\times n}$. Merkitään jokaisella $i \in \{1, \ldots, n\}$

$$v_i = Ae_i = a_{1i}e_1 + \cdots + a_{ni}e_n = \sum_{j=1}^{n} a_{ji}e_j.$$

Koska $\mathrm{vol}_{\mathbb{R}^{n\times 1}}$ on alternoiva multilineaarikuvaus lauseen 5.3.20 nojalla, niin lauseen 5.3.16 perusteella pätee

$$\begin{aligned}
\mathrm{vol}_{\mathbb{R}^{n\times 1}}(Ae_1, \ldots, Ae_n) &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \ldots, v_n) \\
&= \left(\sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \right) \mathrm{vol}_{\mathbb{R}^n}(e_1, \ldots, e_n) \\
&= \left(\sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \right) = \det A.
\end{aligned}$$

Toiseksi viimeisessä askeleessa käytettiin tietoa, että $\mathrm{vol}_{\mathbb{R}^{n\times 1}}(e_1, \ldots, e_n) = 1$. □

5.5 Determinantin perusominaisuuksia

Todistetaan nyt lemman 5.4.4 ja lauseen 5.3.20 avulla kaksi determinantin perusominaisuuksista eli sarakelineaarisuus ja alternoivuus. Sarakelineaarisuuden voi muotoilla monin tavoin. Seuraavassa muotoilussa yksi matriisin A sarakeista on korvattu kahden muun vektorin lineaarikombinaatiolla. Tätä tulosta hyödynnetään usein muodossa, jossa matriisin sarakkeeseen lisätään toisen sarakkeen monikerta.

Lause 5.5.1. *Olkoon $A \in \mathbb{R}^{n\times n}$ neliömatriisi, jonka sarakkeet ovat $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$ sekä olkoot $v, w \in \mathbb{R}^{n\times 1}$ ja $a, b \in \mathbb{R}$. Tällöin jokaisella $i \in \{1, \ldots, n\}$ pätee*

$$\begin{aligned}
&\det \begin{bmatrix} v_1 & \cdots & v_{i-1} & (av+bw) & v_{i+1} & \cdots & v_n \end{bmatrix} \\
&\quad = a \det \begin{bmatrix} v_1 & \cdots & v_{i-1} & v & v_{i+1} & \cdots & v_n \end{bmatrix} \\
&\qquad + b \det \begin{bmatrix} v_1 & \cdots & v_{i-1} & w & v_{i+1} & \cdots & v_n \end{bmatrix}.
\end{aligned}$$

[2] Itseasiassa seuraavan lemman tulos voitaisiin ottaa determinantin määritelmäksi.

Todistus. Olkoon $B \in \mathbb{R}^{n\times n}$ matriisi $B = \begin{bmatrix} v_1 & \cdots & v_{i-1} & (av+bw) & v_{i+1} & \cdots & v_n \end{bmatrix}$. Tällöin lemman 5.4.4 nojalla pätee

$$\begin{aligned}\det B &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(Be_1, Be_2, \ldots, Be_n)\\ &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \ldots, v_{i-1}, av+bw, v_{i+1}, \ldots, v_n).\end{aligned}$$

Tilavuusmuodon $\mathrm{vol}_{\mathbb{R}^{n\times 1}}$ multilineaarisuuden (lause 5.3.20) nojalla pätee

$$\begin{aligned}\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \ldots, v_{i-1}, av+bw, v_{i+1}, \ldots, v_n) &= a\ \mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)\\ &\quad + b\ \mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n).\end{aligned}$$

Väite seuraa nyt käyttämällä jälleen lemmaa 5.4.4. □

Huomautus 5.5.2. *Determinantin sarakelineaarisuuden voi todistaa myös suoraan determinantin kaavasta* (5.6).

Toinen determinantin perusominaisuuksista on, että matriisin kahden sarakkeiden paikan vaihtaminen vaihtaa determinantin merkin. Tämä seuraa suoraan tilavuusmuodon $\mathrm{vol}_{\mathbb{R}^{n\times 1}}$ alternoivuudesta.

Lause 5.5.3. *Olkoon $A \in \mathbb{R}^{n\times n}$ matriisi, jonka sarakkeet ovat $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$. Tällöin kaikilla $1 \le i < j \le n$ pätee*

$$\det \begin{bmatrix} v_1 & \cdots & v_{i-1} & v_j & v_{i+1} & \cdots v_{j-1} & v_i & v_{j+1} & \cdots & v_n \end{bmatrix} = -\det A.$$

Todistus. Lemman 5.4.4 ja lemman 5.3.6 nojalla

$$\begin{aligned}&\det \begin{bmatrix} v_1 & \cdots & v_{i-1} & v_j & v_{i+1} & \cdots v_{j-1} & v_i & v_{j+1} & \cdots & v_n \end{bmatrix}\\ &= \mathrm{vol}_{\mathbb{R}^n}(Ae_1, \ldots, Ae_{i-1}, Ae_j, Ae_{i+1}, \ldots, Ae_{j-1}, Ae_i, Ae_{j+1}, \ldots, Ae_n)\\ &= -\mathrm{vol}_{\mathbb{R}^n}(Ae_1, \ldots, Ae_{i-1}, Ae_i, Ae_{i+1}, \ldots, Ae_{j-1}, Ae_j, Ae_{j+1}, \ldots, Ae_n) = -\det A.\end{aligned}$$

□

5.6 Tulomatriisin determinantti on determinanttien tulo

Yksi determinantin tärkeitä ominaisuuksia on, että kahden neliömatriisin $A \in \mathbb{R}^{n\times n}$ ja $B \in \mathbb{R}^{n\times n}$ tulon determinantti on determinanttien $\det A$ ja $\det B$ tulo. Kirjataan tämä tulos lauseeksi.

Lause 5.6.1. *Olkoot $A \in \mathbb{R}^{n\times n}$ ja $B \in \mathbb{R}^{n\times n}$ neliömatriiseja. Tällöin*

$$\det(AB) = (\det A)(\det B).$$

Todistus. Todistetaan lause käyttämällä lausetta 5.3.16. Määritellään funktio $\omega\colon \mathbb{R}^{n\times 1} \times \cdots \times \mathbb{R}^{n\times 1} \to \mathbb{R}$ kaavalla

$$\omega(v_1, \ldots, v_n) = \mathrm{vol}_{\mathbb{R}^n}(Av_1, \ldots, Av_n)$$

kaikilla $v_1, \ldots, v_n \in \mathbb{R}^n$. Koska $v \mapsto Av$ on lineaarikuvaus ja $\mathrm{vol}_{\mathbb{R}^n}$ on alternoiva multilineaarikuvaus, niin ω on alternoiva multilineaarikuvaus. Lisäksi

$$\omega(e_1, \ldots, e_n) = \mathrm{vol}_{\mathbb{R}^n}(Ae_1, \ldots, Ae_n) = \det A$$

ja

$$\det(AB) = \mathrm{vol}_{\mathbb{R}^n}((AB)e_1, \ldots, (AB)e_n) = \mathrm{vol}_{\mathbb{R}^n}(A(Be_1), \ldots, A(Be_n)) = \omega(Be_1, \ldots, Be_n).$$

Olkoot $b_1, \ldots, b_n \in \mathbb{R}^{n \times 1}$ matriisin $B = [b_{ji}]$ sarakkeet, eli $B = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}$. Tällöin lauseen 5.3.16 perusteella

$$\begin{aligned}
\omega(Be_1, \ldots, Be_n) &= \omega(b_1, \ldots, b_n) \\
&= \left(\sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} \right) \omega(e_1, \ldots, e_n) \\
&= (\det B)\omega(e_1, \ldots, e_n) = (\det B)(\det A).
\end{aligned}$$

Näin ollen

$$\det(AB) = (\det A)(\det B).$$

□

5.7 Matriisin transpoosin determinantti

Matriisilla A ja sen transpoosilla A^t on sama determinantti.

Lause 5.7.1. *Jokaisella neliömatriisilla $A \in \mathbb{R}^{n \times n}$ pätee*

$$\det A^t = \det A.$$

Erikoistapauksessa $n = 2$ todistus on suora lasku.

Erikoistapauksen $n = 2$ todistus. Olkoon

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

Tällöin

$$\det A^t = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - cb = \det A.$$

Näin ollen väite pätee. □

Yleisessä tilanteessa $n \geq 3$ todistus on hieman konstikas kirjoittaa tarkasti. Käsitellään siis tässä ainoastaan todistuksen idea. Tarkka todistus on annettu liitteessä C.1.

Lauseen 5.7.1 todistuksen idea. Huomiossa 5.4.2 todettiin, että determinantin kaavassa käydään läpi kaikki mahdolliset tavat valita jokaisesta sarakkeesta alkio sellaisella tavalla, että samalta riviltä ei koskaan valita kahta alkiota. Selvästi tämä vastaa sitä, että jokaiselta riviltä olisi valittu alkio sellaisella tavalla, että jokaisesta sarakkeesta valitaan täsmälleen yksi alkio. Tämä tarkoittaa yleisen $n \times n$-matriisin $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ tapauksessa sitä, että determinantin määritelmässä olevan summan

$$\sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \tag{5.7}$$

sijaan oltaisiinkin laskettu summa

$$\sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{5.8}$$

Koska matriisi A^t saadaan matriisista vaihtamalla rivit ja sarakkeet päittäin, eli $(A^t)_{ji} = A_{ij} = a_{ij}$, niin havaitaan, että

$$\begin{aligned}\sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} &= \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) (A^t)_{\sigma(1)1} (A^t)_{\sigma(2)2} \cdots (A^t)_{\sigma(n)n} \\ &= \det(A^t).\end{aligned}$$

Näin ollen jäljellä on osoittaa, että kaavojen (5.7) ja (5.8) summat antavat saman tuloksen, eli että

$$\det A = \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Tämä on totta, mutta vaatii argumentin. Tarkka perustelu liitteessä C.1. □

5.8 Determinantti ja kääntyvyys

Kirjataan ensin tärkeä huomio, että matriisin determinantti on nolla, jos sen sarakkeet ovat lineaarisesti riippuvia. Tämä tulos on suora seuraus vastaavasta tuloksesta alternoiville multilineaarikuvauksille.

Lause 5.8.1. *Olkoon $A \in \mathbb{R}^{n\times n}$ neliömatriisi. Tällöin $\det A = 0$, jos matriisin A sarakkeet ovat lineaarisesti riippuvia.*

Todistus. Oletuksen nojalla jono $(Ae_1, \ldots, Ae_n)$ on sidottu. Näin ollen lauseiden 5.3.20 ja 5.3.7 nojalla $\mathrm{vol}_{\mathbb{R}^n}(Ae_1, \ldots, Ae_n) = 0$. Nyt lemman 5.4.4 nojalla saadaan

$$\det A = \mathrm{vol}_{\mathbb{R}^{n\times 1}}(Ae_1, \ldots, Ae_n) = 0.$$

□

Osoitetaan nyt tärkein determinantteja koskeva tulos, josta koko tarina aloitettiin: determinantti kertoo onko neliömatriisi kääntyvä.

Lause 5.8.2. *Neliömatriisi* $A \in \mathbb{R}^{n \times n}$ *on kääntyvä, jos ja vain jos* $\det A \neq 0$.

Todistus. Oletetaan ensin, että matriisi A on kääntyvä ja osoitetaan, että $\det A \neq 0$. Koska $\det I = 1$, niin determinanttien tulokaavan (lause 5.6.1) nojalla

$$(\det A)(\det A^{-1}) = \det(AA^{-1}) = \det I = 1.$$

Näin ollen $\det A \neq 0$.

Oletetaan nyt, että matriisi A ei ole kääntyvä ja osoitetaan, että $\det A = 0$. Koska A ei ole kääntyvä, niin kääntyvyyden karakterisointilauseen (lause 3.10.1) nojalla matriisin A sarekkeet ovat lineaarisesti riippuvia toisistaan. Väite seuraa nyt suoraan lauseesta 5.8.1. □

5.9 Determinantin kehityskaava

Perustellaan nyt luvun lopuksi determinanttien niin sanotut *kehityskaavat* joiden avulla voidaan laskea $n \times n$-matriisien determinantteja palauttamalla ne $(n-1) \times (n-1)$-matriisien determinantteihin. Tätä varten esitellään seuraava merkintä.

Määritelmä 5.9.1. *Olkoon* $A = [a_{ji}] \in \mathbb{R}^{n \times n}$ *neliömatriisi ja olkoot* $k, \ell \in \{1, \dots, n\}$. *Matriisi* $A|_{kl}$ *on se* $(n-1) \times (n-1)$*-matriisi, joka saadaan poistamalla matriisista* A *rivi* k *ja sarake* ℓ*, eli* $A|_{k\ell}$ *on matriisi*

$$A|_{kl} = \begin{bmatrix} a_{11} & \cdots & a_{1(l-1)} & a_{1(l+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{(k-1)1} & \cdots & a_{(k-1)(l-1)} & a_{(k-1)(l+1)} & \cdots & a_{(k-1)n} \\ a_{(k+1)1} & \cdots & a_{(k+1)(l-1)} & a_{(k+1)(l+1)} & \cdots & a_{(k+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(l-1)} & a_{n(l+1)} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}.$$

Huomautus 5.9.2. *Usein matriisia* $A|_{k\ell}$ *merkitään lyhyesti* $A_{k\ell}$*, mutta tämä merkintä on varattu tässä esityksessä merkinnäksi matriisin* A *rivin* k *sarakkeen* ℓ *alkiolle.*

Esimerkki 5.9.3. *Olkoon*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \in \mathbb{R}^{3\times3}.$$

Nyt matriisi $A|_{11}$ *saadaan poistamalla matriisista* A *ensimmäinen rivi ja ensimmäinen sarake, eli lihavoidut alkiot*

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{2} & 3 & 4 \\ \mathbf{3} & 4 & 5 \end{bmatrix}.$$

Vastaavasti matriisi $A|_{32}$ saadaan poistamalla kolmas (eli alin) rivi ja toinen sarake, eli lihavoidut alkiot

$$\begin{bmatrix} 1 & \mathbf{2} & 3 \\ 2 & \mathbf{3} & 4 \\ \mathbf{3} & \mathbf{4} & \mathbf{5} \end{bmatrix}.$$

Tässä tapauksessa matriisit $A|_{k\ell} \in \mathbb{R}^{2\times 2}$ ovat siis seuraavat 9 matriisia:

$$\begin{aligned} A|_{11} &= \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}, & A|_{12} &= \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}, & A|_{13} &= \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, \\ A|_{21} &= \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}, & A|_{22} &= \begin{bmatrix} 1 & 3 \\ 3 & 5 \end{bmatrix}, & A|_{23} &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \\ A|_{31} &= \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, & A|_{32} &= \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, & A|_{33} &= \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}. \end{aligned}$$

Seuraavaa lausetta kutsutaan determinantin kehityskaavaksi sarakkeen ℓ suhteen.

Lause 5.9.4 (Determinantin kehityskaava sarakkeen suhteen)**.** *Olkoon $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ neliömatriisi. Olkoon $\ell \in \{1, \dots, n\}$. Tällöin*

$$\det A = \sum_{k=1}^{n} (-1)^{k+\ell} a_{k\ell} \det A|_{k\ell}.$$

Käsitellään vielä lauseen käyttöä havainnollistava esimerkki ennen lauseen todistamista.

Esimerkki 5.9.5. *Olkoon*

$$A = [a_{ji}] = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$$

Lauseen 5.9.4 perusteella matriisin A determinanttia voi laskea minkä tahansa sarakkeen suhteen. Lasketaan determinantti ensin kehittämällä se kolmannen sarakkeen suhteen, koska siinä on eniten nollia, eli valitaan $\ell = 3$. Tällöin

$$\begin{aligned} \det A &= \sum_{k=1}^{3} (-1)^{k+3} a_{k3} \det A|_{k3} \\ &= (-1)^{1+3} a_{13} \det A|_{13} + (-1)^{2+3} a_{23} \det A|_{23} + (-1)^{3+3} a_{33} \det A|_{33} \\ &= 1 \cdot 0 \cdot \det A|_{13} + (-1) \cdot 0 \cdot \det A|_{23} + 1 \cdot 1 \cdot \det A|_{33} \\ &= \det A|_{33} \\ &= \det \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = 1 \cdot 1 - 0 \cdot 2 = 1. \end{aligned}$$

Lasketaan nyt sama determinantti kehittämällä se keskimmäisen sarakkeen suhteen, eli valitaan $\ell = 2$. Tällöin

$$
\begin{aligned}
\det A &= \sum_{k=1}^{3} (-1)^{k+2} a_{k2} \det A|_{k2} \\
&= (-1)^{1+2} a_{12} \det A|_{12} + (-1)^{2+2} a_{22} \det A|_{22} + (-1)^{3+2} a_{32} \det A|_{32} \\
&= (-1) \cdot 0 \cdot \det A|_{12} + 1 \cdot 1 \cdot \det A|_{22} + (-1) \cdot 2 \cdot \det A|_{32} \\
&= \det A|_{22} - 2 \det A|_{32} \\
&= \det \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} - 2 \det \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \\
&= (1 \cdot 1 - 0 \cdot 3) - 2\,(1 \cdot 0 - 0 \cdot 2) = 1.
\end{aligned}
$$

Viimeinen tapaus, eli deteriminantin $\det A$ *laskeminen kehittämällä se ensimmäisen sarakkeen suhteen, jätetään lukijalle harjoitustehtäväksi.*

Huomautus 5.9.6. *Koska matriisilla ja sen transpoosilla on sama determinantti, voidaan matriisin determinantti kehittää myös jokaisen rivin suhteen. Tämä tulos jätetään harjoitustehtäväksi.*

5.9.1 Deteriminantin kehityskaavan todistus

Determinantin kehityskaavan (eli lauseen 5.9.4) todistus kokoaa kaiken kehittämämme teorian yhteen todistukseen. Käytännössä tämä tarkoittaa, että todistus on pitkä ja polveileva. Lukijaa kehotetaan sivuuttamaan todistus ensimmäisellä lukukerralla ja palaamaan siihen vasta determinantin tultua tutuksi harjoitusten myötä.

Kirjataan ennen lauseen 5.9.4 todistusta kaksi aputulosta. Ensimmäisen lemman todistus on mielekkäintä todistaa permutaatioiden avulla. Tämän vuoksi todistus on siirretty liitteeseen C.1. Lemma sanoo, että mikäli $n \times n$-matriisin $B = [b_{ji}]$ viimeisellä rivillä on ainoastaan yksi nollasta poikkeava kerroin $b_{nn} = 1$, niin matriisin B determinantin arvo ei riipu alimmasta rivistä eikä viimeisestä sarakkeesta. Formaalisti tämä voidaan kirjoittaa seuraavasti.

Lemma 5.9.7. *Olkoon $B = [b_{ji}] \in \mathbb{R}^{n \times n}$ sellainen matriisi, että $b_{ni} = 0$ kaikilla $i < n$ ja $b_{nn} = 1$. Tällöin*

$$\det B = \det B|_{nn}.$$

Toinen aputulos on erikoistapaus determinaatin merkin muuttumisesta sarekkeita ja rivejä permutoitaessa siten, että matriisin ℓ:s sarake siirretään viimeiseksi.

Lemma 5.9.8. *Olkoon $B = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} \in \mathbb{R}^{n \times n}$. Tällöin jokaisella $\ell \in \{1, \ldots, n\}$ pätee*

$$\det \begin{bmatrix} b_1 & \cdots & b_{\ell-1} & b_{\ell+1} & \cdots & b_n & b_\ell \end{bmatrix} = (-1)^{n-\ell} \det B.$$

Huomautus 5.9.9. *Koska $\det B^t = \det B$, niin vastaava tulos pätee myös siirettäessä matriisin B rivi ℓ viimeiseksi riviksi.*

Lemman 5.9.8 todistus. Koska kahden sarakevektorin paikan vaihtaminen (eli permutoiminen) vaihtaa deteriminantin merkkiä (lause 5.5.3), niin havaitaan, että

$$\begin{aligned}\det B &= \det \begin{bmatrix} b_1 & \cdots & b_{\ell-1} & b_\ell & b_{\ell+1} & \cdots & b_n \end{bmatrix} \\ &= (-1) \det \begin{bmatrix} b_1 & \cdots & b_{\ell-1} & b_{\ell+1} & b_\ell & \cdots & b_n \end{bmatrix} \\ &= (-1)(-1) \det \begin{bmatrix} b_1 & \cdots & b_{\ell-1} & b_{\ell+1} & b_{\ell+2} & b_\ell & \cdots & b_n \end{bmatrix}.\end{aligned}$$

Näin ollen

$$\det B = (-1)^{n-\ell} \det \begin{bmatrix} b_1 & \cdot & \cdots & b_{\ell-1} & b_{\ell+1} & \cdots & b_n & b_\ell \end{bmatrix}.$$

□

Lauseen 5.9.4 todistus. Olkoon $A = [a_{ji}] \in \mathbb{R}^{n \times n}$ ja $\ell \in \{1, \ldots, n\}$. Osoitetaan, että

$$\det A = \sum_{k=1}^{n} (-1)^{k+\ell} a_{k\ell} \det A|_{k\ell}.$$

Olkoot $a_1, \ldots, a_n \in \mathbb{R}^{n \times 1}$ matriisin A sarakkeet, eli $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$. Olkoon lisäksi $a_i = a_{1i} e_1 + \cdots + a_{nj} e_n$ jokaisella $i = 1, \ldots, n$.

Havaitaan ensin, että lemman 5.9.8 perusteella

$$\det A = (-1)^{n-\ell} \det \begin{bmatrix} a_1 & \cdot & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & a_\ell \end{bmatrix}.$$

Toisaalta determinantin sarakelineaarisuuden (lause 5.5.1) nojalla pätee

$$\begin{aligned}&\det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & a_\ell \end{bmatrix} \\ &= \det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & \sum_{k=1}^{n} a_{k\ell} e_k \end{bmatrix} \\ &= \sum_{k=1}^{n} a_{k\ell} \det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & e_k \end{bmatrix}.\end{aligned}$$

Yhdistämällä nämä havainnot saadaan, että

$$\det A = \sum_{k=1}^{n} (-1)^{n-\ell} a_{k\ell} \det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & e_k \end{bmatrix}.$$

Tarkastellaan nyt matriiseja $\det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & e_k \end{bmatrix}$ jokaisella indeksillä $k \in \{1, \ldots, n\}$ erikseen. Havaitaan aluksi, että

$$\begin{aligned}&\det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & e_k \end{bmatrix} \\ &= \det \begin{bmatrix} a_{11} & \cdots & a_{1(\ell-1)} & a_{1(\ell+1)} & \cdots & a_{1n} & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{(k-1)1} & \cdots & a_{(k-1)(\ell-1)} & a_{(k-1)(\ell+1)} & \cdots & a_{(k-1)n} & 0 \\ a_{k1} & \cdots & a_{k(\ell-1)} & a_{k(\ell+1)} & \cdots & a_{kn} & 1 \\ a_{(k+1)1} & \cdots & a_{(k+1)(\ell-1)} & a_{(k+1)(\ell+1)} & \cdots & a_{(k+1)n} & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{n(\ell-1)} & a_{n(\ell+1)} & \cdots & a_{nn} & 0 \end{bmatrix}.\end{aligned}$$

Näin ollen matriisin sarakelineearisuuden ja lauseen 5.8.1 nojalla, voidaan determinantin arvoa muuttamatta jokaisesta sarakkeesta $a_1, \dots, a_{\ell-1}, a_{\ell+1}, a_n$ vähentää vektorin e_k suuntainen komponentti, eli sarakeesta a_i vektori $a_{ki}e_k$. Näin ollen

$$\begin{aligned}
&\det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & e_k \end{bmatrix} \\
&= \det \begin{bmatrix}
a_{11} & \cdots & a_{1(\ell-1)} & a_{1(\ell+1)} & \cdots & a_{1n} & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots \\
a_{(k-1)1} & \cdots & a_{(k-1)(\ell-1)} & a_{(k-1)(\ell+1)} & \cdots & a_{(k-1)n} & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 & 1 \\
a_{(k+1)1} & \cdots & a_{(k+1)(\ell-1)} & a_{(k+1)(\ell+1)} & \cdots & a_{(k+1)n} & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots \\
a_{n1} & \cdots & a_{n(\ell-1)} & a_{n(\ell+1)} & \cdots & a_{nn} & 0
\end{bmatrix}.
\end{aligned}$$

Sovelletaan nyt lemmaa 5.9.8 tämän matriisin riveiin eli tämän matriisin transpoosin sarakkeisiin. Tällöin saadaan

$$\begin{aligned}
&\det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & e_k \end{bmatrix} \\
&= (-1)^{n-k} \det \begin{bmatrix}
a_{11} & \cdots & a_{1(\ell-1)} & a_{1(\ell+1)} & \cdots & a_{1n} & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots \\
a_{(k-1)1} & \cdots & a_{(k-1)(\ell-1)} & a_{(k-1)(\ell+1)} & \cdots & a_{(k-1)n} & 0 \\
a_{(k+1)1} & \cdots & a_{(k+1)(\ell-1)} & a_{(k+1)(\ell+1)} & \cdots & a_{(k+1)n} & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots \\
a_{n1} & \cdots & a_{n(\ell-1)} & a_{n(\ell+1)} & \cdots & a_{nn} & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 & 1
\end{bmatrix}.
\end{aligned}$$

Sovelletaan nyt lemmaa 5.9.7 saatuun matriisiin, jolloin saadaan

$$\begin{aligned}
&\det \begin{bmatrix}
a_{11} & \cdots & a_{1(\ell-1)} & a_{1(\ell+1)} & \cdots & a_{1n} & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots \\
a_{(k-1)1} & \cdots & a_{(k-1)(\ell-1)} & a_{(k-1)(\ell+1)} & \cdots & a_{(k-1)n} & 0 \\
a_{(k+1)1} & \cdots & a_{(k+1)(\ell-1)} & a_{(k+1)(\ell+1)} & \cdots & a_{(k+1)n} & 0 \\
\vdots & & \vdots & \vdots & & \vdots & \vdots \\
a_{n1} & \cdots & a_{n(\ell-1)} & a_{n(\ell+1)} & \cdots & a_{nn} & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 & 1
\end{bmatrix} \\
&= \det \begin{bmatrix}
a_{11} & \cdots & a_{1(\ell-1)} & a_{1(\ell+1)} & \cdots & a_{1n} \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{(k-1)1} & \cdots & a_{(k-1)(\ell-1)} & a_{(k-1)(\ell+1)} & \cdots & a_{(k-1)n} \\
a_{(k+1)1} & \cdots & a_{(k+1)(\ell-1)} & a_{(k+1)(\ell+1)} & \cdots & a_{(k+1)n} \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{n1} & \cdots & a_{n(\ell-1)} & a_{n(\ell+1)} & \cdots & a_{nn}
\end{bmatrix} \\
&= \det A|_{kl}.
\end{aligned}$$

Yhdistämällä saadut tulokset saadaan

$$\begin{aligned}
\det A &= \sum_{k=1}^{n} (-1)^{n-\ell} a_{k\ell} \det \begin{bmatrix} a_1 & \cdots & a_{\ell-1} & a_{\ell+1} & \cdots & a_n & e_k \end{bmatrix} \\
&= \sum_{k=1}^{n} (-1)^{n-\ell} a_{k\ell} (-1)^{n-k} \det A|_{kl} = \sum_{k=1}^{n} (-1)^{2n-(\ell+k)} a_{k\ell} \det A|_{kl} \\
&= \sum_{k=1}^{n} (-1)^{\ell+k} a_{k\ell} \det A|_{kl},
\end{aligned}$$

missä viimeisessä askeleessa on käytetty tietoa, että

$$(-1)^{2n-(\ell+k)} = (-1)^{2n}(-1)^{-(\ell+k)} = \frac{1}{(-1)^{\ell+k}} = (-1)^{\ell+k}.$$

Tämä päättää (pitkän) todistuksen. $\square$

Luku 6

Ominaisarvot ja ominaisvektorit

6.1 Motivointi

Tämä motivointi on poikkeuksellinen, koska katse suunnataan pitkälle tulevaisuuteen.

Suurin osa tähänastisista aiheista on käsitelly tavalla tai toisella yhtälöryhmän $[A \mid b]$ ratkaisemista. Ensimmäisessä ja toisessa luvussa tärkeässä roolissa olivat matriisin A rivit. Luvussa 3 ongelma tulkittiin matriisin sarakeiden avulla. Luvussa 4 mukaan tuotiin geometriaa ja matriisi itse tulkittiin uudelleen esimerkiksi QR-hajotelman avulla. Determinantit antoivat toisen lähestymistavan tarkastella matriisien sarakkeiden lineaarista riippumattomuutta.

Näkökulmaa voidaan kuitenkin vaihtaa ja kysyä voidaanko näillä ideoilla ymmärtää suoraan itse matriiseja paremmin. Tämä on mahdollista ja tässä tärkein työkalu tulee olemaan avaruuden $\mathbb{R}^{n\times 1}$ kanta. Paljastuu, että teoria tulee olemaan kaikkein vaikuttavin neliömatriisien tilanteessa. Heuristisesti syy tähän on se, että tällöin voidaan etsiä yhtä kantaa, joka sopii sekä yhtälön $Ax = b$ ratkaisujen $x \in \mathbb{R}^{n\times 1}$ että vakiovektoreiden $b \in \mathbb{R}^{n\times 1}$ esittämiseen matriisin A tapauksessa. Käytännön tasolla tämä tarkoittaa kannan muodostamista matriisin ominaisvektoreista ja tähän aiheeseen tutustuminen aloitetaan tässä luvussa.

Aloitetaan motivoivalla esimerkillä.

Esimerkki 6.1.1. *Oletetaan, että $A \in \mathbb{R}^{n\times n}$ on sellainen matriisi, että on olemassa avaruuden $\mathbb{R}^{n\times 1}$ sellainen kanta $(v_1, \dots, v_n)$, että jokaisella $i \in \{1, \dots, n\}$ pätee*

$$Av_i = \lambda_i v_i,$$

jollain luvulla $\lambda_i \in \mathbb{R}$. Ratkaistaan nyt yhtälö

$$Ax = b.$$

Koska $(v_1, \dots, v_n)$ on kanta, niin vektorilla $b \in \mathbb{R}^{n\times 1}$ on yksikäsitteiset koordinaatit tässä kannassa, eli $b = c_1v_1 + \dots + c_nv_n$. Vastaavasti, jos $x \in \mathbb{R}^{n\times 1}$, niin $x = y_1v_1 + \dots + y_nv_n$. Koska $Av_i = \lambda_i v_i$ jokaisella $i \in \{1, \dots, n\}$, niin saadaan yhtälö

$$\begin{aligned} c_1v_1 + \dots + c_nv_n = b = Ax = A(y_1v_1 + \dots + y_nv_n) &= y_1Av_1 + \dots + y_nAv_n \\ &= y_1\lambda_1v_1 + \dots + y_n\lambda_nv_n, \end{aligned}$$

eli jokaisella $i \in \{1, \ldots, n\}$ yhtälö $c_i = y_i \lambda_i$.

Tässä kannassa $(v_1, \ldots, v_n)$ yhtälön $Ax = b$ ratkaiseminen vastaa siis sellaisen yhtälön

$$Dy = c$$

ratkaisemista, missä $y = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^t$, $c = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix}^t$ ja $D = [d_{ji}] \in \mathbb{R}^{n \times n}$ on matriisi, jolle pätee $d_{ii} = \lambda_i$ jokaisella $i \in \{1, \ldots, n\}$ ja $d_{ji} = 0$, kun $j \neq i$.[1]

Ainoa jäljellä oleva kysymys tässä tilanteessa on selvittää vektorin b kertoimet $c_1, \ldots, c_n$ kannassa $(v_1, \ldots, v_n)$, jotka voidaan selvittää kertolaskulla $c = P^{-1}b$, missä $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. Jos yhtälön ratkaisu x halutaan ilmaista tavallisessa kannassa, niin vastaus saadaan kertolaskulla $x = Py$.

Tarkastellaan nyt vastavaa konkreettista esimerkkiä.

Esimerkki 6.1.2. *Olkoot*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \text{ ja } b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

ja ratkaistaan yhtälö

$$Ax = b.$$

Tarkastellaan vektoreita

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ ja } v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Vektorit v_1 ja v_2 ovat kohtisuorassa toisiaan vastaan, joten ne muodostavat avaruuden $\mathbb{R}^{2 \times 1}$ kannan. Lisäksi

$$Av_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3v_1$$

ja

$$Av_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = v_2.$$

Ratkaistaan nyt yhtälö

$$Ax = b,$$

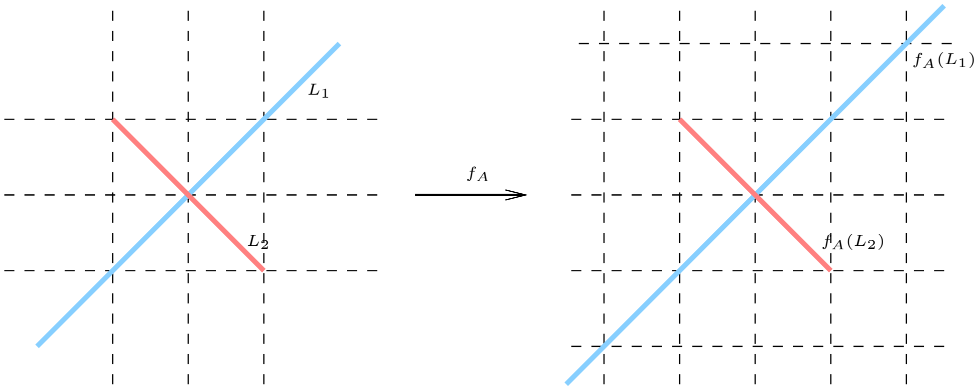
missä

$$b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Vektorin b kertoimet kannassa (v_1, v_2) ovat $(2, -1)$, eli $b = 2v_1 - v_2$. Näin ollen yhtälön ratkaisulle $x = y_1 v_1 + y_2 v_2$ saadaan yhtälö

$$2v_1 - v_2 = b = A(y_1 v_1 + y_2 v_2) = y_1 3 v_1 + y_2 v_2.$$

Koska vektorit v_1 ja v_2 ovat lineaarisesti riippumattomia, niin saadaan, että $y_1 = 2/3$ ja $y_2 = -1$.



Kuva 6.1: Esimerkin 6.1.2 vektoreiden v_1 ja v_2 virittämien suorien kuvat kuvauksessa f_A.

Laajennetaan hieman näkökulmaa ja tarkastellaan edellisen esimerkin matriisia $A \in \mathbb{R}^{2\times 2}$ vastaavaa kuvausta $f_A \colon \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}$, $x \mapsto Ax$. Geometrisesti tämä kuvaus pitää vektorin v_1 virittämän suoran L_1 paikallaan ja skaalaa sen vektoreita luvulla 3. Vektorin v_2 virittämän suoran L_2 vektorit kuvaus f_A puolestaan pitää paikallaan; katso kuva 6.1.

Matriisi A käyttäytyy kannassa (v_1, v_2) aivan samoin kuin matriisia

$$B = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix};$$

vastaava kuvaus $f_B \colon \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}$, $x \mapsto Bx$, standardikannassa. Matriisit A ja B liittyvätkin toisiinsa kaavalla

$$A = PBP^{-1}.$$

missä $P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ kuten edellä.

Huomaa, että tässä käytettiin lauseen 3.10.1 antamaa tietoa, että P on kääntyvä. Kaavan $A = PBP^{-1}$ voi osoittaa todeksi suoralla laskulla, mutta tähän seikkaan palataan myöhemmin yleisessa teoriassa.

Yleisesti neliömatriiseja $A \in \mathbb{R}^{n\times n}$ ja $B \in \mathbb{R}^{n\times n}$ kutsutaan *similaareiksi*, jos on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{n\times n}$. jolle pätee $A = PBP^{-1}$. Kääntyvää matriisia P puolestaan kutsutaan *kannanvaihtomatriisiksi*, koska matriisilla P kertominen vie kannan $(e_1, \ldots, e_n)$ kannaksi $(Pe_1, \ldots, Pe_n)$.

Koska jokainen avaruuden $\mathbb{R}^{n\times 1}$ kanta määrittelee kääntyvän matriisin, on kannavaihtomatriiseja runsaasti. Siten herää kysymys, että miten tunnistaa similaarit matriisit ja että onko annetun matriisin kanssa similaarien matriisien joukossa jossain mielessä paras matriisi.

Teoria, jota lähdetään kehittämään kertoo, esimerkin 6.1.2 matriisin A tapauksessa matriisi B on paras mahdollinen, koska matriisi A voidaan ilmaista venytysten avulla

[1]Tälläistä matriisia tullaan kutsumaan jatkossa diagonaalimatriisiksi.

sopivassa kannassa ja matriisi B ilmaisee nämä venytykset. Sanotaan, että matriisi A on *diagonalisoituva*.

Herää kysymys, että onko jokaisella neliömatriisilla $A \in \mathbb{R}^{n \times n}$ tälläinen ominaisuus, eli että onko aina olemassa sellainen kanta, jossa matriisi A voidaan esittää venytyksinä kanta-alkioista. Vastaus on, että näin ei ole. Paljastuu, että esimerkiksi matriiseilla

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ ja } \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

ei ole tälläistä ominaisuutta.

Tämä puolestaan herättää monia kysymyksiä: Milloin matriisi on similaari diagonaalimatriisin kanssa? Miten tämä selvitetään? Jos ei ole, niin mitä vaihtoehtoja silloin saadaan? Entä sitten?

Todellinen vastaus näihin kysymyksiin on erittäin syvällinen tulos, että matriisille voidaan aina löytää ns. normaalimuoto, eli avaruus $\mathbb{R}^{n \times 1}$ voidaan hajottaa sellaisiin aliavaruuksiin, joihin rajoitettuna matriisi käyttyy kuin skaalaus tai kierto. Jossain mielessä tämän kysymyksen selvittämiseen käytetään tämä luvun lisäksi myös kurssit Lineaarialgebra ja matriisilaskenta II ja III. Tähän projektiin sisältyvät sellaiset tulokset kuin symmetrisen matriisin spektraalilause, neliömatriisin singulaariarvohajotelma ja Jordanin normaalimuoto. Lähes kaikki lineaarialgebran sovellukset perustuvat näihin normaalimuotoja koskeviin tuloksiin.

Tätä pitkällistä projektia varten matriiseja siirrytään tulkitsemaan kurssilla Lineaarialgebra ja matriisilaskenta II geometrisemmin lineaarikuvauksina. Ensimmäinen vaihe tässä siirtymässä on kuitenkin matriisin ominaisarvojen ja ominaisvektoreiden määritteleminen.

6.2 Määritelmä

Määritelmä 6.2.1. *Luku $\lambda \in \mathbb{R}$ on matriisin $A \in \mathbb{R}^{n \times n}$* ominaisarvo*, jos on olemassa sellainen nollasta eroava vektori $x \in \mathbb{R}^{n \times 1}$, että*

$$Ax = \lambda x. \tag{6.1}$$

Nollasta poikkeavaa vektoria $x \in \mathbb{R}^{n \times 1}$, joka toteuttaa yhtälön (6.1) *kutsutaan matriisin A* ominaisvektoriksi *ominaisarvolla λ.*

Analysoimalla yhtälöä (6.1) havaitaan, että yhtälö toteutuu, jos ja vain jos

$$(A - \lambda I)x = 0, \tag{6.2}$$

missä I on $(n \times n)$-identiteettimatriisi. Tämä seuraa suoraan havainnoista, että $Ax = \lambda x$ on yhtäpitävää yhtälön $Ax = \lambda Ix$ eli yhtälön $Ax - \lambda Ix = 0$ kanssa. Näin ollen $Ax = \lambda x$, jos ja vain jos yhtälö (6.2) toteutuu. Näin on havaittu, että luku $\lambda \in \mathbb{R}$ on matriisin A ominaisarvo, jos ja vain jos yhtälöllä $(A - \lambda I)x = 0$ on nollasta poikkeavia eli epätriviaaleja ratkaisuja. Koska epätriviaalien ratkaisujen olemassaolo on yhtäpitävää sen kanssa, että matriisi $A - \lambda I$ ei ole kääntyvä, niin on havaittu seuraava tulos, joka kirjataan lemmaksi.

Lemma 6.2.2. *Luku $\lambda \in \mathbb{R}$ on matriisin $A \in \mathbb{R}^{n\times n}$ ominaisarvo, jos ja vain jos matriisi $A - \lambda I$ ei ole kääntyvä.*

Todistus. Kirjoitetaan edellä tehty päättely vielä tarkemmin uudelleen.

Oletetaan, että λ on matriisin A ominaisarvo. Tällöin on olemassa sellainen vektori $x \in \mathbb{R}^{n\times 1}$, $x \neq 0$, että $Ax = \lambda x$. Tällöin

$$(A - \lambda I)x = Ax - \lambda I x = Ax - \lambda x = 0.$$

Näin ollen matriisi $A - \lambda I$ ei ole kääntyvä lauseen 3.10.1 perusteella.

Oletetaan nyt, että matriisi $A - \lambda I$ ei ole kääntyvä. Tällöin lauseen 3.10.1 perusteella yhtälöllä $(A - \lambda I)x = 0$ on epätriviaali ratkaisu, eli on olemassa sellainen nollasta poikkeava vektori $x \in \mathbb{R}^{n\times 1}$, että $(A - \lambda I)x = 0$. Tällöin

$$Ax = (A - \lambda I + \lambda I)x = (A - \lambda I)x + \lambda I x = 0 + \lambda x = \lambda x.$$

□

Yleensä lemma 6.2.2 muotoillaan determinantin avulla. Kirjaamme myös tämän muotoilun lauseeksi sen käyttökelpoisuuden vuoksi.

Lause 6.2.3. *Luku $\lambda \in \mathbb{R}$ on neliömatriisin $A \in \mathbb{R}^{n\times n}$ ominaisarvo, jos ja vain jos* $\det(A - \lambda I) = 0$.

Todistus. Lemman 6.2.2 nojalla λ on matriisin A ominaisarvo, jos ja vain jos matriisi $A - \lambda I$ ei ole kääntyvä. Toistaalta lauseen 5.8.2 nojalla matriisi $A - \lambda I$ on kääntyvä, jos ja vain jos $\det(A - \lambda I) \neq 0$. □

Lausetta 6.2.3 voidaan tulkita kahdella tavalla. Ensinnäkin se sanoo, että annetun matriisin kaikki ominaisarvot voidaan selvittää yhdestä determinattiin liittyvästä yhtälöstä. Toisaalta se sanoo myös, että ominaisarvot voi selvittää ilman, että tarvitsee etsiä ominaisarvoja vastaavia ominaisvektoreita.[2]

Tarkastellaan nyt kolmea eri esimerkkiä ominaisarvojen olemassaolosta.

Esimerkki 6.2.4. *Olkoon*

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

Etsitään tämän matriisin ominaisarvot $\lambda \in \mathbb{R}$ ratkaisemalla yhtälö $\det(A - \lambda I) = 0$. *Koska*

$$A - \lambda I = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix},$$

niin

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda) - 1 \cdot 0 = (1-\lambda)(2-\lambda).$$

Näin ollen yhtälön $\det(A - \lambda I) = 0$ *ratkaisut ovat* $\lambda = 1$ *ja* $\lambda = 2$.

[2] Tämä on kaksiteräinen miekka, kuten tulemme huomaamaan.

Esimerkki 6.2.5. *Olkoon nyt*

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Etsitään tämän matriisin ominaisarvot $\lambda \in \mathbb{R}$ *ratkaisemalla yhtälö* $\det(B - \lambda I) = 0$. *Nyt*

$$\det(B-\lambda I) = \det \begin{bmatrix} 1-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda)-1 = \big(2 - \lambda - 2\lambda + \lambda^2\big)-1 = 1-3\lambda+\lambda^2.$$

Yhtälön

$$\lambda^2 - 3\lambda + 1 = 0$$

ratkaisut löydetään helposti käsin käyttämällä yleistä toisen asteen ratkaisukaavaa tai neliöksi täydentämällä:

$$0 = \lambda^2 - 3\lambda + 1 = \lambda^2 - 2\frac{3}{2}\lambda + (\frac{3}{2})^2 - (\frac{3}{2})^2 + 1 = (\lambda - \frac{3}{2})^2 - \frac{9}{4} + 1.$$

Näin ollen

$$(\lambda - \frac{3}{2})^2 = \frac{5}{4}$$

eli

$$\lambda = \frac{3}{2} \pm \frac{\sqrt{5}}{2}.$$

Molemmissa esimerkeissä 2×2-matriisilla oli kaksi erisuurta ominaisarvoa. Tämä ei ole välttämätöntä: niitä voi olla vain yksi tai ei lainkaan.

Esimerkki 6.2.6. *Olkoon*

$$C = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Tällöin $C = 2I$, *joten*

$$\det(C - \lambda I) = \det(2I - \lambda I) = \det\left((2-\lambda)I\right) = (2-\lambda)^2 \det I = (2-\lambda)^2.$$

Näin ollen $\lambda = 2$ *on ainoa ominaisarvo.*

Esimerkki 6.2.7. *Olkoon*

$$D = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Tällöin

$$\det(D - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = (-\lambda)^2 - (-1)1 = \lambda^2 + 1 > 0.$$

Näin ollen yhtälöllä $\det(D - \lambda I) = 0$ *ei ole (reaalisia) ratkaisuja eli matriisilla* D *ei ole ominaisarvoja.*

Huomautus 6.2.8. *Edellä olevat esimerkit johtavat päätelmään, että matriisin $A \in \mathbb{R}^{n\times n}$ ominaisarvot $\lambda \in \mathbb{R}$ selvittävä yhtälö* $\det(A-\lambda I) = 0$ *on itseasiassa polynomiyhtälö tuntemattoman λ suhteen. Tämä on oikea havainto, sillä pätee*

$$\det(A - \lambda I) = (-1)^n\lambda^n + (-1)^{n-1}c_{n-1}\lambda^{n-1} + \cdots - c_1\lambda + \det A,$$

missä $c_1, \ldots, c_{n-1} \in \mathbb{R}$. Emme tarvitse tätä havaintoa tässä luvussa, joten todistus jätetään kiinnostuneelle lukijalle. Tapaus $n = 2$ voidaan ratkaista helposti käsin ja yleinen tapaus seuraa esimerkiksi determinantin kehityskaavasta induktiolla. Todettakoon kuitenkin seuraava huomio. Koska funktio $\lambda \mapsto$ $\det(A - \lambda I)$ *on asteen n polynomi, niin sillä on korkeintaan n nollakohtaa. Näin ollen matriisilla A on korkeintaan n ominaisarvoa.*

Esimerkki 6.2.9. *Esimerkin 6.2.7 matriisi D on erikoistapaus hieman yleisemmästä kiertomatriisien tapauksesta. Olkoot $a, b \in \mathbb{R}$ sellaisia lukuja, että $a^2 + b^2 = 1$ ja olkoon*

$$R = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Matriisi D on tälläinen matriisi parametreilla $a = 0$ ja $b = -1$. Muita esimerkkejä tälläisistä matriiseista ovat esimerkiksi

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \text{ ja } \begin{bmatrix} \sqrt{2}/\sqrt{3} & -1/\sqrt{3} \\ 1/\sqrt{3} & \sqrt{2}/\sqrt{3} \end{bmatrix}.$$

Oletetaan nyt, että $b \neq 0$. Koska jokaisella $\lambda \in \mathbb{R}$ pätee

$$\det(R - \lambda I) = \det \begin{bmatrix} a-\lambda & -b \\ b & a-\lambda \end{bmatrix} = (a-\lambda)(a-\lambda) - b(-b) = (a-\lambda)^2 + b^2 \geq b^2 > 0,$$

niin yhtälöllä $\det(R - \lambda I) = 0$ *ei ole (reaalisia) ratkaisuja ja matriisilla R ei siis tässä tapauksessa ole ominaisarvoja. Huomaa, että tapauksessa $b = 0$ pätee* $\det(R - \lambda I) = (a - \lambda)^2$ *ja saadaan, että $\lambda = a$ on matriisin R ominaisarvo.*

Huomautus 6.2.10. *Esimerkin 6.2.9 matriiseja R kutsutaan tason $\mathbb{R}^2$* kiertomatriiseiksi *(tai lyhyesti* kierroiksi*). Syy tähän on seuraava. Olkoot $a, b \in \mathbb{R}$ sellaisia, että $a^2 + b^2 = 1$. Koska piste (a, b) on tason yksikköympyrällä, niin on olemassa sellainen kulma $\theta \in \mathbb{R}$, että* $\cos\theta = a$ *ja* $\sin\theta = b$*. Matriisi R voidaan siis kirjoittaa muodossa*

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Kulmaa θ kutsutaan kiertokulmaksi*, koska matriisin R määrittelemä kuvaus $f_R \colon \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}$, $x \mapsto Rx$ kiertää tasoa $\mathbb{R}^2$ vastapäivään kulman θ verran.*

6.2.1 Ominaisvektoreiden löytäminen

Kuten edellä käsiteltiin matriisin $A \in \mathbb{R}^{n\times n}$ ominaisarvot $\lambda \in \mathbb{R}$ vastaavat determinanttiyhtälön $\det(A - \lambda I) = 0$ ratkaisuja. Kun matriisin A ominaisarvot $\lambda_1, \dots, \lambda_m$ on näin selvitetty, voidaan vastaavat ominaisvektorit selvittää yhtälöstä $(A - \lambda_k I)x = 0$. Koska kyseessä on tuttu matriisiyhtälön ratkaiseminen, aloitetaan esimerkeillä.

Esimerkki 6.2.11. *Olkoon*

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

kuten esimerkissä 6.2.4. Matriisin A ominaisarvot ovat siis $\lambda_1 = 1$ ja $\lambda_2 = 2$. Etsitään vastaavat ominaisvektorit.

Koska

$$A - \lambda_1 I = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},$$

niin havaitaan, että yhtälön $(A - \lambda_1 I)x = 0$ ratkaisut ovat

$$x = \begin{bmatrix} t \\ 0 \end{bmatrix} = te_1,$$

missä $t \in \mathbb{R}$. Näin ollen matriisin A ominaisarvoa λ_1 vastaavat ominaisvektorit ovat $\{te_1 : t \in \mathbb{R} \setminus \{0\}\}$.

Vastaavasti

$$A - \lambda_2 I = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Näin ollen ominaisarvoa λ_2 vastaavat ominaisvektorit saadaan yhtälöstä

$$\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Näin ollen ominaisarvoa λ_2 vastaaville ominaisvektoreille $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ pätee

$$-x_1 + x_2 = 0$$

eli $x_2 = x_1$. Näin ollen ominaisvektorit tässä tapauksessa ovat $\{t(e_1 + e_2) : t \in \mathbb{R} \setminus \{0\}\}$.

Huomautus 6.2.12. *Matriisin $A \in \mathbb{R}^{n\times n}$ ominaisarvot ja ominaisvektorit voidaan siis löytää kaksivaiheisella prosessilla:*

- *Etsitään polynomiyhtälön $\det(A - \lambda I) = 0$ ratkaisut $\lambda = \lambda_1, \dots, \lambda_m$.*
- *Etsitään yhtälöryhmän $(A - \lambda_k I)x = 0$ nollasta poikkeavat ratkaisut $x \in \mathbb{R}^{n\times 1}$, jokaisella ominaisarvolla λ_k.*

Tämä prosessi on kätevä pienillä matriiseilla, mutta muuttuu laskennallisesti haastavaksi suurilla matriiseilla.

6.3 Ominaisavaruudet

Edellä käsitellyt esimerkit johdattavat ajatukseen, että matrisiin $A \in \mathbb{R}^{n\times n}$ ominaisarvoihin liittyvät ominaisvektorit muodostavat avaruuden $\mathbb{R}^{n\times 1}$ aliavaruuden, ainakin kunhan nollavektori lisätään ominaisvektoreiden joukkoon. Tämä on todellakin näin ja tällä on aliavaruudella on matriisiin liittyvä geometrinen merkitys, johon viitattiin motivointiluvussa. Aloitetaan kirjaamalla ylös, mitä ominaisvektoreiden muodostamalla aliavaruudella tarkoitetaan.

Olkoon $A \in \mathbb{R}^{n\times n}$ matriisi ja määritellään jokaisella matriisin A ominaisarvolla $\lambda \in \mathbb{R}$ joukko

$$E(\lambda, A) = \{x \in \mathbb{R}^{n\times 1} \colon Ax = \lambda x\}.$$

Huomautus 6.3.1. *Huomaa, että joukko $E(\lambda, A)$ koostuu matriisin A ominaisarvoa λ vastaavista ominaisvektoreista ja nollavektorista 0.*

Lemma 6.3.2. *Olkoon $A \in \mathbb{R}^{n\times n}$ matriisi ja $\lambda \in \mathbb{R}$ matriisin A ominaisarvo. Tällöin osajoukko $E(\lambda, A) \subset \mathbb{R}^{n\times 1}$ on avaruuden $\mathbb{R}^{n\times 1}$ aliavaruus.*

Todistus. Olkoot $x, y \in E(\lambda, A)$ ja $a, b \in \mathbb{R}$. Osoitetaan, että $z = ax + by \in E(\lambda, A)$. Tällöin

$$A(ax + by) = A(ax) + A(by) = aAx + bAy = a\lambda x + b\lambda y = \lambda(ax + by).$$

Näin ollen $ax + by \in E(\lambda, A)$. Joukko $E(\lambda, A)$ on siis aliavaruus. □

Määritelmä 6.3.3. *Matriisin $A \in \mathbb{R}^{n\times n}$ ominaisarvoon $\lambda \in \mathbb{R}$ liittyvää aliavaruutta $E(\lambda, A)$ kutsutaan* matriisin A ominaisarvoa λ vastaavaksi ominaisavaruudeksi*.*

Huomautus 6.3.4. *Ominaisavaruudella on seuraava geometrinen tulkinta. Kaikilla $x \in E(\lambda, A)$ siis pätee $Ax = \lambda x$. Näin ollen kuvaus $f_A \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$, $x \mapsto Ax$, kuvaa aliavaruuden $E(\lambda, A)$ itseensä, eli $f_A(E(\lambda, A)) \subset E(\lambda, A)$. Näin ollen kuvauksen f_A rajoittuma aliavaruuteen voidaan kirjoittaa muodossa $f_A|_{E(\lambda,A)} \colon E(\lambda, A) \to E(\lambda, A)$ ja sillä on kaava $x \mapsto \lambda x$.*

Huomautus 6.3.5. *Koska λ on matriisin A ominaisarvo, niin tiedetään, että on olemassa nollavektorista poikkeava vektori $x \in \mathbb{R}^{n\times 1}$, jolle pätee $Ax = \lambda x$. Näin ollen ominaisarvoa λ vastaava aliavaruus $E(\lambda, A)$ ei ole koskaan nolla-avaruus $\{0\}$.*

Huomautus 6.3.6. *Aliavaruus $E(\lambda, A)$ voi olla koko avaruus $\mathbb{R}^{n\times 1}$. Esimerkiksi matriisin $A = 2I \in \mathbb{R}^{n\times n}$ ainoa ominaisarvo on 2, mutta ominaisavaruus $E(2, A)$ on koko avaruus $\mathbb{R}^{n\times 1}$, koska jokainen vektori $x \in \mathbb{R}^{n\times 1}$ toteuttaa tässä tapauksessa yhtälön $Ax = 2x$.*

Tarkastellaan nyt kahta esimerkkiä.

Esimerkki 6.3.7. *Olkoon*

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

Tällöin

$$\begin{aligned}
\det(A-\lambda I) = \det \begin{bmatrix} 1-\lambda & 0 & 2 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{bmatrix} &= (1-\lambda)\det \begin{bmatrix} 1-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix} + \det \begin{bmatrix} 0 & 2 \\ 1-\lambda & 0 \end{bmatrix} \\
&= (1-\lambda)(1-\lambda)(2-\lambda) - 2(1-\lambda) \\
&= (1-\lambda)\left((1-\lambda)(2-\lambda) - 2\right) \\
&= (1-\lambda)\left(2-\lambda-2\lambda+\lambda^2-2\right) \\
&= (1-\lambda)\left(\lambda^2-3\lambda\right) \\
&= \lambda(1-\lambda)\left(\lambda-3\right).
\end{aligned}$$

Näin ollen matriisilla A on kolme ominaisarvoa 0, 1 *ja* 3. *Näitä ominaisarvoja vastaavat ominaisvektorit ovat muotoa tv_1, tv_2 ja tv_3, missä $v_1 = 2e_1 - e_3$, $v_2 = e_2$ ja $v_3 = e_1 + e_3$. Matriisin A ominaisavaruudet ovat siis*

$$E(0,A) = \mathrm{Sp}(v_1),\ \ E(1,A) = \mathrm{Sp}(v_2) \text{ ja } E(3,A) = \mathrm{Sp}(v_3).$$

Esimerkki 6.3.8. *Olkoon*

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

Tällöin

$$\begin{aligned}
\det(A-\lambda I) = \det \begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{bmatrix} &= (1-\lambda)\det \begin{bmatrix} 2-\lambda & 0 \\ 0 & 1-\lambda \end{bmatrix} + \det \begin{bmatrix} 0 & 1 \\ 2-\lambda & 0 \end{bmatrix} \\
&= (1-\lambda)(2-\lambda)(1-\lambda) - (2-\lambda) \\
&= (2-\lambda)\left((1-\lambda)^2 - 1\right) \\
&= (2-\lambda)\left(1-2\lambda+\lambda^2-1\right) \\
&= \lambda(2-\lambda)\left(\lambda-2\right) = \lambda(\lambda-2)^2.
\end{aligned}$$

Näin ollen matriisilla A on kaksi ominaisarvoa 0 *ja* 2. *Ominaisarvoa* 0 *vastaavat ominaisvektorit ovat muotoa tv_1, missä $v_1 = e_1 - e_3$. Vastaavasti ominaisarvoa* 2 *vastaavat ominaisvektorit ovat muotoa $tv_2 + sv_3$, missä $v_2 = e_2$ ja $v_3 = e_1 + e_3$. Näin ollen matriisin A ominaisavaruudet ovat*

$$E(0,A) = \mathrm{Sp}(v_1) \text{ ja } E(2,A) = \mathrm{Sp}(v_2, v_3).$$

6.4 Matriisin diagonalisoiminen ominaisarvojen avulla

Palataan nyt motivoivan luvun kysymykseen, että mitä matriisista voidaan sanoa, jos omainaisvektoreista voitaisiin muodostaa kanta. Tämä kysymys motivoi diagonaalimatriisin käsitteen.

Määritelmä 6.4.1. *Neliömatriisin* $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ *alkioita* $a_{11}, \dots, a_{nn}$ *kutsutaan* matriisin A diagonaalialkioksi *ja jonoa* $(a_{11}, \dots, a_{nn})$ matriisin A diagonaaliksi.

Määritelmä 6.4.2. *Neliömatriisi* $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ *on* diagonaalimatriisi, *jos matriisin* A *nollasta poikkeavat alkiot ovat diagonaalilla, eli* $a_{ji} = 0$ *kaikilla* $j \neq i$.

Aloitetaan esimerkillä ominaisarvojen ja diagonaalimatriisien yhteydestä.

Esimerkki 6.4.3. *Olkoon* $D = [d_{ji}] \in \mathbb{R}^{n\times n}$ *diagonaalimatriisi, jonka diagonaalialkiot ovat* $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, *eli* $d_{ii} = \lambda_i$ *jokaisella* $i \in \{1, \dots, n\}$. *Tällöin* $De_i = \lambda_i e_i$ *jokaisella* $i \in \{1, \dots, n\}$. *Näin ollen diagonaalialkiot ovat matriisin* D *ominaisarvoja ja standardikannan alkiot ominaisvektoreita.*

Seuraava tulos kertoo edellisen esimerkin yleisen muodon: matriisi voidaan ilmaista diagonaalimatriisin avulla, jos sen ominaisvektoreista voidaan muodostaa kanta. Jotta tämä yhteys voidaan muotoilla tarkasti, niin kannattaa palauttaa mieleen lauseen 3.10.1 tulos, että $(v_1, \dots, v_n)$ on avaruuden $\mathbb{R}^{n\times 1}$ kanta, jos ja vain jos matriisi $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ on kääntyvä.

Lause 6.4.4. *Olkoon* $A \in \mathbb{R}^{n\times n}$ *sellainen matriisi, että on olemassa avaruuden* $\mathbb{R}^{n\times 1}$ *kanta* $(v_1, \dots, v_n)$, *joka koostuu matriisin* A *ominaisvektoreista. Olkoot* $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ *ja* $D \in \mathbb{R}^{n\times n}$ *diagonaalimatriisi*

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

missä $\lambda_i \in \mathbb{R}$ *on vektoria* v_i *vastaava matriisin* A *ominaisarvo. Tällöin*

$$A = PDP^{-1}.$$

Todistus. Jokaisella $i \in \{1, \dots, n\}$ pätee $Pe_i = v_i$ ja siten $P^{-1}v_i = e_i$. Toisaalta $De_i = \lambda_i e_i$ jokaisella $i \in \{1, \dots, n\}$. Koska λ_i on vektoria v_i vastaava matriisin A ominaisarvo, niin $Av_i = \lambda_i v_i$. Näin ollen

$$P^{-1}APe_i = P^{-1}Av_i = P^{-1}(\lambda_i v_i) = \lambda_i P^{-1}v_i = \lambda_i e_i$$

jokaisella $i \in \{1, \dots, n\}$, eli

$$P^{-1}AP = \begin{bmatrix} P^{-1}APe_1 & \cdots & P^{-1}APe_n \end{bmatrix} = \begin{bmatrix} \lambda_1 e_1 & \cdots & \lambda_n e_n \end{bmatrix} = D.$$

Tällöin

$$A = PDP^{-1}.$$

□

Matriiseja, jotka voidaan kirjoittaa kääntyvän matriisin ja diagonaalimatriisin avulla, kutsutaan diagonalisoituviksi.

Määritelmä 6.4.5. *Neliömatriisia $A \in \mathbb{R}^{n \times n}$ sanotaan* diagonalisoituvaksi, *jos on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{n \times n}$ ja diagonaalimatriisi $D \in \mathbb{R}^{n \times n}$, että $A = PDP^{-1}$.*

Huomautus 6.4.6. *Kurssilla Lineaarialgebra ja matriisilaskenta II tulkitaan lineaarikuvausten teorian avulla, että yhtälö $A = PDP^{-1}$ tarkoittaa matriisin A esittämistä kannassa matriisin P sarakkeiden määräämässä kannassa. Tämä vastaa jälleen yhteeen motivointiluvussa tehtyyn tulkintaan.*

Kirjataan vielä tulos, joka yleistää esimerkin 6.4.3 huomion toisella tavalla. Yhdessä lauseen 6.4.4 kanssa tämä lause antaa tuloksen, että matriisi on diagonalisoituva jos ja vain jos matriisin ominaisvektoreista voidaan muodostaa kanta.

Lause 6.4.7. *Olkoon $A \in \mathbb{R}^{n \times n}$ sellainen matriisi, että on olemassa kääntyvä matriisi $P \in \mathbb{R}^{n \times n}$ ja diagonaalimatriisi $D \in \mathbb{R}^{n \times n}$, että $A = PDP^{-1}$. Tällöin matriisin P sarakkeet ovat matriisin A ominaisvektoreita ja matriisin D diagonaalialkiot näitä ominaisvektoreita vastaavia ominaisarvoja.*

Todistus. Olkoon $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ ja olkoon $(\lambda_1, \ldots, \lambda_n)$ matriisin D diagonaali. Tällöin

$$Av_i = PDP^{-1}v_i = PD(P^{-1}v_i) = PDe_i = P(\lambda_i e_i) = \lambda_i Pe_i = \lambda_i v_i.$$

Tämä päättää todistuksen. □

Korollaari 6.4.8. *Matriisi A on diagonalisoituva, jos ja vain jos on olemassa avaruuden $\mathbb{R}^{n \times 1}$ kanta $(v_1, \ldots, v_n)$, jonka vektorit ovat matriisin A ominaisvektoreita.*

Todistus. Ehdon riittävyys seuraa lauseesta 6.4.4. Ehdon välttämättömyys puolestaan lauseesta 6.4.7. □

Osa II

Lineaarialgebra ja matriisilaskenta II

Lukijalle

Tällä kurssilla siirrytään tarkastelemaan sarakeiden ja matriisien sijaan yleisiä vektoriavaruuksia ja niiden välisiä lineaarikuvauksia.

Matemaattisesta näkökulmasta yleisen teorian edut ovat kiistattomat. Vektoriavaruuden määritelmä antaa sellaisen yleisyyden, joka on sekä riittävän tehokas konseptualisoimaan yhtälöryhmien ratkaisemiseen tarvittava käsitteistö kuten virittäminen ja vapaus että samalla riittävän yleinen laajentamaan tämän käsitteistön sarakeavaruuksia yleisempiin tilanteisiin. Tämä muuttaa matriisilaskennan yleisemmäksi lineaarialgebraksi.

Konkreettisia esimerkkejä tämän teorian sovelluskohteista ovat esimerkiksi ääriarvo-ongelmat usemman muuttujan analyysissä – eli vektorianalyysissä – sekä derivoinnin ja integroinnin ymmärtäminen lineaarisina operaatioina. Näitä lineaarialgebran sovelluksia yhden ja useamman muuttujan analyysiin puolestaan hyödynnetään laajalti modernissa matematiikassa aina todennäköisyyslaskennasta differentiaaligeometriaan ja algebrallisesta topologiasta finanssimatematiikkaan. Yhteistä näille sovelluksille on, että lineaarialgebrallista käsitteistöä hyödynnetään siirtämään tarkastelu yksittäisistä funktioista, tai muista vektoreina ymmärrettävistä objekteista, funktioiden avaruuksiin ja niiden ominaisuuksiin. Lineaarialgebraa voidaan soveltaa, kun tarkasteltavilla funktioiden avaruuksilla on ns. lineaarista rakennetta eli tilanteissa, joissa funktioiden lineaarikombinaatiot kuuluvat samaan avaruuteen.

Pelkkä matematiikan sisäinen kauneus ei kuitenkaan riitä selittämään käsiteltävän materiaalin hyödyllisyyttä. Kurssin Lineearialgebra ja matriisilaskenta I antamia matriisilaskennan menetelmiä voidaan soveltaa tilanteissa, joissa tarkasteltava kysymys on jo valmiiksi muotoiltu sarakkeiden tai matriisien muotoon. Sarakkeiksi ja matriiseiksi muotoiltu tieto vaatii kuitenkin aina tulkinnan. Yleensä annettujen lukujen merkitys on selitetty viereisessä sarakkeessa, joka tarkoittaa, että tieto on annettu muodossa, jossa vierekkäiset sarakkeet ovat kuin kuvauksen lähtö- ja maaliavaruus.

Matemaattisesti katsoen onkin mielekästi käsitellä sarakkeiden sijaan funkioita, sillä funkion käsite antaa systemaattisen tavan kuvata tarkasteltavaa informaatiota. Tällöin tarkasteltavat funktiot kuuluvat vektoriavaruuteen $\mathscr{F}(X)$, jolla tarkoitetaan annetun joukon X reaaliarvoisten funktioiden $f\colon X \to \mathbb{R}$ joukkoa. Funktioavaruus $\mathscr{F}(X)$ on hyvin yleinen käsite, joten sen avulla voidaan kuvailla ja käsitellä hyvin moninaisia ilmiöitä.

Esimerkiksi taulukkolaskentaohjelman kenttiin syötetyt luvut muodostavat funktion $f\colon X \to \mathbb{R}$, jossa X on täytettyjen kenttien joukko ja jokaisella $x \in X$ luku $f(x)$ on kentässä oleva arvo. Helpoimmassa muodossaan nämä kentät sisältyvät yhteen taulukon riviin tai sarakkeeseen, jolloin funktion voi ajatella suoraan rivi- tai sarakevektorina. Mikään ei kuitenkaan estä tarkastelemasta kenttiä eri taulukoista tai yleisemmin tietokannan tauluja, jolloin funktion tulkinta sarakkeena vaatii valintoja.

Toisaalta, jos joukko X muodostuu esimerkiksi osakesalkkuun valituista osakkeista, niin funktio $f\colon X \to \mathbb{R}$ voi kuvata esimerkiksi osakkeen arvoa tai salkussa olevien osakkeiden määrää. Tässä tapauksessa funktioavaruus $\mathscr{F}(X)$ sisältää esimerkiksi kaikki joukon X osakkeisiin liittyvät eri tunnuslukuja kuvaavat funktiot. Koska $\mathscr{F}(X)$ on vektoriavaruus eikä pelkkä joukko, niin on mahdollista tutkia esimerkiksi salkun arvon optimointia hyödyntäen joukon $\mathscr{F}(X)$ lineaarista rakennetta.

Koska funktioavaruuden $\mathscr{F}(X)$ alkiot ovat funkioita eivätkä abstrakteja vektoreita, on niillä luonnollisia operaatioita, jotka voidaan tulkita lineaarialgebran keinoin. Esimerkiksi tarkasteltessa suunnattua verkkoa $\Gamma = (V, E)$ voidaan tarkastella sekä funktioita verkon Γ solmujen joukolla V eli avaruutta $\mathscr{F}(V)$ että funktioita verkon Γ kaarien joukolla E eli avaruutta $\mathscr{F}(E)$. Koska verkon kaaret kertovat verkon solmujen yhteydet, voidaan tarkastella funktion $f\colon V \to \mathbb{R}$ muutosta kaaria pitkin. Tämä muutos on luonnollisinta tulkita funktiona $\nabla f\colon E \to \mathbb{R}$

$$(\nabla f)(e) = f(y) - f(x)$$

missä $e = (x, y)$ on kaari solmusta x solmuun y. Tällä $f \mapsto \nabla f$ on kuvaus funktioavaruudesta $\mathscr{F}(V)$ funktioavaruuteen $\mathscr{F}(E)$ eli kuvaus $\nabla\colon \mathscr{F}(V) \to \mathscr{F}(E)$. Tämä kuvaus on itseasiassa lineaarikuvaus ja, mikäli verkko Γ on äärellinen, niin tämä kuvaus voidaan tulkita matriisina.

Tämä kurssi keskittyy äärellisulotteisten vektoriavaruuksien ja niiden välisten lineaarikuvausten teoriaan. Yksi tämän teorian päätuloksista on, että äärellisulotteisella vektoriavaruudella on kanta. Tämän abstraktin tuloksen seuraukset ovat moninaiset. Erityisesti se sanoo, että äärellisen joukon X tapauksessa edellä esitelty vektoriavaruus $\mathscr{F}(X)$ voidaan samastaa sarakeavaruuden kanssa. Tämä tarkoittaa, että funktio $f\colon X \to \mathbb{R}$, joka tulkitaan vektorina, voidaan samastaa sarakevektorin kanssa, kunhan avaruudelle $\mathscr{F}(X)$ valitaan kanta. Yleisemmin tällä kurssilla osoitetaan kuinka äärellisulotteisen lineaarialgebran kysymykset voidaan palauttaa ratkaistavaksi kurssin Lineaarialgebra ja matriisilaskenta I menetelmillä.

Luku 7

Vektoriavaruudet ja lineaarikuvaukset

Lineaarialgebra ja matriisilaskenta I kurssilla havaittiin, että sarakeavaruuden $\mathbb{R}^{n\times 1}$ ja matriisiavaruuksien $\mathbb{R}^{m\times n}$ alkioita voitiin laskea yhteen ja kertoa vakiolla samoilla periaatteilla kuin avaruuden $\mathbb{R}^n$. Itseasiassa sarakeavaruuden $\mathbb{R}^{n\times 1}$ ja avaruuden $\mathbb{R}^n$ välillä ei tehty muuta kuin merkinnällinen ero. Avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^{n\times 1}$ alkiosta puhuttiin vektoreina, mutta matriisien $\mathbb{R}^{m\times n}$ kohdalla ei tälläistä käsitettä käytetty.

Luvussa 3 havaittiin myös, että avaruuksilla $\mathbb{R}^n$ ja $\mathbb{R}^{n\times 1}$ on osajoukkoja, jotka ovat suljettuja yhteenlaskun ja skalaarikertolaskun suhteen, eli niillä on samankaltaisia ominaisuuksia kuin avaruuksilla $\mathbb{R}^n$ ja $\mathbb{R}^{n\times 1}$.

Tässä luvussa määritellään yleisen vektoriavaruuden käsite, joka antaa oikean kontekstin puhua kaikista näistä ilmiöistä. Paljastuu, että $\mathbb{R}^n$, $\mathbb{R}^{n\times 1}$ ja $\mathbb{R}^{m\times n}$ sekä avaruuksien $\mathbb{R}^{n\times 1}$ aliavaruudet ovat kaikki vektoriavaruuksia. Luvusta 3 tutut käsitteet kuten vapaus, virittäminen, kanta ja dimensio yleistyvät kaikille vektoriavaruuksille. Näitä asioita käsitellään luvussa 8.

Vektoriavaruuksien välisiä luonnollisia kuvauksia kutsutaan lineaarikuvauksiksi. Luvusta 6 tuttu kuvaus $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times n}$, $x \mapsto Ax$, eli matriisilla kertominen, on esimerkki lineaarikuvauksesta. Paljastuu, että kaikki äärellisulotteisten vektoriavaruuksien väliset kuvaukset voidaan esittää matriisilla kertomisen avulla. Näitä asioita puolestaan käsitellään tarkasti luvussa 9.

7.1 Reaaliset vektoriavaruudet

Aloitetaan yleisen vektoriavaruuden määritelmästä. Haluttaessa olla hieman tarkempia tässä yhteydessä puhutaan reaalisista vektoriavaruuksista, koska määritelmässä käytetyt skalaarit ovat reaalilukuja,

Määritelmä 7.1.1. *Kolmikko* $(V, +, \cdot)$ *on* (reaalinen) vektoriavaruus*, jos* V *on epätyhjä joukko ja* $+\colon V \times V \to V$ *sekä* $\cdot\colon \mathbb{R} \times V \to V$ *ovat sellaisia funktiota, että*

1. $u + v = v + u$ *kaikilla* $u, v \in V$*,*

2. $(u+v)+w = u+(v+w)$ *ja* $(ab)v = a(bv)$ *kaikilla* $u, v, w \in V$ *ja* $a, b \in \mathbb{R}$*,*

3. on olemassa sellainen alkio $0 \in V$*, että* $v + 0 = v$ *kaikilla* $v \in V$*,*

4. jokaisella $v \in V$ *on olemassa sellainen alkio* $w \in V$*, että* $v + w = 0$*,*

5. $1v = v$ *jokaisella* $v \in V$*,*

6. $a(u+v) = au+av$ *ja* $(a+b)v = av+bv$ *kaikilla* $a, b \in \mathbb{R}$ *ja* $u, v \in V$*.*

Mikäli kolmikko $(V, +, \cdot)$ *on vektoriavaruus, niin joukon* V *alkiota kutsutaan* vektoreiksi*.*

Useampi huomio on paikallaan.

Huomautus 7.1.2. *Kun yleisen vektoriavaruuden määritelmän näkee ensimmäistä kertaa, herää kysymys, että onko määritelmässä mitään järkeä, sillä nämä ominaisuudet ovat aina voimassa, kun vektoreilla lasketaan. Ero aiempaan on siinä, että aiemmin vektoreiden yhteenlasku ja skalaarikertolasku on määritelty jollain kaavalla, jonka avulla voidaan tarkistaa, että nämä ominaisuudet ovat voimassa. Määritelmässä 7.1.1 puolestaan annetaan vaatimus, mitä ominaisuuksia sellaisten (jollain kaavoilla määriteltyjen) funktioiden tulee toteuttaa, joita halutaan kutsua vektoreiden yhteenlaskuksi ja skaarikertolaskuksi. Merkinnät* $+$ *ja* $\cdot$ *on valittu juuri sitä seikkaa silmällä pitäen, että näillä laskutoimituksilla laskeminen tapahtuu juuri samoin kuin tutuissa tapauksissa.*

Huomautus 7.1.3. *Ehtoja* (1)-(6) *kutsutaan myös vektoriavaruuksien* aksioomiksi*.*

Huomautus 7.1.4. *Yleensä ei merkitä* $(V, +, \cdot)$*, vaan sanotaan lyhyesti, että* V *on vektoriavaruus, jonka yhteenlasku ja skaalarikertolasku ovat* $+\colon V \times V \to V$ *ja* $\cdot\colon \mathbb{R} \times V \to V$*. Funkioita* $+$ *ja* $\cdot$ *kutsutaan myös* vektoriavaruuden V laskutoimituksiksi*. Laskutoimitusta* $+\colon V \times V \to V$ *kutsutaan vektoriavaruuden* V yhteenlaskuksi *ja laskutoimitusta* $\cdot\colon \mathbb{R} \times V \to V$ skalaarikertolaskuksi*.*

Huomautus 7.1.5. *Vektoria* $0 \in V$ *kutsutaan tilanteesta riippuen* yhteenlaskun neutraalialkioksi*,* avaruuden V nollavektoriksi *tai* avaruuden V origoksi*. Ehdossa* (4) *vektoria* w *kutsutaan vektorin* v vastavektoriksi*. Ehdon* (6) *perusteella tämä vastavektori* $-v$ *on vektori* $(-1)v$*.*

Huomautus 7.1.6. *Formaalisti vektoreiden* $v_1, \ldots, v_n \in V$ *summaa pitäisi merkitä*

$$(\cdots((v_1 + v_2) + v_3) + \cdots + v_{n-1}) + v_n.$$

Ehdosta (2) *kuitenkin seuraa, että vektoreiden* $v_1, \ldots, v_n$ *summausjärjestystä voidaan muuttaa. Koska kaikki summausjärjestykset antavat saman lopputuloksen, ei järjestystä tarvitse merkitä suluilla vaan voidaan merkitä lyhyesti* $v_1 + \cdots + v_n$ *kuten sarakevektorienkin tapauksessa tehtiin. Lisäksi ehdon* (1) *perusteella summan termit* $v_1 + \cdots + v_n$ *voi järjestää mielivaltaiseen järjestykseen summan siitä muuttumatta.*

Huomautus 7.1.7. *Kuten aiemmin vektoreiden* $u, v \in V$ *summaa merkitään* $u + v$ *eikä* $+(u, v)$*, kuten olisi formaalisti oikein. Vastaavasti skalaarikertolaskun merkkiä* $\cdot$ *ei käytetä konkreettisissa laskuissa.*

Kuten huomautuksessa 7.1.6 todettiin, vektoriavaruuden aksioomat sallivat vektoreiden summat ja yleisemmin lineaarikombinaatiot aivan kuten sarakeavaruuksien tapauksessa. Kirjataan tämä vektoriavaruuksien perusmääritelmä vielä tarkasti.

Määritelmä 7.1.8. *Vektoriavaruuden V vektori $v \in V$ on* vektoreiden $v_1, \dots, v_k \in V$ lineaarikombinaatio*, jos on olemassa sellaiset luvut $a_1, \dots, a_k \in \mathbb{R}$, että*

$$v = a_1 v_1 + \cdots + a_k v_k.$$

7.1.1 Perusesimerkkejä vektoriavaruuksista

Seuraaviin esimerkkeihin on koottu tuttuja avaruuksia, jotka ovat vektoriavaruuksia määritelmän 7.1.1 mielessä. Lisää esimerkkejä vektoriavaruuksista, kuten polynomien avaruudet ja yleisemmät funktioavaruudet, annetaan luvussa 7.5.

Esimerkki 7.1.9. *Sarakeavaruus $\mathbb{R}^{n\times 1}$ on vektoriavaruus sen tavallisilla laskutoimituksilla $+\colon \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ ja $\cdot\colon \mathbb{R} \times \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$, jotka on määritelty kaavoilla*

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

ja $\cdot\colon \mathbb{R} \times \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$,

$$a \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} ax_1 \\ \vdots \\ ax_n \end{bmatrix}$$

kaikilla

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$$

ja $a \in \mathbb{R}$.

Esimerkki 7.1.10. *Sarakeavaruuden $\mathbb{R}^{n\times 1}$ alivaruus $V \subset \mathbb{R}^{n\times 1}$ on vektoriavaruus.*

Esimerkki 7.1.11. *Matriisiavaruus $\mathbb{R}^{m\times n}$ varustettuna matriisien yhteenlaskulla ja skalaarikertolaskulla on vektoriavaruus. Vektoriavaruuden $\mathbb{R}^{m\times n}$ nolla-alkio on matriisi $0 \in \mathbb{R}^{m\times n}$, joka kaikki kertoimet ovat nollia. Matriisin $A = [a_{ji}]$ vastamatriisi (eli vastavektori) on matriisi $-A = [-a_{ji}]$.*

Esimerkki 7.1.12. *Avaruus $\mathbb{R}^n$ on vektoriavaruus tutuilla laskutoimituksilla $+\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ ja $\cdot\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. Erityisesti $\mathbb{R}$ eli $\mathbb{R}^1$ on (reaalinen) vektoriavaruus. Samoin origon kautta kulkevat suorat ja tasot ovat vektoriavaruuksia.*

7.2 Lineaarikuvaukset

Vektoriavaruuksien väliset luonnolliset kuvaukset ovat sellaisia, jotka kunnioittavat yhteenlaskua ja skalaarikertolaskua. Tämä tarkoittaa seuraavaa.

Määritelmä 7.2.1. *Olkoot V ja W vektoriavaruuksia. Kuvaus $f\colon V \to W$ on* lineaarikuvaus avaruudelta V avaruudelle W, *jos kaikilla $v, v' \in V$ ja $a \in \mathbb{R}$ pätee sekä*

1. *$f(v + v') = f(v) + f(v')$ ja*
2. *$f(av) = af(v)$.*

Muutama kommentti on taas paikallaan.

Huomautus 7.2.2. *Määritelmässä käytettiin yleistä tapaa, että yhteenlaskun ja skalaarikertolaskun kohdalla ei tehty eroa suoritetaanko lasku avaruudessa V vai avaruudessa W. Ehdossa* (1) *yhteenlasku $v + v'$ suoritetaan avaruudessa V ja $v + v' \in V$. Toisaalta $f(v)$ ja $f(v')$ ovat avaruuden W vektoreita, joten yhteenlasku $f(v) + f(v')$ suoritetaan avaruudessa W. Vastaavat huomiot voidaan tehdä ehdossa* (2)*.*

Huomautus 7.2.3. *Olkoon $f\colon V \to W$ lineaarikuvaus vektoriavaruuksien V ja W välillä. Koska jokaisella $v \in V$ pätee $f(0v) = 0f(v) = 0$, niin havaitaan, että erityisesti $f(0) = 0$. Vastaavasti $f(-v) = f((-1)v) = (-1)f(v) = -f(v)$ jokaisella $v \in V$. Näin ollen jokainen lineaarikuvaus kuvaa nollavektorin nollavektoriksi ja vastavektorin vastavektoriksi.*

Huomautus 7.2.4. *Lineaarikuvauksen ehdot 1 ja 2 voidaan yhdistää yhdeksi ehdoksi: kaikilla $v, v' \in V$ ja $a \in \mathbb{R}$ pätee $f(av + v') = af(v) + f(v')$.*

Kuvaus, joka on annettu kaavalla, osoitetaan lineaarikuvaukseksi tarkistamalla lineaarikuvauksen määritelmän ehdot. Yleensä tämä tehdään tarkistamalla ehdot samanaikaisesti kuten edellisessä huomautuksessa todetaan.

Esimerkki 7.2.5. *Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ kuvaus $(x, y) \mapsto (4x+2y, y-x, x+y)$. Osoitetaan, että f on lineaarikuvaus. Olkoot $(x, y), (x', y') \in \mathbb{R}^2$ ja $a \in \mathbb{R}$. Tällöin*

$$\begin{aligned}
f(a(x,y) + (x',y')) &= f(ax + x', ay + y') \\
&= (4(ax + x') + 2(ay + y'), (ay + y') - (ax + x'), ax + x' + ay + y') \\
&= (a(4x + 2y) + 4x' + 2y', a(y - x) + (y' - x'), a(x + y) + (x' + y')) \\
&= a(4x + 2y, y - x, x + y) + (4x' + 2y', y' - x', x' + y') \\
&= af(x,y) + f(x',y').
\end{aligned}$$

Kuvaus f on siis lineaarinen.

7.2.1 Kaksi perusesimerkkiä lineaarikuvauksista

Annetaan nyt kaksi esimerkkiä tämän kurssin kannalta tärkeistä lineaarikuvauksista: lineaarikuvaukset sarakeavaruuksien välillä ja sarakeavaruudesta yleiseen vektoriavaruuteen.

Tärkein esimerkki lineaarikuvauksesta on matriisien yhteydessä esitelty lineaarikuvaus $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$, $x \mapsto Ax$, missä $A \in \mathbb{R}^{m\times n}$. Kirjataan tämä ensimmäiseksi esimerkiksi.

Esimerkki 7.2.6. *Olkoon $A \in \mathbb{R}^{m\times n}$ ja tarkastellaan kuvausta $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times n}$, $x \mapsto Ax$. Matriisitulon perusominaisuuksien perustella kaikilla $x, y \in \mathbb{R}^{n\times 1}$ ja $a \in \mathbb{R}$ pätee sekä*

$$f_A(x+y) = A(x+y) = Ax + Ay = f_A(x) + f_A(y) \tag{7.1}$$

että

$$f_A(ax) = A(ax) = aAx = af_A(x).$$

Näin ollen f_A on lineaarikuvaus vektoriavaruudelta $\mathbb{R}^{n\times 1}$ vektoriavaruudelle $\mathbb{R}^{m\times 1}$. Huomaa, että kaavassa (7.1) *yhteenlasku $x+y$ suoritetaan avaruudessa $\mathbb{R}^{n\times 1}$ ja yhteenlasku $f_A(x) + f_A(y)$ avaruudessa $\mathbb{R}^{m\times 1}$.*

Kirjataan jatkoa varten edellisen esimerkin kuvaus f_A yleiseksi merkinnäksi.

Merkintä 7.2.7. *Olkoon $A \in \mathbb{R}^{m\times n}$. Lineaarikuvaus $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ on kuvaus $x \mapsto Ax$.*

Toinen esimerkki saadaan sarakeavaruuden ja vektoriavaruuden välisistä lineaarikuvauksista. Esimerkin tulos jätetään harjoitustehtäväksi.

Esimerkki 7.2.8. *Olkoon V vektoriavaruus ja olkoot $v_1, \ldots, v_n \in V$. Tällöin kuvaus $f_{(v_1,\ldots,v_n)}\colon \mathbb{R}^{n\times 1} \to V$,*

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto x_1v_1 + \cdots + x_nv_n$$

on lineaarikuvaus.

Kirjataan tämänkin esimerkin kuvauksen merkintä $f_{(v_1,\ldots,v_n)}$ yleiseksi merkinnäksi jatkoa varten

Merkintä 7.2.9. *Olkoon V vektoriavaruus ja $(v_1, \ldots, v_n)$ jono avaruuden V vektoreita. Lineaarikuvaus $f_{(v_1,\ldots,v_n)}\colon \mathbb{R}^{n\times 1} \to V$ on kuvaus*

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto x_1v_1 + \cdots + x_nv_n.$$

Huomautus 7.2.10. *Mikäli V on sarakeavaruus $\mathbb{R}^{m\times 1}$ ja $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$, ovat kuvaukset f_A ja $f_{(v_1,\ldots,v_n)}$ (tietysti) sama kuvaus kurssin* Lineaarialgebra ja matriisilaskenta I *perusteella.*

Huomautus 7.2.11. *Luvussa 9.4 osoitetaan, että jokainen lineaarikuvaus $f\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ saadaan matriisitulosta eli $f = f_A$ jollain $A \in \mathbb{R}^{m\times n}$, ja että jokainen lineaarikuvaus $f\colon \mathbb{R}^{n\times 1} \to V$ on lineaarikuvaus $f = f_{(v_1,\ldots,v_n)}$ jollain avaruuden V jonolla $(v_1, \ldots, v_n)$.*

7.2.2 Lisää esimerkkejä lineaarikuvauksista

Matriisitulosta saadaan myös haastavampia esimerkkejä.

Esimerkki 7.2.12. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi ja tarkastellaan kuvausta $f\colon \mathbb{R}^{n\times k} \to \mathbb{R}^{m\times k}$, $B \mapsto AB$. Kuten edellisessäkin esimerkissä matriisiavaruudet $\mathbb{R}^{n\times k}$ ja $\mathbb{R}^{m\times k}$ ovat vektoriavaruuksia. Lisäksi matriisitulon ominaisuuksien perusteella kaikilla $B, C \in \mathbb{R}^{n\times k}$ ja $a \in \mathbb{R}$ pätee*

$$f(B + C) = A(B + C) = AB + AC = f(B) + f(C)$$

ja

$$f(aB) = A(aB) = aAB = af(B).$$

Kuvaus f on siis lineaarikuvaus avaruudelta $\mathbb{R}^{n\times k}$ avaruudelle $\mathbb{R}^{m\times k}$.

Toinen tärkeä luokka lineaarikuvauksia ovat *lineaariset funktionaalit* eli lineaarikuvaukset vektoriavaruudelta reaaliluvuille. Huomaa, että tässä hyödynnetään tulkintaa, että $\mathbb{R}$ on itsessään vektoriavaruus.

Esimerkki 7.2.13. *Useamman muuttujan analyysissä*[1] *(eli vektorianalyysissä) funktiota $\mathrm{pr}_i\colon \mathbb{R}^n \to \mathbb{R}$, $(x_1, \ldots, x_n) \mapsto x_i$, missä $i \in \{1, \ldots, n\}$, kutsutaan* avaruuden $\mathbb{R}^n$ i:nneksi koordinaattiprojektioksi. *Osoitetaan, että funktio pr_i on lineaarikuvaus.*

Olkoot $x = (x_1, \ldots, x_n)$ ja $y = (y_1, \ldots, y_n)$. Tällöin $\mathrm{pr}_i(x) = x_i$, $\mathrm{pr}_i(y) = y_i$ ja

$$\mathrm{pr}_i(x + y) = \mathrm{pr}_i(x_1 + y_1, \ldots, x_n + y_n) = x_i + y_i = \mathrm{pr}_i(x) + \mathrm{pr}_i(y).$$

Lisäksi jokaisella $a \in \mathbb{R}$ pätee

$$\mathrm{pr}_i(ax) = \mathrm{pr}_i(ax_1, \ldots, ax_n) = ax_i = a\mathrm{pr}_i(x).$$

Näin ollen pr_i on lineaarikuvaus.

Myös keskiarvo voidaan tulkita lineaarikuvauksena. Seuraavan esimerkin yksityiskohdat jätetään harjoitustehtäväksi.

Esimerkki 7.2.14. *Keskiarvo funktio $\kappa\colon \mathbb{R}^n \to \mathbb{R}$, $(x_1, \ldots, x_n) \mapsto \frac{1}{n}(x_1 + \cdots + x_n)$, on lineaarikuvaus.*

[1] Eikä pelkästään siellä.

Luvussa 4 sanottiin pistetulon olevan lineaarinen molempien argumenttiensa suhteen, ja luvussa 5 käytettiin terminogiaa, että tilavuusmuoto on lineaarinen argumenttiensa suhteen. Nämä pistetulon ja tilavuusmuodon ominaisuudet vastaavat täsmälleen lineaarikuvaukselta vaadittuja ominaisuuksia, kuten seuraavat esimerkit osoittavat.

Esimerkki 7.2.15. *Olkoon $w \in \mathbb{R}^{n\times 1}$. Määritellään funktio $g_w\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$ kaavalla $v \mapsto v \cdot w$. Osoitetaan, että g_w on lineaarikuvaus.*

Olkoot $v, v' \in \mathbb{R}^{n\times 1}$ ja $a \in \mathbb{R}$. Tällöin

$$g_w(v + v') = (v + v') \cdot w = v \cdot w + v' \cdot w = g_w(v) + g_w(v')$$

ja

$$g_w(av) = (av) \cdot w = a(v \cdot w) = ag_w(v).$$

Näin ollen g_w on lineaarikuvaus.

Vastaavasti osoitetaan, että funktio $\widetilde{g}_w\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$, $v \mapsto w \cdot v$, on lineaarikuvaus.

Esimerkki 7.2.16. *Olkoon $\omega\colon \mathbb{R}^{n\times 1} \times \cdots \times \mathbb{R}^{n\times 1} \to \mathbb{R}$ multilineaarikuvaus ja olkoot $w_1, \ldots, w_n \in \mathbb{R}^{n\times 1}$. Tällöin jokaisella $i \in \{1, \ldots, n\}$ funktio $f\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$, $v \mapsto \omega(w_1, \ldots, w_{i-1}, v, w_{i+1}, \ldots, w_n)$, on lineaarikuvaus.*

Erityisesti funktio $D_i\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$,

$$v \mapsto \det \begin{bmatrix} w_1 & \cdots & w_{i-1} & v & w_{i+1} & \cdots & w_n \end{bmatrix},$$

on lineaarikuvaus.

7.2.3 Lineaarikuvaukset ja lineaarikombinaatiot

Lineaarikuvausten perusominaisuus on, että lineaarikuvaus kuvaa lineaarikombinaatiot lineaarikombinaatioiksi. Tämä ominaisuus itseasiassa karakterisoi lineaarikuvaukset. Kirjataan tämä tulos lemmaksi.

Lemma 7.2.17. *Olkoot V ja W vektoriavaruuksia ja $f\colon V \to W$ kuvaus. Tällöin f on lineaarikuvaus, jos ja vain jos jokaisella $k \in \mathbb{Z}_+$ ja kaikilla $v_1, \ldots, v_k \in V$ ja $a_1, \ldots, a_k \in \mathbb{R}$ pätee*

$$f\left(a_1v_1 + \cdots + a_kv_k\right) = a_1f(v_1) + \cdots + a_kf(v_k).$$

Todistus. Oletetaan, että f on lineaarikuvaus ja osoitetetaan, että kaikilla $v_1, \ldots, v_k \in V$ ja $a_1, \ldots, a_k \in \mathbb{R}$ pätee

$$f\left(a_1v_1 + \cdots + a_kv_k\right) = a_1f(v_1) + \cdots + a_kf(v_k).$$

Todistetaan väite induktiolla. Jos $k = 1$, niin tällöin $f(a_1v_1) = a_1f(v_1)$ ehdon 2 perusteella. Jos $k = 2$, niin tällöin $f(a_1v_1+a_2v_2) = f(a_1v_1)+f(a_2v_2) = a_1f(v_1)+a_2f(v_2)$ ehtojen 1 ja 2 perusteella.

Oletetaan, että väite pätee luvulla $k \in \mathbb{Z}_+$. Olkoot $v_1, \ldots, v_{k+1} \in V$ vektoreita ja $a_1, \ldots, a_{k+1} \in \mathbb{R}$ lukuja. Merkitään $w = a_1v_1 + \cdots + a_kv_k \in V$. Tällöin

$$f(a_1v_1 + \cdots + a_{k+1}v_{k+1}) = f(a_1v_1 + \cdots + a_kv_k + a_{k+1}v_{k+1}) = f(w + a_{k+1}v_{k+1}).$$

Koska

$$f(w + a_{k+1}v_{k+1}) = f(w) + f(a_{k+1}v_{k+1}) = f(w) + a_{k+1}f(v_{k+1})$$

ehtojen 1 ja 2 perusteella ja

$$f(w) = f(a_1v_1 + \cdots + a_kv_k) = a_1f(v_1) + \cdots + a_kf(v_k)$$

induktio-oletuksen nojalla, niin

$$f(a_1v_1 + \cdots + a_{k+1}v_{k+1}) = a_1f(v_1) + \cdots + a_kf(v_k) + a_{k+1}f(v_{k+1}).$$

Tämä päättää induktioaskeleen todistuksen.

Oletetaan nyt, että kaikilla $v_1, \ldots, v_k \in V$ ja $a_1, \ldots, a_k \in \mathbb{R}$ pätee

$$f\left(a_1v_1 + \cdots + a_kv_k\right) = a_1f(v_1) + \cdots + a_kf(v_k).$$

Olkoot $v, v' \in V$ ja $a \in \mathbb{R}$. Tällöin oletuksen nojalla

$$f(v + v') = f(v) + f(v')$$

ja

$$f(av) = af(v).$$

Näin ollen f on lineaarikuvaus. □

7.2.4 Lineaarikuvausten yhdistäminen

Lineaarikuvausten tärkeä perusominaisuus on, että kahden lineaarikuvauksen yhdistetty kuvaus on lineaarikuvaus. Todistus on suoraviivainen harjoitus yhdistetyn kuvauksen ja lineaarikuvauksen määritelmään liittyen, joten se jätetään harjoitustehtäväksi.

Lause 7.2.18. *Olkoot $f\colon V \to W$ ja $g\colon W \to U$ lineaarikuvauksia avaruudelta V avaruudelle W ja avaruudelta W avaruudelle U. Tällöin yhdistetty kuvaus $g \circ f\colon V \to U$, $v \mapsto g(f(v))$, on lineaarikuvaus.*

Lineaarikuvauksille pätee myös, että bijektiivisen lineaarikuvauksen käänteiskuvaus on lineaarikuvaus.

Lause 7.2.19. *Olkoon $f\colon V \to W$ bijektiivinen lineaarikuvaus. Tällöin käänteiskuvaus $f^{-1}\colon W \to V$ on lineaarikuvaus.*

Todistus. Olkoot $w, w' \in W$ ja $a \in \mathbb{R}$. Osoitetaan, että

$$f^{-1}(aw + w') = af^{-1}(w) + f^{-1}(w').$$

Käänteiskuvauksen määritelmästä ja kuvauksen f lineaarisuudesta seuraa, että

$$f(f^{-1}(aw + w')) = aw + w' = af(f^{-1}(w)) + f(f^{-1}(w')) = f(af^{-1}(w) + f^{-1}(w')).$$

Koska f on bijektio, niin se on erityisesti injektio ja siten

$$f^{-1}(aw + w') = af^{-1}(w) + f^{-1}(w').$$

□

7.3 Aliavaruudet

Määritellään nyt vektoriavaruuden aliavaruuden käsite.

Määritelmä 7.3.1. *Osajoukko $W \subset V$ on vektoriavaruuden V* aliavaruus*, jos*

1. *avaruuden V nollavektori $0 \in V$ kuuluu joukkoon W,*
2. *kaikilla $v, w \in W$ pätee $v + w \in W$ ja*
3. *kaikilla $a \in \mathbb{R}$ ja $v \in W$ pätee $av \in W$.*

Tämä aliavaruuden määritelmä vastaa täysin luvussa 3 määritelmää 3.2.1. Ainoa ero on, että W on vektoriavaruuden V osajoukko, eikä sarakeavaruuden $\mathbb{R}^{n\times 1}$ osajoukko. Kaikki tutut aliavaruudet, kuten $\{0\}$ ja koko avaruus V, ovat siis aliavaruuksia myös tässä mielessä.

Seuraavan esimerkin osajoukko $W \subset \mathbb{R}^{2\times 3}$ on helppo osoittaa aliavaruudeksi suoraan määritelmästä. Yksityiskohdat jätetään harjoitustehtäväksi.

Esimerkki 7.3.2. *Olkoon*

$$W = \left\{ \begin{bmatrix} x & 0 & y \\ 0 & z & 0 \end{bmatrix} \in \mathbb{R}^{2\times 3} \colon x, y, z \in \mathbb{R} \right\}.$$

Tällöin W on vektoriavaruuden $\mathbb{R}^{2\times 3}$ aliavaruus.

Samalla perustelulla kuin luvussa 3 voidaan osoittaa, että annettujen vektoreiden kaikkien lineaarikombinaatioiden joukko on aliavaruus. Koska todistus on sama, se jätetään harjoitustehtäväksi.

Lemma 7.3.3. *Olkoon V vektoriavaruus ja $v_1, \ldots, v_k \in V$ vektoreita. Tällöin osajoukko*

$$\{a_1 v_1 + \cdots + a_k v_k \in V \colon a_1, \ldots, a_k \in \mathbb{R}\} \subset V$$

on avaruuden V aliavaruus.

Määritelmä 7.3.4. *Vektoreiden $v_1, \ldots, v_k \in V$ kaikkien lineaarikombinaatioiden joukkoa*

$$\mathrm{Sp}(v_1, \ldots, v_k) = \{a_1 v_1 + \cdots + a_k v_k \in V \colon a_1, \ldots, a_k \in \mathbb{R}\}$$

kutsutaan vektoreiden $v_1, \ldots, v_k$ virittämäksi aliavaruudeksi.

Virittäjiä käyttämällä on helppo todeta, että esimerkiksi origon kautta kulkevat suorat ovat aina aliavaruuksia.

Esimerkki 7.3.5. *Olkoon V vektoriavaruus ja $v \in V$ nollasta poikkeava vektori, eli $v \neq 0$. Tällöin*

$$W = \{tv \in V \colon t \in \mathbb{R}\}$$

on avaruuden V aliavaruus, origon kautta kulkeva vektorin v suuntainen suora.

Huomautus 7.3.6. *Koska* $\mathbb{R}^n$ *on vektoriavaruus, niin lineaarikombinaatioiden avulla on helppo tunnistaa tutut osajoukot, kuten origon kautta kulkevat suorat ja tasot, aliavaruuksiksi. Nämä aliavaruudet vaikuttavat hyvin samankaltaisilta kuin sarakeavaruuden* $\mathbb{R}^{n\times 1}$ *aliavaruudet. Tämä ei ole sattumaa.*

Kuten luvussa 3, myös tässä yleisemmässä vektoriavaruuksien tilanteessa osajoukko $W \subset V$ voidaan todistaa aliavaruudeksi käyttäen aliavaruuskriteeriota. Todistus sivuutetaan, koska se on täysin analoginen lemman 3.2.8 todistukseen.

Lemma 7.3.7. *Olkoon* V *vektoriavaruus. Epätyhjä osajoukko* $W \subset V$ *on aliavaruus, jos ja vain jos* $av + w \in W$ *kaikilla* $v, w \in W$ *ja* $a \in \mathbb{R}$.

Esimerkki 7.3.8. *Olkoon* $x \in \mathbb{R}^{n\times 1}$ *vektori ja määritellään matriisiavaruuden* $\mathbb{R}^{m\times n}$ *osajoukko*

$$W = \{A \in \mathbb{R}^{m\times n} \colon Ax = 0\}.$$

Osoitetaan aliavaruuskriteerion avulla, että W *on avaruuden* $\mathbb{R}^{m\times n}$ *aliavaruus. Koska nollamatriisille* $0 \in \mathbb{R}^{m\times n}$ *pätee* $0x = 0$, *niin joukko* W *on epätyhjä. Olkoot nyt* $A, B \in W$ *ja* $a \in \mathbb{R}$. *Tällöin matriisille* $aA + B$ *pätee*

$$(aA + B)x = a(Ax) + Bx = a \cdot 0 + 0 = 0,$$

eli $aA + B \in W$. *Näin ollen* W *on avaruuden* $\mathbb{R}^{3\times 2}$ *aliavaruus*[2].

7.3.1 Esimerkki: Matriisin riviavaruus

Luvussa 3 matriisille $A \in \mathbb{R}^{m\times n}$ määriteltiin sarakeavaruus $\mathrm{Col}(A) \in \mathbb{R}^{m\times 1}$ ja nolla-avaruus $\mathrm{Null}(A) \in \mathbb{R}^{n\times 1}$. Kolmas matriiseihin liittyvä aliavaruus on matriisin riviavaruus, joka on riviavaruuden $\mathbb{R}^{1\times n}$ aliavaruus.

Määritelmä 7.3.9. *Matriisiavaruutta* $\mathbb{R}^{1\times n}$ *kutsutaan* riviavaruudeksi.

Koska matriisiavaruudet $\mathbb{R}^{m\times n}$ ovat vektoriavaruuksia, niin erityisesti myös $\mathbb{R}^{1\times n}$ on vektoriavaruus. Näin ollen voidaan tarkastella sen aliavaruuksia.

Määritelmä 7.3.10. *Matriisin*

$$A = \begin{bmatrix} - & w_1 & - \\ & \vdots & \\ - & w_m & - \end{bmatrix} \in \mathbb{R}^{m\times n}$$

riviavaruus $\mathrm{Row}(A) \subset \mathbb{R}^{1\times n}$ *on aliavaruus*

$$\mathrm{Row}(A) = \mathrm{Sp}(w_1, \ldots, w_m) = \{y_1 w_1 + \cdots + y_m w_m \in \mathbb{R}^{1\times n} \colon y_1, \ldots, y_m \in \mathbb{R}\},$$

missä $w_1, \ldots, w_n \in \mathbb{R}^{1\times n}$ *ovat matriisin* A *rivit.*

[2] Aliavaruutta W kutsutaan kirjallisuudessa vektorin x *annihilaattoriksi.*

Matriisin $A \in \mathbb{R}^{m\times n}$ riviavaruus ei siis ole kummankaan sarakeavaruuden $\mathbb{R}^{m\times 1}$ tai $\mathbb{R}^{n\times 1}$ aliavaruus vaan vektoriavaruuden $\mathbb{R}^{1\times n}$ aliavaruus. Kuten sarakeavaruuden $\mathrm{Col}(A)$ tapauksessa, myös aliavaruus $\mathrm{Row}(A)$ voidaan ilmaista matriisitulon avulla. Käsitellään asiaa esimerkin avulla ennen yleistä tulosta.

Esimerkki 7.3.11. *Olkoon*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 5}.$$

Tällöin matriisin A rivit ovat $w_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}$ ja $w_2 = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \end{bmatrix}$.

Vektori $w \in \mathbb{R}^{1\times 5}$ kuuluu aliavaruuteen $\mathrm{Row}(A)$, jos ja vain jos on olemassa sellaiset $y_1, y_2 \in \mathbb{R}$, että

$$w = y_1 w_1 + y_2 w_2,$$

missä

$$y_1 w_1 + y_2 w_2 = \begin{bmatrix} y_1 + 2y_2 & y_1 + 2y_2 & y_1 + 2y_2 & y_1 + 2y_2 & y_1 + 2y_2 \end{bmatrix}.$$

Toisaalta matriisitulon laskusääntöjen mukaan

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} y_1 + 2y_2 & y_1 + 2y_2 & y_1 + 2y_2 & y_1 + 2y_2 & y_1 + 2y_2 \end{bmatrix}.$$

Näin ollen $w \in \mathrm{Row}(A)$, jos ja vain jos on olemassa sellainen rivivektori $\begin{bmatrix} w_1 & w_2 \end{bmatrix} \in \mathbb{R}^{1\times 2}$, että

$$w = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}$$

Kirjataan nyt yleinen tulos, joka antaa tulkinnan äskeiselle esimerkille.

Lause 7.3.12. *Olkoon $A \in \mathbb{R}^{m\times n}$. Tällöin*

$$\mathrm{Row}(A) = \{y^t A \in \mathbb{R}^{1\times n} \colon y \in \mathbb{R}^{m\times 1}\}.$$

Todistus. Olkoot $w_1, \ldots, w_m$ matriisin A rivit, eli

$$A = \begin{bmatrix} - & w_1 & - \\ & \vdots & \\ - & w_m & - \end{bmatrix}.$$

Matriisitulon määritelmän perusteella, jokaisella

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^{m\times 1}$$

pätee

$$y_1w_1 + \cdots + y_mw_m = \begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix} \begin{bmatrix} - & w_1 & - \\ & \vdots & \\ - & w_m & - \end{bmatrix} = y^tA.$$

Näin ollen $w \in \mathrm{Row}(A)$, eli $w = y_1w_1 + \cdots + y_mw_m$ jollain $y_1, \ldots, y_m \in \mathbb{R}$, jos ja vain jos $w = y^tA$, jollain $y \in \mathbb{R}^{m\times 1}$. $\square$

Huomautus 7.3.13. *Matriisin $A \in \mathbb{R}^{m\times n}$ riviavaruuden $\mathrm{Row}(A)$ voisi esittää myös matriisin A transpoosin A^t sarakeavaruuden avulla. Tämä seuraa havainnosta, että $y^tA = (A^ty)^t$ jokaisella $y \in \mathbb{R}^{m\times 1}$. Näin ollen $w \in \mathrm{Row}(A)$ jos ja vain jos $w^t \in \mathrm{Col}(A^t)$.*

Huomautus 7.3.14. *Matriisin $A \in \mathbb{R}^{m\times n}$ rivi ja sarakeavaruuden välillä on toinenkin yllättävä yhteys, nimittäin pätee, että $\dim \mathrm{Row}(A) = \dim \mathrm{Col}(A)$. Tämän voi todistaa joko elementaaristi palauttamalla matriisi A supistettuun porrasmuotoon tai abstraktimmin. Tähän palataan myöhemmin.*

7.3.2 Aliavaruus on vektoriavaruus

On teoreettisesti tärkeä huomio, että aliavaruus on vektoriavaruus itsessään. Tämä perustuu havaintoon, että vektoriavaruuden $(V, +, \cdot)$ laskutoimitusten $+\colon V \times V \to V$ ja $\cdot\colon \mathbb{R} \times V \to V$ rajoittumat joukkoon, eli funktiot $+|_{W\times W}\colon W \times W \to V$ ja $\cdot|_{\mathbb{R}\times W}\colon \mathbb{R} \times W \to V$ ovat aliavaruuden määritelmän nojalla sellaisia, että niiden maalijoukoksi voidaan rajoittaa avaruuden V sijasta aliavaruus W. Sanotaan, että aliavaruus W *perii yhteenlaskun ja skalaarikertolaskun avaruudelta* V. Kirjataan tämä huomio lauseeksi.

Lause 7.3.15. *Olkoon $(V, +, \cdot)$ vektoriavaruus ja $W \subset V$ aliavaruus. Tällöin W on vektoriavaruus niillä laskutoimituksilla, jotka W perii avaruudelta V.*

Todistus. Merkitään avaruuden V laskutoimitusten $+\colon V \times V \to V$ ja $\cdot\colon \mathbb{R} \times V \to V$ rajoittumia joukkoon W lyhyesti samoilla merkinnöillä. Koska jokainen joukon W alkio on myös joukon V alkio, niin selvästi nämä uudet laskutoimitukset toteuttavat määritelmän 7.1.1 ehdot. Selvästi vektoriavaruuden V nolla-alkio $0 \in V$ voidaan valita yhteenlaskun neutraalialkioksi avaruudessa W ja vektorin $v \in W$ vastavektori on sama vektori $-v$ kuin avaruudessa V. $\square$

Tehdään sama päättely vielä tutussa sarakeavaruuden tapauksessa.

Esimerkki 7.3.16. *Avaruuden $\mathbb{R}^{n\times 1}$ aliavaruus $V \subset \mathbb{R}^{n\times 1}$ on itsessään vektoriavaruus. Tämä havaitaan toteamalla, että ehdot 1, 2, 5 ja 6 seuraavat suoraan sarakeavaruuden $\mathbb{R}^{n\times 1}$ laskutoimitusten ominaisuuksista. Koska $0 \in V$, niin yhteenlaskun $+\colon V \times V \to V$ neutraalialkioksi voidaan (tietysti) valita tämä nollavektori 0. Ehto 3 siis toteutuu. Ehto 4 seuraa oletuksesta, että V on aliavaruus.*

Mainittakoon vielä lopuksi, että koska vektoriavaruuden aliavaruus on itsessään vektoriavaruus, niin voidaan kysyä onko lineaarikuvauksen rajoittuma[3] aliavaruuteen myös lineaarikuvaus. Näin todellakin on. Vaikka todistus on ainoastaan määritelmien sisällön ymmärtämistä, kirjataan tulos kuitenkin lauseeksi.

Lause 7.3.17. *Olkoon $f\colon V \to W$ lineaarikuvaus vektoriavaruudesta V vektoriavaruuteen W ja olkoon $U \subset V$ aliavaruus. Tällöin kuvaus*

$$f|_U \colon U \to W,\ u \mapsto f(u),$$

on lineaarikuvaus.

Todistus. Koska $U \subset V$, niin kaikilla $u, u' \in U$ ja $a \in \mathbb{R}$ pätee

$$(f|_U)(au + u') = f(au + u') = af(u) + f(u') = a(f|_U)(u) + (f|_U)(u').$$

Koska U on vektoriavaruus, niin $f|_U$ on vektoriavaruuksien välinen lineaarikuvaus. □

Huomautus 7.3.18. *Usein kuvauksen $f\colon V \to W$ lähtöjoukon rajoittamisen yhteydessä rajoitetaan myös kuvauksen maalijoukkoa rajoittuman kuvaan eli kuvauksen $f|_U \colon U \to W$ sijasta siirrytäänkin tarkastelemaan kuvausta $f|_U \colon U \to f(U)$. Koska molempia tapoja käytetään rinnakkain, ei ole suositeltavaa merkitä rajoittumaa lyhyesti symbolilla $f|_U$ edes tapauksessa $f|_U \colon U \to W$ vaan lähtö- ja maalijoukko on mielekästi aina ilmaista tarkasti. Yksi käytännön syy tähän on se, että lähtö- ja maalijoukon valinnalla on merkitystä esimerkiksi kuvauksen injektiivisyyden ja surjektiivisuuden kannalta.*

7.4 Lineaarikuvauksen kuva ja ydin ovat aliavaruuksia

Jokaiseen lineaarikuvaukseen liittyy kaksi luonnollista aliavaruutta: ydin ja kuva. Käsitteinä nämä vastaavat matriisin nolla-avaruutta ja sarakeavaruutta. Tätä yhteyttä käsitellään tarkemmin luvussa 9.5.

Määritelmä 7.4.1. *Vektoriavaruuksien V ja W välisen lineaarikuvauksen $f\colon V \to W$* ydin $\ker(f)$ *on joukko*

$$\ker(f) = \{v \in V \colon f(v) = 0\}$$

ja kuva $\operatorname{im}(f)$ *on joukko*

$$\operatorname{im}(f) = \{f(v) \in W \colon v \in V\}.$$

Huomautus 7.4.2. *Lineaarikuvauksen f ydin on siis origon alkukuva $f^{-1}(0)$.*

Lause 7.4.3. *Vektoriavaruuksien V ja W välisen lineaarikuvauksen $f\colon V \to W$ ydin $\ker(f)$ on avaruuden V aliavaruus ja kuva $\operatorname{im}(f)$ on avaruuden W aliavaruus.*

[3] Palautetaan mieleen, että kuvauksen $f\colon X \to Y$ rajoittuma osajoukkoon $A \subset X$ on kuvaus $f|_A \colon A \to Y$, joka on määritelty kaavalla $x \mapsto f(x)$.

Huomautus 7.4.4. *Ytimen* $\ker(f)$ *osoittaminen aliavaruudeksi on analoginen matriisin A nolla-avaruuden* $\text{Null}(A)$ *aliavaruudeksi osoittamisen kanssa. Vertaa tätä todistusta lemman 3.3.1 todistukseen. Matriisin A sarakeavaruus* $\text{Col}(A)$ *todettiin aliavaruudeksi luvussa 3.5 hyödyntämällä virittämisen käsitettä, jota meillä ei vielä tässä vaiheessa ole käytössä. Tämän vuoksi kuva* $\text{im}(f)$ *osoitetaan aliavaruudeksi suoraan määritelmästä. Vektoriavaruuden virittämistä käsitellään seuraavassa luvussa.*

Yleisemmin aliavaruuden kuva lineaarikuvauksessa on aliavaruus. Samoin aliavaruuden alkukuva on aliavaruus. Tämän yleisemmän tuloksen todistus on olennaisesti sama kuin lauseen 7.4.3, joten se jätetään harjoitustehtäväksi.

Lemma 7.4.5. *Olkoon $f\colon V \to W$ lineaarikuvaus vektoriavaruudesta V vektoriavaruuteen W ja olkoot $U \subset V$ ja $P \subset W$ aliavaruuksia. Tällöin*

$$f(U) = \{f(u) \in W : u \in U\} \subset W$$

ja

$$f^{-1}(P) = \{v \in V : f(v) \in P\} \subset V$$

ovet aliavaruuksia.

Näiden kommenttien jälkeen olemme valmiit todistamaan lauseen 7.4.3.

Lauseen 7.4.3 todistus. Osoitetaan ensin, että ydin $\ker(f)$ on avaruuden V aliavaruus. Käytetään osoittamiseen aliavaruuskriteeriota. Koska $f(0) = 0$, niin joukko $\ker(f)$ on epätyhjä. Olkoot nyt $v, v' \in \ker(f)$ ja $a \in \mathbb{R}$. Koska f on lineaarikuvaus ja $f(v) = f(v') = 0$, niin

$$f(av + v') = f(av) + f(v') = af(v) + f(v') = 0 + 0 = 0.$$

Näin ollen $\ker(f)$ on aliavaruus.

Osoitetaan nyt, että $\text{im}(f)$ on avaruuden W aliavaruus. Koska $f(0) = 0$, niin $0 \in \text{im}(f)$ ja joukko $\text{im}(f)$ on epätyhjä. Olkoot nyt $w, w' \in \text{im}(f)$ ja $a \in \mathbb{R}$. Osoitetaan, että $aw + w' \in \text{im}(f)$.

Joukon $\text{im}(f)$ määritelmän nojalla on olemassa sellaiset $v, v' \in V$, että $f(v) = w$ ja $f(v') = w'$. Tällöin kuvauksen f lineaarisuuden nojalla

$$f(av + v') = f(av) + f(v') = af(v) + f(v') = aw + w'.$$

Näin ollen $aw + w' \in \text{im}(f)$, joten $\text{im}(f)$ on aliavaruus. □

Kuvauksen ydin ja kuva on kiinnostavia siitä syystä, että ne karakterisoivat lineaarikuvauksen injektiivisyyden ja surjektiivisuuden.

Lause 7.4.6. *Olkoon $f\colon V \to W$ lineaarikuvaus avaruudesta V avaruuteen W. Tällöin*

1. *kuvaus f on injektio, jos ja vain jos* $\ker(f) = \{0\}$, *ja*
2. *kuvaus f on surjektio, jos ja vain jos* $\text{im}(f) = W$.

Todistus. Näistä jälkimmäinen väite on selvä suoraan määritelmästä. Riittää siis todistaa ensimmäinen väite.

Oletetaan ensin, että f on injektio. Koska $f(0) = 0$ ja f on injektio, niin $f^{-1}(0) = \{0\}$. Koska $\ker(f) = f^{-1}(0)$, niin $\ker(f) = \{0\}$.

Oletetaan nyt, että $\ker(f) = \{0\}$, ja osoitetaan, että f on injektio. Oletetaan, että $v, v' \in V$ ovat sellaisia, että $f(v) = f(v')$. Koska f on lineaarinen, niin $f(v - v') = f(v) - f(v') = 0$. Näin ollen $v - v' \in \ker(f)$, eli $v - v' = 0$. Näin ollen $v = v'$ ja kuvaus f on injektio. □

Huomautus 7.4.7. *Lauseen 7.4.6 paljastaa analogian lineaarikuvauksen f alkukuvajoukkojen $f^{-1}(w)$ rakenteen ja yhtälön $Ax = b$ ratkaisujoukon välillä. Kuten matriisiyhtälölle, lineaarikuvaukselle pätee seuraava ominaisuus. Olkoon $f\colon V \to W$ lineaarikuvaus ja $w \in \operatorname{im}(f)$. Olkoon myös $v \in V$ sellainen vektori, että $f(v) = w$. Tällöin*

$$f^{-1}(w) = \{v + v' \in V : v' \in \ker(f)\}. \tag{7.2}$$

Luvussa 3 esiteltyä merkintätapaa käyttäen voidaan siis kirjoittaa, että

$$f^{-1}(w) = v + \ker(f). \tag{7.3}$$

7.5 Esimerkki: Funktioavaruudet

Tarkastellaan nyt funktioavaruuksia motivoivana esimerkkinä vektoriavaruuksista.

7.5.1 Jatkuvien funktioiden avaruus $C(]a, b[)$

Konkreettisuuden vuoksi tarkastellaan aluksi differentiaalilaskennasta tuttuja esimerkkejä. Olkoon $]a, b[\subset \mathbb{R}$ reaaliakselin väli, missä $-\infty < a < b < \infty$. Olkoon $C(]a, b[)$ kaikkien jatkuvien funktioiden $]a, b[\to \mathbb{R}$ joukko eli

$$C(]a, b[) = \{f\colon]a, b[\to \mathbb{R} \mid f \text{ on jatkuva}\}.$$

Jatkuvia funktiota koskeva perustulos on, että jatkuvien funktioiden $f\colon]a, b[\to \mathbb{R}$ ja $g\colon]a, b[\to \mathbb{R}$ summafunkio

$$f + g\colon]a, b[\to \mathbb{R}, \quad x \mapsto f(x) + g(g) \tag{7.4}$$

on myös jatkuva. Esimerkki summafunktiosta kuvassa 7.1.

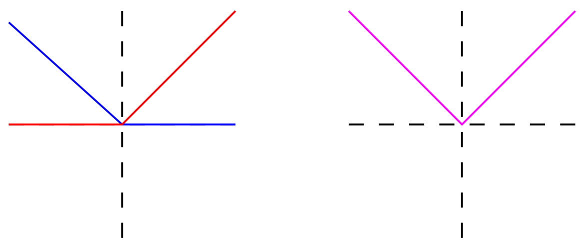
Myös jatkuvan funktion $f\colon]a, b[\to \mathbb{R}$ ja luvun $\lambda \in \mathbb{R}$ tulo, eli funktio

$$(\lambda f)\colon]a, b[\to \mathbb{R}, \quad x \mapsto \lambda f(x), \tag{7.5}$$

on myös jatkuva.

Nämä kaavat määrittelevät jatkuvien funktioiden $]a, b[\to \mathbb{R}$ joukossa yhteenlaskun ja skalaarikertolaskun

$$+\colon C(]a, b]) \times C(]a, b[) \to C(]a, b[),\ (f, g) \mapsto f + g,$$
$$\cdot\colon \mathbb{R} \times C(]a, b[) \to C(]a, b[),\ (a, f) \mapsto af.$$



Kuva 7.1: Funktioiden $f_1 \colon \mathbb{R} \to \mathbb{R}$, $x \mapsto (|x| - x)/2$, ja $f_2 \colon \mathbb{R} \to \mathbb{R}$, $x \mapsto (|x| + x)/2$, summa $f_1 + f_2$ on itseisarvofunktio $f \colon \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$. Kuvassa vasemmalla funktioiden f_1 ja f_2 graafit. Funktion f graafi oikealla.

Reaalilukujen yhteen- ja kertolaskun ominaisuudet siirtyvät suoraan laskutoimitusten $+$ ja $\cdot$ ominaisuuksiksi ja on helppo havaita, että vektoriavaruuden aksioomat (1)-(6) ovat voimassa.

Osoitetaan ensimmäinen ehto aksioomasta (2). Olkoot $f, g, h \in C(]a, b[)$. Tällöin jokaisella $x \in]a, b[$ pätee

$$((f+g)+h)(x) = (f+g)(x)+h(x) = f(x)+g(x)+h(x) = f(x)+(g+h)(x) = (f+(g+h))(x).$$

Näin ollen $(f + g) + h = f + (g + h)$. Muut vektoriavaruuden aksioomat tarkistetaan samoin. Kolmikko $(C(]a, b[), +, \cdot)$ on siis vektoriavaruus.

Huomautus 7.5.1. *Tässä esimerkissä on käsitelty konkreettisuuden vuoksi jatkuvia funktioita rajoitetulla välillä $]a, b[$. Samat argumentit pätevät kaikilla reaaliakselin väleillä $\Delta \subset \mathbb{R}$ ja erityisesti reaaliakselilla itsellään eli joukko $C(\Delta)$ varustettuna kaavoilla 7.4 ja 7.5 määritellyillä laskutoimituksilla on vektoriavaruus. Erityisesti $C([a, b])$, $C([a, b[)$, $C(]a, b])$, $C(]a, \infty[)$, $C([a, \infty[)$, $C(]-\infty, b[)$, $C(]-\infty, b])$ ja $C(\mathbb{R})$ ovat vektoriavaruuksia kaikilla $a < b$.*

7.5.2 Derivoituvien funktioiden avaruus

Olkoon $]a, b[$ reaaliakselin väli kuten edellisessä esimerkissä. Differentiaalilaskennassa osoitetaan, että jos funktiot $f, g \colon]a, b[\to \mathbb{R}$ ovat kaikkialla derivioituvia, niin silloin myös niiden summafunktio $f + g \colon]a, b[\to \mathbb{R}$ on kaikkialla derivoituva. Differentiaalilaskennassa osoitetaan myös, että vakiolla kertominen ei muuta derivoituvuutta eli jokaisella $\lambda \in \mathbb{R}$ funktio $\lambda f \colon]a, b[\to \mathbb{R}$ on derivoituva.

Määritellään nyt joukko

$$D(]a, b[) = \{f \colon]a, b[\to \mathbb{R} \mid f \text{ on derivoituva}\}.$$

Yllä tehdyt huomiot osoittavat, että funktio $+ \colon D(]a, b[) \times D(]a, b[) \to D(]a, b[)$, missä

$$(f + g)(x) = f(x) + g(x)$$

kaikilla $x \in]a, b[$, on hyvin määritelty laskutoimitus joukossa $D(]a, b[)$. Vastaavasti havaitaan, että skalaarikertolasku $\cdot \colon \mathbb{R} \times D(]a, b[) \to \mathbb{R}$, joka on määrtelty kaavalla

$$(\lambda f)(x) = \lambda f(x)$$

kaikilla $x \in]a, b[$, on hyvin määritelty.

Koska laskutoimitukset on määritelty samoilla kaavoilla kuin jatkuvien funktioiden tapauksessa on suoraviivaista tarkastaa, että kolmikko $(D(]a, b[), +, \cdot)$ on vektoriavaruus,

Huomautus 7.5.2. *Jälleen käsittely voidaan laajentaa kaikille väleille $\Delta \subset \mathbb{R}$, kunhan derivaatan määritelmä päätepisteissä ymmärretään toispuoleisena derivaattana, eli joukko $D(\Delta)$ on vektoriavaruus, kun yhteen- ja skalaarikertolasku määritellään kaavoilla 7.4 ja 7.5.*

7.5.3 Funktioavaruudet $\mathscr{F}(X)$

Esimerkit $C(\Delta)$ ja $D(\Delta)$ ovat erikoistapauksia yleisemmistä funktioavaruuksista. Olkoot X mielivaltainen joukko ja $\mathscr{F}(X)$ kaikkien funktioiden $f \colon X \to \mathbb{R}$ joukko. Tällöin $C(\Delta)$ ja $D(\Delta)$ ovat joukon $\mathscr{F}(\Delta)$ osajoukkoja.

Kaikkien funktioiden joukko $\mathscr{F}(X)$ on siis joukko

$$\mathscr{F}(X) = \{f \colon X \to \mathbb{R} \mid f \text{ on funktio}\}.$$

ja joukon $\mathscr{F}(X)$ alkiot ovat siis funkiota.

Esimerkki 7.5.3. *Olkoon $X =]0, 1[$ reaaliakselin $\mathbb{R}$ väli. Tällöin funktio $f \colon]0, 1[\to \mathbb{R}$, $x \mapsto x^2 + 1$, on joukon $\mathscr{F}(X)$ alkio eli $f \in \mathscr{F}(X)$. Myös funktio $g \colon]0, 1[\to \mathbb{R}$, joka on määritelty kaavalla*

$$g(x) = \begin{cases} -1, & x \leq 1/2 \\ x + 1, & x > 1/2 \end{cases}$$

on joukon $\mathscr{F}(X)$ alkio eli $g \in \mathscr{F}(X)$.

Avaruudet $\mathscr{F}(X)$ antavat siis yleisen kontekstin tutummille funktioavaruuksille $C(\Delta)$ ja $D(\Delta)$. Tämän materiaalin kannalta tärkeimmät esimerkit ovat kuitenkin funktioavaruuksia $\mathscr{F}(X)$, missä X on äärellinen joukko. Syy tähän on se, että nämä funktioavaruudet tulevat antamaan esimerkkejä äärellisulotteisista vektoriavaruuksista.

Esimerkki 7.5.4. *Olkoon $X = \{a, b, c\}$ joukko, jossa on kolme alkiota eli $a \neq b$, $a \neq c$ ja $b \neq c$. Tällöin esimerkiksi funktio $f \colon X \to \mathbb{R}$,*

$$x \mapsto \begin{cases} 3, & x = a \\ -3, & x = b \\ 3, & x = c \end{cases}$$

ja funktio $g \colon X \to \mathbb{R}$, joka on määritelty kaavalla

$$g(x) = \begin{cases} -8, & x = a, b \\ 1, & x = c \end{cases}$$

kuuluvat joukkoon $\mathscr{F}(X)$ eli joukkoon $\mathscr{F}(\{a, b, c\})$.

Edellisessä esimerkissä funktiot f ja g määriteltiin pisteittäin käyttäen hyväksi joukon alkioille annettuja symboleja. Jos joukon X alkiot numeroidaan, voidaan määritellä funktioita hyödyntäen tätä numerointia.

Esimerkki 7.5.5. *Olkoon $X = \{p_1, \ldots, p_n\}$ joukko, jossa on n alkiota. Tällöin sekä funktio $f\colon X \to \mathbb{R}$, $p_k \mapsto (-1)^k$, että funktio $g\colon X \to \mathbb{R}$, joka on määritelty kaavalla*

$$g(x) = \begin{cases} 2k^2 + 3, & x = p_k,\ k \text{ on pariton} \\ 0, & x = p_k,\ k \text{ on parillinen} \end{cases},$$

kuuluvat joukkoon $\mathscr{F}(X)$.

Huomautus 7.5.6. *Edellisissä esimerkeissä on esitelty vaihtoehtoisia tapoja esittää funktioita, joita tullaan käyttämään jatkossa. Tässä materiaalissa oletetaan, että funktion käsite ja yllä käytetyt merkinnät ovat tuttuja sillä tasolla kuin niitä on käsitelty* Johdatus yliopistomatematiikkaan kurssilla. *Jatkuvien, derivoituvien tai integroituvien funktioiden teoriaa ei tällä kurssilla tarvita lukiomatematiikan laajuutta enempää.*

Jokaisessa funktioavaruudessa on kaksi erityistä osajoukkoa funktioita: vakiofunktiot ja (pisteen) karakteristiset funktiot. Molemmilla tulee olemaan oma roolinsa jatkossa. Vakiofunktiot antavat esimerkin funktioavaruuden aliavaruudesta ja karakteristiset funktiot antavat luonnollisen kannan funktioavaruudelle $\mathscr{F}(X)$, jos X on äärellinen joukko. Molempia aiheita käsitellään tarkemmin, kun näiden käsitteiden yhteydessä.

Määritelmä 7.5.7. *Olkoon X joukko. Funktio $f\colon X \to \mathbb{R}$ on vakiofunktio, jos on olemassa sellainen luku $\lambda \in \mathbb{R}$, että $f(x) = \lambda$ jokaisella $x \in X$.*

Määritelmä 7.5.8. *Olkoon X joukko ja $A \subset X$ osajoukko. Funktio $f\colon X \to \mathbb{R}$ on joukon A karakteristinen funktio, jos*

$$f(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

Huomautus 7.5.9. *Aihepiiristä riippuen joukon karakteriselle funktiolle voidaan antaa erityinen merkintä. Joukon $A \subset X$ karakterisista funktioista voidaan merkitä esimerkiksi merkinnällä $\chi_A\colon X \to \mathbb{R}$ eli $\chi_A(x) = 1$, jos $x \in A$ ja $\chi_A(x) = 0$, jos $x \notin A$.*

Äärellisen joukon X tapauksessa ns. pisteen karakteriset funktiot muodostavat tärkeän luokan karakterisia funktiota. Vaikka kyseessä on joukon karakterisen funktion erikoistapaus, annetaan määritelmä kuitenkin myös formaalisti.

Määritelmä 7.5.10. *Olkoon X joukko ja $p \in X$. Funkiota $\chi_p\colon X \to \mathbb{R}$,*

$$x \mapsto \begin{cases} 1, & x = p \\ 0, & x \neq 0 \end{cases}$$

kutsutaan pisteen p karakteristiseksi funktioksi.

Laskutoimitukset funktioavaruuksissa

Joukko $\mathscr{F}(X)$ funktioiden yhteen- ja skaalarikertolaskulla varustettuna on vektoriavaruus eli kun yhteenlaskuksi $+\colon \mathscr{F}(X) \times \mathscr{F}(X) \to \mathscr{F}(X)$ ja skalaarikertolaskuksi $\cdot\colon \mathbb{R} \times \mathscr{F}(X) \to \mathscr{F}(X)$ valitaan kaavoja 7.4 ja 7.5 vastaavat laskutoimitukset eli laskutoimitukset, jotka on määritelty kaavoilla

$$(f+g)(x) = f(x) + g(x)$$

ja

$$(\lambda f)(x) = \lambda f(x)$$

kaikilla $f, g \in \mathscr{F}(X)$, $\lambda \in \mathbb{R}$ ja $x \in X$. Näitä laskutoituksia kutsutaan yleisesti *funktioiden (luonnollisiksi) yhteen- ja skalaarikertolaskuiksi.* Huomaa, että jokaisella $x \in X$ alkiot $f(x)$ ja $g(x)$ ovat lukuja ja laskutoimitukset $f(x)+g(x)$ ja $\lambda f(x)$ ovat reaalilukujen yhteen- ja kertolasku.

Huomautus 7.5.11. *Avaruuden $\mathscr{F}(X)$ määritelmä olisi voitu antaa vieläkin yleisemmin muodossa $\mathscr{F}(X, V)$, missä V on vektoriavaruus. Tämä yleistys jätetään kiinnostuneelle lukijalle.*

Huomautus 7.5.12. *Avaruudet $C(\Delta)$ ja $D(\Delta)$ ovat avaruuden $\mathscr{F}(\Delta)$ aliavaruuksia. Lisäksi $D(\Delta)$ on avaruuden $C(\Delta)$ aliavaruus.*

Esimerkkinä funktioavaruuden $\mathscr{F}(X)$ aliavaruudesta osoitetaan, että vakiofunktiot muodostavat aina avaruuden $\mathscr{F}(X)$ aliavaruuden.

Esimerkki 7.5.13. *Olkoon X joukko ja $\mathscr{V}(X) = \{f\colon X \to \mathbb{R} \mid f \text{ on vakiofunktio}\}$ avaruuden $\mathscr{F}(X)$ osajoukko. Osoitetaan, että $\mathscr{V}(X)$ on aliavaruus. Käytetään aliavaruus kriteerioita. Huomaa, että selvästi nollafunktio kuuluu joukkoon $\mathscr{V}(X)$.*

Olkoot $f, g \in \mathscr{V}(X)$ ja $a \in \mathbb{R}$. Koska f ja g ovat vakiofunktioita, niin on olemassa sellaiset vakiot $c_f \in \mathbb{R}$ ja $c_g \in \mathbb{R}$, että $f(x) = c_f$ ja $g(x) = c_g$ jokaisella $x \in X$. Koska

$$(af+g)(x) = af(x) + g(x) = ac_f + c_g.$$

jokaisella $x \in X$, niin $af + g$ on vakiofunktio. Osajoukko $\mathscr{V}(X)$ on siis aliavaruus.

Äärellisen joukon reaaliarvoisten funktioiden avaruus

Tarkastellaan nyt esimerkinomaisesti tilannetta funktioavaruutta $\mathscr{F}(X)$, kun X on äärellinen joukko. Osa yleisistä väitteistä jätetään harjoitustehtäviksi.

Esimerkki 7.5.14. *Olkoon $X = \{a, b, c\}$ kolmen alkion joukko ja olkoot $f\colon X \to \mathbb{R}$ ja $g\colon X \to \mathbb{R}$ funktioita kuten esimerkissä 7.5.4. Määritetään funktiot $f+g\colon X \to \mathbb{R}$ ja $4f\colon X \to \mathbb{R}$.*

Funktion $f+g\colon X \to \mathbb{R}$ arvot ovat

$$(f+g)(a) = f(a) + g(a) = 3 + (-8) = -5$$
$$(f+g)(b) = f(b) + g(b) = -3 + (-8) = -11$$

ja

$$(f+g)(c) = f(c) + g(c) = 3 + 1 = 4.$$

Näin ollen funktio $f + g\colon X \to \mathbb{R}$ on määritelty kaavalla

$$x \mapsto \begin{cases} -5, & x = a \\ -11, & x = b \\ 4, & x = c \end{cases}$$

Vastaavasti saadaan, että funktio $4f\colon X \to \mathbb{R}$ on funktio

$$x \mapsto \begin{cases} 12, & x = a \\ -12, & x = b \\ 12, & x = c \end{cases}$$

Edellinen esimerkki saattaa vaikuttaa liiankin konkreettiselta. Sen varsinainen motivaatio on kuitenkin toimia pohjustuksena huomiolle, että äärellisen joukon reaaliarvoiset funktiot ovat karakterististen funktioiden lineaarikombinaatioita.

Esimerkki 7.5.15. *Olkoon $X = \{a, b, c\}$ kolmen alkion joukko ja olkoot $f\colon X \to \mathbb{R}$ ja $g\colon X \to \mathbb{R}$ funktioita kuten esimerkissä 7.5.4. Tällöin molemmat funktiot f ja g voidaan kirjoittaa karakterististen funktioiden χ_a, χ_b ja χ_c lineaarikombinaatioina.*

Tarkastellaan funktiota f ja etsitään sellaiset luvut $\lambda_a, \lambda_b, \lambda_c \in \mathbb{R}$, että

$$f = \lambda_a\chi_a + \lambda_b\chi_b + \lambda_c\chi_c. \tag{7.6}$$

Palautetaan mieleen, että kaksi funkioita ovat samat, jos niillä on sama lähtö- ja maaliavaruus sekä sama arvo jokaisessa lähtöjoukon pisteessä. Näin ollen yhtälö (7.6) toteutuu, jos ja vain jos jokaisella $x \in X$ pätee

$$f(x) = \lambda_a\chi_a(x) + \lambda_b\chi_b(x) + \lambda_c\chi_c(x).$$

Käydään läpi kaikki tapaukset.

Jos $x = a$, niin karakteristisen funktion määritelmän mukaan $\chi_a(a) = 1$, $\chi_b(a) = 0$ ja $\chi_c(a) = 0$. Näin ollen saadaan yhtälö

$$f(a) = \lambda_a \cdot 1 + \lambda_b \cdot 0 + \lambda_c \cdot 0 = \lambda_a.$$

Vastaavasti tapauksissa $x = b$ ja $x = c$ saadaan yhtälöt

$$f(b) = \lambda_b$$

ja

$$f(c) = \lambda_c.$$

Näin on havaittu, että yhtälön (7.6) toteutumisen välittämätön ehto on, että $\lambda_a = f(a)$, $\lambda_b = f(b)$ ja $\lambda_c = f(c)$. Toisaalta käymällä tapaukset samalla tavalla läpi, havaitaan, että

$$((f(a)\chi_a + f(b)\chi_b + f(c)\chi_c)(x) = f(x)$$

jokaisella $x \in X$. Näin ollen

$$f = f(a)\chi_a + f(b)\chi_b + f(c)\chi_c = 3\chi_a - 3\chi_b + 3\chi_c.$$

Vastaavasti

$$g = -8\chi_a - 8\chi_b + \chi_c.$$

Huomautus 7.5.16. *Mikäli edellinen esimerkki ei ollut entuudestaan tuttu kannattaa tässä vaiheessa hieman pysähtyä. Esimerkistä huomaa selvästi konseptuaalisen eron itse funktion f ja sen arvon $f(x)$ pisteessä $x \in X$ kanssa. Tämä ero näkyy erityisen selvästi niissä kohdissa, joissa tarkastellaan toistaalta funktioiden yhtäsuuruutta ja funktioiden arvojen yhtäsuuruutta jokaisessa pisteessä. Toisaalta sama ero näkyy myös, kun luvut λ_a, λ_b ja λ_c on määritetty ja todettu, että ne ovat funktion f arvot pisteissä a, b ja c.*

Lineaarialgebrallinen tulkinta tästä ilmiöstä on, että esimerkissä funktio f on vektoriavaruuden $\mathscr{F}(X)$ alkio ja funktion f arvot pisteissä a, b ja c tulevat olemaan vektorin f koordinaatit kannassa (χ_a, χ_b, χ_c).

Lineaarikuvaukset äärellisen joukon reaaliarvoisten funktioiden avaruudelta

Funktioavaruuksiin $\mathscr{F}(X)$ liittyy luonnollisia lineaarikuvauksia, joista yleisin on ns. evaluaatiofunktionaali.

Esimerkki 7.5.17. *Olkoon X joukko, $p \in X$ ja $e_p \colon \mathscr{F}(X) \to \mathbb{R}$ kuvaus*

$$f \mapsto f(p).$$

Evaluaatiofunktionaali e_p liittää siis jokaiseen joukon X reaaliarvoiseen funktioon $f \colon X \to \mathbb{R}$ tämän funktion arvon pisteessä p. Käyttäen toista merkintaa, evaluaatiofunktionaali $e_p \colon \mathscr{F}(X) \to \mathbb{R}$ on siis kuvaus, joka on määritelty kaavalla

$$e_p(f) = f(p)$$

jokaisella $f \in \mathscr{F}(X)$.

Huomaa, että e_p on siis kuvaus, jonka argumenttina ei ole piste p vaan joukon $\mathscr{F}(X)$ alkio eli funktio $f \mapsto X \to \mathbb{R}$.

Osoitetaan, että $e_p \colon \mathscr{F}(X) \to \mathbb{R}$ on lineaarikuvaus. Tulee siis osoittaa, että

$$e_p(f+g) = e_p(f) + e_p(g) \tag{7.7}$$

kaikilla $f, g \in \mathscr{F}(X)$ ja että

$$e_p(af) = ae_p(f) \tag{7.8}$$

jokaisella $f \in \mathscr{F}(X)$ ja $a \in \mathbb{R}$.

Olkoot $f, g \in \mathscr{F}(X)$. Tällöin

$$e_p(f+g) = (f+g)(p) = f(p) + g(p) = e_p(f) + e_p(g)$$

Näin ollen yhtälö (7.7) on voimassa. Yhtälö (7.8) osoitetaan samoin.

Äärellisen joukon X tapauksessa tutut operaatiot, kuten keskiarvo, antavat esimerkkejä lineaarikuvauksista.

Esimerkki 7.5.18. *Olkoon* $X = \{p_1, \ldots, p_n\}$ *äärellinen joukko, jossa on* n *alkiota. Määritellään kuvaus* $\kappa\colon \mathscr{F}(X) \to \mathbb{R}$ *kaavalla*

$$\kappa(f) = \frac{f(p_1) + \cdots + f(p_n)}{n}$$

kaikilla $f \in \mathscr{F}(X)$.

Osoitetaan, että κ *on lineaarikuvaus. Olkoot* $f, g \in \mathscr{F}(X)$ *ja* $a \in \mathbb{R}$*. Tällöin*

$$\begin{aligned}
\kappa(af + g) &= \frac{(af + g)(p_1) + \cdots + (af + g)(p_n)}{n} \\
&= \frac{(af(p_1) + g(p_1)) + \cdots (af(p_n) + g(p_n))}{n} \\
&= \frac{(af(p_1) + \cdots + af(p_n)) + (g(p_1) + \cdots + g(p_n))}{n} \\
&= a\frac{f(p_1) + \cdots + f(p_n)}{n} + \frac{g(p_1) + \cdots + g(p_n)}{n} \\
&= a\kappa(f) + \kappa(g).
\end{aligned}$$

Kuvaus κ *on siis lineaarikuvaus.*

Tarkastellaan viimeisenä esimerkkinä lineaarikuvausta, joka vähentää funktiosta sen keskiarvon. Tämä antaa esimerkin lineaarikuvauksesta funktioavaruudesta itseensä.

Esimerkki 7.5.19. *Olkoon* $X = \{p_1, \ldots, p_n\}$ *äärellinen joukko, jossa on* n *alkiota ja olkoon* $\kappa\colon \mathscr{F}(X) \to \mathbb{R}$ *keskiarvokuvaus kuten edellisessä esimerkissä. Määritellään kuvaus* $\sigma\colon \mathscr{F}(X) \to \mathscr{F}(X)$ *kaavalla*

$$(\sigma(f))(x) = f(x) - \kappa(f)$$

kaikilla $f \in \mathscr{F}(X)$ *ja* $x \in X$.

Huomaa, että funktio $f \in \mathscr{F}(X)$ *on kuvauksen* σ *argumentti. Koska kuvauksen* σ *arvo* $\sigma(f)$ *pisteessä* $f \in \mathscr{F}(X)$ *on funktio* $\sigma(f)\colon X \to \mathbb{R}$*, niin* $\sigma(f)$ *määritellään antamalla sen arvo jokaisessa joukon* X *pisteessä.*

Osoitetaan nyt, että σ *on lineaarinen. Olkoot* $f, g \in \mathscr{F}(X)$*. Tällöin* $\sigma(f+g)\colon X \to \mathbb{R}$ *on funktio, joka pisteessä* $x \in X$ *saa arvon*

$$\sigma(f + g)(x) = (f + g)(x) - \kappa(f + g).$$

Koska $\kappa\colon \mathscr{F}(X) \to \mathbb{R}$ *on lineaarinen, niin* $\kappa(f + g) = \kappa(f) + \kappa(g)$*. Näin ollen*

$$\begin{aligned}
\sigma(f + g)(x) &= (f + g)(x) - \kappa(f + g) \\
&= f(x) + g(x) - (\kappa(f) + \kappa(g)) \\
&= (f(x) - \kappa(f)) + (g(x) - \kappa(g)) \\
&= \sigma(f)(x) + \sigma(g)(x) \\
&= (\sigma(f) + \sigma(g))(x)
\end{aligned}$$

jokaisella $x \in X$ eli

$$\sigma(f+g) = \sigma(f) + \sigma(g).$$

Vastaavasti osoitetaan, että

$$\sigma(af) = a\sigma(f)$$

kaikilla $f \in \mathscr{F}(X)$ ja $a \in \mathbb{R}$. Kuvaus $\sigma\colon \mathscr{F}(X) \to \mathscr{F}(X)$ on siis lineaarinen.

Tarkastellaan vielä kuvausten κ ja σ ytimiä esimerkkeinä avaruuden $\mathscr{F}(X)$ aliavaruuksista.

Esimerkki 7.5.20. *Olkoon $X = \{p_1, \ldots, p_n\}$ äärellinen joukko, jossa on n alkiota, ja olkoon $\kappa\colon \mathscr{F}(X) \to \mathbb{R}$ lineaarikuvaus kuten esimerkissä 7.5.18.*

Olkoon $f \in \ker\kappa$. Tällöin

$$0 = \kappa(f) = f(p_1) + \cdots + f(p_n).$$

Näin ollen

$$f(p_n) = -(f(p_1) + \cdots + f(p_{n-1})).$$

Ydin $\ker\kappa$ on siis aliavaruus

$$\ker\kappa = \{f\colon X \to \mathbb{R} \mid f(p_n) = -(f(p_1) + \cdots + f(p_{n-1}))\}.$$

Esimerkki 7.5.21. *Olkoon $X = \{p_1, \ldots, p_n\}$ äärellinen joukko, jossa on n alkiota, ja olkoon $\sigma\colon \mathscr{F}(X) \to \mathbb{R}$ lineaarikuvaus kuten esimerkissä 7.5.19.*

Olkoon $f \in \ker\sigma$. Tällöin

$$0 = \sigma(f)(x) = f(x) - \kappa(f)$$

jokaisella $x \in X$. Näin ollen

$$f(x) = \kappa(f)$$

jokaisella $x \in X$. Koska $\kappa(f)$ on luku, niin f on vakiofunktio.

Ydin $\ker\sigma$ on siis vakiofunktioiden aliavaruus

$$\ker\sigma = \{f\colon X \to \mathbb{R} \mid f \text{ on vakio}\}.$$

7.5.4 Polynomien vektoriavaruudet

Polynomit muodostavat derivoituvien funktioiden aliavaruuksia.

Olkoon $\Delta \subset \mathbb{R}$ reaaliakselin väli. Funktio $p\colon \Delta \to \mathbb{R}$ on *polynomi*, jos on olemassa $n \in \mathbb{N}$ ja sellaiset luvut $a_0, \ldots, a_n \in \mathbb{R}$, että

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

jokaisella $x \in \mathbb{R}$. Lukuja $a_0, \ldots, a_n$ kutsutaan polynomin p kertoimiksi. Lisäksi lukua $n \in \mathbb{N}$ sanotan polynomin p *asteeksi*, jos $a_n \neq 0$ ja $a_m = 0$ kaikilla $m > n$.

Olkoon $\mathscr{P}(\Delta)$ kaikkien polynomien $p\colon \Delta \to \mathbb{R}$ joukko ja olkoon $\mathscr{P}_n(\Delta)$ kaikkien korkeintaan astetta $n \in \mathbb{N}$ olevien polynomien $p\colon \Delta \to \mathbb{R}$ joukko.

Mikäli p ja q ovat polynomeja joukossa $\mathscr{P}(\Delta)$, niin selvästi niiden summa $p+q$ on polynomi ja jokaisella $\lambda \in \mathbb{R}$ funktio λp on polynomi. Tässä yhteenlasku ja skalaarikertolasku ovat edellä käsitellyt funktioiden yhteen- ja skalaarikertolasku. Käymällä läpi vektoriavaruuden aksioomat (1)–(6) havaitaan, että $\mathscr{P}$ ja jokainen $\mathscr{P}_n$ ovat vektoriavaruuksia. Tämä voidaan kuitenkin havaitaan suoraviivaisemmin toteamalla, että joukot $\mathscr{P}$ ja $\mathscr{P}_n$ ovat derioituvien funktioiden avaruuden $D(\Delta)$ aliavaruuksia.

Jatkossa reaaliakselin polynomifunkitioiden avaruuksista $\mathscr{P}(\mathbb{R})$ ja $\mathscr{P}_n(\mathbb{R})$ tullaan käyttämään lyhyempiä merkintöjä $\mathscr{P} = \mathscr{P}(\mathbb{R})$ ja $\mathscr{P}_n = \mathscr{P}_n(\mathbb{R})$.

Polynomien yhteen- ja skalaarikertolasku kertoimittain

Polynomien yhteen- ja skalaarikertolasku voidaan suorittaa kertoimittain seuraavasti. Olkoot $p\colon \Delta \to \mathbb{R}$, $x \mapsto a_n x^n + \cdots + a_0$, ja $q\colon \Delta \to \mathbb{R}$, $x \mapsto b_n x^n + \cdots + b_0$, polynomeja sekä $\lambda \in \mathbb{R}$. Tällöin $p+q\colon \Delta \to \mathbb{R}$ on polynomi $x \mapsto (a_n+b_n)x^n + \cdots + (a_0+b_0)$, sillä jokaisella $x \in \Delta$ pätee

$$\begin{aligned}
(p+q)(x) &= p(x) + q(x) \\
&= (a_n x^n + \cdots + a_0) + (b_n x^n + + \cdots + b_0) \\
&= a_n x^n + \cdots + a_0 + b_n x^n + + \cdots + b_0 \\
&= (a_n x^n + b_n x^n) + \cdots + (a_0 + b_0) \\
&= (a_n + b_n) x^n + \cdots + (a_0 + b_0).
\end{aligned}$$

Vastaavasti osoitetaan, että $\lambda p\colon \Delta \to \mathbb{R}$ on polynomi $x \mapsto \lambda a_n x^n + \cdots + \lambda a_0$.

Huomautus 7.5.22. *Polynomin määritelmässä mikään ei suoraan sano, että polynomin kertoimet tai aste olisivat yksikäsitteisiä. Näin tulee kuitenkin olemaan.*

7.5.5 Derivaatta lineaarikuvauksena

Avaruutta $D(\Delta)$ yleisemmin käytössä on jatkuvasti derivoituvien funktioiden avaruus

$$C^1(\Delta) = \{f\colon \Delta \to \mathbb{R} \mid f \in D(\Delta),\ f' \in C(\Delta)\}.$$

Otetaan tunnettuna, että $C^1(\Delta)$ on vektoriavaruus tavallisilla funktioiden yhteen- ja skalaarikertolaskulla.

Differentiaalilaskennassa osoitetaan, että derivoituville funktioille $f, g \in \Delta \to \mathbb{R}$ pätee $(f+g)'(x) = f'(x) + g'(x)$ ja $(\lambda f)'(x) = \lambda f'(x)$ kaikilla $x \in \Delta$. Näin ollen $(f+g)' = f' + g'$ ja $(\lambda f)' = \lambda f'$ kaikilla $f, g \in D(\Delta)$ ja $\lambda \in \mathbb{R}$. Kuvaus

$$D\colon D(\Delta) \to \mathscr{F}(\Delta),\ f \mapsto f',$$

on siis lineaarikuvaus.

Yleisemmin käytetty on saman kaavan määrittelemä lineaarikuvaus

$$D\colon C^1(\Delta) \to C(\Delta),\ f \mapsto f'.$$

Koska polynomit ovat jatkuvasti derivoituvia, niin sama kaava määrittelee myös lineaarikuvaukset

$$D\colon \mathscr{P}(\Delta) \to \mathscr{P}(\Delta),\ p \mapsto p',$$

ja

$$D\colon \mathscr{P}_n(\Delta) \to \mathscr{P}_n(\Delta),\ p \mapsto p',$$

jokaisella $n \in \mathbb{N}$.

7.5.6 Integraali lineaarikuvauksena

Integraalilaskennassa osoitetaan, että suljetulla ja rajoitetulla välillä $[a,b]$ määritellyt jatkuvat funktiot ovat integroituvia ja että integroituvien funktioiden summat ja tulot vakioiden kanssa ovat integroituvia. Lisäksi

$$\int_a^b f(x) + g(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x.$$

ja

$$\int_a^b \lambda f(x)\,\mathrm{d}x = \lambda \int_a^b g(x)\,\mathrm{d}x.$$

Välin $[a,b]$ yli integroiminen määrittelee siis lineaarikuvauksen

$$\widetilde{I}\colon C([a,b]) \to \mathbb{R},\ f \mapsto \int_a^b f(x)\,\mathrm{d}x,$$

jatkuvien funktioiden avaruudelta reaaliluvuille.[4]

Integraalilaskennassa osoitetaan myös, että jatkuvalla funktiolla $f\colon [a,b] \to \mathbb{R}$ on integraalifunktio $F\colon [a,b] \to \mathbb{R}$, joka on määritelty kaavalla

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

kaikilla $x \in [a,b]$, ja jolla on ominaisuus $F'(x) = f(x)$ kaikilla $x \in [a,b]$. Näin ollen $F \in C^1([a,b])$.

Integraalin ominaisuuksista seuraa, että kuvaus $I\colon C([a,b]) \to C^1([a,b])$, joka liittää funktioon f sen integraalifunktion F, eli kuvaus I on määritelty kaavalla

$$(If)(x) = \int_a^x f(t)\,\mathrm{d}t,$$

[4] Itse asiassa jatkuvien funktioiden avaruus voidaan tässä vaihtaa integroituvien funktioiden avaruuteen, joka on myös vektoriavaruus funktioiden yhteen- ja skalaarikertolaskun suhteen.

missä $x \in [a, b]$, määritelty kuvaus on lineaarinen. Osoitetaan vielä kuvauksen I lineaarisuus. Olkoot $f, g \in C([a, b])$ ja $\lambda \in \mathbb{R}$. Tällöin jokaisella $x \in [a, b]$ pätee

$$\begin{aligned}
\left(I(\lambda f + g)\right)(x) &= \int_a^x (\lambda f + g)(t) \, \mathrm{d}t \\
&= \int_a^x \lambda f(t) + g(t) \, \mathrm{d}t \\
&= \lambda \int_a^x f(t) \, \mathrm{d}t + \int_a^x g(t) \, \mathrm{d}t \\
&= \lambda (If)(x) + (Ig)(x) = (\lambda If + Ig)(x).
\end{aligned}$$

Näin ollen

$$I(\lambda f + g) = \lambda If + Ig.$$

Kuvaus I on siis lineaarinen.

Luku 8

Äärellisulotteisten vektoriavaruuksien teoria

Tähän lukuun on kerätty äärellisesti viritettyjen vektoriavaruuksien teorian tärkeimmät tulokset:

- jokaisella äärellisesti viritetyllä vektoriavaruudella on kanta,
- jokaisella äärellisesti viritetyllä vektoriavaruudella on dimensio,
- jokainen virittävä jono sisältää kannan ja
- jokainen vapaa jono voidaan laajentaa kannaksi.

Lisäksi käsitellään näiden tulosten tärkeimpiä seurauksia.

8.1 Virittäminen, vapaus ja kanta

Virittämisen ja vapauden käsitteet määritellään yleisille vektoriavaruuksille kuten luvussa 3 aliavaruuksien tapauksessa. Paljastuu, että monet tulokset todistetaan tässä yleisemmässä tilanteessa aivan samoin kuin luvun 3 tapauksessa.

Määritelmä 8.1.1. *Vektoriavaruus V on* äärellisesti viritetty, *jos on olemassa $k \in \mathbb{N}$ ja sellaiset vektorit $v_1, \dots, v_k \in V$, että $\mathrm{Sp}(v_1, \dots, v_k) = V$.*

Huomautus 8.1.2. *Jatkossa käytetään konventiota, että nolla-avaruus $\{0\}$ on tyhjän jonon virittämä, eli $\{0\} = \mathrm{Sp}(\emptyset)$. Tämän voi tulkita tarkoittavan, että aliavaruuden $\{0\}$ virittämiseen ei tarvita vektoreita, vaikkakin on myös totta, että $\{0\} = \mathrm{Sp}(0)$. Tässä esityksessä syy tälle konventiolle on, että nollavektorin muodostama yhden alkion mittainen jono (0) ei ole kohta annettavan määritelmän mukaan vapaa jono.*

Esimerkki 8.1.3. *Olkoon $\Delta \subset \mathbb{R}$ reaaliakselin väli ja $n \in \mathbb{N}$. Olkoot myös $p_k \colon \Delta \to \mathbb{R}$, $x \mapsto x^k$, jokaisella $k \in \{0, \dots, n\}$, missä p_0 on vakiofunktio $x \mapsto 1$. Tällöin $\mathscr{P}_n(\Delta) = \mathrm{Sp}(p_0, \dots, p_n)$. Näin ollen $\mathscr{P}_n(\Delta)$ on äärellisesti viritetty.*

Esimerkki 8.1.4. *Olkoon X äärellinen epätyhjä joukko, jossa on n pistettä. Määritellään jokaisella $p \in X$ funktio $\chi_p \colon X \to \mathbb{R}$ kaavalla*

$$\chi_p(x) = \begin{cases} 1, & x = p \\ 0, & x \neq p. \end{cases}$$

Tällöin jokaisella $f \colon X \to \mathbb{R}$ pätee

$$f = \sum_{p \in X} f(p)\chi_p.$$

(Tarkka todistus jätetään harjoitustehtäväksi.) Näin ollen $\mathscr{F}(X) = \mathrm{Sp}(\chi_{p_1}, \ldots, \chi_{p_n})$, missä $\{p_1, \ldots, p_n\} = X$. Erityisesti $\mathscr{F}(X)$ on äärellisesti viritetty.

Lauseessa 3.4.3 osoitettiin avaruuden $\mathbb{R}^{n\times 1}$ tapauksessa, että $\mathrm{Sp}(v_1, \ldots, v_k) \subset \mathbb{R}^{n\times 1}$ on aliavaruus kaikilla $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$. Sama todistus pätee sanasta sanaan, kun avaruus $\mathbb{R}^{n\times 1}$ vaihdetaan yleiseen avaruuteen V. Tämän vuoksi todistusta ei tässä toisteta. Kirjataan tämä tulos kuitenkin lemmaksi jatkoa varten.

Lemma 8.1.5. *Olkoot V vektoriavaruus ja olkoot $w_1, \ldots, w_m \in V$. Tällöin osajoukko $W = \mathrm{Sp}(w_1, \ldots, w_m)$ on avaruuden V aliavaruus.*

Huomautus 8.1.6. *Itseasiassa myös lauseen 3.4.7 tulos, että $\mathrm{Sp}(v_1, \ldots, v_k)$ on pienin aliavaruus, joka sisältää vektorit $v_1, \ldots, v_k$, on voimassa. Jälleen todistus on aivan sama kuin avaruuden $\mathbb{R}^{n\times 1}$ tilanteessa.*

Lineaarinen riippumattomuus ja jonon vapaus määritellään kuten sarakevaruuden tapauksessa.

Määritelmä 8.1.7. *Vektoriavaruuden V vektorit $v_1, \ldots, v_k$ ovat* lineaarisesti riippumattomia*, jos kaikilla $a_1, \ldots, a_k \in \mathbb{R}$ yhtälöstä*

$$a_1 v_1 + \cdots + a_k v_k = 0$$

seuraa $a_1 = a_2 = \cdots = a_n = 0$.

Määritelmä 8.1.8. *Avaruuden V vektoreiden jonoa $(v_1, \ldots, v_k)$ kutsutaan* vapaaksi*, jos vektorit $v_1, \ldots, v_k$ ovat lineaarisesti riippumattomia. Jonoa, joka ei ole vapaa, kutsutaan sidotuksi.*

Äärellisesti viritetyn avaruuden kanta voidaan nyt määritellä seuraavasti.

Määritelmä 8.1.9. *Äärellisesti viritetyn vektoriavaruuden V vektoreiden $v_1, \ldots, v_k$ jono $(v_1, \ldots, v_k)$ on* avaruuden V kanta*, jos vektorit $v_1, \ldots, v_k$ ovat lineaarisesti riippumattomia ja virittävät avaruuden V eli $V = \mathrm{Sp}(v_1, \ldots, v_k)$.*

Huomautus 8.1.10. *Huomautuksen 8.1.2 perusteella voidaan tulkita, että myös avaruudella $\{0\}$ on kanta, ns. tyhjä kanta $(\emptyset)$. Huomaa, että (0) ei ole avaruuden $\{0\}$ kanta, koska jono (0) ei ole vapaa.*

Kootaan nyt muutamia esimerkkejä vektoriavaruuksien kannoista. Esimerkkien yksityiskohdat jätetään harjoitustehtäviksi.

Esimerkki 8.1.11. *Avaruuden* $\mathbb{R}^{n\times 1}$ *standardikanta* $(e_1,\dots,e_n)$ *ja avaruuden* $\mathbb{R}^n$ *standardikanta* $(e_1,\dots,e_n)$ *ovat kantoja.*

Esimerkki 8.1.12. *Olkoon* $(v_1,\dots,v_n)$ *avaruuden* $\mathbb{R}^{n\times 1}$ *kanta. Tällöin jono* $(v_1^t,\dots,v_n^t)$ *on avaruuden* $\mathbb{R}^{1\times n}$ *kanta.*

Esimerkki 8.1.13. *Alkeismatriisien yhteydessä (luvussa 2.5) esitellyt matriisit* E^{pr} *muodostavat matriisiavaruuden* $\mathbb{R}^{m\times n}$ *kannan. Tarkemmin sanottuna jokaisella* $m\ge 1$ *ja* $n\ge 1$ *jono* $(E^{11},\dots,E^{1n},E^{21},\dots,E^{mn})$ *on avaruuden* $\mathbb{R}^{m\times n}$ *kanta. Tätä kantaa kutsutaaan avaruuden* $\mathbb{R}^{m\times n}$ *standardikannaksi.*

Esimerkki 8.1.14. *Esimerkin 8.1.3 jono* $(p_0,\dots,p_n)$ *on avaruuden* $\mathscr{P}_n(\Delta)$ *kanta.*

Esimerkki 8.1.15. *Esimerkin 8.1.4 jono* $(\chi_{p_1},\dots,\chi_{p_n})$ *on avaruuden* $\mathscr{F}(X)$ *kanta.*

Huomautus 8.1.16. *Tarkkaavaiset lukijat jo huomasivatkin, että määritelmässä kanta määritellään ainoastaan äärellisesti viritetyille vektoriavaruuksille. Itseasiassa kanta voidaan määritellä kaikille vektoriavaruuksille, myös sellaisille, jotka eivät ole äärellisesti viritettyjä, ja yleinen tulos on että aivan jokaisella vektoriavaruudella on kanta. Tosin, jos avaruus ei ole äärellisesti viritetty, niin tällöin avaruuden kanta ei vastaa sitä intuitiota, joka saatiin luvussa 3. Koska tämä käsittely johdattaa täysin toiseen aiheeseen, määritellään kanta tässä yhteydessä ainoastaan äärellisesti viritetyille avaruuksille.*

Yleisten vektoriavaruuksien kohdalla kannan merkitys on, että kannan valinnan jälkeen voidaan määritellä vektorin koordinaatit tässä kannassa. Heuristisesti avaruuden V kannan valinta kiinnittää avaruudelle V koordinaatiston.

Määritelmä 8.1.17. *Olkoon* $(v_1,\dots,v_k)$ *vektoriavaruuden* V *kanta. Vektorin* $v\in V$ koordinaatit $(a_1,\dots,a_k)$ kannassa $(v_1,\dots,v_k)$ *ovat ne luvut* $a_1,\dots,a_k\in\mathbb{R}$*, joille pätee*

$$v = a_1v_1+\cdots+a_kv_k.$$

Vektorin koordinaatit kannassa ovat yksikäsitteiset ja tämä ominaisuus määrää kannan. Tämä osoitetaan aivan kuten lemmassa 3.7.3, joten todistus jätetään lukijalle harjoitustehtäväksi.

Lause 8.1.18. *Olkoon* V *äärellisesti viritetty vektoriavaruus. Tällöin jono* $(v_1,\dots,v_k)$ *on avaruuden* V *kanta, jos ja vain jos jokaisella avaruuden* V *vektorilla on yksikäsitteiset koordinaatit eli jokaisella* $v\in V$ *on olemassa yksikäsitteiset sellaiset luvut* $a_1,\dots,a_k\in\mathbb{R}$*, että* $v=a_1v_1+\cdots+a_kv_k$*.*

Huomautus 8.1.19. *Vaikka koordinaattien määritelmä vastaa tässä tapauksessa täysin luvussa 3 annettua määritelmää, on sillä syvällinen tulkinta.*

Luvussa 3 tarkasteltiin vektoreita, jotka olivat sarakeavaruuden $\mathbb{R}^{n\times 1}$ *alkiota. Näin ollen ne olivat jo alunperin ilmoitettu reaalilukujen jonoina. Sama reaalilukujen jono* *toistui, kun vektorit kirjoitettiin standardikannassa* $(e_1, \ldots, e_n)$. *Tässä tapauksssa aliavaruuden* $V \subset \mathbb{R}^{n\times 1}$ *yleinen kanta* $(v_1, \ldots, v_k)$ *siis antoi ainoastaan uudet luonnollisemmat koordinaatit, joissa esittää vektoreita.*

Yleisen vektoriavaruuden V kohdalla vektorit eivät välttämättä ole lukujonoja, joten niillä ei ole luonnollisia koordinaatteja. Kannan olemassaolo antaa siis tavan samastaa vektori yksikäsitteisen lukujonon kanssa.

Huomautus 8.1.20. *Herääkin ajatus, voiko äärellisulotteisen vektoriavaruuden V vektorit siis jotenkin samastaa sarakeavaruuden kanssa lineaarikuvauksella käyttäen koordinaatteja. Näin juurikin tullaan tekemään ja tätä käsitellään luvussa 9.3.*

8.1.1 Vapaa jono on aina korkeintaan yhtä pitkä kuin virittävä

Seuraava lause on teoreettisesti katsoen tämän luvun tärkeimpiä tuloksia. Se sanoo, että äärellisesti viritetyn avaruuden vapaassa jonossa on aina korkeintaan yhtä monta alkiota kuin virittävässä jonossa, eli lineaarisesti riippumattomia vektoreita ei voi koskaan löytää lukumäärällisesti enempää kuin mitä tarvitaan avaruuden virittämiseen. Tämä pätee riippumatta siitä, mitä lineaarisesti riippumattomia vektoreita tai mitä virittäviä vektoreita tarkastellaan.

Lause 8.1.21. *Olkoon V vektoriavaruus, joka on vektoreiden* $w_1, \ldots, w_m \in V$ *virittämä. Tällöin jokaiselle vapaalle jonolle* $(v_1, \ldots, v_k)$ *pätee* $k \leq m$.

Lauseen todistusta varten kirjataan aputulos, joka sanoo, että sellaisen vektorin u_j, joka on jonon $(u_1, \ldots, u_m)$ muiden vektoreiden lineaarikombinaatio, poistaminen jonosta ei muuta virittyvää aliavaruutta.

Lemma 8.1.22. *Olkoon V vektoriavaruus ja olkoot* $v_1, \ldots, v_k \in V$ *vektoreita. Jos jono* $(v_1, \ldots, v_k)$ *on sidottu, niin on olemassa sellainen indeksi* $1 \leq j \leq k$, *että* $v_j \in \mathrm{Sp}(v_1, \ldots, v_{j-1})$. *Lisäksi*

$$\mathrm{Sp}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k) = \mathrm{Sp}(v_1, \ldots, v_k).$$

Todistus. Koska jono $(v_1, \ldots, v_k)$ on sidottu ja jono (v_1) on vapaa, niin on olemassa pienin sellainen indeksi $j \in \{1, \ldots, k\}$, että jono $(v_1, \ldots, v_j)$ on sidottu. Tällöin on olemassa sellaiset luvut $a_1, \ldots, a_j \in \mathbb{R}$, jotka eivät kaikki ole nollia, että

$$a_1v_1 + \cdots + a_{j-1}v_{j-1} + a_jv_j = 0.$$

Koska jono $(v_1, \ldots, v_{j-1})$ on vapaa, niin $a_j \neq 0$. Näin ollen

$$v_j = \frac{1}{a_j}\left(a_1v_1 + \cdots + a_{j-1}v_{j-1}\right),$$

eli $v_j \in \mathrm{Sp}(v_1, \ldots, v_{j-1})$.

Osoitetaan nyt toinen väite. Lemman 8.1.5 perusteella $\mathrm{Sp}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k) \subset \mathrm{Sp}(v_1, \ldots, v_k)$. Toisaalta jokaisella $i \in \{1, \ldots, k\}$ pätee $v_i \in \mathrm{Sp}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k)$. Näin ollen lemman 8.1.5 perusteella pätee $\mathrm{Sp}(v_1, \ldots, v_k) \subset \mathrm{Sp}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k)$. □

Lauseen 8.1.21 todistus. Väite todistetaan osoittamalla, että jonon $(w_1, \dots, w_m)$ vektoreita voidaan vaihtaa yksitellen vektoreihin $(v_1, \dots, v_k)$ viritettävän aliavaruuden siitä muuttumatta, eli osoitetaan, että on olemassa sellaiset indeksit $1 \leq i_1 < \cdots < i_\ell \leq m$, missä $\ell = m - k$, joille pätee

$$V = \mathrm{Sp}(v_1, \dots, v_k, w_{i_1}, \dots, w_{i_\ell}).$$

Tällöin $\ell \geq 0$, joten $m \geq k$. Todistus on rakenteeltaan induktio.

Koska $\mathrm{Sp}(w_1, \dots, w_m) = V$, niin jono $(v_1, w_1, \dots, w_m)$ on sidottu. Sovelletaan nyt lemmaa 8.1.22 jonoon $(v_1, w_1, \dots, w_m)$. Koska jono (v_1) on vapaa, niin se ei ole sidottu ja siten lemman 8.1.22 perusteella on olemassa sellainen $j \in \{1, \dots, m\}$, että $w_j \in \mathrm{Sp}(v_1, w_1, \dots, w_{j-1})$ ja $\mathrm{Sp}(v_1, w_1, \dots, w_{j-1}, w_{j+1}, \dots, w_m) = \mathrm{Sp}(v_1, w_1, \dots, w_m) = V$. Merkitään jonolla $(i_{11}, \dots, i_{1(m-1)})$ jonoa $(1, \dots, j-1, j+1, \dots, m)$.

Oletetaan nyt, että $1 \leq r < k$ on sellainen luku, että on olemassa sellainen jonon $(1, \dots, m)$ osajono $(i_{r1}, \dots, i_{r(m-r)})$, että

$$V = \mathrm{Sp}(v_1, \dots, v_r, w_{i_{r1}}, \dots, w_{i_{r(m-r)}})$$

Tällöin jono $(v_1, \dots, v_{r+1}, w_{i_{r1}}, \dots, w_{i_{r(m-r)}})$ on sidottu. Merkitään $u_p = v_p$ jokaisella $1 \leq p \leq r+1$ ja $u_{r+1+p} = w_{i_{rp}}$ jokaisella $1 \leq p \leq m - r$.

Koska jono $(u_1, \dots, u_{m+1})$ on sidottu, niin lemman 8.1.22 perusteella on olemassa sellainen indeksi $j \in \{1, \dots, m+1\}$, että $u_j \in \mathrm{Sp}(u_1, \dots, u_{j-1})$ ja että

$$\mathrm{Sp}(u_1, \dots, u_{j-1}, u_{j+1}, \dots u_{m+1}) = \mathrm{Sp}(u_1, \dots, u_{m+1}).$$

Osoitetaan, että $j > r+1$. Jos näin ei ole, niin $u_j = v_j$ ja $v_j \in \mathrm{Sp}(v_1, \dots, v_{j-1})$. Tämä on ristiriita, koska $(v_1, \dots, v_j)$ on vapaa jono. Näin ollen $j > r+1$, eli $u_j = w_{i_\ell}$ jollain $1 \leq \ell \leq m - r$. Näin ollen

$$V = \mathrm{Sp}(v_1, \dots, v_r, v_{r+1}, w_{i_{r1}}, \dots, w_{i_{r(\ell-1)}}, w_{i_{r(\ell+1)}}, \dots, w_{r(m-r)}).$$

Uudeksi jonoksi $(i_{(r+1)1}, \dots, i_{(r+1)(m-(r+1))})$ voidaan siis valita jono

$$(i_{r1}, \dots, i_{r(\ell-1)}, i_{r(\ell+1)}, \dots, i_{r(m-r)}).$$

Tämä päättää induktioaskeleen ja siten todistuksen. □

8.1.2 Maksimaalinen vapaa jono on virittävä

Lauseen 8.1.21 merkitys on siinä, että se antaa ylärajan äärellisesti viritetyn vektoriavaruuden vapaan jonon pituudelle – jokaista virittävää jonoa voidaan käyttää antamaan tälläinen yläraja. Paljastuu, että jos avaruuden vapaassa jonossa on yhtä monta alkiota kuin jossain virittävässä jonossa, niin myös vapaa jono on virittävä. Tämä on merkittävä havainto, koska virittävän jonon ja vapaan jonon vektoreilla ei ole päällisin puolin mitään tekemistä toistensa kanssa.

Lause 8.1.23. *Olkoon V äärellisesti virittyvä vektoriavaruus ja olkoon $(w_1, \dots, w_k)$ avaruuden V virittävä jono. Tällöin jokainen avaruuden V vapaa jono $(v_1, \dots, v_k)$ on virittävä.*

Huomautus 8.1.24. *Lauseen muotoilu voi herättää kysymyksiä. Miksi ei oteta suoraan lyhintä virittävää jonoa ja kirjoiteta väitettä tämän pituuden avulla. Tätä pituutta kutsutaan avaruuden dimensioksi ja siihen palataan seuraavassa luvussa.*

Todistus. Tehdään vastaoletus, että $V \neq \mathrm{Sp}(v_1, \dots, v_k)$. Tällöin on olemassa sellainen vektori $v \in V$, että $v \notin \mathrm{Sp}(v_1, \dots, v_k)$.

Osoitetaan, että jono $(v_1, \dots, v_k, v)$ on vapaa. Olkoot $a_1, \dots, a_k, a \in \mathbb{R}$ sellaisia lukuja, että

$$a_1 v_1 + \dots + a_k v_k + av = 0.$$

Osoitetaan, että $a = 0$. Tehdään vastaoletus, että näin ei ole. Tällöin

$$v = \frac{1}{a}(a_1 v_1 + \dots + a_k v_k) \in \mathrm{Sp}(v_1, \dots, v_k).$$

Tämä on ristiriita vektorin v valinnan perusteella. Näin ollen $a = 0$. Koska $a = 0$ ja jono $(v_1, \dots, v_k)$ on vapaa, niin $a_1 = \dots = a_k = 0$. Näin ollen jono $(v_1, \dots, v_k, v)$ on vapaa.

Tällöin jono $(v_1, \dots, v_k, v)$ on sellainen vapaa jono, joka on pidempi kuin virittävä jono $(w_1, \dots, w_k)$. Tämä on ristiriita lauseen 8.1.21 kanssa. Vastaoletus $V \neq \mathrm{Sp}(v_1, \dots, v_k)$ on siis väärä ja $V = \mathrm{Sp}(v_1, \dots, v_k)$. Jono $(v_1, \dots, v_k)$ virittää siis avaruuden V. □

Tulosta voidaan ajatella myös toisinpäin: Jos virittävässä jonossa on yhtä monta alkiota kuin jossain vapaassa jonossa, niin virittävä jono on jo itsessään vapaa.

Lause 8.1.25. *Olkoon V äärellisesti virittyvä vektoriavaruus ja olkoon $(v_1, \dots, v_k)$ avaruuden V vapaa jono. Tällöin jokainen avaruuden V virittävä jono $(w_1, \dots, w_k)$ on vapaa.*

Todistus. Tehdään vastaoletus, että $(w_1, \dots, w_k)$ ei ole vapaa. Lemman 8.1.22 perusteella on olemassa sellainen indeksi $j \in \{1, \dots, k\}$, että

$$\mathrm{Sp}(w_1, \dots, w_{j-1}, w_{j+1}, \dots, w_k) = \mathrm{Sp}(w_1, \dots, w_k).$$

Tällöin $(w_1, \dots, w_{j-1}, w_{j+1}, \dots, w_k)$ on virittävä jono, joka on lyhyempi kuin vapaa jono $(v_1, \dots, v_k)$. Tämä on ristiriita, joten jono $(w_1, \dots, w_k)$ on vapaa. □

8.1.3 Vektoriavaruuksien kantalause

Lineaarialgebran päätulos äärellisesti viritettyille vektoriavaruuksille muotoiltuna on seuraava ns. kantalause, joka sanoo, että jokaisella vektoriavaruudella on kanta.[1]

Lause 8.1.26 (Vektoriavaruuksien kantalause). *Äärellisesti viritetyllä vektoriavaruudella on kanta.*

[1] Lukijan ei tarvitse huolestua rajoituksesta, että avaruus on äärellisesti viritetty. Jokaisella vektoriavaruudella on kanta. Tätä yleisempää tulosta ei kuitenkaan tässä yhteydessä todisteta.

Tämä tulos seuraa hieman tarkemmasta tuloksesta, että jokaisesta virittävästä jonosta voidaan löytää kanta.

Lause 8.1.27. *Olkoon V vektoriavaruus ja $(v_1, \dots, v_k)$ avaruuden V virittävä jono. Tällöin on olemassa $d \in \{1, \dots, k\}$ ja sellaiset indeksit $1 \le i_1 < \dots < i_d \le k$, että $(v_{i_1}, \dots, v_{i_d})$ on avaruuden V kanta.*

Todistus. Olkoon $d \in \{1, \dots, k\}$ pienin sellainen luku, että on olemassa sellaiset indeksit $1 \le i_1 < \dots < i_d \le k$, että

$$\mathrm{Sp}(v_{i_1}, \dots, v_{i_d}) = \mathrm{Sp}(v_1, \dots, v_k).$$

Osoitetaan, että $(v_{i_1}, \dots, v_{i_d})$ on vapaa jono. Jos näin ei ole, niin $(v_1, \dots, v_{i_d})$ on sidottu ja lemman 8.1.22 on olemassa sellainen indeksi $j \in \{1, \dots, d\}$, että

$$\mathrm{Sp}(v_{i_1}, \dots, v_{i_{j-1}}, v_{i_{j+1}}, \dots, v_{i_d}) = \mathrm{Sp}(v_{i_1}, \dots, v_{i_d}).$$

Tämä on ristiriita luvun d minimaalisuuden kanssa. Näin ollen $(v_{i_1}, \dots, v_{i_d})$ on vapaa jono. □

Todistetaan nyt vektoriavaruuksien kantalause eli lause 8.1.26.

Lauseen 8.1.26 todistus. Olkoon V äärellisesti viritetty vektoriavaruus. Jos $V = \{0\}$, niin tällöin avaruudella V on tyhjä kanta (). Voidaan siis olettaa, että $V \neq \{0\}$. Tällöin määritelmän mukaan on olemassa sellaiset vektorit $v_1, \dots, v_k \in V$, että $V = \mathrm{Sp}(v_1, \dots, v_k)$. Lauseen 8.1.27 perusteella on olemassa $d \in \{1, \dots, k\}$ ja sellaiset indeksit $1 \le i_1 < \dots < i_d \le k$, että $(v_{i_1}, \dots, v_{i_d})$ on avaruuden V kanta. □

8.2 Vektoriavaruuden dimensio

Äärellisesti viritetyn vektoriavaruuden dimensio tullaan määritelemään kanta-alkioiden määrän avulla. Näin ollen tulee ensin osoittaa, että äärellisesti viritetyn vektoriavaruuden kaikissa kannoissa on yhtä monta alkiota.

Lause 8.2.1. *Olkoon V äärellisesti viritetty vektoriavaruus. Tällöin jokaisessa avaruuden V kannassa on yhtä monta alkiota.*

Todistus. Koska V on äärellisesti viritetty, niin sillä on kanta lauseen 8.1.26 perusteella. Olkoot nyt $(v_1, \dots, v_k)$ ja $(w_1, \dots, w_m)$ avaruuden V kantoja. Koska jono $(v_1, \dots, v_k)$ on vapaa ja $(w_1, \dots, w_m)$ on virittävä, niin lauseen 8.1.21 perusteella $k \le m$. Vastaavasti päätellään, että $m \le k$. Näin ollen $m = k$. Jokaisessa avaruuden V kannassa on siis yhtä monta alkiota. □

Määritelmä 8.2.2. *Äärellisesti viritetyn vektoriavaruuden V* dimensio $\dim V$ *on alkioiden määrä avaruuden V kannassa. Äärellisesti viritettyä vektoriavaruutta kutsutaan* äärellisulotteiseksi*. Vektoriavaruutta, jonka dimensio on n, kutsutaan* n-ulotteiseksi*.*

Esimerkki 8.2.3. *Avaruuden* $\mathbb{R}^{n\times 1}$ *dimensio on* n*, koska standardikannassa* $(e_1,\dots,e_n)$ *on* n *alkiota. Vastaavasti myös avaruuden* $\mathbb{R}^n$ *dimensio on* n*, koska sen standardikannassa* $(e_1,\dots,e_n)$ *on myös* n *alkiota, missä* $e_i = (\delta_{1i},\dots,\delta_{ni})$ *jokaisella* $i \in \{1,\dots,n\}$*. Tässä* δ_{ji} *on sama Kroneckerin delta kuin luvussa 2.*

Esimerkki 8.2.4. *Suora*

$$L = \{tv \in \mathbb{R}^n : t \in \mathbb{R}\}$$

missä $v \neq 0$*, on* 1*-ulotteinen. Kannaksi voidaan valita mikä tahansa vektori* tv*, missä* $t \neq 0$*.*

Esimerkki 8.2.5. *Taso*

$$P = \{tv + sw \in \mathbb{R}^n : t, s \in \mathbb{R}\}$$

missä vektorit v *ja* w *ovat lineaarisesti riippumattomia, on* 2*-ulotteinen. Kannaksi voidaan valita esimerkiksi* (v, w)*.*

Lauseiden 3.8.5 ja B.0.3 tulokset voidaan nyt ilmaista dimension avulla. Kirjataan nämä esimerkeiksi aliavaruuksien dimensioista.

Esimerkki 8.2.6. *Olkoon* $A \in \mathbb{R}^{m\times n}$ *matriisi. Lauseen 3.8.5 perusteella sarakeavaruudella* $\mathrm{Col}(A)$ *on kanta, jossa on yhtä monta alkiota kuin matriisin* A *supistettua porrasmuotoa* B *vastaavassa yhtälöryhmässä* $\left[B \mid 0\right]$ *on sidottuja muuttujia. Näin ollen* $\dim \mathrm{Col}(A)$ *on yhtälöryhmän* $\left[B \mid 0\right]$ *sidottujen muuttujien lukumäärä.*

Esimerkki 8.2.7. *Olkoon* $A \in \mathbb{R}^{m\times n}$ *matriisi. Lauseen B.0.3 perusteella nolla-avaruudella* $\mathrm{Null}(A)$ *on kanta, jossa on yhtä monta alkiota kuin matriisin* A *supistettua porrasmuotoa* B *vastaavassa yhtälöryhmässä* $\left[B \mid 0\right]$ *on vapaita muuttujia. Näin ollen* $\dim \mathrm{Null}(A)$ *on yhtälöryhmän* $\left[B \mid 0\right]$ *vapaiden muuttujien lukumäärä.*

8.2.1 Aliavaruuden dimensio on korkeintaan koko avaruuden dimensio

Tarkasteltaessa äärellisulotteisen vektoriavaruuden aliavaruuksia, on intuitiivisesti selvää, että myös aliavaruudet ovat äärellisulotteisia ja että niiden dimensio on korkeintaan koko avaruuden dimensio. Todistetaan tämä intuitiivinen havainto nyt tarkasti.

Lause 8.2.8. *Olkoon* V *äärellisulotteinen vektoriavaruus. Tällöin jokainen aliavaruus* $W \subset V$ *on äärellisulotteinen ja* $\dim W \leq \dim V$*. Lisäksi jokainen aliavaruuden* W *kanta* $(v_1,\dots,v_k)$ *voidaan laajentaa avaruuden* V *kannaksi* $(v_1,\dots,v_k,v_{k+1},\dots,v_m)$*.*

Todistus. Osoitetaan ensin, että W on äärellisesti viritetty.

Osoitetaan ensin, että jokaisessa avaruuden W vapaassa jonossa $(w_1,\dots w_k)$ on korkeintaan $\dim V$ jäsentä. Tehdään vastaoletus, että $k > \dim V$. Tällöin $(w_1,\dots,w_k)$ on avaruuden V vapaa jono, joka on pidempi kuin jokin virittävä jono. Tämä on ristiriita. Näin ollen $k \leq \dim V$.

Koska W on äärellisesti viritetty, niin se on äärellisulotteinen. Tämä päättää ensimmäisen väitteen todistuksen.

Olkoon nyt $(v_1, \ldots, v_k)$ avaruuden W kanta. Jos $W = V$, niin $(v_1, \ldots, v_k)$ on avaruuden V kanta. Oletetaan nyt, että $W \neq V$.

Olkoon nyt $m \in \{k, \ldots, n\}$ suurin sellainen luku, että on olemassa avaruuden V vapaa jono $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_m)$. Kiinnitetään nyt sellaiset vektorit $v_{k+1}, \ldots, v_m \in V$, että jono $(v_1, \ldots, v_m)$ on vapaa. Tällöin $(v_1, \ldots, v_m)$ on avaruuden $\mathrm{Sp}(v_1, \ldots, v_m)$ kanta.

Osoitetaan, että $\mathrm{Sp}(v_1, \ldots, v_m) = V$. Jos $\mathrm{Sp}(v_1, \ldots, v_m) \neq V$, niin on olemassa sellainen vektori $v_{m+1} \in V \setminus \mathrm{Sp}(v_1, \ldots, v_m)$. Tällöin joukon $\mathrm{Sp}(v_1, \ldots, v_m)$ määritelmän perusteella, v_{m+1} ei ole vektoreiden $v_1, \ldots, v_m$ lineaarikombinaatio. Näin ollen jono $(v_1, \ldots, v_{m+1})$ on vapaa. Tämä on ristiriita luvun m valinnan perusteella. Näin ollen $\mathrm{Sp}(v_1, \ldots, v_m) = V$. □

Lauseella 8.2.8 on mielenkiintoinen korollaari, joka sanoo, että mikäli äärellisulotteisen vektoriavaruuden aliavaruudella on sama dimensio kuin koko avaruudella, niin se on jo itsessään koko avaruus.

Korollaari 8.2.9. *Olkoon V äärellistulotteinen vektoriavaruus ja $W \subset V$ sellainen aliavaruus, että* $\dim W = \dim V$. *Tällöin $W = V$.*

Todistus. Olkoon $(w_1, \ldots, w_k)$ aliavaruuden W kanta. Tällöin $(w_1, \ldots, w_k)$ voidaan laajentaa avaruuden V kannaksi $(w_1, \ldots, w_k, w_{k+1}, \ldots, w_m)$. Koska $m = \dim V = \dim W = k$, niin $(w_1, \ldots, w_k)$ on avaruuden V kanta, eli $W = \mathrm{Sp}(w_1, \ldots, w_k) = V$. □

Huomautus 8.2.10. *Edellisen todistuksen voi kirjoittaa monin eri tavoin käyttäen tämän ja edellisten lukujen tuloksia.*

8.3 Aliavaruuksien summa ja suora summa

Aliavaruuksien summa ja suora summa ovat käsitteitä, jotka tulevat toistumaan kurssin aikana useasti. Tässä luvussa suoraa summaa tarvitaan aliavaruuksien dimensiolauseen todistamiseen. Myöhemmissä luvuissa aliavaruuksien suoraa summaa käytetään ortogonaaliprojektioiden ja lineaarikuvausten ominaisarvojen yhteydessä.

Määritelmä 8.3.1. *Vektoriavaruuden V aliavaruuksien U ja W* summa *on aliavaruus*

$$U + W = \{u + w \in V : u \in U,\ w \in W\}.$$

Määritelmä 8.3.2. *Vektoriavaruuden V aliavaruuksien U ja W summa on* suora, *jos $U \cap W = \{0\}$. Tällöin merkitään $U \oplus W = U + W$.*

Summan ja suoran summan eroa on havainnollistettu kuvassa 8.1.

Näiden kahden käsitteen tulkita on seuraava. Aliavaruuksien U ja W summa $U + W$ on se aliavaruus, jonka alkiot voidaan kirjoittaa aliavaruuksien U ja W alkioiden summana. Aliavaruuksien summa on suora, jos avaruuden $U + W$ alkiot voidaan kirjoittaa yksikäsitteisesti aliavaruuksien U ja W alkioiden summana. Perustellaan tämä jälkimmäinen väite tarkasti.

Kuva 8.1: Vasemmalla: Aliavaruudet U ja W, joiden summa ei ole suora. Oikealla: Aliavaruudet U ja W, joiden summa on suora

Lause 8.3.3. *Olkoon V vektoriavaruus ja olkoot U ja W sen aliavaruuksia. Tällöin $U + W = U \oplus W$, jos ja vain jos jokaisella $v \in U + W$ on olemassa yksikäsitteiset sellaiset vektorit $u \in U$ ja $w \in W$, että $v = u + w$.*

Todistus. Oletetaan ensin, että summa $U + W$ on suora, eli että $U \cap W = \{0\}$. Olkoon $v \in U + W$ ja oletetaan, että $u, u' \in U$ ja $w, w' \in W$ ovat sellaisia, että $v = u + w$ ja $v = u' + w'$.

Osoitetaan, että $u = u'$ ja $v = v'$. Koska $u + w = v = u' + w'$, niin $u - u' = w' - w$. Koska U ja W ovat aliavaruuksia, niin $u - u' \in U$ ja $w' - w \in W$. Näin ollen $u - u' \in U \cap W$ ja $w' - w \in U \cap W$. Koska $U \cap W = \{0\}$, niin $u - u' = 0$ ja $w' - w = 0$, eli $u = u'$ ja $w = w'$.

Oletetaan nyt, että jokaisella $v \in U + W$ on olemassa yksikäsitteiset sellaiset vektorit $u \in U$ ja $w \in W$, että $v = u + w$. Osoitetaan, että $U \cap W = \{0\}$.

Olkoon $v \in U \cap W$. Valitaan $u = v \in U$ ja $w = 0 \in W$. Tällöin $v = u + w$. Toisaalta voidaan valita $u' = 0 \in U$ ja $w' = v \in W$, jolloin $v = u' + w'$. Koska $u = u'$ ja $w = w'$, niin $v = u = u' = 0$. $\square$

Muutama kommentti on paikallaan.

Huomautus 8.3.4. *Aliavaruuksien U ja W summa $U + W$ todellakin on aliavaruus. Tämä havaitaan seuraavasti. Olkoot $v, v' \in U + W$. Tällöin $v = u + w$ ja $v' = u' + w'$ joillain $u, u' \in U$ ja $w, w' \in W$. Näin ollen jokaisella $a \in \mathbb{R}$ pätee*

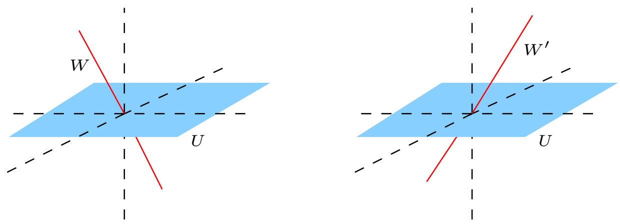
$$av + v' = a(u + w) + u' + w' = (au + u') + (aw + w'),$$

missä $au + u' \in U$ ja $aw + w' \in W$. Näin ollen $av + v' \in U + W$.

Huomautus 8.3.5. *Aliavaruuksien U ja W leikkaus $U \cap W$ on myös aliavaruus. Tämä havaitaan seuraavasti. Olkoot $v, v' \in U \cap W$ ja $a \in \mathbb{R}$. Tällöin $v, v' \in U$. Koska U on aliavaruus, niin $av + v' \in U$. Vastaavasti $av + v' \in W$. Näin ollen $av + v' \in U \cap W$. Osajoukko $U \cap W$ on siis aliavaruus.*

Huomautus 8.3.6. *Vaikka summaa $U + W$ kutsutaan suoraksi summaksi tapauksessa $U \cap W = \{0\}$, niin tällä ei ole mitään tekemistä kohtisuoruuden kanssa, mikä on omalla tavallaan valitettavaa.*

Huomautus 8.3.7. *Huomautettakoon, että suoran summan tekijät eivät ole yksikäsitteisiä eli yhtälöistä $V = U \oplus W$ ja $V = U \oplus W'$ ei seuraa, että $W = W'$ Katso kuva 8.2.*



Kuva 8.2: Taso U ja sellaiset suorat $W \neq W'$ avaruudessa $\mathbb{R}^3$, joille pätee $\mathbb{R}^3 = U \oplus W$ ja $\mathbb{R}^2 = U \oplus W'$.

Lauseen 8.3.3 avulla lause 8.2.8 voidaan tulkita seuraavasti.

Korollaari 8.3.8. *Olkoon V äärellisulotteinen vektoriavaruus ja $W \subset V$ aliavaruus. Tällöin on olemassa sellainen aliavaruus $U \subset V$, että $V = W \oplus U$.*

Todistus. Olkoon $(v_1, \ldots, v_k)$ aliavaruuden W kanta. Lauseen 8.2.8 perusteella kanta jono $(v_1, \ldots, v_k)$ voidaan jatkaa avaruuden V kannaksi $(v_1, \ldots, v_m)$. Olkoon $U = \mathrm{Sp}(v_{k+1}, \ldots, v_m)$. Tällöin

$$W + U = \mathrm{Sp}(v_1, \ldots, k) + \mathrm{Sp}(v_{k+1}, \ldots, v_m) = \mathrm{Sp}(v_1, \ldots, v_m) = V.$$

Osoitetaan nyt, että $W \cap U = \{0\}$. Olkoon $v \in W \cap U$. Tällöin on olemassa sellaiset luvut $a_1, \ldots, a_m \in \mathbb{R}$, että $v = a_1v_1 + \cdots + a_kv_k$ ja $v = a_{k+1}v_{k+1} + \cdots + a_mv_m$. Koordinaattien yksikäsitteisyyden vuoksi $a_1 = \cdots = a_k = a_{k+1} = \cdots = a_m = 0$. Näin ollen $v = 0$. ☐

8.3.1 Esimerkkejä suorista summista

Tarkastellaan nyt kahta sarakeavaruuksiin perustuvaa esimerkkiä, jotka havainnollistavat suoraa summaa sarakeavaruuden $\mathbb{R}^{4\times1}$ aliavaruuksien avulla. Seuraavissa esimerkeissä

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \; v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \text{ ja } v_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

ovat avaruuden $\mathbb{R}^{4\times1}$ sarakevektoreita ja

$$V = \mathrm{Sp}(v_1, v_2, v_3) \subset \mathbb{R}^{4\times1}$$

niiden virittämä aliavaruus. Vektorit v_1, v_2, v_3 ovat toisiaan vastaan kohtisuorassa, joten ne ovat lineaarisesti riippumattomia. Näin ollen (v_1, v_2, v_3) on aliavaruuden V kanta.

Esimerkki 8.3.9. *Tarkastellaan aliavaruuksia*

$$U = \mathrm{Sp}(v_1, v_2) \subset V \text{ ja } W = \mathrm{Sp}(v_3).$$

Tällöin

$$V = \mathrm{Sp}(v_1, v_2, v_3) = \mathrm{Sp}(v_1, v_2) + \mathrm{Sp}(v_3) = U + W.$$

Osoitetaan, että tämä summa on itseasiassa suora eli

$$V = U \oplus W.$$

Olkoon $v \in V$ ja olkoot (x_1, x_2, x_3) vektorin v koordinaatit kannassa (v_1, v_2, v_3). Olkoot $u = x_1v_1 + x_2v_2 \in U$ ja $w = x_3v_3 \in W$. Tällöin $v = u + w$. Osoitetaan vielä, että tämä on ainoa tapa esittää vektori v aliavaruuksien U ja W vektoreiden summana. Olkoot $u' \in U$ ja $w' \in W$ sellaisia vektoreita, että $v = u' + w'$. Tällöin $u' = y_1v_1 + y_2v_2$ ja $w' = y_3v_3$, missä $y_1, y_2, y_3 \in \mathbb{R}$.[2] Tällöin

$$v = u' + w' = y_1v_1 + y_2v_2 + y_3v_3$$

eli (y_1, y_2, y_3) ovat vektorin v koordinaatit kannassa (v_1, v_2, v_3). Vektorien koordinaattien yksikäsitteisyyden nojalla saadaan, että $y_1 = x_1$, $y_2 = x_2$ ja $y_3 = x_3$. Näin ollen $u' = u$ ja $w' = w$.

Koska jokaisella vektorilla on yksikäsitteinen esitys aliavaruuksien U ja W vektoreiden summana, niin summa $U + W$ on suora eli $U \oplus V$.

Tarkastellaan nyt esimerkkiä, jossa summa ei ole suora.

Esimerkki 8.3.10. *Olkoot*

$$U = \mathrm{Sp}(v_1, v_2) \text{ ja } W = \mathrm{Sp}(v_2, v_3).$$

Tällöin $V = U + W$, mutta summa ei ole suora. Perustellaan tämä kahdella tavalla. Vektori v_2 voidaan kirjoittaa sekä muodossa $v_2 = u+w$, missä $u = v_2 \in U$ ja $w = 0 \in W$, ja $v_2 = u' + w'$, missä $u' = 0 \in U$ ja $w' = v_2 \in W$. Näin ollen summa $U + W$ ei ole suora lauseen 8.3.3 perusteella.

Toisaalta

$$U \cap W = \mathrm{Sp}(v_2) \neq \{0\},$$

joten summa $U + W$ ei ole suora suoran summan määritelmän perusteella.

Huomautus 8.3.11. *Myös useamman aliavaruuden suora summa voidaan määritellä. Tätä käsitellään liitteessä D.*

[2] Huomaa, että (y_1, y_2) ovat itseasiassa vektorin u' koordinaatit aliavaruuden U kannassa (v_1, v_2) ja $(y.3)$ on vektorin w' koordinaatti kannassa (v_3).

8.3.2 Projektiot ja suorat summat

Projektiot ovat lineaarikuvauksia, jotka liittyvät läheisesti suoriin summiin. Aloitetaan motivoivalla tuloksella ennen varsinaista määritelmää.

Lause 8.3.12. *Olkoon V vektoriavaruus ja olkoot U ja W sellaisia aliavaruuksia, että $V = U \oplus W$. Tällöin on olemassa yksikäsitteiset sellaiset lineaarikuvaukset $\pi_U\colon V \to V$ ja $\pi_W\colon V \to V$, joille pätee $\operatorname{im} \pi_U = U$, $\operatorname{im} \pi_W = W$ ja*

$$v = \pi_U(v) + \pi_W(v) \tag{8.1}$$

jokaisella $v \in V$.

Todistus. Määritellään aluksi kuvaukset π_U ja π_V. Olkoon $v \in V$. Tällöin on olemassa yksikäsitteiset sellaiset vektorit $u \in U$ ja $w \in W$, että $v = u + w$. Asetetaan $\pi_U(v) = u$ ja $\pi_W(v) = w$. Koska $u \in U$ ja $w \in W$ olivat yksikäsitteisiä, niin kuvaukset $\pi_U\colon V \to V$ ja $\pi_W\colon V \to V$ ovat hyvin määriteltyjä. Lisäksi $\pi_U(v) + \pi_W(v) = v$.

Kuvaukset π_U ja π_W ovat myös yksikäsitteiset sellaiset kuvaukset, että $\pi_U(v) + \pi_W(v) = v$ jokaisella $v \in V$. Tämä seuraa siitä, että $V = U \oplus W$. Perustellaan tämä väite kuitenkin vielä tarkasti. Olkoot $\widetilde{\pi}_U\colon V \to V$ ja $\widetilde{\pi}_W\colon V \to V$ sellaiset kuvaukset, että jokaisella $v \in V$ pätee $\widetilde{\pi}_U(v) \in U$, $\widetilde{\pi}_W(v) \in W$ ja $\widetilde{\pi}_U(v) + \widetilde{\pi}_W(v) = v$. Osoitetaan, että $\widetilde{\pi}_U = \pi_U$ ja että $\widetilde{\pi}_W = \pi_W$. Riittää osoittaa, että $\widetilde{\pi}_U(v) = \pi_U(v)$ ja $\widetilde{\pi}_W(v) = \pi_W(v)$ jokaisella $v \in V$.

Olkoon $v \in V$ ja olkoot $u \in U$ ja $w \in W$ ne yksikäsitteiset vektorit, että $v = u + w$. Vektoreiden u ja w yksikäsitteisyydestä seuraa, että $\widetilde{\pi}_U(v) = u$ ja $\pi_W(v) = w$. Koska kuvausten π_U ja π_W määritelmän mukaan $\pi_U(v) = u$ ja $\pi_W(v) = w$, niin $\widetilde{\pi}_U(v) = \pi_U(v)$ ja $\widetilde{\pi}_W(v) = \pi_W(v)$. Näin ollen $\widetilde{\pi}_U = \pi_U$ ja $\widetilde{\pi}_W = \pi_W$. Kuvaukset π_U ja π_W ovat siis yksikäsitteisiä.

Osoitetaan nyt, että kuvaukset π_U ja π_W ovat lineaarisia. Olkoot $v, v' \in V$ ja $a \in \mathbb{R}$. Osoitetaan, että $\pi_U(av + v') = a\pi_U(v) + \pi_U(v')$ ja $\pi_W(av + v') = a\pi_W(v) + \pi_W(v')$.

Koska $v = \pi_U(v) + \pi_W(v)$ ja $v' = \pi_U(v') + \pi_W(v')$, niin

$$av + v' = a(\pi_U(v) + \pi_W(v)) + \pi_U(v') + \pi_W(v') = (a\pi_U(v) + \pi_U(v')) + (a\pi_W(v) + \pi_W(v')).$$

Toisaalta

$$av + v' = \pi_U(av + v') + \pi_W(av + v'),$$

missä $\pi_U(av + v') \in U$ ja $\pi_W(av + v') \in W$.

Koska $a\pi_U(v) + \pi_U(v') \in U$ ja $a\pi_W(v) + \pi_W(v') \in W$, niin lauseen 8.3.3 nojalla

$$\pi_U(av + v') = a\pi_U(v) + \pi_U(v')$$

ja

$$\pi_W(av + v') = a\pi_W(v) + \pi_W(v').$$

Kuvaukset π_U ja π_W ovat siis lineaarisia. □

Huomautus 8.3.13. *Kaava* (8.1) *voidaan myös kirjoittaa muodossa* $\mathrm{id}_V = \pi_U + \pi_W$.

Kuvauksia $\pi_U \colon V \to V$ ja $\pi_V \colon V \to V$ kutsutaan suoran summan $V = U \oplus W$ *tekijäprojektioiksi.*[3] Näitä kuvauksia kutsutaan projektioiksi, koska ne ovat projektioita seuraavan määritelmän mielessä.

Määritelmä 8.3.14. *Lineaarikuvaus $\pi \colon V \to V$ on* projektio aliavaruudelle $W \subset V$, *jos $\pi \circ \pi = \pi$ ja $\operatorname{im} \pi = W$.*

Seuraava lemma jätetään harjoitustehtäväksi.

Lemma 8.3.15. *Suoran summan $V = U \oplus W$ tekijäprojektiot $\pi_U \colon V \to V$ ja $\pi_W \colon V \to V$ ovat projektiota.*

Projektioilla on kaksi yleistä ominaisuutta.

Lause 8.3.16. *Olkoon $\pi \colon V \to V$ projektio. Tällöin*

1. $\pi|_{\operatorname{im} \pi} = \operatorname{id}$ *ja*

2. $V = \ker \pi \oplus \operatorname{im} \pi$.

Todistus. Olkoon $v \in \operatorname{im} \pi$. Tällöin on olemassa sellainen $v' \in V$, että $\pi(v') = v$. Tällöin $\pi(v) = \pi(\pi(v')) = \pi(v') = v$. Näin ollen $\pi|_{\operatorname{im} \pi} = \operatorname{id}$. Tämä todistaa ensimmäisen väitteen.

Olkoon nyt $v \in V$ ja olkoon $v' = v - \pi(v) \in V$. Koska $\pi(v') = \pi(v - \pi(v)) = \pi(v) - \pi(\pi(v)) = \pi(v) - \pi(v) = 0$, niin $v' \in \ker \pi$. Näin ollen $v = v' + \pi(v) \in \ker \pi + \operatorname{im} \pi$. Osoitetaan nyt, että summa on suora. Olkoon $v \in \ker \pi \cap \operatorname{im} \pi$. Koska $v \in \operatorname{im} \pi$, niin on olemassa sellainen $v' \in V$, että $\pi(v') = v$. Koska $v \in \ker \pi$, niin $v = \pi(v') = \pi(\pi(v')) = \pi(v) = 0$. Näin ollen $\ker \pi + \operatorname{im} \pi = \ker \pi \oplus \operatorname{im} \pi$. □

8.4 Aliavaruuksien dimensiolause

Aliavaruuksien dimensiolause sanoo, että aliavaruuksien U ja V summan $U + V$ dimensio voidaan laskea aliavaruuksien U, V ja $U \cap V$ dimensioista.

Lause 8.4.1 (Aliavaruuksien dimensiolause). *Olkoot U ja W vektoriavaruuden V äärellisesti viritettyjä aliavaruuksia. Tällöin*

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Osoitetaan tätä tulosta varten, että aliavaruuksien suoran summan dimensio on aliavaruuksien dimensioiden summa. Aloitetaan todistamalle lemma virittämisen ja aliavaruuksien summan välisestä yhteydestä.

Lemma 8.4.2. *Olkoot $U = \operatorname{Sp}(u_1, \ldots, u_k)$ ja $W = \operatorname{Sp}(w_1, \ldots, w_m)$ vektoriavaruuden V äärellisesti viritettyjä aliavaruuksia. Tällöin $U + W = \operatorname{Sp}(u_1, \ldots, u_k, w_1, \cdots, w_m)$.*

[3] Tässä sana “tekijä” tulee englanninkielisestä termistä “factor”.

Todistus. Olkoon $v \in U + W$. Tällöin $v = u + w$, missä $u \in U$ ja $w \in W$. Näin ollen $v = a_1 u_1 + \cdots + a_k u_k + b_1 w_1 + \cdots + b_m w_m$ joillain $a_1, \ldots, a_k, b_1, \ldots, b_m \in \mathbb{R}$. Näin ollen $U + W \subset \mathrm{Sp}(u_1, \ldots, u_k, w_1, \ldots, w_m)$.

Koska $U \subset U + W$ ja $W \subset U + W$, niin jokaisella $i \in \{1, \ldots, k\}$ pätee $u_i \in U + W$ ja jokaisella $j \in \{1, \ldots, m\}$ pätee $w_j \in U + W$. Näin ollen $\mathrm{Sp}(u_1, \ldots, u_k, w_1, \ldots, w_m) \subset U + W$. □

Aliavaruuksien dimensionlause suorille summille seuraa suoraan seuraavasta lauseesta, joka karakterisoi summan suoruuden kantojen avulla.

Lause 8.4.3. *Olkoon V vektoriavaruus ja olkoot U ja W avaruuden V äärellisulotteisia aliavaruuksia. Olkoot lisäksi $(u_1, \ldots, u_k)$ on aliavaruuden U kanta ja $(w_1, \ldots, w_m)$ aliavaruuden W kanta. Tällöin aliavaruuksien U ja W summa $U + W$ on suora, jos ja vain jos $(u_1, \ldots, u_k, v_1, \ldots, v_m)$ on aliavaruuden $U + W$ kanta*

Todistus. Oletetaan ensin, että summa $U + W$ on suora. Riitää osoittaa, että jokaisella $v \in U \oplus W$ on yksikäsitteiset sellaiset luvut $x_1, \ldots, x_m, y_1, \ldots, y_m \in \mathbb{R}$, että

$$v = x_1 u_1 + \cdots + x_k u_k + y_1 v_1 + \cdots + y_m v_m.$$

Olkoon $v \in U + W$. Olkoot $x_1, \ldots, x_k, y_1, \ldots, y_m \in \mathbb{R}$ ja $x'_1, \ldots, x'_k, y'_1, \ldots, y'_m \in \mathbb{R}$ sellaisia lukuja, että

$$v = x_1 u_1 + \cdots + x_k u_k + y_1 v_1 + \cdots + y_m v_m$$

ja

$$v = x'_1 u_1 + \cdots + x'_k u_k + y'_1 v_1 + \cdots + y'_m v_m.$$

Merkitään $u = x_1 u_1 + \cdots + x_k u_k \in U$, $w = y_1 v_1 + \cdots + y_m v_m \in W$, $u' = x'_1 u_1 + \cdots + x'_k u_k$ ja $w' = y'_1 v_1 + \cdots + y'_m v_m$.

Tällöin $v = u + w$ ja $v = u' + w'$. Koska summa $U + W$ on suora, niin lauseen 8.3.3 mukaan $u = u'$ ja $w = w'$. Koska $(u_1, \ldots, u_k)$ on alivaruuden U kanta ja $(w_1, \ldots, w_m)$ on aliavaruuden W kanta, niin $x_i = x'_i$ ja $y_j = y'_j$ kaikilla $i \in \{1, \ldots, k\}$ ja $j \in \{1, \ldots, m\}$.

Oletetaan nyt, että $(u_1, \ldots, u_k, w_1, \ldots, w_m)$ on aliavaruuden $U + W$ kanta. Osoitetaan, että summa $U + W$ on suora. Olkoon $v \in U \cap W$. Tällöin on olemassa sellaiset luvut $x_1, \ldots, x_k, y_1, \ldots, y_m \in \mathbb{R}$, että $v = x_1 u_1 + \cdots + x_k u_k$ ja $v = y_1 w_1 + \cdots + y_m w_m$. Näin ollen vektorilla v on kannassa $(u_1, \ldots, u_k, w_1, \ldots, w_m)$ sekä koordinaatit $(x_1, \ldots, x_k, 0, \ldots, 0)$ että koordinaatit $(0, \ldots, 0, y_1, \ldots, y_m)$. Koordinaattien yksikäsitteisyyden nojalla $x_1 = \cdots = x_k = 0$ ja $y_1 = \cdots = y_m = 0$. Näin ollen $v = 0$ ja $U \cap W = \{0\}$. Summa $U + W$ on siis suora. □

Kuten edellä mainittiin lauseesta 8.4.3 seuraa suoraan aliavaruuksien dimensiolause suorien summien erikoistapauksessa. Kirjataan se korollaariksi.

Korollaari 8.4.4. *Olkoon V vektoriavaruus ja olkoot U ja W avaruuden V sellaisia äärellisulotteisia alaivaruuksia, joiden summa $U + W$ on suora. Tällöin*

$$\dim(U \oplus W) = \dim U + \dim W.$$

Todistus. Olkoon $(u_1, \ldots, u_k)$ aliavaruuden U kanta ja $(w_1, \ldots, w_m)$ aliavaruuden W kanta. Tällöin lauseen 8.4.3 perusteella $(u_1, \ldots, u_k, w_1, \ldots, w_m)$ on aliavaruuden $U \oplus W$ kanta. Näin ollen

$$\dim(U \oplus W) = k + m = \dim U + \dim W.$$

□

Näiden valmistelujen jälkeen ollaan valmiita todistamaan aliavaruuksien dimensiolause.

Lauseen 8.4.1 todistus. Olkoot U ja W avaruuden V äärellisulotteisia aliavaruuksia. Tällöin $U \cap W$ on aliavaruuksien U ja W yhteinen aliavaruus.

Olkoon $(v_1, \ldots, v_\ell)$ aliavaruuden $U \cap W$ kanta. Täydennetään $(v_1, \ldots, v_\ell)$ aliavaruuden U kannaksi $(v_1, \ldots, v_\ell, u_1, \ldots, u_k)$ ja aliavaruuden W kannaksi $(v_1, \ldots, v_\ell, w_1, \ldots, w_m)$. Merkitään $U' = \operatorname{Sp}(u_1, \ldots, u_k)$ ja $W' = \operatorname{Sp}(w_1, \ldots, w_m)$. Tällöin lauseen 8.4.3 perusteella

$$U = (U \cap W) \oplus U'$$

ja

$$W = (U \cap W) \oplus W'.$$

Näin ollen korollaarin 8.4.4 perusteella

$$\dim W = \dim(U \cap W) + \dim W'.$$

Tarkastellaan nyt summaa $U + W$. Koska $U \cap W \subset U$, niin

$$U + W = U + \big((U \cap W) + W'\big) = U + (U \cap W) + W' = U + W',$$

Osoitetaan, että summa $U + W'$ on suora. Riittää osoittaa, että $U \cap W' = \{0\}$. Olkoon $v \in U \cap W'$. Koska $W' \subset W$, niin tällöin $v \in U \cap W$. Koska $W = (U \cap W) \oplus W'$, niin $(U \cap W) \cap W' = \{0\}$. Näin ollen $v = 0$ ja $U \cap W' = \{0\}$. Summa $U + W'$ on siis suora ja korollaarin 8.4.4 perusteella

$$\dim(U \oplus W') = \dim U + \dim W'.$$

Yhdistämällä saadut yhtälöt saadaan

$$\dim(U + W) = \dim(U \oplus W') = \dim U + \dim W' = \dim U + \dim W - \dim(U \cap W).$$

Tämä päättää todistuksen. □

Huomautus 8.4.5. *Lauseen 8.4.1 voi todistaa myös osoittamalla, että aliavaruuksien U ja W kannoista saatava jono $(v_1, \ldots, v_\ell, u_1, \ldots, u_k, w_1, \ldots, w_m)$ on aliavaruuden $U + W$ kanta.*

Luku 9

Lineaarikuvausten teoria

Tässä luvussa käsitellään lineaarikuvausten yleistä teoriaa. Tässä teoriassa ei oikeastaan ole kuin yksi lause eli lineaarikuvausten kantalause. Lause sanoo, että *lineaarikuvaus määräytyy täysin kuvauksen arvoista kanta-alkioilla.* Sen seuraukset ovat moninaiset.

Ensimmäinen seuraus on, että *kaikki äärellisulotteiset vektoriavaruudet, joilla on sama dimensio, ovat keskenään samanlaisia.* Erityisesti *n-ulotteinen vektoriavaruus voidaan samastaa sarakeavaruuden $\mathbb{R}^{n\times 1}$ kanssa.* Jos haluaa kuulostaa pompöösiltä, voi sanoa, että *dimensio on äärellisulotteisen vektoriavaruuden ainoa invariantti.*

Kuten voi arvata, termeille *samanlainen* ja *samaistaa* tulee antaa tarkka matemaattinen sisältö. Vektoriavaruuksien samastus perustuu isomorfian käsitteelle ja sitä käsitellään tarkemmin luvussa 9.3.

Toinen lineaarikuvausten kantalauseen seuraus on, että *äärellisulotteisten vektoriavaruuksien väliset lineaarikuvaukset vastaavat matriiseja.* Itseasiassa tämä analogia tehdään todella tarkaksi:

> Jos lineaarikuvauksen $f\colon V \to W$ lähtö- ja maaliavaruudet samastataan isomorfismeilla sarakeavaruuksien $\mathbb{R}^{n\times 1}$ ja $\mathbb{R}^{m\times 1}$ kanssa, niin samat isomorfimit samastavat itse lineaarikuvauksen f lineaarikuvauksen $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ kanssa.

Matriisia A tullaan kutsumaan kuvauksen f esitysmatriisiksi. Esitysmatriiseja käsitellään luvussa 9.5.

9.1 Lineaarikuvaukset, vapaus, virittäminen ja kannat

Lineaarikuvausten teorian yksi rakennuspalikka on havainto, että bijektiiviset lineaarikuvaukset kuvaavat kannan kannaksi. Tämä jakautuu kahteen osaan: injektiiviset lineaarikuvaukset kuvaavat vapaat jonot vapaiksi jonoiksi ja surjektiiviset lineaarikuvaukset kuvaavat virittävät jonot virittäviksi jonoiksi. Nämä havainnot ovat mielenkiintoisia, koska ne yhdistävät vektoriavaruuksiin liittyvät käsitteet kuvauksiin liittyviin käsitteisiin. Itseasiassa ominaisuus, että lineaarikuvaus kuvaa kannan kannaksi karakterisoi bijektiiviset lineaarikuvaukset. Kirjataan nyt nämä havainnot tarkoiksi tuloksiksi.

Lause 9.1.1. *Olkoon $f: V \to W$ lineaarikuvaus vektoriavaruudesta V vektoriavaruuteen W ja olkoon $(v_1, \ldots, v_n)$ avaruuden V vapaa jono. Jos kuvaus f on injektio, niin $(f(v_1), \ldots, f(v_n))$ on avaruuden W vapaa jono.*

Todistus. Oletetaan, että $f: V \to W$ on injektio, ja osoitetaan, että jono $(f(v_1), \ldots, f(v_n))$ on vapaa.

Olkoot $a_1, \ldots, a_n \in \mathbb{R}$ sellaisia, että

$$a_1 f(v_1) + \cdots + a_n f(v_n) = 0.$$

Tällöin

$$f(a_1 v_1 + \cdots + a_n v_n) = a_1 f(v_1) + \cdots + a_n f(v_n) = 0.$$

Koska f on injektio, niin

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Koska jono $(v_1, \ldots, v_n)$ on vapaa, niin $a_1 = \cdots = a_n = 0$. Näin ollen jono $(f(v_1, \ldots, f(v_n))$ on vapaa. $\square$

Lause 9.1.2. *Olkoon $f: V \to W$ lineaarikuvaus vektoriavaruudesta V vektoriavaruuteen W ja olkoon $(v_1, \ldots, v_n)$ jono, joka virittää avaruuden V. Jos f on surjektio, niin jono $(f(v_1), \ldots, f(v_n))$ virittää avaruuden W.*

Todistus. Olkoon, että $f: V \to W$ on surjektio, ja osoitetaan, että $W = \mathrm{Sp}(f(v_1), \ldots, f(v_n))$.

Olkoon $w \in W$. Koska f on surjektio, niin on olemassa sellainen $v \in V$, että $f(v) = w$. Koska $V = \mathrm{Sp}(v_1, \ldots, v_n)$, niin on olemassa sellaiset luvut $a_1, \ldots, a_n \in \mathbb{R}$, että $v = a_1 v_1 + \cdots + a_n v_n$. Tällöin

$$a_1 f(v_1) + \cdots + a_n f(v_n) = f(a_1 v_1 + \cdots + a_n v_n) = f(v) = w.$$

Näin ollen $W \subset \mathrm{Sp}(f(v_1), \cdots, f(v_n))$ eli jono $(f(v_1), \ldots, f(v_n))$ virittää avaruuden W. $\square$

Korollaari 9.1.3. *Olkoon $f: V \to W$ lineaarikuvaus vektoriavaruudesta V vektoriavaruuteen W ja olkoon $(v_1, \ldots, v_n)$ avaruuden V kanta. Jos f on bijektio, niin $(f(v_1), \ldots, f(v_n))$ on avaruuden W kanta.*

Todistus. Oletetaan, että f on bijektio. Tällöin f on sekä injektio että surjektio. Koska jono $(v_1, \ldots, v_n)$ on kanta, niin se on sekä vapaa jono että virittävä jono. Näin ollen lauseiden 9.1.1 ja 9.1.2 perusteella, jono $(f(v_1), \ldots, f(v_n))$ on sekä vapaa että virittävä. $\square$

Tämän luvun päätulos on, että korollaari 9.1.3 pätee myös toiseen suuntaan, eli että lineaarikuvaus, joka kuvaa kannan kannaksi, on bijektio. Yhdistetään tämä tulos korollaarin 9.1.3 kanssa seuraavaksi lauseeksi.

Lause 9.1.4. *Olkoon $f: V \to W$ lineaarikuvaus vektoriavaruudesta V vektoriavaruuteen W ja olkoon $(v_1, \ldots, v_n)$ avaruuden V kanta. Tällöin f on bijektio, jos ja vain jos $(f(v_1), \ldots, f(v_n))$ on avaruuden W kanta.*

Todistus. Korollaarissa 9.1.3 on osoitettu, että jos f on bijektio, niin $(f(v_1), \ldots, f(v_n))$ on avaruuden W kanta. Näin ollen riittää siis osoittaa, että jos $(f(v_1), \ldots, f(v_n))$ on avaruuden W kanta, niin f on bijektio.

Osoitetaan ensin, että f on injektio. Riittää osoittaa, että $\ker f = \{0\}$. Olkoon $v \in \ker f$ eli olkoon $v \in V$ sellainen vektori, että $f(v) = 0$. Olkoot $(a_1, \ldots, a_n)$ vektorin v koordinaatit kannassa $(v_1, \ldots, v_n)$. Tällöin

$$f(v) = f(a_1 v_1 + \cdots + a_n v_n) = a_1 f(v_1) + \cdots + a_n f(v_n).$$

Näin ollen $(a_1, \ldots, a_n)$ ovat vektorin $f(v)$ koordinaatit kannassa $(f(v_1), \ldots, f(v_n))$. Koska $f(v) = 0$, niin $a_1 = \ldots = a_n = 0$. Näin ollen $v = 0$. Kuvaus f on siis injektio.

Osoitetaan nyt, että f on surjektio. Olkoon $w \in W$. Koska $(f(v_1), \ldots, f(v_n))$ on avaruuden W kanta, niin vektorilla w on koordinaatit $(a_1, \ldots, a_n)$ kannassa $(f(v_1), \ldots, f(v_n))$. Olkoon $v = a_1 v_1 + \cdots + a_n v_n \in V$. Tällöin

$$f(v) = f(a_1 v_1) + \cdots + f(a_n v_n) = a_1 f(v_1) + \cdots + a_n f(v_n) = w.$$

Kuvaus f on siis surjektio. Tämä päättää todistuksen. □

Huomautus 9.1.5. *Näissä todistuksissa käytetiin ainoastaan lineaarikuvausten ominaisuutta, että ne kuvaavat lineaarikombinaatiot lineaarikombinaatioiksi. Tämä ominaisuus yhdistettiin suoraan vapauden ja virittämisen määritelmiin.*

Näiden tulosten lisäksi on mahdollista muotoilla muitakin vastaavia tuloksia. Kirjataan ylös vielä yksi yleinen lemma, jonka todistus jätetään harjoitustehtäviksi.

Lemma 9.1.6. *Olkoon $f\colon V \to W$ lineaarikuvaus vektoriavaruudesta V vektoriavaruuteen W. Avaruuden V jono $(v_1, \ldots, v_n)$ on vapaa, jos jono $(f(v_1), \ldots, f(v_n))$ on vapaa.*

Tässä vaiheessa on hyvä muistaa, että vapaan jonon kuvajonon ei tarvitse olla vapaa ja että virittävän jonon kuvajonon ei tarvitse olla virittävä kuten seuraava yleinen esimerkki osoittaa.

Esimerkki 9.1.7. *Olkoon $A \in \mathbb{R}^{m \times n}$ ja $f_A\colon \mathbb{R}^{n \times 1} \to \mathbb{R}^{m \times 1}$. Olkoon myös $(e_1, \ldots, e_n)$ avaruuden $\mathbb{R}^{n \times 1}$ standarikanta. Jos $m < n$, niin tällöin jono $(f(e_1), \ldots, f(e_n)))$ eli jono $(Ae_1, \ldots, Ae_n)$ ei ole vapaa. Jos taas $m > n$, niin tällöin $(f(e_1), \ldots, f(e_n))$ ei ole virittävä.*

Palautetaan vielä mieleen, että lineaarikuvauksen ei tarvitse olla injektio eikä surjektio.

Esimerkki 9.1.8. *Olkoon*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

ja $f_A\colon \mathbb{R}^{2 \times 1} \to \mathbb{R}^{2 \times 1}$, $x \mapsto Ax$. Tällöin f_A ei ole injektio eikä surjekio.

9.2 Lineaarikuvausten dimensiolause

Lineaarikuvausten dimensiolauseeksi sanotaan tulosta, että lineaarikuvauksen ytimen ja kuvan dimensioiden summa on lähtöavaruuden dimensio.[1]

Lause 9.2.1 (Lineaarikuvausten dimensiolause)**.** *Olkoon V äärellisulotteinen vektoriavaruus ja $f\colon V \to W$ lineaarikuvaus vektoriavaruuteen W. Tällöin*

$$\dim V = \dim \ker(f) + \dim \operatorname{im}(f).$$

Todistuksen idea on esittää avaruus V kahden aliavaruuden $\ker f$ ja U suorana summana, missä aliavaruudella U on sellaiset ominaisuudet, että $f(U) = \operatorname{im} f$ ja $\dim U = \dim \operatorname{im} f$. Tällöin ominaisuudesta $V = \ker f \oplus U$ seuraa suoran summan dimensiolauseen (Korollaari 8.4.4) perusteella, että $\dim V = \dim \ker f + \dim U = \dim \ker f + \dim \operatorname{im} f$. Heuristisesti voidaan ajatella, että aliavaruus $U \subset V$ vastaa aliavaruuden $\operatorname{im} f \subset W$ kopiota avaruudessa V. Tehdään tämä vastaavus nyt tarkasti ja todistetaan tämän jälkeen lause 9.2.1.

Lemma 9.2.2. *Olkoon V äärellisulotteinen vektoriavaruus ja $f\colon V \to W$ lineaarikuvaus vektoriavaruuteen W. Tällöin aliavaruus $\operatorname{im} f \subset W$ on äärellisulotteinen ja on olemassa sellainen aliavaruus $U \subset V$, että $f(U) = \operatorname{im} f$ ja $\dim U = \dim \operatorname{im} f$. Lisäksi kuvauksen f rajoittuma $f|_U\colon U \to f(U)$ on bijektio.*

Todistus. Osoitetaan ensin, että $\operatorname{im} f$ on äärellisulotteinen. Koska V on äärellisulotteinen, niin sillä on kanta. Olkoon $(v_1, \ldots, v_n)$ avaruuden V kanta. Tällöin $\operatorname{im} f = \operatorname{Sp}(f(v_1), \ldots, f(v_n))$. Näin ollen $\operatorname{im} f$ on äärelllisesti viritetty. Näin ollen se on äärellisulotteinen.

Olkoon nyt $(w_1, \ldots, w_m)$ aliavaruuden $\operatorname{im} f$ kanta. Valitaan jokaisella $i \in \{1, \ldots, m\}$ sellainen vektori $u_i \in V$, että $f(u_i) = w_i$. Olkoon $U = \operatorname{Sp}(u_1, \ldots, u_m)$. Koska jono $(u_1, \ldots, u_m)$ on vapaa, niin se on aliavaruuden U kanta lemman 9.1.6 nojalla. Näin ollen $\dim U = m = \dim \operatorname{im} f$.

Osoitetaan vielä, että $f(U) = \operatorname{im} f$. Olkoon $w \in \operatorname{im} f$ ja olkoot $(a_1, \ldots, a_m)$ vektorin w koordinaatit kannassa $(w_1, \ldots, w_m)$. Tällöin $u = a_1u_1 + \cdots + a_mu_m \in U$ ja

$$f(u) = f(a_1u_1 + \cdots + a_mu_m) = a_1f(u_1) + \cdots + a_mf(u_m) = a_1w_1 + \cdots + a_nw_n = w.$$

Näin ollen $\operatorname{im} f \subset f(U)$. Koska $\dim f(U) = \dim \operatorname{im} f$, niin $f(U) = \operatorname{im} f$.

Osoitetaan nyt viimeinen väite, että $f|_U$ on bijektio. Aliavaruus U on itsessään vektoriavaruus ja $(u_1, \ldots, u_m)$ on tämän avaruuden kanta. Koska $f(U) = \operatorname{im} f$, niin rajoittumakuvaus $f|_U\colon U \to f(U)$ kuvaa avaruuden U kannan $(u_1, \ldots, u_m)$ avaruuden $f(U)$ kannaksi $(w_1, \ldots, w_m)$. Näin ollen $f|_U\colon U \to f(U)$ on lauseen 9.1.4 perusteella bijektio. □

Lauseen 9.2.1 todistus. Olkoon aliavaruus $U \subset V$ kuten lemmassa 9.2.2. Osoitetaan, että

$$V = \ker f \oplus U.$$

[1] Tätä lausetta kutsutaan myös lineaarikuvausten peruslauseeksi.

Osoitetaan ensin, että
$$V = \ker f + U.$$
Riittää siis osoittaa, että jokainen $v \in V$ voidaan esittää muodossa $v = v' + u$, missä $v' \in \ker f$ ja $u \in U$. Olkoon $v \in V$. Koska $f(U) = \operatorname{im} f$ ja $f(v) \in \operatorname{im} f$, niin on olemassa sellainen $u \in U$, että $f(u) = f(v)$. Olkoon nyt $v' = v - u$. Selvästi $v = v' + u$. Toisaalta $f(v') = f(v - u) = f(v) - f(u) = 0$ eli $v' \in \ker f$. Näin ollen $V = \ker f + U$.

Osoitetaan nyt, että $\ker f \cap U = \{0\}$. Olkoon $v \in \ker f \cap U$. Koska $v \in \ker f$, niin $f(v) = 0$. Näin ollen $(f|_U)(v) = f(v) = 0$. Koska kuvaus $f|_U$ on injektio, niin näin ollen $v = 0$.

Korollaarin 8.4.4 nojalla
$$\dim V = \dim(\ker f \oplus U) = \dim \ker f + \dim U = \dim \ker f + \dim \operatorname{im} f.$$
Tämä päättää todistuksen. □

9.2.1 Lineaarikuvauksen bijektiivisyyslause

Lineaarikuvausten yleisessä teoriassa tämä tulos käsitellään yleensä isomorfismien yhteydessä. Kirjataan tämä lineaarikuvausten bijektiivisyys lause kuitenkin itsenäisesti, koska se on luvussa 3 esitetyn neliömatriisien kääntyvyyslauseen (lause 3.10.1) suora vastine lineaarikuvausten teoriassa.

Palautetaan mieleen, että neliömatriisien kääntyvyyslauseen todistus perustui yhtälöryhmän supistettuun porrasmuotoon. Lineaarikuvausten tapauksessa tulos seuraa (lähes) suoraan lineaarikuvausten dimensiolauseesta (lause 9.2.1. Tässä tapauksessa on luonnollisinta kirjata ensin tulos, että saman ulotteisten vektoriavaruuksien välinen lineaarikuvaus on injektiivinen, jos ja vain jos se on surjektiivinen.

Lause 9.2.3. *Olkoot V ja W n-ulotteisia vektoriavaruuksia ja olkoon $f\colon V \to W$ lineaarikuvaus. Tällöin f on injektio, jos ja vain jos f on surjektio.*

Todistus. Oletetaan, että f on injektio. Tällöin $\ker f = \{0\}$, joten $\dim \ker f = 0$. Näin ollen
$$\dim \operatorname{im} f = \dim V = n = \dim W.$$
Näin ollen korollaarin 8.2.9 perusteella $\operatorname{im} f = W$. Näin ollen f on surjektio.

Oletetaan nyt, että f on surjektio. Tällöin $\operatorname{im} f = W$, joten
$$\dim \ker f = \dim V - \dim \operatorname{im} f = \dim V - \dim W = 0.$$
Näin ollen $\ker f = \{0\}$ ja f on injektio. □

Seuraava korollaari on haluttu neliömatriisien kääntyvyyslauseen (lause 3.10.1) vastine lineaarikuvauksille.

Korollaari 9.2.4. *Olkoot V ja W n-ulotteisia vektoriavaruuksia ja olkoon $f\colon V \to W$ lineaarikuvaus. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. f *on bijektio,*

2. f *on injektio ja*

3. f *on surjektio.*

9.3 Isomorfismit

Edellisen luvun lopuksi tarkasteltiin bijektiivisiä lineaarikuvauksia. Näitä kutsutaan vektoriavaruuksien teoriassa isomorfismeiksi.

Määritelmä 9.3.1. *Lineaarikuvaus* $f\colon V \to W$ *vektoriavaruudelta* V *vektoriavaruudelle* W *on* isomorfismi, *jos se on bijektiivinen. Vektoriavaruudet* V *ja* W *ovat isomorfisia, jos on olemassa isomorfismi* $f\colon V \to W$.

Aloitetaan muutamalla huomiolla.

Huomautus 9.3.2. *Termi isomorfisuus viittaa samanlaisuuteen. Samoin kuin bijektion voidaan ajatella samastavan joukot, tässä teoriassa isomorfismit samastavat vektoriavaruudet. Tätä ei pidä kuitenkaan ajatella niin, että isomorfiset avaruudet ovat sama avaruus. Analogia joukkoihin tässä yhteydessä on, että kaikki äärelliset joukot, joissa on sama määrä alkioita ovat bijektiivisiä keskenään, mutta eivät sama joukko.*[2] *Samalla tavalla kohta havaitaan, että äärellisulotteisten avaruuksien dimensio määrää näiden avaruuksien isomorfisuuden.*

Huomautus 9.3.3. *Koska isomorfismilla tarkoitetaan lineaarista bijektiota, niin samastaminen voidaan nyt tulkita seuraavasti: Avaruudet* V *ja* W *ovat isomorfisia, jos ne voidaan samastaa joukkoina sellaisella bijektiolla* $f\colon V \to W$*, että* f *samastaa myös avaruuksien* V *ja* W *laskutoimitukset eli kolmikot* $(V, +, \cdot)$ *ja* $(W, +, \cdot)$*. Formaalisti tällä tarkoitetaan sitä, että esimerkiksi yhteenlasku* $+\colon W \times W \to W$ *voidaan kirjoittaa yheenlaskun* $+\colon V \to V \to V$ *ja bijektion* f *avulla muodossa*

$$w + w' = f(f^{-1}(w) + f^{-1}(w')).$$

Huomautus 9.3.4. *Kirjallisuudessa avaruuksien isomorfisuutta merkitään esimerkiksi symbolilla* $\cong$ *eli merkitään* $V \cong W$*, jos avaruudet* V *ja* W *ovat isomorfisia. Näissä luentomuistiinpanoissa tätä merkintää ei käytetä kuin tässä luvussa.*

Kirjataan nyt ylös havainto, että isomorfismin käänteiskuvaus on isomorfismi ja isomorfismien yhdiste on isomorfismi.

Lause 9.3.5. *Olkoot* $f\colon V \to W$ *ja* $g\colon W \to U$ *isomorfismeja. Tällöin*

1. $f^{-1}\colon W \to V$ *ja*

2. $g \circ f\colon V \to U$

[2] Esimerkiksi maatilalla voi olla kolme eläintä ja kerrostalossa kolme rappukäytävää.

ovat isomorfismeja. Lisäksi identtinen kuvaus $\mathrm{id}_V : V \to V$ *on isomorfismi.*

Todistus. Koska f ja g ovat bijektioita, niin myös f^{-1} ja $g \circ f$ ovat bijektioita. Koska f ja g ovat lineaarisia, niin myös f^{-1} ja $g \circ f$ ovat lineaarisia lauseiden 7.2.19 ja 7.2.18 nojalla. Koska identtinen kuvaus id_V on selvästi lineaarinen bijektio, on se isomorfismi. $\square$

Huomautus 9.3.6. *Lauseen 9.3.5 tulkinta on seuraava. Jos on annettu vektoriavaruuksien joukko* $\mathscr{V}$*, niin isomorfisuus määrittelee ekvivalenssirelaation joukkoon* $\mathscr{V}$*, koska kaikille* $V, W, U \in \mathscr{V}$ *pätee*

1. $V \cong V$,

2. $W \cong V$, *jos* $V \cong W$, *ja*

3. $V \cong U$, *jos* $V \cong W$ *ja* $W \cong U$.

Tämä ominaisuus suoraa lauseen 9.3.5 havainnoista. Yksityiskohdat jätetään harjoitustehtäväksi.

Ensimmäinen äärellisulotteisten vektoriavaruuksien isomorfisuutta koskeva havainto on, että jokainen äärellisulotteinen avaruus on isomorfinen samanulotteisen sarakeavaruuden kanssa. Tarvittava isomorfismi löydetään helposti kannan avulla. Kirjataan tämä tärkeä havainto lauseeksi. Palautetaan mieleen, että lineaarikuvaus $f_{(v_1,\dots,v_n)} : \mathbb{R}^{n\times 1} \to V$ on määritelty kaavalla

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto x_1 v_1 + \cdots + x_n v_n$$

Lause 9.3.7. *Olkoon* V *äärellisulotteinen vektoriavaruus ja* $(v_1, \dots, v_n)$ *avaruuden* V *kanta. Tällöin lineaarikuvaus* $\Phi_{(v_1,\dots,v_n)} = f_{(v_1,\dots,v_n)} : \mathbb{R}^{n\times 1} \to V$ *on isomorfismi. Erityisesti avaruudet* $\mathbb{R}^{n\times 1}$ *ja* V *ovat isomorfisia.*

Todistus. Osoitetaan, että lineaarikuvaus $\Phi_{(v_1,\dots,v_n)}$ on bijektio osoittamalla, että se on injektio ja surjektio.

Olkoon

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \ker \Phi.$$

Tällöin

$$0 = \Phi(x) = x_1 v_1 + \cdots + x_n v_n.$$

Koska $(v_1, \dots, v_n)$ on kanta, niin $x_1 = \cdots = x_n = 0$ eli $x = 0$. Näin ollen Φ on injektio.

Olkoon nyt $v \in V$ ja olkoot $(y_1, \dots, y_n)$ vektorin v koordinaatit kannassa $(v_1, \dots, v_n)$. Olkoon

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

Tällöin

$$\Phi_{(v_1,\ldots,v_n)}(y) = y_1v_1 + \cdots + y_nv_n = v.$$

Näin ollen $\Phi_{(v_1,\ldots,v_n)}$ on surjektio. $\square$

Huomautus 9.3.8. *Koska* $\dim V = \dim \mathbb{R}^{n\times 1}$, *niin korollaarin 9.2.4 nojalla riittää osoittaa, että* $\Phi_{(v_1,\ldots,v_n)}$ *on injektio tai surjektio. Todistuksessa korostuu nyt kuitenkin enemmän isomorfismin* $\Phi_{(v_1,\ldots,v_n)}$ *ominaisuus, että se kuvaa sarakevektorin* x *sille vektorille* v*, jonka koordinaatit kannassa* $(v_1,\ldots,v_n)$ *ovat verktorin* x *kertoimet. Erityisesti se kuvaa standardikannan* $(e_1,\ldots,e_n)$ *kannaksi* $(v_1,\ldots,v_n)$*.*

Kirjataan lauseessa 9.3.7 käytetty merkintä yleiseksi merkinnäksi. Huomaa, että tämä merkintä laajentaa merkintää 7.2.9.

Merkintä 9.3.9. *Olkoon* V *äärellisulotteinen vektoriavaruus ja* $(v_1,\ldots,v_n)$ *avaruuden* V *kanta. Isomorfismi* $\Phi_{(v_1,\ldots,v_n)}\colon \mathbb{R}^{n\times 1} \to V$ *on kuvaus* $\Phi_{(v_1,\ldots,v_n)} = f_{(v_1,\ldots,v_n)}$*.*

Koska vektoriavaruuksien isomorfisuus on ekvivalenssirelaatio, niin lauseesta 9.3.7 seuraa, että äärellisulotteiset vektoriavaruudet ovat isomorfisia keskenään, jos ja vain jos niillä on sama dimensio.

Lause 9.3.10. *Äärellisulotteiset vektoriavaruudet* V *ja* W *ovat isomorfisia, jos ja vain jos* $\dim V = \dim W$*.*

Todistus. Oletetaan ensin, että $n = \dim V = \dim W$. Tällöin lauseen 9.3.7 perusteella on olemassa isomorfismi $\Phi_{(v_1,\ldots,v_n)}\colon \mathbb{R}^{n\times 1} \to V$ ja isomorfismi $\Phi_{(w_1,\ldots,w_n)}\colon \mathbb{R}^{n\times 1} \to W$, missä $(v_1,\ldots,v_n)$ on avaruuden V kanta ja $(w_1,\ldots,w_n)$ on avaruuden W kanta. Nyt lauseen 9.3.5 perusteella yhdistetty kuvaus $\Phi_{(w_1,\ldots,w_n)} \circ \Phi^{-1}_{(v_1,\ldots,v_n)}\colon V \to W$ on isomorfismi. Avaruudet V ja W ovat siis isomorfisia.

Oletetaan nyt, että avaruudet V ja W ovat isomorfisia ja osoitetaan, että $\dim V = \dim W$. Olkoon $f\colon V \to W$ isomorfismi. Tällöin $\dim W = \dim \operatorname{im} f$. Koska f on injektio, niin $\dim \ker f = 0$. Näin ollen lineaarikuvausten dimensiolauseen (lause 9.2.1) nojalla

$$\dim V = \dim \ker f + \dim \operatorname{im} f = \dim \operatorname{im} f = \dim W.$$

Tämä päättää todistuksen. $\square$

Lauseen 9.3.10 todistus osoittaa, että lauseen väite ei ole optimaalinen vaan, että isomorfisten vektoriavaruuksien välisistä isomorfismeista voidaan sanoa paljon enemmänkin. Huomataan nimittäin, että isomorfismien $\Phi_{(v_1,\ldots,v_n)}$ ja $\Phi_{(w_1,\ldots,w_n)}$ määritelmissä valittiin avaruudelle V kanta $(v_1,\ldots,v_n)$ ja avaruudelle W kanta $(w_1,\ldots,w_n)$. Suora lasku osoittaa, että isomorfismi $\Phi_{(w_1,\ldots,w_n)} \circ \Phi^{-1}_{(v_1,\ldots,v_n)}\colon V \to W$ vie kannan $(v_1,\ldots,v_n)$ kannalle $(w_1,\ldots,w_n)$. Kirjataan tämä havainto kannan siirrymisestä lemmaksi. Tähän havaintoon palataan luvussa 9.6, jossa käsitellään kannanvaihtomatriiseja.

Lemma 9.3.11. *Olkoot V ja W n-ulotteisia vektoriavaruuksia ja olkoot $(v_1, \ldots, v_n)$ avaruuden V kanta sekä $(w_1, \ldots, w_n)$ avaruuden W kanta. Tällöin isomorfismille $\Psi = \Phi_{(w_1,\ldots,w_n)} \circ \Phi^{-1}_{(v_1,\ldots,v_n)} \colon V \to W$ pätee*

$$\Psi(x_1v_1 + \cdots + x_nv_n) = x_1w_1 + \cdots + x_nw_n$$

kaikilla $x_1, \ldots, x_n \in \mathbb{R}$. Erityisesti $\Psi(v_i) = w_i$ jokaiselle $i \in \{1, \ldots, n\}$.

Todistus. Olkoot $x_1, \ldots, x_n \in \mathbb{R}$ ja

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

Tällöin

$$\begin{aligned} \Psi(x_1v_1 + \cdots + x_nv_n) &= \Psi(\Phi_{(v_1,\ldots,v_n)}(x)) \\ &= (\Phi_{(w_1,\ldots,w_n)} \circ \Phi^{-1}_{(v_1,\ldots,v_n)})(\Phi_{(v_1,\ldots,v_n)}(x)) \\ &= \Phi_{(w_1,\ldots,w_n)}(x) = x_1w_1 + \cdots + x_nw_n. \end{aligned}$$

Tämä todistaa ensimmäisen väitteen. Toinen väite seuraa välittömästi. □

Esimerkki 9.3.12. *Olkoon*

$$V = \{t(1,1,1) + s(1,0,-1) \in \mathbb{R}^3 \colon t, s \in \mathbb{R}\}$$

origon kautta kulkeva taso avaruudessa $\mathbb{R}^3$. Tarkemmin $V = \mathrm{Sp}((1,1,1),(1,0,-1))$ ja vektorit $(1,1,1)$ ja $(1,0,-1)$ muodostavat aliavaruuden V kannan. Näin ollen V on avaruuden $\mathbb{R}^3$ 2-ulotteinen aliavaruus ja siten itsessään 2-ulotteinen vektoriavaruus. Osoitetaan, että V on isomorfinen avaruuden $\mathbb{R}^2$ kanssa.

Olkoon (e_1, e_2) avaruuden $\mathbb{R}^2$ standardikanta. Kantalauseen 9.4.2 nojalla on olemassa yksikäsitteinen lineaarikuvaus $f \colon \mathbb{R}^2 \to V$, jolle pätee $f(e_1) = (1,1,1)$ ja $f(e_2) = (1,0,-1)$. Tällä kuvauksella on itseasiassa kaava

$$(x_1, x_2) \mapsto x_1(1,1,1) + x_2(1,0,-1).$$

Koska f on lineaarikuvaus, niin riittää osoittaa, että f on bijektio.

Osoitetaan ensin, että f on surjektio. Olkoon $v \in V$. Tällöin $v = t(1,1,1) + s(1,0,-1)$, joillakin $t, s \in \mathbb{R}$. Nyt $f(t,s) = v$. Näin ollen f on surjektio.

Osoitetaan nyt, että f on injektio. Lauseen 7.4.6 perusteella riittää osoittaa, että $\ker(f) = \{0\}$. Oletetaan, että $x = (x_1, x_2) \in \mathbb{R}^2$ on sellainen vektori, että $f(x_1, x_2) = 0$ eli että

$$x_1(1,1,1) + x_2(1,0,-1) = 0.$$

Koska $((1,1,1),(1,0,-1))$ on avaruuden V kanta, niin $x_1 = x_2 = 0$. Kuvaus f on siis injektio.

9.4 Lineaarikuvausten kantalause

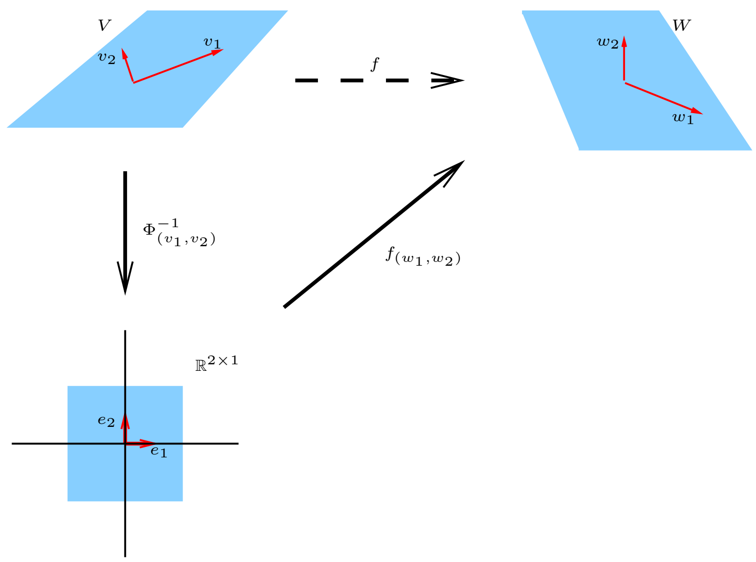
Lineaarikuvausten kantalausee sanoo, että lineaarikuvaus määräytyy täysin kuvauksen arvoista kanta-alkiolla. Lauseessa on kaksi osaa: olemassaolo ja yksikäsitteisyys. Olemassaoloväite sanoo:

> Jos on annettu äärellisulotteinen vektoriavaruus V ja sen kanta $(v_1, \dots, v_n)$ sekä vektoriavaruus W ja vektorit $w_1, \dots, w_n \in W$, niin on olemassa sellainen lineaarikuvaus $f\colon V \to W$, että $f(v_i) = w_i$ jokaisella $i \in \{1, \dots, n\}$.

Lauseen yksikäsitteisyysväite puolestaan sanoo:

> Tälläisiä lineaarikuvauksia $f\colon V \to W$ on täsmälleen yksi ja että kuvauksella f on kaava
> $$f(x_1 v_1 + \dots + x_n v_n) = x_1 w_1 + \dots + x_n w_n$$
> kaikilla $x_1, \dots, x_n \in \mathbb{R}$.

Olemassaolotuloksen voi todistaa lähtien annetusta kaavasta. Tällöin joudutaan kuitenkin perustelemaan, että kuvaus f on hyvin määritelty, eli että se ei riipu kannan $(v_1, \dots, v_n)$ valinnasta. Tämä tekninen vaihe voidaan sivuuttaa, kun huomataan, että kuvaus f saadaan yhdistettynä kuvauksena $f_{(w_1,\dots,w_n)} \circ \Phi^{-1}_{(v_1,\dots,v_n)} \colon V \to W$ kuten kuvassa 9.1.



Kuva 9.1: Yhdistetty kuvaus $f = f_{(w_1,\dots,w_n)} \circ \Phi^{-1}_{(v_1,\dots,v_n)} \colon V \to W$ havainnollistettuna esimerkkitapauksessa $n = \dim V = 2$.

Aloitetaan kantalauseen todistus osoittamalla, että tämä yhdistetty kuvaus on haluttu kuvaus f.

Lemma 9.4.1. . *Olkoot V äärellisulotteinen vektoriavaruus ja $(v_1, \dots, v_n)$ avaruuden V kanta. Olkoot myös W vektoriavaruus ja $w_1, \dots, w_n \in W$. Tällöin yhdistetty kuvaus $f = f_{(w_1,\dots,w_n)} \circ \Phi^{-1}_{(v_1,\dots,v_n)} \colon V \to W$ on lineaarikuvaus, jolle pätee $f(v_i) = w_i$.*

Todistus. Koska f on lineaarikuvausten yhdiste, niin se on lineaarikuvaus. Näin ollen riittää osoittaa, että $f(v_i) = w_i$ jokaisella $i \in \{1, \dots, n\}$.

Olkoon $i \in \{1, \dots, n\}$. Tällöin

$$f(v_i) = (f_{(w_1,\dots,w_n)} \circ \Phi^{-1}_{(v_1,\dots,v_n)})(v_i) = f_{(w_1,\dots,w_n)}(\Phi^{-1}_{(v_1,\dots,v_n)}(v_i)) = f_{(w_1,\dots,w_n)}(e_i) = w_i.$$

Tämä päättää todistuksen. □

Todistetaan nyt lineaarikuvausten kantalause.

Lause 9.4.2 (Lineaarikuvausten kantalause)**.** *Olkoon V n-ulotteinen vektoriavaruus ja olkoon $(v_1, \dots, v_n)$ avaruuden V kanta. Olkoot myös W vektoriavaruus ja $w_1, \dots, w_n \in W$. Tällöin on olemassa täsmälleen yksi sellainen lineaarikuvaus $f \colon V \to W$, jolle pätee $f(v_i) = w_i$ kaikilla $i \in \{1, \dots, n\}$. Kuvaukselle f pätee lisäksi*

$$f(x_1v_1 + \cdots + x_nv_n) = x_1w_1 + \cdots + x_nw_n \tag{9.1}$$

kaikilla $x_1, \dots, x_n \in \mathbb{R}$.

Todistus. Lemman 9.4.1 nojalla lineaarikuvaus $f = f_{(w_1,\dots,w_n)} \circ \Phi^{-1}_{(v_1,\dots,v_n)} \colon V \to W$ toteuttaa ehdon $f(v_i) = w_i$ jokaisella $i \in \{1, \dots, n\}$.

Osoitetaan nyt, että kuvaus f toteuttaa yhtälön 9.1. Olkoot $x_1, \dots, x_n \in \mathbb{R}$. Tällöin

$$f(x_1v_1 + \cdots + x_nv_n) = x_1f(v_1) + \cdots + x_nf(v_n) = x_1w_1 + \cdots + x_nw_n.$$

Osoitetaan vielä yksikäsitteisyys. Olkoon $g \colon V \to W$ lineaarikuvaus, jolle pätee $g(v_i) = w_i$ jokaisella $i \in \{1, \dots, n\}$. Osoitetaan, että $g(v) = f(v)$ jokaisella $v \in V$. Tällöin $f = g$.

Olkoon $v \in V$ ja $(x_1, \dots, x_n)$ vektorin v koordinaatit kannassa $(v_1, \dots, v_n)$. Tällöin

$$g(v) = g(x_1v_1 + \cdots + x_nv_n) = x_1g(v_1) + \cdots + x_ng(v_n) = x_1w_1 + \cdots + x_nw_n = f(v).$$

Tämä päättää todistuksen. □

Osoitetaan kantalauseen sovelluksena, että esimerkkien 7.2.6 ja 7.2.8 lineaarikuvausten lisäksi ei ole muita lineaarikuvauksia sarakeavaruuksista vektoriavaruuksiin.

Korollaari 9.4.3. *Olkoon $f \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ lineaarikuvaus. Tällöin $f = f_A$, missä $A = \begin{bmatrix} f(e_1) & \cdots & f(e_n) \end{bmatrix} \in \mathbb{R}^{m\times n}$.*

Todistus. Olkoon $A = \begin{bmatrix} f(e_1) & \cdots & f(e_n) \end{bmatrix}$. Tällöin

$$f_A(e_i) = Ae_i = f(e_i)$$

jokaisella $i \in \{1, \dots, n\}$. Näin ollen kantalauseen yksikäsitteisyys osan nojalla $f = f_A$. □

Korollaari 9.4.4. *Olkoon V vektoriavaruus ja $f\colon \mathbb{R}^{n\times 1} \to V$ lineaarikuvaus. Tällöin $f = f_{(v_1,\ldots,v_n)}$, missä $v_i = f(e_i)$ jokaisella $i \in \{1,\ldots,n\}$.*

Todistus. Kuvauksen $f_{(v_1,\ldots,v_n)}$ määritelmän nojalla

$$f_{(v_1,\ldots,v_n)}(e_i) = v_i = f(e_i)$$

jokaisella $i \in \{1,\ldots,n\}$. Näin ollen $f = f_{(v_1,\ldots,v_n)}$. □

Huomautus 9.4.5. *Korollaarin 9.4.3 voi todistaa käyttämättä kantalausetta suoraan matriisitulosta. Korollaari 9.4.4 puolestaan on kantalauseen yksikäsitteisyysväitteen uudelleenmuotoilu.*

9.5 Lineaarikuvauksen esitysmatriisi

Korollaareista 9.4.3 ja 9.4.4 saattoi herätä ajatus, että korollaaria 9.4.3 ei olisi tarvinnut kirjata omaksi tulokseksen, koska se on korollaarin 9.4.4 erikoistapaus. Syy korollaarin 9.4.3 korostamiseen johtuu siitä, että kaikki äärellisulotteisten vektoriavaruuksien väliset lineaarikuvaukset voidaan palauttaa matriisituloon lineaarikuvauksen esitysmatriisin avulla.

Määritelmä 9.5.1. *Matriisi $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ on* lineaarikuvauksen $f\colon V \to W$ esitysmatriisi avaruuksien V ja W kantojen $(v_1,\ldots,v_n)$ ja $(w_1,\ldots,w_m)$ suhteen, *jos*

$$f(v_i) = a_{1i}w_1 + \cdots + a_{mi}w_m$$

jokaisella $i \in \{1,\ldots,n\}$.

Huomautus 9.5.2. *Huomaa, että esitysmatriisin määritelmän mukaan vektorin $f(v_i)$ kertoimet kannassa $(w_1,\ldots,w_m)$ muodostavat matriisin A i:nnen sarakkeen.*

Esimerkki 9.5.3. *Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ lineaarikuvaus $(x,y) \mapsto (y, x+3y, 2x)$. Etsitään kuvauksen f esitysmatriisi standardikannasta (e_1,e_2) standardikantaan (e_1,e_2,e_3). Koska*

$$f(e_1) = f(1,0) = (0,1,2) = e_2 + 2e_3 = 0e_1 + 1e_2 + 2e_3$$

ja

$$f(e_2) = f(0,1) = (1,3,0) = e_1 + 3e_2 = 1e_1 + 3e_2 + 0e_3$$

niin esitysmatriisi on

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 3 \\ 2 & 0 \end{bmatrix}.$$

Esimerkki 9.5.4. *Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ lineaarikuvaus $(x,y) \mapsto (y, x+3y, 2x)$. Pidetään tunnettuna, että (v_1,v_2), missä $v_1 = e_1 + e_2$ ja $v_2 = e_1 - e_2$, on avaruuden $\mathbb{R}^2$ kanta ja että (w_1,w_2,w_3), missä $w_1 = e_1 + e_3$, $w_2 = e_2$ ja $w_3 = e_1 - e_3$, on avaruuden $\mathbb{R}^3$ kanta. Etsitään kuvauksen f esitysmatriisi kannasta (v_1,v_2) kantaan (w_1,w_2,w_3).*

Ratkaistaan ensin esitysmatriisin A kertoimet a_{11}, a_{21}, ja a_{31} yhtälöstä

$$f(v_1) = f(1,1) = (1,4,2) = a_{11}w_1 + a_{21}w_2 + a_{31}w_3.$$

Koska

$$\begin{aligned} a_{11}w_1 + a_{21}w_2 + a_{31}w_3 &= a_{11}(e_1 + e_3) + a_{21}e_2 + a_{31}(e_1 - e_3) \\ &= (a_{11} + a_{31})e_1 + a_{21}e_2 + (a_{11} - a_{31})e_3 \\ &= (a_{11} + a_{31}, a_{21}, a_{11} - a_{31}), \end{aligned}$$

niin saadaan yhtälöt

$$\begin{cases} a_{11} + a_{31} & = 1 \\ a_{22} & = 4 \\ a_{11} - a_{31} & = 2 \end{cases}$$

Ratkaisemalla yhtälöryhmä esimerkiksi matriisimuodossa, saadaan yhtälöryhmän ainoaksi ratkaisuksi $a_{11} = 3/2$, $a_{21} = 4$, ja $a_{31} = -1/2$. Näin ollen matriisin A ensimmäinen sarake on

$$\begin{bmatrix} 3/2 \\ 4 \\ -1/2 \end{bmatrix}.$$

Matriisin A toinen sarake saadaan kertoimista a_{12}, a_{22} ja a_{32}, jotka toteuttavat yhtälön

$$f(v_2) = f(1,-1) = (-1,-2,2) = a_{12}w_1 + a_{22}w_2 + a_{32}w_3.$$

Ratkaisemalla yhtälöryhmä kuten edellä saadaan, että matriisin A toinen sarake on

$$\begin{bmatrix} 1/2 \\ -2 \\ -3/2 \end{bmatrix}$$

Kuvauksen f esitysmatriisi kannasta (v_1, v_2) kantaan (w_1, w_2, w_3) on siis

$$A = \begin{bmatrix} 3/2 & 1/2 \\ 4 & -2 \\ -1/2 & -3/2 \end{bmatrix}.$$

Yleensä ollaan kiinnostuneita lineaarikuvauksen esitysmatriisista. Toisaalta annetun matriiisin avulla voidaan määritellä lineaarikuvauksia, joilla on halutut ominaisuudet. Tarkastellaan tästä tilanteesta kahta esimerkkiä.

Esimerkki 9.5.5. *Olkoon $A = [a_{ji]}] \in \mathbb{R}^{m\times n}$ matriisi. Tällöin $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$, $x \mapsto Ax$, on lineaarikuvaus, jonka esitysmatriisi standardikannasta $(e_1, \ldots, e_n)$ standardikantaan $(e_1, \ldots, e_m)$ on A, koska matriisitulon määritelmän nojalla*

$$f_A(e_i) = Ae_i = a_{1i}e_1 + \cdots + a_{mi}e_m$$

jokaisella $i \in \{1, \ldots, n\}$.

Esimerkki 9.5.6. *Olkoon*

$$A = [a_{ji}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 3 \end{bmatrix} \in \mathbb{R}^{3\times 2}.$$

Pidetään jälleen tunnettuna, että (v_1, v_2), missä $v_1 = e_1 + e_2$ ja $v_2 = e_1 - e_2$, on avaruuden $\mathbb{R}^2$ kanta ja että (w_1, w_2, w_3), missä $w_1 = e_1 + e_3$, $w_2 = e_2$ ja $w_3 = e_1 - e_3$, on avaruuden $\mathbb{R}^3$ kanta.

Koska (v_1, v_2) on avaruuden $\mathbb{R}^2$ kanta, niin kantalauseen nojalla on olemassa yksikäsitteinen sellainen lineaarikuvaus $f\colon \mathbb{R}^2 \to \mathbb{R}^3$, että

$$f(v_1) = a_{11}w_1 + a_{21}w_2 + a_{31}w_3 = 1w_1 + 0w_2 + 2w_3 = (3, 0, -1)$$

ja

$$f(v_2) = a_{21}w_1 + a_{22}w_2 + a_{32}w_3 = 0w_1 + 1w_2 + 3w_3 = (3, 1, -3).$$

Tällöin A on lineaarikuvauksen f esitysmatriisi kannasta (v_1, v_2) kantaan (w_1, w_2, w_3).

Lasketaan lineaarikuvaukselle f vielä konkreettinen kaava. Koska $e_1 = (1/2)(v_1+v_2)$ ja $e_2 = (1/2)(v_1 - v_2)$, niin

$$f(e_1) = (1/2)f(v_1) + (1/2)f(v_2) = \frac{1}{2}(3, 0, -1) + \frac{1}{2}(3, 1, -3)0 = \frac{1}{2}(6, 1, -4) = (3, 1/2, 3)$$

ja

$$f(e_2) = \frac{1}{2}f(v_1) - (1/2)f(v_2) = \frac{1}{2}(3, 0, -1) - \frac{1}{2}(3, 1, -3) = \frac{1}{2}(0, -1, 2) = (0, -1/2, 1).$$

Näin ollen

$$f(x, y) = xf(e_1) + yf(e_2) = x(3, 1/2, 3) + y(0, -1/2, 1) = (3x, (x - y)/2, 3x + y)$$

kaikilla $(x, y) \in \mathbb{R}^2$.

Huomautus 9.5.7. *Koska vektorin koordinaatit kannassa ovat yksikäsitteiset ja lineaarikuvaus $f\colon V \to W$ määräytyy kanta-alkioiden kuvista, on lineaarikuvauksen f esitysmatriisi A avaruuksien V ja W kantojen $(v_1, \ldots, v_n)$ ja $(w_1, \ldots, w_m)$ suhteen yksikäsitteinen. Tarkemmin sanottuna esitysmatriisin A sarakevektorin Ae_i kertoimet saadaan vektorin $f(v_i)$ koordinaateista $(a_{1i}, \ldots, a_{mi})$ kannassa $(w_1, \ldots, w_m)$.*

Huomautus 9.5.8. *Kuten esimerkit osoittavat, kantojen muuttuessa myös kuvavektoreiden koordinaatit muuttuvat. Näin ollen lineaarikuvauksen $f\colon V \to W$ esitysmatriisi A riippuu kantojen $(v_1, \ldots, v_n)$ ja $(w_1, \ldots, w_m)$ valinnasta. Tämä tarkoittaa sitä, että matriisi muuttuu, jos jompikumpi kannoista vaihdetaan johonkin toiseen kantaan.*

Huomautus 9.5.9. *Useissa lineaarialgebran esityksissä tehdään merkinnällinen ero lineaarikuvausten esitysmatriisien ja yhtälöryhmien teoriasta tuttujen matriisien välille. Tällöin käytetään jotain symbolia kertomaan, mihin kantoihin saatu esitysmatriisi liittyy. Tässä esityksessä näin ei kuitenkaan tehdä, koska lyhennysmerkinnät saattavat hämärtää kannan valintaan liittyvien esitysmatriisien ja tuttujen sarakeavaruuksiin liittyvien matriisien suhdetta. Näin saattaa erityisesti käydä, kun tarkastellaan sarakeavaruuksien välisten lineaarikuvausten esitysmatriiseja. Lyhennys merkinnän sijaan tekstissä korostetaan aina esitysmatriisin yhteydessä, mihin kantoihin se liittyy.*

9.5.1 Lineaarikuvauksen suhde esitysmatriisin antamaan lineaarikuvaukseen

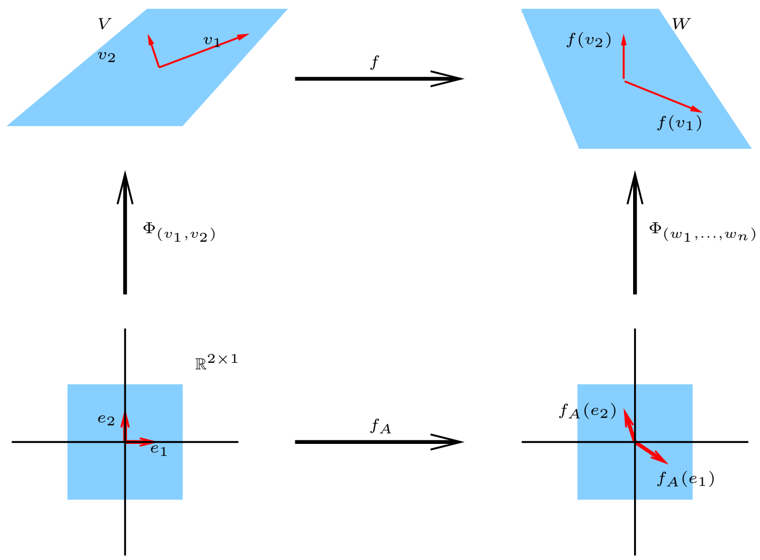
Tärkein lineaarikuvausten esitysmatriiseihin liittyvä huomio on seuraava:

> Lineaarikuvauksen esitysmatriisi muuttaa lineaarikuvauksiin liittyvät konkreettiset laskut tutuiksi matriisioperaatioksi.

Tällä huomatuksella tarkoitetaan sitä, että kunhan pidetään kirjaa, minkä kantojen suhteen esitysmatriisi A on laskettu, lineaarikuvausta $f\colon V \to W$ koskevat kysymykset muuttuvat kuvausta f_A koskeviksi kysymyksiksi. On kuitenkin huomattava, että avaruudet V ja W eivät ole yleensä ole sarakeavaruuksia, joten lineaarikuvaus $f\colon V \to W$ ei yleisesti voi edes olla sama kuvaus kuin matriisin A määrittelemä kuvaus f_A. Tämä on kuitenkin lähes totta, sillä riittää samastaa avaruudet V ja W isomorfismien $\Phi_{(v_1,\ldots,v_n)}\colon \mathbb{R}^{n\times 1} \to V$ ja $\Phi_{(w_1,\ldots,w_m)}\colon \mathbb{R}^{m\times 1} \to W$ avulla sarakeavaruuksien kanssa, jonka jälkeen kuvausten f ja f_A välille saadaan tarkka yhteys.

Kuvaillaan tämä yhteys ensin sanallisesti:

> Jos sarakevektori x kuvataan avaruuteen V isomorfismilla $\Phi_{(v_1,\ldots,v_n)}$ ja saatu vektori $v = \Phi_{(v_1,\ldots,v_n)}(x)$ kuvataan edelleen kuvauksella f, niin tämä on sama asia kuin kertoa kuvata sarakevektori x esitysmatriisia A vastaavalla kuvauksella f_A ja tämän jälkeen kuvata saatu vektori Ax avaruuteen W isomorfismilla $\Phi_{(w_1,\ldots,w_n)}$. Katso kuva 9.2.



Kuva 9.2: Kuvauksen $f\colon V \to W$ ja sen esitysmatriisiin $A \in \mathbb{R}^{m\times n}$ kuvauksen $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ välinen suhde esitettynä vektoreiden e_1 ja e_2 avulla tapauksessa $n = m = 2$.

Kirjataan tämä tulos nyt tarkasti.

Lause 9.5.10. *Olkoot V ja W äärellisulotteisia vektoriavaruuksia. Olkoot lisäksi jono $(v_1, \dots, v_n)$ avaruuden V kanta ja jono $(w_1, \dots, w_k)$ avaruuden W kanta. Tällöin lineaarikuvauksen $f: V \to W$ esitysmatriisille A kannoissa $(v_1, \dots, v_n)$ ja $(w_1, \dots, w_n)$ pätee*

$$f \circ \Phi_{(v_1,\dots,v_n)} = \Phi_{(w_1,\dots,w_k)} \circ f_A.$$

Tarkastellaan lauseen väitettä vielä kuvallisesti ennen formaalia todistusta. Kuvaukset $f: V \to W$ ja $f_A: \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ yhdessä isomorfismien $\Phi_{(v_1,\dots,v_n)}: \mathbb{R}^{n\times 1} \to V$ ja $\Phi_{(w_1,\dots,w_n)}: \mathbb{R}^{m\times 1} \to W$ kanssa voidaan koota yhteen symboliseen kuvaan kuten kuvassa 9.2.

Huomautus 9.5.11. *Kuva 9.2 herättää kysymyksen eikö tätä symbolista kuvaa voisi kirjoittaa jotenkin yksinkertaisemmin ja formaalimmin. Yksi vastaus tähän on ns. kuvauskaavio. Kuvan 9.2 tapauksessa kuvauskaavio on*

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{R}^{n\times 1} \arrow[u, "\Phi_{(v_1,\dots,v_n)}"] \arrow[r, "f_A"] & \mathbb{R}^{m\times 1} \arrow[u, "\Phi_{(w_1,\dots,w_n)}"'] \end{tikzcd}$$

Kuvauskaavio ei sisällä informaatiota yksittäisistä vektoreista, mutta sisältää kaikki tarkasteltavat kuvaukset. Kuvauskaaviossa nuolien seuraaminen tarkoittaa kuvausten yhdistämistä. Lisäksi, ellei toisin mainita, kuvauskaavio oletetaan aina kommutoivan eli että kaaviosta muodostettavat kahden avaruuden väliset erilaiset yhdistetyt kuvaukset ovat aina itseasiassa sama kuvaus. Tämän kaavion tapauksessa tämä tarkoittaa, että

$$f \circ \Phi_{(v_1,\dots,v_n)} = \Phi_{(w_1,\dots,w_n)} \circ f_A,$$

mikä on lauseen 9.5.10 väite.

Lauseen 9.5.10 todistus. Olkoon $i \in \{1, \dots, n\}$. Tällöin

$$\begin{aligned}
(\Phi_{(w_1,\dots,w_m)} \circ f_A)(e_i) &= (\Phi_{(w_1,\dots,w_m)}(f_A(e_i)) \\
&= \Phi_{(w_1,\dots,w_m)}(Ae_i) \\
&= \Phi_{(w_1,\dots,w_m)}(a_{1i}e_1 + \cdots + a_{mi}e_m) \\
&= a_{1i}w_1 + \cdots + a_{mi}w_m \\
&= f(v_i) \\
&= f(\Phi_{(v_1,\dots,v_n)}(e_i) \\
&= (f \circ \Phi_{(v_1,\dots,v_n)})(e_i).
\end{aligned}$$

Näin ollen $\Phi_{(w_1,\dots,w_m)} \circ f_A = f \circ \Phi_{(v_1,\dots,v_n)}$ lineaarikuvausten kantalauseen nojalla. □

Huomautus 9.5.12. *Lause 9.5.10 voidaan kirjoittaa myös muodossa*

$$f_A = \Phi^{-1}_{(w_1,\dots,w_k)} \circ f \circ \Phi_{(v_1,\dots,v_n)}$$

tai

$$f = \Phi_{(w_1,\ldots,w_k)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)}.$$

Näistä ensimmäinen voidaan kirjoittaa myös muodossa

$$\left(\Phi^{-1}_{(w_1,\ldots,w_k)} \circ f \circ \Phi_{(v_1,\ldots,v_n)}\right)(x) = Ax$$

kaikilla $x \in \mathbb{R}^{n\times 1}$.

Yllä korostettiin, että yksi lauseen 9.5.10 merkityksistä on se, että kuvausten esitysmatriiseilla voi laskea kuin tavallisilla matriiseilla. Osoitetaan tästä esimerkkinä, että yhdistetyn kuvauksen esitysmatriisi on esitysmatriisien tulo, kun kannat on valittu oikein.

Esimerkki 9.5.13. *Olkoot* V, W *ja* U *äärellisulotteisia vektoriavaruuksia,* $(v_1,\ldots,v_n)$ *averuuden* V *kanta,* $(w_1,\ldots,w_m)$ *avaruuden* W *kanta ja* $(u_1,\ldots,u_k)$ *avaruuden* U *kanta. Olkoot nyt* $f\colon V \to W$ *ja* $g\colon W \to U$ *lineaarikuvauksia. Tällöin* $g \circ f\colon V \to U$ *on lineaarikuvaus.*

Olkoot nyt $A \in \mathbb{R}^{m\times n}$ *on kuvauksen* f *esitysmatriisi kannasta* $(v_1,\ldots,v_n)$ *kantaan* $(w_1,\ldots,w_m)$ *ja* $B \in \mathbb{R}^{k\times m}$ *on kuvauksen* g *esitysmatriisi kannasta* $(w_1,\ldots,w_m)$ *kantaan* $(u_1,\ldots,u_k)$*. Osoitetaan, että* $BA \in \mathbb{R}^{k\times n}$ *on kuvauksen* $g \circ f$ *esitysmatriisi kannasta* $(v_1,\ldots,v_n)$ *kantaan* $(u_1,\ldots,u_k)$*.*

Koska

$$f = \Phi_{(w_1,\ldots,w_k)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)}.$$

ja

$$g = \Phi_{(u_1,\ldots,u_k)} \circ f_B \circ \Phi^{-1}_{(w_1,\ldots,w_m)},$$

niin

$$\begin{aligned} g \circ f &= \left(\Phi_{(u_1,\ldots,u_k)} \circ f_B \circ \Phi^{-1}_{(w_1,\ldots,w_m)}\right) \circ \left(\Phi_{(w_1,\ldots,w_k)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)}\right) \\ &= \Phi_{(u_1,\ldots,u_k)} \circ f_B \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)} \\ &= \Phi_{(u_1,\ldots,u_k)} \circ f_{BA} \circ \Phi^{-1}_{(v_1,\ldots,v_n)}. \end{aligned}$$

Näin ollen BA *on haluttu esitysmatriisi. Tätä voidaan myös havainnollistaa kaaviona:*

$$\begin{tikzcd} V \arrow[r, "f"] & W \arrow[r, "g"] \arrow[d, shift right, "\Phi^{-1}"'] & U \\ \mathbb{R}^{n\times 1} \arrow[u, "\Phi_{(v_1,\ldots,v_n)}"] \arrow[r, "f_A"] & \mathbb{R}^{m\times 1} \arrow[r, "f_B"] \arrow[u, shift right, "\Phi_{(w_1,\ldots,w_m)}"'] & \mathbb{R}^{k\times 1} \arrow[u, "\Phi_{(u_1,\ldots,u_k)}"'] \end{tikzcd}$$

missä kuvaus $\Phi^{-1}\colon W \to \mathbb{R}^{m\times 1}$ *on lyhennysmerkintä kuvauksen* $\Phi_{(w_1,\ldots,w_m)}$ *käänteiskuvaukselle eli* $\Phi^{-1} = \Phi^{-1}_{(w_1,\ldots,w_m)}$.

9.5.2 Lineaarikuvauksen esitysmatriisin löytäminen käytännössä

Tarkastellaan nyt esimerkkien avulla, kuinka lineaarikuvaukselle löydetään esitysmatriisi. Ennen esitysmatriisin etsimistä tulee varmistua kahdesta seikasta:

(-1) Pitää varmistaa, että annettu kuvaus $f\colon V \to W$ on lineaarikuvaus.

(0) Avaruuksilla V ja W pitää olla kannat $(v_1, \dots, v_n)$ ja $(w_1, \dots, w_m)$, joiden suhteen esitysmatriisi määritetään. Jos kantoja ei ole annettuja, niin ne tulee valita. Avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^{n\times 1}$ kohdalla valinta on usein standardikanta.

Näiden ennakkovalmistelujen jälkeen prosessi on kaksivaiheinen:

(1) Selvitetään vektoreiden $f(v_i)$ koordinaatit $(a_{1i}, \dots, a_{mi})$ kannassa $(w_1, \dots, w_m)$. Usein tämä johtaa yhtälöryhmän ratkaisemiseen.

(2) Muodostetaan saaduista koordinaateista sarakevektori $a_i \in \mathbb{R}^{m\times 1}$ ja sarakevektoreista matriisi $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$.

Tarkastellaan ensin sellaisia esimerkkejä, joissa esitysmatriisin voi lukea suoraan kanta-alkioden kuvavektoreista. Ensimmäinen on esimerkistä 7.2.5 tuttu lineaarikuvaus ja seuraavat kaksi liittyvät derivaattakuvaukseen.

Esimerkki 9.5.14. *Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ kuvaus $(x, y) \mapsto (4x + 2y, y - x, x + y)$ kuten esimerkissä 7.2.5. Etsitään kuvauksen f esitysmatriisi standardikantojen (e_1, e_2) ja (e_1, e_2, e_3) suhteen.*

Koska

$$f(e_1) = f(1, 0) = (4, -1, 1) = 4e_1 - e_2 + e_3$$

ja

$$f(e_2) = f(0, 1) = (2, 1, 1) = 2e_1 + e_2 + e_3,$$

niin kuvauksen f esitysmatriisi $A \in \mathbb{R}^{3\times 2}$ standardikannoissa on

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

Esimerkki 9.5.15. *Olkoon $\mathscr{P}_3 \subset \mathscr{F}(\mathbb{R}, \mathbb{R})$ korkeintaan astetta 3 olevian polynomien avaruus ja olkoon $D\colon \mathscr{P}_3 \to \mathscr{P}_3$ derivointi. Polynomit $p_k\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^k$, missä $k \in \{0, \dots, 3\}$ muodostavat avaruuden $\mathscr{P}_n$ kannan, eli $\dim \mathscr{P}_3 = 4$.*

Jokaisella $k \in \{1, \dots, 3\}$ funktion $x \mapsto x^k$ derivaattafunktio on $x \mapsto kx^{k-1}$. Näin ollen $Dp_k = kp_{k-1}$. Koska vakiofunktion $x \mapsto 1$ derivaattafunktio on $x \mapsto 0$, niin $Dp_1 = 0$.

Olkoon nyt $p\colon \mathbb{R} \to \mathbb{R}$ polynomi, joka on määritelty kaavalla $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$. Tällöin $p = a_0p_0 + a_1p_1 + a_2p_2 + a_3p_3$ ja

$$Dp = D(a_0p_0 + a_1p_1 + a_2p_2 + a_3p_3) = a_0Dp_0 + a_1Dp_1 + a_2Dp_2 + a_3Dp_3 = 0 + a_1 + 2a_2p_1 + 3a_3p_2.$$

Lineaarikuvauksen $D\colon \mathscr{P}_n \to \mathscr{P}_n$ *esitysmatriisi* $A \in \mathbb{R}^{4\times 4}$ *kannasta* $(p_0,\dots,p_n)$ *kantaan* $(p_0,\dots,p_n)$ *on siis*

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Esimerkki 9.5.16. *Olkoon jokaisella* $k \in \mathbb{Z}$ $s_k\colon [0,2\pi] \to \mathbb{R}$ *funktio* $x \mapsto \sin(kx)$ *ja* $c_k\colon [0,2\pi] \to \mathbb{R}$ *funktio* $x \mapsto \cos(kx)$. *Oletetaan tunnetuksi, että funktiot* (c_1,s_1,c_2,s_2) *ovat lineaarisesti riippumattomia.*[3] *Olkoon*

$$V = \mathrm{Sp}(c_1,s_1,c_2,s_2) \subset C^1([0,2\pi]).$$

Tällöin $\dim V = 4$ *ja* (c_1,s_1,c_2,s_2) *on avaruuden* V *kanta.*

Koska $Ds_k = kc_k$ *ja* $Dc_k = -ks_k$ *jokaisella* $k \in \mathbb{Z}$, *niin saadaan, että lineaarikuvauksen* $D\colon V \to V$ *esitysmatriisi* A *kannassa* (c_1,s_1,c_2,s_2) *on*

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

Edellisissä esimerkeissä kanta-alkioiden kuvien koordinaatit maaliavaruuden kannassa oli helppo lukea suoraan kuvauksen kaavasta eikä koordinaattien etsimistä varten tarvinnut ratkaista yhtälöryhmää. Yleisesti esitysmatriisin etsiminen johtaa kuitenkin yhtälöryhmän ratkaisemiseen. Tarkastellaan yhtä tyypillistä esimerkkiä.

Esimerkki 9.5.17. *Olkoon* $M \in \mathbb{R}^{m\times n}$ *matriisi ja* $f_M\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$. *Olkoon nyt* $V \subset \mathbb{R}^{n\times 1}$ *aliavaruus, olkoon* $W \subset \mathbb{R}^{m\times 1}$ *sellainen aliavaruus, että* $f(V) \subset W$, *ja olkoon* f *kuvauksen* f_M *rajoittuma aliavaruudelta* V *aliavaruudelle* W *eli* $f = (f_M)|_V\colon V \to W$, $x \mapsto Mx$. *Olkoon* $(v_1,\dots,v_k)$ *aliavaruuden* V *kanta ja olkoon* $(w_1,\dots,w_\ell)$ *avaruuden* W *kanta.*

Esitysmatriisin määritelmän mukaan kuvauksen f *esitysmatriisin* $A = \begin{bmatrix} a_1 & \cdots & a_k \end{bmatrix} \in \mathbb{R}^{\ell\times k}$ *sarakkeet*

$$a_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}$$

toteuttavat yhtälön

$$f(v_i) = a_{1i}w_1 + \cdots + a_{\ell i}w_\ell.$$

Koska

$$f(v_i) = f_M(v_i) = Mv_i$$

[3] Tämä on helpointa osoittaa käyttäen kohtisuoruutta eli luvun 10 tietoja.

ja

$$a_{1i}w_1 + \cdots + a_{\ell i}w_\ell = \begin{bmatrix} w_1 & \cdots & w_\ell \end{bmatrix} \begin{bmatrix} a_{1i} \\ \vdots \\ a_{\ell i} \end{bmatrix},$$

niin sarakevektori $a_i \in \mathbb{R}^{\ell \times 1}$ *toteuttaa siis yhtälön*

$$Ba_i = Mv_i,$$

missä $B = \begin{bmatrix} w_1 & \cdots & w_\ell \end{bmatrix}$. *Näin ollen matriisin* A *sarakeet voidaan lukea yhtälöryhmän*

$$\begin{bmatrix} B \mid Mv_1 & \cdots Mv_n \end{bmatrix}$$

supistetusta porrasmuodosta.

9.5.3 Sarakeavaruuksien välisen lineaarikuvauksen $f_M \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ esitysmatriisi annettujen kantojen suhteen

Käytännön laskuissa usein esiintyy tapaus, jossa haluataan etsiä lineaarkuvaukselle $f_M \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$, $x \mapsto Mx$ esitysmatriisi A annettujen kantojen suhteen. Tätä kutsutaan usein lyhyesti *matriisin esittämiseksi annetuissa kannoissa.* Tämä terminologia on periaatteessa väärin. Kysymys ei ole matriisin esittämisestä annetuissa kannoissa vaan matriisiin liittyvän lineaarikuvauksen esittämisestä annetuissa kannoissa. Tarkalleen ottaen kysymys on siis seuraavasta lauseesta.

Lause 9.5.18. *Olkoot* $(v_1, \ldots, v_n)$ *sarakevaruuden* $\mathbb{R}^{n\times 1}$ *kanta,* $(w_1, \ldots, w_m)$ *sarakeavaruuden* $\mathbb{R}^{m\times 1}$ *kanta ja* $M \in \mathbb{R}^{m\times n}$ *matriisi. Tällöin lineaarikuvauksen* $f_M \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ *esitysmatriisi* $A \in \mathbb{R}^{m\times n}$ *kantojen* $(v_1, \ldots, v_n)$ *ja* $(w_1, \ldots, w_m)$ *suhteen on matriisi*

$$A = Q^{-1}MP$$

missä $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ *ja* $Q = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix}$.

Todistus. Tässä tilanteessa isomorfismi $\Phi_{(v_1,\ldots,v_n)}$ on kuvaus f_P ja isomorphismi $\Phi_{(w_1,\ldots,w_m)}$ on kuvaus f_Q. Näin ollen esitysmatriisin A määritelmän perusteella

$$f_A = f_Q^{-1} \circ f_M \circ f_P = f_{Q^{-1}MP}.$$

Näin ollen $A = Q^{-1}MP$. □

9.6 Kannanvaihtomatriisi

Kuten edellä on todettu, lineaarikuvauksen $f \colon V \to W$ esitysmatriisi riippuu avaruuksille V ja W valituista kannoista. Eri kantoihin liittyvien esitysmatriisien välinen suhde on samankaltainen kuin lauseessa 9.5.18 lineaarikuvauksen $f_M \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ eri esitysmatriisien suhde matriisiin $A \in \mathbb{R}^{m\times n}$.

Aloitetaan kannanvaihtomatriisin yleisellä määritelmällä, perusominaisuudella ja siihen liittyvällä esimerkillä.

Määritelmä 9.6.1. *Matriisi $P \in \mathbb{R}^{n\times n}$ on* kannanvaihtomatriisi vektoriavaruuden V kannasta $(v_1,\dots,v_n)$ kantaan $(v'_1,\dots,v'_n)$, *jos*

$$f_P = \Phi^{-1}_{(v'_1,\dots,v'_n)} \circ \Phi_{(v_1,\dots,v_n)}.$$

Huomautus 9.6.2. *Käyttäen edellisen luvun terminologiaa, huomataan, että kannanvaihtomatriisi P on identtisen kuvauksen id_V esitysmatriisi kannasta $(v_1,\dots,v_n)$ kantaan $(v'_1,\dots,v'_n)$, sillä*

$$\Phi^{-1}_{(v'_1,\dots,v'_n)} \circ \mathrm{id} \circ \Phi_{(v_1,\dots,v_n)} = \Phi^{-1}_{(v'_1,\dots,v'_n)} \circ \Phi_{(v_1,\dots,v_n)} = f_P.$$

Termi kannanvaihtomatriisi on hieman hämäävä, koska matriisi P ei vei kantavektoreita $v_1,\dots,v_n$ kantavektoreiksi $v'_1,\dots,v'_n$. Tämä jo siitä yksinkertaisesta syystä, että vektorit kuuluvat vektoriavaruuteen V, jonka ei edes tarvitse olla sarakeavaruus. Termi kannanvaihtomatriisi liittyy pikemminkin koordinaatteihin. Kannanvaihtomatriisi P vie vektorin $v \in V$ koordinaatit kannassa $(v'_1,\dots,v'_n)$ koordinaateille kannassa $(v_1,\dots,v_n)$. Kirjataan tämä lauseeksi.

Lause 9.6.3. *Olkoon $v \in V$ ja olkoot $(x_1,\dots,x_n)$ vektorin v koordinaatit kannassa $(v_1,\dots,v_n)$ ja $(x'_1,\dots,x'_n)$ koordinaatit kannassa $(v'_1,\dots,v'_n)$. Olkoon $P \in \mathbb{R}^{n\times n}$ kannanvaihtomatriisi avaruuden V kannasta $(v_1,\dots,v_n)$ kantaan $(v'_1,\dots,v'_n)$. Tällöin*

$$\begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} = P \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Todistus. Koska

$$\Phi_{(v_1,\dots,v_n)}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = v \quad \text{ja} \quad \Phi_{(v'_1,\dots,v'_n)}\left(\begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}\right) = v,$$

niin

$$\begin{aligned}\begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} &= \Phi^{-1}_{(v'_1,\dots,v'_n)}(v) = \Phi^{-1}_{(v'_1,\dots,v'_n)}\left(\Phi_{(v_1,\dots,v_n)}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)\right) \\ &= \left(\Phi^{-1}_{(v'_1,\dots,v'_n)} \circ \Phi_{(v_1,\dots,v_n)}\right)\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = P\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.\end{aligned}$$

□

Huomautus 9.6.4. *Jos määritelmässä 9.6.1 tarkasteltaisiin isomorfismin $\Phi = \Phi^{-1}_{(v'_1,\dots,v'_n)} \circ \Phi_{(v_1,\dots,v_n)} \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ sijaan isomorfismia $\Phi' = \Phi_{(v'_1,\dots,v'_n)} \circ \Phi^{-1}_{(v_1,\dots,v_n)} \colon V \to V$, niin tällöin $\Phi'(v_i) = v'_i$ jokaisella $i \in \{1,\dots,n\}$. Tämä isomorfismi vie avaruuden V kannan $(v_1,\dots,v_n)$ kannalle $(v'_1,\dots,v'_n)$, mutta siihen ei liity luonnollista matriisia P toisin kuin isomorfismiin Φ.*

Ensimmäisiin kannanvaihtomatriiseihin tutustuttiin itseasiassa jo luvussa 2: jokainen kääntyvä matriisi on kannanvaihtomatriisi.

Esimerkki 9.6.5. *Olkoon* $(v_1, \ldots, v_n)$ *sarakeavaruuden* $\mathbb{R}^{n\times 1}$ *kanta. Tällöin* $\Phi_{(v_1,\ldots,v_n)} = f_P$*, missä* $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$*. Näin ollen kannanvaihtomatriisi standardikannasta* $(e_1, \ldots, e_n)$ *kantaan* $(v_1, \ldots, v_n)$ *on* P^{-1}*. Jos* $(v_1', \ldots, v_n')$ *on jokin toinen sarakeavaruuden* $\mathbb{R}^{n\times 1}$ *kanta, niin tällöin kannanvaihtomatriisi kannasta* $(v_1, \ldots, v_n)$ *kantaan* $(v_1', \ldots, v_n')$ *on* $Q = (P')^{-1}P$*, missä* $P' = \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix}$*.*

9.6.1 Kannanvaihtomatriisi ja lineaarikuvausten esitysmatriisit

Tässä luvussa todistetaan lauseen 9.5.18 yleinen versio kaikille lineaarikuvauksille. Aloitetaan merkinnöistä.

Olkoon $f\colon V \to W$ lineaarikuvaus äärellisulotteisten vektoriavaruuksien V ja W välillä. Olkoot $(v_1, \ldots, v_n)$ ja $(v_1', \ldots, v_n')$ avaruuden V kantoja ja olkoot $(w_1, \ldots, w_m)$ ja $(w_1', \ldots, w_m')$ avaruuden V kantoja. Olkoot vielä $A \in \mathbb{R}^{m\times n}$ kuvauksen f esitysmatriisi kannasta $(v_1, \ldots, v_n)$ kantaan $(w_1, \ldots, w_m)$ ja $A' \in \mathbb{R}^{m\times n}$ kuvauksen f esitysmatriisi kannasta $(v_1', \ldots, v_n')$ kantaan $(w_1', \ldots, w_m')$. Olkoot myös $P \in \mathbb{R}^{n\times n}$ kannanvaihtomatriisi avaruuden V kannasta $(v_1, \ldots, v_n)$ kantaan $(v_1', \ldots, v_n')$ sekä $Q \in \mathbb{R}^{n\times n}$ avaruuden W kannasta $(w_1, \ldots, w_m)$ kantaan $(w_1', \ldots, w_m')$.

Palautetaan mieleen, että esitysmatriisien A ja A' kuvaukset $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ ja $f_{A'}\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ liittyvät siis kuvaukseen $f\colon V \to W$ diagrammien

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{R}^{n\times 1} \arrow[u, "\Phi_{(v_1,\ldots,v_n)}"] \arrow[r, "f_A"] & \mathbb{R}^{m\times 1} \arrow[u, "\Phi_{(w_1,\ldots,w_n)}"'] \end{tikzcd} \qquad \text{ja} \qquad \begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{R}^{n\times 1} \arrow[u, "\Phi_{(v_1',\ldots,v_n')}"] \arrow[r, "f_{A'}"] & \mathbb{R}^{m\times 1} \arrow[u, "\Phi_{(w_1',\ldots,w_n')}"'] \end{tikzcd}$$

mukaisesti. Koska

$$f_{A'} = \Phi^{-1}_{(w_1',\ldots,w_n')} \circ f \circ \Phi_{(v_1',\ldots,v_n')}$$

ja

$$f = \Phi_{(w_1,\ldots,w_n)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)},$$

niin

$$\begin{aligned} f_{A'} &= \Phi^{-1}_{(w_1',\ldots,w_n')} \circ \Big(\Phi_{(w_1,\ldots,w_n)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)}\Big) \circ \Phi_{(v_1',\ldots,v_n')} \\ &= \Big(\Phi^{-1}_{(w_1',\ldots,w_n')} \circ \Phi_{(w_1,\ldots,w_n)}\Big) \circ f_A \circ \Big(\Phi^{-1}_{(v_1,\ldots,v_n)} \circ \Phi_{(v_1',\ldots,v_n')}\Big) \\ &= \Big(\Phi^{-1}_{(w_1',\ldots,w_n')} \circ \Phi_{(w_1,\ldots,w_n)}\Big) \circ f_A \circ \Big(\Phi^{-1}_{(v_1',\ldots,v_n')} \circ \Phi_{(v_1,\ldots,v_n)}\Big)^{-1} \\ &= f_Q \circ f_A \circ f_P^{-1}. \end{aligned}$$

eli

$$f_{A'} \circ f_P = f_Q \circ f_A.$$

Kuvaukset f_A ja $f_{A'}$ liittyvät siis toisiinsa kaavion

$$\begin{tikzcd} \mathbb{R}^{n\times 1} \arrow[r, "f_{A'}"] & \mathbb{R}^{m\times 1} \\ \mathbb{R}^{n\times 1} \arrow[u, "f_P"] \arrow[r, "f_A"] & \mathbb{R}^{m\times 1} \arrow[u, "f_Q"'] \end{tikzcd} \tag{9.2}$$

mukaisesti. Kootaan nyt havainnot lauseeksi.

Lause 9.6.6 (Kannanvaihtomatriisi)**.** *Olkoot $(v_1, \dots, v_n)$ ja $(v'_1, \dots, v'_n)$ avaruuden V kantoja sekä olkoot $(w_1, \dots, w_m)$ ja $(w'_1, \dots, w'_m)$ avaruuden W kantoja. Tällöin lineaarikuvauksen $f\colon V \to W$ esitysmatriisia $A \in \mathbb{R}^{m\times n}$ kantojen $(v_1, \dots, v_n)$ ja $(w_1, \dots, w_m)$ suhteen ja esitysmatriisia $A' \in \mathbb{R}^{m\times n}$ kantojen $(v'_1, \dots, v'_n)$ ja $(w'_1, \dots, w'_m)$ yhdistää yhtälö*

$$A' = QAP^{-1},$$

missä $Q \in \mathbb{R}^{m\times m}$ on kannanvaihtomatriisi kannasta $(w_1, \dots, w_n)$ kantaan $(w'_1, \dots, w'_n)$ ja $P \in \mathbb{R}^{n\times n}$ kannanvaihtomatriisi kannasta $(v_1, \dots, v_n)$ kantaan $(v'_1, \dots, v'_n)$.

Todistus. Koska $f_P = \Phi^{-1}_{(v'_1,\dots,v'_n)} \circ \Phi_{(v_1,\dots,v_n)}$, $f_Q = \Phi^{-1}_{(w'_1,\dots,w'_n)} \circ \Phi_{(w_1,\dots,w_n)}$, niin

$$\begin{aligned} f_{A'} &= \left(\Phi^{-1}_{(w'_1,\dots,w'_n)} \circ \Phi_{(w_1,\dots,w_n)}\right) \circ f_A \circ \left(\Phi^{-1}_{(v_1,\dots,v_n)} \circ \Phi_{(v'_1,\dots,v'_n)}\right) \\ &= f_Q \circ f_A \circ (f_P)^{-1} \\ &= f_Q \circ f_A \circ f_{P^{-1}} = f_{QAP^{-1}}. \end{aligned}$$

Näin ollen

$$A' = QAP^{-1}.$$

□

9.6.2 Kommentti

Tämän luvun tuloksista on helppo arvata, että esitys- ja kannanvaintomatriiseihin liittyvissä kaavoissa on helppo tehdä virheitä. Tämän vuoksi kaavoja ei kannata opetella ulkoa vaan aina kannattaa piirtää kaavio ja muistaa seuraavat ohjeet:

- Ainoastaan sarakeavaruuksien väliset kuvaukset on määritelty suoraan matriisitulona.
- Yleisessä tilanteessa lineaarikuvauksen $f\colon V \to W$ esitysmatriisi saadaan yhdistämällä kuvaukseen f isomorfismit $\Phi_{(v_1,\dots,v_n)}\colon \mathbb{R}^{n\times 1} \to V$ ja $\Phi_{(w_1,\dots,w_m)}\colon \mathbb{R}^{m\times 1} \to W$.
- Avaruuksien $V = \mathbb{R}^{n\times 1}$ ja $W = \mathbb{R}^{m\times 1}$ tapauseksssa isomorfismit $\Phi_{(v_1,\dots,v_n)}\colon \mathbb{R}^{n\times 1} \to V$ ja $\Phi_{(w_1,\dots,w_m)}\colon \mathbb{R}^{m\times 1} \to W$ vastaavat matriiseilla $\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n\times n}$ ja $\begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \in \mathbb{R}^{m\times m}$ kertomista.

Tarkastelemalla luvun asioita tästä näkökulmasta havaitaan, että kaikki esitysmatriiseihin liittyvät ominaisuudet voidaan päätellä seuraamalla seuraavan kaavion nuolia ja ottamalla yhdisteitä vastavista kuvauksista:

$$\begin{tikzcd}
\mathbb{R}^{n\times 1} \arrow[rrr, dashed, "f_{A'}"] \arrow[dr, "\Phi'_V"] & & & \mathbb{R}^{m\times 1} \arrow[dl, "\Phi'_W"'] \\
& V \arrow[r, "f"] & W & \\
\mathbb{R}^{n\times 1} \arrow[rrr, dashed, "f_A"'] \arrow[uu, "f_P"] \arrow[ur, "\Phi_V"'] & & & \mathbb{R}^{m\times 1} \arrow[uu, "f_Q"'] \arrow[ul, "\Phi_W"]
\end{tikzcd}$$

Kaaviossa on käytetty lyhyennysmerkintöjä $\Phi_V = \Phi_{(v_1,\dots,v_n)}$, $\Phi'_V = \Phi_{(v'_1,\dots,v'_n)}$, $\Phi_W = \Phi_{(w_1,\dots,w_m)}$ ja $\Phi'_W = \Phi_{(w'_1,\dots,w'_m)}$.

Esimerkiksi esitysmatriisi A saadaan yhdistystä kuvauksesta $f_A = \Phi_W^{-1} \circ f \circ \Phi_V$ ja esitysmatriisien A' ja A välinen yhteys yhdistetystä kuvauksesta $f_A = f_Q^{-1} \circ f_{A'} \circ f_P = f_{Q^{-1}AP}$, joka siis vastaa matriisien kertolaskua.

Jos tässä kaaviossa avaruudet V ja W ovat sarakeavaruuksia ja kuvaus f on kuvaus $f = f_M$, missä $M \in \mathbb{R}^{m\times n}$, niin tällöin kaavio saa muodon

$$\begin{tikzcd}
\mathbb{R}^{n\times 1} \arrow[rrr, dashed, "f_{A'}"] \arrow[dr, "\Phi'_V"] & & & \mathbb{R}^{m\times 1} \arrow[dl, "\Phi'_W"'] \\
& \mathbb{R}^{n\times 1} \arrow[r, "f_M"] & \mathbb{R}^{m\times 1} & \\
\mathbb{R}^{n\times 1} \arrow[rrr, dashed, "f_A"'] \arrow[uu, "f_P"] \arrow[ur, "\Phi_V"'] & & & \mathbb{R}^{m\times 1} \arrow[uu, "f_Q"'] \arrow[ul, "\Phi_W"]
\end{tikzcd}$$

Tässä tapauksessa siis kaikki kaavion kuvaukset vastaavat matriisilla kertomista. Huomaa myös, että jos tätä kaavioita käytetään tilanteessa, jossa sarakeavaruuksiin valitaan toiset kannat, niin kaaviosta tarvitaan vain kuvaukseen f_A liittyvää osaa.

Luku 10

Sisätuloavaruudet

Luvussa 4 käsiteltiin sarakeavaruuden $\mathbb{R}^{n\times 1}$ geometrian ja lineaarialgebran välistä yhteyttä pistetulon käsitteen kautta. Tässä luvussa annetaan yleiselle vektoriavaruudelle geometria pistetuloa vastaavan sisätulon käsitteen avulla. Sisätulon avulla määritellään vektoreiden kohtisuoruus ja vektorin pituus. Tällä on modernin matematiikan kannalta suuri merkitys. Esimerkiksi funktioavaruuksien sisätulojen avulla voidaan määritellä ns. Hilbertin avaruudet, jotka ovat tärkeässä roolissa sekä sovelluksissa (kuten fysiikassa tai lineaaristen osittaisdifferentiaaliyhtälöiden teoriassa) että teoreettisesti. Esimerkiksi Fourier-sarjojen teoria perustuu Hilbertin avaruuksien teoriaan.

Tämän luvun tavoitteena on tutustua äärellisulotteisten sisätuloavaruuksien teoriaan. Teoreettisesti tärkein huomio on, että edellisissä luvuissa löydetty yhteys sarakeavaruuksiin $\mathbb{R}^{n\times 1}$ voidaan laajentaa pistetuloa ja yleistä sisätuloa koskevaksi:

> Siinä missä äärellisulotteisen vektoriavaruuden V kannan $(v_1,\ldots,v_n)$ valinta samastaa avaruuden V sarakeavaruuden $\mathbb{R}^{n\times 1}$ kanssa, niin ortonormaalin kannan valinta samastaa äärellisulotteisen sisätuloavaruuden $(V,\langle\cdot,\cdot\rangle)$ pistetuloavaruuden $(\mathbb{R}^{n\times 1},\cdot)$ kanssa.

Tämä tulos muotoillaan tarkasti isometrioiden avulla.

Tässä luvussa määritellään sisätulon lisäksi kohtisuoruus ja ortonormaalit kannat sisätuloavaruuksissa sekä todistetaan Cauchyn–Schwarzin epäyhtälö. Lisäksi laajennetaan kohtisuoruuden käsittelyä ja tarkastellaan aliavaruuksien kohtisuoria aliavaruuksia. Lopuksi tarkastellaan sisätuloavaruuksia yleisempiä normiavaruuksia.

10.1 Määritelmiä

Kuten pistetulo, vektoriavaruuden sisätulo on symmetrinen bilineaarikuvaus avaruudelta reaaliluvuille, joka toteuttaa ehdon, että nollasta poikkeavan vektorin sisätulon itsensä kanssa on positiivinen. Kirjotetaan tämä nyt tarkasti.

Määritelmä 10.1.1. *Olkoon V vektoriavaruus. Funktio $\langle\cdot,\cdot\rangle\colon V\times V\to\mathbb{R}$ on* avaruuden V sisätulo*, jos kaikilla $v,v',w\in V$ ja $a\in\mathbb{R}$ pätee*

1. $\langle av+v', w\rangle = a\langle v, w\rangle + \langle v', w\rangle$,

2. $\langle v, w\rangle = \langle w, v\rangle$,

3. $\langle v, v\rangle \geq 0$ *ja*

4. $\langle v, v\rangle = 0$, *jos ja vain jos* $v = 0$.

Esimerkki 10.1.2. *Sarakeavaruuden* $\mathbb{R}^{n\times 1}$ *pistetulo* $\cdot\colon \mathbb{R}^{n\times 1}\times\mathbb{R}^{n\times 1}\to\mathbb{R}$ *on esimerkki sisätulosta. Yleisemmin matriisiavaruuden* $\mathbb{R}^{m\times n}$ *pistetulo* $\cdot\colon \mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}\to\mathbb{R}$, *joka on määritelty kaavalla*

$$[a_{ji}]\cdot[b_{ji}] = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ji}b_{ji}$$

kaikilla $A = [a_{ji}] \in \mathbb{R}^{m\times n}$ *ja* $B = [b_{ji}] \in \mathbb{R}^{m\times n}$, *on sisätulo. Aivan kuten sarakevektoreiden kohdalla, tämä pistetulo voidaan kirjoittaa toisinkin seuraavasti*

$$A\cdot B = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ji}b_{ji} = \sum_{i=1}^{m}\sum_{j=1}^{n}(A^t)_{ij}B_{ji} = \sum_{i=1}^{n}(A^tB)_{ii}.$$

Diagonaalialkioiden summaa kutsutaan matriisin jäljeksi. Tähän palataan myöhemmin.

Yksi erittäin helppo mutta tärkeä huomio sisätuloon liittyen on, että sisätulon rajoittuma aliavaruuteen on sisätulo. Tämän huomion seurauksena saadaan huomattavasti (abstrakteja) esimerkkejä sisätuloavaruuksista. Kirjataan tämä havainto lemmaksi.

Lemma 10.1.3. *Olkoon* $(V, \langle\cdot,\cdot\rangle)$ *sisätuloavaruus ja* $W \subset V$ *aliavaruus. Tällöin funktio* $\langle\cdot,\cdot\rangle_W\colon W\times W\to\mathbb{R}$, *joka on määritelty kaavalla*

$$\langle w, w'\rangle_W = \langle w, w'\rangle$$

on sisätulo aliavaruudessa W, *eli* $(W, \langle\cdot,\cdot\rangle_W)$ *on sisätuloavaruus.*

Todistus. Olkoot $v, v', w \in W$ ja $a \in \mathbb{R}$. Koska $W \subset V$, niin $v, w, w' \in V$. Koska $\langle\cdot,\cdot\rangle$ on sisätulo, niin

$$\langle av+v', w\rangle_W = \langle av+v', w\rangle = a\langle v, w\rangle + \langle v', w\rangle = a\langle v, w\rangle_W + \langle v', w\rangle_W.$$

Muut kohdat todistetaan vastaavasti. □

Esimerkki 10.1.4. *Tyypillinen sovellus edelliselle lemmalle on sarakeavaruuden* $\mathbb{R}^{n\times 1}$ *pistetulon* $\cdot$ *rajoittuma aliavaruuteen* $V \subset \mathbb{R}^{n\times 1}$.

Vektoreiden kohtisuoruus ja pituus määritellään sisätuloavaruudessa $(V, \langle\cdot,\cdot\rangle)$ kuten vastaavat käsitteet sarakeavaruuden pistetulolle.

Määritelmä 10.1.5. *Sisätuloavaruuden* $(V, \langle\cdot,\cdot\rangle)$ *vektorit* v *ja* w *ovat* toisiaan vastaan kohtisuorassa, *jos*

$$\langle v, w\rangle = 0.$$

Määritelmä 10.1.6. *Vektorin* $v \in V$ *pituus eli normi* $\|v\|$ *sisätuloavaruudessa* $(V, \langle\cdot,\cdot\rangle)$ *on luku*

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

Kuten pistetulonkin tapauksessa, myös sisätuloavaruuksien tilanteessa keskenään toisiaan vastaan kohtisuorat vektorit ovat lineaarisesti riippumattomia.

Lause 10.1.7. *Olkoon* $(V, \langle\cdot,\cdot\rangle)$ *sisätuloavaruus ja olkoot* $v_1, \ldots, v_n \in V$ *toisiaan vastaan kohtisuorassa olevia vektoreita, joista yksikään ei ole nollavektori. Tällöin jono* $(v_1, \ldots, v_n)$ *on vapaa.*

Todistus. Olkoot $a_1, \ldots, a_n \in \mathbb{R}$ sellaisia lukuja, että

$$a_1v_1 + \cdots + a_nv_n = 0.$$

Tällöin jokaisella $j \in \{1, \ldots, n\}$ pätee

$$0 = \langle a_1v_1 + \cdots + a_nv_n, v_j\rangle = a_1\langle v_1, v_j\rangle + \cdots + a_n\langle v_n, v_j\rangle = a_j\langle v_j, v_j\rangle = a_j\|v_j\|^2.$$

Koska $\|v_j\| > 0$ jokaisella $j \in \{1, \ldots, n\}$, niin $a_1 = \cdots = a_n = 0$. $\square$

10.2 Extra: Esimerkki funktioavaruuksien sisätulosta

Tulevissa luvuissa tullaan osoittamaan, että äärellisulotteiset vektoriavaruudet voidaan samastaa pistetuloavaruuksien $(\mathbb{R}^{n\times 1}, \cdot)$ kanssa. Tästä tuloksesta voi saada vaikutelman, että sisätuloavaruuden käsite tässä yleisyydessä on tarpeeton. Sisätulolla on kuitenkin tärkeä rooli ääretönulotteisten funktioavaruuksien teoriassa. Tarkastellaan nyt yhtä tärkeää esimerkkiä.

Olkoon $[a, b] \subset \mathbb{R}$ väli ja $C([a, b])$ jatkuvien funktioiden $f\colon [a, b] \to \mathbb{R}$ vektoriavaruus. Määritellään $\langle\cdot,\cdot\rangle_{L^2}\colon C([a, b]) \times C([a, b]) \to \mathbb{R}$ kaavalla

$$\langle f, g\rangle_{L^2} = \frac{1}{b-a}\int_a^b f(x)g(x)\,\mathrm{d}x$$

kaikilla $f, g \in C([a, b])$.

Selvästi funktio $\langle\cdot,\cdot\rangle_{L^2}$ on symmetrinen eli että $\langle f, g\rangle_{L^2} = \langle g, f\rangle_{L^2}$. Lisäksi $\langle\cdot,\cdot\rangle_{L^2}$ on lineaarinen molempien argumenttien suhteen, sillä integraalin perusominaisuuksien nojalla kaikilla $f_1, f_2, g \in C([a, b])$ ja $\lambda \in \mathbb{R}$ pätee

$$\begin{aligned}
\langle \lambda f_1 + f_2, g\rangle_{L^2} &= \frac{1}{b-a}\int_a^b (\lambda f_1(x) + f_2(x))g(x)\,\mathrm{d}x \\
&= \lambda\frac{1}{b-a}\int_a^b f_1(x)g(x)\,\mathrm{d}x + \frac{1}{b-a}\int_a^b f_2(x)g(x)\,\mathrm{d}x \\
&= \lambda\langle f_1, g\rangle_{L^2} + \langle f_2, g\rangle_{L^2}.
\end{aligned}$$

Koska $f(x)^2 \geq 0$ kaikilla $x \in [a,b]$, niin lisäksi saadaan

$$\langle f, f \rangle_{L^2} = \frac{1}{b-a} \int_a^b f(x)^2 \, \mathrm{d}x \geq 0.$$

kaikilla $f \in C([a,b])$. Huomaa, että tähän mennessä ei ole vielä käytetty funktioiden jatkuvuutta muutoin kuin takaamaan, että integraali on olemassa. Jatkuvuutta tarvitaan kuitenkin osoittamaan, että $\langle f, f \rangle = 0$, jos ja vain jos f on nollafunktio $x \mapsto 0$. Koska todistuksessa tarvitaan jatkuvuuden määritelmää ja tulos käsitellään integraalilaskennan kursseilla, sivuutetaan tämä todistus tässä yhteydessä.

Näin saatu ns. L^2-sisätulo $\langle \cdot, \cdot \rangle_{L^2} \colon C([a,b]) \times C([a,b]) \to \mathbb{R}$ antaa siis mahdollisuuden puhua kohtisuorista funktioista ja funktioiden pituudesta. Yksi tämän havainnon konkreettisimpia sovelluksia on Fourier-teoria, jota käytetään esimerkiksi signaalinkäsittelyssä.

Yksi tapa esitellä Fourier-teoria on sanoa, että siinä tarkastellaan 2π-periodisten[1] jatkuvien funktioiden $f \colon \mathbb{R} \to \mathbb{R}$ sarjakehitelmiä trigonometristen funktioiden suhteen eli halutaan esittää funktio f sarjana

$$f(x) = \sum_{k=-\infty}^{\infty} a_k \cos(kx) + b_k \sin(kx),$$

missä $a_k, b_k \in \mathbb{R}$.

Fourier-teorian peruslause sanoo, että jatkuvalle 2π-periodiselle funktiolle $f \colon \mathbb{R} \to \mathbb{R}$ pätee

$$f(x) = \sum_{k=-\infty}^{\infty} \langle c_k, f \rangle_{L^2} c_k + \langle s_k, f \rangle_{L^2} s_k,$$

missä $c_k \colon \mathbb{R} \to \mathbb{R}$ on funktio $x \mapsto \sqrt{2}\cos(kx)$ ja $s_k \colon \mathbb{R} \to \mathbb{R}$ on funktio $x \mapsto \sqrt{2}\sin(kx)$. Huomaa, että kaavan sisätulot $\langle c_k, f \rangle_{L^2}$ ja $\langle s_k, f \rangle_{L^2}$ ovat siis avaruuden $C([0,2\pi])$ sisätuloja

$$\langle c_k, f \rangle_{L^2} = \frac{1}{2\pi} \int_0^{2\pi} \cos(kx) f(x) \, \mathrm{d}x$$

ja

$$\langle s_k, f \rangle_{L^2} = \frac{1}{2\pi} \int_0^{2\pi} \sin(kx) f(x) \, \mathrm{d}x.$$

Tämä Fourier-teorian peruslause on hyvin syvällinen ja itseasiassa paljon yleisempi kuin tässä esitetty. Sitä käsitellään tarkemmin Fourier-teorian kurssilla syventävissä opinnoissa. Teoria perustuu havaintoon, että funktiot c_k ja s_ℓ ovat kaikilla $k \in \mathbb{Z}$ ja $\ell \in \mathbb{Z} \setminus \{0\}$ keskenään kohtisuorassa toisiaan vastaan olevia yksikkövektoreita avaruudessa $(C([0,2\pi]), \langle \cdot, \cdot \rangle_{L^2})$.

[1] Funktio $f \colon \mathbb{R} \to \mathbb{R}$ on θ-periodinen periodilla $\theta \in \mathbb{R}$, jos $f(x+\theta) = f(x)$ kaikilla $x \in \mathbb{R}$.

10.3 Ortonormaali kanta

Kohtisuoruuden ja pituuden avulla voidaan määritellä ortonormaalikanta.

Määritelmä 10.3.1. *Sisätuloavaruuden* $(V, \langle\cdot,\cdot\rangle)$ *kanta* $(v_1, \ldots, v_n)$ *on* ortonormaali, *jos*

1. $\langle v_j, v_i\rangle = 0$ *kaikilla* $j \neq i$ *ja*
2. $\|v_j\| = 1$ *kaikilla* $j \in \{1, \ldots, n\}$.

Huomautus 10.3.2. *Palautetaan mieleen, että itseasiassa pituuden määritelmää ei tarvita ortonormaalin kannan määritelmässä, sillä kanta* $(v_1, \ldots, v_n)$ *on ortonormaali, jos ja vain jos* $\langle v_j, v_i\rangle = \delta_{ji}$ *kaikilla* $j, i \in \{1, \ldots, n\}$.

Kirjataan vielä ortonormaaliin kantaan liittyvä yleisempi ortonormaalin jonon määritelmä.

Määritelmä 10.3.3. *Sisätuloavaruuden* $(V, \langle\cdot,\cdot\rangle)$ *vektoreiden jono* $(v_1, \ldots, v_k)$ *on* ortonormaali, *jos ehdot* (1) *ja* (2) *ovat voimassa.*

Luvussa 8.1 osoitettiin, että jokaisella äärellisesti viritetyllä vektoriavaruudella on kanta. Sisätuloavaruuksien teoriassa vastaava lause on, että jokaisella äärellisesti viritetyllä sisätuloavaruudella on ortonormaalikanta. Tämä tulos ei ole varsinaisesti uusi, vaan versio sarakeavaruuksista tutun Gram–Schmidt ortonormeerausprosessista. Tuloksen voi todistaa muokaamalla lauseen 4.3.9 todistusta. Tässä tulos kuitenkin todistetaan toisella todistuksella, joka antaa suoran todistuksen tälle hieman heikommalle olemassaolo väitteelle.

Lause 10.3.4 (Gram–Schmidt yleisille sisätuloavaruuksille)**.** *Jokaisella äärellisesti viritetyllä sisätuloavaruudella* $(V, \langle\cdot,\cdot\rangle)$ *on ortonormaali kanta.*

Todistus. Todistetaan väite induktiolla dimension suhteen. Jos $(V, \langle\cdot,\cdot\rangle)$ on nollaulotteinen sisätuloavaruus, niin tyhjä kanta () on myös ortonormaali kanta. Olkoon $(V, \langle\cdot,\cdot\rangle)$ yksiulotteinen sisätuloavaruus, eli $\dim V = 1$. Tällöin on olemassa sellainen $v \in V$, että $v \neq 0$. Valitaan nyt $w = (1/\|v\|)v \in V$. Tällöin (w) on avaruuden V ortonormaali kanta, sillä $\|w\| = (1/\|v\|)\|v\| = 1$.

Tehdään induktio-oletus, että $n \in \mathbb{N}$ on sellainen luku, että jokaisella n-ulotteisella sisätuloavaruudella on ortonormaali kanta.

Olkoon $(V, \langle\cdot,\cdot\rangle)$ $(n+1)$-ulotteinen sisätuloavaruus ja $(v_1, \ldots, v_{n+1})$ sen kanta. Olkoon $W = \mathrm{Sp}(v_1, \ldots, v_n)$ avaruuden V aliavaruus. Koska $(v_1, \ldots, v_n)$ on avaruuden W kanta ja sisätulon $\langle\cdot,\cdot\rangle$ rajoittuma $\langle\cdot,\cdot\rangle|_W$ aliavaruuteen W on sisätulo, niin $(W, \langle\cdot,\cdot\rangle|_W)$ on n-ulotteinen sisätuloavaruus. Näin ollen sillä on induktio-oletuksen nojalla ortonormaali kanta $(w_1, \ldots, w_n)$. Huomaa, että $w_1, \ldots, w_n \in V$ ja että $\langle w_j, w_i\rangle = \langle w_j, w_i\rangle|_W = \delta_{ji}$ kaikilla $j, i \in \{1, \ldots, n\}$.

Olkoon nyt

$$u = \langle v_{n+1}, w_1\rangle w_1 + \cdots + \langle v_{n+1}, w_n\rangle w_n \in W.$$

Osoitetaan, että vektori $v_{n+1} - u$ on kohtisuorassa jokaista vektoria $w_1, \dots, w_n$ vastaan. Olkoon $j \in \{1, \dots, n\}$. Tällöin

$$\begin{aligned}
\langle v_{n+1} - u, w_j \rangle &= \langle v_{n+1}, w_j \rangle - \langle u, w_j \rangle \\
&= \langle v_{n+1}, w_j \rangle - \langle \langle v_{n+1}, w_1 \rangle w_1 + \cdots + \langle v_{n+1} w_n, w_n \rangle w_n, w_j \rangle \\
&= \langle v_{n+1}, w_j \rangle - \sum_{i=1}^{n} \langle \langle v_{n+1}, w_i \rangle w_i, w_j \rangle \\
&= \langle v_{n+1}, w_j \rangle - \langle v_{n+1}, w_j \rangle = 0.
\end{aligned}$$

Vektorit $v_{n+1} - u$ ja w_j ovat siis kohtisuorassa toisiaan vastaan jokaisella $j \in \{1, \dots, n\}$.

Koska $v_{n+1} \notin \mathrm{Sp}(v_1, \dots, v_n)$ ja $u \in \mathrm{Sp}(v_1, \dots, v_n)$, niin $v_{n+1} - u \neq 0$ ja erityisesti $\|v_{n+1} - u\| > 0$. Olkoon

$$w_{n+1} = \frac{v_{n+1} - u}{\|v_{n+1} - u\|}.$$

Osoitetaan nyt, että $(w_1, \dots, w_{n+1})$ on avaruuden V ortonormaali kanta. Koska $v_{n+1} = \|v_{n+1} - u\| w_{n+1} + u$, niin $v_{n+1} \in \mathrm{Sp}(w_1, \dots, w_{n+1})$. Näin ollen

$$V = \mathrm{Sp}(v_1, \dots, v_{n+1}) = \mathrm{Sp}(w_1, \dots, w_{n+1}).$$

Koska $(w_1, \dots, w_{n+1})$ on avaruuden V minimaalinen virittäjäjono, niin se on kanta.

Osoitetaan vielä, että kanta $(w_1, \dots, w_{n+1})$ on ortonormaali. Koska $(w_1, \dots, w_n)$ on avaruuden $(W, \langle \cdot, \cdot \rangle|_W)$ ortonormaali kanta, niin $\langle w_j, w_i \rangle = \delta_{ji}$ kaikilla $j, i \in \{1, \dots, n\}$. Koska $\|w_{n+1}\| = 1$ ja

$$\langle w_{n+1}, w_j \rangle = \frac{1}{\|v_{n+1} - u\|} \langle v_{n+1} - u, w_j \rangle = 0$$

kaikilla $j \in \{1, \dots, n\}$, niin $(w_1, \dots, w_{n+1})$ on ortonormaali kanta. Tämä päättää induktioaskeleen todistuksen. $\square$

Kuten edellä todettiin käyttämällä luvusta 3 tutun lauseen 4.3.9 todistusta, saadaan hieman vahvempi ortonormeeraustulos. Seuraavan väitteen todistus jätetään kiinnostuneelle lukijalle harjoitustehtäväksi.[2] Huomaa, että lausen muotoilussa on käytetty lauseen 10.1.7 tietoa, että ortonormaali jono on vapaa.

Lause 10.3.5. *Olkoon $(V, \langle \cdot, \cdot \rangle)$ äärellisulotteinen sisätuloavaruus ja $(v_1, \dots, v_n)$ sellainen avaruuden V kanta, että osajono $(v_1, \dots, v_k)$ on ortonormaali. Tällöin on olemassa sellainen avaruuden V ortonormaalikanta $(w_1, \dots, w_n)$, että $w_i = v_i$ jokaisella $i \in \{1, \dots, k\}$ ja että $\mathrm{Sp}(w_1, \dots, w_j) = \mathrm{Sp}(v_1, \dots, v_j)$ jokaisella $j \in \{1, \dots, n\}$.*

Tämän lauseen voi myös kirjata konseptuaalisemmin.

Korollaari 10.3.6. *Olkoon $(V, \langle \cdot, \cdots \rangle)$ äärellisulotteinen sisätuloavaruus. Tällöin jokainen ortonormaali jono voidaan laajentaa avaruuden V ortonormaaliksi kannaksi.*

[2] Riittää vaihtaa lauseen 4.3.9 todistuksen merkinnät.

Kirjataan vielä ylös luvusta 4 pistetuloa koskeva tuttu tulos, joka omalta osaltaan perustelee ortonormaalin kannan hyödyllisyyden. Lisää syitä ortonormaalin kannan käsitteelle saadaan isometrioiden yhteydessä. Seuraavan lauseen voi todistaa aivan kuten lauseen 4.3.7. Annetaan kuitenkin hieman lyhyempi vaihtoehtoinen todistus.

Lause 10.3.7. *Olkoon $(v_1, \dots, v_n)$ sisätuloavaruuden $(V, \langle \cdot, \cdot \rangle)$ ortonormaali kanta. Tällöin jokaisella $v \in V$ pätee*

$$v = \langle v_1, v \rangle v_1 + \cdots + \langle v_n, v \rangle v_n.$$

Todistus. Olkoon $v = x_1 v_1 + \cdots + x_n v_n \in V$. Koska kanta $(v_1, \dots, v_n)$ on ortonormaali, niin jokaisella $i \in \{1, \dots, n\}$ pätee

$$\langle v_i, v \rangle = \langle v_i, x_1 v_1 + \cdots + x_n v_n \rangle = x_1 \langle v_i, v_1 \rangle + \cdots + x_n \langle v_i, v_n \rangle = x_i \langle v_i, v_i \rangle = x_i.$$

Väite seuraa. $\square$

Toinen saman tyyppinen vektorin koordinaatteja koskeva tuttu tulos on, että vektorin pituuden voi laskea sen koordinaateista ortonormaalissa kannassa. Kirjataan tämäkin lauseeksi.

Lause 10.3.8. *Olkoon $(v_1, \dots, v_n)$ sisätuloavaruuden $(V, \langle \cdot, \cdot \rangle)$ ortonormaali kanta. Tällöin jokaisella $v = x_1 v_1 + \cdots + x_n v_n \in V$ pätee*

$$\|v\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Todistus. Koska

$$\|v\|^2 = \langle x_1 v_1 + \cdots + x_n v_n, x_1 v_1 + \cdots + x_n v_n \rangle = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j x_i \langle v_j, v_i \rangle = \sum_{j=1}^{n} x_j^2,$$

niin väite seuraa. $\square$

10.4 Cauchyn–Schwarzin epäyhtälö

Seuraava epäyhtälö on yksi modernin matematiikan tärkeimmistä epäyhtälöistä.

Lause 10.4.1 (Cauchy–Schwarz[3])**.** *Olkoon $(V, \langle \cdot, \cdot \rangle)$ sisätuloavaruus. Tällöin kaikilla $v, w \in V$ pätee*

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

Epäyhtälössä pätee yhtäsuuruus, jos ja vain jos vektorit v ja w ovat lineaarisesti riippuvia.[4]*.*

[3]Tunnetaan myös nimillä *Schwarzin epäyhtälö*, *Schwarzin lemma* ja *Cauchyn–Schwarzin–Bunjakovskin epäyhtälö*.

[4]Joskus sanotaan, että yhtäsuuruus pätee, jos vektorit v ja w ovat saman suuntaiset. Väitteessä käytetty muotoilu korostaa mahdollisuutta, että toinen vektoreista on nollavektori.

Huomautus 10.4.2. *Cauchyn ja Schwarzin epäyhtälön merkitys perustuu ns. Hilbert avaruuksien eli täydellisten sisätuloavaruuksien hyödyllisyyteen matematiikassa ja sen sovellusaloilla. Tyypillisesti tarvittavat Hilbert avaruudet ovat ääretönulotteisia funktioavaruuksia kuten Fourier-teoriassa esiintyvä neliöintegroituvien funktioiden avaruus* $L^2([0,2\pi])$ *tai osittaisdifferentiaaliyhtälöiden teoriassa esiintyvä Sobolev avaruus* $H^1(\mathbb{R}^n)$. *Neliöintegroituvien funktioiden avaruuden* $L^2([0,2\pi])$ *tapauksessa Cauchyn–Schwarzin epäyhtälö saa muodon*

$$\int_0^{2\pi} f(x)g(x)\,\mathrm{d}x \le \left(\int_0^{2\pi} f(x)^2\,\mathrm{d}x\right)^{1/2}\left(\int_0^{2\pi} g(x)^2\,\mathrm{d}x\right)^{1/2}$$

missä $f, g \in L^2([0,2\pi])$.

Näissä luentomuistiinpanoissa Cauchyn–Schwarzin epäyhtälöä käytetään osoittamaan, että pituusfunktio $\|\cdot\|\colon V \to [0,\infty[$, $v \mapsto \sqrt{\langle v, v\rangle}$, *totettaa normin ehdot. Tähän palataan luvussa 10.7.*

Lauseen 10.4.1 todistus. Voidaan olettaa, että $v \neq 0$ ja $w \neq 0$. Tarkastellaan suoraa

$$\ell = \{v + tw \in V : t \in \mathbb{R}\}$$

avaruudessa V ja etsitään vektori $u = v + t_0 w \in \ell$, joka on kohtisuorassa suoran ℓ suuntavektoria w vastaan, eli ratkaistaan yhtälö

$$\langle v + t_0 w, w\rangle = 0.$$

Koska

$$\langle v + t_0 w, w\rangle = \langle v, w\rangle + t_0\langle w, w\rangle = \langle v, w\rangle + t_0\|w\|^2,$$

niin saadaan

$$t_0 = -\frac{\langle v, w\rangle}{\|w\|^2}.$$

Haluttu vektori on siis

$$u = v - \frac{\langle v, w\rangle}{\|w\|^2} w \in \ell.$$

Huomaa, että $\|w\| > 0$, koska $w \neq 0$.

Tarkastellaan nyt pisteen $u = v + t_0 w \in \ell$ etäisyyttä origosta. Seuraava lasku sisältää tämän todistuksen oleellisimman havainnoin. Koska

$$\begin{aligned}
0 \le \|v + t_0 w\|^2 &= \langle v + t_0 w, v + t_0 w\rangle = \|v\|^2 + 2t_0\langle v, w\rangle + t_0^2\|w\|^2 \\
&= \|v\|^2 + t_0\left(2\langle v, w\rangle + t_0\|w\|^2\right) = \|v\|^2 + t_0\left(2\langle v, w\rangle - \langle v, w\rangle\right) \\
&= \|v\|^2 + t_0\langle v, w\rangle = \|v\|^2 - \frac{\langle v, w\rangle}{\|w\|^2}\langle v, w\rangle = \|v\|^2 - \left(\frac{\langle v, w\rangle}{\|w\|}\right)^2,
\end{aligned}$$

niin

$$\left|\frac{\langle v, w\rangle}{\|w\|}\right| \le \|v\|.$$

Näin ollen
$$|\langle v, w\rangle| \leq \|v\|\|w\|.$$

Tarkastellaan nyt milloin epäyhtälö toteutuu yhtäsuuruutena. Oletetaan aluksi, että vektorit v ja w ovat lineaarisesti riippuvia ja osoitetaan, että tässä tapauksessa Cauchyn–Schwarzin epäyhtälö toteuttuu yhtälönä. Koska Havaitaan aluksi, että mikäli $v = 0$ tai $w = 0$, niin selvästi $\langle v, w\rangle = 0$ ja $\|v\|\|w\| = 0$. Näin ollen epäyhtälö toteutuu tässä tilanteessa yhtälönä. Voidaan siis olettaa, että molemmat vektorit ovat nollasta poikkeavia. Koska v ja w ovat lineaarisesti riippuvia, niin $v = tw$, jollain $t \in \mathbb{R}$. Näin ollen $t_0 = -\frac{\langle v,w\rangle}{\|w\|^2} = -t\frac{\|w\|^2}{\|w\|^2} = t$ ja $u = v + t_0 w = v - tw = 0$. Näin ollen edellä tehdyn laskun nojalla $0 = \|v\|^2 - \left(\frac{\langle v,w\rangle}{\|w\|}\right)^2$ eli $|\langle v, w\rangle| = \|v\|\|w\|$. Cauchyn–Schwarzin epäyhtälö toteutuu siis tässä tilanteessa yhtälönä.

Osoitetaan nyt, että mikäli Cauchyn-Schwarzin epäyhtälö totetuu yhtälönä, niin tällöin vektorit v ja w ovat lineaarisesti riippuvia. Oletetaan siis, että vektoreille v ja w pätee $|\langle v, w\rangle| = \|v\|\|w\|$. Jos $w = 0$, niin tällöin vektorit ovat lineaarisesti riippuvia. Voidaan siis olettaa, että $w \neq 0$. Koska Cauchyn–Scwarzin epäyhtälö toteutuu yhtälönä, niin edellä tehdyn laskun nojalla $u = v - \frac{\langle v,w\rangle}{\|w\|^2} w$, missä on nollavektori. Näin ollen $v = t_0 w$ ja vektorit v ja w ovat lineaarisesti riippuvia. □

Huomautus 10.4.3. *Todistus osoittaa, että Schwarzin lemman taustalla on geometrinen havainto, että affiinin suoran (todistuksessa suora ℓ) etäisyys origosta saadaan etsimällä se suoran piste (tässä $v + t_0 w$), jonka suuntavektori on kohtisuorassa affiinin suoran suuntavektoria (tässä w) vastaan. Suoran ℓ etäisyyttä origosta ei tarvittu todistuksessa, mutta tämä havainto seuraa vertaamalla pisteen $u = v + t_0 v$ ja yleisen pisteen $v + tw \in \ell$ etäisyyttä origosta. Tämä vertailu on voidaan tehdä neliöksi täydentämällä, sillä*

$$\begin{aligned}
\|v + tw\|^2 &= \langle v + tw, v + tw\rangle = \|v\|^2 + 2t\langle v, w\rangle + t^2\|w\| \\
&= \|v\|^2 + 2t\|w\|\frac{\langle v, w\rangle}{\|w\|} + (t\|w\|)^2 \\
&= \|v\|^2 - \left(\frac{\langle v, w\rangle}{\|w\|}\right)^2 + \left(\frac{\langle v, w\rangle}{\|w\|} + t\|w\|\right)^2 \\
&\geq \|v\|^2 - \left(\frac{\langle v, w\rangle}{\|w\|}\right)^2 = \|v + t_0 w\|^2.
\end{aligned}$$

Näin ollen piste $v + t_0 w \in \ell$ antaa suoran ℓ etäisyyden origosta.

Huomautus 10.4.4. *Mikäli sekä v että w ovat nollasta poikkeavia vektoreita, Schwarzin epäyhtälö voidaan kirjoittaa muodossa*

$$\left|\left\langle \frac{v}{\|v\|}, \frac{w}{\|w\|}\right\rangle\right| \leq 1,$$

missä $v/\|v\|$ ja $w/\|w\|$ ovat yksikkövektoreita. Erityisesti Schwarzin lemma siis sanoo, että yksikkövektoreiden sisätulo on itseisarvoltaan korkeintaan yksi.

Lisäksi vektoreiden v ja w välinen kulma $\angle(v,w) \in [0, 2\pi[$ voidaan määritellä kaavalla

$$\cos\angle(v,w) = \left\langle \frac{v}{\|v\|}, \frac{w}{\|w\|} \right\rangle$$

Schwarzin lemma siis osoittaa, että sisätulo on pistetulon luonnollinnen yleistys.

10.5 Aliavaruuksien kohtisuoruus

Koska sisätulo on lineaarinen molempien argumenttiensa suhteen, niin voidaan tehdä seuraava havainto: jos vektori $v \in V$ on kohtisuorassa sekä vektoria $w \in V$ että vektoria $w' \in V$ vastaan, niin v on kohtisuorassa kaikkia vektoreiden w ja w' lineaarikombinaatioita vastaan, sillä kaikilla $a, b \in \mathbb{R}$ pätee

$$\langle v, aw + bw' \rangle = a\langle v, w \rangle + b\langle v, w' \rangle = 0.$$

Tämä osoittaa, että kaikki vektoria v vastaan kohtisuorat vektorit muodostavat aliavaruuden. Tämä havainto johtaa itseasiassa aliavaruuden kohtisuoran komplementin käsitteeseen. Määritellään tätä varten ensin kohtisuorat aliavaruudet.

Määritelmä 10.5.1. *Sisätuloavaruuden $(V, \langle\cdot,\cdot\rangle)$ aliavaruudet W ja W' ovat* toisiaan vastaan kohtisuorassa*, jos kaikilla $w \in W$ ja $w' \in W'$ pätee $\langle w, w' \rangle = 0$.*

Esimerkki 10.5.2. *Sarakeavaruuden $(\mathbb{R}^{2\times 1}, \cdot)$ suorat $L = \mathrm{Sp}(e_1+e_2)$ ja $L' = \mathrm{Sp}(e_1-e_2)$ ovat toisiaan vastaan kohtisuorassa. Tämä havaitaan seuraavasti. Havaitaan ensin, että*

$$(e_1 + e_2) \cdot (e_1 - e_2) = e_1 \cdot e_1 - e_1 \cdot e_2 + e_2 \cdot e_1 - e_2 \cdot e_2 = 1 - 0 + 0 - 1 = 0.$$

Olkoot nyt $w \in L$ ja $w' \in L'$. Tällöin $w = t(e_1 + e_2)$ ja $w' = t'(e_1 - e_2)$ joillain $t, t' \in \mathbb{R}$. Näin ollen

$$w \cdot w' = t(e_1 + e_2) \cdot t'(e_1 - e_2) = tt'(e_1 + e_2) \cdot (e_1 - e_2) = 0.$$

10.5.1 Kohtisuoruuden tarkistaminen virittäjistä

Esimerkissä 10.5.2 käytettiin strategiaa, että ensin tarkasteltiin suorien L ja L' virittäjien kohtisuoruutta ja tämän jälkeen tämän avulla osoitettiin kohtisuoruus kaikille suorien L ja L' vektoreiden välillä. Tämä strategia toimii yleisesti kaikille virittäjäjoukoille ja erityisesti kannoille.

Lause 10.5.3. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus. Tällöon avaruuden V aliavaruudet $W = \mathrm{Sp}(w_1, \ldots, w_k)$ ja $W' = \mathrm{Sp}(w_1', \ldots, w_m')$ ovat toisiaan vastaan kohtisuorassa, jos ja vain jos $\langle w_j, w_i' \rangle = 0$ kaikilla $j \in \{1, \ldots, k\}$ ja $i \in \{1, \ldots, m\}$.*

Todistus. Oletetaan, että W ja W' ovat toisiaan vastaan kohtisuoria aliavaruuksia. Olkoot $j \in \{1, \ldots, k\}$ ja $i \in \{1, \ldots, m\}$. Koska $w_j \in W$ ja $w_i' \in W'$, niin kohtisuorien aliavaruuksien määritelmän mukaisesti $\langle w_j, w_i' \rangle = 0$.

Oletetaan nyt, että $\langle w_j, w'_i\rangle = 0$ kaikilla $j \in \{1, \ldots, k\}$ ja $i \in \{1, \ldots, m\}$. Osoitetaan, että W ja W' ovat toisiaan vastaan kohtisuorassa. Olkoot $w \in W$ ja $w' \in W'$. Tällöin on olemassa sellaiset luvut $a_1, \ldots, a_k \in \mathbb{R}$ ja $b_1, \ldots, b_m \in \mathbb{R}$, että $w = a_1w_1 + \cdots + a_kw_k$ ja $w' = b_1w'_1 + \cdots + b_mw'_m$. Tällöin

$$\begin{aligned}\langle w, w'\rangle &= \langle a_1w_1 + \cdots + a_kw_k, b_1w'_1 + \cdots + b_mw'_m\rangle \\ &= \sum_{j=1}^{k}\sum_{i=1}^{m}\langle a_jw_j, b_iw_i\rangle \\ &= \sum_{j=1}^{k}\sum_{i=1}^{m} a_jb_j\langle w_j, w_i\rangle = 0.\end{aligned}$$

Näin ollen W ja W' ovat toisiaan vastaan kohtisuorassa. □

10.5.2 Aliavaruuden kohtisuora komplementti

Aliavaruuden kohtisuoralla komplementilla tarkoitetaan aliavaruutta, joka sisältää kaikki annetun aliavaruuden vektoreita vastaan kohtisuorat vektorit. Määritellään ensin tämä joukko tarkasti ja osoitetaan tämän jälkeen, että kyseinen joukko on todellakin aliavaruus.

Määritelmä 10.5.4. *Sisätuloavaruuden* $(V, \langle\cdot,\cdot\rangle)$ *aliavaruuden* $W \subset V$ kohtisuora-komplementti $W^\perp \subset V$ *on osajoukko*

$$W^\perp = \{v \in V : \langle v, w\rangle = 0, \ w \in W\}.$$

Lause 10.5.5. *Sisätuloavaruuden* $(V, \langle\cdot,\cdot\rangle)$ *aliavuuden* $W \subset V$ *kohtisuorakomplementti* $W^\perp$ *on avaruuden* V *aliavaruus.*

Todistus. Riittää osoittaa, että $W^\perp$ toteuttaa aliavaruusehdon. Olkoot $v, v' \in W^\perp$ ja $a \in \mathbb{R}$. Tällöin jokaisella $w \in W$ pätee

$$\langle av + v', w\rangle = a\langle v, w\rangle + \langle v', w\rangle = 0.$$

Näin ollen $W^\perp$ on aliavaruus. □

Esimerkki 10.5.6. *Esimerkin 10.5.2 suora* L' *on itseasiassa suoran* L *kohtisuora-komplementti* $L^\perp$ *eli* $L' = L^\perp$. *Tämän voi päätellä seuraavasti.*

Esimerkissä 10.5.2 osoitettiin, että kaikki aliavaruuden L' *vektorit ovat kohtisuorassa aliavaruuden* L *vektoreita vastaan. Näin ollen* $L' \subset L^\perp$.

Inkluusion $L^\perp \subset L'$ *osoittaminen voidaan tehdä seuraavasti. Havaitaan ensin, että vektorit* $v_1 = e_1 + e_2$ *ja* $v_2 = e_1 - e_2$ *muodostavat avaruuden* $\mathbb{R}^{2\times 1}$ *kannan. Olkoon nyt* $v = av_1 + bv_2 \in L^\perp$. *Tällöin ehdosta* $\langle v, v_1\rangle = 0$ *seuraa, että*

$$0 = \langle av_1 + bv_2, v_1\rangle = a\langle v_1, v_1\rangle + b\langle v_2, v_1\rangle = a\langle v_1, v_1\rangle,$$

eli $a = 0$, *koska* $v_1 \neq 0$. *Näin ollen* $v \in L'$.

Koska ainoastaan nollavektori on kohtisuorassa itseään vastaan, niin aliavaruudella W ja sen komplementilla $W^\perp$ ei ole muita yhteisiä vektoreita, kuin nollavektori. Näin ollen kunhan osoitetaan, että kaikki avaruuden V vektorit voidaan kirjoittaa aliavaruuksien W ja $W^\perp$ vektoreiden summina, niin saadaan että itseasiassa $V = W \oplus W^\perp$. Osoitetaan tämä nyt tarkasti.

Lause 10.5.7. *Olkoon $(V, \langle \cdot, \cdot \rangle)$ äärellisulotteinen sisätuloavaruus. Tällöin jokaisella aliavaruudella $W \subset V$ pätee*

$$V = W \oplus W^\perp.$$

Todistus. Koska V on äärellisulotteinen vektoriavaruus, niin näin ollen myös sen aliavaruus W on äärellisulotteinen. Näin ollen sillä on ortonormaalikanta sisätulon $\langle \cdot, \cdot \rangle$ suhteen.[5] Olkoon $(w_1, \dots, w_k)$ aliavaruuden W ortonormaalikanta. Laajennetaan nyt jono $(w_1, \dots, w_k)$ avaruuden V kannaksi $(w_1, \dots, w_k, v_{k+1}, \dots, v_n)$. Nyt Gram–Schmidt ortonormeerausprosessin (lause 10.3.5) perusteella on olemassa sellainen avaruuden V ortonormaali kanta $(w_1, \dots, w_k, w_{k+1}, \dots, w_n)$, että

$$\mathrm{Sp}(w_1, \dots, w_j) = \mathrm{Sp}(w_1, \dots, w_k, v_{k+1}, \dots, v_j)$$

kaikilla $j \in \{k+1, \dots, n\}$. Olkoon nyt $W' = \mathrm{Sp}(w_{k+1}, \dots, w_n)$. Koska $(w_1, \dots, w_n)$ on avaruuden V kanta, niin $V = W + W'$ ja $W \cap W' = \{0\}$. Näin ollen $V = W \oplus W'$.

Riittää siis osoittaa, että $W' = W^\perp$. Koska kanta $(w_1, \dots, w_n)$ on ortonormaali, niin W ja W' ovat toisiaan vastaan kohtisuorassa. Näin ollen $W' \subset W^\perp$. Osoitetaan nyt, että $W^\perp \subset W'$. Olkoon $w \in W^\perp$. Koska $(w_1, \dots, w_n)$ on kanta, niin

$$w = a_1 w_1 + \cdots + a_n w_n$$

joillain $a_1, \dots, a_n \in \mathbb{R}$. Koska $w \in W^\perp$, niin $\langle w, w_i \rangle = 0$ kaikilla $i \in \{1, \dots, k\}$. Näin ollen

$$0 = \langle w, w_i \rangle = \langle a_1 w_1 + \cdots + a_n w_n, w_i \rangle = a_i \langle w_i, w_i \rangle = a_i$$

jokaisella $i \in \{1, \dots, k\}$ ja siten

$$w = a_{k+1} w_{k+1} + \cdots + a_n w_n \in W'.$$

Tämä osoittaa, että $W^\perp \subset W'$, mikä päättää todistuksen. □

Kirjataan vielä ylös lauseen 10.5.7 tärkeä seuraus: aliavaruuden kohtisuoran komplementin kohtisuora komplementti on alkuperäinen aliavaruus.

Lause 10.5.8. *Olkoon $(V, \langle \cdot, \cdot \rangle)$ äärellisulotteinen sisätuloavaruus ja $W \subset V$ aliavaruus. Tällöin*

$$(W^\perp)^\perp = W.$$

[5] Oikeammin tietysti sisätulon pitäisi ensin rajoittaa aliavaruuteen W, mutta tätä eroa ei tehdä.

Todistus. Osoitetaan ensin, että $W \subset (W^{\perp})^{\perp}$. Olkoon $w \in W$. Tällöin jokaisella $w' \in W^{\perp}$ pätee

$$\langle w, w' \rangle = 0.$$

eli $w \in (W^{\perp})^{\perp}$. Näin ollen $W \subset (W^{\perp})^{\perp}$.

Osoitetaan nyt, että $(W^{\perp})^{\perp} \subset W$. Olkoon $v \in (W^{\perp})^{\perp}$. Lauseen 10.5.7 nojalla $V = W \oplus W^{\perp}$ eli on olemassa sellaiset vektorit $w \in W$ ja $w'' \in W^{\perp}$, että $v = w + w''$. Koska $v \in (W^{\perp})^{\perp}$ ja $w'' \in W^{\perp}$, niin

$$0 = \langle w'', v \rangle = \langle w'', w + w'' \rangle = \langle w'', w \rangle + \langle w'', w'' \rangle = \langle w'', w'' \rangle.$$

Näin ollen $w'' = 0$ ja $v = w \in W$. Tämä päättää inkluusion $(W^{\perp})^{\perp} \subset W$ todistuksen. □

10.6 Isometriat

10.6.1 Motivointi

Tämän luvun ensimmäinen motivaatio on seuraava ortogonaalimatriisien geometrinen ominaisuus. Jos $m \geq n$ ja $U \in \mathbb{R}^{m \times n}$ on ortogonaalimatriisi, niin tällöin siihen liittyvä lineaarikuvaus $f_U : \mathbb{R}^{n \times 1} \to \mathbb{R}^{m \times 1}$, $x \mapsto Ux$, on pituuden säilyttävä kuvaus, eli isometria. Erityisesti f_U on injektio. Jos lisäksi $m = n$, niin tällöin f_U on isomorfismi. Jos puolestaan $m < n$, niin tällöin kuvaus f_U voidaan tulkita projektioksi avaruuden $\mathbb{R}^{m \times 1}$ aliavaruudelle $\mathbb{R}^{n \times 1}$.

Tässä luvussa käsitellään tapaukset $m > n$ ja $m = n$. Näistä tapaus $m = n$, eli ortogonaalisten neliömatriisien tapaus on lopulta kaikkein mielenkiintoisin, koska se liittyy kannanvaihtoon ortogonaalisilla matriiseilla ja tätä kautta matriisien normaalimuotojen teoriaan. Projektiot, eli tapaus $m < n$, on kiinnostava muista syistä ja se käsitellään erikseen luvussa 8.3.2.

10.6.2 Määritelmiä

Tässä luvussa isometrialla tarkoitetaan sisätuloavaruuksien välistä sisätulon säilyttävää säilyttävää lineaarikuvausta.

Määritelmä 10.6.1. *Olkoot $(V, \langle \cdot, \cdot \rangle_V)$ ja $(W, \langle \cdot, \cdot \rangle_W)$ sisätuloavaruuksia. Lineaarikuvaus $f : V \to W$ on* isometria*, jos*

$$\langle f(v), f(v') \rangle_W = \langle v, v' \rangle_V$$

kaikilla $v, v' \in V$.

Määritelmästä voi suoraan päätellä, että isometriat säilyttävät vektoreiden kohtisuoruuden ja pituuden. Isometria kuvaa ortonormaalin kannan aina jonoksi ortonormaaleja vektoreita. Koska tämä ehto karaketerisoi isometriat, niin kirjataan se lauseeksi seuraavasti.

Lause 10.6.2. *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *sisätuloavaruuksia ja* $(v_1, \ldots, v_n)$ *avaruuden* V *ortonormaali kanta. Tällöin lineaarikuvaus* $f \colon V \to W$ *on isometria, jos ja vain jos kuvajono* $(f(v_1), \ldots, f(v_n))$ *on ortonormaali.*

Todistus. Oletetaan ensin, että f on isometria. Olkoon $i, j \in \{1, \ldots, n\}$. Koska f on isometria, niin

$$\langle f(v_j), f(v_i)\rangle_W = \langle v_j, v_i\rangle_V = \delta_{ji}.$$

Oletetaan nyt, että kuvajono $(f(v_1), \ldots, f(v_n))$ on ortonormaali. Osoitetaan, että f on isometria. Olkoot $v, v' \in V$. Olkoot $(x_1, \ldots, x_n)$ vektorin v koordinaatit kannassa $(v_1, \ldots, v_n)$ ja olkoot $(x'_1, \ldots, x'_n)$ vektorin v' koordinaatit kannassa $(v_1, \ldots, v_n)$. Tällöin

$$\begin{aligned}\langle f(v), f(v')\rangle &= \langle f(x_1v_1 + \cdots + x_nv_n), f(x'_1v_1 + \cdots + x'_nv_n)\rangle \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} x_jx'_i\langle f(v_j), f(v_i)\rangle = \sum_{j=1}^{n}\sum_{i=1}^{n} x_jx'_i\delta_{ji} = \sum_{j=1}^{n} x_jx'_j = \langle v, v'\rangle.\end{aligned}$$

Kuvaus f on siis isometria. □

Huomautus 10.6.3. *Koska isometrian* $f \colon V \to W$ *ei tarvitse olla surjektio, niin kuvajonon* $(f(v_1), \ldots, f(v_n))$ *ei tarvitse olla avaruuden* W *ortonormaalikanta, vaikka* $(v_1, \ldots, v_n)$ *olisi avaruuden* V *ortonormaalikanta. Jos kuitenkin* $\dim V = \dim W$*, niin tällöin* $(f(v_1), \ldots, f(v_n))$ *on avaruuden* W *ortonormaali kanta, sillä jono* $(f(v_1), \ldots, f(v_n))$ *on vapaa lauseen 10.1.7 perusteella.*

.

Tämä määritelmä voitaisiin antaa usealla eri tavalla kuten seuraava lemma osoittaa.

Lemma 10.6.4. *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *sisätuloavaruuksia sekä* $\|\cdot\|_V \colon V \to [0, \infty[$ *ja* $\|\cdot\|_W \colon W \to [0, \infty[$ *näitä sisätuloja vastaavat normit. Olkoon myös* $f \colon V \to W$ *lineaarikuvaus. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. f *on isometria,*
2. $\|f(v)\|_W = \|v\|_V$ *kaikilla* $v \in V$ *ja*
3. $\|f(v) - f(v')\|_W = \|v - v'\|_W$ *kaikilla* $v, v' \in V$.

Huomautus 10.6.5. *Kommentti lemman tulkinnasta on paikallaan. Ensimmäinen ehto* (1)*, eli isometrian määritelmä sanoo, että isometria säilyttää sisätulon. Toinen ehto* (2) *sanoo, että isometria säilyttää pituudet. Kolmas ehto* (3) *sanoo, että isometria säilyttää etäisyydet. Mielenkiintoista on, että pituuden säilyttäminen on yhtäpitävää sisätulon, eli heuristisesti kulmien, säilymisen kanssa.*

Lemman 10.6.4 todistus. Osoitetaan ensin, että ehdot (2) ja (3) ovat yhtäpitäviä. Tämän jälkeen osoitetaan, että ehdot (1) ja (2) ovat yhtäpitäviä.

Oletetaan, että $\|f(v)\|_W = \|v\|_V$ kaikilla $v \in V$. Olkoot $v, v' \in V$. Tällöin

$$\|f(v) - f(v')\|_W = \|f(v - v')\|_W = \|v - v'\|_V.$$

Oletetaan nyt, että $\|f(v) - f(v')\|_W = \|v - v'\|_V$ kaikilla $v, v' \in V$. Tällöin

$$\|f(v)\|_W = \|f(v) - f(0)\|_W = \|v - 0\|_V = \|v\|_V$$

kaikilla $v \in V$. Tämä osoittaa ehtojen (2) ja (3) yhtäpitävyyden.

Osoitetaan nyt kaksi ensimmäistä ehtoa yhtäpitäviksi. Oletetaan, että f on isometria ja olkoon $v \in V$. Tällöin

$$\|f(v)\|_W = \sqrt{\langle f(v), f(v)\rangle_W} = \sqrt{\langle v, v\rangle_V} = \|v\|_V.$$

Oletetaan nyt, että $\|f(v)\|_W = \|v\|_V$ kaikilla $v \in V$. Olkoot $v, v' \in V$. Tällöin

$$\|f(v)-f(v')\|_W^2 = \langle f(v)-f(v'), f(v)-f(v')\rangle_W = \|f(v)\|_W^2 - 2\langle f(v), f(v')\rangle_W + \|f(v')\|_W^2.$$

Vastaavasti

$$\|v - v'\|_V^2 = \|v\|_V^2 - 2\langle v, v'\rangle_V + \|v'\|_V^2.$$

Koska $\|f(v)\|_W = \|v\|_V$, $\|f(v')\|_W = \|v'\|_V$ ja $\|f(v-v')\|_W = \|v-v'\|_W$, niin saadaan

$$\langle f(v), f(v')\rangle_W = \langle v, v'\rangle_V.$$

Kuvaus f on siis isometria. Tämä osoittaa ehtojen (1) ja (2) yhtäpitävyyden. □

Lauseella 10.6.4 kaksi suoraa seurausta. Ensimmäinen seuraus on, että isometriat ovat injektioita.

Korollaari 10.6.6. *Olkoon $f\colon V \to W$ isometria sisätuloavaruudesta $(V, \langle\cdot,\cdot\rangle_V)$ sisätuloavaruuteen $(W, \langle\cdot,\cdot\rangle_W)$. Tällöin f on injektio.*

Todistus. Olkoon $v \in \ker f$. Tällöin $\|v\|_V = \|f(v)\|_W = 0$. Näin ollen $v = 0$ ja $\ker f = \{0\}$. Kuvauksen f injektiivisyys seuraa nyt lauseesta 7.4.6. □

Toinen lauseen 10.6.4 suora seuraus on, että isometrioiden yhdisteet ovat isometrioita ja bijektiivisten isometrioiden, eli isomorfisten isometrioiden, käänteiskuvaukset ovat isometrioita. Tämän korollaarin todistus jätetään harjoitustehtäväksi.

Korollaari 10.6.7. *Olkoot $f\colon V \to W$ ja $g\colon W \to U$ isometrioita sisätuloavaruuksien $(V, \langle\cdot,\cdot\rangle_V)$, $(W, \langle\cdot,\cdot\rangle_W)$ ja $(U, \langle\cdot,\cdot\rangle_U)$ välillä. Tällöin yhdistetty kuvaus $g \circ f\colon V \to U$ on isometria. Jos lisäksi $f\colon V \to W$ on isomorfismi, niin $f^{-1}\colon W \to V$ on isometria.*

10.6.3 Isometriset isomorfismit

Kuten luvussa 9.3 havaittiin kannan valita samastaa äärellisulotteisen vektoriavaruuden sarakeavaruuden kanssa. Tässä luvussa osoitetaan, että ortonormaalin kannan valita samastaa äärellisulotteisen sisätuloavaruuden $(V, \langle\cdot,\cdot\rangle)$ pistetuloavaruuden $(\mathbb{R}^{n\times 1}, \cdot)$ kanssa.

Kuten on helppo arvata, halutun isomorfismin antaa jälleen kuvaus $\Phi_{(u_1,\ldots,u_n)}\colon \mathbb{R}^{n\times 1} \to V$. Osoitetaan ensin, että tämä isomorfismi on isometria ortonormaalin kannan tapauksessa. Tulos seuraa suoraan lauseesta 10.6.2, mutta annetaan silti suora todistus.

Lemma 10.6.8. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus ja $(u_1,\dots,u_n)$ ortonormaalikanta. Tällöin isomorfismi $\Phi_{(u_1,\dots,u_n)}\colon \mathbb{R}^{n\times 1}\to V$ on isometria.*

Todistus. Olkoot

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n\times 1} \text{ ja } y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

Koska $(u_1,\dots,u_n)$ on ortonormaalikanta, niin $u_i\cdot u_j = \delta_{ij}$ kaikilla $i,j\in\{1,\dots,n\}$. Näin ollen

$$\begin{aligned}\langle \Phi_{(u_1,\cdots,u_n)}(x), \Phi_{(u_1,\dots,u_n)}(y)\rangle &= \langle x_1u_1+\cdots+x_nu_n, y_1u_1+\cdots+y_nu_n\rangle \\ &= x_1y_1+\cdots+x_ny_n = x\cdot y.\end{aligned}$$

Kuvaus $\Phi_{(u_1,\dots,u_n)}$ on siis isometria. □

Lauseesta 10.6.2 saadaan myös seuraava isometristen isomorfismien karakterisointi.

Lause 10.6.9. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ sisätuloavaruuksia. Tällöin isomorfismi $\Phi\colon V\to W$ on isometria, jos ja vain jos Φ kuvaa avaruuden V ortonormaalin kannan $(v_1,\dots,v_n)$ avaruuden W ortonormaaliksi kannaksi $(\Phi(v_1),\dots,\Phi(v_n))$.*

Todistus. Oletetaan, että Φ on isometrinen isomorfismi ja olkoon $(v_1,\dots,v_n)$ avaruuden V ortonormaali kanta. Koska Φ on isomorfismi, niin $(\Phi(v_1),\dots,\Phi(v_n))$ on avaruuden W kanta. Lauseen 10.6.2 nojalla $(\Phi(v_1),\dots,\Phi(v_n))$ on myös avaruuden W ortonormaali jono. Näin ollen se on ortonormaali kanta.

Oletetaan nyt, että Φ avaruuden V ortonormaalin kannan $(v_1,\dots,v_n)$ avaruuden W ortonormaaliksi kannaksi $(\Phi(v_1),\dots,\Phi(v_n))$. Koska Φ kuvaa kannan kannaksi, niin se on isomorfismi. Lisäksi lauseen 10.6.2 nojalla Φ on isometria. □

10.6.4 Ortogonaalimatriisit ja isometriat

On yleinen tulos, että isometriat ja ortogonaalimatriisit vastaavat toisiaan. Kirjataan tästä tuloksesta ensin sarakeavaruuksien versio.

Lause 10.6.10. *Olkoon $A\in\mathbb{R}^{m\times n}$ matriisi. Tällöin kuvaus $f_A\colon \mathbb{R}^{n\times 1}\to\mathbb{R}^{m\times 1}$ on isometria avaruudesta $(\mathbb{R}^{n\times 1},\cdot)$ avaruuteen $(\mathbb{R}^{m\times 1},\cdot)$, jos ja vain jos A on ortogonaalimatriisi.*

Todistus. Oletetaan ensin, että f_A on isometria. Olkoon $A = \begin{bmatrix} v_1 & \cdots & v_n\end{bmatrix}$. Tällöin

$$v_j\cdot v_i = (Ae_j)\cdot(Ae_i) = f_A(e_j)\cdot f_A(e_i) = e_j\cdot e_i$$

kaikilla $j,i\in\{1,\dots,n\}$. Näin ollen vektorit $v_1,\dots,v_n$ ovat toisiaan vastaan kohtisuorassa olevia yksikkövektoreita. Näin ollen $A^tA = I$ eli A on ortogonaalimatriisi.

Oletetaan nyt, että $A\in\mathbb{R}^{m\times n}$ on ortogonaalimatriisi. Olkoot $x,y\in\mathbb{R}^{n\times 1}$. Tällöin

$$f_A(x)\cdot f_A(y) = Ax\cdot Ay = (Ax)^tAy = x^tA^tAy = x^tIy = x^ty = x\cdot y.$$

Kuvaus f_A on siis isometria. □

Yleisesti lineaarikuvaus äärellisulotteisten sisätuloavaruuksien välillä on isometria, jos ja vain jos sen esitysmatriisi ortonormaaleissa kannoissa on ortogonaalimatriisi. Seuraava lause jätetään harjoitustehtäväksi.

Lause 10.6.11. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ äärellisulotteisia sisätuloavaruuksia. Tällöin lineaarikuvaus $f\colon V \to W$ on isometria, jos ja vain jos sen esitysmatriisi $A \in \mathbb{R}^{m\times n}$ avaruuden V ortonormaalista kannasta $(v_1, \dots, v_n)$ avaruuden W ortonormaaliin kantaan $(w_1, \dots, w_m)$ on ortogonaalimatriisi.*

10.7 Normiavaruudet

Sisätulon $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{R}$ määräämää pituusfunktiota $\|\cdot\|\colon V \to [0, \infty[$ kutsutaan *sisätulon induksoimaksi (tai määräämäksi) normiksi.* Tämä on erikoistapaus yleisemmästä normin käsitteestä.

Määritelmä 10.7.1. *Olkoon V vektoriavaruus. Funktio $\|\cdot\|\colon V \to [0, \infty[$ on avaruuden V* normi*, jos*

1. *kaikilla $v, w \in V$ pätee $\|v + w\| \leq \|v\| + \|w\|$,*
2. *kaikilla $v \in V$ ja $a \in \mathbb{R}$ pätee $\|av\| = |a|\,\|v\|$ ja*
3. *$\|v\| = 0$, jos ja vain jos $v = 0$.*

Paria $(V, \|\cdot\|)$, missä V on vektoriavaruus ja $\|\cdot\|$ on normi, kutsutaan normiavaruudeksi.

Huomautus 10.7.2. *Ehtoa* (1) *kutsutaan* kolmioepäyhtälöksi*, koska ehdosta seuraa, että kolmen vektorin $u, v, w \in V$ määräämän kolmion sivut toteuttavat ehdon, että jokainen sivu on pituudeltaan korkeintaan kahden muun sivun pituuksien summa.*

Sisätulon määräämä normi on yleisen normin erikoistapaus siinä mielessä, että kolmioepäyhtälön lisäksi normi toteuttaa Pythagoraan lauseen: jos vektorit $v, w \in V$ ovat kohtisuorassa toisiaan vastaan, niin

$$\|v + w\|^2 = \langle v + w, v + w\rangle = \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle = \|v\|^2 + \|w\|^2.$$

Tähän palataan seuraavassa luvussa.

Osoitetaan, että sisätulon $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{R}$ määräämä funktio $\|\cdot\|\colon V \to [0, \infty[$ todella on normi.

Lause 10.7.3. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus. Tällöin funktio $\|\cdot\|\colon V \to [0, \infty[$, $v \mapsto \sqrt{\langle v, v\rangle}$, on avaruuden V normi.*

Lauseen todistuksen vaativin kohta on kolmioepäyhtälön todistus, joka perustuu Cauchyn–Schwarzin epäyhtälöön.

Lauseen 10.7.3 todistus. Osoitetaan ensin kolmioepäyhtälö. Olkoot $v, w \in V$. Tällöin

$$\begin{aligned}\|v+w\|^2 &= \langle v+w, v+w\rangle = \|v\|^2 + 2\langle v,w\rangle + \|w\|^2 \\ &\le \|v\|^2 + 2|\langle v,w\rangle| + \|w\|^2 \\ &\le \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2 .\end{aligned}$$

Näin ollen

$$\|v+w\| \le \|v\| + \|w\|.$$

Kolmioepäyhtälö on näin osoitettu.

Normin ehto (2) seuraa lähes välittömästi. Olkoot $v \in V$ ja $a \in \mathbb{R}$. Tällöin

$$\|av\| = \sqrt{\langle av, av\rangle} = \sqrt{a^2\langle v,v\rangle} = |a|\|v\|.$$

Osoitetaan vielä normin ehto (3). Olkoon $v \in V \setminus \{0\}$. Koska $\langle v,v\rangle > 0$, niin

$$\|v\| = \sqrt{\langle v,v\rangle} > 0.$$

Toisaalta $\|0\| = \sqrt{\langle 0,0\rangle} = 0$. Näin ollen $\|v\| = 0$, jos ja vain jos $v = 0$. □

10.7.1 Yleiset normit; suunnikassääntö

Vektoriavaruuksilla on myös sellaisia normeja, jotka eivät liity sisätuloon. Tälläisten normien avulla voidaan puhua vektoreiden pituudesta tilanteissa, joissa etäisyys ei liity Pythagoraan lauseeseen eli etäisyyteen ei liity luonnollista kohtisuoruuden käsitettä.

Esimerkkejä normeista ovat seuraavat tason $\mathbb{R}^2$ ns. ℓ_1- ja ℓ_∞-normit.

Esimerkki 10.7.4. *Funktio* $\|\cdot\|_1 \colon \mathbb{R}^2 \to \mathbb{R}$*, joka on määritelty kaavalla*

$$\|(x_1,x_2)\|_1 = \| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \|_1 = |x_1| + |x_2|$$

kaikilla $(x_1,x_2) \in \mathbb{R}^2$*, on avaruuden* $\mathbb{R}^2$ *normi.*

Esimerkki 10.7.5. *Funktio* $\|\cdot\|_\infty \colon \mathbb{R}^2 \to \mathbb{R}$*, joka on määritelty kaavalla*

$$\|(x_1,x_2)\|_\infty = \| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \|_\infty = \max\{|x_1|, |x_2|\}$$

kaikilla $(x_1,x_2) \in \mathbb{R}^2$*, on avaruuden* $\mathbb{R}^2$ *normi.*

Siinä missä normit ℓ_1 ja ℓ_∞ voidaan osoittaa normeiksi suorilla laskuilla, yleisempie ℓ_p-normien kohdalla tarvitaan hiukan konkaavien funktioiden teoriaa. Funktioiden $\|\cdot\|_1$ ja $\|\cdot\|_\infty$ osoittaminen normeiksi jätetään harjoitustehtäväksi. Seuraavan esimerkin funktiot $\|\cdot\|_p$ osoitetaan normeiksi reaalianalyysin kurssilla syventävissä opinnoissa.

Esimerkki 10.7.6. *Olkoon* $1 < p < \infty$. *Tällöin funktio* $\|\cdot\|_p \colon \mathbb{R}^2 \to \mathbb{R}$, *joka on määritelty kaavalla*

$$\|(x_1, x_2)\|_p = (|x_1|^p + |x_2|^p)^{1/p}$$

kaikilla $(x_1, x_2) \in \mathbb{R}^2$, *on avaruuden* $\mathbb{R}^2$ *normi.*

Yksikään näistä normeista ei liity sisätuloon, koska ne eivät toteuta ns. *suunnikassääntöä*. Kirjataan tämä sääntö lauseeksi. Todistus on suora lasku ja se jätetään harjoitustehtäväksi.

Lause 10.7.7. *Olkoon* $(V, \langle\cdot,\cdot\rangle)$ *sisätulo avaruus ja* $\|\cdot\| \colon V \to [0, \infty[$ *sisätuloon* $\langle\cdot,\cdot\rangle$ *liittyvä normi, eli funktio, jolle pätee* $\|v\| = \sqrt{\langle v, v\rangle}$ *kaikilla* $v \in V$. *Tällöin kaikilla* $v, w \in V$ *pätee*

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

Luku 11

Lineaariset operaattorit

Lineaarikuvausta $f\colon V \to V$ vektoriavaruudesta V itseensä kutsutaan *lineaariseksi operaattoriksi* tai lyhyesti *operaattoriksi*. Lineaarisilla operaattoreilla on tässä teoriassa sama rooli kuin neliömatriiseilla matriisien teoriassa, sillä niistä voidaan kysyä neliömatriiseista tuttuja kysymyksiä:

1. Mitkä ovat operaattorin $f\colon V \to V$ ominaisarvot, ominaisvektorit ja ominaisavaruudet?

2. Jos avaruus V on äärellisulotteinen, mikä on operaattorin $f\colon V \to V$ determinantti.

Kysymys ominaisarvoista käsitellään tässä luvussa niin, että aluksi annetaan yleiset määritelmät käyttäen vektoriavaruuksien teoriaa ja tämän jälkeen ominaisarvojen ja -vektoreiden löytäminen palautetaan matriisien teoriaan.

Lineaarikuvauksen determinantin käsittelylle on kaksi mahdollista tapaa. Näistä selvästi suositumpi ja suoraviivaisempi on seuraava. Valitaan avaruudelle V kanta ja lasketaan lineaarikuvauksen $f\colon V \to V$ esitysmatriisin determinantti. Tässä määritelmässä tulee ensin varmistaa, että esitysmatriisin determinantin arvo ei riipu valitusta kannasta. Tätä varten on luonnollista tarkastella esitysmatriiseja, jotka määräytyvät yhdestä avaruuden kannasta.

Toinen – ja ei niin suosittu tapa – on havaita, että samoin kuin sarakeavaruudella $\mathbb{R}^{n\times 1}$, avaruudella V on tilavauusmuoto vol_V, jonka avulla lineaarikuvauksen determinantti voidaan määritellä. Etuna tässä lähestymistavassa on, että tällöin saadaan myös tulkinta ristitulolle.

Huomautus 11.0.1. *Yleisenä huomiona sanottakoon, että lineaaristen operaattorien ja neliömatriisen teorian välillä on myös pieniä eroavaisuuksia, josta kannattaa olla tietoinen ja joihin palataan jatkossa. Tarkasteltaessa lineaarikuvausta $f\colon V \to W$ saman dimensioisten vektoriavaruuksien V ja W välillä, tämän lineaarikuvauksen esitysmatriisi A on neliömatriisi. Näin ollen on mahdollista tarkastella matriisin A ominaisarvoja ja ominaisvektoreita, vaikka määritelmällisesti kuvauksella f ei ole ominaisarvoja tai ominaisvektoreita. Kuvauksen f deteriminantin tarina on hieman erilainen. Selvästi* *matriisille A voidaan laskea determinantti. Koska avaruus V ja W ovat eri avaruuksia ja matriisin A määrittelyyn tarvitaan kaksi eri kantaa, tulee matriisin A determinantti tulkita oikein. Helpointa tämä on tehdä käyttäen tilavuusmuotoja* vol_V *ja* vol_W*.*

11.1 Lineaarioperaattorin esitysmatriisi; similaarit matriisit

Kuten edellisestä kommenteista voi päätellä lineaarisen operaattorien esitysmatriisit määritellään käyttäen vain yhtä kantaa eli kuvauksen lähtö- ja maaliavaruudessa on sama kanta. Tämä eroaa yleisten lineaarikuvausten tilateesta, jossa lähtö- ja maaliavaruussa voi olla, ja usein täytyykin olla, eri kanta.

Syy tähän käytäntöön selviää tarkemmin seuraavissa luvuissa. Otetaan kuitenkin jo tässä vaiheessa käyttöön tarvittava terminologia.

Määritelmä 11.1.1. *Olkoon V n-ulotteinen vektoriavaruus ja $(v_1, \ldots, v_n)$ avaruuden V kanta. Matriisi $A \in \mathbb{R}^{n\times n}$ on* lineaarikuvauksen $f\colon V \to V$ esitysmatriisiksi kannassa *$(v_1, \ldots, v_n)$, jos A on kuvauksen f esitysmatriisi kantojen $(v_1, \ldots, v_n)$ ja $(v_1, \ldots, v_n)$ suhteen, eli*

$$f \circ \Phi_{(v_1,\ldots,v_n)} = \Phi_{(v_1,\ldots,v_n)} \circ f_A.$$

Esimerkki 11.1.2. *Jos V on n-ulotteinen vektoriavaruus, niin identtisen kuvauksen* $\mathrm{id}\colon V \to V$ *esitysmatriisi kannasta $(v_1, \ldots, v_n)$ riippumatta on aina identtinen matriisi $I \in \mathbb{R}^{n\times n}$. Samoin nollakuvauksen $0\colon V \to V$, $v \mapsto 0$, esitysmatriisi on aina nollamatriisi $0 \in \mathbb{R}^{n\times n}$.*

Seuraava yleinen havainto seuraa suoraan lauseesta 9.6.6. Annetaan sille kuitenkin suora todistus.

Lause 11.1.3. *Olkoon V äärellisulotteinen vektoriavaruus sekä olkoot $(v_1, \ldots, v_n)$ ja $(v'_1, \ldots, v'_n)$ avaruuden V kantoja. Tällöin lineaarikuvauksen $f\colon V \to V$ esitysmatriisille A kannan suhteen $(v_1, \ldots, v_n)$ ja esitysmatriisille A' kannan suhteen $(v'_1, \ldots, v'_n)$ pätee*

$$A' = PAP^{-1},$$

missä $P \in \mathbb{R}^{n\times n}$ kannanvaihtomatriisi kannasta $(v_1, \ldots, v_n)$ kantaan $(v'_1, \ldots, v'_n)$.

Todistus. Matriisi $P \in \mathbb{R}^{n\times n}$ on määritelty kaavalla $f_P = \Phi^{-1}_{(v'_1,\ldots,v'_n)} \circ \Phi_{(v_1,\ldots,v_n)}$. Lisäksi $f_{A'} = \Phi^{-1}_{(v'_1,\ldots,v'_n)} \circ f \circ \Phi_{(v'_1,\ldots,v'_n)}$ ja $f = \Phi_{(v_1,\ldots,v_n)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)}$. Näin ollen

$$\begin{aligned}
f_{A'} &= \Phi^{-1}_{(v'_1,\ldots,v'_n)} \circ \left(\Phi_{(v_1,\ldots,v_n)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)}\right) \circ \Phi_{(v'_1,\ldots,v'_n)} \\
&= \left(\Phi^{-1}_{(v'_1,\ldots,v'_n)} \circ \Phi_{(v_1,\ldots,v_n)}\right) \circ f_A \circ \left(\Phi^{-1}_{(v_1,\ldots,v_n)} \circ \Phi_{(v'_1,\ldots,v'_n)}\right) \\
&= f_P \circ f_A \circ (f_P)^{-1} = f_{PAP^{-1}}.
\end{aligned}$$

Näin ollen $A' = PAP^{-1}$. □

Lineaarikuvauksen $f\colon V \to V$ esitysmatriisit yhden kannan suhteen ovat siis similaareja keskenään.

Määritelmä 11.1.4. *Neliömatriisit $A \in \mathbb{R}^{n\times n}$ ja $B \in \mathbb{R}^{n\times n}$ ovat* similaareja*, jos on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{n\times n}$, että*

$$A = PBP^{-1}.$$

11.2 Lineaarioperaattorin ominaisarvot

Lineaarisen operaattorin ominaisarvot, ominaisvektorit ja ominaisavaruudet määritellään kuten neliömatriisin tapauksessa.

Määritelmä 11.2.1. *Olkoon V vektoriavaruus. Luku $\lambda \in \mathbb{R}$ on lineaarikuvauksen $f\colon V \to V$* ominaisarvo*, jos on olemassa sellainen nollasta poikkeava vektori $v \in V$, $v \neq 0$, että*

$$f(v) = \lambda v.$$

Vektoria v kutsutaan tällöin kuvauksen f ominaisvektoriksi ominaisarvolla λ. Ominaisarvoa $\lambda \in \mathbb{R}$ vastaava ominaisavaruus *on osajoukko*

$$E(\lambda, f) = \{v \in V\colon f(v) = \lambda v\} \subset V.$$

Esimerkki 11.2.2. *Identtisen kuvauksen* $\mathrm{id}\colon V \to V$*, $v \mapsto v$, ainoa on ominaisarvo on $\lambda = 1$. Kaikki nollasta poikkeavat vektorit $v \in V$, $v \neq 0$, ovat ominaisvektoreita.*

Esimerkki 11.2.3. *Olkoon $A \in \mathbb{R}^{n\times n}$ neliömatriisi. Lineaarikuvauksen $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ ominaisarvot ja ominaisvektorit ovat samat kuin matriisin A.*

Seuraava tulos isometrioiden ja projektioiden ominaisarvoista jätetään harjoitustehtäväksi.

Esimerkki 11.2.4. *Isometrian ominaisarvo ovat itseisarvoltaan* 1*. Projektion ominaisarvo on joko* 0 *tai* 1*. Huomaa, että isometrialla ei tarvitse olla ominaisarvoja, kun taas projektiolla on aina vähintää yksi ominaisarvo.*

Annetaan vielä konkreettinen esimerkki ominaisarvoista.

Esimerkki 11.2.5. *Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ lineaarikuvaus $(x, y) \mapsto (y, x)$. Tällöin standardikannan (e_1, e_2) alkioille pätee*

$$f(e_1 + e_2) = f((1, 0) + (0, 1)) = f(1, 1) = (1, 1) = e_1 + e_2$$

ja

$$f(e_1 - e_2) = f((1, 0) - (0, 1)) = f(1, -1) = (-1, 1) = -e_1 + e_2 = -(e_1 - e_2).$$

Näin ollen vektori $w_1 = e_1 + e_2 \in \mathbb{R}^2$ on operaattorin f ominaisvektori ominaisarvolla $\lambda_1 = 1$ ja $w_2 = e_1 - e_2 \in \mathbb{R}^2$ kuvauksen f ominaisvektori ominaisarvolla $\lambda_2 = -1$.

Tämä konkreettinen esimerkki herättää useita kysymyksiä: Ovatko tässä kaikki kuvauksen f ominaisarvot, entä ominaisvektorit? Onko olemassa suoraviivaista tapaa selvittää ne kaikki?

Vastaukset näihin kysymyksiin tässä tapauksessa ovat seuraavat: Tässä ovat kaikki ominaisarvot. Koska ominaisarvoja vastaavat ominaisvektorit muodostavat aliavaruuden, on olemassa muitakin ominaisvektoreita. Lisäksi ominaisarvot ja ominaisvektorit voidaan selvittää esitysmatriisin avulla.

Seuraavassa vastataan näihin kysymyksiin yleisesti käännetyssä järjestyksessä. Aloitetaan ominaisarvojen ja ominaisvektoreiden löytämisellä esitysmatriisin avulla.

Lause 11.2.6. *Olkoon V äärellisulotteinen vektoriavaruus, $(v_1, \dots, v_n)$ avaruuden V kanta, $f\colon V \to V$ lineaarikuvaus ja A kuvauksen f esitysmatriisi kannassa $(v_1, \dots, v_n)$. Tällöin kuvauksella f ja sen esitysmatriisilla A on samat ominaisarvot. Lisäksi vektori $x \in \mathbb{R}^{n\times 1}$ on matriisin A ominaisvektori, jos ja vain jos $v = \Phi_{(v_1,\dots,v_n)}(x) \in V$ on lineaaarikuvauksen $f\colon V \to V$ ominaisvektori.*

Todistus. Olkoon $\lambda \in \mathbb{R}$ matriisin A ominaisarvo ja $x \in \mathbb{R}^{n\times 1}$ sitä vastaava ominaisvektori. Tällöin $f_A(x) = Ax = \lambda x$. Olkoon nyt $v = \Phi_{(v_1,\dots,v_n)}(x) \in V$. Koska $x \neq 0$, niin $v \neq 0$. Koska

$$f(v) = \Phi_{(v_1,\dots,v_n)}(f_A(x)) = \Phi_{(v_1,\dots,v_n)}(\lambda x) = \lambda \Phi_{(v_1,\dots,v_n)}(x) = \lambda v,$$

niin v on kuvauksen f ominaisvektori. Luku λ on siis kuvauksen f ominaisarvo.

Vastaavasti osoitetaan, että kuvauksen f ominaisarvot ovat kuvauksen f_A ominaisarvoja ja että vektori $x = \Phi^{-1}_{(v_1,\dots,v_n)}(v)$ on matriisin A ominaisvektori, jos $v \in V$ on kuvauksen f ominaisvektori. □

Kaksi kommenttia on paikallaan.

Huomautus 11.2.7. *Lause olisi voitu kirjoittaa myös muotoon, että kuvauksilla f ja f_A on samat ominaisarvot.*

Huomautus 11.2.8. *On selvää, että lineaarikuvauksella $f\colon V \to V$ ja sen esitysmatriisilla A kannassa $(v_1, \dots, v_n)$ ei voi olla samoja ominaisvektoreita, joska lineaarikuvauksen ominaisvektorit ovat avaruuden V vektoreita ja matriisin ominaisvektorit sarakeavaruuden $\mathbb{R}^{n\times 1}$ vektoreita.*

Näitä huomiota tärkeämpi on kuitenkin huomio, että lauseen kannalta on oleellista, että esitysmatriisi riippuu vain yhdestä kannasta. Annetaan tähän liittyen laskennallisesti helppo, mutta ilmiötä kuvaava esimerkki, kolmessa vaiheessa. Ensimmäinen vaihe havainnollistaa samalla, että lauseen 11.2.6 avulla löydetään kaikki operaattorin ominaisarvot.

Esimerkki 11.2.9. *Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ esimerkin 11.2.5 kuvaus $(x, y) \mapsto (y, x)$. Koska $f(e_1) = e_2$ ja $f(e_2) = e_1$, niin kuvauksen f esitysmatriisi standardikannassa (e_1, e_2) on*

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

Lasketaan nyt matriisin A ominaisarvot luvusta 6 tutulla menetelmällä. Koska

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = (-\lambda)^2 - 1 = \lambda^2 - 1$$

niin matriisin A ominaisarvot ovat $\lambda_1 = 1$ ja $\lambda_2 = -1$. Lauseen 11.2.6 perusteella nämä luvut ovat esimerkin 11.2.5 operaattorin f ainoat ominaisarvot.

Todetaan vielä, että matriisin A ominaisavaruudet ovat

$$E(1, A) = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

ja

$$E(-1, A) = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right).$$

Etsitään nyt tämän avulla kuvauksen f ominaisvektorit.

Operaattorin f ominaisarvoa 1 vastaava ominaisvektori $v \in \mathbb{R}^2 \setminus \{0\}$ on lauseen 11.2.6 mukaan sellainen vektori, että $v = \Phi_{(e_1,e_2)}(x)$, missä $x \in E(1, A)$ on ominaisvektori eli

$$x = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

jollain $t \in \mathbb{R} \setminus \{0\}$. Näin ollen

$$\begin{aligned} v = \Phi_{(e_1,e_2)}\left(t \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) &= t\Phi_{(e_1,e_2)}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \\ &= t\Phi_{(e_1,e_2)}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + t\Phi_{(e_1,e_2)}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= te_1 + te_2 = t(e_1 + e_2) = tw_1 \end{aligned}$$

eli

$$E(1, f) = \mathrm{Sp}(w_1)$$

kuten esimerkki 11.2.5 vihjasi. Vastaavasti päätellään, että

$$E(-1, f) = \mathrm{Sp}(w_2).$$

Esimerkki 11.2.10. *Tarkastellaan nyt esimerkin 11.2.5 kuvausta $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (y, x)$, kannassa (w_1, w_2), missä $w_1 = e_1 + e_2$ ja $w_2 = e_1 - e_2$. Esimerkin 11.2.5 perusteella tiedetään, että vektorit w_1 ja w_2 kuvauksen f ominaisvektoreita ominaisarvoille 1 ja -1 eli $f(w_1) = w_1$ ja $f(w_2) = -w_2$. Näin ollen operaattorin f esitysmatriisi kannassa (w_1, w_2) on*

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

Lasketaan nyt tämän matriisin A ominaisarvot

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} 1 - \lambda & 0 \\ 0 & -1 - \lambda \end{bmatrix} = (1 - \lambda)(-1 - \lambda) = -(1 - \lambda)(1 + \lambda).$$

Näin ollen matriisin A ominaisarvot ovat jälleen $\lambda_1 = 1$ ja $\lambda_2 = -1$. Tässä tapausessa ominaisavaruudet ovat

$$E(1, A) = \mathrm{Sp}\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$$

ja

$$E(-1, A) = \mathrm{Sp}\left(\begin{bmatrix}0\\1\end{bmatrix}\right).$$

Lasketaan nyt uudelleen operaattorin f ominaisavaruudet tästä esitysmatriisista. Olkoon $v \in E(1, f)$. Tällöin $v = \Phi_{(w_1,w_2)}(x)$, missä $x \in E(1, A)$ on ominaisvektori eli

$$x = t\begin{bmatrix}1\\0\end{bmatrix}$$

jollain $t \in \mathbb{R} \setminus \{0\}$. Näin ollen

$$v = \Phi_{(w_1,w_2)}\left(t\begin{bmatrix}1\\0\end{bmatrix}\right) = t\Phi_{(w_1,w_2)}\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = tw_1$$

eli

$$E(1, f) = \mathrm{Sp}(w_1)$$

kuten jo edellisessä esimerkissä havaittiin. Vastaavasti osoitetaan, että

$$E(-1, f) = \mathrm{Sp}(w_2).$$

Lause 11.2.6 sanoo, että yllä tehdyt esimerkit eivät ole sattumia, vaan yllä lasketut esimerkin 11.2.5 operaattorin $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ ominaisarvot tai ominaisavaruudet eivät riipu käytetystä esitysmatriisista eli kannan (v_1, v_2) valinnasta. Laskut osoittavat, että esitysmatriisin ominaisavaruudet kuitenkin riippuvat esitysmatriisista ja siten avaruudelle V valitusta kannasta (v_1, v_2).

Lasketaan nyt tässä konkreettisessa tilanteessa kaksi esimerkkiä, jotka osoittavat, että kuvauksen f lähtö- ja maaliavaruuteen todellakin on valittava sama kanta, jotta tämä ilmiö on voimassa.

Esimerkki 11.2.11. *Tarkastellaan nyt esimerkin 11.2.5 kuvausta $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (y, x)$, kannasta (v_1, v_2) standardikantaan (e_1, e_2), missä $v_1 = e_1 + e_2$ ja $v_2 = e_1 - e_2$ kuten edellisessä esimerkissä. Koska*

$$f(v_1) = v_1 = e_1 + e_2$$

ja

$$f(v_2) = -v_2 = -e_1 + e_2,$$

niin kuvauksen f esitysmatriisi kannasta (v_1, v_2) kantaan (e_1, e_2) on

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

Tällöin yhtälöstä

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 - (-1) = (1-\lambda)^2 + 1.$$

havaitaan, että tällä matriisilla A ei edes ole ominaisarvoja!

Esimerkki 11.2.12. *Tarkastellaan nyt esimerkin 11.2.5 kuvausta $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (y, x)$, standardikannasta (e_1, e_2) kantaan $(2e_1, 3e_2)$. Koska*

$$f(e_1) = e_2 = (1/3)(3e_2)$$

ja

$$f(e_2) = e_1 = (1/2)(2e_1)$$

niin kuvauksen f esitysmatriisi standardikannasta (e_1, e_2) kantaan $(2e_1, 3e_2)$ on

$$A = \begin{bmatrix} 0 & 1/2 \\ 1/3 & 0 \end{bmatrix}.$$

Tällöin yhtälöstä

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1/2 \\ 1/3 & -\lambda \end{bmatrix} = (-\lambda)^2 - 1/6 = \lambda^2 - 1/6$$

havaitaan, että esitysmatriisin A ominaisarvot ovat $\lambda_1 = 1/\sqrt{6}$ ja $\lambda_2 = -1/\sqrt{6}$. Näillä luvuilla ei ole mitään tekemistä lineaarikuvauksen f ominaisarvojen kanssa.

11.3 Ominaisavaruudet

Edellä käsiteltiin ominaisarvojen laskemista. Siirrytään nyt käsittelemään ominaisarvojen ja ominaisvektoreiden geometrista merkitystä.

Aloitetaan havainnosta, että operaattorin ominaisavaruus on aliavaruus, aivan kuten neliömatriisien tapauksessa. Seuraavan lemman voi todistaa äärellisulotteisten avaruuksien välisille operaattoreille esitysmatriisin avulla. Annetaan kuitenkin lyhyt suora todistus.

Lemma 11.3.1. *Olkoon $f\colon V \to V$ lineaarikuvaus ja $\lambda \in \mathbb{R}$ sen ominaisarvo. Tällöin $E(\lambda, f)$ on vektoriavaruuden V aliavaruus.*

Todistus. Olkoot $v, v' \in E(\lambda, f)$ ja $a \in \mathbb{R}$. Tällöin

$$f(av + v') = f(av) + f(v') = af(v) + f(v') = a\lambda v + \lambda v' = \lambda(av + v').$$

Näin ollen $av + v' \in E(\lambda, f)$. Osajoukko $E(\lambda, f)$ on siis aliavaruus. □

Lauseen 11.2.6 tulos voidaan nyt tulkita seuraavasti.

Korollaari 11.3.2. *Olkoon V äärellisulotteinen vektoriavaruus ja $(v_1, \dots, v_n)$ sen kanta. Olkoon $f \colon V \to V$ lineaarikuvaus ja A sen esitysmatriisi kannassa $(v_1, \dots, v_n)$. Tällöin jokaisella lineaarikuvauksen f ominaisarvolla $\lambda \in \mathbb{R}$ pätee*

$$E(\lambda, f) = \Phi_{(v_1,\dots,v_n)}(E(\lambda, A)).$$

Todistus. Olkoon $\lambda \in \mathbb{R}$ lineaarikuvauksen f ominaisarvo. Lauseen 11.2.6 perusteella $v \in E(\lambda, f)$, jos ja vain jos $v = \Phi_{(v_1,\dots,v_n)}(x)$ jollain $x \in E(\lambda, A)$. □

Ominaisavaruuden todellinen merkitys on seuraava.

Lause 11.3.3. *Olkoon $f \colon V \to V$ lineaarioperaattori ja $\lambda \in \mathbb{R}$ sen ominaisarvo. Tällöin $f(E(\lambda, f)) \subset E(\lambda, f)$ eli rajoittumakuvaus $f|_{E(\lambda,f)} \colon E(\lambda, f) \to E(\lambda, f)$ on hyvin määritelty. Lisäksi rajoittumakuvauksella $f|_{E(\lambda,f)} \colon E(\lambda, f) \to E(\lambda, f)$ on kaava $v \mapsto \lambda v$.*

Todistus. Olkoon $v \in E(\lambda, f)$. Tällöin $f(v) = \lambda v$. Koska $E(\lambda, f)$ on aliavaruus, niin $\lambda v \in E(\lambda, f)$ eli $f(v) \in E(\lambda, f)$. Näin ollen $f(E(\lambda, f)) \subset E(\lambda, f)$ ja kuvaus $f|_{E(\lambda,f)} \colon E(\lambda, f) \to E(\lambda, f)$ on hyvin määritelty. Lisäksi havaittiin, että jokaisella $v \in E(\lambda, f)$ pätee

$$f|_{E(\lambda,f)}(v) = f(v) = \lambda v$$

eli kuvauksen $f|_{E(\lambda,f)}$ kaava on $v \mapsto \lambda v$. □

Lause 11.3.3 siis sanoo, että ominaisavaruuteen rajoitettuna, operaattori käyttäytyy kuin helpoin mahdollinen lineaarikuvaus eli skaalaus. Se myös sanoo, että jos operaattori rajoitetaan ominaisavaruuteen, niin siitä ei tule mielivaltaista lineearikuvausta, vaan rajoittamalla samalla myös maaliavaruutta, saadaan operaattori.

Huomautus 11.3.4. *Lauseen 11.3.3 todistusta voi analysoida hieman pidemmälle: rajoittuma $f|_{E(\lambda,f)} \colon E(\lambda, f) \to E(\lambda, f)$ on isomorfismi, jos ja vain jos $\lambda \neq 0$. Tämä havainto jätetään harjoitustehtäväksi.*

11.3.1 Ominaisvektoreiden lineaarinen riippumattomuus

Lineaaristen operaattorien ominaisavaruuksien tärkein ominaisuus on, että eri ominaisarvoja vastaavat ominaisvektorit ovat toisistaan lineaarisesti riippumattomia. Tämä tulos voidaan muotoilla suoran summan käsitettä käyttäen muodossa, että ominaisavaruuksien summa on suora. Aloitetetaan kuitenkin muotoilusta, jossa ominaisvektorit muodostavat vapaan jonon.

Lause 11.3.5. *Olkoon $f \colon V \to V$ operaattori, olkoot $\lambda_1 < \lambda_2 < \dots < \lambda_k$ operaattorin f ominaisarvoja sekä olkoot $v_i \in E(\lambda_i, f)$ ominaisvektoreita jokaisella $i \in \{1, \dots, k\}$. Tällöin jono $(v_1, \dots, v_k)$ on vapaa.*

Huomaa, että väitteen muotoilussa on olennaista, että vektorit $v_1, \dots, v_k$ kuuluvat eri ominaisavaruuksiin ja ovat nollasta poikkeavia.

Lauseen 11.3.5 todistus. Tehdään vastaoletus, että jono $(v_1, \dots, v_k)$ ei ole vapaa. Tällöin on olemassa pienin sellainen luku $1 < m \leq k$, että jono $(v_1, \dots, v_m)$ ei ole vapaa. Koska tällöin jono $(v_1, \dots, v_{m-1})$ on vapaa, niin on olemassa sellaiset luvut $b_1, \dots, b_{m-1} \in \mathbb{R}$, että
$$v_m = b_1 v_1 + \cdots + b_{m-1} v_{m-1}.$$
Koska $v_i \in E(\lambda_i, f)$ jokaisella $i \in \{1, \dots, m\}$, niin
$$\begin{aligned}\lambda_m v_m = f(v_m) &= f(b_1 v_1 + \cdots + b_{m-1} v_{m-1})\\ &= b_1 f(v_1) + \cdots + b_{m-1} f(v_{m-1})\\ &= b_1 \lambda_1 v_1 + \cdots + b_{m-1} \lambda_{m-1} v_{m-1}.\end{aligned}$$
Osoitetaan nyt, että $\lambda_m \neq 0$. Tehdään vastaoletus, että $\lambda_m = 0$. Tällöin $\lambda_1 < \cdots < \lambda_{m-1} < 0$ ja
$$b_1 \lambda_1 v_1 + \cdots b_{m-1} \lambda_{m-1} v_{m-1} = 0.$$
Koska jono $(v_1, \dots, v_{m-1})$ on vapaa, niin $b_1 = \cdots = b_{m-1} = 0$. Näin ollen $v_m = 0$, mikä on ristiriita, koska v_m on ominaisvektori. Näin ollen $\lambda_m \neq 0$.

Koska $\lambda_m \neq 0$, niin
$$v_m = b_1 \frac{\lambda_1}{\lambda_m} v_1 + \cdots + b_{m-1} \frac{\lambda_{m-1}}{\lambda_m} v_{m-1}.$$
Koska jono $(v_1, \dots, v_{m-1})$ on vapaa ja $v_m \in \mathrm{Sp}(v_1, \dots, v_{m-1})$, niin vektorilla v_m on yksikäsitteiset koordinaatit aliavaruuden $\mathrm{Sp}(v_1, \dots, v_{m-1})$ kannassa $(v_1, \dots, v_{m-1})$. Näin ollen
$$b_i = b_i \frac{\lambda_i}{\lambda_m}$$
jokaisella $i \in \{1, \dots, m-1\}$. Koska $\lambda_i \neq \lambda_m$ jokaisella $i \in \{1, \dots, m-1\}$, niin $b_1 = \cdots = b_{m-1} = 0$. Tällöin $v_m = 0$, mikä on jälleen ristiriita. Tämä päättää todistuksen. □

Eräs tämän lauseen seuraus on, että äärellisulotteisen vektoriavaruuden V operaattorin $f \colon V \to V$ ominaisarvojen määrä on korkeintaan avaruuden V dimensio.

Korollaari 11.3.6. *Olkoon V äärellisulotteinen vektoriavaruus ja $f \colon V \to V$ lineaarinen operaattori. Tällöin operaattorilla f on korkeintaan* $\dim V$ *erisuurta ominaisarvoa.*

Todistus. Olkoon
$$\Lambda(f) = \{\lambda \in \mathbb{R} \colon \lambda \text{ on operaattorin } f \text{ ominaisarvo}\}$$
operaattorin f kaikkien ominaisarvojen joukko. Osoitetaan, että tässä joukossa on korkeintaan $\dim V$ alkiota.

Tehdään vastaoletus, että joukossa $\Lambda(f)$ on $k > \dim V$ erisuurta alkiota. Tällöin joukon $\Lambda(f)$ määritelmän nojalla operaattorilla $f \colon V \to V$ on k erisuurta ominaisarvoa $\lambda_1 < \cdots < \lambda_k$. Olkoon $v_i \in E(\lambda_i, f)$ ominaisvektori jokaisella $i \in \{1, \dots, k\}$. Tällöin jono $(v_1, \dots, v_k)$ on vapaa. Tämä on ristiriita, koska lauseen 8.1.21 nojalla jokaisessa avaruuden V vapaassa jonossa on korkeintaan $\dim V$ alkiota.

Näin ollen joukosta $\Lambda(f)$ voidaan valita korkeintaan $\dim V$ erisuurta alkiota. Näin ollen joukossa $\Lambda(f)$ on korkeintaan $\dim V$ alkiota. □

Lauseen 11.3.5 voi tulkita suorien summien avulla seuraavasti. Lauseen todistus ja tarvittavat määritelmät on annettu liitteessä D.

Lause 11.3.7 (Ominaisavaruuksien summa on suora)**.** *Olkoon $f\colon V \to V$ lineaarinen operaattori ja olkoot $\lambda_1 < \cdots < \lambda_k$ sen ominaisarvoja. Tällöin ominaisavaruuksien summa $E(\lambda_1, f) + \cdots + E(\lambda_k, f)$ on suora, eli*

$$E(\lambda_1, f) + \cdots + E(\lambda_k, f) = E(\lambda_1, f) \oplus \cdots \oplus E(\lambda_k, f).$$

11.4 Lineaarioperaattorin determinantti

Lauseen 11.1.3 perusteella lineaarioperaattorin esitysmatriisit eri kantojen suhteen ovat similaareja eli tarkemmin: Jos $(v_1, \ldots, v_n)$ ja $(v_1', \ldots, v_n')$ ovat avaruuden V kantoja ja matriisit $A \in \mathbb{R}^{n\times n}$ ja $A' \in \mathbb{R}^{n\times n}$ lineaarioperaattorin $f\colon V \to V$ vastaavat esitysmatriisit, niin tällöin

$$A' = PAP^{-1},$$

missä P on kannanvaihtomatriisi kannasta $(v_1, \ldots, v_n)$ kantaan $(v_1', \ldots, v_n)$. Koska

$$\det A' = \det(PAP^{-1}) = (\det P)(\det A)(\det P)^{-1} = \det A,$$

niin esitysmatriisin determinantti ei riipu valitusta kannasta.

Näin ollen äärellisulotteisen vektoriavaruuden V lineaarisen operaattorin $f\colon V \to V$ determinantti voidaan määritellä sen esitysmatriisin $A \in \mathbb{R}^{n\times n}$ determinanttina.

Määritelmä 11.4.1. *Äärellisulotteisen vektoriavaruuden V lineaarisen operaattorin $f\colon V \to V$* determinantti $\det f$ *on sen esitysmatriisin A determinantti* $\det A$.

Huomaa, että jälleen sekä määritelmässä, että sitä edeltävässä laskussa esitysmatriisi lasketaan yhden kannan suhteen.

Esimerkki 11.4.2. *Identtisen kuvauksen $\mathrm{id}_V\colon V \to V$ determinantti on 1 ja nollakuvauksen $0\colon V \to V$ determinantti on 0.*

Esimerkki 11.4.3. *Jos lineaarikuvauksen $f\colon V \to V$ esitysmatriisi on diagonalisoituva eli on olemassa sellainen avaruuden V kanta $(v_1, \ldots, v_n)$, että tässä kannassa kuvauksen f esitysmatriisi A on diagonaalimatriisin, niin tällöin $\det f$ on matriisin A diagonaalialkioiden tulo.*

Kuten tulomatriisin determinantti on determinanttien tulo, lineaarioperaattorien yhdisteen determinantti on determinanttien tulo. Tämä seuraa suoraan havainnosta, että operaattoreiden yhdisteen esitysmatriisi on esitysmatriisien tulo. Todistus jätetään harjoitustehtäväksi. Kirjataan tämä tulos kuitenkin lauseeksi.

Lause 11.4.4. *Olkoot $f\colon V \to V$ ja $g\colon V \to V$ lineearioperaattoreita äärellisulotteiselta vektoriavaruudelta V itselleen. Tällöin $\det(f \circ g) = (\det f)(\det g)$.*

Vastaavasti saadaan karakterisointi operaattorin kääntyvyydelle determinantin avulla. Myös tämä tulos seuraa suoraan matriisien vastaavasta lauseesta ja jätetään harjoitustehtäväksi.

Lause 11.4.5. *Lineaarioperaattori $f\colon V \to V$ äärellisulotteiselta vektoriavaruudelta V itselleen on kääntyvä, jos ja vain jos $\det f \neq 0$. Tällöin $\det f^{-1} = (\det f)^{-1}$.*

11.5 Lineaarioperaattorin jälki

Vaikka operaattorin determinantti on määritelty esitysmatriisin avulla, determinantti liittää operaattoriin $f\colon V \to V$ luvun $\det f$, joka riippuu ainoastaan kuvauksesta itsestään eikä yksittäisestä esitysmatriisista. Toinen tälläinen karaketeristinen luku on operaattorin jälki. Aloitetaan määritelmästä matriiseille.

Määritelmä 11.5.1. *Neliömatriisin $A = [a_{ji}] \in \mathbb{R}^{n\times n}$* jälki *(engl.* trace*) on matriisin A diagonaalialkioiden summa, eli*

$$\mathrm{tr}(A) = a_{11} + \cdots + a_{nn}.$$

Kuten determinantille myös matriisin jäljelle pätee tulokaava: tulomatriisin jälki ei riipu tulontekijöiden järjestyksestä.

Lause 11.5.2. *Olkoot $A, B \in \mathbb{R}^{n\times n}$. Tällöin $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.*

Todistus. Koska

$$\mathrm{tr}(AB) = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} A_{ik}B_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n} B_{ki}A_{ik} = \sum_{k=1}^{n}(BA)_{kk} = \mathrm{tr}(BA),$$

niin väite pätee. □

Korollaari 11.5.3. *Olkoot $A \in \mathbb{R}^{n\times n}$ ja $A' \in \mathbb{R}^{n\times n}$ similaareja matriiseja. Tällöin $\mathrm{tr}A = \mathrm{tr}A'$.*

Todistus. Koska A ja A' ovat similaareja, niin on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{n\times n}$, että $A' = PAP^{-1}$. Näin ollen

$$\mathrm{tr}A' = \mathrm{tr}(PAP^{-1}) = \mathrm{tr}((PA)P^{-1}) = \mathrm{tr}(P^{-1}(PA)) = \mathrm{tr}A.$$

□

Matriisin jälki antaa siis välttämättömän, mutta ei riittävän, ehdon matriisin similaariudelle.

Esimerkki 11.5.4. *Matriisit*

$$A = \begin{bmatrix} 2 & 4 \\ -1 & 1 \end{bmatrix} \text{ ja } A' = \begin{bmatrix} -1 & 6 \\ 3 & 3 \end{bmatrix}$$

eivät ole similaareja, koska $\mathrm{tr}A = 3 \neq 2 = \mathrm{tr}A'$.

Esimerkki 11.5.5. *Matriisit*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ ja } A' = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

eivät ole similaareja, vaikka $\mathrm{tr} A = 1 = \mathrm{tr} A'$*, sillä* A' *on kääntyvä, mutta matriisi* A *ei ole.*

Koska lineaarioperaattorin esitysmatriiseilla on sama jälki, voidaan operaattorin jälki määritellä esitysmatriisin jälkenä.

Määritelmä 11.5.6. *Lineaarioperaattorin* $f\colon V \to V$ jälki $\mathrm{tr} f$ *on sen (jonkin) esitysmatriisin* A *jälki* $\mathrm{tr} A$.

Kirjataan nyt muutama matriisin jälkeen liittyvä perustulos lauseeksi. Toisin kuin determinantin tapauksessa nämä tulokset ovat hyvin suoraviivaisia todistaa ja ne jätetään harjoitustehtäviksi.

Lause 11.5.7. *Kuvaus* $\mathrm{tr}\colon \mathbb{R}^{n\times n} \to \mathbb{R}$*,* $A \mapsto \mathrm{tr} A$*, on lineaarikuvaus eli kaikilla* $A, B \in \mathbb{R}^{n\times n}$ *ja* $a \in \mathbb{R}$ *pätee*

$$\mathrm{tr}(A + aB) = \mathrm{tr}(A) + a\,\mathrm{tr}(B).$$

Lisäksi $\mathrm{tr}(A^t) = \mathrm{tr}(A)$.

Lineaarikuvauksen jälkeen palataan matriisin karakteristisen polynomin yhteydessä kurssilla *Lineaarialgebra ja matriisilaskenta III*. Annetaan sille kuitenkin tässä vaiheessa yksi mielenkiintoinen sovellus, joka sanoo, että projektion $\pi\colon V \to V$ kuvan dimension voi laskea sen esitysmatriisien jäljistä.

Lause 11.5.8. *Olkoon* $\pi\colon V \to V$ *projektio äärellisulotteiselta vektoriavaruudelta itselleen. Tällöin*

$$\mathrm{tr}\pi = \dim \operatorname{im} \pi.$$

Todistus. Olkoon $(v_1, \ldots, v_n)$ sellainen avaruuden V kanta, että $(v_1, \ldots, v_k)$ on aliavaruuden $\operatorname{im} \pi$ kanta ja että $(v_{k+1}, \ldots, v_n)$ on aliavaruuden $\ker \pi$ kanta. Projektion π esitysmatriisi kannassa $(v_1, \ldots, v_n)$ on diagonaalimatriisi $A = \begin{bmatrix} e_1 & \cdots e_k & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}$. Näin ollen

$$\mathrm{tr}\pi = \mathrm{tr} A = k = \dim \operatorname{im} \pi.$$

□

11.6 Extra: Determinantti tilavuuden muutoksena

Kuvauksen $f\colon V \to V$ determinantin määritelmä ei *a priori* liity tilavuuden muutokseen. Annetaan nyt determinantille tämä tulkinta tilavuusmuodon avulla. Tulkintaa varten tarvitaan jälleen alternoivan multilineaarikuvauksen käsite.

Määritelmä 11.6.1. *Funktio* $\omega\colon V \times \cdots \times V \to \mathbb{R}$ *on* n*-ulotteisen avaruuden* V alternoiva multilineaarikuvaus*, jos kaikilla* $w, w', w_1, \ldots, w_n \in V$ *ja* $a \in \mathbb{R}$ *pätee*

1. $\omega(w_1,\dots,aw+w',\dots,w_n) = a\omega(w_1,\dots,w,\dots,w_n) + \omega(w_1,\dots,w',\dots,w_n)$ *ja*

2. $\omega(w_1,\dots,w,\dots,w,\dots,w_n) = 0$.

Helpoin tapa muodostaa avaruuden V tilavuusmuoto on valita avaruudelle V kanta $(v_1,\dots,v_n)$ ja määritellä funktio $\mathrm{vol}_{(v_1,\dots,v_n)}\colon V\times\cdots\times V\to\mathbb{R}$ kaavalla

$$\mathrm{vol}_{(v_1,\dots,v_n)}(w_1,\dots,w_n) = \mathrm{vol}_{\mathbb{R}^n}(\Phi^{-1}_{(v_1,\dots,v_n)}(w_1),\dots,\Phi^{-1}_{(v_1,\dots,v_n)}(w_n))$$

kaikilla $w_1,\dots,w_n\in V$. Funktio $\mathrm{vol}_{(v_1,\dots,v_n)}$ on alternoiva multilineaarikuvaus, jolle pätee

$$\mathrm{vol}_{(v_1,\dots,v_n)}(v_1,\dots,v_n) = 1.$$

Nämä todistukset jätetään harjoitustehtäväksi.

Määritelmä 11.6.2. *Funktio* $\mathrm{vol}_{(v_1,\dots,v_n)}\colon V\times\cdots\times V\to\mathbb{R}$ *on avaruuden* V tilavuusmuoto kannan $(v_1,\dots,v_n)$ suhteen.

Tilavuusmuodon $\mathrm{vol}_{(v_1,\dots,v_n)}$ avulla lineaarikuvauksen $f\colon V\to V$ determinantti voidaan tulkita vektoreiden $w_1,\dots,w_n\in V$ virittämän suunnikkaan tilavuuden muutoksena kuvauksessa f.

Lause 11.6.3. *Olkoon* V *vektoriavaruus,* $(v_1,\dots,v_n)$ *avaruuden* V *kanta ja* $f\colon V\to V$ *lineaarikuvaus. Tällöin*

$$\mathrm{vol}_{(v_1,\dots,v_n)}(f(w_1),\dots,f(w_n)) = (\det f)\mathrm{vol}_{(v_1,\dots,v_n)}(w_1,\dots,w_n)$$

kaikilla $w_1,\dots,w_n\in V$.

Todistus. Olkoon $A\in\mathbb{R}^{n\times n}$ kuvauksen f esitysmatriisi kannassa $(v_1,\dots,v_n)$. Tällöin $\Phi^{-1}_{(v_1,\dots,v_n)}\circ f = f_A\circ\Phi^{-1}_{(v_1,\dots,v_n)}$.

Olkoot $w_1,\dots,w_n\in V$ ja $y_i = \Phi^{-1}_{(w_1,\dots,w_n)}(w_i)$ jokaisella $i\in\{1,\dots,n\}$. Tällöin

$$\begin{aligned}\mathrm{vol}_{(v_1,\dots,v_n)}(w_1,\dots,w_n) &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(\Phi^{-1}_{(w_1,\dots,w_n)}(w_1),\dots,\Phi^{-1}_{(w_1,\dots,w_n)}(w_n))\\ &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(y_1,\dots,y_n) = \det\begin{bmatrix}y_1 & \cdots & y_n\end{bmatrix}.\end{aligned}$$

Nyt tulomatriisin determinantin kaavan (lause 5.6.1) nojalla

$$\begin{aligned}\mathrm{vol}_{(v_1,\dots,v_n)}(f(w_1),\dots,f(w_n)) &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(\Phi^{-1}_{(v_1,\dots,v_n)}(f(w_1)),\dots,(\Phi^{-1}_{(v_1,\dots,v_n)}(f(w_1)))\\ &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(f_A(\Phi^{-1}_{(v_1,\dots,v_n)}(w_1)),\dots,f_A(\Phi^{-1}_{(v_1,\dots,v_n)}(w_1)))\\ &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(f_A(y_1),\dots,f_A(y_n))\\ &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(Ay_1,\dots,Ay_n)\\ &= \det\left(A\begin{bmatrix}y_1 & \cdots & y_n\end{bmatrix}\right)\\ &= (\det A)\left(\det\begin{bmatrix}y_1 & \cdots & y_n\end{bmatrix}\right)\\ &= (\det f)\mathrm{vol}_{(v_1,\dots,v_n)}(w_1,\dots,w_n).\end{aligned}$$

□

Huomautus 11.6.4. *Lauseen 5.6.1 todistus paljastaa, että molempien todistusten taustalla on itseasiassa lauseen 5.3.16 todistus. Lause 5.3.16 pätee täsmälleen samalla todistuksella myös avaruuden V alternoiville multilineaarikuvauksille.*

Yleisesti ottaen avaruuden V eri kannat määrittelevät eri tilavuusmuodot. Olennaisesti sama todistus kuin lauseelle 11.6.3 antaa kannanvaihdon aiheuttaman tilavuusmuodon muutoksen.

Lause 11.6.5. *Olkoon V vektoriavaruus ja olkoot $(v_1, \ldots, v_n)$ ja $(v'_1, \ldots, v'_n)$ sen kantoja. Tällöin*

$$\mathrm{vol}_{(v'_1,\ldots,v'_n)}(w_1, \ldots, w_n) = (\det P)\mathrm{vol}_{(v_1,\ldots,v_n)}(w_1, \ldots, w_n)$$

kaikilla $w_1, \ldots, w_n \in V$, missä P on kannanvaihtomatriisi kannasta $(v_1, \ldots, v_n)$ kantaan $(v'_1, \ldots, v'_n)$.

Todistus. Olkoot $w_1, \ldots, w_n \in V$. Tällöin

$$\begin{aligned}
\mathrm{vol}_{(v'_1,\ldots,v'_n)}(w_1, \ldots, w_n) &= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(\Phi^{-1}_{(v'_1,\ldots,v'_n)}(w_1), \ldots, (\Phi^{-1}_{(v'_1,\ldots,v'_n)}(w_1)) \\
&= \mathrm{vol}_{\mathbb{R}^{n\times 1}}(f_P(\Phi^{-1}_{(v_1,\ldots,v_n)}(w_1)), \ldots, f_P(\Phi^{-1}_{(v_1,\ldots,v_n)}(w_1))) \\
&= (\det P)\mathrm{vol}_{\mathbb{R}^{n\times 1}}(\Phi^{-1}_{(v_1,\ldots,v_n)}(w_1), \ldots, \Phi^{-1}_{(v_1,\ldots,v_n)}(w_1)) \\
&= (\det P)\mathrm{vol}_{(v_1,\ldots,v_n)}(w_1, \ldots, w_n).
\end{aligned}$$

$\square$

Kannanvaihtomatriisin determantin merkin avulla määritellään ovatko avaruuden kannat samoin suunnistettuja.

Määritelmä 11.6.6. *Avaruuden V kannat $(v_1, \ldots, v_n)$ ja $(v'_1, \ldots, v'_n)$ ovat* samoin suunnistetut*, jos kannasta $(v_1, \ldots, v_n)$ kantaan $(v'_1, \ldots, v'_n)$ määritellyn kannanvaihtomatriisin P determinantin* $\det P$ *merkki on positiivinen. Jos determinantin merkki on negatiivinen, niin sanotaan, että kannat ovat* vastakkaisesti suunnistetut*.*

Lauseen 11.6.5 mukaan samoin suunnistetut ortonormaalit kannat määräävät saman tilavuusmuodon.

Korollaari 11.6.7. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus ja olkoot $(u_1, \ldots, u_n)$ ja $(u'_1, \ldots, u'_n)$ sen samoin suunnistettuja ortonormaaleja kantoja. Tällöin*

$$\mathrm{vol}_{(u_1,\ldots,u_n)}(w_1, \ldots, w_n) = \mathrm{vol}_{(u'_1,\ldots,u'_n)}(w_1, \ldots, w_n)$$

kaikilla $w_1, \ldots, w_n \in V$.

Todistus. Kannanvaihtomatriisi P on ortogonaalimatriisi, jonka determinantti on positiivinen, eli $\det P = 1$. $\square$

Huomautus 11.6.8. *Määritelmä 11.6.6 antaa tulkinnan, esimerkiksi fysiikasta tutulle,* vasemman ja oikean käden koordiaatistolle*: Avaruuden* $\mathbb{R}^3$ *kantaan* (v_1, v_2, v_3) *liittyvä koordinaatisto on* vasemman käden koordinaatisto*, jos kanta* (v_1, v_2, v_3) *on samoin suunnistettu standardikannan* (e_1, e_2, e_3) *kanssa. Vastaavasti koordinaatisto on* oikean käden koordinaatisto*, jos kanta* (v_1, v_2, v_3) *on vastakkaisesti suunnistettu standarikannan kanssa.*

Huomautus 11.6.9. *Lause 11.6.3 motivoi saman ulotteisten vektoriavaruuksien lineaarikuvausten determinantin määritelmän.*

Olkoot V *ja* W n*-ulotteisia vektoriavaruuksia ja olkoot* $(v_1, \ldots, v_n)$ *avaruuden* V *kanta ja* $(w_1, \ldots, w_n)$ *avaruuden* W *kanta. Olkoon nyt* $f \colon V \to W$ *lineaarikuvaus ja* A *sen esitysmatriisi kannasta* $(v_1, \ldots, v_n)$ *kantaan* $(w_1, \ldots, w_n)$*. Olennaisesti samalla todistuksella kuin lauseessa 11.6.3 voidaan osoittaa, että*

$$\mathrm{vol}_{(w_1,\ldots,w_n)}(f(v_1'), \ldots, f(v_n')) = (\det A)\mathrm{vol}_{(v_1,\ldots,v_n)}(v_1', \ldots, v_n')$$

kaikilla $v_1', \ldots, v_n' \in V$*. Erityisesti*

$$\det A = \mathrm{vol}_{(w_1,\ldots,w_n)}(f(v_1), \ldots, f(v_n)).$$

Esitysmatriisin A *determinantti kertoo siis vektoreiden* $v_1', \ldots, v_n'$ *virittämän suunnikkaan tilavauuden muutoksen kuvauksessa* f*, kun tilavuutta mitataan tilavuusmuodoilla* $\mathrm{vol}_{(v_1,\ldots,v_n)}$ *ja* $\mathrm{vol}_{(w_1,\ldots,w_n)}$*. Determinantti* $\det A$ *riippuu tilavuusmuodoista eli valituista kannoista. Näin ollen sanotaan, että* $\det A$ *on* kuvauksen f determinantti $\det f$ avaruudesta $(V, \mathrm{vol}_{(v_1,\ldots,v_n)})$ avaruuteen $(W, \mathrm{vol}_{(w_1,\ldots,w_n)}$.

11.7 Sovellus: Ristitulo

Kolmiulotteisessa sisätuloavaruudessa $(\mathbb{R}^{3\times1}, \cdot)$ kaksiulotteisen aliavaruuden (eli origon kautta kulkevan tason) $W \subset \mathbb{R}^{3\times1}$ kohtisuorakomplementti $W^\perp$ on yksiulotteinen aliavaruus (eli origon kautta kulkeva suora). Tämä tarkoittaa sitä, että kaikki tasoa W vastaan kohtisuorat vektorit ovat toistensa skaalauksia.

Koska tason W virittäjät w_1 ja w_2 yhdessä suoralle $W^\perp$ valitun virittäjän $v \in W^\perp$ kanssa muodostavat avaruuden V kannan, niin suoralle $W^\perp$ voidaan valita sellainen virittäjä $v \in W^\perp$, että kanta (w_1, w_2, v) on ns. *positiivisesti suunnistettu* eli että $\det \begin{bmatrix} w_1 & w_2 & v \end{bmatrix} > 0$.

Tässä kuvaillussa menetelmässä suoran $W^\perp$ virittäjän löytämiselle tulee ensin ratkaista yhtälöryhmä aliavaruuden $W^\perp$ löytämiseksi ja tämän jälkeen tehdä valinta suoran $W^\perp$ virittäjän $v \in W^\perp$ suhteen. Avaruuden $\mathbb{R}^{3\times1}$ ristitulo antaa suoran menetelmän geometrisesti mielekkään tavan valita virittäjä käyttäen ainoastaan tason W virittäjävektoreita w_1 ja w_2.

Määritelmä 11.7.1. *Funktiota* $\times\colon \mathbb{R}^{3\times 1}\times\mathbb{R}^{3\times 1}\to\mathbb{R}^{3\times 1}$, *joka on määritelty kaavalla*

$$\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\right)\mapsto\begin{bmatrix}\det\begin{bmatrix}x_2 & y_2\\x_3 & y_3\end{bmatrix}\\-\det\begin{bmatrix}x_1 & y_1\\x_3 & y_3\end{bmatrix}\\\det\begin{bmatrix}x_1 & y_1\\x_2 & y_2\end{bmatrix}\end{bmatrix},$$

kutsutaan avaruuden $\mathbb{R}^{3\times 1}$ ristituloksi *(engl.* cross product*).*

Lähdetään nyt osoittamaan, että ristitulolla on halutut ominaisuudet. Vektorin $w_1\times w_2$ kohtisuoruus vektoreita w_1 ja w_2 vastaan seuraa seuraavasta havainnosta.

Lause 11.7.2. *Olkoot* $w_1,w_2,w_3\in\mathbb{R}^{3\times 1}$*. Tällöin*

$$(w_1\times w_2)\cdot w_3=\det\begin{bmatrix}w_1 & w_2 & w_3\end{bmatrix}.$$

Todistus. Olkoot

$$w_1=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix},\ w_2=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}\ \text{ja}\ w_3=\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix}.$$

Tällöin

$$\begin{aligned}(w_1\times w_2)\cdot w_3&=\begin{bmatrix}\det\begin{bmatrix}a_2 & b_2\\a_3 & b_3\end{bmatrix}\\-\det\begin{bmatrix}a_1 & b_1\\a_3 & b_3\end{bmatrix}\\\det\begin{bmatrix}a_1 & b_1\\a_2 & b_2\end{bmatrix}\end{bmatrix}\cdot\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix}\\&=c_1\det\begin{bmatrix}a_2 & b_2\\a_3 & b_3\end{bmatrix}-c_2\det\begin{bmatrix}a_1 & b_1\\a_3 & b_3\end{bmatrix}+c_3\det\begin{bmatrix}a_1 & b_1\\a_2 & b_2\end{bmatrix}\\&=\det\begin{bmatrix}c_1 & a_1 & b_1\\c_2 & a_2 & b_2\\c_3 & a_3 & b_3\end{bmatrix}=\det\begin{bmatrix}a_1 & b_1 & c_1\\a_2 & b_2 & c_2\\a_3 & b_3 & c_3\end{bmatrix}=\det\begin{bmatrix}w_1 & w_2 & w_3\end{bmatrix}.\end{aligned}$$

□

Korollaari 11.7.3. *Olkoot* $w_1,w_2\in\mathbb{R}^{3\times 1}$*. Tällöin vektori* $w_1\times w_2$ *on kohtisuorassa vektoreita* w_1 *ja* w_2 *vastaan eli* $w_1\times w_2\in\operatorname{Sp}(w_1,w_2)^\perp$*.*

Todistus. Väite seuraa suoraan lauseesta 11.7.2, sillä

$$(w_1\times w_2)\cdot w_1=\det\begin{bmatrix}w_1 & w_2 & w_1\end{bmatrix}=0$$

ja

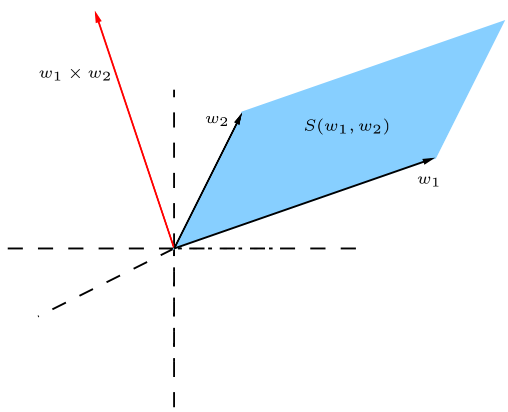
$$(w_1\times w_2)\cdot w_2=\det\begin{bmatrix}w_1 & w_2 & w_2\end{bmatrix}=0.$$

□

Jäljellä on osoittaa, että $w_1 \times w_2$ virittää yksiulotteisen aliavaruuden $\mathrm{Sp}(w_1, w_2)^\perp$, jos vektorit w_1 ja w_2 ovat lineaarisesti riippumattomia. Koska riittää osoittaa, että $|w_1 \times w_2| > 0$, niin on luonnollisinta osoittaa, että vektorin $w_1 \times w_2$ pituudella on geometrinen tulinta:

> Vektorin $w_1 \times w_2$ pituus on vektoreiden w_1 ja w_2 virittämän 2-ulotteisen suunnikkaan $S(w_1, w_2) = \{tw_1 + sw_2 \in \mathbb{R}^{3\times 1} : t, s \in [0,1]\}$ pinta-ala.

Vaikka tätä tulkintaa on helppo havainnollistaa symbolisella kuvalla (kuten kuva 11.1), niin tässä tulkinnassa on selvä heikkous: *Mitä tarkoittaa 3-ulotteisen avaruuden $\mathbb{R}^{3\times 1}$ 2-ulotteisen suunnikkaan pinta-ala?*



Kuva 11.1: Vektoreiden w_1 ja w_2 ristitulo $w_1 \times w_2$ ja suunnikas $S(w_1, s_2)$.

Vältetään jälleen pinta-alan käsitteen tarkkaa määrittelemistä ja annetaan sen sijaan ristitulolle kaava 2-ulotteisessa aliavaruudessa W ortonormaalin kannan avulla. Palataan tämän jälkeen tämän lauseen pinta-alatulkintaan.

Lause 11.7.4. *Olkoon $W \subset \mathbb{R}^{3\times 1}$ 2-ulottinen aliavaruus, (u_1, u_2) aliavaruuden W ortonormaali kanta ja $u_3 \in W^\perp$ sellainen yksikkövektori, että* $\det \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} > 0$. *Tällöin kaikilla $w_1 = au_1 + bu_2 \in W$ ja $w_2 = cu_1 + du_2$ pätee*

$$w_1 \times w_2 = (ad - bc)u_3.$$

Huomautus 11.7.5. *Koska kanta (u_1, u_2, u_3) on ortonormaali, niin matriisi $U = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$ on ortogonaalinen. Koska ortogonaalimatriisin determinantti on joko 1 tai -1, niin oletuksesta seuraa, että* $\det U = 1$.

Osoitetaan lauseen 11.7.4 todistusta varten, että ristitulo $\times : \mathbb{R}^{3\times 1} \times \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$ on alternoiva bilineaarikuvaus. Kirjataan tämä tulos lyhyesti seuraavaan muotoon. Koska väite seuraa suoraan 2×2-matriisien determinantin vastaavista ominaisuuksista, todistus jätetään harjoitustehtäväksi.

Lemma 11.7.6. *Olkoot $w, w', w'' \in \mathbb{R}^{3\times 1}$ ja $a \in \mathbb{R}$. Tällöin*

1. $(aw + w') \times w'' = a(w \times w'') + w' \times w''$ *ja*

2. $w \times w = 0.$

Erityisesti

$$w' \times w = -w \times w'.$$

Näiden valmistelujen jälkeen voidaan siirtyä lauseen 11.7.4 todistukseen.

Lauseen 11.7.4 todistus. Olkoot $w_1 = au_1 + bu_2 \in W$ ja $w_2 = cu_1 + du_2$. Tällöin lemman 11.7.6 perusteella

$$\begin{aligned} w_1 \times w_2 &= (au_1 + bu_2) \times (cu_1 + du_2) = au_1 \times (cu_1 + du_2) + bu_2 \times (cu_1 + du_2) \\ &= ac(u_1 \times u_1) + ad(u_1 \times u_2) + bc(u_2 \times u_1) + bd(u_2 \times u_2) \\ &= ad(u_1 \times u_2) + bc(u_2 \times u_1) = (ad - bc)(u_1 \times u_2). \end{aligned}$$

Näin ollen riittää osoittaa, että $u_1 \times u_2 = u_3$. Korollaarin 11.7.3 perusteella $u_1 \times u_2 \in \mathrm{Sp}(u_1, u_2)^\perp = \mathrm{Sp}(u_3)$ eli $u_1 \times u_2 = ((u_1 \times u_2) \cdot u_3)\, u_3$. Lauseen 11.7.2 perusteella

$$(u_1 \times u_2) \cdot u_3 = \det \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = 1.$$

Näin ollen $u_1 \times u_2 = u_3$. Tämä päättää todistuksen. □

Lauseen 11.7.4 voi tulkita kahdella eri tavalla. Annetaan ensin tulkinta kuvauksen determinanttina.

Huomautus 11.7.7. *Olkoot* $w_1 = au_1 + bu_2, w_2 = cu_1 + du_2 \in W$ *ja olkoon* $f\colon W \to W$ *lineaarioperaattori, jolle pätee sekä* $f(u_1) = w_1$ *että* $f(u_2) = w_2$*. Tällöin operaattorin* f *esitysmatriisi kannassa* (u_1, u_2) *on*

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

Näin ollen $\det f = \det A = ad - bc$*, eli*

$$w_1 \times w_2 = f(u_1) \times f(u_2) = (\det f) u_3.$$

Tulkitaan nyt lause 11.7.4 avaruuden W tilavuusmuodon $\mathrm{vol}_{(u_1,u_2)}$ avulla.

Huomautus 11.7.8. *Koska* $\mathrm{vol}_{(u_1,u_2)}$ *on alternoiva ja bilineaarinen, niin sama lasku kuin lauseen 11.7.4 todistuksessa antaa*

$$\begin{aligned} \mathrm{vol}_{(u_1,u_2)}(w_1, w_2) &= \mathrm{vol}_{(u_1,u_2)}(au_1 + bu_2, cu_1 + du_2) \\ &= ac\mathrm{vol}_{(u_1,u_2)}(u_1, u_1) + ad\mathrm{vol}_{(u_1,u_2)}(u_1, u_2) \\ &\quad + bc\mathrm{vol}_{(u_1,u_2)}(u_2, u_1) + bd\mathrm{vol}_{(u_1,u_2)}(u_2, u_2) \\ &= (ad - bc)\mathrm{vol}_{(u_1,u_2)}(u_1, u_2). \end{aligned}$$

Koska $\mathrm{vol}_{(u_1,u_2)}(u_1, u_2) = 1$*, niin*

$$\mathrm{vol}_{(u_1,u_2)}(w_1, w_2) = ad - bc.$$

Näin ollen $ad - bc$ *on suunnikkaan* $S(w_1, w_2)$ *pinta-ala tasossa* W *(luonnollisen) tilavuusmuodon* $\mathrm{vol}_{(u_1,u_2)}$ *suhteen.*

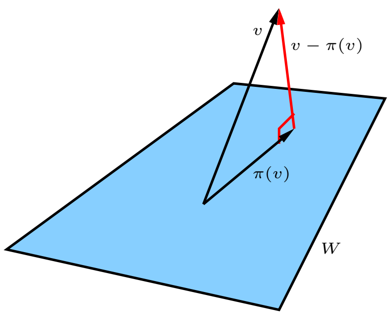
Luku 12

Ortogonaaliprojektioista adjungaatteihin

Tässä luvussa käsitellään ortogonaaliprojektioita, niiden sovelluksia ja yhteyttä lineaarikuvausten adjungaatteihin. Ortogonaaliprojektion tärkeimmät konkreettiset sovellukset liittyvät geometriaan: pisteen etäisyyden määrittäminen aliavaruudesta ja siihen liittyvä pienimmän neliösumman ongelma. Nämä ongelmat esiintyvät tyypillisesti sarakeavaruuden aliavaruuksien tapauksessa eli ovat matriisien sarakeavaruuksien kysymyksiä. Tämän vuoksi on luonnollista myös ratkaista nämä ongelmat tässä kontekstissa. Ratkaisut herättävät kuitenkin kysymyksiä yleisistä geometrisista periaatteista ratkaisujen takana. Tämä johtaa adjungaattien yleiseen teoriaan.

12.1 Ortogonaaliprojektiot

Palautetaan mieleen luvusta 8.3.2, että lineaarikuvaus $\pi\colon V \to V$ on projektio aliavaruudelle $W \subset V$, jos $\pi \circ \pi = \pi$ ja $\operatorname{im} \pi = W$. Jos vektoriavaruuden V sijaan tarkastellaan sisätuloavaruutta $(V, \langle\cdot,\cdot\rangle)$, voidaan ehdon $\pi \circ \pi = \pi$ sijaan vaatia kohtisuoruutta aliavaruutta W vastaan kuten kuvassa 12.1.



Kuva 12.1: Vektorin $v \in V$ kuva $\pi(v) \in W$ ortogonaaliprojektiossa $\pi\colon V \to V$ aliavaruudelle W.

Näin saavutaan seuraavaan ortogonaaliprojektion määritelmään.

Määritelmä 12.1.1. *Sisätuloavaruuden* $(V, \langle\cdot,\cdot\rangle)$ *lineaarikuvaus* $\pi\colon V \to V$ *on* ortogonaaliprojektio*, jos jokaisella* $v \in V$ *erotusvektori* $v - \pi(v)$ *on kohtisuorassa kuva-avaruutta* $\operatorname{im}\pi$ *vastaan, eli kaikilla* $v \in V$ *ja* $w \in \operatorname{im}\pi$ *pätee* $\langle v - \pi(v), w\rangle = 0$*. Lisäksi sanotaan, että* π *on* ortogonaaliprojektio aliavaruudelle $W \subset V$*, jos* $\operatorname{im}\pi = W$*.*

Varmistetaan vielä, että ortogonaaliprojektio todellakin on projektio.

Lemma 12.1.2. *Olkoon* $(V, \langle\cdot,\cdot\rangle)$ *sisätuloavaruus ja* $\pi\colon V \to V$ *ortogonaaliprojektio. Tällöin* π *on projektio eli että* $\pi \circ \pi = \pi$*.*

Todistus. Olkoon $W = \operatorname{im}\pi$. Olkoot $v \in V$ ja $w = \pi(v) \in V$. Riittää osoittaa, että $\|\pi(w) - w\| = 0$.

Koska $w \in W$ ja $\pi(w) \in W$, niin

$$\|\pi(w) - w\|^2 = \langle \pi(w) - w, \pi(w) - w\rangle = \langle \pi(w) - w, \pi(w)\rangle - \langle \pi(w) - w, w\rangle = 0.$$

Näin ollen $\pi(w) = w$ eli $\pi(\pi(v)) = \pi(v)$. Kuvaus π on siis projektio. □

12.1.1 Ortogonaaliprojektion olemassaolo ja yksikäsitteisyys

Seuraava lause sanoo, että sisätuloavaruuden jokaiselle äärellisulotteiselle aliavaruudelle on yksikäsitteinen ortogonaaliprojektio ja että haluttu ortogonaaliprojektio π_W voidaan määritellä aliavaruuden W ortonormaalin kannan avulla.

Lause 12.1.3. *Olkoon* $(V, \langle\cdot,\cdot\rangle)$ *sisätuloavaruus ja* $W \subset V$ *äärellisulotteinen aliavaruus. Tällöin on olemassa täsmälleen yksi ortogonaaliprojektio* $\pi_W\colon V \to V$ *aliavaruudelle* W*. Lisäksi, jos* $(w_1, \ldots, w_m)$ *on aliavaruuden* W *ortonormaali kanta, niin*

$$\pi_W(v) = \langle w_1, v\rangle w_1 + \cdots + \langle w_m, v\rangle w_m.$$

kaikilla $v \in V$*.*

Todistus. Olkoon $(w_1, \ldots, w_m)$ aliavaruuden W ortonormaali kanta. Määritellään kuvaus $\pi_W\colon V \to V$ kaavalla

$$\pi_W(v) = \langle w_1, v\rangle w_1 + \cdots + \langle w_m, v\rangle w_m.$$

Selvästi π_W on lineaarinen ja $\operatorname{im}\pi_W = W$. Osoitetaan, että tämä kuvaus on haluttu ortogonaaliprojektio aliavaruudelle W.

Jatketaan ensin vapaa jono $(w_1, \ldots, w_m)$ avaruuden V ortonormaaliksi kannaksi $(w_1, \ldots, w_n)$. Olkoon $v \in V$. Tällöin

$$\begin{aligned} v &= (\langle w_1, v\rangle w_1 + \cdots + \langle w_m, v\rangle w_m) + (\langle w_{m+1}, v\rangle w_{m+1} + \cdots + \langle w_n, v\rangle w_n) \\ &= \pi_W(v) + (\langle w_{m+1}, v\rangle w_{m+1} + \cdots + \langle w_n, v\rangle w_n)\,. \end{aligned}$$

Näin ollen

$$v - \pi_W(v) = \langle w_{m+1}, v\rangle w_{m+1} + \cdots + \langle w_n, v\rangle w_n.$$

eli $v - \pi_W(v) \in \mathrm{Sp}(w_{m+1}, \dots, w_n)$. Koska $(w_1, \dots, w_n)$ on ortonormaali kanta, niin vektori $v - \pi_W(v)$ on kohtisuorassa jokaista vektoria $w \in \mathrm{Sp}(w_1, \dots, w_m) = \operatorname{im} \pi$ vastaan. Kuvaus π_W on siis ortogonaaliprojektio aliavaruudelle W.

Osoitetaan vielä ortogonaaliprojektion yksikäsitteisyys. Olkoon $\pi \colon V \to V$ ortogonaaliprojektio aliavaruudelle W. Osoitetaan, että $\|\pi(v) - \pi_W(v)\| = 0$ jokaisella $v \in V$. Olkoon $v \in V$. Koska $\pi(v) - \pi_W(v) \in W$, niin

$$\begin{aligned}\|\pi(v) - \pi_W(v)\|^2 &= \langle \pi(v) - \pi_W(v), \pi(v) - \pi_W(v) \rangle \\ &= \langle (\pi(v) - v) - (\pi_W(v) - v), \pi(v) - \pi_W(v) \rangle \\ &= \langle \pi(v) - v, \pi(v) - \pi_W(v) \rangle - \langle \pi_W(v) - v, \pi(v) - \pi_W(v) \rangle = 0.\end{aligned}$$

□

Ortogonaaliprojektion yksikäsitteisyyden voi osoittaa myös ainakin kahdella vaihtoehtoisella tavalla. Ortogonaaliprojektio $\pi \colon V \to V$ aliavaruudelle W kuvaa pisteen yksikäsitteiselle lähimmälle pisteelle aliavaruuteen W. Tämä tulos todistetaan lauseessa 12.1.6.

Toisaalta voidaan myös muokata edellä annettua todistusta ja osoittaa, että vektorit $v - \pi(v)$ ja $v - \pi_W(v)$ eivät voi olla samanaikaisesti kohtisuorassa aliavaruutta W vastaan, jos $\pi(v) \neq \pi_W(v)$. Tämä todistus motivoi seuraavan huomion, joka karakterisoi ortogonaaliprojektiot kaikkien projektioiden joukossa.

Lemma 12.1.4. *Olkoon $(V, \langle \cdot, \cdot \rangle)$ sisätuloavaruus ja $\pi \colon V \to V$ projektio. Tällöin π on ortogonaaliprojektio, jos ja vain jos ydin $\ker \pi$ on kohtisuorassa kuvaa $\operatorname{im} \pi$ vastaan.*

Todistus. Oletetaan ensin, että π on ortogonaaliprojektio. Olkoon $v \in \ker \pi$. Tällöin $v - \pi(v) = v$, joten kaikilla $w \in \operatorname{im} \pi$ pätee $\langle v, w \rangle = \langle v - \pi(v), w \rangle = 0$. Näin ollen ydin $\ker \pi$ on kohtisuorassa kuvaa $\operatorname{im} \pi$ vastaan.

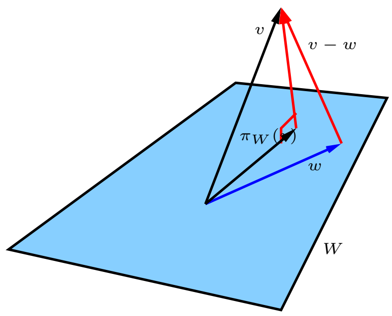
Oletetaan nyt, että ydin $\ker \pi$ on kohtisuorassa kuvaa vastaan. Osoitetaan, että π on ortogonaaliprojektio. Olkoot $v \in V$ ja $w \in \operatorname{im} \pi$. Koska π on projektio, niin $\pi(\pi(v)) = \pi(v)$. Näin ollen $\pi(v - \pi(v)) = \pi(v) - \pi(\pi(v)) = \pi(v) - \pi(v) = 0$ eli $v - \pi(v) \in \ker \pi$. Näin ollen $v - \pi(v)$ on oletuksen nojalla kohtisuorassa vektoria w vastaan. Kuvaus π on siis ortogonaaliprojektio. □

Huomautus 12.1.5. *Lauseen 12.1.3 todistus on siitä mielenkiintoinen, että projektio π_W on yksikäsitteinen, mutta sen kaava annetaan ortonormaaliin kannan avulla, joka ei ole yksikäsitteinen. Tämä näennäinen ristiriita tulee kuitenkin ymmärretyksi, kun havaitaan, että ortonormaalin kannan rooli todistuksessa on antaa avaruuden W ortokomplementti $W^\perp$ ja tämän jälkeen osoittaa, että $v - \pi(v) \in W^\perp$ jokaisella $v \in V$. Tähän palataan vielä.*

12.1.2 Pisteen etäisyys aliavaruudesta

Ortogonaaliprojektion määritelmä annetaan kohtisuoruuden avulla, mutta sillä on syvempi merkitys:

Jos $\pi_W \colon V \to V$ on ortogonaaliprojektio aliavaruudelle W, niin kuvavektori $\pi_W(v)$ on se aliavaruuden W vektori, joka on lähimpänä vektoria v; katso kuva 12.2.



Kuva 12.2: Vektoreiden $v \in V$ ja $w \in W$ etäisyys $\|v-w\|$ on vähintään yhtäsuuri kuin vektoreiden v ja $\pi_W(v)$ etäisyys $\|v - \pi_W i(v)\|$.

Kirjataan tämä tulos lauseeksi sen tärkeyden vuoksi.

Lause 12.1.6. *Olkoon* $(V, \langle\cdot,\cdot\rangle)$ *sisätuloavaruus ja* $\pi_W \colon V \to V$ *ortogonaaliprojektio aliavaruudelle* W*. Tällöin kaikilla* $v \in V$ *ja* $w \in W \setminus \{\pi(v)\}$ *pätee*

$$\|v - \pi_W(v)\| < \|v - w\|.$$

Erityisesti kaikilla $v \in V$ *pätee*

$$\|v - \pi_W(v)\| = \min\{\|v-w\| \colon w \in W\}.$$

Todistus. Olkoon $v \in V$ ja $w \in W$. Merkitään $u = \pi_W(v) - w$. Koska π_W on ortogonaaliprojektio aliavaruudelle W ja $u \in W$, niin $\langle v - \pi_W(v), u\rangle = 0$. Näin ollen

$$\begin{aligned}
\|v-w\|^2 &= \|v - \pi_W(v) + \pi_W(v) - w\|^2 \\
&= \|v - \pi_W(v) + u\|^2 \\
&= \langle v - \pi_W(v) + u, v - \pi_W(v) + u\rangle \\
&= \|v - \pi_W(v)\|^2 + 2\langle v - \pi_W(v), u\rangle + \|u\|^2 \\
&= \|v - \pi_W(v)\|^2 + \|u\|^2.
\end{aligned}$$

Jos $w \neq \pi_W(v)$, niin $u \neq 0$. Tällöin

$$\|v-w\|^2 = \|v - \pi_W(v)\|^2 + \|u\|^2 > \|v - \pi_W(v)\|^2.$$

Tämä todistaa väitteen. $\square$

Lauseiden 12.1.6 ja 12.1.3 perusteella voidaan määritellä pisteen etäisyys aliavaruudesta.

Määritelmä 12.1.7. *Olkoon* $(V, \langle\cdot,\cdot\rangle)$ *sisätuloavaruus. Pisteen* $v \in V$ *etäisyys äärellisulotteisesta aliavaruudesta* $W \subset V$ *on* $\|v - \pi_W(v)\|$*, missä* $\pi_W \colon V \to V$ *on ortgonaaliprojektio aliavaruudelle* W*.*

12.2 Ortogonaaliprojektio matriisin sarakeavaruudelle I

Tärkein erikoistapaus ortogonaaliprojektioista on ortogonaaliprojektio $\pi_{\mathrm{Col}(A)}\colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{m\times 1}$ matriisin $A \in \mathbb{R}^{m\times n}$ sarakeavaruudelle $\mathrm{Col}(A) \subset \mathbb{R}^{m\times 1}$. Huomaa, että sisätulo tässä tapauksessa on avaruuden $\mathbb{R}^{m\times 1}$ pistetulo.

Tässä erikoistapauksessa ortogonaaliprojektiolle $\pi_{\mathrm{Col}(A)}$ voidaan antaa konkreettinen kaava matriisin A avulla. Huomaa, että A ei kuitenkaan ole projektion $\pi_{\mathrm{Col}(A)}$ matriisi.

Lause 12.2.1. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi, jonka sarakkeet ovat lineaarisesti riippumattomia. Tällöin ortogonaaliprojektio $\pi_{\mathrm{Col}(A)}\colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{m\times 1}$ matriisin A sarakeavaruudelle* $\mathrm{Col}(A)$ *on kuvaus*

$$y \mapsto A\left(A^t A\right)^{-1} A^t y. \tag{12.1}$$

Lause on erittäin mielenkiintoinen. Sen voi todistaa näennäisesti kolmella eri tavalla. Ensimmäinen tapa on käyttää yleistä teoriaa. Tiedetään, että projektio on ortogonaaliprojektio, jos ydin $\ker\pi$ on kohtisuorassa kuvaa $\mathrm{im}\,\pi$ vastaan, ja että ortogonaaliprojektio aliavaruudelle on yksikäsitteinen. Näin ollen riittää osoittaa, että matriisi A^tA on kääntyvä, jolloin kuvaus $\pi\colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{m\times 1}$, $y \mapsto A(A^tA)^{-1}A^ty$, on määritelty, ja tarkistaa, että π on sellainen projektio aliavaruudelle $\mathrm{Col}(A)$, jonka ydin on kohtisuorassa kuvaa vastaan. Tämä todistus ei anna lainkaan intuitiota siitä, miten kaava (12.1) on löydetty tai mihin se liittyy.

Toinen tapa on johtaa kaava yhtälöistä, jotka seuraavat kohtisuoruudesta. Tämä geometrinen todistus johtaa ns. normaaliyhtälöön ja yhdistää ortogonaaliset projektiot geometriseen ongelmaan pisteen etäisyyteen aliavaruudesta. Tämä tapa selittää transpoosit kaavassa (12.1), mutta ei anna konseptuaalista yhteyttä kuvauksen $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ ja projektion $\pi_{\mathrm{Col}(A)}$ välille.

Viimeinen tapa on tarkastella matriiseja A ja A^t lineaarikuvauksina $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ ja $f_{A^t}\colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{n\times 1}$ ja liittää ortogonaaliprojektio $\pi_{\mathrm{Col}(A)}$ näiden kuvausten ominaisuuksiin. Tämä puolestaan johtaa adjungaatin käsitteeseen ja on muita tapoja teoreettisempi.

Kirjataan ennen ensimmäisen todistuksen aloittamista yleiseksi lemmaksi tulos, että matriisi A^tA on kääntyvä, jos matriisin A sarakkeet ovat lineaarisesti riippumattomia..

Lemma 12.2.2. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi, jonka sarakkeet ovat lineaarisesti riippumattomia. Tällöin $A^tA \in \mathbb{R}^{n\times n}$ on kääntyvä.*

Todistus. Olkoon $x \in \mathrm{Null}(A^tA)$. Tällöin

$$0 = x^tA^tAx = (Ax)^t \cdot (Ax)$$

eli $Ax = 0$. Koska matriisin A sarakkeet ovat lineaarisesti riippumattomia, niin $x = 0$. Näin ollen $\mathrm{Null}(A^tA) = \{0\}$ eli matriisi A^tA on kääntyvä. $\square$

Lauseen 12.2.1 ensimmäinen todistus. Lemman 12.2.2 perusteella matriisi A^tA on kääntyvä ja näin ollen kuvaus $\pi\colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{m\times 1}$, $y \mapsto A(A^tA)^{-1}A^ty$, on hyvin määritelty.

Osoitetaan ensin, että π on projektio. Olkoon $y \in \mathbb{R}^{m\times 1}$. Tällöin

$$\begin{aligned}\pi(\pi(y)) &= A(A^tA)^{-1}A^t\left(A(A^tA)^{-1}A^ty\right)\\ &= A(A^tA)^{-1}(A^tA)(A^tA)^{-1}A^ty = A(A^tA)^{-1}A^ty = \pi(y).\end{aligned}$$

Kuvaus π on siis projektio.

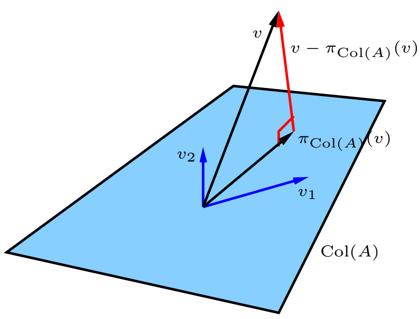
Osoitetaan nyt, että ydin π on kohtisuorassa kuvaa $\operatorname{im}\pi$ vastaan. Olkoot $v \in \ker\pi$ ja $w \in \operatorname{im}\pi$. Olkoon lisäksi $v' \in V$ sellainen vektori, että $\pi(v') = w$. Nyt tulomatriisin transpoosin kaavan perusteella

$$\begin{aligned}v \cdot w = v \cdot \pi(v') &= v \cdot A(A^tA)^{-1}A^tv' = v^t\left(A(A^tA)^{-1}A^t\right)v'\\ &= \left((A(A^tA)^{-1}A^t)^tv\right)^tv' = \left((A(A^tA)^{-1}A^t)v\right)^tv' = \pi(v)\cdot v' = 0.\end{aligned}$$

Näin ollen lemman 12.1.4 perusteella π on ortogonaaliprojektio aliavaruudelle $\mathrm{Col}(A)$. Lauseen 12.1.3 nojalla $\pi_{\mathrm{Col}(A)} = \pi$. □

12.2.1 Lauseen 12.2.1 geometrinen todistus

Olkoon $A = \begin{bmatrix} v_1 & \cdots & v_n\end{bmatrix}$ matriisi, jonka sarakkeet ovat lineaarisesti riippumaattomia. Aloitetaan lauseen 12.2.1 todistaminen tarkastelemalla kuvaa 12.3.



Kuva 12.3: Vektorin $v \in V$ kuva $\pi_{\mathrm{Col}(A)}(v) \in W$ ortogonaaliprojektiossa $\pi_{\mathrm{Col}(A)}\colon V \to V$, missä $A = \begin{bmatrix} v_1 & v_2\end{bmatrix}$.

Koska vektorit $v_1, \ldots, v_n$ ovat lineaarisesti riippumattomia, niin jono $(v_1, \ldots, v_n)$ on aliavaruuden $\mathrm{Col}(A)$ kanta. Koska vektori $v - \pi_{\mathrm{Col}(A)}(v)$ on kohtisuorassa aliavaruutta $\mathrm{Col}(A)$ vastaan, niin erityisesti

$$(v - \pi_{\mathrm{Col}(A)}(v)) \cdot v_i = 0$$

jokaisella $i \in \{1, \ldots, n\}$. Merkitään $w = \pi_{\mathrm{Col}(A)}(v)$.

Nyt sarakeavaruuden määritelmän nojalla $w = Ax$ jollain $x \in \mathbb{R}^{n\times 1}$. Koska lisäksi $v_i = Ae_i$ jokaisella $i \in \{1, \ldots, \}$, niin jokaisella $i \in \{1, \ldots, n\}$ saadaan yhtälö

$$(v - Ax) \cdot Ae_i = 0.$$

Koska

$$(v - Ax) \cdot Ae_i = (v - Ax)^t Ae_i = (A^t(v - Ax))^t e_i = (A^t(v - Ax)) \cdot e_i$$

niin havaitaan, että vektori $A^t(v - Ax) \in \mathbb{R}^{n\times 1}$ on nollavektori, sillä

$$A^t(v - Ax) = (A^t(v - Ax) \cdot e_1)e_1 + \cdots + (A^t(v - Ax) \cdot e_n)e_n = 0.$$

Näin ollen

$$A^t v = A^t Ax. \tag{12.2}$$

Koska lemman 12.2.2 perusteella matriisi A^tA on kääntyvä, niin saadaan

$$x = (A^tA)^{-1}A^t v.$$

Näin ollen

$$\pi(v) = w = Ax = A(A^tA)^{-1}A^t v.$$

Tämä päättää lauseen 12.2.1 toisen todistuksen.

Muutama huomio on paikallaan.

Huomautus 12.2.3. *Tätä todistusta kutsutaan tässä yhteydessä geometriseksi todistukseksi, koska se perustuu kuvan 12.3 havaintoon, että vektori $v - \pi_{\mathrm{Col}(A)}(v)$ on kohtisuorassa kantavektoreita $v_1, \ldots, v_n$ vastaan. Tämä johtaa yhtälöön* (12.2)*, joka ratkaistaan käyttäen matriisin A^tA kääntyvyyttä. Tässä yhteydessä yhtälö* (12.2) *ja lemma 12.2.2 eivät kuitenkaan saa geometrista tulkintaa. Näitä asioita käsitellään lauseen 12.2.1 kolmannessa todistuksessa.*

12.3 Sovellus: Pienimmän neliösumman ongelma

Ennen lauseen 12.2.1 kolmatta todistusta, sovelletaan ortogonaaliprojektioita pienimmän neliösumman ratkaisun olemassaoloon. Esitellään ensin ratkaistavana oleva kysymys.

Klassisesti lineaarisen yhtälöryhmän ratkaiseminen vastaa seuraavaa kysymystä.

Ongelma. *Olkoon $y \in \mathbb{R}^{m\times 1}$ ja $A \in \mathbb{R}^{m\times n}$. Ratkaise yhtälö*

$$Ax = y. \tag{12.3}$$

Jos vektori ei kuulu sarakeavaruuteen $\mathrm{Col}(A)$, niin yhtälöllä (12.3) ei ole ratkaisua. Yleisesti siis yhtälöllä $Ax = y$ ei voi olettaa olevan ratkaisuja.

Jos $\mathrm{Col}(A) \neq \mathbb{R}^{m\times 1}$, niin onkin luonnollista kysyä, mikä vektori $x \in \mathbb{R}^{n\times 1}$ antaa parhaimman *approksimatiivisen ratkaisun*, eli millä vektorilla $\hat{x} \in \mathbb{R}^{n\times 1}$ vektoreiden y ja $A\hat{x}$ etäisyys on mahdollisimman pieni. Uusi ongelma on siis seuraava:

Ongelma. *(Pienimmän neliösumman ongelma) Olkoon $y \in \mathbb{R}^{m\times 1}$ ja $A \in \mathbb{R}^{m\times n}$. Etsi vektori $\hat{x} \in \mathbb{R}^{n\times 1}$, jolle pätee*

$$\|y - A\hat{x}\| = \min\{\|y - Ax\| \colon x \in \mathbb{R}^{n\times 1}\}.$$

Koska etäisyyden $\|y - Ax\|$ minimi saadaan samalla vektorilla kuin etäisyyden neliön $\|y - Ax\|^2$, niin tätä approksimatiivista ratkaisua kutsutaan myös *pienimmän neliösumman ratkaisuksi*[1].

Koska $Ax \in \mathrm{Col}(A)$ kaikilla $x \in \mathbb{R}^{n\times 1}$, niin pienin mahdollinen etäisyys on lauseen 12.1.6 perusteella $\|y - \pi_{\mathrm{Col}(A)}(y)\|$. Näin ollen haettu approksimatiivinen ratkaisu on siis vektori $\hat{x} \in \mathbb{R}^{n\times 1}$, joka ratkaisee yhtälön

$$A\hat{x} = \pi_{\mathrm{Col}(A)}(y).$$

Jos matriisin A sarakkeet ovat lineaarisesti riippumattomia, tämä yhtälö osataan ratkaista yksikäsitteisesti projektion $\pi_{\mathrm{Col}(A)}$ matriisiesityksen avulla. Kirjataan tämä tulos seuraavassa muodossa.

Lause 12.3.1. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi, jonka sarakkeet ovat lineaarisesti riippumattomia ja $y \in \mathbb{R}^{m\times 1}$. Tällöin vektorille $\hat{x} \in \mathbb{R}^{n\times 1}$ pätee*

$$\|y - A\hat{x}\| = \min\{\|y - Ax\| \colon x \in \mathbb{R}^{n\times 1}\}$$

jos ja vain jos

$$A^tA\hat{x} = A^ty.$$

Huomautus 12.3.2. *Koska matriisi A^tA on kääntyvä, niin jokaisella $y \in \mathbb{R}^{m\times 1}$ pienimmän neliösumman ratkaisu yhtälölle $Ax = y$ on siis $\hat{x} = (A^tA)^{-1}A^ty$.*

Huomautus 12.3.3. *Pienimmän neliösumman ratkaisu $\hat{x}$ on siis yhtäpitävästi lauseen 12.2.1 geometrisessa todistuksessa esiintyneen yhtälön*

$$A^tAx = A^ty$$

ratkaisu. Kirjallisuudessa tätä yhtälöä kutsutaan yhtälön $Ax = y$ normaaliyhtälöksi. *Lauseen 12.2.1 geometrinen todistus selittää tämän alkuperän.*

Lauseen 12.3.1 todistus. Olkoon $\pi_{\mathrm{Col}(A)} \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ ortogonaalinen projektio aliavaruudelle $\mathrm{Col}(A)$. Lauseen 12.1.6 nojalla $\pi_{\mathrm{Col}(A)}(y)$ on yksikäsitteinen vektori, jolle pätee

$$\|y - \pi_{\mathrm{Col}(A)}(y)\| = \min\{\|y - Ax\| \colon x \in \mathbb{R}^{n\times 1}\}.$$

Koska $\pi_{\mathrm{Col}(A}(y) \in \mathrm{Col}(A)$ ja matriisin A sarakkeet ovat lineaarisesti riippumattomia, niin on olemasssa yksikäsitteinen sellainen vektori $\hat{x} \in \mathbb{R}^{n\times 1}$, että $A\hat{x} = \pi_{\mathrm{Col}(A)}(y)$.

Lauseen 12.2.1 perusteella

$$A\hat{x} = \pi_{\mathrm{Col}(A)}(y) = A(A^tA)^{-1}A^ty.$$

Näin ollen vektorin $\hat{x}$ yksikäsitteisyyden nojalla,

$$\hat{x} = (A^tA)^{-1}A^ty.$$

Väite on näin todistettu. □

[1] Kirjoita etäisyys $\|y - Ax\|^2$ koordinaatien avulla termin "neliösumma" selittämiseksi.

12.4 Ortogonaaliprojektio matriisin sarakeavaruudelle II

Tarkastellaan nyt ortogonaaliprojektiota $\pi_{\mathrm{Col}(A)}\colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{m\times 1}$ yleisen matriisin $A \in \mathbb{R}^{m\times n}$ tapauksessa.

Havaitaan ensin, että koska matriisin A sarakkeita ei ole oletettu lineaarisesti riippumattomiksi, niin lemman 12.2.2 tulos ei ole voimassa, eli että matriisi A^tA ei ole kääntyvä.

Esimerkki 12.4.1. *Olkoon*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Tällöin

$$A^tA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

ei ole kääntyvä.

Näin ollen ortogonaaliprojektiota $\pi_{\mathrm{Col}(A)}$ ei voi määritellä kaavalla (12.1). Projektiolle löydetään kuitenkin kaava valitsemalla matriisin A sarakkeista kanta ja siirtymällä tarkastelemaan kantavektoreiden määrittelemää matriisia. Tehdään tämä nyt tarkasti.

Olkoon $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ ja $k = \dim \mathrm{Col}(A)$. Tällöin lauseen 3.8.5 nojalla on olemassa indeksit $1 \leq i_1 < \cdots < i_k \leq n$, että $(v_{i_1}, \ldots, v_{i_k})$ on sarakeavaruuden $\mathrm{Col}(A)$ kanta. Olkoon $P = \begin{bmatrix} v_{i_1} & \cdots & v_{i_k} \end{bmatrix} \in \mathbb{R}^{m\times k}$. Tällöin $\mathrm{Col}(A) = \mathrm{Col}(P)$, joten $\pi_{\mathrm{Col}(A)} = \pi_{\mathrm{Col}(P)}$ ortogonaaliprojektion yksikäsitteisyyden nojalla. Koska matriisin P sarakkeet ovat lineaarisesti riippumattomia, niin

$$\pi_{\mathrm{Col}(A)}(y) = \pi_{\mathrm{Col}(P)}(y) = P(P^tP)^{-1}P^ty$$

jokaisella $y \in \mathbb{R}^{n\times 1}$.

Avaruuden $\mathrm{Col}(A)$ ortonormaalin kannan avulla saadaan kuitenkin vielä parempi tulos.

Lause 12.4.2. *Olkoot $A \in \mathbb{R}^{m\times n}$ matriisi, $(u_1, \ldots, u_k)$ avaruuden* $\mathrm{Col}(A)$ *ortonormaalikanta ja $Q = \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} \in \mathbb{R}^{m\times k}$. Tällöin*

$$\pi_{\mathrm{Col}(A)}(y) = QQ^ty.$$

Todistus. Koska $\pi_{\mathrm{Col}(A)} = \pi_{\mathrm{Col}(Q)}$, matriisin Q sarakkeet lineaarisesti riippumattomia ja $Q^tQ = I$, niin

$$\pi_{\mathrm{Col}(A)}(y) = \pi_{\mathrm{Col}(Q)}(y) = Q(Q^tQ)^{-1}Q^ty = QQ^ty$$

jokaisella $y \in \mathbb{R}^{n\times 1}$. □

Huomautus 12.4.3. *Matriisi Q voidaan lukea myös matriisin A QR-hajotelmasta $A = QR$, jos ortogonaalimatriisille Q pätee* $\mathrm{Col}(A) = \mathrm{Col}(Q)$. *Tälläinen QR-hajotelma syntyy esimerkiksi Gram–Schmidt ortonormeeraus prosessin tuloksena.*

12.4.1 Normaaliyhtälö yleisessä tapauksessa

Tarkastellaan vielä normaaliyhtälön muotoa tilanneessa, jossa matriisin $A \in \mathbb{R}^{m\times n}$ sarakkeet eivät ole lineaarisesti riippumattomia.

Koska matriisi A^tA ei ole kääntyvä, niin voidaan pitää hieman yllättävänä sitä seikkaa, että normaaliyhtälöllä

$$A^tAx = A^ty \tag{12.4}$$

on ratkaisu $x \in \mathbb{R}^{n\times 1}$ jokaisella $y \in \mathbb{R}^{m\times 1}$. Tämä ratkaisu ei kuitenkaan ole tässä tapauksessa yksikäsitteinen, kuten kohta huomataan.

Tämä tulos on luonnollista kirjata seuraavassa muodossa, että matriiseilla A^tA ja A^t on sama sarakeavaruus.

Lause 12.4.4. *Olkoon $A \in \mathbb{R}^{m\times n}$. Tällöin*

$$\mathrm{Col}(A^tA) = \mathrm{Col}(A^t).$$

Todistus. Osoitetaan ensin, että $\mathrm{Col}(A^tA) \subset \mathrm{Col}(A^t)$. Olkoon $v \in \mathrm{Col}(A^tA)$. Tällöin on olemassa sellainen $x \in \mathbb{R}^{n\times 1}$, että $v = A^tAx$. Olkoon $y = Ax \in \mathbb{R}^{n\times 1}$. Näin ollen $v = A^ty$ eli $v \in \mathrm{Col}(A^t)$.

Osoitetaan nyt, että $\mathrm{Col}(A^t) \subset \mathrm{Col}(A^tA)$. Olkoon $v \in \mathrm{Col}(A^t)$. Olkoon $y \in \mathbb{R}^{m\times 1}$ sellainen vektori, että $A^ty = v$ ja olkoon $x \in \mathbb{R}^{n\times 1}$ sellainen vektori, että $Ax = \pi_{\mathrm{Col}(A)}(v) \in \mathrm{Col}(A)$. Olkoot myös vektorit $v_1, \ldots, v_n \in \mathbb{R}^{m\times 1}$ matriisin sarakkeet eli $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$.

Koska vektori $y - \pi_{\mathrm{Col}(A)}(y)$ on kohtisuorassa aliavaruutta $\mathrm{Col}(A)$ vastaan, niin se on kohtisuorassa matriisin A sarakkeita vastaan. Näin ollen

$$(y - \pi_{\mathrm{Col}(A)}(y)) \cdot v_i = 0.$$

Koska $v_i = Ae_i$, niin

$$A^t(y - \pi_{\mathrm{Col}(A)}(y)) \cdot e_i = 0$$

jokaisella $i \in \{1, \ldots, n\}$. Näin ollen $A^t(y - \pi_{\mathrm{Col}(A)}(y)) = 0$ ja

$$v = A^ty = A^t(y - \pi_{\mathrm{Col}(A)}(y)) + A^t\pi_{\mathrm{Col}(A)}(y) = A^t\pi_{\mathrm{Col}(A)}(y) = A^tAx.$$

Tämä päättää todistuksen. □

Kaksi kommenttia tästä todistuksesta on paikallaan.

Huomautus 12.4.5. *Ensinnäkin todistus on oleellisesti sama kuin lauseen 12.2.1 geometrinen todistus. Ainoa ero on näkökulmassa: Lauseen 12.2.1 todistuksessa haetaan projektiolle kaavaa ja tässä todistuksessa halutaan osoittaa, että vektori $v - \pi_{\mathrm{Col}(A)}(v)$ on matriisin A^t nolla-avaruudessa.*

Kirjataan nyt normaaliyhtälöä koskeva tulos seurauksena.

Korollaari 12.4.6. *Olkoon $A \in \mathbb{R}^{m\times n}$. Tällöin jokaisella vektorilla $y \in \mathbb{R}^{m\times 1}$ on olemassa sellainen vektori $x \in \mathbb{R}^{n\times 1}$, että*

$$A^tAx = A^ty.$$

Todistus. Koska $A^ty \in \mathrm{Col}(A^t)$ ja $\mathrm{Col}(A^tA) = \mathrm{Col}(A^t)$, niin on olemassa sellainen $x \in \mathbb{R}^{n\times 1}$, että $A^ty = A^tAx$. □

12.4.2 Pienimmän neliösumman ongelma yleisessä tapauksessa

Tarkastellaan vielä pienimmän neliösumman ongelmaa kolmesta eri näkökulmasta yleisen matriisin tapauksessa. Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi. Olkoon nyt $(u_1, \ldots, u_k)$ aliavaruuden $\mathrm{Col}(A)$ ortonormaali kanta. Koska matriisin A sarakkeita ei oletettu lineaarisesti riippumattomiksi, niin $k \leq n$.

Olkoon $Q = \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} \in \mathbb{R}^{n\times k}$. Tällöin $\mathrm{Col}(A) = \mathrm{Col}(Q)$ ja $\pi_{\mathrm{Col}(A)} = \pi_{\mathrm{Col}(Q)}$. Näin ollen vektori $\hat{x} \in \mathbb{R}^{n\times 1}$ on pienimmän neliösumman ongelman

$$\|y - A\hat{x}\| = \min\{\|y - Ax\| \colon x \in \mathbb{R}^{n\times 1}\}$$

ratkaisu, jos ja vain jos

$$A\hat{x} = \pi_{\mathrm{Col}(A)}(y) = \pi_{\mathrm{Col}(Q)}(y) = QQ^t y.$$

Jos matriisin A sarakkeet eivät ole lineaarisesti riippumattomia, niin tällä yhtälöllä on äärettömästi ratkaisuja. Tarkemmin sanottuna, jos $\hat{x}$ on yksi yhtälön ratkaisu, niin kaikki joukon

$$\hat{x} + \mathrm{Null}(A) = \{\hat{x} + v \in \mathbb{R}^{n\times 1} \colon v \in \mathrm{Null}(A)\}$$

ovat yhtälön ratkaisuja.

Pienimmän neliösumman ongelma voidaan aina ratkaista normaaliyhtälön avulla. Huomaa, että normaaliyhtälöllä on aina ratkaisu lauseen 12.4.4 perusteella.

Lause 12.4.7. *Olkoon $A \in \mathbb{R}^{m\times n}$ ja $y \in \mathbb{R}^{m\times 1}$. Tällöin vektori $\hat{x} \in \mathbb{R}^{n\times 1}$ toteuttaa normaaliyhtälön*

$$A^t Ax = A^t y,$$

jos ja vain jos $\hat{x}$ on pienimmän neliösumman ongelman

$$\|y - A\hat{x}\| = \min\{\|y - Ax\| \colon x \in \mathbb{R}^{n\times 1}\}$$

ratkaisu.

Todistus. Oletetaan, ensin $A^t A\hat{x} = A^t y$. Osoitetaan, että $\hat{x}$ on pienimmän neliösumman ongelman ratkaisu. Riittää osoittaa, että erotusvektori $y - A\hat{x}$ on kohtisuorassa aliavaruutta $\mathrm{Col}(A)$ vastaan. Koska matriisin $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ sarakkeet sisältävät aliavaruuden $\mathrm{Col}(A)$ kannan, niin riittää osoittaa, että $v_i \cdot (y - A\hat{x}) = 0$ jokaisella $i \in \{1, \ldots, n\}$ eli että $A^t(y - A\hat{x}) = 0$. Koska $A^t A\hat{x} = A^t y$, niin

$$A^t(y - A\hat{x}) = A^t y - A^t A\hat{x} = A^t y - A^t y = 0.$$

Näin ollen $\hat{x}$ on pienimmän neliösumman ongelman ratkaisu.

Oletetaan nyt, että $\hat{x} \in \mathbb{R}^{n\times 1}$ on pienimmän neliösumman ongelman ratkaisu. Osoitetaan, että $A^t Ax = A^t y$. Koska $y - A\hat{x}$ on kohtisuorassa aliavaruutta $\mathrm{Col}(A)$ vastaan, niin $A^t(y - A\hat{x}) = 0$. Näin ollen $A^t A\hat{x} = A^t y$. □

Esimerkki 12.4.8. *Olkoot*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 2} \text{ ja } y = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^{3\times 1}.$$

Tällöin $A = \begin{bmatrix} e_1 & e_1 \end{bmatrix}$ *ja* $\mathrm{Col}(A) = \mathrm{Sp}(e_1)$. *Näin ollen* $\pi_{\mathrm{Col}(A)}(y) = ae_1$. *Koska*

$$A^t A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ ja } A^t y = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ a \end{bmatrix}$$

niin normaaliyhtälön $A^t Ax = A^t y$ *ratkaisut ovat*

$$x = t \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} a \\ 0 \end{bmatrix},$$

missä $t \in \mathbb{R}$. *Huomaa, että*

$$\mathrm{Null}(A) = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} \right).$$

Käyttäen matriisin A QR-hajotelmaa voidaan pienimmän neliösumman ongelma ratkaista seuraavassa muodossa.

Lause 12.4.9. *Olkoon* $A \in \mathbb{R}^{m\times n}$ *matriisi ja olkoon* $A = QR$ *matriisin* A *QR-hajotelma, jossa* $Q \in \mathbb{R}^{m\times k}$ *on sellainen ortogonaalimatriisi, että* $\mathrm{Col}(A) = \mathrm{Col}(Q)$, *ja* $R \in \mathbb{R}^{k\times n}$ *on yläkolmiomatriisi. Olkoon* $y \in \mathbb{R}^{m\times 1}$. *Tällöin vektorille* $\hat{x} \in \mathbb{R}^{n\times 1}$ *pätee*

$$\|y - A\hat{x}\| = \min\{\|y - Ax\| \colon x \in \mathbb{R}^{n\times 1}\}, \tag{12.5}$$

jos ja vain jos

$$R\hat{x} = Q^t y.$$

Todistus. Oletetaan, että $\hat{x} \in \mathbb{R}^{n\times 1}$ on pienimmän neliösumman ongelman (12.5). Tällöin $A\hat{x} = \pi_{\mathrm{Col}(A)}$. Koska $\mathrm{Col}(A) = \mathrm{Col}(Q)$, niin $\pi_{\mathrm{Col}(A)} = \pi_{\mathrm{Col}(Q)}$. Näin ollen $A\hat{x} = QQ^t y$. Koska $A = QR$, niin

$$R\hat{x} = Q^t QR\hat{x} = Q^t A\hat{x} = Q^t(QQ^t y) = (Q^t Q)Q^t y = Q^t y.$$

Toisaalta, jos $R\hat{x} = Q^t y$, niin

$$A\hat{x} = QR\hat{x} = QQ^t y = \pi_{\mathrm{Col}(A)}(y).$$

Näin ollen $\hat{x}$ on pienimmän neliösumman ongelman ratkaisu. □

12.5 Lineaarikuvauksen adjungaatti

Lause 12.4.4 on syvällisempi kuin ensi näkemältä uskoisi. Siinä kiteytyy ortogonaaliprojektion $\pi_{\mathrm{Col}(A)} \colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{m\times 1}$ suhde matriisin A ja A^t määräämiin kuvauksiin $f_A \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ ja $f_{A^t} \colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{n\times 1}$.

Lineaarikuvausten f_A ja f_{A^t} avulla kirjoitettuna yhtälö $\mathrm{Col}(A^tA) = \mathrm{Col}(A^t)$ voidaan tulkita muodossa $\mathrm{im}(f_{A^t} \circ f_A) = \mathrm{im}\, f_{A^t}$. Tästä voidaan päätellä, että $\ker f_{A^t} \cap \mathrm{im}\, f_A = \{0\}$. Aliavaruudet $\ker f_{A^t}$ ja $\mathrm{im}\, f_A$ ovat itseasiassa toisiaan vastaan kohtisuorassa, sillä lauseen 12.4.4 todistuksesta voidaan puolestaan pienellä lisätyölllä päätellä, että $\ker f_{A^t} = \ker \pi_{\mathrm{Col}(A)}$.

Nämä kaikki havainnot kuvauksista f_A ja f_{A^t} liittyvät yleiseen teoriaan, että f_{A^t} on kuvauksen f_A adjungaatti. Ortogonaaliprojektiot puolestaan ovat itseadjungoituja lineaarikuvauksia.

12.5.1 Lineaarikuvauksen adjungaatti ja matriisin transpoosi

Määritelmä 12.5.1. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ sisätuloavaruuksia. Lineaarikuvaus $f^* \colon W \to V$ on lineaarikuvauksen $f \colon V \to W$* adjungaatti*, jos kaikilla $v \in V$ ja $w \in W$ pätee*

$$\langle f(v), w\rangle_W = \langle v, f^*(w)\rangle_V.$$

Esimerkki 12.5.2. *Lineaarikuvauksen $f_A \colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ adjungaatti on lineaarikuvaus $f_{A^t} \colon \mathbb{R}^{m\times 1} \to \mathbb{R}^{n\times 1}$, kun avaruuksissa $\mathbb{R}^{n\times 1}$ ja $\mathbb{R}^{m\times 1}$ sisätulona on pistetulo, sillä kaikilla $x \in \mathbb{R}^{n\times 1}$ ja $y \in \mathbb{R}^{m\times 1}$ pätee*

$$f_A(x) \cdot y = (Ax) \cdot y = (Ax)^t y = x^t A^t y = x^t(A^t y) = x \cdot (A^t y) = x \cdot f_{A^t}(y).$$

Adjungaatin $f^* \colon W \to V$ perusominaisuus on, että sen ydin $\ker f^*$ on kohtisuorassa alkuperäisen kuvauksen $f \colon V \to W$ kuvaa $\mathrm{im}\, f$ vastaan.

Lemma 12.5.3. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ sisätuloavaruuksia sekä olkoon $f \colon V \to W$ lineaarikuvaus ja $f^* \colon W \to V$ sen adjungaatti. Tällöin*

$$\ker f^* = (\mathrm{im}\, f)^\perp.$$

Todistus. Olkoon $w \in \ker f^*$. Tällöin jokaisella $v \in V$ pätee

$$\langle w, f(v)\rangle = \langle f^*(w), v\rangle = 0.$$

Näin ollen w on kohtisuorassa aliavaruutta $\mathrm{im}\, f$ vastaan eli $\ker f^* \subset (\mathrm{im}\, f^*)^\perp$.

Oletetaan nyt, että $w \in (\mathrm{im}\, f)^\perp$. Tällöin

$$\langle f^*(w), f^*(w)\rangle = \langle f(f^*(w)), w\rangle = 0.$$

Näin ollen $f^*(w) = 0$ eli $w \in \ker f^*$. □

Äärellisulotteisten sisätuloavaruuksien välisen lineaarikuvauksen adjungaatti saadaan esitysmatriisin transpoosista, kunhan esitysmatriisi on muodostettu ortnormaaleja kantoja käyttäen. Tämä osoittaa samalla, että lineaarikuvauksen adjungaatti on yksikäsitteinen.

Lause 12.5.4. *Olkoot $(V, \langle \cdot, \cdot \rangle_V)$ ja $(W, \langle \cdot, \cdot \rangle_W)$ äärellisulotteisia sisätuloavaruuksia. Tällöin jokaisella lineaarikuvauksella $f\colon V \to W$ on yksikäsitteinen adjungaatti $f^*\colon W \to V$. Lisäksi, jos A on kuvauksen f esitysmatriisi ortonormaalista kannasta $(v_1, \ldots, v_n)$ ortonormaaliin kantaan $(w_1, \ldots, w_m)$, niin matriisi A^t on kuvauksen f^* esitysmatriisi ortonormaalista kannasta $(w_1, \ldots, w_m)$ ortonormaaliin kantaan $(v_1, \ldots, v_n)$.*

Lauseen väiteen toinen osa voidaan kiteyttää kaavioiden avulla seuraavasti

jos
$$\begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{R}^{n\times 1} \arrow[u, "\Phi_{(v_1,\ldots,v_n)}"] \arrow[r, "f_A"] & \mathbb{R}^{m\times 1} \arrow[u, "\Phi_{(w_1,\ldots,w_n)}"'] \end{tikzcd}$$
niin
$$\begin{tikzcd} V & W \arrow[l, "f^*"'] \\ \mathbb{R}^{n\times 1} \arrow[u, "\Phi_{(v_1,\ldots,v_n)}"] & \mathbb{R}^{m\times 1} \arrow[u, "\Phi_{(w_1,\ldots,w_n)}"'] \arrow[l, "f_{A^t}"'] \end{tikzcd}$$

kun kannat $(v_1, \ldots, v_n)$ ja $(w_1, \ldots, w_m)$ ovat ortonormaaleja. Koska kuvaus f^* voidaan määritellä kaavion avulla, niin riittää osoittaa, että näin saatu lineaarikuvaus on kuvauksen f yksikäsitteinen adjungaatti.

Lauseen 12.5.4 todistus. Asetetaan kuvaus $f^*\colon W \to V$ kaavalla

$$f^* = \Phi_{(v_1,\ldots,v_n)} \circ f_{A^t} \circ \Phi^{-1}_{(w_1,\ldots,w_m)}$$

ja osoitetaan, että

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle$$

kaikilla $v \in V$ ja $w \in W$.

Olkoon $v \in V$ ja $w \in W$. Olkoot myös $x = \Phi^{-1}_{(v_1,\ldots,v_n)}(v) \in \mathbb{R}^{n\times 1}$ ja $y = \Phi^{-1}_{(w_1,\ldots,w_m)}(w) \in \mathbb{R}^{m\times 1}$. Tällöin

$$\begin{aligned} \langle v, f^*(w) \rangle_V &= \langle v, (\Phi_{(v_1,\ldots,v_n)} \circ f_{A^t} \circ \Phi^{-1}_{(w_1,\ldots,w_m)})(w) \rangle_V \\ &= \langle \Phi_{(v_1,\ldots,v_n)}(x), \Phi_{(v_1,\ldots,v_n)}(f_{A^t}(y)) \rangle_V = x \cdot f_{A^t}(y). \end{aligned}$$

Koska

$$\Phi_{(w_1,\ldots,w_m)} \circ f_A = f \circ \Phi_{(v_1,\ldots,v_n)},$$

niin

$$\begin{aligned} \langle f(v), w \rangle_W &= \langle f(\Phi_{(v_1,\ldots,v_n)}(x)), \Phi_{(w_1,\ldots,w_m)}(y) \rangle_W \\ &= \langle (f \circ \Phi_{(v_1,\ldots,v_n)})(x)), \Phi_{(w_1,\ldots,w_m)}(y) \rangle_W \\ &= \langle (\Phi_{(w_1,\ldots,w_m)} \circ f_A)(x)), \Phi_{(w_1,\ldots,w_m)}(y) \rangle_W \\ &= f_A(x) \cdot y = x \cdot f_{A^t}(y) = \langle v, f^*(w) \rangle_V. \end{aligned}$$

Osoitetaan nyt, että f^* on kuvauksen $f\colon V \to W$ ainoa adjungaatti. Olkoon $g\colon W \to V$ sellainen lineaarikuvaus, että

$$\langle f(v), w\rangle_W = \langle v, g(w)\rangle_V$$

kaikilla $v \in V$ ja $w \in W$. Tällöin jokaisella $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, n\}$ pätee

$$\langle v_i, g(w_j)\rangle_V = \langle f(v_i), w_j\rangle_W = \langle v_i, f^*(w_j)\rangle_V.$$

Näin ollen jokaisella $j \in \{1, \ldots, m\}$ vektoreilla $g(w_j)$ ja $f^*(w_j)$ on samat koordinaatit kannassa $(v_1, \ldots, v_n)$, joten $g(w_j) = f^*(w_j)$ jokaisella $j \in \{1, \ldots, m\}$. Kantalauseen nojalla $g = f^*$. □

12.5.2 Itseadjungoidut lineaarikuvaukset ja symmetriset matriisit

Lineearikuvausta sanotaan itseadjungoiduksi, jos se on oma adjungaattinsa.

Määritelmä 12.5.5. *Lineaarikuvaus $f\colon V \to V$ sisätuloavaruudesta $(V, \langle\cdot,\cdot\rangle)$ itseensä on* itseadjungoitu, *jos $f^* = f$.*

Itsedjungoituja kuvauksia on runsaasti, sillä lineaarikuvauksen ja sen adjungaatin yhdistetty kuvaus on itseadjungoitu.

Lemma 12.5.6. *Olkoot $(V, \lambda\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ sisätuloavaruuksia sekä olkoot $f\colon V \to W$ lineaarikuvaus ja $f^*\colon W \to V$ sen adjungaatti. Tällöin $f^* \circ f\colon V \to V$ on itseadjungoitu.*

Todistus. Olkoot $v, v' \in V$. Tällöin

$$\langle (f^*\circ f)(v), v'\rangle_V = \langle f^*(f(v)), v'\rangle_V = \langle f(v), f(v')\rangle_W = \langle v, f^*(f(v'))\rangle_V = \langle v, (f^*\circ f)(v')\rangle_V.$$

Näin ollen $(f^* \circ f)^* = f^* \circ f$. □

Korollaari 12.5.7. *Olkoon $A \in \mathbb{R}^{m\times n}$. Tällöin kuvaus $f_{A^tA}\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ on itseadjungoitu.*

Todistus. Väite seuraa suoraan havainnoista $(f_A)^* = f_{A^t}$ ja $f_{A^tA} = f_{A^t} \circ f_A$. □

Edellisen korollaarin pohjalta voi tulla johtopäätökseen, että itseadjungoidut lineaarikuvaukset vastaavat symmetrisiä matriiseja.

Määritelmä 12.5.8. *Matriisi $A \in \mathbb{R}^{n\times n}$ on symmetrinen, jos $A^t = A$.*

Tämä tulos on muotoiltu tarkasti seuraavassa lauseessa. Koska tulos seuraa lähes suoraan lauseesta 12.5.4, se jätetään harjoitustehtäväksi.

Lause 12.5.9. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus. Lineaarikuvaus $f\colon V \to V$ on itseadjungoitu, jos ja vain jos jokaisella avaruuden V ortonormaalilla kannalla $(v_1, \ldots, v_n)$ kuvauksen f esitysmatriisi A kannasta $(v_1, \ldots, v_n)$ kantaan $(v_1, \ldots, v_n)$ on symmetrinen matriisi.*

Lemmasta 12.5.3 seuraa suoraan, että itseadjungoidun operaattorin kuva on kohtisuorassa ydintä vastaan. Kirjataan tämä tärkeä tulos lauseeksi ja sen muotoilu matriiseille korollaariksi.

Lause 12.5.10. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus ja $f\colon V \to V$ itseadjungoitu lineaarikuvaus. Tällöin* $\ker f = (\operatorname{im} f)^{\perp}$*.*

Todistus. Koska $f^* = f$, niin lemman 12.5.3 nojalla

$$\ker f = \ker f^* = (\operatorname{im} f)^{\perp}.$$

□

Korollaari 12.5.11. *Olkoon $A \in \mathbb{R}^{n\times n}$ symmetrinen matriisi. Tällöin* $\operatorname{Null}(A) = (\operatorname{Col}(A))^{\perp}$*. Erityisesti*

$$\mathbb{R}^{n\times 1} = \operatorname{Col}(A) \oplus \operatorname{Null}(A).$$

12.5.3 Ortogonaaliprojektiot ja itseadjungoidut kuvaukset

Lauseen 12.5.10 avulla havaitaan, että itseadjungoidut projektiot vastaavat täsmälleen ortogonaaliprojektioita.

Lause 12.5.12. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus ja kuvaus $\pi\colon V \to V$ projektio. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *π on ortogonaaliprojektio,*
2. *π on itseadjungoitu ja*
3. $\ker \pi = (\operatorname{im} \pi)^{\perp}$.

Todistus. Oletetaan, että π on ortogonaaliprojektio. Olkoot $v, v' \in V$. Tällöin

$$\langle \pi(v), v'\rangle = \langle \pi(v), v' - \pi(v')\rangle + \langle \pi(v), \pi(v')\rangle = \langle \pi(v), \pi(v')\rangle.$$

Vastaavasti $\langle v, \pi(v')\rangle = \langle \pi(v), \pi(v')\rangle$. Näin ollen

$$\langle \pi(v), v'\rangle = \langle v, \pi(v')\rangle$$

kaikilla $v, v' \in V$, eli $\pi^* = \pi$. Tämä todistaa ensimmäisen implikaation (1) $\Rightarrow$ (2).

Jos π on itseadjungoitu, niin tällöin lauseen 12.5.10 nojalla $\ker \pi = (\operatorname{im} \pi)^{\perp}$. Tämä todistaa seuraavan implikaation (2) $\Rightarrow$ (3).

Oletetaan nyt, että $\ker \pi = (\operatorname{im} \pi)^{\perp}$ ja osoitetaan, että π on ortogonaaliprojektio. Olkoon $v \in V$ ja merkitään $w = v - \pi(v)$. Koska π on projektio, niin $\pi(w) = \pi(v - \pi(v)) = \pi(v) - \pi(\pi(v)) = \pi(v) - \pi(v) = 0$ eli $w \in \ker \pi$. Koska $\ker \pi = (\operatorname{im} \pi)^{\perp}$ oletuksen nojalla, niin $\langle v - \pi(v), \pi(v)\rangle = 0$. Kuvaus π on siis ortogonaaliprojektio. Tämä päättää todistuksen. □

Yhdistämällä adjungaatin $f^*\colon W \to V$ perusominaisuuden, että sen ydin on kohtisuorassa kuvauksen $f\colon V \to W$ kuvaa vasten, ja ortogonaaliprojektion ytimen kohtisuoruuden omaa kuvaansa vasten saadaan seuraava mielenkiintoinen korollaari.

Korollaari 12.5.13. *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *äärellisulotteisia sisätuloavaruuksia ja* $f\colon V \to W$ *lineaarikuvaus. Tällöin*

$$\ker f^* = \ker \pi_{\operatorname{im} f}.$$

Todistus. Väite seuraa suoraan lauseesta 12.5.12 ja lemmasta 12.5.3, sillä

$$\ker \pi_{\operatorname{im} f} = (\operatorname{im} \pi_{\operatorname{im} f})^{\perp} = (\operatorname{im} f)^{\perp} = \ker f^*.$$

□

12.5.4 Yleisen teorian seuraukset

Lemman 12.5.3 suorana seurauksena saadaan tulos, että lineaarikuvauksen $f\colon V \to W$ ja sen adjungaatin $f^*\colon W \to V$ kuvien dimensio on sama, jos W on äärellisulotteinen.

Lause 12.5.14. *Olkoon* $f\colon V \to W$ *lineaarikuvaus sisätuloavaruudesta* $(V, \langle\cdot,\cdot\rangle_V)$ *äärellisulotteiseen sisätuloavaruuteen* $(W, \langle\cdot,\cdot\rangle_W)$*. Tällöin*

$$\dim \operatorname{im} f^* = \dim \operatorname{im} f.$$

Todistus.

$$\dim \operatorname{im} f^* = \dim W - \dim \ker f^* = \dim W - \dim(\operatorname{im} f)^{\perp} = \dim \operatorname{im} f.$$

□

Tällä abstraktilla lauseella on konkreettinen korollaari.

Korollaari 12.5.15. *Olkoon* $A \in \mathbb{R}^{m\times n}$ *matriisi. Tällöin*

$$\dim \operatorname{Col}(A^t) = \dim \operatorname{Col}(A).$$

Erityisesti

$$\dim \operatorname{Row}(A) = \dim \operatorname{Col}(A).$$

Huomautus 12.5.16. *Tämän korollaarin todistusta on pantattu aina luvusta 3 asti. Jos matriisin* A *riviavaruus* $\operatorname{Row}(A)$ *olisi määritelty luvussa 3, niin tämä tulos olisi voitu todistaa myös havainnosta, että sidottuja muuttujia vastaavat rivit antavat riviavaruuden* $\operatorname{Row}(A)$ *kannan. On kuitenkin mielenkiintoista havaita, että nyt tämä algebrallinen tulos voitiin todistaa geometrisin keinoin.*

Lause 12.5.14 seuraa myös seuraavasta yleisemmästä tuloksesta, joka perustuu korollaariin 12.5.13. Tämä tulos on lauseen 12.4.4 vastine lineaarikuvauksille. Huomaa, että seuraavassa tuloksessa kuvauksen $f\colon V \to W$ lähtöavaruuden V ei tarvitse olla äärellisulotteinen eli kuvauksella f ei tarvitse olla esitysmatriisia.

Lause 12.5.17. *Olkoon $f \colon V \to W$ lineaarikuvaus sisätuloavaruudesta $(V, \langle\cdot,\cdot\rangle_V)$ äärellisulotteiseen sisätuloavaruuteen $(W, \langle\cdot,\cdot\rangle_W)$. Tällöin*

$$\operatorname{im}(f^* \circ f) = \operatorname{im} f^*.$$

Todistus. Selvästi $\operatorname{im}(f^* \circ f) \subset \operatorname{im} f^*$. Osoitetaan, että $\operatorname{im} f^* \subset \operatorname{im}(f^* \circ f)$.

Olkoot $v \in \operatorname{im} f^*$ ja $w \in W$ sellainen, että $f^*(w) = v$. Koska $\pi_{\operatorname{im} f}(w - \pi_{\operatorname{im} f}(w)) = \pi_{\operatorname{im} f}(w) - \pi_{\operatorname{im} f}(\pi_{\operatorname{im} f}(w)) = \pi_{\operatorname{im} f}(w) - \pi_{\operatorname{im} f}(w) = 0$, niin $w - \pi_{\operatorname{im} f}(w) \in \ker \pi_{\operatorname{im} f}$. Koska korollaarin 12.5.13 perusteella $\ker f^* = \ker \pi_{\operatorname{im} f}$, niin $f^*(w - \pi_{\operatorname{im} f}(w)) = 0$. Näin ollen

$$v = f^*(w) = f^*(w - \pi_{\operatorname{im} f}(w)) + f^*(\pi_{\operatorname{im} f}(w)) = f^*(\pi_{\operatorname{im} f}(w)).$$

Koska $\pi_{\operatorname{im} f}(w) \in \operatorname{im} f^*$, niin $v \in \operatorname{im}(f^* \circ f)$. $\square$

12.5.5 Ortogonaaliprojektion kaava adjungaateilla

Annetaan nyt kolmas todistus lauseelle 12.2.1. Aloitetaan lemman 12.2.2 vastineella.

Lemma 12.5.18. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ sisätuloavaruuksia sekä $f \colon V \to W$ lineaarikuvaus ja $f^* \colon W \to V$ sen adjungaatti. Tällöin* $\ker(f^* \circ f) = \ker f$. *Erityisesti $f^* \circ f \colon V \to V$ on isomorfismi, jos f on injektio ja V on äärellisulotteinen.*

Todistus. Selvästi $\ker f \subset \ker(f^* \circ f)$. Olkoon nyt $v \in \ker(f^* \circ f)$. Tällöin

$$0 = \langle (f^* \circ f)(v), v\rangle_V = \langle f(v), f(v)\rangle_W.$$

Näin ollen $f(v) = 0$ eli $v \in \ker f$.

Jos V on äärellisulotteinen ja f on injektio, niin tällöin myös $f^* \circ f$ on injektio avaruudesta V itseensä. Näin ollen se on myös surjektio kuvausten dimensiolauseen nojalla ja siten isomorfismi. $\square$

Lause 12.5.19. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ äärellisulotteisia sisätuloavaruuksia sekä $f \colon V \to W$ injektiivinen lineaarikuvaus. Tällöin ortogonaaliprojektio $\pi_{\operatorname{im} f} \colon W \to W$ aliavaruudelle* $\operatorname{im} f$ *on yhdistetty kuvaus*

$$\pi_{\operatorname{im} f} = f \circ (f^* \circ f)^{-1} \circ f^*.$$

Todistus. Olkoon $\pi \colon W \to W$ kuvaus $\pi = f \circ (f^* \circ f)^{-1} \circ f^*$. Riittää osoittaa, että π on projektio ja että $\ker \pi = (\operatorname{im} \pi)^\perp$.

Koska

$$\begin{aligned}\pi \circ \pi &= \big(f \circ (f^* \circ f)^{-1} \circ f^*\big) \circ \big(f \circ (f^* \circ f)^{-1} \circ f^*\big)\\ &= f \circ (f^* \circ f)^{-1} \circ (f^* \circ f) \circ (f^* \circ f)^{-1} \circ f^*\\ &= f \circ (f^* \circ f)^{-1} \circ f^* = \pi,\end{aligned}$$

niin π on projektio.

Osoitetaan nyt, että π on itseadjungoitu. Olkoot $w, w' \in W$. Tällöin

$$\begin{aligned}
\langle \pi(w), w' \rangle_W &= \langle \left(f \circ (f^* \circ f)^{-1} \circ f^* \right)(w), w' \rangle_W = \langle \left((f^* \circ f)^{-1} \circ f^* \right)(w), f^*(w') \rangle_W \\
&= \langle f^*(w), ((f^* \circ f)^{-1} \circ f^*)(w') \rangle_W = \langle w, (f \circ (f^* \circ f)^{-1} \circ f^*)(w') \rangle_W \\
&= \langle w, \pi(w') \rangle_W .
\end{aligned}$$

Näin ollen π on itseadjungoitu.

Lauseen 12.5.12 nojalla operaattori $\pi \colon W \to W$ on on siis ortogonaaliprojektio aliavaruudelle $\operatorname{im} f$ ja siten $\pi = \pi_W$ ortogonaaliprojektion yksikäsitteisyyden nojalla. $\square$

Osa III

Lineaarialgebra ja matriisilaskenta III

Luku 13

Kompleksiset vektoriavaruudet

Aiemmissa luvuissa kaikki vektoriavaruudet ovat olleet ns. reaalisia vektoriavaruuskia eli avaruuden V skalaarikertolasku on ollut $\cdot\colon \mathbb{R} \times V \to V$. Tässä luvussa määritellään kompleksiset vektoriavaruudet. Ainoa muutos on, että kompleksiselle vektoriavaruudelle V skalaarikertolasku on muotoa $\cdot\colon \mathbb{C} \times V \to V$. Vaikka kompleksiset vektoriavaruudet eivät liity suoraan moniin lineaarialgebran sovelluksiin, on niillä tärkeä rooli lineaario-peraattoreiden teoriassa, kuten kohta tullaan huomaamaan.

Luku alkaa palauttamalla mieleen kompleksiluvut, mutta sitä ennen käsitellään esimerkkiä, joka osoittaa, että kompleksilukuja ei voida välttää, vaikka tarkasteltaisiin ainostaan reaalisia matriiseja.

Huomautus 13.0.1. *Tässä luvussa keskitytään pääosin esittelemään aiemmista luvuista tuttuja määritelmiä ja tuloksia kompleksisten vektoriavaruuksien tapauksessa. Tulosten todistukset ovat analogisia aiemmin todistetuille tuloksille ja todistukset tämän vuoksi sivuutetaan. Joidenkin tulosten kohdalla tuloksesta annetaan hieman yleisempi versio. Tällöin todistus jätetään kiinnostuneelle lukijalle harjoitustehtäväksi.*

13.1 Motivointi: kompleksiset ominaisarvot

Tarkastellaan kiertomatriisia

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

missä $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$, eli $\theta \neq k\pi$ kaikilla $k \in \mathbb{Z}$.

Matriisia A vastaava lineaarikuvaus $f_A\colon \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}$, $x \mapsto Ax$, on perusesimerkki lineaarikuvauksesta, jolla ei ole ominaisvektoreita. Tämän voi jopa havaita konkreettisesti havaitsemalla, että lineaarikuvaus f_A kiertää tasoa $\mathbb{R}^{2\times 1}$ eikä näin ollen pidä yhtään suoraa paikallaan. Tarkempi perustelu tälle faktalle on, että jos kuvauksella f_A olisi ominaisvektori $x \in \mathbb{R}^{2\times 1}$, niin sillä olisi myös ominaisarvo $\lambda \in \mathbb{R}$. Palautetaan ensin mieleen, miksi ominaisarvoa ei ole.

Merkintöjen helpottamiseksi, merkitään $a = \cos\theta$ ja $b = \sin\theta$. Tarkasteltava matriisi on siis

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

missä $a^2 + b^2 = 1$. Lisäksi $b \neq 0$, koska $\theta \neq k\pi$.

Jos nollasta poikkeava vektori $x \in \mathbb{R}^{2\times 1}$ on lineaarikuvauksen f_A ominaisvektori ominaisarvolla $\lambda \in \mathbb{R}$, niin $f_A(x) = \lambda x$, eli $Ax = \lambda x$.

Nollasta poikkeava vektori x toteuttaa siis yhtälön

$$(A - \lambda I)\, x = 0$$

eli matriisi $A - \lambda I$ ei ole kääntyvä. Tämä tarkoittaa, että

$$\det(A - \lambda I) = 0.$$

Koska

$$\det(A - \lambda I) = \det \begin{bmatrix} a-\lambda & -b \\ b & a-\lambda \end{bmatrix} = (a-\lambda)^2 - (-b^2) = (a-\lambda)^2 + b^2,$$

niin saadaan yhtälö

$$(a-\lambda)^2 = -b^2$$

ominaisarvolle $\lambda \in \mathbb{R}$.

Koska $b \neq 0$, niin $b^2 > 0$. Näin ollen

$$(a-\lambda)^2 < 0,$$

joka on ristiriita, koska λ oletettiin reaaliluvuksi. (Harjoitustehtävä: Millaista matriisia tarkastellaan, jos $b = 0$? Entä jos $a = 0$?)

13.1.1 Kiertomatriisin kompleksiset ominaisarvot

Tarkastellaan nyt yhtälöä

$$(a-\lambda)^2 = -b^2 \tag{13.1}$$

ilman tulkintaa, että yhtälöllä on yhteys ominaisarvoihin. Jos $b = 0$, niin yhtälön ratkaisu on $\lambda = a$. Oletetaan nyt, että $b \neq 0$.

Käytetään nyt tietoa, että kompleksilukujen joukossa yhtälöllä $z^2 = -b^2$ on kaksi ratkaisua

$$z = ib$$

ja

$$z = -ib.$$

Näin ollen jos sallitaan, että luku λ on kompleksinen, saadaan yhtälöt

$$a - \lambda = ib \quad \text{ja} \quad a - \lambda = -ib$$

eli

$$\lambda = a - ib \quad \text{ja} \quad \lambda = a + ib.$$

Näin saadaan, että alkuperäisellä yhtälöllä (13.1) on kaksi kompleksista ratkaisua.

Huomautus 13.1.1. *Huomaa, että ratkaisut $a-ib$ ja $a+ib$ ovat toistensa kompleksikonjugaatteja. Tämä ei ole sattumaa. Jatkossa osoitetaan, että reaalikertoimisen matriisin aidosti kompleksiset ominaisarvot esiintyvät pareittain.*

13.1.2 Kiertomatriisin ominaisvektorit?

Edellisen luvun havaintojen perusteella on selvää, että tapauksessa $b \neq 0$ luvut $\lambda = a+ib$ ja $\lambda = a - ib$ ratkaisevat yhtälön $\det(A - \lambda I) = 0$, mutta eivät kuteinkaan ole matriisin A ominaisarvoja, koska ne eivät ole reaalilukuja. Lisäksi syy tähän on se, että matriisilla A ei ole näitä lukuja vastaavia ominaisvektoreita sarakeavaruudessa $\mathbb{R}^{2\times 1}$.

Samastetaan nyt sarakeavaruus $\mathbb{R}^{2\times 1}$ avaruuden $\mathbb{R}^2$ kanssa ja tarkastellaan kuvausta $f_A \colon \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}$ lineaarikuvauksena tasosta $\mathbb{R}^2$ itseensä eli tarkastellaan kuvausta $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ kuvausta $(x, y) \mapsto (ax - by, bx + ay)$.

Vaihdetaan nyt avaruuden $\mathbb{R}^2$ tilalle avaruus $\mathbb{C}^2$ laskutoimituksilla

$$(z_1, w_1) + (z_2, w_2) = (z_1 + z_2, w_1 + w_2)$$

ja

$$\alpha(z, w) = (\alpha z, \alpha w)$$

kaikilla $(z_1, w_1), (z_2, w_2), (z, w) \in \mathbb{C}^2$ ja $\alpha \in \mathbb{C}$, varustettuna ja tarkastellaan kuvauksen f sijasta kuvausta

$$\varphi \colon \mathbb{C}^2 \to \mathbb{C}^2, \quad (z, w) \mapsto (az - bw, bz + aw).$$

Kiinnitetään nyt

$$\lambda = a + ib$$

ja tarkastellaan onko yhtälöllä

$$\varphi_A((z, w)) = \lambda(z, w) \tag{13.2}$$

eli yhtälöllä

$$(az - bw, bz + aw) = (\lambda z, \lambda w).$$

ratkaisuja (z, w) joukossa $\mathbb{C}^2$.

Koska avaruuden $\mathbb{C}^2$ alkiot (z_1, w_1) ja (z_2, w_2) ovat sama alkio, jos ja vain jos $z_1 = z_2$ ja $w_1 = w_2$, niin saadaan yhtälöpari

$$\begin{cases} az & -bw & = \lambda z \\ bz & +aw & = \lambda w \end{cases}$$

eli yhtälöpari

$$\begin{cases} (a - \lambda)z & -bw & = 0 \\ bz & +(a - \lambda)w & = 0. \end{cases}$$

Koska $\lambda = a + ib$, niin oikeasti olemme tarkastelemassa kahta yhtälöparia

$$\begin{cases} -ibz & -bw & = 0 \\ bz & -ibw & = 0. \end{cases}$$

Koska $b \neq 0$, niin tämä yhtälöpari on ekvivalentti yhtälöparin

$$\begin{cases} iz & +w & = 0 \\ z & -iw & = 0 \end{cases}$$

kanssa. Koska $i^2 = -1$, niin yhtälöparin ensimmäinen rivi antaa

$$z = -\frac{w}{i} = -\frac{iw}{i^2} = iw,$$

joka on sama ehto kuin minkä yhtälöparin toinen rivi antaa. Näin ollen yhtälön (13.2) ratkaisujoukko on

$$\{(iw, w) \in \mathbb{C}^2 \colon w \in \mathbb{C}\}.$$

Tarkastellaan tätä ratkaisujoukkoa vielä lähemmin. Avaruuden $\mathbb{C}^2$ laskutoimitusten avulla saadaan, että

$$V = \{(iw, w) \in \mathbb{C}^2 \colon w \in \mathbb{C}\} = \{w(i, 1) \in \mathbb{C}^2 \colon w \in \mathbb{C}\}.$$

Koska

$$\varphi((z, w)) = \lambda(z, w)$$

kaikilla $(z, w) \in V$, niin annettujen laskutoimitusten avulla on helppo havaita, että kaikilla $\alpha \in \mathbb{C}$ ja $(z_1, w_1), (z_2, w_2) \in V$ pätee $\alpha(z_1, w_1) + (z_2, w_2) \in V$.

Näin ollen, voidaan tulkita, että V on itseasiassa avaruuden $\mathbb{C}^2$ (kompleksisesti) 1-ulotteinen aliavaruus ja kuvauksen φ ominaisarvoa λ vastaava ominaisavaruus. Tämä on se tulkinta, joka seuraavaksi formalisoidaan.

13.2 Kompleksiluvut

Joissain esityksissä kompleksiluvut esitellään sanomalla, että kompleksiluvut ovat lukuja $z = x + iy$, missä $x \in \mathbb{R}$ ja $y \in \mathbb{R}$ ja sanomalla, että kompleksiluvuille pätee reaalilukujen tavalliset yhteen- ja kertolasku sillä erotuksella, että $i^2 = -1$. Tämä tarkoittaa sitä, että jos $z = x + iy$ ja $w = x' + iy'$ ovat kompleksilukuja, niin voidaan laskea

$$z + w = (x + iy) + (x' + iy') = x + iy + x' + iy' = (x + x') + i(y + y')$$

ja

$$\begin{aligned} zw &= (x + iy)(x' + iy') = xx' + ixy' + iyx' + (iy)(iy') \\ &= xx' + i(xy' + yx') + i^2(yy') = xx' - yy' + i(xy' + yx'). \end{aligned}$$

Kompleksiluku $z = x + iy$ määräytyy siis täysin lukuparista $(x, y) \in \mathbb{R}^2$. Ero kompleksilukujen ja tason $\mathbb{R}^2$ vektoreiden välillä on siis ainoastaan kertolaskussa. Määritellään nyt kompleksiluvut tarkasti.

Määritelmä 13.2.1. Kompleksiluvut $\mathbb{C}$ *ovat kolmikko* $(\mathbb{R}^2, +, \cdot)$, *missä* $+\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ *on tavallinen vektoreiden yhteenlasku ja* $\cdot\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ *on laskutoimitus*

$$(x, y)(x', y') = (xx' - yy', xy' + x'y)$$

kaikilla $(x, y), (x', y') \in \mathbb{R}^2$.

Jatkossa merkinnällä $\mathbb{C}$ tarkoitetaan sekä joukkoa $\mathbb{R}^2$ että kolmikkoa $(\mathbb{R}^2, +, \cdot)$. Joukon $\mathbb{C}$ alkioita kutsutaan *kompleksiluvuiksi.* Kompleksilukua $i = e_2$ kutsutaan *imaginääriyksiköksi* Lisäksi otetaan käyttöön merkintä $x(1, 0) = x$ kaikilla $x \in \mathbb{R}$, jonka avulla kompleksiluku $z = (x, y) \in \mathbb{C}$ voidaan kirjoittaa sen klassisessa esitysmuodossa

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

Käyttämällä näitä merkintöjä saadaan tuttu kaava

$$i^2 = e_2e_2 = (0, 1)(0, 1) = (0 - 1, 0 + 0) = (-1, 0) = -e_1 = -1.$$

Olkoon $z = x + iy \in \mathbb{C}$. Lukua x kutsutaan kompleksluvun z *reaaliosaksi* $\Re z$ ja lukua y luvun z *imaginääriosaksi* $\Im z$ eli $z = \Re z + i\Im z$. Huomaa, että kompleksiluvun imaginääriosa on tyypillistä kirjoittaa muodossa iy, vaikka i on vektori ja y on reaaliluku.

Huomautus 13.2.2. *Edellä esitellyllä merkinnällä on syvällinen tulkinta. Jokainen reaaliluku* $x \in \mathbb{R}$ *voidaan tulkita kompleksilukuna* $x = x + i0 \in \mathbb{C}$. *Käytännössä tämä tarkoittaa, että reaaliluvut tulkitaan kompleksilukujen osajoukoksi eli* $\mathbb{R} \subset \mathbb{C}$. *Näin tehtäessä reaaliluvut* $\mathbb{R}$ *itseasiassa samastetaan tason* $\mathbb{R}^2$ x*-akselin eli suoran* $\{(x, 0) \in \mathbb{R}^2 \colon x \in \mathbb{R}\}$ *kanssa. Koska tässä tulkinnassa reaalilukujen laskutoimitukset sopivat yhteeen kompleksilukujen laskutoimitusten kanssa, niin algebran kieltä käyttäen sanotaan, että reaaliluvut ovat kompleksilukujen (eräs) alikunta. Alikuntia käsitellään tarkemmin syventävissä opinnoissa.*

Näin määriteltynä kompleksiluvut $\mathbb{C}$ ovat algebrallisessa mielessä *kunta*, eli niille pätee reaaliluvuilta tutut ominaisuudet

1. $(z + w) + u = z + (w + u)$ ja $(zw)u = z(wu)$ kaikilla $z, w, u \in \mathbb{C}$,
2. $z + w = w + z$ ja $zw = wz$ kaikilla $z, w \in \mathbb{C}$,
3. luvuille $0 = (0, 0)$ ja $1 = (1, 0)$ pätee $z + 0 = z$ ja $1z = z$ kaikilla $z \in \mathbb{C}$,
4. jokaisella $z \in \mathbb{C}$ on olemassa yksikäsitteinen $w \in \mathbb{C}$, jolle pätee $z + w = 0$ ja jota merkitään $w = -z$,
5. jokaisella $z \in \mathbb{C} \setminus \{0\}$ on olemassa yksikäsitteinen $w \in \mathbb{C}$, jolle pätee $zw = 1$, ja jota merkitään $w = 1/z$, sekä
6. $z(w + u) = zw + zu$ kaikilla $z, w, u \in \mathbb{C}$.

Näistä ominaisuuksista haastavin on käänteisluvun $1/z$ olemassaolo, eli kohta (5). Huomaa, että merkintä käänteisluvun merkintä $1/z$ on perusteltu vasta, kun käänteisluku on osoitettu yksikäsitteiseksi. Selkeyden vuoksi merkintä on jo esitelty kunta-aksioomien yhteydessä yllä. Tätä kompleksilukujen ominaisuutta käsitellään tarkemmin alla. Muut kompleksilukujen ominaisuudet oletetaan jatkossa tunnetuiksi. Käänteisluvun olemassaolo ja yksikäsitteisyys (eli ominaisuus (5)) osoitetaan käyttäen kompleksikonjugaatin ja normin käsiteitä. Seuraavasssa todistuksessa oletetaan, että ehdot 1 ja 3 pätevät.

Määritelmä 13.2.3. *Kompleksiluvun $z = x + iy \in \mathbb{C}$* kompleksikonjugaatti $\overline{z} \in \mathbb{C}$ *on luku $\overline{z} = x - iy$.*

Määritelmä 13.2.4. *Kompleksiluvun $z \in \mathbb{C}$* itseisarvo $|z| \in \mathbb{R}$ *on luku $|z| = \sqrt{z\overline{z}}$.*

Huomautus 13.2.5. *Huomaa, että kompleksiluvun $z = x+iy$ itseisarvo on itseasiassa vektorin $(x, y) \in \mathbb{R}^2$ euklidinen pituus (eli normi), sillä*

$$|z|^2 = z\overline{z} = (x+iy)\overline{(x+iy)} = (x+iy)(x-iy) = x^2-xiy+iyx-i^2y^2 = x^2+y^2 = |(x,y)|^2.$$

Lisäksi kompleksiluvuille $z = \Re z \in \mathbb{C}$, joiden imaginääriosa on nolla, itseisarvo $|z|$ on reaaliluvun $\Re z$ itseisarvo.

Huomattakoon vielä, että erityisesti kompleksiluku z on nollasta poikkeava eli $z \neq 0$, jos ja vain jos $|z| > 0$.

Lemma 13.2.6. *Olkoon $z \in \mathbb{C}$ nollasta poikkeava kompleksiluku. Tällöin luvulle $w = \overline{z}/|z|^2$ pätee $zw = 1$ eli*

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2}.$$

Lisäksi w on yksikäsitteinen sellainen kompleksiluku, jolle pätee $zw = 1$.

Todistus. Koska

$$zw = z\frac{\overline{z}}{|z|^2} = \frac{z\overline{z}}{|z|^2} = \frac{|z|^2}{|z|^2} = 1,$$

väitteen ensimmäinen osa pätee.

Oletetaan nyt, että $w' \in \mathbb{C}$ olisi sellainen kompleksiluku, jolle pätee $zw' = 1$. Tällöin

$$w = w1 = w(zw') = (wz)w' = w'$$

ehtojen 1 ja 3 perusteella. □

Huomautus 13.2.7. *Kompleksikonjugaatti on tärkeä kompleksiluvuille karakteristinen käsite, jonka perusominaisuudet $\overline{zw} = \overline{z}\,\overline{w}$ ja $\overline{\overline{z}} = z$ on hyvä tuntea. Lineaarialgebrallisesti kompleksikonjugaatti $\tau\colon \mathbb{C} \to \mathbb{C}$, $z \mapsto \overline{z}$, on lineaarikuvaus $\tau\colon \mathbb{R}^2 \to \mathbb{R}^2$, jolle pätee $e_1 \mapsto e_1$ ja $e_2 \mapsto -e_2$.*

Huomautus 13.2.8. *Useissa kompleksilukujen perusesityksissä käsitellään määritelmän yhteydessä kompleksiluvun argumenttia ja polaariesitystä sekä niihin liittyviä de Moivrén kaavaa ja eksponenttifunkiota. Tässä esityksessä näitä asioita ei käsitellä, koska ne eivät ole keskeisessä roolissa jatkossa.*

13.3 Kompleksiset vektoriavaruudet

Kompleksiset vektoriavaruudet eli kompleksikertoimiset vektoriavaruudet ovat itseasiassa todella helppo asia:

> Vaihda kaikissa lineaarialgebran tutuissa käsitteissä reaaliluvut kompleksilukuihin. Tutut tulokset ja menetelmät pätevät.

Kirjataan nyt hieman formaalimmin ja tarkemmin tuon yleisperiaatteen seuraukset. Kompleksinen vektoriavaruus määritellään aivan kuten reaalinen vektoriavaruus.

Määritelmä 13.3.1. *Joukko V laskutoimituksilla $+\colon V \times V \to V$ ja $\cdot\colon \mathbb{C} \times V \to V$ on* kompleksinen vektoriavaruus, *jos*

1. *kaikilla $v, w, u \in V$ pätee $v + w = w + v$ ja $(v + w) + u = v + (w + u)$,*
2. *on olemassa nollavektori $0 \in V$, jolle pätee $v + 0 = v$ kaikilla $v \in V$,*
3. *jokaisella vektorilla $v \in V$ on vastavektori $-v \in V$, jolle pätee $v + (-v) = 0$,*
4. *kaikilla vektoreilla $v, w \in V$ ja $a \in \mathbb{C}$ pätee $a(v + w) = av + aw$,*
5. *jokaisella $v \in V$ ja kaikilla $a, b \in \mathbb{C}$ pätee $(a + b)v = av + bv$ ja $a(bv) = (ab)v$ ja*
6. *$1v = v$ jokaisella $v \in V$.*

Esimerkki 13.3.2. *Ensimmäinen esimerkki kompleksisesta vektoriavaruudesta on avaruutta $\mathbb{R}^n$ vastaava avaruus $\mathbb{C}^n$ eli kompleksilukujonojen $(z_1, \dots, z_n)$ joukko varustettuna laskutoimituksilla*

$$(z_1, \dots, z_n) + (w_1, \dots, w_n) = (z_1 + w_1, \dots, z_n + w_n)$$

ja

$$a(z_1, \dots, z_n) = (az_1, \dots, az_n)$$

kaikilla $(z_1, \dots, z_n), (w_1, \dots, w_n) \in \mathbb{C}^n$ ja $a \in \mathbb{C}$.

Avaruudet $\mathbb{C}^{n\times 1}$ ja yleisemmin $\mathbb{C}^{m\times n}$ kaikilla $m, n \in \mathbb{N}$ määritellään kuin reaalisessa tapauksessa avaruudet $\mathbb{R}^{n\times 1}$ ja $\mathbb{R}^{m\times n}$.

Esimerkki 13.3.3. *Olkoon X joukko ja olkoon $\mathscr{F}_{\mathbb{C}}(X)$ kaikkien kompleksiarvoisten funktioiden $f\colon X \to \mathbb{C}$ joukko. Tällöin $\mathscr{F}_{\mathbb{C}}(X)$ on vektoriavaruus $(\mathscr{F}_{\mathbb{C}}(X), +, \cdot)$, missä laskutoimitukset $+\colon \mathscr{F}_{\mathbb{C}}(X) \times \mathscr{F}_{\mathbb{C}}(X) \to \mathscr{F}_{\mathbb{C}}(X)$, $(f, g) \mapsto f + g$, ja $\cdot\colon \mathbb{C} \times \mathscr{F}_{\mathbb{C}}(X) \to \mathscr{F}_{\mathbb{C}}(X)$, $(a, f) \mapsto af$ on määritelty kaikilla $f, g \in \mathscr{F}_{\mathbb{C}}(X)$ ja $a \in \mathbb{C}$ kaavoilla*

$$(f + g)(x) = f(x) + g(x)$$

ja

$$(af)(x) = af(x)$$

kaikilla $x \in X$.

Vastaavalla tavalla määritellään kompleksisesti derivoituvien funktioiden (eli kompleksianalyyttisten funktioiden) avaruus ja jatkuvien funktioiden avaruudet.

Esimerkki 13.3.4. *Funktio $p\colon \mathbb{C} \to \mathbb{C}$ on (kompleksinen) polynomi, jos on olemassa sellaiset vakiot $a_0, \dots, a_n \in \mathbb{C}$, että $p(z) = a_n z^n + \cdots + a_1 z + a_0$ kaikilla $z \in \mathbb{C}$. Kompleksisten polynomien joukko $\mathscr{P} = \{p\colon \mathbb{C} \to \mathbb{C} \mid p \text{ on polynomi}\}$ on vektoriavaruus, kun yhteenlaskuksi ja skalaarikertolaskuksi valitaan funktioiden yhteen- ja skalaarikertolasku.*

Kuten reaalisten vektoriavaruuksien tapauksessa, aliavaruudet antavat lisää esimerkkejä kompleksisista vektoriavaruuksista.

Määritelmä 13.3.5. *Kompleksisen vektoriavaruuden V osajoukko $W \subset V$ on* aliavaruus, *jos*

1. $0 \in W$,
2. $w + w' \in W$ *kaikilla* $w, w' \in W$ *ja*
3. $aw \in W$ *kaikilla* $w \in W$ *ja* $a \in \mathbb{C}$.

Esimerkki 13.3.6. *Olkoon $\mathscr{P}_n$ kaikkien korkeintaan astetta n olevien kompleksisten polynomien joukko, eli $p \in \mathscr{P}_n$, jos $p\colon \mathbb{C} \to \mathbb{C}$ on funktio $p(z) = a_n z^n + \cdots + a_1 z + a_0$. Tällöin $\mathscr{P}_n$ on avaruuden $\mathscr{P}$ aliavaruus.*

Esimerkki 13.3.7. *Olkoon $q\colon \mathbb{C} \to \mathbb{C}$ polynomi. Polynomi $p \in \mathscr{P}$ on jaollinen polynomilla q, jos on olemassa sellainen polynomi $f \in \mathscr{P}$, että $p = qf$. Olkoon $\mathscr{P}(q)$ kaikkien sellainen polynomien joukko, jotka ovat jaollisia polynomilla q eli $\mathscr{P}(q) = \{qf \in \mathscr{P} : f \in \mathscr{P}\}$. Merkitään myös $\mathscr{P}_n(q) = \mathscr{P}(q) \cap \mathscr{P}_n$. Tällöin $\mathscr{P}(q)$ on avaruuden $\mathscr{P}$ aliavaruus ja $\mathscr{P}_n(q)$ on sekä avaruuden $\mathscr{P}(q)$ että avaruuden $\mathscr{P}_n$ aliavaruus.*

13.3.1 Lineaarikombinaatiot, virittäminen ja vapaus

Vektoriavaruuksien peruskäsitteet kuten lineaarikombinaatio, virittäminen ja vapaus määritellään kompleksisten vektoriavaruuksien tapauksessa aivan kuten reaalisten vektoriavaruuksien tapauksessa. Ainoa ero on, että reaalilukujen sijaan käytetään kompleksilukuja.

Määritelmä 13.3.8. *Kompleksisen vektoriavaruuden V vektori $v \in V$ on vektoreiden $v_1, \dots, v_n \in V$* lineaarikombinaatio, *jos on olemassa sellaiset luvut $a_1, \dots, a_n \in \mathbb{C}$, että*

$$v = a_1 v_1 + \cdots + a_n v_n.$$

Määritelmä 13.3.9. *Olkoon V kompleksikertoiminen vektoriavaruus. Jonon $(v_1, \dots, v_n)$* virittämä aliavaruus $\mathrm{Sp}(v_1, \dots, v_n)$ *on*

$$\mathrm{Sp}(v_1, \dots, v_n) = \{a_1 v_1 + \cdots + a_n v_n \in V : a_1, \dots, a_n \in \mathbb{C}\}.$$

Määritelmä 13.3.10. *Jono $(v_1, \dots, v_n)$ kompleksisen vektoriavaruuden V vektoreita* virittää *avaruuden V, jos jokainen vektori $v \in V$ on vektoreiden $v_1, \dots, v_n$ lineaarikombinaatio eli $V = \mathrm{Sp}(v_1, \dots, v_n)$. Lisäksi tällöin sanotaan, että avaruus V on äärellisesti viritetty.*

Määritelmä 13.3.11. *Jono* $(v_1, \ldots, v_n)$ *kompleksisen vektoriavaruuden* V *vektoreita* on vapaa*, jos yhtälön*

$$a_1v_1 + \cdots + a_nv_n = 0,$$

missä $a_1, \ldots, a_n \in \mathbb{C}$*, ainoa ratkaisu on* $a_1 = a_2 = \cdots = a_n = 0$*.*

13.3.2 Kanta ja dimensio

Kompleksisen vektoriavaruuden kanta määritellään kuten reaalisessa tapauksessa.

Määritelmä 13.3.12. *Jono* $(v_1, \ldots, v_n)$ *kompleksisen vektoriavaruuden* V *vektoreita on* avaruuden V kanta*, jos jono* $(v_1, \ldots, v_n)$ *on vapaa ja virittää avaruuden* V*.*

Esimerkki 13.3.13. *Avaruuden* $\mathbb{C}^n$ *standardikanta on* $(e_1, \ldots, e_n)$*, missä jokainen* e_i *on vektori* $e_i = (e_{1i}, \ldots, e_{ni})$*, missä* $e_{ji} = \delta_{ji}$ *kaikilla* $j \in \{1, \ldots, n\}$*. Jono* $(e_1, \ldots, e_n)$ *on kanta, koska jokaisella* $v = (z_1, \ldots, z_n) \in \mathbb{C}^n$ *pätee*

$$v = (z_1, \ldots, z_n) = z_1e_1 + \cdots + z_ne_n$$

ja luvut $z_1, \ldots, z_n$ *ovat selvästi yksikäsitteisiä.*

Avaruuden $\mathbb{C}^{n\times 1}$ *standardikanta* $(e_1, \ldots, e_n)$ *määritellään analogisesti.*

Esimerkki 13.3.14. *Määritellään jokaisella* $k \in \mathbb{N}$ *polynomi* $p_k \colon \mathbb{C} \to \mathbb{C}$ *kaavalla* $z \mapsto z^k$*. Vastaavasti kuin reaalisessa tapauksessa pätee, että jokaisella* $n \in \mathbb{N}$ *jono* $(p_0, \ldots, p_n)$ *on polynomiavaruuden* $\mathscr{P}_n$ *kanta. Erityisesti saadaan, että jokaisella* $p \in \mathscr{P}$ *kaavan*

$$p(z) = a_0 + a_1z + \cdots + a_nz^n$$

kertoimet $a_0, \ldots, a_n \in \mathbb{C}$ *ovat yksikäsitteiset. Tästä saadaan seurauksena, että polynomin* $p \in \mathscr{P}$ aste $\deg p \in \mathbb{N}$ *eli korkeimman nollasta poikkeavan kertoimen indeksi, on hyvin määritelty.*

Luvusta 8 tutut tulokset pätevät myös äärellisesti viritetyille kompleksisille vektoriavaruuksille.

Lause 13.3.15. *Jokaisella äärellisesti viritetyllä kompleksisella vektoriavaruudella on kanta.*

Lause 13.3.16. *Äärellisesti viritetyn kompleksisen vektoriavaruuden jokaisessa kannassa on yhtä monta alkiota.*

Näin ollen jokainen äärellisesti viritetty vektoriavaruus on äärellisulotteinen.

Määritelmä 13.3.17. *Äärellisesti viritetyn kompleksisen vektoriavaruuden* V (kompleksinen) dimensio $\dim V$ *on kannan* $(v_1, \ldots, v_n)$ *alkioiden lukumäärä* n*.*

Esimerkki 13.3.18. *Avaruuden* $\mathbb{C}^n$ *dimensio* $\dim \mathbb{C}^n$ *kompleksisena vektoriavaruutena on* n*. Samoin* $\dim \mathbb{C}^{n\times 1} = n$*. Polynomien* $\mathscr{P}_n$ *avaruuden dimensio on* $\dim \mathscr{P}_n = n + 1$*.*

Myös kompleksisten vektoriavaruuksien tapauksessa kanta määrittelee koordinaatit.

Lemma 13.3.19. *Olkoon* V *kompleksinen vektoriavaruus. Tällöin jono* $(v_1, \ldots, v_n)$ *on avaruuden* V *kanta, jos ja vain jos jokaisella* $v \in V$ *on yksikäsitteiset sellaiset luvut* $a_1, \ldots, a_n \in \mathbb{C}$*, joille pätee*

$$v = a_1v_1 + \cdots + a_nv_n.$$

13.3.3 Aliavaruuksien summa ja suora summa

Kuten reaalisten vektoriavaruuksien tapauksessa, aliavaruuksien leikkaus ja summa ovat aliavaruuksia.

Määritelmä 13.3.20. *Kompleksisen vektoriavaruuden V aliavaruuksien U ja W* summa *$U + W$ on aliavaruus*

$$U + W = \{u + w \in V : u \in U,\ w \in W\}.$$

Summa $U + W$ on suora*, jos $U \cap W = \{0\}$. Aliavaruuksin U ja W suoraa summaa merkitään $U \oplus W$.*

Kuten reaalisssa tapauksessa äärelisulotteisten aliavaruuksien U ja W summa on suora, jos niiden kannat voidaan yhdistää aliavaruuden $U + W$ kannaksi.

Lause 13.3.21. *Olkoon V kompleksinen vektoriavaruus sekä olkoot $(u_1, \dots, u_n)$ aliavaruuden U kanta ja $(w_1, \dots, w_m)$ aliavaruuden W kanta. Tällöin summa $U+W$ on suora, jos ja vain jos $(u_1, \dots, u_n, w_1, \dots, w_m)$ on aliavaruuden $U + W$ kanta.*

Korollaari 13.3.22. *Olkoon V kompleksinen vektoriavaruus ja olkoot U ja W sellaisia äärellisulotteisia aliavaruuksia, joiden summa on suora. Tällöin*

$$\dim(U \oplus W) = \dim U + \dim W.$$

Aliavaruuksien dimensiolause pätee kompleksisille avaruuksille samalla todistuksella kuin reaalisessa tapauksessa.

Lause 13.3.23 (Aliavaruuksien dimensiolause kompleksisille avaruuksille)**.** *Olkoon V kompleksinen vektoriavaruus ja olkoot U ja W sen äärellisulotteisia alivaruuksia. Tällöin $U + W$ on äärellisulotteinen ja*

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Useamman aliavaruuden suora summa

Useamman aliavaruuden suora summa määritellään induktiivisesti.

Määritelmä 13.3.24. *Avaruuden V aliavaruuksien $W_1, \dots, W_m$ summa on* suora*, jos*

$$(W_1 + \cdots + W_{k-1}) \cap W_k = \{0\}$$

kaikilla $k \in \{2, \dots, m\}$. Tällöin merkitään $W_1 \oplus \cdots \oplus W_m = W_1 + \cdots + W_m$.

Lause 13.3.25. *Olkoon V kompleksinen vektoriavaruus ja olkoot $W_1, \dots, W_m$ sellaisia äärellisulotteisia aliavaruuksia, joiden summa on suora. Tällöin*

$$\dim(W_1 \oplus \cdots \oplus W_m) = \dim W_1 + \cdots + \dim W_m.$$

Todistus. Määritelmästä 13.3.24 seuraa, että $W_1 \oplus \cdots \oplus W_m = (W_1 \oplus \cdots \oplus W_{m-1}) \oplus W_m$. Näin ollen korollaarin 13.3.22 perusteella

$$\dim(W_1 \oplus \cdots \oplus W_m) = \dim(W_1 \oplus \cdots \oplus W_{m-1}) + \dim W_m.$$

Väite seuraa nyt induktiolla lukumäärän m suhteen. □

Lauseen 8.3.3 vastine useamman aliavaruuden suoran summan tapauksessa on seuraava.

Lause 13.3.26. *Olkoon V kompleksinen vektoriavaruus ja olkoot $W_1, \ldots, W_m$ avaruuden V äärellisulotteisia aliavaruuksia. Tällöin summa $W_1 + \cdots + W_m$ on suora, jos ja vain jos jokaisella $v \in V$ on olemassa yksikäsitteiset sellaiset vektorit $w_1 \in W_1, \ldots, w_m \in W_m$, että $v = w_1 + \cdots + w_m$.*

Seuraava lause on todistettu liitteessä D (lause D.1.5).

Lause 13.3.27. *Olkoon V äärellisulotteinen kompleksinen vektoriavaruus. Olkoot lisäksi $W_1, \ldots, W_m$ avaruuden V aliavaruuksia ja olkoon $(w_{1j}, \ldots, w_{d_j j})$ aliavaruuden W_j kanta jokaisella $j \in \{1, \ldots, m\}$. Tällöin summa $W_1 + \cdots + W_m$ on suora, jos ja vain jos jono $(w_{11}, \ldots, w_{d_1 1}, w_{12}, \ldots, w_{d_m m})$ on aliavaruuden $W_1 + \cdots + W_m$ kanta.*

13.3.4 Suoraan summaan liittyvät projektiot

Luvun 8.3.2 tulos, että avaruuden suoraan summaan $V = U \oplus W$ liittyy luonnolliset projektiot $\pi_U \colon V \to U$ ja $\pi_W \colon V \to W$, pätee myös useamman summattavan tapauksessa. Palautetaan tätä varten mieleen projektion määritelmä.

Määritelmä 13.3.28. *Operaattori $\pi \colon V \to V$ on* projektio avaruuden V aliavaruudelle *$W \subset V$, jos $\operatorname{im} \pi = W$ ja $\pi \circ \pi = \pi$.*

Suoraan summaan $V = V_1 \oplus \cdots \oplus V_m$ liittyvät luonnolliset projektiot $\pi_{V_i} \colon V \to V$ määritellään seuraavasti. Lauseen todistus on lähes identtinen lauseen 8.3.12 kanssa.

Lause 13.3.29. *Olkoon V vektoriavaruus ja $V = V_1 \oplus \cdots \oplus V_m$. Tällöin jokaisella $j \in \{1, \ldots, m\}$ on olemassa yksikäsitteinen sellainen projektio $\pi_{V_j} \colon V \to V$ aliavaruudelle V_j, että*

$$v = \pi_{V_1}(v) + \cdots + \pi_{V_m}(v)$$

jokaisella $v \in V$.

Huomautus 13.3.30. *Kuten lauseen 8.3.12 todistus osoittaa, projektiot $\pi_{V_1}, \ldots, \pi_{V_m}$ ovat yksikäsitteisiä, koska jokaisella $v \in V$ on yksikäsitteinen esitys aliavaruuksien $V_1, \ldots, V_m$ vektoreiden summana. Tarkemmin sanottuna lauseen 13.3.26 nojalla jokaiselle projektiolle π_{V_j} pätee $\pi_{V_j}(v) = v_j$, missä vektorit $v_1 \in V_1, \ldots, v_m \in V_m$ ovat ne yksikäsitteiset vektorit, joille pätee $v = v_1 + \cdots + v_m$.*

13.4 Kompleksiset lineaarikuvaukset

Määritelmä 13.4.1. *Kuvaus $f: V \to W$ kompleksisten vektoriavaruuksien V ja W välillä on* (kompleksi)lineaarinen, *jos kaikilla $v, w \in V$ ja $a \in \mathbb{C}$ pätee*

$$f(av + w) = af(v) + f(w).$$

Kuten reaalisessa tapauksessa lineaarikuvaus määräytyy yksikäsitteisesti kanta-alkioiden kuvista.

Lause 13.4.2. *Olkoot V ja W kompleksisia vektoriavaruuksia ja $(v_1, \ldots, v_n)$ avaruuden V kanta. Olkoot myös $w_1, \ldots, w_n \in W$. Tällöin on olemassa täsmälleen yksi sellainen lineaarikuvaus $f: V \to W$, jolle pätee $f(v_i) = w_i$ kaikilla $i \in \{1, \ldots, n\}$.*

Kirjataan vielä lineaarikuvausten peruslause kompleksisessa tapauksessa. Jälleen todistus on sama kuin reaalisessa tapauksessa.

Lause 13.4.3. *Olkoot V ja W kompleksisia vektoriavaruuksia ja olkoon $f: V \to W$ lineaarikuvaus. Jos V on äärellisulotteinen, niin $\operatorname{im} f$ on avaruuden W äärellisulotteinen aliavaruus ja*

$$\dim V = \dim \ker f + \dim \operatorname{im} f.$$

13.4.1 Kompleksiset isomorfismit ja kompleksisten avaruuksien samastaminen

Määritelmä 13.4.4. *Kompleksinen lineaarikuvaus $f: V \to W$ on* isomorfismi, *jos f on bijektio.*

Lause 13.4.5. *Olkoon V äärellisulotteinen kompleksinen vektoriavaruus ja $(v_1, \ldots, v_n)$ avaruuden V kanta. Tällöin kuvaus*

$$\Psi_{(v_1,\ldots,v_n)}: \mathbb{C}^{n\times 1} \to V, \qquad \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \mapsto z_1v_1 + \cdots + z_nv_n,$$

on isomorfismi.

Korollaari 13.4.6. *Jokainen äärellisulotteinen kompleksikertoiminen vektoriavaruus on isomorfinen jonkin avaruuden $\mathbb{C}^{n\times 1}$ kanssa, eli tarkemmin: Jos V on n-ulotteinen kompleksinen vektoriavaruus, niin on olemassa lineaarinen isomorfismi $\Psi: \mathbb{C}^{n\times 1} \to V$.*

13.5 Kompleksiset matriisit

Kompleksikertoimiset matriisit määritellään kuten reaaliset matriisit eli matriisi $A \in \mathbb{C}^{m\times n}$ on *taulukko*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

missä $a_{ji} \in \mathbb{C}$ kaikilla $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, n\}$.

Kompleksikertoimisten matriisien osat kuten rivit ja sarakkeet sekä laskutoimitukset kuten yhtenlasku ja matriisitulo määritellään kuten reaalisessa tapauksessa. Erityisesti siis käänteismatriisi on määritelty kuten aiemmin, kuten myös transpoosi. Huomaa, että identiteettimatiriisi $I \in \mathbb{C}^{n\times n}$ on matriisi $I = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix}$, missä $(e_1, \ldots, e_n)$ on avaruuden $\mathbb{C}^{n\times 1}$ standardikanta.

Kompleksikojugaatin käsite on hyödyllistä laajentaa matriiseille ja sarekevektoreille. Matriisin

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{C}^{m\times n}$$

kompleksikonjugaatti on matriisi

$$\overline{A} = \overline{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}} = \begin{bmatrix} \overline{a_{11}} & \cdots & \overline{a_{1n}} \\ \vdots & & \vdots \\ \overline{a_{m1}} & \cdots & \overline{a_{mn}} \end{bmatrix} \in \mathbb{C}^{m\times n}.$$

Sarakevektorin $z \in \mathbb{C}^{n\times 1}$ kompleksikonjugaatti $\overline{z} \in \mathbb{C}^{n\times 1}$ on tämän määritelmän erikoistapaus.

Kompleksikertoiminen matriisi $A \in \mathbb{C}^{m\times n}$ määrittelee lineaarikuvauksen $\varphi_A \colon \mathbb{C}^{n\times 1} \to \mathbb{C}^{m\times 1}$ kaavalla $z \mapsto Az$, jossa Az on matriisin A ja sarakevektorin

$$z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^{n\times 1}$$

matriisitulo.

Luvussa 9.5 käsitelty lineaarikuvausten esitysmatriisilause yleistyy suoraan kompleksisille lineaarikuvauksille.

Lause 13.5.1 (Kompleksisen lineaarikuvauksen esitysmatriisi)**.** *Olkoon $\varphi \colon V \to W$ äärellisulotteisten kompleksisten vektoriavaruuksien V ja W välinen lineaarikuvaus sekä olkoon $(v_1, \ldots, v_n)$ avaruuden V kanta ja olkoon $(w_1, \ldots, w_m)$ avaruuden W kanta. Tällöin on olemassa yksikäsitteinen sellainen matriisi $A \in \mathbb{C}^{m\times n}$, että*

$$\Psi_{(w_1,\ldots,w_m)} \circ \varphi_A = \varphi \circ \Psi_{(v_1,\ldots,v_n)}.$$

13.5.1 Kompleksisen neliömatriisin kääntyvyys ja determinantti

Kompleksisille neliömatriiseille pätee samat tulokset kuin reaalisille neliömatriiseille. Näistä olennaisimmat on koottu seuraavaan lauseeseen.

Lause 13.5.2. *Olkoon $A \in \mathbb{C}^{n\times n}$ neliömatriisi. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *matriisi A on kääntyvä,*

2. matriisin A sarakkeet virittävät avaruuden $\mathbb{C}^{n\times 1}$,

3. matriisin A sarakkeet ovat lineaarisesti riippumattomia,

4. lineaarikuvaus $\varphi_A\colon \mathbb{C}^{n\times 1}\to\mathbb{C}^{n\times 1}$ on isomorfismi ja

5. $\det A \neq 0$.

Kommentoidaan lyhyesti ehtojen yhtäpitävyyttä. Ehtojen 1–3 yhtäpitävyys seuraa samoista päättelyistä kuin lauseessa 3.10.1. Ehtojen 1 ja 4 seuraa lineaarikuvausten kantalauseesta ja jätetään harjoitustehtäväksi.

Tarkastellaan nyt tarkemmin ehtojen 1 ja 5 yhtäpitävyyttä. Tätä varten tulee ensin määritellä kompleksisen neliömatriisin determinantti. Kuten arvata saattaa, se määritellään täsmälleen samalla kaavalla kuin reaalisessa tapauksessa.

Määritelmä 13.5.3. *Neliömatriisin $A=[a_{ji}]\in\mathbb{C}^{n\times n}$* determinantti *on kompleksiluku*

$$\det A = \sum_{\sigma\in\mathrm{Sym}(n)} \mathrm{sign}(\sigma) a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}. \tag{13.3}$$

Kompleksisessa tapauksessa determinantin käsite voidaan johtaa avaruuden $\mathbb{C}^{n\times 1}$ tilavuusmuodosta täsmälleen samoin perusteluin kuin reaalisessa tapauksessa. Kunnan vaihtuminen reaalisesta kompleksiseen ei myöskään vaikuta muihin luvussa 5 argumentteihin. Tämän vuoksi kompleksisen neliömatriisin determinantilla on täsmälleen samat ominaisuudet kuin reaalisen neliömatriisin determinantilla.

Lause 13.5.4 (Sarakelineaarisuus). *Olkoon $A=\begin{bmatrix} v_1 & \cdots & v_n\end{bmatrix}\in\mathbb{C}^{n\times n}$ neliömatriisi ja $v,w\in\mathbb{C}^{n\times 1}$. Tällöin jokaisella $1\le k\le n$ pätee*

$$\begin{aligned}
&\det\begin{bmatrix} v_1 & \cdots v & v_{k-1} & v+w & v_{k+1} & \cdots & v_n\end{bmatrix}\\
&= \det\begin{bmatrix} v_1 & \cdots v & v_{k-1} & v & v_{k+1} & \cdots & v_n\end{bmatrix}\\
&+ \det\begin{bmatrix} v_1 & \cdots v & v_{k-1} & w & v_{k+1} & \cdots & v_n\end{bmatrix}.
\end{aligned}$$

Lause 13.5.5 (Alternoivuus). *Olkoon $A=\begin{bmatrix} v_1 & \cdots & v_n\end{bmatrix}\in\mathbb{C}^{n\times n}$ sellainen neliömatriisi, että $v_i=v_j$ jollakin $j\neq i$. Tällöin* $\det A = 0$.

Lause 13.5.6 (Determinantin transpoosi-invarianssi). *Olkoon $A\in\mathbb{C}^{n\times n}$ neliömatriisi. Tällöin* $\det A^t = \det A$.

Lause 13.5.7 (Tulomatriisin determinantti). *Olkoot $A\in\mathbb{C}^{n\times n}$ ja $B\in\mathbb{C}^{n\times n}$ neliömatriiseja. Tällöin* $\det(AB)=(\det A)(\det B)$.

Huomautus 13.5.8. *Erityisesti käytännön laskuissa kompleksisen neliömatriisin determinantti voidaan laskea kehityskaavalla rivin tai sarakkeen suhteen, vaikka tätä ei edellä kirjattu uudeksi tulokseksi.*

Neliömatriisin kääntyvyyden ja determinantin välinen yhteys seuraa näistä perustuloksista kuten reaalisessa tuloksessa. Perustelu on seuraava. Determinantin sarakelineaarisuudesta ja alternoivuudesta saadaan seuraava korollaari.

Korollaari 13.5.9. *Olkoon* $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{C}^{n\times n}$ *neliömatriisi, jonka sarakkeet ovat lineaarisesti riippuvia. Tällöin* $\det A = 0$.

Toisaalta determinantin tulokaavasta ja tiedosta $\det I = 1$ saadaan seuraava korollaari.

Korollaari 13.5.10. *Olkoon* $A \in \mathbb{C}^{n\times n}$ *kääntyvä matriisi. Tällöin* $\det A^{-1} = (\det A)^{-1}$. *Erityisesti* $\det A \neq 0$.

Lauseen 13.5.2 ehtojen 1 ja 5 yhtäpitävyys seuraa nyt näistä havainnoista. Tarkemmat yksityiskohdat jätetään harjoitustehtäväksi.

13.5.2 Reaalisista matriiseista kompleksisiin

Reaalikertoimiset matriisit $\mathbb{R}^{m\times n}$ voidaan tulkita luonnollisella tavalla kompleksikertoimisina matriiseina, sillä jokainen $x \in \mathbb{R}$ on myös kompleksiluku $x + i\cdot 0 \in \mathbb{C}$, eli $\mathbb{R} \subset \mathbb{C}$.

Koska $\mathbb{R} \subset \mathbb{C}$, niin reaalikertoiminen matriisi $A \in \mathbb{R}^{m\times n}$ on samalla myös kompleksikertoiminen matriisi $A \in \mathbb{C}^{m\times n}$, eli $\mathbb{R}^{m\times n} \subset \mathbb{C}^{m\times n}$. Kuten jatkossa huomataan, tätä havaintoa tarvitaan, kun tarkastellaan matriisin ominaisarvoja.

Huomautus 13.5.11. *Huomaa, että matriisin* $A \in \mathbb{R}^{m\times n}$ *tulkitseminen kompleksikertoimisena tarkoittaa, että ajattelemme sitä sekä* $\mathbb{R}$*-lineaarikuvauksen* $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ *että* $\mathbb{C}$*-lineaarikuvauksen* $\varphi_A\colon \mathbb{C}^{n\times 1} \to \mathbb{C}^{m\times 1}$ *esitysmatriisina.*

13.5.3 Kompleksinen vektoriavaruus reaalisena avaruutena

Koska kompleksiluvut ovat avaruuden $\mathbb{R}^2$ vektoreita, niin

$$\mathbb{C}^n = \mathbb{C} \times \cdots \times \mathbb{C} = \mathbb{R}^2 \times \cdots \times \mathbb{R}^2 = \mathbb{R}^{2n}$$

eli jokainen avaruuden $\mathbb{C}^n$ vektori $(z_1, \ldots, z_n)$, missä $z_j = x_j + iy_j$ jokaisella $j \in \{1, \ldots, n\}$, voidaan tulkita vektorina $(x_1, y_1, \ldots, x_n, y_n)$. Formaalisti tämä tulkinta tarkoittaa bijektiota

$$\rho\colon \mathbb{C}^n \to \mathbb{R}^{2k}, \quad (x_1 + iy_1, \ldots, x_n + iy_n) \mapsto (x_1, y_1, \ldots, x_n, y_n).$$

Tämä bijektio puolestaan on lineaarinen isomorfismi, jos avaruus $\mathbb{C}^n$ tulkitaan reaalisena vektoriavaruuteena, jonka laskutoimitukset ovat $+\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^n$ ja $\cdot\colon \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ eli skalaarikertolasku rajoitetaan vain reaaliluvuille. Näin saadaan, että kompleksinen vektoriavaruus $\mathbb{C}^n$, joka on kompleksista dimensiota n, on samalla myös reaalinen vektoriavaruus, jonka (reaalinen) dimensio on $2n$.

Vastaava päättely voidaan tehdä kaikille kompleksisille vektoriavaruuksille: Kompleksinen vektoriavaruus V on samalla myös reaalinen vektoriavaruus. Lisäksi, jos V on äärellisulotteinen, niin sen dimensio reaalisena vektoriavaruutena on kaksi kertaa sen dimensio kompleksisena avaruutena.

Myös kompleksiset lineaarikuvaukset $f\colon V \to W$ kompleksisten vektoriavaruuksien välillä voidaan tulkita vastaavalla tavalla (reaalisina) lineaarikuvauksina (reaalisten) vektoriavaruuksien välillä.

13.6 Kompleksiset sisätuloavaruudet

Sisätulo yleistyy kompleksikertoimiseen avaruuteen lähes sellaisenaan. Ainoa poikkeus on, että kompleksisissa avaruuksissa sisätulo on kompleksiarvoinen ja konjugaattisymmetrinen.

Määritelmä 13.6.1. *Olkoon V kompleksinen vektoriavaruus. Funktio $\langle\cdot,\cdot\rangle\colon V\times V\to\mathbb{C}$ on* avaruuden V sisätulo*, jos kaikilla $u,v,w\in V$ ja $\lambda\in\mathbb{C}$ pätee*

1. *(positiivisuus) $\langle v,v\rangle\in\mathbb{R}$ ja $\langle v,v\rangle\geq 0$,*
2. *(definiittisyys) $\langle v,v\rangle=0$, jos ja vain jos $v=0$,*
3. *(additiivisuus) $\langle u+v,w\rangle=\langle u,w\rangle+\langle v,w\rangle$,*
4. *(homogeenisuus) $\langle\lambda v,w\rangle=\lambda\langle v,w\rangle$ ja*
5. *(konjugaattisymmetrisyys) $\langle w,v\rangle=\overline{\langle v,w\rangle}$.*

Paria $(V,\langle\cdot,\cdot\rangle)$ kutsutaan kompleksiseksi sisätuloavaruudeksi*.*

Esimerkki 13.6.2. *Funktio $\cdot\colon\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C}$,*

$$(z_1,\ldots,z_n)\cdot(w_1,\ldots,w_n)=z_1\overline{w_1}+\cdots+z_n\overline{w_n},$$

on avaruuden $\mathbb{C}^n$ sisätulo. Kuten reaalisessa tapauksessa kutsutaan pistetuloksi. Sarakeavaruuden $\mathbb{C}^{n\times 1}$ pistetulo määritellään vastaavasti.

Kuten reaalisessa tapauksessa, funktio $\|\cdot\|\colon V\to\mathbb{R}$, $v\mapsto\sqrt{\langle v,v\rangle}$, on normi. Näin ollen myös kompleksisissa avaruuksissa voidaan määritellä ortonormaali kanta.

Määritelmä 13.6.3. *Kompleksisen sisätuloavaruuden $(V,\langle\cdot,\cdot\rangle)$ kanta $(v_1,\ldots,v_n)$ on* ortonormaali*, jos $\langle v_j,v_i\rangle=\delta_{ji}$ kaikilla $j,i\in\{1,\ldots,n\}$.*

Gram–Schmidtin lauseen todistus toimii sellaisenaan myös kompleksisille sisätuloavaruuksille, joten jokaisella äärellisulotteisella kompleksisella sisätuloavaruudella on ortonormaali kanta.

Lause 13.6.4 (Gram–Schmidt)**.** *Olkoon V äärellisulotteinen kompleksinen sisätuloavaruus. Tällöin on olemassa avaruuden V ortonormaali kanta.*

Aliavaruuden kohtisuorakomplementti määritellään kuten reaalisessakin tilanteessa. Kohtisuorakomplementti osoitetaan aliavaruudeksi kuten reaalisessa tapauksessa.

Määritelmä 13.6.5. *Kompleksisen sisätuloavaruuden $(V,\langle\cdot,\cdot\rangle)$ (kompleksisen) aliavaruuden $W\subset V$* kohtisuorakomplementti *on aliavaruus*

$$W^\perp=\{v\in V\colon\langle v,w\rangle=0 \text{ kaikilla } w\in W\}.$$

Kuten reaalisten aliavaruuksien tilanteessa, aliavaruuden $W \subset V$ kohti suorakomplementti $W^\perp$ voidaan selvittää aliavaruuden W virittäjien avulla. Kirjoitetaan tämä havainto esimerkin muotoon.

Esimerkki 13.6.6. *Olkoon $W = \mathrm{Sp}(w_1, \dots, w_k) \subset \mathbb{C}^{n\times 1}$ aliavaruus. Tällöin sisätulon lineaariisuuden nojalla $v \in V$ on kohtisuorassa aliavaruutta W vastaan, jos ja vain jos $w_i \cdot v = 0$ jokaisella $i \in \{1, \dots, k\}$. Koska*

$$w_i \cdot v = w_i^t \overline{v} = \overline{\overline{w_i}^t v}$$

jokaisella $i \in \{1, \dots, k\}$, niin saadaan, että $v \in W^\perp$, jos ja vain jos $\overline{A}^t v = 0$, missä $A = \begin{bmatrix} w_1 & \cdots & w_k \end{bmatrix}$.

13.7 Unitaariset ja hermiittiset matriisit

Edellinen esimerkki nostaa esiin matriisin konjugaattitranspoosin roolin kompleksisessa teoriassa. Sen avulla voidaan määritellä symmetrisen matriisin ja ortogonaalisen matriisin vastineet.

Määritelmä 13.7.1. *Matriisin $A \in \mathbb{C}^{m\times n}$* konjugaattitranspoosi *on matriisi $A^* \in \mathbb{C}^{n\times m}$, joka on määritelty kaavalla*

$$A^* = \overline{A}^T,$$

eli kaavalla

$$(A^*)_{ji} = \overline{A_{ij}}.$$

kaikilla $j \in \{1, \dots, n\}$ ja $i \in \{1, \dots, n\}$.

Palautetaan mieleen, että reaalisen matriisin $A \in \mathbb{R}^{m\times n}$ tapauksessa transpoosille pätee

$$(Av) \cdot w = v \cdot (A^t w)$$

kaikilla $v \in \mathbb{R}^{n\times 1}$ ja $w \in \mathbb{R}^{m\times 1}$. Tämä ominaisuus selittää konjugaatin konjugaattitranspoosin käsitteessä. Seuraavan lemman todistus jätetään harjoitustehtäväksi.

Lemma 13.7.2. *Olkoon $A \in \mathbb{C}^{m\times n}$. Tällöin kaikilla $v \in \mathbb{C}^{n\times 1}$ ja $w \in \mathbb{C}^{m\times 1}$ pätee*

$$(Av) \cdot w = v \cdot (A^* w).$$

Konjugaattitranspoosin avulla voidaan määritellä symmetristen ja ortgonaalisten matriisien kompleksiset vastineet. Symmetrisen matriisien kompleksinen vastine on hermiittinen matriisi.

Määritelmä 13.7.3. *Matriisi $U \in \mathbb{C}^{n\times n}$ on* hermiittinen*, jos $U^* = U$.*

Ortogonaalisen matriisin kompleksinen vastine on puolestaan unitaarinen matriisi.

Määritelmä 13.7.4. *Matriisi* $U \in \mathbb{C}^{m\times n}$ *on* unitaarinen, *jos* $U^*U = I$.

Kuten reaalisessa tapauksessa unitaarimatriisin $U \in \mathbb{C}^{m\times n}$ sarakkeet ovat lineaarisesti riippumattomia. Erityisesti unitaarisen neliömatriisin $U \in \mathbb{C}^{n\times n}$ sarekkeet muodostavat avaruuden $\mathbb{C}^{n\times 1}$ ortonormaalin kannan. Kirjataan tämä lauseeksi, jonka todistus jätetään harjoitustehtäväksi.

Lause 13.7.5. *Olkoon* $U = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \in \mathbb{C}^{n\times n}$ *matriisi. Tällöin* U *on unitaarinen, jos ja vain jos* $(u_1, \ldots, u_n)$ *on avaruuden* $(\mathbb{C}^{n\times n}, \cdot)$ *ortonormaalikanta. Lisäksi tällöin* U *on kääntyvä ja* $U^{-1} = U^*$.

Unitaarimatriiseilla on sama rooli kompleksisten isometrioiden teoriassa kuin mikä ortogonaalimatriiseilla on reaalisten isometrioiden teoriassa.

Määritelmä 13.7.6. *Lineaarikuvaus* $f\colon V \to W$ *kompleksisisesta sisätuloavaruudesta* $(V, \langle\cdot,\cdot\rangle_V)$ *kompleksiseen sisätuloavaruuteen* $(W, \langle\cdot,\cdot\rangle_W)$ *on* isometria, *jos*

$$\langle f(v), f(v')\rangle_W = \langle v, v'\rangle_V$$

kaikilla $v, v' \in V$.

Lause 13.7.7. *Olkoon* $(v_1, \ldots, v_n)$ *kompleksisen sisätuloavaruuden* $(V, \langle\cdot,\cdot\rangle_V)$ *kanta ja* $(w_1, \ldots, w_m)$ *kompleksisen sisätuloavaruuden* $(W, \langle\cdot,\cdot\rangle_W)$ *kanta. Tällöin lineaarikuvaus* $f\colon V \to W$ *on isometria, jos ja vain jos sen esitysmatriisi* $U \in \mathbb{C}^{m\times n}$ *kannasta* $(v_1, \ldots, v_n)$ *kantaan* $(w_1, \ldots, w_m)$ *on unitaarimatriisi.*

13.8 Kompleksiset operaattorit

Kuten reaalisten vektoriavaruuksien tapauksessa, kompleksista lineaarikuvausta avaruudesta itseensä kutsutaan operaattoriksi. Kompleksisen operaattorin ominaisarvot, ominaisvektorit ja ominaisavaruudet määritellään kuten reaalisessa tapauksessa.

Määritelmä 13.8.1. *Kompleksista lineaarikuvausta* $\varphi\colon V \to V$ *kompleksisesta vektoriavaruudesta itseseensä kutsutaan* (kompleksiseksi) operaattoriksi.

Määritelmä 13.8.2. *Nollasta poikkeava vektori* $v \in V$ *on kompleksisen operaattorin* $\varphi\colon V \to V$ ominaisvektori ominaisarvolla $\lambda \in \mathbb{C}$, *jos* $\varphi(v) = \lambda v$.

Määritelmä 13.8.3. *Kompleksisen operaattorin* $\varphi\colon V \to V$ *ominaisarvoa* $\lambda \in \mathbb{C}$ *vastaava* ominaisavaruus *on*

$$E(\lambda, \varphi) = \{v \in V : \varphi(v) = \lambda v\}.$$

Kompleksisen operaattorin ominaisarvot ja ominaisvektorit löydetään kuten reaalisessa tapauksessa. Kirjataan tätä varten vielä formaalisti neliömatriisin $A \in \mathbb{C}^{n\times n}$ ominaisarvon ja ominaisvektorin määritelmä.

Määritelmä 13.8.4. *Nollasta poikkeava vektori* $z \in \mathbb{C}^{n\times 1}$ *on matriisin* $A \in \mathbb{C}^{n\times n}$ ominaisvektori ominaisarvolla $\lambda \in \mathbb{C}$, *jos* $Az = \lambda z$.

Kompleksisen neliömatriisin ominaisarvot voidaan selvittää determinanttiyhtälöstä kuten reaalisessa tapauksessa. Kirjataan tämä lemmaksi. Todistus jätetään lukijalle hajoitustehtäväksi.

Lemma 13.8.5. *Luku $\lambda \in \mathbb{C}$ on matriisin $A \in \mathbb{C}^{n\times n}$ ominaisarvo, jos ja vain jos $\det(A - \lambda I) = 0$.*

Samalla argumentilla kuin luvussa 11.2 havaitaan, että kompleksisen operaattorin ominaisarvot ovat samat kuin sen esitysmatriisin ominaisarvot, kunhan operaattori esitetään käyttäen yhtä kantaa.

Lause 13.8.6. *Olkoon $(v_1, \dots, v_n)$ avaruuden V kanta, $\varphi\colon V \to V$ operaattori ja A operaattorin φ esitysmatriisi kannassa $(v_1, \dots, v_n)$. Tällöin luku $\lambda \in \mathbb{C}$ on operaattorin φ ominaisarvo, jos ja vain jos se on esitysmatriisin A ominaisarvo.*

13.9 Ositetut matriisit

Ositettuja matriiseja sivuttiin luvussa 2, kun matriisi kirjoitettiin joko sarakkeiden tai rivien avulla.

Esimerkki 13.9.1. *Olkoon*

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix}.$$

Tällöin matriisin A sarakkeet ovat

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \; v_2 = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \; v_3 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}, \; v_3 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}, \; \textit{ja } v_4 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}.$$

Tällöin

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} & \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} & \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}.$$

Vastaavasti matriisin A rivit ovat

$$w_1 = \begin{bmatrix} 1 & 1 & 2 & 2 \end{bmatrix},$$
$$w_2 = \begin{bmatrix} 1 & 1 & 2 & 2 \end{bmatrix}$$

ja

$$w_3 = \begin{bmatrix} 3 & 3 & 4 & 4 \end{bmatrix}.$$

Tällöin

$$A = \begin{bmatrix} \begin{bmatrix} 1 & 1 & 2 & 2 \end{bmatrix} \\ \begin{bmatrix} 1 & 1 & 2 & 2 \end{bmatrix} \\ \begin{bmatrix} 3 & 3 & 4 & 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}.$$

Yleisemmin matriisi voidaan osittaa alimatriiseihin.

Esimerkki 13.9.2. *Edellisen esimerkin matriisi voidaan osittaa esimerkiksi seuraavasti*

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 3 & 3 \end{bmatrix} & \begin{bmatrix} 2 & 2 \\ 2 & 2 \\ 4 & 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \\ \begin{bmatrix} 3 & 3 \end{bmatrix} & \begin{bmatrix} 4 & 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 3 & 3 & 4 \end{bmatrix} & \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} A_3 & B_3 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{bmatrix} \\ \begin{bmatrix} 3 & 3 & 4 & 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} A_4 \\ B_4 \end{bmatrix}.$$

Määritelmä 13.9.3. *Matriisin $A \in \mathbb{C}^{m \times n}$ ositus on sellainen taulukko*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rp} \end{bmatrix}, \tag{13.4}$$

matriiseja A_{ji}, että on olemassa sellaiset luvut $D_1, \ldots, D_r \in \mathbb{N}$ ja $d_1, \ldots, d_p \in \mathbb{N}$, että $A_{ji} \in \mathbb{C}^{D_j \times d_i}$. Taulukkoa (13.4) kutsutaan ositetuksi matriisiksi *ja ja matriiseja A_{ji}* lohkoiksi.

Yksi perustelu tälle määritelmälle on seuraava algebrallinen matriisien kertolaskuun liittyvä syy. Toinen perustelu tälle määritelmälle annetaan suorien summien ja kuvauksien avulla luvun lopussa. Koska tulos seuraa suoraan matriisitulon määritelmästä, se jätetään harjoitustehtäväksi.

Lause 13.9.4. *Olkoot*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rp} \end{bmatrix} \text{ ja } B = \begin{bmatrix} B_{11} & \cdots & B_{1q} \\ \vdots & & \vdots \\ B_{p1} & \cdots & B_{pq} \end{bmatrix}$$

ositettuja matriiseja. Jos jokaisella $j \in \{1, \ldots, r\}$ ja $i \in \{1, \ldots, q\}$ tulomatriisi $A_{jk}B_{ki}$ on määritelty jokaisella $k \in \{1, \ldots, p\}$. Tällöin tulomatriisi AB on ositettu matriisi

$$AB = \begin{bmatrix} C_{11} & \cdots & C_{1q} \\ \vdots & & \vdots \\ C_{r1} & \cdots & C_{rq} \end{bmatrix},$$

missä

$$C_{ji} = A_{j1}B_{1i} + A_{j2}B_{2i} + \cdots + A_{jp}B_{pi}.$$

Esimerkki 13.9.5. *Olkoon*

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \\ \begin{bmatrix} 3 & 3 \end{bmatrix} & \begin{bmatrix} 4 & 4 \end{bmatrix} \end{bmatrix}.$$

Tällöin

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \\ \begin{bmatrix} 3 & 3 \end{bmatrix} & \begin{bmatrix} 4 & 4 \end{bmatrix} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \\ \begin{bmatrix} 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 & 4 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \end{bmatrix}.$$

Ositettujen matriisien yhteenlasku tapahtuu aivan kuten voisi olettaa.

Lemma 13.9.6. *Olkoot*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rp} \end{bmatrix} \in \mathbb{C}^{n\times n} \quad ja \quad B = \begin{bmatrix} B_{11} & \cdots & B_{1p} \\ \vdots & & \vdots \\ B_{r1} & \cdots & B_{rp} \end{bmatrix} \in \mathbb{C}^{n\times n}$$

osoitettuja matriiseja. Jos matriisiyhteenlaskut $A_{ji} + B_{ji}$ ovat hyvin määriteltyjä, niin

$$aA + B = \begin{bmatrix} aA_{11} + B_{11} & \cdots & aA_{1p} + B_{1p} \\ \vdots & & \vdots \\ aA_{r1} + B_{r1} & \cdots & aA_{rp} + B_{rp} \end{bmatrix}$$

kaikilla $a \in \mathbb{C}$.

Huomautus 13.9.7. *Huomaa, että on helppo osittaa matriisit A ja B sellaisilla tavoilla, että yhteenlaskua ei voi suorittaa lohkojen avulla.*

13.9.1 Ositetun matriisin kääntyvyys

Koska ositettu matriisi on matriisi myös tavallisessa mielessä, niin sen kääntyvyyttä koskevat samat lauseet kuin tavallisia matriiseja. Näistä esimerkkeinä lauseet 2.5.1 ja 3.10.1.

Ositetun matriisin

$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} \in \mathbb{R}^{n\times n}$$

käänteismatriisille ei ole yleistä kaavaa, joka perustuisi matriiseihin A, B, C ja D. Tähän on monia syitä. Yksi sellainen on, että yhdenkään alimatriisin ei tarvitse olla kääntyvä vaikka M olisi kääntyvä. Tälläinen tapaus on matriisi

$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Ositetun matriisin kääntyvyys voidaan kuitenkin joissain tapauksissa lukea suoraan lohkoista. Aloitetaan tarkastelu ositetuista diagonaalimatriiseista.

Määritelmä 13.9.8. *Ositettu matriisi*

$$M = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rp} \end{bmatrix}$$

on ositettu diagonaalimatriisi, *jos jokainen matriiseista A_{jj} on neliömatriisi ja $A_{ji} = 0$ kaikilla $j \neq i$.*

Esimerkki 13.9.9. *Olkoot A_1 ja A_2 neliömatriiseja. Tällöin matriisi*

$$M = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

on ositettu diagonaalimatriisi.

Seuraava lemma seuraa suoraan esimerkiksi lauseesta 2.5.1.

Lemma 13.9.10. *Olkoon*

$$M = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

on ositettu diagonaalimatriisi. Tällöin M on kääntyvä, jos ja vain jos matriisit A_1 ja A_2 ovat kääntyviä. Mikäli M on kääntyvä, niin

$$M^{-1} = \begin{bmatrix} A_1^{-1} & 0 \\ 0 & A_2^{-1} \end{bmatrix}.$$

Todistus. Oletetaan, että $M \in \mathbb{C}^{n \times n}$ on kääntyvä ja osoitetaan, että matriisit $A_1 \in \mathbb{C}^{n_1 \times n_1}$ ja $A_2 \in \mathbb{C}^{n_2 \times n_2}$ ovat kääntyviä.

Olkoot $v_1 \in \mathbb{C}^{n_1 \times 1}$ ja $v_2 \in \mathbb{C}^{n_2 \times 1}$. Tällöin

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{C}^{n \times 1}.$$

Koska M on kääntyvä, niin $Mv = 0$ jos ja vain jos $v = 0$. Koska

$$Mv = \begin{bmatrix} A_1 v_1 \\ A_2 v_2 \end{bmatrix},$$

niin näin ollen saadaan sekä $A_1 v_1 = 0$, jos ja vain jos $v_1 = 0$, että $A_2 v_2 = 0$, jos ja vain jos $v_2 = 0$. Näin ollen sekä A_1 että A_2 ovat kääntyviä.

Oletetaan nyt, että A_1 ja A_2 ovat kääntyviä. Tällöin vastaavalla argumentilla havaitaan, että M on kääntyvä.

Molemmissa tapauksissa käänteismatriisin M^{-1} kaava seuraa ositettujen matriisin tulokaavasta. □

Vastaavalla tavalla voidaan tarkastella ositettujen yläkolmiomatriisien kääntyvyyttä. Koska tätä termiä ei tarvita tämän luvun jälkeen, ei sitä erikseen määritellä tai käytetä seuraavassa lemmassa. Lemman todistus jätetään harjoitustehtäväksi.

Lemma 13.9.11. *Olkoon*
$$M = \begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}.$$
ositettu matriisi, jonka lohkot A_1 ja A_2 ovat neliömatriiseja. Tällöin M on kääntyvä, jos ja vain jos A_1 ja A_2 ovat kääntyviä. Mikäli M on kääntyvä, niin
$$M^{-1} = \begin{bmatrix} A_1^{-1} & -A_1^{-1}BA_2^{-1} \\ 0 & A_2^{-1} \end{bmatrix}.$$

Huomautus 13.9.12. *Vastaava tulos pätee luonnollisesti myös ositetuille alakolmiomatriiseille*
$$M = \begin{bmatrix} A_1 & 0 \\ B & A_2 \end{bmatrix}.$$

Annetaan nyt välttämätön ja riittävä ehto ositetun matriisin
$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$$
kääntyvyydelle erikoistapauksessa, että lohko A on kääntyvä. Yleinen tapaus seuraa tästä erikoistapauksesta kannanvaihdoilla.

Lause 13.9.13. *Olkoon*
$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} \in \mathbb{R}^{n\times n}$$
ositettu matriisi, jonka alimatriisi $A \in \mathbb{R}^{k\times k}$ on kääntyvä neliömatriisi. Tällöin matriisi M on kääntyvä, jos ja vain jos matriisi $E = D - BA^{-1}C$ on kääntyvä. Lisäksi tällöin
$$M^{-1} = \begin{bmatrix} (A^{-1} + A^{-1}CE^{-1}BA^{-1}) & (-A^{-1}CE^{-1}) \\ (-E^{-1}BA^{-1}) & E^{-1} \end{bmatrix}. \tag{13.5}$$

Lauseen ytimessä on ositetun matriisin hajoittaminen tuloksi, joka tehdään ensin. Lemman todistus on suora lasku, joten se jätetään lukijalle.

Lemma 13.9.14. *Olkoon*
$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} \in \mathbb{R}^{n\times n}$$
ositettumatriisi, jonka alimatriisi $A \in \mathbb{R}^{k\times k}$ on kääntyvä neliömatriisi. Tällöin
$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ BA^{-1} & D - BA^{-1}C \end{bmatrix} \begin{bmatrix} I & C \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

Huomautus 13.9.15. *Puuttuvan matriisin löytää myös kertomalla matriisia M ensin matriisilla*
$$\begin{bmatrix} A^{-1} & 0 \\ 0 & I \end{bmatrix}$$
ja sitten matriisilla
$$\begin{bmatrix} I & -C \\ 0 & I \end{bmatrix}.$$

Lauseen 13.9.13 todistus. Olkoon

$$M' = \begin{bmatrix} I & 0 \\ BA^{-1} & D - BA^{-1}C \end{bmatrix}.$$

Lemman 13.9.14 nojalla

$$M = \begin{bmatrix} I & 0 \\ BA^{-1} & D - BA^{-1}C \end{bmatrix} \begin{bmatrix} I & C \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} = M' \begin{bmatrix} I & C \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix},$$

missä matriisit

$$\begin{bmatrix} I & C \\ 0 & I \end{bmatrix} \quad \text{ja} \quad \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

ovat kääntyviä.

Oletetaan ensin, että matriisi $D - BA^{-1}C$ on kääntyvä. Huomautuksen 13.9.12 perusteella matriisi M' on tällöin kääntyvä. Näin ollen matriisi M on kääntyvien matriisien tulona kääntyvä.

Oletetaan nyt, että matriisi M on kääntyvä. Koska matriisit

$$\begin{bmatrix} I & C \\ 0 & I \end{bmatrix} \quad \text{ja} \quad \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

ovat kääntyviä, niin matriisi M' on kääntyvä. Koska matriisi M' on kääntyvä, niin lemman 13.9.11 perusteella matriisi $D - BA^{-1}C$ on kääntyvä.

Osoitetaan vielä, että käänteismatriisilla M^{-1} on haluttu kaava (13.5). Koska matriisi M on kääntyvä, niin myös $E = D - BA^{-1}C$ on kääntyvä. Näin ollen

$$\begin{aligned} M^{-1} = \begin{bmatrix} A & C \\ B & D \end{bmatrix}^{-1} &= \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} I & C \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ BA^{-1} & D - BA^{-1}C \end{bmatrix}^{-1} \\ &= \begin{bmatrix} A^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -C \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ BA^{-1} & E \end{bmatrix}^{-1}, \end{aligned}$$

missä

$$\begin{aligned} \begin{bmatrix} I & 0 \\ BA^{-1} & E \end{bmatrix}^{-1} &= \left(\begin{bmatrix} I & 0 \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & E \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} I & 0 \\ 0 & E \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ BA^{-1} & I \end{bmatrix}^{-1} \\ &= \begin{bmatrix} I & 0 \\ 0 & E^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -BA^{-1} & I \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ -E^{-1}BA^{-1} & E^{-1} \end{bmatrix}. \end{aligned}$$

Näin ollen pätee

$$\begin{aligned}
M^{-1} &= \begin{bmatrix} A^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -C \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ BA^{-1} & E \end{bmatrix}^{-1} \\
&= \begin{bmatrix} A^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -C \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -E^{-1}BA^{-1} & E^{-1} \end{bmatrix} \\
&= \begin{bmatrix} A^{-1} & -A^{-1}C \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -E^{-1}BA^{-1} & E^{-1} \end{bmatrix} \\
&= \begin{bmatrix} (A^{-1} + A^{-1}CE^{-1}BA^{-1}) & -A^{-1}CE^{-1} \\ -E^{-1}BA^{-1} & E^{-1} \end{bmatrix}
\end{aligned}$$

□

Huomautus 13.9.16. *Edellisestä todistuksesta sanottakoon, että mikäli olettaa matriisin E kääntyvyyden, niin matrisiin M kääntyvyyden voi helposti todistaa suoralla kertolaskulla. Tämä ei kuitenkaan juurikaan selitä, miksi väite on totta ja miksi matriisin E kääntyvyys on välttämätön ehto matriisin M kääntyvyydelle.*

Huomautus 13.9.17. *Kuten ennen lauseen 13.9.13 väitettä mainittiin, mikäli ositetun matriisin*

$$M = \begin{bmatrix} A & C \\ B & D \end{bmatrix}$$

jokin lohkoista B, C tai D on kääntyvä, voidaan kysymys matriisin M kääntyvyydestä palauttaa lauseen 13.9.13 tilanteeseen sopivalla kannanvaihdolla. Tällöin tietysti lauseessa esiintyvä matriisi E muuntuu kannanvaihtoa vastaavalla tavalla. (Harjoitustehtävä)

13.9.2 Ositetun matriisin lohkojen yhteys lineaarikuvauksiin

Aloitetaan tarkastelu konkreettisesta esimerkistä.

Esimerkki 13.9.18. *Olkoon $f\colon \mathbb{R}^4 \to \mathbb{R}^4$ lineaarikuvaus $(x, y, z, w) \mapsto (2x+3y, y, w, z)$. Tällöin kuvauksen f esitysmatriisi standardikannassa on*

$$A = \begin{bmatrix} 2 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Kuvauksen f esitysmatriisi on siis ositettu matriisi

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

missä

$$A_1 = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} \quad \textit{ja} \quad A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Toisin kuin matriisi A matriisit A_1 ja A_2 ovat kaksiulotteisten vektoriavaruuksien välisten lineaarikuvausten $f_1\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x_1, x_2) \mapsto (2x_1 + 3x_2, x_2)$ ja $f_2\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x_1, x_2) \mapsto (x_2, x_1)$ esitysmatriiseja. Huomaa, että tässä tarkoituksella muutettiin koordinaattien symboleita.

Edellinen esimerkki herättää kysymyksen kuinka lineaarikuvaukset $f_1\colon \mathbb{R}^2 \to \mathbb{R}^2$ ja $f_2\colon \mathbb{R}^2 \to \mathbb{R}^2$ liittyvät lineaarikuvaukseen f, jonka lähtö- ja maaliavaruudet ovat neliulotteisia. Selvitetään yhteys kahdessa vaiheessa.

Esimerkki 13.9.19. *Olkoot $f\colon \mathbb{R}^4 \to \mathbb{R}^4$, $f_1\colon \mathbb{R}^2 \to \mathbb{R}^2$ ja $f_2\colon \mathbb{R}^2 \to \mathbb{R}^2$ lineaarikuvauksia kuten edellisessä esimerkissä.*

Olkoon nyt $\iota_1\colon \mathbb{R}^2 \to \mathbb{R}^4$ lineaarikuvaus $(x_1, x_2) \mapsto (x_1, x_2, 0, 0)$ ja olkoon $\iota_2\colon \mathbb{R}^2 \to \mathbb{R}^4$ lineaarikuvaus $(x_1, x_2) \mapsto (0, 0, x_1, x_2)$.

Koska yhdistetty kuvaus $f \circ \iota_1\colon \mathbb{R}^2 \to \mathbb{R}^4$ on kuvaus $(x_1, x_2) \mapsto (2x_1 + 3x_2, x_2, 0, 0)$, niin saadaan, että kuvauksen $f \circ \iota_1$ esitysmatriisi standardikannasta standardikantaan on

$$\begin{bmatrix} 2 & 3 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A_1 \\ 0 \end{bmatrix}$$

eli kuvauksen f esitysmatriisin kaksi ensimmäistä saraketta. Vastaavasti saadaan, että yhdistetyn kuvauksen $f \circ \iota_2\colon \mathbb{R}^2 \to \mathbb{R}^4$ eli kuvauksen $(x_1, x_2) \mapsto (0, 0, x_2, x_1)$ esitysmatriisi on

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ A_2 \end{bmatrix}$$

eli kuvauksen f esitysmatriisin kaksi viimeistä saraketta.

Edellä on selvitetty kuinka matriisin A sarakkeet voidaan tunnistaa lineaarikuvausten avulla. Edellisestä esimerkistä voidaan tehdä huomio, että lineaarikuvauksen ι_1 kuva on aliavaruus $V_1 = \{(x_1, x_2, 0, 0) \in \mathbb{R}^4 \colon x_1, x_2 \in \mathbb{R}\}$ ja lineaarikuvauksen ι_2 kuva on aliavaruus $V_2 = \{(0, 0, x_3, x_4) \in \mathbb{R}^4 \colon x_3, x_4 \in \mathbb{R}\}$, joille pätee $\mathbb{R}^4 = V_1 \oplus V_2$.

Käytetään nyt projektioita matriisien A_1 ja A_2 erottamiseksi sarakkeista.

Esimerkki 13.9.20. *Olkoot $f\colon \mathbb{R}^4 \to \mathbb{R}^4$, $f_1\colon \mathbb{R}^2 \to \mathbb{R}^2$, $f_2\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\iota_1\colon \mathbb{R}^2 \to \mathbb{R}^4$ ja $\iota_2\colon \mathbb{R}^2 \to \mathbb{R}^4$ lineaarikuvauksia kuten edellisessä esimerkissä.*

Olkoot $p_1\colon \mathbb{R}^4 \to \mathbb{R}^2$, $(y_1, y_1, y_3, y_4) \mapsto (y_1, y_2)$ ja $p_2\colon \mathbb{R}^4 \to \mathbb{R}^2$, $(y_1, y_2, y_3, y_4) \mapsto (y_3, y_4)$, lineaarikuvauksia. Tällöin

$$(p_1 \circ f \circ \iota_1)(x_1, x_2) = p_1(f(x_1, x_2, 0, 0)) = p_1(2x_1 + 3x_2, x_2, 0, 0) = (2x_1 + 3x_2, x_2) = f_1(x_1, x_2)$$

ja

$$(p_2 \circ f \circ \iota_2)(x_1, x_2) = p_2(f(0, 0, x_1, x_2)) = p_2(0, 0, x_2, x_1) = (x_2, x_1) = f_2(x_1, x_2)$$

kaikilla $x_1, x_2 \in \mathbb{R}$. Näin ollen

$$f_1 = p_1 \circ f \circ \iota_1$$

ja

$$f_2 = p_2 \circ f \circ \iota_2.$$

Matriisit A_1 ja A_2 ovat siis yhdistettyjen kuvausten $p_1 \circ f \circ \iota_1$ ja $p_2 \circ f \circ \iota_2$ esitysmatriisit. Vastaavat laskut osoittavat, että kuvaukset $p_j \circ f \circ \iota_i$ ovat nollakuvauksia, jos $j \neq i$.

Edellisestä esimerkistä voi myös huomata, että mikäli aliavaruudet V_1 ja V_2 ovat kuten edellä, niin $\ker p_1 = V_2$ ja $\ker p_2 = V_1$. Lisäksi rajoittumat $p_1|_{V_1} \colon V_1 \to \mathbb{R}^2$ ja $p_2|_{V_2} \colon V_2 \to \mathbb{R}^2$ ovat isomorfismeja. Tämä ei ole sattumaa.

Tästä esimerkistä voi päätellä, että myös yleisesti ositetun matriisin lohkot liittyvät sellaisiin lineaarikuvauksiin, jotka liittyvät alkuperäisen lineaarikuvauksen lähtö- ja maaliavaruuden esityksiin suorina summina. Tarkastellaan tätä nyt yleisten ositettujen matriisien tapauksessa.

Yleinen tapaus

Tarkastellaan nyt tätä ilmiötä yleisesti tarkastelemalla ositettua matriisia

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rp} \end{bmatrix} \in \mathbb{C}^{m \times n},$$

missä $A_{ji} \in \mathbb{C}^{D_j \times d_i}$ kaikilla $j \in \{1, \ldots, r\}$ ja $i \in \{1, \ldots, p\}$.

Määritellään kantalauseen avulla kuvaukset

$$\iota_i \colon \mathbb{C}^{d_i \times 1} \to \mathbb{C}^{n \times 1}, e_k \mapsto e_{(d_1 + \cdots + d_{i-1}) + k},$$

ja

$$p_j \colon \mathbb{C}^{m \times 1} \to \mathbb{C}^{D_j \times 1}, e_\ell \mapsto \begin{cases} e_{\ell - (D_1 + \cdots + D_{j-1})}, & D_1 + \cdots + D_{j-1} < \ell \leq D_1 + \cdots + D_j \\ 0, & \text{muutoin.} \end{cases}$$

jokaisella $i \in \{1, \ldots, p\}$ ja $j \in \{1, \ldots, r\}$. Huomaa, että jälleen standardikannan alkiota merkitään samoilla symboleilla kaikissa sarekeavaruudessa. Erityisesti siis

$$p_j\left(\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}\right) = \begin{bmatrix} y_{D_1 + \cdots + D_{j-1} + 1} \\ \vdots \\ y_{D_1 + \cdots + D_j} \end{bmatrix}$$

jokaisella $y \in \begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix}^t$.

Tällöin kaikilla $j \in \{1, \ldots, r\}$ ja $i \in \{1, \ldots, p\}$ yhdistetty kuvaus $p_j \circ f_A \circ \iota_i$ kuvaa sarakeavaruuden $\mathbb{C}^{d_i \times 1}$ sarakeavaruuteen $\mathbb{C}^{D_j \times 1}$.

Lemma 13.9.21. *Olkoon $A \in \mathbb{C}^{m\times n}$ ja olkoot $A_{ji} \in \mathbb{C}^{D_j\times d_j}$ kaikilla $j \in \{1,\dots,r\}$ ja $i \in \{1,\dots,p\}$. Tällöin*

$$p_j \circ \varphi_A \circ \iota_i = \varphi_{A_{ji}}$$

jos ja vain jos

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rp} \end{bmatrix}. \tag{13.6}$$

Todistus. Oletetaan ensin, että A on ositettu matriisi (13.6). Huomataan aluksi, että

$$(A_{ji})_{\ell k} = A_{(D_1+\cdots+D_{j-1}+\ell)(d_1+\cdots+d_{i-1}+k)}$$

kaikilla $j \in \{1,\dots,r\}$, $i \in \{1,\dots,p\}$, $k \in \{1,\dots,d_i\}$ ja $\ell \in \{1,\dots,D_j\}$.

Olkoot nyt $j \in \{1,\dots,r\}$, $i \in \{1,\dots,p\}$, $k \in \{1,\dots,d_i\}$ ja $\ell \in \{1,\dots,D_j\}$. Tällöin

$$\begin{aligned} e_\ell \cdot (p_j \circ \varphi_A \circ \iota_i)(e_k) &= e_\ell \cdot p_j(\varphi_A(e_{d_1+\cdots+d_{i-1}+k})) \\ &= e_\ell \cdot p_j\left(\begin{bmatrix} a_{1(d_1+\cdots+d_{i-1}+k)} \\ \vdots \\ a_{m(d_1+\cdots+d_{i-1}+k)} \end{bmatrix}\right) \\ &= e_\ell \cdot \begin{bmatrix} a_{(D_1+\cdots+D_{j-1}+1)(d_1+\cdots+d_{i-1}+k)} \\ \vdots \\ a_{(D_1+\cdots+D_j)(d_1+\cdots+d_{i-1}+k)} \end{bmatrix} \\ &= a_{(D_1+\cdots+D_{j-1}+\ell)(d_1+\cdots+d_{i-1}+k)} = (A_{ji})_{\ell k}. \end{aligned}$$

Näin ollen

$$p_j \circ \varphi_A \circ \iota_i = \varphi_{A_{ji}}.$$

Oletetaan nyt, että $p_j \circ \varphi_A \circ \iota_i = \varphi_{A_{ji}}$ kaikilla $j \in \{1,\dots,r\}$ ja $i \in \{1,\dots,p\}$. Olkoot nyt $i_0 \in \{1,\dots,n\}$ ja $j_0 \in \{1,\dots,m\}$. Tällöin on olemassa yksikäsitteiset sellaiset $j \in \{1,\dots,r\}$, $\ell \in \{1,\dots,D_j\}$, $i \in \{1,\dots,p\}$ ja $k \in \{1,\dots,d_i\}$ että $j_0 = D_1+\cdots+D_{j-1}+\ell$ ja $i_0 = d_1+\cdots+d_{i-1}+k$. Nyt vastaava lasku kuin yllä osoittaa, että

$$a_{j_0 i_0} = e_{j_0} \cdot \varphi_A(e_{i_0}) = e_{j_0} \cdot (p_j \circ \varphi_A \circ \iota_i)(e_k) = e_\ell \cdot \varphi_{A_{ji}}(e_k) = (A_{ji})_{\ell k}.$$

Näin ollen A on haluttu ositettu matriisi. $\square$

Tarkastellaan luvun lopuksi esitysmatriisien ja aliavaruuksien suorien summien välistä yhteyttä. Aloitetaan kuvauksia $\iota_i \colon \mathbb{C}^{d_i\times 1} \to \mathbb{C}^{n\times 1}$ ja $p_j \colon \mathbb{C}^{D_j\times 1} \to \mathbb{C}^{m\times 1}$ vastaavilla kuvauksilla.

Olkoot V ja W äärellisulotteisia vektoriavaruuksia ja olkoot $V = V_1 \oplus \cdots \oplus V_p$ ja $W = W_1 \oplus \cdots \oplus W_r$ suoria summia. Määritellään jokaisella $i \in \{1,\dots,p\}$ lineaarikuvaus

$$\iota_{V_i} \colon V_i \to V, v \mapsto v,$$

jota kutsutaan aliavaruuden V_i *inkluusio kuvaukseksi*. Määritellään myös kaikilla $j \in \{1, \dots, r\}$ yksikäsitteiset lineaarikuvaukset $p_{W_j}\colon W \to W_j$, joille pätee

$$w = p_{W_1}(w) + \dots + p_{W_r}(w).$$

kaikilla $w \in W$. Huomaa, että p_{W_j} on oleellisesti suoraan summaan $W = W_1 \oplus \dots \oplus W_r$ liittyvä projektio $\pi_{W_j}\colon W \to W$.

Lause 13.9.22. . *Olkoot V ja W äärellisulotteisia vektoriavaruuksia sekä olkoot $V = V_1 \oplus \dots \oplus V_p$ ja $W = W_1 \oplus \dots \oplus W_r$ suoria summia. Olkoot lisäksi $(v_{11}, \dots, v_{d_1 1}, v_{21}, \dots, v_{d_p p})$ ja $(w_{11}, \dots, w_{D_1 1}, w_{21}, \dots, w_{D_r r})$ avaruuksien V ja W sellaisia kantoja, että $(v_{1i}, \dots, v_{d_i i})$ on aliavaruuden V_i kanta ja $(w_{1j}, \dots, w_{D_j j})$ aliavaruuden W_j kanta jokaisella $i \in \{1, \dots, p\}$ ja $j \in \{1, \dots, r\}$. Olkoon lisäksi $\varphi\colon V \to W$ lineaarikuvaus ja $A_{ji} \in \mathbb{C}^{D_j \times d_i}$ lineaarikuvauksen $\varphi_{ji} = p_{W_j} \circ \varphi \circ \iota_{V_i}\colon V_i \to W_j$ esitysmatriisi annetuissa kannoissa. Tällöin kuvauksen φ esitysmatriisi on ositettu matriisi*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rp} \end{bmatrix}.$$

Todistus. Olkoon $(v_1, \dots, v_n)$ annettu avaruuden V kanta ja $(w_1, \dots, w_m)$ annettu avaruuden W kanta. Merkitään $\Psi_V = \Psi_{(v_1, \dots, v_n)}\colon \mathbb{C}^{n \times 1} \to V$ ja $\Psi_W = \Psi_{(w_1, \dots, w_m)}\colon \mathbb{C}^{m \times 1} \to W$. Merkitään myös $\Psi_{V_i} = \Psi_{(v_{1i}, \dots, v_{d_i i})}\colon \mathbb{C}^{d_i \times 1} \to V_i$ ja $\Psi_{W_j} = \Psi_{(w_{1j}, \dots, w_{D_j j})}\colon \mathbb{C}^{D_j \times 1} \to W_j$ jokaisella $i \in \{1, \dots, p\}$ ja $j \in \{1, \dots, r\}$. Tällöin

$$\Psi_V \circ \iota_i = \iota_{V_i} \circ \Psi_{V_i}$$

ja

$$p_{W_j} \circ \Psi_W = \Psi_{W_j} \circ p_j$$

jokaisella $i \in \{1, \dots, p\}$ ja $j \in \{1, \dots, r\}$. Näin ollen

$$\begin{aligned} \varphi_{A_{ji}} &= \Psi_{W_j}^{-1} \circ p_{W_j} \circ \varphi \circ \iota_{V_i} \circ \Psi_{V_i} \\ &= p_j \circ \Psi_W^{-1} \circ \varphi \circ \Psi_V \circ \iota_i \\ &= p_j \circ \varphi_A \circ \iota_i. \end{aligned}$$

Väite seuraa nyt lemmasta 13.9.21. □

Luku 14

Operaattorin yläkolmioesitys

14.1 Kompleksisella operaattorilla on ominaisarvo

Tärkein syy tarkastella kompleksia vektoriavaruuksia on, että jokaisella kompleksisella lineaarikuvauksella $V \to V$ on ominaisarvo.

Lause 14.1.1. *Olkoon V äärellisulotteinen kompleksinen vektoriavaruus, jolle* $\dim V \neq 0$*, ja olkoon $\varphi \colon V \to V$ lineaarikuvaus. Tällöin on lineaarikuvauksella f on ominaisarvo $\lambda \in \mathbb{C}$.*

Yleinen tulos seuraa suoraan vastaavasta tuloksesta neliömatriiseille, kun sitä sovelletaan kuvauksen φ esitysmatriisin. Näin ollen riittää todistaa seuraava lause.

Lause 14.1.2. *Jokaisella neliömatriisilla $A \in \mathbb{C}^{n \times n}$ on ominaisarvo $\lambda \in \mathbb{C}$.*

Kuten aiemmin on todettu. Nämä ovat erittäin syvällisiä tuloksia ja ne eivät päde reaalisten vektoriavaruuksien välisille lineaarikuvauksille tai reaalisille matriiseille. Tulokset perustuvat kompleksilukujen ominaisuuteen, että jokaisella vähintään astetta yksi olevalla kompleksikertoimisella polynomilla on juuri. Tätä tulosta kutsutaan *algebran peruslauseeksi.*

Lause 14.1.3 (Algebran peruslause)**.** *Olkoon $p \colon \mathbb{C} \to \mathbb{C}$, $z \mapsto a_n z^n + \cdots + a_1 z + a_0$, kompleksinen polynomi (eli $a_1, \ldots, a_n \in \mathbb{C}$), joka on vähintään astetta* 1 *eli* $\deg p \geq 1$.

Jos polynomin p aste $\deg p$ *on vähintään yksi, niin polynomilla p on juuri, eli jos $n \geq 1$ ja $a_n \neq 0$, niin on olemassa sellainen $\lambda \in \mathbb{C}$ että $p(\lambda) = 0$.*

Algebran peruslauseesta ja polynomien jakoalgoritmista seuraa, että polynomi p voidaan jakaa ensimmäisen asteen tekijöihin. Kirjataan tulos nyt kahdessa vaiheessa. Ensimmäinen vaihe jätetään harjoitustehtäväksi.

Lause 14.1.4. *Olkoon $p \colon \mathbb{C} \to \mathbb{C}$, $z \mapsto a_n z^n + \cdots + a_1 z + a_0$, vähintään astetta yksi oleva kompleksinen polynomi eli* $\deg p \geq 1$. *Tällöin on olemassa astetta* $(\deg p) - 1$ *oleva polynomi $q \colon \mathbb{C} \to \mathbb{C}$, jolle pätee $p(z) = (z - \lambda) q(z)$ kaikilla $z \in \mathbb{C}$.*

Korollaari 14.1.5. *Olkoon $p\colon \mathbb{C} \to \mathbb{C}$ komplesinen astetta $n \geq 1$ oleva polynomi. Tällöin on olemassa $a \in \mathbb{C}$ ja sellaiset $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, että*

$$p(z) = a(z - \lambda_1) \cdots (z - \lambda_n) = a \prod_{j=1}^{n} (z - \lambda_j).$$

Todistus. Olkoon p asteen 1 polynomi. Tällöin p on funktio $z \mapsto a_1 z + a_0$, missä $a_1 \neq 0$ ja $a_0 \in \mathbb{C}$. Koska $a_1 z + a_0 = a_1(z - (-a_0/a_1))$ kaikilla $z \in \mathbb{C}$, niin väite pätee, kun $n = 1$.

Oletetaan, että väite pätee asteen $n \in \mathbb{N}$ polynomeille ja olkoon p asteen $n + 1$ polynomi. Lauseen 14.1.4 nojalla on olemassa sellainen luku $\lambda \in \mathbb{C}$ ja sellainen asteen n polynomi q, että $p(z) = (z - \lambda)q(z)$ kaikilla $z \in \mathbb{C}$. Väite seuraa soveltamalla induktio-oletusta voidaan nyt soveltaa polynomiin q. □

Huomautus 14.1.6. *Korollaari pätee myös asteen $n = 0$ polynomeille, kun funktio*

$$z \mapsto \prod_{j=1}^{0} (z - \lambda_j)$$

tulkitaan vakiofunktioksi $z \mapsto 1$.

Huomautus 14.1.7. *Algebran peruslause ja kompleksikertoimisten polynomien teoria otetaan tällä kurssilla annettuna. Aihetta on käsitelty tarkemmin esimerkiksi Axlerin erinomaisessa kirjassa [1], jota kirjoittaja suosittelee tässäkin yhteydessä kaikille.*

Todistetaan nyt ominaisarvon olemassaoloa kompleksisessa tapauksessa determinanttien avulla.

Lauseen 14.1.2 todistus determinanteilla. Olkoon $A \in \mathbb{C}^{n \times n}$ neliömatriisi. Tarkastellaan funktiota $p\colon \mathbb{C} \to \mathbb{C}$,

$$\lambda \mapsto \det(A - \lambda I).$$

Kuten reaalisessakin tapauksessa, determinantin kehityskaavojen perusteella, funktio p on asteen n polynomi

$$\lambda \mapsto (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1 \lambda + a_0,$$

missä $a_{n-1}, \ldots, a_0 \in \mathbb{C}$ ja $a_0 = \det A$.

Algebran peruslauseen nojalla polynomilla p on juuri λ, eli

$$\det(A - \lambda I) = 0.$$

Näin ollen, kuten reaalisten matriisien tapauksessa, matriisi $A - \lambda I$ ei ole kääntyvä ja yhtälöllä

$$(A - \lambda I)v = 0$$

on epätriviaali ratkaisu $v \in \mathbb{C}^{n \times 1}$. Koska

$$Av = \lambda v,$$

niin λ on matriisin A ominaisarvo. □

Todistetaan nyt lause 14.1.1.

Lauseen 14.1.1 todistus. Olkoon $\varphi\colon V \to V$ operaattori ja olkoon $(v_1, \dots, v_n)$ avaruuden V kanta. Tällöin lauseen 14.1.2 operaattorin φ esitysmatriisilla A kannassa $(v_1, \dots, v_n)$ on ominaisarvo $\lambda \in \mathbb{C}$. Lauseen 13.8.6 perusteella λ on kuvauksen φ ominaisarvo. $\square$

Huomautus 14.1.8. *Lause 14.1.1 voidaan todistaa myös ilman esitysmatriisiin siirtymistä ja determinanttia hyödyntäen kompleksisten polynomien tekijöihin jakoa. Tämä todistus on esitetty liitteessä E.1.*

14.2 Kompleksisen operaattorin yläkolmioesitys

Edellisen luvun olemassaolotuloksesta seuraa, että äärellisulotteisen kompleksisen vektoriavaruuden V kompleksisella operaattorilla $f\colon V \to V$ on esitysmatriisi, joka on yläkolmiomatriisi. Yhtäpitävästi tämä tarkoittaa, jokainen kompleksinen neliömatriisi voidaan kannanvaihdolla saattaa yläkolmiomatriisin muotoon. Palautetaan ensin mieleen yläkolmiomatriisin käsite neliömatriisien tapauksessa.

Määritelmä 14.2.1. *Neliömatriisi $A = [a_{ji}] \in \mathbb{C}^{n\times n}$ on* yläkolmiomatriisi, *jos matriisin A diagonaalin alapuoliset kertoimet ovat nollia, eli $a_{ji} = 0$ kaikilla $j > i$, eli*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Neliömatriisi $A \in \mathbb{C}^{n\times n}$ on alakolmiomatriisi, *jos A^t on yläkolmiomatriisi.*

Huomautus 14.2.2. *Matriisitulon ominaisuuksien perusteella yläkolmiomatriiseilla on seuraava ominaisuus: Matriisi $A \in \mathbb{C}^{n\times n}$ on yläkolmiomatriisi, jos ja vain jos $Ae_j \in \mathrm{Sp}(e_1, \dots, e_j)$ kaikilla $j \in \{1, \dots, n\}$. Vastaavasti matriisi A on alakolmiomatriisi, jos ja vain jos $Ae_j \in \mathrm{Sp}(e_j, \dots, e_n)$ jokaisella $j \in \{1, \dots, n\}$.*

Huomautus 14.2.3. *Palautetaan mieleen, että yläkolmiomatriisien tulot ja käänteismatriisit ovat yläkolmiomatriiseja. Samoin alakolmiomatriisien tulot ja käänteismatriisit ovat alakolmiomatriiseja.*

Aloitetaan tuloksesta matriiseille, joka sanoo, että jokainen kompleksinen neliömatriisi on similaari yläkolmiomatriisin kanssa.

Lause 14.2.4. *Olkoon $A \in \mathbb{C}^{n\times n}$ matriisi. Tällöin on olemassa sellainen kääntyvä matriisi $P \in \mathbb{C}^{n\times n}$ ja sellainen yläkolmiomatriisi $T \in \mathbb{C}^{n\times n}$, että $A = PTP^{-1}$.*

Operaattorihin sovellettuna tämä tulos antaa seuraavan korollaarin.

Korollaari 14.2.5. *Olkoon V n-ulotteinen kompleksinen vektoriavaruus ja $f\colon V \to V$ lineaarioperaattori. Tällöin on olemassa sellainen avaruuden V kanta $(v_1, \dots, v_n)$, että kuvauksen f esitysmatriisi tässä kannassa on yläkolmiomatriisi.*

Todistus. Olkoon $(w_1, \dots, w_n)$ avaruuden V kanta ja A operaattorin f esitysmatriisi kannan $(w_1, \dots, w_n)$ suhteen. Lauseen 14.2.4 nojalla on olemassa sellainen kääntyvä matriisi $P \in \mathbb{C}^{n\times n}$ ja yläkolmiomatriisi $T \in \mathbb{C}^{n\times n}$, että $A = PTP^{-1}$. Olkoon nyt $\Psi_{(w_1,\dots,w_n)} \colon \mathbb{C}^{n\times 1} \to V$ isomorfismi $e_i \mapsto v_i$. Määritellään nyt $v_i = \Psi_{(v_1,\dots,v_n)}(Pe_i)$ jokaisella $i \in \{1, \dots, n\}$. Tällöin $(v_1, \dots, v_n)$ on avaruuden V kanta ja lineaarikuvausten kantalauseen nojalla isomorfismi $\Psi_{(v_1,\dots,v_n)}$ on yhdistetty kuvaus $\Psi_{(v_1,\dots,v_n)} = \Psi_{(w_1,\dots,w_n)} \circ \varphi_P$. Näin ollen operaattorin f esitysmatriisi kannassa $(v_1, \dots, v_n)$ on T. □

Huomautus 14.2.6. *Korollaarille 14.2.5 voidaan antaa myös suora todistus, joka perustuu invariantin aliavaruuden käsitteelle, käyttämättä lausetta 14.2.4. Tätä on käsitelty liitteessä E.2.*

Todistetaan seuraavaksi lause 14.2.4. Todistuksen merkintöjen helpottamiseksi todistuksessä syntyvät apumatriisit on esitetty ositettuina matriiseina. Todistuksen voi myös kirjoittaa käyttäen invariantteja aliavaruuksia. Tämä tapa on käsitelty liitteessä E.2.

Lauseen 14.2.4 todistus. Todistetaan väite induktiolla. Väite selvästi pätee, kun $n = 1$. Oletetaan, että $n \in \mathbb{N}$ on sellainen luku, että väite pätee luvulla $n - 1$.

Olkoon $A \in \mathbb{C}^{n\times n}$. Lauseen 14.1.1 perusteella matriisilla A on ominaisarvo $\lambda \in \mathbb{C}$ ja sitä vastaava ominaisvektori $v \in \mathbb{C}^{n\times 1}$. Koska yhden nollasta poikkeavan vektorin mittainen jono (v) on vapaa, voidaan se jatkaa avaruuden $\mathbb{C}^{n\times 1}$ kannaksi $(v, w_1, \dots, w_{n-1})$. Olkoon nyt $\tilde{P} = \begin{bmatrix} v & w_1 & \cdots & w_{n-1} \end{bmatrix} \in \mathbb{C}^{n\times n}$. Koska $(v, w_1, \dots, w_{n-1})$ on kanta, niin matriisi $\tilde{P}$ on kääntyvä. Lisäksi

$$\tilde{P}^{-1} A \tilde{P} e_1 = \tilde{P}^{-1} A v = \tilde{P}^{-1}(\lambda v) = \lambda \tilde{P}^{-1} v = \lambda e_1.$$

Näin ollen $\tilde{P}^{-1} A \tilde{P}$ on ositettu matriisi

$$\tilde{P}^{-1} A \tilde{P} = \begin{bmatrix} \lambda & y^t \\ 0 & \tilde{A} \end{bmatrix},$$

missä $y \in \mathbb{C}^{(n-1)\times 1}$ on sarakevektori ja $\tilde{A} \in \mathbb{C}^{(n-1)\times(n-1)}$ neliömatriisi. Huomaa, että transpoosi tekee sarakevektorista y rivivektorin y^t.

Induktio-oletuksen nojalla on olemassa sellainen kääntyvä matriisi $\tilde{Q} \in \mathbb{C}^{(n-1)\times(n-1)}$ ja sellainen yläkolmiomatriisi $\tilde{R} \in \mathbb{C}^{(n-1)\times(n-1)}$, että

$$\tilde{A} = \tilde{Q} \tilde{R} \tilde{Q}^{-1}.$$

Määritellään nyt matriisi $P \in \mathbb{C}^{n\times n}$ kaavalla

$$P = \tilde{P} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{bmatrix}.$$

Tällöin

$$
\begin{aligned}
P^{-1}AP &= \left(\tilde{P}\begin{bmatrix}1 & 0\\ 0 & \tilde{Q}\end{bmatrix}\right)^{-1} A \left(\tilde{P}\begin{bmatrix}1 & 0\\ 0 & \tilde{Q}\end{bmatrix}\right)\\
&= \begin{bmatrix}1 & 0\\ 0 & \tilde{Q}\end{bmatrix}^{-1} \tilde{P}^{-1}A\tilde{P}\begin{bmatrix}1 & 0\\ 0 & \tilde{Q}\end{bmatrix}\\
&= \begin{bmatrix}1 & 0\\ 0 & \tilde{Q}^{-1}\end{bmatrix}\begin{bmatrix}\lambda & y^t\\ 0 & \tilde{A}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & \tilde{Q}\end{bmatrix}\\
&= \begin{bmatrix}1 & 0\\ 0 & \tilde{Q}^{-1}\end{bmatrix}\begin{bmatrix}\lambda & y^t\tilde{Q}\\ 0 & \tilde{A}\tilde{Q}\end{bmatrix}\\
&= \begin{bmatrix}\lambda & (\tilde{Q}^t y)^t\\ 0 & \tilde{Q}^{-1}\tilde{A}\tilde{Q}\end{bmatrix} = \begin{bmatrix}\lambda & (\tilde{Q}^t y)^t\\ 0 & \tilde{R}\end{bmatrix}.
\end{aligned}
$$

Koska matriisi

$$
T = \begin{bmatrix}\lambda & (\tilde{Q}^t y)^t\\ 0 & \tilde{R}\end{bmatrix}
$$

on yläkolmiomatriisi ja

$$
A = PTP^{-1},
$$

niin tämä päättää induktioaskeleen ja siten todistuksen. □

Esimerkki 14.2.7. *Olkoon*

$$
A = \begin{bmatrix}-1 & 2\\ -2 & 3\end{bmatrix}.
$$

Etsitään matriisille A yläkolmioesitys seuraamalla todistuksen algoritmia. Aloitetaan laskemalla matriisin A ominaisarvot. Koska

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

niin matriisin A ainoa ominaisarvo on $\lambda = 1$. Koska

$$
A-\lambda I = \begin{bmatrix}-2 & 2\\ -2 & 2\end{bmatrix},
$$

niin $(A-\lambda I)v = 0$, jos ja vain jos $v \in \mathrm{Sp}(e_1+e_2)$. Näin ollen tätä ominaisarvoa vastaava ominaisavaruus on

$$
E(1, A) = \mathrm{Sp}\left(\begin{bmatrix}1\\ 1\end{bmatrix}\right).
$$

Olkoon nyt $v = e_1 + e_2$ ja täydennetään jono (v) avaruuden $\mathbb{C}^{2\times 1}$ kannaksi. Koska jono (v, e_1, e_2) virittää avaruuden $\mathbb{C}^{2\times 1}$, niin se sisältää avaruuden $\mathbb{C}^{2\times 1}$ kannan. Asettamalla $M = \begin{bmatrix}v & e_1 & e_2\end{bmatrix}$. Havaitaan, että haluttu kanta löytyy viemällä matriisi M supistettuun porrasmuotoon. Tässä kyseisessä tapauksessa on kuitenkin selvää, että (v) voidaan täydentää kannaksi (v, e_1).

Olkoon nyt

$$Q = \begin{bmatrix} v & e_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Tällöin

$$Q^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}.$$

Näin ollen

$$Q^{-1}AQ = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

Koska matriisi $Q^{-1}AQ$ on yläkolmiomatriisi, niin voidaan lopettaa, sillä nyt $R = Q^{-1}AQ$ ja Q toteuttavat ehdon

$$QRQ^{-1} = Q(Q^{-1}AQ)Q^{-1} = A.$$

Muutama huomio on paikallaan.

Huomautus 14.2.8. *Yleisen $n \times n$-matriisin A yläkolmioesitys lasketaan vastaavalla tavalla. Algoritmia voi nopeuttaa seuraavasti.*

Ratkaistaan matriisin A ominaisarvot $\lambda_1, \ldots, \lambda_k$ ja etsitään näitä ominaisarvoja vastaavien ominaisavaruuksien $E(\lambda_i, A)$ kannat. Olkoon $(v_1, \ldots, v_k)$ sellainen vapaa jono ominaisvektoreita, että ne virittävät aliavaruuden $E(\lambda_1, A) \oplus \cdots \oplus E(\lambda_k, A)$. Täydennetään jono $(v_1, \ldots, v_k)$ avaruuden $\mathbb{C}^{n\times 1}$ kannaksi $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ etsimällä matriisin $\begin{bmatrix} v_1 & \cdots & v_k & e_1 & \cdots & e_n \end{bmatrix}$ sarakeavaruuden kanta. Olkoon nyt $\tilde{P} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. Tällöin matriisi $\tilde{P}^{-1}A\tilde{P}$ on ositettumatriisi

$$\begin{bmatrix} D & E \\ 0 & \tilde{A} \end{bmatrix}$$

missä D on ominaisarvojen diagonaalimatriisi. Mikäli matriisi $\tilde{A}$ ei ole yläkolmiomatriisi, niin toistetaan sama prosessi matriisille $\tilde{A}$, kunnes yläkolmiomuoto on löydetty. Näin saadaan $\tilde{A} = \tilde{Q}R\tilde{Q}^{-1}$ missä R on yläkolmiomatriisi. Nyt voidaan asettaa

$$P = \tilde{P} \begin{bmatrix} I & 0 \\ 0 & \tilde{Q} \end{bmatrix}$$

ja saadaan

$$A = P \begin{bmatrix} D & E\tilde{Q} \\ 0 & R \end{bmatrix} P^{-1}.$$

Esimerkki 14.2.9. *Olkoon*

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

Jälleen $\lambda = 1$ on matriisin A ainoa ominaisarvo ja sen ominaisavaruus on $E(1, A) = \mathrm{Sp}(e_1 + e_2, e_3)$. Jono $(e_1 + e_2, e_3)$ voidaan täydentää kannaksi $(e_1 + e_2, e_3, e_1)$. Tällöin matriisi

$$Q = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

on kääntyvä. Lisäksi

$$\begin{aligned} R = Q^{-1}AQ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

on yläkolmiomatriisi. Näin ollen

$$A = QRQ^{-1}$$

on matriisin A haluttu yläkolmioesitys.

14.2.1 Yläkolmiomatriiseista invariantteihin aliavaruuksiin

Lause 14.2.4 herättää kysymyksen, että mitä matriisin yläkolmioesitys kertoo operaattoreista. Kysymyksen tarkempi pohdinta johdattaa kysymään, mitä yläkolmiomatriisit oikeasti tarkoittavat. Tämä kysymys johtaa invarianttin aliavaruuden käsitteeseen, jota on käsitelty tarkemmin litteessä E.2.

Seuraava konkreettinen esimerkki paljastaa tilanteesta kaiken olennaisen.

Esimerkki 14.2.10. *Olkoon*

$$R = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}.$$

Koska $f_R(e_i) = v_i$ jokaisella $i \in \{1, \ldots, 3\}$. Tällöin

$$\begin{aligned} \varphi_R(e_1) &= 2e_1 \in \mathrm{Sp}(e_1), \\ \varphi_R(e_2) &= -e_1 + e_2 \in \mathrm{Sp}(e_1, e_2) \text{ ja} \\ \varphi_R(e_3) &= e_1 + e_3 \in \mathrm{Sp}(e_1, e_2, e_3). \end{aligned}$$

Esimerkki havainnollistaa, että yläkolmiomatriisin $R \in \mathbb{C}^{n \times n}$ karakteristinen ominaisuus on, että $Re_i \in \mathrm{Sp}(e_1, \ldots, e_i)$ kaikilla $i \in \{1, \ldots, n\}$. Tämä tarkoittaa, että lineaarikuvauksella $\varphi_R \colon \mathbb{C}^{n \times 1} \to \mathbb{C}^{n \times 1}$ on ominaisuus, että jokaisella $i \in \{1, \ldots, n\}$ pätee

$$\varphi_R(\mathrm{Sp}(e_1, \ldots, e_i)) \subset \mathrm{Sp}(e_1, \ldots, e_i).$$

Tarkastellaan nyt yleistä operaattoria : $\mathbb{C}^{n\times 1} \to \mathbb{C}^{n\times 1}$. Tällöin $\varphi = \varphi_A$ jollakin $A \in \mathbb{C}^{n\times n}$. Lauseen 14.2.4 nojalla $A = PTP^{-1}$, missä P on kääntyvä ja T on yläkolmiomatriisi. Näin ollen voidaan sanoa, että operaattorin φ esitysmatriisi kannassa $(Pe_1, \dots, Pe_n)$ on yläkolmiomatriisi. Olemme oleellisesti todistaneet seuraavan yleisen tuloksen, joka on todistettu liitteessä E.2.

Lause 14.2.11 (Yläkolmioesityksen operaattoritulkinta)**.** *Olkoon $\varphi\colon V \to V$ äärellisulotteisen kompleksisen vektoriavaruuden V operaattori. Tällöin on olemassa sellainen avaruuden V kanta $(v_1, \dots, v_n)$, jossa operaattorin φ esitysmatriisi on yläkolmiomatriisi.*

Tämän lauseen varsinainen tulos on kuitenkin seuraava aliavaruuksia koskeva huomio, joka sanoo, että kannalla $(v_1, \dots, v_n)$ on sellainen erityinen ominaisuus, että aliavaruudet $\mathrm{Sp}(v_1, \dots, v_i)$ kuvautuvat itsensä sisään kuvauksessa φ.

Korollaari 14.2.12. *Olkoon $\varphi\colon V \to V$ äärellisulotteisen kompleksisen vektoriavaruuden V operaattori. Tällöin on olemassa sellainen avaruuden V kanta $(v_1, \dots, v_n)$, että*

$$\varphi(\mathrm{Sp}(v_1, \dots, v_i)) \subset \mathrm{Sp}(v_1, \dots, v_i)$$

jokaisella $i \in \{1, \dots, n\}$.

Tälläisiä aliavaruuksia kutsutaan invarianteiksi aliavaruuksiksi.

Määritelmä 14.2.13. *Avaruuden V aliavaruus $W \subset V$ on operaattorin $\varphi\colon V \to V$* invariantti *aliavaruus, jos $\varphi(W) \subset W$.*

Matriisin yläkolmioesitys antaa siis seuraavan tuloksen.

Korollaari 14.2.14. *Jokainen äärellisulotteinen kompleksinen vektoriavaruus V voidaan kirjoittaa kasvavana jonona operaattorin $\varphi\colon V \to V$ invarianteista aliavaruuksista.*

Invariantit aliavaruudet ovat olennainen käsite operaattorien teoriassa myös siitä syystä, että operaattorin rajoittuma invarianttiin aliavaruuteen on hyvin määritelty operaattori. Kirjataan tämä lähes triviaali havainto lauseeksi sen tärkeyden vuoksi.

Lause 14.2.15. *Olkoon V vektoriavaruus ja $W \subset V$ operaattorin $\varphi\colon V \to V$ invariantti aliavaruus. Tällöin rajoittuma $\varphi|_W\colon W \to W$ on hyvin määritelty operaattori.*

Todistus. Koska W on operaattorin invariantti aliavaruus, niin $\varphi(W) \subset W$. Näin ollen rajoittuman $\varphi|_W\colon W \to V$ maaliavaruus voidaan rajoittaa avaruuteen W. Näin ollen $\varphi|_W\colon W \to W$ on hyvin nääritelty operaattori. □

14.2.2 Reaalisen matriisin yläkolmioesitys

Lause E.2.1 ei päde reaalisten vektoriavaruuksien tapauksessa. Syy tähän on se, että yläkolmiomatriisilla $T \in \mathbb{R}^{n\times n}$ on vähintään yksi ominaisvektori ja tämä ominaisuus säilyy kannanvaihdossa. Näin ollen esimerkiksi kiertomatriiseille ei ole yläkolmioesitystä. Yläkolmioesitys kuitenkin on olemassa, mikäli polynomi $t \mapsto \det(A - tI)$ faktoroituu ensimmäisen asteen tekijöihin eli polynomilla $t \mapsto \det(A - tI)$ on n reaalista juurta moninkerrat huomioiden.

Lause 14.2.16. *Olkoon $A \in \mathbb{R}^{n\times n}$ sellainen matriisi, että polynomi $t \mapsto \det(A-tI)$ faktoroituu ensimmäisen asteen tekijöihin. Tällöin on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{n\times n}$ ja yläkolmiomatriisi $T \in \mathbb{R}^{n\times n}$, että $A = PTP^{-1}$.*

Todistus. Koska $t \mapsto \det(A - tI)$ faktoroituu ensimmäisen asteen tekijöihin, niin on olemassa sellaiset luvut $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, että

$$\det(A - tI) = c(t - \lambda_1)\cdots(t - \lambda_n),$$

missä $c \in \mathbb{R}$. Tällöin luvut $\lambda_1, \dots, \lambda_n$ ovat matriisin A ominaisarvoja. Olkoon nyt $v_1 \in \mathbb{R}^{n\times 1}$ matriisin A sellainen ominaisvektori, jonka ominaisarvo on λ_1.

Merkitään $A_0 = A$ ja olkoon nyt $P_1 \in \mathbb{R}^{n\times n}$ sellainen kääntyvä matriisi, jonka ensimmäinen sarake on v_1. Tällöin $P_1^{-1}A_0P_1$ on sellainen matriisi, että

$$P_1^{-1}A_0P_1e_1 = P_1^{-1}Av_i = \lambda_1 P_1^{-1}v_i = \lambda_1 e_1.$$

eli matriisin $P_1^{-1}A_0P_1$ ensimmäinen sarake on $\lambda_1 e_1$.

Olkoon nyt $A_1 \in \mathbb{R}^{(n-1)\times(n-1)}$ matriisi, joka saadaan matriisista $P_1^{-1}A_0P_1$ poistamalla ensimmäinen rivi ja sarake, eli

$$P_1^{-1}A_0P_1 = \begin{bmatrix} \lambda_1 & b_1 \\ 0 & A_1 \end{bmatrix},$$

missä $b_1 \in \mathbb{R}^{1\times(n-1)}$.

Koska

$$\det(P_1^{-1}A_0P_1 - tI) = \det(P_1^{-1}(A_0 - tI)P_1) = \det(P_1^{-1})\det(A_0 - tI)\det(P_1) = \det(A_0 - tI)$$

ja

$$\det(P_1^{-1}A_0P_1 - tI) = (\lambda_1 - t)\det(A_1)$$

niin saadaan, että luvut $\lambda_2, \dots, \lambda_n$ ovat polynomin $t \mapsto \det(A_1 - tI)$ juuria (monikerrat huomioiden). Näin ollen väite seuraa induktiolla. □

14.3 Schurin lause

Lause 14.2.4 esitetään yleensä vahvemmassa muodossa, että matriisi P voidaan valita unitaariseksi. Tätä tulosta kutsutaan Schurin lauseeksi. Schurin lause seuraa lähes suoraan kompleksisesta Gram–Schmidt menetelmästä (eli kompleksisesta QR-hajotelmasta). Kirjataan tarvittavat tulokset.

Lemma 14.3.1. *Olkoon $A \in \mathbb{C}^{n\times n}$ kääntyvä matriisi. Tällöin on olemassa sellaiset unitaarimatriisi $Q \in \mathbb{C}^{n\times n}$ ja yläkolmiomatriisi $R \in \mathbb{C}^{n\times n}$, että $A = QR$.*

Todistus. Koska matriisin A sarakkeet muodostavat avaruuden $\mathbb{C}^{n\times 1}$ kannan, niin Gram–Schmidt algoritmi tuottaa sellaisen ortonormaalin kannan $(u_1, \dots, u_n)$, että

$$Q = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \in \mathbb{C}^{n\times n}$$

on unitaarimatriisi, jolle pätee, että matriisi $Q^{-1}A$ on yläkolmiomatriisi. (Harjoitustehtävä). □

Lause 14.3.2 (Schurin lause)**.** *Olkoon $A \in \mathbb{C}^{n\times n}$ matriisi. Tällöin on olemassa sellaiset unitaarimatriisi $U \in \mathbb{C}^{n\times n}$ ja yläkolmiomatriisi $T \in \mathbb{C}^{n\times n}$, että $A = UTU^*$.*

Todistus. Korollaarin 14.2.5 perusteella on olemassa sellainen kääntyvä matriisi $P \in \mathbb{C}^{n\times n}$ ja yläkolmiomatriisi $Y \in \mathbb{C}^{n\times n}$, että $A = PYP^{-1}$. Lemman 14.3.1 nojalla on olemassa sellainen unitaarimatriisi $Q \in \mathbb{C}^{n\times n}$ ja yläkolmiomatriisi $R \in \mathbb{C}^{n\times n}$, että $P = QR$. Tällöin

$$A = (QR)Y(QR)^{-1} = QRYR^{-1}Q^{-1} = Q(RYR^{-1})Q^*.$$

Koska R ja Y ovat yläkolmiomatriiseja, niin $T = RYR^{-1}$ on yläkolmiomatriisi. □

Huomautus 14.3.3. *Schurin lauseen todistus paljastaa, että mikäli matriisille $A \in \mathbb{C}^{n\times n}$ on yläkolmioesitys jossain kannassa, niin sille on yläkolmioesitys ortonormaalissa kannassa. Vastaava tulos pätee myös reaalisille matriiseille.*

14.3.1 Schurin lauseen reaalinen vastine

Schurin lauseella on myös reaalinen vastine.

Lause 14.3.4 (Reaalinen Schurin lause)**.** *Olkoon $A \in \mathbb{R}^{n\times n}$ sellainen matriisi, että sillä on n reaalista ominaisarvoa (kertaluvun mukaan laskettuna). Tällöin on olemassa sellaiset ortogonaalimatriisi $O \in \mathbb{R}^{n\times n}$ ja yläkolmiomatriisi $T \in \mathbb{R}^{n\times n}$, että $A = OTO^t$.*

Todistus. Lauseen 14.2.16 nojalla on olemassa sellaiset kääntyvä matriisi $P \in \mathbb{R}^{n\times n}$ ja yläkolmiomatriisi $Y \in \mathbb{R}^{n\times n}$, että $A = PYP^{-1}$. QR-hajotelman nojalla on olemassa sellainen ortogonaalimatriisi $O \in \mathbb{R}^{n\times n}$ ja yläkolmiomatriisi $R \in \mathbb{R}^{n\times n}$, että $P = OR$. Nyt voidaan valita $T = RYR^{-1}$. □

Luku 15

Operaattorin diagonalisoituvuus

Yksi tärkeimpiä neliömatriiseihin ja siten operaattoreihin liittyviä kysymyksiä on, että voidaanko annettu matriisi saattaa kannanvaihdolla diagonaalimatriisin muotoon. Reaalisille neliömatriiseille tämä kysymys on siis seuraava:

> Olkoon $A \in \mathbb{R}^{n\times n}$. Onko olemassa sellaista avaruuden $\mathbb{R}^{n\times 1}$ kantaa $(v_1, \dots, v_n)$, että kuvauksen $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times n}$ esitysmatriisi kannassa $(v_1, \dots, v_n)$ on diagonaalimatriisi? Toisin sanoen, onko olemassa sellaista kantaa $(v_1, \dots, v_n)$ ja diagonaalimatriisia $D \in \mathbb{R}^{n\times n}$, että $A = PDP^{-1}$, missä $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n\times n}$?

Kompleksiselle neliömatriisille kysymys on analoginen; riittää vaihtaa matriisiavaruus $\mathbb{R}^{n\times n}$ avaruuteen $\mathbb{C}^{n\times n}$ ja sarakeavaruus $\mathbb{R}^{n\times 1}$ avaruuteen $\mathbb{C}^{n\times 1}$.

On tärkeää huomata, että matriisin diagonalisoituvuus ei ole automaattinen ominaisuus. Kuten luvussa 6 huomattiin, diagonalisoituvuus on yhteydessä matriisin ominaisarvoihin. Näin ollen kaikki matriisit eivät diagonalisoidu! Tätä huomiota on tärkeää peilata vasten edellisen luvun tuloksia matriisin yläkolmioesityksestä. Edellisessä luvussa osoitetettiin, että jos $A \in \mathbb{C}^{n\times n}$ on kompleksinen neliömatriisi, niin tällöin on olemassa sellainen avaruuden $\mathbb{C}^{n\times 1}$ kanta $(v_1, \dots, v_n)$, että $A = PTP^{-1}$, missä $T \in \mathbb{C}^{n\times n}$ on yläkolmiomatriisi ja $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. Koska tämän tuloksen todistus perustui matriisin A ominaisarvon löytämiseen, voi tulla ajatukseen, että tämällä perusteella jokainen kompleksinen neliömatriisi pystytään diagonalisoimaan, jos vain ollaan riittävän fiksuja. Todistuksen tarkempi analyysi kuitenkin osoittaa, että näin ei todellakaan ole. Edellä mainittu tärkeä huomio on siis, että edes kaikkia kompleksisia matriiseja ei voida diagonalisoida.

Tämä herättää kysymyksen, että mitkä matriisit sitten ovat diagonalisoituvia? Ensimmäinen vastaus on saatu jo luvussa 6: *Matriisi $A \in \mathbb{C}^{n\times n}$ voidaan diagonalisoida, jos ja vain jos sen ominaisvektoreista voidaan muodostaa avaruuden $\mathbb{C}^{n\times 1}$ kanta.*

Vaikka vastaus on kattava, ei se ole kuitenkaan täysin tyydyttävä: diagonalisoitutuuvuuden selvittämiseksi matriisi pitää oleellisesti kokeilla diagonalisoida!

Mielenkiintoinen ja samalla erittäin hyödyllinen vastaus saadaan symmetristen matriisien tapauksessa:

Symmetrinen neliömatriisi on aina diagonalisoituva.

Tämä tulos on erittäin hyvä. Se ei ole diagonalisoituvuuden karakterisointi, mutta se antaa selkeän ehdon milloin matriisi on diagonalisoituva. Lisäksi tarkastettava ehto, eli symmetrisyys, on sellainen, että sen tarkastamiseksi ei tarvitse laskea mitään.

Kuten väitteen muotoilusta voi huomata, se on mielenkiintoinen siinä mielessä, että ei ole eritelty ovatko matriisit reaalisia tai kompleksisia. Erityinen hienous on siinä, että tuloksen tarkempi muotoilu sanoo, että kompleksiset symmetriset matriisit ovat kompleksisesti diagonalisoituvia ja reaaliset symmetriset matriisit reaalisesti diagonalisoituvia. Tuloksen tarkka muotoilu sanoo siis eristyisesti, että reaalisella symmetrisellä neliömatriisilla on aina ominaisarvoja! Nämä tulokset todistetaan luvussa 15.4.

15.1 Diagonalisoituvuuden karakterisointi ominaisavaruuksien avulla

Palautetaan mieleen, että matriisi $A = [a_{ji}] \in \mathbb{C}^{n\times n}$ on *diagonaalimatriisi*, jos $a_{ji} = 0$ kaikilla $j \neq i$, missä $j, i \in \{1, \dots, n\}$.

Esimerkki 15.1.1. *Matriisi*

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

on diagonaalimatriisi. Selvästi matriisi

$$B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

on yläkolmiomatriisi, mutta ei diagonaalimatriisi.

Kirjataan nyt diagonalisoituvuus sekä neliömatriisien että operaattorien kielellä. Aloitetaan matriiseista.

Määritelmä 15.1.2. *Neliömatriisi $A \in \mathbb{C}^{n\times n}$ on* (kompleksisesti) diagonalisoituva*, jos on olemassa sellainen kääntyvä matriisi $P \in \mathbb{C}^{n\times n}$, että $A = PDP^{-1}$, missä $D \in \mathbb{C}^{n\times n}$ on diagonaalimatriisi.*

Vastaavasti, neliömatriisi $A \in \mathbb{R}^{n\times n}$ on (reaalisesti) diagonalisoituva*, jos on olemassa sellainen kääntyvä matriisi $P \in \mathbb{R}^{n\times n}$, että $A = PDP^{-1}$, missä $D \in \mathbb{R}^{n\times n}$ on diagonaalimatriisi.*

Huomautus 15.1.3. *Yleensä ei korosteta neliömatriisin diagonalisoituvuuden reaalisuutta tai kompleksisuutta, koska merkitys on yleensä asiayhteydestä selvä. Määritelmän tasolla on kuitenkin hyvä huomata, että reaalikertomisen matriisin diagonalisoituvuutta voidaan tarkastella kummaltakin kannalta.*

Operaattoreille diagonalisoituvuus voidaan määritellä seuraavasti.

Määritelmä 15.1.4. *Äärellisulotteisen vektoriavaruuden V operaattori $f: V \to V$ on* diagonalisoituva*, jos on olemassa sellainen avaruuden V kanta, jossa operaattorin f esitysmatriisi on diagonaalimatriisi.*

Huomautus 15.1.5. *Määritelmä 15.1.4 olisi voitu antaa näennäisesti heikommassakin mutta yhtäpitävässä muodossa seuraavasti: Äärellisulotteisen vektoriavaruuden V operaattori $f: V \to V$ on diagonalisoituva, jos on olemassa sellainen avaruuden V kanta, jossa operaattorin f esitysmatriisi on diagonalisoituva.*

Koska esitysmatriisin diagonalisoituvuus vastaa kannanvaihtoa sarakeavaruudessa ja siten kannanvaihtoa avaruudessa V nämä ehdot ovat yhtäpitävät. Yksityiskohdat jätetään harjoitustehtäväksi.

Palautetaan mieleen, että diagonalisoinnissa $A = PDP^{-1}$ matriisin $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ sarakkeet ovat matriisin A ominaisvektoreita. Tämä seuraa tutusta laskusta

$$Av_i = PDP^{-1}v_i = PDe_i = P(\lambda_i e_i) = \lambda_i Pe_i = \lambda_i v_i,$$

missä v_i on matriisin P i:s sarakevektori ja λ_i diagonaalimatriisin D i:s diagonaalialkio.

Annetaan nyt tähän huomioon perustuva esimerkki sellaisesta matriisista, jota ei voi diagonalisoida.

Esimerkki 15.1.6. *Olkoon*

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \in \mathbb{C}^{2\times 2}.$$

Determinanttiyhtälön $\det(A - \lambda I) = 0$ *perusteella matriisin A ainoa ominaisarvo on $\lambda = 0$. Tätä ominaisarvoa vastaavat ominaisvektorit saadaan siis yhtälöstä $Ax = 0$. Ominaisavaruus $E(0, A)$ voidaan selvittää ratkaisemalla vastaava yhtälöryhmä tai päättelemällä seuraavasti. Koska $Ae_1 = e_2$ ja $Ae_2 = 0$, niin* $\dim \mathrm{Col}(A) \geq 1$ *ja* $\dim \mathrm{Null}(A) \geq 1$. *Koska dimensiolauseen nojalla* $\dim \mathrm{Col}(A) + \dim \mathrm{Null}(A) = \dim \mathbb{C}^{2\times 1} = 2$, *niin* $\dim \mathrm{Col}(A) = \dim \mathrm{Null}(A) = 1$. *Näin ollen* $\mathrm{Null}(A) = \mathrm{Sp}(e_2)$.

Matriisilla A ei siis ole kahta lineaarisesti riippumatonta ominaisvektoria. Ominaisvektoreista ei siis voi muodostaa kantaa. Näin ollen ei ole olemassa sellaista kääntyvää matriisia P, että $A = PDP^{-1}$, missä D on diagonaalimatriisi.

Annataan tässä yhteydessä myös esimerkki diagonalisoituvasta matriisista.

Esimerkki 15.1.7. *Esimerkin 15.1.1 matriisi B on diagonalisoituva. Tämä seuraa siitä, että sillä on kaksi ominaisarvoa 2 ja 1, joiden ominaisavaruudet ovat $E(2, B) =$* $\mathrm{Sp}\{e_1\}$ *ja $E(1, B) =$* $\mathrm{Sp}\{e_1 - e_2\}$.

Selvästi $(e_1, e_1 - e_2)$ on avaruuden $\mathbb{R}^{2\times 1}$ (tai avaruuden $\mathbb{C}^{2\times 1}$) kanta. Tässä kannassa lineaarikuvaus $\varphi_B: \mathbb{C}^{2\times 1} \to \mathbb{C}^{2\times 1}$, $z \mapsto Bz$, saa matriisin

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

15.1.1 Ominaisavaruuksien summa on suora

Täydennetään nyt luvussa 11.3 käsiteltyjä tuloksia.

Palautetaan mieleen, että operaattorin $f\colon V \to V$ eri ominaisarvoihin $\lambda \neq \mu$ liittyvät ominaisavaruudet $E(\lambda, f)$ ja $E(\mu, f)$ eivät leikkaa epätriviaalisti, eli että $E(\lambda, f) \cap E(\mu, f) = \{0\}$. Tässä tilanteessa tämä havaitaan seuraavasti. Olkoon $v \in E(\lambda, f) \cap E(\mu, f)$. Tällöin $\lambda v = f(v) = \mu v$. Koska $\lambda \neq \mu$, niin $v = 0$. Väite siis seuraa. Yleisemmin lauseen 11.3.5 nojalla tiedetään seuraavaa:

> Olkoon $f\colon V \to V$ operaattori, olkoot $\lambda_1 < \lambda_2 < \cdots < \lambda_k$ operaattorin f ominaisarvoja sekä olkoot $v_i \in E(\lambda_i, f)$ ominaisvektoreita jokaisella $i \in \{1, \ldots, k\}$. Tällöin jono $(v_1, \ldots, v_k)$ on vapaa.

Kuten luvussa 11.3 huomautettiin, tästä seuraa, että ominaisavaruuksien summa on suora. Reaalisten vektoriavaruuksien tulos, eli liitteen D lause D.2.1, pätee sellaisenaan myös kompleksisille vektoriavaruuksille.

Lause 15.1.8. *Olkoot V kompleksinen vektoriavaruus, $f\colon V \to V$ operaattori ja olkoot $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ operaattorin f erisuuria ominaisarvoja. Tällöin*

$$E(\lambda_1, f) + \cdots + E(\lambda_m, f) = E(\lambda_1, f) \oplus \cdots \oplus E(\lambda_m, f).$$

Koska ominaisavaruuksin summa on suora, niin niiden summan dimensio on dimensioiden summa. Kirjataan tästä huomiosta seuraava korollaariksi.

Korollaari 15.1.9. *Olkoot V äärellisulotteinen kompleksinen vektoriavaruus ja $f\colon V \to V$ operaattori. Jos $\lambda_1, \ldots, \lambda_m$ ovat operaattorin f erisuuria ominaisarvoja, niin*

$$\dim E(\lambda_1, f) + \cdots + \dim E(\lambda_m, f) \leq \dim V.$$

Koska jokainen ominaisavaruus on vähintään yksi ulotteinen, saadaan toinen todistus tutulle tulokselle ominaisarvojen määrästä.

Korollaari 15.1.10. *Olkoot V äärellisulotteinen vektoriavaruus ja $f\colon V \to V$ operaattori. Tällöin operaattorilla f on korkeintaan* $\dim V$ *erisuurta ominaisarvoa.*

15.1.2 Diagonalisoituvuuden yleisiä karakterisaatioita

Seuraava lause yhdistää kaikki aiemmat tulkinnat operaattorin diagonalisoituvuudesta ominaisarvojen ja ominaisavaruuksien avulla karakterisointiin invarianttien aliavaruuksien avulla.

Lause 15.1.11. *Olkoon V äärellisulotteinen vektoriavaruus ja $f\colon V \to V$ operaattori. Olkoot $\lambda_1, \ldots, \lambda_m$ operaattorin f kaikki erisuuret ominaisarvot. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *operaattori f on diagonalisoituva,*

2. *avaruudella V on kanta, joka koostuu operaattorin f ominaisvektoreista,*

3. *on olemassa sellaiset avaruuden V 1-ulotteiset aliavaruudet $U_1, \ldots, U_n$, jotka ovat invariantteja operaattorin f suhteen ja joille pätee $V = U_1 \oplus \cdots \oplus U_n$,*

4. *$V = E(\lambda_1, f) \oplus \cdots \oplus E(\lambda_m, f)$ ja*

5. $\dim V = \dim E(\lambda_1, f) + \cdots + \dim E(\lambda_m, f)$.

Todistus. Ehdot (1) ja (2) ovat selvästi yhtäpitäviä. (Harjoitustehtävä).

Lauseiden 15.1.8 ja 13.3.25 perusteella ehdot (4) ja (5) ovat yhtäpitäviä. Lisäksi implikaatio (4) $\Rightarrow$ (2) seuraa lauseesta 13.3.27.

Oletetaan, että ehto (2) pätee ja olkoon $(v_1, \ldots, v_n)$ sellainen avaruuden V kanta, jonka alkiot ovat kuvauksen f ominaisvektoreita. Osoitetaan, että ehto (3) on voimassa.

Olkoon $U_j = \mathrm{Sp}\{v_j\}$ jokaisella $j \in \{1, \ldots, n\}$. Olkoon $j \in \{1, \ldots, n\}$. Koska v_j on ominaisvektori, niin $f(v_j) = \lambda_j v_i$ jollakin $\lambda_i \in \mathbb{C}$. Näin ollen $f(av_j) = a\lambda_j v \in U_j$ kaikilla $a \in \mathbb{C}$, eli $fU_j \subset U_j$. Näin ollen aliavaruudet $U_1, \ldots, U_n$ ovat invariantteja operaattorin f suhteen. Koska $(v_1, \ldots, v_n)$ on avaruuden V kanta, niin

$$V = \mathrm{Sp}\{v_1\} \oplus \cdots \oplus \mathrm{Sp}\{v_n\} = U_1 \oplus \cdots \oplus U_n.$$

Näin ollen ehdosta (2) seuraa ehto (3).

Se, että ehdosta (3) seuraa ehto (2), jätetään harjoitustehtäväksi. Tämä osoittaa, että ehdot (1), (2) ja (3) ovat yhtäpitäviä. Jäljellä on siis osoittaa implikaatio (2) $\Rightarrow$ (4).

Oletetaan, että ehto (2) pätee ja olkoon $(v_1, \ldots, v_n)$ avaruuden V kanta, joka koostuu operaattorin f ominaisvektoreista. Koska jokainen avaruuden V vektori on näin ollen lineaarikombinaatio operaattorin f ominaisvektoreista, niin saadaan

$$V = E(\lambda_1, f) + \cdots + E(\lambda_m, f).$$

Lauseen 15.1.8 nojalla tämä summa on suora. Ehto (4) seuraa. □

Korollaari 15.1.12. *Olkoon V äärellisulotteinen vektoriavaruus. Jos operaattorilla $f \colon V \to V$ on $\dim V$ eri ominaisarvoa, niin f on diagonalisoituva.*

Todistus. Oletuksen nojalla avaruudella V on operaattorin f ominaisvektoreista koostuva kanta. Operaattori f on siis diagonalisoituva lauseen 15.1.11 nojalla. □

Huomautus 15.1.13. *Lukija voi jäädä tässä vaiheessa miettimään, miksi invarianteista aliavaruuksista puhuttiin tässä yhteydessä, koska niillä ei ole kovin suurta roolia. Invarianttien aliavaruuksien merkitys kuitenkin paljastuu, kun palauttaa mieleen matriisin yläkolmioesitykseen liittyvät tulokset.*

Yläkolmiomatriisit voidaan karakterisoida seuraavasti: Matriisi A on yläkolmiomatriisi, jos ja vain jos jokaisella $i \in \{1, \ldots, n\}$ pätee $Ae_i \in \mathrm{Sp}(e_1, \ldots, e_i)$. Tämän ehdon puolestaan voi kirjoittaa muodossa, että matriisi A on yläkolmiomatriisi, jos ja vain jos jokaisella $i \in \{1, \ldots, n\}$ alaivaruus $\mathrm{Sp}(e_1, \ldots, e_i)$ on operaatorin $f_A \colon \mathbb{R}^{n \times 1} \to \mathbb{R}^{n \times 1}$ invariantti aliavaruus.

Näin ollen invariantit aliavaruudet antavat mielekkän käsitteen verrata matriisin yläkolmioesitystä ja diagonalisoituvuutta. Asiasta inspiroitunut lukija ohjataan jälleen Axlerin [1] kirjan pariin.

15.2 Symmetrisen neliömatriisin ominaisarvot

Tässä luvussa aloitetaan symmetristen matriisien spektraalilauseen todistus:

> Matriisi $A \in \mathbb{R}^{n \times n}$ on symmetrinen, jos ja vain jos on olemassa ortogonaalimatriisi P, että PAP^t on diagonaalimatriisi.

Tässä tulos on kirjattu reaalisille matriiseille. Todistus joka seuraavissa luvuissa annetaan on mielenkiintoinen siinä suhteessa, että siinä käytetään kompleksisten matriisien teoriaa. Itseasiassa tämä todistus on yksi tärkeä motivaatio komplekisisten matriisien ja vektoriavaruuksien teorian esittelylle.

Palautetaan mieleen luvuista 12.5.1 ja 13.6 symmetrisen ja hermiittisen matriisin määritelmät:

> Neliömatriisi $A \in \mathbb{R}^{n \times n}$ on *symmetrinen*, jos $A^t = A$, ja kompleksinen neliömatriisi $A \in \mathbb{C}^{n \times n}$ on *hermiittinen*, jos $A^* = A$, missä $A^* = \overline{A}^t$.

Seuraava lause on helposta todistuksestaan huolimatta koko teorian ydin: Vaikka hermiittiset matriisit ovat kompleksikertoimisia, niin niiden ominaisarvot ovat aina reaalisia!

Lause 15.2.1. *Hermiittisen matriisin ominaisarvot ovat reaalisia.*

Todistus. Olkoon $A \in \mathbb{C}^{n \times n}$ hermiittinen matriisi sekä olkoot $\lambda \in \mathbb{C}$ sen ominaisavo ja $v \in \mathbb{C}^{n \times n}$ sitä vastaava ominaisvektori. Tällöin lemman 13.7.2 perusteella

$$\lambda|v|^2 = \lambda(v \cdot v) = (Av) \cdot v = v \cdot (A^* v) = v \cdot (Av) = v \cdot (\lambda v) = \overline{\lambda}(v \cdot v) = \overline{\lambda}|v|^2.$$

Koska v on ominaisvektori, niin $|v|^2 > 0$. Näin ollen $\lambda = \overline{\lambda}$ eli $\lambda \in \mathbb{R}$. $\square$

Edellisestä lauseesta saadaan, että reaalisella polynomilla $x \mapsto \det(A - xI)$ on aina juuri, jos $A^t = A$. Seuraava lause sanoo, että tällöin on myös olemassa tätä juurta vastaava reaalinen sarakevektori, joka on matriisin A ominaisvektori!

Lause 15.2.2. *Olkoon $A \in \mathbb{R}^{n \times n}$ symmetrinen matriisi. Tällöin matriisilla A on ominaisarvo $\lambda \in \mathbb{R}$.*

Todistus. Tulkitaan A kompleksisena matriisina $A \in \mathbb{C}^{n \times n}$. Kompleksisena matriisina sillä on lauseen 14.1.2 nojalla ominaisarvo $\lambda \in \mathbb{C}$. Koska matriisin A kertoimet ovat reaalisia ja A on symmetrinen, niin kompleksisena matriisina A on unitaarinen. Näin ollen $\lambda \in \mathbb{R}$. Osoitetaan nyt, että λ on reaalisen matriisin A ominaisarvo eli että on olemassa sellainen nollasta poikkeava sarakevektori $x \in \mathbb{R}^{n \times 1}$, että $Ax = \lambda x$.

Koska $\lambda \in \mathbb{R}$ on matriisin A ominaisarvo, kun A tulkitaan komplekisena matriisina, niin on olemassa sellainen sarakevektori $z \in \mathbb{C}^{n \times 1}$, että $Az = \lambda z$. Olkoot nyt $x \in \mathbb{R}^{n \times 1}$ ja $y \in \mathbb{R}^{n \times 1}$ sellaisia sarakevektoreita, että $z = x + iy$. Koska $\lambda \in \mathbb{R}$, niin

$$Ax + iAy = A(x + iy) = Az = \lambda z = \lambda(x + iy) = \lambda x + i\lambda y.$$

Näin ollen $Ax = \lambda x$ ja $Ay = \lambda y$ eli vektorit $x \in \mathbb{R}^{n \times 1}$ ja $y \in \mathbb{R}^{n \times 1}$ ovat reaalisen matriisin $A \in \mathbb{R}^{n \times n}$ ominaisvektoreita ominaisarvolla $\lambda \in \mathbb{R}$. $\square$

Lause 15.2.3. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ reaalinen sisätuloavaruus ja $f\colon V \to V$ itseadjungoitu operaattori. Tällöin operaattorilla f on ominaisarvo.*

Todistus. Olkoon $v_1, \ldots, v_n$ avaruuden V kanta ja olkoon A operaattorin f esitysmatriisi tässä kannassa. Koska f on itseadjungoitu, niin A on symmetrinen, eli $A^t = A$. Lauseen 15.2.3 perusteella matriisilla A on ominaisarvo. Näin ollen operaattorilla f on ominaisarvo. □

Huomautus 15.2.4. *Lauseen 15.2.3 voi todistaa käyttämättä kompleksisten vektoriavaruuksien teoriaa. Kiinnostunut lukija ohjataan jälleen Axlerin hienon kirjan [1] pariin.*

15.3 Invariantin aliavaruuden kohtisuora komplementti

Spektralilauseen todistuksessa hyödynnetään kahta itseadjungoitujen operaattorien perusominaisuutta: itseadjungoidun operaattorin invariantin aliavaruuden komplementti on itsessään invariantti ja operaattorin rajoittuma tähän aliavaruuteen on operaattori. Kirjataan nämä tulokset tässä vaiheessa lauseiksi.

Lause 15.3.1. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus, $f\colon V \to V$ itseadjungoitu operaattori ja $W \subset V$ operaattorin f invariantti aliavaruus. Tällöin $W^\perp$ on operaattorin f invariantti aliavaruus eli $fW^\perp \subset W^\perp$.*

Todistus. Olkoon $v \in W^\perp$. Halutaan osoittaa, että $f(v) \in W^\perp$, eli että $\langle f(v), w\rangle = 0$ kaikilla $w \in W$. Olkoon $w \in W$. Koska f on itseadjungoitu ja $fW \subset W$, niin

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle = \langle v, f(w)\rangle = 0.$$

Näin ollen $f(v) \in W^\perp$. □

Kirjataan ylös myös havainto, että itseadjungoidun operaattorin rajoittuma invarianttiin aliavaruuteen on itseadjungoitu. Huomaa, että sisätulon rajoittuma aliavaruuteen on sisätulo siinä aliavaruudessa.

Lemma 15.3.2. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus, $f\colon V \to V$ itseadjungoitu operaattori ja $W \subset V$ operaattorin f invariantti aliavaruus. Tällöin $f|_W\colon W \to W$ on sisätuloavaruuden $(W, \langle\cdot,\cdot\rangle)$ itseadjungoitu operaattori.*

Todistus. Olkoot $w, w' \in W$. Tällöin

$$\langle (f|_W)(w), w'\rangle = \langle f(w), w'\rangle = \langle w, f^*(w')\rangle = \langle w, f(w')\rangle = \langle w, (f|_W)(w')\rangle.$$

□

15.4 Spektraalilause

Tässä luvussa todistetaan spektraalihajotelma reaalisille itseadjungoiduille operaattoreille ja reaalisille symmetrisille matriiseille. Kompleksinen tapaus on käsitelty esimerkiksi Axlerin kirjan [1] luvussa 7.B.

Lause 15.4.1 (Reaalinen spektraalilause)**.** *Olkoon $(V, \langle \cdot, \cdot \rangle)$ reaalinen sisätuloavaruus ja $f : V \to V$ operaattori. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *f on itseadjungoitu,*
2. *avaruudella V on operaattorin f ominaisvektoreista koostuva ortonormaalikanta ja*
3. *on olemassa avaruuden V ortonormaalikanta, jossa operaattorin f esitysmatriisi on diagonaalimatriisi.*

Kirjataan spektraalilause myös matriisien tapauksessa. Ominaisuuksien johtaminen jätetään harjoitustehtäväksi.

Korollaari 15.4.2. *Olkoon $A \in \mathbb{R}^{n \times n}$ neliömatriisi. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *A on symmetrinen,*
2. *avaruudella $\mathbb{R}^{n \times 1}$ on matriisin A ominaisvektoreista koostuva ortonormaali kanta,*
3. *on olemassa sellainen ortogonaalimatriisi $P \in \mathbb{R}^{n \times n}$ ja sellainen diagonaalimatriisi $D \in \mathbb{R}^{n \times n}$, että*
$$A = PDP^t.$$

Lisäksi, tässä tapauksessa, matriisin D diagonaalielementit ovat matriisin A ominaisarvoja.

Huomautus 15.4.3. *Palautetaan mieleen, että mikäli $D = [d_{ji}] \in \mathbb{R}^{n \times n}$ on diagonaalimatriisi ja $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n \times n}$ on kääntyvä, niin matriisin D diagonaalialkiot ovat täsmälleen matriisin $A = PDP^t \in \mathbb{R}^{n \times n}$ ominaisarvot. Olkoon $1 \leq i \leq n$. Tällöin*

$$Av_i = PDP^{-1}v_i = PDe_i = P(d_{ii}e_i) = d_{ii}Pe_i = d_{ii}v_i.$$

Koska matriisin P sarakkeet ovat matriisin A ominaisvektoreita ja muodostavat avaruuden $\mathbb{R}^{n \times 1}$ kannana, niin matriisilla A ei ole muita ominaisarvoja.

Huomautus 15.4.4. *Koska matriisin D diagonaalialkiot ovat täsmälleen matriisin A ominaisarvot ja matriisin P sarakkeet ovat matriisin A vastaavat ominaisvektorit, niin* *tarvittaessa uudeelleen järjestämällä matriisin P sarakkeet voidaan matriisi D esittää ositettuna matriisina*

$$D = \begin{bmatrix} \lambda_1 I_{d_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{d_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k I_{d_k} \end{bmatrix},$$

missä $\lambda_1 > \cdots > \lambda_k$ ovat matriisin A ominaisarvot, $I_{d_j} \in \mathbb{R}^{d_j \times d_j}$ on identiteetti matriisi ja $d_j = \dim E(\lambda_j, A)$ jokaisella $j \in \{1, \ldots, k\}$.

Spektraalilause on merkittävä tulos: Se antaa täydellisen karakterisoinnin itseadjungoiduille operaattoreille ja symmetrisille matriiseille. Sen seurauksena tiedetään, että reaalisella symmetrisellä $n \times n$-matriisilla on n lineaarisesti riippumatonta ominaisvektoria ja että se on diagonalisoituva. Annetaan tästä ilmiöstä nyt konkreettinen esimerkki.

Esimerkki 15.4.5. *Olkoon*

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

Selvästi A on symmetrinen matriisi. Sen ominaisarvot voidaan selvittää yhtälöstä

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = (2-\lambda)^2 - 1.$$

Näin ollen matriisin A ominaisarvot ovat $\lambda_1 = 3$ ja $\lambda_2 = 1$. Näitä ominaisarvoja vastaavat ominaisvektorit ovat

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ ja } v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Nyt

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

on ortgonaalimatriisi, jolle pätee

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^t.$$

Huomaa, että ominaisarvojen järjestys eivätkä niitä vastaavat ominaisvektorit ole yksikäsitteisiä. Näin ollen myöskään matriisit P tai D eivät ole yksikäsitteisiä. Esimerkiksi myös matriisille

$$Q = \begin{bmatrix} -v_2 & v_1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

pätee

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right)^t.$$

Huomautus 15.4.6. *Yleensä sanotaan, että itseadjungoidulla operaattorilla* $f\colon V \to V$ *(tai ekvivalentisti symmetrisellä matriisilla) on* $\dim V$ *ominaisarvoa, kun ominaisarvot lasketaan geometrisen multiplisiteetin mukaan. Tällä tarkoitetaan sitä, että ominaisarvoille* $\lambda_1 > \cdots > \lambda_k$ *pätee*

$$\dim E(\lambda_1, f) + \cdots + \dim E(\lambda_k, f) = \dim V.$$

Dimensiota $\dim E(\lambda_j, f)$ *kutsutaan* ominaisarvon λ_j geometriseksi multiplisiteetiksi.

Huomautus 15.4.7. *Symmetrisen matriisin* A *potenssit voidaan helposti laskea diagonaaliesityksestä:*

$$A^2 = AA = PDP^tPDP^t = PD(P^tP)DP^t = PDDP^t = P(D^2)P^t.$$

Vastaavasti kaikilla $k \in \mathbb{N}$ *saadaan*

$$A^k = P(D^k)P^t,$$

missä matriisi D^k *on diagonaalimatriisi, jonka diagonaalielementit ovat matriisin* D *diagonaalielementtien potensseja. Jos matriisin* A *ominaisarvot ovat nollasta poikkeavia, niin lisäksi saadaan kaava*

$$A^{-1} = PD^{-1}P^t$$

matriisin A *käänteismatriisille, missä* D^{-1} *on diagonaalimatriisi, jonka diagonaalielementit ovat matriisin* D *diagonaalielementtien käänteislukuja.*

Spektraalilauseen todistuksessa tarvittavat tärkeimmät aputulokset ovat ominaisarvojen olemassaolo (Lause 15.2.3) ja invariantin aliavaruuden komplementin invarianttius (Lause 15.3.1).

Lauseen 15.4.1 todistus. Osoitetaan implikaatiot: (3) $\Rightarrow$ (1), (1) $\Rightarrow$ (2) ja (2) $\Rightarrow$ (3). Näistä ensimmäinen ja viimeinen implikaatio ovat helppoja.

Oletetaan, että oletus (3) pätee ja olkoon $(v_1, \ldots, v_n)$ sellainen avaruuden V ortonormaalikanta, että $f = \Phi \circ f_D \circ \Phi^{-1}$, missä $\Phi\colon \mathbb{R}^{n\times 1} \to V$ on isomorfismi $\Phi = \Phi_{(v_1,\ldots,v_n)}$ eli isomorfismi $e_i \mapsto v_i$, ja D on diagonaalimatriisi. Koska $D^t = D$, niin $f^* = f$, eli (1) pätee.

Oletetaan nyt, että (2) pätee ja olkoon $(v_1, \ldots, v_n)$ operaattorin f ominaisvektoreista koostuva ortonormaalikanta ja olkoot $\lambda_1, \ldots, \lambda_n$ vastaavat ominaisarvot. Olkoon $A \in \mathbb{R}^{n\times n}$ kuvauksen f esitysmatriisi tässä kannassa. Tällöin

$$\Phi(Ae_i) = \Phi(f_A(e_i)) = f(\Phi(e_i)) = f(v_i) = \lambda_i v_i = \lambda_i \Phi(e_i) = \Phi(\lambda_i e_i).$$

Näin ollen $Ae_i = \lambda_i e_i$. Matriisi A on siis ominaisarvoista $\lambda_1, \ldots, \lambda_n$ koostuva diagonaalimatriisi, eli (3) pätee.

Oletetaan lopuksi, että (1) on voimassa ja osoitetaan, että (2) pätee. Todistetaan väite induktiolla avaruuden V dimension suhteen. Väite pätee selvästi 1-ulotteisille avaruuksille. Olkoon nyt $n \in \mathbb{N}$ sellainen, että väite pätee kaikille avaruuksille W, joiden dimensio on korkeintaan n.

Olkoon V $(n+1)$-ulotteinen avaruus ja olkoon $f: V \to V$ itseadjungoitu operaattori. Lauseen 15.2.3 nojalla operaattorilla f on ominaisarvo $\lambda \in \mathbb{R}$. Olkoon $u \in V$ ominaisarvoa λ vastaava ominaisvektori, jolle pätee $\|u\| = 1$, ja olkoon $U = \mathrm{Sp}\{u\}$. Koska $fU \subset U$, niin lauseen 15.3.1 nojalla $fU^\perp \subset U^\perp$. Koska $\dim U > 0$, niin $\dim U^\perp < \dim V$. Koska $f|_{U^\perp}: U^\perp \to U^\perp$ on itseadjungoitu operaattori sisätuloavaruudessa $(U^\perp, \langle\cdot,\cdot\rangle)$ lemman 15.3.2 nojalla, niin induktio-oletuksesta seuraa, että on olemassa avaruuden $U^\perp$ ortonormaalikanta $(u_1, \ldots, u_n)$, joka koostuu operaattorin $f|_{U^\perp}$ ominaisvektoreista. Tällöin $(u, u_1, \ldots, u_n)$ on avaruuden V ortonormaalikanta, joka koostuu operaattorin f ominaisvektoreista. □

15.4.1 Spektraalilauseen ominaisavaruustulkinta

Spektraalilauseen mukaan mukaan avaruus V voidaan kirjoittaa suorana summana itseadjungoidun operaattorin $f: V \to V$ ominaisavaruuksista. Kirjataan tämä muotoilu omaksi tulokseekseen.

Lause 15.4.8 (Spektraalilauseen toinen muotoilu). *Olkoon $(V, \langle\cdot,\cdot\rangle)$ äärellisulotteinen sisätuloavaruus, $f: V \to V$ itseadjungoitu operaattori ja $\lambda_1 > \cdots > \lambda_k$ operaattorin f kaikki ominaisarvot. Tällöin operaattori $f: V \to V$ on itseadjungoitu, jos ja vain jos avaruus V on operaattorin f ominaisavaruuksien suora summa ja ominaisavaruudet ovat kohtisuorassa toisiaan vastaan eli*

$$V = E(\lambda_1, f) \oplus \cdots \oplus E(\lambda_k, f)$$

missä $\lambda_1 > \cdots > \lambda_k$ ovat operaattorin f ominaisarvot ja ominaisavaruudet $E(\lambda_i, f)$ ja $E(\lambda_j, f)$ ovat toisiaan vastaan kohtisuorassa kaikilla $i \neq j$.

Todistus. Oletetaan, että f on itsedjungoitu operaattori ja olkoot $\lambda_1 > \cdots > \lambda_k$ operaattorin f ominaisarvot. Tällöin spektraalilauseen (lause 15.4.1) mukaan on olemassa avaruuden V ortonormaali kanta $(v_1, \ldots, v_n)$, joka koostuu operaattorin f ominaisvektoreista. Järjestämällä kantavektorit tarvittaessa uudelleen, voidaan olettaa, että on olemassa sellaiset luvut $1 = i_1 < \cdots < i_k < i_{k+1} = n+1$, että jonon $(v_1, \ldots, v_n)$ osajonossa $(v_{i_j}, \ldots, v_{i_j - 1})$ on täsmälleen ominaisarvoa λ_j vastaavat jonon $(v_1, \ldots, v_n)$ vektorit.

Koska $E(\lambda_j, f) \supset \mathrm{Sp}(v_{i_j}, \ldots, v_{i_{j+1}-1})$ jokaisella $j \in \{1, \ldots, k\}$, niin

$$\begin{aligned} E(\lambda_1, f) \oplus \cdots \oplus E(\lambda_k, f) &\supset \mathrm{Sp}(v_{i_1}, \ldots, v_{i_2 - 1}) \oplus \cdots \oplus \mathrm{Sp}(v_{i_k}, \ldots, v_n) \\ &= \mathrm{Sp}(v_1, \ldots, v_n) = V. \end{aligned}$$

Oletetaan nyt, että $V = E(\lambda_1, f) \oplus \cdots \oplus E(\lambda_k, f)$, missä $\lambda_1 > \cdots > \lambda_k$ ovat operaattorin f ominaisarvot. Valitaan nyt jokaiselle $E(\lambda_j, f)$ ortonormaalikanta, $(v_{1j}, \ldots, v_{d_j, j})$, missä $d_j = \dim E(\lambda_j, f)$. Koska ominaisavaruudet $E(\lambda_j, f)$ ovat toisiaan vastaan kohtisuorassa, niin $(v_{11}, \ldots, v_{d_1 1}, v_{12}, \ldots, v_{d_k k})$ on avaruuden V ortonormaalikanta, joka koostuu operaattorin f ominaisvektoreista. Spektraalilauseen nojalla f on itseadjungoitu. □

Korollaari 15.4.9. *Olkoon $A \in \mathbb{R}^{n \times n}$ neliömatriisi ja olkoot $\lambda_1 > \cdots > \lambda_k$ matriisin A kaikki ominaisarvot. Tällöin A on symmetrinen, jos ja vain jos*

$$\mathbb{R}^{n \times 1} = E(\lambda_1, A) \oplus \cdots \oplus E(\lambda_k, A)$$

ja aliavaruudet $E(\lambda_i, A)$ ja $E(\lambda, A)$ ovat toisiaan vastaan kohtisuorassa kaikilla $i \neq j$.

Luku 16

Symmetrisen matriisin definiittisyys ja semidefiniittisyys

Tämä luku aloitetaan tarkastelemalla positiivisesti semidefiniittejä matriiseja ja niitä vastaavia positiivisia operaattoreita. Positiivisesti semidefiniitillä matriisilla on hyvin määritelty neliöjuuri, jonka avulla todistetaan positiivisesti semidefiniittien operaattorien polaarihajotelma. Teorian sovelluksena käsitellään Choleskyn hajotelmaa.

16.1 Symmetrisen neliömatriisin definiittisyys

Yksi tapa luokitella symmetriset matriiisit on tarkastella niiden definiittisyyttä.

Määritelmä 16.1.1. *Symmetrinen matriisi $A \in \mathbb{R}^{n\times n}$ on*

- positiivisesti semidefiniitti*, jos $x^tAx \geq 0$ kaikilla vektoreilla $x \in \mathbb{R}^{n\times 1}$,*
- positiivisesti definiitti*, jos $x^tAx > 0$ kaikilla nollasta poikkeavilla vektoreilla $x \in \mathbb{R}^{n\times 1}$,*
- negatiivisesti semidefiniitti*, jos $x^tAx \leq 0$ kaikilla vektoreilla $x \in \mathbb{R}^{n\times 1}$,*
- nagatiivisesti definiitti*, jos $x^tAx < 0$ kaikilla nollasta poikkeavilla vektoreilla $x \in \mathbb{R}^{n\times 1}$ ja*
- indefiniitti*, jos A ei ole positiivisesti tai negatiivisesti semidefinitiitti.*

Tärkein esimerkki semidefiniitistä matriisista on seuraava.

Esimerkki 16.1.2. *Olkoon $E \in \mathbb{R}^{m\times n}$ matriisi. Tällöin matriisi $A = E^tE \in \mathbb{R}^{n\times n}$ on positiivisesti semidefiniitti. Tämä seuraa huomiosta, että kaikilla $x \in \mathbb{R}^{n\times 1}$ pätee*

$$x^tAx = x^tE^tEx = (Ex)^t(Ex) = Ex \cdot Ex \geq 0.$$

Huomautus 16.1.3. *Jos matriisi $A \in \mathbb{R}^{n\times n}$ on negatiivisesti (semi)definiitti, niin tällöin $B = -A$ on on positiviisesti (semi)definiitti. Näin ollen usein tarkastellaan ainoastaan positiivisesti (semi)definiittejä matriiseja.*

Huomautus 16.1.4. *Symmetrinen matriisi A on indefiniitti, jos on olemassa sellaiset $x \in \mathbb{R}^{n\times n}$ ja $y \in \mathbb{R}^{n\times n}$, että $x^tAx > 0$ ja $y^tAy < 0$.*

Huomautus 16.1.5. *Symmetristen neliömatriisien luokittelussa esiintyvää funktiota $q_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$,*

$$x \mapsto x^tAx,$$

kutsutaan matriisin A neliömuodoksi *(engl.* quadratic form*). Näin ollen neliömetriisien luokittelu (semi)definiitteihin ja indefiniitteihin neliömatriiseihin on itseasiassa näiden neliömuotojen luokittelu.*

Neliömuodoilla on tärkeä rooli klassisessa geometriassa, esimerkiksi kartioleikkausten teoriassa, mutta myös pääakselien teoriassa. Liitteessä F.2 raapaistaan hieman tämän aiheen pintaa.

16.2 Esimerkkejä: Sarakevaruuksien sisätulot

Tämän kurssin kannalta tärkeimpiä esimerkkejä ovat sarakeavaruuksien sisätulot. Niitä voidaan konstruoida neliömatriisien avulla.

Lause 16.2.1. *Olkoon $A \in \mathbb{R}^{n\times n}$ positiivisesti definiitti neliömatriisi. Tällöin funktio $\langle\cdot,\cdot\rangle\colon \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1} \to \mathbb{R}$, joka on määritelty kaavalla*

$$\langle x, y\rangle = x^tAy$$

kaikilla $x, y \in \mathbb{R}^{n\times 1}$, on sisätulo sarakeavaruudessa $\mathbb{R}^{n\times 1}$.

Huomautus 16.2.2. *Koska x^tAy on vektoreiden x ja Ay pistetulo, niin lauseen 16.2.1 sisätulo $\langle\cdot,\cdot\rangle$ voidaan myös kirjoittaa pistetulon avulla muodossa $\langle x, y\rangle = x \cdot Ay$ kaikilla $x, y \in \mathbb{R}^{n\times 1}$.*

Lauseen 16.2.1 todistus. Matriisitulon ja transpoosin ominaisuuksien perusteella kaikilla $x, x', y \in \mathbb{R}^{n\times 1}$ ja $a \in \mathbb{R}$ pätee

$$\langle ax + x', y\rangle = (ax + x')^tAy = (ax^t + (x')^t)Ay = ax^tAy + (x')^tAy = a\langle x, y\rangle + \langle x', y\rangle.$$

Lisäksi kaikilla $x, y \in \mathbb{R}^{n\times 1}$ pätee

$$\langle y, x\rangle = y^tAx = y^tA^tx = (Ay)^tx = (x^tAy)^t = x^tAy = \langle x, y\rangle.$$

Huomaa, että x^tAy on luku eli 1×1-matriisi.

Koska $x^tAx > 0$ kaikilla $x \neq 0$ ja $x^tAx = 0$ nollavektorille $x = 0$, niin $\langle x, x\rangle \geq 0$ kaikilla $x \in \mathbb{R}^{n\times 1}$ ja $\langle x, x\rangle = 0$ jos ja vain jos $x = 0$.

Näin ollen $\langle\cdot,\cdot\rangle$ on sisätulo. □

Itseasiassa kaikki sarakeavaruuden $\mathbb{R}^{n\times 1}$ sisätulot voidaan kirjoittaa lauseen 16.2.1 ehdot täyttävän matriisin ja pistetulon avulla.

Lause 16.2.3. *Olkoon* $\langle\cdot,\cdot\rangle$ *sarakeavaruuden* $\mathbb{R}^{n\times 1}$ *sisätulo. Tällöin on olemassa sellainen matriisi* $A \in \mathbb{R}^{n\times n}$, *että*
$$\langle x, y\rangle = x^t Ay.$$
Lisäksi matriisille A *pätee sekä* $A^t = A$ *että* $x^t Ax > 0$ *kaikilla* $x \in \mathbb{R}^{n\times 1}$.

Todistus. Olkoon $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ matriisi
$$a_{ji} = \langle e_j, e_i\rangle,$$
missä $(e_1, \ldots, e_n)$ on avaruuden $\mathbb{R}^{n\times 1}$ standardikanta. Osoitetaan, että A on haluttu matriisi.

Olkoot $x = x_1e_1 + \cdots + x_ne_n \in \mathbb{R}^{n\times 1}$ ja $y = y_1e_1 + \cdots + y_ne_n \in \mathbb{R}^{n\times 1}$. Tällöin
$$\begin{aligned}\langle x, y\rangle &= \sum_{j=1}^{n}\sum_{i=1}^{n}\langle x_je_j, y_ie_i\rangle = \sum_{j=1}^{n}\sum_{i=1}^{n} x_jy_i\langle e_j, e_i\rangle \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} x_jy_ia_{ji} = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ji}y_i = \sum_{j=1}^{n} x_j(Ay)_j = x^t Ay.\end{aligned}$$

Matriisin A halutut ominaisuudet seuraavat sisätulon $\langle\cdot,\cdot\rangle$ ominaisuuksista seuraavasti. Koska kaikilla $j, i \in \{1, \ldots, n\}$ pätee
$$a_{ij} = \langle e_i, e_j\rangle = \langle e_j, e_i\rangle = a_{ji},$$
niin $A^t = A$. Vastaavasti kaikilla nollasta poikkeavilla vektoreilla $x \in \mathbb{R}^{n\times 1}$ pätee
$$x^t Ax = \langle x, x\rangle > 0.$$
□

Matriisin osoittaminen positiivisesti definiitiksi on helpointa tehdä käyttäen spektraalilausetta. Tähän palataan myöhemmin tässä luvussa. Annetaan kuitenkin tässä vaiheessa jo esimerkki mielenkiintoisesta positiivisesti definiitistä 2×2-matriisista.

Esimerkki 16.2.4. *Olkoon*
$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$
esimerkin 15.4.5 matriisi.

Osoitetaan, että $x^t Ax > 0$ *kaikilla nollaista poikkeavilla vektoreilla* $x \in \mathbb{R}^{n\times 1}$. *Olkoon* $x = x_1e_1 + x_2e_2$. *Tällöin*
$$\begin{aligned}x^t Ax &= \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix} \\ &= x_1(2x_1 + x_2) + x_2(x_1 + 2x_2) \\ &= 2x_1^2 + x_1x_2 + x_1x_2 + 2x_2^2.\end{aligned}$$

Koska

$$\begin{aligned} 2x_1^2 + x_1x_2 + x_1x_2 + 2x_2^2 &= x_1^2 + x_2^2 + x_1^2 + 2x_1x_2 + x_2^2 \\ &\geq x_1^2 + x_2^2 + x_1^2 - 2|x_1||x_2| + x_2^2 \\ &= x_1^2 + x_2^2 + (|x_1| - |x_2|)^2 \geq x_1^2 + x_2^2 > 0, \end{aligned}$$

niin $x^tAx > 0$.

16.3 Positiivisesti semidefiniitit matriisit ja positiiviset operaattorit

On luontevaa ajatella, että matriisin definiittisyyden tai semidefiniittisyyden voi selvittää suoraan matriisin kertoimista. Näin saadaankin riittävä ehto. Koska jokaiselle matriisille $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ pätee

$$a_{ii} = e_i^t A e_i$$

kaikilla $i \in \{1, \ldots, n\}$, niin matriisin definiittisyys tai semidefiniittisyys kertoo diagonaalialkioiden merkin. Tämä tieto ei kuitenkaan ole riittävä esimerkiksi definiittisyyden ja semidefiniittisyyden erottamiseksi.

Esimerkki 16.3.1. *Matriisi*

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

on positiivisesti semi-definiitti matriisi, jonka diagonaalialkiot ovat positiivisia.

Spektraalilauseen avulla on helppo havaita, että operaattori on positiviinen, jos ja vain jos sen ominaisarvot ovat ei-negatiivisia.

Lause 16.3.2. *Olkoon $A \in \mathbb{R}^{n\times n}$ symmetrinen matriisi. Tällöin A on positiivisesti semidefiniitti, jos ja vain jos matriisin A ominaisarvot ovat ei-negatiivisia. Vastaavasti matriisi A on positiivisesti definiitti, jos ja vain jos matriisin A ominaisrvot ovat positiivisia.*

Todistus. Oletetaan, että A on positiviisesti semidefiniitti. Olkoon $\lambda \in \mathbb{R}$ matriisin A ominaisarvo ja $x \in \mathbb{R}^{n\times 1}$ vastaava ominaisvektori. Tällöin

$$0 \leq x^tAx = x^t(\lambda x) = \lambda x^t x.$$

Koska $x^tx > 0$, niin $\lambda \geq 0$.

Oletetaan nyt, että A on symmetrinen matriisi, jonka ominaiarvot ovat ei-negatiivisia. Tällöin on olemassa sellainen ortogonaalimatriisi $P \in \mathbb{R}^{n\times n}$ ja sellainen diagonaalimatriisi $D \in \mathbb{R}^{n\times n}$, jonka diagonaalialkiot $\lambda_1, \ldots, \lambda_n$ ovat matriisin A ominaisarvoja, että $A = PDP^t$. Lisäksi matriisin P sarakkeet $(v_1, \ldots, v_n)$ muodostavat avaruuden $\mathbb{R}^{n\times 1}$ kannan. Olkoon nyt $x \in \mathbb{R}^{n\times 1}$ ja olkoot $a_i, \ldots, a_n \in \mathbb{R}$ sellaisia, että

$x = a_1 v_1 + \cdots + a_n v_n$. (Itseasiassa $a_i = x \cdot v_i$ jokaisella $i \in \{1, \ldots, \}$.) Olkoon $a \in \mathbb{R}^{n\times 1}$ sarakevektori $a = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^t$. Tällöin $x = Pa$ ja saadaan

$$x^t Ax = (Pa)^t PDP^t (Pa) = a^t P^t PDP^t Pa = a^t Da = \lambda_1 a_1^2 + \cdots + \lambda_n a_n^2 \geq 0.$$

Matriisi A on siis positiivisesti semidefiniitti.

Vastaavat päättelyt antavat väitteen toisen osan, että matriisi A on positiivisesti definiitti, jos ja vain jos matriisin A ominaisrvot ovat positiivisia. □

Esimerkki 16.3.3. *Esimerkin 15.4.5 matriisi*

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

on positiivisesti definiittinen.

Esimerkki 16.3.4. *Olkoon*

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

Koska

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = -\lambda(2-\lambda) - 1 = \lambda^2 - 2\lambda - 1 = (\lambda - 1)^2 - 2$$

niin matriisin A ominaisarvot ovat $\lambda_1 = 1 + \sqrt{2}$ ja $\lambda_2 = 1 - \sqrt{2}$. Matriisi A ei siis ole positiivisesti semidefiniittinen, vaan itseasiassa indefiniittinen.

Edellinen lause yhdessä spektraalilauseen kanssa karakterisoi positiivisesti semidefiniitin matriisin kääntyvyyden.

Korollaari 16.3.5. *Olkoon $A \in \mathbb{R}^{n\times n}$ positiivisesti semidefiniitti matriisi. Tällöin A on kääntyvä, jos ja vain jos A on positiivisesti definiitti.*

16.3.1 Positiiviset operaattorit

Positiivisesti semidefiniittit matriisit vastaavat (terminologialtaan harhaanjohtavasti) positiivisia operattoreita.

Määritelmä 16.3.6. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus. Lineaarioperaattori $f\colon V \to V$ on* positiivinen*, jos f on itseadjungoitu ja $\langle f(v), v\rangle \geq 0$ kaikilla $v \in V$.*

Lause 16.3.7. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus. Tällöin operaattori $f\colon V \to V$ on positiivinen, jos ja vain jos on olemassa avaruuden V kanta, jossa kuvauksen f matriisi on positiviisesti semidefiniitti.*

Todistus. Olkoon $(v_1, \ldots, v_n)$ avaruuden V kanta ja $\Phi_{(v_1,\ldots,v_n)} \colon \mathbb{R}^{n\times 1} \to V$ isomorfismi $e_i \mapsto v_i$. Tällöin $f = \Phi_{(v_1,\ldots,v_n)} \circ f_A \circ \Phi^{-1}_{(v_1,\ldots,v_n)}$, missä $A \in \mathbb{R}^{n\times n}$ on kuvauksen f esitysmatriisi tässä kannassa. Koska $\Phi_{(v_1,\ldots,v_n)}$ on isometria, niin jokaisella $x \in \mathbb{R}^{n\times 1}$ pätee

$$x^t Ax = x \cdot f_A(x) = \langle \Phi_{(v_1,\ldots,v_n)}(x), \Phi_{(v_1,\ldots,v_n)}(f_A(x))\rangle = \langle \Phi_{(v_1,\ldots,v_n)}(x), f(\Phi_{(v_1,\ldots,v_n)}(x))\rangle.$$

Näin ollen f on positiviinen, jos ja vain jos A on positiivisesti semidefiniitti. □

Tyypillisin esimerkki positiivisesta operaattorista saadaan symmetrisoimalla operaattori omalla adjungaatillaan. Tämä vastaa matriisin kertomista omalla transpoosillaan. Kirjatataan tämä havainto tärkeytensä vuoksi omaksi lemmakseen.

Lemma 16.3.8. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ sisätuloavaruuksia ja $f \colon V \to W$ lineaarikuvaus. Tällöin $f^*f = f^* \circ f \colon V \to V$ on positiivinen operaattori.*

Todistus. Osoitetaan ensin, että f^*f on itseadjungoitu. Olkoot $v, v' \in V$. Tällöin adjungaatin määritelmän nojalla

$$\langle f^*f(v), v'\rangle_V = \langle f(v), f(v')\rangle_W = \langle v, f^*f(v')\rangle_V.$$

Näin ollen $(f^*f)^* = f^*f$ eli f^*f on itseadjungoitu.

Osoitetaan nyt, että f^*f on positiivinen operaattori. Olkoon $v \in V$. Tällöin

$$\langle f^*(f(v)), v\rangle_V = \langle f(v), f(v)\rangle_W = \|f(v)\|_W^2 \geq 0.$$

□

Helpoin esimerkki positiivisesta operaattorista saadaan, kun tarkastellaan skaalausta vakiolla. Huomaa, että seuraava esimerkki ei riipu lainkaan valitusta sisätulosta.

Esimerkki 16.3.9. *Operaattori $f \colon V \to V$, $v \mapsto \lambda v$, on positiivinen kaikilla $\lambda > 0$. Tämä seuraa suoraan havainnosta, että*

$$\langle f(v), w\rangle = \langle \lambda v, w\rangle = \langle v, \lambda w\rangle = \langle v, f(w)\rangle$$

ja

$$\langle f(v), v\rangle = \langle \lambda v, v\rangle = \lambda\|v\|^2 \geq 0$$

kaikilla $v \in V$ ja $w \in V$.

Tätä esimerkkiä voidaan viedä pidemmälle ja todeta, että operaattori on positiivinen, jos ja vain jos se on itseadjungoitu ja sen ominaisarvot ovat ei-negatiivisia. Tämän tuloksen voi todistaa käyttämällä vastaavia semidefiniittien matriisien tuloksia. Todistetaan se kuitenkin suoraan spektraalilausetta (lause 15.4.1) käyttäen.

Lause 16.3.10. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus. Tällöin operaattori $f \colon V \to V$ on positiivinen, jos ja vain jos on olemassa avaruuden V ortonormaali kanta $(v_1, \ldots, v_n)$, joka koostuu operaattorin f ominaisvektoreista ja operaattorin f ominaisarvot ovat ei-negatiivisia.*

Todistus. Oletetaan, että f on positiviinen operaattori. Koska f on itseadjungoitu, niin spektraalilauseen nojalla on olemassa avaruuden V ortonormaali kanta $(v_1, \ldots, v_n)$, joka koostuu operaattorin f ominaisvektoreista. Olkoon nyt $\lambda_i \in \mathbb{R}$ ominaisvektoria v_i vastaava ominaisarvo. Koska f on positiiviinen, niin

$$0 \leq \langle f(v_i), v_i \rangle = \langle \lambda_i v_i, v_i \rangle = \lambda_i \|v_i\|.$$

Näin ollen $\lambda_i \geq 0$. Koska operaattorilla ei ole muita ominaisarvoja, operaattorin f ominaisarvot ovat ei-negatiivisia.

Oleteaan nyt, että f on sellainen operaattori, että on olemassa avaruuden V kanta $(v_1, \ldots, v_n)$, joka koostuu operaattorin f ominaisvektoreista ja että ominaisvektoria v_i vastaava ominaisarvo λ_i on ei-negatiivinen.

Spektraalilauseen mukaan f on itseadjungoitu. Riittää siis osoittaa, että $\langle f(v), v \rangle \geq 0$ jokaisella $v \in V$. Olkoon $v = a_1 v_1 + \cdots + a_n v_n \in V$. Tällöin

$$f(v) = a_1 f(v_1) + \cdots + a_n f(v_n) = a_1 \lambda_1 v_1 + \cdots + a_n \lambda_n v_n.$$

Koska kanta $(v_1, \ldots, v_n)$ on ortonormaali, niin

$$\langle f(v), v \rangle = \langle a_1 \lambda_1 v_1 + \cdots + a_n \lambda_n v_n, a_1 v_1 + \cdots + a_n v_n \rangle = \lambda_1 a_1^2 + \cdots + \lambda_n a_n^2 \geq 0.$$

Tämä päättää todistuksen. $\square$

16.4 Positiivisesti semidefiniitin neliömatriisin neliöjuuri

Matriisin neliöjuuren määritelmä on seuraava.

Määritelmä 16.4.1. *Matriisi $B \in \mathbb{R}^{n \times n}$ on matriisin $A \in \mathbb{R}^{n \times n}$ neliöjuuri, jos $B^2 = A$.*

Neliömatriisilla ei tarvitse olla yhtään neliöjuurta tai niitä voi olla useita.

Esimerkki 16.4.2. *Osoitetaan, että matriisilla*

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

ei ole neliöjuurta. Olkoon

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2 \times 2}.$$

Tällöin

$$B^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{bmatrix} = \begin{bmatrix} a^2 + bc & b(a + d) \\ c(a + d) & bc + d^2 \end{bmatrix}.$$

Yhtälöstä $A = B^2$ seuraa välittömästi, että joko $b = 0$ tai $a + d = 0$. Jos $b = 0$, niin tällöin $a = 0$ ja $d = 0$ eli $a + d = 0$. Tämä on kuitenkin ristiriidassa vaatimuksen $c(a + d) = 1$ kanssa.

Esimerkki 16.4.3. *Identtisen matriisin*

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

neliöjuuria ovat

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ ja } \begin{bmatrix} 0 & 1/2 \\ 2 & 0 \end{bmatrix}.$$

Positiivisesti semidefiniiteilla matriiseilla on aina neliöjuuri ja näiden neliöjuurien joukossa on yksikäsitteinen neliöjuuri, joka on positiivisesti semidefiniittinen. Tämän tuloksen olemassaolo osa on helppo seuraus spektraalilauseesta. Yksikäsitteisyys on hieman haastavampi ja perustellaan seuraavassa luvussa. Aloitetaan olemassaolo tuloksen todistus teknisellä huomiolla.

Lemma 16.4.4. *Olkoon $P \in \mathbb{R}^{n\times n}$ ortogonaalimatriisi ja*

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \in \mathbb{R}^{n\times n}$$

diagonaalimatriisi, jonka diagonaalialkiot ovat ei-negatiivisia. Tällöin matriisille $B = PDP^t$ pätee

$$B^2 = PD^2P^t.$$

Lisäksi matriisin D diagonaalialkiot ovat matriisin B ominaisarvoja, matriisin D^2 diagonaalialkiot ovat matriisin B^2 ominaisarvoja ja matriisin P sarakkeet ovat matriisien B ja B^2 ominaisvektoreita.

Todistus. Koska $P^tP = I$, niin

$$B^2 = PDP^tPDP^t = PDDP^t = PD^2P^t.$$

Osoitetaan nyt, että matriisin P sarakkeet ovat molempien matriisien ominaisvektoreita. Olkoon $P = [v_1 \cdots v_n]$. Tällöin jokaisella $i \in \{1, \dots, n\}$ pätee

$$Bv_i = (PDP^t)v_i = PD(P^tv_i) = PDe_i = P(\lambda_ie_i) = \lambda_iPe_i = \lambda_iv_i$$

eli matriisin P sarakkeet ovat matriisin B ominaisvektoreita ja matriisin D diagonaalialkiot ovat niitä vastaavat ominaisarvot. Vastavasti

$$B^2v_i = BBv_i = B(\lambda_iv_i) = \lambda_iBv_i = \lambda_i^2v_i.$$

□

Lause 16.4.5. *Olkoon $A \in \mathbb{R}^{n\times n}$ positiivisesti semidefiniitti matriisi. Tällöin on olemassa positiivisesti semidefiniitti matriisi $B \in \mathbb{R}^{n\times n}$, jolle pätee $B^2 = A$.*

Todistus. Spektraalilauseen nojalla matriisilla A on ominaisarvot $\lambda_1, \ldots, \lambda_n$ sekä sellainen ortogonaalimatriisi $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n \times n}$, joka koostuu matriisin A ominaisvektoreista, ja sellainen diagonaalimatriisi $D \in \mathbb{R}^{n \times n}$, jonka diagonaalialkiot ovat matriisin A ominaisarvoja, että $A = PDP^t$. Merkitään

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Lauseen 16.3.2 perusteella ominaisarvot $\lambda_1, \cdots, \lambda_n$ ovat ei-negatiivisia. Näin ollen matriisi

$$D^{1/2} = \begin{bmatrix} \lambda_1^{1/2} & & \\ & \ddots & \\ & & \lambda_n^{1/2} \end{bmatrix}$$

on hyvin määritelty, jolle matriisitulon määritelmän nojalla pätee

$$D^{1/2} D^{1/2} = D.$$

Näin ollen matriisi $B = PD^{1/2}P^t$ on positiivisesti semidefiniitti matriisi, jolle pätee lemman 16.4.4 perusteella pätee $B^2 = A$. □

Lauseen 16.4.5 todistus antaa itseasiassa metodin positiivisesti semidefiniitin neliöjuuren löytäselle. Kirjoitetaan tämä metodi vielä esimerkin muotoon.

Esimerkki 16.4.6. *Olkoon*

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

esimerkin 15.4.5 matriisi. Tällöin

$$A = UDU^t,$$

missä

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

on ortogonaalimatriisi ja

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

Matriisilla A on siis positiviisesti definiittinen neliöjuuri

$$\begin{aligned} B = UD^{1/2}U^t &= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^t \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} \sqrt{3} & 1 \\ \sqrt{3} & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & \sqrt{3}+1 \end{bmatrix}. \end{aligned}$$

Tarkistetaan vielä saatu tulos

$$\begin{aligned} B^2 &= \left(\frac{1}{2}\begin{bmatrix}\sqrt{3}+1 & \sqrt{3}-1\\ \sqrt{3}-1 & \sqrt{3}+1\end{bmatrix}\right)\left(\frac{1}{2}\begin{bmatrix}\sqrt{3}+1 & \sqrt{3}-1\\ \sqrt{3}-1 & \sqrt{3}+1\end{bmatrix}\right)\\ &= \frac{1}{4}\begin{bmatrix}\sqrt{3}+1 & \sqrt{3}-1\\ \sqrt{3}-1 & \sqrt{3}+1\end{bmatrix}\begin{bmatrix}\sqrt{3}+1 & \sqrt{3}-1\\ \sqrt{3}-1 & \sqrt{3}+1\end{bmatrix}\\ &= \frac{1}{4}\begin{bmatrix}(\sqrt{3}+1)^2+(\sqrt{3}-1)^2 & 2(\sqrt{3}^2-1^2)\\ 2(\sqrt{3}^2-1^2) & (\sqrt{3}-1)^2+(\sqrt{3}+1)^2\end{bmatrix}\\ &= \frac{1}{4}\begin{bmatrix}8 & 4\\ 4 & 8\end{bmatrix} = A. \end{aligned}$$

Kuten edellä mainittiin, positiivisesti semidefiniitillä matriisilla on täsmälleen yksi positiivisesti semidefiniitti neliöjuuri. Edellisen esimerkin metodi antaa siis aina tämän halutun neliöjuuren. Tämä tulos on luonnollista osoittaa käyttäen ominaisavaruuksia. Siksi käsittely vaihdetaan operaattorien kielelle.

16.4.1 Positiivisen operaattorin positiivinen neliöjuuri

Positiivisesti semidefiniitin matriisin neliöjuurta vastaa positiivisen operaattorin neliöjuuri.

Määritelmä 16.4.7. *Olkoon V vektoriavaruus. Operaattori $g\colon V \to V$ on operaattorin $f\colon V \to V$* neliöjuuri*, jos $g^2 = g \circ g = f$.*

Lauseen 16.4.5 perusteella positiviisella operaattorilla on siis aina neliöjuuri. Osoitetaan nyt, että näiden neliöjuurien joukossa on täsmälleen yksi, joka on positiivinen operaattori.

Lause 16.4.8. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus ja $f\colon V \to V$ positiivinen operaattori. Tällöin on olemassa yksikäsitteinen positiivinen operaattori $g\colon V \to V$, jolle pätee $f = g^2 = g \circ g\colon V \to V$.*

Aloitetaan aputuloksella, että operaattorilla $f\colon V \to V$, $v \mapsto \lambda v$, yksikäsitteinen positiivinen neliöjuuri, jos $\lambda > 0$.

Lemma 16.4.9. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus ja $\lambda > 0$. Tällöin operaattorin $f\colon V \to V$, $v \mapsto \lambda v$, ainoa positiivinen neliöjuuri on operaattori $g\colon V \to V$, $v \mapsto \sqrt{\lambda}v$.*

Todistus. Koska $g^2(v) = g(g(v)) = g(\sqrt{\lambda}v) = \sqrt{\lambda}^2 v = \lambda v = f(v)$ kaikilla $v \in V$, niin g on operaattorin f neliöjuuri. Selvästi operaattori g on myös positiivinen.

Olkoon nyt $h\colon V \to V$ sellainen positiivinen operaattori, että $h^2 = f$. Osoitetaan, että $h = g$.

Koska h on positiivinen operaattori, niin on olemassa avaruuden V ortonormaalikanta $(v_1, \ldots, v_n)$, joka koostuu operaattorin h ominaisvektoreista. Olkoon jokaisella

$i \in \{1, \ldots, n\}$ luku $\mu_i \in \mathbb{R}$ ominaisvektoria v_i vastaava ominaisarvo eli $g(v_i) = \mu_i v_i$ jokaisella $i \in \{1, \ldots, n\}$. Koska jokaisella $i \in \{1, \ldots, n\}$ pätee toisaalta

$$f(v_i) = \lambda v_i$$

ja toisaalta

$$f(v_i) = h^2(v_i) = \mu_i^2 v_i,$$

niin $\mu_i = \sqrt{\lambda}$ jokaisella $i \in \{1, \ldots, n\}$.

Osoitetaan, että tästä seuraa $h = g$. Olkoon $v = a_1 v_1 + \cdots + a_n v_n \in V$. Tällöin

$$h(v) = a_1 h(v_1) + \cdots + a_n h(v_n) = a_1 \sqrt{\lambda} v_i + \cdots + a_n \sqrt{\lambda} v_n = \sqrt{\lambda} v = g(v).$$

Tämä päättää todistuksen. □

Edellinen todistus vihjaa, että lauseen 16.4.8 väite seuraa, jos osoitetaan, että positiivisella operaattorilla ja sen positiivisella neliöjuurella on samat ominaisavaruudet. Kirjataan tämä tulos lemmaksi. Huomaa, että lemman todistuksen oleellinen vaihe on huomio, positiivisen operaattorin rajoittuma ominaisavaruuteen on positiivinen operaattori.

Lemma 16.4.10. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ sisätuloavaruus, $f\colon V \to V$ positiivinen operaattori, $h\colon V \to V$ operaattorin f positiivinen neliöjuuri ja $\lambda \geq 0$ operaattorin f ominaisarvo. Tällöin $\sqrt{\lambda}$ on operaattorin h ominaisarvo ja*

$$E(\sqrt{\lambda}, h) = E(\lambda, f).$$

Todistus. Koska jokaisella $v \in E(\sqrt{\lambda}, h)$ pätee $f(v) = h(h(v)) = h(\sqrt{\lambda} v) = \lambda v$, niin $E(\sqrt{\lambda}, h) \subset E(\lambda, f)$. Näin ollen on jäljellä osoittaa, että $\sqrt{\lambda}$ on operaattorin h ominaisarvo ja sisältyminen toiseen suuntaan

Koska h on positiivinen operaattori, niin spektraalilauseen nojalla on olemassa sellainen avaruuden V ortonormaali kanta $(v_1, \ldots, v_n)$, joka koostuu operaattorin h ominaisvektoreista. Olkoon μ_i ominaisvektoria v_i vastaava ominaisarvo. Olkoon nyt $v = a_1 v_1 + \cdots + a_n v_n \in E(\lambda, f)$. Tällöin toisaalta

$$f(v) = \lambda v = a_1 \lambda v_1 + \cdots + a_n \lambda v_n$$

ja toisaalta

$$f(v) = h^2(v) = a_1 \mu_1^2 v_1 + \cdots + a_n \mu_n^2 v_n.$$

Näin ollen jokaisella $i \in \{1, \ldots, n\}$ pätee joko $a_i = 0$ tai $\mu_i = \sqrt{\lambda}$. Näin ollen $v \in E(\sqrt{\lambda}, h)$. □

Kirjataan lauseen 16.4.8 todistusta varten vielä yksi lemma, joka on versio lauseesta 16.4.5 lineaarikuvauksille. Lemman todistus jätetään lukijalle harjoitustehtäväksi.

Lemma 16.4.11. *Olkoon $(V, \langle\cdot,\cdot\rangle)$ äärellisulotteinen sisätuloavaruus, $f\colon V \to V$ positiivinen operaattori ja olkoot $\lambda_1 < \cdots < \lambda_n$ sen ominaisarvot. Tällöin kuvaus $g\colon V \to V$, joka on määritelty kaavalla*

$$g(v_1 + \cdots + v_n) = \sqrt{\lambda_1}v_1 + \cdots + \sqrt{\lambda_n}v_n$$

kaikilla $v_i \in E(\lambda_i, f)$ ja $i \in \{1, \ldots, n\}$, on hyvin määritelty positiviinen operaattori, joka on operaattorin f neliöjuuri.

Lauseen 16.4.8 todistus. Olkoon $f\colon V \to V$ positiivinen operaattori ja olkoot $\lambda_1 < \cdots < \lambda_n$ sen ominaisarvot. Merkitään $V_i = E(\lambda_i, f)$ jokaisella $i \in \{1, \ldots, n\}$. Lemman 16.4.11 perusteella operaattori $g\colon V \to V$, joka on määritelty kaavalla

$$g(v_1 + \cdots + v_n) = \sqrt{\lambda_1}v_1 + \cdots + \sqrt{\lambda_n}v_n$$

kaikilla $v_i \in V_i$, on hyvin määritelty operaattorin f positiivinen neliöjuuri.

Olkoon $h\colon V \to V$ operaattorin f positiivinen neliöjuuri. Osoitetaan, että $h = g$. Olkoon $i \in \{1, \ldots, n\}$. Lemman 16.4.10 nojalla $E(\sqrt{\lambda_i}, h) = E(\lambda_i, f) = V_i$. Näin ollen $h|_{V_i}\colon V_i \to V_i$ on operaattorin $f|_{V_i}\colon V_i \to V_i$ positiivinen neliöjuuri. Lemman 16.4.9 nojalla $h|_{V_i}$ on kuvaus $v \mapsto \sqrt{\lambda_i}v_i$ eli $h(v) = g(v)$ jokaisella $v \in V_i$. Väite seuraa nyt kuvausten g ja h lineaarisuudesta. □

Huomautus 16.4.12. *Lemmoihin tuskastunut lukija voi jäädä miettimään, että voisiko lauseen 16.4.8 todistaa suoraviivaisemminkin. Kyllä voi. Lemmoissa 16.4.9 ja 16.4.10 tehdyt laskut voi yhdistää ja osoittaa suoraan, että jos $(v_1, \ldots, v_n)$ on avaruuden V ortonormaali kanta, joka koostuu operaattorin f ominaisvektoreista, niin $h(v_i) = \sqrt{\lambda_i}v_i$, missä λ_i on ominaisvektoria v_i vastaava ominaisarvo. Tämä suora todistus kuitenkin hämärtää hieman neliöjuurioperaattorin rakennetta.*

16.5 Sovellus: Choleskyn hajotelma

Sovelletaan nyt neliöjuurien teoriaa positiivisesti semi-definiittien matriisien esittämiseen kolmiomatriisien avulla. Cholesky hajotelmalla tarkoitetaan positiivisesti semi-definitiin matriisin $A \in \mathbb{R}^{n\times n}$ esittämistä kahden kolmiomatriisin tulona

$$A = R^tR$$

missä R on yläkolmiomatriisi, jonka diagonaalielementit ovat ei-negatiivisia. Koska Choleskyn hajotelma kirjoitetaan usein yläkolmiomatriisin R sijaan alakolmiomatriisin $T = R^t$ avulla muodossa

$$A = TT^t,$$

niin jatkossa käytetään tätä esitystapaa Choleskyn hajotelmasta. Choleskyn hajotelman tarkka määritelmä on seuraava.

Määritelmä 16.5.1. *Positiivisesti semi-definiitin matriisin $A \in \mathbb{R}^{n\times n}$* Choleskyn hajotelma *on $A = TT^t$, missä $T \in \mathbb{R}^{n\times m}$ on alakolmiomatriisi, jolle pätee $t_{jj} \geq 0$ kaikilla $j = 1, \ldots, n$.*

Huomautus 16.5.2. *Määritelmässä oletetaan, että matriisi A on positiivisesti semidefiniitti. Tämä ei ole lisäoletus matriisille A, sillä kaikki matriisit, jotka voidaan kirjoittaa muodossa TT^t, jollakin alakolmiomatriisilla T, ovat positiivisesti semidefiniittejä. Tämä seuraa suoraan (tutusta) huomioista, että kaikilla $x \in \mathbb{R}^{n\times 1}$ pätee*

$$(TT^t)x \cdot x = T(T^t x) \cdot x = (T^t x) \cdot (T^t x) \geq 0.$$

Choleskyn hajotelman hyödyllisyys liittyy yhtälöryhmien numeeriseen ratkaisemiseen. Tarkastellaan matriisiyhtälöä $Ax = y$, missä $A \in \mathbb{R}^{n\times n}$. Oletetaan, että tunnetaan matriisin A Cholesky hajotelma, eli että $A = TT^t$, missä $T \in \mathbb{R}^{n\times n}$ on alakolmiomatriisi. Tällöin

$$T(T^t x) = Ax = y$$

eli voidaan aloittaa ratkaisemalla yhtälö

$$Tz = y$$

Tämä yhtälöryhmä voidaan ratkaista tehokkaalla eliminointimenetelmällä kuten seuraava esimerkki osoittaa. Tämän jälkeen riittää ratkaista yhtälö $T^t x = z$ samalla menetelmällä.

Esimerkki 16.5.3. *Olkoot*

$$T = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \quad ja \quad y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Tällöin yhtälöstä

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

seuraa välittömästi, että $z_1 = 1/2$, $z_2 = 2 - z_1 = 3/2$ ja $z_3 = \frac{1}{2}(3 - z_2) = 3/4$.

Seuraava lause on luvun päätulos. Palautetaan todista varten mieleen, että matriisin $A \in \mathbb{R}^{m\times n}$ aste $\operatorname{rank}(A)$ on sen sarakeavaruuden dimensio eli $\operatorname{rank}(A) = \dim \operatorname{Col}(A)$.

Lause 16.5.4. *Positiivisesti semidefiniitillä matriisilla $A \in \mathbb{R}^{n\times n}$ on Choleskyn hajotelma.*

Todistus. Lauseen 16.4.5 nojalla matriisilla A on olemassa yksikäsitteinen positiviisesti semidefiniitti neliöjuuri B. Olkoon $k = \operatorname{rank}(B)$.

Olkoon nyt ortogonaalimatriisi $Q \in \mathbb{R}^{n\times k}$ ja yläkolmiomatriisi $R \in \mathbb{R}^{k\times n}$ matriisin B QR-hajotelma, eli $B = QR$. Koska B on symmetrinen matriisin, niin

$$R^t R = R^t Q^t Q R = (QR)^t QR = B^t B = B^2 = A.$$

Valitaan nyt $T = R^t$. □

Huomautus 16.5.5. *Huomaa, että matriisin A positiivisesti semidefiniitille neliöjuurelle B pätee* $\operatorname{rank}(A) = \operatorname{rank}(B)$. *Huomaa myös, että mikäli $k = \operatorname{rank}(B) < n$, niin edellisen lauseen todistus antaa alakolmiomatriisin T, joka ei ole neliömatriisi vaan $n \times k$-matriisi. Huomaa kuitenkin, että lisäämällä matriisiin Q nollasarakkeita ja matriisiin R nollarivejä, molemmat matriisit voidaan täydentää neliömatriiseiksi Q' ja R', joille pätee, että $B = Q'R'$ ja $(R')^t R' = A$. Tällöin $T = (R')^t$ on Cholesky hajotelman mukainen alakolmiomatriisi. (Harjoitustehtävä)*

Positiivisesti definiittien matriisien Choleskyn hajotelma on yksikäsitteinen ja alakolmiomatriisi on neliömatriisi. Todistetaan tämä seuraavaksi.

Lause 16.5.6. *Positiivisesti definiitin matriisin $A \in \mathbb{R}^{n\times n}$ Choleskyn hajotelma $A = TT^t$ on yksikäsitteinen.*

Todistus. Oletetaan, että T ja L ovat sellaisia alakolmiomatriiseja, että $A = TT^t$ ja $A = LL^t$. Osoitetaan, että $T = L$.

Koska A on positiviisesti definiitti, niin se on kääntyvä korollaarin 16.3.5 perusteella. Näin ollen myös matriisit T ja L ovat kääntyviä. Koska L ja T ovat alakolmiomatriiseja, niin myös matriisit L^{-1} ja $L^{-1}T$ ovat alakolmiomatriiseja.

Koska $TT^t = A = LL^t$, niin $L^{-1}T = L^t(T^t)^{-1}$. Näin ollen transpoosin ominaisuuksien perusteella

$$L^{-1}T = L^t(T^t)^{-1} = L^t(T^{-1})^t = (T^{-1}L)^t.$$

Koska $T^{-1}L$ on alakolmiomatriisi, niin $(T^{-1}L)^t$ on yläkolmiomatriisi. Matriisi $L^{-1}T$ on siis sekä alakolmio- että yläkolmiomatriisi. Näin ollen $L^{-1}T$ on diagonaalimatriisi. Merkitään $D = L^{-1}T$, jolloin $T = LD$. Nyt

$$LL^t = A = TT^t = (LD)(LD)^t = LDD^tL^t.$$

Koska matriisi L on kääntyvä, niin $DD^t = I$. Koska D on diagonaalimatriisi, niin $D = I$. Näin ollen matriisin D jokainen diagonaalialkio on joko 1 tai -1. Koska matriisien T ja L diagonaalialkiot ovat positiviisia, niin $D = I$ ja $T = L$. $\square$

Positiivisesti semidefiniitin matriisin Choleskyn hajotelman ei tarvitse olla yksikäsitteinen, kuten seuraava (helppo) esimerkki osoittaa.

Esimerkki 16.5.7. *Osoitetaan, että matriisilla*

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

on useampi Choleskyn hajotelma ratkaisemalla matriisin T kertoimet suoraan suoraan yhtälöstä $A = TT^t$.

Olkoon

$$T = \begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}$$

Tällöin

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix} \begin{bmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{bmatrix} = \begin{bmatrix} a^2 & ab & ad \\ ab & b^2 + c^2 & bd + ce \\ ad & bd + ce & d^2 + e^2 + f^2 \end{bmatrix}.$$

Näin ollen matriisi T on yhtälön $A = TT^t$ ratkaisu, jos ja vain jos $a = 1$, $b = 0$, $c = 0$, $d = 0$ ja $e^2 + f^2 = 1$. Näin ollen yhtälöllä $A = TT^t$ on useampia ratkaisuja.

Huomautus 16.5.8. *Lukijalle voi herätä onko suoraviivaisempaa tapaa löytää annetun matriisin Choleskyn hajotelma kuin käyttää neliöjuuren QR-hajotelmaa. Kaksi tälläistä tapaa on annettu liitteessä F.4.*

Luku 17

Lineaarikuvauksen polaari- ja singulaariarvohajotelmat

17.1 Polaarihajotelma

Polaarihajotelma on positiivisten operaattorien teorian sovellus yleisten lineaarikuvausten teoriaan. Polaarihajotelmalla tarkoitetaan sisätuloavaruuksien välisen lineaarikuvauksen $f\colon V \to W$ esittämistä positiivisen operaattorin $g\colon V \to V$ ja isometrian $s\colon \operatorname{im} g \to W$ yhdisteenä eli $f = s \circ g$. Tuloksen voi siis tulkita sanomalla, että jokainen operaattori on isometriaa vailla positiivinen. Matriisien kielellä tämä tarkoittaa erityisesti sitä, että jokaisella $A \in \mathbb{R}^{m\times n}$, missä $m \geq n$, on olemassa sellainen positiivisesti semidefiniitti matriisi $B \in \mathbb{R}^{n\times n}$ ja sellainen ortogonaalimatriisi $O \in \mathbb{R}^{m\times n}$, että

$$A = OB.$$

Polaarihajotelman positiivinen operaattori g on itseasiassa operaattorin $f^*f\colon V \to V$ yksikäsitteinen positiivinen neliöjuuri. Otetaan tätä varten käyttöön merkintä operaattorin f^*f positiiviselle neliöjuurelle.

Määritelmä 17.1.1. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ äärellisulotteisia sisätuloavaruuksia ja $f\colon V \to W$ lineaarikuvaus. Merkitään $\sqrt{f^*f}\colon V \to V$ sitä yksikäsitteistä operaattoria, joka on operaattorin $f^* \circ f\colon V \to V$ positiivinen neliöjuuri, eli positiivista operaattoria, jolle pätee $\sqrt{f^*f} \circ \sqrt{f^*f} = f^* \circ f$.*

Lineaarikuvausten polaarihajotelma voidaan muotoilla seuraavasti.

Lause 17.1.2 (Polaarihajotelma I)**.** *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ äärellisulotteisia sisätuloavaruuksia ja olkoon $f\colon V \to W$ lineaarikuvaus. Tällöin on olemassa sellainen isometria $s\colon \operatorname{im}\sqrt{f^*f} \to W$, että*

$$f = (s \circ \pi) \circ \sqrt{f^*f},$$

missä $\pi\colon V \to V$ on ortogonaaliprojektio operaattorin $\sqrt{f^*f}$ kuvalle $\operatorname{im}\sqrt{f^*f}$.

Sovelluksissa on kuitenkin helpompaa jakaa polaarilause tapauksiin $\dim V \leq \dim W$ ja $\dim V > \dim W$. Tapauksessa $\dim V \leq \dim W$ ei tarvita projektiota $\pi\colon V \to V$, vaan tulos voidaan muotoilla seuraavasti.

Lause 17.1.3 (Polaarihajotelma II)**.** *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *sellaisia äärellisulotteisia sisätuloavaruuksia, että* $\dim V \leq \dim W$*, ja olkoon* $f\colon V \to W$ *lineaarikuvaus. Tällöin on olemassa sellainen isometria* $S\colon V \to W$*, että*

$$f = S \circ \sqrt{f^*f}.$$

Tätä polaarihajotelman toista versiota puoltaa seuraava polaarilauseen muotoilu matriiseille.

Lause 17.1.4 (Matriisien polaarihajotelma)**.** *Olkoon* $A \in \mathbb{R}^{m\times n}$ *matriisi, missä* $m \geq n$*. Tällöin on olemassa sellainen ortogonaali matriisi* $O \in \mathbb{R}^{m\times n}$*, että*

$$A = O(A^tA)^{1/2}.$$

Todistus. Sovelletaan lausetta 17.1.3 kuvaukseen $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$, $x \mapsto Ax$.

Tällöin $f_A = S \circ \sqrt{f_A^* f_A}$, missä $S\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ on isometria. Koska S on lineaarikuvaus sarakeavaruudelta sarakeavaruudelle, niin on olemassa sellainen matriisi $O \in \mathbb{R}^{m\times n}$, että $S = f_O$. Koska S on isometria, eli kuvaus f_O on isometria, niin matriisi O on ortogonaalimatriisi. Koska

$$f_A = S \circ \sqrt{f_A^* f_A} = f_O \circ f_{\sqrt{A^tA}} = f_{O\sqrt{A^tA}},$$

niin

$$A = O\sqrt{A^tA}.$$

□

Huomautus 17.1.5. *Lauseelle 17.1.4 voidaan myös muotoilla versio tapauksessa* $m < n$ *käyttäen lausetta 17.1.2. Tällöin matriisi* O *ei kuitenkaan ole ortgonaalimatriisi vaan yhdistetyn kuvauksen* $s \circ \pi\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ *matriisi.*

17.1.1 Polaarihajotelman löytäminen

Vaikka lauseen 17.1.4 väite ei kerro kerro kuinka matriisi O löydetään, sen voi kuitenkin ratkaista matriiseista A ja $(A^tA)^{1/2}$ kuten seuraava esimerkki osoittaa. Huomaa, että matriisin $(A^tA)^{1/2}$ selvittäminen palautuu spektraalilauseeseen.

Esimerkki 17.1.6. *Olkoon*

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

Etsitään matriisin A *polaarihajotelma* $A = OB$*. Havaitaan ensin, että*

$$A^tA = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Esimerkin 16.4.6 perusteella tämän matriisin positiivisesti definiittinen neliöjuuri on

$$B = \frac{1}{2}\begin{bmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & \sqrt{3}+1 \end{bmatrix}.$$

Ortogonaalimatriisi $O = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \in \mathbb{R}^{3\times 2}$ *ratkaistaan nyt yhtälöstä*

$$A = OB.$$

Tässä tapauksessa matriisi B *on kääntyvä, joten*

$$O = AB^{-1}.$$

Esimerkki 17.1.7. *Olkoon*

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

Tällöin

$$A^tA = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

Tällöin

$$A^tA = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right).$$

Näin ollen

$$(A^tA)^{1/2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Nyt yhtälö

$$A = OB$$

toteutuu kaikilla ortogonaalimatriiseilla $O \in \mathbb{R}^{3\times 2}$*, joilla pätee*

$$O(e_1 + e_2) = e_1 + e_2.$$

Esimerkiksi voidaan valita

$$O = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Huomaa, että yhtälö $A = OB$ *on yhtäpitävä yhtälön* $A^t = B^tO^t$ *kanssa. Näin ollen matriisin* O^t *sarakkeet* $\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$ *määräytyvät yhtälöstä*

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$$

eli yhtälöistä

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} v_1, \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} v_2 \quad ja \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} v_3.$$

17.1.2 Polaarihajotelman todistus

Polaarihajotelman todistuksen ytimessä on kaksi havaintoa. Ensimmäinen havainto on, että jokainen vektori $v \in V$ kuvautuu kuvauksissa f ja $\sqrt{f^*f}$ saman pituisiksi, eli $\|f(v)\| = \|\sqrt{f^*f}(v)\|$. Tämä on välttämätön ehto isometrian S olemassaololle. Toinen havainto on, että on olemassa sellainen avaruuden S isometria $S\colon V \to V$, jolle pätee $f = S \circ \sqrt{f^*f}$. Kirjataan nämä havainnot erillisiksi lemmoiksi.

Lemma 17.1.8. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ sisätuloavaruuksia ja $f\colon V \to W$ lineaarikuvaus. Tällöin*

$$\|f(v)\|_W = \|\sqrt{f^*f}(v)\|_V$$

kaikilla $v \in V$.

Todistus. Olkoon $v \in V$. Koska $f^* \circ f = \sqrt{f^*f} \circ \sqrt{f^*f}$, niin

$$\begin{aligned} \|f(v)\|_W^2 &= \langle f(v), f(v)\rangle_W = \langle v, f^*(f(v))\rangle_V = \langle v, \sqrt{f^*f}\sqrt{f^*f}(v)\rangle_V \\ &= \langle \sqrt{f^*f}(v), \sqrt{f^*f}(v)\rangle_V = \|\sqrt{f^*f}(v)\|_V^2. \end{aligned}$$

□

Kirjataan nyt yleinen tulos isometrioista, jonka erikoistapausta sovelletaan polaarihajotelman todistuksessa.

Lemma 17.1.9. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$, $(W, \langle\cdot,\cdot\rangle_W)$ ja $(U, \langle\cdot,\cdot\rangle_U)$ äärellisulotteisia sisätuloavaruuksia. Olkoot lisäksi $f\colon V \to W$ ja $g\colon V \to U$ sellaiset lineaarikuvaukset, että $\|f(v)\|_W = \|g(v)\|_U$ jokaisella $v \in V$. Tällöin on olemassa sellainen isometria $s\colon \operatorname{im} g \to \operatorname{im} f$, että $f = s \circ g$.*

Todistus. Olkoon $(v_1, \ldots, v_n)$ sellainen avaruuden V kanta, että $(g(v_1), \ldots, g(v_k))$ on aliavaruuden $\operatorname{im} g \subset U$ kanta jollain $1 \le k \le n$. Lineaarikuvausten kantalauseen nojalla, on olemassa yksikäsitteinen sellainen lineaarikuvaus $s\colon \operatorname{im} g \to \operatorname{im} f$, että $s(g(v_i)) = f(v_i)$ jokaisella $i \in \{1, \ldots, k\}$. Osoitetaan, että s on isometria.

Olkoot $u = a_1 g(v_1) + \cdots + a_k g(v_k) \in \operatorname{im} g$ ja $v = a_1 v_1 + \cdots + a_k v_k$. Tällöin $u = g(v)$ ja

$$s(u) = a_1 s(g(v_1)) + \cdots + a_k s(g(v_k)) = a_1 f(v_1) + \cdots + a_k f(v_k) = f(v).$$

Näin ollen oletuksen nojalla

$$\|s(u)\|_W = \|f(v)\|_W = \|g(v)\|_U = \|u\|_U.$$

Kuvaus s on siis isometria.

Osoitetaan vielä, että $f = s \circ g$. Havaitaan ensin, että $\ker g \subset \ker f$, sillä jokaisella $v \in \ker g$ pätee $\|f(v)\|_W = \|g(v)\|_U = 0$. Olkoon nyt $v \in V$ ja $u = f(v) = a_1 u_1 + \cdots + a_k u_k \in U$. Tällöin $g(a_1 v_1 + \cdots + a_k v_k) = u = g(v)$, joten $v - (a_1 v_1 + \cdots + a_k v_k) \in \ker g \subset \ker f$. Näin ollen

$$\begin{aligned} f(v) &= f(a_1 v_1 + \cdots + a_k v_k) = a_1 f(v_1) + \cdots + a_k f(v_k) \\ &= a_1 h(u_1) + \cdots + a_k h(u_k) = h(u) = h(f(v)). \end{aligned}$$

Tämä päättää todistuksen. □

Lauseen 17.1.2 eli polaarihajotelman I todistus. Olkoon $f\colon V \to W$ lineaarikuvaus. Tällöin lemman 17.1.8 perusteella $\|f(v)\|_W = \|\sqrt{f^*f}(v)\|_V$ kaikilla $v \in V$. Olkoon nyt $\pi\colon V \to V$ ortogonaaliprojektio aliavaruudellle $\operatorname{im}\sqrt{f^*f}$. Koska $(\pi \circ \sqrt{f^*f})(v) = \sqrt{f^*f}(v)$ jokaisella $v \in V$, niin lemman 17.1.9 perusteella on olemassa sellainen isometria $s\colon \operatorname{im}\sqrt{f^*f} \to W$, että $f = s \circ (\pi \circ \sqrt{f^*f})$. Tämä päättää todistuksen. □

Lauseen 17.1.3 todistus perustuu seuraavaan lemmaan, joka jätetään harjoitustehtäväksi.

Lemma 17.1.10. *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *sellaisia äärellisulotteisia sisätuloavaruuksia, että* $\dim V \leq \dim W$, *ja olkoon* $U \subset V$ *aliavaruus. Tällöin jokainen isometria* $s\colon U \to W$ *voidaan jatkaa isometriaksi* $S\colon V \to W$ *eli* S *on sellainen isometria, että* $S|_U = s$.

Lauseen 17.1.3 eli polaarihajotelman II todistus. Olkoon $f\colon V \to W$ lineaarikuvaus. Tällöin lemman 17.1.8 perusteella $\|f(v)\|_W = \|\sqrt{f^*f}(v)\|_V$ kaikilla $v \in V$. Näin ollen lemman 17.1.9 perusteella on olemassa sellainen isometria $s\colon \operatorname{im}\sqrt{f^*f} \to W$, että $f = s \circ \sqrt{f^*f}$. Koska $\dim V \leq \dim W$, niin lemma 17.1.10 perusteella on olemassa sellainen isometria $S\colon V \to W$, että $S|_{\operatorname{im}\sqrt{f^*f}} = s$. Näin ollen $f = S \circ \sqrt{f^*f}$. Tämä päättää todistuksen. □

17.2 Singulaariarvohajotelma

Tässä luvussa todistetaan lineaarioperaattoreiden singulaariarvohajotelma:

> Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi, missä $m \geq n$, ja $s_1, \ldots, s_n$ sen singulaariarvot. Tällöin on olemassa sellaiset ortogonaaliset matriisit $P \in \mathbb{R}^{m\times n}$ ja $Q \in \mathbb{R}^{n\times n}$ sekä diagonaalimatriisi $\Sigma \in \mathbb{R}^{n\times n}$, että
>
> $$A = P\Sigma Q^t.$$

Singulaariarvohajotelmaa on luonnollista verrata symmetrisien matriisien spektraalilauseeseen (lause 15.4.1): *Olkoon* $A \in \mathbb{R}^{n\times n}$ *symmetrinen matriisi. Tällöin on olemassa sellainen ortogonaalimatriisi* $Q \in \mathbb{R}^{n\times n}$ *ja diagonaalimatriisi* $D \in \mathbb{R}^{n\times n}$, *että* $A = QDQ^t$.

Vaikka näissä väitteissä on monia samankaltaisuuksia, niin niiden välillä on valtava ero. Spektraalilause sanoo, että symmetrisen matriisin $A \in \mathbb{R}^{n\times n}$ esitysmatriisi ominaisvektoreiden määräämässä kannassa on diagonaalimatriisi. Kuvauksille tämä tarkoittaa, että itseadjungoidun operaattorin esitysmatriisi on diagonaalimatriisi. Kuten hyvin tiedetään, diagonaalimatriisin D diagonaalielementit ovat matriisin A ominaisarvoja.

Singulaariarvohajotelman tulkinta on toinen. Koska matriisi A ei ole neliömatriisi, ei voi puhua sen ominaisarvoista tai ominaisvektoreista. On kuitenkin huomattavaa, että voidaan puhua diagonaalimatriisin Σ ominaisarvoista ja ominaisvektoreista. Toisaalta huomataan, että matriisit P ja Q ovat ortogonaalisia eli kuvauksina tulkittuina ne määrittelevät isometriat $f_P\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ ja $f_Q\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$. Singulaariarvohajotelman tulkinta kuvauksina on siis, että matriisia A vastaava lineaarikuvaus $f_A\colon \mathbb{R}^{n\times 1} \to$

$\mathbb{R}^{m\times 1}$ voidaan kirjoittaa yhdisteenä kolmesta kuvauksesta $f_A = f_P \circ f_\Sigma \circ f_Q^{-1}$, missä kuvaukset f_P ja f_Q^{-1} ovat isometrioita eli eivät muuta pituuksia ja kuvauksen f_A venytykset on koodattu kuvaukseen f_Σ.

Näin ollen herää kysymys, kuinka matriisit P, Q ja Σ löydetään. Paljastuu, että singulaariarvohajotelma seuraa suoraan polaarihajotelmasta ja että matriisi Σ on matriisin $(A^tA)^{1/2}$ ominaisarvojen matriisi. Aloitetaan nyt singulaariarvojen määritelmällä.

Määritelmä 17.2.1. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi. Matriisin $(A^tA)^{1/2}$ ominaisarvoa $\lambda \in \mathbb{R}$ kutsutaan matriisin A* singulaariarvoiksi*. Singulaariarvon $\lambda \in \mathbb{R}$* (geometrinen) kertaluku *$\mu_\lambda \in \mathbb{N}$ on $\mu_\lambda = \dim E(\lambda, (A^tA)^{1/2})$.*

Huomautus 17.2.2. *Syy singulaariarvon geometrisen kertaluvun määritelmälle on seuraava huomio. Olkoon $A \in \mathbb{R}^{m\times n}$ ja olkoot $\lambda_1 < \ldots < \lambda_k$ matriisin $(A^tA)^{1/2} \in \mathbb{R}^{n\times n}$ kaikki ominaisarvot. Tällöin spektraalilauseen nojalla*

$$\mu_{\lambda_1} + \cdots + \mu_{\lambda_k} = \dim E(\lambda_1, (A^tA)^{1/2}) + \cdots + \dim E(\lambda_k, (A^tA)^{1/2}) = n$$

eli $(m \times n)$-matriisin A singulaariarvojen geometristen kertalukujen summa on n.

Lause 17.2.3 (Singulaariarvohajotelma (engl. *singular value decomposition*)**.** *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi, missä $m \geq n$. Tällöin on olemassa sellaiset ortogonaalimatriisit $P \in \mathbb{R}^{m\times n}$ ja $Q \in \mathbb{R}^{n\times n}$ sekä matriisin A singulaariarvoista koostuva diagonaalimatriisi $\Sigma \in \mathbb{R}^{n\times n}$, että*

$$A = P\Sigma Q^t.$$

Todistus. Matriisien polaarihajotelman (Lause 17.1.4) nojalla on olemassa sellainen ortogonaalimatriisi $O \in \mathbb{R}^{m\times n}$, että $A = O(A^tA)^{1/2}$. Koska matriisi A^tA on symmetrinen, niin spektraalilauseen (Korollaari 15.4.2) nojalla on olemassa matriisin A^tA ominaisarvoista koostuva diagonaalimatriisi $D \in \mathbb{R}^{n\times n}$ ja ominaisvektoreista koostuva ortogonaalimatriisi $Q \in \mathbb{R}^{n\times n}$, että $A^tA = QDQ^t$. Lemman 16.4.4 nojalla $(A^tA)^{1/2} = QD^{1/2}Q^t$. Näin ollen

$$A = O(A^tA)^{1/2} = OQD^{1/2}Q^t.$$

Tällöin $P = OQ \in \mathbb{R}^{m\times n}$ ja $\Sigma = D^{1/2}$ ovat halutut matriisit. □

Huomautus 17.2.4. *Singulaariarvohajotelman väitteessä tai todistuksessa ei korostettu singulaariarvojen geometrista kertalukua. Koska diagonaalimatriisi Σ on matriisin $(A^tA)^{1/2}$ spektraalihajotelman $(A^tA)^{1/2} = QD^{1/2}Q^t$ matriisi $D^{1/2}$, niin matriisin A singulaariarvot toistuvat matriisissa Σ niiden geometrisen kertaluvun mukaisesti.*

Huomautus 17.2.5 (Algoritmi singulaariarvohajotelman löytämiselle)**.** *Lauseen 17.2.3 todistusta voidaan suoraviivaistaa ja antaa seuraava algoritmi singulaariarvohajotelman löytämiselle. Olkoon $A \in \mathbb{R}^{m\times n}$.*

1. *Laske A^tA ja sen ominaisarvot $\lambda_1 > \cdots > \lambda_k$.*

2. *Etsi matriisin A^tA spektralihajotelma $A^tA = QDQ^t$, missä $Q \in \mathbb{R}^{n\times n}$ on matriisin A^tA ominaisvektoreista muodostuva ortogonaalimatriisi ja $D \in \mathbb{R}^{n\times n}$ ominaisarvoista koostuva diagonaalimatriisi.*

3. *Määrittele* $\Sigma = D^{1/2}$.

4. *Ratkaise matriisi* $P \in \mathbb{R}^{m\times n}$ *yhtälöstä* $A = P\Sigma Q^t$ *esimerkiksi ratkaisemalla sen transpoosi* P^t *yhtälöstä* $\left[Q\Sigma \mid A^t\right]$.

Huomaa, että matriisin P *voi myös ratkaista polaarihajotelmaa käyttäen seuraavasti:*

1. *Muodosta matriisin* A^tA *spektralihajotelma* $A^tA = QDQ^t$.

2. *Laske matriisi* $(A^tA)^{1/2} = QD^{1/2}Q^t$.

3. *Etsi matriisin* A *polaarihajotelman* $A = O(A^tA)^{1/2}$ *ortogonaalimatriisi* $O \in \mathbb{R}^{m\times n}$ *ratkaisemalla matriisi* O^t *yhtälöstä* $\left[(A^tA)^{1/2} \mid A^t\right]$.

4. *Nyt* $P = OQ$ *ja* $\Sigma = D^{1/2}$.

Huomautus 17.2.6. *Singulaariarvohajotelman todistus ja edellä annettu algoritmi korostavat, että hajotelman matriisi* Q *koostuu matriisin* A^tA *ominaisvektoreista. Yllä oleva algoritmi ei anna tälläistä tulkintaa matriisille* P. *Suora lasku kuitenkin paljastaa, että matriisin* P *sarakkeet ovat matriisin* $AA^t \in \mathbb{R}^{m\times m}$ *ominaisvektoreita. Tämä havaitaan seuraavasti. Koska matriisi* $Q \in \mathbb{R}^{n\times n}$ *on ortogonaalinen, niin singulaariarvohajotelman* $A = P\Sigma Q^t$ *perusteella saadaan*

$$AA^t = (P\Sigma Q^t)(P\Sigma Q^t)^t = P\Sigma Q^t(Q^t)^t\Sigma^t P^t = P\Sigma Q^t Q\Sigma^t P^t = P\Sigma^2 P^t.$$

Näin ollen matriisit P *ja* $D = \Sigma^2$ *muodostavat matriisin* AA^t *spektraalihajotelman. Matriisin* P *sarakkeet ovat siis matriisin* AA^t *ominaisvektoreita.*

Huomautus 17.2.7. *Edellisestä huomautuksesta voi tulla mieleen, että matriisin* P *voisi laskea suoraan käyttämällä ainoastaan spektraalihajotelmaa matriisille* AA^t. *Tämä ei kuitenkaan ole totta, sillä spektraalihajotelman* $AA^t = PDP^t$ *matriisit* P *ja* D *eivät ole yksikäsitteisiä.*

Huomautus 17.2.8. *Huomautetaankoon kuitenkin vielä lopuksi, että jos* $A = P\Sigma Q^t$ *on matriisin* A *spektraalihajotelma, eli että* P *ja* Q *ovat ortogonaalimatriiseja ja* Σ *on diagonaalimatriisi, niin tällöin*

$$A^tA = (P\Sigma Q^t)^t(P\Sigma Q^t) = Q\Sigma P^t P\Sigma Q^t = Q\Sigma^2 Q^t,$$

eli matriisin Q *sarakkeet ovat tällöin matriisin* A^tA *ominaisvektoreita ja matriisin* Σ^2 *diagonaalialkiot matriisin* A^tA *ominaisarvoja.*

17.2.1 Lineaarikuvauksen singulaariarvohajotelma

Lineaarikuvausten singulaariarvohajotelmaa voi lähestyä kahdesta näkökulmasta. tärkein huomio kuitenkin on, että lineaarikuvauksen singulaariarvot voidaan määritellä ilman matriiseja.

Määritelmä 17.2.9. *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *äärellisulotteisia sisätuloavaruuksia. Luku* $\lambda \in \mathbb{R}$ *on lineaarikuvauksen* $f\colon V \to W$ singulaariarvo*, jos se on operaattorin* $\sqrt{f^*f}\colon V \to V$ *ominaisarvo.*

Kuten matriisien tapauksessa saadaan seuraava yleinen tulos lineaarikuvauksille. Huomaa, että tässä muotoilussa kuvauksen $f\colon V \to W$ lähtö- ja maaliavaruudet ovat äärellisulotteisia, mutta lauseessa ei ole lisärajoitusta $\dim V \leq \dim W$.

Lause 17.2.10 (Sisätuloavaruuksien välisten lineaarikuvausten singulaariarvohajotelma)**.** *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *äärellisulotteisia sisätuloavaruuksia sekä* $f\colon V \to W$ *lineaarikuvaus. Tällöin on olemassa sellainen lineaarikuvaus* $h\colon \mathbb{R}^{n\times 1} \to W$ *ja sellainen isometria* $g\colon \mathbb{R}^{n\times 1} \to V$*, että*

$$f = h \circ f_\Sigma \circ g^*$$

missä $\Sigma \in \mathbb{R}^{n\times n}$ *on diagonaalimatriisi, jonka diagonaalialkiot ovat kuvauksen* f *singulaariarvoja, ja että* $h|_{\mathrm{Col}(\Sigma)}\colon \mathrm{Col}(\Sigma) \to W$ *on isometria.*

Todistus. Lineaarikuvausten polaarihajotelman (lause 17.1.2) nojalla on olemassa sellainen isometria $s\colon \operatorname{im}\sqrt{f^*f} \to W$, että $f = (s \circ \pi) \circ \sqrt{f^*f}$, missä $\pi\colon V \to V$ on ortgonaaliprojektio aliavaruudelle $\operatorname{im}\sqrt{f^*f}$. Koska operaattori f^*f on itseadjungoitu, niin spektraalilauseen nojalla on olemassa sellainen avaruuden V ortonormaalikanta $(v_1, \dots, v_n)$ että operaattorin f^*f esitysmatriisi tässä kannassa on diagonaalimatriisi $D \in \mathbb{R}^{n\times n}$. Tällöin $D^{1/2}$ on operaattorin $\sqrt{f^*f}$ esitysmatriisi tässä kannassa. Merkitään $\Sigma = D^{1/2}$.

Olkoon nyt $g = \Phi_{(v_1,\dots,v_n)}\colon \mathbb{R}^{n\times n} \to V$ kantaan $(v_1, \dots, v_n)$ liittyvä isometrinen isomorfismi. Tällöin $g^* = g^{-1}$, joten

$$\sqrt{f^*f} = g \circ f_\Sigma \circ g^{-1} = g \circ f_\Sigma \circ g^*.$$

Olkoot nyt $h = s \circ \pi \circ g\colon \mathbb{R}^{n\times 1} \to W$. Tällöin

$$f = (s \circ \pi) \circ \sqrt{f^*f} = s \circ g \circ f_\Sigma \circ g^* = h \circ f_\Sigma \circ g^*.$$

Osoitetaan vielä, että $h|_{\mathrm{Col}(\Sigma)}\colon \mathrm{Col}(\Sigma) \to W$ on isometria. Olkoot $x, x' \in \mathrm{Col}(\Sigma)$. Koska vektorit $g(x)$ ja $g(x')$ kuuluvat avaruuteen $\operatorname{im}\sqrt{f^*f}$, niin $\pi(g(x)) = g(x)$ ja $\pi(g(x')) = g(x')$. Näin ollen

$$\begin{aligned}
\langle h(x), h(x')\rangle_W &= \langle (s(\pi(g(x))), s(\pi(g(x')))\rangle_W \\
&= \langle \pi(g(x)), \pi(g(x'))\rangle_V \\
&= \langle g(x), g(x')\rangle_V = x \cdot x'.
\end{aligned}$$

Kuvaus $h|_{\mathrm{Col}(\Sigma)}\colon \mathrm{Col}(\Sigma) \to W$ on siis isometria. □

Korollaari 17.2.11. *Olkoot* $(V, \langle\cdot,\cdot\rangle_V)$ *ja* $(W, \langle\cdot,\cdot\rangle_W)$ *äärellisulotteisia sisätuloavaruuksia sekä* $f\colon V \to W$ *lineaarikuvaus. Tällöin on olemassa sellainen avaruuden* V *ortonormaalikanta* $(v_1, \dots, v_n)$ *ja avaruuden* W *ortonormaalikanta* $(w_1, \dots, w_m)$*, että* $f(v_i) = \sigma_i w_i$ *jokaisella* $i \in \{1, \dots, n\}$*, missä* σ_i *on kuvauksen* f *singulaariarvo.*

Todistus. Olkoon $g\colon \mathbb{R}^{n\times 1} \to V$ isometria ja $h\colon \mathbb{R}^{n\times 1} \to W$ lineaarikuvaus kuten lauseessa 17.2.10 eli

$$f = h \circ f_\Sigma \circ g^*,$$

missä $\Sigma \in \mathbb{R}^{n\times n}$ on diagonaalimatriisi $\Sigma = \begin{bmatrix}\sigma_1 e_1 & \cdots & \sigma_n e_n\end{bmatrix}$, jonka diagonaalialkiot ovat kuvauksen f singulaariarvoja, ja $h|_{\mathrm{Col}(\Sigma)}\colon \mathrm{Col}(\Sigma) \to W$ on isometria. Koska g on isomorfismi, niin $(v_1, \dots, v_n)$, missä $v_i = g(e_i)$ jokaisella $i \in \{1, \dots, n\}$, on avaruuden V ortonormaali kanta. Lisäksi $(h(e_1), \dots, h(e_n))$ on avaruuden W ortonormaali jono, joka voidaan laajentaa ortonormaaliksi kannaksi $(w_1, \dots, w_m)$, missä $w_i = h(e_i)$ jokaisella $i \in \{1, \dots, n\}$. Tällöin $f(v_i) = f(g(e_i)) = h(f_\Sigma(e_i)) = \sigma_i h(e_i) = \sigma_i w_i$ jokaisella $i \in \{1, \dots, n\}$. $\square$

Edellisen korollaarin voi myös kirjata seuraavasti. Korollaarin todistus jätetään harjoitustehtäväksi.

Korollaari 17.2.12. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ n-ulotteinen sisätuloavaruus ja $(W, \langle\cdot,\cdot\rangle_W)$ m-ulotteinen sisätuloavaruus sekä $f\colon V \to W$ lineaarikuvaus. Tällöin on olemassa sellainen avaruuden V ortonormaali kanta $(v_1, \dots, v_n)$ ja avaruuden W ortonormaali kanta $(w_1, \dots, w_m)$, että*

$$f(v) = \sigma_1\langle v_1, v\rangle w_1 + \cdots + \sigma_n\langle v_n, v\rangle w_n \tag{17.1}$$

kaikilla $v \in V$, missä $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n$ ovat lineaarikuvauksen f singulaariarvoja.

Edellä ollaan sivuutettu lineaarikuvauksen singulaariarvojen laskeminen esitysmatriisista. Seuraava tulos jätetään harjoitustehtäväksi.

Lause 17.2.13. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ n-ulotteinen sisätuloavaruus ja $(W, \langle\cdot,\cdot\rangle_W)$ m-ulotteinen sisätuloavaruus sekä $f\colon V \to W$ lineaarikuvaus. Olkoot lisäksi $(v_1, \dots, v_n)$ avaruuden V ortonormaali kanta ja $(w_1, \dots, w_m)$ avaruuden W ortonormaali kanta. Tällöin kuvauksen f esitysmatriisilla A kannasta $(v_1, \dots, v_n)$ kantaan $(w_1, \dots, w_m)$ on samat singulaariarvot kuin kuvauksella f.*

17.2.2 Singulaariarvot ja venytys

Tarkastellaan nyt singulaariarvojen geometrista merkitystä. Koska matriisin $A \in \mathbb{R}^{m\times n}$ singulaariarvot $\sigma_1, \dots, \sigma_k$ ovat positiviisesti semidefiniitin matriisin $(A^tA)^{1/2}$ ominaisarvoja, niin ne ovat aina ei-negatiivisia ja niitä on korkeintaan n kappaletta eli $1 \leq k \leq n$.

Suurimmalla singulaariarvolla on seuraava merkitys.

Lause 17.2.14. *Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi ja olkoon $\sigma \in \mathbb{R}$ matriisin A suurin singulaariarvo. Tällöin kaikilla vektoreilla $v \in \mathbb{R}^{n\times 1}$ pätee*

$$\|Av\| \leq \sigma\|v\|$$

ja yhtälössä pätee yhtäsuuruus, jos ja vain jos $v \in E(\sigma, (A^tA)^{1/2})$.

Todistus. Hyödynnetään korollaaria 17.2.12. Olkoon $(v_1, \ldots, v_n)$ avaruuden $\mathbb{R}^{n\times 1}$ ortonormaalikanta ja olkoon $(w_1, \ldots, w_m)$ avaruuden $\mathbb{R}^{m\times 1}$ ortonormaalikanta kuten korollaarissa 17.2.12 eli $f_A(v_i) = \sigma_i w_i$ jokaisella $i = 1, \ldots, n$, missä $\sigma_i \in \mathbb{R}$ on kuvauksen f_A singuaaliarvo.

Olkoon nyt $v = \sum_{i=1}^n t_i v_i$ vektorin v esitys kannassa $(v_1, \ldots, v_n)$. Tällöin $Av = \sum_{i=1}^n t_i A v_i = \sum_{i=1}^n t_i \sigma_i w_i$. Näin ollen

$$\|Av\|^2 = \sum_{i=1}^n (t_i\sigma_i)^2 = \sum_{i=1}^n t_i^2\sigma_i^2 \leq \sigma^2 \sum_{i=1}^n t_i^2 = \sigma^2\|v\|^2 = (\sigma\|v\|)^2.$$

Ensimmäinen väite seuraa.

Oletetaan nyt, että $v \in \mathbb{R}^{n\times 1}$ on sellainen vektori, että $\|Av\| = \sigma\|v\|^2$. Tällöin edellisessä laskussa on yhtäsuuruus ja jokaisella indeksillä $i \in \{1, \ldots, n\}$ pätee joko $t_i = 0$ tai $\sigma_i = \sigma$. Näin ollen v on sellaisten kantavektoriden v_i lineaarikombinaatio, joilla $\sigma_i = \sigma$. Koska $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$ ovat ortonormaaleja kantoja, niin

$$(A^tA)v_i \cdot v_j = Av_i \cdot Av_j = (\sigma_i w_i)\cdot(\sigma_j w_j) = \sigma_i\sigma_j.$$

Näin ollen $(A^tA)v_i = \sigma_i^2 v_i$ eli v_i on matriisin A^tA ominaisvektori. Siten se on myös matriisin $(A^tA)^{1/2}$ ominaisvektori. Näin on päätelty, että yhtäsuuruudesta $\|Av\| = \sigma\|v\|$ seuraa $v \in E(\sigma, (A^tA)^{1/2})$.

Oletetaan nyt, että $v \in E(\sigma, (A^tA)^{1/2})$. Tällöin $v \in E(\sigma^2, A^2A)$ ja

$$\|Av\|^2 = (Av)\cdot(Av) = (A^tA)v \cdot v = \sigma^2 v \cdot v = \sigma^2\|v\|^2$$

eli $\|Av\| = \sigma\|v\|$.

Lineaarikuvauksille tämä tulos voidaan tulkita seuraavasti. Yksityiskohdat jätetään harjoitustehtäväksi.

Lause 17.2.15. *Olkoot $(V, \langle\cdot,\cdot\rangle_V)$ ja $(W, \langle\cdot,\cdot\rangle_W)$ äärellisulotteisia sisätuloavaruuksia sekä olkoot $f\colon V \to W$ lineaarikuvaus ja σ sen suurin singulaariarvo. Tällöin kaikilla vektoreilla $v \in V$ pätee*

$$\|f(v)\|_W \leq \sigma\|v\|_V$$

ja yhtälössä pätee yhtäsuuruus, jos ja vain jos $v \in E(\sigma, \sqrt{f^*f})$.

Muiden singulaariarvojen tulkinta

Lauseen 17.2.14 todistusta analysoimalla saadaan tulkinta myös muille singulaariarvoille. Tässä analyysissa on mielekkäintä hyödyntää lauseen 17.2.14 tulkitaan lineaarikuvauksille eli lausetta 17.2.15.

Olkoon $A \in \mathbb{R}^{m\times n}$ matriisi ja $f_A\colon \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ sitä vastaava lineaarikuvaus. Olkoon $\sigma_1 \geq 0$ matriisin A suurin singulaariarvo ja olkoot $\sigma_1 > \sigma_2 > \cdots > \sigma_k$ matriisin A singulaariarvot laskevassa järjestyksessä. Havaitaan aluksi, että lauseen 17.2.15 perusteella

$$\|f_A(x)\| \leq \sigma_1\|x\|$$

kaikilla $x \in \mathbb{R}^{n\times 1}$ ja että yhtälössä pätee yhtäsuuruus, jos ja vain jos $x \in E(\sigma_1, (A^tA)^{1/2})$.

Koska matriisin $(A^tA)^{1/2}$ ominaisavaruudet ovat kohtisuorassa toisiaan vastaan havaitaan nyt, että

$$\mathbb{R}^{n\times 1} = E(\sigma_1, (A^tA)^{1/2}) \oplus (E(\sigma_1, (A^tA)^{1/2}))^{\perp},$$

missä

$$(E(\sigma_1, (A^tA)^{1/2}))^{\perp} = E(\sigma_2, (A^tA)^{1/2}) \oplus \cdots \oplus E(\sigma_k, (A^tA)^{1/2}).$$

Rajoitetaan kuvaus f_A avaruuteen $(E(\sigma_1, (A^tA)^{1/2}))^{\perp}$ eli tarkastellaan kuvausta $f_A^1 = f_A|_{E(\sigma_1,(A^tA)^{1/2}))^{\perp}} \colon E(\sigma_1, (A^tA)^{1/2}))^{\perp} \to \mathbb{R}^{m\times 1}$. Tällöin kuvauksen f_A^1 suurin singulaariarvo on σ_2 ja

$$\|f_A^1(x)\| \leq \sigma_2\|x\|$$

kaikilla $x \in (E(\sigma_1, (A^tA)^{1/2}))^{\perp}$. Lisäksi $\|f_A^1(x)\| = \sigma_2\|x\|$, jos ja vain jos $x \in E(\sigma_2, (A^tA)^{1/2})$.

Näin jatkaen saadaan kaikille matriisin A singulaariarvoille tulkinta kuvauksen f_A sopivan rajoittuman suurinpana singulaariarvona.

17.2.3 Matriisin operaattorinormi

Lauseen 17.2.14 nojalla voidaan määritellä matriisin ja siten lineaarikuvauksen ns. operaattorinormi.

Määritelmä 17.2.16. *Matriisin $A \in \mathbb{R}^{m\times n}$ suurinta singulaariarvoa kutsutaan sen* operaattorinormiksi $\|A\|$.

Lause 17.2.14 voidaan nyt muotoilla seuraavasti. Kirjataan tämä uudelleentulkinta korollaariksi.

Korollaari 17.2.17. *Olkoon $A \in \mathbb{R}^{m\times n}$. Tällöin*

$$\|A\| = \max\{\|Ax\| \colon x \in \mathbb{R}^{n\times 1},\ \|x\| = 1\}.$$

Huomautus 17.2.18. *Monissa esityksissä korollaarin 17.2.17 tulos otetaan operaattorinormin määritelmäksi. Näin tehtäessä tulee ensin osoittaa, että maksimi todellakin on olemassa. Maksimin olemassaolo seuraa topologisesta havainnosta, että pistetulon määräämä etäisuus avaruuteen $\mathbb{R}^{n\times 1}$ tekee avaruudesta $\mathbb{R}^{n\times 1}$ lokaalisti kompaktin metrisen avaruuden. Tätä käsitellään tarkemmin kurssilla Topologia IB.*

Todettakoon vielä matriisin operaattorinormi on todellakin normi matriisien avaruudessa.

Lause 17.2.19. *Funktio $\|\cdot\| \colon \mathbb{R}^{m\times n} \to [0,\infty[$, $A \mapsto \|A\|$, missä $\|A\|$ on matriisin A suurin singulaariarvo, on normi matriisiavaruudessa $\mathbb{R}^{m\times n}$.*

Todistus. Olkoot $A, B \in \mathbb{R}^{m\times n}$. Tällöin avaruuden $\mathbb{R}^{m\times 1}$ pistetulon määräämän normin kolmioepäyhtälön perusteella jokaisella $x \in \mathbb{R}^{n\times 1}$ pätee

$$\|(A+B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\|\|x\| + \|B\|\|x\|.$$

Näin ollen

$$
\begin{aligned}
\|A+B\| &= \max\{\|(A+B)x\| \colon x \in \mathbb{R}^{n\times 1},\ \|x\| = 1\} \\
&= \max\{\|Ax+Bx\| \colon x \in \mathbb{R}^{n\times 1},\ \|x\| = 1\} \\
&\le \max\{\|A\|\|x\| + \|B\|\|x\| \colon x \in \mathbb{R}^{n\times 1},\ \|x\| = 1\} \\
&= \|A\| + \|B\|.
\end{aligned}
$$

Näin ollen funktio $\|\cdot\| \colon \mathbb{R}^{m\times n} \to [0,\infty[$ toteuttaa kolmioepäyhtälön.

Olkoot nyt $A \in \mathbb{R}^{m\times n}$ ja $c \in \mathbb{R}$. Tällöin

$$
\begin{aligned}
\|cA\| &= \max\{\|cAx\| \colon x \in \mathbb{R}^{n\times 1},\ \|x\| = 1\} \\
&= \max\{|c|\|Ax\| \colon x \in \mathbb{R}^{n\times 1},\ \|x\| = 1\} \\
&= |c| \max\{\|Ax\| \colon x \in \mathbb{R}^{n\times 1},\ \|x\| = 1\} \\
&= |c|\|A\|.
\end{aligned}
$$

Oletetaan nyt, että $A \in \mathbb{R}^{m\times n}$ on sellainen matriisi, että $\|A\| = 0$. Tällöin $Ax = 0$ kaikilla yksikkövektoreilla $x \in \mathbb{R}^{n\times 1}$. Näin ollen $Ax = 0$ kaikilla vektoreilla $x \in \mathbb{R}^{n\times 1}$ eli $A = 0$.

Tämä päättää todistuksen. $\square$

Luku 18

Jordanin normaalimuoto

Tämän, ja samalla viimeisen, luvun tavoite on todistaa kompleksisen operaattorin Jordanin normaalimuodon olemassaolo. Kuten hyvin tiedetään, kaikkia $n \times n$-neliömatriiseja ei voi diagonalisoida, koska avaruus $\mathbb{R}^{n\times 1}$ ei ole matriisin ominaisavaruuksien suora summa. Jordanin lause sanoo, että paras tulos joka voidaan saavuttaa on, että voidaan löytää sellainen avaruuden $\mathbb{R}^{n\times 1}$ kanta, jossa annettu matriisi voidaan kirjoittaa ositettuna diagonaalimatriisina.

Jordanin lauseen antama ositettu matriisi perustuu alkuperäisen matriisin yleistettyihin ominaisavaruuksiin ja matriisille saatava esitys on hyvin lähellä diagonalisoituvien matriisien esitystä. Annetaan tulosta varten tarvittava Jordanin lohkon määritelmä.

Merkintä 18.0.1. *Olkoon* $N_n = [a_{ji}] \in \mathbb{C}^{n\times n}$ *matriisi*

$$a_{ji} = \begin{cases} 1, & i = j-1 \\ 0, & \textit{muutoin.} \end{cases}$$

Määritelmä 18.0.2. *Yläkolmiomatriisi* $J_{n,\lambda} = [a_{ji}] \in \mathbb{C}^{n\times n}$ *on* Jordanin lohko parametrilla $\lambda \in \mathbb{C}^1$, *jos* $J_{n,\lambda} = \lambda I_n + N_n$, *missä* $I_n \in \mathbb{C}^{n\times n}$ *on identiteetti matriisi.*

Huomautus 18.0.3. *Yleisesti matriisi* $J_{n,\lambda}$ *on siis matriisi*

$$J_{n,\lambda} = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}.$$

Erityisesti tapauksissa $n = 2, 3, 4$ *saadaan matriisit*

$$J_{2,\lambda} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \ J_{3,\lambda} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \ \textit{ja } J_{4,\lambda} = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}.$$

[1]eli puhekielessä *Jordan blokki.*

Usein matriisin $J_{n,\lambda}$ kokoa ei tarvitse erikseen korostaa. Tällöin merkitään lyhyesti $J_\lambda = J_{n,\lambda}$.

Jordanin lause kompleksisille neliömatriiseille on seuraava tulos, jota kutsutaan matriisin Jordanin normaalimuodoksi.

Lause 18.0.4. *Jordanin normaalimuoto Olkoon $A \in \mathbb{C}^{n\times n}$. Tällöin on olemassa sellainen kääntyvä matriisi $P \in \mathbb{C}^{n\times n}$, että*

$$A = PJP^{-1},$$

missä J on Jordan lohkoista koostuva ositettu diagonaalimatriisi

$$J = \begin{bmatrix} J_{\mu_1} & & & \\ & J_{\mu_2} & & \\ & & \ddots & \\ & & & J_{\mu_\ell} \end{bmatrix}$$

missä $\mu_1, \dots, \mu_\ell$ ovat matriisin A ominaisarvoja.

Huomautus 18.0.5. *Lauseen muotoilusta voi tulla käsitys, että luvut $\mu_1, \dots, \mu_\ell$ ovat matriisin A erisuuria ominaisarvoja. Todistus kuitenkin osoittaa, että näin ei ole vaan, että yksittäinen ominaisarvo voi esiintyä jonossa $(\mu_1, \dots, \mu_\ell)$ useamman kerran.*

Huomautus 18.0.6. *Jos matriisin A Jordanin normaalimuodossa matriisin J Jordanin lohkot J_{μ_j} ovat 1×1-matriiseja, niin tällöin matriisi A on määritelmän mukaan diagonalisoituva. Toisaalta, jos A on diagonalisoituva, niin kuten kohta havaitaan matriisi J on matriisin A ominaisarvojen diagonaalimatriisi. Matriisin Jordanin normaalimuoto on siis tässä mielessä neliömatriisin diagonalisoituvuuden yleistys, kuten edellä jo todettiin.*

Jordanin lauseen muotoilu herättää kysymyksen, mihin matriisin J Jordanin lohkot oikeasti liittyvät. Vastaus on, että Jordanin lohkot liittyvät matriisin yleistettyihin ominaisavaruuksiiin kuten diagonaalimatriisit liittyvät matriisin ominaisavaruuksiin. Aloitetaan lauseen todistaminen siis yleistetyistä ominaisavaruuksista.

18.1 Yleistetyt ominaisavaruudet

Aloitetaan tarkastelemalla tuttuja ominaisarvoja uudesta kulmasta. Olkoon $A \in \mathbb{C}^{n\times n}$ neliömatriisi ja $\lambda \in \mathbb{C}$ sen ominaisarvo. Tällöin ominaisavaruus $E(A, \lambda)$ on itseasiassa matriisin $A - \lambda I$ nolla-avaruus eli

$$E(A, \lambda) = \{v \in \mathbb{C}^{n\times 1} \colon Av = \lambda v\} = \{v \in \mathbb{C}^{n\times 1} \colon (A - \lambda I)v = 0\} = \mathrm{Null}(A - \lambda I).$$

Matriisin A yleistetty ominaisavaruus puolestaan on sellaisten vektoreiden $v \in \mathbb{C}^{n\times 1}$ aliavaruus, että $(A - \lambda I)^k v = 0$ jollain $k \in \mathbb{N}$. Tarkemmin sanottuna määritelmä on seuraava.

Määritelmä 18.1.1. *Vektori* $v \in \mathbb{C}^{n\times 1}\setminus\{0\}$ *on matriisin* $A \in \mathbb{C}^{n\times n}$ ominaisarvoa $\lambda \in \mathbb{C}$ vastaava yleistetty ominaisvektori*, jos on olemassa sellainen* $k \in \mathbb{N}$*, että* $(A-\lambda I)^k v = 0$*. Aliavaruutta*

$$G(\lambda, A) = \{v \in \mathbb{C}^{n\times 1} \colon \textit{on olemassa sellainen } k \in \mathbb{N}, \textit{ että } (A-\lambda I)^k v = 0\}.$$

kutsutaan matriisin A ominaisarvoa λ vastaavaksi yleistetyksi ominaisavaruudeksi*.*

Muutama kommentti ja tulkinta on paikallaan.

Huomautus 18.1.2. *Vaikka matriisilla* $A \in \mathbb{C}^{n\times n}$ *on yleistettyjä ominaisvektoreita, ei sillä ole erillisiä yleistettyjä ominaisarvoja, vaan yleistetyt ominaisvektorit liittyvät aina matriisin ominaisarvoihin. Huomaa myös, että ominaisvektorit ovat yleistettyjä ominaisvektoreit, sillä jokaisella ominaisarvolla* $\lambda \in \mathbb{C}$ *pätee* $E(\lambda, A) \subset G(\lambda, A)$

Väite, että yleistetty ominaisavaruus todella on aliavaruus, vaatii lyhyen todistuksen.

Lemma 18.1.3. *Olkoon* $A \in \mathbb{C}^{n\times n}$ *neliömatriisi ja* $\lambda \in \mathbb{C}$ *sen yleistetty ominaisarvo. Tällöin yleistetty ominaisavaruus* $G(\lambda, A)$ *on avaruuden* $\mathbb{C}^{n\times 1}$ *aliavaruus.*

Todistus. Olkoot $v, w \in G(\lambda, A)$ ja $a \in \mathbb{C}$. Tällöin on olemassa sellaiset $k \in \mathbb{N}$ ja $\ell \in \mathbb{N}$, että $(A-\lambda I)^k v = 0$ ja $(A-\lambda I)^\ell w = 0$. Olkoon nyt $p = \max\{k, \ell\}$. Tällöin $(A-\lambda I)^p(av) = a(A-\lambda I)^{p-k}(A-\lambda I)^k v = 0$ ja vastaavasti $(A-\lambda I)^p w = (A-\lambda I)^{p-\ell}(A-\lambda I)^\ell w = 0$. Näin ollen $(A-\lambda I)^p(av + w) = 0$ eli $av + w \in G(\lambda, A)$. □

Tarkastellaan nyt karakteristista esimerkkiä, joka antaa tulkinan yleistetyille ominaisarvoille ja oiminaisvektoreille.

Esimerkki 18.1.4. *Olkoon*

$$A = J_{2,\lambda} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

missä $\lambda \in \mathbb{C}$*.*

Matriisin A *ainoa ominaisarvo on* λ *ja* $E(\lambda, A) = \mathrm{Sp}(e_1)$*. Toisaalta standardikannan* (e_1, e_2) *toiselle vektorille* e_2 *pätee*

$$(A-\lambda I)e_2 = Ae_e - \lambda e_2 = e_1 + \lambda e_2 - \lambda e_2 = e_1,$$

joten

$$(A-\lambda I)^2 e_2 = (A-\lambda I)(A-\lambda I)e_2 = (A-\lambda)e_1 = 0.$$

Näin ollen $e_2 \in G(0, A)$ *ja*

$$G(0, A) = \mathrm{Sp}(e_1, e_2) = \mathbb{R}^{2\times 1}.$$

Esimerkki 18.1.5. *Olkoon nyt*

$$A = J_{3,\lambda} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Jälleen λ on matriisin A ainoa ominaisarvo ja $E(\lambda, A) = \mathrm{Sp}(e_1)$. Nyt samat laskut kuin edellisessä esimerkissä osoittavat, että

$$(A - \lambda I)e_2 = e_1$$

ja

$$(A - \lambda I)e_3 = e_2.$$

Näin ollen

$$(A - \lambda I)^2 e_2 = 0$$

ja

$$(A - \lambda I)^3 e_3 = 0.$$

Näin ollen

$$G(\lambda, A) = \mathrm{Sp}(e_1, e_2, e_3) = \mathbb{R}^{3\times 1}.$$

Kuten arvata saattaa, edellisten esimerkkien tulokset ovat täysin yleisiä ja ne selittävät osaltaan Jordanin normaalimuodon Jordanin lohkoja. Seuraavan lemman yksityiskohdat jätetään harjoitustehtäväksi.

Lemma 18.1.6. *Olkoon $\lambda \in \mathbb{C}$. Tällöin $E(\lambda, J_{n,\lambda}) = \mathrm{Sp}(e_1)$ ja*

$$(J_{n,\lambda} - \lambda I)e_j = e_{j-1}$$

jokaisella $j \in \{2, \ldots, n\}$. Erityisesti

$$G(\lambda, J_{n,\lambda}) = \mathbb{R}^{n\times 1}.$$

Huomautus 18.1.7. *Koska $J_{n,\lambda} - \lambda I = N_n$, niin edellisen lemman olisi voinut myös kirjoittaa muodossa $E(0, N_n) = \mathrm{Sp}(e_1)$, $N_n e_j = e_{j-1}$ jokaisella $j \in \{2, \ldots, n\}$ ja $G(0, N_n) = \mathbb{R}^{n\times 1}$.*

Kirjataan vielä yksi perusesimerkki.

Esimerkki 18.1.8. *Olkoon*

$$A = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

Oletetaan ensin, että $\lambda_1 \neq \lambda_2$. Tällöin matriisilla A on kaksi ominaisarvoa λ_1 ja λ_2. Näitä ominaisarvoja vastaavat ominaisavaruudet ovat $E(\lambda_1, A) = \mathrm{Sp}(e_1)$ ja $E(\lambda_2, A) = \mathrm{Sp}(e_3)$. Kuten edellisissä esimerkeissä havaitaan nyt, että $e_2 \in G(\lambda_1, A)$ ja $e_3 \in G(\lambda_2, A)$. Näin ollen $G(\lambda_1, A) = \mathrm{Sp}(e_1, e_2)$ ja $G(\lambda_2, A) = \mathrm{Sp}(e_3, 3_4)$.

Oletetaan nyt, että $\lambda_1 = \lambda_2$. Tällöin $E(\lambda_1, A) = \mathrm{Sp}(e_1, e_3)$ ja $G(\lambda_1, A) = \mathbb{R}^{4\times 1}$.

18.1.1 Yleistetyn ominaisavaruuden dynaaminen tulkinta

Esimerkeistä 18.1.4 ja 18.1.5 herää kysymys onko yleistetyllä ominaisavaruudella dynaamista tulkitaa. Kyllä on. Tarkastellaan lineaarikuvausta $\varphi_{A-\lambda I}\colon \mathbb{C}^{n\times 1} \to \mathbb{C}^{n\times 1}$, $v \mapsto (A-\lambda I)v$. Koska matriisien kertolasku vastaa kuvausten yhdistämistä, niin jokaisella $k \in \mathbb{N}$ saadaan

$$\varphi_{(A-\lambda I)^k} = \varphi_{A-\lambda I} \circ \cdots \circ \varphi_{A-\lambda I} = (\varphi_{A-\lambda I})^k.$$

Näin ollen

> Vektori v kuuluu yleistettyyn ominaisavaruuteen $G(\lambda, A)$, jos ja vain jos vektorin v *rata* $(v, \varphi_{A-\lambda I)}(v), \varphi^2_{A-\lambda I}(v), \ldots, \varphi^k_{A-\lambda I}(v), \ldots)$ sisältää nollavektorin eli vektorin v rata on $(v, \varphi_{A-\lambda I)}(v), \varphi^2_{A-\lambda I}(v), \ldots, \varphi^{k-1}_{A-\lambda I}(v), 0, 0, \ldots)$.

18.1.2 Operaattorin yleistetty ominaisavaruus

Operaattorin yleistetyt ominaisvektorit määrittellään arvattavalla tavalla. Määritelmää varten tehdään kuitenkin seuraavat havainnot. Jos $f\colon V \to V$ on operaattori ja $\lambda \in \mathbb{C}$ sen ominaisarvo, niin $(f - \lambda\mathrm{id})\colon V \to V$ on avaruuden V operaattori. Lisäksi, mikäli $g\colon V \to V$ on operaattori, niin $g^k\colon V \to V$ on operaattori $g^k = g\circ\cdots\circ g\colon V \to V$, missä kuvaus g on yhdistetty itseensä k kertaa.

Määritelmä 18.1.9. *Vektori $v \in V\setminus\{0\}$ on operaattorin $f\colon V \to V$* ominaisarvoa $\lambda \in \mathbb{C}$ vastaava yleistetty ominaisvektori*, jos on olemassa sellainen $k \in \mathbb{N}$, että $(f-\lambda\mathrm{id})^k(v) = 0$. Aliavaruutta*

$$G(\lambda, f) = \{v \in V\colon \text{ on olemassa sellainen } k \in \mathbb{N}, \text{ että } (f-\lambda\mathrm{id})^k(v) = 0\}$$

kutsutaan operaattorin f ominaisarvoa λ vastaavaksi yleistetyksi ominaisavaruudeksi*.*

Koska operaattorin yleistetty ominaisavaruus osoitetaan aliavaruudeksi kuten matriisien tapauksessa, todistus sivuutetaan. Operaattorin $f\colon V \to V$ yleistetyn ominaisavaruuden $G(\lambda, f)$ määritelmästä havaitaan suoraan, että se on operaattorin $f - \lambda\mathrm{id}\colon V \to V$ invariantti aliavaruus.

Lemma 18.1.10. *Olkoon V vektoriavaruus, $f\colon V \to V$ operaattori ja $\lambda \in \mathbb{C}$ operaattorin f ominaisarvo. Tällöin $G(\lambda, f)$ on operaattorien $(f-\lambda\mathrm{id})$ ja f invariantti aliavaruus eli*

$$(f-\lambda\mathrm{id})(G(\lambda, f)) \subset G(\lambda, f)$$

ja

$$f(G(\lambda, f)) \subset G(\lambda, f).$$

Todistus. Olkoon $v \in G(\lambda, f)$. Tällöin on olemassa sellainen $k \in \mathbb{N}$, että $(f-\lambda\mathrm{id})^k(v) = 0$. Näin ollen $(f-\lambda\mathrm{id})(v) = (f-\lambda\mathrm{id})^{k+1}(v) = 0$ eli $(f-\lambda\mathrm{id})(v) \in G(\lambda, f)$. Tämä todistaa ensimmäisen väitteen.

Olkoon jälleen $v \in G(\lambda, f)$. Koska $G(\lambda, f)$ on aliavaruus, niin edellisen kohdan perusteella pätee $f(v) = (f - \lambda + id)(v) + \lambda v \in G(\lambda, f)$. Näin ollen $G(\lambda f)$ on operaattorin f invarianttialiavaruus. $\square$

18.2 Yleistetty ominaisavaruus on nolla-avaruus

Tässä luvussa osoitetaan, että yleistetty ominaisavaruus voidaan aina esittää nolla-avaruutena.

Lause 18.2.1. *Olkoon $A \in \mathbb{C}^{n\times n}$ neliömatriisi ja $\lambda \in \mathbb{C}$ sen ominaisarvo. Tällöin*

$$G(\lambda, A) = \mathrm{Null}(A - \lambda I)^n.$$

Tulos on mielenkiintoinen kahdesta syystä. Ensinnäkin se erityisesti sanoo, että jokaisella $v \in G(\lambda, A)$ pätee $(A-\lambda I)^n v = 0$. Tämä tarkoittaa, että yleistetyn ominaisvektorin määritelmässä jokaisella $v \in G(\lambda, A)$ voidaan aina valita $k = n$. Tämä puolestaan sanoo, että matriisin yleistetty ominaisavaruus $G(\lambda, A)$ voidaan löytää algoritmisesti etsimällä nolla-avaruudelle $\mathrm{Null}(A - \lambda I)^n$ kanta.

Esimerkki 18.2.2. *Olkoon*

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Tällöin

$$0 = \det(A - \lambda I) = (3-\lambda)(2-\lambda)^2$$

joten $\lambda_1 = 3$ ja $\lambda_2 = 2$ ovat matriisin A ainoat ominaisarvo.

Etsitään yleistetty ominaisavaruus $G(\lambda_1, A)$. Koska

$$(A - \lambda_1 I)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

ja

$$(A - \lambda_1 I)^3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

niin

$$G(\lambda_1, A) = \mathrm{Null}(A - \lambda_1 I)^3 = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right).$$

Etsitään nyt yleistetty ominaisavaruus $G(\lambda_2, A)$. Koska

$$(A - \lambda_2 I)^2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

ja

$$(A - \lambda_2 I)^3 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

niin

$$G(\lambda_2, A) = \mathrm{Null}(A - \lambda_2 I) = \mathrm{Sp}(e_2, e_3).$$

Huomautus 18.2.3. *Huomaa, että edellisessä esimerkissä ei laskettu lainkaan ominaisavaruuksia $E(\lambda_1, A)$ ja $E(\lambda_2, A)$. On kuitenkin hyödyllistä huomata, että tässä kyseisessä tapauksessa $G(\lambda_1, A) = E(\lambda_1, A)$ ja $G(\lambda_2, A) \neq E(\lambda_2, A)$. Ominaisarvon λ_1 tapauksessa huomio seuraa suoraan havainnosta, että $1 \leq \dim E(\lambda_1, A) \leq \dim G(\lambda_1, A) = 1$. Ominaisarvon λ_2 kohdalla on mielenkiintoista myös havaita, että $\dim E(\lambda_2, A) = 1$, $E(\lambda_2, A) = \mathrm{Sp}(e_2)$ ja että $(A - \lambda_2 I)e_3 = e_2$.*

Lauseen 18.2.1 todistus perustuu seuraavien aliavaruuksien tarkasteluun.

Merkintä 18.2.4. *Olkoon $A \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}$ matriisin A yleistetty ominaisarvo ja $k \in \mathbb{N}$. Merkitään*

$$G_k(\lambda, A) = \{v \in \mathbb{C}^{n\times 1} \colon (A - \lambda I)^k v = 0\} = \mathrm{Null}(A - \lambda I)^k.$$

Huomautus 18.2.5. *Suoraan määritelmästä seuraa, että $G_1(\lambda, A) = E(\lambda, A)$ ja että $G_k(\lambda, A) \subset G(\lambda, A)$. Koska jokainen $v \in G(\lambda, A)$ kuuluu määritelmän nojalla johonkin aliavaruuksista $G_k(\lambda, A)$, niin*

$$\bigcup_{k=1}^{\infty} G_k(\lambda, A) = G(\lambda, A).$$

Lisäksi

$$G_1(\lambda, A) \subset G_2(\lambda, A) \subset \cdots \subset G_k(\lambda, A) \subset \cdots \subset G(\lambda, A).$$

Vaikka edellä annetaan ymmärtää, että eri aliavaruuksia $G_k(\lambda, A)$ voisi periaatteessa olla ääretön määrä, niin näin ei kuitenkaan ole, koska paljastuu, että $G_k(\lambda, A)$ on aina avaruuden $G_{k+1}(\lambda, A)$ aito aliavaruus, jos $G_k(\lambda, A) \neq G(\lambda, A)$. Kirjataan nyt tämä tulos lemmaksi seuraavassa muodossa, joka heuristisesti sanoo, että jos aliavaruuksien jono $G_1(\lambda, A) \subset G_2(\lambda, A) \subset \cdots$ pysähtyy hetkellisesti, niin se pysähtyy lopullisesti.

Lemma 18.2.6. *Olkoon $A \in \mathbb{C}^{n\times n}$ ja $m \in \mathbb{N}$. Jos $G_m(\lambda, A) = G_{m+1}(\lambda, A)$, niin $G(\lambda, A) = G_m(\lambda, A)$.*

Todistus. Osoitetaan, että $G_k(\lambda, A) \subset G_m(\lambda, A)$ jokaisella $k > m$. Tällöin

$$G(\lambda, A) = \bigcup_{k=1}^{m} G_k(\lambda, A) = \bigcup_{k=1}^{\infty} G_k(\lambda, A) \cup \bigcup_{k=m}^{\infty} G_k(\lambda, A) \subset G_m(\lambda, A) \cup G_m(\lambda, A) = G_m(\lambda, A).$$

Todistetaan väite induktiolla luvun $k > m$ suhteen. Jos $k = m + 1$, niin väite pätee oletuksen nojalla. Oletetaan nyt, että väite pätee luvulla $k > m$ ja osoitetaan, että $G_{k+1}(\lambda, A) \subset G_m(\lambda, A)$.

Olkoon $v \in G_{k+1}(\lambda, A)$. Tällöin $(A - \lambda I)^k (A - \lambda I)v = (A - \lambda I)^{k+1} v = 0$. Näin ollen $(A - \lambda I)v \in G_k(A, \lambda)$. Koska $G_k(\lambda, A) \subset G_m(\lambda, A)$, niin $(A - \lambda I)^m (A - \lambda I)v = 0$. Näin ollen $v \in G_{m+1}(\lambda, A) = G_m(\lambda, A)$. $\square$

Edellisestä lemmasta seuraa, että jonossa $G_1(\lambda, A) \subset G_2(\lambda, A) \subset \cdots$ voi olla korkeintaa n eri suurta aliavaruutta. Todistetaan nyt tarkasti hieman tarkempi tulos.

Lause 18.2.7. *Olkoon $A \in \mathbb{C}^{n\times n}$. Tällöin*

$$G(\lambda, A) = G_m(\lambda, A),$$

missä $m = \dim G(\lambda, A) \leq n$.

Todistus. Olkoon $m = \dim G(\lambda, A) \leq n$. Tehdään vastaoletus, että $G_m(\lambda, A) \neq G(\lambda, A)$. Tällöin lemman 18.2.6 nojalla $G_k(\lambda, A) \neq G_{k+1}(\lambda, A)$ jokaisella $k \in \{1, \ldots, m\}$. Näin ollen $\dim G_k(\lambda, A) < \dim G_{k+1}(\lambda, A)$ jokaisella $k \in \{1, \ldots, m\}$. Näin ollen

$$\dim G(\lambda, A) \geq \dim G_{m+1}(\lambda, A) \geq m + \dim G_1(\lambda, A) = m + \dim E(\lambda, A) > m = \dim G(\lambda, A).$$

Tämä on ristiriita, joten $G_m(\lambda, A) = G(\lambda, A)$. □

Lause 18.2.1 seuraa lähes välittömästi lauseesta 18.2.7.

Lauseen 18.2.1 todistus. Merkitään $m = \dim G(\lambda, A)$. Olkoon $v \in G(\lambda, A)$. Koska $G(\lambda, A) = G_m(\lambda, A)$, niin

$$(A - \lambda I)^n v = (A - \lambda I)^{n-m}(A - \lambda I)^m v = 0.$$

Näin ollen $v \in \mathrm{Null}(A - \lambda I)^n$.

Olkoon nyt $v \in \mathrm{Null}(A - \lambda, I)^n$. Tällöin $(A - \lambda I)^n v = 0$ eli $v \in G(\lambda, A)$. Tämä todistaa ensimmäisen väitteen. □

Edelliset lauseet voi todistaa äärellisulotteisten avaruuksien operaattoreille joko suoraan samoilla todistuksilla kuin neliömatriiseille tai hyödyntämällä seuraavaa lausetta, joka sanoo kuinka operaattorin ja sen esitysmatriisin yleistettyt ominaisavaruudet vastaavat toisiaan. Koska tämän lauseen todistus on hyvin samankaltainen kuin korollaarin 11.3.2

Lause 18.2.8. *Olkoon V äärellisulotteinen vektoriavaruus ja $(v_1, \ldots, v_n)$ sen kanta. Olkoon $f\colon V \to V$ lineaarikuvaus ja A sen esitysmatriisi kannassa $(v_1, \ldots, v_n)$. Tällöin jokaisella lineaarikuvauksen f ominaisarvolla $\lambda \in \mathbb{C}$ pätee*

$$G(\lambda, f) = \Psi_{(v_1,\ldots,v_n)}(G(\lambda, A)).$$

Käyttäen kumpaa tahansa argumenttia lause 18.2.1 voidaan kirjata seuraavassa muodossa.

Korollaari 18.2.9. *Olkoot V äärellisulotteinen vektoriavaruus, $f\colon V \to V$ operaattori ja $\lambda \in \mathbb{C}$ operaattorin f ominaisarvo. Tällöin*

$$G(\lambda, f) = \ker(f - \lambda \mathrm{id})^n$$

18.2.1 Yleistetyn ominaisavaruuden komplementti

Tässä luvussa todistetaan seuraava hieman yllättäväkin tulos.

Lause 18.2.10. *Olkoon $A \in \mathbb{C}^{n\times n}$. Tällöin*

$$\mathbb{C}^{n\times 1} = G(\lambda, A) \oplus \mathrm{Col}(A - \lambda I)^n.$$

Tulos on yllättävä, koska yleisesti neliömatriisin $A \in \mathbb{C}^{n\times n}$ nolla- ja sarakeavaruuden summa ei ole suora eikä $\mathbb{C}^{n\times 1}$. Huomaa, että lauseen 18.2.1 perusteella pätee $G(\lambda, A) = \mathrm{Null}(A - \lambda I)^n$.

Esimerkki 18.2.11. *Olkoon*

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

Tällöin $\mathrm{Null}(A) = \mathrm{Sp}(e_1)$ *ja* $\mathrm{Col}(A) = \mathrm{Sp}(e_1)$ *eli* $\mathrm{Null}(A) + \mathrm{Col}(A) = \mathrm{Sp}(e_1)$*. Toisaalta* $A^2 = 0$*, joten* $\mathrm{Null}(A^2) = \mathbb{C}^{2\times 1}$ *ja* $\mathrm{Col}(A) = \{0\}$*.*

Lauseen 18.2.10 tuloksen tekee vieläkin yllättävämmäksi se, että se ei liity yleistettyihin ominaisavaruuksiin muutoin kuin lauseen 18.2.1 välityksellä, eli se voidaan muotoilla täysin yleisesti tuloksena operaattoreille. Koska tulosta käytetään jatkossa operaattoreille, niin kirjataan se siksi tässä muodossa. Seuraavassa lauseessa merkinnällä $f^n \colon V \to V$ tarkoitetaan operaattorin $f \colon V \to V$ n-kertaista yhdistettyä kuvausta itsensä kanssa eli $f^n = f \circ f^{n-1}$ ja $f^0 = \mathrm{id}$.

Lause 18.2.12. *Olkoon V äärellisulotteinen vektoriaaruus ja $n = \dim V$. Tällöin operaattorille $f \colon V \to V$ pätee*

$$V = \ker f^n \oplus \operatorname{im} f^n.$$

Lisäksi sekä $\ker f^n$ *että* $\operatorname{im} f^n$ *ovat operaattorin f invariantteja aliavaruuksia eli lineaarikuvaukset* $f|_{\ker f^n} \colon \ker f^n \to \ker f^n$ *ja* $f|_{\operatorname{im} f^n} \colon \operatorname{im} f^n \to \operatorname{im} f^n$ *ovat hyvin määriteltyjä operaattoreita.*

On mielenkiintoista havaita, että tämän lauseen todistus on hyvin samankaltainen kuin lauseen 18.2.7. Muotoillaan tätä varten lemma.

Lemma 18.2.13. *Olkoon V äärellisulotteinen vektoriavaruus, $n = \dim V$ ja $f \colon V \to V$ operaattori. Tällöin* $\ker f^n = \ker f^m$ *kaikilla $m \geq n$.*

Todistus. Osoitetaan ensin, että jos $\ker f^k = \ker f^{k+1}$ jollain $k \in \mathbb{N}$, niin tällöin $\ker f^k = \ker f^m$ kaikilla $m \geq k$. Oletetaan, että väite pätee luvulla $m \in \mathbb{N}$ eli $\ker f^k = \ker f^m$. Olkoon $v \in \ker f^{m+1}$. Tällöin $f(v) \in \ker f^m$. Koska $\ker f^m = \ker f^k$, niin $f(v) \in \ker f^k$ ja $v \in \ker f^{k+1} = \ker f^k$. Näin ollen $\ker f^k = \ker f^m$ kaikilla $m \geq k$.

Osoitetaan nyt, että luvulla $n = \dim V$ pätee $\ker f^n = \ker f^{n+1}$. Tehdään nyt vastaoletus, että $\ker f^n \neq \ker f^{n+1}$. Tällöin $\ker f^k \neq \ker f^{k+1}$ jokaisella $1 \leq k \leq n$. Koska $\ker f \subset \ker f^2 \subset \cdots$, niin $\dim\ker f^{n+1} \geq 1 + \dim\ker f^n \geq 2 + \dim\ker f^{n-1} \geq n + 1 > \dim V$. Tämä on ristiriita, joten $\ker f^n = \ker f^{n+1}$ ja siten $\ker f^n = \ker f^m$ jokaisella $m \geq n$. $\square$

Lauseen 18.2.12 todistus. Koska

$$\dim V = \dim \ker f^n + \dim \operatorname{im} f^n,$$

niin ensimmäisen väitteen todistamiseksi riittää osoittaa, että $\ker f^n \cap \operatorname{im} f^n = \{0\}$. Olkoon $v \in \ker f^n \cap \operatorname{im} f^n$. Koska $v \in \operatorname{im} f^n$, niin on olemassa sellainen $w \in V$, että $f^n(w) = v$. Koska $v \in \ker f^n$, niin $f^{2n}(w) = f^n(v) = 0$. Näin ollen $w \in \ker f^{2n} = \ker f^n$ ja $v = f^n(w) = 0$.

Osoitetaan nyt toinen väite eli aliavaruuksien $\ker f^n$ ja $\operatorname{im} f^n$ invarianttius. Molemmat seuraavat lähes suoraan huomiosta, että $f \circ f^n = f^{n+1} = f^n \circ f$.

Olkoon $v \in \ker f^n$. Tällöin $f^n(f(v)) = f(f^n(v)) = 0$ eli $f(v) \in \ker f^n$. Näin ollen $f(\ker f^n) \subset \ker f^n$. Olkoon nyt $v \in \operatorname{im} f^n$. Tällöin on olemassa sellainen $w \in V$, että $f^n(w) = v$. Näin ollen $f(v) = f(f^n(w)) = f^n(f(w)) \in \operatorname{im} f^n$ eli $f(\operatorname{im} f^n) \subset \operatorname{im} f^n$.

□

18.3 Yleistettyjen ominaisavaruuksien summa

Tässä luvussa osoitetaan yllättävä tulos, että yleistettyjen ominaisavaruuksien summa on koko avaruus. Tarkemmin tällä tarkoitetaan seuraavaa.

Lause 18.3.1. *Olkoon $A \in \mathbb{C}^{n\times n}$ neliömatriisi ja kaikki $\lambda_1, \ldots, \lambda_k$ matriisin A erisuuret ominaisarvot. Tällöin*

$$\mathbb{C}^{n\times 1} = G(\lambda_1, A) \oplus \cdots \oplus G(\lambda_k, A).$$

Tätä tulosta kannattaa välittömästi verrata ominaisavaruuksia koskeviin tuloksiin. Kuten on tunnettua, matriisin $A \in \mathbb{C}^{n\times n}$ ominaisavaruuksien summa on koko avaruus $\mathbb{C}^{n\times 1}$, jos ja vain jos avaruudella $\mathbb{C}^{n\times 1}$ on matriisin A ominaisvektoreista koostuva kanta eli että matriisi A on diagonalisoituva. Lause 18.3.1 voidaan tulkita sanomalla, että avaruudella $\mathbb{C}^{n\times 1}$ on aina matriisin A yleistetyistä ominaisvektoreista koostuva kanta. Kirjataan tämä (hämmästyttävä) tulos korollaariksi.

Korollaari 18.3.2. *Olkoon $A \in \mathbb{C}^{n\times n}$ neliömatriisi. Tällöin avaruudella $\mathbb{C}^{n\times 1}$ on matriisin A yleistetyistä ominaisvektoreista koostuva kanta.*

Todistus. Olkoot $\lambda_1, \ldots, \lambda_k$ matriisin A kaikki erisuuret ominaisarvot. Olkoon jokaisella $i \in \{1, \ldots, n\}$ jono $(v_{1i}, \ldots, v_{d_i i})$ aliavaruuden $G(\lambda_i, A)$ kanta, missä $d_i = \dim G(\lambda_i, A)$. Tällöin lauseen 18.3.1 nojalla jono $(v_{11}, \ldots, v_{d_1 1}, d_{21}, \ldots, v_{d_k k})$ on avaruuden $\mathbb{C}^{n\times 1}$ kanta.

□

Lauseen 18.3.1 todistus jakautuu kahteen osaan. Ensin osoitetaan omana tuloksenaan, että yleistettyjen ominaisavaruuksien summa on aina suora. Tämän jälkeen osoitetaan, että yleistettyjen ominaisavaruuksien summa on koko avaruus $\mathbb{C}^{n\times 1}$.

18.3.1 Yleistettyjen ominaisavaruuksien summa on suora

Hieman yllättäen lauseen 18.2.7 tulosta voidaan hyödyntää osoittamaan, että yleistettyjen ominaisavaruuksien summa on suora. Kirjataan tulos tämä seuraavasti.

Lause 18.3.3. *Olkoon $A \in \mathbb{C}^{n\times n}$ ja olkoot $\lambda_1, \dots, \lambda_k$ kaikki matriisin A erisuuret ominaisarvot. Tällöin*

$$G(\lambda_1, A) + \cdots + G(\lambda_k, A) = G(\lambda_1, A) \oplus \cdots \oplus G(\lambda_k, A).$$

Tulosta varten tarvitaan triviaali – mutta oleellinen – havainto.

Lemma 18.3.4. *Olkoot $A \in \mathbb{C}^{n\times n}$ ja $\lambda, \mu \in \mathbb{C}$. Tällöin*

$$(A - \lambda I)(A - \mu I) = (A - \mu I)(A - \lambda I).$$

Todistus. Koska

$$\begin{aligned}(A - \lambda I)(A - \mu I) &= A(A - \mu I) - \lambda I(A - \mu I)\\ &= A^2 - \mu I - \lambda A + \lambda\mu I\\ &= A^2 - \mu I - \lambda A + \mu\lambda I\\ &= (A - \mu I)(A - \lambda I),\end{aligned}$$

väite on todistettu.

□

Lauseen 18.3.3 todistus. Olkoon $v_i \in G(\lambda_i, A)$ nollasta poikkeava vektori jokaisella $i \in \{1, \dots, k\}$ ja olkoot $a_1, \dots, a_n \in \mathbb{C}$ sellaisia lukuja, että

$$a_1 v_1 + \cdots + a_k v_k = 0.$$

Lauseen 18.2.7 perusteella $G(\lambda_1, A) = G_{m_1}(\lambda_1, A)$, missä $m_1 = \dim G(\lambda_1, A)$. Näin ollen $(A - \lambda_1 I)^{m_1} v_1 = 0$ ja on olemassa suurin sellainen luku $m < m_1$, että $(A - \lambda_1 I)^m v_1 \neq 0$.

Osoitetaaan, että $a_1 = 0$. Olkoon $w = (A - \lambda_1 I)^m v_1$. Koska $(A - \lambda_1 I)w = (A - \lambda_1 I)^{m+1} v_1 = 0$, niin $Aw = \lambda_1 w_1$ eli $w \in E(\lambda_1, A)$. Koska

$$(A - \lambda_j I)w = Aw - \lambda_j w = \lambda_1 w - \lambda_j w = (\lambda_1 - \lambda_j)w$$

jokaisella $j \in \{2, \dots, k\}$, niin

$$(A - \lambda_j I)^n w = (\lambda_1 - \lambda_j)^n w$$

jokaisella $j \in \{2, \dots, k\}$.

Olkoon nyt

$$B = (A - \lambda_2 I)^n \cdots (A - \lambda_k I)^n.$$

ja

$$\mu = (\lambda_1 - \lambda_2)^n \cdots (\lambda_1 - \lambda_k)^n \neq 0.$$

Tällöin
$$Bw = \mu w.$$

Lemman 18.3.4 perusteella matriisin B tekijät $(A-\lambda_j I)^n$ voidaan uudeelleen järjestää haluttuun järjestykseen. Koska $(A - \lambda_j I)^n v_j = 0$ jokaisella $j \in \{2, \dots, k\}$, niin
$$Bv_j = 0$$
kaikilla $j \in \{2, \dots, k\}$. Näin ollen
$$\begin{aligned}0 &= (A - \lambda_1 I)^m B\,(a_1 v_1 + \cdots + a_k v_k)\\ &= (A - \lambda_1 I)^m B(a_1 v_1)\\ &= a_1 B(A - \lambda_1 I)^m v_1 = a_1 B w_1 = a_1 \mu w_1.\end{aligned}$$
Koska $w_1 \neq 0$ ja $\mu \neq 0$, niin $a_1 = 0$.

Vastaavasti osoitetaan, että $a_j = 0$ kaikilla $j \in \{1, \dots k\}$. Tämä päättää todistuksen. □

18.3.2 Yleistetyt ominaisavaruudet virittävät koko avaruuden

Todistetaan nyt lauseen 18.3.1 todistuksen toinen puoli eli että yleistetyt ominaisavaruudet virittävät koko avaruuden. Koska todistus kirjoitetaan operaattorien avulla, kirjataan myös tulos operaattoreille.

Lause 18.3.5. *Olkoot V äärellisulotteinen vektoriavaruus, $f\colon V \to V$ operaattori ja $\lambda_1, \dots, \lambda_k$ kaikki operaattorin f erisuuret ominaisarvot. Tällöin*
$$V = G(\lambda_1, f) \oplus \cdots \oplus G(\lambda_k, f).$$

Todistus. Todistetaan väite induktiolla dimension $n = \dim V$ suhteen. Selvästi väite pätee jos $\dim V = 1$, sillä tällöin operaattorilla $f\colon V \to V$ on täsmälleen yksi ominaisarvo $\lambda_1 \in \mathbb{C}$ ja $G(\lambda_1, f) = E(\lambda_1, f) = V$. Oletetaan nyt, että $n \in \mathbb{N}$ on sellainen luku, että väite pätee kaikilla äärellisulotteisilla vektoriavaruuksilla V ja operaattoreilla $f\colon V \to V$, joilla $\dim V < n$. Osoitetaan, että väitä pätee luvulla n.

Olkoon V n-ulotteinen vektoriavaruus, $f\colon V \to V$ operaattori ja $\lambda_1, \dots, \lambda_k$ kaikki operaattorin f erisuuret ominaisarvot.

Olkoon $f_1 = f - \lambda_1 \mathrm{id}\colon V \to V$. Tällöin korollaarin 18.2.9 perusteella $G(\lambda_1, f) = \ker(f - \lambda_1 \mathrm{id})^n = \ker f_1^n$ ja lauseen 18.2.12 perusteella
$$V = \ker f_1^n \oplus \operatorname{im} f_1^n.$$

Merkitään nyt $W = \operatorname{im} f_1^n$ ja osoitetaan, että
$$W \subset G(\lambda_2, f) \oplus \cdots \oplus G(\lambda_k, f).$$

Koska W on operaattorin f_1 invariantti aliavaruus, niin $g = f_1|_W\colon W \to W$ on hyvin määritelty operaattori. Koska $G(\lambda_1, f) \supset E(\lambda_1, f) \neq \{0\}$, niin $\dim W < n$. Näin ollen
$$W = G(\mu_1, g) \oplus \cdots \oplus G(\mu_\ell, g)$$

missä $\mu_1, \ldots, \mu_\ell$ ovat kaikki operaattorin g erisuuret ominaisarvot.

Osoitetaan, että $\lambda_1 + \mu_j$ on operaattorin f ominaisarvo jokaisella $j \in \{1, \ldots, \ell\}$ ja että $G(\mu_j, g) \subset G(\lambda_1 + \mu_j, f)$ jokaisella $j \in \{1, \ldots, k\}$.

Olkoon $v \in E(\mu_j, g)$ nollasta poikkeava vektori. Tällöin

$$(f - (\lambda_1 + \mu_j)\mathrm{id})(v) = (g - \mu_j \mathrm{id})(v) = 0.$$

Näin ollen $\lambda_1 + \mu_j$ on operaattorin f ominaisarvo. Olkoon nyt $v \in G(\mu_j, g)$. Tällöin $(g - \mu_j \mathrm{id})^n(v) = 0$. Koska $g(v) = (f - \lambda_1 \mathrm{id})(v)$, niin $(f - (\lambda_1 + \mu_j)\mathrm{id})^n(v) = 0$ eli $v \in G(\lambda_1 + \mu_j, f)$. Näin ollen $G(\mu_j, g) \subset G(\lambda_1 + \mu_j, f)$.

Koska $\lambda_1, \ldots, \lambda_k$ ovat kaikki operaattorin f ominaisarvot, niin

$$G(\mu_j, g) \subset G(\lambda_1, f) \oplus G(\lambda_2, f) \oplus \cdots \oplus G(\lambda_k, f)$$

kaikilla $j \in \{1, \ldots, \ell\}$. Näin ollen

$$W \subset G(\lambda_1, f) \oplus G(\lambda_2, f) \oplus \cdots \oplus G(\lambda_k, f).$$

Koska $W \cap G(\lambda_1, f) = \{0\}$, niin

$$W \subset G(\lambda_2, f) \oplus \cdots \oplus G(\lambda_k, f).$$

Tämä päättää todistuksen. $\square$

18.4 Neliömatriisin Jordanin normaalimuoto

Kuten nyt on helppo arvata, että matriisin $A \in \mathbb{C}^{n \times n}$ Jordanin normaalimuodon lohkot liittyvät matriisin yleistettyihin ominaisavaruuksiin $G(\lambda, A)$ ja matriiseihin $A - \lambda I$, missä $\lambda \in \mathbb{C}$ on matriisin A ominaisarvo. Esitellään tätä varten nilpotentin operaattorin ja matriisin käsitteet.

Määritelmä 18.4.1. *Operaattori $f \colon V \to V$ on* nilpotentti*, jos on olemassa sellainen luku $k \in \mathbb{N}$, että $f^k = 0$. Vastaavasti neliömatriisi $A \in \mathbb{C}^{n \times n}$ on* nilpotentti*, jos operaattori $\varphi_A \colon \mathbb{C}^{n \times 1} \to \mathbb{C}^{n \times 1}$, $v \mapsto Av$, on nilpotentti eli on olemassa sellainen $k \in \mathbb{N}$, että $A^k = 0$.*

Ei ole sattumaa, että Jordanin normaalimuodon yhteydessä esiteltiin matriisit $N_n \in \mathbb{C}^{n \times n}$. Ne ovat kanonisia esimerkkejä nilpontenteista matriiseista.

Esimerkki 18.4.2. *Huomaa, että $N_n e_1 = 0$ ja että $N_n e_i = e_{i-1}$ kaikilla $i \in \{2, \ldots, n\}$. Näin ollen $N_n^n = 0$. Esimerkiksi tapauksessa $n = 3$ sama voidaan havaita konkreettisesti laskemalla kaikki matriisitulot vaiheittain, sillä*

$$N_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$N_3^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

ja

$$N_3^3 = N_3^1 N_3^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

Huomautus 18.4.3. *Huomaa, että Jordan lohkot $J_{n,\lambda} = \lambda I_n + N_n$ on nilpotentti ainoastaan tapauksessa $\lambda = 0$. Tämä seuraa havainnosta, että λ on matriisin $J_{n,\lambda}$ ominaisarvo ja nilponentilla matriisilla ei ole nollasta poikkeavia ominaisarvoja.*

Vaikka kaikki matriisit – ja siten kaikki operaattorit – eivät ole nilponentteja, on mielenkiinoista huomata, että jokaiseen operaattoriin liittyy luonnollisella tavalla kokoelma nilpontentteja operaattoreita. Tarkemmin sanottuna jokaisella operaattorilla $f\colon V \to V$ ja sen ominaisarvolla λ operaattorin $f - \lambda\mathrm{id}$ rajoittuma yleistettyyn aliavaruuteen $G(\lambda, f)$ on nilpotentti. Muotoillaan tämä lähes suoraan korollaarista 18.2.9 seuraava tulos uudelleen lauseeksi sen tärkeyden vuoksi.

Lause 18.4.4. *Olkoon V äärellisulotteinen vektoriavaruus, $f\colon V \to V$ operaattori ja $\lambda \in \mathbb{C}$ operaattorin f ominaisarvo. Tällöin operaattori $(f - \lambda\mathrm{id})|_{G(\lambda,f)}\colon G(\lambda, f) \to G(\lambda, f)$ on hyvin määritelty nilpotentti operaattori.*

Todistus. Koska $G(\lambda, f)$ on lemman 18.1.10 nojalla operaattorin $(f - \lambda\mathrm{id})$ invariantti aliavaruus, niin operaattori $f_1 = (f - \lambda\mathrm{id})|_{G(\lambda,f)}\colon G(\lambda, f) \to G(\lambda, f)$ on hyvin määritelty. □

Seuraava lause (lause 18.4.8) osoittaa, että jokaisella nilpontentilla operaattorilla $f\colon V \to V$ on olemassa sellainen avaruuden V kanta, että tässä kannassa operaattorin f esitysmatriisi on ositettu matriisi, joka koostuu matriiseista N_n. Koska lauseen yleinen muoto on vaikea tulkintainen, tarkastellaan ensin kahta erikoistapausta ja lasketaan niissä matriisin Jordanin normaalimuoto.

Esimerkki 18.4.5. *Olkoon*

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

kuten esimerkissä 18.2.2. Matriisin A ominaisarvot ovat $\lambda_1 = 3$ ja $\lambda_2 = 2$ sekä näitä vastavat ominaisavaruudet ovat

$$G(\lambda_1, A) = \mathrm{Sp}(e_1 + e_2)$$

ja

$$G(\lambda_2, A) = \mathrm{Sp}(e_2, e_3).$$

Ominaisarvon λ_1 tapauksessa yleistetty ominaisavaruus $G(\lambda_1, A)$ on ominaisvaruus $E(\lambda_1, A)$. Näin ollen lineaarikuvaus $f_1 = (f_A - \lambda_1\mathrm{id})|_{G(\lambda_1,A)}\colon G(\lambda_1, A) \to G(\lambda_1, A)$ on *nolla-kuvaus ja vastaava matriisi on siis nollamatriisi. Lineaarikuvauksen* $f_A|_{G(\lambda_1,A)} \to G(\lambda_1, A) \to G(\lambda_1, A)$ *matriisi kannassa* $(e_1 + e_2)$ *puolestaan on* $J_{\lambda_1} = [\lambda_1] \in \mathbb{R}^{1\times 1}$.

Ominaisarvon λ_2 *tapauksessa yleistetty ominaisavaruus* $G(\lambda_2, A)$ *ei ole ominaisavaruus* $E(\lambda_2, A)$ *vaan* $E(\lambda_2, A) = \mathrm{Sp}(e_2)$. *Olkoon nyt* $f_2 = (f_A - \lambda_2 \mathrm{id})|_{G(\lambda_2,A)} \colon G(\lambda_2, A) \to G(\lambda_2, A)$. *Tällöin lineaarikuvaukselle* f_2 *pätee esimerkin 18.2.2 perusteella sekä*

$$f_2(e_2) = (f_A - \lambda_2 \mathrm{id})(e_2) = (A - \lambda_2 I)e_2 = 0$$

että

$$f_2(e_3) = (A - \lambda_2 I)e_3 = e_2.$$

Näin ollen lineaarikuvauksella f_2 *on kannassa* (e_2, e_3) *matriisi*

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Lineaarikuvauksen $f_A|_{G(\lambda_2,A)} \colon G(\lambda_2, A) \to G(\lambda_2, A)$ *matriisi on siis*

$$J_{\lambda_2} = \begin{bmatrix} \lambda_2 & 1 \\ 0 & \lambda_2 \end{bmatrix}.$$

Esimerkki 18.4.6. *Olkoon*

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Esimerkin 18.4.5 perusteella kuvauksen f_A *matriisi kannassa* $(e_1 + e_2, e_2, e_3)$ *on*

$$J = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Matriisilla A *on siis Jordanin hajotelma*

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

Tarkastellaan vielä toista versiota samasta esimerkistä.

Esimerkki 18.4.7. *Olkoon*

$$M = \begin{bmatrix} A & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A \end{bmatrix} \in \mathbb{R}^{9\times 9}$$

ositettu neliömatriisi.

Esimerkin 18.4.5 perusteella matriisin M *ominaisarvot ovat* $\lambda_1 = 3$, $\lambda_2 = 2$ *ja* $\lambda_3 = 0$. *Lisäksi*

$$G(\lambda_1, M) = E(\lambda_1, M) = \mathrm{Sp}(e_1 + e_2, e_7 + e_8),$$

$$G(\lambda_2, M) = \mathrm{Sp}(e_2, e_3, e_8, e_9)$$

ja

$$G(\lambda_3, M) = E(\lambda_3, M) = \mathrm{Sp}(e_4, e_5, e_6).$$

Nyt avaruuden $\mathbb{R}^{9\times 1}$ *kannassa*

$$(e_1 + e_2, e_7 + e_8, e_2, e_3, e_8, e_9, e_4, e_5, e_6)$$

kuvauksella $f_M \colon \mathbb{R}^{9\times 1} \to \mathbb{R}^{9\times 1}$ *on esitysmatriisi*

$$J = \begin{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} & & & \\ & \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} & & \\ & & \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} & \\ & & & 0 \end{bmatrix}.$$

Muotoillaan nyt Jordanin normaali muodon yleinen tapaus nilpotentin operaattorin, kuten operaattorin $(f - \lambda \mathrm{id})|_{G(\lambda,f)} \colon G(\lambda, f) \to G(\lambda, f)$, tapauksessa. Yleisen operaaattorin tapaus käsitellään tämän jälkeen.

Lause 18.4.8. *Olkoon V äärellisulotteinen vektoriavaruus ja $f \colon V \to V$ nilpotentti operaattori. Tällöin on olemassa sellaiset luvut $m \in \mathbb{N}$ ja $k \in \mathbb{N}$ sekä sellaiset luvut $m_1, \ldots, m_k \in \mathbb{N}$ että sellaiset vektorit $v_1, \ldots, v_k$, että jono*

$$(f^{m_1}(v_1), \cdots, f(v_1), v_1, f^{m_2}(v_2), \ldots, f(v_2), v_2, \ldots, f^{m_k}(v_k), \ldots, f(v_k), v_k, v_{k+1}, \ldots, v_m)$$

on avaruuden V kanta, jolle pätee

$$f^{m_1+1}(v_1) = \cdots = f^{m_k+1}(v_k) = 0$$

ja

$$f(v_{k+1}) = \cdots = f(v_m) = 0.$$

Erityisesti operaattorin f esitysmatriisi tässä kannassa on ositettu diagonaalimatriisi

$$N = \begin{bmatrix} N_{m_1} & & & \\ & \ddots & & \\ & & N_{m_k} & \\ & & & 0 \end{bmatrix} \in \mathbb{C}^{n\times n},$$

missä $n = \dim V$.

Todistus. Todistus on jälleen induktio avaruuden V dimension suhteen. Väite pätee selvästä tapauksessa $\dim V = 1$, sillä tällöin f on nilpotenttiuiden vuoksi nollakuvaus.

Oletetaan nyt, että väite $n \in \mathbb{N}$ on sellainen luku, että väite pätee kaikilla avaruuksilla V ja nilpotenteilla operaattoreilla $f \colon V \to V$, joilla $\dim V < n$. Osoitetaan, että väitä pätee luvulla n.

Olkoon V n-ulotteinen vektoriavaruus ja $f \colon V \to V$ nilpotentti operaattori. Koska on olemassa sellainen $k \in \mathbb{N}$, että $f^k = 0$, niin $\ker f \neq 0$. Koska $\dim \ker f > 0$, niin $\dim \operatorname{im} f = \dim V - \dim \ker f < \dim V$. Näin ollen $\operatorname{im} f \subset V$ on aito aliavaruus. Koska operaattorin kuva on aina invariantti aliavaruus, niin $f_1 = f|_{\operatorname{im} f} \colon \operatorname{im} f \to \operatorname{im} f$ on hyvin määritelty opeaattori. Koska $f_1^k(v) = f^k(v) = 0$ jokaisella $v \in \operatorname{im} f$, niin f_1 on nilpotentti operaattori. Soveltamalla induktio-oletusta avaruuteen $\operatorname{im} f$ ja operaattoriin f_1 saadaan, että on olemassa sellainen luku $k \in \mathbb{N}$, sellaiset luvut $\ell_1, \ldots, \ell_k \in \mathbb{N}$ ja sellaiset vektorit $w_1, \ldots, w_k \in \operatorname{im} f$, luvut $\ell_1, \ldots, \ell_k$, että jono

$$(f_1^{\ell_1}(w_1), \ldots, f_1(w_1), w_1, \ldots, f_1^{\ell_k}(w_k), \ldots, f_1(w_k), w_k) \tag{18.1}$$

on avaruuden $\operatorname{im} f$ kanta ja että

$$f_1^{\ell_1+1}(w_1) = \cdots = f_1^{\ell_k+1}(w_k) = 0.$$

Koska vektorit $w_1, \ldots, w_k$ kuuluvat operaattorin f kuvaan, niin jokaisella $j \in \{1, \ldots k\}$ on olemassa sellainen vektori $v_j \in V$, että $f(v_j) = w_j$.

Osoitetaan, että jono

$$(f^{\ell_1+1}(v_1), \ldots, f^2(v_1), f(v_1), v_1, \ldots, f^{\ell_k+1}(v_k), \ldots, f^2(v_k), f(v_k), v_k) \tag{18.2}$$

eli jono

$$(f_1^{\ell_1}(w_1), \ldots, f_1(w_1), w_1, v_1, \ldots, f_1^{\ell_k}(w_k), \ldots, f_1(w_k), w_k, v_k)$$

on avaruuden V kanta. Tällöin voidaan valita $m_j = \ell_j + 1$, jolloin saadaan $f^{m_j+1}(v_j) = f^{m_j}(w_j) = f_1^{\ell_j+1}(w_j) = 0$.

Osoitetaan ensin, että jonon vektorit ovat lineaarisesti riippumattomia. Olkoot a_{ji} sellaiset luvut, missä $j \in \{0, \ell_1 + 1\}$ jokaisella $i \in \{1, \ldots, k\}$, että

$$0 = a_{(\ell_1+1)1} f^{\ell_1+1}(v_1) + \cdots + a_{01} v_1 + \cdots + a_{\ell_k+1,k} f^{\ell_k+1}(v_k) + \cdots + a_{0k} v_k.$$

Tällöin

$$\begin{aligned}
0 &= f\Big(a_{(\ell_1+1)1} f^{\ell_1+1}(v_1) + \cdots + a_{01} v_1 + \cdots + a_{(\ell_k+1)k} f^{\ell_k+1}(v_k) + \cdots + a_{0k} v_k\Big) \\
&= a_{(\ell_1+1)1} f(f^{\ell_1+1}(v_1)) + \cdots + a_{01} f(v_1) + \cdots + a_{(\ell_k+1)k} f(f^{\ell_k+1}(v_k)) + \cdots + a_{0k} f(v_k) \\
&= a_{(\ell_1+1)1} f^{\ell_1+1}(f(v_1)) + \cdots + a_{01} f(v_1) + \cdots + a_{(\ell_k+1)k} f^{\ell_k+1}(f(v_k)) + \cdots + a_{0k} f(v_k) \\
&= a_{(\ell_1+1)1} f^{\ell_1+1}(w_1) + \cdots + a_{01} w_1 + \cdots + a_{(\ell_k+1)k} f^{\ell_k+1}(w_k) + \cdots + a_{0k} w_k \\
&= a_{\ell_1 1} f^{\ell_1}(w_1) + \cdots + a_{01} w_1 + \cdots + a_{\ell_k k} f^{\ell_k}(w_k) + \cdots + a_{0k} w_k.
\end{aligned}$$

Koska jono (18.1) on avaruuden $\operatorname{im} f$ kanta, niin $a_{ji} = 0$ kaikilla $j \in \{0, \ldots, \ell_i\}$ ja $i \in \{1, \ldots, k\}$. Näin ollen

$$\begin{aligned}
0 &= a_{(\ell_1+1)1} f^{\ell_1+1}(v_1) + \cdots + a_{01} v_1 + \cdots + a_{\ell_k+1,k} f^{\ell_k+1}(v_k) + \cdots + a_{0k} v_k \\
&= a_{(\ell_1+1)1} f^{\ell_1+1}(v_1) + \cdots + a_{(\ell_k+1)k} f^{\ell_k+1}(v_k) \\
&= a_{(\ell_1+1)1} f^{\ell_1}(w_1) + \cdots + a_{(\ell_k+1)k} f^{\ell_k}(w_k).
\end{aligned}$$

Käyttämällä uudelleen tietoa, että jono (18.1) on avaruuden $\operatorname{im} f$ kanta, saadaan

$$a_{(\ell_1+1)1} = \cdots = a_{(\ell_k+1)k} = 0.$$

Jono (18.2) on siis vapaa.

Olkoon $W \subset V$ jonon (18.2) virittämä aliavaruus. Jos $W = V$, niin (18.2) on haluttu kanta. Tarkastellaan nyt vielä tapausta, että $W \neq V$. Täydennetään jono (18.2) avaruuden V kannaksi

$$(f^{\ell_1+1}(v_1), \dots, f^2(v_1), f(v_1), v_1, \dots, f^{\ell_k+1}(v_k), \dots, f^2(v_k), f(v_k), v_k, w_{k+1}, \dots, w_m). \tag{18.3}$$

Jäljellä on osoittaa, että vektorit $w_{k+1}, \dots, w_m$ voidaan korvata operaattorin f ytimeen $\ker f$ kuuluvilla vektoreilla $v_{k+1}, \dots, v_m$.

Olkoon $p \in \{k+1, \dots, m\}$. Koska $f(w_p) \in \operatorname{im} f$, niin on olemassa sellaiset luvut $a_{ji} \in \mathbb{C}$, missä $j \in \{1, \dots, \ell_i\}$ ja $i \in \{1, \dots, k\}$, että

$$f(w_p) = a_{\ell_1 1} f^{\ell_1}(w_1) + \cdots + a_{01} w_1 + \cdots + a_{\ell_k k} f^{\ell_k}(w_k) + \cdots + a_{0k} w_k.$$

Olkoon nyt

$$u_p = a_{\ell_1 1} f^{\ell_1}(v_1) + \cdots + a_{01} v_1 + \cdots + a_{\ell_k k} f^{\ell_k}(v_k) + \cdots + a_{0k} v_k \in W.$$

Tällöin $f(w_p) = f(u_p)$. Näin ollen $v_p = w_p - u_p \in \ker f$.

Koska $u_p \in W$ jokaisella $p \in \{k+1, \dots, m\}$, niin

$$W + \operatorname{Sp}(w_{k+1}, \dots, w_m) = W + \operatorname{Sp}(v_{k+1} + u_{k+1}, \dots, v_m + u_m) = W + \operatorname{Sp}(v_{k+1}, \dots, v_m).$$

Näin ollen

$$V = W + \operatorname{Sp}(v_{k+1}, \dots, v_m)$$

ja jono

$$(f^{\ell_1+1}(v_1), \dots, f^2(v_1), f(v_1), v_1, \dots, f^{\ell_k+1}(v_k), \dots, f^2(v_k), f(v_k), v_k, v_{k+1}, \dots, v_m)$$

virittää avaruuden V. Koska tässä jonossa on yhtä monta alkiota kuin avaruuden V kannassa (18.3), niin se on avaruuden V kanta. Tämä kanta toteuttaa vaaditut ehdot. Tämä päättää induktiotodistuksen. □

Lauseen 18.4.8 todistuksesta voidaan poimia seuraava algoritmi Jordanin normaalimuotoon liittyvälle kannan valinnalle.

Huomautus 18.4.9 (Yleistetyn ominaisavaruuden kannan valinta). *Olkoon $A \in \mathbb{C}^{n \times n}$ neliömatriisi ja $\lambda \in \mathbb{C}$ sen ominaisarvo. Sanotaan, että vektorin $v \in G(\lambda, A)$ nilpotenssiaste on sellainen luku $m \in \mathbb{N}$, että $(A - \lambda I)^m v \neq 0$ ja $(A - \lambda I)^{m+1} v = 0$. Määritellään myös $A_\lambda = A - \lambda I$.*

Etsitään aliavaruuden $G(\lambda, A)$ kanta seuraavasti:

1. *Etsi pienin sellainen $d \in \mathbb{N}$, että $\operatorname{Null}(A - \lambda I)^{d+1} = \operatorname{Null}(A - \lambda I)^d$.*

2. *Etsi kanta* $(u_1, \ldots, u_m)$ *avaruudelle* $\mathrm{Null}(A - \lambda I)^d$.

3. *Järjestä kanta-alkiot* $(u_1, \ldots, u_m)$ *sellaiseen järjestykseen, että vektorilla* u_1 *on suurin nilpotenssiaste vektoreista* $u_1, \ldots, u_m$.

4. *Lisää kantaan alkiot* $(A_\lambda^{d-1} u_1, \ldots, A_\lambda u_1, u_1)$.

5. *Poista jonosta* $(u_2, \ldots, u_m)$ *ne alkiot, jotka kuuluvat jonon* $(A_\lambda^{d-1} u_1, \ldots, A_\lambda u_1, u_1)$ *virittämään aliavaruuteen* $\mathrm{Sp}(A_\lambda^{d-1} u_1, \ldots, A_\lambda u_1, u_1)$.

6. *Mikäli jonoon* $(u_2, \ldots, u_m)$ *jää alkiota, toista prosessi jäljelle jääneille alkioille alkaen vaiheesta 3.*

Jordanin normaalimuoto voidaan nyt kirjata seuraavassa muodossa.

Lause 18.4.10. *Jordanin normaalimuoto operaattoreille Olkoon* V *äärellisulotteinen kompleksinen vektoriavaruus ja* $f\colon V \to V$ *operaattori. Tällöin on olemassa sellainen avaruuden* V *kanta* $(w_1, \ldots, w_n)$*, että operaattorin* f *esitysmatriisi kannassa* $(w_1, \ldots, w_m)$ *on ositettu diagonaalimatriisi*

$$J = \begin{bmatrix} J_{\mu_1,n_1} & & & \\ & J_{\mu_2,n_2} & & \\ & & \ddots & \\ & & & J_{\mu_\ell,n_\ell} \end{bmatrix},$$

missä luvut $\mu_1, \ldots, \mu_\ell$ *ovat operaattorin* f *ominaisarvoja. Lisäksi operaattorin* f *ominaisarvo* λ_i *toistuu matriisin* J *diagonaalilla* $d_i = \dim G(\lambda_i, f)$ *kertaa.*

Todistus. Lauseen 18.3.5 perusteella

$$V = G(\lambda_1, V) \oplus \cdots \oplus G(\lambda_k, V)$$

missä $\lambda_1, \ldots, \lambda_k$ ovat operaattorin f erisuuret ominaisarvot.

Olkoon $i \in \{1, \ldots, k\}$ ja $d_i = \dim G(\lambda_i, f)$. Lauseen 18.4.4 perusteella operaattori $(f - \lambda_i \mathrm{id})\colon G(\lambda_i, f) \to G(\lambda_i, f)$ on nilpotentti. Näin ollen lauseen 18.4.8 perusteella on olemassa sellainen aliavaruuden $G(\lambda_i, f)$ kanta $(w_{1i}, \ldots, w_{d_i i})$, että operaattorin $f - \lambda_i \mathrm{id}$ esitysmatriisi tässä kannassa on

$$N_i = \begin{bmatrix} N_{m_1} & & & \\ & \ddots & & \\ & & N_{m_{k_i}} & \\ & & & 0 \end{bmatrix} \in \mathbb{C}^{d_i \times d_i}.$$

Koska operaattorin $\lambda_i \mathrm{id}$ esitysmatriisi tässäkin kannassa on $\lambda_i I$, niin saadaan, että operaattorin

$$f = (f - \lambda_i \mathrm{id}) + \lambda_i \mathrm{id}$$

esitysmatriisi tässä aliavaruuden $G(\lambda_i, f)$ kanssa on $N_i + \lambda_i I$ eli

$$J_i = \begin{bmatrix} J_{\lambda_i, m_1} & & & \\ & \ddots & & \\ & & J_{\lambda_i, m_{k_i}} & \\ & & & \lambda_i I \end{bmatrix} \in \mathbb{C}^{d_i \times d_i}.$$

Olkoon

$$(w_{11}, \dots, w_{d_1 1}, w_{12}, \dots, w_{d_k k})$$

avaruuden V kanta. Operaattorin f esitysmatriisi tässä kanssa on

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}$$

Koska ositettujen diagonaalimatriisien ositettu diagonaalimatriisi on ositettu diagonaalimatriisi, niin väite seuraa. □

Huomautus 18.4.11. *On tärkeää huomata, että ositetun diagonaalimatriisin J lohkot eivät vastaa operaattorin f yleistettyjä ominaisavaruuksia vaan yleistettyjen ominaisavaruuksien aliavaruuksia.*

Soveltamalla lausetta 18.4.10 operaattoriin $\varphi_A \colon \mathbb{C}^{n\times 1} \to \mathbb{C}^{n\times 1}$ saadaan todistus luvun alussa kirjatulle lauseelle.

Jordanin normaalimuoto matriiseille (lause 18.0.4). *Olkoon $A \in \mathbb{C}^{n\times n}$. Tällöin on olemassa sellainen kääntyvä matriisi $P \in \mathbb{C}^{n\times n}$, että*

$$A = PJP^{-1},$$

missä J on Jordan lohkoista koostuva ositettu diagonaalimatriisi

$$J = \begin{bmatrix} J_{\mu_1} & & & \\ & J_{\mu_2} & & \\ & & \ddots & \\ & & & J_{\mu_\ell} \end{bmatrix}$$

missä $\mu_1, \dots, \mu_\ell$ ovat matriisin A ominaisarvoja.

Kootaan vielä lopuksi yhteen ratkaisumenetelmä matriisin Jordanin normaalimuodolle.

Huomautus 18.4.12. *Matriisin $A \in \mathbb{C}^{n\times n}$ Jordanin normalimuoto voidaan selvittää seuraavasti:*

1. *Etsitään matriisin A ominaisarvot $\lambda_1, \dots, \lambda_k$.*

2. *Etsitään jokaiselle yleistetylle ominaisavaruudelle $G(\lambda_i, A)$ lauseen 18.4.8 mukainen kanta $(v_{i,1}, \ldots, v_{i,d_i})$, missä $d_i = \dim G(\lambda_i, A)$; katso huomautus 18.4.9.*

3. *Yhdistetään saadut kannat avaruuden $\mathbb{C}^{n\times 1}$ kannaksi ja muodostetaan kantavektoreista kääntyvä matriisi $P \in \mathbb{C}^{n\times n}$.*

4. *Matriisi J voidaan nyt laskea kaavasta $J = P^{-1}AP$ tai muodostaa lauseen 18.4.8 antamista lohkoista J_λ.*

18.5 Sovellus determinanttiin ja jälkeen

Yksi Jordanin normaalmuotoon liittyvistä tuloksesta on se, että matriisin determinantti on ominaisarvojen tulo ja että matriisin jälki on ominaisarvojen summa. Näihin tuloksiin ei itseasiassa tarvita Jordanin normaalimoudon todistamista vaan ne seuravat samoista aiemmin tässä luvussa todistetuista tuloksista samalla todistuksella.

Lause 18.5.1. *Olkoon $A \in \mathbb{C}^{n\times n}$ ja $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$. Tällöin*

$$\det A = \lambda_1^{d_1} \cdots \lambda_k^{d_k}$$

ja

$$\mathrm{tr}A = d_1\lambda_1 + \cdots + d_k\lambda_k,$$

missä $d_i = \dim G(\lambda_i, A)$ jokaisella $i \in \{1, \ldots, k\}$.

Todistus. Olkoon $J \in \mathbb{C}^{n\times n}$ matriisin A Jordanin normaalimuoto ja $P \in \mathbb{C}^{n\times n}$ sellainen kääntyvä matriisi, että

$$A = PJP^{-1}.$$

Koska Jordan lohkon J_{μ_i,n_i} determinantti on $\mu_i^{n_i}$ ja jälki on $n_i\mu_i$, niin

$$\det A = \det J = \mu_1^{n_1} \cdots \mu_\ell^{n_\ell}$$

ja

$$\mathrm{tr}A = \mathrm{tr}J = n_1\mu_1 + \cdots + n_\ell\mu_\ell.$$

Koska jokainen matriisin A ominaisarvo λ_i toistuu matriisin J diagonaalilla täsmälleen d_i kertaa, niin

$$\mu_1^{n_1} \cdots \mu_\ell^{n_\ell} = \lambda_1^{d_1} \cdots \lambda_k^{d_k}$$

ja

$$n_1\mu_1 + \cdots + n_\ell\mu_\ell = d_1\lambda_1 + \cdots + d_k\lambda_k.$$

□

Huomautus 18.5.2. *Edellä esitettyä todistusta voi pitää hieman epätyydyttävänä. Toisen todistuksen tällä tulokselle saa tarkastelemalla yleinen operaattorin $\varphi\colon V \to V$ yleistettyjä ominaisavaruuksia $G(\lambda_i, f)$ ja operaattorin $\varphi|_{G(\lambda_i,f)}\colon G(\lambda_i, f) \to G(\lambda_i, f)$ esitysmatriisia kuten Jordanin normaalimuodon todistuksessa.*

Huomautus 18.5.3. *Tämä lause esitetään usein muodossa, että matriisin* $A \in \mathbb{C}^{n\times n}$ *determinantti on sen ominaisarvojen tulo algebrallisen kertaluku huomioon ottaen. Ominaisarvon* $\lambda \in \mathbb{C}$ *algebrallisella kertaluvulla tarkoitetaan tässä yhteydessä polynomin* $t \mapsto \det(A - tI)$ *nollakohdan astetta kohdassa* $t = \lambda$.

Polynomia $p_A \colon \mathbb{C} \to \mathbb{C}$ *kutsutaan matriisin* A karakteristiseksi polynomiksi. *Edellä lausuttu tulos siis sanoo, että polynomin* p_A *nollakohdan aste kohdassa* λ *on yleistetyn ominaisavaruuden* $G(\lambda, A)$ *dimensio. Tätä hienoa algebrallista faktaa ei käsitellä tässä yhteydessä tätä enempää. Kiinnostunut lukija ohjataan vaikkapa Axlerin kirjan pariin.*

Osa IV

Liitteet

Liite A

Liite: Supistetun porrasmuodon yksikäsitteisyys

Tässä liitteessä osoitetaan, että matriisin supistettu porrasmuoto on yksikäsitteinen.

Lause A.0.1. *Matriisin $A \in \mathbb{R}^{m \times n}$ supistettu porrasmuoto B on yksikäsitteinen, eli jos B ja B' ovat matriisin A supistettuja porrasmuotoja, niin $B = B'$.*

Huomautus A.0.2. *On mielenkiintoinen yksityiskohta, että kääntyvämatriisi $P \in \mathbb{R}^{m \times m}$, joka muodostuu tulona rivioperaatioiden määräämistä alkeismatriisien ei puolestaan ole yksikäsitteinen. Syy tähän on se, että matriisissa B voi olla nollarivejä. Yksi esimerkki tälläisestä tapauksesta on matriisi*

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix},$$

jonka supistettu porrasmuoto on

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},$$

mutta jolle pätee sekä

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

että

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

Lauseen A.0.1 todistus. Olkoon $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \in \mathbb{R}^{m \times n}$ ja olkoot $B = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}$ ja $B' = \begin{bmatrix} b'_1 & \cdots & b'_n \end{bmatrix}$ matriisin A supistettuja porrasmuotoja.

Koska B ja B' saadaan matriisista A rivioperaatioilla, niin on olemassa sellaiset kääntyvät $m \times m$-matriisit P ja P', että $A = PB$ ja $A = P'B'$. Olkoon nyt $Q = (P')^{-1}P$. Tällöin $B' = QB$.

Olkoot nyt $1 \leq i_1 < \cdots < i_k \leq n$ matriisin B sidottuja muuttujia vastaavien sarakkeiden indeksit. Koska sarakkeet $(b_{i_1}, \ldots, b_{i_k})$ muodostavat sarakeavaruuden $\mathrm{Col}(B)$ kannan, niin ne ovat lineaarisesti riippumattomia. Koska matriisi Q on kääntyvä, niin myös vastaavat matriisin B' sarakkeet $b'_{i_1}, \ldots, b'_{i_k}$ ovat lineaarisesti riippumattomia. Näin ollen sarakkeita $b_{i_1}, \ldots, b_{i_k}$ vastaavat muuttujat ovat vapaita matriisin B' yhtälöryhmässä $\big[B' \bigm| 0\big]$. Olkoon nyt $i \notin \{i_1, \ldots, i_k\}$. Tällöin sarake b_i on sarakkeiden $b_{i_1}, \ldots, b_{i_k}$ lineaarikombinaatio. Näin ollen myös sarake b'_i on vastaavien sarakkeiden $b'_{i_1}, \ldots, b'_{i_k}$ lineaarikombinaatio. Näin ollen saraketta i vastaava muuttuja on vapaa muuttuja molemmissa matriiseissa B ja B'. Matriiseilla B ja B' on siis samat sidotut ja vapaat muuttujat. Eritysesti siis $b'_{i_\ell} = b_{i_\ell} = e_\ell$ kaikilla $1 \leq \ell \leq k$.

Olkoon nyt $i \notin \{i_1, \ldots, i_k\}$. Tällöin b_i on sarakkeiden $b_{i_1}, \ldots, b_{i_k}$ lineaarikombinaatio, eli on olemassa sellaiset luvut $a_1, \ldots, a_k \in \mathbb{R}$ että

$$b_i = a_1 b_{i_1} + \cdots + a_k b_{i_k}.$$

Tällöin

$$b'_i = Q b_i = a_1 Q b_{i_1} + \cdots + a_k Q b_{i_k} = a_1 b'_{i_1} + \cdots + a_k b'_{i_k} = a_1 e_1 + \cdots + a_k e_k = b_i.$$

Matriisit B ja B' ovat siis sama matriisi. $\square$

Liite B

Liite: Algoritmi nolla-avaruuden kannan löytämiselle

Aloitetaan nolla-avaruuden kannan etsiminen havainnolla, että matriisilla ja sen supistetulla porrasmuodolla on sama nolla-avaruus.

Lause B.0.1. *Olkoot $A \in \mathbb{R}^{m\times n}$ matriisi ja $B \in \mathbb{R}^{m\times n}$ matriisin A supistettu porrasmuoto. Tällöin* $\mathrm{Null}(A) = \mathrm{Null}(B)$*.*

Todistus. Koska B on matriisin A supistettu porrasmuoto, niin yhtälöryhmä $\left[A \mid 0\right]$ saadaan muotoon $\left[B \mid 0\right]$ rivioperaatiolla. Näin ollen $A = PB$, missä $P \in \mathbb{R}^{m\times m}$ on kääntyvä matriisi.

Osoitetaan ensin, että $\mathrm{Null}(A) \subset \mathrm{Null}(B)$. Olkoon $x \in \mathrm{Null}(A)$. Tällöin $P(Bx) = Ax = 0$. Koska P on kääntyvä, niin $Bx = 0$. Näin ollen $x \in \mathrm{Null}(B)$.

Osoitetaan nyt, että $\mathrm{Null}(B) \subset \mathrm{Null}(A)$. Olkoon $x \in \mathrm{Null}(B)$. Tällöin $Bx = 0$. Näin ollen $Ax = P(Bx) = P0 = 0$, eli $x \in \mathrm{Null}(A)$. $\square$

Lauseen B.0.1 merkitys on siinä, että etsittäessä kantaa matriisin A nolla-avaruudelle, voidaan suoraan siirtyä tarkastelemaan vastaavaa supistetussa porrasmuodossa olevaa matriiisia.

Tehdään tämä ensin konkreettisen esimerkin tapauksessa. Tämä esimerkki on muunnelma esimerkistä 3.9.1.

Esimerkki B.0.2. *Olkoon*

$$B = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \end{bmatrix} = \begin{bmatrix} 1 & -4 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

edellisestä luvusta tuttu matriisi.

Nolla-avaruus $\mathrm{Null}(B)$ *koostuu niistä vektoreista*

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

joilla pätee $Bx = 0$. *Kun tämä yhtälö kirjoitetaan yhtälöryhmän muodossa saadaan yhtälöryhmä*

$$\begin{cases} x_1 & -4x_2 & & +2x_4 & = & 0 \\ & & x_3 & -x_4 & = & 0 \\ & & & 0 & = & 0 \end{cases} \tag{B.1}$$

Tiedetään, että muuttujat x_1 *ja* x_3 *ovat sidottuja ja että muuttujat* x_2 *ja* x_4 *ovat vapaita.*

Valitsemalla $x_2 = 1$ *ja* $x_4 = 0$ *saadaan ratkaisu*

$$w_2 = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

ja vastaavasti valitsemalla $x_2 = 0$ *ja* $x_4 = 1$ *saadaan ratkaisu*

$$w_4 = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Koska $\mathrm{Null}(B)$ *on aliavaruus, niin lisäksi tiedetään, että kaikki vektoreiden* w_2 *ja* w_4 *lineaarikombinaatiot ovat yhtälön* $Bx = 0$ *ratkaisuja, eli* $\mathrm{Sp}(w_1, w_2) \subset \mathrm{Null}(B)$.

Vektorit (w_2, w_4) *ovat lineaarisesti riippumattomia, koska täsmälleen toisella vektoreista* w_2 *ja* w_4 *on nollasta poikkeava kerroin riveillä* 2 *ja* 4.

Osoitetaan vielä, että $\mathrm{Null}(B) = \mathrm{Sp}(w_1, w_2)$. *Olkoon*

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \in \mathrm{Null}(B).$$

Tällöin

$$y = y_2 \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y_4 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} y_1 - 4y_2 + 2y_4 \\ 0 \\ y_3 - y_4 \\ 0 \end{bmatrix}$$

Halutaan osoittaa, että viimeinen summattava on nollavektori.

Koska $y, w_1, w_2 \in \mathrm{Null}(B)$ *ja* $\mathrm{Null}(B)$ *on aliavaruus, niin*

$$\begin{bmatrix} y_1 - 4y_2 + 2y_4 \\ 0 \\ y_3 - y_4 \\ 0 \end{bmatrix} \in \mathrm{Null}(B),$$

eli yhtälöryhmän (B.1) *ratkaisu. Näin ollen yhtälöryhmän ensimmäisen rivin perusteella*

$$(y_1 - 4y_2 + 2y_4) - 4 \cdot 0 + 2 \cdot 0 = 0$$

eli $y_1 - 4y_2 + 2y_4 = 0$. *Vastaavasti yhtälöryhmän* (B.1) *toisen rivin perusteella* $y_3 - y_4 = 0$. *Näin ollen*

$$\begin{bmatrix} y_1 - 4y_2 + 2y_4 \\ 0 \\ y_3 - y_4 \\ 0 \end{bmatrix} = 0.$$

Näin on päätelty, että $y \in \mathrm{Sp}(w_2, w_4)$.

Koska (w_2, w_4) *on aliavaruuden* $\mathrm{Null}(B)$ *vapaa virittäjäjono, niin se on kanta.*

Analysoidaan tätä konkreettista esimerkkiä. Esimerkissä lähdettiin tarkastelemaan yhtälöryhmän vapaita muuttujia. Yhtälöryhmässä on yhteensä n muuttujaa ja jokaista vapaata muuttujaa x_ℓ kohti valittiin sellainen avaruuden $\mathbb{R}^{n\times 1}$ vektori w_ℓ, että sen kerroin rivillä ℓ on 1, jonka muut vapaita muuttujia vastaavat kertoimet ovat nollia. Tällöin yhtälöryhmän avulla voitiin ratkaista vektorin w_ℓ ne kertoimet, jotka olivat sidottuja muuttujia vastaavilla riveillä.

Koska vektorissa w_ℓ on täsmälleen yksi nollasta poikkeava vapaata muuttujaa vastaava kerroin, niin havaittiin, että näin saadut vektori ovat lineaarisesti riippumattomia. Tämän jälkeen osoitetiin, että kaikki nolla-avaruuden vektorit saadaan näiden vektoreiden lineaarikombinaationa. Näin pääteltiin, että nolla-avaruudella on sellainen kanta, jossa on yhtä monta vektoria, kuin yhtälöryhmässä on vapaita muutujia. Tämä on yleinen tulos, joka voidaan kirjoittaa hieman tarkemmin seuraavasti.

Lause B.0.3. *Olkoon* $B \in \mathbb{R}^{m\times n}$ *supistetussa porrasmuodossa oleva matriisi, olkoot* $1 \le j_1 < \cdots < j_k \le n$ *yhtälöryhmän* $\left[B \mid 0\right]$ *vapaiden muutujien indeksit sekä olkoot* $1 \le i_1 < \cdots < i_{n-k} \le n$ *yhtälöryhmän* $\left[B \mid 0\right]$ *sidottujen muutujien indeksit. Tällöin on olemassa aliavaruuden* $\mathrm{Null}(B)$ *sellainen kanta* $(w_{j_1}, \ldots, w_{j_k})$, *että jokaisella* $\ell \in \{1, \ldots, k\}$ *pätee* $w_{j_\ell} = e_{j_\ell} + u_\ell$, *missä* $u_\ell \in \mathrm{Sp}(e_{i_1}, \ldots, e_{i_{n-k}})$.

Ennen lauseen todistusta on hyvä tarkastella, miten tulosta voi soveltaa.

Esimerkki B.0.4. *Olkoon* $A \in \mathbb{R}^{m\times n}$. *Selvitetään ensin matriisin* A *supistettu porrasmuoto* $B = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}$. *Supistusta porrasmuodosta voidaan selvittää vapaiden muuttujen indeksit* $1 \le j_1 < \cdots < j_k \le n$ *ja sidottujen muuttujien indeksit* $1 \le i_1 < \cdots < i_{n-k} \le n$.

Olkoon nyt $\ell \in \{1, \ldots, k\}$. *Lauseen nojalla on olemassa sellainen vektori* $u_\ell \in \mathrm{Sp}(e_{i_1}, \ldots, e_{i_{n-k}})$, *että*

$$B(e_{j_\ell} + u_\ell) = 0$$

eli että

$$Bu_\ell = -Be_{j_\ell} = -b_{j_\ell}.$$

Tämä yhtälöryhmä voidaan ratkaista ja ratkaisuista valita sellainen, jolla on nollasta poikkeava kerroin ainoastaan riveillä $1 \le i_1 < \cdots < i_{n-k} \le n$, *eli joka kuuluu avaruuteen* $\mathrm{Sp}(e_{i_1}, \ldots, e_{i_{n-k}})$. *Huomaa, että tällöin on valittu sellainen ratkaisu, jonka kaikki vapaita muuttujia vastaavat kertoimet ovat nollia.*

Koska matriisi B on supistetussa porrasmuodossa, niin itseasiassa käytännön tilanteessa vektorit voidaan oleellisesti lukea matriisista B. Tarkastellaan vielä konkreettista tilannetta.

Esimerkki B.0.5. *Tarkastellaan supistetussa porrasmuodossa olevaa matriisia*

$$B = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 4}.$$

Tällöin sidottujen muuttujien indeksit ovat 1 *ja* 3 *sekä vapaiden muuttujien indeksit* 2 *ja* 4. *Näin ollen etsitään sellaisia vektoreita* $u_1, u_2 \in \mathrm{Sp}(e_1, e_3) \subset \mathbb{R}^{4\times 1}$, *joille* $w_2 = e_2 + u_1$ *ja* $w_4 = e_4 + u_2$ *muodostavat avaruuden* $\mathrm{Null}(B)$ *kannan. Vektorit* u_1 *ja* u_2 *ovat siis muotoa*

$$u_1 = \begin{bmatrix} x_1 \\ 0 \\ x_3 \\ 0 \end{bmatrix} \text{ ja } u_2 = \begin{bmatrix} y_1 \\ 0 \\ y_3 \\ 0 \end{bmatrix}$$

ja niille pätee yhtälöt $Bu_1 = -b_2$ *ja* $Bu_2 = -b_4$. *Halutaan siis ratkaista matriisiyhtälö*

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ 0 & 0 \\ x_3 & y_3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ 0 & -1 \end{bmatrix}.$$

Koska tämä yhtälö vastaa yhtälöä

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_3 & y_3 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ 0 & -1 \end{bmatrix},$$

niin

$$\begin{bmatrix} x_1 & y_1 \\ x_3 & y_3 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ 0 & -1 \end{bmatrix}.$$

Halutut vektorit ovat siis

$$u_1 = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ ja } u_2 = \begin{bmatrix} -3 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

Vaikka lauseen B.0.3 todistus on huomattavan tekninen, se luonnollisesti heijastelee näitä havaintoja. Todistuksen voi ensimmäisellä lukukerralla halutessaan sivuuttaa.

Lauseen B.0.3 todistus. Olkoon

$$B = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}.$$

Ennen vektoreiden $w_{j_1}, \ldots, w_{j_k}$ valitsemista, tehdään kaksi havaintoa. Koska matriisi B on supistetussa porrasmuodossa, niin sidottuja muuttujia vastaaville indekseille i_ℓ pätee

$$Be_{i_\ell} = e_\ell.$$

Toisaalta vapaata muuttujaa vastaavalle sarekkeelle b_{j_ℓ} pätee

$$Be_{j_\ell} = b_{j_\ell} = b_{1j_\ell}e_1 + \cdots + b_{(n-k)j_\ell}e_{n-k}.$$

Näin ollen jokaista vapaata muuttujaa vastaavalla indeksillä j_ℓ pätee

$$\begin{aligned} Be_{j_\ell} &= b_{1j_\ell}e_1 + \cdots + b_{(n-k)j_\ell}e_{n-k} \\ &= b_{1j_\ell}Be_{i_1} + \cdots + b_{(n-k)j_\ell}Be_{i_{n-k}} \\ &= B\left(b_{1j_\ell}e_{i_1} + \cdots + b_{(n-k)j_\ell}e_{i_{n-k}}\right). \end{aligned}$$

Huomaa, että edellisissä laskuissa $e_1, \ldots, e_{n-k} \in \mathbb{R}^{m\times 1}$ ja $e_{i_1}, \ldots, e_{i_{n-k}} \in \mathbb{R}^{n\times 1}$.

Määritellään jokaisella $\ell \in \{1, \ldots, k\}$ vektorit

$$u_\ell = -\left(b_{1j_\ell}e_{i_1} + \cdots + b_{(n-k)j_\ell}e_{i_{n-k}}\right) \in \mathrm{Sp}(e_{i_1}, \ldots, e_{i_{n-k}}).$$

ja

$$w_{j_\ell} = e_{j_\ell} + u_\ell.$$

Tällöin

$$Bw_{j_\ell} = Be_{j_\ell} + Bu_\ell = Be_{j_\ell} - Be_{j_\ell} = 0,$$

eli $w_{j_\ell} \in \mathrm{Null}(B)$.

Osoitetaan nyt, että $(w_{j_1}, \ldots, w_{j_k})$ on aliavaruuden $\mathrm{Null}(B)$ kanta. Osoitetaan ensin, että jono $(w_{j_1}, \ldots, w_{j_k})$ on vapaa. Olkoot $a_1, \ldots, a_k \in \mathbb{R}$ sellaisia lukuja, että

$$a_1w_{j_1} + \cdots + a_kw_{j_k} = 0.$$

Nyt termien uudelleen ryhmittelyn jälkeen saadaan

$$a_1e_{j_1} + \cdots + a_ke_{j_k} + (a_1u_1 + \cdots + a_ku_k) = 0.$$

Koska vektorit $u_1, \ldots, u_k$ kuuluvat aliavaruuteen $\mathrm{Sp}(e_{i_1}, \ldots, e_{i_{n-k}})$, niin saadaan, että $a_1u_1 + \cdots + a_ku_k \in \mathrm{Sp}(e_{i_1}, \ldots, e_{i_{n-k}})$. Näin ollen on olemassa sellaiset luvut $c_1, \ldots, c_{n-k} \in \mathbb{R}$, että

$$a_1u_1 + \cdots + a_ku_k = c_1e_{i_1} + \cdots + c_{n-k}e_{i_{n-k}}.$$

Tällöin saadaan yhtälö

$$a_1 e_{j_1} + \cdots + a_k e_{j_k} + c_1 e_{i_1} + \cdots + c_{n-k} e_{i_{n-k}} = 0.$$

Koska jonossa $(e_{j_1}, \ldots, e_{j_k}, e_{i_1}, \ldots, e_{i_{n-k}})$ on standardikannan $(e_1, \ldots, e_n)$ jäsenet uudelleen järjestettynä, niin jono $(e_{j_1}, \ldots, e_{j_k}, e_{i_1}, \ldots, e_{i_{n-k}})$ on kanta. Näin ollen $a_1 = \cdots = a_k = c_1 = \cdots = c_{n-k} = 0$. Erityisesti $a_1 = \cdots = a_k = 0$ ja jono $(w_{j_1}, \ldots, w_{j_k})$ on vapaa.

Osoitetaan nyt, että jono $(w_{j_1}, \ldots, w_{j_k})$ virittää aliavaruuden $\mathrm{Null}(B)$. Aloitetaan havainnolla. Olkoon

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathrm{Null}(B)$$

sellainen vektori, että $y_{j_\ell} = 0$ kaikilla $\ell \in \{1, \ldots, k\}$. Koska yhtälöryhmä B on supistetussa porrasmuodossa, niin tällöin myös sidottuja muuttujia vastaavat termit ovat nollia, eli $y_{i_1} = \cdots = y_{i_{n-k}} = 0$. Näin ollen $y = 0$.

Olkoon nyt

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathrm{Null}(B)$$

Tällöin

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x - (x_{j_1} w_{j_1} + \cdots + x_{j_\ell} w_{j_\ell}) \in \mathrm{Null}(B)$$

on vektori, jolle pätee $y_{j_\ell} = 0$ jokaisella $\ell \in \{1, \ldots, k\}$. Näin ollen $y = 0$, eli

$$x = x_{j_1} w_{j_1} + \cdots + x_{j_\ell} w_{j_\ell}.$$

Näin ollen jono $(w_{j_1}, \ldots, w_{j_k})$ virittää aliavaruuden $\mathrm{Null}(A)$. $\square$

Liite C

Liite: Determinanttien teorian lisäsivut

C.1 Permutaatiot ja transpositiot

Tässä luvussa täydennetään luvussa 5.3.2 käsiteltyä permutaatioiden teoriaa. Ensin osoitetaan permutaation merkin tulokaava (lause C.1.4) ja tämän jälkeen, että jokaisen permutaation voi esittää transpositioiden tulona. Palautetaan aluksi mieleen käsitteet ja merkinnät.

Määritelmä C.1.1. *Joukon* $\{1,\ldots,n\}$ *bijektiota* $\sigma\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ *kutsutaan* permutaatioksi*. Permutaatio* $\sigma\colon \{1,\ldots,n\} \to \{1,\ldots n\}$ *on* transpositio*, jos on olemassa sellaiset luvut* $i, j \in \{1,\ldots,n\}$*,* $i \neq j$*, että* $\sigma(i) = j$*,* $\sigma(j) = i$ *ja* $\sigma(k) = k$ *kaikilla* $k \in \{1,\ldots,n\} \setminus \{i,j\}$*.*

Joukon $\{1,\ldots,n\}$ permutaatioiden joukkoa

$$\mathrm{Sym}(n) = \{\sigma\colon \{1,\ldots,n\} \to \{1,\ldots,n\} \colon \sigma \text{ on permutaatio}\}$$

kutsutaan joukon $\{1,\ldots,n\}$ *symmetriaryhmäksi*, koska kahden permutaation yhdiste on permutaatio ja permutaation käänteiskuvaus on permutaatio. Näin ollen joukon $\{1,\ldots,n\}$ permutaatiot muodostavat ryhmän algebrallisessa mielessä.

Permutaation merkki on permutaatioon liittyvä luku 1 tai -1, joka kuvastaa samanaikaisesti kahta permutaation $\sigma\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ ominaisuutta:

- sitä vaihtaako σ lukujen $1 \leq i < j \leq n$ järjestyksen parillisen vai parittoman monta kertaa ja
- sitä tarvitaanko permutaation σ esittämiseen parillinen vai pariton määrä transpositioita.

Yleensä määritelmäksi valitaan jälkimmäinen ominaisuus. Luvussa 5 valittiin kuitenkin ensimmäinen määritelmä, koska se on sekä määritelmällisesti helpompi että teoreettisesti mielenkiintoisempi.

Määritelmä C.1.2. *Permutaation* $\sigma \in \mathrm{Sym}(n)$ *merkki* $\mathrm{sign}(\sigma) \in \{-1, +1\}$ *on*

$$\mathrm{sign}(\sigma) = (-1)^k,$$

missä $k \in \mathbb{N}$ *on niiden lukuparien* (i, j), *missä* $1 \leq i < j \leq n$, *lukumäärä, joille pätee* $\sigma(j) < \sigma(i)$.

Huomautus C.1.3. *Suoraan määritelmästä havaitaan, että transposition merkki on aina* -1 *ja että identtisen permutaation merkki on* 1.

Osoitetaan nyt, että permutaatioiden yhdisteen merkki on merkkien tulo.

Lause C.1.4. *Olkoot* $\sigma, \tau \in \mathrm{Sym}(n)$. *Tällöin* $\mathrm{sign}(\sigma \circ \tau) = \mathrm{sign}(\sigma)\mathrm{sign}(\tau)$.

Todistus. Ennen varsinaisen väitteen todistamista tehdään havainto. Olkoon $n \in \mathbb{N}$ ja $P \colon \mathbb{R}^n \to \mathbb{R}$ polynomi

$$P(x_1, \ldots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

kaikilla $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Olkoon $\sigma \in \mathrm{Sym}(n)$ permutaatio ja määritellään polynomi $P_\sigma \colon \mathbb{R}^n \to \mathbb{R}$ kaavalla

$$P_\sigma(x_1, \ldots, x_n) = P(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

kaikilla $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Tällöin

$$P_\sigma(x_1, \ldots, x_n) = \prod_{1 \leq i < j \leq n} (x_{\sigma(j)} - x_{\sigma(i)})$$

kaikilla $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Olkoot $1 \leq i < j \leq n$. Tällöin joko $\sigma(i) < \sigma(j)$ tai $\sigma(i) > \sigma(j)$. Jos $\sigma(i) < \sigma(j)$, niin monomi $(x_{\sigma(j)} - x_{\sigma(i)})$ on polynomin P tulontekijä. Jos puolestaan $\sigma(i) > \sigma(j)$, niin monomi $-(x_{\sigma(i)} - x_{\sigma(j)})$ on polynomin P tulontekijä. Havainnon perusteella

$$P_\sigma(x_1, \ldots, x_n) = \mathrm{sign}(\sigma) P(x_1, \ldots, x_n),$$

kaikilla $x_1, \ldots, x_n$, eli polynomeina

$$P_\sigma = \mathrm{sign}(\sigma) P.$$

Olkoot nyt $\sigma, \tau \in \mathrm{Sym}(n)$ permutaatioita. Koska

$$\begin{aligned}
\mathrm{sign}(\sigma \circ \tau) P(x_1, \ldots, x_n) &= P(x_{\sigma(\tau(1))}, \ldots, x_{\sigma(\tau(n))}) \\
&= P_\tau(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \\
&= \mathrm{sign}(\tau) P(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \\
&= \mathrm{sign}(\tau) P_\sigma(x_1, \ldots, x_n) \\
&= \mathrm{sign}(\tau)\mathrm{sign}(\sigma) P(x_1, \ldots, x_n)
\end{aligned}$$

kaikilla $(x_1, \ldots, x_n) \in \mathbb{R}^n$, niin

$$\mathrm{sign}(\sigma \circ \tau) = \mathrm{sign}(\sigma)\mathrm{sign}(\tau).$$

□

Edellisestä lauseesta seuraa suoraan permutaation merkin toinen tulkinta.

Korollaari C.1.5. *Olkoon $\sigma \in \mathrm{Sym}(n)$ permutaatio ja $\alpha_1, \ldots, \alpha_\ell$ sellaisia transpositioita, että $\sigma = \alpha_1 \circ \cdots \circ \alpha_\ell$. Tällöin*

$$\mathrm{sign}(\sigma) = (-1)^\ell.$$

Huomautus C.1.6. *Lause C.1.4 osoittaa, että permutaation merkki* $\mathrm{sign}(\cdot)$ *määrittelee ryhmähomomorfismin* $\mathrm{sign}\colon \mathrm{Sym}(n) \to \mathbb{F}_2$*, missä* $\mathbb{F}_2 = \{-1, 1\}$ *on kahden alkion multipilikatiivinen ryhmä. Näitä asioita käsitellään kurssilla* Algebralliset rakenteet I & II.

Huomautus C.1.7. *Lauseen C.1.4 todistus saattaa vaikuttaa täysin hihasta vedetyltä. (Se ei ole kirjoittajan keksimä!) Sillä on kuitenkin syvällinen tulkinta.*

Joukon $\mathrm{Sym}(n)$ *permutaatiot permutoivat luonnolisella tavalla avaruuden* $\mathbb{R}^n$ *koordinaatteja.*[1] *Näin ollen jokainen permutaatio* $\sigma \in \mathrm{Sym}(n)$ *itseasiassa määrittelee kuvauksen avaruuden* $\mathbb{R}^n$ *polynomien avaruudelta itseensä kaavalla* $Q \mapsto Q_\sigma$*, missä* Q *on avaruuden* $\mathbb{R}^n$ *polynomi ja* Q_σ *on määritelty vastaavalla tavalla kuin todistuksessa polynomi* P_σ*.*

Koska joukko $\mathrm{Sym}(n)$ *on kuvausten yhdistämisen suhteen ryhmä, niin edellä olevat kuvaukset* $Q \mapsto Q_\sigma$ *määrittelevät ryhmän* $\mathrm{Sym}(n)$ *toiminnan polynomien avaruuteen. Koska polynomille* P *pätee, että joko* $P_\sigma = P$ *tai* $P_\sigma = -P$*, niin paljastuu, että ryhmä* $\mathrm{Sym}(n)$ *toimii aliavaruuteen* $\mathrm{Sp}\{P\}$ *kuten ryhmä* $\mathbb{Z}_2$ *ja toiminta määräytyy permutaation merkistä. Aiheesta lisää maisteriopinnoissa tai hakusanalla* group action *internetistä.*

C.1.1 Permutaatiot transpositioiden yhdisteenä

Osoitetaan nyt, että jokainen permutaatio on transpositioiden yhdiste. Yhden alkion joukon $\{1\}$ tapaus jätetään lukijalle.

Lemma C.1.8. *Olkoon $n \geq 2$. Jokainen joukon $\{1, \ldots, n\}$ permutaatio $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ voidaan kirjoittaa yhdistettynä kuvauksena transpositioista.*

Todistus. Todistetaan väite induktiolla. Oletetaan ensin, että $n = 2$ ja että $\sigma\colon \{1, 2\} \to \{1, 2\}$ on permutaatio. Tällöin, joko $\sigma = \mathrm{id}$ tai σ on transpoosi $\tau\colon \{1, 2\} \to \{1, 2\}$, jolle pätee $\tau(1) = 2$ ja $\tau(2) = 1$. Koska $\mathrm{id} = \tau \circ \tau$, niin väite pätee.

Oletetaan, että väite pätee luvulla $n \in \mathbb{N}$. Olkoon $\sigma\colon \{1, \ldots, n+1\} \to \{1, \ldots n+1\}$ permutaatio. Olkoon nyt $i = \sigma(n+1)$ ja olkoon $\tau\colon \{1, \ldots, n+1\} \to \{1, \ldots, n+1\}$ transpoosi $\tau(i) = n+1$. Tällöin $\tau \circ \sigma$ on permutaatio, jolle pätee $(\tau \circ \sigma)(n+1) = \tau(\sigma(n+1)) = \tau(i) = n+1$. Näin ollen $\rho\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, $j \mapsto (\tau \circ \sigma)(j)$, on joukon $\{1, \ldots, n\}$ permutaatio.

Induktio-oletuksen mukaan on olemassa sellaiset transpositiot $\tau_1, \ldots, \tau_k$, että $\rho = \tau_1 \circ \cdots \circ \tau_k$. Laajennetaan transposititiot $\tau_1, \ldots, \tau_k$ joukon $\{1, \ldots, n+1\}$ transpositioiksi

[1] Itseasiassa kurssin Lineaarialgebra ja matriisilaskenta II kielellä sanotaan, että jokaisella permutatiolla σ kuvaus $(x_1, \ldots, x_n) \mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ on avaruuden $\mathbb{R}^n$ lineaarinen isomorfismi.

asettamalla $\tau_j(n+1) = n+1$ jokaisella $j \in \{1,\dots,k\}$. Näin ollen näille uusille transpositioille $\tau_1,\dots,\tau_k$ (joista käytetään samaa merkintää) pätee

$$\tau \circ \tau_1 \circ \tau_k = \sigma.$$

□

C.2 Tilavuusmuoto on alternoiva multilineaarikuvaus

Tässä luvussa osoitetaan, että avaruuden $\mathbb{R}^{n\times 1}$ tilavuusmuoto $\mathrm{vol}_{\mathbb{R}^{n\times 1}}$ on alternoiva multilineaarikuvaus.

Lause 5.3.20. *Funktio* $\mathrm{vol}_{\mathbb{R}^{n\times 1}} \colon \mathbb{R}^{n\times 1} \times \cdots \times \mathbb{R}^{n\times 1} \to \mathbb{R}$ *on alternoiva multilineaarikuvaus, jolle pätee* $\mathrm{vol}_{\mathbb{R}^{n\times 1}}(e_1,\dots,e_n) = 1$.

Todistus. Osoitetaan ensin multilineaarisuus. Olkoon $v_1,\dots,v_n \in \mathbb{R}^{n\times 1}$, $v, w \in \mathbb{R}^n$ ja $a, b \in \mathbb{R}$. Merkitään lisäksi $v_j = v_{j1}e_1 + \cdots + v_{jn}e_n$ jokaisella $j \in \{1,\dots,n\}$ sekä $v = v_1e_1 + \cdots + v_ne_n$ ja $w = w_1e_1 + \cdots + w_ne_n$.

Tarkastellaan nyt indeksiä $j \in \{1,\dots,n\}$. Tällöin

$$\begin{aligned}
&\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1,\dots,v_{j-1}, av+bw, v_{j+1},\dots,v_n)\\
&= \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}(av_{\sigma(j)} + bw_{\sigma(j)})v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}.
\end{aligned}$$

Purkamalla sisimmät sulut saadaan jokaisella permutaatiolla σ yhtälö

$$\begin{aligned}
&\mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}(av_{\sigma(j)} + bw_{\sigma(j)})v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}\\
&= \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}(av_{\sigma(j)})v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}\\
&\quad + \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}(bw_{\sigma(j)})v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}\\
&= a\mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}v_{\sigma(j)}v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}\\
&\quad + b\mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}w_{\sigma(j)}v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}.
\end{aligned}$$

Ottamalla huomioon summa yli perumutaatioiden saadaan

$$\begin{aligned}
&\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1,\dots,v_{j-1}, av+bw, v_{j+1},\dots,v_n)\\
&= \sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}(av_{\sigma(j)} + bw_{\sigma(j)})v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}\\
&= a\sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}v_{\sigma(j)}v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}\\
&\quad + b\sum_{\sigma\in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(j-1),j-1}w_{\sigma(j)}v_{\sigma(j+1),j+1}\cdots v_{\sigma(n)n}\\
&= a\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1,\dots,v_{j-1}, v, v_{j+1},\dots,v_n)\\
&\quad + b\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1,\dots,v_{j-1}, w, v_{j+1},\dots,v_n).
\end{aligned}$$

Multilineaarisuus on näin osoitettu.

Osoitetaan nyt, että funktio $\mathrm{vol}_{\mathbb{R}^{n\times 1}}$ on alternoiva. Tätä varten oletetaan, että vektorit $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$ ovat sellaisia, että $v_p = v_r$, joillain idekseillä $p < r$. Olkoon $\tau\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ transpositio, joka vaihtaa indeksien p ja r paikan, eli $\tau(p) = r$, $\tau(r) = p$ ja $\tau(j) = j$ kaikilla $j \neq p, r$. Tällöin (p, r) on ainoa pari joukossa $N(\tau)$, joten $\mathrm{sign}(\tau) = 1$.

Jaetaan permutaatiot $\mathrm{Sym}(n)$ kahteen joukkoon käyttäen apumerkintää. Olkoot

$$S^{<} = \{\sigma \in \mathrm{Sym}(n)\colon \sigma(p) < \sigma(r)\}$$

ja

$$S^{>} = \{\sigma \in \mathrm{Sym}(n)\colon \sigma(p) > \sigma(r)\}.$$

Koska $p < r$, niin $\sigma(p) \neq \sigma(r)$ jokaisella $\sigma \in \mathrm{Sym}(n)$. Näin ollen jokainen permutaatio joukossa $\mathrm{Sym}(n)$ kuuluu täsmälleen toiseen joukoista $S^{<}$ ja $S^{>}$. Lisäksi, jos $\sigma \in S^{>}$, eli $\sigma(p) > \sigma(r)$, niin $\sigma \circ \tau \in S^{<}$. Tämä seuraa siitä, että

$$(\sigma \circ \tau)(p) = \sigma(\tau(p)) = \sigma(r) < \sigma(p) = \sigma(\tau(r)) = (\sigma \circ \tau)(p).$$

Vastaavasti, jos $\sigma \in S^{<}$, niin $\sigma \circ \tau \in S^{>}$.

Käyttäen näitä merkintöjä saadaan, että

$$\begin{aligned}
\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \ldots, v_n) &= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1} \cdots v_{\sigma(n)n} \\
&= \sum_{\xi \in S^{<}} \mathrm{sign}(\xi) v_{\xi(1)1} \cdots v_{\xi(n)n} + \sum_{\zeta \in S^{>}} \mathrm{sign}(\zeta) v_{\sigma(1)1} \cdots v_{\zeta(n)n} \\
&= \sum_{\xi \in S^{<}} \mathrm{sign}(\xi) v_{\xi(1)1} \cdots v_{\xi(n)n} + \sum_{\xi \in S^{<}} \mathrm{sign}(\xi \circ \tau) v_{(\xi\circ\tau)(1)1} \cdots v_{(\xi\circ\tau)(n)n},
\end{aligned}$$

missä viimeisessä yhtäsuuruudessa on käytetty tietoa, että joukot $S^{<}$ ja $S^{>}$ vastaavat toisiaan vastaavuudella, että permutaatio $\xi \in S^{>}$ vastaa permuaatiota $\zeta = \xi \circ \tau \in S^{<}$.[2]

Analysoidaan nyt summan jälkimmäistä termiä. Käytetään nyt hyödyksi tietoa, että $v_r = v_p$. Tästä seuraa, että $v_{\tau(j)} = v_j$ jokaisella $j \in \{1, \ldots, n\}$, eli $v_{\sigma(\tau(j))j} = v_{\sigma(\tau(j))\tau(j)}$ jokaisella $j \in \{1, \ldots, n\}$. Toisaalta, koska τ on permutaatio, niin jokaisella $\sigma \in S^{<}$, tulo

$$v_{\sigma(\tau(1))\tau(1)} \cdots v_{\sigma(\tau(n))\tau(n)}$$

on tulo

$$v_{\sigma(1)1} \cdots v_{\sigma(n)n},$$

[2]Tämä hieman epämääräinen "vastaavuus" tarkoittaa tietysti sitä, että kuvaus $S^{<} \to S^{>}$, $\xi \mapsto \xi \circ \tau$ on bijektio.

jonka tulon tekijät on uudelleen järjestetty. Yhdistämällä nämä kaksi havaintoa, saadaan seuraava havainto

$$
\begin{aligned}
\sum_{\xi \in S^<} \mathrm{sign}(\xi \circ \tau) v_{(\xi\circ\tau)(1)1} \cdots v_{(\xi\circ\tau)(n)n} &= \sum_{\xi \in S^<} \mathrm{sign}(\xi)\mathrm{sign}(\tau) v_{(\xi\circ\tau)(1)1} \cdots v_{(\xi\circ\tau)(n)n} \\
&= \sum_{\xi \in S^<} \mathrm{sign}(\xi)\mathrm{sign}(\tau) v_{(\xi\circ\tau)(1)\tau(1)} \cdots v_{(\xi\circ\tau)(n)\tau(n)} \\
&= \mathrm{sign}(\tau) \sum_{\xi \in S^<} \mathrm{sign}(\xi) v_{(\xi(1)1} \cdots v_{(\xi(n)n} \\
&= -\sum_{\xi \in S^<} \mathrm{sign}(\xi) v_{(\xi(1)1} \cdots v_{(\xi(n)n}.
\end{aligned}
$$

Näin ollen

$$
\begin{aligned}
\mathrm{vol}_{\mathbb{R}^n}(v_1, \ldots, v_n) &= \sum_{\xi \in S^<} \mathrm{sign}(\xi) v_{\xi(1)1} \cdots v_{\xi(n)n} + \sum_{\xi \in S^<} \mathrm{sign}(\xi \circ \tau) v_{(\xi\circ\tau)(1)1} \cdots v_{(\xi\circ\tau)(n)n} \\
&= \sum_{\xi \in S^<} \mathrm{sign}(\xi) v_{\xi(1)1} \cdots v_{\xi(n)n} - \sum_{\xi \in S^<} \mathrm{sign}(\xi) v_{(\xi(1)1} \cdots v_{(\xi(n)n} \\
&= 0.
\end{aligned}
$$

Tämä todistaa, että funktio $\mathrm{vol}_{\mathbb{R}^n}$ on alternoiva.

Osoitetaan vielä, että $\mathrm{vol}_{\mathbb{R}^{n\times 1}}(e_1, \ldots, e_n) = 1$. Tämä seuraa havainnosta, että $e_{ji} \neq 0$ jos vain vain jos $j = i$. Näin ollen $e_{\sigma(1)1} \cdots e_{\sigma(n)n} \neq 0$ ainoastaan identtisellä permutaatiolla id. Näin ollen

$$
\mathrm{vol}_{\mathbb{R}^{n\times 1}}(e_1, \ldots, e_n) = \mathrm{sign}(\mathrm{id}) e_{11} \cdots e_{nn} = 1.
$$

□

C.3 Alternoivien multilineaarikuvausten ominaisuudet

Tässä luvussa todistetaan luvun 5.3.3 lauseet 5.3.16 ja 5.3.17.

Lauseen 5.3.16 todistus on kaksiosoinen lasku. Ensimmäisessä vaiheessa hyödynnetään vektoreiden $v_1, \ldots, v_n$ esitystä standardikannassa ja käytetään multilineaariisuutta tuomaan lineaarikombinaatiot $v_i = v_{1i}e_1 + \cdots + v_{ni}e_n$ ulos multilineaarifunktiosta ω. Toisessa vaiheessa hyödynnetään alternoivuutta havaitaan, että ensimmäisessä vaiheessa saadut summat yksinkertaistuvat. Tässä osassa todistusta tarvitaan permutaation merkin tulokaavaa ja permutaatioden esittämistä transpositioiden tulona.

Lause 5.3.16. *Olkoon* $\omega\colon \mathbb{R}^{n\times 1} \times \cdots \times \mathbb{R}^{n\times 1} \to \mathbb{R}$ *alternoiva multilineaarikuvaus. Tällöin kaikilla* $v_1, \ldots, v_n \in \mathbb{R}^n$*, missä* $v_i = v_{1i}e_1 + \cdots + v_{ni}e_n$*, pätee*

$$
\omega(v_1, \ldots, v_n) = \left(\sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1} \cdots v_{\sigma(n)n} \right) \omega(e_1, \ldots, e_n). \tag{5.5}
$$

Lauseen todistus on kaksiosoinen lasku. Ensimmäisessä vaiheessa hyödynnetään vektoreiden $v_1, \ldots, v_n$ esitystä standardikannassa ja käytetään multilineaariisuutta tuomaan lineaarikombinaatiot $v_i = v_{1i}e_1 + \cdots + v_{ni}e_n$ ulos multilineaarifunktiosta ω. Toisessa vaiheessa hyödynnetään alternoivuutta havaitaan, että ensimmäisessä vaiheessa saadut summat yksinkertaistuvat. Tässä osassa todistusta tarvitaan permutaation merkin tulokaavaa ja permutaatioden esittämistä transpositioiden tulona.

Lauseen 5.3.16 todistus. Koska

$$v_i = v_{1i}e_1 + \cdots v_{ni}e_n = \sum_{j=1}^{n} v_{ji}e_j \in \mathbb{R}^n$$

jokaisella $i \in \{1, \ldots, n\}$, niin

$$\begin{aligned}
\omega(v_1, \ldots, v_n) &= \omega\left(\sum_{j=1}^{n} v_{j1}e_j, \sum_{j=1}^{n} v_{j2}e_j, \ldots, \sum_{j=1}^{n} v_{jn}e_j\right) \\
&= \omega\left(\sum_{j_1=1}^{n} v_{j_1 1}e_{j_1}, \sum_{j_2=1}^{n} v_{j_2 2}e_{j_2}, \ldots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right),
\end{aligned}$$

missä toistuva summamuuttuja j on uudelleen nimetty muuttujiksi $j_1, j_2, \ldots, j_n$. Koska ω on multilineaarinen, niin saadaan toistamalla multilineaarisuutta n kertaa saadaan

$$\begin{aligned}
\omega(v_1, \ldots, v_n) &= \omega\left(\sum_{j_1=1}^{n} v_{j_1 1}e_{j_1}, \sum_{j_2=1}^{n} v_{j_2 2}e_{j_2}, \ldots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \sum_{j_1=1}^{n} v_{j_1 1}\omega\left(e_{j_1}, \sum_{j_2=1}^{n} v_{j_2 2}e_{j_2}, \ldots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} v_{j_1 1}v_{j_2 2}\omega\left(e_{j_1}, e_{j_2}, \sum_{j_3=1}^{n} v_{j_3 3}e_{j_3}, \ldots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \sum_{j_1, j_2=1}^{n} v_{j_1 1}v_{j_2 2}\omega\left(e_{j_1}, e_{j_2}, \sum_{j_3=1}^{n} v_{j_3 3}e_{j_3}, \ldots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \cdots \\
&= \sum_{j_1, \ldots, j_n=1}^{n} v_{j_1 1}v_{j_2 2}v_{j_n n}\omega\left(e_{j_1}, e_{j_2}, \ldots, e_{j_n}\right).
\end{aligned}$$

Huomaa, että viimeissä summassa kaikki muuttujat $j_1, \ldots, j_n$ käyvät läpi arvot $1, \ldots, n$.

Tarkastellaan nyt termejä $\omega\left(e_{j_1}, e_{j_2}, \ldots, e_{j_n}\right)$. Olkoot $j_1, \ldots, j_n \in \{1, \ldots, n\}$. Jos kahdella muuttujista $j_1, \ldots, j_n$ on sama arvo $k \in \{1, \ldots, n\}$, niin tällöin vektori e_k toistuu jonossa $(e_{j_1}, \ldots, e_{j_n})$ ainakin kahdesti. Koska ω on alternoiva, niin tällöin $\omega(e_{j_1}, \ldots, e_{j_n}) = 0$.

Näin ollen termi $\omega(e_{j_1}, \dots, e_{j_n})$ on nollasta poikkeava ainoastaan, jos muuttujat $j_1, \dots, j_n$ saavat kaikki eri arvoja. Näin ollen saatu summa voidaan kirjoittaa muodossa

$$\begin{aligned}\omega(v_1, \dots, v_n) &= \sum_{j_1,\dots,j_n=1}^{n} v_{j_1 1} v_{j_2 2} v_{j_n n} \omega\left(e_{j_1}, e_{j_2}, \dots, e_{j_n}\right) \\ &= \sum_{(j_1,\dots,j_n)} v_{j_1 1} v_{j_2 2} v_{j_n n} \omega\left(e_{j_1}, e_{j_2}, \dots, e_{j_n}\right),\end{aligned}$$

missä summa otetaan yli kaikkien sellisten lukujonojen $(j_1, \dots, j_n)$, missä $j_k \neq j_\ell$ kaikilla $k \neq \ell$.

Tälläisen lukujonot $(j_1, \dots, j_n)$ vastaavat yksikäsitteisesti joukon permutaatioita. Tämä havaitaan seuraavasti. Olkoon $(j_1, \dots, j_n)$ lukujono, jossa $j_k \in \{1, \dots, n\}$ jokaisella k ja $j_k \neq j_\ell$ kaikilla $k \neq \ell$. Määritellään kuvaus $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ kaavalla $k \mapsto j_k$. Tällöin σ on bijektio, eli permutaatio. Lisäksi kaksi eri jonoa $(j_1, \dots, j_n)$ ja $(j'_1, \dots, j'_n)$ määrittelevät eri bijektiot. Lisäksi, jos $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ on bijektio, niin $(\sigma(1), \dots, \sigma(n))$ on vaatimukset täyttävä lukujono. (Yksityiskohdat jätetään harjoitustehtäväksi.)

Näin ollen summa voidaan ottaa jonojen sijaan yli permutaatioiden ja saadaan

$$\begin{aligned}\omega(v_1, \dots, v_n) &= \sum_{j_1,\dots,j_n=1}^{n} v_{j_1 1} v_{j_2 2} v_{j_n n} \omega\left(e_{j_1}, e_{j_2}, \dots, e_{j_n}\right) \\ &= \sum_{\sigma \in \mathrm{Sym}(n)} v_{\sigma(1)1} v_{\sigma(2)2} v_{\sigma(n)n} \omega\left(e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}\right).\end{aligned}$$

Koska lauseen 5.3.17 perusteella

$$\omega(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \mathrm{sign}(\sigma)\omega(e_1, \dots, e_n),$$

niin näin saadaan

$$\begin{aligned}\omega(v_1, \dots, v_n) &= \sum_{\sigma \in \mathrm{Sym}(n)} v_{\sigma(1)1} v_{\sigma(2)2} v_{\sigma(n)n} \omega\left(e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}\right) \\ &= \sum_{\sigma \in \mathrm{Sym}(n)} v_{\sigma(1)1} v_{\sigma(2)2} v_{\sigma(n)n} \mathrm{sign}(\sigma) \omega\left(e_1, e_2, \dots, e_n\right) \\ &= \left(\sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1} v_{\sigma(2)2} v_{\sigma(n)n} \right) \omega\left(e_1, e_2, \dots, e_n\right).\end{aligned}$$

Tämä päättää todistuksen. $\square$

Osoitetaan nyt lause 5.3.17 eli että alternoivan multilineaarikuvauksen arvot muuttuvat permutaation merkillä argumentteja permutoitaessa. Tämä tulos seuraa itseasiassa suoraan permutaation esittämisestä transpositioiden avulla ja alternoivuudesta.

Lause 5.3.17. *Olkoon* $\omega\colon \mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1}\to\mathbb{R}$ *alternoiva multilineaarikuvaus, olkoot* $v_1,\dots,v_n\in\mathbb{R}^n$ *vektoreita ja* $\sigma\colon\{1,\dots,n\}\to\{1,\dots n\}$ *permutaatio. Tällöin*

$$\omega(v_{\sigma(1)},\dots,v_{\sigma(n)})=\operatorname{sign}(\sigma)\omega(v_1,\dots,v_n).$$

Todistus. Alternoivuuden perusteella

$$\omega(w_{\tau(1)},\dots,w_{\tau(n)})=(-1)\omega(w_{\tau(1)},\dots,w_{\tau_n})$$

kaikilla vektoreilla $w_1,\dots,w_n\in\mathbb{R}^{n\times 1}$ ja transpositioilla $\tau\colon\{1,\dots,n\}\to\{1,\dots,n\}$.

Olkoot $\tau_1,\dots,\tau_m$ sellaisia transpositioita, että $\sigma=\tau_1\circ\cdots\circ\tau_m$. Tällöin induktiolla saadaan

$$\omega(v_{(\tau_1\circ\cdots\circ\tau_m)(1)},\dots,v_{(\tau_1\circ\cdots\circ\tau_m)(n)})=(-1)^m\omega(v_1,\dots,v_n)=\operatorname{sign}(\sigma)\omega(v_1,\dots,v_n).$$

□

C.4 Determinantin kehityskaavan lemman todistus

Lemma 5.9.7. *Olkoon* $B=[b_{ji}]\in\mathbb{R}^{n\times n}$ *sellainen matriisi, että* $b_{ni}=0$ *kaikilla* $i<n$ *ja* $b_{nn}=1$. *Tällöin*

$$\det B=\det B_{nn}.$$

Todistus. Olkoon

$$S=\{\sigma\in\operatorname{Sym}(n)\colon \sigma(n)=n\}.$$

Tällöin jokaisella $\sigma\in S$ pätee, että rajoittuma $\sigma|_{\{1,\dots,n-1\}}\colon\{1,\dots,n-1\}\to\{1,\dots,n-1\}$ on joukon $\{1,\dots,n-1\}$ permutaatio. Lisäksi joukon S permutaatiot vastaavat yksikäsitteisesti permutaatioita $\operatorname{Sym}(n-1)$, jossa vastaavuus saadaan laajentamalla permutaatio $\sigma\in\operatorname{Sym}(n)$ permutaatioksi $\hat\sigma\colon\{1,\dots,n\}\to\{1,\dots,n\}$ kaavalla $\hat\sigma(k)=\sigma(k)$, jos $k<n$, ja $\hat\sigma(n)=n$.[3]

Olkoon nyt $\sigma\in\operatorname{Sym}(n)$. Jos $\sigma(n)\neq n$, niin $b_{\sigma(n)n}=0$ ja $\operatorname{sign}(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n}=0$. Näin ollen

$$\begin{aligned}\det B&=\sum_{\sigma\in\operatorname{Sym}(n)}\operatorname{sign}(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n}\\&=\sum_{\sigma\in S}\operatorname{sign}(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n}\\&=\sum_{\sigma\in S}\operatorname{sign}(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n-1)n}\cdot 1\\&=\sum_{\sigma\in\operatorname{Sym}(n-1)}\operatorname{sign}(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n-1)n}\\&=\det B_{nn}.\end{aligned}$$

□

[3] Tämä vastaavuus on itseasiassa bijektio $\operatorname{Sym}(n-1)\to S$.

C.5 Cramerin sääntö ja käänteismatriisin kaava

Yksi determinantin (teoreettisista) sovelluksista on kaava matriisin käänteismatriisille. Tarkka väite on seuraava.

Lause C.5.1. *Olkoon $A \in \mathbb{R}^{n\times n}$ kääntyvä matriisi. Tällöin*

$$A^{-1} = \frac{1}{\det A}\mathrm{adj}A,$$

missä $\mathrm{adj}A \in \mathbb{R}^{n\times n}$ *on matriisi, jonka alkiot ovat* $(\mathrm{adj}A)_{ji} = (-1)^{j+i}\det(A|_{ij})$.

Matriisia $\mathrm{adj}A$ kutsutaan matriisin A *(klassiseksi) adjungaatiksi* ja yllä matriisi $A|_{ij}$ on determinantin kehityskaavan yhteydessä esitelty matriisi, joka saadaan matriisista A poistamalla i:s rivi ja j:s sarake.

Lauseen C.5.1 todistuksen ytimessä on niin kutsuttu Cramerin sääntö, joka kertoo kuinka yhtälön $Ax = b$ ratkaisun x kertoimet saadaan sellaisten matriisien determinanteista, jotka saadaan vaihtamalla sarakevektori b matriisin A sarakkeiden tilalle. Määritellään nämä matriisit Cramerin säännön todistusta varten.

Olkoon $A = [v_1 \cdots v_n] \in \mathbb{R}^{n\times n}$ matriisi, jonka sarakkeet ovat $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$, ja olkoon $b \in \mathbb{R}^{n\times 1}$ sarakevektori. Tällöin jokaisella $i \in \{1, \ldots, n\}$ merkitään

$$A_i(b) = \begin{bmatrix} v_1 & \cdots & v_{i-1} & b & v_{i+1} & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Cramerin sääntö sanoo, että yhtälön $Ax = b$ ratkaisu x voidaan laskea matriisien $A_1(b), \ldots, A_n(b)$ determinanteista. Väite on seuraava.

Lemma C.5.2. *Olkoon $A \in \mathbb{R}^{n\times n}$ kääntyvä matriisi ja $b \in \mathbb{R}^{n\times 1}$. Tällöin yhtälön $Ax = b$ yksikäsitteisen ratkaisun $x = [x_1 \cdots x_n]^t \in \mathbb{R}^{n\times 1}$ kertoimet ovat*

$$x_i = \frac{\det A_i(b)}{\det A}$$

jokaisella $i \in \{1, \ldots, n\}$.

Todistus. Olkoon $e_1, \ldots, e_n$ avaruuden $\mathbb{R}^{n\times 1}$ standardikanta ja olkoon $i \in \{1, \ldots, n\}$. Koska

$$I_i(x) = [e_1 \cdots e_{i-1}\ x\ e_{i+1} \cdots e_n],$$

niin

$$\begin{aligned} AI_i(x) &= A[e_1 \cdots e_{i-1}\ x\ e_{i+1} \cdots e_n] \\ &= [Ae_1 \cdots Ae_{i+1}\ Ax\ Ae_{i+1} \cdots Ae_n] \\ &= [a_1 \cdots a_{i-1}\ b\ a_{i+1} \cdots a_n] = A_i(b). \end{aligned}$$

Koska

$$\det I_i(x) = x_i,$$

niin

$$\det(A)x_i = \det(A)\det(I_i(x)) = \det(AI_i(x)) = \det(A_i(b)).$$

Koska A on kääntyvä, niin $\det(A) \neq 0$ ja

$$x_i = \frac{\det(A_i(b))}{\det(A)}.$$

□

Todistetaan nyt Cramerin säännön avulla käänteismatriisin kaava.

Lauseen C.5.1 todistus. Merkitään $A^{-1} = [v_1 \cdots v_n] \in \mathbb{R}^{n\times n}$, missä $v_i = [v_{1i} \cdots v_{ni}]^t \in \mathbb{R}^{n\times 1}$ jokaisella $i \in \{1, \ldots, n\}$.

Koska $AA^{-1} = I = [e_1 \cdots e_n]$, niin $Av_i = e_i$ jokaisella $i \in \{1, \ldots, n\}$. Cramerin kaavan (lemma C.5.2) nojalla

$$v_{ji} = \frac{\det A_j(e_i)}{\det A}.$$

jokaisella $j \in \{1, \ldots, n\}$.

Kehitetään matriisin $A_j(e_i)$ determinantti sarakkeen j suhteen, jolloin saadaan

$$\det(A_j(e_i)) = (-1)^{j+i} \det(A|_{ij}) = \operatorname{adj}(A)_{ji}.$$

Näin ollen

$$A^{-1} = \frac{1}{\det A} \operatorname{adj}(A).$$

□

C.6 Hadamardin epäyhtälö

Joissain sovelluksissa halutaan arvioida matriisin determinantin suuruutta. Yleensä tämä tehdään Hadamardin epäyhtälön avulla.

Lause C.6.1 (Hadamardin epäyhtälö)**.** *Olkoon* $A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n\times n}$ *neliömatriisi, missä* $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$ *ovat matriisin* A *sarakkeet. Tällöin*

$$|\det A| \leq |v_1| \cdots |v_n|. \tag{C.1}$$

Lisäksi, jos vektorit $v_1, \ldots, v_n$ *ovat nollasta poikkeavia, niin epäyhtälössä* (C.1) *pätee yhtälö, jos ja vain jos matriisin* A *sarakkeet ovat toisiaan vastaan kohtisuorassa.*

Hadamardin epäyhtälö on suora sovellus QR-hajotelmasta, kahdesta determinantin perusominaisuudesta sekä yläkolmiomatriiseja koskevasta huomiosta. Tarvittavat aputulokset kirjataan lemmoiksi ja jätetään lukijalle harjoitustehtäviksi.

Lemma C.6.2. *Olkoon* $U \in \mathbb{R}^{n\times n}$ *ortogonaalimatriisi. Tällöin* $|\det U| = 1$.

Lemma C.6.3. *Olkoon* $R = [r_{ji}] \in \mathbb{R}^{n\times n}$ *yläkolmiomatriisi. Tällöin* $\det R = r_{11} \cdots r_{nn}$.

Lemma C.6.4. *Olkoon* $R = [r_{ji}] \in \mathbb{R}^{n\times n}$ *sellainen yläkolmiomatriisi, että sen sarakkeet ovat nollasta poikkeavia ja toisiaan vastaan kohtisuorassa. Tällöin* R *on diagonaalimatriisi, eli* $R = \begin{bmatrix} r_{11}e_1 & \cdots & r_{nn}e_n \end{bmatrix}$.

Lauseen C.6.1 todistus. Lauseen 4.6.1 nojalla on olemassa sellaiset ortogonaalimatriisi $Q \in \mathbb{R}^{n\times n}$ ja yläkolmiomatriisi $R = [r_{ji}] \in \mathbb{R}^{n\times n}$, että

$$A = QR.$$

Olkoot nyt $w_1, \ldots, w_n \in \mathbb{R}^{n\times 1}$ matriisin R sarakkeet, eli $R = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix}$. Tällöin

$$A = QR = Q\begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} Qw_1 & \cdots & Qw_n \end{bmatrix},$$

eli $v_i = Qw_i$ jokaisella $i \in \{1, \ldots, n\}$.

Koska Q on ortogonaalimatriisi, niin

$$|Qw_i|^2 = (Qw_i) \cdot (Qw_i) = w_i^t Q^t Q w_i = w_i^t w_i = w_i \cdot w_i = |w_i|^2$$

jokaisella $i \in \{1, \ldots, n\}$. Toisaalta

$$|r_{ii}| \leq \sqrt{r_{1i}^2 + \cdots + r_{ni}^2} = |w_i|$$

jokaisella $i \in \{1, \ldots, n\}$. Lisäksi tässä epäyhtälössä pätee yhtäsuuruus, jos ja vain jos $r_{ji} = 0$ kaikilla $j \neq i$, eli $w_i = r_{ii} e_i$.

Koska

$$\det A = \det(QR) = (\det Q)(\det R),$$

niin lemmojen C.6.2 ja C.6.3 perusteella saadaan

$$\begin{aligned} |\det A| = |\det R| &= |r_{11} \cdots r_{nn}| \\ &= |r_{11}| \cdots |r_{nn}| \\ &\leq |w_1| \cdots |w_n| \\ &= |v_1| \cdots |v_n|. \end{aligned}$$

Tämä päättää Hadamardin epäyhtälön (C.1) todistuksen.

Osoitetaan nyt, että jos vektorit $v_1, \ldots, v_n$ ovat nollasta poikkeavia, niin epäyhtälössä pätee yhtäsuuruus, jos ja vain jos matriisin A sarakkeet ovat toisiaan vastaan kohtisuorassa.

Oletetaan ensin, että epäyhtälössä pätee yhtäsuuruus. Tällöin $w_i = r_{ii} e_i$ jokaisella $i \in \{1, \ldots, n\}$, eli R on diagonaalimatriisi. Näin ollen

$$Ae_k \cdot Ae_\ell = v_k \cdot v_\ell = (Qw_k) \cdot (Qw_\ell) = w_k \cdot w_\ell = (r_{kk} e_k) \cdot (r_{\ell\ell} e_\ell) = r_{kk} r_{\ell\ell} e_k \cdot e_\ell = 0$$

kaikilla $k \neq \ell$.

Oletaan nyt, että matriisin A sarakkeet ovat kohtisuorassa toisiaan vastaan. Tällöin kaikilla $k \neq \ell$ pätee

$$0 = v_k \cdot v_\ell = (Qw_k) \cdot (Qw_\ell) = w_k \cdot w_\ell$$

eli matriisin R sarakkeet ovat toisiaan vastaan kohtisuorassa. Koska $|w_i| = |v_i|$ jokaisella $i \in \{1, \ldots, n\}$, niin lemman C.6.4 perusteella matriisi R on diagonaalimatriisi, eli $R = \begin{bmatrix} r_{11} e_1 & \cdots & r_{nn} e_n \end{bmatrix}$. Nyt $|w_i| = |r_{ii}|$ jokaisella $i \in \{1, \ldots, n\}$, joten epäyhtälössä (C.1) pätee yhtäsuuruus. $\square$

Liite D

Liite: Suoran summan lisäsivut

D.1 Useamman aliavaruuden suora summa

Useamman aliavaruuden summa on helppo määritellä käyttäen määritelmän 8.3.1 ideaa.

Määritelmä D.1.1. *Vektoriavaruuden V aliavaruuksien $W_1, \ldots, W_k \subset V$* summa *on aliavaruus*

$$W_1 + \cdots + W_k = \{w_1 + \cdots + w_k \in V : w_i \in W_i \text{ jokaisella } i \in \{1, \ldots, k\}\}.$$

Jälleen on suoraviivaista todistaa, että $W_1 + \cdots + W_k$ todellakin on aliavaruus ja tämä jätetään harjoitustehtäväksi.

Useamman aliavaruuden suora summa määritellään seuraavasti.

Määritelmä D.1.2. *Vektoriavaruuden V aliavaruuksien $W_1, \ldots, W_k \subset V$ summa $W_1 + \cdots + W_k$ on* suora*, jos*

$$(W_1 + \cdots + W_{\ell-1}) \cap W_\ell = \{0\}$$

jokaisella $\ell \in \{2, \ldots, k\}$. Tällöin merkitään $W_1 \oplus \cdots \oplus W_k = W_1 + \cdots + W_k$.

Huomautus D.1.3. *Koska $W_1 + \cdots + W_\ell = (W_1 + \cdots + W_{\ell-1}) + W_\ell$, niin tämä määritelmä voidaan tulkita siten, että aliavaruuksien $W_1, \ldots, W_k$ summa on suora, jos $W_1 + \cdots + W_\ell = (W_1 + \cdots + W_{\ell-1}) \oplus W_\ell$ jokaisella $\ell \in \{2, \ldots, k\}$.*

Seuraava lause antaa kuitenkin hyödyllisemmän tulkinnan, joka sanoo, että aliavaruuksien summa on suora, jos ja vain jos jokaisella vektorilla on yksikäsitteinen esitys aliavaruuksien vektoreiden summana. Tämä tulos on analoginen lauseen 8.3.3 kanssa ja sitä voidaan käyttää summan suoruuden tarkastamiseen, kuten lauseen jälkeen käsiteltävät esimerkit havainnollistavat.

Lause D.1.4. *Olkoon V vektoriavaruus. Aliavaruuksien $W_1, \ldots, W_k \subset V$ summa $W_1 + \cdots + W_k$ summa on suora, jos ja vain jos jokaisella $w \in W_1 + \cdots + W_k$ on olemassa sellainen yksikäsitteinen vektori $w_i \in W_i$ jokaisella $i \in \{1, \ldots, k\}$, että $w = w_1 + \cdots + w_k$.*

Todistus. Oletetaan, että summa $W_1+\cdots+W_k$ on suora. Olkoon nyt $w \in W_1+\cdots+W_k$. Tällöin jokaisella $i \in \{1,\ldots,k\}$ on olemassa sellainen vektori $w_i \in W_i$, että $w = w_1+\cdots+w_k$. Olkoot nyt jokaisella $i \in \{1,\ldots,k\}$ $w_i' \in W_i$ sellainen vektori, että $w = w_1'+\cdots+w_k'$. Osoitetaan, että $w_i = w_i'$ jokaisella $i \in \{1,\ldots,k\}$.

Merkitään

$$v_{k-1} = w_1 + \cdots + w_{k-1} \in W_1 + \cdots + W_{k-1}$$

ja

$$v_{k-1}' = w_1' + \cdots + w_{k-1}' \in W_1 + \cdots + W_{k-1}.$$

Tällöin $w = v_{k-1} + w_k$ ja $w = v_{k-1}' + w_k'$. Koska summa $(W_1 + \cdots + W_{k-1}) + W_k$ on suora, niin lauseen 8.3.3 nojalla $v_{k-1} = v_{k-1}'$ ja $w_k = w_k'$.

Vastavasti osoitetaan, että $w_\ell = w_\ell'$ jokaisella $\ell = k-1,\ldots,1$. Näin ollen vektorit $w_1,\ldots,w_k$ ovat yksikäsitteisiä.

Oletetaan nyt, että jokaisella $w \in W_1+\cdots+W_k$ on olemassa yksikäsitteiset sellaiset vektorit $w_i \in W_i$, että $w = w_1+\cdots+w_k$. Osoitetaan, että $(W_1+\cdots+W_{\ell-1}) \cap W_\ell = \{0\}$ jokaisella $\ell \in \{2,\ldots,k\}$.

Tehdään vastaoletus, että $(W_1+\cdots+W_{\ell-1}) \cap W_\ell \neq \{0\}$ jollain $\ell \in \{2,\ldots,k\}$. Olkoon $w \in (W_1 + \cdots + W_{\ell-1}) \cap W_\ell$ nollasta poikkeava vektori. Tällöin $w = w_1 + \cdots + w_{\ell-1}$ jollain $w_i \in W_i$ jokaisella $i \in \{1,\ldots,\ell\}$. Lisäksi $w \in W_\ell$. Merkitään $w_\ell = -w$. Tällöin

$$0 = w_1 + \cdots + w_{\ell-1} + w_\ell.$$

Koska $0 = w_1'+\cdots+w_\ell'$, missä $w_1' = \cdots = w_\ell' = 0$, niin oletuksen nojalla $w_1 = \cdots = w_\ell = 0$. Tämä on ristiriita, koska $0 = w_\ell = -w \neq 0$ vektorin w valinnan nojalla. Vastaoletus on siis väärä ja $(W_1 + \cdots + W_{\ell-1}) \cap W_\ell = \{0\}$ jokaisella $\ell \in \{2,\ldots,k\}$. $\square$

Lause D.1.5. *Olkoon V kompleksinen vektoriavaruus ja olkoot $W_1,\ldots,W_m$ avaruuden V äärellisulotteisia aliavaruuksia. Olkoon $(w_{j1},\ldots,w_{jd_j})$ aliavaruuden W_j kanta jokaisella $j \in \{1,\ldots,m\}$. Tällöin summa $W_1 + \cdots + W_m$ on suora, jos ja vain jos jono $(w_{11},\ldots,w_{1d_1},w_{21},\ldots,w_{md_m})$ on aliavaruuden $W_1 + \cdots + W_m$ kanta.*

Todistus. Oletetaan ensin, että summa $W_1 + \cdots + W_m$ on suora. Riittää osoittaa, että jono

$$(w_{11},\ldots,w_{1d_1},w_{21},\ldots,w_{md_m})$$

on vapaa.

Olkoot $a_{ji} \in \mathbb{R}$ kaikilla $j \in \{1,\ldots,m\}$ ja $i \in \{1,\ldots,d_j\}$ sellaisia lukuja, että

$$\sum_{j=1}^{m}\sum_{i=1}^{d_j} a_{ji}w_{ji} = 0.$$

Olkoon nyt

$$w_j = \sum_{i=1}^{d_j} a_{ji}w_{ji} \in W_j$$

jokaisella $j \in \{1, \ldots, m\}$.

Koska

$$w_1 + \cdots + w_m = 0,$$

niin lauseen D.1.4 perusteella $w_1 = \cdots = w_m = 0$. Koska jono $(w_{j1}, \ldots, w_{jd_j})$ on kanta jokaisella $j \in \{1, \ldots, m\}$, niin saadaan $a_{ji} = 0$ kaikilla indekseillä j ja i. Jono on siis vapaa.

Oletetaan nyt, että jono $(w_{11}, \ldots, w_{1d_1}, w_{21}, \ldots, w_{md_m})$ on kanta. Olkoon $w \in W_1 + \cdots + W_m$. Tällöin on olemassa yksikäsitteisen sellaiset luvut a_{ji}, missä $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, d_j\}$, että

$$w = \sum_{j=1}^{m} \sum_{i=1}^{d_j} a_{ji} w_{ji}.$$

Näin ollen on olemassa yksikäsitteiset sellaiset vektorit $w_j \in W_j$, että $w = w_1 + \cdots + w_m$. Summa $W_1 + \cdots + W_m$ on siis suora. □

Huomautus D.1.6. *Yksi lauseen D.1.4 seurauksista on, että aliavaruuksien $W_1, \ldots, W_k$ summan $W_1 + \cdots + W_k$ suoruus ei riipu aliavaruuksien järjestyksestä.*

Tarkastellaan nyt jälleen sarakeavaruuden $\mathbb{R}^{4\times 1}$ vektoreiden

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \text{ ja } v_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

virittämää aliavaruutta

$$V = \mathrm{Sp}(v_1, v_2, v_3) \subset \mathbb{R}^{4\times 1}.$$

Kuten edellisessä luvussa todettiin, (v_1, v_2, v_3) on aliavaruuden V kanta.

Esimerkki D.1.7. *Olkoot*

$$W_1 = \mathrm{Sp}(v_1), \; W_2 = \mathrm{Sp}(v_2) \textit{ ja } W_3 = \mathrm{Sp}(v_3)$$

avaruuden V aliavaruuksia.

Luvun 3 perusteella jokainen aliavaruus W_1, W_2, W_3 on yksiulotteinen eli suora. Tarkastelemalla aliavaruutta V vektoriavaruutena havaitaan, että

$$V = \mathrm{Sp}(v_1, v_2, v_3) = \mathrm{Sp}(v_1) + \mathrm{Sp}(v_2) + \mathrm{Sp}(v_3).$$

Koska (v_1, v_2, v_3) on aliavaruuden V kanta, niin tämä summa on itseasiassa suora summa, eli

$$V = \mathrm{Sp}(v_1) \oplus \mathrm{Sp}(v_2) \oplus \mathrm{Sp}(v_3) = W_1 \oplus W_2 \oplus W_3.$$

Osoitetaan tämä nyt tarkasti käyttäen luvun 3 tuloksia ja lausetta 8.3.3.

Olkoon $v \in V$. Koska (v_1, v_2, v_3) on aliavaruuden V kanta, niin vektorilla v on yksikäsitteiset koordinaatit (x_1, x_2, x_3), eli $v = x_1v_1 + x_2v_2 + x3_v3$. Asetetaan nyt $w_1 = x_1v_1 \in W_1$, $w_2 = x_2v_2 \in W_2$ ja $w_3 = x_3v_3 \in W_3$. Nyt

$$v = x_1v_1 + x_2v_2 + x_3v_3 = w_1 + w_2 + w_3.$$

Koska koordinaait ovat yksikäsitteisiä, niin myös vektorit w_1, w_2, w_3 ovat yksikäsitteisiä. Osoitetaan tämä kuitenkin vielä tarkasti. Olkoot $w_1' \in W_1$, $w_2' \in W_2$ ja $w_3' \in W_3$ sellaisia vektoreita, että $v = w_1' + w_2' + w_3'$. Aliavaruuksien W_1, W_2, W_3 määritelmien perusteella, jokaisella $i \in \{1, 2, 3\}$ on olemassa sellainen luku $y_i \in \mathbb{R}$, että $w_i' = y_iv_i$. Tällöin

$$v = w_1' + w_2' + w_3' = y_1v_1 + y_2v_2 + y_3v_3.$$

Vektorin v koordinaattien yksikäsitteisyyden nojalla saadaan, että $y_1 = x_1$,$y_2 = x_2$ ja $y_3 = x_3$. Näin ollen $w_i' = w_i$ jokaisella $i\{1, 2, 3\}$.

Kirjataan vielä lyhyesti ylös yksi esimerkki samasta teemasta. Esimerkin tarkemmat pohdinnat jätetään kiinostuneelle lukijalle.

Esimerkki D.1.8. *Olkoot $U = \mathrm{Sp}(v_1, v_2)$ ja $W = \mathrm{Sp}(v_3)$ kuten esimerkissä 8.3.9, eli tiedetään, että $V = U \oplus W$. Summaan voidaan ottaa mukaan nolla-avaruuksia $\{0\}$ summan suoruuden siitä muuttumatta, eli esimerkiksi*

$$V = U \oplus \{0\} \oplus W.$$

Tämä seuraa havainnosta, että jokainen $v \in V$ voidaan kirjoittaa yksikäsitteisesti muodossa

$$v = u + 0 + w,$$

missä $u \in U$, $0 \in \{0\}$ ja $w \in W$.

D.2 Ominaisavaruuksien summa on suora

Osoitetaan nyt tämän teorian sovelluksena, että lineaarisen operaattorin ominaisavaruuksien summa on suora.

Lause D.2.1 (Ominaisavaruuksien summa on suora)**.** *Olkoon $f\colon V \to V$ lineaarinen operaattori ja olkoot $\lambda_1 < \cdots < \lambda_k$ sen ominaisarvoja. Tällöin ominaisavaruuksien summa $E(\lambda_1, f) + \cdots + E(\lambda_k, f)$ on suora, eli*

$$E(\lambda_1, f) + \cdots + E(\lambda_k, f) = E(\lambda_1, f) \oplus \cdots \oplus E(\lambda_k, f).$$

Tämän ominaisavaruuksia koskevan tuloksen voi muotoilla myös konkreettisemmin muodossa, että eri ominaisarvoja vastaavat ominaisvektorit ovat lineaarisesti riippumattomia. Tämä muoto tuloksesta on kirjattu seuraavaan lauseeseen ja lause 11.3.7 todistetaan heti tämän muotoilun jälkeen.

Lauseen 11.3.7 todistus. Olkoon $v \in E(\lambda_1, f) + \cdots + E(\lambda_k, f)$ ja olkoot vektorit $v_i, v_i' \in E(\lambda_i, f)$ jokaisella $i \in \{1, \dots, k\}$ sellaisia, että

$$v = v_1 + \cdots + v_k$$

ja

$$v = v_1' + \cdots + v_k'.$$

Tällöin

$$(v_1 - v_1') + \cdots + (v_k - v_k') = 0.$$

Osoitetaan, että $v_i - v_i' = 0$ jokaisella $i \in \{1, \dots, k\}$. Tehdään vastaolotus, että näin ole vaan, että on olemassa sellainen indeksi $i \in \{1, \dots, k\}$, että $v_i - v_i' \neq 0$. Olkoot nyt $1 \leq j_1 < \cdots < j_m \leq k$ kaikki sellaiset indeksit $i \in \{1, \dots, k\}$, joilla $v_i - v_i' \neq 0$. Koska $v_i - v_i' \in E(\lambda_i, f)$ jokaisella $i \in \{1, \dots, k\}$, niin soveltamalla lausetta 11.3.5 ominaisarvoihin $\lambda_{j_1} < \ldots < \lambda_{j_m}$ ja ominaisvektoreihin $w_1, \dots, w_m$, missä $w_i = v_{j_i} - v_{j_i}'$ jokaisella $i \in \{1, \dots, m\}$, saadaan ristiriita. Näin ollen $v_i - v_i' = 0$ jokaisella $i \in \{1, \dots, k\}$.

Koska jokaisella vektorilla v on yksikäsitteinen esitys summassa $E(\lambda_1, f) + \cdots + E(\lambda_k, f)$, niin summa $E(\lambda_1, f) + \cdots + E(\lambda_k, f)$ on lauseen D.1.4 perusteella suora. $\square$

Liite E

Liite: Kompleksisten ominaisarvojen lisäsivut

E.1 Kompleksisen ominaisarvon olemassaolo ilman determinanttia

Tässä liitteessa annetaan toinen todistus kompleksisen ominaisarvon olemassaololle, eli lauseelle 14.1.1. Aihetta on käsitelty tarkemmin Axlerin kirjassa [1].

Todistus perustuu seuraavaan määritelmään.

Määritelmä E.1.1. *Olkoon $f\colon V \to V$ lineaarioperaattori ja $p\colon \mathbb{C} \to \mathbb{C}$ polynomi $z \mapsto a_n z^n + \cdots + a_1 z + a_0$. Tällöin $p(f)\colon V \to V$ on kaavalla*

$$v \mapsto a_n f^n(v) + \cdots + a_1 f(v) + a_0,$$

määritelty lineaarioperaattori, missä $f^n(v) = f(f^{n-1}(v))$ kaikilla $n \geq 1$ ja $f^0 = \mathrm{id}_V$.

Tarvitsemme myös pienen lemman.

Lemma E.1.2. *Olkoon $f\colon V \to V$ lineaarioperaattori ja olkoot $p\colon \mathbb{C} \to \mathbb{C}$ ja $q\colon \mathbb{C} \to \mathbb{C}$ polynomeja. Merkitään myöt $pq\colon \mathbb{C} \to \mathbb{C}$ näiden polynomien tuloa, eli funktiota $z \mapsto p(z)q(z)$. Tällöin*

$$pq(f) = p(f) \circ q(f) = q(f) \circ p(f).$$

Todistus. Olkoon $p\colon \mathbb{C} \to \mathbb{C}$ polynomi $z \mapsto \sum_{j=0}^{m} a_j z^j$ ja $q\colon \mathbb{C} \to \mathbb{C}$ polynomi $z \mapsto \sum_{k=0}^{n} b_k z^k$. Tällöin

$$pq(z) = \left(\sum_{j=0}^{m} a_j z^j\right)\left(\sum_{k=0}^{n} b_k z^k\right) = \sum_{j=0}^{m}\left(a_j z^j \sum_{k=0} n b_k z^k\right) = \sum_{j=0}^{m}\sum_{k=0}^{n} a_j b_k z^{j+k}.$$

Vastaavalla laskulla saadaan, että $qp = pq$.

Koska f on lineaarikuvaus, niin kaikilla $v \in V$ pätee

$$\begin{aligned} pq(f)(v) &= \sum_{j=0}^{m}\sum_{k=0}^{n} a_j b_k f^{j+k}(v) = \sum_{j=0}^{m}\sum_{k=0}^{n} a_j b_k f^j(f^k(v)) \\ &= \sum_{j=0}^{m} a_j f^j\left(\sum_{k=0}^{n} b_k f^{j+k}(v)\right) = \sum_{j=0}^{m} a_j f^j(Q(f)(v)) \\ &= p(f)\,(q(f)(v)) = (p(f)\circ q(f))(v). \end{aligned}$$

Koska $pq = qp$, niin $pq(f) = qp(f)$. Näin ollen

$$p(f)\circ q(f) = (pq)(f) = (qp)(f) = q(f)\circ p(f).$$

$\square$

Edellisen lemman merkitys on seuraava. Olkoon $p\colon \mathbb{C} \to \mathbb{C}$ polynomi $z \mapsto a_0 + a_1 z + \cdots + a_n z^n$. Tekijöihin jaon perusteella p voidaan kirjoittaa muodossa $p(z) = a(z-\lambda_1)\cdots(z-\lambda_n) = ap_1(z)\cdots p_n(z)$, missä jokainen $p_i\colon \mathbb{C} \to \mathbb{C}$ on polynomi $z \mapsto z-\lambda_i$. Lemman E.1.2 nojalla kaikilla $f\colon V \to V$ pätee

$$p(f)(v) = a_0 + a_1 f(v) + \cdots + a_n f^n(v) = a\,(p_1(f)\circ\cdots\circ p_n(f))\,(v),$$

missä $p_i(f)\colon V \to V$ on operaattori $p_i(f) = f - \lambda_i \mathrm{id}_V$.

Lauseen 14.1.1 vaihtoehtoinen todistus. Olkoon $n = \dim V$ ja olkoon $v \in V \setminus \{0\}$ (jokin) vektori. Koska jonossa

$$(v, f(v), f^2(v), \ldots f^n(v))$$

on $n+1$ jäsentä, niin se ei ole lineaarisesti riippumaton. Näin ollen on olemassa sellaiset kertoimet $a_0, \ldots, a_n \in \mathbb{C}$, joista jokin on nollasta poikkeava, että

$$a_0 v + a_1 f(v) + \cdots + a_n f^n(v) = 0.$$

Havaitaan ensin, että jokin kertoimista $a_1, \ldots, a_n$ ei ole nolla. Jos näin pätisi, niin silloin $0 = a_0 v$. Tällöin $a_0 = 0$, sillä $v \neq 0$. Tämä on ristiriita, koska jokin kertoimista $a_0, \ldots, a_n$ ei ole nolla. Olemme siis päätelleet, että jokin kertoimista $a_1, \ldots, a_n$ ei ole nolla.

Olkoon $P\colon \mathbb{C} \to \mathbb{C}$ polynomi $z \mapsto a_0 + a_1 z + \cdots + a_n z^n$. Kompleksisten polynomien tekijöihin jaon (Korollaari 14.1.5) nojalla on olemassa $a \in \mathbb{C}$ ja sellaiset $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, että

$$p(z) = a(z-\lambda_1)\cdots(z-\lambda_n).$$

Lemman E.1.2 nojalla

$$0 = a_0 + a_1 f(v) + \cdots + a_n f^n(v) = a(f - \lambda_1 \mathrm{id}_V)\circ\cdots\circ(f - \lambda_n \mathrm{id}_V)(v).$$

Koska $v \neq 0$, niin päätellään, että on olemassa sellainen $i \in \{1, \ldots, n\}$, että $w = ((f - \lambda_{i+1}\mathrm{id}_V)\cdots\circ\cdots(f - \lambda_n \mathrm{id}_V))\,(v) \neq 0$, mutta $(f - \lambda_i \mathrm{id}_V)(w) = 0$. Näin ollen jokin operaattoreista $f - \lambda_i \mathrm{id}_V$ ei ole injektio ja on olemassa sellainen vektori $\tilde{v} \in V$, että $f(v) = \lambda_i \tilde{v}$. Operaattorilla f on siis ominaisarvo λ_i. $\square$

E.2 Kompleksisen operaattorin yläkolmioesitys

Annetaan nyt suora todistus korollaarille 14.2.5 invarianttien aliavaruuksien avulla.

Lause E.2.1. *Olkoon V n-ulotteinen kompleksinen vektoriavaruus ja $f\colon V \to V$ lineaarioperaattori. Tällöin on olemassa sellainen avaruuden V kanta $(v_1, \dots, v_n)$, että kuvauksen f esitysmatriisi tässä kannassa on yläkolmiomatriisi.*

Määritelmä E.2.2. *Kompleksisen vektoriavaruuden V aliavaruus W on* invariantti operaattorin $f\colon V \to V$ suhteen, *jos $fW \subset W$.*

Invariantin aliavaruuden merkitys on se, että operaattori $f\colon V \to V$ voidaan rajoittaa aliavaruuteen U eli että kuvaus $f|_U\colon U \to U$ on hyvin määritelty operaattori.

Esimerkki E.2.3. *Triviaalit esimerkkit invarianteista aliavaruuksista ovat koko avaruus V ja nolla-avaruus $\{0\}$. Muita esimerkkejä operaattoreihin liittyvistä invarianteista aliavaruuksista ovat ydin ja kuva. Tämän luvun kannalta tärkein esimerkki operaattorin $f\colon V \to V$ invariantista aliavaruudesta on ominaisarvoon $\lambda \in \mathbb{C}$ liittyvä ominaisavaruus $E(\lambda, f) \subset V$.*

Seuraava lemma kertoo kuinka yläkolmiomatriisit ja invariantit aliavaruudet liittyvät toisiinsa. Tätä lemmaa voidaan ajatella huomautuksen 14.2.2 vastineena invarianteille aliavaruuksille. Huomaa, että lemma pätee sekä kompleksisille että reaalisille vektoriavaruuksille.

Lemma E.2.4. *Olkoon V vektoriavaruus, $(v_1, \dots, v_n)$ avaruuden V kanta ja $f\colon V \to V$ operaattori. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *lineaarikuvauksen f esitysmatriisi kannassa $(v_1, \dots, v_n)$ on yläkolmiomatriisi,*
2. *$f(v_j) \in \mathrm{Sp}(v_1, \dots, v_j)$ kaikilla $j \in \{1, \dots, n\}$ ja*
3. *$\mathrm{Sp}(v_1, \dots, v_j)$ on operaattorin f invariantti aliavaruus jokaisella $j \in \{1, \dots, n\}$.*

Todistus. Oletetetaan, että operaattorin f esitysmatriisi A kannassa $(v_1, \dots, v_n)$ on yläkolmiomatriisi. Tällöin jokaisella $j \in \{1, \dots, n\}$ pätee

$$f(v_j) = f(\Psi_{(v_1,\dots,v_n)}(e_j)) = f(v_j) = \Psi_{(v_1,\dots,v_n)}(Ae_j).$$

Koska A on yläkolmiomatriisi, niin $Ae_j \in \mathrm{Sp}(e_1, \dots, e_j)$. Näin ollen $f(v_j) \in \mathrm{Sp}(v_1, \dots, v_j)$.

Oletetaan nyt, että $f(v_j) \in \mathrm{Sp}\{v_1, \dots, v_j\}$ kaikilla $j \in \{1, \dots, n\}$. Olkoon $v \in \mathrm{Sp}(v_1, \dots, v_j)$. Tällöin on olemassa sellaiset luvut $a_1, \dots, a_j \in \mathbb{C}$, että $v = a_1v_1 + \dots + a_jv_j$. Näin ollen $f(v) = a_1f(v_1) + \dots + a_jf(v_j) \in \mathrm{Sp}(v_1, \dots, v_j)$. Aliavaruus $\mathrm{Sp}(v_1, \dots, v_j)$ on siis operaattorin f invariantti aliavaruus jokaisella $j \in \{1, \dots, n\}$.

Oletetaan nyt, että $\mathrm{Sp}(v_1, \dots, v_j)$ on operaattorin f invariantti aliavaruus jokaisella $j \in \{1, \dots, n\}$. Olkoon $A \in \mathbb{C}^{n\times n}$ operaattorin f esitysmatriisi kannassa $(v_1, \dots, v_j)$. Koska

$$\Psi_{(v_1,\dots,v_j)}(\varphi_A(e_j)) = f(\Psi_{(v_1,\dots,v_j)}(e_j)) = f(v_j) \in \mathrm{Sp}(v_1, \dots, v_j)$$

ja $\mathrm{Sp}(v_1, \dots, v_j) = \Psi_{(v_1,\dots,v_n)}(\mathrm{Sp}(e_1, \dots, e_j))$, niin $Ae_j = \varphi_A(e_j) \in \mathrm{Sp}(e_1, \dots, e_j)$. Matriisi A on siis yläkolmiomatriisi. $\square$

Lauseen E.2.1 todistuksen ytimessä on seuraava huomio. Väitteessä käytetään termiä *avaruuden V aito aliavaruus*, jolla tarkoitetaan aliavaruutta $U \subset V$, joka ei ole koko avaruus V. Seuraava lemma pätee jälleen sekä reaalisille että kompleksille vektoriavaruuksille.

Lemma E.2.5. *Olkoon V vektoriavaruus, $f\colon V \to V$ operaattori, jolla on ominaisarvo λ. Tällöin lineaarikuvauksen $f_\lambda\colon V \to V$, $v \mapsto f(v) - \lambda v$, kuva $U_\lambda = \operatorname{im} f_\lambda$ on kuvauksen f invariantti aliavaruus. Lisäksi, jos V on äärellisulotteinen, niin $\operatorname{im} f_\lambda$ on aito aliavaruus.*

Todistus. Olkoon $v \in U_\lambda$. Tällöin

$$f(v) = f(v) - \lambda v + \lambda v = f_\lambda(v) + \lambda v.$$

Koska U_λ on aliavaruus ja $v \in U_\lambda$, niin $\lambda v \in U_\lambda$. Koska $f_\lambda(v) \in U_\lambda$ ja $\lambda v \in U_\lambda$, niin $f(v) \in U_\lambda$. Näin ollen $f(\operatorname{im} f_\lambda) \subset U_\lambda$.

Oleteaan nyt, että V on äärellisulotteinen. Koska $f_\lambda(v) = f(v) - \lambda v = 0$ jokaisella $v \in E(\lambda, f)$, niin $E(\lambda, f) \subset \ker f_\lambda$. Näin ollen $\dim \ker f_\lambda > 0$. Lauseen 13.4.3 perusteella $\dim U_\lambda = \dim \operatorname{im} f_\lambda < \dim V$. Näin ollen U_λ on avaruuden V aito aliavaruus. □

Huomautus E.2.6. *Tämä lemma voidaan tulkita matriiseilla seuraavasti. Olkoon $A \in \mathbb{C}^{n\times n}$ neliömatriisi ja $\lambda \in \mathbb{C}$ sen ominaisarvo. Tällöin jokaisella $z \in \mathbb{C}^{n\times 1}$ pätee*

$$A(A - \lambda I)z = A^2 z - \lambda A z = (A - \lambda I)Az.$$

eli $A(\mathrm{Col}(A - \lambda I)) \subset \mathrm{Col}(A - \lambda I)$.

Huomautus E.2.7. *Avaruuden V äärellisulotteisuus on välttämätön oletus lemman E.2.5 jälkimmäisessä väitteessä. Derivointioperaattori $D\colon \mathscr{P} \to \mathscr{P}$ on esimerkki operaattorista, jolle väitteen jälkimmäinen osa ei päde.*

Lauseen E.2.1 todistus. Todistetaan väite induktiolla avaruuden V dimension suhteen. Selvästi väite pätee tapauksessa $n = 1$, sillä yksiulotteisen kompleksisen vektoriavaruuden V operaattorilla $f\colon V \to V$ on aina ominaisarvo.

Olkoon $n \in \mathbb{N}$. Oletetaan, että väite pätee kaikilla kompleksisilla vektoriavaruuksilla, joiden dimensio on alle n. Olkoon V n-ulotteinen kompleksinen vektoriavaruus ja $f\colon V \to V$ operaattori. Lauseen 14.1.1 nojalla operaattorilla f on ominaisarvo $\lambda \in \mathbb{C}$.

Olkoon nyt $f_\lambda\colon V \to V$ operaattori $v \mapsto f(v) - \lambda v$, kuten lemmassa E.2.5. Tällöin $U = \operatorname{im} f_\lambda$ on operaattorin f invariantti aliavaruus. Lisäksi U on avaruuden V aito aliavaruus eli $\dim U < \dim V$. Olkoon $k = \dim U$.

Koska $fU \subset U$, niin rajoittumakuvaus $f|_U\colon U \to U$ on hyvin määrilty, eli rajoittuma $f|_U$ on avaruuden U operaattori. Koska $\dim U < \dim V = n$, niin induktio-oletuksen nojalla on olemassa sellainen avaruuden U kanta $(u_1, \ldots, u_k)$, missä $k = \dim U$, jossa operaattorin $f|_U$ esitysmatriisi on yläkolmiomatriisi.

Olkoon nyt $(v_1, \ldots, v_n)$ sellainen avaruuden V kanta, että $v_j = u_j$ kaikilla $j \leq k$. Osoitetaan, että operaattorin f esitysmatriisi tässä kannassa on yläkolmiomatriisi. Lemman E.2.4 nojalla riittää osoittaa, että $f(v_j) \in \mathrm{Sp}\{v_1, \ldots, v_j\}$ jokaisella $1 \leq j \leq n$.

Oletetaan ensin, että $1 \leq j \leq k$. Koska operaattorin $f|_U : U \to U$ esitysmatriisi kannassa $(u_1, \ldots, u_k)$ on yläkolmiomatriisi, niin lemman E.2.4 nojalla jokaisella $j \in \{1, \ldots, k\}$ pätee

$$f(v_j) = f(u_j) = (f|_U)(u_j) \in \mathrm{Sp}(u_1, \ldots, u_j) = \mathrm{Sp}(v_1, \ldots, v_j).$$

Oletetaan nyt, että $k < j \leq n$. Tällöin

$$f(v_j) = (f - \lambda \mathrm{id}_V)(v_j) + \lambda v_j = f_\lambda(v_j) + \lambda v_j.$$

Koska $f_\lambda(v_j) \in U = \mathrm{Sp}(v_1, \ldots, v_k)$, niin

$$f(v_j) \in \mathrm{Sp}(v_1, \ldots, v_k, v_j) \subset \mathrm{Sp}\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_j\}.$$

Operaattorilla f on siis yläkolmiomatriisi esitys kannassa $(v_1, \ldots, v_n)$. $\square$

Liite F

Liite: Semidefiniittisyyden lisäsivut

F.1 Yleisten vektoriavaruuksien sisätulot ovat pistetuloja

Tarkastellaan nyt millaisia ovat yleisien vektoriavaruuksien sisätulot. Lauseen 9.3.7 perusteella tiedetään, että jokainen äärellisulotteinen vektoriavaruus on isomorfinen sarakeavaruuden kanssa. Tästä seuraa, että jokainen äärellisulotteisen vektoriavaruuden sisätulo vastaa jotain sarakeavaruuden sisätuloa, eli positiividefiniittiä neliömatriisia. Tulkinta siis on, että äärellisulotteisten vektoriavaruuksien sisätulot eivät ole yleisesti kummempia kuin edellisessä luvussa esitetyt matriisit. Tähän huomioon palataan vielä myöhemmin, mutta tehdään nyt tuloksen ensimmäinen versio.

Lause F.1.1. *Olkoon $(V, \langle \cdot, \cdot \rangle)$ sisätuloavaruus ja $\Phi\colon \mathbb{R}^{n\times 1} \to V$ isomorfismi. Tällöin on olemassa sellainen positiivisesti definiitti matriisi $A \in \mathbb{R}^{n\times n}$, jolle pätee*

$$\langle v, w \rangle = \Phi^{-1}(v) \cdot A\Phi^{-1}(w).$$

kaikilla $v, w \in \mathbb{R}^{n\times 1}$.

Todistus perustuu havaintoon, että isomorfismin Φ avulla voidaan avaruuden V sisätulo $\langle \cdot, \cdot \rangle$ siirtää sarakeavaruuteen $\mathbb{R}^{n\times 1}$, eli määritellä sarakeavaruuteen uusi sisätulo $\langle \cdot, \cdot \rangle_\Phi$ valitun isomorfismin Φ avulla. Muotoillaan tämä lemmaksi.

Lemma F.1.2. *Olkoon $(V, \langle \cdot, \cdot \rangle)$ sisätuloavaruus ja $\Phi\colon \mathbb{R}^{n\times 1} \to V$ isomorfismi. Tällöin funktio $\langle \cdot, \cdot \rangle_\Phi\colon \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1} \to \mathbb{R}$, joka on määritelty kaavalla*

$$\langle x, y \rangle_\Phi = \langle \Phi(x), \Phi(y) \rangle$$

kaikilla $x, y, \in \mathbb{R}^{n\times 1}$, on sisätulo sarakeavaruudessa $\mathbb{R}^{n\times 1}$.

Lauseen F.1.1 todistus. Olkoot $x, x', y \in \mathbb{R}^{n\times 1}$ ja $a \in \mathbb{R}$. Tällöin

$$\begin{aligned}\langle ax + x', y\rangle_\Phi &= \langle \Phi(ax + x'), \Phi(y)\rangle \\ &= \langle a\Phi(x) + \Phi(x'), \Phi(y)\rangle \\ &= a\langle \Phi(x), \Phi(y)\rangle + \langle \Phi(x'), Phi(y)\rangle \\ &= a\langle x, y\rangle_\Phi + \langle x', y\rangle_\Phi.\end{aligned}$$

Vastaavasti osoitetaan, että $\langle x, y\rangle_\Phi = \langle y, x\rangle_\Phi$ ja että $\langle x, x\rangle_\Phi \geq 0$.

Osoitetaan nyt, että $\langle x, x\rangle_W = 0$, jos ja vain jos $x = 0$. Jos $x = 0$, niin $\langle x, x\rangle_W = \langle \Phi(x), \Phi(x)\rangle = \langle 0, 0\rangle = 0$. Toisaalta, jos $x \neq 0$, niin $\Phi(x) \neq 0$, koska Φ on isomorfismi. Näin ollen $\langle x, x\rangle_W = \langle \Phi(x), \Phi(x)\rangle > 0$, koska $\langle \cdot, \cdot\rangle$ on sisätulo. Tämä päättää todistuksen. □

Huomautus F.1.3. *Huomio jonka voi yllä olevasta laskusta tehdä on, että todistus on olennaisesti sama kuin edellisen luvun todistukset matriiseille. Todistuksessa tarvitaan isomorfismin lineaarisuutta ja sitä, että* $\ker \Phi = \{0\}$.

Todistetaan nyt lause F.1.1.

Lauseen F.1.1 todistus. Lemman F.1.1 nojalla funktio $\langle \cdot, \cdot\rangle_\Phi \colon \mathbb{R}^{n\times 1} \times \mathbb{R}^{n\times 1} \to \mathbb{R}$, joka on määritelty kaavalla

$$\langle x, y\rangle_\Phi = \langle \Phi(x), \Phi(y)\rangle$$

kaikilla $x, y \in \mathbb{R}^{n\times 1}$, on sisätulo. Näin ollen lauseen 16.2.1 perusteella on olemassa sellainen positiivisesti definiitti matriisi $A \in \mathbb{R}^{n\times n}$, että

$$\langle x, y\rangle_\Phi = x \cdot Ay$$

kaikilla $x, y \in \mathbb{R}^{n\times 1}$.

Olkoot nyt $v, w \in V$. Tällöin $v = \Phi(\Phi^{-1}(v))$ ja $w = \Phi(\Phi^{-1}(w))$, joten

$$\langle v, w\rangle = \langle \Phi(\Phi^{-1}(v)), \Phi(\Phi^{-1}(w))\rangle = \langle \Phi^{-1}(v), \Phi^{-1}(w)\rangle_\Phi = \Phi^{-1}(v) \cdot A\Phi^{-1}(w).$$

□

F.2 Neliömuodot ja toisen asteen käyrät

Symmetristen neliömatriisien luokittelussa esiintyvää funktiota $q_A \colon \mathbb{R}^{n\times 1} \to \mathbb{R}$,

$$x \mapsto x^t Ax,$$

missä $A \in \mathbb{R}^{n\times n}$ on siis symmetrinen neliömatriisi, kutsutaan matriisin A *neliömuodoksi* (engl. *quadratic form*). Näin ollen neliömetriisien luokittulu (semi)definiitteihin ja indefiniitteihin neliömatriiseihin on itseasiassa näiden neliömuotojen luokittelu.

Neliömuodoilla on tärkeä rooli klassisessa geometriassa, esimerkiksi kartioleikkausten teoriassa, mutta myös pääakselien teoriassa. Raapaistaan nyt hieman pintaa tästä asiasta.

Ennen varsinaista esimerkkiä, palautetaan mieleen neliömatriisien yhteys sisätuloihin.

Esimerkki F.2.1. *Olkoon $A \in \mathbb{R}^{n\times n}$ positiivisesti definiitti matriisi. Tällöin funktio $\langle\cdot,\cdot\rangle\colon \mathbb{R}^{n\times 1}\times\mathbb{R}^{n\times 1}\to\mathbb{R}$, joka on määritelty kaavalla*

$$\langle x, y\rangle = x^t A y$$

kaikilla $x, y \in \mathbb{R}^{n\times 1}$, on sisätulo avaruudessa $\mathbb{R}^{n\times 1}$. (Harjoitustehtävä) Lisäksi tämän sisätulon normille $\|\cdot\|\colon \mathbb{R}^{n\times 1}\to\mathbb{R}$ pätee

$$\|x\| = \sqrt{x^t A x} = \sqrt{q_A(x)}.$$

Toisaalta, jos $\langle\cdot,\cdot\rangle\colon \mathbb{R}^{n\times 1}\to\mathbb{R}$ on sisätulo, niin on olemassa sellainen symmetrinen positiivisesti definiitti matriisi $A \in \mathbb{R}^{n\times n}$, että $x^t A y = \langle x, y\rangle$ kaikilla $x, y \in \mathbb{R}^{n\times 1}$. (Harjoitustehtävä)

Tarkastellaan nyt yhteyttä toisen asteen käyriin ja aloitetaan koordinaattiakselien suuntaisista ellipseistä.

Esimerkki F.2.2. .

Olkoot $a_1, \dots, a_n > 0$ reaalilukuja ja olkoon

$$E = \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon \frac{x_1^2}{a_1} + \dots + \frac{x_n^2}{a_n} \le 1\}$$

(Piirrä tapaus $n = 2$ ja $a_1 = a_2 = 1$ sekä tapaus $n = 2$ ja $a_1 = 1$ sekä $a_2 = 2$.)

Olkoon $A \in \mathbb{R}^{n\times n}$ diagonaalimatriisi

$$A = \begin{bmatrix} \frac{1}{a_1} & & \\ & \ddots & \\ & & \frac{1}{a_n} \end{bmatrix}.$$

Tällöin jokaisella $x = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^t \in \mathbb{R}^{n\times 1}$ pätee

$$q_A(x) = x^t A x = [x_1 \cdots x_n] \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \frac{x_1^2}{a_1} + \dots + \frac{x_n^2}{a_n}.$$

Näin ollen

$$E = \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon q_A(\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^t) \le 1\}.$$

Huomautus F.2.3. *Huomaa, että edellisessä esimerkissä, matriisi A määrittelee sisätulon, joten joukko E voidaan kirjoittaa myös tätä sisätuloa vastaavan normin $\|\cdot\|$ avulla muodossa*

$$E = \{(x_1, \dots, x_n) \in \mathbb{R}^n \colon \|[x_1 \cdots x_n]^t\| = \sqrt{q_A([x_1 \cdots x_n]^t)} \le 1\}.$$

Joukko E on siis kaikkien niiden pisteiden $x \in \mathbb{R}^n$ joukko, joiden etäisyys origosta on alle 1 normin $(x_1, \dots, x_n) \mapsto \sqrt{x^t A x}$ määräämässä metriikassa. (Tässä vektori $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ on samastettu sarakevektorin $x = [x_1, \dots, x_n]^t$ kanssa.)

Tarkastellaan nyt samaa ilmiötä toisinpäin.

Esimerkki F.2.4. *Olkoon $A \in \mathbb{R}^{n\times n}$ symmetrinen positiivisesti definiitti neliömatriisi ja tarkastellaan joukkoa*

$$E_A = \{x \in \mathbb{R}^{n\times 1} \colon q_A(x) \leq 1\}.$$

Spektraalilauseen nojalla on olemassa sellainen ortonormaalimatriisi $P \in \mathbb{R}^{n\times n}$ ja diagonaalimatriisi $D \in \mathbb{R}^{n\times n}$, että $PDP^t = A$. Tarkastellaan nyt neliömuotoa q_A ja esitetään se neliömuodon q_D avulla.

Olkoon $x \in \mathbb{R}^{n\times n}$. Tällöin

$$q_A(x) = x^t PDP^t x = (P^t x)^t D(P^t x) = q_D(P^t x).$$

Näin ollen

$$\begin{aligned} E_A &= \{x \in \mathbb{R}^{n\times 1} \colon q_D(P^t x) \leq 1\} \\ &= \{x \in \mathbb{R}^{n\times 1} \colon q_D(P^{-1} x) \leq 1\} \\ &= \{Py \in \mathbb{R}^{n\times 1} \colon y \in \mathbb{R}^{n\times 1},\ q_D(y) \leq 1\}. \end{aligned}$$

Määritellään nyt

$$E_D = \{y \in \mathbb{R}^{n\times 1} \colon q_D(y) \leq 1\}.$$

Nyt

$$E_A = PE_D.$$

eli joukko E_A on joukon E_D kuva kuvauksessa $f_P \colon y \mapsto Py$.

Koska D on diagonaalimatriisi, niin saadaan, että

$$q_D([y_1 \cdots y_n]^t) = d_1 y_1^2 + \cdots + d_n y_n^2$$

missä $d_1, \ldots, d_n > 0$ ovat matriisin D diagonaalialkiot, eli

$$D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}.$$

Näin ollen E_D on ellipsi avaruudessa $\mathbb{R}^{n\times 1}$, joka edellisen esimerkin merkinnöin vastaa ellipsiä

$$\{(y_1, \ldots, y_n) \in \mathbb{R}^n \colon \frac{y_1^2}{1/d_1} + \cdots + \frac{y_n^2}{1/d_n} \leq 1\}.$$

Koska matriisi P on ortogonaalimatriisi, niin kuvaus f_P on isometria, eli vektoreilla y ja Py on sama pituus pistetulon määräämässä metriikassa kaikilla $y \in \mathbb{R}^{n\times 1}$. Näin ollen myös joukko E_A on ellipsi. Siinä missä E_D on ellipsi, jonka akselit ovat koordinaattiakselien suuntaisia, ellipsin E_A akselit ovat matriisin P sarakevektoreiden suuntaisia. Näitä suuntia kutsutaan ellipsin E_A pääakseleiksi.

Määritelmä F.2.5. *Olkoon $A \in \mathbb{R}^{n\times n}$ symmetrinen matriisi ja olkoot $P \in \mathbb{R}^{n\times n}$ sellainen ortogonaalimatriisi ja $D \in \mathbb{R}^{n\times n}$ sellainen diagonaalimatriisi, että $A = PDP^{-1}$. Neliömuodon q_A* pääekseliesitys *on neliömuoto $q_D : \mathbb{R}^{n\times 1} \to \mathbb{R}$. Lisäksi matriisin P sarakkeita kutsutaan neliömuodon q_A* pääakseleiksi.

Huomaa, että yllä esitetty määritelmä ei rajoitu positiivisesti definiitteihin neliömatriiseihin vaan on yleinen symmetristen matriisien määritelmä. Samat tarkastelut jotka tehtiin yllä positiivisesti definiiteille matriiseille on mahdollista tehdä yleisille symmetrisille matriiseille, sillä erotuksella, että yhteyttä sisätuloon ei tässä tapauksessa enää ole. Tarkastellaan kahta esimerkkiä.

Esimerkki F.2.6. *Olkoot $a, b > 0$ ja*

$$H = \{(x_1, x_2) \in \mathbb{R}^2 : \frac{x_1^2}{a} - \frac{x_2^2}{b} = 1\}$$

Tällöin H on hyperbeli avaruudessa $\mathbb{R}^2$.

Olkoon

$$A = \begin{bmatrix} \frac{1}{a} & 0 \\ 0 & -\frac{1}{b} \end{bmatrix}.$$

Tällöin

$$H = \{(x_1, x_2) \in \mathbb{R}^2 : q_A([x_1 x_2]^t) = 1\}.$$

Esimerkki F.2.7. *Olkoon*

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Tällöin $x \mapsto x^t M x$ on neliömuoto

$$[x_1 \cdots x_4] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1^2 + x_2^2 + x_3^2 - x_4^2.$$

Matriisi M ei (selvästi) määrittele sisätuloa, mutta määrittelee niin sanotun Minkowskin sisätulon *avaruuteen $\mathbb{R}^4$. Minkowskin sisätuloa käytetään yleisessä suhteellisuusteoriassa. Huomaa, että toisin kuin edellisissä esimerkeissä, tässä tapauksessa*

$$\{x \in \mathbb{R}^{r\times 1} : q_M(x) = 0\}$$

ei ole piste, vaan paraboloidi avaruudessa $\mathbb{R}^{4\times 1}$.

F.3 Useamman muuttujan toisen asteen polynomiyhtälön ratkaiseminen neliömuotojen avulla

Matriisien neliömuotojen hyödyllisyys seuraa havainnosta, että toisen asteen polynomi yhtälöt voidaan ratkaista neliömuotojen avulla. Tarkastellaan pääakseleiden sovelluksena yleistä toisen asteen polynomia $p\colon \mathbb{R}^n \to \mathbb{R}$ ja yhtälön

$$p(x) = 0 \tag{F.1}$$

ratkaisuja eli ratkaisujoukkoa

$$R_p = \{x \in \mathbb{R}^n \colon p(x) = 0\}.$$

Huomautus F.3.1. *Tason $\mathbb{R}^2$ erikoistapauksessa tämä vastaa toisen asteen yhtälön*

$$c_{11}x_1^2 + c_{12}x_1x_2 + c_{22}x_2^2 + b_1x_1 + b_2x_2 + a = 0,$$

missä $c_{11}, c_{12}, c_{22}, b_1, b_2, a \in \mathbb{R}$, ratkaisujen etsimistä.

Yleinen tulos jakautuu tapauksiin, jotka riippuvat polynomista p.
Tarkastellaan tämän vuoksi ongelmaa neljässä vaiheessa:

- Ensimmäisessä yhtälö $p(x) = 0$ kirjoitetaan neliömuodon ja sisätulon avulla.
- Toisessa neliömuoto kirjoitetaan pääakselimuodossa.
- Kolmannessa vaiheessa sisätuloon liittyvät termit yhdistetään neliömuodon termeihin, mikäli neliömuodon pääakseliesitys sen mahdollistaa. Tämä vaihe on neliöksitäydentämisen korkeampiulotteinen vastine.
- Jäljelle jäävät tapaukset käsitellään esimerkin muodossa ja jätetään harjoitustehtäviksi.

Ensimmäinen vaihe on siis polynomin tulkitseminen neliömuodon ja sisätulon avulla.

Lemma F.3.2. *Olkoon $p\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$ toisen asteen polynomi. Tällöin on olemassa sellainen symmetrinen neliömatriisi $A \in \mathbb{R}^{n\times n}$, sellainen vektori $\mathbb{R}^{n\times 1}$ ja sellainen luku $a \in \mathbb{R}$, että*

$$p(x) = q_A(x) + b \cdot x + a.$$

Todistus. Olkoon

$$p(x_1, \dots, x_n) = \sum_{1\le i\le j} c_{ij}x_ix_j + \sum_{i=1}^{n} b_ix_i + a$$

tarkasteltava polynomi. Olkoon $q\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$ neliömuoto

$$(x_1, \dots, x_n) = \sum_{1\le i\le j} c_{ij}x_ix_j.$$

Olkoon nyt $A = [a_{ji}] \in \mathbb{R}^{n\times n}$ matriisi, jolle pätee $a_{ji} = (c_{ji} + c_{ij})/2$ kaikilla $j, i \in \{1, \ldots, n\}$. Tällöin A on symmetrinen ja $q = q_A$. Olkoon lisäksi $b = [b_1 \cdots b_n]^t \in \mathbb{R}^{n\times 1}$. Tällöin

$$p(x) = q(x) + b \cdot x + a = q_A(x) + b \cdot x + a$$

jokaisella $x \in \mathbb{R}^{n\times 1}$. □

Toinen vaihe on yleisen neliömuodon q_A pääakseliesitykseen siirtyminen. Tämä antaa meille matriisin P. Huomaa, että seuraavan lemman todistus on oleellisesti tehty esimerkissä F.2.2.

Lemma F.3.3. *Olkoon $p\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$ toisen asteen polynomi, joka on muotoa*

$$p(x) = q_A(x) + b \cdot x + a$$

jollakin symmetrisellä matriisilla $A \in \mathbb{R}^{n\times n}$, jollakin vektorilla $b \in \mathbb{R}^{n\times 1}$ ja jollaikin $a \in \mathbb{R}$. Tällöin on olemassa sellainen ortogonaalimatriisi $P \in \mathbb{R}^{n\times n}$ ja diagonaalimatriisi $D \in \mathbb{R}^{n\times n}$, että

$$p(Py) = q_D(y) + (P^t b) \cdot y + a.$$

kaikilla $y \in \mathbb{R}^{n\times 1}$.

Todistus. Olkoon $y \in \mathbb{R}^{n\times 1}$. Tällöin

$$\begin{aligned} p(Py) &= q_A(Py) + b \cdot Py + a \\ &= (Py)^t A(Py) + b^t Py + a \\ &= y^t P^t P D P^t P y + (P^t b)^t y + a \\ &= y^t D y + (P^t b) \cdot y + a = q_D(y) + (P^t b) \cdot y + a. \end{aligned}$$

□

Edellisen lemman varsinainen hyöty paljastuu seuraavasta lemmasta, joka vastaa neliöksi täydentämistä. Huomaa, että neliöksi ei voi täydentää sellaisten muuttujien y_i suhteen, joiden toisen potenssin kerroin on nolla.

Lemma F.3.4. *Olkoon $D \in \mathbb{R}^{n\times n}$ diagonaalimatriisi, jonka diagonaalialkiot ovat $d_1, \ldots, d_n \in \mathbb{R}$ ovat kaikki nollasta poikkeavia ja olkoon $b = [b_1 \cdots b_n]^t \in \mathbb{R}^{n\times 1}$. Tällöin on olemassa sellainen vektori $\tilde{b} \in \mathbb{R}$ ja vakio $\tilde{a} \in \mathbb{R}$, että*

$$q_D(y) + b \cdot y = q_D(y - \tilde{b}) + \tilde{a}$$

kaikilla $y \in \mathbb{R}^{n\times 1}$.

Todistus. Olkoon $y = [y_1 \cdots y_n]^t \in \mathbb{R}^{n\times 1}$. Koska $d_i \neq 0$ kaikilla $i \in \{1, \ldots, n\}$, niin

$$\begin{aligned}
q_D(y) + b\cdot y &= d_1y_1^2 + \cdots + d_ny_n^2 + b_1y_1 + \cdots + b_ny_n \\
&= (d_1y_1^2 + b_1y_1) + \cdots + (d_ny_n^2 + b_ny_n) \\
&= d_1\left(y_1^2 + \frac{b_1}{d_1}y_1\right) + \cdots + d_n\left(y_n^2 + \frac{b_n}{d_n}y_n\right) \\
&= d_1\left(\left(y_1 + \frac{b_1}{2d_1}\right)^2 - \left(\frac{b_1}{2d_1}\right)^2\right) + \cdots + d_n\left(\left(y_n + \frac{b_n}{2d_n}\right)^2 - \left(\frac{b_n}{2d_n}\right)^2\right) \\
&= d_1\left(y_1 + \frac{b_1}{2d_1}\right)^2 + \cdots + d_n\left(y_n + \frac{b_n}{2d_n}\right)^2 - \left(d_1\left(\frac{b_1}{2d_1}\right)^2 + \cdots + d_n\left(\frac{b_n}{2d_n}\right)^2\right)
\end{aligned}$$

Merkitään $\tilde{b}_i = b_i/(2d_i)$ jokaisella $i \in \{1, \ldots, n\}$ ja $\tilde{b} = [\tilde{b}_1 \cdots \tilde{b}_n]^t \in \mathbb{R}^{n\times 1}$. Olkoon myös

$$\tilde{a} = -\left(d_1\left(\frac{b_1}{2d_1}\right)^2 + \cdots + d_n\left(\frac{b_n}{2d_n}\right)^2\right).$$

Tällöin

$$q_D(y) + b\cdot y = q_D(y + \tilde{b}) + \tilde{a}.$$

□

Kirjataan nyt yleinen tulos, jonka nämä lemmat yhdessä todistavat.

Lause F.3.5. *Olkoon $p\colon \mathbb{R}^{n\times 1} \to \mathbb{R}$ toisen asteen polynomi. Mikäli*

$$p(x) = q_A(x) + b\cdot y + a$$

jollain neliömatriisilla $A \in \mathbb{R}^{n\times n}$, vektorilla $b \in \mathbb{R}^{n\times 1}$ ja luvulla $a \in \mathbb{R}$ ja matriisin A spektraalihajotelma $A = PDP^t$ on sellainen, että matriisin D diagonaalialkiot $d_1, \ldots, d_n$ ovat kaikki nollasta poikkeavia, niin

$$p(x) = 0$$

jos ja vain jos vektorille $y = [y_1 \cdots y_n]^t = Px$ pätee

$$d_1\left(y_1 - \frac{b_1}{2d_1}\right)^2 + \cdots + d_n\left(y_n - \frac{b_1}{2d_1}\right)^2 + a = \sum_{j=1}^{n}\left(\frac{b_i}{2d_i}\right)^2.$$

Huomautus F.3.6. *Lauseen sanoma on, että yleinen toisen asteen polynomiyhtälö $p(x) = 0$ avaruudessa $\mathbb{R}^n$ voidaan tässä tapauksessa saattaa avaruuden $\mathbb{R}^n$ ortogonaalikuvauksella, jota edustaa matriisi P, yhtälöksi, jossa ei ole toisen kertaluvun ristitermejä, eli muotoon*

$$d_1(y_1 - c_1)^2 + \cdots + d_n(y_n - c_n)^2 + a = e.$$

Yhden muuttujan tapauksessa tämä vastaa yhtälöä

$$d(y-c)^2 + a = 0 \quad eli \quad (y-c)^2 = \frac{e-a}{d},$$

jonka ratkaisut ovat

$$y = c \mp \sqrt{\frac{e-a}{d}},$$

jos $(e-a)/d \geq 0$.

Lukijaa voi tässä vaiheessa alkaa vaivaamaan, että mitä tapahtuu, jos spektraalihajotelmassa matriisi D sisältää nollia, eli että nolla on matriisin A ominaisrvo. Tällöin polynomin p ratkaisujoukko on paraabeli (tai sen yleistys). Käsitellään tämä yhden esimerkin valossa.

Esimerkki F.3.7. *Olkoon* $p\colon \mathbb{R}^{3\times 1} \to \mathbb{R}$ *neliömuoto*

$$p([x_1 x_2 x_3]^t) = x_1^2 - x_2^2 - x_3$$

kaikilla $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^t \in \mathbb{R}^{3\times 1}$.

Olkoot

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

ja $b = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^t \in \mathbb{R}^{3\times 1}$. *Tällöin*

$$p(x) = q_A(x) + b \cdot x.$$

Koska A *on jo diagonaalinen, niin spektraalihajotelmassa* $P = I$ *ja* $D = A$.

Nyt yhtälö

$$p(x) = 0$$

vastaa yhtälöä

$$x_3 = x_1^2 - x_2^2.$$

avaruudessa $\mathbb{R}^3$. *(Piirrä kuva.)*

F.4 Choleskyn hajotelman löytäminen algoritmisesti

Tässä liitteessä käsitellään algoritminen menetelmä positiivisesti semidefiniitin matriisin Choleskyn hajotelman löytämiselle.

Todistetaan ratkaisualgoritmia varten aputulos.

Lemma F.4.1. *Olkoon*

$$A = \begin{bmatrix} a_{11} & b^t \\ b & \tilde{A} \end{bmatrix} \in \mathbb{R}^{n\times n},$$

positiivisesti semidefiniitti matriisi, missä $\tilde{A} \in \mathbb{R}^{(n-1)\times(n-1)}$ *on symmetrinen matriisi ja* $b \in \mathbb{R}^{(n-1)\times 1}$ *on sarakevektori. Jos* $a_{11} = 0$, *niin* $b = 0$. *Jos* $a_{11} \neq 0$, *niin matriisi* $\tilde{A} - \frac{1}{a_{11}} bb^t$ *on positiviisesti semidefiniitti.*

Todistus. Tarkastellaan ensin tapausta $a_{11} = 0$. Osoitetaan väite induktiolla dimension n suhteen. Aloitetaan ensimmäisestä epätriviaalista tapauksesta $n = 2$. Huomaa, että tapauksessa $n = 1$, vektori b on ns. tyhjä sarake.

Koska $a_{11} = 0$, niin

$$A = \begin{bmatrix} 0 & b \\ b & d \end{bmatrix},$$

missä $b, d \in \mathbb{R}$. Koska A on positiivisesti semidefiniitti, niin jokaisella $x \in \mathbb{R}$ pätee

$$0 \leq \begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} 0 & b \\ b & d \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} b & xb + d \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = xb + xb + d = 2xb + d.$$

Näin ollen $b = 0$.

Oletetaan nyt, että väite pätee kaikille positiivisesti semidefiniiteille $(n-1) \times (n-1)$-matriiseille. Olkoon nyt A positiivisesti semidefiniitti $n \times n$-matriisi

$$A = \begin{bmatrix} 0 & v^t & b' \\ v & \tilde{A}' & w^t \\ b' & w & d \end{bmatrix},$$

missä $v \in \mathbb{R}^{(n-2)\times 1}$ ja $w \in \mathbb{R}^{1\times(n-2)}$ ovat vektoreita, $b', d \in \mathbb{R}$ lukuja ja $\tilde{A}' \in \mathbb{R}^{(n-2)\times(n-2)}$ matriisi.

Koska matriisin A $(n-1)$:s pääminori

$$A_{n-1} = \begin{bmatrix} 0 & v^t \\ v & \tilde{A}' \end{bmatrix},$$

on positiivisesti semidefiniitti, niin induktio-oletuksen nojalla $v = 0$. Näin ollen

$$A = \begin{bmatrix} 0 & 0 & b' \\ 0 & \tilde{A}' & w^t \\ b' & w & d \end{bmatrix}.$$

Koska A on positiivisesti semidefiniitti, niin

$$0 \leq \begin{bmatrix} x & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & b' \\ 0 & \tilde{A}' & w^t \\ b' & w & d \end{bmatrix} \begin{bmatrix} x \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = 2xb' + d.$$

Näin ollen $b' = 0$ ja induktioaskel on todistettu.

Osoitetaan nyt, että tapauksessa $a_{11} \neq 0$ matriisi $\tilde{B} = \tilde{A} - \frac{1}{a_{11}} bb^t$ on positiivisesti semidefiniitti. Olkoon $x \in \mathbb{R}^{(n-1)\times 1}$ ja

$$y = \begin{bmatrix} -\frac{1}{a_{11}} b^t x \\ x \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

Koska A on positiivisesti semidefiniitti, niin

$$\begin{aligned}
0 \leq y^t A y &= \left[\begin{array}{c|c} -\frac{1}{a_{11}} b^t x & x^t \end{array}\right] \left[\begin{array}{cc} a_{11} & b^t \\ b & \tilde{A} \end{array}\right] \left[\begin{array}{c} -\frac{1}{a_{11}} b^t x \\ x \end{array}\right] \\
&= \left[\begin{array}{c|c} -b^t x + x^t b & -\frac{1}{a_{11}} b^t x b^t + x^t \tilde{A} \end{array}\right] \left[\begin{array}{c} -\frac{1}{a_{11}} b^t x \\ x \end{array}\right] \\
&= \frac{1}{a_{11}} (b^t x)^2 - \frac{1}{a_{11}} x^t b b^t x - \frac{1}{a_{11}} b^t x b^t x + x^t \tilde{A} x \\
&= x^t \tilde{A} x - \frac{1}{a_{11}} x^t (b b^t) x = x^t \tilde{B} x.
\end{aligned}$$

Väite on näin osoitettu. □

Choleskyn hajotelman iteratiivinen ratkaiseminen perustuu seuraavaan induktio todistukseen. Selvästi Choleskyn hajotelma voidaan löytää jokaiselle positiivisesti semidefiniitille 1×1-matriisille. Oletetaan, että Cholesky-hajotelma osataan löytää jokaiselle positiivisesti semidefiniitille $(n-1) \times (n-1)$-matriisille.

Olkoon

$$A = \left[\begin{array}{cc} a_{11} & b^t \\ b & \tilde{A} \end{array}\right] \in \mathbb{R}^{n \times n}$$

positiivisesti semidefiniitti matriisi, missä $\tilde{A} \in \mathbb{R}^{(n-1)\times(n-1)}$ on symmetrinen positiivisesti semidefiniitti matriisi ja $b \in \mathbb{R}^{(n-1)\times 1}$ on sarakevektori.

Olkoon nyt

$$T = \left[\begin{array}{cc} a & 0 \\ c & \tilde{T} \end{array}\right] \in \mathbb{R}^{n \times n}$$

alakolmiomatriisi, missä $\tilde{T} \in \mathbb{R}^{(n-1)\times(n-1)}$ on alakolmiomatriisi, jolla on ei-negatiivinen diagonaali, $c \in \mathbb{R}^{1\times(n-1)}$ on rivivektori ja $a \in \mathbb{R}$ on ei-negatiivinen reaaliluku.

Koska

$$TT^t = \left[\begin{array}{cc} a & 0 \\ c & \tilde{T} \end{array}\right] \left[\begin{array}{cc} a & c^t \\ 0 & \tilde{T}^t \end{array}\right] = \left[\begin{array}{cc} a^2 & ac^t \\ ac & cc^t + \tilde{T}\tilde{T}^t \end{array}\right],$$

niin matriisi T ratkaisee yhtälön

$$A = TT^t,$$

jos ja vain jos matriisi $\tilde{T}$, vektori c ja luku a toteuttavat yhtälöt

$$\begin{cases} a^2 &= a_{11} \\ ac &= b \\ cc^t + \tilde{T}\tilde{T}^t &= \tilde{A} \end{cases}$$

Käsitellään kaksi eri tapausta.

Tapaus $a_{11} = 0$. Tässä tapauksessa lemman F.4.1 nojalla pätee $b = 0$. Koska $\tilde{A}$ on positiivisesti definiitti $(n-1) \times (n-1)$-matriisi, niin on olemassa sellainen alakolmiomatriisi $\tilde{T}$, että $\tilde{T}\tilde{T}^t = \tilde{A}$. Näin ollen voidaan valita $a = 0$, $c = 0$ ja yhtälön $\tilde{A} = \tilde{T}\tilde{T}^t$ ratkaisu $\tilde{T}$.

Tapaus $a_{11} > 0$. Valitaan nyt $a = \sqrt{a_{11}}$ ja $c = \frac{1}{a}b$. Koska lemman F.4.1 perusteella matriisi $\tilde{A} - \frac{1}{a_{11}}bb^t$ on positiivisesti semidefiniitti, niin induktio-oletuksen nojalla on olemassa sellainen alakolmiomatriisi $\tilde{T}$, että $\tilde{A} - \frac{1}{a_{11}}bb^t = \tilde{T}\tilde{T}^t$.

Tämä päättää induktioaskeleen todistuksen.

Huomautus F.4.2. *Tässä iteratiivisessa ratkaisussa riittää siis ratkaista luku a ja vektori c matriisin A tapauksessa ja siirtyä sen jälkeen ratkaisemaan vastaavat luku ja vektori matriisille $\tilde{A}$ jne.*

Huomautus F.4.3. *Lukija on saattanut jo huomata, että olemme yllä antaneet konstruktiivisen todistuksen positiivisesti semidefiniitin matriisin Choleskyn hajotelman olemassaololle.*