Lineaarialgebra ja matriisilaskenta I–III

Pekka Pankka

22. marraskuuta 2020

Lukijalle

Lineaarialgebra ja matriisilaskenta I–III 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 opintojen kurssiin Lineaarialgebra ja matriisilaskenta III.

Kurssi Lineaarialgebra ja matriisilaskenta I sijoittuu matematiikan perusopintoiin ja sen kurssi Lineaarialgebra ja matriisilaskenta II matematiikan aineopintoihin. Kurssi Lineaarialgebra ja matriisilaskenta III on puolestaan sijoittunut tilastotieteen aineopintojen. 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 kattamaan ainoastaan kurssien Lineaarialgebra ja matriisilaskenta I ja II aiheita ja matriisin normaalimuotoihin liittyviä tuloksia käsitellään rajoitetusti. Kolmen periodin mittainen kokonaisuus mahdollistaa kuitenkin aihepiirin kattavan käsittelyn. Tämä on tärkeää erityisesti teorian hyödyntämistä ajatellen, sillä teorian tärkeät sovellukset perustuvat neliömatriisien normaalimuodoista kumpuaviin syvällisiin ominaisuuksiin.

Ensimmäinen kurssi Lineaarialgebra ja matriisilaskenta I on tässä esityksessä korostetusti matriisien teoriaan keskittyvä kurssi. Tästä on hyötyä sekä kurssin materiaalin konkreettisessa soveltamisessa että teoreettisesti, koska aliavaruuksiien teorialle saadaan luonnollinen käyttökohde matriisien nolla- ja sarakeavaruuksista, joka yhtyy yhtälöryhmien ratkaisemiseen. Aliavaruuksien yleinen dimensioteoria on tässä esityksessä siirretty seuraavalle kurssille Lineaarialgebra ja matriisilaskenta II.

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 myös tutustumista muihin lineaarialgebraa käsittelevien kirjoihin ja luentomuistiinpanoihin.

Helsingissä lokakuussa 2020
Pekka Pankka

Kirjallisuutta

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

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

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

[4] J. Möttönen. Lineaari algebra ja matriisilaskenta III. 2019.

[5] L. Oinonen and J. Rämö. Johdatus lineaarialgebraan – osa II. 2016.

Sisältö

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 normaalimuoto.
- 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

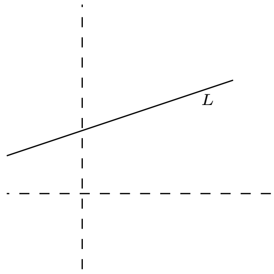
Kun puhekielessä puhutaan yhtälöstä $y = ax + b$, niin silloin tarkoitetaan niitä tason $\mathbb{R}^2$ pisteitä (x, y), jotka toteuttavat yhtälön $y = ax + b$ eli joukkoa

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

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.

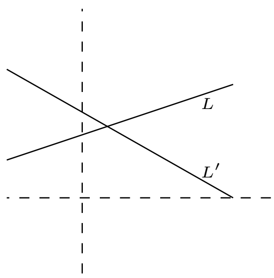
Tälläisissä tilanteissa tarkastellaan esimerkiksi kahta suoraa

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

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*

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

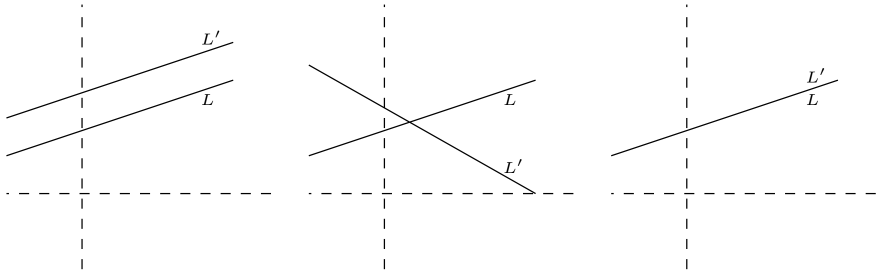
eli yhtälön $(5/2)x = 2$*. Tästä saadaan, että* $x = 4/5$ *ja* $y = 2x + 3 = 2(4/5) + 3 = 8/5 + 3 = 23/5$*.*

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 yhtälöillä. Tarkemmin tämä havainto on seuraava. Olkoon

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

[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.

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.

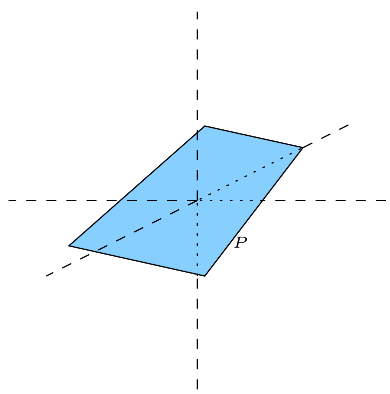
Avaruudessa $\mathbb{R}^3$ jokaisella pisteellä on kolme koordinaattia, joita tässä luvussa merkitään (x, y, z). Tarkastellaan nyt yhtälöä

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

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

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

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 koordinaattiakeselien suuntaiset tasot.

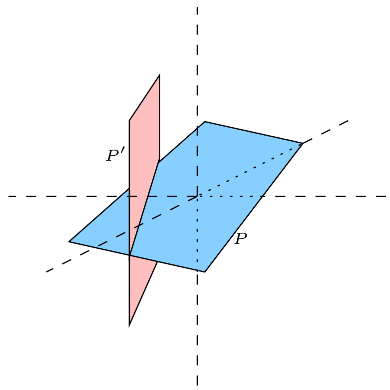
Tarkastellaan nyt kahden avaruuden $\mathbb{R}^3$ tason

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

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 *kutstutaan 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 $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, n\}$ sekä $b_j \in \mathbb{R}$ kaikilla $j \in \{1, \ldots, m\}$. Vektoria $(x_1, \ldots, x_n) \in \mathbb{R}^n$, 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*

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

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

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

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* $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$ *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 $0 = (0, \ldots, 0) \in \mathbb{R}^n$ 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). Mmääritelmässä 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 $a_1, \dots, a_n$ ovat nollia, niin tällöin yhtälöllä (1.10) ei ole ratkaisuja, eli

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

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 $a_1, \dots, a_n \in \mathbb{R}$ 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 $a_1, \dots, a_n \in \mathbb{R}$ 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

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

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

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

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

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

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)**.**

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

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

$$\sim \begin{cases} x_1 & & +2x_3 & = -5 & \\ & & -4x_3 & = 63/2 & | \cdot (-1/4) \\ & x_2 & -x_3 & = 5/2 & \end{cases}$$

$$\underset{R_2 \leftrightarrow R_3}{\sim} \begin{cases} x_1 & & +2x_3 & = & -5 & | -2R_3 \\ & x_2 & -x_3 & = & 5/2 & | +R_3 \\ & & x_3 & = & -63/8 & \end{cases}$$

$$\sim \begin{cases} x_1 & & & = & 86/8 \\ & x_2 & & = & -43/8 \\ & & x_3 & = & -63/8 \end{cases}$$

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 normaalimuodon, 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 $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, n\}$ 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 $p = 1, \ldots, n$, niin $a_{rp} = 0$ kaikilla $r > j$ ja $p \in \{1, \ldots, n\}$ 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 normaalimuodossa 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 silla ei ole ratkaisuja.*

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

Esimerkki 1.4.11. *Mikään yhtälöryhmistä*

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

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

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 todistuksen voi sivuuttaa tässä vaiheessa ja palata siihen luvun 3 jälkeen.

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*

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

vastaavassa supistetussa porrasmuodossa olevassa yhtälöryhmässä

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

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

Lauseen 1.4.12. Olkoon

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

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 rivioperaationde 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

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

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*

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

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. Todistukseen on luonnollista palata 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 normaalimuoto 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 $(x_1, , sx_n)$ 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*

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

Ratkaistaan esimerkin 1.4.5 yhtälöryhmä käyttäen tätä merkintää

Esimerkki 2.1.2. *Kirjoitetaan yhtälöryhmä*

$$\left\{\begin{array}{ccccl} x_1 & -x_2 & +2x_3 & -x_4 & = 3 \\ x_1 & +x_2 & & +x_4 & = 3 \end{array}\right.$$

muodossa

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

ja sovelletaan esimerkissä 1.4.4 käytettyjä merkintöjä ratkaisussa. Tällöin saadaan

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

Alkuperäinen yhtälöryhmä on siis ekvivalentti supistetussa porrasmuodossa olevan yhtälöryhmän

$$\left\{\begin{array}{ccccl} x_1 & & +x_3 & & = 3 \\ & x_2 & -x_3 & +x_4 & = 0 \end{array}\right.$$

kanssa. Näin ollen

$$\left\{\begin{array}{rcl} x_1 & = & -t+3 \\ x_2 & = & +t-s \\ x_3 & = & t \\ x_4 & = & s \end{array}\right.$$

missä $t, s \in \mathbb{R}$.

Yhtälöryhmän matriisimerkintä herättää ajatuksen, että onko näitä merkintöjä käytten 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*

$$\left\{\begin{array}{ccccl} x_1 & -x_2 & +2x_3 & -x_4 & = 3 \\ x_1 & +x_2 & & +x_4 & = 3 \end{array}\right. \tag{2.2}$$

ja samoilla kertoimilla määriteltyä yhtälöryhmää

$$\begin{cases} y_1 & -y_2 & +2y_3 & -y_4 & = & 0 \\ y_1 & +y_2 & & +y_4 & = & -2 \end{cases} \tag{2.3}$$

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 kasitteisiin, jossa merkinnässä (2.1) viivalla erotetut puolet tulkitaan omina matriiseinaan ja yhtälöryhmän kertoimet $x_1, \ldots, x_n$ merkitätään 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* $j \in \{1, \dots, m\}$ *ja* $i \in \{1, \dots, n\}$, *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 sarakkaista ja riveistä. Tällä tarkoitetaan seuraavia alkuperäiseen matriisiin liittyviä matriisija.

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. Matriiisi 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 2.3.2. 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

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

niin

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

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 $e_1, \dots, e_n$ 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 $e_1, \dots, e_n$.

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 $i \in \{1, \dots, n\}$. Tällöin

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

□

Matriisitulo ja rivivektorit

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

Merkintä 2.3.14. *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 2.3.15. *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 2.3.16. *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 2.3.17. *Olkoot $A \in \mathbb{R}^{m\times n}$ ja $B \in \mathbb{R}^{n\times k}$. Tällöin*

$$AB = \begin{bmatrix} - & A_1 B & - \\ & \vdots & \\ - & A_m B & - \end{bmatrix}.$$

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 2.3.18. *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 $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots p\}$. 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 2.3.19. *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 2.3.20. *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* $f_M : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ *kaavalla* $x \mapsto Mx$.

Olkoot nyt $A \in \mathbb{R}^{m\times n}$ *ja* $B \in \mathbb{R}^{n\times k}$ *matriiseja. Tällöin kuvaukset* $f_A : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ *ja* $f_B : \mathbb{R}^{k\times 1} \to \mathbb{R}^{n\times 1}$ *voidaan yhdistää kuvaukseksi* $f_A \circ f_B : \mathbb{R}^{k\times 1} \to \mathbb{R}^{m\times 1}$*, 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 $I \in\in \mathbb{R}^{n\times n}$ identiteetti matriisi. 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. $\square$

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 $j, i \in \{1, \dots, n\}$ *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* $1 \le p \le n$ *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 tulomatriisit $D_{p,\lambda}A$ *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

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

Lisäksi

$$D_{p,\lambda}A = AD_{p,\lambda}.$$

Huomautus 2.5.8. *Väitteen jälkimmäinen huomio on itseasiassa yleisemmän huomion erikoistapaus ja pätee kaikille diagonaalimatriiseille. Tähän ilmiöön palataan myöhemmin.*

Lemman 2.5.7 todistus. Koska

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

jokaisella $j \in \{1, \dots, n\}$, niin lemman 2.3.17 perusteella

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

Käytetään jälkimmäiseen huomioon matriisin $D_{p,\lambda}$ sarake-esitystä. Olkoon $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$. Tällöin

$$\begin{aligned} AD_{p,\lambda} &= A \begin{bmatrix} e_1 & \cdots & e_{p-1} & \lambda e_p & e_{p+1} & \cdots & e_n \end{bmatrix} \\ &= \begin{bmatrix} Ae_1 & \cdots & Ae_{p-1} & A(\lambda e_p) & Ae_{p+1} & \cdots & Ae_n \end{bmatrix} \\ &= \begin{bmatrix} a_1 & \cdots & a_{p-1} & \lambda a_p & a_{p+1} & \cdots & e_n \end{bmatrix} \\ &= D_{p,\lambda} A. \end{aligned}$$

Tämä päättää todistuksen. □

Matriisin rivien järjestyksen vaihtaminen matriisitulon avulla

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

Merkintä 2.5.9. *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 2.5.10. *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 2.5.11. *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 2.5.12. *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

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

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

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

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

Lemma 2.5.13. *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 $1 \leq p < r \leq m$ pätee

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

Todistus. Olkoot $1 \leq p < r \leq m$. Tällöin lemman 2.3.17 nojalla

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

$\square$

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 $1 \leq p < r \leq m$ 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 $p, r \in \{1, \ldots, n\}$, joilla $p \neq r$.

Merkintä 2.5.14. *Olkoot $p, r \in \{1, \ldots n\}$, $p \neq r$. 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 2.5.15. *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 2.5.16. *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 2.5.17. *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

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

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

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 2.5.18. *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

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

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 2.5.19. *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

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

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

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

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) $\Rightarrow$ (4) $\Rightarrow$ (1). Yhdessä jo todistettujen implikaatioiden kanssa tämä todistaa oletusten yhtäpitävyyden.

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

$$Ax = 0$$

ja yhtälöryhmä $\left[A \mid 0\right]$ saadaan rivioperaatioilla supistettuun normaalimuotoon $\left[I \mid 0\right]$, 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

$$\left[A \mid 0\right] \sim \left[B_1A \mid 0\right] \sim \left[B_2B_1A \mid 0\right] \sim \cdots \sim \left[B_k \cdots B_2E_1A \mid 0\right],$$

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) $\Rightarrow$ (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 perustella voidaan kuitenkin havaita, että yhtälöiden $Ax = b$ ja $Ax = 0$ ratkaisuilla on jotakin yhteistä. Tarkemmin sanottuna kaikki yhtälön $Ax = b$ ratkaisut saadaan muodostamalla yhdestä ainosta 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 $y' = y + z$. $\square$

Yhtälöä $Ax = 0$ kutsutaan *yhtälöä $Ax = b$ vastavaksi 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 suoralasku:

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

$\square$

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. 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: 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*

$$L = \{tu \in \mathbb{R}^{2\times 1} \colon t \in \mathbb{R}\}.$$

Tällöin L on origon kautta kulkeva suora, sillä $0u = 0$.

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*

$$u = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \ \tilde{u} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \in \mathbb{R}^{3\times 1}$$

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 $0u + 0\tilde{s} = 0$.

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

$$av = a(t_1 u + s_1 \tilde{u}) = (at_1)u + (as_1)\tilde{u} \in P.$$

Näin ollen P on aliavaruus.

3.2 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 sovelluskohdteet 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 on helpo 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. $\square$

3.3 Esimerkki: Matriisin nolla-avaruus

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 likäksi 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.

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

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*

$$x_1 + 2x_2 = 3$$

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

Tarkastellaan nyt homogeenista yhtälöryhmää $\left[A \mid 0\right]$. *Tämän yhtälöryhmän supistettu porrasmuoto on*

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

Yhtälöryhmän $\left[A \mid 0\right]$ *ratkaisut* (x_1, x_2) *ovat muotoa siis*

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

eli

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

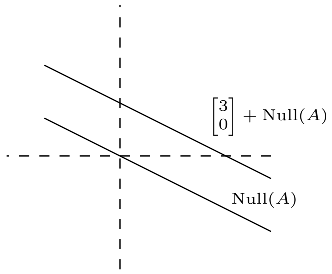
missä $t \in \mathbb{R}$. *Näin ollen*

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

Yhdistämällä tämä jo saatuun tulokseen havaitaan, että yhtälöryhmän $\left[A \mid b\right]$ *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 3.1



Kuva 3.1: Esimerkin 3.3.6 matriisin A nolla-avaruus Null(A) ja suora $\begin{bmatrix}3\\0\end{bmatrix} + \mathrm{Null}(A)$ sarakeavaruudessa $\mathbb{R}^{2\times 1}$.

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

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

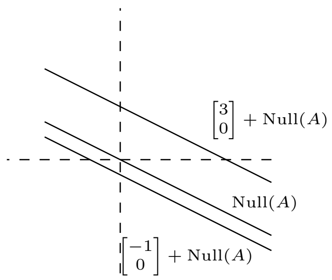
Näin ollen $x + y$ on yhtälöryhmän $\left[A \mid b\right]$ ratkaisu.

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

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

eli $y \in \mathrm{Null}(A)$. □

Lauseen 3.3.4 varsinainen merkitys on siinä, että se paljastaa yhtälöryhmäien $\left[A \mid b\right]$ ja $\left[A \mid b'\right]$ ratkaisujoukot ovat 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.



Kuva 3.2: 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 $\mathbb{R}^{2\times 1}$.

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 $\begin{bmatrix} A \mid b' \end{bmatrix}$ *ratkaisujoukko on*
$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} + \mathrm{Null}(A).$$
Molempien yhtälöryhmien raktkaisujoukot on havainnollistettu kuvassa 3.2.

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 $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ lineaarikombinaatioksi*, jos on olemassa sellaiset luvut* $a_1, \ldots, a_k \in \mathbb{R}$*, että*
$$v = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k.$$

Vektoreiden $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ kaikkien lineaarikombinaatioidan joukkoa merkitään
$$\mathrm{Sp}(v_1, \ldots, v_k) = \{a_1 v_1 + \cdots + a_k v_k \in \mathbb{R}^{n\times 1} \colon a_1, \ldots, a_k \in \mathbb{R}\}.$$

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

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

Olkoot $v, w \in \mathrm{Sp}(v_1, \ldots, v_k)$ ja $a \in \mathbb{R}$. Lineaarikombinaatioiden joukon määritelmän perusteella, on olemassa sellaiset luvut $a_1, \ldots, a_k \in \mathbb{R}$ ja $b_1, \ldots, b_k \in \mathbb{R}$, että $v = a_1v_1 + \cdots + a_kv_k$ ja $w = b_1v_1 + \cdots + b_kv_k$. Järjestelemällä vektoreihin $v_1, \ldots, v_k$ liittyviä termejä saadaan

$$\begin{aligned} av + w &= a\,(a_1v_1 + \cdots + a_kv_k) + (b_1v_1 + \cdots + b_kv_k) \\ &= (aa_1)v_1 + \cdots + (aa_k)v_k + b_1v_1 + \cdots + b_kv_k \\ &= (aa_1 + b_1)v_1 + \cdots + (aa_k + b_k)v_k. \end{aligned}$$

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

$$av + w = c_1v_1 + \cdots + c_kv_k \in \mathrm{Sp}(v_1, \ldots, v_k).$$

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

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 pieninpänä 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 $v \in V$. $\square$

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 Matriisin sarakeavaruus

Kun tarkastellaan vektoreiden $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ virittämää aliavaruutta $\mathrm{Sp}(v_1, \ldots, v_n)$, tarkastellaan samalla näiden vektoreiden jonoa. Tämän jonon $(v_1, \ldots, v_n)$ tarkasteleminen puolestaan oleellisesti tarkoittaa, että tarkastellaan matriisia $\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. Aliavaruutta $\mathrm{Sp}(v_1, \ldots, v_n)$ 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*

$$\mathrm{Col}(A) = \mathrm{Sp}(v_1, \ldots, v_k) \subset \mathbb{R}^{n\times 1}.$$

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 $v_1, \ldots, v_k$ lineaarikombinaatio, eli on olemassa sellaiset luvut $a_1, \ldots, a_k \in \mathbb{R}$, että

$$v = a_1v_1 + \cdots + a_kv_k.$$

Merkitään nyt

$$x = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} \in \mathbb{R}^{k\times 1}.$$

Tällöin

$$v = a_1v_1 + \cdots + a_kv_k = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} = Ax.$$

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

$$Ax = x_1v_1 + \cdots + x_kv_k \in \mathrm{Col}(A),$$

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. $\square$

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 $\begin{bmatrix} A \mid b \end{bmatrix}$ on ratkaisu.

Olkoon nyt $b \in \mathbb{R}^{n\times 1}$ sellainen vektori, että yhtälöryhmällä $\begin{bmatrix} A \mid b \end{bmatrix}$ 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
$$b = Ax = x_1 v_1 + \cdots + x_k v_k \in \mathrm{Sp}(v_1, \ldots, v_k) = \mathrm{Col}(A).$$
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 kappletta 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* $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ *lineaarisesti riippumattomia vektoreita ja* $v \in \mathrm{Sp}(v_1, \ldots, v_k)$*. Tällöin on olemassa sellaiset yksikäsitteiset luvut* $x_1, \ldots, x_k \in \mathbb{R}$*, että*

$$v = x_1v_1 + \ldots + x_kv_k.$$

Todistus. Olkoon $v \in \mathrm{Sp}(v_1, \ldots, v_k)$. Aliavaruuden $\mathrm{Sp}(v_1, \ldots, v_k)$ määritelmän nojalla on olemassa sellaiset luvut $x_1, \ldots, x_k \in \mathbb{R}$, että

$$v = x_1v_1 + \cdots + x_kv_k.$$

Oletetaan nyt, että $y_1, \ldots, y_k \in \mathbb{R}$ ovat sellaisia lukuja, että

$$v = y_1v_1 + \cdots + y_kv_k.$$

Tällöin uudelleen järjestelemällä summat saadaan

$$(y_1 - x_1)v_1 + \cdots + (y_k - x_k)v_k = (y_1v_1 + \cdots + y_kv_k) - (x_1v_1 + \cdots + x_kv_k) = v - v = 0.$$

Koska vektorit $v_1, \ldots, v_k$ ovat lineaarisesti riippumattomia, niin $y_1-x_1 = \cdots = y_k-x_k = 0$, eli $y_i = x_i$ kaikilla $i = 1, \ldots, k$. □

3.6.1 Lineaarisen riippumattomuuden selvittäminen

Luonnollisin tapa selvittää vektoreiden $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ 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 $v_1, \ldots, v_k$ 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* $v_1, \ldots, v_k$ *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 $v_1, \ldots, v_k \in \mathbb{R}^{n\times 1}$ 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* $v_1, \dots, v_k \in \mathbb{R}^{n \times 1}$ *vektoreita. Jono* $(v_1, \dots, v_k)$ *on sidottu, jos ja vain jos on olemassa sellainen* $j \in \{1, \dots, k\}$ *ja luvut* $b_1, \dots, b_{j-1} \in \mathbb{R}$*, että*

$$v_j = b_1 v_1 + \dots + b_{j-1} v_{j-1},$$

eli $v_j \in \mathrm{Sp}(v_1, \dots, v_{j-1})$*.*

Todistus. Oletetaan ensin, että jono $(v_1, \dots, v_k)$ on sidottu, eli että vektorit $v_1, \dots, v_k$ eivät ole lineaarisesti riippumattomia. Tällöin on olemassa sellaiset luvut $a_1, \dots, a_k \in \mathbb{R}$, että

$$a_1 v_1 + \dots + a_k v_k = 0$$

ja että $a_i \neq 0$ jollain $i \in \{1, \dots, k\}$. Näin ollen on olemassa suurin sellainen indeksi $j \in \{1, \dots, k\}$, että $a_j \neq 0$, eli

$$a_1 v_1 + \dots + a_{j-1} v_{j-1} + a_j v_j = 0.$$

Vähentämällä vektori $a_j v_j$ yhtälön molemmilta puolilta saadaan, että

$$a_1 v_1 + \dots + a_{j-1} v_{j-1} = -a_j v_j.$$

Koska $a_j \neq 0$, niin saadaan, että

$$v_j = -\frac{a_1}{a_j} v_1 - \dots - \frac{a_{j-1}}{a_j} v_{j-1}.$$

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

$$v_j = b_1 v_1 + \dots + b_{j-1} v_{j-1}.$$

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

$$\begin{aligned} a_1 v_1 + \dots + a_k v_k &= a_1 v_1 + \dots + a_{j-1} v_{j-1} + a_j v_j + a_{j+1} v_{j+1} + \dots + a_k v_k \\ &= (-b_1) v_1 + \dots + (-b_{j-1}) v_{j-1} + (b_1 v_1 + \dots + b_{j-1} v_{j-1}) \\ &= (b_1 - b_1) v_1 + \dots + (b_{j-1} - b_{j-1}) v_{j-1} = 0. \end{aligned}$$

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* $1 \leq i_1 < \dots < i_d \leq n$*, on vapaa.*

Todistus. Olkoot $a_1, \dots, a_d \in \mathbb{R}$ sellaisia, että

$$a_1 v_{i_1} + a_2 v_{i_2} + \cdots + a_d v_{i_d} = 0.$$

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

$$b_1 v_1 + b_2 v_2 + \cdots + b_k v_k = a_1 v_{i_1} + a_2 v_{i_2} + \cdots + a_d v_{i_d} = 0.$$

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

$$a_1 v_1 + \cdots + a_k v_k = 0.$$

Tällöin

$$a_1 A v_1 + \cdots + a_k A v_k = A(a_1 v_1 + \cdots + a_k v_k) = A0 = 0.$$

Koska jono $(Av_1, \dots, Av_k)$ on vapaa, niin $a_1 = \cdots = a_k = 0$. Näin ollen jono $(v_1, \dots, v_k)$ on osoitettu vapaaksi. □

3.7 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 $v_1, \dots, v_k$ lineaarikombinaationa (kohta 1)

$$v = a_1 v_1 + \cdots + a_k v_k$$

ja että tämä lineaarikombinaatio on yksikäsitteinen (kohta 2), eli että luvut $a_1, \dots, a_k \in \mathbb{R}$ ovat yksikäsitteisiä. Lineaarikombinaation yksikäsitteisyys seuraa lauseesta 3.6.2. Toisaalta, jos jokainen vektori $v \in V$ voidaan kirjoittaa yksikäsitteisenä lineaarikombinaationa vektoreista $v_1, \dots, v_k$, niin tällöin jono $(v_1, \dots, v_k)$ on aliavaruuden V kanta. Kirjataan tämä havainto lemmaksi.

Lemma 3.7.3. *Olkoot* $v_1, \dots, v_k \in V$*. Tällöin jono* $(v_1, \dots, v_k)$ *on aliavaruuden* V *kanta, jos ja vain jos jokaisella* $v \in V$ *on olemassa yksikäsitteiset sellaiset luvut* $a_1, \dots, a_k \in \mathbb{R}$*, että* $v = a_1 v_1 + \cdots + a_k v_k$*.*

Todistus. Oletetaan, että $(v_1, \dots, v_k)$ on aliavaruuden V kanta. Olkoon $v \in V$. Koska $V = \mathrm{Sp}(v_1, \dots, v_k)$, niin on olemassa sellaiset luvut $x_1, \dots, x_k \in \mathbb{R}$, että $v = x_1 v_1 + \cdots + x_k v_k$. Koska jono $(v_1, \dots v_k)$ on vapaa, niin lauseen 3.6.2 perusteella luvut $x_1, \dots, x_k$ ovat yksikäsitteisiä.

Oletetaan nyt, että jokaisella $v \in V$ on olemassa yksikäsitteiset sellaiset luvut $a_1, \dots, a_k \in \mathbb{R}$, että $v = a_1 v_1 + \cdots + a_k v_k$ ja osoitetaan, että $(v_1, \dots, v_k)$ on aliavaruuden V kanta. Koska $V = \mathrm{Sp}(v_1, \dots, v_k)$, riittää osoittaa, että jono $(v_1, \dots, v_k)$ on vapaa. Olkoot $a_1, \dots, a_k \in \mathbb{R}$ sellaisia lukuja, että $a_1 v_1 + \cdots + a_k v_k = 0$. Koska $0 v_1 + \cdots + 0 v_k = 0$, niin lukujen $a_1, \dots, a_k$ yksikäsitteisyyden nojalla $a_1 = \cdots = a_k = 0$. Jono $(v_1, \dots, v_k)$ on näin ollen vapaa. □

Määritelmä 3.7.4. *Olkoon* $V \subset \mathbb{R}^{n \times 1}$ *aliavaruus ja* $(v_1, \dots, v_k)$ *aliavaruuden* V *kanta. Lukujen* $a_1, \dots, a_k \in \mathbb{R}$ *jonoa* $(a_1, \dots, a_k)$ *kutsutaan vektorin* $v \in V$ kertoimiksi tai koordinaateiksi kannassa $(v_1, \dots, v_k)$*, jos*

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

Esimerkki 3.7.5. *Avaruuden $\mathbb{R}^{n\times 1}$ standardikanta $(e_1,\dots,e_n)$ 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

$$v = x_1e_1 + \cdots + x_ne_n,$$

missä luvut $x_1,\dots,x_n$ ovat selvästi yksikäsitteisiä. Standardikanta $(e_1,\dots,e_n)$ 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

$$v = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1e_1 + x_2e_2$$

voidaan kirjoittaa yksikäsitteisesti vektoreiden v_1 ja v_2 lineaarikombinaationa. Tällöin tulemme samalla selvittäneeksi vektorin $v = x_1e_1 + x_2e_2$ kertoimet kannassa (v_1,v_2).

Ratkaistaan tätä varten luvut y_1 ja y_2 yhtälöstä

$$x_1e_1 + x_2e_2 = y_1v_1 + y_2v_2. \tag{3.2}$$

Koska

$$\begin{aligned} y_1v_1 + y_2v_2 &= y_1(e_1+e_2) + y_2(e_1 - e_2) \\ &= y_1e_1 + y_1e_2 + y_2e_1 - y_2e_2 \\ &= (y_1+y_2)e_1 + (y_1-y_2)e_2, \end{aligned}$$

niin yhtälö (3.2) *on yhtäpitävä yhtälön*

$$x_1e_1 + x_2e_2 = (y_1 + y_2)e_1 + (y_1 - y_2)e_2$$

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

$$\begin{cases} y_1 & = \frac{1}{2}\left(x_1 + x_2\right) \\ y_2 & = \frac{1}{2}\left(x_1 - x_2\right), \end{cases}$$

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 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* $\begin{bmatrix} B \mid 0 \end{bmatrix}$ *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 $\begin{bmatrix} B \mid 0 \end{bmatrix}$ *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_2) *on vapaan jonon* (e_1, e_2, e_3, e_4) *osajono* (e_1, e_3)*. Näin ollen* (v_1, v_2) *on vapaa jono lemman 3.6.6 perusteella. Koska* (v_1, v_2) *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 $\begin{bmatrix} B \mid 0 \end{bmatrix}$ 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 $\begin{bmatrix} B \mid 0 \end{bmatrix}$ 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 $\begin{bmatrix} B \mid 0 \end{bmatrix}$ 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

$$B = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_{11} & \cdots & v_{m1} \\ \vdots & & \vdots \\ v_{m1} & \cdots & v_{mn} \end{bmatrix}.$$

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 $(b_{i_1}, \ldots, b_{i_d})$ on vapaa jono, niin $a_1 = \cdots = a_d = 0$. Näin ollen jono $(Pb_1, \ldots, Pb_{i_d})$ on vapaa.

Jono $(Pb_1, \ldots, Pb_n)$ on siis aliavaruuden $\mathrm{Col}(A)$ vapaa virittäjä jono, eli kanta. $\square$

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 Sovellus: 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 kanta avaruudelle $\mathrm{Null}(A)$. Käsitellään tämän jälkeen yleinen tapaus palauttamalla kysymys matriisin supistettuun porrasmuotoon.

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 *virttävät 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 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* $\mathrm{Null}(A)$ *kanta.*

Edellisen esimerkin metodia voi soveltaa kaikkissa tilanteissa, joissa matriisin A nolla-avaruus $\mathrm{Null}(A)$ 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 $\mathrm{Null}(A)$ on sellainen kanta, jonka alkioiden määrä on täsmälleen yhtälöryhmän $\left[A \mid 0\right]$ supistetun porrasmuodon $\left[B \mid 0\right]$ vapaiden muuttujien määrä.

3.9.1 Algoritmi nolla-avaruuden kannan löytämiselle; vapaiden muuttujien rooli

Aloitetaan nolla-avaruuden kannan etsiminen havainnolla, että matriisilla ja sen supistetulla porrasmuodolla on sama nolla-avaruus.

Lause 3.9.2. *Olkoot* $A \in \mathbb{R}^{m \times n}$ *matriisi ja* $B \in \mathbb{R}^{m \times n}$ *matriisin* A *supistettu porrasmuoto. Tällöin* $\mathrm{Null}(A) = \mathrm{Null}(B)$*.*

[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.

Todistus. Koska B on matriisin A supistettu porrasmuoto, niin yhtälöryhmä $\begin{bmatrix} A \mid 0 \end{bmatrix}$ saadaan muotoon $\begin{bmatrix} B \mid 0 \end{bmatrix}$ 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 $x \in \mathrm{Null}(A)$. □

Lauseen 3.9.2 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 3.9.3. *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*

$$\left\{ \begin{array}{rrrrcl} x_1 & -4x_2 & & +2x_4 & = & 0 \\ & & x_3 & -x_4 & = & 0 \\ & & & 0 & = & 0 \end{array} \right. \tag{3.3}$$

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 (3.3) *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* (3.3) *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 3.9.4. *Olkoon* $B \in \mathbb{R}^{m\times n}$ *supistetussa porrasmuodossa oleva matriisi, olkoot* $1 \leq j_1 < \cdots < j_k \leq n$ *yhtälöryhmän* $\left[B \mid 0\right]$ *vapaiden muutujien indeksit sekä olkoot* $1 \leq i_1 < \cdots < i_{n-k} \leq n$ *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 todistus on hyvä tarkastella, miten lausetta voi soveltaa.

Esimerkki 3.9.5. *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* $1 \leq j_1 < \cdots < j_k \leq n$ *ja sidottujen muuttujien indeksit* $1 \leq i_1 < \cdots < i_{n-k} \leq n$*.*

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 $1 \leq i_1 < \cdots < i_{n-k} \leq n$*, 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 3.9.6. *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 3.9.4 todistus on huomattavan tekninen, se luonnollisesti heijastelee näitä havaintoja. Todistuksen voi ensimmäisellä lukukerralla halutessaan sivuuttaa.

Lauseen 3.9.4 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 $w_{j_1}, \dots, w_{j_k}$ 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 $e_1, \dots, e_{n-k} \in \mathbb{R}^{m\times 1}$ ja $e_{i_1}, \dots, e_{i_{n-k}} \in \mathbb{R}^{n\times 1}$.

Määritellään jokaisella $\ell \in \{1, \dots, k\}$ vektorit

$$u_\ell = -\left(b_{1j_\ell} e_{i_1} + \cdots + b_{(n-k)j_\ell} e_{i_{n-k}}\right) \in \mathrm{Sp}(e_{i_1}, \dots, e_{i_{n-k}}).$$

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 $(w_{j_1}, \dots, w_{j_k})$ on aliavaruuden $\mathrm{Null}(B)$ kanta. Osoitetaan ensin, että jono $(w_{j_1}, \dots, w_{j_k})$ on vapaa. Olkoot $a_1, \dots, a_k \in \mathbb{R}$ sellaisia lukuja, että

$$a_1 w_{j_1} + \cdots + a_k w_{j_k} = 0.$$

Nyt termien uudelleen ryhmittelyn jälkeen saadaan

$$a_1 e_{j_1} + \cdots + a_k e_{j_k} + (a_1 u_1 + \cdots + a_k u_k) = 0.$$

Koska vektorit $u_1, \dots, u_k$ kuuluvat aliavaruuteen $\mathrm{Sp}(e_{i_1}, \dots, e_{i_{n-k}})$, niin saadaan, että $a_1 u_1 + \cdots + a_k u_k \in \mathrm{Sp}(e_{i_1}, \dots, e_{i_{n-k}})$. Näin ollen on olemassa sellaiset luvut $c_1, \dots, c_{n-k} \in \mathbb{R}$, että

$$a_1 u_1 + \cdots + a_k u_k = c_1 e_{i_1} + \cdots + c_{n-k} e_{i_{n-k}}.$$

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 $(e_{j_1}, \dots, e_{j_k}, e_{i_1}, \dots, e_{i_{n-k}})$ on standardikannan $(e_1, \dots, e_n)$ jäsenet uudelleen järjestettynä, niin jono $(e_{j_1}, \dots, e_{j_k}, e_{i_1}, \dots, e_{i_{n-k}})$ 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 $(w_{j_1}, \dots, w_{j_k})$ on vapaa.

Osoitetaan nyt, että jono $(w_{j_1}, \dots, w_{j_k})$ 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 $\ell \in \{1, \dots, k\}$. 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 $\mathrm{Null}(A)$. $\square$

3.9.2 Kommentteja

Yhteenvetona edellisistä luvuista voidaan sanoa, että $\begin{bmatrix} A & 0 \end{bmatrix}$ supistetun porrasmuodon sidotut muuttujat antavat sarakeavaruuden $\mathrm{Col}(A)$ kanta-alkioiden määrän ja vapaat muuttujat antavat nolla-avaruuden $\mathrm{Null}(A)$ 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 $\mathrm{Col}(A)$ ja $\mathrm{Null}(A)$ 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 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 tärkeämpi niistä korollaariksi.

Korollaari 3.10.2. *Olkoon $(v_1, \ldots, v_n)$ jono avaruuden $\mathbb{R}^{n \times 1}$ vektoreita. Tällöin seuraavat ehdot ovat yhtäpitäviä:*

1. *$(v_1, \ldots, v_n)$ on vapaa,*

2. $(v_1, \ldots, v_n)$ *virittää avaruuden* $\mathbb{R}^{n\times 1}$ *ja*

3. $(v_1, \ldots, v_n)$ *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. □

Toinen huomio liittyy yhtälöiden ratkaisemiseen.

Huomautus 3.10.3. *Huomaa, että 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

$$b = x_1v_1 + \cdots + x_nv_n = Ax,$$

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

$$b = Ay = y_1v_1 + \cdots + y_nv_n,$$

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. □

3.11 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 $1 \leq k \leq n$, 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 *on sarakeavaruuden* $\mathrm{Col}(A)$ *dimensio* $\dim \mathrm{Col}(A)$.

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 $1 \leq k \leq n$ 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 $1 \leq i \leq k$. Jos $\mathrm{Sp}(v_1, \ldots, v_k) = V$ tai $k = n$, niin lopetetetaan vektoreiden valitseminen. Jos $k < n$ ja $\mathrm{Sp}(v_1, \ldots, v_k) \neq V$, niin tällöin $V \setminus \mathrm{Sp}(v_1, \ldots, v_k) \neq \emptyset$ 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

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

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 ainoastaaan 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 $w_i = A u_i$. 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

$$a_1 u_1 + \cdots + a_m u_m = 0.$$

Tällöin

$$a_1 w_1 + \cdots + a_m w_m = a_1 A u_1 + \cdots + a_m A u_m = A(a_1 u_1 + \cdots + a_m u_m) = 0.$$

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:

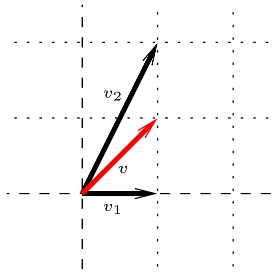
Ongelma. *Voiko vektorin pituuden laskea aina sen kertoimista annetussa kannassa kaavalla* (4.1)*?*

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.

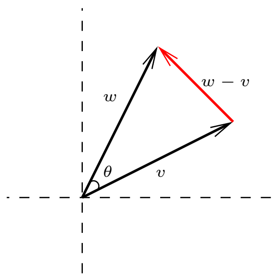
Tästä havaitaan, että pisteen $(1,1)$ *etäisyyttä origosta ei siis voi laskea vektorin* v *kertoimista kannassa* (v_1, v_2)*, sillä*

$$\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{1}{4}} = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}} \neq \sqrt{2}.$$

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* $\mathbb{R}^n$ *standarikannan 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* $k \in \{1, \dots n\}$*. 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

$$u = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad v = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \text{ ja } w = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

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. $\square$

Pistetulon bilineaarisuutta käytetään konkreettisissa laskuissa jatkuvasti ilman erillistä mainintaa. Tämä pätee eritysesti, kun vektorit on ilmoitettu standardikannassa, kuten seuravassa 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 $v_1, \ldots, v_k \in \mathbb{R}^n$ nollasta poikkeavia vektoreita, jotka ovat kohtisuorassa toisiaan vastaan eli $v_i \cdot v_j = 0$ kaikilla $i \neq j$. Tällöin jono $(v_1, \ldots, v_k)$ on vapaa.*

Todistus. Olkoot $a_1, \ldots, a_k \in \mathbb{R}$ sellaisia reaalilukuja, että

$$a_1 v_1 + \cdots + a_k v_k = 0.$$

Osoitetaan, että $a_j = 0$ jokaisella $j \in \{1, \ldots, k\}$.

Olkoon $j \in \{1, \ldots, k\}$. 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 $j \in \{1, \ldots, k\}$, niin saadaan, että $a_j = 0$ jokaisella $j \in \{1, \ldots, k\}$, 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* $\mathbb{R}^{n\times 1}$ *standarikannan* $(e_1, \ldots, e_n)$ *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 3.10.2 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 Vektorin kertoimet ortonormaalissa kannassa

Todistetaan seuraavaksi, että vektorin kertoimet ortonormaalissa kannassa voidaan laskea pistetulon avulla. Tämä tarkoittaa sitä, että vektorin kertoimet 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 $(u_1, \ldots, u_k)$ aliavaruuden V ortonormaali kanta. Jos vektori $v \in V$ on kohtisuorassa kaikkia kantavektoreita $u_1, \ldots, u_k$ vastaan, niin $v = 0$.*

Todistus. Oletetaan, että $v \in V$ on sellainen vektori, että $u_i \cdot v = 0$ kaikilla $i \in \{1, \ldots, k\}$. Koska $(u_1, \ldots, u_k)$ on aliavaruuden V kanta, niin on olemassa sellaiset yksikäsitteliset luvut $a_1, \ldots, a_k \in \mathbb{R}$, että

$$v = a_1 u_1 + \cdots + a_k u_k.$$

Olkoon nyt $i \in \{1, \ldots, k\}$. Oletuksen nojalla $v \cdot u_i = 0$. Koska $(u_1, \ldots, u_k)$ 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 $i \in \{1, \ldots, k\}$, 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 $i \in \{1, \ldots, n\}$.

Olkoon $i \in \{1, \ldots, k\}$. 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 $(v_1, \ldots, v_k)$ aliavaruuden V sellainen kanta, että vektorit $v_1, \ldots, v_m$ ovat kohtisuorassa toisiaan vastaan olevia yksikkövektoreita jollain $0 \leq m \leq k$. Tällöin on olemassa sellainen aliavaruuden V ortonormaalikanta $(u_1, \ldots, u_k)$, että $u_i = v_i$ jokaisella $1 \leq i \leq m$ ja että jokaisella $m + 1 \leq \ell \leq k$ pätee*

$$\mathrm{Sp}(u_1, \ldots, u_\ell) = \mathrm{Sp}(v_1, \ldots, v_\ell). \tag{4.5}$$

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–Schidt 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 $u_{m+1} \in \mathrm{Sp}(u_1, \dots, u_m, v)$, niin tämän todistamiseksi riittää osoittaa, että $v \in \mathrm{Sp}(u_1, \dots, u_{m+1})$. Koska $u_{m+1}, w \in \mathrm{Sp}(u_1, \dots, u_{m+1})$ ja

$$|v - w| u_{m+1} + w = |v - w| \frac{v - w}{|v - w|} + w = v - w + w = v,$$

niin $v \in \mathrm{Sp}(u_1, \dots, u_{m+1})$.

Jäljellä on osoittaa, että vektori u_{m+1} on kohtisuorassa kaikkia vektoreita $u_1, \dots, u_m$ vastaan. Olkoon $1 \leq j \leq m$. Koska jonon $(u_1, \dots, u_m)$ 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 $(v_1, \dots, v_k)$ aliavaruuden V kanta ja olkoon $0 \leq m \leq k$ sellainen luku, että jonon $(v_1, \dots v_m)$ 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 $(u_1, \dots, u_k)$ on aliavaruuden V ortonormaali kanta. Voidaan siis olettaa, että $m < k$.

Muodostetaan ortonormaali kanta $(u_1, \dots, u_k)$ 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 $(u_1, \dots, u_{m+1})$ vektorit ovat toisiaan vastaan kohtisuorassa olevia yksikkövektoreita, joille pätee

$$\mathrm{Sp}(u_1, \dots, u_{m+1}) = \mathrm{Sp}(v_1, \dots, v_{m+1}).$$

Oletetaan nyt, että $m \leq i < k$ on sellainen indeksi, että on olemassa jono $(u_1, \dots, u_i)$ toisiaan vastaan kohtisuorassa olevia yksikkövektoreita, joille pätee

$$\mathrm{Sp}(u_1, \dots, u_\ell) = \mathrm{Sp}(v_1, \dots, v_\ell)$$

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. $\square$

Huomautus 4.3.13. *Lauseen 4.3.9 todistus paljastaa, että ortonormaalikanta löydetään selkeän algortimin avulla. Tämän vuoksi lausetta 4.3.9 kutsutaan* kannan Gram–Schmidt ortonormeeraus prosessiksi. *Edellisessä todistuksessa induktio askel vastaa tätä algoritmia. Algortmin 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 3.10.2 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 $(v_1, \ldots, v_n)$ 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 $(u_1, \ldots, u_n)$, että*

$$U = PA,$$

missä $U = \begin{bmatrix}u_1 & \cdots & u_n\end{bmatrix} \in \mathbb{R}^{n\times n}$ ja matriisi $A \in \mathbb{R}^{n\times n}$ on yläkolmiomatriisi.

Todistus. Olkoon $(u_1, \ldots, u_n)$ avaruuden $\mathbb{R}^n$ ortonormaalikanta kuten lauseessa 4.3.9, eli että

$$\mathrm{Sp}(u_1, \ldots, u_\ell) = \mathrm{Sp}(v_1, \ldots, v_\ell) \tag{4.6}$$

jokaisella $1 \leq \ell \leq n$.

Tarkastellaan ortonormaalin kannan $(u_1, \ldots, u_n)$ kantavektoria u_i. Ehdon (4.6) perusteella tiedetään, että $v_i \in \mathrm{Sp}(u_1, \ldots, u_i)$, eli että vektori v_i on kantavektoreiden $u_1, \ldots, u_i$ lineaarikombinaatio

$$v_i = a_{1i}u_1 + \cdots + a_{ii}u_i.$$

Vektorin v_i esitys kannassa $(u_1, \ldots, u_n)$ 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 $i = 1, \ldots, n$. Tällöin

$$v_i = Ua_i$$

jokaisella $i = 1, \ldots, n$, eli

$$P = UA,$$

missä

$$A = [a_{ji}] = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Sarakevektoreiden a_i määritelmän perusteella, $a_{ji} = 0$ kaikilla $j > i$. Näin ollen matriisi A on yläkolmiomatriisi. $\square$

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 $i \in \{1,\ldots,n\}$ ja $j \in \{1,\ldots,m\}$ 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 $j \in \{1, \ldots, m\}$ ja $i \in \{1, \ldots, k\}$ pätee

$$(AC)^t_{ji} = (AC)_{ij} = \sum_{\ell=1}^{n} a_{i\ell} c_{\ell j} = \sum_{\ell=1}^{n} (A^t)_{\ell i} (C^t)_{j\ell} = \sum_{\ell=1}^{n} (C^t)_{j\ell} (A^t)_{\ell i} = (C^t A^t)_{ji}.$$

Näin ollen

$$(AC)^t = C^t A^t.$$

□

Tärkein syy kiinnostuksestamme transpoosiin liittyy siis huomioon, että pistetulo voidaan transpoosin avulla kirjoittaa matriisitulona:

$$v \cdot w = v^t w. \tag{4.7}$$

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*

$$(Ax) \cdot y = x \cdot A^t y$$

Todistus. Yhtälön (4.7) ja lemman 4.4.4 kohdan (4) perusteella

$$(Ax) \cdot y = (Ax)^t y = x^t A^t y = x^t (A^t y) = x \cdot (A^t y).$$

□

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 $v_1, \ldots, v_m \in \mathbb{R}^{n \times 1}$ vastaan.*

Koska $v_i \cdot v = 0$ jokaisella $i \in \{1, \ldots, m\}$, niin $v_i^t v = 0$ jokaisella $i \in \{1, \ldots, m\}$. 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

$$Av = \begin{bmatrix} v_1^t v \\ \vdots \\ v_m^t v \end{bmatrix} = 0.$$

Näin ollen v on kohtisuorassa vektoreita $v_1, \ldots, v_m$ 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

$$\begin{array}{cc} & \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 3 & -5 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \end{array}.$$

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 $v_1, \ldots, v_n$ 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:

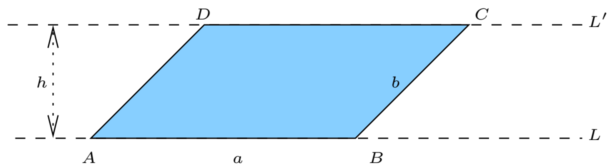
Ongelma. *Onko mahdollista selvittää neliömatriisin* $A \in \mathbb{R}^{n \times n}$ *sarakkeiden lineaarinen riippumattomuus suoralla laskulla?*

Vastaus tähän kysymykseen on tietysti myönteinen ja, kuten luvussa 6 paljastuu, tällä ominaisuudella on tärkeä (teoreettinen) merkitys selvitettäessä matriisien ominaisuuksia.

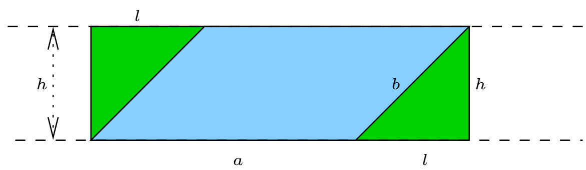
Kysymyksen ratkaisussa tarvittavien käsitteiden motivoimiseksi tarkastellaan aluksi suunnikkaan pinta-alan ja vektoreiden lineaarisen riippumattomuuden välistä yhteyttä. Pinta-alaa käsitellään tässä yhteydessä epäformaalisti lukio-opinnoista tutulla intuitiivisella tavalla. Sen avulla voidaan kuitenkin 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'.

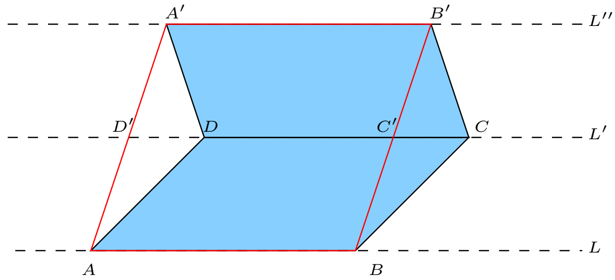


Kuva 5.2: Sinisen suunnikkaan alle on piiretty virheä suorakulmio, jonka pinta-ala on $(a + l)h$, missä $l = \sqrt{b^2 - h^2}$. Koska vihreiden kolmioiden yhteenlaskettu pinta-ala on $2 \cdot (lh/2) = lh$, niin sinisen suunnikkaan pinta-ala on $(a + l)h - lh = ah$.

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

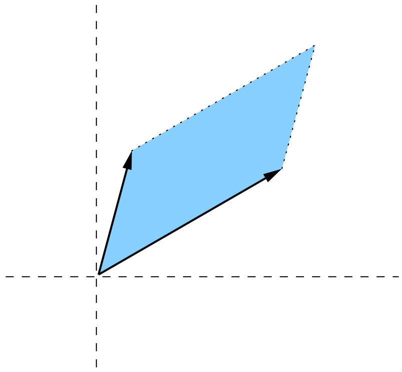
$$u + S(v,w) = \{u + u' \in \mathbb{R}^2 \colon u' \in S(v,w)\} = \{u + tv + sw \in \mathbb{R}^2 \colon t, s \in [0,1]\} \subset \mathbb{R}^2.$$

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.

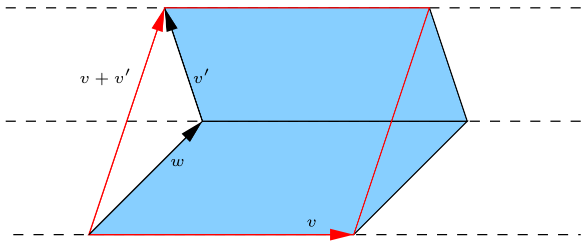
Käyttäen näitä merkintöjä voidaan kuvan 5.3 suunnikkaat $ABCD$ ja $A'B'CD$ ilmaista vektoreilla $v, w, v' \in \mathbb{R}^2$ kuten kuvassa 5.5.

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

$$\mathrm{Ala}(v + v', w) = \mathrm{Ala}(v,w) + \mathrm{Ala}(v',w). \tag{5.1}$$



Kuva 5.4: Vektoreiden v ja w virittämä suunnikas $S(v, w)$.



Kuva 5.5: Sunnikas $ABB'A'$ on ilmastu vektoreiden v ja w avulla ja suunnikas $ABCD$ vektoreiden v' ja w avulla.

Koska pinta-ala ei riipu vektoreiden järjestyksestä, niin vastaavasti pätee

$$\mathrm{Ala}(w, v + v') = \mathrm{Ala}(w, v) + \mathrm{Ala}(w, v').$$

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.

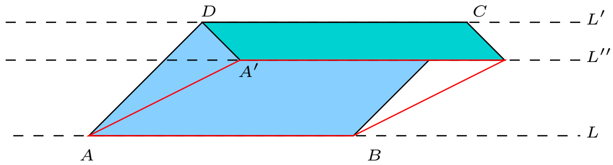
Miksi olemme kiinnostuneita pinta-alasta lineaarialgebran yhteydessä? Tärkein huomio on, että tason vektorit v ja w ovat lineaarisesti riippuvia, jos ja vain jos ne ovat samalla suoralla. Näin ollen vektorit v ja w ovat lineaarisesti riippuvia, jos ja vain jos suunnikkaan pinta-ala $S(v, w)$ on nolla!

Tämä havainto yleistyy itseasiassa kaikkiin avaruuksiin $\mathbb{R}^{n\times 1}$ muodossa, että vektorit $v_1, \ldots, v_n \in \mathbb{R}^{n\times 1}$ ovat lineaarisesti riippumattomia, jos ja vain jos niiden virittämän suunnikkaan

$$S(v_1, \ldots, s_n) = \{t_1 v_1 + \cdots + t_n v_n \in \mathbb{R}^{n\times 1} : t_1, \ldots, t_n \in [0, 1]\}$$

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 skalauksen 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 on 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 määritelmästä 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 $\mathrm{vol}_{\mathbb{R}^{2\times 1}} : \mathbb{R}^{2\times 1} \times \mathbb{R}^{2\times 1} \to \mathbb{R}$, 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* $\mathrm{vol}_{\mathbb{R}^2} : \mathbb{R}^{2\times 1} \times \mathbb{R}^{2\times 1} \to \mathbb{R}$*, 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. $\square$

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. $\square$

5.3 Johdatus $n \times n$-neliömatriisin determinanttiin

Lukijalle

Tämän luvun voi 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* $\omega\colon \mathbb{R}^{n\times 1}\times\cdots\times\mathbb{R}^{n\times 1}\to\mathbb{R}$, $(v_1,\ldots,v_k)\mapsto\omega(v_1,\ldots,v_n)$, *on* multilineaarikuvaus, *jos jokaisella* $i\in\{1,\ldots,n\}$ *ja kaikilla vektoreilla* $v,w,v_1,\ldots,v_n\in\mathbb{R}^{n\times 1}$ *ja vakioilla* $a\in\mathbb{R}$ *pätee sekä*

$$\begin{aligned}\omega(v_1,\ldots,v_{i-1},v+w,v_{i+1},\ldots,v_n)&=\omega(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n)\\&\quad+\omega(v_1,\ldots,v_{i-1},w,v_{i+1},\ldots,v_n)\end{aligned}$$

että

$$\omega(v_1,\ldots,v_{i-1},av_i,v_{i+1},\ldots,v_n)=a\omega(v_1,\ldots,v_{i-1},v_i,v_{i+1},\ldots,v_n).$$

Muutama huomautus on paikallaan.

Huomautus 5.3.2. *Kuvausta* $(v_1,\ldots,v_n)\mapsto\omega(v_1,\ldots,v_n)$ *sanotaan multilineaariseksi, jos se on siis lineaarinen erikseen jokaisen muuttujan* v_i *suhteen. Huomaa, että edellisessä määritelmässä ehdot voidaan kirjoittaa lyhyesti muodossa, että jokaisella* $i\in\{1,\ldots,n\}$ *kuvaus* $v\mapsto\omega(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n)$ *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 voidaan myös tarkastella yleisempiä multilineaarikuvauksia, joiden argumenttina on jonkin muu määrä avaruuden* $\mathbb{R}^{n\times 1}$ *vektoreita. 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* $v,v_1,\ldots,v_n\in\mathbb{R}^n$ *pätee*

$$\omega(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_{j-1},v,v_{j+1},\ldots,v_n)=0.$$

Huomautus 5.3.5. *Yllä oleva määritelmä voidaan tulkita sanomalla, että alternoiva multilineaarikuvaus saa arvon nolla, jos jokin vektori esiintyy kahdesti argumenttina olevassa jonossa* $(v_1,\ldots,v_n)$. *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* $v_1,\ldots,v_n\in\mathbb{R}^{n\times 1}$ *ja* $1\le i<j\le n$ *pätee*

$$\omega(v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_{j-i},v_i,v_{j+1},\ldots v_n)=-\omega(v_1,\ldots,v_n).$$

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

$$\begin{aligned}\omega(v_1,\ldots,v_n)&=\omega(v_1,\ldots,v_{i-1},v_i,v_{i+1},\ldots,v_n)\\&=\omega\left(v_1,\ldots,v_{i-1},-\frac{1}{a_i}\sum_{\substack{j=1\\ j\neq i}}^{n}a_jv_j,v_{i+1},\ldots,v_n\right)\\&=-\frac{1}{a_j}\sum_{\substack{j=1\\ j\neq i}}^{n}a_j\omega(v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_n).\end{aligned}$$

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

$$\omega(v_1,\ldots,v_n)=-\frac{1}{a_j}\sum_{\substack{j=1\\ j\neq i}}^{n}a_j\omega(v_1,\ldots,v_{i-1},v_j,v_{i+1},\ldots,v_n)=0.$$

□

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 tapauksssa ja tehdään luvussa 5.8.*

5.3.2 Permutaatiot

Permutaatio on toinen nimi bijektiolle.

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 $1 \le i < j \le n$, 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, kirjoitataan nämä määritelmät tarkemmin seuraavasti.

Olkoon $\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\}$ permutaatio. Merkitään

$$N(\sigma) = \#\{(i,j)\colon 1 \le i < j \le n,\ \sigma(i) > \sigma(j)\} \in \mathbb{N},$$

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* $1 \le i < j \le 3$, *on täsmälleen kolme kappletta, 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

$$\operatorname{sign}(\sigma) = (-1)^{N(\sigma)} = (-1)^2 = 1.$$

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* $\sigma(4) = 1$, *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* $\operatorname{sign}(\sigma) = (-1)^{N(\sigma)} = -1$.

Esimerkki 5.3.15. *Identtisen permutaation* $\operatorname{id}\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$, *eli permutaation, jolle pätee* $\operatorname{id}(i) = i$ *kaikilla* $i \in \{1,\ldots,n\}$, *merkki on aina* 1, *sillä* $N(\operatorname{id}) = 0$.

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 B.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*

$$\omega(v_1,\ldots,v_n) = \left(\sum_{\sigma\in\operatorname{Sym}(n)} \operatorname{sign}(\sigma)v_{\sigma(1)1}\cdots v_{\sigma(n)n}\right)\omega(e_1,\ldots,e_n). \tag{5.5}$$

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*

$$\omega(v_{\sigma(1)},\ldots,v_{\sigma(n)}) = \operatorname{sign}(\sigma)\omega(v_1,\ldots,v_n).$$

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 B.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 B.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.

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 permutatiota* $\sigma \in \mathrm{Sym}(3)$. *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 tapauset:*

$$\begin{bmatrix} \mathbf{a_{11}} & a_{12} & a_{13} \\ a_{21} & \mathbf{a_{22}} & a_{23} \\ a_{31} & a_{32} & \mathbf{a_{33}} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{a_{11}} & a_{12} & a_{13} \\ a_{21} & a_{22} & \mathbf{a_{23}} \\ a_{31} & \mathbf{a_{32}} & a_{33} \end{bmatrix},$$
$$\begin{bmatrix} a_{11} & \mathbf{a_{12}} & a_{13} \\ \mathbf{a_{21}} & a_{22} & a_{23} \\ a_{31} & a_{32} & \mathbf{a_{33}} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & \mathbf{a_{13}} \\ \mathbf{a_{21}} & a_{22} & a_{23} \\ a_{31} & \mathbf{a_{32}} & a_{33} \end{bmatrix},$$
$$\begin{bmatrix} a_{11} & \mathbf{a_{12}} & a_{13} \\ a_{21} & a_{22} & \mathbf{a_{23}} \\ \mathbf{a_{31}} & a_{32} & a_{33} \end{bmatrix} \quad \textit{ja} \quad \begin{bmatrix} a_{11} & a_{12} & \mathbf{a_{13}} \\ a_{21} & \mathbf{a_{22}} & a_{23} \\ \mathbf{a_{31}} & a_{32} & a_{33} \end{bmatrix}.$$

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 sarkelineaarisuus 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*

$$\begin{aligned}
&\det \begin{bmatrix} v_1 & \cdots & v_{i-1} & (av+bw) & v_{i+1} & \cdots & v_n \end{bmatrix} \\
&= a \det \begin{bmatrix} v_1 & \cdots & v_{i-1} & v & v_{i+1} & \cdots & v_n \end{bmatrix} \\
&\quad + b \det \begin{bmatrix} v_1 & \cdots & v_{i-1} & w & v_{i+1} & \cdots & v_n \end{bmatrix}.
\end{aligned}$$

[2]Itseasiassa lemman tulos olisi voitu myös 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 $v_1, \dots, v_n \in \mathbb{R}^n$. Koska $v \mapsto Av$ on lineaarikuvaus ja $\mathrm{vol}_{\mathbb{R}^n}$ on alternoiva multilineaarikuvaus, niin ω on alternoiva multilineaarikuvaus. Lisäksi

$$\omega(e_1, \dots, e_n) = \mathrm{vol}_{\mathbb{R}^n}(Ae_1, \dots, Ae_n) = \det A$$

ja

$$\det(AB) = \mathrm{vol}_{\mathbb{R}^n}((AB)e_1, \dots, (AB)e_n) = \mathrm{vol}_{\mathbb{R}^n}(A(Be_1), \dots, A(Be_n)) = \omega(Be_1, \dots, Be_n).$$

Olkoot $b_1, \dots, b_n \in \mathbb{R}^{n\times 1}$ matriisin $B = [b_{ji}]$ sarakkeet, eli $B = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}$. Tällöin lauseen 5.3.16 perusteella

$$\begin{aligned}\omega(Be_1, \dots, Be_n) &= \omega(b_1, \dots, b_n) \\ &= \left(\sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} \right) \omega(e_1, \dots, e_n) \\ &= (\det B)\omega(e_1, \dots, e_n) = (\det B)(\det A).\end{aligned}$$

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 B.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 B.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 karaketerisointilauseen (lause 3.10.1) nojalla matriisin A sarekkeet eivät ole lineaarisesti riippumattomia. 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*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

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 sarekkeen 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

$$\begin{aligned} \det A &= \sum_{k=1}^{3} (-1)^{k+3} a_{k3} \det A|_{k3} \\ &= (-1)^{1+3} a_{13} \det A|_{13} + (-1)^{2+3} a_{23} \det A|_{23} + (-1)^{3+3} a_{23} \det A|_{23} \\ &= 1 \cdot 0 \cdot \det A|_{13} + (-1) \cdot 0 \cdot \det A|_{23} + 1 \cdot 1 \cdot \det A|_{23} \\ &= \det A|_{23} \\ &= \det \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = 1 \cdot 1 - 0 \cdot 2 = 1. \end{aligned}$$

Lasketaan nyt sama determinantti kehittämällä se keskimmäisen sarekkeen 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 sarekkeen 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 B.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 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 $\ell \in \{1,\dots,n\}$. Osoitetaan, että

$$\det A = \sum_{k=1}^{n} (-1)^{k+\ell} a_{k\ell} \det A|_{k\ell}.$$

Olkoot $a_1,\dots,a_n \in \mathbb{R}^{n\times 1}$ 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 $i = 1,\dots,n$.

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 $k \in \{1,\dots,n\}$ 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 $a_1, \ldots, a_{\ell-1}, a_{\ell+1}, a_n$ 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 $\left[A \mid b\right]$ 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 vakio vektoreiden $b \in \mathbb{R}^{n\times 1}$ esittämiseen matriisin A tapauksessa. Käytännön tasolla tämä tarkoittaa kannan muodostamista matriisin ominaisvektoreista ja tähän aiheeseen tutustuminen aloitetaan tässä luvussa.

Aloitetaan motivoivasta hypoteettisesta esimerkistä.

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 $b = c_1v_1 + \cdots + c_nv_n$. Vastaavasti, jos $x \in \mathbb{R}^{n\times 1}$, niin $x = y_1v_1 + \cdots + y_nv_n$. Koska $Av_i = \lambda_i v_i$ jokaisella $i \in \{1, \dots, n\}$, niin saadaan yhtälö

$$\begin{aligned} c_1v_1 + \cdots + c_nv_n = b = Ax = A(y_1v_1 + \cdots + y_nv_n) &= y_1Av_1 + \cdots + y_nAv_n \\ &= y_1\lambda_1v_1 + \cdots + y_n\lambda_nv_n, \end{aligned}$$

eli jokaisella $i \in \{1, \dots, n\}$ *yhtälö* $c_i = y_i \lambda_i$.

Tässä kannassa $(v_1, \dots, v_n)$ *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* $i \in \{1, \dots, n\}$ *ja* $d_{ji} = 0$, *kun* $j \neq i$.[1]

Ainoa jäljellä oleva kysymys tässä tilanteessa on selvittää vektorin b *kertoimet* $c_1, \dots, c_n$ *kannassa* $(v_1, \dots, v_n)$, *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

$$Ax = b,$$

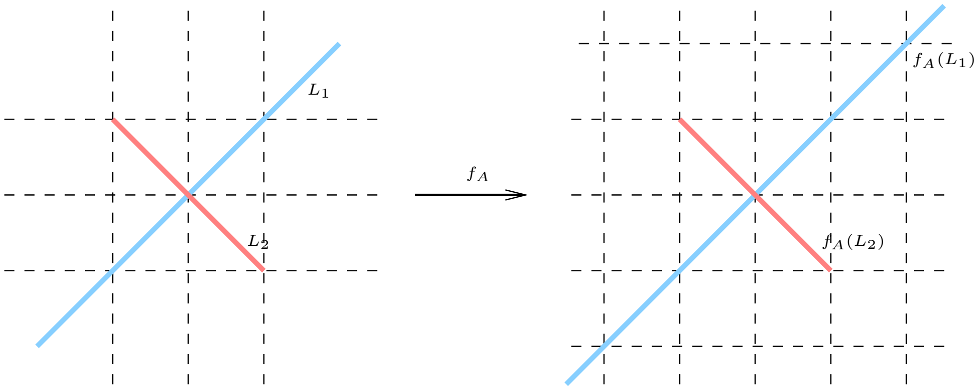
missä

$$b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

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 puolestan 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 aliavauuksiin, joihin rajoitettuna matriisi käyttyy kuin skaalaus tai kierto. Jossain mielessä tämän kysymyksen selvittämiseen käytetään tämän luku, mutta myös koko 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 määrittelmeninen.

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 identtinen $(n \times n)$-matriisi. 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

$$(A - \lambda I)x = Ax - \lambda Ix = Ax - \lambda x = 0.$$

Näin ollen matriisi $A - \lambda I$ ei ole kääntyvä lauseen 3.10.1 perusteela.

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

$$Ax = (A - \lambda I + \lambda I)x = (A - \lambda I)x + \lambda Ix = 0 + \lambda x = \lambda x.$$

□

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 $\det(A - \lambda I) = 0$. □

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 λ 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 kaksi teräinen miekka, kuten tulemme huomaamaan.

Esimerkki 6.2.5. *Olkoon nyt*

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

Etsitään tämän matriisin ominaisarvot λ *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

$$\det(D - \lambda I) = \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = (-\lambda)^2 - (-1)1 = \lambda^2 + 1 > 0.$$

Näin ollen yhtälöllä $\det(D - \lambda I) = 0$ *ei ole 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* λ *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 ratkaisuja. Matriisilla* R *ei siis tässä tapauksessa ole lainkaan 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 vastaavat determinanttiyhtälön $\det(A-\lambda I)=0$ ratkaisuja. Kun matriisin A ominaisarvot $\lambda_1,\ldots,\lambda_m$ on näin selvitetty, voidaan vastaavat ominaisvektorit selvittää yhtälöstä $(A-\lambda_k)x=0$. 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

$$A-\lambda_1 I=\begin{bmatrix}1 & 1\\ 0 & 2\end{bmatrix}-\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix},$$

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 $\lambda=\lambda_1,\ldots,\lambda_m$.*
- *Etsitään yhtälöryhmän $(A-\lambda_k I)x=0$ epätriviaalit ratkaisut x, 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 $\mathbb{R}^{n\times n}$, 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

$$
\begin{aligned}
\det(A-\lambda I) = \begin{bmatrix} 1-\lambda & 0 & 2 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{bmatrix} &= (1-\lambda)\det\begin{bmatrix} 1-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix} + \det\begin{bmatrix} 0 & 2 \\ 1-\lambda & 0 \end{bmatrix} \\
&= (1-\lambda)(1-\lambda)(2-\lambda) - 2(1-\lambda) \\
&= (1-\lambda)\left((1-\lambda)(2-\lambda) - 2\right) \\
&= (1-\lambda)\left(2-\lambda-2\lambda+\lambda^2-2\right) \\
&= (1-\lambda)\left(\lambda^2-3\lambda\right) \\
&= \lambda(1-\lambda)\left(\lambda-3\right).
\end{aligned}
$$

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

$$
\begin{aligned}
\det(A-\lambda I) = \begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{bmatrix} &= (1-\lambda)\det\begin{bmatrix} 2-\lambda & 0 \\ 0 & 1-\lambda \end{bmatrix} + \det\begin{bmatrix} 0 & 1 \\ 2-\lambda & 0 \end{bmatrix} \\
&= (1-\lambda)(2-\lambda)(1-\lambda) - (2-\lambda) \\
&= (2-\lambda)\left((1-\lambda)^2 - 1\right) \\
&= (2-\lambda)\left(1-2\lambda+\lambda^2-1\right) \\
&= \lambda(2-\lambda)\left(\lambda-2\right) = \lambda(\lambda-2)^2.
\end{aligned}
$$

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, että edellisen esimerkin yleisen muodon: matriisi voidaan ilmaista diagonaalimatriisin avulla, jos sen ominaisvektoreista voidaan muodostaa kanta. Jotta tämä yhteys voidaan muotoilla tarkisti, 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 ominaisarvoja.*

Todistus. Ehdon riittävyys seuraa lauseesta 6.4.4. Ehdon välttämättömyys puolestaan lauseesta 6.4.7. □

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 $b'_{i_1}, \dots, b'_{i_k}$ ovat lineaarisesti riippumattomia. Näin ollen sarakkeita $b_{i_1}, \dots, b_{i_k}$ vastaavat muuttujat ovat vapaita matriisin B' yhtälöryhmässä $\left[B' \mid 0\right]$. Olkoon nyt $i \notin \{i_1, \dots, i_k\}$. Tällöin sarake b_i on sarakkeiden $b_{i_1}, \dots, b_{i_k}$ lineaarikombinaatio. Näin ollen myös sarake b'_i on vastaavien sarakkeiden $b'_{i_1}, \dots, b'_{i_k}$ 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 $i \notin \{i_1, \dots, i_k\}$. Tällöin b_i on sarakkeiden $b_{i_1}, \dots, b_{i_k}$ lineaarikombinaatio, eli on olemassa sellaiset luvut $a_1, \dots, a_k \in \mathbb{R}$ että

$$b_i = a_1 b_{i_1} + \cdots + a_k b_{i_k}.$$

Tällöin

$$b'_i = Qb_i = a_1 Qb_{i_1} + \cdots + a_k Qb_{i_k} = a_1 b'_{i_1} + \cdots + a_k b'_{i_k} = a_1 e_1 + \cdots + a_k e_k = b_i.$$

Matriisit B ja B' ovat siis sama matriisi. $\square$

Liite B

Liite: Determinanttien teorian lisäsivut

B.1 Permutaatiot ja transpositiot

Tässä luvussa täydennetään luvussa 5.3.2 käsiteltyä permutaatioiden teoriaa. Ensin osoitetaan permutaation merkin tulokaava (lause B.1.4) ja tämän jälkeen, että jokaisen permutaation voi esittää transpositioiden tulona. Palautetaan aluksi mieleen käsitteet ja merkinnät.

Määritelmä B.1.1. *Joukon* $\{1, \dots, n\}$ *bijektiota* $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ *kutsutaan* permutaatioksi. *Permutaatio* $\sigma\colon \{1, \dots, n\} \to \{1, \dots n\}$ *on* transpositio*, jos on olemassa sellaiset luvut* $i, j \in \{1, \dots, n\}$*,* $i \neq j$*, että* $\sigma(i) = j$*,* $\sigma(j) = i$ *ja* $\sigma(k) = k$ *kaikilla* $k \in \{1, \dots, n\} \setminus \{i, j\}$*.*

Joukon $\{1, \dots, n\}$ permutaatioiden joukkoa

$$\mathrm{Sym}(n) = \{\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\} : \sigma \text{ on permutaatio}\}$$

kutsutaan joukon $\{1, \dots, n\}$ *symmetriaryhmäksi*, koska kahden permutaation yhdiste on permutaatio ja permutaation käänteiskuvaus on permutaatio. Näin ollen joukon $\{1, \dots, n\}$ permutaatiot muodostavat ryhmän algebrallisessa mielessä.

Permutaation merkki on permutaatioon liittyvä luku 1 tai -1, joka kuvastaa samanaikaisesti kahta permutaation $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ 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 B.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 B.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 B.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 B.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 B.1.6. *Lause B.1.4 osoittaa, että permutaation merkki* sign$(\cdot)$ *määrittelee ryhmähomomorfismin* sign: $\mathrm{Sym}(n) \to \mathbb{F}_2$, *missä $\mathbb{F}_2 = \{-1, 1\}$ on kahden alkion multipilikatiivinen ryhmä. Näitä asioita käsitellään kurssilla* Algebralliset rakenteet I & II.

Huomautus B.1.7. *Lauseen B.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.*

B.1.1 Permutaatiot transpositioiden yhdisteenä

Osoitetaan nyt, että jokainen permutaatio on transpositioiden yhdiste. Yhden alkion joukon $\{1\}$ tapaus jätetään lukijalle.

Lemma B.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

$$\tau \circ \tau_1 \circ \tau_k = \sigma.$$

□

B.2 Funktio $\mathrm{vol}_{\mathbb{R}^{n\times 1}}$ 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

$$\begin{aligned}
&\mathrm{vol}_{\mathbb{R}^{n\times 1}}(v_1, \dots, v_{j-1}, av + bw, v_{j+1}, \dots, v_n) \\
&\quad = \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sign}(\sigma) v_{\sigma(1)1} \cdots v_{\sigma(j-1),j-1}(av_{\sigma(j)} + bw_{\sigma(j)}) v_{\sigma(j+1),j+1} \cdots v_{\sigma(n)n}.
\end{aligned}$$

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

$$\begin{aligned}
\sum_{\xi\in S^<} \operatorname{sign}(\xi\circ\tau) v_{(\xi\circ\tau)(1)1}\cdots v_{(\xi\circ\tau)(n)n} &= \sum_{\xi\in S^<} \operatorname{sign}(\xi)\operatorname{sign}(\tau) v_{(\xi\circ\tau)(1)1}\cdots v_{(\xi\circ\tau)(n)n} \\
&= \sum_{\xi\in S^<} \operatorname{sign}(\xi)\operatorname{sign}(\tau) v_{(\xi\circ\tau)(1)\tau(1)}\cdots v_{(\xi\circ\tau)(n)\tau(n)} \\
&= \operatorname{sign}(\tau) \sum_{\xi\in S^<} \operatorname{sign}(\xi) v_{(\xi(1)1}\cdots v_{(\xi(n)n} \\
&= -\sum_{\xi\in S^<} \operatorname{sign}(\xi) v_{(\xi(1)1}\cdots v_{(\xi(n)n}.
\end{aligned}$$

Näin ollen

$$\begin{aligned}
\operatorname{vol}_{\mathbb{R}^n}(v_1,\ldots,v_n) &= \sum_{\xi\in S^<} \operatorname{sign}(\xi) v_{\xi(1)1}\cdots v_{\xi(n)n} + \sum_{\xi\in S^<} \operatorname{sign}(\xi\circ\tau) v_{(\xi\circ\tau)(1)1}\cdots v_{(\xi\circ\tau)(n)n} \\
&= \sum_{\xi\in S^<} \operatorname{sign}(\xi) v_{\xi(1)1}\cdots v_{\xi(n)n} - \sum_{\xi\in S^<} \operatorname{sign}(\xi) v_{(\xi(1)1}\cdots v_{(\xi(n)n} \\
&= 0.
\end{aligned}$$

Tämä todistaa, että funktio $\operatorname{vol}_{\mathbb{R}^n}$ on alternoiva.

Osoitetaan vielä, että $\operatorname{vol}_{\mathbb{R}^{n\times 1}}(e_1,\ldots,e_n) = 1$. 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

$$\operatorname{vol}_{\mathbb{R}^{n\times 1}}(e_1,\ldots,e_n) = \operatorname{sign}(\mathrm{id}) e_{11}\cdots e_{nn} = 1.$$

□

B.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 permutaatioiden 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*

$$\omega(v_1,\ldots,v_n) = \left(\sum_{\sigma\in\operatorname{Sym}(n)} \operatorname{sign}(\sigma) v_{\sigma(1)1}\cdots v_{\sigma(n)n}\right)\omega(e_1,\ldots,e_n). \tag{5.5}$$

Lauseen todistus on kaksiosoinen lasku. Ensimmäisessä vaiheessa hyödynnetään vektoreiden $v_1, \dots, v_n$ esitystä standardikannassa ja käytetään multilineaariisuutta tuomaan lineaarikombinaatiot $v_i = v_{1i}e_1 + \dots + v_{ni}e_n$ 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 $i \in \{1, \dots, n\}$, niin

$$\begin{aligned}
\omega(v_1, \dots, v_n) &= \omega\left(\sum_{j=1}^{n} v_{j1}e_j, \sum_{j=1}^{n} v_{j2}e_j, \dots, \sum_{j=1}^{n} v_{jn}e_j\right) \\
&= \omega\left(\sum_{j_1=1}^{n} v_{j_1 1}e_{j_1}, \sum_{j_2=1}^{n} v_{j_2 2}e_{j_2}, \dots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right),
\end{aligned}$$

missä toistuva summamuuttuja j on uudelleen nimetty muuttujiksi $j_1, j_2, \dots, j_n$. Koska ω on multilineaarinen, niin saadaan toistamalla multilineaarisuutta n kertaa saadaan

$$\begin{aligned}
\omega(v_1, \dots, v_n) &= \omega\left(\sum_{j_1=1}^{n} v_{j_1 1}e_{j_1}, \sum_{j_2=1}^{n} v_{j_2 2}e_{j_2}, \dots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \sum_{j_1=1}^{n} v_{j_1 1}\omega\left(e_{j_1}, \sum_{j_2=1}^{n} v_{j_2 2}e_{j_2}, \dots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} v_{j_1 1}v_{j_2 2}\omega\left(e_{j_1}, e_{j_2}, \sum_{j_3=1}^{n} v_{j_3 3}e_{j_3}, \dots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \sum_{j_1, j_2=1}^{n} v_{j_1 1}v_{j_2 2}\omega\left(e_{j_1}, e_{j_2}, \sum_{j_3=1}^{n} v_{j_3 3}e_{j_3}, \dots, \sum_{j_n=1}^{n} v_{j_n n}e_{j_n}\right) \\
&= \cdots \\
&= \sum_{j_1, \dots, j_n=1}^{n} v_{j_1 1}v_{j_2 2}v_{j_n n}\omega\left(e_{j_1}, e_{j_2}, \dots, e_{j_n}\right).
\end{aligned}$$

Huomaa, että viimeissä summassa kaikki muuttujat $j_1, \dots, j_n$ käyvät läpi arvot $1, \dots, n$.

Tarkastellaan nyt termejä $\omega\left(e_{j_1}, e_{j_2}, \dots, e_{j_n}\right)$. Olkoot $j_1, \dots, j_n \in \{1, \dots, n\}$. Jos kahdella muuttujista $j_1, \dots, j_n$ on sama arvo $k \in \{1, \dots, n\}$, niin tällöin vektori e_k toistuu jonossa $(e_{j_1}, \dots, e_{j_n})$ ainakin kahdesti. Koska ω on alternoiva, niin tällöin $\omega(e_{j_1}, \dots, e_{j_n}) = 0$.

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. □

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

$$\omega(v_{(\tau_1\circ\cdots\circ\tau_m)(1)},\dots,v_{(\tau_1\circ\cdots\circ\tau_m)(n)})=(-1)^m\omega(v_1,\dots,v_n)=\operatorname{sign}(\sigma)\omega(v_1,\dots,v_n).$$

□

B.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 $\operatorname{Sym}(n-1)\to S$.