

Jokke Häsä, Terhi Kaarakka, Lotta Oinonen, Heikki Orelma ja
Johanna Rämö

VEKTORIT JA MATRIISIT

ALKUSANAT

Tämä oppimateriaali on johdatus vektorien ja matriisien perusominaisuuksiin, teoriaan ja sovelluksiin.

Materiaalin pohjana ovat olleet Tampereen yliopiston moniste ”Johdatus vektori- matriisilaskentaan” sekä Helsingin yliopiston moniste ”Johdatus lineaarialgebraan, Osa 1”. Ensin mainitun kirjoittajia ovat olleet Terhi Kaarakka ja Heikki Orelma, ja jälkimmäisen Jokke Häsä, Lotta Oinonen ja Johanna Rämö. Nämä kaksi monistetta on yhdistänyt täksi materiaaliksi Johanna Rämö.

Materiaali on koostettu niin, että sen lukuja voi käyttää monessa eri järjestyksessä. Esimerkiksi suoria ja tasoja koskevan luvun voi käydä vasta myöhemmässä vaiheessa kurssia ja lineaariset yhtälöryhmät käsitellä ennen matriiseja. Matriisien sarake- ja nolla-avaruuksia käsittelevät alaluvut puolestaan sijaitsevat usean eri luvun lopussa, mutta ne voi myös käsitellä yhtenä kokonaisuutena.

Tampere, 19. lokakuuta 2020

Jokke Häsä, Terhi Kaarakka, Lotta Oinonen, Heikki Orelma ja Johanna Rämö

SISÄLLYSLUETTELO

1	Vektorit	1
1.1	Mitä vektorit ovat?	1
1.2	Vektoriavaruus $\mathbb{R}^n$	3
1.3	Pistetulo ja normi	12
1.4	Projektio suoralle	20
1.5	Ristitulo	25
2	Sovellus: Suorat ja tasot	32
2.1	Suorat	32
2.2	Tasot	40
3	Matriisit	47
3.1	Mitä matriisit ovat?	47
3.2	Matriisien laskutoimituksia	48
3.3	Matriisit kuvauksina	58
3.4	Erityisiä matriiseja ja matriisien laskusääntöjä	61
3.5	Matriisin transpoosi	64
3.6	Käänteismatriisi	66
3.7	Sarake- ja nolla-avaruus	71
4	Lineaariset yhtälöryhmät	74
4.1	Ongelmia ja ratkaisuja	74
4.2	Lineaarisen yhtälöryhmän määritelmä	75
4.3	Yhtälönratkaisun idea	79
4.4	Gaussin–Jordanin eliminointimenetelmä	84
4.5	Yhtälöryhmästä matriisiyhtälöksi	97
4.6	Alkeismatriisit	99
4.7	Sarake- ja nolla-avaruus yhtälönratkaisussa	107
5	Lineaarialgebraa	111
5.1	Yhtälöitä vektoriavaruuksissa	111
5.2	Virittäminen	112
5.3	Aliavaruudet	116
5.4	Lineaarinen riippumattomuus	122
5.5	Kanta ja dimensio	130
5.6	Sarake- ja nolla-avaruuden kanta ja dimensio	139
6	Determinantti ja ominaisarvot	146
6.1	Pienten matriisien determinantit	146

6.2	Determinantin yleinen määritelmä ja kehityskaavat	147
6.3	Determinantin ominaisuuksia	150
6.4	Ominaisarvon määritelmä	156
6.5	Ominaisarvojen löytäminen	163
6.6	Diagonalisointi	166
7	Sovellus: Pienimmän neliösumman menetelmä	173
7.1	Kun ratkaisut puuttuvat	173
7.2	Pienimmän neliösumman menetelmä	173
7.3	Polynomin sovittaminen pisteistöön	178

1 VEKTORIT

1.1 Mitä vektorit ovat?

Pohdi 1.1.1

Mikä seuraavista luonnehdinnoista kuvaa mielestäsi *parhaiten* vektoria?

1. suure, jolla on suunta ja pituus
2. $a\mathbf{i} + b\mathbf{j}$
3. nuoli
4. origosta lähtevä nuoli
5. koordinaatiston piste
6. (a, b)
7. $\begin{bmatrix} a \\ b \end{bmatrix}$
8. geenitekniikan työväline

Vastasitpa edelliseen tehtävään mitä tahansa, oli vastauksesi oikein. Kaikilla edellä kuvatuilla tavoilla voi kuvata vektoria. Eri yhteyksissä vektoreita käsitellään eri tavalla. Ennen kuin ryhdymme käyttämään vektoreita, onkin tärkeää sopia, mitä vektorit tässä materiaalissa tarkoittavat.

Katso video ”[Mitä ovat vektorit?](#)”.

Lukiomatematiikassa vektorit esitetään olioina, joilla on suunta ja pituus. Tason ja kolmiulotteisen avaruuden vektorit kirjoitetaan yleensä yksikkövektorien $\mathbf{i}$, $\mathbf{j}$ ja $\mathbf{k}$ avulla. Eräs tason vektori voisi olla vaikkapa $\mathbf{v} = -3\mathbf{i} + 2\mathbf{j}$ ja eräs kolmiulotteisen avaruuden vektori $\mathbf{w} = -5\mathbf{i} + \sqrt{2}\mathbf{j} + 101\mathbf{k}$. Kaikki informaatio sisältyy vektorien $\mathbf{i}$, $\mathbf{j}$ ja $\mathbf{k}$ edessä oleviin kertoimiin. Vektorin ilmaisemiseksi riittää siis listata nuo kertoimet. Edellä mainitun vektorin $\mathbf{v}$ voisi kirjoittaa vaikkapa muodossa $(-3, 2)$. Vektorin $\mathbf{w}$ voisi puolestaan kirjoittaa muodossa $(-5, \sqrt{2}, 101)$.

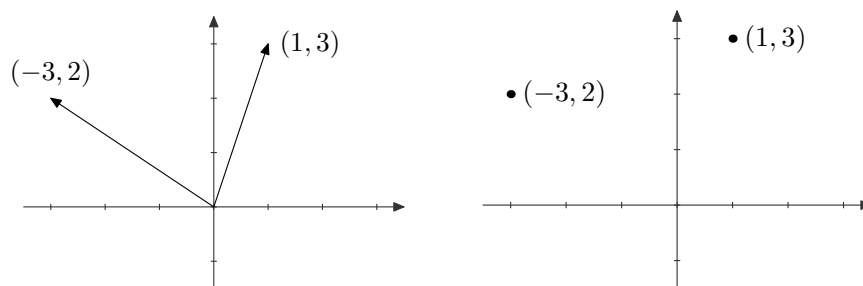
Vektoreita voi merkitä lukulistoina useilla eri tavoilla ja merkintätapa saattaa vaihdella tilanteen mukaan. Olet jo törmännyt kahteen erilaiseen esitystapaan tässä materiaalissa

sekä videossa:

$$(-3, 2) \text{ ja } \begin{bmatrix} -3 \\ 2 \end{bmatrix}.$$

Tästä lähin nuo kaksi merkintää tarkoittavat tässä materiaalissa samaa asiaa ja niiden välillä vaihdellaan edestakaisin saumattomasti.

Monet vektorien ominaisuuksista on helpompi ymmärtää, jos niitä havainnollistaa geometrisesti. Vektoria voikin ajatella origosta lähtevänä nuolena. Toisaalta origosta lähtevän nuolen kohdalla kaikki informaatio sisältyy siihen pisteeseen, jossa nuolen pää on. Nuolen varsi on tavallaan turha. Origosta lähtevä nuoli voidaan siis samaistaa kärkipisteensä kanssa, eli vektorin voi ajatella olevan sama asia kuin piste. Kuvassa 1.1.1 on esitetty vektorit $(-3, 2)$ ja $(1, 3)$ näillä kahdella eri havainnollistustavalla.



Kuva 1.1.1. Vektoreita $(1, 3)$ ja $(-3, 2)$ voi havainnollistaa origosta lähtevänä nuolena tai koordinaatiston pisteinä.

Vektoreita merkitään tässä tekstissä lihavoiduilla kirjaimilla. Voidaan esimerkiksi kirjoittaa $\mathbf{v} = (-4, -10)$. Vektoreille on eri yhteyksissä, aloilla ja oppilaitoksissa käytössä erilaisia merkintöjä kuten $\underline{v}$, $\bar{v}$ ja $\vec{v}$. Tällaiset merkinnät ovat erityisen käteviä käsinkirjoitetussa tekstissä, jossa lihavoitua on vaikea tehdä. Tämän materiaalin kuvat ja videot on tuotettu konteksteissa, joissa vektoreita on tapana merkitä viivalla tai nuolella. Siksi niissä näkyy erilaisia merkintöjä kuin leipätekstissä.

Listamerkintää voi yleistää paljon pidemmälle kuin tason tai kolmiulotteisen avaruuden vektoreihin. Voidaan vaikkapa kirjoittaa 4-ulotteisen avaruuden vektori $(1991, 71, 175, 2)$ tai 6-ulotteisen avaruuden vektori $(13, 252, 46, 168, 164, 110)$. Näistä vektoreista ei enää oikein pysty piirtämään kuvaa koordinaatistoon. Ne voivat kuitenkin sisältää reaali-ilmaan liittyvää informaatiota. Ensimmäinen vektoreista voi vaikkapa sisältää henkilön syntymävuoden, painon, pituuden ja sisarusten lukumäärän. Jälkimmäinen vektori voi vaikkapa kertoa, kuinka monta ihmistä on 6-kerroksisen kauppakeskuksen kussakin kerroksessa jollain ajanhetkellä. Kun tiedot koodataan tähän tapaan vektoreiksi, niiden käsittelemiseen voi sen jälkeen hyödyntää tällä kurssilla opittavia menetelmiä ja tuloksia. Näitä tuloksia voidaan käyttää esimerkiksi silloin, kun selvitetään, miten suomalaisten pituus on yhteydessä heidän syntymävuoteensa, tai kun ennustetaan, miten ihmiset liikkuvat kauppakeskuksen kerroksesta toiseen.

1.2 Vektoriavaruus $\mathbb{R}^n$

Edellä hahmoteltiin, millä eri tavoilla vektoreita voi ajatella. Matematiikassa on kuitenkin tapana sopia käsitteille määritelmät eli sopimukset, jotka kertovat, mitä käsitteellä tarkoitetaan. Määritelmät saattavat vaihdella kontekstista toiseen. Esimerkiksi yhdellä kurssilla saattaa olla käytössä yhdenlainen vektorin määritelmä ja toisella kurssilla toisenlainen. Määritelmän esittäminen takaakin, että kaikki tietävät, mitä käsitteellä tässä yhteydessä tarkalleen ottaen tarkoitetaan.

Määritelmä 1.2.1

Oletetaan, että $n \in \{1, 2, 3, \dots\}$. *Vektori* on n -jono

$$(v_1, v_2, \dots, v_n),$$

missä $v_1, v_2, \dots, v_n \in \mathbb{R}$.

Tässä materiaalissa vektori määritellään siis listana lukuja. Kuten edellä todettiin, lukulistamerkitöjä on useita erilaisia. Tässä materiaalissa käytetään kahta erilaista tapaa:

$$(v_1, v_2, \dots, v_n) \quad \text{ja} \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Vektoreita voidaan merkitä kummalla tahansa näistä tavoista.

Geometrinen tulkinta tulee olemaan jatkuvasti läsnä vektoreita käsiteltäessä, sillä sen avulla pystyy havainnollistamaan vektoreiden ominaisuuksia. Vektoreita voi ajatella origosta lähtevinä nuolina tai koordinaatiston pisteinä. Se, millainen havainnollistamistapa on paras, riippuu siitä, mitä ollaan tekemässä. Yksittäistä vektoria tutkittaessa origosta lähtevä nuoli on yleensä hyödyllisin. Vektoreiden muodostamia joukkoja tutkittaessa puolestaan pistetulkinta on hyödyllinen.

Vektorit muodostavat vektoriavaruuksia. Esimerkiksi muotoa (a, b) olevat vektorit muodostavat vektoriavaruuden $\mathbb{R}^2$ ja muotoa (a, b, c) olevat vektorit muodostavat vektoriavaruuden $\mathbb{R}^3$.

Määritelmä 1.2.2

Vektoriavaruus $\mathbb{R}^n$ on joukko

$$\{(v_1, v_2, \dots, v_n) \mid v_1, v_2, \dots, v_n \in \mathbb{R}\}.$$

Usein termin vektoriavaruus sijasta käytetään lyhyesti ilmaisua *avaruus*.

Vektorin $\mathbf{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ *komponenteiksi* kutsutaan lukuja $v_1, v_2, \dots, v_n$. Esimerkiksi vektorin $\mathbf{v} = (4, -1)$ ovat komponentit ovat 4 ja -1 . Vektorin $\mathbf{u} = (\frac{1}{2}, -\sqrt{50}, 0)$ komponentit ovat puolestaan $\frac{1}{2}$, $-\sqrt{5}$ ja 0. Huomaa, että komponenttien järjestys on vektoreissa oleellinen. Esimerkiksi vektori $(101, 7)$ on eri asia kuin vektori $(7, 101)$. Sovimme, että ellei toisin mainita, vektorin $\mathbf{v}$ komponentteja merkitään symboleilla $v_1, v_2, \dots, v_n$.

Pohdi 1.2.3

Eräässä pikkukaupungissa on liikenneympyrä, johon johtaa neljä tietä. Kaupungin liikennesuunnittelijat seuraavat liikenneympyrän liikennevirtoja. Joka päivä kaupungin järjestelmiin tallentuu liikenneympyrään eri teitä pitkin saapuvien autojen lukumäärä. Liikennemäärät tallennetaan vektoriin, jossa on neljä komponenttia. Kukin komponentti vastaa yhtä liikenneympyrään saapuvista teistä. Eräänä maanantaina liikennettä kuvasi vektori $\mathbf{v} = (150, 400, 200, 80)$. Tiistaina liikennettä kuvasi vektori $\mathbf{w} = (130, 350, 150, 120)$.

1. Suunnittelijat haluavat tutkia yhden päivän sijasta alkuviikon liikennevirtoja. Miten muodostetaan vektori, joka kuvaa maanantaina ja tiistain yhdenlaskettuja liikennevirtoja?
2. Seuraavan viikon maanantaina kaikki liikennevirrat kaksinkertaistuivat. Miten muodostetaan vektori, joka kuvaa seuraavan viikon maanantain liikennevirtoja?

Vektoreille voidaan määritellä erilaisia laskutoimituksia. Käsitellään ensin yhteenlaskua. Yhteenlasku suoritetaan lisäämällä yhteenlaskettavien vektorien komponentit yhteen.

Määritelmä 1.2.4

Oletetaan, että $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{w} \in \mathbb{R}^n$. Tällöin

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix}.$$

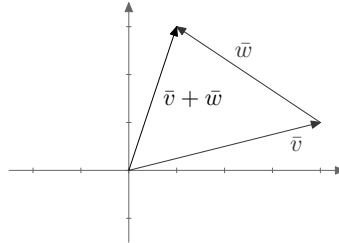
Laskutoimitusta nimitetään vektorien *yhteenlaskuksi*.

Esimerkiksi vektoreiden $\mathbf{v} = (4, 1)$ ja $\mathbf{w} = (-3, 2)$ summa on

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} + \begin{bmatrix} -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 + (-3) \\ 1 + 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Yhteenlaskua voidaan havainnollistaa geometrisesti (kuva 1.2.1). Nyt on hyödyllistä ajatella vektoreita suuntajanoina, joiden paikalla ei ole merkitystä. Ainoastaan suunta ja pituus

merkitsevät. Tällöin vektoreita voi liikutella koordinaatistossa. Vektorien summa nähdään asettamalla vektoreita vastaavat suuntajanat peräkkäin niin, että jälkimmäinen vektori alkaa siitä, mihin ensimmäinen päättyi. Summavektorin alkupiste on ensimmäisen vektorin alkupiste ja päätepiste jälkimmäisen vektorin päätepiste.



Kuva 1.2.1. Vektorit $\mathbf{v}$ ja $\mathbf{w}$ sekä niiden summa $\mathbf{v} + \mathbf{w}$.

Vektoreita voidaan kertoa reaaliluvuilla.

Määritelmä 1.2.5

Oletetaan, että $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{w} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$. Tällöin

$$c\mathbf{v} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}.$$

Laskutoimitusta kutsutaan *skalaarikertolaskuksi*.

Tutkitaan vaikkapa vektoria $\mathbf{v} = (4, 1)$. Nyt

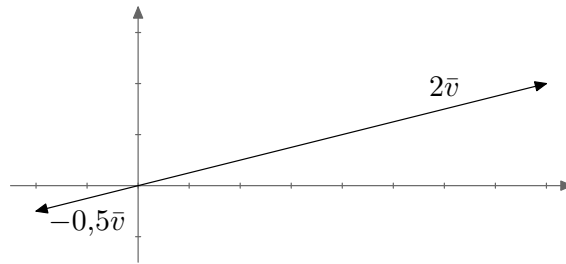
$$2\mathbf{v} = 2 \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 \\ 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

ja

$$-\frac{1}{2}\mathbf{v} = -\frac{1}{2} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \cdot 4 \\ -\frac{1}{2} \cdot 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -\frac{1}{2} \end{bmatrix}.$$

Skalaarimonikerrat $2\mathbf{v}$ ja $-\frac{1}{2}\mathbf{v}$ on piirretty kuvaan 1.2.2. Huomataan, että skalaarikertolasku venyttää tai kutistaa vektoria. Toisin sanoen se *skaalaa* vektoria. Vektorin suunta säilyy samana, jos kerroin on positiivinen ja kääntyy vastakkaiseksi, jos kerron on negatiivinen.

Vektorien yhteydessä reaalilukuja kutsutaan *skalaareiksi*. Jos $\mathbf{v} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$, vektoria $c\mathbf{v}$ nimitetään vektorin $\mathbf{v}$ *skalaarimonikerraksi*. Nimitys juontuu siitä, että reaaliluvulla kertominen skaalaa vektoria.



Kuva 1.2.2. Skalaarimonikerrat $2\mathbf{v}$ ja $-\frac{1}{2}\mathbf{v}$.

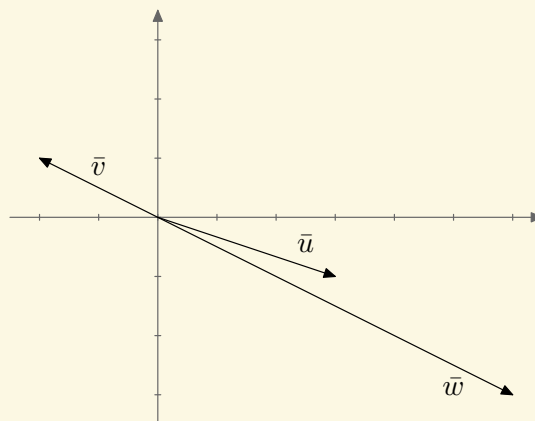
Skalaarikertolasku säilyttää (tai kääntää vastakkaiseksi) vektorin suunnan. Otetaan tämä havainto yleisten vektorien yhdensuuntaisuuden määritelmäksi.

Määritelmä 1.2.6

Vektoriavaruuden $\mathbb{R}^n$ vektorit $\mathbf{v}$ ja $\mathbf{w}$ ovat *yhdensuuntaiset*, jos $\mathbf{v} = r\mathbf{w}$ jollakin vakiolla $r \in \mathbb{R} \setminus \{0\}$. Tällöin merkitään $\mathbf{v} \parallel \mathbf{w}$.

Esimerkki 1.2.7

Tutkitaan vektoreita $\mathbf{v} = (-2, 1)$, $\mathbf{w} = (6, -3)$ ja $\mathbf{u} = (3, -1)$. Ne on esitetty kuvassa 1.2.3.



Kuva 1.2.3. Vektorit $\mathbf{v}$ ja $\mathbf{w}$ ovat yhdensuuntaiset. Vektorit $\mathbf{v}$ ja $\mathbf{u}$ eivät ole yhdensuuntaisia.

Kuvan perusteella vektorit $\mathbf{v} = (-2, 1)$ ja $\mathbf{w} = (6, -3)$ ovat yhdensuuntaiset. Tämä voidaan osoittaa täsmällisesti huomaamalla, että

$$\mathbf{v} = (-2, 1) = -\frac{1}{3}(6, -3) = -\frac{1}{3}\mathbf{w}.$$

Osoitetaan sitten, että vektorit $\mathbf{v} = (-2, 1)$ ja $\mathbf{u} = (3, -1)$ eivät ole yhdensuuntaiset. Tehdään tämä niin kutsutulla epäsuoralla todistuksella. Oletetaan vastoin väitettä, että vektorit ovat yhdensuuntaiset. Tavoitteena on päätyä ristiriitaan. Oletuksen mukaan

on olemassa $r \in \mathbb{R}$, jolle pätee $\mathbf{v} = r\mathbf{u}$. Tästä seuraa, että

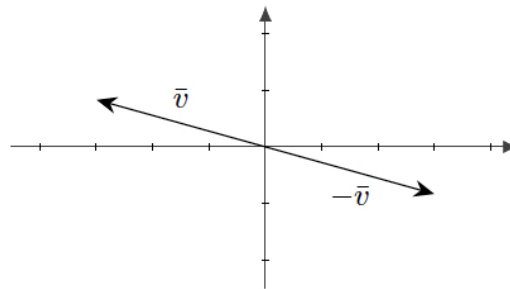
$$\underbrace{(-2, 1)}_{\mathbf{v}} = r \underbrace{(3, -1)}_{\mathbf{u}} = (3r, -r).$$

Siispä $-2 = 3r$ ja $1 = -r$. Ensimmäisen yhtälön mukaan $r = -2/3$, mutta toisen yhtälön mukaan $r = -1$. Tämä on mahdotonta. Olettamalla, että vektorit $\mathbf{v}$ ja $\mathbf{u}$ ovat yhdensuuntaiset, päädyttiin ristiriitaan. Siten vektorit eivät ole yhdensuuntaiset.

Määritelmä 1.2.8

Vektorin $\mathbf{v}$ *vastavektori* on skalaarimonikerta $(-1)\mathbf{v}$. Sitä merkitään $-\mathbf{v}$.

Esimerkiksi vektorin $\mathbf{v} = (-3, 5/6)$ vastavektori on $-\mathbf{v} = (3, -5/6)$. Näitä on havainnollistettu kuvassa 1.2.4. Vastavektori $-\mathbf{v}$ on yhtä pitkä kuin $\mathbf{v}$ ja osoittaa vastakkaiseen suuntaan.



Kuva 1.2.4. Vektori $\mathbf{v}$ ja sen vastavektori $-\mathbf{v}$.

Summan ja vastavektorin avulla voidaan määritellä vektorien erotus.

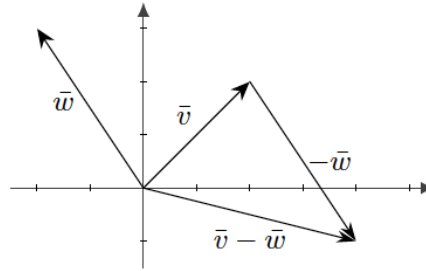
Määritelmä 1.2.9

Vektoreiden $\mathbf{v}$ ja $\mathbf{w}$ *erotus* on summa $\mathbf{v} + (-\mathbf{w})$. Sitä merkitään $\mathbf{v} - \mathbf{w}$.

Esimerkiksi vektorien $\mathbf{v} = (2, 2)$ ja $\mathbf{w} = (-2, 3)$ erotus on

$$\mathbf{v} - \mathbf{w} = (2, 2) - (-2, 3) = (2, 2) + (-1)(-2, 3) = (2, 2) + (2, -3) = (4, -1).$$

Vektoreiden erotus on erikoistapaus vektorien summasta, ja erotuksen voikin määrittää kuvasta samaan tapaan kuin summan (kuva 1.2.5). Määritelmän mukaan vektorien $\mathbf{v}$ ja $\mathbf{w}$ erotus $\mathbf{v} - \mathbf{w}$ saadaan laskemalla yhteen vektori $\mathbf{v}$ ja vastavektori $-\mathbf{w}$. Piirroksessa tämä tarkoittaa sitä, että jälkimmäisen vektorin suunta on käännettävä ennen yhteenlaskua.



Kuva 1.2.5. Vektorit $\mathbf{v}$ ja $\mathbf{w}$ sekä niiden erotus $\mathbf{v} - \mathbf{w}$.

Määritelmä 1.2.10

Vektoria $(0, 0, \dots, 0)$ kutsutaan *nollavektoriksi*. Sille käytetään merkintää $\mathbf{0}$.

Pohdi 1.2.11

Marty McFlyn leijulaudalla pystyy kulkemaan vektorien $(3, -2)$ ja $(4, 0)$ suuntaisesti (eteen ja taaksepäin). Loun kahvila sijaitsee pisteessä $(-4, 4)$.

1. Mitä yhtälö $(-4, 4) = -2(3, -2) + (1/2)(4, 0)$ kertoo Martyn leijulaudan toiminnasta? Mitä vektorien edessä olevat kertoimet kuvaavat?
2. Havainnollista kohdan 1 yhtälöä kuvan avulla.

Lineaarikombinaatioita on havainnollistettu videon ”[Lineaarikombinaatiot, aliavaruus, kantavektori](#)” alussa.

Kun yhteenlasku ja skalaarikertolasku yhdistetään, saadaan aikaiseksi *lineaarikombinaatioita*. Esimerkiksi vektoreiden $\mathbf{v} = (-5, 6)$ ja $\mathbf{w} = (3, 6)$ eräs lineaarikombinaatio on

$$-2\mathbf{v} + (1/3)\mathbf{w} = -2(-5, 6) + (1/3)(3, 6) = (10, 12) + (1, 2) = (11, 14).$$

Määritelmä 1.2.12

Vektori $\mathbf{w} \in \mathbb{R}^n$ on vektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{R}^n$ *lineaarikombinaatio* eli *lineaariyhdistelmä*, jos on olemassa sellaiset reaaliluvut $a_1, a_2, \dots, a_k$, että

$$\mathbf{w} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_k\mathbf{v}_k.$$

Pohdi 1.2.13

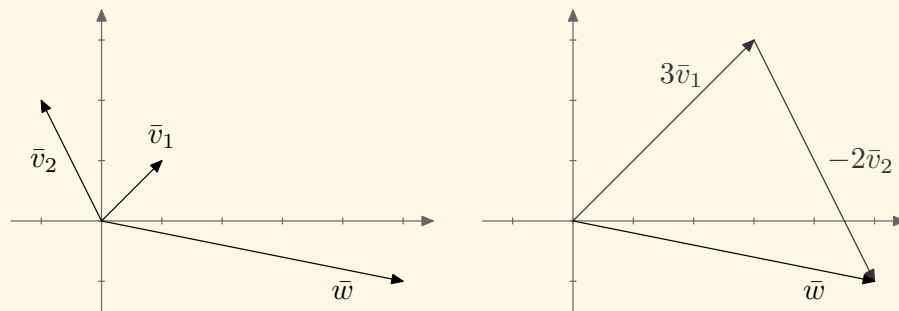
Jasminilla on taikamatto, jota voi ohjata vektorien $(0, 1, -3)$ sekä $(-4, 2, 2)$ suuntaisesti eteen- tai taaksepäin. Jasmin lähtee Bagdadista, jonka koordinaatit ovat $(0, 0, 0)$. Hän haluaa matkustaa naapurikaupunkiin, jonka koordinaatit ovat $(-8, 5, 4)$. Onnistuuko matka Jasminin taikamatolla?

Esimerkki 1.2.14

Merkitään $\mathbf{v}_1 = (1, 1)$, $\mathbf{v}_2 = (-1, 2)$ ja $\mathbf{w} = (5, -1)$. Tutkitaan, onko $\mathbf{w}$ vektoreiden $\mathbf{v}_1$ ja $\mathbf{v}_2$ lineaarikombinaatio. Hetken pohdinnan jälkeen huomataan, että

$$3\mathbf{v}_1 - 2\mathbf{v}_2 = 3(1, 1) - 2(-1, 2) = (3, 3) - (-2, 4) = (5, -1) = \mathbf{w}.$$

Siten $\mathbf{w}$ on vektoreiden $\mathbf{v}_1$ ja $\mathbf{v}_2$ lineaarikombinaatio.



Kuva 1.2.6. Vektori $\mathbf{w}$ on vektoreiden $\mathbf{v}_1$ ja $\mathbf{v}_2$ lineaarikombinaatio.

Edellisessä esimerkissä arvattiin, mitkä kertoimien a_1 ja a_2 pitää olla, jotta pätsi $\mathbf{w} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2$. Läheskään aina arvaaminen ei onnistu. Tällöin tarvitaan tietoa yhtälöryhmien ratkaisemisesta kuten seuraava esimerkki osoittaa.

Esimerkki 1.2.15

Tutkitaan, onko vektori $\mathbf{w} = (-2, 3, 2, -1)$ vektoreiden

$$\mathbf{v}_1 = (0, -1, 2, 1), \quad \mathbf{v}_2 = (2, 0, 1, -1) \quad \text{ja} \quad \mathbf{v}_3 = (4, 2, 2, 0)$$

lineaarikombinaatio. On siis selvitettävä, onko olemassa reaalilukuja x_1, x_2, x_3 , joille pätee

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{w}.$$

Sijoittamalla annetut vektorit yllä olevaan yhtälöön saadaan

$$x_1(0, -1, 2, 1) + x_2(2, 0, 1, -1) + x_3(4, 2, 2, 0) = (-2, 3, 2, -1)$$

ja laskemalla kerto- ja yhteenlaskut auki yhtälö voidaan vielä muuttaa muotoon

$$(2x_2 + 4x_3, -x_1 + 2x_3, 2x_1 + x_2 + 2x_3, x_1 - x_2) = (-2, 3, 2, -1).$$

Vektorit ovat samat täsmälleen silloin, kun niiden kaikki komponentit ovat samat. Kun tarkastellaan jokaista komponenttia erikseen, saatua vektoryhtälöä vastaa yhtälöryh-

mä

$$\begin{cases} 2x_2 + 4x_3 = -2 \\ -x_1 + 2x_3 = 3 \\ 2x_1 + x_2 + 2x_3 = 2 \\ x_1 - x_2 = -1 \end{cases}$$

Miten tällainen yhtälöryhmä ratkaistaan? Ennen voidaan syventyä enemmän lineaarikombinaatioihin, tarvitaan tietoa yhtälönratkaisusta. Yhtälöryhmiin tutustutaan [myöhemmin](#).

Vektoreita, joiden lineaarikombinaatioita käsitellään usein, ovat

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Jokainen avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa näiden vektoreiden avulla. Esimerkiksi avaruudessa $\mathbb{R}^3$ vektori

$$\mathbf{v} = \begin{bmatrix} 5 \\ -11 \\ 6/7 \end{bmatrix}$$

Voidaan kirjoittaa muodossa

$$\mathbf{v} = \begin{bmatrix} 5 \\ -11 \\ 6/7 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 11 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 6/7 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 5\mathbf{e}_1 - 11\mathbf{e}_2 + 6/7\mathbf{e}_3.$$

Yleisesti jokainen avaruuden $\mathbb{R}^n$ vektori $\mathbf{v}$ voidaan kirjoittaa näiden vektoreiden avulla:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + v_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \dots + v_n\mathbf{e}_n.$$

Vektoreita $\mathbf{e}_1, \dots, \mathbf{e}_n$ kutsutaan avaruuden $\mathbb{R}^n$ *luonnolliseksi kannaksi*.

Voidaan osoittaa, että vektoriavaruuden $\mathbb{R}^n$ vektoreille pätevät koulusta tutut laskusäännöt.

Lause 1.2.16

Oletetaan, että $\mathbf{v}, \mathbf{w}, \mathbf{u} \in \mathbb{R}^n$ ja $a, b \in \mathbb{R}$. Tällöin pätee:

1. $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ (vaihdannaisuus)
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (liitännäisyys)
3. $\mathbf{v} + \mathbf{0} = \mathbf{v}$
4. $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
5. $a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w}$ (osittelulaki)
6. $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$ (osittelulaki)
7. $a(b\mathbf{v}) = (ab)\mathbf{v}$
8. $1\mathbf{v} = \mathbf{v}$
9. $0\mathbf{v} = \mathbf{0}$.

Matematiikassa lause tarkoittaa väitettä, joka voidaan perustella todeksi nojautumalla määritelmiin ja aikaisemmin todeksi osoitettuihin väitteisiin.

Todistus. Todistetaan esimerkin vuoksi kohta 1 ja jätetään loput kohdat harjoitustehtäviksi. Oletetaan kuten lauseessa, että $\mathbf{v} \in \mathbb{R}^n$ ja $\mathbf{w} \in \mathbb{R}^n$. Tällöin $\mathbf{v} = (v_1, \dots, v_n)$ ja $\mathbf{w} = (w_1, \dots, w_n)$, ja luvut $v_1, \dots, v_n$ ja $w_1, \dots, w_n$ ovat reaalilukuja. Koska reaalilukujen yhteenlasku on vaihdannainen, jokaisella $i \in \{1, \dots, n\}$ pätee $v_i + w_i = w_i + v_i$. Nyt nähdään, että

$$\begin{aligned}\mathbf{v} + \mathbf{w} &= (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n) \\ &= (w_1 + v_1, w_2 + v_2, \dots, w_n + v_n) = \mathbf{w} + \mathbf{v}.\end{aligned}$$

Väite on todistettu. □

Esimerkki 1.2.17

Oletetaan, että $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Tutkitaan vektoria $\mathbf{z} = 3(2\mathbf{u} + \mathbf{v} + \mathbf{w}) - (\mathbf{u} + 2\mathbf{v} + 3\mathbf{w})$ ja sievennetään sitä vektorien laskusääntöjen avulla:

$$\mathbf{z} = 3(2\mathbf{u} + \mathbf{v} + \mathbf{w}) - (\mathbf{u} + 2\mathbf{v} + 3\mathbf{w}) = 6\mathbf{u} + 3\mathbf{v} + 3\mathbf{w} - \mathbf{u} - 2\mathbf{v} - 3\mathbf{w} = 5\mathbf{u} + \mathbf{v}.$$

 Tiivistelmä

- Vektorit ovat järjestettyjä lukulistoja.
- Vektoreita voi havainnollistaa eri tavoin: origosta lähtevänä nuolena tai koordinaatiston pisteinä.
- Vektorit muodostavat vektoriavaruuksia.

- Vektoreita voi laskea yhteen ja kertoa reaaliluvuilla (skalaarikertolasku).
- Kun vektoreita lasketaan yhteen, niitä kannattaa havainnollistaa suuntajanoina, joiden sijainnilla ei ole väliä.
- Lineaarikombinaatioissa yhdistyvät yhteenlasku ja skalaarikertolasku.
- Vektorit ovat yhdensuuntaiset, jos ne saadaan toisistaan skalaarilla kertomalla.

1.3 Pistetulo ja normi

Avaruuksissa $\mathbb{R}^2$ ja $\mathbb{R}^3$ on totuttu puhumaan vektorien pituuksista ja vektoreiden välisestä kulmista. Näiden käsitteiden yleistäminen korkeampiulotteisiin avaruuksiin ei välttämättä onnistu pelkästään geometrisen intuition avulla. Kuitenkin esimerkiksi Pythagoraan lauseen voidaan ajatella toimivan kaikissa ulottuvuuksissa samalla tavalla. Kyseinen lause, samoin kuin muutkin vektoreiden pituuksiin ja kulmiin liittyvät käsitteet, voidaan ilmaista pistetulon avulla, ja pistetulo puolestaan voidaan laskea avaruudessa $\mathbb{R}^n$, oli n miten suuri tahansa.

Määritelmä 1.3.1

Vektoreiden $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$ ja $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{R}^n$ *pistetulo* on

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n.$$

Vektorien pistetulo on aina reaaliluku eli skalaari. Pistetuloa kutsutaankin myös *skalaarituloksi*.

Esimerkki 1.3.2

Vektorien $\mathbf{v} = (3, -2, 0)$ ja $\mathbf{w} = (1, -2, \sqrt{3})$ *pistetulo* on

$$\mathbf{v} \cdot \mathbf{w} = 3 \cdot 1 + (-2)(-2) + 0 \cdot \sqrt{3} = 7.$$

Tulemme näkemään, että pistetulo kertoo, osoittavatko vektorit samaan suuntaan vai eri suuntaan. Jos vektorien pistetulo on positiivinen, vektorit osoittavat samaan suuntaan. Jos pistetulo on nolla, vektorit ovat kohtisuorassa toisiaan vastaan. Negatiivinen pistetulo puolestaan tarkoittaa, että vektorit osoittavat eri suuntaan. Tähän palataan tarkemmin vektorien välisen kulman määritelmän yhteydessä ([määritelmä 1.3.16](#)).

Pistetulolle voidaan johtaa laskusääntöjä.

Lause 1.3.3

Oletetaan, että $\mathbf{v}, \mathbf{w}, \mathbf{u} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$. Tällöin

1. $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

$$2. \mathbf{v} \cdot (\mathbf{w} + \mathbf{u}) = \mathbf{v} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{u}$$

$$3. (c\mathbf{v}) \cdot \mathbf{w} = c(\mathbf{v} \cdot \mathbf{w}).$$

Todistus. Todistetaan vain kohta 2 ja jätetään loput kohdat harjoitustehtäviksi. Merkitään $\mathbf{v} = (v_1, \dots, v_n)$, $\mathbf{w} = (w_1, \dots, w_n)$ ja $\mathbf{u} = (u_1, \dots, u_n)$. Nyt nähdään, että

$$\begin{aligned} \mathbf{v} \cdot (\mathbf{w} + \mathbf{u}) &= (v_1, \dots, v_n) \cdot (w_1 + u_1, w_2 + u_2, \dots, w_n + u_n) \\ &= v_1(w_1 + u_1) + v_2(w_2 + u_2) + \dots + v_n(w_n + u_n) \\ &= v_1w_1 + v_1u_1 + v_2w_2 + v_2u_2 + \dots + v_nw_n + v_nu_n \\ &= (v_1w_1 + v_2w_2 + \dots + v_nw_n) + (v_1u_1 + v_2u_2 + \dots + v_nu_n) \\ &= \mathbf{v} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{u}. \end{aligned}$$

Tässä käytettiin reaalilukujen yhteenlaskun ja kertolaskun osittelulakia. □

Kohdan 2 osittelulaki pätee myös toisin päin: $(\mathbf{v} + \mathbf{w}) \cdot \mathbf{u} = \mathbf{v} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{u}$. Tämä seuraa kohdasta 1, jonka mukaan pistetulo on vaihdannainen:

$$(\mathbf{v} + \mathbf{w}) \cdot \mathbf{u} \stackrel{1}{=} \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) \stackrel{2}{=} \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \stackrel{1}{=} \mathbf{v} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{u}.$$

Samaan tapaan myös kohta 3 voidaan kääntää muotoon $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$.

Seuraava lause osoittaa, että vektorin pistetulo itsensä kanssa on aina epänegatiivinen. Ainoastaan nollavektorin pistetulo itsensä kanssa on nolla.

Lause 1.3.4

Oletetaan, että $\mathbf{v} \in \mathbb{R}^n$. Tällöin

1. $\mathbf{v} \cdot \mathbf{v} \geq 0$
2. $\mathbf{v} \cdot \mathbf{v} = 0$, jos ja vain jos $\mathbf{v} = \mathbf{0}$.

Todistus. Todistetaan kohdat yksi kerrallaan.

1. Nähdään, että $\mathbf{v} \cdot \mathbf{v} = v_1^2 + v_2^2 + \dots + v_n^2 \geq 0 + 0 + \dots + 0 = 0$, sillä reaaliluvun neliö on aina epänegatiivinen. Tämä todistaa väitteen.
2. "⇒": Oletetaan, että $\mathbf{v} \cdot \mathbf{v} = 0$. Tällöin $v_1^2 + v_2^2 + \dots + v_n^2 = 0$. Koska jokainen yhteenlaskettava on epänegatiivinen, täytyy yhteenlaskettavien olla nollija. Toisin sanoen $v_i^2 = 0$ kaikilla $i \in \{1, \dots, n\}$. Tästä seuraa, että $v_i = 0$ kaikilla $i \in \{1, \dots, n\}$. Siten $\mathbf{v} = (0, 0, \dots, 0) = \mathbf{0}$.

"⇐": Oletetaan, että $\mathbf{v} = \mathbf{0}$. Nyt $\mathbf{v} \cdot \mathbf{v} = 0^2 + 0^2 + \dots + 0^2 = 0$. □

Esimerkki 1.3.5

Luonnollisen kannan vektoreille on voimassa

$$\mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1, & \text{kun } i = j \\ 0, & \text{kun } i \neq j. \end{cases}$$

Jos merkitään

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

Voidaan kirjoittaa $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. Funktiota δ_{ij} kutsutaan *Kroneckerin delta-symboliksi* tai lyhyesti *Kroneckerin deltaksi*.

Vektorin normi

Pistetulon avulla voidaan määritellä avaruuden $\mathbb{R}^n$ vektorin normi eli pituus. [Lauseen 1.3.4](#) nojalla pätee $\mathbf{v} \cdot \mathbf{v} \geq 0$, joten seuraavassa määritelmässä juurrettava on epänegatiivinen, kuten kuuluu olla.

Määritelmä 1.3.6

Vektorin $\mathbf{v} \in \mathbb{R}^n$ normi eli pituus on

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

Määritelmästä seuraa, että $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$, kun pistetulo lasketaan auki. Lisäksi $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$.

Esimerkki 1.3.7

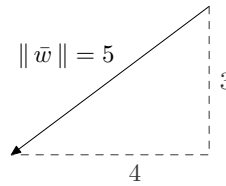
Vektorin $\mathbf{v} = (1/2, 3, -2, 0)$ normi on

$$\|\mathbf{v}\| = \sqrt{(1/2)^2 + 3^2 + (-2)^2 + 0^2} = \sqrt{\frac{53}{4}} = \frac{\sqrt{53}}{2}.$$

Tasossa normia voi havainnollistaa Pythagoraan lauseen avulla. Kuvaan 1.3.1 on piirretty vektori $\mathbf{w} = (-4, -3)$. Pythagoraan lausetta käyttäen sen pituudeksi saadaan

$$\sqrt{3^2 + 4^2} = \sqrt{25} = 5.$$

Pituuden geometrinen tulkinta antaa siis saman tuloksen kuin normin määritelmä.



Kuva 1.3.1. Vektorin $\mathbf{w} = (-4, -3)$ normi eli pituus on 5.

Seuraava lause ilmaisee normien avulla sen, että vektorin pituus on aina epänegatiivinen ja nollavektori on ainoa vektori, jonka pituus on nolla.

Lause 1.3.8

Oletetaan, että $\mathbf{v} \in \mathbb{R}^n$. Tällöin

1. $\|\mathbf{v}\| \geq 0$
2. $\|\mathbf{v}\| = 0$, jos ja vain jos $\mathbf{v} = \mathbf{0}$.

Todistus. Tulokset seuraavat suoraan normin määritelmästä, neliöjuuren ominaisuuksista ja lauseesta 1.3.4.

1. Määritelmän mukaan $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. Neliöjuuren arvo on aina epänegatiivinen, joten $\|\mathbf{v}\| \geq 0$.
2. Huomataan, että $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = 0$, jos ja vain jos juurrettava $\mathbf{v} \cdot \mathbf{v}$ on nolla. Lauseen 1.3.4 nojalla taas $\mathbf{v} \cdot \mathbf{v} = 0$, jos ja vain jos $\mathbf{v} = \mathbf{0}$. Tämä todistaa väitteen. $\square$

Skalaarikertolaskun määritelmän yhteydessä todettiin, että kun vektoria kerrotaan skalaarilla, vektorin suunta pysyy samana tai kääntyy vastakkaiseksi, mutta sen pituutta ”skaalataan”. Asian voi ilmaista myös normin avulla: kun vektoria kerrotaan skalaarilla, vektorin pituus tulee kerrotuksi tuolla samalla skalaarilla. On vain otettava huomioon, että pituus on aina epänegatiivinen, joten skalaarista pitää ottaa itseisarvo. Seuraava lause ilmaisee asian täsmällisesti, mutta sen todistaminen jätetään harjoitustehtäväksi.

Lause 1.3.9

Oletetaan, että $\mathbf{v} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$. Tällöin $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$.

Jos vain vektorin suunnalla on merkitystä, pyritään usein yksinkertaisuuden vuoksi rajoittamaan vektoreihin, joiden pituus on yksi. Tällaisia vektoreita kutsutaan yksikkövektoreiksi.

Määritelmä 1.3.10

Vektori $\mathbf{u} \in \mathbb{R}^n$ on *yksikkövektori*, jos sen normi on yksi eli

$$\|\mathbf{u}\| = 1.$$

Avaruuden $\mathbb{R}^2$ vektorit $(1, 0)$ ja $(0, 1)$ ovat yksikkövektoreita, sillä niiden pituus on yksi. Yleisemmin kaikki luonnollisen kannan vektorit $\mathbf{e}_i \in \mathbb{R}^n$ ovat yksikkövektoreita. On kuitenkin olemassa paljon muitakin yksikkövektoreita, kuten seuraava esimerkki osoittaa.

Esimerkki 1.3.11

Etsitään jokin yksikkövektori, joka on yhdensuuntainen vektorin $\mathbf{v} = (2, -1, 0)$ kanssa. Vektorin $\mathbf{v}$ normi on $\|\mathbf{v}\| = \sqrt{4 + 1} = \sqrt{5}$. Jos vektori $\mathbf{v}$ kerrotaan skalaarilla $1/\sqrt{5}$, saadaan vektori $(1/\sqrt{5})\mathbf{v}$, jonka pituus on lauseen 1.3.9 nojalla

$$\frac{1}{\sqrt{5}} \|\mathbf{v}\| = \frac{1}{\sqrt{5}} \cdot \sqrt{5} = 1.$$

Vektori $(1/\sqrt{5})\mathbf{v}$ on siis yksikkövektori. Lisäksi vektorit $\mathbf{v}$ ja $(1/\sqrt{5})\mathbf{v}$ ovat yhdensuuntaiset, koska ne eroavat vain skalaarikertoimen verran. Eräs tavoiteltu vektori on siis $(1/\sqrt{5})\mathbf{v}$.

Edellistä esimerkkiä mukaillen saadaan seuraava yleinen tulos.

Lause 1.3.12

Oletetaan, että $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Tällöin vektori $\frac{1}{\|\mathbf{v}\|}\mathbf{v}$ on yksikkövektori, joka on samansuuntainen vektorin $\mathbf{v}$ kanssa.

Todistus. Todistus jätetään harjoitustehtäväksi. □

Vektorien välinen kulma ja etäisyys

Pistetulon avulla voidaan määritellä vektorin pituuden lisäksi vektorien välinen kulma. Tarkastellaan aluksi yksinkertaisinta eli suoraa kulmaa.

Määritelmä 1.3.13

Vektorit $\mathbf{v} \in \mathbb{R}^n$ ja $\mathbf{w} \in \mathbb{R}^n$ ovat *kohtisuorassa toisiaan vastaan* eli *ortogonaaliset*, jos $\mathbf{v} \cdot \mathbf{w} = 0$. Tällöin merkitään $\mathbf{v} \perp \mathbf{w}$.

Esimerkki 1.3.14

Tason vektorit $(2, 1)$ ja $(-2, 4)$ ovat kohtisuorassa toisiaan vastaan, sillä

$$(2, 1) \cdot (-2, 4) = 2 \cdot (-2) + 1 \cdot 4 = 0.$$

Voidaan siis merkitä $(2, 1) \perp (-2, 4)$.

Vektorit $(2, 1)$ ja $(3, 4)$ puolestaan eivät ole kohtisuorassa toisiaan, sillä

$$(2, 1) \cdot (3, 4) = 2 \cdot 3 + 1 \cdot 4 = 10 \neq 0.$$

Esimerkki 1.3.15

Määritetään kaikki ne vektorit, jotka ovat kohtisuorassa vektoria $\mathbf{n} = (3, -2)$ vastaan. Vektori $(x, y) \in \mathbb{R}^2$ on kohtisuorassa vektoria $\mathbf{n}$, jos ja vain jos vektorien pistetulo on nolla. Toisin sanoen täytyy päteä

$$3x - 2y = 0.$$

Tämä yhtälö muuntuu muotoon $y = \frac{3}{2}x$. Vektoria $\mathbf{n}$ vastaan kohtisuorat vektorit ovat siis muotoa $\mathbf{x} = \left(x, \frac{3}{2}x\right) = x \left(1, \frac{3}{2}\right)$, missä $x \in \mathbb{R}$. Vektorien kärkipisteet muodostavat origon kautta kulkevan suoran, jonka yhtälö on $y = \frac{3}{2}x$.

Pistetulon avulla voi määrittää vektorien välisen kulman myös yleisemmin kuin suoran kulman tapauksessa.

Määritelmä 1.3.16

Vektorien $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ ja $\mathbf{w} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ välinen *kulma* on se kulma α , jolle pätee $0 \leq \alpha \leq \pi$ ja

$$\cos \alpha = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|}.$$

Esimerkki 1.3.17

Vektorien $\mathbf{v} = (3, -2, 0)$ ja $\mathbf{w} = (1, -2, \sqrt{3})$ välinen kulma α saadaan yhtälöstä

$$\cos \alpha = \frac{7}{\sqrt{13}\sqrt{8}}.$$

Lisäksi täytyy päteä $0 \leq \alpha \leq \pi$. Laskimella saadaan vektorien välisen kulman likiarvoksi $\alpha \approx 0,81$.

Vektorien välinen kulma antaa pistetulolle geometrisen merkityksen. Määritelmän mukaan vektorien $\mathbf{v}$ ja $\mathbf{w}$ pistetulolle pätee

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos \alpha,$$

missä α on vektorien välinen kulma. Tarkastellaan erilaisia tapauksia.

- $0 \leq \alpha < \pi/2$: Tällöin $\cos \alpha > 0$, joten vektorien pistetulo on positiivinen. Mitä pienempi kulma α on, sitä suurempi on kosini ja sitä kautta pistetulo.

- $\alpha = 0$: Tällöin $\cos \alpha = 1$, joten vektorien pistetulo on nolla. Vektorit ovat toisiaan vastaan kohtisuorassa. Vektorien välisen kulman määritelmä on siis linjassa aiemmin esitetyn kohtisuoruuden määritelmän kanssa.
- $\pi/2 < \alpha \leq \pi$: Tällöin $\cos \alpha < 0$, joten vektorien pistetulo on negatiivinen. Mitä suurempi kulma α on, pienempi on kosini ja sitä kautta pistetulo.

Edelliset kohdat voi ilmaista myös toisilla sanoilla. Vektorien välinen kulma kertoo, kuinka paljon vektorit osoittavat samaan suuntaan. Pistetulo on sitä suurempi, mitä enemmän vektorit osoittavat samaan suuntaan ja sitä pienempi, mitä enemmän ne osoittavat eri suuntiin.

Vektorien välisen kulman määritelmä vaatii itse asiassa hiukan lisähuomiota. Ei nimittäin ole itsestään selvää, että määritelmässä mainitun kulman α voi laskea jokaiselle vektorille. Kosinifunktio on määritelty niin, että jokaista lukua $x \in [-1, 1]$ vastaa täsmälleen yksi sellainen kulma α , että $0 \leq \alpha \leq \pi$ ja $\cos \alpha = x$. Ennen kuin määritelmän voi ottaa virallisesti käyttöön, täytyy varmistua siitä, että

$$-1 \leq \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|} \leq 1$$

kaikilla vektoreilla $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Tätä varten tarvitaan muutamia aputuloksia.

Ensimmäinen aputulos on eräs yleisen teorian kannalta tärkeä tulos, jota tässä vaiheessa kuitenkin tarvitaan lähinnä sitä seuraavan lemmän todistamiseen.

Lause 1.3.18 (Cauchyn-Schwarzin epäyhtälö)

Oletetaan, että $\mathbf{v} \in \mathbb{R}^n$ ja $\mathbf{w} \in \mathbb{R}^n$. Tällöin

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

Todistus. Lauseen pystyy todistamaan kätevästi projektion käsitteen avulla, joten todistus esitetään vasta [projektion yhteydessä](#). □

Lemma 1.3.19

Oletetaan, että $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Tällöin

$$-1 \leq \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|} \leq 1.$$

Todistus. Schwarzin epäyhtälön mukaan $|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|$. Tästä seuraa, että

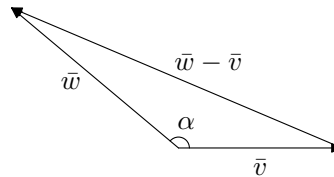
$$-\|\mathbf{v}\| \|\mathbf{w}\| \leq \mathbf{v} \cdot \mathbf{w} \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

Jakamalla näin saadut epäyhtälöt positiivisella luvulla $\|\mathbf{v}\|\|\mathbf{w}\|$ saadaan

$$-1 \leq \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} \leq 1. \quad \square$$

Lemman nojalla tiedetään, että vektorien välinen kulma on mahdollista määrittää kaikilla vektoreilla.

Vaikka määritelmiä ei tarvitsekaan perustella mitenkään, on kuitenkin valaisevaa katsoa, miten vektorien välisen kulman määritelmä vastaa tasossa geometrista käsitystämme.



Kuva 1.3.2. Vektoreiden $\mathbf{v}$ ja $\mathbf{w}$ välinen kulma kosinilauseen näkökulmasta.

Kosinilauseen mukaan kuvan 1.3.2 kolmiossa pätee

$$\|\mathbf{w} - \mathbf{v}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\|\cos \alpha.$$

Toisaalta normin määritelmän ja pistetulon ominaisuuksien nojalla

$$\|\mathbf{w} - \mathbf{v}\|^2 = (\mathbf{w} - \mathbf{v}) \cdot (\mathbf{w} - \mathbf{v}) = \mathbf{w} \cdot \mathbf{w} - \mathbf{w} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2 - 2(\mathbf{v} \cdot \mathbf{w}) + \|\mathbf{w}\|^2.$$

Saadaan siis yhtälö

$$\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\|\cos \alpha = \|\mathbf{v}\|^2 - 2(\mathbf{v} \cdot \mathbf{w}) + \|\mathbf{w}\|^2,$$

josta edelleen

$$\cos \alpha = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}.$$

Vektorien pistetulon ja normin avulla voidaan todistaa koulusta tuttu Pythagoraan lause. Todistus jätetään harjoitustehtäväksi.

Lause 1.3.20 (Pythagoraan lause)

Oletetaan, että vektorit $\mathbf{v} \in \mathbb{R}^n$ ja $\mathbf{w} \in \mathbb{R}^n$ ovat kohtisuorassa toisiaan vastaan. Tällöin

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$$

Kun vektorit $\mathbf{v}$ ja $\mathbf{w}$ tulkitaan tason pisteiksi, niiden välinen etäisyys voidaan määritellä niitä yhdistävän suuntajanan $\mathbf{v} - \mathbf{w}$ pituutena. Tämä taas palautuu vektorin normiin.

Määritelmä 1.3.21

Oletetaan, että $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Vektorien $\mathbf{v}$ ja $\mathbf{w}$ välinen etäisyys on

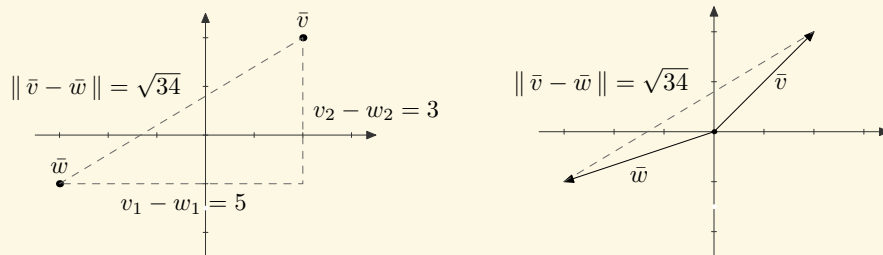
$$d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|.$$

Esimerkki 1.3.22

Vektoreiden $\mathbf{v} = (2, 2)$ ja $\mathbf{w} = (-3, -1)$ välinen etäisyys on

$$d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\| = \|(2 - (-3), 2 - (-1))\| = \|(5, 3)\| = \sqrt{5^2 + 3^2} = \sqrt{34}.$$

Etäisyyttä on havainnollistettu kahdella eri tavalla seuraavassa kuvassa.



Kuva 1.3.3. Vektoreiden $\mathbf{v}$ ja $\mathbf{w}$ välinen etäisyys. Ensimmäisessä kuvassa vektorit on havainnollistettu tason pisteinä, jälkimmäisessä kuvassa origosta lähtevinä nuolina.

Tiivistelmä

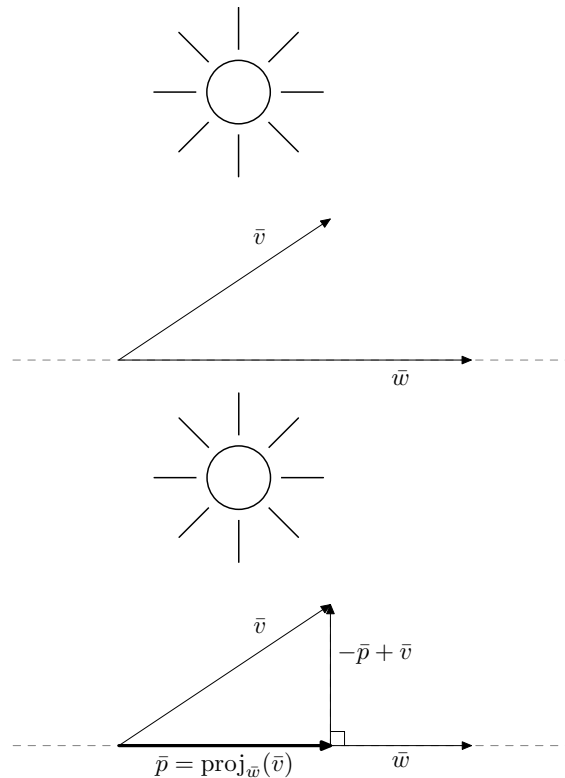
- Vektoreiden pistetulo on operaatio, joka tuottaa kahdesta vektorista tulokseksi reaaliarvon.
- Vektorin normi eli pituus määritellään pistetulon avulla.
- Vektoreiden välinen kulma voidaan määrittää pistetulon avulla.
- Vektorit ovat kohtisuorassa toisiaan vastaan, jos niiden pistetulo on nolla.

1.4 Projektio suoralle

Kohtisuoruuden käsite tarjoaa mahdollisuuden määrittellä vektorin *projektio*. Tätä voidaan ajatella vektorin heittämänä varjona kuvan 1.4.1 mukaan. Projektiio on hyödyllinen työkalu esimerkiksi tietokonegrafiikassa ja geometriassa (ks. [esimerkki 2.1.6](#)). Projektion avulla voidaan hajottaa vektori kahden toisiaan vastaan kohtisuorassa olevan vektorin summaksi, mikä on hyödyllistä fysiikan voimavektoreita käsiteltäessä.

Tarkastellaan varjovertausta tarkemmin. Voidaan ajatella, että vektorin $\mathbf{v}$ projektiio vektorin $\mathbf{w}$ suuntaiselle suoralle on vektorin $\mathbf{v}$ heittäminen varjo, kun aurinko paistaa kohtisuoraan suoraan vastaan kuten kuvassa 1.4.1. Matemaattisesti tämä voidaan ilmaista sanomalla,

että projektio $\mathbf{p}$ on yhdensuuntainen vektorin $\mathbf{w}$ kanssa ja vektorit $\mathbf{w}$ sekä $-\mathbf{p} + \mathbf{v}$ ovat kohtisuorassa toisiaan vastaan.



Kuva 1.4.1. Projektiota voi ajatella vektorin heittämänä varjona.

Määritelmä 1.4.1

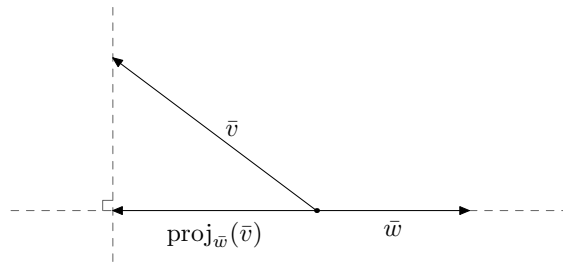
Olkoot $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ ja $\mathbf{w} \neq \mathbf{0}$. Tällöin vektorin $\mathbf{v}$ *projektio* vektorin $\mathbf{w}$ virittämälle suoralle on sellainen vektori $\mathbf{p} \in \mathbb{R}^n$, että

1. vektori $\mathbf{p}$ on yhdensuuntainen vektorin $\mathbf{w}$ kanssa
2. vektori $\mathbf{v} - \mathbf{p}$ on kohtisuorassa vektoria $\mathbf{w}$ vastaan.

Projektiota $\mathbf{p}$ merkitään $\text{proj}_{\mathbf{w}}(\mathbf{v})$.

Vektoria $\text{proj}_{\mathbf{w}}(\mathbf{v})$ kutsutaan myös lyhyesti vektorin $\mathbf{v}$ projektioksi vektorille $\mathbf{w}$.

Vektorin projektion voi määrittää kuvan avulla seuraavasti. Piirretään vektorit $\mathbf{v}$ ja $\mathbf{w}$ alkamaan samasta pisteestä ja piirretään vektorin $\mathbf{w}$ suuntainen suora (ks. kuva 1.4.2). Projektio $\text{proj}_{\mathbf{w}}(\mathbf{v})$ löydetään piirtämällä suora, joka on kohtisuorassa vektorin $\mathbf{w}$ suuntaista suoraa vastaan ja kulkee vektorin $\mathbf{v}$ kärjen kautta.



Kuva 1.4.2. Vektorin $\mathbf{v}$ projektion määrittäminen piirtämällä.

Määritelmän avulla projektioista voi piirtää kuvan, mutta projektion laskeminen määritelmästä lähtien on yleensä hankalaa. Siinä auttaa projektion laskukaava.

Lause 1.4.2

Olkoot $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ ja $\mathbf{w} \neq \mathbf{0}$. Tällöin on olemassa täsmälleen yksi projektio $\text{proj}_{\mathbf{w}}(\mathbf{v})$ ja se saadaan kaavasta

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}.$$

Käyttämällä normin määritelmää projektion kaavan voi kirjoittaa myös muodossa

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|^2} \mathbf{w}.$$

Todistus. Aloitetaan osoittamalla, että vektori

$$\frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}$$

täyttää projektion määritelmässä mainitut ehdot.

Ensinnäkin huomataan, että kerroin $(\mathbf{v} \cdot \mathbf{w})/(\mathbf{w} \cdot \mathbf{w})$ on reaaliluku, sillä vektoreiden pistetulot ovat reaalilukuja. Vektori

$$\frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}$$

on siis vektorin $\mathbf{w}$ skalaarimonikerta. Siten projektion määritelmän ensimmäinen ehto täyttyy.

Tutkitaan sitten projektion määritelmän toista ehtoa. On osoitettava, että vektorit

$$\mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}$$

ja $\mathbf{w}$ ovat kohtisuorassa toisiaan vastaan. Pistetulon laskusääntöjen nojalla

$$\left(\mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} \right) \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} - \left(\frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} \right) \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} (\mathbf{w} \cdot \mathbf{w}).$$

Pistetulosta tulee aina tulokseksi reaaliluku, joten seuraavaksi voidaan käyttää reaaliluku-

jen laskusääntöjä:

$$\mathbf{v} \cdot \mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}}(\mathbf{w} \cdot \mathbf{w}) = \mathbf{v} \cdot \mathbf{w} - \frac{(\mathbf{v} \cdot \mathbf{w})(\mathbf{w} \cdot \mathbf{w})}{\mathbf{w} \cdot \mathbf{w}} = \mathbf{v} \cdot \mathbf{w} - \mathbf{v} \cdot \mathbf{w} = 0.$$

Koska pistetuloksi saatiin nolla, ovat vektorit kohtisuorassa toisiaan vastaan. Siten

$$\frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}$$

on vektorin $\mathbf{v}$ projektio vektorin $\mathbf{w}$ virittämälle suoralle eli

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}.$$

Osoitetaan seuraavaksi, että projektion ehdot täyttäviä vektoreita on vain yksi. Merkitään $\mathbf{p} = \text{proj}_{\mathbf{w}}(\mathbf{v})$. Projektion määritelmän ensimmäisen kohdan perusteella $\mathbf{p} = t\mathbf{w}$ jollakin $t \neq 0$. Määritelmän toisen kohdan mukaan vektori $\mathbf{v} - \mathbf{p}$ on kohtisuorassa vektoria $\mathbf{w}$ vastaan, joten $(\mathbf{v} - \mathbf{p}) \cdot \mathbf{w} = 0$. Pistetuloksi saadaan

$$(\mathbf{v} - \mathbf{p}) \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} - \mathbf{p} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} - t\mathbf{w} \cdot \mathbf{w}.$$

Päädytään yhtälöön $\mathbf{v} \cdot \mathbf{w} - t(\mathbf{w} \cdot \mathbf{w}) = 0$, josta saadaan ratkaistua $t = (\mathbf{v} \cdot \mathbf{w})/(\mathbf{w} \cdot \mathbf{w})$. Nyt siis

$$\mathbf{p} = t\mathbf{w} = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}.$$

Väite on todistettu. □

Esimerkki 1.4.3

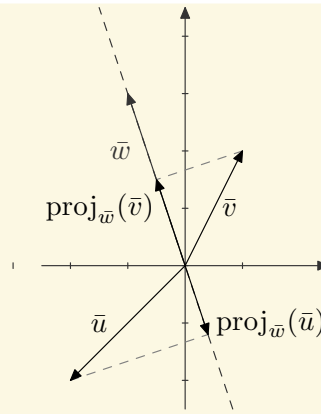
Esimerkiksi vektorin $\mathbf{v} = (1, 2)$ projektio vektorin $\mathbf{w} = (-1, 3)$ virittämälle suoralle on

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} = \frac{5}{10}(-1, 3) = \frac{1}{2}(-1, 3) = \left(-\frac{1}{2}, \frac{3}{2}\right).$$

Vektorin $\mathbf{u} = (-2, -2)$ projektio vektorin $\mathbf{w}$ virittämälle suoralle on puolestaan

$$\text{proj}_{\mathbf{w}}(\mathbf{u}) = \frac{-4}{10}(-1, 3) = -\frac{2}{5}(-1, 3) = \left(\frac{2}{5}, -\frac{6}{5}\right).$$

Projektiot on esitetty kuvassa 1.4.3.



Kuva 1.4.3. Vektoreiden $\mathbf{v}$ ja $\mathbf{u}$ projektiot vektorin $\mathbf{w}$ virittämälle suoralle.

Jos on annettu vektorit $\mathbf{v} \in \mathbb{R}^n$ ja $\mathbf{w} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, vektori $\mathbf{v}$ voidaan kirjoittaa summana vektoreista $\mathbf{w}$ sekä $-\text{proj}_{\mathbf{w}}(\mathbf{v}) + \mathbf{v}$. (Ks. kuva 1.4.1.) Nämä vektorit ovat projektion määrittelyn mukaan kohtisuorassa toisiaan vastaan. Vektori $\mathbf{v}$ voidaan siis hajottaa summa-
maksiksi kahdesta toisiaan vastaan kohtisuorassa olevasta vektorista. Vektoria $-\text{proj}_{\mathbf{w}}(\mathbf{v}) + \mathbf{v}$ kutsutaan vektorin $\mathbf{v}$ kohtisuoraksi komponentiksi vektorin $\mathbf{w}$ suhteen.

Esimerkki 1.4.4

Määritetään vektorin $\mathbf{v} = (2, 7, -11)$ kohtisuora komponentti vektorin $\mathbf{w} = (3, 0, -4)$ suhteen.

Vektorin $\mathbf{v}$ projektiio vektorin $\mathbf{w}$ virittämälle suoralle on

$$\text{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} = \frac{2 \cdot 3 + 7 \cdot 0 + (-11)(-4)}{3^2 + 0^2 + (-4)^2} \mathbf{w} = 2\mathbf{w} = \begin{bmatrix} 6 \\ 0 \\ -8 \end{bmatrix},$$

jolloin kohtisuoraksi komponentiksi saadaan

$$\mathbf{v} - \text{proj}_{\mathbf{w}}(\mathbf{v}) = \mathbf{v} - 2\mathbf{w} = \begin{bmatrix} -4 \\ 7 \\ -3 \end{bmatrix}.$$

Laskemalla voi vielä tarkistaa, että $(2, 7, -11) = (6, 0, -8) + (-4, 7, -3)$, joten $\mathbf{v}$ on tosiaan löydettyjen vektorien summa. Lisäksi $(6, 0, -8) \cdot (-4, 7, -3) = 0$, joten vektorit ovat kohtisuorassa toisiaan vastaan.

Edellisessä luvussa esitettiin Cauchyn-Schwarzin epäyhtälö (lause 1.3.18), mutta todistusta lykättiin myöhemmäksi. Todistetaan epäyhtälö nyt vektoriprojektion avulla.

Todistus. Oletetaan, että $\mathbf{v} \in \mathbb{R}^n$ ja $\mathbf{w} \in \mathbb{R}^n$. Osoitetaan, että

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

Tutkitaan ensin tapaus $\mathbf{w} = \mathbf{0}$. Tässä tapauksessa epäyhtälön molempien puolten lausekkeiden arvot ovat 0, ja epäyhtälö pätee.

Oletetaan sitten, että $\mathbf{w} \neq \mathbf{0}$. Nyt voidaan projisoida vektorin $\mathbf{v}$ vektorin $\mathbf{w}$ virittämälle suoralle. Projektiovektorin pituus on

$$\|\text{proj}_{\mathbf{w}}(\mathbf{v})\| = \left\| \frac{\mathbf{w} \cdot \mathbf{v}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} \right\| = \frac{|\mathbf{w} \cdot \mathbf{v}|}{\|\mathbf{w} \cdot \mathbf{w}\|} \|\mathbf{w}\| = \frac{|\mathbf{w} \cdot \mathbf{v}|}{\|\mathbf{w}\|^2} \|\mathbf{w}\| = \frac{|\mathbf{w} \cdot \mathbf{v}|}{\|\mathbf{w}\|}.$$

Toisaalta

$$\mathbf{v} = \text{proj}_{\mathbf{w}}(\mathbf{v}) + \mathbf{w}.$$

Koska $\text{proj}_{\mathbf{w}}(\mathbf{v})$ ja $\mathbf{w}$ ovat kohtisuorassa toisiaan vasten, saadaan **Pythagoraan lauseen** perusteella

$$\|\mathbf{v}\|^2 = \|\text{proj}_{\mathbf{w}}(\mathbf{v}) + \mathbf{w}\|^2 = \|\text{proj}_{\mathbf{w}}(\mathbf{v})\|^2 + \|\mathbf{w}\|^2.$$

Edelleen $\|\text{proj}_{\mathbf{w}}(\mathbf{v})\|^2 + \|\mathbf{w}\|^2 \geq \|\text{proj}_{\mathbf{w}}(\mathbf{v})\|^2$, joten

$$\|\mathbf{v}\|^2 \geq \|\text{proj}_{\mathbf{w}}(\mathbf{v})\|^2.$$

Koska vektorien normit ovat positiivisia, voidaan päätellä, että $\|\mathbf{v}\| \geq \|\text{proj}_{\mathbf{w}}(\mathbf{v})\|$. On siis osoitettu, että vektorin $\mathbf{v}$ projektion pituus ei koskaan ylitä vektorin $\mathbf{v}$ pituutta. Näin ollen

$$\frac{|\mathbf{v} \cdot \mathbf{w}|}{\|\mathbf{w}\|} \leq \|\mathbf{v}\|,$$

mistä seuraa

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

□

Tiivistelmä

- Kun vektori $\mathbf{v}$ projisoidaan vektorin $\mathbf{w}$ virittämälle suoralle, on tuloksena vektorin $\mathbf{w}$ kanssa yhdensuuntainen vektori.
- Projektion ja vektorin $\mathbf{v}$ erotus on kohtisuorassa vektoria $\mathbf{w}$ vastaan.

1.5 Ristitulo

Avaruuden $\mathbb{R}^3$ vektoreille voidaan määritellä pistetulon lisäksi niin kutsuttu ristitulo. Pistetulosta poiketen ristitulon tulos ei ole reaaliluku vaan avaruuden $\mathbb{R}^3$ vektori. Ristitulosta on hyötyä esimerkiksi silloin, kun tarvitaan vektori, joka on kohtisuorassa jotakin tasoa vastaan.

Ristitulo poikkeaa kurssilla tähän mennessä määritellyistä käsitteistä siinä, että sen määritelmää ei voida yleistää kaikkiin avaruuksiin $\mathbb{R}^n$. Ristitulo on vain avaruuden $\mathbb{R}^3$ laskutoimitus.

Määritelmä 1.5.1

Vektorien $\mathbf{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$ ja $\mathbf{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$ *ristitulo* on vektori

$$\mathbf{v} \times \mathbf{w} = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1).$$

Ristitulon $\mathbf{v} \times \mathbf{w}$ laskemiseen voi käyttää muistisääntöjä, joista tässä esitellään kaksi erilaista.

Esimerkki 1.5.2

Lasketaan vektorien $\mathbf{v} = (1, 2, 3)$ ja $\mathbf{w} = (5, 6, 7)$ ristitulo. Järjestetään taulukoksi vektorit $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$ ja $\mathbf{e}_3 = (0, 0, 1)$ sekä $(1, 2, 3)$ ja $(5, 6, 7)$ seuraavalla tavalla:

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix}$$

Tätä taulukkoa tulkitaan niin, että kutakin vektoria $\mathbf{e}_1$, $\mathbf{e}_2$ ja $\mathbf{e}_3$ kerrotaan taulukolla, jossa ovat kaikki ne luvut, jotka eivät ole kyseisen vektorin alla. Näin saatuja vektoreita summataan ja vähennetään keskenään:

$$\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \mathbf{e}_3.$$

Tässä on tärkeää huomata, että keskimmäisessä termissä on miinusmerkki, muissa ei.

Jäljelle jääviä 2×2 -taulukkoita tulkitaan seuraavasti. Taulukon poikki piirretään vinoviivat. Samalla viivalla olevat alkio kerrotaan keskenään. Jos viiva on lävistäjän suuntainen, tulee tulon eteen plusmerkki ja muutoin miinusmerkki. Lopuksi tulot summataan.

$$\begin{vmatrix} \cancel{a_{11}} & a_{12} \\ \cancel{a_{21}} & \cancel{a_{22}} \end{vmatrix} \begin{matrix} - \\ + \end{matrix}$$

Nyt edellisestä laskusta saadaan

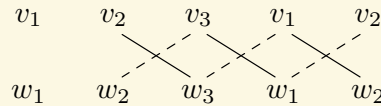
$$(2 \cdot 6 - 3 \cdot 5)\mathbf{e}_1 - (1 \cdot 6 - 3 \cdot 4)\mathbf{e}_2 + (1 \cdot 5 - 2 \cdot 4)\mathbf{e}_3 = -3\mathbf{e}_1 + 6\mathbf{e}_2 - 3\mathbf{e}_3 = (-3, 6, -3).$$

$$\text{Siten } (1, 2, 3) \times (4, 5, 6) = (-3, 6, -3).$$

Tässä esitetty muistisääntö on sama kuin niin kutsutun determinantin laskemisessa käytettävä muistisääntö. Determinanteja käsitellään matriisien yhteydessä.

Esimerkki 1.5.3

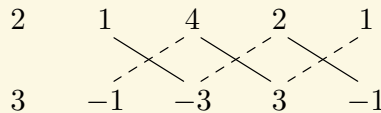
Ristitulon $\mathbf{v} \times \mathbf{w}$ laskemiseen voi käyttää kuvassa 1.5.1 esitettyä muistisääntöä. Yhtenäisellä viivalla yhdistettyjen komponenttien tulosta vähennetään katkoviivalla yhdistettyjen komponenttien tulo.



Kuva 1.5.1. Ristitulon $\mathbf{v} \times \mathbf{w}$ laskeminen.

Merkitään $\mathbf{a} = (2, 1, 4)$ ja $\mathbf{b} = (3, -1, -3)$. Kuvan 1.5.2 perusteella voidaan laskea

$$\mathbf{a} \times \mathbf{b} = (1 \cdot (-3) - 4 \cdot (-1), 4 \cdot 3 - 2 \cdot (-3), 2 \cdot (-1) - 1 \cdot 3) = (1, 18, -5).$$



Kuva 1.5.2. Ristitulon $\mathbf{a} \times \mathbf{b}$ laskeminen.

Eräs ristitulon sovelluksista on, että sen avulla voidaan löytää vektori, joka on kohtisuorassa kahta vektoria vastaan.

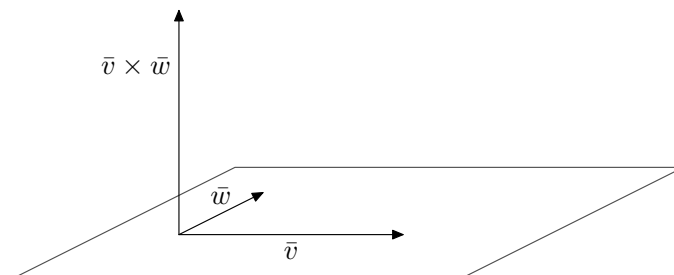
Lause 1.5.4

Oletetaan, että $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Tällöin $(\mathbf{v} \times \mathbf{w}) \perp \mathbf{v}$ ja $(\mathbf{v} \times \mathbf{w}) \perp \mathbf{w}$.

Todistus. Laskemalla huomataan, että

$$\begin{aligned} (\mathbf{v} \times \mathbf{w}) \cdot \mathbf{v} &= (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1) \cdot (v_1, v_2, v_3) \\ &= (v_2w_3 - v_3w_2)v_1 + (v_3w_1 - v_1w_3)v_2 + (v_1w_2 - v_2w_1)v_3 \\ &= v_2w_3v_1 - v_3w_2v_1 + v_3w_1v_2 - v_1w_3v_2 + v_1w_2v_3 - v_2w_1v_3 = 0. \end{aligned}$$

Siten vektorit $(\mathbf{v} \times \mathbf{w})$ ja $\mathbf{v}$ ovat kohtisuorassa toisiaan vastaan. Väitteen toinen osa osoitetaan samalla tavalla. □



Kuva 1.5.3. Ristitulo $\mathbf{v} \times \mathbf{w}$ on kohtisuorassa vektoria $\mathbf{v}$ ja vektoria $\mathbf{w}$ vastaan.

Seuraavassa lauseessa on lueteltu ristituloon liittyviä laskusääntöjä. Erityisesti sääntöihin 1, 5 ja 7 on hyvä kiinnittää huomiota, sillä ne poikkeavat monista tutuista laskusäännöistä. Esimerkiksi säännön 1 mukaan ristitulo ei ole vaihdannainen laskutoimitus.

Lause 1.5.5

Oletetaan, että $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ ja $c \in \mathbb{R}$. Tällöin

1. $\mathbf{v} \times \mathbf{w} = -(\mathbf{w} \times \mathbf{v})$ (antikommutointi)
2. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ (osittelulaki)
3. $(\mathbf{v} + \mathbf{w}) \times \mathbf{u} = \mathbf{v} \times \mathbf{u} + \mathbf{w} \times \mathbf{u}$ (osittelulaki)
4. $c(\mathbf{v} \times \mathbf{w}) = (c\mathbf{v}) \times \mathbf{w} = \mathbf{v} \times (c\mathbf{w})$
5. $\mathbf{v} \times \mathbf{v} = \mathbf{0}$
6. $\mathbf{0} \times \mathbf{v} = \mathbf{0}$ ja $\mathbf{v} \times \mathbf{0} = \mathbf{0}$
7. $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$.

Todistus. Lauseen todistus on suoraviivainen ja käyttää ainoastaan ristitulon määritelmää. Todistus jätetään harjoitustehtäväksi. □

Lause 1.5.6

Jos avaruuden $\mathbb{R}^3$ vektorit ovat yhdensuuntaiset, niiden ristitulo on nolla.

Todistus. Oletetaan, että $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Oletetaan lisäksi, että $\mathbf{v}$ ja $\mathbf{w}$ ovat yhdensuuntaisia. Tällöin on olemassa $c \in \mathbb{R} \setminus \{0\}$, jolle pätee $\mathbf{w} = c\mathbf{v}$. Edellisen lauseen nojalla

$$\mathbf{v} \times \mathbf{w} = \mathbf{v} \times (c\mathbf{v}) = c(\mathbf{v} \times \mathbf{v}) = c \cdot \mathbf{0} = \mathbf{0}.$$

□

Lauseen 1.5.4 perusteella tiedetään, että kahden vektorin ristitulo on kohtisuorassa kumpaakin vektoria vastaan, joten ristitulovektorin suunnalla on vain kaksi mahdollisuutta. Ristitulovektorin pituus puolestaan määräytyy seuraavasta lauseesta.

Lause 1.5.7

Oletetaan, että $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Jos $\mathbf{v} \neq \mathbf{0}$ ja $\mathbf{w} \neq \mathbf{0}$, niin

$$\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\| \|\mathbf{w}\| \sin \alpha,$$

missä α on vektorien $\mathbf{v}$ ja $\mathbf{w}$ välinen kulma.

Todistus. Todistus on melko tekninen, joten sen voi ohittaa ensimmäisillä lukukerroilla.

Normin ja ristitulon määritelmien nojalla

$$\|\mathbf{v} \times \mathbf{w}\|^2 = (v_2w_3 - w_2v_3)^2 + (v_1w_3 - w_1v_3)^2 + (v_1w_2 - w_1v_2)^2.$$

Kun lasketaan lauseke termeittäin auki, saadaan

$$\begin{aligned} \|\mathbf{u} \times \mathbf{v}\|^2 &= v_2^2w_3^2 - 2v_2w_3w_2v_3 + v_2^2v_3^2 + v_1^2w_3^2 - 2v_1w_3w_1v_3 + w_1^2v_3^2 \\ &\quad + v_1^2w_2^2 - 2v_1w_2w_1v_2 + w_1^2v_2^2 \end{aligned}$$

Tämä lauseke on mahdollista kirjoittaa muodossa

$$(v_1^2 + v_2^2 + v_3^2)(w_1^2 + w_2^2 + w_3^2) - (v_1w_1 + v_2w_2 + v_3w_3)^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2.$$

Nyt on osoitettu, että

$$\|\mathbf{v} \times \mathbf{w}\|^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2.$$

Vektorien välisen kulman määritelmän nojalla $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\|\cos(\alpha)$, joten

$$\begin{aligned} \|\mathbf{v} \times \mathbf{w}\|^2 &= \|\mathbf{v}\|^2\|\mathbf{w}\|^2 - \|\mathbf{v}\|^2\|\mathbf{w}\|^2\cos^2(\alpha) = \|\mathbf{v}\|^2\|\mathbf{w}\|^2(1 - \cos^2(\alpha)) \\ &= \|\mathbf{v}\|^2\|\mathbf{w}\|^2\sin^2(\alpha) = (\|\mathbf{v}\|\|\mathbf{w}\|\sin\alpha)^2. \end{aligned}$$

Näin ollen

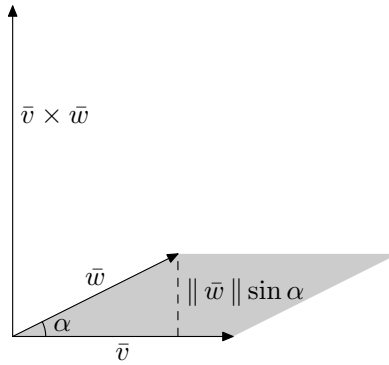
$$\|\mathbf{v} \times \mathbf{w}\|^2 = (\|\mathbf{v}\|\|\mathbf{w}\|\sin\alpha)^2$$

Lisäksi vektorien välisen kulman määritelmän mukaan $0^\circ \leq \alpha \leq 180^\circ$, mistä seuraa, että $\sin\alpha \geq 0$. Lisäksi vektorien normit ovat aina epänegatiivisia. Siten $\|\mathbf{v} \times \mathbf{w}\| \geq 0$ ja $\|\mathbf{v}\|\|\mathbf{w}\|\sin\alpha \geq 0$. Saadusta yhtälöstä voidaan näin ollen päätellä, että

$$\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\|\|\mathbf{w}\|\sin\alpha.$$

Tämä todistaa väitteen. □

Edellisestä lauseesta seuraa, että ristitulovektorin $\mathbf{v} \times \mathbf{w}$ pituus on yhtä suuri kuin vektorien $\mathbf{v}$ ja $\mathbf{w}$ määräämän suunnikkaan ala (kuva 1.5.4). Oletetaan nimittäin, että vektorien $\mathbf{v}$ ja $\mathbf{w}$ välinen kulma on α . Tällöin suunnikkaan korkeus on $\|\mathbf{w}\|\sin\alpha$. Näin suunnikkaan pinta-alaksi saadaan $\|\mathbf{w}\|\sin\alpha \cdot \|\mathbf{v}\| = \|\mathbf{v} \times \mathbf{w}\|$.



Kuva 1.5.4. Ristitulovektorin $\mathbf{v} \times \mathbf{w}$ pituus on yhtä suuri kuin vektorien $\mathbf{v}$ ja $\mathbf{w}$ määrämän suunnikkaan ala.

Ristitulon avulla voidaan määrittää myös suuntaissärmiön tilavuus. Vektoreiden $\mathbf{v}$, $\mathbf{w}$ ja $\mathbf{u}$ määrämän suuntaissärmiön tilavuus on pohjan pinta-alan $\|\mathbf{v} \times \mathbf{w}\|$ ja korkeuden h tulo (kuva 1.5.5). Korkeuden h selvittämiseksi lasketaan vektorin $\mathbf{u}$ projektio ristitulovektorin $\mathbf{v} \times \mathbf{w}$ virittämälle aliavaruudelle:

$$\text{proj}_{\mathbf{v} \times \mathbf{w}}(\mathbf{u}) = \frac{(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}}{(\mathbf{v} \times \mathbf{w}) \cdot (\mathbf{v} \times \mathbf{w})} (\mathbf{v} \times \mathbf{w}).$$

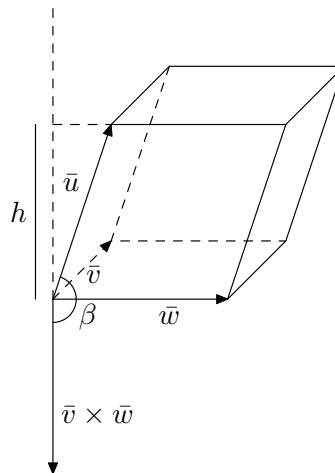
Korkeus h on tämän vektorin pituus eli normi:

$$\begin{aligned} h &= \|\text{proj}_{\mathbf{v} \times \mathbf{w}}(\mathbf{u})\| = \left\| \frac{(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}}{(\mathbf{v} \times \mathbf{w}) \cdot (\mathbf{v} \times \mathbf{w})} (\mathbf{v} \times \mathbf{w}) \right\| = \left| \frac{(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}}{(\mathbf{v} \times \mathbf{w}) \cdot (\mathbf{v} \times \mathbf{w})} \right| \|\mathbf{v} \times \mathbf{w}\| \\ &= \frac{|(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}|}{\|\mathbf{v} \times \mathbf{w}\|^2} \|\mathbf{v} \times \mathbf{w}\| = \frac{|(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}|}{\|\mathbf{v} \times \mathbf{w}\|} \end{aligned}$$

Tilavuudeksi saadaan pohjan pinta-ala kertaa korkeus:

$$\|\mathbf{v} \times \mathbf{w}\| \cdot \frac{|(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}|}{\|\mathbf{v} \times \mathbf{w}\|} = |(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}|$$

Suuntaissärmiön tilavuus on siis niin kutsutun *skalaarikolmitulon* $(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u}$ itseisarvo. [Lauseesta 1.5.5](#) seuraa, että vektorien $\mathbf{v}$, $\mathbf{w}$ ja $\mathbf{u}$ järjestyksellä tässä kaavassa ei ole väliä.



Kuva 1.5.5. Vektoreiden $\mathbf{v}$, $\mathbf{w}$ ja $\mathbf{u}$ määrämän suuntaissärmiön tilavuus.

 Tiivistelmä

1. Avaruudessa $\mathbb{R}^3$ on määritelty vektoreiden ristitulo. Se tuottaa tulokseksi avaruuden $\mathbb{R}^3$ vektorin.
2. Ristitulon avulla on mahdollista löytää kahta vektoria vastaan kohtisuorassa oleva vektori.

2 SOVELLUS: SUORAT JA TASOT

2.1 Suorat

Tässä luvussa sovelletaan vektoreita suorien ja tasojen ilmaisemiseen. Kaikille luvun väitteille ei anneta tarkkoja todistuksia.

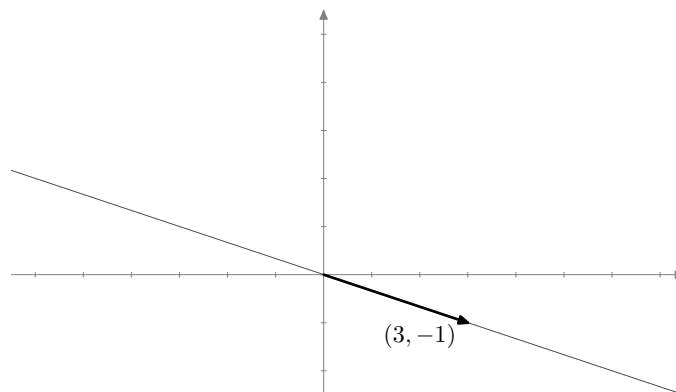
Pohdi 2.1.1

Marty McFly lentää leijulaudallaan, jota voi ohjailla vektorin $(-3, 1)$ suuntaisesti (eteen ja taaksepäin).

1. Marty lähtee kotoaan, joka on koordinaatiston pisteessä $(0, 0)$. Minne kaikkialle Marty voi laudallaan päästä? Hahmottele tilanteesta kuva.
2. Marty lähtee kaverinsa kotoa, jonka koordinaatit ovat $(1, 2)$. Minne kaikkialle Marty voi tällä kertaa laudallaan päästä? Hahmottele tilanteesta kuva.

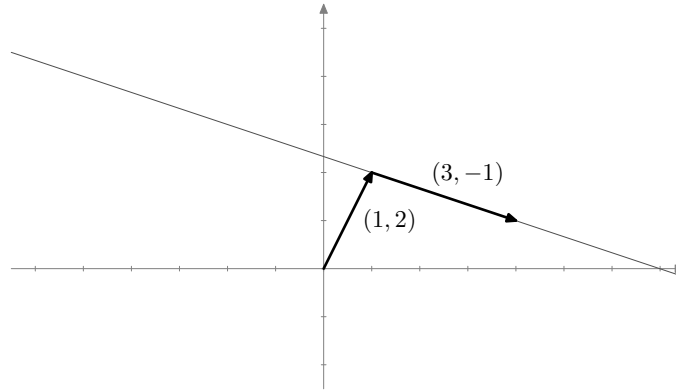
Koulussa suoria kirjoitettiin useimmiten muotoa $y = kx + b$ olevien yhtälöiden avulla. Myös vektoreita voi käyttää suorien ilmaisemiseen. Tällöin monet suoria koskevat ongelmat kuten pisteen etäisyys suorasta muuttuvat helpommiksi käsitellä.

Vektorin $(-3, 1)$ kanssa yhdensuuntaiset vektorit ovat muotoa $t(-3, 1)$, missä t on reaaliluku. Jos vektoreita ajattelee koordinaatiston pisteinä, nämä pisteet sijaitsevat samalla suoralla. Suora kulkee origon kautta kulkevan suoran, joka on yhdensuuntainen vektorin $(-3, 1)$ kanssa.



Kuva 2.1.1. Muotoa $t(-3, 1)$ olevat vektorit muodostavat origon kautta kulkevan suoran.

Jos halutaan muodostaa vektorimerkintöjen avulla suora, joka ei kulje origon kautta, on suora ensin ”siirrettävä” haluttuun paikkaan. Esimerkiksi vektorit, jotka ovat muotoa $(1, 2) + t(3, -1)$, muodostavat suoran, joka on vektorin $(-3, 1)$ kanssa yhdensuuntainen ja kulkee pisteen $(1, 2)$ kautta. Tämän suoran pisteet saadaan lisäämällä vektoriin $(1, 2)$ vektoreita, jotka ovat yhdensuuntaisia vektorin $(3, -1)$ kanssa. Tilannetta on havainnollistettu kuvassa 2.1.2.



Kuva 2.1.2. Suora, jonka alkiot ovat muotoa $(1, 2) + t(3, -1)$, missä $t \in \mathbb{R}$.

Kuvan 2.1.2 suora on joukko $\{(1, 2) + t(3, -1) \mid t \in \mathbb{R}\}$. Suoran voi myös ilmaista yhtälön $\mathbf{x} = (1, 2) + t(3, -1)$ avulla. Tällöin ajatellaan, että suoran alkioita ovat täsmälleen ne vektorit $\mathbf{x}$, jotka ovat muotoa $(1, 2) + t(3, -1)$ jollakin $t \in \mathbb{R}$.

Määritelmä 2.1.2

Vektoriavaruuden $\mathbb{R}^n$ suora on joukko

$$\{\mathbf{p} + t\mathbf{v} \mid t \in \mathbb{R}\},$$

missä $\mathbf{p} \in \mathbb{R}^n$ ja $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Vektoria $\mathbf{p}$ kutsutaan suoran *paikkavektoriksi* ja vektoria $\mathbf{v}$ suoran *suuntavektoriksi*. Suoran yhtälö on

$$\mathbf{x} = \mathbf{p} + t\mathbf{v}.$$

Huomaa, että yllä ei ole annettu mitä tahansa suoran kuvailua, vaan suoran *määritelmä*. Emme siis ole lähteneet esimerkiksi jostakin suoran geometrisestä muotoilusta ja ilmaisseet saman asian vektoreilla, vaan määritelleet suoran käsitteen tämän kurssin tarpeita varten.

Suorien alkiot ovat vektoreita. Vektoreita voi kuitenkin ajatella myös pisteinä, ja suorien tapauksessa tämä ajattelutapa on yleensä hyödyllinen. Sanotaan, että piste (a, b) on suoralla S tai että suora S kulkee pisteen (a, b) kautta, jos $(a, b) \in S$. Vastaavia ilmauksia käytetään vektoriavaruudessa $\mathbb{R}^n$.

Esimerkki 2.1.3

Esimerkiksi joukko

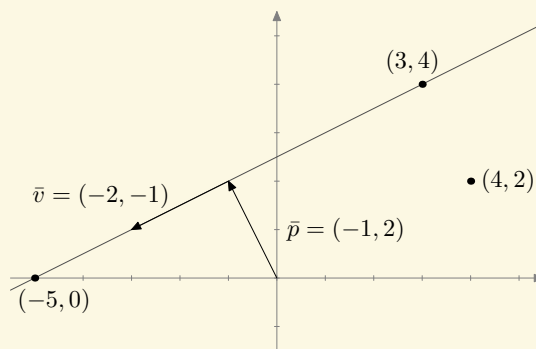
$$S = \{(-1, 2) + t(-2, -1) \mid t \in \mathbb{R}\}$$

on suora vektoriavaruudessa $\mathbb{R}^2$. Tämän suoran yhtälö on

$$\mathbf{x} = (-1, 2) + t(-2, -1),$$

eli suoralla ovat kaikki pisteet, jotka toteuttavat kyseisen yhtälön jollakin reaalityyppillä t .

Suora S on piirretty kuvaan 2.1.3. Määritelmän mukaan mikä tahansa suoran S piste voidaan kirjoittaa summana vektorista $\mathbf{p} = (-1, 2)$ ja jostakin vektorin $\mathbf{v} = (-2, -1)$ skalaarimonikerrasta. Esimerkiksi $(-5, 0) = \mathbf{p} + 2\mathbf{v}$ ja $(3, 4) = \mathbf{p} - 2\mathbf{v}$, joten $(-5, 0) \in S$ ja $(3, 4) \in S$.



Kuva 2.1.3. Suoran S paikkavektori on $\mathbf{p} = (-1, 2)$ ja suuntavektori $\mathbf{v} = (-2, -1)$.

Toisaalta piste $(4, 2)$ ei ole suoralla S . Jos nimittäin $(4, 2) = (-1, 2) + t(-2, -1)$ jollakin $t \in \mathbb{R}$, niin $4 = -1 - 2t$ ja $2 = 2 - t$. Ensimmäisen yhtälön perusteella $t = -5/2$ ja toisen perusteella $t = 0$. Tämä on mahdotonta, joten ei ole olemassa sellaista lukua $t \in \mathbb{R}$, jolle pätee $(4, 2) = (-1, 2) + t(-2, -1)$. Siispä $(4, 2) \notin S$.

Huomaa, että myöskään suuntavektori $(-2, -1)$ ei ole suoran S alkio. Toisin sanoen piste $(-2, -1)$ ei ole suoralla S .

Suoran yhtälön voi myös kirjoittaa *parametrimuodossa*, jossa listataan yhtälöt, jotka pätevät suoran alkioiden komponenteille. Esimerkiksi edellisen esimerkin suoran $\mathbf{x} = (-1, 2) + t(-2, -1)$ parametrimuotoinen yhtälö on

$$\begin{cases} x_1 = -1 - 2t \\ x_2 = 2 - t. \end{cases}$$

Suorien suuntavektoreita käsiteltäessä on hyödyllistä ottaa käyttöön uudenlainen merkintä vektoreille. Oletetaan, että $A = (a_1, a_2, \dots, a_n)$ ja $B = (b_1, b_2, \dots, b_n)$ ovat avaruuden pisteitä. Vektori $\overline{AB}$ on vektori, jota vastaavan suuntajanan alkupiste on A ja päätepiste on

B. Tällöin

$$\overrightarrow{AB} = \begin{bmatrix} b_1 - a_1 \\ b_2 - a_2 \\ \vdots \\ b_n - a_n \end{bmatrix},$$

eli vektorin $\overrightarrow{AB}$ komponentit ovat pisteiden B ja A komponenttien erotukset. Usein vektorille $\overrightarrow{AB}$ näkee myös käytettävän merkintää $\overrightarrow{AB}$.

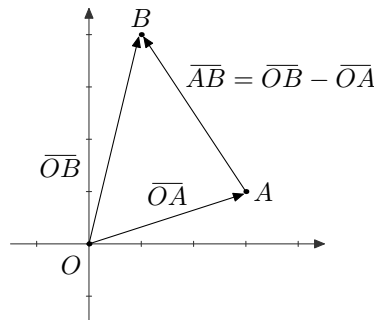
Esimerkiksi tason pisteitä $A = (3, 1)$ ja $B = (1, 4)$ yhdistää vektori

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

Tämä vektori on piirretty kuvaan 2.1.4.

Origoa $(0, 0, \dots, 0)$ on tapana merkitä kirjaimella O . Siten origosta lähtevälle vektorille, jonka kärki on pisteessä A , saadaan merkintä $\overrightarrow{OA}$. Vektoria $\overrightarrow{OA}$ kutsutaan pisteen A paikkavektoriksi.

Lisäksi pätee $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$. Tätä on havainnollistettu kuvassa 2.1.4.



Kuva 2.1.4. Vektorit $\overrightarrow{OA}$, $\overrightarrow{OB}$ ja $\overrightarrow{AB}$.

Ryhdytään seuraavaksi määrittämään suoraa, joka kulkee annettujen pisteiden kautta.

Esimerkki 2.1.4

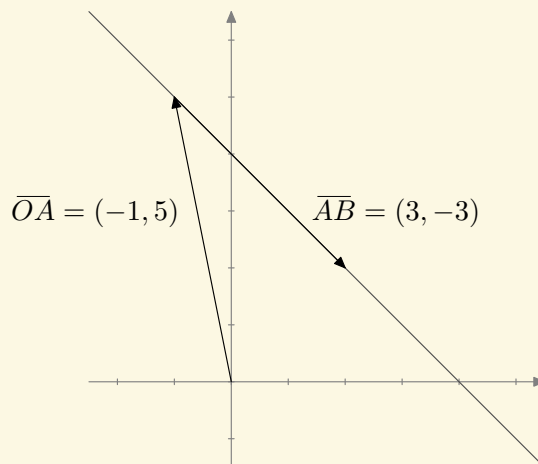
Määritetään pisteiden $A = (-1, 5)$ ja $B = (2, 2)$ kautta kulkeva suora. Tätä varten suoralle täytyy löytää paikkavektori ja suuntavektori. Paikkavektoriksi käy minkä tahansa suoran pisteen paikkavektori, esimerkiksi vektori $\overrightarrow{OA} = (-1, 5)$. Suuntavektoriksi käy mikä tahansa suoran kanssa yhdensuuntainen vektori, esimerkiksi vektori

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (2, 2) - (-1, 5) = (3, -3).$$

Näin saadaan suora $S = \{(-1, 5) + t(3, -3) \mid t \in \mathbb{R}\}$. Sitä vastaava yhtälö on

$$\mathbf{x} = (-1, 5) + t(3, -3).$$

Suora on esitetty kuvassa 2.1.5.



Kuva 2.1.5. Suora $S = \{(-1, 5) + t(3, -3) \mid t \in \mathbb{R}\}$.

Tarkistetaan vielä, että annetut pisteet A ja B todellakin ovat suoralla S . Huomataan, että $(-1, 5) = (-1, 5) + 0(3, -3)$ ja $(2, 2) = (-1, 5) + 1(3, -3)$. Siten suora S kulkee pisteiden A ja B kautta.

Edellistä esimerkkiä mukaillen on aina mahdollista määrittää kahden pisteen kautta kulkeva suora. Yleisemmin voidaan osoittaa, että jos suora halutaan kirjoittaa muodossa

$$\{\mathbf{p} + t\mathbf{v} \mid t \in \mathbb{R}\},$$

paikkavektoriksi $\mathbf{p}$ voidaan valita suoran minkä tahansa pisteen paikkavektori, ja suuntavektoriksi $\mathbf{v}$ voidaan valita mikä tahansa suoran suuntainen vektori. Seuraava esimerkki havainnollistaa asiaa.

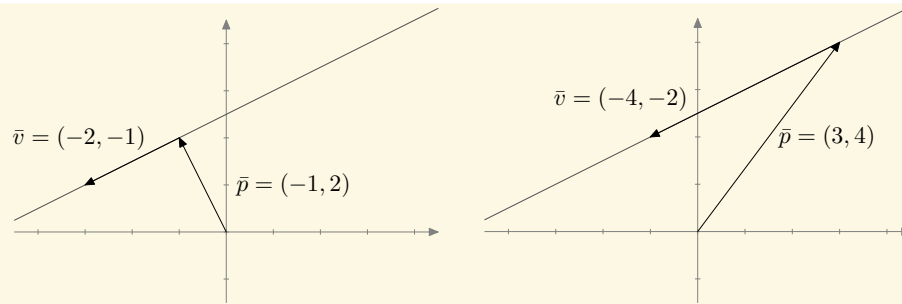
Esimerkki 2.1.5

Tutkitaan [esimerkin 2.1.3](#) suoraa $S = \{(-1, 2) + t(-2, -1) \mid t \in \mathbb{R}\}$. Sille voidaan valita paikkavektoriksi mikä tahansa suoran alkio, vaikkapa vektori $(3, 4) = (-1, 2) - 2(-2, -1)$. Suuntavektoriksi puolestaan kelpaa mikä tahansa vektorin $(-2, -1)$ kanssa yhdensuuntainen vektori, vaikkapa vektori $(-4, -2) = 2(-2, -1)$.

Suora S voidaan siis kirjoittaa muodossa

$$\{(3, 4) + s(-4, -2) \mid s \in \mathbb{R}\}.$$

Vaikka tämä joukko onkin äkkiseltään katsottuna erilainen kuin suoran S alkuperäinen määritelmä, on joukoissa täsmälleen samat alkiot. Asiaa on havainnollistettu kuvassa 2.1.6.



Kuva 2.1.6. Suoran paikkavektori ja suuntavektori eivät ole yksikäsitteisiä, vaan ne voidaan valita monella eri tavalla.

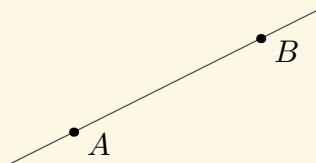
Kun suora ilmaistaan vektorimerkinnöillä, helpottuvat monet laskut. Seuraavassa esimerkissä näytetään, kuinka pisteen etäisyyden suorasta voi määrittää vektoriprojektion avulla.

Esimerkki 2.1.6

Pisteen X etäisyys suorasta $S = \{\mathbf{p} + t\mathbf{v} \mid t \in \mathbb{R}\}$ on kaikkein lyhin välimatka, joka voi olla pisteen X ja suoran S jonkin pisteen välillä. Voidaan osoittaa, että tämä etäisyys on sama kuin sellaisen janan pituus, jonka toinen päätepiste on X ja toinen suoralla S , ja joka muodostaa suoran kulman suoran S kanssa. Etäisyyden määrittämiseen voidaan siten käyttää projektiota.

Määritetään pisteen $X = (4, -1, 9)$ etäisyys suorasta S , joka kulkee pisteiden $A = (2, -3, 5)$ ja $B = (4, 1, 7)$ kautta (ks. kuva 2.1.7).

$X \bullet$



Kuva 2.1.7. Pisteiden A ja B kautta kulkeva suora S .

Määritetään ensin vektori jostakin suoran pisteestä tutkittavaan pisteeseen. Esimerkiksi vektori $\overrightarrow{AX} = \overrightarrow{OX} - \overrightarrow{OA} = (2, 2, 4)$ käy tähän tarkoitukseen. Lisäksi tarvitaan jokin suoran suuntavektori, kuten vaikkapa vektori $\overrightarrow{AB} = (2, 4, 2)$.

Vektorin $\overrightarrow{AX}$ projektio suoralle S on

$$\text{proj}_{\overrightarrow{AB}}(\overrightarrow{AX}) = \frac{\overrightarrow{AX} \cdot \overrightarrow{AB}}{\overrightarrow{AB} \cdot \overrightarrow{AB}} \overrightarrow{AB} = \frac{20}{24} (2, 4, 2) = \frac{5}{6} (2, 4, 2).$$

Erotus $\overrightarrow{AX} - \text{proj}_{\overrightarrow{AB}}(\overrightarrow{AX})$ on projektion määritelmän mukaisesti kohtisuorassa suoraa

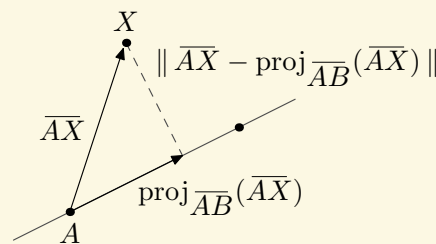
S vastaan. Lasketaan kyseinen erotus:

$$\begin{aligned}\overline{AX} - \text{proj}_{\overline{AB}}(\overline{AX}) &= (2, 2, 4) - \frac{5}{6}(2, 4, 2) = \frac{6}{6}(2, 2, 4) - \frac{5}{6}(2, 4, 2) \\ &= \frac{1}{6}(12 - 10, 12 - 20, 24 - 10) = \frac{1}{6}(2, -8, 14) = \frac{1}{3}(1, -4, 7)\end{aligned}$$

Koska $\overline{AX} - \text{proj}_{\overline{AB}}(\overline{AX})$ on kohtisuorassa suoraa S vastaan, antaa erotusvektorin pituus pisteen X etäisyyden suorasta:

$$\|\overline{AX} - \text{proj}_{\overline{AB}}(\overline{AX})\| = \frac{1}{3}\|(1, -4, 7)\| = \frac{1}{3}\sqrt{1 + 16 + 49} = \frac{1}{3}\sqrt{66}.$$

Siten pisteen X etäisyys suorasta S on $\frac{1}{3}\sqrt{66}$. Asiaa on havainnollistettu kuvassa 2.1.8.



Kuva 2.1.8. Pisteen X etäisyys suorasta S .

Tasossa suoran yhtälölle saadaan vielä yksi uusi ilmaisutapa pistetulon avulla. **Esimerkissä 1.3.15** nähtiin, kuinka vektoria $\mathbf{n} = (3, -2)$ vastaan kohtisuorat vektorit muodostavat origon kautta kulkevan suoran. Tämän suoran alkioit ovat täsmälleen ne vektorit $\mathbf{x}$, jotka toteuttavat yhtälön

$$\mathbf{n} \cdot \mathbf{x} = 0.$$

Vektoria $\mathbf{n}$ kutsutaan suoran normaaliksivektoriksi ja yhtälöä suoran normaalimuotoiseksi yhtälöksi.

Määritelmä 2.1.7

Avaruuden $\mathbb{R}^2$ suoran *normaalivektori* on mitä tahansa nollavektorista poikkeava vektori $\mathbf{n}$, joka on kohtisuorassa suoran suuntavektoria vastaan.

Määritelmänsä mukaan normaalivektori on kohtisuorassa suoran suuntavektoria vastaan. Jos suora kulkee origon kautta, suoran normaalivektori on kohtisuorassa *kaikkia* tason vektoreita vastaan. Suora koostuu täsmälleen niistä vektoreista, jotka ovat kohtisuorassa normaalivektoria vastaan.

Jos suora ei kulje origon kautta, tulee normaalimuotoisesta yhtälöstä hiukan monimutkaisempi kuin edellä. Olkoon suoran paikkavektori $\mathbf{p}$ ja normaali $\mathbf{n}$. Vektori $\mathbf{x}$ on suoran alkio täsmälleen siinä tapauksessa, että

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0.$$

Määritelmä 2.1.8

Avaruuden $\mathbb{R}^2$ suoran yhtälön *normaalimuoto* on

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$$

missä $\mathbf{p}$ on suoran paikkavektori ja $\mathbf{n}$ on suoran normaalivektori.

Kun merkitään $\mathbf{x} = (x, y)$, $\mathbf{n} = (a, b)$ sekä $\mathbf{n} \cdot \mathbf{p} = c$ ja sijoitetaan nämä normaalimuotoon, saadaan suoran yhtälö muotoon

$$ax + by = c.$$

Tätä kutsutaan suoran yhtälön *yleiseksi muodoksi*.

Esimerkki 2.1.9

Kirjoitetaan suoran

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

yhtälö normaalimuodossa.

Aloitetaan etsimällä suoralle jokin normaalivektori. Merkitään normaalivektoria $\mathbf{n} = (a, b)$. Määritelmän perusteella normaalivektorin täytyy olla kohtisuorassa suunta-vektoria vastaan eli

$$3a + 2b = 0.$$

Nyt täytyy keksiä jotkin luvut a ja b , joilla yhtälö toteutuu. Eräs ratkaisu on $a = 2$ ja $b = -3$, jolloin normaaliksi saadaan $\mathbf{n} = (2, -3)$. Suoran yhtälön normaalimuoto on siis nyt

$$\begin{bmatrix} 2 \\ -3 \end{bmatrix} \cdot \left(\mathbf{x} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = 0.$$

 Tiivistelmä

1. Suoran pisteet saadaan lisäämällä paikkavektoriin suuntavektorin skalaarimonikertoja.
2. Tason pisteet saadaan lisäämällä paikkavektoriin suuntavektorien lineaarikombinaatioita.
3. Tason $\mathbb{R}^2$ suoran voi myös kirjoittaa normaalivektorin avulla. Se on vektori, joka on kohtisuorassa suoraa.

2.2 Tasot

Pohdi 2.2.1

Jasmin lentää taikamatollaan, jota voi ohjailla vektorien $(1, 0, 0)$ ja $(0, 1, 0)$ suuntaisesti (eteen ja taaksepäin).

1. Jasmin lähtee kotipalatsistaan, joka on koordinaatiston pisteessä $(0, 0, 0)$. Minne kaikkialle Jasmin voi matollaan päästä? Hahmottele tilanteesta kuva.
2. Jasmin lähtee naapurikaupungista, jonka koordinaatit ovat $(-2, 1, 1)$. Minne kaikkialle Jasmin voi tällä kertaa matollaan päästä? Hahmottele tilanteesta kuva.

Vektoriavaruuksissa voidaan määritellä tasot samaan tapaan kuin suorat. Nyt tarvitaan paikkavektorin lisäksi kaksi suuntavektoria.

Määritelmä 2.2.2

Vektoriavaruuden $\mathbb{R}^n$ *taso* on joukko

$$\{\mathbf{p} + s\mathbf{v} + t\mathbf{w} \mid s, t \in \mathbb{R}\},$$

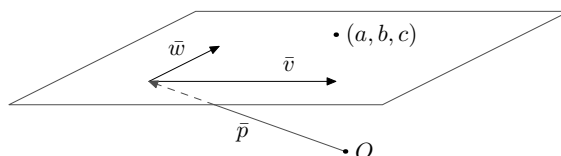
missä $\mathbf{p} \in \mathbb{R}^n$, $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ ja vektorit $\mathbf{v}$ ja $\mathbf{w}$ eivät ole yhdensuuntaiset. Vektoria $\mathbf{p}$ kutsutaan tason *paikkavektoriksi* ja vektoreita $\mathbf{v}$ ja $\mathbf{w}$ tason *suuntavektoreiksi*.

Tason yhtälö on

$$\mathbf{x} = \mathbf{p} + s\mathbf{v} + t\mathbf{w}.$$

Tason suuntavektorit eivät saa olla yhdensuuntaisia. Jos ne nimittäin olisivat yhdensuuntaisia, olisi edellisen määritelmän joukko pelkkä suora.

Kuten suorien tapauksessa, tason alkioita voidaan ajatella pisteinä. Olkoon T vektoriavaruuden $\mathbb{R}^3$ taso. Sanotaan, että piste (a, b, c) on tasossa T tai että taso T kulkee pisteen (a, b, c) kautta, jos $(a, b, c) \in T$. Tasoa ja siinä olevaa pistettä on havainnollistettu kuvassa 2.2.1. Vastaavia ilmauksia käytetään vektoriavaruudessa $\mathbb{R}^n$.



Kuva 2.2.1. Piste (a, b, c) on tasossa T .

Esimerkki 2.2.3

Määritetään pisteiden $A = (0, 1, 0)$, $B = (-1, 3, 2)$ ja $C = (-2, 0, 1)$ kautta kulkeva taso T . Valitaan ensin tason paikkavektori. Esimerkiksi tason pisteen A paikkavektori $\overrightarrow{OA} = (0, 1, 0)$ käy tähän tarkoitukseen. Lisäksi tarvitaan tason suuntaiset suuntavek-

torit:

$$\overline{AB} = \overline{OB} - \overline{OA} = (-1, 2, 2) \quad \text{ja} \quad \overline{AC} = \overline{OC} - \overline{OA} = (-2, -1, 1).$$

Nyt on tarkistettava, että vektorit $\overline{AB}$ ja $\overline{AC}$ eivät ole yhdensuuntaiset, sillä muutoin kyseessä ei ole taso. Tämä jätetään harjoitustehtäväksi lukijalle. Näin saadaan taso

$$T = \{\overline{OA} + s\overline{AB} + t\overline{AC} \mid s, t \in \mathbb{R}\} = \{(0, 1, 0) + s(-1, 2, 2) + t(-2, -1, 1) \mid s, t \in \mathbb{R}\}.$$

Tason yhtälö on

$$\mathbf{x} = (0, 1, 0) + s(-1, 2, 2) + t(-2, -1, 1).$$

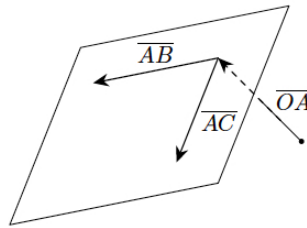
Tarkistetaan vielä, että pisteet A , B ja C tosiaankin ovat tasossa T . Huomataan, että

$$A = (0, 1, 0) + 0(-1, 2, 2) + 0(-2, -1, 1),$$

$$B = (0, 1, 0) + 1(-1, 2, 2) + 0(-2, -1, 1) \quad \text{ja}$$

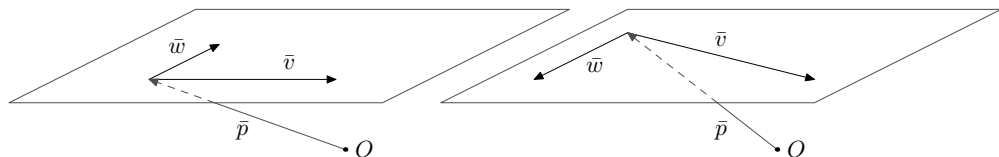
$$C = (0, 1, 0) + 0(-1, 2, 2) + 1(-2, -1, 1).$$

Siten T on taso, joka kulkee pisteiden A , B ja C kautta.



Kuva 2.2.2. Taso T kulkee pisteiden A , B ja C kautta.

Edellisessä käytetty menetelmä toimii yleisemminkin. Jos taso halutaan kirjoittaa muodossa $\{\mathbf{p} + s\mathbf{w} + t\mathbf{v} \mid s, t \in \mathbb{R}\}$, paikkavektoriksi $\mathbf{p}$ voidaan valita tason minkä tahansa pisteen paikkavektori. Suuntavektoreiksi $\mathbf{w}$ ja $\mathbf{v}$ voidaan valita mitkä tahansa tason suuntaiset vektorit, kunhan $\mathbf{w}$ ja $\mathbf{v}$ eivät ole yhdensuuntaiset. Taso on siis mahdollista kirjoittaa usealla eri tavalla joukkona $\{\mathbf{p} + s\mathbf{w} + t\mathbf{v} \mid s, t \in \mathbb{R}\}$. Tilannetta on havainnollistettu kuvassa 2.2.3.

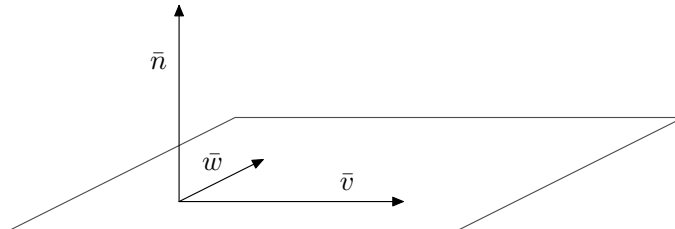


Kuva 2.2.3. Taso voidaan kirjoittaa eri tavoin joukkona $\{\mathbf{p} + s\mathbf{w} + t\mathbf{v} \mid s, t \in \mathbb{R}\}$.

Samaan tapaan kuin avaruuden $\mathbb{R}^2$ suorien tapauksessa, avaruuden $\mathbb{R}^3$ tasot voidaan kirjoittaa normaalivektorin avulla.

Määritelmä 2.2.4

Olkoon T avaruuden $\mathbb{R}^3$ taso. Vektoria, joka on kohtisuorassa tason T suuntavektoreita vastaan, kutsutaan tason *normaalivektoriksi*.



Kuva 2.2.4. Tason normaali $\mathbf{n}$.

Esimerkki 2.2.5

Vektori $(3, 20, 12)$ on tason

$$T = \{s(4, 0, -1) + t(0, 3, -5) \mid s, t \in \mathbb{R}\}$$

normaalivektori, sillä $(3, 20, 12) \cdot (4, 0, -1) = 0$ ja $(3, 20, 12) \cdot (0, 3, -5) = 0$. Vektori $(3, 20, 12)$ on itse asiassa kohtisuorassa kaikkia tason T vektoreita vastaan, sillä

$$\begin{aligned} \begin{bmatrix} 3 \\ 20 \\ 12 \end{bmatrix} \cdot \left(s \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 3 \\ -5 \end{bmatrix} \right) &= \begin{bmatrix} 3 \\ 20 \\ 12 \end{bmatrix} \cdot \left(s \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix} \right) + \begin{bmatrix} 3 \\ 20 \\ 12 \end{bmatrix} \cdot \left(t \begin{bmatrix} 0 \\ 3 \\ -5 \end{bmatrix} \right) \\ &= s \left(\begin{bmatrix} 3 \\ 20 \\ 12 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix} \right) + t \left(\begin{bmatrix} 3 \\ 20 \\ 12 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ -5 \end{bmatrix} \right) \\ &= s \cdot 0 + t \cdot 0 = 0. \end{aligned}$$

Vektori $(3, 20, 12)$ on myös tason

$$U = \{(0, -14, -7) + s(4, 0, -1) + t(0, 3, -5) \mid s, t \in \mathbb{R}\}$$

normaali, sillä se on kohtisuorassa tämänkin tason suuntavektoreita vastaan:

$$(3, 20, 12) \cdot (4, 0, -1) = 0 \quad \text{ja} \quad (3, 20, 12) \cdot (0, 3, -5) = 0.$$

Vektori $(3, 20, 12)$ ei kuitenkaan ole kohtisuorassa mitään tason U vektoria vastaan. (Tämän tarkistaminen jätetään lukijalle.) Erona edelliseen tapaukseen on se, että taso U ei kulje origon kautta.

Määritelmänsä mukaan normaalivektori on kohtisuorassa tason suuntavektoreita vastaan. Jos taso kulkee origon kautta, tason normaalivektori on kohtisuorassa *kaikkia* tason vektoreita vastaan. Taso koostuu täsmälleen niistä vektoreista, jotka ovat kohtisuorassa normaalia vastaan. Toisin sanoen, jos $\mathbf{n}$ on origon kautta kulkevan tason T normaali, piste $\mathbf{x}$ on tasossa T , jos ja vain jos

$$\mathbf{n} \cdot \mathbf{x} = 0.$$

Jos taso ei kulje origon kautta, saa tason yhtälö hiukan toisen muodon. Oletetaan, että T on avaruuden $\mathbb{R}^3$ taso, jonka paikkavektori on $\mathbf{p}$ ja jolla on normaalivektori $\mathbf{n}$. Tällöin $\mathbf{x}$ on tasossa T , jos ja vain jos

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0.$$

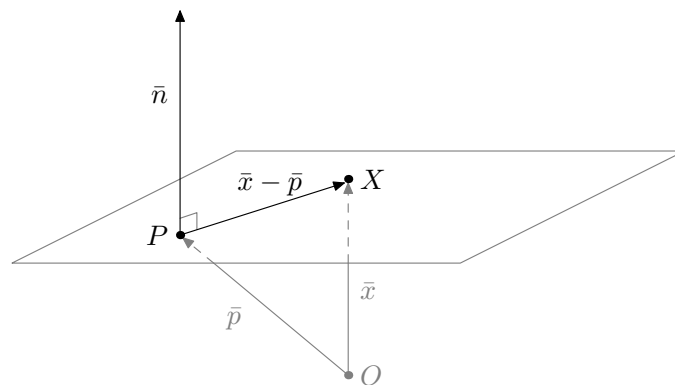
Tilannetta on havainnollistettu kuvassa 2.2.5.

Määritelmä 2.2.6

Avaruuden $\mathbb{R}^3$ tason yhtälön *normaalimuoto* on

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0,$$

missä $\mathbf{p}$ on tason paikkavektori ja $\mathbf{n}$ normaalivektori.



Kuva 2.2.5. Tason T normaalimuotoisen yhtälön havainnollistus.

Esimerkki 2.2.7

Oletetaan, että avaruuden $\mathbb{R}^3$ taso T kulkee pisteen $P = (6, 0, 1)$ kautta ja sillä on normaalivektori $\mathbf{n} = (1, 2, 3)$. Tason T normaalimuotoinen yhtälö on tällöin

$$(1, 2, 3) \cdot (\mathbf{x} - (6, 0, 1)) = 0.$$

Tasossa T ovat siis täsmälleen ne pisteet $\mathbf{x}$, jotka toteuttavat edellä esitetyn yhtälön. Toisin sanoen

$$T = \{\mathbf{x} \in \mathbb{R}^3 \mid (1, 2, 3) \cdot (\mathbf{x} - (6, 0, 1)) = 0\}.$$

Kirjoitetaan yhtälö vielä hiukan toisenlaisessa muodossa. Merkitään $\mathbf{x} = (x, y, z)$,

missä $x, y, z \in \mathbb{R}$. Nyt

$$\begin{aligned}(1, 2, 3) \cdot (\mathbf{x} - (6, 0, 1)) &= (1, 2, 3) \cdot (x - 6, y - 0, z - 1) \\ &= x - 6 + 2y + 3z - 3 \\ &= x + 2y + 3z - 9.\end{aligned}$$

Tason normaalimuotoinen yhtälö saa siis muodon $x + 2y + 3z - 9 = 0$, ja voidaan kirjoittaa $T = \{\mathbf{x} \in \mathbb{R}^3 \mid x + 2y + 3z - 9 = 0\}$.

Edellisestä esimerkkejä mukaillen voidaan osoittaa, että tason yhtälön voi aina muuttaa muotoon

$$ax + by + cz = d,$$

missä $a, b, c, d \in \mathbb{R}$. Tätä kutsutaan tason yhtälön *yleiseksi muodoksi*.

Esimerkki 2.2.8

Määritetään normaalimuotoinen yhtälö tasolle T , joka kulkee pisteiden $A = (0, 1, 0)$, $B = (-1, 3, 2)$ ja $C = (-2, 0, 1)$ kautta. Tätä varten tarvitaan tason T normaalivektori eli vektori, joka on kohtisuorassa tason suuntavektoreita vastaan. Valitaan suuntavektoreiksi $\overline{AB} = (-1, 2, 2)$ ja $\overline{AC} = (-2, -1, 1)$, jolloin normaaliksi käy lauseen 1.5.4 perustella vektorien ristitulo $\overline{AB} \times \overline{AC} = (4, -3, 5)$.

Lisäksi tarvitaan jokin tason paikkavektori, kuten esimerkiksi $\overline{OA} = (0, 1, 0)$. Kun merkitään vielä $\mathbf{x} = \overline{OX} = (x, y, z)$, tason normaalimuotoiseksi yhtälöksi saadaan [esimerkin 2.2.7](#) mukaisesti

$$(\underbrace{\overline{AB} \times \overline{AC}}_{\text{normaalivektori}}) \cdot (\overline{OX} - \overline{OA}) = 0.$$

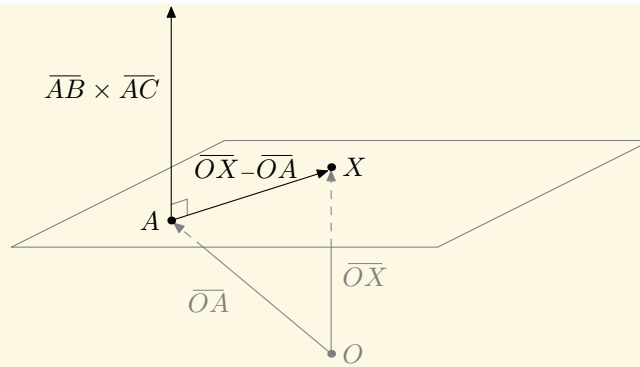
Tilannetta on havainnollistettu kuvassa 2.2.6.

Kun normaalimuotoiseen yhtälöön sijoitetaan luvut, se saadaan muotoon

$$(4, -3, 5) \cdot (x, y - 1, z) = 0.$$

Laskemalla pistetulo saadaan yhtälö lopulliseen muotoon $4x - 3y + 5z + 3 = 0$. Tämä on tason normaalimuotoinen yhtälö. Näin ollen

$$T = \{(x, y, z) \in \mathbb{R}^3 \mid 4x - 3y + 5z + 3 = 0\}.$$



Kuva 2.2.6. Tason T normaalimuotoisen yhtälön määrittäminen, kun tiedetään tason pisteet A , B ja C .

Esimerkki 2.2.9

Pisteen X etäisyys tasosta T on kaikkein lyhin välimatka, joka voi olla pisteen X ja tason T jonkin pisteen välillä. Voidaan osoittaa, että tämä etäisyys on sama kuin sellaisen janan pituus, jonka toinen päätepiste on X ja toinen tasossa T , ja joka muodostaa suoran kulman tason T kanssa. Etäisyyden määrittämiseen voidaan käyttää normaalivektoria ja projektiota.

Tutkitaan edellisen esimerkin tasoa T , joka kulkee pisteiden $A = (0, 1, 0)$, $B = (-1, 3, 2)$ ja $C = (-2, 0, 1)$ kautta ja jolla on normaalivektori $\mathbf{n} = (4, -3, 5)$. Selvitetään, mikä on pisteen $D = (5, 3, -1)$ etäisyys tasosta T .

Valitaan jokin tason T piste kuten piste $A = (0, 1, 0)$. Tutkitaan vektoria $\overline{AD}$. Vektorin $\overline{AD}$ projektio normaalivektorille $\mathbf{n}$ kulkee tasolta pisteeseen D ja on kohtisuorassa tasoa vastaan. Siten projektiovektorin pituus on pisteen D etäisyys tasosta T . Toisin sanoen etäisyys on $\|\text{proj}_{\mathbf{n}}(\overline{AD})\|$.

Projektiovektoriksi saadaan

$$\begin{aligned}\text{proj}_{\mathbf{n}}(\overline{AD}) &= \frac{\overline{AD} \cdot \mathbf{n}}{\mathbf{n} \cdot \mathbf{n}} \mathbf{n} = \frac{(5, 2, -1) \cdot (4, -3, 5)}{(4, -3, 5) \cdot (4, -3, 5)} (4, -3, 5) \\ &= \frac{20 - 6 - 5}{16 + 9 + 25} (4, -3, 5) = \frac{9}{50} (4, -3, 5).\end{aligned}$$

Projektiovektorin pituus on

$$\begin{aligned}\|\text{proj}_{\mathbf{n}}(\overline{AD})\| &= \left\| \frac{9}{50} (4, -3, 5) \right\| = \left| \frac{9}{50} \right| \|(4, -3, 5)\| \\ &= \frac{9}{50} \sqrt{16 + 9 + 25} = \frac{9}{50} \sqrt{50} = \frac{9}{\sqrt{50}}.\end{aligned}$$

Siten pisteen D etäisyys tasosta T on $9/\sqrt{50}$.

Esimerkki 2.2.10

Tutkitaan tasoa $T = \{\mathbf{x} \in \mathbb{R}^3 \mid x + 2y + 3z - 9 = 0\}$ sekä tasoa $U = \{\mathbf{x} \in \mathbb{R}^3 \mid 3x - 2y + 4z + 5 = 0\}$. Halutaan selvittää, mitkä pisteet kuuluvat sekä tasoon T että tasoon U .

Piste (x, y, z) kuuluu molempiin tasoihin, jos ja vain jos se toteuttaa molempien tasojen yhtälöt. Täytyy siis päteä yhtä aikaa $x + 2y + 3z - 9 = 0$ ja $3x - 2y + 4z + 5 = 0$. Saadaan yhtälöryhmä

$$\begin{cases} x + 2y + 3z - 9 = 0 \\ 3x - 2y + 4z + 5 = 0. \end{cases}$$

Tällaisia yhtälöryhmiä opitaan ratkaisemaan **hetken kuluttua**.

 Tiivistelmä

1. Tason pisteet saadaan lisäämällä paikkavektoriin suuntavektorien lineaarikombinaatioita.
2. Avaruuden $\mathbb{R}^3$ tason voi myös kirjoittaa normaalivektorin avulla. Sen on vektori, joka on kohtisuorassa tasoa vastaan.

3 MATRIISIT

3.1 Mitä matriisit ovat?

Pohdi 3.1.1

Ruska ja Tuisku ovat lähdössä ruokaostoksille ja vertailevat hintoja kahdessa lähikaupassaan. Alla olevassa taulukossa on heidän kauppalistansa sekä ruokatavaroiden hinnat eri kaupoissa:

	maitoja	sämpylöitä	jogurtteja		Y-kauppa	T-valinta
Ruska	6	4	4	maito	1,40 e	1,30 e
Tuisku	6	2	3	sämpylä	1,10 e	1,15 e
				jogurtti	0,50 e	0,60 e

1. Kuinka monta jogurttia Tuisku aikoo ostaa?
2. Mitä kaikkea Ruska on ostamassa ja kuinka paljon?
3. Mitkä ovat tuotteiden hinnat T-valinnassa?

Kuten Ruskan ja Tuiskun tapauksesta nähdään toisinaan asiat on kätevää kirjoittaa muistiin taulukkoon. Matematiikassa lukutaulukkoja kutsutaan matriiseiksi. Esimerkiksi Ruskan ja Tuiskun kauppalistan voi kirjoittaa matriisina

$$A = \begin{bmatrix} 6 & 4 & 4 \\ 6 & 2 & 3 \end{bmatrix}$$

ja ruokatavaroiden hinnat matriisiksi

$$B = \begin{bmatrix} 1,40 & 1,30 \\ 1,10 & 1,15 \\ 0,50 & 0,60 \end{bmatrix}.$$

Esimerkiksi Tuiskun ostama jogurttimäärä näkyy matriisissa A rivillä 2 ja sarakkeessa 3.

Ruskan ostoslista näkyy matriisin A ensimmäisellä rivillä. T-valinnan hinnat puolestaan näkyvät matriisin B toisessa sarakkeessa.

Reaali-alkioinen $m \times n$ -matriisi on reaalilukutaulukko, jossa on m riviä ja n saraketta. Esimerkiksi

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

on $m \times n$ -matriisi. Voidaan myös sanoa, että matriisin A koko on $m \times n$. Kaikkien reaalikertoimisten $m \times n$ -matriisien joukkoa merkitään $\mathbb{R}^{m \times n}$. Matriisissa olevia lukuja kutsutaan matriisin *alkioiksi*. Rivillä i sarakkeessa j olevalle alkioille käytetään tässä tekstissä merkintää $A(i, j)$.

Esimerkki 3.1.2

Esimerkiksi

$$B = \begin{bmatrix} 1 & 0 & 5 \\ -3 & 11 & 2 \\ 4 & 0 & 2 \\ 0 & \sqrt{2} & -6 \end{bmatrix}$$

on reaali-alkioinen 4×3 -matriisi eli $B \in \mathbb{R}^{4 \times 3}$. Siinä $B(1, 3) = 5$ ja $B(2, 2) = 11$.

Vektorit ovat matriiseja, joissa on yksi sarake. Toisinaan on hyödyllistä ajatella matriisiä vierekkäin asetetuksi pystyvektoreiksi. Tällöin voidaan kirjoittaa

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix},$$

missä $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathbb{R}^m$. Nyt vektorit $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ kuvaavat matriisin sarakkeita.

Matriisien alkioille on matematiikassa käytössä monenlaisia merkintöjä. Esimerkiksi matriisin alkioille voidaan käyttää merkinnän $A(i, j)$ sijasta merkintää $[A]_{ij}$. Matriisille $A \in \mathbb{R}^{m \times n}$ puolestaan voidaan käyttää merkintää $A = [a_{ij}]_{m \times n}$ tai $A = [a_{ij}]$, joissa on ilmaistu, että matriisin alkiot ovat muotoa a_{ij} .

3.2 Matriisien laskutoimituksia

Matriiseille, kuten vektoreillekin, voidaan määritellä erilaisia laskutoimituksia. Osa niistä muistuttaa läheisesti vektorien vastaavia laskutoimituksia.

Matriisien yhteenlasku

Matriisien *yhteenlasku* määritellään seuraavasti. Olkoot $A, B \in \mathbb{R}^{m \times n}$. Matriisien A ja B summa saadaan laskemalla yhteen matriisien vastinsarakkeet:

$$A + B = \begin{bmatrix} \mathbf{a}_1 + \mathbf{b}_1 & \mathbf{a}_2 + \mathbf{b}_2 & \cdots & \mathbf{a}_n + \mathbf{b}_n \end{bmatrix}.$$

Tuloksena on $m \times n$ -matriisi.

Sama asia voidaan sanoa toteamalla, että summamatriisi saadaan laskemalla yhteen samoissa kohdissa olevat alkiot:

$$(A + B)(i, j) = A(i, j) + B(i, j)$$

kaikilla $i \in \{1, \dots, m\}$ ja $j \in \{1, \dots, n\}$.

Esimerkiksi

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ 0 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1+2 & 2+(-1) \\ 3+0 & 4+1 \\ 5+3 & 6+2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 3 & 5 \\ 8 & 8 \end{bmatrix}.$$

Vain samankokoisia matriiseja voidaan laskea yhteen.

Skalaarikertolasku

Minkä tahansa matriisin voi kertoa reaalityluvulla ja tätä toimitusta kutsutaan *skalaarikertolaskuksi*. Reaalityluvun c ja matriisin $A \in \mathbb{R}^{m \times n}$ ja tulo saadaan kertomalla matriisin sarakkeet luvulla c :

$$cA = \begin{bmatrix} c\mathbf{a}_1 & c\mathbf{a}_2 & \cdots & c\mathbf{a}_n \end{bmatrix}.$$

Saatava tulos on $m \times n$ -matriisi, jota nimitetään matriisin A *skalaarimonikerraksi*.

Sama asia voidaan sanoa toteamalla, että skalaarimatriisi saadaan kertomalla matriisin A jokaista alkioita luvulla c :

$$(cA)(i, j) = c \cdot A(i, j)$$

kaikilla $i \in \{1, \dots, m\}$ ja $j \in \{1, \dots, n\}$.

Esimerkiksi

$$2 \begin{bmatrix} 2 & -1 \\ 0 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 2 & 2 \cdot (-1) \\ 2 \cdot 0 & 2 \cdot 1 \\ 2 \cdot 3 & 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 0 & 2 \\ 6 & 4 \end{bmatrix}.$$

Matriisia $(-1)A$ on tapana merkitä $-A$. Matriisisummaa $A + (-B)$ merkitään $A - B$ ja sitä kutsutaan matriisien A ja B *erotukseksi*.

Matriisin ja vektorin tulo

Pohdi 3.2.1

Opiskelijoiden päivittäinen matkustaminen riippuu edellisestä päivästä seuraavasti. Opiskelijoista, jotka tulivat yliopistolle julkisella liikenteellä, saapuu julkisella liikenteellä seuraavana päivänä 80 % ja polkupyörällä 20 %. Opiskelijoista, jotka tulivat yliopistolle pyörällä, tulee seuraavana päivänä julkisella liikenteellä 30 % ja pyörällä 70 %. Eräänä päivänä yliopistolle tuli julkisella liikenteellä 1000 opiskelijaa ja pyörällä 200 opiskelijaa.

Kuinka monta opiskelijaa tuli yliopistolle julkisella liikenteellä seuraavana päivänä? Entä pyörällä?

Matriiseille voidaan määritellä myös *matriisikertolasku*. Tämä laskutoimitus on hieman monimutkaisempi kuin edellä määritellyt eikä mitään vastaavaa ole olemassa vektoreille.

Ennen kuin tutkitaan kahden matriisin välistä tuloa, tarkastellaan sen erikoistapausta, matriisin ja vektorin välistä kertolaskua. Tehdään se edellä esitetyn pohdintatehtävän avulla. Opiskelijoiden käyttäytymiseen liittyvät tiedot voidaan tallentaa matriisiin

$$A = \begin{bmatrix} 0,8 & 0,3 \\ 0,2 & 0,7 \end{bmatrix}.$$

Tehtävässä kuvatun päivän tilannetta puolestaan kuvaa vektori

$$\mathbf{v} = \begin{bmatrix} 1000 \\ 200 \end{bmatrix}.$$

Seuraavan päivänä julkista liikennettä käytti $0,8 \cdot 1000 + 0,2 \cdot 200$ opiskelijaa. Polkupyörää puolestaan käytti $0,3 \cdot 1000 + 0,7 \cdot 200$ opiskelijaa. Seuraavan päivän tilannetta kuvaa siis vektori

$$\begin{bmatrix} 0,8 \cdot 1000 + 0,2 \cdot 200 \\ 0,3 \cdot 1000 + 0,7 \cdot 200 \end{bmatrix}.$$

Tämä vektori voidaan kirjoittaa muodossa

$$1000 \begin{bmatrix} 0,8 \\ 0,2 \end{bmatrix} + 200 \begin{bmatrix} 0,3 \\ 0,7 \end{bmatrix}.$$

Tämän vektorin sanotaan olevan matriisin A ja vektorin $\mathbf{v}$ tulo. Toisin sanoen

$$A\mathbf{v} = \begin{bmatrix} 0,8 & 0,3 \\ 0,2 & 0,7 \end{bmatrix} \begin{bmatrix} 1000 \\ 200 \end{bmatrix} = 1000 \begin{bmatrix} 0,8 \\ 0,2 \end{bmatrix} + 200 \begin{bmatrix} 0,3 \\ 0,7 \end{bmatrix} = \begin{bmatrix} 860 \\ 340 \end{bmatrix}.$$

Matriisiin ja vektorin tulo lasketaan siis kertomalla vektorin komponenteilla matriisin sarakkeita ja summaamalla näin saadut vektorit yhteen. Opiskelijoiden matkustamista käsittelevän esimerkin tapauksessa tulo kertoo seuraavan päivän tilanteen, kun jonkin tietyn päivän tilanne tiedetään.

Määritelmä 3.2.2

Matriisin $A \in \mathbb{R}^{m \times n}$ ja vektorin $\mathbf{v} \in \mathbb{R}^n$ tulo on

$$A\mathbf{v} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + \cdots + v_n\mathbf{a}_n.$$

Tässä vektorit $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ ovat matriisin A sarakkeet ja reaaliluvut $v_1, v_2, \dots, v_n$ vektorin $\mathbf{v}$ komponentit.

Huomaa, että kaikkia matriiseja ja vektoreita ei voi kertoa keskenään. Jotta matriisilla voi kertoa vektoria, täytyy matriisiin rivissä olla yhtä monta alkioita kuin vektorissa komponentteja.

Esimerkki 3.2.3

Lasketaan matriisin

$$A = \begin{bmatrix} 1 & 4 & -2 \\ 2 & 0 & 3 \\ 0 & 1 & 7 \\ 3 & -3 & 6 \end{bmatrix}$$

ja vektorin

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

tulo.

Matriisin ja vektorin tulon laskemiseksi jokainen matriisin A sarake kerrotaan vastaa-

valla vektorin $\mathbf{v}$ komponentilla ja tulokset lasketaan yhteen:

$$A\mathbf{v} = 1 \begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} 4 \\ 0 \\ 1 \\ -3 \end{bmatrix} + 3 \begin{bmatrix} -2 \\ 3 \\ 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 8 \\ 0 \\ 2 \\ -6 \end{bmatrix} + \begin{bmatrix} -6 \\ 9 \\ 21 \\ 18 \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \\ 23 \\ 15 \end{bmatrix}.$$

Matriisin $A \in \mathbb{R}^{m \times n}$ ja vektorin $\mathbf{v} \in \mathbb{R}^n$ tulo voidaan kirjoittaa myös toisenlaisessa muodossa pistetulon avulla. Olkoot $\mathbf{r}_1, \dots, \mathbf{r}_m$ matriisin A rivit. Nyt

$$A\mathbf{v} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{v} \\ \mathbf{r}_2 \cdot \mathbf{v} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{v} \end{bmatrix},$$

Tulon $A\mathbf{v}$ alkio on siis matriisin A rivien ja vektorin $\mathbf{v}$ pistetuloja.

Esimerkki 3.2.4

Lasketaan [esimerkissä 3.2.3](#) esiintynyt matriisin ja vektorin tulo uudelleen, mutta käytetään tällä kertaa edellä esitettyä laskutapaa. Nyt

$$A\mathbf{v} = \begin{bmatrix} 1 & 4 & -2 \\ 2 & 0 & 3 \\ 0 & 1 & 7 \\ 3 & -3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 4 \cdot 2 - 2 \cdot 3 \\ 2 \cdot 1 + 0 \cdot 2 + 3 \cdot 3 \\ 0 \cdot 1 + 1 \cdot 2 + 7 \cdot 3 \\ 3 \cdot 1 - 3 \cdot 2 + 6 \cdot 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \\ 23 \\ 15 \end{bmatrix}.$$

Kahden matriisin tulo

Pohdi 3.2.5

Palataan Ruskan ja Tuiskun seuraan. He ovat edelleen lähdössä ruokaostoksille ja vertailevat hintoja kahdessa lähikaupassaan. Tässä vielä Ruskan ja Tuiskun kauppalista sekä ruokatavaroiden hinnat eri kaupoissa.

	maitoja	sämpylöitä	jogurtteja		Y-kauppa	T-valinta
Ruska	6	4	4	maito	1,40 e	1,30 e
Tuisku	6	2	3	sämpylä	1,10 e	1,15 e
				jogurtti	0,50 e	0,60 e

Minkälaisilla laskuilla voi laskea seuraavat tiedot?

1. Ruskan ostosten yhteishinta Y-kaupassa
2. Tuiskun ostosten yhteishinta Y-kaupassa
3. Ruskan ostosten yhteishinta T-valinnassa
4. Tuiskun ostosten yhteishinta T-valinnassa

Tähän mennessä on kerrottu matriisilla vektoreita. Vektoria voi ajatella matriisina, jossa on yksi sarake. Siirrytään nyt tutkimaan yleisesti, millainen on kahden matriisin välinen tulo. Olkoon matriisi A edellisen pohdintatehtävän kauppalista ja matriisi B hintataulukko. Toisin sanoen

$$A = \begin{bmatrix} 6 & 4 & 4 \\ 6 & 2 & 3 \end{bmatrix}$$

ja

$$B = \begin{bmatrix} 1,40 & 1,30 \\ 1,10 & 1,15 \\ 0,50 & 0,60 \end{bmatrix}.$$

Matriisien A ja B tulo on sellainen matriisi, joka sisältää tiedon ostosten yhteishinnoista eri kaupoissa.

Y-kaupassa Ruskan ja Tuiskun ostosten yhteishinnat lasketaan seuraavasti:

$$6 \cdot 1,40 + 4 \cdot 1,10 + 4 \cdot 0,50$$

$$6 \cdot 1,40 + 2 \cdot 1,10 + 3 \cdot 0,50.$$

Tätä voi kuvata vektorilla

$$\begin{bmatrix} 6 \cdot 1,40 + 4 \cdot 1,10 + 4 \cdot 0,50 \\ 6 \cdot 1,40 + 2 \cdot 1,10 + 3 \cdot 0,50 \end{bmatrix}.$$

T-valinnassa yhteishinnat puolestaan ovat

$$6 \cdot 1,30 + 4 \cdot 1,15 + 4 \cdot 0,60$$

$$6 \cdot 1,30 + 2 \cdot 1,15 + 3 \cdot 0,60$$

ja niitä voi kuvata vektorilla

$$\begin{bmatrix} 6 \cdot 1,30 + 4 \cdot 1,15 + 4 \cdot 0,60 \\ 6 \cdot 1,30 + 2 \cdot 1,15 + 3 \cdot 0,60 \end{bmatrix}.$$

Tulomatriisissa AB on sarakkeina nämä kaksi vektoria:

$$AB = \begin{bmatrix} 6 \cdot 1,40 + 4 \cdot 1,10 + 4 \cdot 0,50 & 6 \cdot 1,30 + 4 \cdot 1,15 + 4 \cdot 0,60 \\ 6 \cdot 1,40 + 2 \cdot 1,10 + 3 \cdot 0,50 & 6 \cdot 1,30 + 2 \cdot 1,15 + 3 \cdot 0,60 \end{bmatrix} = \begin{bmatrix} 14,80 & 14,80 \\ 12,10 & 11,90 \end{bmatrix}.$$

Matriisista AB nähdään, kummassa kaupassa Ruskan ja Tuiskun kannattaa asioida. Ruskan ostokset ovat tulevat molemmissa kaupoissa yhtä kalliiksi (14,80 euroa). Tuiskun ostokset puolestaan tulevat halvemmiksi T-valinnassa (11,90 euroa).

Tulomatriisin sarakevektorit voi myös kirjoittaa muodossa

$$\begin{bmatrix} 6 \cdot 1,40 + 4 \cdot 1,10 + 4 \cdot 0,50 \\ 6 \cdot 1,40 + 2 \cdot 1,10 + 3 \cdot 0,50 \end{bmatrix} = 1,40 \begin{bmatrix} 6 \\ 6 \end{bmatrix} + 1,10 \begin{bmatrix} 4 \\ 2 \end{bmatrix} + 0,50 \begin{bmatrix} 4 \\ 3 \end{bmatrix} = A\mathbf{b}_1$$

ja

$$\begin{bmatrix} 6 \cdot 1,30 + 4 \cdot 1,15 + 4 \cdot 0,60 \\ 6 \cdot 1,30 + 2 \cdot 1,15 + 3 \cdot 0,60 \end{bmatrix} = 1,30 \begin{bmatrix} 6 \\ 6 \end{bmatrix} + 1,15 \begin{bmatrix} 4 \\ 2 \end{bmatrix} + 0,60 \begin{bmatrix} 4 \\ 3 \end{bmatrix} = A\mathbf{b}_2,$$

missä $\mathbf{b}_1$ ja $\mathbf{b}_2$ ovat matriisin B sarakkeet. Tulomatriisin AB sarakkeet saadaan siis kertomalla matriisilla A matriisin B sarakkeita.

Määritelmä 3.2.6

Matriisien $A \in \mathbb{R}^{m \times n}$ ja $B \in \mathbb{R}^{n \times p}$ tulo on

$$AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}.$$

Tässä vektorit $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ ovat matriisin B sarakkeet. Huomaa, että kaksi matriisia voidaan kertoa keskenään vain, jos ensimmäisessä on yhtä paljon sarakkeita kuin toisessa on rivejä.

Jos matriisi B on vektori eli siinä on vain yksi sarake, on matriisitulon määritelmä sama kuin matriisin ja vektorin välisen tulon määritelmä. Matriisien tulon määritelmä siis yleistää matriisin ja vektorin välistä tuloa.

Esimerkki 3.2.7

Lasketaan matriisien

$$A = \begin{bmatrix} 4 & 5 \\ 6 & 7 \\ -1 & -2 \end{bmatrix} \quad \text{ja} \quad B = \begin{bmatrix} 1 & 2 \\ 8 & 9 \end{bmatrix}$$

tulo. Koska matriisissa A on kaksi saraketta (3×2 -matriisi) ja matriisissa B on vastaavasti kaksi riviä (2×2 -matriisi), matriisit voidaan kertoa keskenään. Tulomatriisi on 3×2 -matriisi.

Lasketaan ensin matriisin A tulot matriisin B sarakkeiden kanssa:

$$A\mathbf{b}_1 = \begin{bmatrix} 4 & 5 \\ 6 & 7 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 8 \end{bmatrix} = 1 \begin{bmatrix} 4 \\ 6 \\ -1 \end{bmatrix} + 8 \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ -1 \end{bmatrix} + \begin{bmatrix} 40 \\ 56 \\ -16 \end{bmatrix} = \begin{bmatrix} 44 \\ 62 \\ -17 \end{bmatrix}$$

ja

$$A\mathbf{b}_2 = \begin{bmatrix} 4 & 5 \\ 6 & 7 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 9 \end{bmatrix} = 2 \begin{bmatrix} 4 \\ 6 \\ -1 \end{bmatrix} + 9 \begin{bmatrix} 5 \\ 7 \\ -2 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \\ -2 \end{bmatrix} + \begin{bmatrix} 45 \\ 63 \\ -18 \end{bmatrix} = \begin{bmatrix} 53 \\ 75 \\ -20 \end{bmatrix}.$$

Nämä vektorit ovat matriisiin AB sarakkeet. Toisin sanoen

$$AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 44 & 53 \\ 62 & 75 \\ -17 & -20 \end{bmatrix}.$$

Esimerkki 3.2.8

Matriisien

$$C = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 6 & 7 \\ 3 & -1 & -2 \end{bmatrix} \quad \text{ja} \quad D = \begin{bmatrix} 1 & 2 & 3 \\ 8 & 9 & 10 \end{bmatrix}$$

tuloa ei voi laskea. Matriisissa C on nimittäin kolme saraketta ja matriisissa D kaksi riviä. Koska nämä lukumäärät eivät ole samat, ei tuloa CD ole määritelty.

Matriisien $A \in \mathbb{R}^{m \times n}$ ja $B \in \mathbb{R}^{n \times p}$ tulo voidaan kirjoittaa myös toisenlaisessa muodossa.

Olkoot $\mathbf{r}_1, \dots, \mathbf{r}_m$ matriisin A rivit. Nyt

$$AB = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{b}_1 & \mathbf{r}_1 \cdot \mathbf{b}_2 & \cdots & \mathbf{r}_1 \cdot \mathbf{b}_p \\ \mathbf{r}_2 \cdot \mathbf{b}_1 & \mathbf{r}_2 \cdot \mathbf{b}_2 & \cdots & \mathbf{r}_2 \cdot \mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_m \cdot \mathbf{b}_1 & \mathbf{r}_m \cdot \mathbf{b}_2 & \cdots & \mathbf{r}_m \cdot \mathbf{b}_p \end{bmatrix}.$$

Matriisin AB rivin i ja sarakkeen j alkio on siis matriisin A rivin i ja matriisin B sarakkeen j pistetulo.

Esimerkki 3.2.9

Lasketaan [esimerkissä 3.2.7](#) esiintyneiden matriisien

$$A = \begin{bmatrix} 4 & 5 \\ 6 & 7 \\ -1 & -2 \end{bmatrix} \quad \text{ja} \quad B = \begin{bmatrix} 1 & 2 \\ 8 & 9 \end{bmatrix}$$

tulo uudelleen, mutta käytetään tällä kertaa edellä esitettyä laskutapaa. Nyt

$$AB = \begin{bmatrix} 4 & 5 \\ 6 & 7 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 8 & 9 \end{bmatrix} = \begin{bmatrix} 4 \cdot 1 + 5 \cdot 8 & 4 \cdot 2 + 5 \cdot 9 \\ 6 \cdot 1 + 7 \cdot 8 & 6 \cdot 2 + 7 \cdot 9 \\ (-1) \cdot 1 + (-2) \cdot 8 & (-1) \cdot 2 + (-2) \cdot 9 \end{bmatrix} = \begin{bmatrix} 44 & 53 \\ 62 & 75 \\ -17 & -20 \end{bmatrix}.$$

Kolmas tapa kirjoittaa tulomatriisi AB on ilmaista kaava, jolla tulomatriisin alkiot saadaan:

$$\begin{aligned} (AB)(i, j) &= A(i, 1)B(1, j) + A(i, 2)B(2, j) + \cdots + A(i, n)B(n, j) \\ &= \sum_{k=1}^n A(i, k)B(k, j) \end{aligned}$$

kaikilla $i \in \{1, \dots, m\}$ ja $j \in \{1, \dots, p\}$. Merkintä

$$\sum_{k=1}^n c_k$$

tarkoittaa summaa $c_1 + c_2 + \cdots + c_n$.

Esimerkki 3.2.10

Matriisikertolasku ei ole vaihdannainen operaatio eli tulon tekijöiden järjestystä ei voi

vaihtaa. Tarkastellaan vaikkapa matriiseja

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

Laskemalla tulo molemmin päin huomataan, että

$$AB = \begin{bmatrix} -4 & 7 \\ -8 & 5 \end{bmatrix} \quad \text{mutta} \quad BA = \begin{bmatrix} 3 & 6 \\ -7 & -2 \end{bmatrix}.$$

Siten $AB \neq BA$.

Kertolaskun avulla matriisista voidaan poimia jokin tietty sarake tai rivi. Palautetaan mieleen, että merkinnällä $\mathbf{e}_i$ tarkoitetaan luonnollisen kannan vektoria $(0, \dots, 0, 1, 0, \dots, 0)$, jossa 1 on i :nnes komponentti.

Lause 3.2.11

Olkoon A $m \times n$ -matriisi. Oletetaan, että $i \in \{1, \dots, n\}$ ja $j \in \{1, \dots, m\}$. Tällöin

1. $A\mathbf{e}_i$ on matriisin A i :s sarake
2. $\mathbf{e}_j^\top A$ on matriisin A j :s rivi.

Todistus. Todistetaan esimerkin vuoksi ensimmäinen kohta. Matriisin ja vektorin tulon määritelmän mukaan

$$A\mathbf{e}_i = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \dots + 0\mathbf{a}_{i-1} + 1\mathbf{a}_i + 0\mathbf{a}_{i+1} + \dots + 0\mathbf{a}_n = \mathbf{a}_i.$$

Toisessa kohdassa merkintä $\mathbf{e}_j^\top$ tarkoittaa *vaakavektoria*

$$\mathbf{e}_j^\top = \begin{bmatrix} 0 & 0 & \dots & 1 & \dots & 0 \end{bmatrix},$$

missä luku 1 esiintyy j :nnessä sarakkeessa. Loput jätetään harjoitustehtäväksi. □

Matrisitulon avulla voidaan määritellä myös neliömatriisin potenssi.

Määritelmä 3.2.12

Oletetaan, että A on $n \times n$ -matriisi ja $k \in \{1, 2, \dots\}$. Tällöin matriisin A k :s *potenssi* on

$$A^k = \underbrace{AA \cdots A}_{k \text{ kpl}}.$$

Tiivistelmä

- Kaksi samankokoista matriisia voi laskea yhteen.
- Matriiseja voi kertoa reaalityyppillä. Tätä kutsutaan skalaarikertolaskuksi.
- Kaksi matriisia voi kertoa keskenään. Tällöin kuitenkin ensimmäisen matriisin sarakkeiden lukumäärän täytyy olla sama kuin toisen matriisin rivien lukumäärä.

3.3 Matriisit kuvauksina

Videossa ”Lineaarikuvaukset ja matriisit” näytetään, kuinka matriiseja voi ajatella kuvauksina. Tätä kautta saadaan myös havainnollistus matriisikertolaskulle.

Matriisia voidaan ajatella kuvauksena, joka kuvaa vektoreita toisiksi vektoreiksi kertolaskun avulla. Esimerkiksi matriisi

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

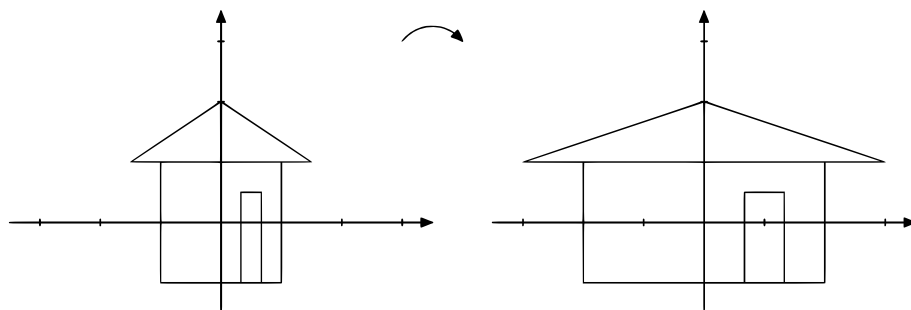
tuottaa kuvauksen, joka kuvaa $\mathbb{R}^2$ vektorin $\mathbf{x}$ vektoriksi $A\mathbf{x}$. Esimerkiksi vektori $\mathbf{x} = (-4, 3)$ kuvautuu vektoriksi

$$A\mathbf{x} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \end{bmatrix} = \begin{bmatrix} -8 \\ 3 \end{bmatrix}.$$

Toisin sanoen $(-4, 3) \mapsto (-8, 3)$. Yleisesti vektori $\mathbf{x} \in \mathbb{R}^2$ kuvautuu kuvauksessa vektoriksi

$$A\mathbf{x} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ x_2 \end{bmatrix}.$$

Toisin sanoen $(x_1, x_2) \mapsto (2x_1, x_2)$. Vektorin ensimmäinen komponentti siis kaksinkertaistuu ja toinen komponentti pysyy samana. Niinpä matriisi A venyttää vektoreita vaakasuunnassa (kuva 3.3.1).



Kuva 3.3.1. Matriisi A venyttää vektoreita vaakasuunnassa.

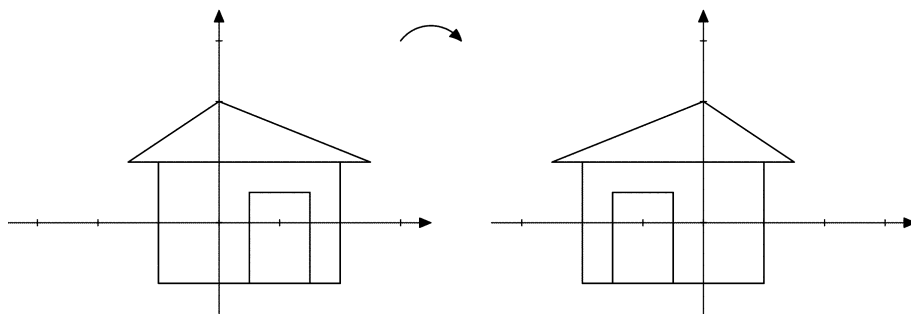
Tutkitaan, sitten millaisen kuvauksen määrää matriisi

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

Vektori $\mathbf{x} \in \mathbb{R}^2$ kuvautuu kuvauksessa vektoriksi

$$B\mathbf{x} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}.$$

Toisin sanoen $(x_1, x_2) \mapsto (-x_1, x_2)$. Matriisi B *peilaa* vektorit x_2 -akselin suhteen (kuva 3.3.2).



Kuva 3.3.2. Matriisi B peilaa vektorit pystyakselin suhteen.

Tutkitaan vielä matriisin

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

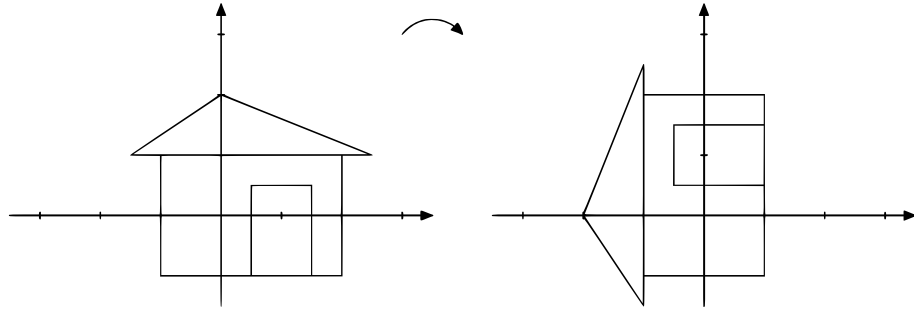
määrää kuvauksia. Vektori $\mathbf{x} \in \mathbb{R}^2$ kuvautuu kuvauksessa vektoriksi

$$C\mathbf{x} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$

Toisin sanoen $(x_1, x_2) \mapsto (-x_2, x_1)$. Matriisi C *kiertää* vektoreita kulman $\pi/2$ verran vastapäivään eli positiiviseen kiertosuuntaan (kuva 3.3.3). Voidaan osoittaa, että matriisin

$$\begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

määrää lineaarikuvaus kiertää vektoreita origon ympäri kulman φ verran (vastapäivään, jos $\varphi > 0$, ja myötäpäivään, jos $\varphi < 0$).



Kuva 3.3.3. Matriisi C kiertää vektoreita kulman $\pi/2$ verran positiiviseen kiertosuuntaan.

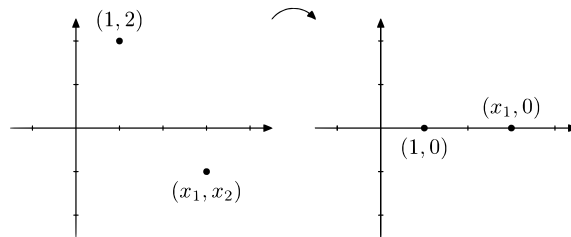
Tarkastellaan lopuksi matriisia

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

Vektori $\mathbf{x} \in \mathbb{R}^2$ kuvautuu kuvauksessa vektoriksi

$$P\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}.$$

Toisin sanoen $(x_1, x_2) \mapsto (x_1, 0)$. Matriisi P projisoi vektorit vaaka-akselille (kuva 3.3.4).

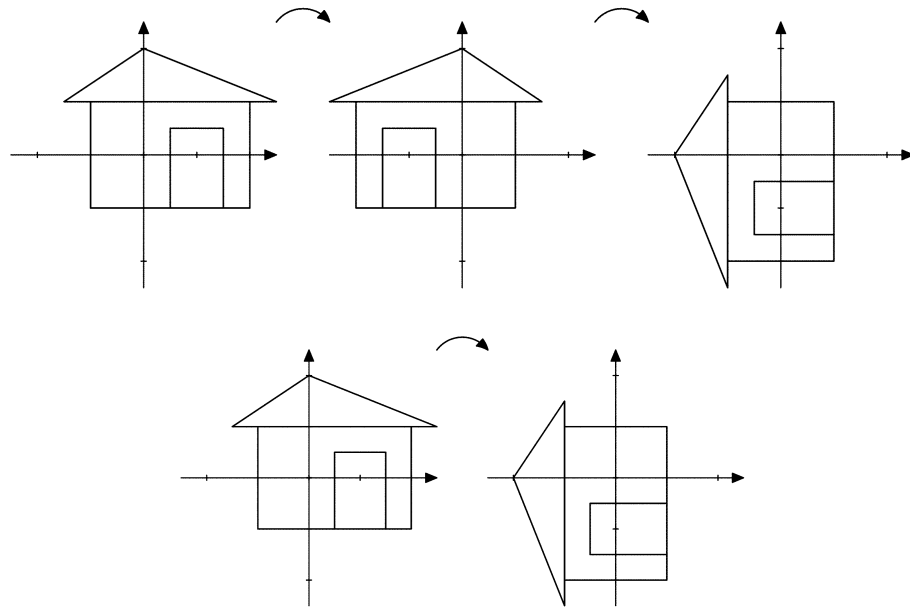


Kuva 3.3.4. Matriisi P projisoi vektorit vaaka-akselille.

Kun matriisia ajatellaan kuvauksena, matriisin sarakkeina ovat luonnollisen kannan vektoreiden kuvavektorit. Esimerkiksi edellä esitetyn matriisin A sarakkeet ovat $A\mathbf{e}_1$ ja $A\mathbf{e}_2$, missä $\mathbf{e}_1 = (1, 0)$ ja $\mathbf{e}_2 = (0, 1)$. Tämä seuraa suoraan [lauseesta 3.2.11](#), jonka nojalla matriisin tulo luonnollisen kannan vektorin kanssa poimii matriisista vastaavan sarakkeen.

Kun matriiseja ajatellaan kuvauksina, matriisitulo vastaa sitä, että kaksi kuvausta tehdään peräkkäin. Toisin sanoen matriisitulo vastaa yhdistettyä kuvausta.

Tutkitaan, mitä matriisien yhdistetty kuvaus näyttää geometrisesti. Tehdään ensin peilaus B ja sitten kierto C . Yhdistetty kuvaus ensin peilaa vektorit pysty-akselin suhteen ja sen jälkeen kiertää niitä kulman $\pi/2$ verran vastapäivään (ks. kuva 3.3.5).



Kuva 3.3.5. Ylemmässä kuvassa näytetään, mitä vektoreille tapahtuu, kun niitä kerrotaan ensin matriisilla B ja sitten matriisilla C . Alemmassa kuvassa sama saadaan aikaan kertomalla vektoreita matriisilla CB .

Tutkitaan sitten, miten yhdistetty kuvaus lasketaan. Kun vektoria $\mathbf{x} \in \mathbb{R}^2$ kuvataan matriisilla B , saadaan vektori $B\mathbf{x}$. Kun tätä tulosvektoria kuvataan matriisilla C saadaan vektori $C(B\mathbf{x})$. Matriisien laskusääntöjen (lause 3.4.1) nojalla tämä on sama asia kuin $(CB)\mathbf{x}$. Toisin sanoen tulomatriisi CB on kuvaus, joka ensin kiertää ja sitten peilaa vektoreita.

Ensi alkuun voi tuntua siltä, että matriisit ovat tulomatriisissa väärin päin. Jos peilaus B tehdään ensin, miksi se on tulossa CB vasta toisena tulon tekijänä? Tämä kuitenkin johtuu tavasta, jolla kuvauksia kirjoitetaan. Kuvaus B kirjoitetaan alkion $\mathbf{x}$ vasemmalle puolelle ($B\mathbf{x}$). Siksi tulossa CB kuvaus B tehdään ensin.

✈️ Tiivistelmä

- Matriiseja voidaan ajatella kuvauksina, jotka kuvaavat vektoreita.
- Matriisien kertolasku vastaa kuvausten yhdistämistä.

3.4 Erityisiä matriiseja ja matriisien laskusääntöjä

Seuraavaksi esitellään muutamia matriiseja tai matriisityyppejä, jotka ansaitsevat hyödyllisyytensä vuoksi oman nimen. Matriiseja, joiden kaikki alkiot ovat nollia, eli jotka ovat muotoa

$$O_{m \times n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{m \times n},$$

kutsutaan *nollamatriiseiksi*. *Ykkösmatriiseja* puolestaan ovat matriisit

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Ykkösmatriiseja kutsutaan myös *yksikkömatriseksi* tai *identiteettimatriiseksi*.

Nollamatriisissa voi olla mikä tahansa määrä rivejä ja sarakkeita. Sen sijaan ykkösmatriisissa on aina yhtä paljon rivejä ja sarakkeita. Jos matriisin tyylistä ei ole epäselvyyttä, saatetaan merkitä yksinkertaisemmin $O_{m \times n} = O$ ja $I_n = I$.

Nollamatriisit käyttäytyvät matriisien yhteenlaskun suhteen samalla tavalla kuin nolla lukujen yhteenlaskussa (tai nollavektori vektorien yhteenlaskussa): sellaisen lisääminen toiseen matriisiin ei muuta tuota toista matriisia mitenkään. Samalla tavoin ykkösmatriisit käyttäytyvät matriisikertolaskussa aivan kuten reaalityttö 1 tavallisessa kertolaskussa. Kaikilla $A \in \mathbb{R}^{m \times n}$ pätee nimittäin

$$I_m A = A \quad \text{ja} \quad A I_n = A.$$

Eri puolilta kerrottaessa on matriisikertolaskun rajoituksen vuoksi käytettävä eri kokoista ykkösmatriiseja.

Neliömatrissi on matrissi, jossa on yhtä monta riviä ja saraketta. Esimerkiksi ykkösmatriisit ovat neliömatriseja.

Neliömatrissin alkio on *lävistäjällä* eli *diagonaalilla*, jos alkion rivin ja sarakkeen numerot ovat samat. Matrissi, jonka kaikki nollasta poikkeavat alkioit ovat lävistäjällä, on *lävistäjämatriisi* eli *diagonaalimatriisi*. Lävistäjämatrissi, jonka kaikki lävistäjäalkioit ovat samoja, on puolestaan *skalaarimatriisi*. Skalaarimatriisit ovat ykkösmatriissin skalaarimonikertoja. Esimerkiksi

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 15 \end{bmatrix}$$

on lävistäjämatriisi ja

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = 2I_3$$

on skalaarimatriisi.

Matriisien laskutoimitukset noudattavat tiettyjä sääntöjä, jotka monessa kohdassa muistuttavat lukujen laskusääntöjä.

Lause 3.4.1

Seuraavat säännöt pätevät matriiseille A , B ja C sekä reaalityyppisille a ja b , jos laskutoimitukset on määritelty:

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $A(BC) = (AB)C$
4. $A(B + C) = AB + AC$
5. $(A + B)C = AC + BC$
6. $(ab)A = a(bA)$
7. $a(AB) = (aA)B = A(aB)$.

Kuten aiemmin on jo mainittu, yleisessä tapauksessa $AB \neq BA$, eli tulon vaihdannaisuus ei päde matriiseilla. Myöskään tulon nollasääntö ei päde matriiseilla: kahden matriisin tulo voi olla nollamatriisi, vaikka kumpikaan tulon tekijöistä ei ole nollamatriisi.

Todistus. Osoitetaan esimerkin vuoksi kohta 4. Oletetaan, että $A \in \mathbb{R}^{m \times n}$ ja $B, C \in \mathbb{R}^{n \times p}$. Nyt $A(B + C)$ ja $AB + AC$ ovat molemmat $m \times p$ -matriiseja. On osoitettava, että kyseisten matriisien alkiot ovat samoja. Olkoot sitä varten $i \in \{1, 2, \dots, m\}$ ja $j \in \{1, 2, \dots, p\}$. Nähdään, että

$$\begin{aligned}
 (A(B + C))(i, j) &= \sum_{k=1}^n A(i, k) \cdot (B + C)(k, j) \\
 &= \sum_{k=1}^n A(i, k) (B(k, j) + C(k, j)) \\
 &= \sum_{k=1}^n (A(i, k)B(k, j) + A(i, k)C(k, j)) \\
 &= \sum_{k=1}^n A(i, k)B(k, j) + \sum_{k=1}^n A(i, k)C(k, j) \\
 &= (AB)(i, j) + (AC)(i, j) = (AB + AC)(i, j).
 \end{aligned}$$

Koska matriisit $A(B + C)$ ja $AB + AC$ ovat samankokoisia ja niillä on täsmälleen samat alkiot, pätee $A(B + C) = AB + AC$. □

Tiivistelmä

- Kun nollamatriisin O laskee yhteen toisen matriisin kanssa, ei tapahdu mitään.

- Kun ykkösmatriisilla I kertoo toista matriisia, ei tapahdu mitään.
- Matriisien laskutoimituksille pätevät monet samat laskusäännöt kuin reaalilukujen laskutoimituksille.
- Joidenkin laskusääntöjen suhteen matriisit kuitenkin käyttäytyvät eri tavalla kuin reaaliluvut. Matriisitulo ei esimerkiksi ole vaihdannainen eikä sille päde tulon nollasääntö.
- Kun sievennät lausekkeita matriisien laskusääntöjen avulla, mieti aina tarkkaan, päteekö soveltamasi laskusääntö!

3.5 Matriisin transpoosi

Määritelmä 3.5.1

Oletetaan, että A on $m \times n$ -matriisi. Sen *transpoosi* A^T on $n \times m$ -matriisi, joka saadaan vaihtamalla matriisin A rivit ja sarakkeet keskenään.

Esimerkiksi matriisin

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 5 & 0 & 1 \end{bmatrix}$$

transpoosi on

$$A^T = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 1 \end{bmatrix}.$$

Pistetuloa voidaan ajatella matriisin ja vektorin tulona. Näin saadaan pistetulolle toisenlainen kirjoitustapa. Jotta vektorit voi kertoa keskenään, täytyy ensimmäisestä vektorista ottaa transpoosi. Oletetaan, että $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Tällöin matriisin ja vektorin tulon määritelmän nojalla

$$\begin{aligned} \mathbf{x}^T \mathbf{y} &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y_1 x_1 + y_2 x_2 + \cdots + y_n x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \end{bmatrix} = \begin{bmatrix} \mathbf{x} \cdot \mathbf{y} \end{bmatrix}. \end{aligned}$$

Kun samastetaan 1×1 -matriisit ja reaaliluvut, saadaan $\mathbf{x}^T \mathbf{y} = \mathbf{x} \cdot \mathbf{y}$.

Määritelmä 3.5.2

Neliömatrisiin A sanotaan olevan *symmetrinen*, jos $A^T = A$.

Esimerkki 3.5.3

Merkitään

$$B = \begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 0 \end{bmatrix}.$$

Tällöin

$$B^T = \begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 0 \end{bmatrix} = B.$$

Siis B on symmetrinen.

Transpoosioperaation käyttäytymistä matriisien laskutoimitusten kanssa valottaa seuraava lause.

Lause 3.5.4

Seuraavat säännöt pätevät matriiseille A ja B sekä reaaliluvulle t , jos laskutoimitukset on määritelty (ts. matriisit ovat sopivaa kokoa):

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(AB)^T = B^T A^T$
4. $(tA)^T = t(A^T)$.

Erityisesti kannattaa huomata tulon tekijöiden järjestyksen vaihtuminen kohdassa 3.

Todistus. Osoitetaan todeksi kohta 3 ja jätetään loput kohdat lukijalle. Olkoot $A \in \mathbb{R}^{m \times n}$ ja $B \in \mathbb{R}^{n \times p}$. Nyt sekä $(AB)^T$ että $B^T A^T$ ovat molemmat $p \times m$ -matriiseja. On osoitettava, että kyseisten matriisien alkiot ovat samoja. Olkoot $i \in \{1, 2, \dots, p\}$ ja $j \in \{1, 2, \dots, m\}$. Nähdään, että

$$\begin{aligned} (AB)^T(i, j) &= (AB)(j, i) = \sum_{k=1}^n A(j, k) \cdot B(k, i) = \sum_{k=1}^n A^T(k, j) \cdot B^T(i, k) \\ &= \sum_{k=1}^n B^T(i, k) \cdot A^T(k, j) = (B^T A^T)(i, j). \end{aligned}$$

Siten $(AB)^T = B^T A^T$. □

3.6 Käänteismatriisi

Matriiseille ei ole määritelty jakolaskua. Joissakin tapauksissa tämä puute voidaan korjata käyttämällä niin sanottuja käänteismatriiseja, jotka toimivat samalla tavalla kuin käänteisluvut tavallisten lukujen kertolaskussa. Toisin sanoen käänteismatriisilla kertominen ajaa saman asian kuin jakaminen. Pian tullaan kuitenkin valitettavasti huomaamaan, että kaikilla matriiseilla ei ole käänteismatriisia.

Esimerkiksi luvun 3 käänteisluku on $\frac{1}{3}$. Luvun ja sen käänteisluvun tulo on yksi, ja tässä tapauksessa siis $3 \cdot \frac{1}{3} = 1$. Matriisilla ja sen käänteismatriisilla on sama ominaisuus: niiden tulo on ykkösmatriisi I . Koska matriisikertolasku ei ole vaihdannainen, täytyy käänteismatriisin määritelmässä varmistaa, että vastaukseksi tulee ykkösmatriisi kerrottiin matriiseja sitten kummin päin tahansa. Siksi määritelmässä on kaksi ehtoa.

Määritelmä 3.6.1

Oletetaan, että $A \in \mathbb{R}^{n \times n}$. Jos on olemassa saman $B \in \mathbb{R}^{n \times n}$, jolle pätee

$$AB = I \quad \text{ja} \quad BA = I,$$

sanotaan, että A on *kääntyvä* ja B on matriisin A *käänteismatriisi*.

Käänteismatriisin määritelmässä rajoitutaan vain neliömatriiseihin. On itse asiassa mahdollista osoittaa, että kääntyvän matriisin ehdot eivät voi mitenkään päteä muille kuin neliömatriiseille. Kannattaa kuitenkin pitää mielessä, että edes kaikilla neliömatriiseilla ei ole käänteismatriisia.

Kääntyviä matriiseja kutsutaan myös *säännöllisiksi* matriiseiksi. Sellaisia matriiseja, joilla ei ole käänteismatriisia, voidaan kutsua *singulaarisiksi*.

Esimerkki 3.6.2

Osoitetaan, että matriisin

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

käänteismatriisi on

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

laskemalla määritelmän vaatimat kertolaskut:

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

ja

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

Koska $AB = I$ ja $BA = I$, matriisi B on matriisin A käänteismatriisi.

Esimerkki 3.6.3

Läheskään kaikilla matriiseilla ei ole käänteismatriisia. Osoitetaan, että vaikkapa matriisilla

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

ei ole käänteismatriisia. Oletetaan vastoin väitettä, että

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

on matriisin A käänteismatriisi. Nyt $AB = I$ eli

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Laskemalla yhtälön vasemmalla puolella oleva matriisitulo, saadaan

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

Nyt matriisien vasemmanpuoleisista sarakkeista nähdään, että $a = 0$ ja toisaalta $a = 1$. Päädytään siis ristiriitaan. Näin ollen matriisilla A ei ole käänteismatriisia.

Lause 3.6.4

Matriisilla on korkeintaan yksi käänteismatriisi.

Todistus. Oletetaan, että matriisilla A on käänteismatriisit B ja B' . Silloin pätee muun muassa $AB' = I$ ja $BA = I$. Saadaan pääteltyä, että

$$B = BI = B(AB') = (BA)B' = IB' = B'.$$

Yllä olevan yhtälöketjun perusteella B ja B' ovat välttämättä sama matriisi. Näin ollen matriisin A käänteismatriiseja ei voi olla enempää kuin yksi. $\square$

Jos matriisi A on kääntyvä, sen käänteismatriisille käytetään merkintää A^{-1} . Huomaa, että merkintää A^{-1} ei voi käyttää ennen kuin on varmistanut, että matriisi A todella on kääntyvä. Seuraava lause auttaa joidenkin matriisien käänteismatriisien löytämisessä.

Lause 3.6.5

Oletetaan, että matriisit A ja B ovat kääntyviä. Tällöin myös matriisit A^{-1} , AB ja A^T ovat kääntyviä, ja niiden käänteismatriisit ovat seuraavat:

1. $(A^{-1})^{-1} = A$
2. $(A^T)^{-1} = (A^{-1})^T$
3. $(AB)^{-1} = B^{-1}A^{-1}$.

Todistus. Lauseen matriisit osoitetaan kääntyviksi näyttämällä kussakin tapauksessa, että väitetty käänteismatriisi todella on matriisin käänteismatriisi.

1. Käänteismatriisin määritelmän mukaan $AA^{-1} = I$ ja $A^{-1}A = I$. Tämä tarkoittaa saman määritelmän mukaan myös sitä, että A on matriisin A^{-1} käänteismatriisi. Siispä A^{-1} on kääntyvä ja voidaan lisäksi merkitä $(A^{-1})^{-1} = A$.
2. Osoitetaan, että matriisin A^T käänteismatriisi on $(A^{-1})^T$. [Lauseen 3.5.4](#) nojalla pätee

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I.$$

Samalla tavalla osoitetaan, että $(A^{-1})^T A^T = I$. Siten $(A^{-1})^T$ on matriisin A^T käänteismatriisi. Tästä seuraa myös, että matriisi A^T on kääntyvä.

3. Matriisia AB koskevan väitteen todistaminen jätetään harjoitustehtäväksi. $\square$

2×2 -matriisin käänteismatriisi

Matriiseille, joiden koko on 2×2 , on olemassa erityinen kaava käänteismatriisin löytämiseksi.

Lause 3.6.6

Matriisi

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

on kääntyvä, jos ja vain jos $ad - bc \neq 0$. Jos matriisi A on kääntyvä, sen käänteismatriisi on

$$\frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Todistus. Oletetaan, että $ad - bc \neq 0$. Merkitään

$$B = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Laskemalla voidaan todeta, että $AB = I$ ja $BA = I$. Siten B on matriisin A käänteismatriisi ja A on kääntyvä.

Oletetaan sitten, että $ad - bc = 0$ ja osoitetaan, että matriisi A ei tällöin ole kääntyvä. Nyt on tutkittavana kaksi eri tapausta: joko $a = 0$ tai $a \neq 0$. Jos $a = 0$, niin oletuksesta seuraa, että $bc = 0$. Siten joko $b = 0$ tai $c = 0$. Tällöin

$$A = \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \quad \text{tai} \quad A = \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}.$$

Kummassakaan tapauksessa ei ole olemassa matriisia B , jolle pätee $AB = I$ ja $BA = I$. Ensimmäisessä tapauksessa tulon AB tulee nimittäin välttämättä nollarivi ja jälkimmäisessä tapauksessa tulon BA tulee välttämättä nollosarake. Siten A ei ole kääntyvä.

Tutkitaan sitten tapaus $a \neq 0$. Nyt $d = bc/a$ ja

$$A = \begin{bmatrix} a & b \\ c & bc/a \end{bmatrix}.$$

Oletetaan, että on olemassa sellainen matriisi

$$B = \begin{bmatrix} x & y \\ z & w \end{bmatrix},$$

että $AB = I$. Tällöin

$$AB = \begin{bmatrix} a & b \\ c & bc/a \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} ax + bz & ay + bw \\ cx + bcw/a & cy + bcw/a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

eli

$$\begin{cases} ax + bz = 1 \\ ay + bw = 0 \\ cx + bz/a = 0 \\ cy + bcw/a = 1. \end{cases}$$

Kolmannen yhtälön perusteella $c(x + bz/a) = 0$. Jos $c = 0$, päädytään samankaltaiseen tilanteeseen kuin silloin, kun $a = 0$. Siten voidaan olettaa, että $c \neq 0$. Tällöin täytyy päteä $x + bz/a = 0$ eli $x = -bz/a$. Toisaalta ensimmäisen yhtälön perusteella $x = (1 - bz)/a$. Nyt $-bz = 1 - bz$, joten $1 = 0$. Tämä on mahdotonta. Siten matriisilla A ei ole käänteismatriisia. $\square$

Suurempien kuin 2×2 -matriisien käänteismatriisien laskemiseksi ei helppoa kaavaa. On kuitenkin olemassa menetelmä, jolla voidaan aina selvittää, onko matriisi kääntövä. Jos matriisiin on kääntövä, voidaan menetelmän avulla lisäksi selvittää sen käänteismatriisi. Tähän menetelmään tutustutaan [myöhemmin](#).

Edellisessä lauseessa esiintynyttä lukua $ad - bc$ kutsutaan matriisin A determinantiksi. Se määrittää, onko matriisi kääntövä.

Määritelmä 3.6.7

Matriisiin

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

determinantti on $\det(A) = ad - bc$.

Matriisin determinanttia voi merkitä myös pystyviivojen avulla:

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Esimerkki 3.6.8

Matriisiin

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

determinantti on $\det(M) = 1 \cdot 4 - (-1) \cdot 2 = 4 + 2 = 6$.

Matriisille, jonka koko on 2×2 , voi käyttää determinantin laskemiseen kuvassa 3.6.1 esitettyä muistisääntöä. Piirretään matriisin poikki vinoviivat. Samalla viivalla olevat alkiot

kerrotaan keskenään. Jos viiva on lävistäjän suuntainen, tulee tulon eteen plusmerkki ja muutoin miinusmerkki. Lopuksi tulot summataan.

$$\begin{array}{c} \begin{array}{|cc|} \hline a_{11} & a_{12} \\ \hline a_{21} & a_{22} \\ \hline \end{array} \\ + \end{array}$$

Kuva 3.6.1. Muistisääntö 2×2 -matriisin determinantin laskemiseksi.

Lauseesta 3.6.6 sekä determinantin määritelmästä seuraa suoraan seuraava tulos.

Lause 3.6.9

Oletetaan, että A on 2×2 -matriisi. Matriisi A on kääntyvä, jos ja vain jos $\det(A) \neq 0$.

Determinantti voidaan määritellä myös suuremmille kuin tyypin 2×2 neliömatriiseille. Näidenkin matriisien tapauksessa determinantti kertoo, onko matriisi kääntyvä vai ei. Tähän palataan myöhemmin.

✈ Tiivistelmä

- Joillekin matriiseille löytyy käänteismatriisi.
- Käänteismatriisit vastaavat käänteislukuja: matriisin ja sen käänteismatriisin tulo on ykkösmatriisi.

3.7 Sarake- ja nolla-avaruus

Määritelmä 3.7.1

Oletetaan, että $A \in \mathbb{R}^{m \times n}$. Matriisin A *sarakeavaruus* on joukko

$$\mathcal{R}(A) = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = A\mathbf{x}, \text{ jollakin } \mathbf{x} \in \mathbb{R}^n\}.$$

Sarake-avaruutta kutsutaan myös *kuva-avaruudeksi*. Jos matriisia ajattelee kuvauksena, koostuu sarakeavaruus kaikista niistä maalijoukon vektoreista, joille kuvautuu jotain. Kyseessä on siis matriisia vastaavan kuvauksen kuvajoukko.

Esimerkki 3.7.2

Tutkitaan sarakeavaruuden käsitettä matriisin

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

avulla. Kertomalla mitä tahansa avaruuden $\mathbb{R}^4$ vektoria matriisilla A , saadaan aikai-

seksi sarakeavaruuden alkio. Koska

$$\begin{bmatrix} 1 & 2 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -8 \end{bmatrix},$$

esimerkiksi vektori $(5, -8)$ on matriisin A sarake-avaruudessa.

Esimerkki 3.7.3

Jos matriisi tulkitsee kuvauksiksi kuten [aiemmassa luvussa](#), voi sarakeavaruuden päätellä kuvan avulla. Esimerkiksi aliluvussa esitetty projektiomatriisi P kuvaa tason vektorit vaaka-akselille ja jokaiselle vaaka-akselin pisteelle projisoituu jotakin. Siten sarakeavaruus on vaaka-akseli, eli $\mathcal{R}(P) = \{(a, 0) \mid a \in \mathbb{R}\}$.

Kiertomatriisin C sarakeavaruus on puolestaan koko avaruus $\mathbb{R}^2$, sillä jokaiselle tason vektorille kuvautuu kiertossa jokin toinen vektori. Siten $\mathcal{R}(C) = \mathbb{R}^2$.

Määritelmä 3.7.4

Oletetaan, että $A \in \mathbb{R}^{m \times n}$. Matriisin A *nolla-avaruus* on joukko

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

Nolla-avaruutta kutsutaan myös *ytimeksi*. Jos matriisia ajattelee kuvauksena, koostuu nolla-avaruus kaikista niistä lähtöjoukon vektoreista, jotka kuvautuvat nollavektorille.

Esimerkki 3.7.5

Esimerkiksi vektori $\mathbf{x} = \left(-\frac{3}{2}, -\frac{1}{4}, 1, 0\right)$ on matriisin

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

nolla-avaruudessa, sillä

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{3}{2} \\ -\frac{1}{4} \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Vektori $\mathbf{v} = (1, 2, 0, 1)$ puolestaan ei ole matriisin A nolla-avaruudessa, sillä

$$A\mathbf{v} = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Esimerkki 3.7.6

Jos matriisi tulkitsee kuvauksiksi kuten [aiemmassa luvussa](#), voi nolla-avaruuden päättellä kuvan avulla. Esimerkiksi aliluvussa esitetty projektiomatriisi P kuvaa nollavektoriksi kaikki pysty akselin pisteet eli muotoa $(0, a)$ olevat vektorit, missä $a \in \mathbb{R}$. Siten $\mathcal{N}(P) = \{(0, a) \mid a \in \mathbb{R}\}$.

Kiertomatriisin C nolla-avaruudessa puolestaan ei ole mitään muuta kuin nollavektori, sillä mikään nollavektorista poikkeava vektori ei kuvaudu kierrossa nollavektorille. Siten $\mathcal{N}(C) = \{(0, 0)\}$.

Jos $m \times n$ -matriisi A esitetään sarakkeidensa $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ avulla ja $\mathbf{x} \in \mathbb{R}^n$, niin

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n.$$

Vektori $\mathbf{x}$ on matriisin A nolla-avaruudessa täsmälleen silloin, kun sen komponentit toteuttavat vektori yhtälön

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}.$$

Vektori $\mathbf{y}$ on puolestaan matriisin A sarakeavaruudessa täsmälleen silloin, kun

$$\mathbf{y} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

jollakin $\mathbf{x} \in \mathbb{R}^n$. Sen täytyy siis olla matriisin A sarakkeiden lineaarikombinaatio. Matriisin A sarakeavaruus koostuukin kaikista sarakkeiden $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ lineaarikombinaatioista.

4 LINEAARISET YHTÄLÖRYHMÄT

4.1 Ongelmia ja ratkaisuja

Tässä luvussa käsitellään lineaarisia yhtälöryhmiä. Lineaarisia yhtälöryhmiä esiintyy niin käytännön elämää koskeissa ongelmissa kuin vektoreita ja matriiseja tutkittaessa.

Pohdi 4.1.1

Kahvialan yrityksellä on kolme eri paahtimoa, joissa kussakin paahdetaan kolmenlaisia kahvipapuja. Oheisessa taulukossa on kuvattu, kukin paahtimon kahvipavuntuotto yhden päivän aikana.

	paahtimo 1	paahtimo 2	paahtimo 3
vaalea paahto	1000 kg	1000 kg	2000 kg
keskipaahto	1000 kg	2000 kg	2000 kg
tumma paahto	2000 kg	1000 kg	0 kg

Johtaja miettii, pystyykö yritys tuottamaan täsmälleen 80 000 kg vaaleaa paahtoa, 100 000 kg keskipaahtoa ja 40 000 kg tummaa paahtoa ja kuinka monta päivää tähän menee. Millaiset yhtälöt johtajan pitää muodostaa, jotta hän voi ratkaista ongelman?

Edellinen pohdintatehtävä tuottaa yhtälöryhmän. Olkoon x_1 paahtimon 1 vaatima päivien määrä, x_2 paahtimon 2 vaatima päivien määrä ja x_3 paahtimon 3 vaatima päivien määrä. Saadaan seuraavanlaiset yhtälöt:

$$1000x_1 + 1000x_2 + 2000x_3 = 80\,000$$

$$1000x_1 + 2000x_2 + 2000x_3 = 100\,000$$

$$2000x_1 + 1000x_2 + 0x_3 = 40\,000.$$

Jotta pystyttäisiin vastaamaan kysymykseen, pitäisi löytää luvut x_1 , x_2 ja x_3 , jotka toteuttavat nämä kaikki kolme yhtälöä yhtä aikaa.

Yhtälöryhmiin on jo törmätty vektorien yhteydessä. [Esimerkissä 1.2.15](#) tutkittiin, onko vektori $\mathbf{w} = (-2, 3, 2, -1)$ vektoreiden

$$\mathbf{v}_1 = (0, -1, 2, 1), \quad \mathbf{v}_2 = (2, 0, 1, -1) \quad \text{ja} \quad \mathbf{v}_3 = (4, 2, 2, 0)$$

lineaarikombinaatio. Tutkimuksessa päädyttiin yhtälöryhmään

$$\begin{cases} 2x_2 + 4x_3 = -2 \\ -x_1 + 2x_3 = 3 \\ 2x_1 + x_2 + 2x_3 = 2 \\ x_1 - x_2 = -1. \end{cases}$$

Yhtälöryhmän ratkaisusta riippui, onko vektori $\mathbf{w}$ annettujen vektoreiden lineaarikombinaatio. [Esimerkissä 2.2.10](#) puolestaan tutkittiin, mitkä pisteet kuuluvat tasojen

$$T = \{\mathbf{x} \in \mathbb{R}^3 \mid x + 2y + 3z - 9 = 0\} \text{ ja } T = \{\mathbf{x} \in \mathbb{R}^3 \mid 3x - 2y + 4z + 5 = 0\}$$

leikkaukseen. Tällön päädyttiin lineaariseen yhtälöryhmään

$$\begin{cases} x + 2y + 3z - 9 = 0 \\ 3x - 2y + 4z + 5 = 0, \end{cases}$$

jonka ratkaisujen joukko kertoo, mikä tasojen leikkaus on.

Tämäntyyppisiä tilanteita esiintyy lineaarialgebrassa jatkuvasti, ja kysymykset voivat olla hyvin monimuotoisia. Esimerkiksi lineaarikombinaation tapauksessa ei itse asiassa tarvita yhtälöryhmän varsinaista ratkaisua, vaan on ainoastaan osoitettava sen *olemassaolo*. Tasojen kohdalla puolestaan haluamme selvittää, minkälaisen joukon ratkaisut muodostavat. Joskus taas olennaista saattaa olla, onko mahdollisia ratkaisuja yksi vai useampia.

4.2 Lineaarisen yhtälöryhmän määritelmä

Muuttujien $x_1, x_2, \dots, x_n$ *lineaarinen yhtälö* on muotoa

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

missä kertoimet $a_1, \dots, a_n$ ja termi b ovat vakioita. Linearisessa yhtälössä tuntemattomia ei siis ole esimerkiksi korotettu potettiin tai kerrottu keskenään.

Lineaarinen yhtälöryhmä on yhtälöryhmä, joka on muotoa

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

missä $a_{11}, \dots, a_{mn}, b_1, \dots, b_m \in \mathbb{R}$. Symbolit $x_1, x_2, \dots, x_n$ ovat yhtälöiden *tuntematto-*

mia. Lukuja $a_{11}, \dots, a_{mn}$ nimitetään yhtälöryhmän *kertoimiksi* ja lukuja $b_1, b_2, \dots, b_m$ *vakioiksi*. Jos tuntemattomia on vähän, niitä voidaan merkitä myös vaikkapa symboleilla x, y, z ja niin edelleen. Lineaariset yhtälöryhmät on tapana järjestää niin, että kaikki tuntemattomat ovat yhtälösuuruusmerkin vasemmalla puolella ja kaikki vakiot yhtäsuuruusmerkin oikealla puolella.

Lineaarisen yhtälöryhmän ratkaiseminen merkitsee sitä, että löydetään kaikki ne luvut, jotka tuntemattomien $x_1, \dots, x_n$ paikalle sijoitettuina toteuttavat yhtä aikaa kaikki yhtälöt.

Tarkastellaan yhtälöryhmää

$$\begin{cases} 3x + 2y + z = 1 \\ -x + 2y = -1 \\ 2x + 4y + z = 0 \end{cases}$$

Kysymyksessä on niin sanottu lineaarinen yhtälöryhmä, koska yhtälöt ovat kaikki ensimmäisen asteen yhtälöitä. Yritetään ratkaista yhtälöryhmä eli löytää sellaiset luvut x, y ja z , että kaikki ryhmän yhtälöt toteutuvat yhtä aikaa.

Aloitetaan ratkaisemalla toisesta yhtälöstä x :

$$-x + 2y = -1 \iff x = 2y + 1.$$

Sijoitetaan sitten saatu x ensimmäiseen yhtälöön, ja ratkaistaan z :

$$3(2y + 1) + 2y + z = 1 \iff 6y + 3 + 2y + z = 1 \iff z = -8y - 2.$$

Sijoitetaan sitten sekä x että z kolmanteen yhtälöön, jotta voitaisiin ratkaista y :

$$2(2y + 1) + 4y - 8y - 2 = 0 \iff 4y + 2 + 4y - 8y - 2 = 0 \iff 0 = 0.$$

Päädyttiin tulokseen $0 = 0$. Miten tämä pitäisi tulkita? Onko ratkaisuja yksi vai useampia? Päteekö yhtälö ehkä kaikilla luvuilla? Selvästihän x ja z kuitenkin riippuvat y :stä, koska ne ratkaistiin yllä y :n lausekkeina. Mutta samalla tavoinhan y :n voitaisiin ajatella riippuvan x :stä ja z :sta. Vai olisiko sijoitus pitänyt tehdä jossain toisessa järjestyksessä?

Esimerkki osoittaa, että yhtälöryhmien monimutkaistuesssa tarvitaan jokin järjestelmällinen menetelmä, jota käyttämällä saadaan aina varmasti jokin vastaus ja pystytään tulkitsemaan vastauksen merkitys. Tässä luvussa esiteltävä Gaussin–Jordanin eliminointimenetelmä redusoi minkä tahansa lineaarisen yhtälöryhmän sellaiseen muotoon, että vastaus löytyy helposti.

Lineaarisen yhtälöryhmän ratkaisemisen kannalta oleellista ovat vain kertoimien ja vakioiden arvot, esimerkiksi tuntemattomien nimityksellä ei ole merkitystä. Kaikki tieto yhtälöryhmästä voidaan tiivistää lukutaulukkoon eli *matriisiin*, jossa luetellaan kaikki kertoimet sekä vakiot. Kun käsitellään yhtälöryhmien sijasta matriiseja, päästään helpom-

malla, sillä tuntemattomia ei tarvitse kirjata ylös.

Esimerkiksi yhtälöryhmän

$$\begin{cases} -4x_1 + \sqrt{3}x_2 + 2x_3 = 4 \\ x_1 + \frac{6}{8}x_3 = 0 \\ 5x_1 + \sqrt{2}x_2 + 11x_3 = -3 \\ -6x_2 - 32x_3 = 4 \end{cases}$$

matriisi on

$$\left[\begin{array}{ccc|c} -4 & \sqrt{3} & 2 & 4 \\ 1 & 0 & \frac{6}{8} & 0 \\ 5 & \sqrt{2} & 11 & -3 \\ 0 & -6 & -32 & 4 \end{array} \right].$$

Selkeyden vuoksi kertoimet on tapana erottaa vakioista pystyviivalla. Viivalla ei kuitenkaan ole matemaattista merkitystä. Huomaa, että matriisiin on kirjoitettava nolla niiden termien kohdalle, jotka puuttuvat yhtälöryhmästä. Kyseisten termien kertoimena on nimittäin nolla.

Edellä kuvattua matriisia kutsutaan lineaarisen yhtälöryhmän *kokonaismatriisiksi*.

Geometrinen tulkinta

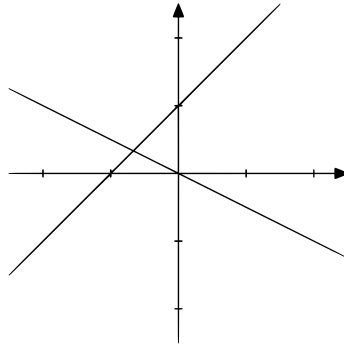
Ennen kuin ryhdytään ratkomaan lineaarisia yhtälöryhmiä algoritmin avulla, tutkitaan, miten niitä voi hahmottaa visuaalisesti. Kun yhtälöryhmässä on kaksi tai kolme tuntematonta, tilannetta voi havainnollistaa analyttisen geometrian avulla. Tutkitaan yhtälöparia

$$\begin{cases} x - y = -1 \\ x + 2y = 0. \end{cases}$$

Näitä kahta yhtälöä voidaan kumpaakin ajatella suorina. Kun yhtälöt kirjoitetaan hiukan toiseen muotoon, on suorista helpompi piirtää kuva:

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

Yhtälöitä vastaavat suorat on esitetty kuvassa 4.2.1.



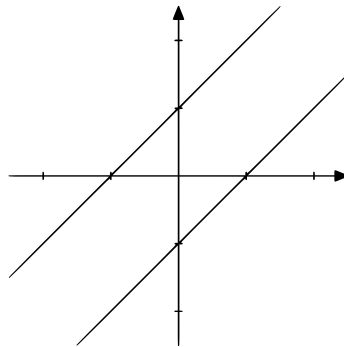
Kuva 4.2.1. Jos yhtälöitä vastaavat suorat risteävät yhdessä pisteessä, yhtälöryhmällä on täsmälleen yksi ratkaisu. Se vastaa suorien leikkauspistettä.

Yhtälön $y = x + 1$ ratkaisuja voidaan ajatella tason pisteinä. Esimerkiksi eräs ratkaisu on $x = 0, y = 1$. Tätä ratkaisua vastaa tason piste $(0, 1)$. Kaikki yhtälön ratkaisut muodostavat muodostavat suoran $y = x + 1$. Samoin toisen yhtälön ratkaisut muodostavat suoran $y = -(1/2)x$. Yhtälöparin ratkaisut toteuttavat molemmat yhtälöt, joten ratkaisut ovat pisteitä, jotka ovat sekä suoralla $y = x + 1$ että suoralla $y = -(1/2)x$. Kuvassa yhtälöryhmän ratkaisut näkyvät siis suorien leikkauskohdassa. Kuvan 4.2.1 perusteella leikkauspiste voisi olla suunnilleen $(-2/3, 1/3)$. Kun yhtälöihin sijoittaa $x = -2/3$ ja $y = 1/3$, näkee, että kyseessä on molempien yhtälöiden ratkaisu. Siten $x = -2/3, y = 1/3$ on tosiaan yhtälöparin ratkaisu.

Tutkitaan sitten yhtälöparia

$$\begin{cases} x - y = -1 \\ x - y = 1. \end{cases}$$

Sen yhtälöitä vastaavat suorat ovat yhdensuuntaisia (ks. kuva 4.2.2). Näillä suorilla ei ole yhtään leikkauspistettä. Siten yhtälöparilla ei kuvan perusteella ole ratkaisuja.

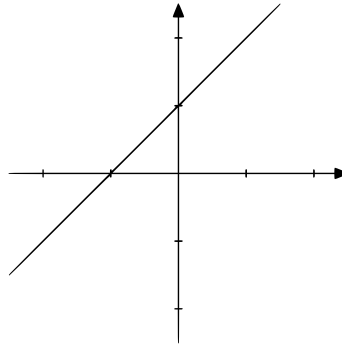


Kuva 4.2.2. Jos yhtälöitä vastaavat suorat eivät leikkaa toisiaan, ei yhtälöryhmällä ole ratkaisuja.

Yhtälöparin

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

molempia yhtälöitä vastaa puolestaan sama suora (ks. kuva 4.2.3). Näin ollen yhtälöitä vastaavat suorat leikkaavat joka ikisessä pisteessään, joten leikkauspisteitä on äärettömän monta. Yhtälöparilla on siis äärettömä monta ratkaisua.



Kuva 4.2.3. Jos yhtälöitä vastaavat suorat ovat sama suora, on niillä äärettömän monta leikkauspistettä. Yhtälöryhmällä on silloin äärettömän monta ratkaisua.

Edellisissä tapauksissa yhtälöparilla oli joko täsmälleen yksi, nolla tai äärettömän monta ratkaisua. Tulemme näkemään, minkä tahansa lineaarisen yhtälöryhmän kohdalla vain nämä kolme vaihtoehtoa ovat mahdollisia. Kun yhtälöryhmän muuttujia on kolme, yhtälöt kuvaavat tasoja. Myös tällöin ratkaisuja voi havainnollistaa kuvan avulla tutkimalla tasojen leikkauksia.

Tiivistelmä

- Lineaarinen yhtälöryhmä koostuu ensimmäisen asteen yhtälöistä
- Yhtälöryhmässä oleva informaatio voidaan kerätä matriisiin, jossa ovat yhtälöryhmän kertoimet ja vakiot.
- Kun tuntemattomia on kaksi, voi yhtälöryhmän yhtälöitä havainnollistaa suorina. Tällöin yhtälöryhmän ratkaisut ovat pisteet, joissa kaikki suorat leikkaavat.

4.3 Yhtälönratkaisun idea

1. asteen yhtälöiden ratkaisu

Ennen kuin paneudutaan yhtälöryhmien ratkaisemiseen, tutkitaan yhtälönratkaisun ideaa yksikertaisemmassa tapauksessa, jossa ratkaistavana on vain yksi yhtälö. Koulussa opitaan ratkaisemaan ensimmäisen asteen yhtälöitä. Sellainen on esimerkiksi $2x - 5 = 4$. Yhtälönratkaisulle opitaan *algoritmi* eli järjestelmällinen sarja operaatioita, joilla yhtälön voi aina ratkaista:

$$2x - 5 = 4$$

$$2x = 4 + 5$$

$$2x = 9$$

$$x = \frac{9}{2}$$

Mutta mitä yhtälönratkaisussa oikeastaan tapahtuu? Miksi oheinen algoritmi tuottaa yhtälön ratkaisun?

Kirjoitetaan yhtälön ratkaisu uudelleen tällä kertaa selittäen välivaiheet ja täsmentäen

merkintöjä:

$$\begin{aligned}
 2x - 5 &= 4 && \parallel \text{Lisätään yhtälön molemmille puolelille 5} \\
 \iff 2x &= 4 + 5 \\
 \iff 2x &= 9 && \parallel \text{Kerrrotaan yhtälön molemmat puolet luvulla } \frac{1}{2} \\
 \iff x &= \frac{9}{2}
 \end{aligned}$$

Ekvivalenssinuoli $\iff$ tarkoittaa, että peräkkäiset yhtälöt ovat *yhtäpitävät*. Ensimmäisestä yhtälöstä seuraa toinen ja toisesta ensimmäinen. Yhtälölle tehdyt operaatiot tuottavat yhtäpitäviä yhtälöitä, sillä operaatiot voi aina peruuttaa: Luvun 5 lisäämisen voi peruuttaa vähentämällä luvun 5. Luvulla $\frac{1}{2}$ kertomisen voi peruuttaa kertomalla luvulla 2.

Yhtäpitävillä yhtälöillä on täsmälleen samat ratkaisut. Yhtälönratkaisualgoritmin ideana on, että yhtälöä muokataan aina vain yksinkertaisempaan muotoon tuottaen yhtäpitäviä yhtälöitä. Uuden yhtälön ratkaisut ovat aina samat kuin edellisen. Alimmalta riviltä ratkaisu on helppo nähdä, sillä se lukee siinä suoraan: $x = \frac{9}{2}$. Tähän samaan ideaan perustuu myös lineaaristen yhtälöryhmien ratkaiseminen.

Lineaaristen yhtälöryhmien ratkaisemisen idea

Seuraavaksi tutustutaan Gaussin–Jordanin menetelmään, jolla voidaan ratkaista mikä tahansa lineaarinen yhtälöryhmä. Periaate on sama kuin 1. asteen yhtälöiden ratkaisussa. Ideana on muokata yhtälöryhmästä uusia yhtälöryhmiä, joilla on samat ratkaisut kuin alkuperäisellä yhtälöryhmällä. Viimeisenä saatu yhtälöryhmä on sellaisessa muodossa, josta sen ratkaisuja koskeviin kysymyksiin on helppo vastata. Koska viimeisen yhtälöryhmän ratkaisut ovat samat kuin alkuperäisen yhtälöryhmän, myös alkuperäisen yhtälöryhmän ratkaisut ja niiden luonne tunnetaan.

Määritelmä 4.3.1

Yhtälöryhmiä kutsutaan *yhtäpitäviksi* eli *ekvivalenteiksi*, jos niillä on täsmälleen samat ratkaisut.

Yhtälöryhmä ratkaistaan muokkaamalla siitä uusia yhtälöryhmiä, jotka ovat ekvivalentteja alkuperäisen yhtälöryhmän kanssa.

Yhtälöryhmiä muokataan niin kutsutuilla alkeisrivimuunnoksilla. Koska matriisien käsitteleminen on helpompaa kuin yhtälöryhmien, tehdään alkeisrivimuunnokset suoraan matriiseille.

Määritelmä 4.3.2

Seuraavat kolme operaatiota ovat *alkeisrivimuunnoksia*:

1. Vaihdetaan kahden rivin paikka matriisissa.
2. Kerrotaan jokin rivi nolasta poikkeavalla reaaliluvulla.
3. Lisätään johonkin riviin jokin toinen rivi reaaliluvulla kerrottuna.

Alkeisrivimuunnoksille käytetään tässä materiaalissa seuraavia lyhennysmerkintöjä.

- $R_i \leftrightarrow R_j$: vaihdetaan rivien i ja j paikat ($i \neq j$).
- aR_i : kerrotaan rivi i luvulla $a \neq 0$.
- $R_i + bR_j$: lisätään riviin i rivi j luvulla b kerrottuna ($i \neq j$).

Esimerkki 4.3.3

Seuraavassa on annettu esimerkit erilaisista alkeisrivimuunnoksista:

$$\begin{bmatrix} -4 & 3 & 4 \\ 1 & 2 & -1 \\ 5 & 3 & 2 \\ 0 & 6 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 5 & 3 & 2 \\ 0 & 6 & 4 \end{bmatrix} \xrightarrow{R_3 - 5R_1} \begin{bmatrix} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 0 & -7 & 7 \\ 0 & 6 & 4 \end{bmatrix} \xrightarrow{-\frac{1}{7}R_3} \begin{bmatrix} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 0 & 1 & -1 \\ 0 & 6 & 4 \end{bmatrix}.$$

Lineaaristen yhtälöryhmien ratkaisu perustuu seuraavaan tulokseen.

Lause 4.3.4

Alkeisrivimuunnoksien tekeminen ei muuta yhtälöryhmän ratkaisuja. Toisin sanoen alkeisrivimuunnoksen tekeminen tuottaa yhtälöryhmän, joka on yhtäpitävä alkuperäisen yhtälöryhmän kanssa.

Todistus. Täsmällinen todistus on melko tekninen, joten sitä ei esitetä tässä. Tutkitaan kuitenkin todistuksen ideaa. Se perustuu siihen, että alkeisrivimuunnokset ovat kääntyviä eli tehdyn alkeisrivimuunnoksen voi aina peruuttaa. alkeisrivimuunnoksen $R_i \leftrightarrow R_j$ voi peruuttaa tekemällä sen uudelleen. alkeisrivimuunnoksen aR_i voi peruuttaa alkeisrivimuunnoksella $\frac{1}{a}R_i$ ja alkeisrivimuunnoksen $R_i + bR_j$ voi peruuttaa alkeisrivimuunnoksella $R_i - bR_j$.

Jos lineaarisesta yhtälöryhmästä muodostetaan alkeisrivimuunnoksen avulla uusi yhtälöryhmä, pätee uusi yhtälöryhmä aina silloin, kun alkuperäinenkin yhtälöryhmä pätee. Siten kaikki alkuperäisen yhtälöryhmän ratkaisut ovat myös uuden yhtälöryhmän ratkaisuja. Toisaalta uudesta yhtälöryhmästä voidaan aina peruuttaa alkuperäiseen yhtälöryhmään jollakin alkeisrivimuunnoksella. Siten kaikki uuden yhtälöryhmän ratkaisut ovat myös alkuperäisen yhtälöryhmän ratkaisuja. Näin ollen molemmilla yhtälöryhmillä on samat ratkaisut. $\square$

$$\begin{array}{ccc}
 \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right] & \xrightarrow[\text{toimituksia}]{\text{alkeisrivi-}} & \left[\begin{array}{cccc|c} c_{11} & c_{12} & \dots & c_{1n} & d_1 \\ c_{21} & c_{22} & \dots & c_{2n} & d_2 \\ & & & \vdots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mn} & d_m \end{array} \right] \\
 \left\{ \begin{array}{l} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{array} \right. & \xleftarrow[\text{ratkaisut}]{\text{samat}} \rightarrow & \left\{ \begin{array}{l} c_{11}x_1 + \dots + c_{1n}x_n = d_1 \\ c_{21}x_1 + \dots + c_{2n}x_n = d_2 \\ \vdots \\ c_{m1}x_1 + \dots + c_{mn}x_n = d_m \end{array} \right.
 \end{array}$$

Kuva 4.3.1. Gaussin–Jordanin eliminointimenetelmän perusta.

Määritelmä 4.3.5

Matriisi A on *riviekvivalentti* matriisiin B kanssa, jos B saadaan matriisista A alkeisrivimuunnoksilla.

Esimerkki 4.3.6

Esimerkissä 4.3.3 matriisista

$$\begin{bmatrix} -4 & 3 & 4 \\ 1 & 2 & -1 \\ 5 & 3 & 2 \\ 0 & 6 & 4 \end{bmatrix}$$

muokattiin alkeisrivimuunnoksella $R_1 \leftrightarrow R_2$ matriisi

$$\begin{bmatrix} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 5 & 3 & 2 \\ 0 & 6 & 4 \end{bmatrix}.$$

Nämä kaksi matriisia ovat riviekvivalentteja. Kaikki muutkin esimerkissä esiintyvät matriisit ovat riviekvivalentteja keskenään. alkeisrivimuunnoksia voidaan ajatella tehtävän myös nolla kappaletta. Siten jokainen matriisi on itsensä kanssa riviekvivalentti.

Nyt lause 4.3.4 voidaan muotoilla toisin sanoin: Jos yhtälöryhmiä vastaavat matriisit ovat riviekvivalentit, yhtälöryhmät ovat yhtäpitävät.

Porrasmatriisit ja redusoidut porrasmatriisit

Yhtälöryhmää ratkaistaessa on tavoitteena muuttaa yhtälöryhmän matriisi alkeisrivimuunnoksilla niin kutsutuksi redusoiduksi porrasmatriisiksi, josta ratkaisut on helppo lukea. Määritellään ensin porrasmatriisi.

Määritelmä 4.3.7

Matriisi on *porrasmatriisi*, jos seuraavat ehdot toteutuvat:

1. mahdolliset nollarivit ovat alimpina
2. kullakin rivillä ensimmäinen nollasta poikkeava alkio, ns. *johtava alkio*, on ylemmän rivin johtavan alkion oikealla puolella.

Seuraavat matriisit ovat porrasmatriiseja. Niiden johtavat alkiot on lihavoitu.

$$\begin{bmatrix} \mathbf{14} & 3 & 2 & 4 \\ 0 & 0 & \mathbf{8} & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & \mathbf{4} & 0 & 0 & -3 \\ 0 & 0 & 0 & -\mathbf{1} & 7 & -11 \\ 0 & 0 & 0 & 0 & \mathbf{1} & -3 \end{bmatrix} \quad \begin{bmatrix} -\mathbf{3} & -41 & 1 & 0 & -3 & 6 \\ 0 & 0 & 0 & 0 & \mathbf{5} & -11 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Porrasmuoto auttaa jo yhtälöryhmän ratkaisemisessa, mutta se ei ole yksikäsitteinen. Kutakin matriisia kohden löytyy nimittäin useampi kuin yksi sen kanssa rivekvivalentti porrasmatriisi. Porrasmatriisi voidaan kuitenkin muokata alkeisrivimuunnosten avulla redusoituun muotoon, joka on kullekin matriisille yksikäsitteinen.

Määritelmä 4.3.8

Matriisi on *redusoitu porrasmatriisi*, jos seuraavat ehdot toteutuvat:

1. matriisi on porrasmatriisi
2. jokaisen rivin johtava alkio on 1 (johtava ykkönen)
3. jokainen johtava alkio on sarakkeensa ainoa nollasta poikkeava alkio.

Redusoituja porrasmatriiseja kutsutaan myös pidemmällä nimellä *redusoitu riviporrasmatriisi*.

Esimerkiksi seuraavat matriisit ovat redusoituja porrasmatriiseja. Johtavat alkiot on jälleen lihavoitu.

$$\begin{bmatrix} \mathbf{1} & 0 & 0 & 4 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & \mathbf{1} & -53 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & -11 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & -3 \end{bmatrix} \quad \begin{bmatrix} \mathbf{1} & 3 & 0 & 0 & -3 & 8 \\ 0 & 0 & \mathbf{1} & -3 & 5 & -11 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Seuraavat matriisit puolestaan eivät ole redusoituja porrasmatriiseja:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{ja} \quad \begin{bmatrix} 1 & 2 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -2 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Niissä on nolasta poikkeavia alkioita johtavien ykkösten sarakkeissa.

Esimerkki 4.3.9

Matriisi

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

on redusoitu porrasmatriisi. Sitä vastaava yhtälöryhmä on

$$\begin{cases} x_1 = 4 \\ x_2 = -2 \\ x_3 = 3 \end{cases}$$

Huomataan, että matriisista näkyy suoraan yhtälöryhmän ratkaisu.

Tiivistelmä

- Alkeisrivimuunnoksia tehdessä matriisi muuttuu, mutta matriisia vastaavan yhtälöryhmän ratkaisut eivät muutu.
- Yhtälöryhmää ratkaistaessa sen matriisi muutetaan alkeisrivitoimituksilla ensin porrasmatriisiksi ja sitten redusoiduksi porrasmatriisiksi.
- Redusoidusta porrasmatriisista ratkaisu on helppo lukea.

4.4 Gaussin–Jordanin eliminointimenetelmä

Tavoitteena on muuttaa yhtälöryhmän matriisi alkeisrivimuunnosten avulla redusoiduksi porrasmatriisiksi, josta ratkaisut näkyvät suoraan. Voidaan osoittaa, että mikä tahansa matriisi voidaan muuttaa redusoiduksi porrasmatriisiksi ja että alkeisrivimuunnosten käyttämisjärjestys ei vaikuta tulokseen. Seuraava esimerkki näyttää, kuinka tämä tehdään.

Esimerkki 4.4.1

Muutetaan matriisi

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

reduoiduksi porrasmatriisiksi. Aloitetaan ensimmäisestä sarakkeesta. Vaihtamalla ensimmäisen ja toisen rivin paikat, saadaan ensimmäisen rivin johtavaksi alkioksi 1:

$$\xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 3 & 2 \\ -1 & -2 & 0 & 3 \end{bmatrix}.$$

Tämän jälkeen johtavan alkion alla olevat alkio on helppo muuttaa nolliksi. Vähennetään ensin toisesta rivistä ensimmäinen rivi luvulla 2 kerrottuna:

$$\xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & -1 & -1 & 0 \\ -1 & -2 & 0 & 3 \end{bmatrix}.$$

Lisätään sitten kolmanteen riviin ensimmäinen rivi luvulla 1 kerrottuna:

$$\xrightarrow{R_3 + R_1} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -2 & 2 & 4 \end{bmatrix}.$$

Nyt ensimmäinen sarake on halutussa muodossa. Siirrytään muokkaamaan toista saraketta. Muutetaan ensin sen johtava alkio ykköseksi, jotta voidaan toimia samoin kuin edellä. Kerrotaan siis toinen rivi luvulla -1 . Saadaan matriisi

$$\xrightarrow{(-1)R_2} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 2 & 4 \end{bmatrix}.$$

Toisen rivin johtavan alkion avulla voidaan muuttaa sen alla oleva alkio nolllaksi. Lisätään kolmanteen riviin toinen rivi luvulla 2 kerrottuna. Saadaan matriisi

$$\xrightarrow{R_3 + 2R_2} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 4 & 4 \end{bmatrix},$$

joka on porrasmatriisi.

Jatketaan muokkaamista niin, että saadaan aikaan redusoitu porrasmatriisi. Muutetaan ensin viimeinenkin johtava alkio ykköseksi:

$$\xrightarrow{\frac{1}{4}R_3} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Muutetaan alimman rivin johtavan alkion avulla kaikki kolmannen sarakkeen muut alkiot nolliksi:

$$\xrightarrow{R_2-R_3} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{R_1-2R_3} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Näin saatu matriisi on redusoitu porrasmatriisi.

Saatu redusoitu porrasmatriisi on eri matriisi kuin se, josta lähdettiin liikkeelle. Matriisit myös vastaavat erilaisia yhtälöryhmiä. Näillä yhtälöryhmillä on kuitenkin samat ratkaisut [lauseen 4.3.4](#) nojalla.

Redusoituu porrasmatriisiin voi päästä monia eri reittejä. On kuitenkin olemassa algoritmi, jonka avulla redusoidun porrasmatriisin voi tuottaa. Varsinkin alussa kannattaa käyttää sitä, sillä sattumanvarainen alkeisrivimuunnosten pyörittäminen menee helposti sotkuiseksi. Esimerkeissä on käytetty tätä algoritmia.

Ohje 4.4.2

Redusoidun porrasmatriisin muodostaminen:

- Porrasmatriisia muodostetaan vasemmalta oikealle ja ylhäältä alaspäin.
- Johtavat alkiot kannattaa useimmiten muuttaa ykkösiksi. Aloita hankkimalla 1. rivin johtavaksi alkioiksi ykkönen.
- Johtavan alkion avulla muutetaan sen alapuolella olevat alkiot nolliksi.
- Siirry sitten 2. riville. Hanki 2. rivin johtavaksi alkioiksi ykkönen ja muuta sen avulla alapuolella olevat alkiot nolliksi. Jatka sitten eteenpäin aina alimmalle riville asti.
- Näin saadaan aikaan porrasmatriisi.
- Redusoitua porrasmatriisia muodostetaan oikealta vasemmalle ja alhaalta ylöspäin.
- Johtavien alkoiden avulla muutetaan niiden yläpuolella olevat alkiot nolliksi. Aloita alimmasta nolasta poikkeavasta rivistä. Muuta johtavan alkion avulla sen

yläpuolella olevat alkiot nolliksi. Siirry sitten toiseksi alimmalle riville. Jatka eteepäin aina ylimmälle riville asti.

- Tee vain yksi alkeisrivimuunnos kerrallaan. Tällöin vältät todennäköisemmin virheet.

Nyt olemme valmiita ratkaisemaan yhtälöryhmiä. Seuraavissa esimerkeissä esiintyy erilaisia tapauksia, joihin yhtälöryhmän ratkaisussa voi päätyä.

Esimerkki 4.4.3

Ratkaistaan yhtälöryhmä

$$\begin{cases} 2x_1 - x_2 + 3x_3 = 2 \\ x_1 \quad \quad + 2x_3 = 1 \\ -x_1 - 2x_2 \quad \quad = 3. \end{cases}$$

Yhtälöryhmän matriisi on

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

Tämä matriisi muutettiin redusoiduksi porrasmatriisiksi [esimerkissä 4.4.1](#):

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

Redusoitua porrasmatriisia vastaava yhtälöryhmä on

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

Koska alkuperäisen yhtälöryhmän ratkaisut ovat [lauseen 4.3.4](#) nojalla samat kuin lopuksi saadun yhtälöryhmän, on yhtälöryhmä ratkaistu. Sen ratkaisu on siis

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

Esimerkki 4.4.4

Ratkaistaan lineaarinen yhtälöryhmä

$$\begin{cases} x + 2y + z = 8 \\ -3x - 6y - 3z = -21 \end{cases}$$

Muutetaan yhtälöryhmän matriisi redusoiduksi porrasmatriisiksi:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ -3 & -6 & -3 & -21 \end{array} \right] \xrightarrow{R_2+3R_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 0 & 0 & 0 & 3 \end{array} \right].$$

Vastaava yhtälöryhmä on

$$\begin{cases} x + 2y + z = 8 \\ 0 = 3. \end{cases}$$

Alin yhtälö on aina epätosi, joten yhtälöryhmällä ei ole ratkaisuja.

Esimerkki 4.4.5

Ratkaistaan lineaarinen yhtälöryhmä

$$\begin{cases} 3x_1 + 3x_2 - 15x_3 = 9 \\ x_1 - 2x_3 = 1 \\ 2x_1 - x_2 - x_3 = 0 \end{cases}$$

Muutetaan yhtälöryhmän matriisi redusoiduksi porrasmatriisiksi:

$$\begin{aligned} & \left[\begin{array}{ccc|c} 3 & 3 & -15 & 9 \\ 1 & 0 & -2 & 1 \\ 2 & -1 & -1 & 0 \end{array} \right] \xrightarrow{\frac{1}{3}R_1} \left[\begin{array}{ccc|c} 1 & 1 & -5 & 3 \\ 1 & 0 & -2 & 1 \\ 2 & -1 & -1 & 0 \end{array} \right] \xrightarrow{R_2-R_1} \left[\begin{array}{ccc|c} 1 & 1 & -5 & 3 \\ 0 & -1 & 3 & -2 \\ 2 & -1 & -1 & 0 \end{array} \right] \\ & \xrightarrow{R_3-2R_1} \left[\begin{array}{ccc|c} 1 & 1 & -5 & 3 \\ 0 & -1 & 3 & -2 \\ 0 & -3 & 9 & -6 \end{array} \right] \xrightarrow{-1 \cdot R_2} \left[\begin{array}{ccc|c} 1 & 1 & -5 & 3 \\ 0 & 1 & -3 & 2 \\ 0 & -3 & 9 & -6 \end{array} \right] \xrightarrow{R_3+3R_2} \left[\begin{array}{ccc|c} 1 & 1 & -5 & 3 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right] \\ & \xrightarrow{R_1-R_2} \left[\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right]. \end{aligned}$$

Saatua matriisia vastaa yhtälöryhmä

$$\begin{cases} x_1 - 2x_3 = 1 \\ x_2 - 3x_3 = 2 \\ 0 = 0 \end{cases}$$

Alin yhtälö $0 = 0$ on aina tosi. Se ei siis anna ratkaisujen kannalta mitään informaatiota ja voidaan unohtaa. Tuntemattomat x_1 ja x_2 riippuvat tuntemattomasta x_3 . Tuntemattomalle x_3 ei puolestaan aseteta mitään rajoitteita, joten se voi olla mikä tahansa reaaliluku. Sanotaan, että x_3 on *vapaa muuttuja*. Merkitään $x_3 = t$, missä $t \in \mathbb{R}$.

Ratkaistaan vielä muut tuntemattomat. Ensimmäinen yhtälö saa nyt muodon $x_1 - 2t = 1$, joten $x_1 = 1 + 2t$. Toinen yhtälö puolestaan on $x_2 - 3t = 2$ eli $x_2 = 2 + 3t$. Siten yhtälöryhmän ratkaisu on

$$\begin{cases} x_1 = 1 + 2t \\ x_2 = 2 + 3t \\ x_3 = t, \end{cases} \quad \text{missä } t \in \mathbb{R}.$$

Ratkaisuja on siis äärettömän monta. Yksittäisiä ratkaisuja saadaan antamalla parametrille t eri arvoja. Esimerkiksi sijoittamalla $t = 1$ saadaan yhdeksi ratkaisuksi $x_1 = 3$, $x_2 = 5$ ja $x_3 = 1$. Sijoittamalla $t = -1$ saadaan toinen ratkaisu $x_1 = -1$, $x_2 = -1$ ja $x_3 = -1$. Jokaisella reaaliluvulla t yhtälöryhmälle saadaan eri ratkaisu.

Edellisessä esimerkissä yhtälöryhmässä oli vapaa muuttuja, jolloin yhtälöryhmällä oli äärettömän monta ratkaisua. Yhtälöryhmässä saattaa olla useitakin vapaita muuttujia. Nämä löytyvät redusoidussa porrasmatriisissa niistä sarakkeista, joissa ei ole lainkaan johtavaa alkioita.

Esimerkki 4.4.6

Lineaarisen yhtälöryhmän matriisi muutettiin alkeisrivimuunnoksilla redusoiduksi porrasmatriisiksi

$$\left[\begin{array}{cccccc|c} 1 & 3 & 0 & 4 & 0 & 0 & 7 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 \end{array} \right].$$

Mikä on yhtälöryhmän ratkaisu?

Johtavat alkioit ovat sarakkeissa 1, 3 ja 6. Muita sarakkeita vastaavat tuntemattomat x_2 , x_4 ja x_5 ovat vapaita muuttujia. Merkitään $x_2 = r$, $x_4 = s$ ja $x_5 = t$, missä $r, s, t \in \mathbb{R}$.

Nyt voidaan kirjoittaa

$$\begin{cases} x_1 + 3r + 4s = 7 \\ x_3 + 2s = 0 \\ x_6 = 3. \end{cases}$$

Tämä yhtälöryhmä on yhtäpitävä yhtälöryhmän

$$\begin{cases} x_1 = 7 - 3r - 4s \\ x_3 = -2s \\ x_6 = 3 \end{cases}$$

kanssa. Yhtälöryhmän ratkaisu on siis

$$\begin{cases} x_1 = 7 - 3r - 4s \\ x_2 = r \\ x_3 = -2s \\ x_4 = s \\ x_5 = t \\ x_6 = 3, \end{cases} \quad \text{missä } r, s, t \in \mathbb{R}.$$

Luvut r , s ja t voidaan valita täysin vapaasti, ja jokainen valinta tuottaa yhtälöryhmän erään ratkaisun.

Kootaan vielä yhteen Gaussin–Jordanin menetelmän vaiheet.

Ohje 4.4.7

Gaussin–Jordanin menetelmä:

1. Kirjoita yhtälöryhmän matriisi.
2. Muuta matriisi alkeisrivimuunnoksilla ensin porrasmatriisiksi ja sitten redusoiduksi porrasmatriisiksi ohjeen 4.4.2 mukaisesti.
3. Lue ratkaisut redusoidusta porrasmatriisista.
 - (a) Jos matriisin alimmalla nollasta poikkeavalla rivillä on ristiriitainen yhtälö, yhtälöryhmällä ei ole ratkaisuja.
 - (b) Jos epätotta yhtälöä ei ole ja jokaisessa viivan vasemmalla puolella olevassa sarakkeessa on johtava alkio, ratkaisuja on täsmälleen yksi.
 - (c) Jos epätotta yhtälöä ei ole ja jostain sarakkeesta puuttuu johtava alkio, ratkaisuja on äärettömän monta. Jokaista saraketta, josta johtava alkio puuttuu, vastaa *vapaa muuttuja*. Vapaan muuttujan arvo voidaan valita vapaasti.

Gauss ei itse kehittänyt nimeään kantavaa menetelmää, vaan sen tunsi jo ainakin Newton

sata vuotta aikaisemmin 1600-luvun loppupuolella. Kiinalaiset puolestaan tunsivat menetelmän jo toisella vuosisadalla eKr. Nimitys ”Gaussin eliminointimenetelmä” tuli käyttöön kuitenkin vasta 1950-luvulla. Tällä nimityksellä tarkoitetaan yleensä nimenomaan porrasmatriisiin tähtäävää menetelmää, ja mikäli halutaan jatkaa redusoituun porrasmatriisiin asti, menetelmää kutsutaan ”Gaussin–Jordanin eliminoinniksi”. Jordan esitti tämän version eliminointimenetelmästä vuonna 1887.

Puetaan vielä täsmällisten lauseiden muotoon Gaussin–Jordanin muutamia menetelmään liittyviä tuloksia. Niiden todistuksia ei kuitenkaan esitetä tässä teknisyytensä vuoksi. Ensinnäkin, jotta Gaussin–Jordanin menetelmä toisimi, täytyy tietää, että redusoidun porrasmatriisin muodostaminen on aina mahdollista.

Lause 4.4.8

Oletetaan, että $A \in \mathbb{R}^{m \times n}$. Tällöin on olemassa yksikäsitteinen matriisi $B \in \mathbb{R}^{m \times n}$, jolle pätevät seuraavat ehdot:

1. A ja B ovat riviekvivalentteja,
2. B on redusoitu porrasmatriisi.

Edellisen lauseen matriisia B kutsutaan matriisiin A *redusoiduksi porrasmuodoksi* tai *redusoiduksi riviporrasmuodoksi*. Sille käytetään merkintää $\text{rref}(A)$. Merkintä tulee englannin kielen ilmaisusta ”reduced row echelon form”.

Yhtälöryhmän ratkaisut luetaan redusoidusta porrasmatriisista. Kuten [ohjeessa 4.4.7](#) todetaan, ratkaisujen lukumäärän suhteen voidaan päätyä vain kolmeen erilaiseen tulokseen.

Lause 4.4.9

Lineaarisella yhtälöryhmällä on joko täsmälleen yksi, ei yhtään tai äärettömän monta ratkaisua.

Tarkastellaan vielä lopuksi hiukan erilaista tapaa kirjoittaa yhtälöryhmän ratkaisu. Lineaarisen yhtälöryhmän ratkaisuja voi ajatella listana lukuja. Ne voi siis kirjoittaa vektorin muodossa.

Esimerkki 4.4.10

[Esimerkissä 4.4.3](#) saatiin yhtälöryhmälle ratkaisu

$$\begin{cases} x_1 = -1 \\ x_2 = -1 \\ x_3 = 1. \end{cases}$$

Ratkaisun voi ilmaista myös muodossa $\mathbf{x} = (-1, -1, 1)$.

Esimerkissä 4.4.5 puolestaan saatiin yhtälöryhmälle ratkaisu

$$\begin{cases} x_1 = 7 - 3r - 4s \\ x_2 = r \\ x_3 = -2s \\ x_4 = s \\ x_5 = t \\ x_6 = 3, \end{cases} \quad \text{missä } r, s, t \in \mathbb{R}.$$

Ratkaisun voi ilmaista myös muodossa

$$\mathbf{x} = \begin{bmatrix} 7 - 3r - 4s \\ r \\ -2s \\ s \\ t \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3 \end{bmatrix} + r \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Määritelmä 4.4.11

Matriisin aste on sen redusoidussa porrasmuodossa esiintyvien johtavien alkioiden lukumäärä. Matriisin A astetta merkitään $\text{rank}(A)$.

Esimerkissä 4.4.1 nähtiin, että matriisin

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

redusoitu porrasmuoto on

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

Koska johtavia alkioita on kolme, on matriisin aste 3.

Porrasmatriisien tulkinta

Usein kiinnostavaa ei ole yhtälöryhmän tarkka ratkaiseminen, vaan se, *kuinka monta* ratkaisua yhtälöryhmällä on. Tällöin ei tarvitse muuttaa yhtälöryhmän matriisia redusoituun

porrasmuotoon, vaan pelkkä porrasmuoto riittää. Kaikkia tässä osiossa esitettyjä väitteitä ei perustella tarkasti, sillä pitkien ja teknisten todistusten kirjoittaminen ei olisi mielekästä.

Kerrataan vielä, miten ratkaisujen lukumäärän voi lukea redusoidusta porrasmatriisista. Olkoon M redusoitu porrasmatriisi.

1. Jos jokin matriisin M viimeisistä riveistä on muotoa $\begin{bmatrix} 0 & \dots & 0 & | & 1 \end{bmatrix}$ (eli rivin johtava ykkönen on pystyviivan oikealla puolella), kyseistä riviä vastaa epätosi yhtälö $0 = 1$. Yhtälöryhmällä ei ole ratkaisua.
2. Oletetaan, että edellinen tapaus ei toteudu. Jos joltakin matriisin M sarakkeelta puuttuu johtava alkio (pystyviivan vasemmalta puolelta), tuota saraketta vastaava muuttuja on vapaa. Yhtälöryhmän ratkaisut voidaan esittää vapaiden muuttujien avulla. Kunkin vapaan muuttujan arvo voidaan valita vapaasti, joten yhtälöryhmällä on ratkaisuja ääretön määrä.
3. Oletetaan, että edelliset tapaukset eivät toteudu. Tällöin matriisin M jokaisessa sarakkeessa pystyviivan vasemmalla puolella on johtava alkio, ja yhtälöryhmällä on täsmälleen yksi ratkaisu.

Esimerkki 4.4.12

Oletetaan, että yhtälöryhmän matriisi on saatu alkeisrivimuunnoksilla muotoon

$$\left[\begin{array}{ccc|c} 4 & -3 & 1 & 0 \\ 0 & -1 & 2 & 2 \\ 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Toiseksi viimeinen rivi vastaa yhtälöä $0 = -4$, joka on aina epätosi. Nyt tiedetään, että alkuperäisellä yhtälöryhmällä ei ole ratkaisuja, eikä redusoiduksi porrasmatriisiksi muuttaminen ole tarpeellista.

Tutkitaan sitten yhtälöryhmää, jonka matriisi on saatu alkeisrivimuunnoksilla porrasmatriisiksi

$$\left[\begin{array}{cccc|c} 4 & -3 & 0 & 1 & 0 \\ 0 & 6 & 2 & 3 & 2 \\ 0 & 0 & 0 & -1 & -4 \end{array} \right].$$

Tässäkin tapauksessa ratkaisujen lukumäärä nähdään suoraan. Koska porrasmatriisissa ei näy yhtälöä, joka olisi epätosi, ei sellaista tule redusoituun porrasmatriisiinkaan. Siten voidaan sanoa suoraan porrasmatriisin perusteella, että yhtälöryhmällä on ratkaisuja. Huomataan vielä, että porrasmatriisin kolmannessa sarakkeessa ei ole johtavaa alkioita. Jos matriisi muutettaisiin redusoiduksi porrasmatriisiksi, ei siinäkään olisi johtavaa alkioita kolmannessa sarakkeessa. Siten yhtälöryhmällä on äärettömän

monta ratkaisua, ja sen näkee suoraan porrasmatriisista.

Tarkastellaan vielä lopuksi yhtälöryhmää, jonka matriisi on saatu alkeisrivimuunnoksilla porrasmatriisiksi

$$\left[\begin{array}{ccc|c} 4 & -3 & 1 & 0 \\ 0 & 6 & 2 & 2 \\ 0 & 0 & 2 & -4 \end{array} \right].$$

Epätosia yhtälöitä ei ole, joten yhtälöryhmällä on ratkaisuja. Koska jokaisessa sarakkeessa viivan vasemmalla puolella on johtava alkio, voidaan päätellä, että vapaita muuttujia ei ole. Siten yhtälöryhmällä on täsmälleen yksi ratkaisu.

Esimerkki 4.4.13

Tarkastellaan yhtälöryhmää

$$\begin{cases} x + y + kz = 1 \\ x + ky + z = 1 \\ kx + y + z = -2. \end{cases}$$

Tutkitaan, miten luvun k arvot vaikuttavat ratkaisujen lukumäärään. Ryhdytään muuttamaan yhtälöryhmän matriisia redusoiduksi porrasmatriisiksi:

$$\begin{aligned} & \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 1 & k & 1 & 1 \\ k & 1 & 1 & -2 \end{array} \right] \xrightarrow{R_2-R_1} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & k-1 & 1-k & 0 \\ k & 1 & 1 & -2 \end{array} \right] \\ & \xrightarrow{R_3-kR_1} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & k-1 & 1-k & 0 \\ 0 & 1-k & 1-k^2 & -2-k \end{array} \right] \xrightarrow{R_3+R_2} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & k-1 & 1-k & 0 \\ 0 & 0 & 2-k-k^2 & -2-k \end{array} \right] \end{aligned}$$

Kaikki alkeisrivimuunnokset voidaan tähän asti tehdä riippumatta siitä, mikä luku k on. Jatkaminen ei kuitenkaan onnistu, sillä toisen rivin alkio $k-1$ saattaa olla nolla, samoin kolmannen rivin alkio $2-k-k^2$. Tarkastellaan näitä tapauksia erikseen.

Oletetaan ensin, että kolmannen rivin alkio $2-k-k^2=0$ eli $k=-2$ tai $k=1$.

- Jos $k = -2$, viimeinen matriisi on

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

Tämä matriisi on porrasmuodossa, joten siitä voidaan päätellä yhtälöryhmän ratkaisujen lukumäärä. Koska epätosia yhtälöitä ei ole, ratkaisuja on olemassa. Tuntematonta x_3 vastaavassa sarakkeessa ei ole johtavaa alkioita, joten x_3 on vapaa muuttuja. Ratkaisuja on siten äärettömän monta.

- Jos $k = 1$, viimeinen matriisi on

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

Havaitaan, että alinta riviä vastaava yhtälö $0 = -3$ on aina epätosi. Siten yhtälöryhmällä ei ole ratkaisuja.

Oletetaan sitten, että toisen rivin alkio $k - 1 = 0$ eli $k = 1$. Tämä tapaus käsiteltiin sattumalta jo edellä.

Tarkastellaan vielä lopuksi tilannetta, jossa sekä toisen rivin alkio $k - 1$ että kolmannen rivin alkio $2 - k - k^2$ ovat nolasta poikkeavia. Tällöin voidaan jatkaa alkeisrivimuunnosten tekemistä. Koska $k - 1 \neq 0$ ja $2 - k - k^2 \neq 0$ saadaan

$$\xrightarrow{\frac{1}{k-1}R_2} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 2 - k - k^2 & -2 - k \end{array} \right] \xrightarrow{\frac{1}{2-k-k^2}R_3} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & \frac{-2-k}{2-k-k^2} \end{array} \right].$$

Saatu matriisi on porrasmuodossa. Koska jokaisessa pystyviivan vasemmalla puolella olevassa sarakkeessa on johtava alkio, yhtälöryhmällä on täsmälleen yksi ratkaisu.

Päädyttiin siis seuraavaan tulokseen. Yhtälöryhmällä on äärettömän monta ratkaisua, jos ja vain jos $k = -2$. Yhtälöryhmällä ei ole ratkaisua, jos ja vain jos $k = 1$. Yhtälöryhmällä on tasan yksi ratkaisu, jos ja vain jos $k \neq 1$ ja $k \neq -2$.

Redusoidut porrasmatriisit ovat teoreettisesti mielenkiintoisia, koska jokaista matriisia vastaa täsmälleen yksi redusoitu porrasmatriisi. Edellä kuitenkin nähtiin, että yhtälöryhmän ratkaisujen lukumäärä on luettavissa jo porrasmatriisista. Itse asiassa yhtälöryhmä voidaan jopa ratkaista porrasmatriisivaiheen avulla.

Esimerkin 4.4.5 yhtälöryhmää vastaava porrasmatriisi oli

$$\left[\begin{array}{ccc|c} 1 & 1 & -5 & 3 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Koska kolmannessa sarakkeessa ei ole johtavaa alkioita, sitä vastaava muuttuja on vapaa. Tähän havaintoon ei tarvita redusoitua porrasmuotoa. Jos merkitään $x_3 = t$, muut muuttujat voidaan ratkaista t :n avulla. Toisen rivin perusteella

$$x_2 - 3t = 2 \iff x_2 = 2 + 3t,$$

ja tämän jälkeen ensimmäisen rivin perusteella

$$x_1 + x_2 - 5t = 3 \iff x_1 + (3t + 2) - 5t = 3 \iff x_1 = 1 + 2t.$$

Porrasmatriisi siis riittää yhtälöryhmän ratkaisemiseen.

Homogeeniset yhtälöryhmät

Määritelmä 4.4.14

Lineaarinen yhtälöryhmä, jonka kaikki vakiot ovat nollia, on nimeltään *homogeeninen yhtälöryhmä*.

Homogeeninen yhtälöryhmä on siis muotoa

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0, \end{cases}$$

missä $a_{11}, \dots, a_{mn} \in \mathbb{R}$. Homogeenisella yhtälöryhmällä on aina ainakin yksi ratkaisu:

$$x_1 = 0, \quad x_2 = 0, \quad \dots \quad x_n = 0.$$

Tätä kutsutaan yhtälöryhmän triviaaliksi ratkaisuksi.

Lause 4.4.15

Jos homogeenisessa yhtälöryhmässä tuntemattomien määrä n on suurempi kuin yhtälöiden määrä m , yhtälöryhmällä on äärettömän monta ratkaisua.

Todistus. Ensinnäkin yhtälöryhmällä on välttämättä ainakin triviaali ratkaisu. Siten ratkaisuja on joko yksi tai äärettömän monta.

Oletuksen mukaan yhtälöryhmän matriisissa on pystyviivan vasemmalla puolella enemmän sarakkeita kuin koko matriisissa on rivejä. Toisaalta johtavia alkioita on enintään yksi joka rivillä. Siten matriisissa on oltava pystyviivan vasemmalla puolella ainakin yksi sarake, jossa ei ole johtavaa alkioita. Näin ollen löytyy vapaa muuttuja, mistä seuraa, että yhtälöryhmällä on äärettömän monta ratkaisua. $\square$

Tiivistelmä

- Gaussin-Jordanin eliminointimentelmällä voi ratkaista lineaarisia yhtälöryhmiä.
- Gaussin-Jordanin eliminointimentelmä antaa tavan, jolla matriisista voi aina muodostaa redusoidun porrasmatriisin. Tästä redusoidusta porrasmatriisista voidaan lukea yhtälöryhmän ratkaisut.
- Lineaarilla yhtälöryhmällä on joko yksi, ei yhtään tai äärettömän monta ratkaisua.
- Yhtälöryhmän ratkaisujen lukumäärän selville saamiseksi riittää muuttaa sitä vastaava matriisi porrasmuotoon.

4.5 Yhtälöryhmästä matriisiyhtälöksi

Aloita katsomalla video ”[Käänteismatriisit, aste, sarakeavaruus ja nolla-avaruus](#)”.

Tässä luvussa esitellään uusi tapa kirjoittaa lineaarinen yhtälöryhmä matriisien avulla käyttäen hyväksi matriisikertolaskua sekä vektoreita. Tällöin yhtälöryhmien ratkaisussa voidaan käyttää avuksi matriisien ominaisuuksia.

Tutkitaan yhtälöryhmää

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 + 5x_4 &= 6 \\ 11x_1 + 12x_2 + 13x_3 + 14x_4 &= 15 \\ -7x_1 - 8x_2 - 9x_3 - 10x_4 &= -11 \end{cases}$$

Yhtälön vasemmalla puolella näkyy jotain samanlaista kuin saadaan tulokseksi matriisien kertolaskusta. Kirjoitetaan yhtälöryhmän kertoimet yhdeksi matriisiksi

$$A = \begin{bmatrix} 1 & 3 & 4 & 5 \\ 11 & 12 & 13 & 14 \\ -7 & -8 & -9 & -10 \end{bmatrix},$$

tuntemattomat toiseksi matriisiksi

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

ja vakiot kolmanneksi

$$\mathbf{b} = \begin{bmatrix} 6 \\ 15 \\ -11 \end{bmatrix}.$$

Nyt

$$A\mathbf{x} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 + 5x_4 \\ 11x_1 + 12x_2 + 13x_3 + 14x_4 \\ -7x_1 - 8x_2 - 9x_3 - 10x_4 \end{bmatrix}$$

eli tulo $A\mathbf{x}$ vastaa yhtälöryhmän vasenta puolta. Siten yhtälöryhmä voidaan kirjoittaa muodossa $A\mathbf{x} = \mathbf{b}$ eli

$$\begin{bmatrix} 1 & 3 & 4 & 5 \\ 11 & 12 & 13 & 14 \\ -7 & -8 & -9 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \\ -11 \end{bmatrix}.$$

Yleisesti yhtälöryhmä

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

voidaan kirjoittaa muodossa $A\mathbf{x} = \mathbf{b}$, missä

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{ja} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Matriisia A kutsutaan yhtälöryhmän *kerroinmatriisiksi*. Jos nyt kerroinmatriisi on neliö-

matriisi, sen kääntyvyys vaikuttaa merkittävästi yhtälöryhmän ratkaisuihin.

Lause 4.5.1

Jos matriisi $A \in \mathbb{R}^{n \times n}$ on kääntyvä ja $\mathbf{b} \in \mathbb{R}^n$, yhtälöllä $A\mathbf{x} = \mathbf{b}$ on täsmälleen yksi ratkaisu, ja se on $\mathbf{x} = A^{-1}\mathbf{b}$.

Todistus. Oletetaan, että matriisi A on kääntyvä. Todistuksessa on kaksi osaa. On osoitettava, että $A^{-1}\mathbf{b}$ on yhtälön ratkaisu ja että muita ratkaisuja ei ole.

Osoitetaan ensin, että $A^{-1}\mathbf{b}$ on yhtälön ratkaisu sijoittamalla se yhtälön vasemmalle puolelle vektorin $\mathbf{x}$ paikalle:

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

Koska tuloksena oli yhtälön oikea puoli, esitetty ratkaisu toteuttaa yhtälön.

Osoitetaan sitten, ettei muita ratkaisuja ole. Oletetaan, että $\mathbf{y}$ on jokin (toinen) ratkaisu. Tällöin $A\mathbf{y} = \mathbf{b}$. Kerrotaan yhtälön molemmat puolet vasemmalta matriisillä A^{-1} , jolloin saadaan

$$A^{-1}A\mathbf{y} = A^{-1}\mathbf{b}.$$

Koska $A^{-1}A = I$, edellinen yhtälö sievenee muotoon $\mathbf{y} = A^{-1}\mathbf{b}$. Kysymyksessä on siis sama ratkaisu kuin edellä. Siten ratkaisuja on vain yksi ja se on $A^{-1}\mathbf{b}$. $\square$

Huomaa, että todistuksessa tarvittiin käänteismatriisin kumpaakin ominaisuutta: $AA^{-1} = I$ ja $A^{-1}A = I$.

Tiivistelmä

- Lineaariset yhtälöryhmät voidaan kirjoittaa matriisiyhtälönä muodossa $A\mathbf{x} = \mathbf{b}$.
- Matriisiyhtälöllä $A\mathbf{x} = \mathbf{b}$ on täsmälleen yksi ratkaisu, jos ja vain jos A on kääntyvä.

4.6 Alkeismatriisit

Myös alkeisrivimuunnokset voi ilmaista matriisikertolaskun avulla. Osoittautuu, että jos matriisia kerrotaan niin kutsutulla alkeismatriisilla, tullaan matriisille tehneeksi jokin alkeisrivimuunnos. Tästä havainnosta tulee olemaan hyötyä kääntyvien matriisien käsitteilyssä.

Määritelmä 4.6.1

Matriisi on *alkeismatriisi*, jos se on saatu ykkösmatriisista yhdellä alkeisrivimuunnoksella.

Esimerkiksi seuraavat matriisit ovat alkeismatriiseja:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Nämä alkeismatriisit on saatu ykkösmatriisista tekemällä sille alkeisrivimuunnokset $-\frac{1}{2}R_3$, $R_2 \leftrightarrow R_4$ ja $R_3 + 3R_1$.

Esimerkki 4.6.2

Osoittautuu, että alkeismatriiseilla kertominen vastaa alkeisrivimuunnosten tekemistä. Tutkitaan tätä edellisen esimerkin alkeismatriisien ja matriisin

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

avulla. Laskemalla nähdään, että E_1 kertoo kolmannen rivin luvulla $-\frac{1}{2}$,

$$E_1 A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -\frac{1}{2}a_{31} & -\frac{1}{2}a_{32} & -\frac{1}{2}a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix},$$

E_2 vaihtaa toisen ja neljännen rivin,

$$E_2 A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{41} & a_{42} & a_{43} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix},$$

ja E_3 lisää kolmanteen riviin luvulla 3 kerrotun ensimmäisen riviin,

$$E_3 A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 3a_{11} + a_{31} & 3a_{12} + a_{32} & 3a_{13} + a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}.$$

Huomataan, että jokaisella alkeismatriisilla kerrottaessa matriisille A tullaan tehneeksi sama alkeisrivimuunnos, jonka avulla alkeismatriisi muodostettiin.

Yksittäinen esimerkki ei takaa, että alkeismatriisilla kertominen vastaa aina alkeisrivimuunnoksen tekemistä. Esimerkin perusteella voi kuitenkin ymmärtää, miksi näin on. Yleisen tapauksen todistaminen ei ole vaikeaa, mutta se on kuitenkin melko työlästä, joten tyydytään mainitsemaan tulos ilman todistusta.

Lemma 4.6.3

Oletetaan, että $A \in \mathbb{R}^{n \times m}$. Olkoon E alkeismatriisi, joka saadaan tekemällä jokin alkeisrivimuunnos ykkösmatriisille I_n . Jos matriisille A tehdään sama alkeisrivimuunnos, tuloksena on matriisi EA .

Lemma tarkoittaa apulausetta. Se on siis pieni tulos, jota voidaan käyttää hyväksi tärkeämpien lauseiden todistamisessa.

Lause 4.6.4

Alkeismatriisit ovat kääntyviä, ja alkeismatriisin käänteismatriisi on myös alkeismatriisi.

Todistus. Tämänkään tuloksen tarkkaa todistusta ei esitetä tässä. Käydään kuitenkin läpi todistuksen idea.

Jokainen alkeisrivimuunnos voidaan peruuttaa toisella alkeisrivimuunnoksella kuten kohta nähdään. Kutsutaan tätä alkeisrivimuunnosta alkuperäisen alkeisrivimuunnoksen *käänteismuunnokseksi*.

Oletetaan, että $a, b \in \mathbb{R}$ ja $a \neq 0$. Jos matriisille tehdään alkeisrivimuunnos $R_i \leftrightarrow R_j$, päästään takaisin alkutilanteeseen tekemällä sama alkeisrivimuunnos uudelleen. Alkeisrivimuunnos $R_i \leftrightarrow R_j$ on siis itsensä käänteismuunnos. Alkeisrivimuunnoksen aR_i käänteismuunnos on puolestaan $\frac{1}{a}R_i$, ja alkeisrivimuunnoksen $R_i + bR_j$ käänteismuunnos on $R_i - bR_j$.

Alkeismatriisin käänteismatriisi saadaan aina käänteismuunnosta vastaavasta alkeismatriisista. Alkeisrivimuunnosta $R_i \leftrightarrow R_j$ vastaava alkeismatriisi on oma käänteismatriisinsa,

alkeisrivimuunnosta aR_i vastaavan alkeismatriisin käänteismatriisi on alkeisrivimuunnosta $\frac{1}{a}R_i$ vastaava alkeismatriisi ja niin edelleen. Alkeisrivimuunnoksen tekeminen vastaa nimittäin edellisen lemmän nojalla alkeismatriisilla kertomista. Esimerkiksi alkeisrivimuunnokset aR_i ja $\frac{1}{a}R_i$ peräkkäin suoritettuina eivät tee matriisille mitään. Siten niitä vastaavien alkeismatriisien tulo on ykkösmatriisi, jolla kertominen ei tee matriisille mitään.

□

Esimerkki 4.6.5

Etsitään alkeismatriisin

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

käänteismatriisi. Matriisi vastaa alkeisrivimuunnosta $R_3 + 3R_1$. Tämän alkeisrivimuunnoksen voi kumota tekemällä alkeisrivimuunnoksen $R_3 - 3R_1$. Sitä vastaava alkeismatriisi on

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

Laskemalla voi vielä varmistaa, että $EF = I$ ja $FE = I$. Siis $E^{-1} = F$.

Lauseessa 4.5.1 todettiin jo kääntyvien matriisien merkitys yhtälöryhmän ratkaisun kannalta. Nyt tuota tulosta voidaan täydentää tarkastelemalla lisäksi alkeisrivimuunnoksia ja niitä vastaavia alkeismatriiseja.

Lause 4.6.6 (Kääntyvien matriisien lause)

Oletetaan, että A on $n \times n$ -neliömatriisi. Seuraavat ehdot ovat yhtäpitäviä:

1. Matriisi A on kääntyvä.
2. Yhtälöllä $A\mathbf{x} = \mathbf{b}$ on täsmälleen yksi ratkaisu kaikilla $\mathbf{b} \in \mathbb{R}^n$.
3. Yhtälöllä $A\mathbf{x} = \mathbf{0}$ on vain triviaali ratkaisu $\mathbf{x} = \mathbf{0}$.
4. Matriisin A on riviekvivalentti ykkösmatriisin kanssa.
5. Matriisi A on alkeismatriisien tulo.
6. Matriisi A ei ole riviekvivalentti minkään nollarivin sisältävän matriisin kanssa.

Muista, että matriisit ovat riviekvivalentit, jos ne voidaan muuttaa alkeisrivimuunnoksilla toisikseen.

Todistus. Osoitetaan väite todistamalla seuraavat päättelyketjut:

$$1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 1 \quad \text{ja} \quad 4 \Rightarrow 6 \Rightarrow 4.$$

Tämän jälkeen tiedetään, että jokainen lauseen kohta on yhtäpitävä toisten kohtien kanssa.

$1 \Rightarrow 2$: Väite on osoitettu [lauseessa 4.5.1](#).

$2 \Rightarrow 3$: Oletetaan, että yhtälöllä $A\mathbf{x} = \mathbf{b}$ on täsmälleen yksi ratkaisu kaikilla $\mathbf{b} \in \mathbb{R}^n$. Tämä pätee myös, jos $\mathbf{b} = \mathbf{0}$. Toisaalta yhtälöllä $A\mathbf{x} = \mathbf{0}$ on aina ratkaisu $\mathbf{x} = \mathbf{0}$. Siten $\mathbf{x} = \mathbf{0}$ on ainoa ratkaisu.

$3 \Rightarrow 4$: Oletetaan, että yhtälöllä $A\mathbf{x} = \mathbf{0}$ on vain ratkaisu $\mathbf{x} = \mathbf{0}$. Merkitään $A(i, j) = a_{ij}$ kaikilla $i, j \in \{1, 2, \dots, n\}$. Yhtälöä $A\mathbf{x} = \mathbf{0}$ vastaava lineaarinen yhtälöryhmä on

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases}$$

Yhtälöryhmässä on sama määrä yhtälöitä ja tuntemattomia. Koska yhtälöryhmällä on täsmälleen yksi ratkaisu $\mathbf{x} = \mathbf{0}$, se on ekvivalentti yhtälöryhmän

$$\begin{cases} x_1 = 0 \\ x_2 = 0 \\ \vdots \\ x_n = 0 \end{cases}$$

kanssa. Tämä tarkoittaa, että matriisi A saadaan alkeisrivimuunnoksilla muutettua ykkösmatriisiksi. Toisin sanottuna A on riviekvivalentti ykkösmatriisin kanssa.

$4 \Rightarrow 5$: Oletetaan, että matriisi A on riviekvivalentti ykkösmatriisin kanssa. Toisin sanoen A voidaan muuttaa alkeisrivimuunnoksilla ykkösmatriisiksi I . Tällöin on olemassa alkeismatriisit $E_1, \dots, E_k$, joilla kertomalla matriisista A saadaan ykkösmatriisi. Päteekin siis

$$E_k \cdots E_1 A = I.$$

Kun yhtälön molemmat puolet kerrotaan vasemmalta matriisilla E_k^{-1} , saadaan uusi yhtälö $E_{k-1} \cdots E_1 A = E_k^{-1}$. Kun tämän yhtälön vasemmat puolet kerrotaan puolestaan matriisilla

E_{k-1}^{-1} , saadaan $E_{k-2} \cdots E_1 A = E_{k-1}^{-1} E_k^{-1}$. Jatkamalla samaan tapaan päädytään yhtälöön

$$A = E_1^{-1} \cdots E_{k-1}^{-1} E_k^{-1}.$$

Koska alkeismatriisin käänteismatriisi on myös alkeismatriisi, on väite todistettu.

$5 \Rightarrow 1$: Oletetaan, että $A = E_1 \cdots E_k$, missä $E_1, \dots, E_k$ ovat alkeismatriiseja. Merkitään $B = E_k^{-1} \cdots E_1^{-1}$. Nyt

$$\begin{aligned} AB &= (E_1 \cdots E_k)(E_k^{-1} \cdots E_1^{-1}) \\ &= E_1 \cdots (E_k E_k^{-1}) \cdots E_1^{-1} \\ &= E_1 \cdots E_{k-1} I E_{k-1}^{-1} \cdots E_1^{-1} \\ &\quad \vdots \\ &= E_1 E_1^{-1} = I. \end{aligned}$$

Siispä $AB = I$. Samalla tavalla nähdään, että $BA = I$. Siten B on matriisin A käänteismatriisi ja A on kääntyvä.

$4 \Rightarrow 6$: Oletetaan, että A on riviekvivalentti ykkösmatriisin kanssa. Tehdään lisäksi vastaoletus, että A on riviekvivalentti jonkin matriisin B kanssa, joka sisältää nollarivin. Tällöin matriisiyhtälöitä $A\mathbf{x} = \mathbf{0}$, $I\mathbf{x} = \mathbf{0}$ ja $B\mathbf{x} = \mathbf{0}$ vastaavilla yhtälöryhmillä on kaikilla samat ratkaisut.

Matriisi B sisältää nollarivin, joten joltakin sen riviltä puuttuu johtava alkio. Koska B sisältää yhtä monta riviä kuin saraketta, myös jostakin sen sarakkeesta puuttuu johtava alkio. Täten yhtälöryhmällä $B\mathbf{x} = \mathbf{0}$ on vapaa muuttuja ja ratkaisuja on ääretön määrä. Kuitenkin yhtälöllä $I\mathbf{x} = \mathbf{0}$ on vain yksi ratkaisu: $\mathbf{x} = \mathbf{0}$. Tämä on ristiriita, joten vastaoletus on väärä ja A ei ole riviekvivalentti nollarivin sisältävän matriisin kanssa.

$6 \Rightarrow 4$: Oletetaan, että A ei ole riviekvivalentti minkään nollarivin sisältävän matriisin kanssa. Olkoon B redusoitu porrasmatriisi, joka on riviekvivalentti matriisin A kanssa. Oletuksen mukaan B ei sisällä nollarivejä, joten sen jokaisella rivillä on johtava alkio. Koska B on neliömatriisi, myös sen jokaisessa sarakkeessa on johtava alkio. Tällöin B on itse asiassa ykkösmatriisi, joten A on riviekvivalentti ykkösmatriisin kanssa. $\square$

Käänteismatriisin määrittäminen

Seuraavaksi esitellään menetelmä, jonka avulla matriisiin kääntyvyyden ja mahdollisen käänteismatriisin voi selvittää. Menetelmä esitellään ensin. Sen jälkeen perustellaan, miksi menetelmä toimii.

Matriisin kääntyvyyden selvittäminen ja käänteismatriisin etsiminen voidaan tehdä yhtä aikaa. Oletetaan, että halutaan selvittää, onko matriisi A kääntyvä. Yhdistetään matriisi A ja ykkösmatriisi I matriisiksi $[A \mid I]$. Tehdään tälle matriisille alkeisrivimuunnoksia, joilla yritetään muuttaa A redusoiduksi porrasmatriisiksi. Jos A saadaan muutettua al-

keisrivimuunnosten avulla ykkösmatriisiksi, on A kääntyvä. Lisäksi matriisin I paikalle muodostuu matriisin A käänteismatriisi.

Havainnollistetaan menetelmää esimerkein.

Esimerkki 4.6.7

Tutkitaan, onko matriisilla

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

käänteismatriisi. Muokataan yhdistettyä matriisia

$$\left[\begin{array}{ccc|ccc} 2 & 4 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array} \right]$$

alkeisrivimuunnoksilla samaan tapaan kuin Gaussin–Jordanin menetelmässä. Tavoitteena on saada vasemmalle puolelle ykkösmatriisi. Muokkaus voi tapahtua esimerkiksi seuraavasti:

$$\begin{aligned} \left[\begin{array}{ccc|ccc} 2 & 4 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array} \right] &\xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 2 & 4 & -2 & 1 & 0 & 0 \end{array} \right] \xrightarrow{R_3 - 2R_1} \\ \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 4 & -10 & 1 & 0 & -2 \end{array} \right] &\xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 4 & -10 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array} \right] \xrightarrow{\frac{1}{4}R_2} \\ \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -\frac{5}{2} & \frac{1}{4} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array} \right] &\xrightarrow{R_2 + \frac{5}{2}R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & 0 & \frac{1}{4} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array} \right] \xrightarrow{R_1 - 4R_3} \\ \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -4 & 1 \\ 0 & 1 & 0 & \frac{1}{4} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array} \right]. \end{aligned}$$

Koska matriisi A saatiin muutettua alkeisrivimuunnoksilla ykkösmatriisiksi, on A

kääntävä. Lisäksi sen käänteismatriisi on

$$\begin{bmatrix} 0 & -4 & 1 \\ \frac{1}{4} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}.$$

Jos matriisi ei ole kääntävä, tulee myös se ilmi matriisia redusoidessa. Jos matriisia voidaan muokata alkeisrivimuunnoksilla niin, että saadaan aikaiseksi nollarivi, ei matriisi ole kääntävä. Toisin sanoen matriisi ei voi olla kääntävä, jos se on riviekvivalentti sellaisen matriisin kanssa, jossa on nollarivi.

Esimerkki 4.6.8

Tutkitaan, onko matriisilla

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

käänteismatriisi. Ryhdytään muokkaamaan yhdistettyä matriisia alkeisrivimuunnoksilla:

$$\begin{aligned} \left[\begin{array}{ccc|ccc} -1 & 0 & 2 & 1 & 0 & 0 \\ 4 & 1 & -3 & 0 & 1 & 0 \\ 3 & 1 & -1 & 0 & 0 & 1 \end{array} \right] &\xrightarrow{(-1)R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 4 & 1 & -3 & 0 & 1 & 0 \\ 3 & 1 & -1 & 0 & 0 & 1 \end{array} \right] &\xrightarrow{R_2-4R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 5 & 4 & 1 & 0 \\ 3 & 1 & -1 & 0 & 0 & 1 \end{array} \right] \\ &\xrightarrow{R_3-3R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 5 & 4 & 1 & 0 \\ 0 & 1 & 5 & 3 & 0 & 1 \end{array} \right] &\xrightarrow{R_3-R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 5 & 4 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{array} \right]. \end{aligned}$$

Koska matriisin B viimeisen rivin paikalle tuli nollarivi, matriisi B ei ole kääntävä.

Perustellaan sitten, miksi menetelmä toimii. Menetelmässä matriisia muokataan alkeisrivimuunnoksilla. Jos näin saadaan aikaiseksi ykkösmatriisi, on alkuperäinen matriisi kääntävä. Tämä perustuu siihen, että jos matriisi A onnistutaan muuttamaan alkeisrivimuunnoksilla ykkösmatriisiksi, niin A on lauseen 4.6.6 nojalla kääntävä eli sillä on käänteismatriisi A^{-1} .

Jos matriisi on kääntävä, käytetyistä alkeisrivimuunnoksista saadaan myös selville, mikä käänteismatriisi on. Oletetaan, että matriisi A on muutettu ykkösmatriisiksi alkeisrivimuunnoksilla, joita vastaavat alkeismatriisit $E_1, \dots, E_k$. Nyt

$$E_k \cdots E_1 A = I.$$

Tällöin käänteismatriisille pätee

$$A^{-1} = IA^{-1} = (E_k \cdots E_1 A)A^{-1} = E_k \cdots E_1 (AA^{-1}) = E_k \cdots E_1 I.$$

Tämä tarkoittaa, että tekemällä ykkösmatriisille I samat alkeisrivimuunnokset kuin tehtiin alunperin matriisille A päädytään käänteismatriisiin A^{-1} . Siis

$$[A \mid I] \rightsquigarrow [I \mid A^{-1}].$$

Siksi käänteismatriisi A^{-1} ilmestyy ykkösmatriisin I paikalle.

Lause 4.6.6 antaa välineet myös sen osoittamiseen, että matriisi ei ole kääntyvä. Lauseen mukaan matriisi A ei ole kääntyvä, jos se on riviekvivalentti sellaisen matriisin kanssa, joka sisältää nollarivin. Tämä voidaan ilmaista myös toisin: matriisi A ei ole kääntyvä, jos siihen saadaan alkeisrivimuunnoksilla muodostettua nollarivi.

Tiivistelmä

- Alkeismatriisilla kertominen vastaa alkeisrivimuunnoksen tekemistä.
- Matriisi on kääntyvä, jos ja vain jos se voidaan muuttaa alkeisrivimuunnoksilla ykkösmatriisiksi.

4.7 Sarake- ja nolla-avaruus yhtälönratkaisussa

Aiemmin määriteltiin matriisin **sarake- ja nolla-avaruudet**. Matriisin A sarakeavaruus on joukko

$$\mathcal{R}(A) = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = A\mathbf{x}, \text{ jollakin } \mathbf{x} \in \mathbb{R}^n\}$$

ja nolla-avaruus joukko

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

Tarkastellaan näihin aliavaruuksiin sisältyvää informaatiota matriisiyhtälön $A\mathbf{x} = \mathbf{b}$ ratkaisun näkökulmasta.

Jos sarakeavaruus $\mathcal{R}(A)$ on tiedossa, voidaan selvittää, onko matriisiyhtälöllä $A\mathbf{x} = \mathbf{b}$ ratkaisuja. Sarakeavaruus $\mathcal{R}(A)$ nimittäin koostuu niistä vektoreista $\mathbf{b}$, joita kohden yhtälöllä $A\mathbf{x} = \mathbf{b}$ on vähintään yksi ratkaisu.

Lause 4.7.1

Olkoon A $m \times n$ -matriisi, sekä $\mathbf{b}$ avaruuden $\mathbb{R}^m$ vektori. Yhtälöllä $A\mathbf{x} = \mathbf{b}$ on ratkaisuja, jos ja vain jos $\mathbf{b} \in \mathcal{R}(A)$

Todistus. Oletetaan ensin, että yhtälöllä $A\mathbf{x} = \mathbf{b}$ on ratkaisuja. Olkoon $\mathbf{v}$ eräs ratkaisu. Tällöin $A\mathbf{v} = \mathbf{b}$. Nyt sarakeavaruuden määritelmän nojalla $\mathbf{b} \in \mathcal{R}(A)$.

Oletetaan sitten, että $\mathbf{b} \in \mathcal{R}(A)$. Nyt sarakeavaruuden määritelmän nojalla on olemassa $\mathbf{v} \in \mathbb{R}^n$, jolle pätee $A\mathbf{v} = \mathbf{b}$. Siten yhtälöllä $A\mathbf{x} = \mathbf{b}$ on ratkaisu. $\square$

Jos tunnetaan matriisin nolla-avaruus sekä jokin yhtälön ratkaisu, saadaan selville kaikki ratkaisut.

Lause 4.7.2

Olkoon A $m \times n$ -matriisi, sekä $\mathbf{b}$ avaruuden $\mathbb{R}^m$ vektori. Olkoon $\mathbf{x}_0$ yhtälön $A\mathbf{x} = \mathbf{b}$ jokin ratkaisu. Tällöin yhtälön ratkaisut ovat täsmälleen muotoa $\mathbf{x}_0 + \mathbf{y}$, missä $\mathbf{y} \in \mathcal{N}(A)$.

Todistus. Osoitetaan, että $\mathbf{v} \in \mathbb{R}^n$ on yhtälön $A\mathbf{x} = \mathbf{b}$ ratkaisu, jos ja vain jos $\mathbf{v} = \mathbf{x}_0 + \mathbf{y}$ jollakin $\mathbf{y} \in \mathcal{N}(A)$.

Oletetaan ensin, että $\mathbf{v} \in \mathbb{R}^n$ on yhtälön $A\mathbf{x} = \mathbf{b}$ ratkaisu. Nyt täytyy löytää nolla-avaruuden $\mathcal{N}(A)$ alkio $\mathbf{y}$, jolle pätee $\mathbf{v} = \mathbf{x}_0 + \mathbf{y}$. Valitaan $\mathbf{y} = \mathbf{v} - \mathbf{x}_0$. Nyt

$$A(\mathbf{v} - \mathbf{x}_0) = A\mathbf{v} - A\mathbf{x}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Siten $\mathbf{y} \in \mathcal{N}(A)$. Lisäksi $\mathbf{v} = \mathbf{x}_0 + \mathbf{y}$.

Oletetaan sitten, että $\mathbf{v} = \mathbf{x}_0 + \mathbf{y}$ jollakin $\mathbf{y} \in \mathcal{N}(A)$. On osoitettava, että $\mathbf{v}$ on yhtälön $A\mathbf{x} = \mathbf{b}$ ratkaisu. Nähdään, että

$$A\mathbf{v} = A(\mathbf{x}_0 + \mathbf{y}) = A\mathbf{x}_0 + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Siten $\mathbf{v}$ on yhtälön $A\mathbf{x} = \mathbf{b}$ ratkaisu. $\square$

Esimerkki 4.7.3

Olkoon

$$A = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix} \quad \text{ja} \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Esitetään yhtälön $A\mathbf{x} = \mathbf{b}$ ratkaisut muodossa $\mathbf{x}_0 + \mathbf{y}$, missä $\mathbf{x}_0$ on yksittäisratkaisu ja $\mathbf{y}$ homogeenisen yhtälön yleinen ratkaisu.

Kirjoitetaan yhtälöryhmä $A\mathbf{x} = \mathbf{b}$ kokonaismatriisina

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

Kun matriisia muokataan alkeisrivimuunnoksilla, saadaan redusoitu porrasmatriisi

$$\left[\begin{array}{cccc|c} 1 & 0 & \frac{3}{2} & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{4} & -\frac{1}{4} & \frac{1}{2} \end{array} \right].$$

Tästä nähdään, että yhtälöryhmän ratkaisu on

$$\begin{cases} x_1 = -(3/2)t - (1/2)s \\ x_2 = 1/2 - (1/4)t + (1/4)s \\ x_3 = t \\ x_4 = s, \end{cases} \quad \text{missä } t, s \in \mathbb{R}.$$

Vektorimuodossa kirjoitettuna ratkaisu on

$$\mathbf{x} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ -\frac{1}{4} \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 0 \\ 1 \end{bmatrix}.$$

Merkitään

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 0 \end{bmatrix} \quad \text{ja} \quad \mathbf{y} = t \begin{bmatrix} -\frac{3}{2} \\ -\frac{1}{4} \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 0 \\ 1 \end{bmatrix}.$$

Osoittautuu, että nämä vektorit toteuttava vaaditut ehdot:

$$A\mathbf{x}_0 = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{b}$$

ja

$$A\mathbf{y} = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2}(3t+s) \\ -\frac{1}{4}(t-s) \\ t \\ s \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}(3t+s) - \frac{1}{2}(t-s) + 2t \\ -\frac{1}{2}(3t+s) + \frac{1}{2}(t-s) + t + s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0},$$

joten haluttu esitys on löydetty.

Koska lineaariset yhtälöryhmät voidaan kirjoittaa matriisiyhtälöinä, voidaan tässä osiossa esitettyjä tuloksia soveltaa lineaarisiin yhtälöryhmiin.

5 LINEAARIALGEBRAA

5.1 Yhtälöitä vektoriavaruuksissa

Aiemmissa luvussa opittiin ratkaisemaan lineaarisia yhtälöryhmiä sekä vastaamaan erilaisiin ratkaisuja koskeviin kysymyksiin. Hyödynnetään näitä tietoja nyt vektoriavaruuden $\mathbb{R}^n$ vektoreiden tutkimiseen.

Osaamme muun muassa vastata [esimerkissä 1.2.15](#) esitettyyn kysymykseen. Esimerkissä haluttiin tietää, onko vektori $\mathbf{w} = (-2, 3, 2, -1)$ vektoreiden

$$\mathbf{v}_1 = (0, -1, 2, 1), \quad \mathbf{v}_2 = (2, 0, 1, -1) \quad \text{ja} \quad \mathbf{v}_3 = (4, 2, 2, 0)$$

linearikombinaatio. Toisin sanoen oli selvitettävä, onko yhtälöllä

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 = \mathbf{w}.$$

ratkaisuja. Tällöin päädyttiin yhtälöryhmään

$$\begin{cases} 2x_2 + 4x_3 = -2 \\ -x_1 + 2x_3 = 3 \\ 2x_1 + x_2 + 2x_3 = 2 \\ x_1 - x_2 = -1. \end{cases}$$

Kun yhtälöryhmän matriisia käsitellään alkeisrivitoimituksilla, saadaan porrasmatriisi

$$\left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & 8 & 6 \\ 0 & 0 & 0 & -2 \end{array} \right].$$

Porrasmatriisissa on epätosi yhtälö, jonka perusteella tiedetään, että ratkaisuja ei ole. Siten $\mathbf{w}$ ei ole vektorien $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ linearikombinaatio.

5.2 Virittäminen

Pohdi 5.2.1

Marty McFly voi ohjata leijulautaansa vektoreilla $(3, -1)$ ja $(-2, -1)$. Tutkitaan, voiko Marty päästä mihin tahansa tason $\mathbb{R}^2$ pisteeseen. Voidaan olettaa, että Marty lähtee origosta $(0, 0)$.

1. Halutaan selvittää, voiko Marty päästä pisteeseen $(b_1, b_2) \in \mathbb{R}^2$. Millaista yhtälöä pitää tutkia?
2. Millainen yhtälöryhmä yhtälöstä saadaan?
3. Onko yhtälöryhmällä ratkaisuja? Vaikuttavatko lukujen b_1 ja b_2 arvot siihen, onko yhtälöryhmällä ratkaisuja?
4. Pääseekö Marty pisteeseen (b_1, b_2) ?
5. Voiko Marty päästä mihin tahansa tason $\mathbb{R}^2$ pisteeseen?

Olemme oppineet selvittämään, onko jokin vektori toisten vektoreiden lineaarikombinaatio. Tässä luvussa tutkitaan, milloin *jokainen* vektoriavaruuden vektori voidaan kirjoittaa joidenkin tiettyjen vektoreiden lineaarikombinaationa. Esimerkiksi avaruuden $\mathbb{R}^2$ jokainen vektori voidaan kirjoittaa vektorien $(1, 0)$ ja $(0, 1)$ lineaarikombinaationa. Avaruuden vektorit ovat nimittäin muotoa (x, y) , missä $x, y \in \mathbb{R}$. Jokainen tätä muotoa oleva alkio voidaan kirjoittaa lineaarikombinaationa

$$(x, y) = x(1, 0) + y(0, 1).$$

Sanotaan, että vektorit $(1, 0)$ ja $(0, 1)$ virittävät avaruuden $\mathbb{R}^2$. Asian voi ilmaista epämuodollisesti sanomalla, että vektoreilla $(1, 0)$ ja $(0, 1)$ pääsee jokaiseen avaruuden $\mathbb{R}^2$ pisteeseen.

Määritelmä 5.2.2

Vektorit $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ *virittävät* vektoriavaruuden $\mathbb{R}^n$, jos jokainen vektoriavaruuden $\mathbb{R}^n$ alkio voidaan kirjoittaa vektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_k$ lineaarikombinaationa.

Virittämisen määritelmä voidaan ilmaista myös toisin sanoin: vektorit $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ virittävät vektoriavaruuden $\mathbb{R}^n$, jos

$$\mathbb{R}^n = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_k\mathbf{v}_k \mid a_1, a_2, \dots, a_k \in \mathbb{R}\}.$$

Vektoreita $\mathbf{v}_1, \dots, \mathbf{v}_k$ kutsutaan vektoriavaruuden $\mathbb{R}^n$ *virittäjiksi*. Huomaa, että määritelmässä on oleellista, että avaruuden virittäjävektorit ovat avaruuden alkioita. Ei voida

esimerkiksi sanoa, että jotkin avaruuden $\mathbb{R}^3$ vektorit virittäisivät avaruuden $\mathbb{R}^2$, sillä $\mathbb{R}^3$ ja $\mathbb{R}^2$ ovat kaksi eri joukkoa, joilla ei ole yhteisiä alkioita.

Samalla tavalla kuin luvun alussa osoitettiin, että vektorit $(1, 0)$ ja $(0, 1)$ virittävät avaruuden $\mathbb{R}^2$, voidaan osoittaa, että vektorit $(1, 0, 0)$, $(0, 1, 0)$ ja $(0, 0, 1)$ virittävät avaruuden $\mathbb{R}^3$. Kyseiset kolme vektoria eivät kuitenkaan ole avaruuden $\mathbb{R}^3$ ainoat virittäjät kuten seuraavasta esimerkistä näkyy.

Esimerkki 5.2.3

Osoitetaan, että vektorit $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (1, 1, 0)$ ja $\mathbf{v}_3 = (1, 0, 0)$ virittävät avaruuden $\mathbb{R}^3$.

On siis osoitettava, että jokainen avaruuden $\mathbb{R}^3$ vektori voidaan esittää vektoreiden $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaationa. Oletetaan tätä varten, että $\mathbf{w} \in \mathbb{R}^3$. Nyt $\mathbf{w} = (w_1, w_2, w_3)$ joillakin $w_1, w_2, w_3 \in \mathbb{R}$.

Jotta vektori $\mathbf{w}$ olisi vektoreiden $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaatio, täytyy löytyä luvut $x_1, x_2, x_3 \in \mathbb{R}$, joille pätee

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 = \mathbf{w}.$$

Osoitetaan, että tällaisia lukuja on olemassa eli että yhtälöllä on ratkaisuja.

Yhtälöä vastaa yhtälöryhmä

$$\begin{cases} x_1 + x_2 + x_3 = w_1 \\ x_1 + x_2 = w_2 \\ x_1 = w_3, \end{cases}$$

jonka matriisi on

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & w_1 \\ 1 & 1 & 0 & w_2 \\ 1 & 0 & 0 & w_3 \end{array} \right].$$

Yhtälöryhmän matriisista saadaan alkeisrivitoimituksilla porrasmatriisi

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & w_1 \\ 0 & -1 & -1 & -w_1 + w_3 \\ 0 & 0 & -1 & -w_1 + w_2 \end{array} \right].$$

Tavoitteena on siis selvittää, onko yhtälöryhmällä ratkaisuja. Tämä voidaan lukea suoraan porrasmatriisista. Koska porrasmatriisissa ei näy epätosia yhtälöitä, yhtälöryhmällä on ratkaisuja. Siten etsityt reaaliluvut x_1 , x_2 ja x_3 ovat olemassa, ja $\mathbf{w}$ on

vektoreiden $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaatio.

Olemme nyt osoittaneet, että jokainen vektoriavaruuden $\mathbb{R}^3$ vektori on vektoreiden $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaatio. Näin ollen kyseiset kolme vektoria virittävät vektoriavaruuden $\mathbb{R}^3$.

Esimerkki 5.2.4

Tutkitaan, virittävätkö vektorit

$$\mathbf{v}_1 = (3, 2, -1), \quad \mathbf{v}_2 = (2, -2, 6) \quad \text{ja} \quad \mathbf{v}_3 = (3, 4, -5)$$

vektoriavaruuden $\mathbb{R}^3$.

Oletetaan, että $\mathbf{w} \in \mathbb{R}^3$. Nyt $\mathbf{w} = (w_1, w_2, w_3)$ joillakin $w_1, w_2, w_3 \in \mathbb{R}$. Tutkitaan, onko $\mathbf{w}$ vektoreiden $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaatio. Täytyy siis selvittää, onko olemassa lukuja $x_1, x_2, x_3 \in \mathbb{R}$, joille pätee

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 = \mathbf{w}.$$

Tätä yhtälöä vastaa yhtälöryhmä

$$\begin{cases} 3x_1 + 2x_2 + 3x_3 = w_1 \\ 2x_1 - 2x_2 + 4x_3 = w_2 \\ -x_1 + 6x_2 - 5x_3 = w_3. \end{cases}$$

Yhtälöryhmän matriisista saadaan alkeisrivitoimituksilla porrasmatriisi

$$\left[\begin{array}{ccc|c} 1 & -6 & 5 & -w_3 \\ 0 & 10 & -6 & w_2 + 2w_3 \\ 0 & 0 & 0 & w_1 - 2w_2 - w_3 \end{array} \right].$$

Porrasmatriisista nähdään, että yhtälöryhmällä on ratkaisuja, jos ja vain jos alinta riviä vastaava yhtälö $0x_1 + 0x_2 + 0x_3 = w_1 - 2w_2 - w_3$ on tosi eli $w_1 - 2w_2 - w_3 = 0$. Siten vektori $\mathbf{w} = (w_1, w_2, w_3)$ on vektoren $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaatio, jos ja vain jos $w_1 - 2w_2 - w_3 = 0$.

Näin ollen esimerkiksi vektori $(1, 0, 0)$ ei ole vektoreiden $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaatio, sillä se ei toteuta edellä saatua yhtälöä. On siis olemassa vektoriavaruuden $\mathbb{R}^3$ vektori, joka ei ole vektoreiden $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarikombinaatio. Siten vektorit $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ eivät viritä avaruutta $\mathbb{R}^3$.

Esimerkki 5.2.5

Tutkitaan, virittävätkö vektorit

$$\mathbf{u}_1 = (1, 1, 0), \quad \mathbf{u}_2 = (1, 0, 1), \quad \mathbf{u}_3 = (0, 1, 1) \quad \text{ja} \quad \mathbf{u}_4 = (-2, 1, 1)$$

avaruuden $\mathbb{R}^3$. Oletetaan, että $\mathbf{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$. On selvittettävä, onko olemassa lukuja $x_1, x_2, x_3, x_4 \in \mathbb{R}$, joille pätee

$$x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + x_3 \mathbf{u}_3 + x_4 \mathbf{u}_4 = \mathbf{w}.$$

Saadaan yhtälöryhmä, jonka matriisi on

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & -2 & w_1 \\ 1 & 0 & 1 & 1 & w_2 \\ 0 & 1 & 1 & 1 & w_3 \end{array} \right].$$

Tästä saadaan alkeisrivitoimituksilla porrasmatriisi

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & -2 & w_1 \\ 0 & 1 & -1 & -3 & w_1 - w_2 \\ 0 & 0 & 1 & 2 & \frac{1}{2}(w_3 + w_2 - w_1) \end{array} \right].$$

Yhtälöryhmällä on ratkaisuja, sillä porrasmatriisissa ei näy epätosia yhtälöitä. Tämä ei riipu mitenkään luvuista w_1, w_2 ja w_3 . Siten $\mathbf{w}$ on vektoreiden $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ ja $\mathbf{u}_4$ lineaarikombinaatio. Näin ollen kyseiset vektorit virittävät avaruuden $\mathbb{R}^3$.

Sekä [esimerkissä 5.2.3](#) että [esimerkissä 5.2.4](#) annetut vektorit virittävät avaruuden. Esimerkit poikkeavat toisistaan siinä suhteessa, että ensin mainitussa jokainen vektori voidaan kirjoittaa virittäjävektoreiden lineaarikombinaationa täsmälleen yhdellä tavalla ja jälkimmäisessä kirjoitustapoja on äärettömän monta. Tällä ei ole väliä, jos vain tutkitaan, virittävätkö vektorit avaruuden. Usein kuitenkin halutaan sellaiset virittäjävektorit, joiden lineaarikombinaatioina avaruuden vektorit voidaan kirjoittaa vain yhdellä tavalla. Tähän palataan [luvussa 5.4](#).

Jokaisessa tämän luvun esimerkissä muodostettiin ensin vektoryhtälö. Tuo yhtälö kirjoitettiin sitten yhtälöryhmänä ja edelleen matriisina. Huomataan, että näin muodostuvan matriisin sarakkeet ovat aina esimerkissä tutkitut virittäjävektorit. Vektoryhtälö

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_k \mathbf{v}_k = \mathbf{w}$$

voidaankin aina kirjoittaa matriisina

$$\left[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_k \mid \mathbf{w} \right].$$

Tämä merkintä tarkoittaa, että sarakkeiden alkiot ovat vektoreissa $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ olevat luvut. Ei ole kuitenkaan tarpeen opetella ulkoa, että virittäjävektorit laitetaan matriisin sarakkeiksi. Jos olet asiasta epävarma, voit aina tarkistaa sen laskemalla kuten esimerkeissä on tehty.

5.3 Aliavaruudet

Pohdi 5.3.1

1. Marty ohjailee leijulautaansa suuntavektoreilla $(-1, 3)$ ja $(3, -9)$. Miltä näyttää se avaruuden $\mathbb{R}^2$ osa, johon Marty voi päästä? Piirrä siitä kuva.
2. Jasmin voi ohjata taikamattoa suuntavektoreilla $(1, 1, 0)$ ja $(0, 0, 2)$. Miltä näyttää se avaruuden $\mathbb{R}^3$ osa, jonka Jasmin voi matollaan saavuttaa? Hahmottamista voi auttaa se, että listaat pisteitä, joihin Jasmin pääsee.

Aliavaruuksia käsitellään videossa ”[Linearikombinaatiot, aliavaruus, kantavektorit](#)”.

Edellisessä luvussa tutkittiin, milloin jotkin tietyt vektorit virittävät vektoriavaruuden $\mathbb{R}^n$. Esimerkiksi vektori $(3, -1)$ ei viritä avaruutta $\mathbb{R}^2$, sillä kaikkia avaruuden $\mathbb{R}^2$ vektoreita ei voida kirjoittaa tämän vektorin lineaarikombinaatioina. Toisaalta vaikkapa vektorit $(-3, 1, 1)$ ja $(2, -1, 2)$ eivät viritä avaruutta $\mathbb{R}^3$, sillä kaikkia avaruuden $\mathbb{R}^3$ vektoreita ei voida kirjoittaa näiden kahden vektorin lineaarikombinaatioina. (Miksi?)

Voidaan kuitenkin ajatella, että vektori $(3, -1)$ virittää avaruuden $\mathbb{R}^2$ sisällä pienemmän avaruuden, joka on vain osa koko avaruudesta $\mathbb{R}^2$. Samalla tavalla vektorit $(-3, 1, 1)$ ja $(2, -1, 2)$ virittävät avaruuden, joka on vain osa vektoriavaruutta $\mathbb{R}^3$. Tällaisia avaruuksia kutsutaan aliavaruuksiksi. Epämuodollisesti ilmaistuna vektoreiden virittämä aliavaruus on se osa avaruutta, johon kyseisillä vektoreilla voi päästä.

Annetaan seuraavaksi täsmällinen määritelmä vektorien virittämälle aliavaruudelle, ja tutkitaan, miltä tällaiset aliavaruudet näyttävät. Vektorien virittämä aliavaruus koostuu kaikista kyseisten vektorien lineaarikombinaatioista.

Määritelmä 5.3.2

Vektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ virittämä *aliavaruus* on joukko

$$\{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k \mid a_1, a_2, \dots, a_k \in \mathbb{R}\}.$$

Tätä joukkoa merkitään $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

Jos $W = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, sanotaan, että vektorit $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ virittävät aliavaruuden W . Vektoreita $\mathbf{v}_1, \dots, \mathbf{v}_k$ kutsutaan aliavaruuden W *virittäjiksi*.

Vektorien virittämää aliavaruutta kutsutaan toisinaan lyhyesti aliavaruudeksi. Tulemme näkemään, että aliavaruudet ovat vektoriavaruuksia toisten vektoriavaruuksien sisässä.

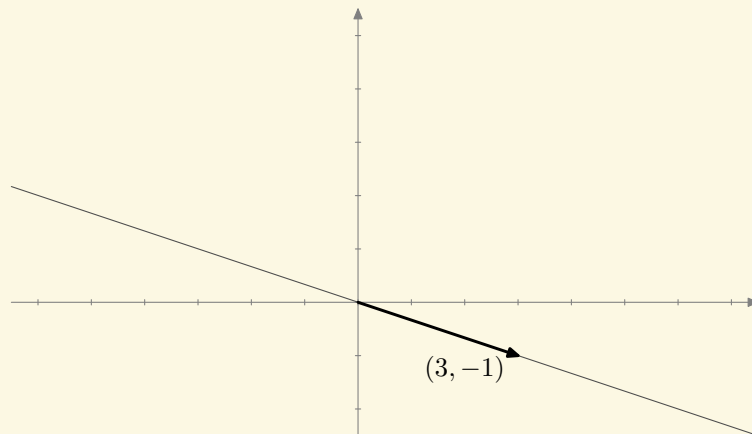
Huomaa, että merkinnässä $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ vektoreiden järjestyksellä ei ole väliä. Tämä johtuu siitä, että vektoreiden yhteenlaskussa summattavien järjestyksellä ei ole väliä. Merkintä span tulee englannin kielen verbistä ”span”, joka tarkoittaa virittämistä tai ulottamista.

Esimerkki 5.3.3

Tarkastellaan vektorin $(3, -1)$ virittämää aliavaruutta $\text{span}\{(-3, 1)\}$. Se koostuu määritelmän mukaan kaikista vektorin $(3, -1)$ lineaarikombinaatioista. Koska vektoreita on vain yksi, ovat lineaarikombinaatiot itse asiassa skalaarimonikertoja. Vektorin $(-3, 1)$ virittämä aliavaruus on

$$\text{span}\{(-3, 1)\} = \{a(3, -1) \mid a \in \mathbb{R}\}.$$

Tutkitaan, miltä aliavaruus $\text{span}\{(-3, 1)\}$ näyttää. Sen alkioita ovat esimerkiksi vektorit $2(3, -1) = (6, -2)$, $(-1/6)(3, -1) = (-1/2, 1/6)$ ja $0(3, -1) = (0, 0)$. Nämä vektorit ovat yhdensuuntaisia vektorin $(-3, 1)$ kanssa. Kun aliavaruuden $\text{span}\{(-3, 1)\}$ alkioita ajatellaan koordinaatiston pisteinä, huomataan pisteiden sijaitsevat samalla suoralla (kuva 5.3.1). Aliavaruus on $\text{span}\{(-3, 1)\}$ siis suora. Koska $(0, 0)$ on tämän suoran alkio, kulkee suora origon kautta.



Kuva 5.3.1. Aliavaruus $\text{span}\{(3, -1)\}$ on origon kautta kulkeva suora.

Esimerkki 5.3.4

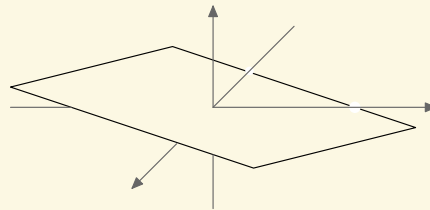
Tarkastellaan seuraavaksi, miltä näyttää vektorien $(-3, 1, 1)$ ja $(2, -1, 2)$ virittämä aliavaruus $\text{span}\{(-3, 1, 1), (2, -1, 2)\}$. Se on vektoriavaruuden $\mathbb{R}^3$ osajoukko. Vekto-

reiden lineaarikombinaatiot muodostavat joukon

$$\text{span}\{(-3, 1, 1), (2, -1, 2)\} = \{a_1(-3, 1, 1) + a_2(2, -1, 2) \mid a_1, a_2 \in \mathbb{R}\}.$$

Sen alkioita ovat esimerkiksi $-4(-3, 1, 1) - 2(2, -1, 2) = (-16, -2, -8)$ ja $-5(-3, 1, 1) + 0(2, -1, 2) = (15, -5, -5)$.

Aliavaruudesta $\text{span}\{(-3, 1, 1), (2, -1, 2)\}$ on hieman vaikeampi hahmotella kuvaa kuin aliavaruudesta $\text{span}\{(-3, 1)\}$. Kaikki vektorien $(-3, 1, 1)$ ja $(2, -1, 2)$ lineaarikombinaatiot ovat kuitenkin samassa tasossa kuin $(-3, 1, 1)$ ja $(2, -1, 2)$. Joukko $\text{span}\{(-3, 1, 1), (2, -1, 2)\}$ muodostaakin avaruuden $\mathbb{R}^3$ tason. Se kulkee origon kautta, sillä $(0, 0, 0)$ on vektorien $(-3, 1, 1)$ ja $(2, -1, 2)$ lineaarikombinaatio.



Kuva 5.3.2. Aliavaruus $\text{span}\{(-3, 1, 1), (2, -1, 2)\}$ on origon kautta kulkeva taso.

Edellä nähtiin, että avaruudessa $\mathbb{R}^3$ vektorien virittämä aliavaruus voi olla origon kautta kulkeva suora tai taso. Vektorien virittämässä aliavaruudessa voi myös olla vain yksi vektori. Nollavektorin virittämä aliavaruus on nimittäin $\text{span}\{\mathbf{0}\} = \{a\mathbf{0} \mid a \in \mathbb{R}\} = \{\mathbf{0}\}$. Tässä aliavaruudessa on siis ainoastaan nollavektori. Myös koko avaruus $\mathbb{R}^3$ on eräiden vektoreiden virittämä aliavaruus:

$$\begin{aligned} \text{span}\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} &= \{a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) \mid a_1, a_2, a_3 \in \mathbb{R}\} \\ &= \{(a_1, a_2, a_3) \mid a_1, a_2, a_3 \in \mathbb{R}\} = \mathbb{R}^3. \end{aligned}$$

Esimerkki 5.3.5

Tutkitaan, miltä näyttävät avaruuden $\mathbb{R}^4$ aliavaruuden

$$\text{span}\{(1, 0, -2, 5), (0, -1, 4, 0), (0, 0, 0, 1)\}$$

alkiot. Määritelmän mukaan

$$\begin{aligned} \text{span}\{(1, 0, -2, 5), (0, -1, 4, 0), (0, 0, 0, 1)\} \\ &= \{a_1(1, 0, -2, 5) + a_2(0, -1, 4, 0) + a_3(0, 0, 0, 1) \mid a_1, a_2, a_3 \in \mathbb{R}\} \\ &= \{(a_1, 0, -2a_1, 5a_1) + (0, -a_2, 4a_2, 0) + (0, 0, 0, a_3) \mid a_1, a_2, a_3 \in \mathbb{R}\} \\ &= \{(a_1, -a_2, -2a_1 + 4a_2, 5a_1 + a_3) \mid a_1, a_2, a_3 \in \mathbb{R}\}. \end{aligned}$$

Aliavaruuden $\text{span}\{(1, 0, -2, 2), (0, -1, 4, 5), (0, 0, 0, 1)\}$ alkio on siis muotoa

$$(a_1, -a_2, -2a_1 + 4a_2, 5a_1 + a_2),$$

missä $a_1, a_2, a_3 \in \mathbb{R}$.

Esimerkki 5.3.6

Joukko $W = \{(4a_1 - 2a_2, 3a_1 - a_2, 5a_1 + a_2) \mid a_1, a_2 \in \mathbb{R}\}$ on eräiden avaruuden $\mathbb{R}^3$ vektorien virittämä aliavaruus. Etsitään tälle aliavaruudelle virittäjävektorit. Toimitaan muuten samoin kuin [esimerkissä 5.3.5](#), mutta käännetään päättelyn suunta:

$$\begin{aligned} W &= \{(4a_1 - 2a_2, 3a_1 - a_2, 5a_1 + a_2) \mid a_1, a_2 \in \mathbb{R}\} \\ &= \{(4a_1, 3a_1, 5a_1) + (-2a_2, -a_2, a_2) \mid a_1, a_2 \in \mathbb{R}\} \\ &= \{a_1(4, 3, 5) + a_2(-2, -1, 1) \mid a_1, a_2 \in \mathbb{R}\} \\ &= \text{span}\{(4, 3, 5), (-2, -1, 1)\}. \end{aligned}$$

Kyseessä on siis vektorien $(4, 3, 5)$ ja $(-2, -1, 1)$ virittämä aliavaruus. Se on origon kautta kulkeva taso.

Toisinaan kaikkia virittäjävektoreita ei tarvita aliavaruuden virittämiseen. Virittäjävektorien joukossa voi siis olla turhia vektoreita. Tutkitaan vektoreiden $(1, 0, 0)$, $(0, 1, 0)$ ja $(1, 1, 0)$ virittämää aliavaruutta $\text{span}\{(1, 0, 0), (0, 1, 0), (1, 1, 0)\}$. Ensinnäkin vektorit $(1, 0, 0)$ ja $(0, 1, 0)$ virittävät xy -tason. Kun joukkoon lisätään lineaarikombinaatioita, joissa on mukana myös $(1, 1, 0)$, ei aikaiseksi saada mitään uutta, sillä myös $(1, 1, 0)$ on samassa xy -tasossa.

Seuraava lause osoittaa, että jos jokin virittäjävektori on toisten virittäjävektoreiden lineaarikombinaatio, se voidaan pudottaa pois virittäjävektoreiden joukosta.

Lause 5.3.7

Oletetaan, että $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{R}^n$ ja lisäksi että $\mathbf{w}$ on vektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaatio. Tällöin

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{w}\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}.$$

Todistus. ” $\subseteq$ ”: Oletetaan, että $\mathbf{v} \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}\}$. Nyt on olemassa reaali-lukukertoimet $a_1, \dots, a_k, a_w \in \mathbb{R}$, joille pätee

$$\mathbf{v} = a_1 \mathbf{v}_1 + \dots + a_k \mathbf{v}_k + a_w \mathbf{w}.$$

Toisaalta koska $\mathbf{w}$ on vektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_k$ lineaarikombinaatio, on olemassa toiset reaalilukukertoimet $c_1, \dots, c_k \in \mathbb{R}$, joille pätee

$$\mathbf{w} = c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k.$$

Sijoitetaan tämä ensimmäiseen yhtälöön, jolloin saadaan

$$\begin{aligned} \mathbf{v} &= a_1 \mathbf{v}_1 + \dots + a_k \mathbf{v}_k + a_w (c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k) \\ &= a_1 \mathbf{v}_1 + \dots + a_k \mathbf{v}_k + a_w c_1 \mathbf{v}_1 + \dots + a_w c_k \mathbf{v}_k \\ &= (a_1 + a_w c_1) \mathbf{v}_1 + \dots + (a_k + a_w c_k) \mathbf{v}_k. \end{aligned}$$

Tästä nähdään, että $\mathbf{v} \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, joten

$$\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}\} \subseteq \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}.$$

” $\supseteq$ ”: Todistuksen toinen osa jätetään harjoitustehtäväksi. □

Tarkastellaan vektoreiden $(-1, 0, -3)$, $(-6, 1, -1)$ ja $(0, 1, 2)$ virittämää aliavaruutta

$$\text{span}\{(-1, 0, -3), (-2, -2, -10), (0, 1, 2)\}.$$

Koska

$$(-2, -2, -10) = 2(-1, 0, -3) - 2(0, 1, 2),$$

voidaan vektori $(-2, -2, -10)$ jättää pois virittäjien joukosta. Toisin sanoen

$$\text{span}\{(-1, 0, -3), (-2, -2, -10), (-5, 1, 2)\} = \text{span}\{(-1, 0, -3), (-5, 1, 2)\}.$$

Seuraavassa esimerkissä tutkitaan, kuinka aliavaruudet ovat pienempiä vektoriavaruuksia toisten vektoriavaruuksien sisässä.

Esimerkki 5.3.8

Tarkastellaan aliavaruutta

$$W = \text{span}\{(-2, -1)\} = \{t(-2, -1) \mid t \in \mathbb{R}\}.$$

Kyseessä on origon kautta kulkeva suora.

Tutkitaan, mitä tapahtuu, kun kaksi aliavaruuden W alkioita lasketaan yhteen. Oletetaan, että $\mathbf{v}, \mathbf{w} \in W$. Tällöin on olemassa sellaiset reaalityöt a ja b , että $\mathbf{v} = a(-2, -1)$ ja $\mathbf{w} = b(-2, -1)$. Nähdään, että

$$\mathbf{v} + \mathbf{w} = a(-2, -1) + b(-2, -1) = (a + b)(-2, -1),$$

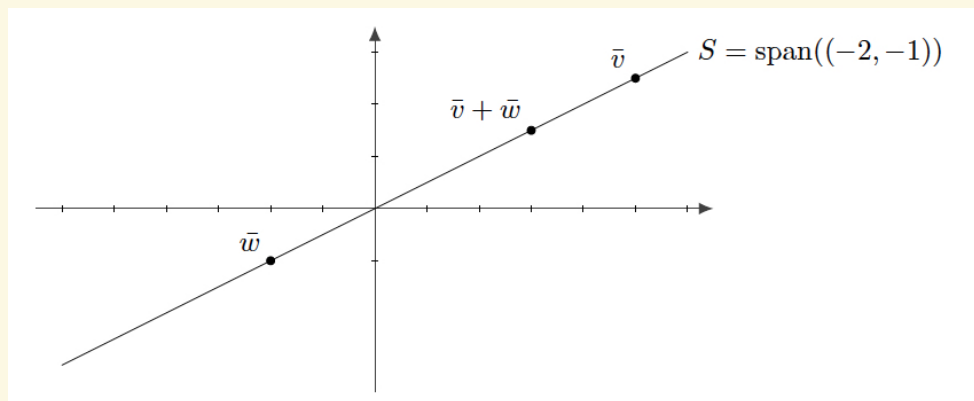
missä $a + b \in \mathbb{R}$. Havaitaan, että summa $\mathbf{v} + \mathbf{w}$ on vektorin $(-2, -1)$ skalaarimonikerta, joten se on aliavaruuden W alkio. Jos lasketaan yhteen mitkä tahansa kaksi aliavaruuden $W = \text{span}\{(-2, -1)\}$ alkioita, on tuloksena siis edelleen aliavaruuden W alkio.

Tarkastellaan sitten aliavaruuden alkioiden skalaarimonikertoja. Oletetaan, että $\mathbf{u} \in W$ ja $k \in \mathbb{R}$. Tällöin on olemassa $c \in \mathbb{R}$, jolle pätee $\mathbf{u} = c(-2, -1)$. Huomataan, että

$$k\mathbf{u} = k(c(-2, -1)) = (kc)(-2, -1),$$

missä $kc \in \mathbb{R}$. Havaitaan, että vektori $k\mathbf{u}$ voidaan kirjoittaa vektorin $(-2, -1)$ skalaarimonikertana, joten $k\mathbf{u} \in W$. Kaikkien aliavaruuden $W = \text{span}\{(-2, -1)\}$ alkioiden skalaarimonikerrat ovat siis edelleen aliavaruuden W alkioita.

Tavallaan suora $W = \text{span}\{(-2, -1)\}$ on oma pieni vektoriavaruutensa avaruuden $\mathbb{R}^2$ sisässä: kun suoran $W = \text{span}\{(-2, -1)\}$ alkioita lasketaan yhteen tai niitä kerrotaan reaaliluvuilla, on tuloksena edelleen aliavaruuden W alkio (ks. kuva 5.3.3). Sama pätee origon kautta kulkeviin tasoihin.



Kuva 5.3.3. Suoran $W = \text{span}\{(-2, -1)\}$ alkioiden $\mathbf{v}$ ja $\mathbf{w}$ summa $\mathbf{v} + \mathbf{w}$ on suoran W alkio.

Edellä tehdyt havainnot voidaan yleistää minkä tahansa vektoreiden virittämälle aliavaruudelle. Jos aliavaruuden kaksi vektoria lasketaan yhteen, on summa edelleen aliavaruudessa. Samoin aliavaruuden vektoreiden skalaarimonikerrat ovat aliavaruudessa. Lisäksi nollavektori kuuluu aina aliavaruuteen.

Lause 5.3.9

Oletetaan, että $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$. Olkoon $W = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$. Tällöin seuraavat väitteet pätevät:

1. Jos $\mathbf{u}, \mathbf{w} \in W$, niin $\mathbf{u} + \mathbf{w} \in W$.
2. Jos $\mathbf{w} \in W$ ja $c \in \mathbb{R}$, niin $c\mathbf{w} \in W$.
3. $\mathbf{0} \in W$.

Todistus. Osoitetaan kohta 1 ja jätetään loput kohdat harjoitustehtäviksi. Oletetaan, että $\mathbf{u}, \mathbf{w} \in W$. Nyt $\mathbf{u} = a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k$ joillakin $a_1, \dots, a_k \in \mathbb{R}$ ja $\mathbf{w} = b_1 \mathbf{v}_1 + \cdots + b_k \mathbf{v}_k$ joillakin $b_1, \dots, b_k \in \mathbb{R}$. Osoitetaan, että summa $\mathbf{u} + \mathbf{w}$ on aliavaruuden W alkio. Huomataan, että

$$\begin{aligned} \mathbf{u} + \mathbf{w} &= (a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k) + (b_1 \mathbf{v}_1 + \cdots + b_k \mathbf{v}_k) \\ &= (a_1 + b_1) \mathbf{v}_1 + \cdots + (a_k + b_k) \mathbf{v}_k. \end{aligned}$$

Koska $\mathbf{u} + \mathbf{w}$ on vektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaatio, pätee $\mathbf{u} + \mathbf{w} \in W$. $\square$

Tässä materiaalissa käsiteltiin vain vektorien virittämiä aliavaruuksia, jotka ovat lineaarikombinaatioista muodostuvia joukkoja. Ne ovat erikoistapaus yleisemmästä aliavaruuden käsitteestä. Vektoriavaruuden osajoukko W on kyseisen vektoriavaruuden *aliavaruus*, jos sille pätevät seuraavat ehdot:

1. $\mathbf{0} \in W$
2. $\mathbf{w} + \mathbf{u} \in W$ kaikilla $\mathbf{w}, \mathbf{u} \in W$
3. $r\mathbf{w} \in W$ kaikilla $r \in \mathbb{R}$ ja $\mathbf{w} \in W$.

Avaruuden $\mathbb{R}^n$ tapauksessa käsitteillä ei ole eroa: jokainen aliavaruus on vektorien virittäjä aliavaruus. Kun vektoriavaruuden käsitettä yleistetään muihinkin kuin avaruuksiin $\mathbb{R}^n$, löytyy aliavaruuksia, jotka eivät ole vektorien virittämiä aliavaruuksia.

Tiivistelmä

- Vektorit virittävät avaruuden, jos jokainen avaruuden vektori voidaan ilmaista niiden lineaarikombinaationa. Epämuodollisesti sanottuna tämä tarkoittaa sitä, että vektoreilla pääsee jokaiseen avaruuden pisteeseen.
- Jos vektorit eivät viritä koko avaruutta, ne virittävät aliavaruuden.
- Avaruuden $\mathbb{R}^3$ aliavaruuksia ovat origon kautta kulkevat suorat ja tasot, joukko $\{\mathbf{0}\}$ sekä koko avaruus $\mathbb{R}^3$.

5.4 Lineaarinen riippumattomuus

Pohdi 5.4.1

1. Martyn leijulautaa pystyy ohjaamaan suuntavektoreilla $(4, -2)$ ja $(2, -1)$. Minne kaikkialle Marty pääsee kulkemaan laudallaan? Jos toinen suuntavektoreista lakkaa toimimasta, miten se vaikuttaa siihen, minne kaikkialle Marty voi päästä?
2. Jasminin taikamattoa pystyy ohjaamaan suuntavektoreilla $(2, 0, 0)$, $(0, 0, -1)$ ja $(2, 0, -1)$. Onko jokin suuntavektoreista turha? Toisin sanoen, voiko Jasmin jättää jonkin suuntavektoreista pois ja päästä silti kaikkiin samoihin paikkoihin?

kuin mihin hän pääsee kolmella vektorilla?

Lineaarista riippumattomuutta käsitellään videossa ”[Lineaarikombinaatiot, aliavaruus, kantavektorit](#)”.

Edellisessä luvussa käsiteltiin vektoreiden virittämiä aliavaruuksia. Aliavaruudessa ovat kaikki ne pisteet, joihin virittäjävektoreilla voi päästä. Joskus osa virittäjävektoreista on turhia. [Lauseen 5.3.7](#) perusteella tiedetään, että jos jokin virittäjävektori on toisten virittäjävektorien lineaarikombinaatio, se ei tuota aliavaruuteen mitään uutta. Esimerkiksi

$$\text{span}\{(0, 2, 1), (-1, -1, 0), (-2, 0, 1)\} = \text{span}\{(0, 2, 1), (-1, -1, 0)\},$$

sillä

$$(-2, 0, 1) = (0, 2, 1) + 2(-1, -1, 0).$$

Usein ollaan kiinnostuneita sellaisista virittäjäjoukoista, jotka ovat minimaalisia, eikä niissä ole yhtään ylimääräisiä vektoreita. Tällaisia virittäjiä kutsutaan *lineaarisesti riippumattomiksi*. Vektorijoukkoa tarkasteltaessa on kuitenkin työlästä tarkistaa, onko jokin vektoreista toisten lineaarikombinaatio. Tätä varten pitäisi käydä läpi yksitellen kaikki joukon vektorit ja tarkistaa, ovatko ne toisten lineaarikombinaatioita. Siksi lineaariselle riippumattomuudelle on valittu hiukan erilainen mutta yhtäpitävä määritelmä, jota on helpompi käyttää. Ryhdytään tutustumaan tähän määritelmään.

Pohdi 5.4.2

Jasmin haastaa Aladdinin kisaan. Tehtävänä on lähteä kotipalatsilta ja palata samaan paikkaan. Ehtona on, että kutakin suuntavektoria saa käyttää vain kerran maton ohjaamiseen. (Yhdellä vektorilla saa kuitenkin aina kulkea niin pitkän matkan kuin haluaa.) Jasminin taikamatossa on suuntavektorit $(2, 0, 0)$, $(0, 0, -1)$ ja $(0, -2, 0)$. Aladdinin taikamatossa on suuntavektorit $(0, 2, 1)$, $(-1, -1, 0)$, $(-2, 0, 1)$.

1. Kirjoita yhtälö, joka vastaa sitä, että Jasmin lähtee kotipalatsilta, käyttää kutakin suuntavektoria kerran maton ohjaamiseen ja palaa takaisin kotipalatsille.
2. Onnistuuko Jasminin matka?
3. Entä onnistuuko Aladdin tekemään saman?
4. Ovatko Jasminin suuntavektorit lineaarisesti riippumattomia? Entä Aladdinin?

Jos virittäjävektoreiden joukossa on turhia vektoreita, tarkoittaa se, että joihinkin avaruuden pisteisiin päästään usealla eri tavalla. Esimerkiksi virittäjävektorien $(0, 2, 1)$, $(-1, -1, 0)$ ja $(-2, 0, 1)$ joukossa on turha vektori. Nyt nollavektorin voi kirjoittaa usealla eri tavalla niiden lineaarikombinaationa:

$$\mathbf{0} = (0, 2, 1) + 2(-1, -1, 0) - (-2, 0, 1)$$

ja

$$\mathbf{0} = 0(0, 2, 1) + 0(-1, -1, 0) - 0(-2, 0, 1).$$

Jos taas virittäjävektorien joukossa ei ole turhia vektoreita, päästään jokaiseen avaruuden pisteeseen vain yhdellä tavalla. Tällöin myös nollavektori saadaan aikaiseksi *ainoastaan* niin, että jokaisen virittäjävektorin kertoimen on nolla. Epämuodollisesti tämän voi ilmaista niin, että jos vektoreilla yrittää kulkea lenkin, joka palaa takaisin lähtöpisteeseen, sen voi toteuttaa ainoastaan pysymällä paikallaan.

Virittäjävektorien joukossa ei ole turhia vektoreita ainoastaan siinä tapauksessa, että nollavektorin voi kirjoittaa niiden lineaarikombinaationa ainoastaan yhdellä tavalla. Otetaan tämä lineaarisen riippumattomuuden määritelmäksi.

Määritelmä 5.4.3

Avaruuden $\mathbb{R}^n$ vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat *lineaarisesti riippumattomia* toisistaan, jos yhtälöllä

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0}$$

on täsmälleen yksi ratkaisu $x_1 = 0, x_2 = 0, \dots, x_k = 0$. (Tässä tuntemattomat $x_1, \dots, x_k$ ovat reaalilukuja.)

Jos vektorit eivät ole lineaarisesti riippumattomia, sanotaan, että ne ovat *lineaarisesti riippuvia* toisistaan.

Lineaarisesta riippumattomuuden ohella käytetään toisinaan myös ilmaisua *vapaa*. Jos vektorit $\mathbf{v}_1, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia, sanotaan, että niiden muodostama jono $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ on vapaa. Jos vektorit eivät ole lineaarisesti riippumattomia, sanotaan, että jono on *sidottu*.

Määritelmässä mainitulla yhtälöllä $x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0}$ on aina ratkaisu $x_1 = 0, x_2 = 0, \dots, x_k = 0$, olivat vektorit lineaarisesti riippumattomia tai ei. Tämä on yhtälön niin kutsuttu *triviaaliratkaisu*, joka on aina olemassa. Lineaarisesti riippumattomien vektoreiden erityisominaisuus on siis se, että yhtälöllä ei ole mitään muita ratkaisuja kuin triviaaliratkaisu.

Esimerkki 5.4.4

Merkitään $\mathbf{v}_1 = (1, 2)$ ja $\mathbf{v}_2 = (-3, -1)$. Tutkitaan, ovatko vektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ lineaarisesti riippumattomia.

Tarkastellaan yhtälöä $x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 = \mathbf{0}$, missä $x_1, x_2 \in \mathbb{R}$. Toisin sanoen tutkittava yhtälö on

$$x_1(1, 2) + x_2(-3, -1) = (0, 0)$$

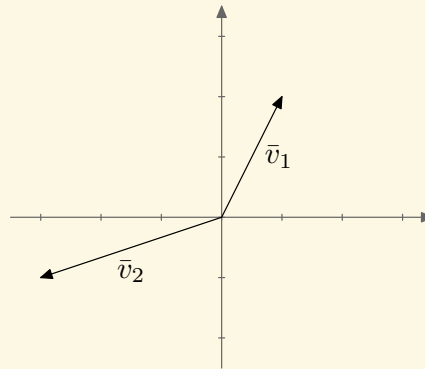
eli

$$\begin{cases} x_1 - 3x_2 = 0 \\ 2x_1 - x_2 = 0. \end{cases}$$

Ratkaistaan tästä x_1 ja x_2 :

$$\left[\begin{array}{cc|c} 1 & -3 & 0 \\ 2 & -1 & 0 \end{array} \right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & 5 & 0 \end{array} \right] \xrightarrow{\frac{1}{5}R_2} \left[\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & 1 & 0 \end{array} \right] \xrightarrow{R_1 + 3R_2} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right].$$

Ainoa ratkaisu on $x_1 = 0$ ja $x_2 = 0$. Vektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ ovat siis lineaarisesti riippumattomia (ks. kuva 5.4.1).



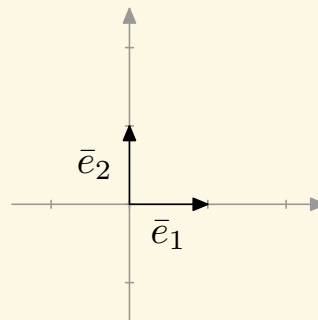
Kuva 5.4.1. Vektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ ovat lineaarisesti riippumattomia.

Esimerkki 5.4.5

Merkitään $\mathbf{e}_1 = (1, 0)$ ja $\mathbf{e}_2 = (0, 1)$. Tutkitaan, ovatko avaruuden $\mathbb{R}^2$ vektorit $\mathbf{e}_1$ ja $\mathbf{e}_2$ lineaarisesti riippumattomia. Tarkastellaan siis yhtälöä

$$x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 = \mathbf{0},$$

missä $x_1, x_2 \in \mathbb{R}$. Toisin sanoen ratkaistava yhtälö on $x_1(1, 0) + x_2(0, 1) = (0, 0)$. Yhtälön vasen puoli sievenee muotoon $x_1(1, 0) + x_2(0, 1) = (x_1, 0) + (0, x_2) = (x_1, x_2)$. Tutkittavana onkin itse asiassa yhtälö $(x_1, x_2) = (0, 0)$. Tämän ainoa ratkaisu on $x_1 = 0$ ja $x_2 = 0$. Näin on osoitettu, että vektorit $\mathbf{e}_1$ ja $\mathbf{e}_2$ ovat lineaarisesti riippumattomia.



Kuva 5.4.2. Vektorit $\mathbf{e}_1$ ja $\mathbf{e}_2$ ovat lineaarisesti riippumattomia.

Esimerkki 5.4.6

Kun vektoreita osoitetaan lineaarisesti riippuviksi, ei välttämättä tarvitse ratkaista yhtälöryhmää. Toisinaan on nimittäin helppo nähdä, minkälaisen kertoimien avulla lineaarikombinaatiosta muodostuu nollavektori.

Merkitään $\mathbf{w}_1 = (2, 1)$ ja $\mathbf{w}_2 = (-4, -2)$. Huomataan, että

$$2\mathbf{w}_1 + \mathbf{w}_2 = (4, 2) + (-4, -2) = \mathbf{0}.$$

Koska vektorien $\mathbf{w}_1$ ja $\mathbf{w}_2$ lineaarikombinaatio on nollavektori, vaikka kertoimet eivät ole nollija, vektorit $\mathbf{w}_1$ ja $\mathbf{w}_2$ ovat määritelmän nojalla lineaarisesti riippuvia.

Esimerkki 5.4.7

Merkitään $\mathbf{v}_1 = (1, 2)$, $\mathbf{v}_2 = (-3, -1)$ ja $\mathbf{v}_3 = (-1, 1)$. Tutkitaan, ovatko vektorit $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ lineaarisesti riippumattomia vai riippuvia. Tarkastellaan yhtälöä

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0},$$

missä $x_1, x_2 \in \mathbb{R}$. Tällöin

$$x_1(1, 2) + x_2(-3, -1) + c_3(-1, 1) = (0, 0)$$

eli komponenteittain

$$\begin{cases} x_1 - 3x_2 - c_3 = 0 \\ 2x_1 - x_2 + c_3 = 0. \end{cases}$$

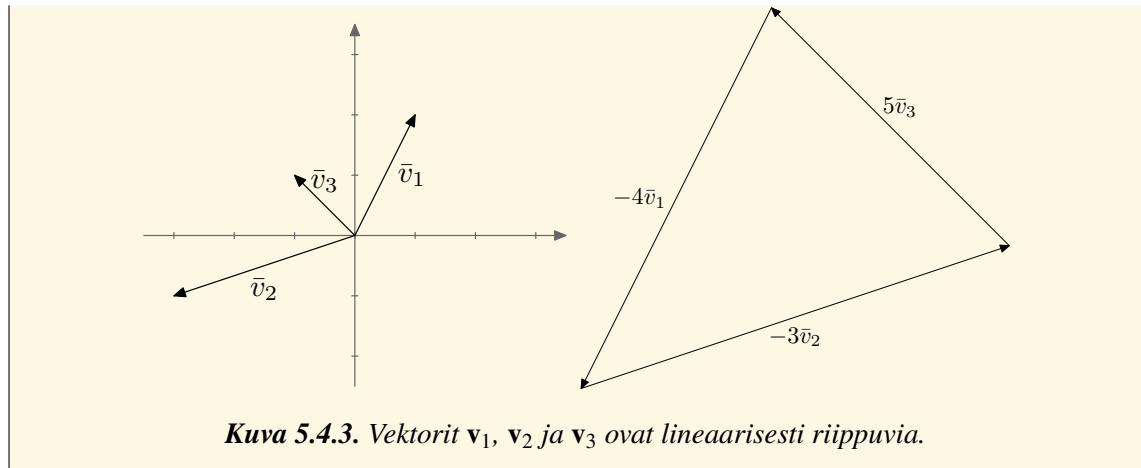
Ratkaistaan tästä x_1 ja x_2 :

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 2 & -1 & 1 & 0 \end{array} \right] &\xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 0 & 5 & 3 & 0 \end{array} \right] &\xrightarrow{\frac{1}{5}R_2} \left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 0 & 1 & 3/5 & 0 \end{array} \right] \\ &\xrightarrow{R_1 + 3R_2} \left[\begin{array}{ccc|c} 1 & 0 & 4/5 & 0 \\ 0 & 1 & 3/5 & 0 \end{array} \right]. \end{aligned}$$

Huomataan, että yhtälöryhmällä on äärettömän monta ratkaisua:

$$\begin{cases} x_1 = -(4/5)t \\ x_2 = -(3/5)t \\ x_3 = t \end{cases} \quad \text{missä } t \in \mathbb{R}.$$

Näin ollen $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ ei ole ainoa ratkaisu. Voidaan valita esimerkiksi $t = 5$, jolloin $c_1 = -4$ ja $c_2 = -3$ ja $c_3 = 5$. Tällöin $-4\mathbf{v}_1 - 3\mathbf{v}_2 + 5\mathbf{v}_3 = \mathbf{0}$. Vektorit $\mathbf{v}_1$, $\mathbf{v}_2$ ja $\mathbf{v}_3$ ovat siis lineaarisesti riippuvia. Tilannetta on havainnollistettu kuvassa 5.4.3.



Määritelmän mukaan vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia, jos yhtälöllä

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0}$$

on täsmälleen yksi ratkaisu $x_1 = 0, x_2 = 0, \dots, x_k = 0$. Näin ollen lineaarisen riippumattomuuden ehdon voi kirjoittaa myös muodossa

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0}, \quad \text{jos ja vain jos} \quad x_1 = 0, x_2 = 0, \dots, x_k = 0.$$

Jos kuitenkin $x_1 = 0, x_2 = 0, \dots, x_k = 0$, niin $x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0}$. Ekvivalenssin toinen suunta on siis aina totta. Siksi lineaarisen riippumattomuuden määritelmä voidaan lyhentää seuraavanlaiseen muotoon:

$$\text{jos} \quad x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0}, \quad \text{niin} \quad x_1 = 0, x_2 = 0, \dots, x_k = 0.$$

Tällaista muotoilua on kätevää käyttää esimerkiksi todistuksissa, joissa ei käsitellä konkreettisia vektoreita.

Kahden vektorin tapauksessa lineaarinen riippumattomuus on helppo tarkistaa. Rittää tutkia, ovatko vektorit yhdensuuntaisia.

Lause 5.4.8

Oletetaan, että $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ ja kumpikaan vektoreista ei ole nollavektori. Tällöin $\mathbf{v}$ ja $\mathbf{w}$ ovat lineaarisesti riippumattomia, jos ja vain jos ne eivät ole yhdensuuntaisia.

Todistus. Todistus jätetään harjoitustehtäväksi. □

Luvun alussa kuvailtiin lineaarisen riippumattomuuden ideaa usealla eri tavalla. Osoitetaan nyt tuloksia, jotka ilmaisevat samat asiat täsmällisesti. Aiemmin todettiin, että lineaarisesti riippumattomien vektorien joukossa ei ole turhia vektoreita. Tämä huomio sisältyy seuraavaan lauseeseen, jonka mukaan vektorit ovat lineaarisesti riippuvia, jos ja vain jos jokin vektoreista voidaan ilmaista toisten lineaarikombinaationa.

Lause 5.4.9

Oletetaan, että $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{R}^n$ ja $k \geq 2$. Vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippuvia, jos ja vain jos jollakin $j \in \{1, 2, \dots, k\}$ vektori $\mathbf{v}_j$ on vektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_k$ lineaarikombinaatio.

Todistus. Muotoa ”jos ja vain jos” oleva väite todistetaan kahdessa osassa. Ensin oletetaan väitteen ensimmäisen osan olevan totta ja osoitetaan, että tällöin jälkimmäinen osa pätee. Tätä todistuksen vaihetta merkitään usein symbolilla ” $\Rightarrow$ ”. Sitten oletetaan jälkimmäisen osan olevan totta ja osoitetaan, että ensimmäinen osa pätee. Tätä todistuksen vaihetta merkitään symbolilla ” $\Leftarrow$ ”. Ryhdytään todistamaan väitettä.

” $\Rightarrow$ ”: Oletetaan, että vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippuvia. On siis olemassa reaalityyppiset $c_1, \dots, c_k$, joilla pätee

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k = \mathbf{0},$$

ja lisäksi $c_j \neq 0$ jollakin $j \in \{1, 2, \dots, k\}$. Nyt

$$c_j \mathbf{v}_j = -c_1 \mathbf{v}_1 - \dots - c_{j-1} \mathbf{v}_{j-1} - c_{j+1} \mathbf{v}_{j+1} - \dots - c_k \mathbf{v}_k$$

ja edelleen

$$\mathbf{v}_j = -\frac{c_1}{c_j} \mathbf{v}_1 - \dots - \frac{c_{j-1}}{c_j} \mathbf{v}_{j-1} - \frac{c_{j+1}}{c_j} \mathbf{v}_{j+1} - \dots - \frac{c_k}{c_j} \mathbf{v}_k.$$

Siis $\mathbf{v}_j$ on vektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_k$ lineaarikombinaatio.

” $\Leftarrow$ ”: Oletetaan sitten, että $\mathbf{v}_j$ on vektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_k$ lineaarikombinaatio jollakin $j \in \{1, 2, \dots, k\}$. Nyt on olemassa sellaiset $c_1, \dots, c_{j-1}, c_{j+1}, \dots, c_k \in \mathbb{R}$, että

$$\mathbf{v}_j = c_1 \mathbf{v}_1 + \dots + c_{j-1} \mathbf{v}_{j-1} + c_{j+1} \mathbf{v}_{j+1} + \dots + c_k \mathbf{v}_k.$$

Tästä seuraa, että

$$\mathbf{0} = c_1 \mathbf{v}_1 + \dots + c_{j-1} \mathbf{v}_{j-1} + (-1) \mathbf{v}_j + c_{j+1} \mathbf{v}_{j+1} + \dots + c_k \mathbf{v}_k.$$

Koska kerroin -1 ei ole nolla, vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippuvia. □

Esimerkki 5.4.10

Tarkastellaan vektoriavaruuden $\mathbb{R}^3$ vektoreita $\mathbf{v}_1 = (1, -1, 0)$, $\mathbf{v}_2 = (1, 1, 0)$, $\mathbf{v}_3 = (0, 0, 2)$ ja $\mathbf{v}_4 = (3, -1, 0)$. Näillä pätee muun muassa

$$2\mathbf{v}_1 + \mathbf{v}_2 + 0\mathbf{v}_3 - \mathbf{v}_4 = (2, -2, 0) + (1, 1, 0) + (0, 0, 0) - (3, -1, 0) = (0, 0, 0),$$

joten vektorit $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ ja $\mathbf{v}_4$ ovat lineaarisesti riippuvia. Edellisen lauseen perus-

teella jokin vektoreista voidaan kirjoittaa toisten lineaarikombinaationa. Yllä olevasta yhtälöstä nähdäänkin, että

$$\mathbf{v}_2 = -2\mathbf{v}_1 + \mathbf{v}_4.$$

Kaikkia vektoreita ei kuitenkaan välttämättä voida kirjoittaa toisten lineaarikombinaationa. Esimerkiksi ei ole olemassa sellaisia lukuja a , b ja c , että pätsi

$$\mathbf{v}_3 = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_4.$$

(Tämän täsmällinen todistaminen jätetään lukijalle.)

Luvun alussa todettiin, että lineaarinen riippumattomuus tarkoittaa sitä, että aliavaruuden vektorit voidaan kirjoittaa virittäjävektoreiden lineaarikombinaatioina vain yhdellä tavalla. Osoitetaan tämä väite.

Lause 5.4.11

Jos vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia, niiden virittämän aliavaruuden $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ alkio voidaan kirjoittaa täsmälleen yhdellä tavalla virittäjävektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaatioina.

Todistus. Oletetaan, että vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia. Merkitään $W = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$. Osoitetaan, että jokainen aliavaruuden W alkio voidaan kirjoittaa täsmälleen yhdellä tavalla virittäjävektoreiden $\mathbf{v}_1, \dots, \mathbf{v}_k$ lineaarikombinaationa.

Ensinnäkin aliavaruuden määritelmän perusteella jokainen aliavaruuden W alkio voidaan kirjoittaa virittäjävektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaationa. Osoitetaan, että vektorit voidaan kirjoittaa lineaarikombinaationa *vain yhdellä tavalla*.

Oletetaan, että alkio $\mathbf{w} \in W$ voidaan kirjoittaa lineaarikombinaationa

$$\mathbf{w} = a_1\mathbf{v}_1 + \dots + a_k\mathbf{v}_k \tag{5.4.1}$$

ja lineaarikombinaationa

$$\mathbf{w} = b_1\mathbf{v}_1 + \dots + b_k\mathbf{v}_k \tag{5.4.2}$$

joillakin $a_1, \dots, a_k, b_1, \dots, b_k \in \mathbb{R}$. Nyt $a_1\mathbf{v}_1 + \dots + a_k\mathbf{v}_k = b_1\mathbf{v}_1 + \dots + b_k\mathbf{v}_k$, joten

$$a_1\mathbf{v}_1 + \dots + a_k\mathbf{v}_k - (b_1\mathbf{v}_1 + \dots + b_k\mathbf{v}_k) = \mathbf{0}.$$

Vektorien yhteenlaskun ja skalaarikertolaskun ominaisuuksien perusteella pätee

$$(a_1 - b_1)\mathbf{v}_1 + \dots + (a_k - b_k)\mathbf{v}_k = \mathbf{0}.$$

Vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia oletuksen nojalla, joten yllä olevasta yhtälöstä seuraa, että kaikki kertoimet ovat nollija: $a_1 - b_1 = 0, \dots, a_k - b_k = 0$. Siten

$a_1 = b_1, \dots, a_k = b_k$. Näin ollen tutkitut lineaarikombinaatiot (5.4.1) ja (5.4.2) ovatkin itse asiassa samanlaiset (niissä on samat kertoimet). Siksi vektoria $\mathbf{w}$ ei voida kirjoittaa usealla eri tavalla virittäjävektoreiden lineaarikombinaationa. $\square$

Lause 5.4.12

Oletetaan, että $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbb{R}^m$, missä $n \in \{1, 2, \dots\}$. Jos $n > m$, niin vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ ovat lineaarisesti riippuvia.

Todistus. Merkitään $\mathbf{v}_k = (v_{1k}, v_{2k}, \dots, v_{mk})$ kaikilla $k \in \{1, \dots, n\}$. Nyt yhtälöä

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n = \mathbf{0}$$

vastaavaksi yhtälöryhmäksi saadaan

$$\begin{cases} v_{11}x_1 + v_{12}x_2 + \dots + v_{1n}x_n = 0 \\ v_{21}x_1 + v_{22}x_2 + \dots + v_{2n}x_n = 0 \\ \vdots \\ v_{m1}x_1 + v_{m2}x_2 + \dots + v_{mn}x_n = 0. \end{cases}$$

Tässä homogeenisessa yhtälöryhmässä on enemmän tuntemattomia kuin yhtälöitä. Siten yhtälöryhmällä on [lauseen 4.4.15](#) nojalla äärettömän monta ratkaisua. Koska löytyy muitakin ratkaisuja kuin triviaaliratkaisu, vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ ovat lineaarisesti riippuvia. $\square$

Tiivistelmä

- Vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{R}^n$ ova lineaarisesti riippumattomia, jos nollavektori voidaan kirjoittaa niiden lineaarikombinaationa vain yhdellä tavalla.
- Jos aliavaruuden virittäjät ovat lineaarisesti riippumattomia, joukossa ei ole turhia virittäjiä.
- Vektorit ovat lineaarisesti riippumattomia, jos ja vain jos mikään vektoreista ei ole toisten lineaarikombinaatio.

5.5 Kanta ja dimensio

Jos avaruuden virittäjävektorit ovat lineaarisesti riippumattomia, niitä kutsutaan avaruuden kannaksi. Kantan vektorit muodostavat ihanteellisen virittäjävektorien joukon, sillä niiden joukossa ei ole turhia virittäjiä ja siksi jokainen avaruuden vektori voidaan kirjoittaa kannan vektorien lineaarikombinaationa täsmälleen yhdellä tavalla.

Määritelmä 5.5.1

Olkoot $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{R}^n$. Vektorijoukko $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ on vektoriavaruuden $\mathbb{R}^n$ kanta, jos seuraavat ehdot pätevät:

1. vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ virittävät avaruuden $\mathbb{R}^n$.
2. vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ ovat lineaarisesti riippumattomia.

Esimerkki 5.5.2

Luvun 5.2 alussa osoitettiin, että vektorit $\mathbf{e}_1 = (1, 0)$ ja $\mathbf{e}_2 = (0, 1)$ virittävät avaruuden $\mathbb{R}^2$. Vektorit $\mathbf{e}_1$ ja $\mathbf{e}_2$ ovat lisäksi lineaarisesti riippumattomia **esimerkin 5.4.5** perusteella. Siten $\{\mathbf{e}_1, \mathbf{e}_2\}$ on avaruuden $\mathbb{R}^2$ kanta.

Vastaavasti avaruudella $\mathbb{R}^n$ on kanta

$$\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}.$$

Tässä $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$, missä luku 1 on vektorin i :s komponentti. Kantaa kutsutaan avaruuden $\mathbb{R}^n$ *luonnolliseksi kannaksi*. Tässä materiaalissa vektoriavaruuden $\mathbb{R}^n$ luonnollista kantaa merkitään symbolilla $\mathcal{E}_n$.

Seuraava lause osoittaa, että vektorit muodostavat avaruuden kannan, jos ja vain jos avaruuden vektorit voidaan kirjoittaa niiden lineaarikombinaatioina täsmälleen yhdellä tavalla. Kannan määritelmässä on kaksi ehtoa, joista ensimmäinen koskee virittämistä ja toinen lineaarista riippumattomuutta. Virittäminen on yhtäpitävää sen kanssa, että avaruuden vektorit *voidaan* kirjoittaa annettujen vektoreiden lineaarikombinaatioina. Lineaarinen riippumattomuus puolestaan on yhtäpitävää sen kanssa, että avaruuden vektorit voidaan kirjoittaa annettujen vektorien lineaarikombinaatioina *täsmälleen yhdellä tavalla*.

Lause 5.5.3

Joukko $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ on vektoriavaruuden $\mathbb{R}^n$ kanta, jos ja vain jos jokainen avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa täsmälleen yhdellä tavalla vektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaationa.

Todistus. " $\Rightarrow$ ": Oletetaan, että $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ on vektoriavaruuden $\mathbb{R}^n$ kanta. Nyt

$$\mathbb{R}^n = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}.$$

Lisäksi vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia. Siten **lauseen 5.4.11** perusteella jokainen avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa täsmälleen yhdellä tavalla vektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaationa.

” $\Leftarrow$ ”: Oletetaan sitten, että jokainen avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa täsmälleen yhdellä tavalla vektorien $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaationa. Osoitetaan, että joukko $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ on avaruuden $\mathbb{R}^n$ kanta.

Koska avaruuden $\mathbb{R}^n$ vektorit voidaan kirjoittaa vektorien $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaationa, virittävät vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ avaruuden $\mathbb{R}^n$.

On vielä osoitettava, että vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia. Sitä varten tutkitaan yhtälöä

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_k \mathbf{v}_k = \mathbf{0},$$

missä $x_1 = 0, \dots, x_k \in \mathbb{R}$. Tiedetään, että ainakin

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + \dots + 0\mathbf{v}_k = \mathbf{0}.$$

Koska nollavektori $\mathbf{0}$ on vektoriavaruuden $\mathbb{R}^n$ alkio, se voidaan oletuksen mukaan kirjoittaa virittäjävektorien lineaarikombinaationa täsmälleen yhdellä tavalla. Siksi pätee $x_1 = 0, \dots, x_k = 0$. Siis vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia. $\square$

Kannan määritelmää käytettäessä täytyy tarkistaa kaksi ehtoa, virittäminen ja lineaarinen riippumattomuus. Niissä molemmissa muodostetaan yhtälöryhmän, jonka kertoimet ovat samat. Tällöin tullaan tehneeksi ylimääräistä työtä. [Lausetta 5.5.3](#) käytettäessä puolestaan tarvitsee tarkistaa vain yksi ehto, joka yhdistää virittämisen ja lineaarisen riippumattomuuden.

Esimerkki 5.5.4

Merkitään $\mathbf{w}_1 = (2, -1)$, $\mathbf{w}_2 = (1, 3)$. Osoitetaan [lauseen 5.5.3](#) avulla, että $\{\mathbf{w}_1, \mathbf{w}_2\}$ on avaruuden $\mathbb{R}^2$ kanta. Oletetaan, että $\mathbf{v} \in \mathbb{R}^2$. Ratkaistaan yhtälö $x_1 \mathbf{w}_1 + x_2 \mathbf{w}_2 = \mathbf{v}$ eli yhtälö $x_1(2, -1) + x_2(1, 3) = (v_1, v_2)$. Sitä vastaa yhtälöryhmä

$$\begin{cases} 2x_1 + x_2 = v_1 \\ -x_1 + 3x_2 = v_2. \end{cases}$$

Kun yhtälöryhmän matriisia muokataan alkeisrivitoimituksilla, saadaan matriisi

$$\left[\begin{array}{cc|c} 1 & -3 & -v_2 \\ 0 & 7 & v_1 + 2v_2 \end{array} \right].$$

Koska tässä porrasmatriisissa ei ole epätosia yhtälöitä ja jokaisessa sarakkeessa viivan vasemmalla puolella on johtava alkio, yhtälöryhmällä on täsmälleen yksi ratkaisu riippumatta vektorista $\mathbf{v} \in \mathbb{R}^2$.

Siten jokainen avaruuden $\mathbb{R}^2$ vektori voidaan kirjoittaa vektorien $\mathbf{w}_1$ ja $\mathbf{w}_2$ lineaarikombinaationa täsmälleen yhdellä tavalla. Näin ollen $\{\mathbf{w}_1, \mathbf{w}_2\}$ on avaruuden $\mathbb{R}^2$ kanta.

Aliavaruuden kanta

Samalla tavalla kuin vektoriavaruudelle määriteltiin kanta voidaan määritellä kanta aliavaruudelle. Aiempi kannan käsite on uuden määritelmän erikoistapaus, sillä koko avaruus $\mathbb{R}^n$ on aina itsensä aliavaruus.

Määritelmä 5.5.5

Oletetaan, että W on vektoriavaruuden $\mathbb{R}^n$ aliavaruus ja että $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k \in W$. Joukko $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ on aliavaruuden W *kanta*, jos seuraavat ehdot pätevät:

1. vektorit $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$ virittävät aliavaruuden W
2. vektorit $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$ ovat lineaarisesti riippumattomia.

Esimerkki 5.5.6

Tutkitaan avaruuden $\mathbb{R}^3$ aliavaruutta $W = \text{span}\{(3, 9, -3), (1, 5, 1)\}$ ja etsitään sille kanta. Vektorit $(3, 9, -3)$ ja $(1, 5, 1)$ virittävät avaruuden W . Lisäksi nämä kaksi vektoria ovat lineaarisesti riippumattomia, sillä ne eivät ole yhdensuuntaisia (ks. [lause 5.4.8](#)). Siten vektorit $(3, 9, -3)$ ja $(1, 5, 1)$ muodostavat aliavaruuden W kannan.

Esimerkki 5.5.7

Etsitään aliavaruudelle $W = \text{span}\{((3, -1, 5), (2, 1, 3), (0, -5, 1))\}$ kanta. Aloitetaan tutkimalla, sattuvatko annetut virittäjävektorit muodostamaan kannan. Tätä varten täytyy tutkia, ovatko ne lineaarisesti riippumattomia. Tarkastellaan siis yhtälöä

$$x_1(3, -1, 5) + x_2(2, 1, 3) + x_3(0, -5, 1) = (0, 0, 0), \quad (5.5.1)$$

missä $x_1, x_2, x_3 \in \mathbb{R}$. Yhtälöstä saadaan yhtälöryhmä, jota vastaa matriisi

$$\left[\begin{array}{ccc|c} 3 & 2 & 0 & 0 \\ -1 & 1 & -5 & 0 \\ 5 & 3 & 1 & 0 \end{array} \right].$$

Tästä saadaan alkeisrivitoimituksilla matriisi

$$\left[\begin{array}{ccc|c} 1 & -1 & 5 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Porrasmatriisista nähdään, että yhtälöryhmällä on äärettömän monta ratkaisua, joten vektorit $(3, -1, 5)$, $(2, 1, 3)$ ja $(0, -5, 1)$ eivät ole lineaarisesti riippumattomia. Siten ne eivät muodosta kantaa aliavaruudelle.

Matriisin avulla voidaan kuitenkin löytää eräs aliavaruuden kanta. Muuttuja x_3 on vapaa, joten sen arvo voidaan valita vapaasti, ja valinta määrää muuttujien x_1 ja x_2 arvot. Siispä yhtälössä (5.5.1) voi kerroin x_3 saada minkä tahansa arvon, esimerkiksi arvon 1. Tällöin saadaan yhtälö

$$a_1(3, -1, 5) + a_2(2, 1, 3) + (0, -5, 1) = (0, 0, 0),$$

missä a_1 ja a_2 ovat joitakin reaalilukuja. Tästä seuraa edelleen, että $(0, -5, 1) = -a_1(3, -1, 5) - a_2(2, 1, 3)$.

Nyt tiedetään, että $(0, -5, 1)$ on toisten virittäjävektorien lineaarikombinaatio, joten lauseen 5.3.7 perusteella

$$W = \text{span}\{(3, -1, 5), (2, 1, 3)\}.$$

Koska vektorit $(3, -1, 5)$ ja $(2, 1, 3)$ eivät ole yhdensuuntaisia, ne ovat lineaarisesti riippumattomat lauseen 5.4.8 nojalla. Siten $\{(3, -1, 5), (2, 1, 3)\}$ on aliavaruuden W kanta.

Samaa menetelmää käyttäen voidaan aina löytää vektorien virittämälle aliavaruudelle kanta.

Avaruudella voi olla monta erilaista kantaa. Esimerkiksi avaruudella $\mathbb{R}^2$ on luonnollinen kanta $\{(1, 0), (0, 1)\}$, mutta edellisessä esimerkissä nähtiin, että myös $\{(2, -1), (1, 3)\}$ on avaruuden $\mathbb{R}^2$ kanta. Huomataan, että kummassakin kannassa on sama määrä vektoreita. Seuraava lause osoittaa, että jokaisessa avaruuden kannassa on aina sama määrä vektoreita.

Lause 5.5.8

Oletetaan, että W on vektoriavaruuden $\mathbb{R}^n$ aliavaruus. Jokaisessa aliavaruuden W kannassa on yhtä monta vektoria.

Todistus. Oletetaan, että $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_j)$ ja $\mathcal{C} = (\mathbf{w}_1, \dots, \mathbf{w}_k)$ ovat molemmat aliavaruuden W kantoja. Pyritään osoittamaan, että $j = k$. Tehdään se osoittamalla, että muut vaihtoehdot $j < k$ ja $k < j$ johtavat ristiriitaan.

Oletetaan, että $j < k$. Tarkastellaan yhtälöä

$$x_1 \mathbf{w}_1 + \dots + x_k \mathbf{w}_k = \mathbf{0}. \quad (5.5.2)$$

Koska $\mathcal{B}$ on aliavaruuden W kanta, voidaan kaikki kannan $\mathcal{C}$ vektorit kirjoittaa kannan $\mathcal{B}$

vektorien lineaarikombinaatioina:

$$\mathbf{w}_1 = a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1j}\mathbf{v}_j$$

$$\mathbf{w}_2 = a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{2j}\mathbf{v}_j$$

$$\vdots$$

$$\mathbf{w}_k = a_{k1}\mathbf{v}_1 + a_{k2}\mathbf{v}_2 + \cdots + a_{kj}\mathbf{v}_j$$

joillakin $a_{11}, \dots, a_{kj} \in \mathbb{R}$. Sijoittamalla nämä yhtälöön (5.5.2) muodostuu yhtäpitävä yhtälö

$$\begin{aligned} x_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1j}\mathbf{v}_j) + x_2(a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{2j}\mathbf{v}_j) \\ + \cdots + x_k(a_{k1}\mathbf{v}_1 + a_{k2}\mathbf{v}_2 + \cdots + a_{kj}\mathbf{v}_j) = \mathbf{0}, \end{aligned}$$

josta saadaan edelleen ryhmittelemällä

$$\begin{aligned} (x_1a_{11} + x_2a_{21} + \cdots + x_ka_{k1})\mathbf{v}_1 + (x_1a_{12} + x_2a_{22} + \cdots + x_ka_{k2})\mathbf{v}_2 \\ + \cdots + (x_1a_{1j} + x_2a_{2j} + \cdots + x_ka_{kj})\mathbf{v}_j = \mathbf{0}. \end{aligned}$$

Joukko $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_j)$ on kanta, joten sen vektorit ovat lineaarisesti riippumaton. Siten edellinen yhtälö toteutuu, jos ja vain jos kaikki kertoimet ovat nollia:

$$\begin{cases} x_1a_{11} + x_2a_{21} + \cdots + x_ka_{k1} = 0 \\ x_1a_{12} + x_2a_{22} + \cdots + x_ka_{k2} = 0 \\ \qquad \qquad \qquad \vdots \qquad \qquad \vdots \\ x_1a_{1j} + x_2a_{2j} + \cdots + x_ka_{kj} = 0 \end{cases}$$

Kyseessä on homogeeninen yhtälöryhmä, jossa tuntemattomien määrä k on suurempi kuin yhtälöiden määrä j . [Lauseen 4.4.15](#) mukaan yhtälöryhmällä on muitakin ratkaisuja kuin triviaaliratkaisu $x_1 = 0, \dots, x_k = 0$. Siis jonon $C = (\mathbf{w}_1, \dots, \mathbf{w}_k)$ vektorit ovat lineaarisesti riippuvia. Tämä on ristiriita, sillä C on aliavaruuden W kanta.

Tapaus $j > k$ käsitellään vastaavasti. Tällöinkin päädytään ristiriitaan. Täytyy siis päteä $j = k$. □

Dimensio

Lukijalla on todennäköisesti jonkinlainen käsitys siitä, mitä ulottuvuudella eli dimensiolla tarkoitetaan. Esimerkiksi avaruus $\mathbb{R}^2$ on kaksiulotteinen samoin kaikki avaruuden $\mathbb{R}^3$ tasot. Avaruus $\mathbb{R}^3$ puolestaan on kolmiulotteinen.

Dimensio määritellään täsmällisesti kannan avulla: avaruuden dimensio on sen kannassa olevien vektorien lukumäärä. Dimension määrittelemisen tällä tavalla ei ole aivan suoraviivaista, sillä yhdellä avaruudella voi olla monta erilaista kantaa. Periaatteessa eri

kannoissa voisi olla eri määrä vektoreita, eikä silloin olisi selvää, mikä niistä määräisi avaruuden dimension. [Lauseen 5.5.8](#) perusteella kuitenkin tiedetään, että jokaisessa kannassa on aina yhtä monta vektoria. Siten dimension määritelmä on järkevä.

Määritelmä 5.5.9

Avaruuden dimensio on sen kannassa olevien vektorien lukumäärä.

Jos avaruuden dimensio on n , sanotaan, että avaruus on n -ulotteinen. Sovitaan, että avaruuden $\{0\}$ dimensio on nolla.

[Esimerkin 5.5.2](#) perusteella vektoriavaruuksella $\mathbb{R}^n$ on aina luonnollinen kanta

$$\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}.$$

Tässä kannassa on n vektoria. Siten avaruuden $\mathbb{R}^n$ dimensio on n , eli $\mathbb{R}^n$ on n -ulotteinen.

[Esimerkissä 5.5.6](#) tutkittiin aliavaruutta $W = \text{span}\{(3, 9, -3), (1, 5, 1)\}$, joka on taso. Esimerkissä nähtiin, että aliavaruudella W on kanta $\{(3, 9, -3), (1, 5, 1)\}$, jossa on kaksi vektoria. Aliavaruuden dimensio on siis 2.

Esimerkki 5.5.10

Määritetään avaruuden $\mathbb{R}^3$ aliavaruuden

$$W = \text{span}\{((1, 3, -1), (-3, -9, 3))\}$$

dimensio. Vektorit $(1, 3, -1)$ ja $(-3, -9, 3)$ kylläkin virittävät avaruuden W , mutta ne eivät muodosta aliavaruuden kantaa. Huomataan nimittäin, että $3(1, 3, -1) = (-3, -9, 3)$, joten vektorit ovat yhdensuuntaisia ja siksi ne eivät [lauseen 5.4.8](#) mukaan ole lineaarisesti riippumattomia kuten kantavektorien pitäisi olla.

Koska toinen virittäjävektori on ensimmäisen skalaarimonikerta, [lauseen 5.3.7](#) nojalla

$$W = \text{span}\{((1, 3, -1), (-3, -9, 3))\} = \text{span}\{((1, 3, -1))\}.$$

Yksi nollavektorista poikkeava vektori on lineaarisesti riippumaton. Näin ollen aliavaruuden W kannan muodostaa vektori $(1, 3, -1)$, joten $\dim(W) = 1$.

Koordinaatit

Pohdi 5.5.11

Jasmin lentää taikamatollaan. Hän ohjaa mattoa suuntavektoreilla $\mathbf{v}_1 = (0, 2, -1)$, $\mathbf{v}_2 = (1, 2, 0)$ ja $\mathbf{v}_3 = (1, 0, 2)$. Jasmin pitää kirjata matkoistaan. Jotta kirjoittamista

olisi mahdollisimman vähän, Jasmin listaa vain käytetyt vektorimäärät.

1. Matkapäiväkirjassa on matka $-5, 3, 1$. Mihin pisteeseen Jasmin matkusti?
2. Jasmin matkustaa pisteeseen $(10, 10, 10)$. Mitä hän kirjoittaa matkapäiväkirjaansa?

Kun avaruuden vektori kirjoitetaan kannan vektorien lineaarikombinaationa, kertoimia kutsutaan vektorin koordinaateiksi.

Määritelmä 5.5.12

Oletetaan, että $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ on aliavaruuden W kanta. Oletetaan, että $\mathbf{u} \in W$. Vektorin $\mathbf{u}$ koordinaateiksi kannan $\mathcal{B}$ suhteen kutsutaan reaalityyppisiä lukuja $b_1, \dots, b_k$, joilla

$$\mathbf{u} = b_1 \mathbf{v}_1 + \dots + b_k \mathbf{v}_k.$$

Vektorin koordinaatit jonkin tietyn kannan suhteen ovat yksikäsitteiset, sillä vektori voidaan lauseen 5.5.3 mukaan kirjoittaa vain yhdellä tavalla kannan alkioiden lineaarikombinaationa. Vektorilla on siis kunkin tietyn kannan suhteen vain yhdet koordinaatit. Eri kantojen suhteen saman vektorin koordinaatit voivat tietenkin olla erilaisia.

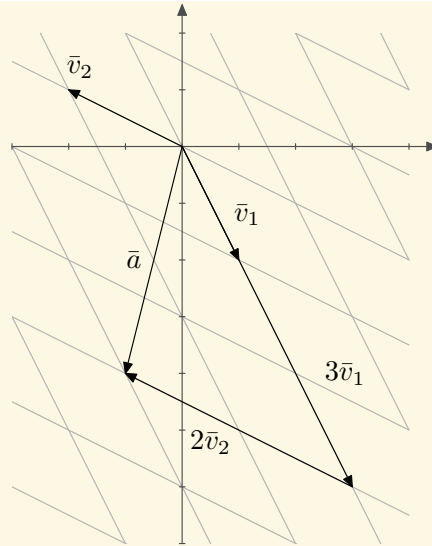
Esimerkki 5.5.13

Luonnollisen kannan suhteen koordinaatit on helppo määrittää. Esimerkiksi vektorin $\mathbf{a} = (-1, -4)$ koordinaatit avaruuden $\mathbb{R}^2$ luonnollisen kannan $\mathcal{E}_2 = \{\mathbf{e}_1, \mathbf{e}_2\}$ suhteen ovat -1 ja -4 , sillä $\mathbf{a} = (-1, -4) = (-1)(1, 0) + (-4)(0, 1) = (-1)\mathbf{e}_1 + (-4)\mathbf{e}_2$.

Tutkitaan sitten erästä toista avaruuden $\mathbb{R}^2$ kantaa. Merkitään $\mathbf{v}_1 = (1, -2)$, $\mathbf{v}_2 = (-2, 1)$. Nyt $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ on avaruuden $\mathbb{R}^2$ kanta, minkä osoittaminen jätetään lukijalle. Määritetään vektorin $\mathbf{a} = (-1, -4)$ koordinaatit kannan $\mathcal{B}$ suhteen. On ratkaistava yhtälö $\mathbf{a} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2$ eli yhtälö

$$(-1, -4) = x_1(1, -2) + x_2(-2, 1).$$

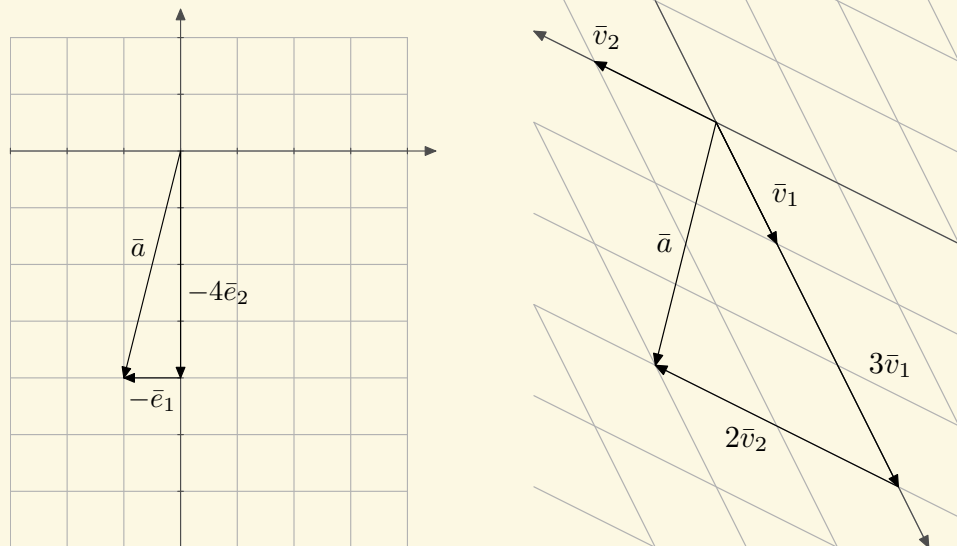
Tästä yhtälöstä saadaan tuttuun tapaan yhtälöryhmä, jonka ratkaisuksi paljastuu $x_1 = 3$, $x_2 = 2$. Näin ollen $\mathbf{a} = 3\mathbf{v}_1 + 2\mathbf{v}_2$ eli vektorin $\mathbf{a}$ koordinaatit kannan $\mathcal{B}$ suhteen ovat 3 ja 2 (ks. kuva 5.5.1).



Kuva 5.5.1. Vektori $\mathbf{a} = (-1, -4)$ ilmaistuna kannan $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ vektoreiden lineaarikombinaationa. Vektorin koordinaatit kannan $\mathcal{B}$ suhteen ovat 3 ja 2.

Vektorin $\mathbf{a} = (-1, -4)$ koordinaatteja kannan $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ suhteen on havainnollistettu kuvassa 5.5.1. Jotta päästään pisteeseen $(-1, -4)$ on mentävä 3 ruutua vektorin $\mathbf{v}_1$ suuntaan ja 2 ruutua vektorin $\mathbf{v}_2$ suuntaan. Siten vektorin $\mathbf{a} = (-1, -4)$ koordinaatit kannan $\mathcal{B}$ suhteen ovat 3 ja 2.

Kun käytetään jotakin muuta kuin luonnollista kantaa, vääntyy koordinaatisto vinoon. Kuvassa 5.5.2 vasemmalla vektori $\mathbf{a}$ on piirretty luonnollista kantaa $\mathcal{E}_2 = \{\mathbf{e}_1, \mathbf{e}_2\}$ vastaavaan tavalliseen koordinaatistoon ja oikealla kantaa $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ vastaavaan koordinaatistoon.



Kuva 5.5.2. Vektori $\mathbf{a}$ luonnollisen kannan $\mathcal{E}_2$ määräämässä koordinaatistossa ja kannan $\mathcal{B}$ määräämässä koordinaatistossa. Kun käytetään jotakin muuta kuin luonnollista kantaa, vääntyy koordinaatisto vinoon.

Koordinaattien tapauksessa järjestys on tärkeä. Koordinaatit luetellaan samassa järjestyksessä kuin kannan vektorit. Palataan vielä edelliseen esimerkkiin, jossa tutkittiin kantaa

$\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ edellisessä esimerkissä vektori, jonka koordinaatit kannan $\mathcal{B}$ suhteen ovat 2 ja -3 , on

$$2\mathbf{v}_1 - 3\mathbf{v}_2 = 2(1, -2) - 3(-2, 1) = (8, -7).$$

Vektori, jonka koordinaatit kannan $\mathcal{B}$ suhteen ovat -3 ja 2 , on puolestaan

$$-3\mathbf{v}_1 + 2\mathbf{v}_2 = -3(1, -2) + 2(-2, 1) = (-5, 8).$$

Se, missä järjestyksessä koordinaatit listataan, on siis oleellista.

Kanta on tässä materiaalissa määritelty joukkona. Tarkalleen ottaen joukkojen tapauksessa alkioiden järjestyksellä ei ole väliä. Kannan tapauksessa kuitenkin ajatellaan, että kannan alkoilla on tietty järjestys, jotta koordinaateista puhuminen on mahdollista.

Video ”[Change of basis](#)” käsittelee vektorien esittämistä eri kantojen suhteen. Videon alku käsittelee tämän alaluvun asioita, ja loppupuolella käsitellään kurssin ulkopuolelle jäävää aihetta, kannanvaihtomatriiseja.

Tiivistelmä

- Avaruuden kanta on vektorijoukko, jonka vektorit on lineaarisesti riippumattomia ja virittävät avaruuden.
- Jokainen avaruuden vektori voidaan kirjoittaa täsmälleen yhdellä tavalla kanta-vektoreiden lineaarikombinaationa. Tätä ominaisuutta voi käyttää, kun todistaa vektorijonon kannaksi.
- Avaruuden ulottuvuus eli dimensio on sen kantavektoreiden lukumäärä.
- Suoran dimensio on yksi ja tason kaksi.
- Vektorin $\mathbf{u}$ koordinaatit jonkin kannan suhteen ovat ne skalaarikertoimet, joilla kannan vektoreista voi muodostaa vektorin $\mathbf{u}$.

5.6 Sarake- ja nolla-avaruuden kanta ja dimensio

[Aiemmin](#) määriteltiin matriisin sarake- ja nolla-avaruudet. Matriisin A sarakeavaruus on joukko

$$\mathcal{R}(A) = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = A\mathbf{x} \text{ jollakin } \mathbf{x} \in \mathbb{R}^n\}$$

ja nolla-avaruus joukko

$$\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

Tähän mennessä sarake- ja nolla-avaruutta on sovellettu yhtälönratkaisussa. Aliavaruuksien, kantojen ja dimensioiden teoria tarjoaa uutta tietoa sarake- ja nolla-avaruuksista ja sitä kautta yhtälönratkaisusta.

Aiemmin todettiin, että matriisin A sarakeavaruus koostuu kaikista sarakkeiden $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$

lineaarikombinaatioista. Tästä seuraa, että sarakeavaruus on sarakkeiden virittämä aliavaruus eli

$$\mathcal{R}(A) = \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}.$$

Lauseen 5.3.9 perusteella saadaan seuraava tulos.

Lause 5.6.1

Olkoon A $m \times n$ -matriisi. Tällöin sarakeavaruudelle $\mathcal{R}(A)$ pätevät seuraavat ehdot:

1. $\mathbf{v} + \mathbf{w} \in \mathcal{R}(A)$ kaikilla $\mathbf{v}, \mathbf{w} \in \mathcal{R}(A)$
2. $c\mathbf{v} \in \mathcal{R}(A)$ kaikilla $c \in \mathbb{R}$ ja $\mathbf{v} \in \mathcal{R}(A)$
3. $\mathbf{0} \in \mathcal{R}(A)$.

Toisin sanoen $\mathcal{R}(A)$ on avaruuden $\mathbb{R}^m$ aliavaruus. Sama tulos voidaan osoittaa nolla-avaruudelle.

Lause 5.6.2

Olkoon A $m \times n$ -matriisi. Tällöin nolla-avaruudelle $\mathcal{N}(A)$ pätevät seuraavat ehdot:

1. $\mathbf{v} + \mathbf{w} \in \mathcal{N}(A)$ kaikilla $\mathbf{v}, \mathbf{w} \in \mathcal{N}(A)$
2. $c\mathbf{v} \in \mathcal{N}(A)$ kaikilla $c \in \mathbb{R}$ ja $\mathbf{v} \in \mathcal{N}(A)$
3. $\mathbf{0} \in \mathcal{N}(A)$.

Todistus. Todistus jätetään harjoitustehtäväksi. □

Lineaaristen yhtälöryhmien tapauksessa todettiin, että niillä on joko täsmälleen yksi, yhtään tai äärettömästi ratkaisuja. Lineaarinen yhtälöryhmä voidaan muuttaa matriisiyhtälöksi. Tällöin väite voidaan todistaa nolla-avaruuden käsitteen avulla.

Lause 5.6.3

Olkoon A $m \times n$ -matriisi, sekä $\mathbf{b}$ avaruuden $\mathbb{R}^m$ vektori. Tällöin täsmälleen yksi seuraavista väitteistä on voimassa yhtälölle $A\mathbf{x} = \mathbf{b}$.

1. Yhtälöllä ei ole ratkaisuja.
2. Yhtälöllä on täsmälleen yksi ratkaisu.
3. Yhtälöllä on äärettömän monta ratkaisua.

Todistus. Riittää osoittaa, että ratkaisujen löytyessä niitä on joko yksi tai ääretön määrä.

Oletetaan, että $\mathbf{x}_0$ on eräs yhtälön ratkaisu. Nyt on kaksi vaihtoehtoa: joko matriisin $\mathcal{N}(A) = \{\mathbf{0}\}$ tai nolla-avaruudessa on muitakin vektoreita kuin nollavektori. Ensimmäisessä tapauksessa homogeenisen yhtälön $A\mathbf{x} = \mathbf{0}$ ainoa ratkaisu on $\mathbf{0}$, joten yhtälön $A\mathbf{x} = \mathbf{b}$

ratkaisuksi saadaan $\mathbf{x}_0 + \mathbf{0} = \mathbf{x}_0$. Yhtälöllä on siis yksikäsitteinen ratkaisu. Jälkimmäisessä tapauksessa nolla-avaruudesta löydetään vektori $\mathbf{y} \neq \mathbf{0}$. Koska $\mathcal{N}(A)$ on aliavaruus, pätee $r\mathbf{y} \in \mathcal{N}(A)$ kaikilla $r \in \mathbb{R}$. Täten jokainen vektori $\mathbf{x}_0 + r\mathbf{y}$, missä $r \in \mathbb{R}$, on yhtälön $A\mathbf{x} = \mathbf{b}$ ratkaisu. Yhtälöllä on siis äärettömän monta ratkaisua. $\square$

Tämän luvun lopussa osoitetaan, että sarake- ja nolla-avaruuksien dimensiot liittyvät toisiinsa. Tämän todistamiseen tarvitaan tietoa sarake- ja nolla-avaruuksien kannoista.

Seuraava lause osoittaa, että matriisiin redusoitu porrasmuoto kertoo, mitkä matriisin sarakkeista muodostavat sarake-avaruuden kannan. Todistus mukailee aiempaa [esimerkkiä 5.5.7](#), jossa itse asiassa tultiin etsineeksi kanta sarakeavaruudelle.

Lause 5.6.4

Eräs matriisin A sarakeavaruuden $\mathcal{R}(A)$ kanta muodostuu niistä matriisin A sarakkeista, joissa on johtavat alkioit redusoidussa porrasmuodossa.

Todistus. Olkoot $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ matriisin A ne sarakevektorit, joiden kohdille johtavat ykköset sijoittuvat redusoidussa porrasmuodossa $\text{rref}(A)$. Osoitetaan, että ne muodostavat sarakeavaruuden kannan. On siis osoitettava, että vektorit ovat lineaarisesti riippumattomia ja virittävät sarakeavaruuden.

Lineaarisen riippumattomuuden osoittamiseksi tutkitaan yhtälöryhmää

$$\left[\begin{array}{cccc|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k & \mathbf{0} \end{array} \right].$$

Koska redusoidussa porrasmuodossa vektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ kohdille sijoittuvat ykköset, on jokaisessa sarakkeessa johtava alkio. Siten yhtälöryhmällä on täsmälleen yksi ratkaisu. Tästä seuraa, että vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomat.

Osoitetaan sitten, että vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ virittävät sarakeavaruuden. Oletetaan, että $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_l$ ovat loput matriisin A sarakkeista. Osoitetaan, että ne ovat vektoreiden $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ lineaarikombinaatioita. Yhtälöryhmällä

$$\left[\begin{array}{cccc|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k & \mathbf{u}_j \end{array} \right]$$

on ratkaisu kaikilla $j \in \{1, 2, \dots, l\}$ aivan kuten edelläkin. Näin ollen [lauseen 5.3.7](#) nojalla

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{u}_1\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}.$$

Samaan tapaan jatkamalla saadaan

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_l\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}.$$

Koska $\mathcal{R}(A) = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_l\}$, on osoitettu, että vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ virittävät sarakeavaruuden.

Näin ollen vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ muodostavat sarakeavaruuden kannan. □

Seuraava esimerkki osoittaa, kuinka nolla-avaruuden kannan voi löytää.

Esimerkki 5.6.5

Etsitään kanta matriisiin

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

nolla-avaruudelle. On siis ratkaistava yhtälöryhmä $A\mathbf{x} = \mathbf{0}$. Kirjoitetaan yhtälöryhmä kokonaismatriisina

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

Kun matriisia muokataan alkeisrivimuunnoksilla, saadaan matriisi

$$\left[\begin{array}{cccc|c} 1 & 0 & \frac{3}{2} & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{4} & -\frac{1}{4} & 0 \end{array} \right].$$

Tästä nähdään, että yhtälöryhmän ratkaisu on

$$\begin{cases} x_1 = -(3/2)t - (1/2)s \\ x_2 = (1/4)t + (1/4)s \\ x_3 = t \\ x_4 = s, \end{cases} \quad \text{missä } t, s \in \mathbb{R}.$$

Vektorimuodossa kirjoitettuna ratkaisu on

$$\mathbf{x} = t \begin{bmatrix} -\frac{3}{2} \\ -\frac{1}{4} \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 0 \\ 1 \end{bmatrix},$$

missä $t, s \in \mathbb{R}$. Huomaa, että ratkaisussa näkyviä vektoreita on yhtä monta kuin vapaita muuttujia.

Nyt tiedetään, että

$$\mathcal{N}(A) = \text{span}\{(-3/2, -1/4, 1, 0), (-1/2, 1/4, 0, 1)\}.$$

Koska vektorit $(-3/2, -1/4, 1, 0)$ ja $(-1/2, 1/4, 0, 1)$ ovat lineaarisesti riippumattomia, ne muodostavat nolla-avaruuden kannan. Huomaa, että lineaarinen riippumattomuus seuraa siitä, että kussakin vektorissa on yksi, ja muissa vektoreissa samassa

kohtaa on nolla.

Esimerkissä esitetty strategia pätee yleisesti mille tahansa matriisille. Tulosta ei kuitenkaan todisteta täsmällisesti tässä materiaalissa teknisyytensä vuoksi.

Tiivistettynä matriisin A nolla- ja sarakeavaruuksien kannat voidaan löytää seuraavasti:

1. Etsi matriisin A redusoitu porrasmuoto $\text{rref}(A)$.
2. Valitse matriisista A ne sarakkeet, joiden kohdalle matriisissa $\text{rref}(A)$ sijoittuu johtava alkio. Nämä pystyvektorit muodostavat sarakeavaruuden $\mathcal{R}(A)$ kannan.
3. Etsi redusoidun porrasmuodon avulla homogeenisen yhtälön $A\mathbf{x} = \mathbf{0}$ ratkaisu ja esitä se vektoreiden lineaarikombinaationa. Lineaarikombinaation vektorit muodostavat nolla-avaruuden $\mathcal{N}(A)$ kannan.

Esimerkki 5.6.6

Tutkitaan matriisia $A = \begin{bmatrix} 1 & 2 & 4 & 2 \\ 2 & 3 & 1 & -1 \\ -1 & -1 & 3 & 3 \end{bmatrix}$ ja etsitään avaruuksille $\mathcal{R}(A)$ ja $\mathcal{N}(A)$ kannat.

Matriisin A redusoitu porrasmuoto on

$$\begin{bmatrix} 1 & 0 & -10 & -8 \\ 0 & 1 & 7 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Koska johtavat alkiot ovat kahdessa ensimmäisessä sarakkeessa, eräs sarakeavaruuden $\mathcal{R}(A)$ kanta on

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} \right\}.$$

Redusoidusta porrasmuodosta nähdään myös, että homogeenisen yhtälöryhmän $A\mathbf{x} = \mathbf{0}$ ratkaisu on

$$\begin{cases} x_1 = 10t + 8s \\ x_2 = -7t - 5s \\ x_3 = t \\ x_4 = s, \end{cases}$$

missä $t, s \in \mathbb{R}$. Vektorimuodossa kirjoitettuna ratkaisu on

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} 10 \\ -7 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix},$$

joten eräs nolla-avaruuden $\mathcal{N}(A)$ kanta on

$$\left\{ \begin{bmatrix} 10 \\ -7 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Matriisin aste kertoo, kuinka monta johtavaa alkioita on matriisin redusoidussa porrasmuodossa. Tämän vuoksi matriisin aste on sama kuin sarakeavaruuden dimensio.

Lause 5.6.7

Olkoon A matriisi. Matriisin aste $\text{rank}(A)$ on sarakeavaruuden dimensio. Toisin sanoen $\text{rank}(A) = \dim(\mathcal{R}(A))$.

Todistus. Matriisin A sarakeavaruuden kanta saadaan valitsemalla matriisista A ne sarakkeet, joita redusoidussa porrasmuodossa vastaa johtava alkio. Kannassa on siis yhtä monta vektoria kuin redusoidussa porrasmuodossa johtavia alkioita. Avaruuden dimensio on kannan vektorien lukumäärä ja matriisin aste puolestaan on johtavien alkioiden lukumäärä redusoidussa porrasmuodossa. Siten sarakeavaruuden dimensio ja matriisin aste ovat samat. $\square$

Esimerkissä 5.6.6 sarakeavaruuden kannassa oli kaksi vektoria, joten $\dim \mathcal{R}(A) = 2$. Siis $\text{rank}(A) = 2$. Toisaalta esimerkin perusteella $\dim \mathcal{N}(A) = 2$. Koska sarakeavaruuden kantaan valitaan ne sarakkeet, joiden kohdalle tulee vapaa muuttuja, ja nolla-avaruuden kantaan puolestaan valikoituu yhtä monta vektoria kuin on vapaita muuttujia, on näiden kahden kannan vektorien yhteenlaskettu lukumäärä sama kuin matriisin sarakkeiden lukumäärä.

Edellinen huomio on ilmaistu matriisien *dimensiolauseessa*, joka liittyy yhteen sarakeavaruuden dimension eli matriisin asteen ja nolla-avaruuden dimension.

Lause 5.6.8 (Dimensiolause)

Jos A on $m \times n$ -matriisi, niin $\dim(\mathcal{N}(A)) + \text{rank}(A) = n$.

Nyt voidaan laajentaa kääntyvien matriisien lausetta 4.6.6 uusilla kohdilla.

Lause 5.6.9 (Laajennettu kääntyvien matriisien lause)

Oletetaan, että A on $n \times n$ -neliömatriisi. Seuraavat ehdot ovat yhtäpitäviä:

1. Matriisi A on kääntyvä.
2. Yhtälöllä $A\mathbf{x} = \mathbf{b}$ on täsmälleen yksi ratkaisu kaikilla $\mathbf{b} \in \mathbb{R}^n$.
3. Yhtälöllä $A\mathbf{x} = \mathbf{0}$ on vain triviaali ratkaisu $\mathbf{x} = \mathbf{0}$.
4. Matriisi A on riviekvivalentti ykkösmatriisin kanssa.
5. Matriisi A on alkeismatriisien tulo.
6. Matriisi A ei ole riviekvivalentti minkään nollarivin sisältävän matriisin kanssa.
7. $\text{rank}(A) = n$.
8. Matriisin A nolla-avaruuden dimensio on nolla, eli $\mathcal{N}(A) = \{\mathbf{0}\}$.
9. Matriisin A sarakkeet ovat lineaarisesti riippumattomia.

Todistus. Ensimmäiset kuusi kohtaa on jo osoitettu yhtäpitäviksi, joten riittää osoittaa, että loput kohdat ovat näiden kanssa yhtäpitäviä. Tehdään tämä todistamalla ekvivalenssi $7 \Leftrightarrow 8$ ja implikaatioketju

$$4 \Rightarrow 7 \Rightarrow 9 \Rightarrow 4.$$

$7 \Leftrightarrow 8$: **Dimensiolauseen 5.6.8** nojalla $\text{rank}(A) + \dim \mathcal{N}(A) = n$, joten $\text{rank}(A) = n$ jos ja vain jos $\dim \mathcal{N}(A) = 0$. Kohdat 7 ja 8 ovat siis yhtäpitävät.

$4 \Rightarrow 7$: Oletetaan, että A on riviekvivalentti ykkösmatriisin kanssa. Tällöin $\text{rref}(A) = I$. Koska matriisissa $\text{rref}(A)$ on n johtavaa alkioita, on **lauseen 5.6.4** perusteella matriisin A sarakeavaruuden dimensio n . Näin ollen **lauseen 5.6.7** nojalla $\text{rank}(A) = n$.

$7 \Rightarrow 9$: Oletetaan, että $\text{rank}(A) = n$. Tällöin **lauseen 5.6.7** nojalla matriisin A sarakeavaruuden dimensio on n , joten sarakeavaruuden kannassa on n vektoria. Kantaan kuuluvat siis kaikki matriisin A sarakkeet. Siten matriisin A sarakkeet ovat lineaarisesti riippumattomia.

$9 \Rightarrow 4$: Oletetaan, matriisin A sarakkeet ovat lineaarisesti riippumattomia. Tällöin ne muodostavat sarakeavaruuden kannan. **Lauseen 5.6.4** perusteella jokaisessa matriisin $\text{rref}(A)$ sarakkeessa on johtava alkio. Toisaalta $\text{rref}(A)$ on neliömatriisi ja redusoitu porrasmatriisi. Ainoa vaihtoehto on, että $\text{rref}(A) = I$. Siten A on riviekvivalentti ykkösmatriisin kanssa. □

6 DETERMINANTTI JA OMINAISARVOT

6.1 Pienten matriisien determinantit

Katso video ”[Determinantti](#)”.

Neliömatriisille voidaan laskea luku, joka kertoo muun muassa, onko matriisi kääntyvä vai ei. Tätä lukua kutsutaan matriisin determinantiksi. Determinantilla on monia sovelluksia, mutta tässä yhteydessä tarkastelemme erityisesti sen yhteyttä matriisin kääntyvyyteen.

Determinantteihin tutustuttiin jo aiemmin 2×2 -matriisin determinantin [määritelmän](#) yhteydessä. Tässä luvussa määritelmä laajennetaan koskemaan kaikkia neliömatriiseja. Tarkastellaan ensin pienten matriisien determinantteja ja esitetään sitten yleinen determinantin määritelmä. Determinantti voidaan laskea vain neliömatriisille.

Palautetaan mieleen, että matriisin

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

determinantti on $\det(A) = ad - bc$. Determinanttia voi merkitä myös pystyviivojen avulla:

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Määritetään sitten 3×3 -matriisiin determinantti. Tämä tapahtuu pienempien, niin kutsuttujen alideterminanttien avulla. Matriisin

$$B = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

determinantti on

$$\det(B) = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

Kukin alideterminantti saadaan poistamalla alkuperäisestä matriisista se rivi ja sarake, jolla alideterminantin edessä oleva kerroin sijaitsee.

Esimerkki 6.1.1

Matriisin

$$N = \begin{bmatrix} 2 & 3 & 8 \\ -4 & 5 & 9 \\ 0 & -7 & -8 \end{bmatrix}$$

determinantti on

$$\begin{aligned} \det(N) &= \begin{vmatrix} 2 & 3 & 8 \\ -4 & 5 & 9 \\ 0 & -7 & -8 \end{vmatrix} = 2 \begin{vmatrix} 5 & 9 \\ -7 & -8 \end{vmatrix} - 3 \begin{vmatrix} -4 & 9 \\ 0 & -8 \end{vmatrix} + 8 \begin{vmatrix} -4 & 5 \\ 0 & -7 \end{vmatrix} \\ &= 2 \cdot (5 \cdot (-8) - 9 \cdot (-7)) - 3 \cdot (-4 \cdot (-8) - 9 \cdot 0) + 8 \cdot (-4 \cdot (-7) - 5 \cdot 0) \\ &= 174. \end{aligned}$$

Tiivistelmä

- Neliömatriisille voidaan laskea determinantti, joka on matriisin alkioden perusteella määritetty reaaliluku.

6.2 Determinantin yleinen määritelmä ja kehityskaavat

Määritellään seuraavaksi determinantti yleisesti kaikille neliömatriiseille. Determinantti voidaan määritellä monin eri tavoin. Tähän lukuun on valittu eräs melko yksinkertainen tapa. Kaikki luvussa esitellyt tulokset voitaisiin johtaa valitusta määritelmästä, mutta useimmat todistukset olisivat niin työläitä, että niistä tyydytään antamaan vain perusidea. Käytettäessä hieman kehittyneempiä määritelmiä monet todistukset helpottuisivat huomattavasti, mutta itse määritelmän esittäminen olisi työläämpää.

Yleinen määritelmä mukailee tapaa, jolla laskettiin 3×3 -matriisin determinantti.

Määritelmä 6.2.1

Olkoon A jokin $n \times n$ -matriisi. Merkitään

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Jos $n = 1$, niin $\det(A) = a_{11}$. Jos taas $n > 1$, niin

$$\det(A) = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det(A_{1j}),$$

missä A_{1j} on matriisi, joka on saatu matriisista A poistamalla ensimmäinen rivi ja j :s sarake.

Määritelmässä olevat kertoimet a_{1j} otetaan matriisin ensimmäiseltä riviltä. Sanotaan, että determinantti on tällöin *kehitetty ensimmäisen rivin suhteen*. Yhtä hyvin voidaan käyttää muita rivejä tai jopa muita sarakkeita.

Lause 6.2.2

Oletetaan, että $A \in \mathbb{R}^{n \times n}$, ja merkitään $A(i, j) = a_{ij}$ kaikilla $i, j \in \{1, \dots, n\}$.

1. Olkoon $i \in \{1, \dots, n\}$. Tällöin

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}),$$

missä A_{ij} on matriisi, joka on saatu matriisista A poistamalla i :s rivi ja j :s sarake. Kyseessä on kehitys rivin i suhteen.

2. Olkoon $j \in \{1, \dots, n\}$. Tällöin

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}),$$

missä A_{ij} on matriisi, joka on saatu matriisista A poistamalla i :s rivi ja j :s sarake. Kyseessä on kehitys sarakkeen j suhteen.

Todistus. Hahmotellaan todistuksen ideaa. Lause voidaan todistaa tarkastelemalla, millaiseen lausekkeeseen [määritelmän 6.2.1](#) kehityskaava lopulta johtaa. Syntyvä lauseke on summa tuloista

$$\pm a_{1k_1} a_{2k_2} \cdots a_{nk_n},$$

missä $k_1, \dots, k_n$ ovat sarakkeiden indeksit jossakin järjestyksessä. Esimerkiksi tyypin 3×3 matriisin determinantin laskeminen johtaa näillä merkinnöillä lausekkeeseen

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Kunkin termin etumerkki määräytyy siitä, onko sarakkeiden järjestys $k_1, \dots, k_n$ alkupe-
räisen järjestyksen $1, \dots, n$ niin sanottu parillinen vai pariton permutaatio. Tämän ha-
vainnon jälkeen on suoraviivaista tarkistaa, että jokainen kehityskaava johtaa itse asiassa
täsmälleen samaan lausekkeeseen. $\square$

Kehityskaavojen etumerkkien vaihtelu (eli kaavoissa muotoa $(-1)^{i+j}$ oleva kerroin) saadaan shakkilautaa muistuttavasta kuviosta

$$\begin{bmatrix} + & - & + & - & \dots \\ - & + & - & + & \\ + & - & + & - & \\ \vdots & & & & \end{bmatrix}.$$

Matriisin tilalle ajatellaan plus- ja miinusmerkeistä koostuva ruudukko, jonka vasemmassa yläkulmassa on plusmerkki. Jos matriisin alkion kohdalla on plusmerkki, tulee kehityskaavassa alkion eteen plusmerkki. Vastaavasti jos alkion kohdalla on miinusmerkki, tulee kehityskaavaankin miinusmerkki. Kunkin alkion omaa etumerkkiä ei myöskään sovi unohtaa.

Esimerkki 6.2.3

Toisinaan voi säästää vaivaa, jos valitsee viisaasti rivin tai sarakkeen, jonka suhteen determinantin kehittää. Lasketaan matriisin

$$D = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 0 \\ 0 & 0 & -1 & 0 \\ 8 & 9 & -4 & -2 \end{bmatrix}$$

determinantti kehittämällä se aluksi kolmannen rivin suhteen:

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 0 \\ 0 & 0 & -1 & 0 \\ 8 & 9 & -4 & -2 \end{vmatrix} &= 0 \cdot \begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 0 \\ 9 & -4 & -2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 5 & 7 & 0 \\ 8 & -4 & -2 \end{vmatrix} \\ &\quad + (-1) \cdot \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 0 \\ 8 & 9 & -2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 8 & 9 & -4 \end{vmatrix} \\ &= - \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 0 \\ 8 & 9 & -2 \end{vmatrix} \end{aligned}$$

Saatu 3×3 -determinantti voidaan kehittää kolmannen sarakkeen suhteen:

$$\begin{aligned}
 - \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 0 \\ 8 & 9 & -2 \end{vmatrix} &= - \left(4 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 2 \\ 8 & 9 \end{vmatrix} + (-2) \cdot \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} \right) \\
 &= -(4 \cdot (45 - 48) - 0 - 2 \cdot (6 - 10)) = -(-12 - 0 + 8) = -(-4) = 4.
 \end{aligned}$$

Siis $\det(D) = 4$.

6.3 Determinantin ominaisuuksia

Kuten [lauseessa 3.6.9](#) todettiin, determinantin merkitys näkyy siinä, että se kertoo matriisin kääntyvyydestä.

Lause 6.3.1

Oletetaan, että A on $n \times n$ -matriisi. Matriisi A on kääntyvä, jos ja vain jos $\det(A) \neq 0$.

Lauseen todistus on esitetty tämän luvun [loppupuolella](#).

Esimerkki 6.3.2

[Esimerkissä 6.1.1](#) saatiin matriisiin

$$N = \begin{bmatrix} 2 & 3 & 8 \\ -4 & 5 & 9 \\ 0 & -7 & -8 \end{bmatrix}$$

determinantiksi 174. Koska determinantti ei ole nolla, matriisi N on kääntyvä.

Esimerkki 6.3.3

Tutkitaan, onko vektorijono $((2, 1, -1), (0, 1, -3), (-2, 1, -5))$ avaruuden $\mathbb{R}^3$ kanta. Kyseessä on kanta, mikäli jokainen vektori $\mathbf{w} \in \mathbb{R}^3$ voidaan ilmaista yksikäsitteisesti annettujen vektorien lineaarikombinaationa. Tutkittava yhtälöryhmä voidaan ilmaista matriisimuodossa yhtälönä $A\mathbf{x} = \mathbf{w}$, missä

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

Yhtälöryhmällä on yksikäsitteinen ratkaisu, jos ja vain jos kerroinmatriisi A on kään-

tyvä (lause 4.6.6). Toisaalta lauseen 6.3.1 nojalla A on kääntövä, jos ja vain jos sen determinantti on nollasta poikkeava. Lasketaan determinantti:

$$\begin{aligned}\det(A) &= \begin{vmatrix} 2 & 0 & -2 \\ 1 & 1 & 1 \\ -1 & -3 & -5 \end{vmatrix} = 2(1 \cdot (-5) - 1 \cdot (-3)) - 0 - 2(1 \cdot (-3) - 1 \cdot (-1)) \\ &= 2 \cdot (-2) - 2 \cdot (-2) = 0.\end{aligned}$$

Koska determinantti on 0, matriisi A ei ole kääntövä. Tästä syystä tutkittavan yhtälöryhmän ratkaisua ei ole olemassa tai se ei ole yksikäsitteinen. Annettu vektorijono ei siis muodosta kantaa.

Determinantin merkitystä voidaan tulkita visuaalisesti kuvausten avulla. Jos matriisia ajatellaan kuvauksena, determinantti kertoo, kuinka paljon kuvaus muuttaa kuvioiden pinta-aloja. Jos matriisilla kuvaa tason yksikköneliötä, tuloksena on suunnikas. Seuraavan lauseen mukaan determinantin itseisarvo on tuon suunnikkaan pinta-ala.

Lause 6.3.4

Oletetaan, että $A \in \mathbb{R}^{2 \times 2}$. Tällöin $|\det(A)|$ on vektoreiden $A\mathbf{e}_1$ ja $A\mathbf{e}_2$ määrittämän suunnikkaan pinta-ala.

Todistus. Oletetaan, että

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

missä $a, b, c, d \in \mathbb{R}$. Palautetaan mieleen, että $\mathbf{e}_1 = (1, 0)$ ja $\mathbf{e}_2 = (0, 1)$. Matriisikertolaskun määritelmän mukaan $A\mathbf{e}_1 = (a, c)$ ja $A\mathbf{e}_2 = (b, d)$. Ajatellaan tasoa $\mathbb{R}^2$ tason $\mathbb{R}^3$ osajoukkona siten, että tason $\mathbb{R}^2$ vektori (a, c) vastaa vektoria $(a, c, 0)$ ja (b, d) vektoria $(b, d, 0)$. Nyt voidaan hyödyntää ristituloa suunnikkaan pinta-alan laskemiseen. Suunnikkaan pinta-ala on ristituloa käsittelevän luvun perusteella

$$\begin{aligned}\|(a, c, 0) \times (b, d, 0)\| &= \|(c \cdot 0 - 0 \cdot d, 0 \cdot b - a \cdot 0, ad - cb)\| \\ &= \|(0, 0, ad - cb)\| \\ &= \sqrt{0^2 + 0^2 + (ad - cb)^2} \\ &= |ad - cb| = |\det(A)|.\end{aligned}$$

□

Edellisen lauseen avulla voi hahmottaa uudella tavalla, miksi determinantti liittyy matriisin kääntyvyyteen. Jos determinantti on nolla, matriisi kuvaa tason yksikköneliön suunnikkaaksi, jonka pinta-ala on nolla. Tuloksena on siis vain viiva. Tällaisessa kuvauksessa

eri vektorit kuvautuvat samaksi vektoriksi, joten kuvausta ei ole mahdollista peruuttaa. Kuvauksella ei siis ole käänteiskuvausta eikä matriisilla käänteismatriisiä.

Kolmiulotteisessa avaruudessa yksikkökuutio kuvautuu matriisilla kuvattaessa suuntaissärmiöksi. Tuon suuntaissärmiön tilavuus on determinantin itseisarvo.

Lause 6.3.5

Oletetaan, että $A \in \mathbb{R}^{3 \times 3}$. Tällöin $|\det(A)|$ on vektoreiden $A\mathbf{e}_1$, $A\mathbf{e}_2$ ja $A\mathbf{e}_3$ määrittämän suuntaissärmiön tilavuus.

Todistus. Todistus jätetään harjoitustehtäväksi. □

Tarkastellaan seuraavaksi, miten determinantti suhtautuu matriisien alkeisrivimuunnoksiin sekä laskutoimituksiin. Seuraavan tuloksen voi todistaa pienille matriiseille suoraan laskemalla, ja yleisen tapauksen voi johtaa determinanttien kehityskaavoista.

Lause 6.3.6

Oletetaan, että A on neliömatriisi. Jos matriisi B saadaan matriisista A

1. vaihtamalla kaksi riviä keskenään, niin $\det(B) = -\det(A)$.
2. kertomalla jokin rivi nollasta poikkeavalla reaaliluvulla t , niin $\det(B) = t \det(A)$.
3. lisäämällä johonkin riviin jokin toinen rivi reaaliluvulla k kerrottuna, niin $\det(B) = \det(A)$.

Koska determinantit voidaan kehittää yhtä hyvin sarakkeiden kuin rivienkin suhteen, transponointi ei vaikuta matriisin determinanttiin:

Lause 6.3.7

Oletetaan, että A on neliömatriisi. Tällöin $\det(A^T) = \det(A)$.

Lauseesta 6.3.7 seuraa, että determinantin sarakkeet käyttäytyvät täsmälleen samalla tavalla kuin sen rivit. Lauseesta 6.3.6 saadaan siis seuraavat muistisäännöt:

1. Jos matriisin kaksi riviä (tai saraketta) vaihtaa keskenään, determinantin etumerkki muuttuu:

$$\begin{vmatrix} 3 & 2 & 7 \\ 1 & 6 & 0 \\ 5 & 8 & 4 \end{vmatrix} = - \begin{vmatrix} 1 & 6 & 0 \\ 3 & 2 & 7 \\ 5 & 8 & 4 \end{vmatrix}.$$

2. Jos matriisin rivillä (tai sarakkeessa) kaikilla alkioilla on yhteinen nollasta poikkeava tekijä, tuon yhteisen tekijän voi ottaa determinantin eteen kertoimeksi:

$$\begin{vmatrix} 1 & 6 & 0 \\ 3 & 2 & 7 \\ 5 & 8 & 4 \end{vmatrix} = 2 \cdot \begin{vmatrix} 1 & 3 & 0 \\ 3 & 1 & 7 \\ 5 & 4 & 4 \end{vmatrix}.$$

3. Jos matriisin riviin (tai sarakkeeseen) lisätään jokin toinen rivi (tai sarake) vakiolla kerrottuna, matriisin determinantti ei muutu:

$$\begin{vmatrix} 1 & 3 & 0 \\ 3 & 1 & 7 \\ 5 & 4 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 0 \\ 0 & -8 & 7 \\ 5 & 4 & 4 \end{vmatrix}.$$

Lauseesta 6.3.6 seuraa myös, että joidenkin matriisien determinantti on helppo määrittää.

Lause 6.3.8

Oletetaan, että A on neliömatriisi. Tällöin

1. jos matriisissa A on nollarivi (nollasarake), niin $\det(A) = 0$
2. jos matriisissa A on kaksi samaa riviä (samaa saraketta), niin $\det(A) = 0$
3. jos A on kolmiomatriisi eli kaikki lävistäjän alapuoliset tai yläpuoliset alkio ovat nollia, niin matriisin A determinantti on lävistäjäalkioiden tulo.

Todistus. Todistetaan vain rivejä koskevat väitteet. Sarakkeita koskevat väitteet voidaan käsitellä samalla tavalla.

1. Kerrotaan matriisin nollarivi luvulla -1 , jolloin matriisi ei muutu. Lauseen 6.3.6 mukaan pätee siis $\det(A) = -1 \cdot \det(A)$, josta seuraa, että $\det(A) = 0$.
2. Vaihdetaan matriisin samanlaiset rivit keskenään, jolloin matriisi ei muutu. Lauseen 6.3.6 mukaan $\det(A) = -\det(A)$, joten $\det(A) = 0$.
3. Tulos nähdään suoraan kehittämällä matriisi rivi riviltä alkaen ylimmästä tai alimmasta rivistä, jolla on vain yksi nollasta poikkeava alkio.

□

Edeltäviä lauseita voidaan käyttää hyväksi determinantin laskemisessa. Kun matriisi muutetaan porrasmatriisiksi alkeisrivimuunnosten avulla, determinantti muuttuu lauseessa 6.3.6 kuvatulla tavalla. Porrasmatriisin determinantti voidaan puolestaan määrittää lauseen 6.3.8 avulla.

Esimerkki 6.3.9

Lasketaan matriisin

$$\begin{bmatrix} 1 & -2 & 0 & 1 \\ -2 & 4 & 1 & 0 \\ 2 & 0 & -2 & 1 \\ 3 & 2 & -1 & -5 \end{bmatrix}$$

determinantti muuntamalla se vaiheittain porrasmatriisiksi. Tarvittavat alkeisrivimuunnokset ovat seuraavat:

$$\begin{aligned} &\begin{bmatrix} 1 & -2 & 0 & 1 \\ -2 & 4 & 1 & 0 \\ 2 & 0 & -2 & 1 \\ 3 & 2 & -1 & -5 \end{bmatrix} \xrightarrow{R_2+2R_1} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 2 & 0 & -2 & 1 \\ 3 & 2 & -1 & -5 \end{bmatrix} \xrightarrow{R_3-2R_1} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 3 & 2 & -1 & -5 \end{bmatrix} \xrightarrow{R_4-3R_1} \\ &\begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 0 & 8 & -1 & -8 \end{bmatrix} \xrightarrow{R_4-2R_3} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 0 & 0 & 3 & -6 \end{bmatrix} \xrightarrow{R_4-3R_2} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 0 & 0 & 0 & -12 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \\ &\begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 4 & -2 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -12 \end{bmatrix}. \end{aligned}$$

Tuloksena on yläkolmiomatriisi, jonka determinantti on lävistäjäalkioiden tulo eli tässä tapauksessa -48 . [Lauseen 6.3.6](#) perusteella ainoastaan viimeinen alkeisrivimuunnos muutti matriisin determinanttia, ja sekin aiheutti vain etumerkin muutoksen. Siispä alkuperäisen matriisin determinantti oli 48 .

Edellisten tulosten avulla voidaan nyt todistaa myös kääntyvän matriisin determinanttiin liittyvä [lause 6.3.1](#). Tehdään se tarkastelemalla alkeisrivimuunnoksia.

Todistus. Osoitetaan, että neliömatriisi A on kääntyvä, jos ja vain jos $\det(A) \neq 0$. Tiedetään, että matriisi A on kääntyvä täsmälleen silloin, kun se voidaan muuttaa alkeisrivimuunnoksilla ykkösmatriisiksi, muuten porrasmatriisiin tulee nollarivi ([lause 4.6.6](#)). [Lauseen 6.3.8](#) nojalla ykkösmatriisin determinantti on 1 ja nollarivin omaavan matriisin determinantti on 0 . Toisaalta [lauseesta 6.3.6](#) seuraa, että jokainen alkeisrivimuunnos säilyttää determinantin nollana tai nollasta poikkeavana sen mukaan, mitä se oli alun perin. Täten matriisi A on kääntyvä täsmälleen silloin, kun sen determinantti on nollasta

poikkeava. □

Tarkastellaan vielä matriisien laskutoimituksien vaikutusta determinanttiiin.

Lause 6.3.10

Oletetaan, että A ja B ovat neliömatriiseja. Tällöin $\det(AB) = \det(A) \det(B)$.

Todistus. Hahmotellaan todistuksen ideaa. Jos jompikumpi tai molemmat matriiseista A ja B eivät ole kääntyviä, myöskään niiden tulo ei ole kääntyvä. Tällöin väite pätee [lauseen 6.3.1](#) nojalla. Oletetaan sitten, että A ja B ovat kääntyviä ja kirjoitetaan ne alkeismatriisien tuloina:

$$A = E_1 \cdots E_r \quad \text{ja} \quad B = F_1 \cdots F_s.$$

[Lauseesta 6.3.6](#) seuraa, että jos E on alkeismatriisi, jokaiselle neliömatriisille M pätee

$$\det(EM) = \det(E) \det(M).$$

Käyttämällä tätä havaintoa toistuvasti yllä esitettyihin tuloihin, nähdään, että

$$\det(A) = \det(E_1) \cdots \det(E_r) \quad \text{ja} \quad \det(B) = \det(F_1) \cdots \det(F_s).$$

Toisaalta $AB = E_1 \cdots E_r F_1 \cdots F_s$, ja samalla tavoin kuin edellä saadaan

$$\det(AB) = \det(E_1) \cdots \det(E_r) \cdot \det(F_1) \cdots \det(F_s).$$

Väite seuraa tästä. □

Lause 6.3.11

Oletetaan, että neliömatriisi A on kääntyvä. Tällöin

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Todistus. Oletuksen mukaan matriisi A on kääntyvä, joten sillä on käänteismatriisi A^{-1} . Edellisen lauseen nojalla pätee

$$\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1.$$

Toisaalta [lauseen 6.3.1](#) mukaan $\det(A) \neq 0$, sillä A on kääntyvä. Jakamalla puolittain saadaan $\det(A^{-1}) = 1/\det(A)$. □

Tiivistelmä

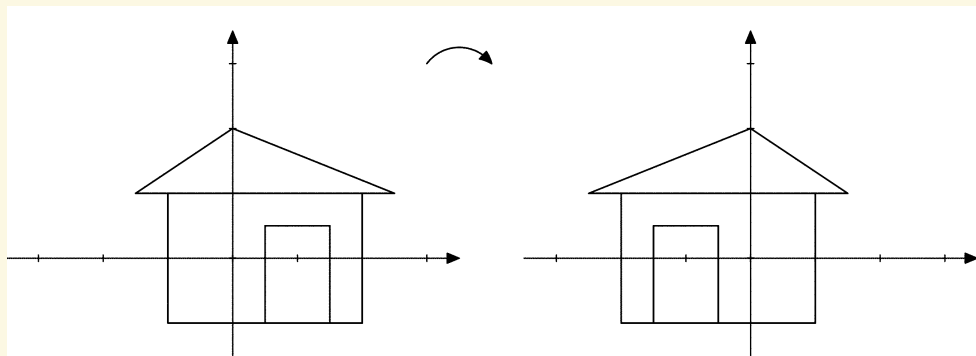
- Matriisi on kääntyvä täsmälleen siinä tapauksessa, että sen determinantti ei ole nolla.
- Determinantille pätee erilaisia laskusääntöjä, joiden avulla determinanttien laskemista voidaan helpottaa.
- Esimerkiksi tulon determinantti on determinanttien tulo ja käänteismatriisin determinantti on determinantin käänteisluku.

6.4 Ominaisarvon määritelmä

Voit ryhtyä tutustumaan ominaisarvioihin ja -vektoreihin katsomalla videon ”[Ominaisvektorit ja ominaisarvot](#)”.

Pohdi 6.4.1

Kun eräällä 2×2 -matriisilla B kerrotaan tason $\mathbb{R}^2$ vektoreita, ne peilautuvat vaak akselin suhteen (kuva 6.4.1).



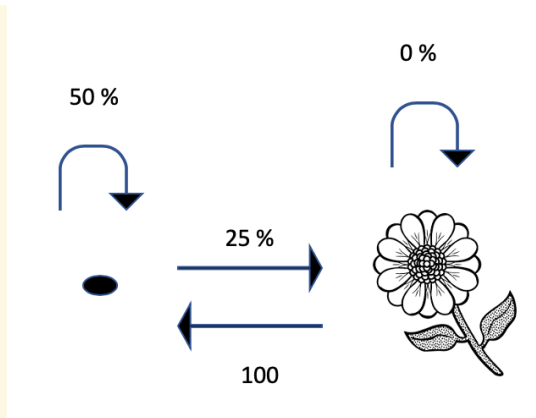
Kuva 6.4.1. Matriisi B peilaa vektorit vaak akselin suhteen.

Pohdi kuvan avulla seuraavia kysymyksiä:

1. Mitkä vektorit pysyvät paikoillaan (eli tulevat kerrotuiksi skalaarilla 1), kun niitä kerrotaan matriisilla B ?
2. Mitkä vektorit kuvautuvat vastavektoreikseen (eli tulevat kerrotuiksi skalaarilla -1), kun niitä kerrotaan matriisilla B ?

Pohdi 6.4.2

Eräs kukkakasvi lisääntyy siementen avulla. Jokainen kukinto tuottaa noin 100 siementä. Seuraavana vuonna 50 % siemenistä itää kukintokasveiksi ja 25 % säilyy siemenä seuraavaan vuoteen. Nämä itämättömät siemenet voivat itää vielä seuraavana vuonna. Kukintoja ei säily seuraavalle vuodelle.



Kuva 6.4.2. Kasvin vuosittaista elinkiertoa kuvaava kaavio.

Kasvipopulaation tilaa jonakin tiettyinä vuonna voidaan kuvata vektorilla $\mathbf{x} = (s, k)$, missä s on siementen lukumäärä ja k on kukintojen lukumäärä. Esimerkiksi tila $(3000, 40)$ tarkoittaa, että populaatiossa on 3000 siementä ja 40 kukintoa. Toisaalta populaation vuosittaista muutosta voidaan kuvata matriisilla

$$A = \begin{bmatrix} 0,25 & 100 \\ 0,5 & 0 \end{bmatrix}.$$

Jos jonakin vuonna populaation tila on $\mathbf{x} = (s, k)$, seuraavana vuonna se on

$$A\mathbf{x} = \begin{bmatrix} 0,25 & 100 \\ 0,5 & 0 \end{bmatrix} \begin{bmatrix} s \\ k \end{bmatrix} = \begin{bmatrix} 0,25s + 100k \\ 0,5s \end{bmatrix}.$$

Seuraavan vuoden tila saadaan siis kertomalla matriisilla A .

1. Halutaan löytää niin sanottu tasapainotila, josta populaation tila ei muutu seuraavana vuonna. Toisin sanoen on löydettävä sellainen tila (s, k) että myös seuraavana vuonna populaation tila on (s, k) . Millaista yhtälöä on ryhdyttävä ratkomaan? (Yhtälöä ei tarvitse ratkaista.)
2. Halutaan löytää tila, josta lähdettäessä populaation koko kaksinkertaistuu. Toisin sanoen on löydettävä sellainen tila (s, k) että myös seuraavana vuonna sekä siemeniä että kukintoja on kaksinkertainen määrä. Millaista yhtälöä on ryhdyttävä ratkomaan? (Yhtälöä ei tarvitse ratkaista.)

Matriisin ja vektorin kertolasku on melko monimutkainen operaatio, mutta toisinaan se sievenee hyvin helppoon muotoon. Tutkitaan vaikkapa matriisia

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

ja vektoria $(1, 2)$. Niiden tulo on

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}.$$

Toisin sanoen $A\mathbf{v} = 5\mathbf{v}$. Matriisilla A kertominen vastaa vektorin $(1, 2)$ tapauksessa skalaarilla viisi kertomista. Sanotaan, että matriisilla A on ominaisarvo 5, johon liittyy ominaisvektori $(1, 2)$.

Kaikilla vektoreilla kertolasku ei kuitenkaan saa näin yksinkertaista muotoa. Esimerkiksi

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

joten vektori $(1, -1)$ ei tule kerrotuksi luvulla 5 eikä millään muullakaan reaalityluvulla.

Matriisin ominaisarvoista puhutaan siis silloin, kun matriisilla kertominen vaikuttaa johonkin vektoriin samalla tavalla kuin skalaarilla kertominen. Tuo vektori on silloin matriisin ominaisvektori ja vastaava skalaari on matriisin ominaisarvo.

Määritelmä 6.4.3

Oletetaan, että A on $n \times n$ -neliömatriisi. Luku $\lambda \in \mathbb{R}$ on matriisin A *ominaisarvo*, jos on olemassa sellainen vektori $\mathbf{v} \in \mathbb{R}^n$, että

$$\mathbf{v} \neq \mathbf{0} \text{ ja } A\mathbf{v} = \lambda\mathbf{v}.$$

Vektoria $\mathbf{v}$, joka toteuttaa yllä mainitut ehdot, kutsutaan ominaisarvoon λ liittyväksi *ominaisvektoriksi*.

Edellinen määritelmä on sekä ominaisarvon että ominaisvektorin määritelmä. Ominaisarvoa ei voida määritellä ilman ominaisvektoreita eikä ominaisvektoreista voida puhua mainitsematta, mihin ominaisarvoon ne liittyvät.

Nollavektorin ei haluta olevan ominaisvektori, sillä jos niin olisi, kaikki reaalityluvut olisivat kaikkien matriisien ominaisarvoja, koska $A\mathbf{0} = \lambda\mathbf{0}$ kaikilla $\lambda \in \mathbb{R}$.

Esimerkki 6.4.4

Matriisilla

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

on ominaisarvo 4, johon liittyy ominaisvektori $\mathbf{v}_1 = (1, 1)$. Tämä nähdään laskemalla

matriisin A ja vektorin $\mathbf{v}_1$ tulo:

$$A\mathbf{v}_1 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4\mathbf{v}_1.$$

Tilannetta on havainnollistettu kuvassa 6.4.3.

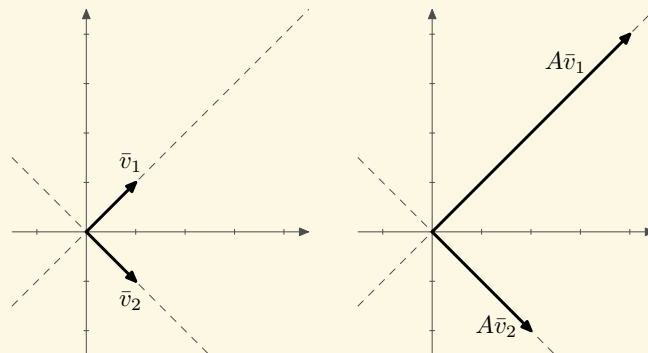
Samaa ominaisarvoa voi vastata useampi eri ominaisvektori. Esimerkiksi $3\mathbf{v}_1 = (3, 3)$ on myös matriisin A ominaisarvoa 4 vastaava ominaisvektori, sillä

$$A(3\mathbf{v}_1) = 3(A\mathbf{v}_1) = 3(4\mathbf{v}_1) = 12\mathbf{v}_1 = 4(3\mathbf{v}_1).$$

Matriisilla A on ominaisarvon 4 lisäksi toinenkin ominaisarvo. Jos nimittäin valitaan $\mathbf{v}_2 = (1, -1)$, saadaan

$$A\mathbf{v}_2 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2\mathbf{v}_2.$$

Siten myös luku 2 on matriisin A ominaisarvo ja $\mathbf{v}_2 = (1, -1)$ on yksi siihen liittyvä ominaisvektori. (Matriisin ominaisarvot opetellaan etsimään [seuraavassa luvussa](#)).



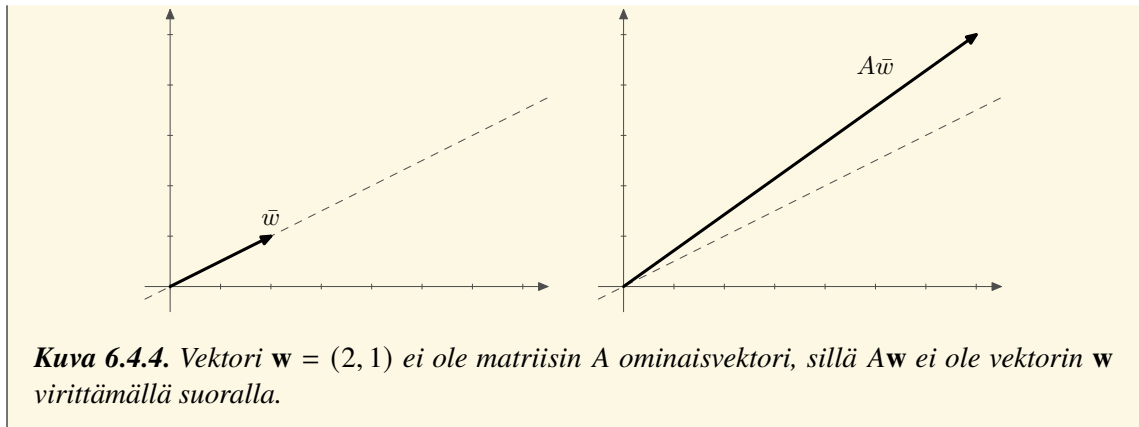
Kuva 6.4.3. Vektori $\mathbf{v}_1 = (1, 1)$ on matriisin A ominaisvektori, sillä $A\mathbf{v}_1 = 4\mathbf{v}_1$ on vektorin $\mathbf{v}_1$ virittämällä suoralla. Samoin vektori $\mathbf{v}_2 = (1, -1)$ on matriisin A ominaisvektori, sillä $A\mathbf{v}_2 = 2\mathbf{v}_2$ on vektorin $\mathbf{v}_2$ virittämällä suoralla.

Kun matriisilla kertoo ominaisvektoria $\mathbf{v}$, tuloksena on vektorin $\mathbf{v}$ skalaarimonikerta. Toisin sanoen tulos on vektorin $\mathbf{v}$ virittämällä suoralla (ks. kuva 6.4.3).

Tutkitaan vielä lopuksi, onko vektori $\mathbf{w} = (2, 1)$ matriisin A ominaisvektori.

$$A\mathbf{w} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}.$$

Nähdään, että $A\mathbf{w}$ ei ole vektorin $\mathbf{w}$ skalaarimonikerta, joten $\mathbf{w}$ ei ole matriisin A ominaisvektori. Tätä on havainnollistettu kuvassa 6.4.4.



Kuten edellinen esimerkki osoittaa, matriisilla voi olla useampi kuin yksi ominaisarvo. Kuhunkin ominaisarvoon liittyy useita ominaisvektoreita. Kaikkien tietyä ominaisarvoa vastaavien ominaisvektorien joukko muodostaa niin kutsutun ominaisvaruuden. Siihen otetaan mukaan myös nollavektori, vaikka se ei olekaan ominaisvektori.

Määritelmä 6.4.5

Oletetaan, että matriisilla $A \in \mathbb{R}^{n \times n}$ on ominaisarvo $\lambda \in \mathbb{R}$. Ominaisarvoa λ vastaava ominaisvaruus on joukko

$$\{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}.$$

Seuraava esimerkki näyttää, miten tiettyyn ominaisarvoon liittyvä ominaisvaruus eli kaikki ominaisarvoa vastaavat ominaisvektorit löydetään.

Esimerkki 6.4.6

Jatketaan edellistä esimerkkiä ja etsitään kaikki matriisin A ominaisarvoa 4 vastaavat ominaisvektorit. Määritetään siis ominaisarvoa 4 vastaava ominaisvaruus.

On ratkaistava yhtälöstä $A\mathbf{v} = 4\mathbf{v}$ tuntematon $\mathbf{v}$. Yhtälö saadaan muotoon

$$A\mathbf{v} - 4\mathbf{v} = \mathbf{0}.$$

Tästä yhtälöstä haluttaisiin nyt ottaa yhteiseksi tekijäksi $\mathbf{v}$, mutta se ei onnistu, sillä A on matriisi ja 4 on reaaliluku, eikä niitä voi vähentää toisistaan. Huomataan kuitenkin, että skalaarimatriisilla $4I$ kertominen vaikuttaa vektoriin $\mathbf{v}$ samalla tavalla kuin luvulla 4 kertominen:

$$4I\mathbf{v} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 4v_1 + 0 \\ 0 + 4v_2 \end{bmatrix} = \begin{bmatrix} 4v_1 \\ 4v_2 \end{bmatrix} = 4\mathbf{v}.$$

Nyt yhtälö saadaan muotoon $A\mathbf{v} - 4I\mathbf{v} = \mathbf{0}$, josta seuraa

$$(A - 4I)\mathbf{v} = \mathbf{0}.$$

Sijoitetaan yhtälöön matriisi A :

$$\left(\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \right) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Nyt yhtälö sievenee muotoon

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Päädetään siis ratkaisemaan yhtälöryhmä

$$\begin{cases} -v_1 + v_2 = 0 \\ v_1 - v_2 = 0 \end{cases}$$

Muutetaan yhtälöryhmän matriisi porrasmuotoon:

$$\left[\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array} \right] \xrightarrow{R_2+R_1} \left[\begin{array}{cc|c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right] \xrightarrow{(-1) \cdot R_1} \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array} \right].$$

Merkitään $v_2 = t$. Tällöin $v_1 = v_2 = t$. Siten yhtälön ratkaisu on

$$\begin{cases} v_1 = t \\ v_2 = t, \end{cases} \text{ missä } t \in \mathbb{R}.$$

Ominaisvektorit ovat siis muotoa (t, t) , missä $t \in \mathbb{R}$. Siten ominaisarvoa 4 vastaava ominaisavaruus on

$$\{(t, t) \mid t \in \mathbb{R}\}$$

Ominaisavaruus on mahdollista kirjoittaa myös muodossa

$$\{t(1, 1) \mid t \in \mathbb{R}\} = \text{span}\{(1, 1)\}.$$

Tästä nähdään, että kyseessä vektorin $(1, 1)$ virittämä aliavaruus eli origon kautta kulkeva suora.

Ominaisarvon määritelmästä seuraa, että kun ominaisavaruudessa olevia vektoreita kertoo matriisilla A , ne skaalautuvat nelinkertaisiksi.

Jos matriisille A löytyy yksikin ominaisvektori, sillä on välttämättä äärettömän monta ominaisvektoria. Jokainen ominaisvektorin $\mathbf{v}$ skalaarimonikerta nollavektoria lukuunottamatta on nimittäin myös ominaisvektori, sillä $A(c\mathbf{v}) = c(A\mathbf{v}) = c(\lambda\mathbf{v}) = \lambda(c\mathbf{v})$ kaikilla $c \in \mathbb{R}$.

Ominaisarvoa vastaavien ominaisvektorien ei kuitenkaan tarvitse kaikkien olla toistensa skalaarimonikertoja kuten seuraava esimerkki osoittaa.

Esimerkki 6.4.7

Tutkitaan matriisia

$$A = \begin{bmatrix} 7 & 8 & -2 \\ -3 & -3 & 1 \\ 9 & 12 & -2 \end{bmatrix}.$$

Tällä matriisilla on ominaisarvo 1, jota vastaa ominaisvektori $(1, 0, 3)$, sillä

$$\begin{bmatrix} 7 & 8 & -2 \\ -3 & -3 & 1 \\ 9 & 12 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}.$$

Toisaalta

$$\begin{bmatrix} 7 & 8 & -2 \\ -3 & -3 & 1 \\ 9 & 12 & -2 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}.$$

joten myös $(-4, 3, 0)$ on ominaisarvoa 1 vastaava ominaisvektori. Vektorit $(1, 0, 3)$ ja $(-4, 3, 0)$ eivät kuitenkaan ole toistensa skalaarimonikertoja. Samaa ominaisarvoa vastaavien ominaisvektorien ei siis tarvitse olla yhdensuuntaisia.

Tutkitaan vielä tarkemmin, miltä ominaisarvoa 1 vastaava ominaisavaruuks näyttää. Laskemalla samaan tapaan kuin esimerkissä 6.4.6 saadaan ominaisvektoria 1 vastaavaksi ominaisavaruuksiksi

$$\{(s - 4t, 3t, 3s) \mid s, t \in \mathbb{R}\}.$$

Tämä joukko voidaan kirjoittaa muodossa

$$\begin{aligned} \{(s - 4t, 3t, 3s) \mid s, t \in \mathbb{R}\} &= \{(s, 0, 3s) + (-4t, 3t, 0) \mid s, t \in \mathbb{R}\} \\ &= \{s(1, 0, 3) + t(-4, 3, 0) \mid s, t \in \mathbb{R}\} \\ &= \text{span}\{(1, 0, 3), (-4, 3, 0)\}. \end{aligned}$$

Ominaisvektorit muodostavat siis tason, joka kulkee origon kautta.

Kun ominaisavaruuksissa olevia vektoreita kertoo matriisilla A ne tulevat kerrotuiksi

skalaarilla 1. Vektoreille ei siis tapahdu mitään.

Tiivistelmä

- Matriisin ominaisarvot on mahdollista löytää determinantin avulla.

6.5 Ominaisarvojen löytäminen

Ominaisarvojen ja ominaisvektoreiden löytäminen perustuu yhtälön

$$A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0} \quad (6.5.1)$$

ratkaisemiseen. Ominaisvektoria $\mathbf{v}$ ei kuitenkaan voida ratkaista ennen kuin tunnetaan ominaisarvo λ . Sen löytämiseksi muutetaan yhtälö hieman toiseen muotoon samaan tapaan kuin [esimerkissä 6.4.6](#). Ensinnäkin huomataan, että $\lambda\mathbf{v} = \lambda I\mathbf{v}$, missä I on yksikkömatriisi. Näin ollen

$$A\mathbf{v} - \lambda\mathbf{v} = A\mathbf{v} - \lambda I\mathbf{v} = (A - \lambda I)\mathbf{v}.$$

Nyt yhtälö (6.5.1) tulee muotoon

$$(A - \lambda I)\mathbf{v} = \mathbf{0}. \quad (6.5.2)$$

Yhtälöä (6.5.2) vastaa homogeeninen yhtälöryhmä, joten sillä on aina triviaaliratkaisu $\mathbf{v} = \mathbf{0}$. Tämä ei kuitenkaan kelpaa ominaisvektoriksi, joten tavoitteena on löytää jokin epätriviaali ratkaisu. [Lauseen 4.6.6](#) nojalla yhtälöllä on epätriviaaleja ratkaisuja täsmälleen silloin, kun kerroinmatriisi $A - \lambda I$ ei ole kääntyvä. Toisaalta [lauseen 6.3.1](#) mukaan neliömatriisi ei ole kääntyvä täsmälleen silloin, kun sen determinantti on 0. Näin saadaan seuraava lause.

Lause 6.5.1

Reaaliluku λ on neliömatriisin A ominaisarvo, jos ja vain jos

$$\det(A - \lambda I) = 0.$$

Lauseke $\det(A - \lambda I)$ on eräs muuttujan λ polynomi. Sitä nimitetään matriisin A *karaktelistiseksi polynomiksi*. Edellinen lause voidaan siis muotoilla myös niin, että matriisin A ominaisarvot ovat sen karakteristisen polynomin nollakohdat.

Esimerkki 6.5.2

Määritetään matriisin

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

ominaisarvot ja niitä vastaavat ominaisavaruudet. Lähdetään liikkeelle laskemalla lauseessa 6.5.1 mainittu determinantti:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1 - \lambda & 2 \\ 3 & 2 - \lambda \end{vmatrix} = (1 - \lambda)(2 - \lambda) - 6 \\ &= 2 - \lambda - 2\lambda + \lambda^2 - 6 = \lambda^2 - 3\lambda - 4. \end{aligned}$$

Matriisin A ominaisarvot ovat lauseen 6.5.1 nojalla yhtälön $\lambda^2 - 3\lambda - 4 = 0$ ratkaisut. Toisen asteen yhtälön ratkaisukaavan mukaan tarkasteltava yhtälö toteutuu, jos ja vain jos $\lambda = 4$ tai $\lambda = -1$. Siten matriisin A ominaisarvot ovat $\lambda_1 = 4$ ja $\lambda_2 = -1$.

Määritetään vielä näihin ominaisarvoihin liittyvät ominaisavaruudet. Kumpaakin ominaisarvoa vastaavat omat ominaisvektorinsa. Tarkastellaan ensin ominaisarvoa $\lambda_1 = 4$. Tällöin ratkaistavana on yhtälö $(A - 4I)\mathbf{v} = \mathbf{0}$. Ratkaistavaksi saadaan siis yhtälöryhmä, jota vastaa matriisi

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

Kun yhtälöryhmä ratkaistaan Gaussin-Jordanin eliminointimenetelmällä, ratkaisun nähdään olevan

$$\begin{cases} x_1 = (2/3)t \\ x_2 = t, \end{cases} \quad \text{missä } t \in \mathbb{R}.$$

Ominaisarvoa 4 ominaisavaruus on siis

$$\{((2/3)t, t) \mid t \in \mathbb{R}\}.$$

Tarkastellaan sitten ominaisarvoa $\lambda_2 = -1$. Nyt ratkaistavana oleva yhtälö on $(A + I)\mathbf{v} = \mathbf{0}$. Sen ratkaisuksi saadaan

$$\begin{cases} x_1 = -t \\ x_2 = t, \end{cases} \quad \text{missä } t \in \mathbb{R}.$$

Ominaisarvoa -1 vastaava ominaisavaruus on siis

$$\{(-t, t) \mid t \in \mathbb{R}\}.$$

Edellisen nojalla matriisin $A \in \mathbb{R}^{n \times n}$ ominaisarvoa λ vastaava ominaisavaruus voidaan

kirjoittaa muodossa

$$\{\mathbf{v} \in \mathbb{R}^n \mid (A - \lambda I)\mathbf{v} = \mathbf{0}\}.$$

Tämä joukko on itse asiassa matriisin $A - \lambda I$ nolla-avaruus eli $\mathcal{N}(A - \lambda I)$. Näin ollen lauseen 5.6.2 nojalla ominaisavaruus on aliavaruus eli se sisältää kaikkien vektoreidensa summat ja skalaarimonikerrat.

Matriisilla voi olla vain äärellisen monta ominaisarvoa.

Lause 6.5.3

Jos A on $n \times n$ -matriisi, sillä on korkeintaan n ominaisarvoa.

Todistus. Koska A on $n \times n$ -matriisi, sen karakteristinen polynomi on korkeintaan astetta n . Karakteristinen polynomi on siis muotoa $c_0 + c_1\lambda + \dots + c_n\lambda^n$, missä $c_0, \dots, c_n \in \mathbb{R}$. Voidaan osoittaa, että yhtälöllä

$$c_0 + c_1\lambda + \dots + c_n\lambda^n = 0$$

on enintään n eri ratkaisua. Näin ollen matriisilla A on enintään n eri ominaisarvoa. $\square$

Joidenkin matriisien ominaisarvojen löytäminen onnistuu helposti. Jos matriisi A on kolmiomatriisi eli kaikki sen lävistäjän alapuoliset tai yläpuoliset alkioit ovat nollia, niin myös $A - \lambda I$ on kolmiomatriisi. Tällöin sen determinantti $\det(A - \lambda I)$ on lävistäjäalkioiden tulo lauseen 6.3.8 nojalla. Näin ollen kolmiomatriisin A karakteristinen polynomin nollakohdat saadaan yhtälöstä

$$(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda) = 0.$$

Lauseen 6.5.1 nojalla saadaan tästä seuraava tulos.

Lause 6.5.4

Oletetaan, että neliömatriisi A on kolmiomatriisi eli että kaikki sen lävistäjän alapuoliset tai yläpuoliset alkioit ovat nollia. Tällöin matriisin A ominaisarvot ovat sen lävistäjän alkioit.

Esimerkki 6.5.5

Porrasmatriisi

$$A = \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 4 & -2 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -12 \end{bmatrix}$$

on kolmiomatriisi, joten sen ominaisarvot ovat lävistäjän alkioit. Siis matriisin A

ominaisarvot ovat $\lambda_1 = 1$, $\lambda_2 = 4$ ja $\lambda_3 = -12$.

Tiivistelmä

- Kun matriisilla kertoo ominaisvektoria, vektori tulee kerrotuksi skalaarilla. Tätä skalaaria kutsutaan ominaisarvoksi.
- Matriisin ominaisvektorit pysyvät virittämällään suoralla, kun niitä kerrotaan matriisilla.

6.6 Diagonalisointi

Edellä nähtiin, että kolmiomatriisien ja siten myös lävistäjämatriisien ominaisarvot voidaan lukea suoraan matriisista. Tässä kappaleessa tutustutaan menetelmään, jonka avulla tietyntyyppiset neliömatriisit saadaan muutettua lävistäjämatriiseiksi, joilla on samat ominaisarvot kuin alkuperäisillä matriiseilla. Matriiseja, joilla tämä menetelmä toimii, kutsutaan diagonalisoituviksi.

Määritelmä 6.6.1

Neliömatriisi $A \in \mathbb{R}^{n \times n}$ on *diagonalisoituva*, jos on olemassa kääntyvä matriisi $P \in \mathbb{R}^{n \times n}$ ja lävistäjämatriisi $D \in \mathbb{R}^{n \times n}$, joille pätee

$$P^{-1}AP = D.$$

Esimerkki 6.6.2

Esimerkin 6.4.4 matriisi

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

on diagonalisoituva. Valitsemalla

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

ja etsimällä esimerkiksi lauseen 3.6.6 avulla sen käänteismatriisi

$$P^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

saadaan

$$P^{-1}AP = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 4 & 4 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 8 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}.$$

Määritelmässä 6.6.1 vaadittu lävistäjämatriisi on siis

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}.$$

Esimerkissä 6.6.2 matriisi P vain tupsahti jostakin. Vertaamalla [esimerkkiin 6.4.4](#) huomataan kuitenkin, että matriisin D lävistäjäalkiot ovat matriisin A ominaisarvot, ja matriisin P sarakkeet ovat jotkin niistä vastaavat ominaisvektorit. Seuraava lause osoittaa, että näin on aina, jos matriisi on diagonalisoituva.

Lause 6.6.3

Neliömatriisi $A \in \mathbb{R}^{n \times n}$ on diagonalisoituva, jos ja vain jos sillä on n lineaarisesti riippumattonta ominaisvektoria. Tällöin

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix},$$

missä matriisin $P \in \mathbb{R}^{n \times n}$ sarakkeet ovat matriisin A lineaarisesti riippumattomia ominaisvektoreita ja $\lambda_1, \dots, \lambda_n$ ovat niitä vastaavat ominaisarvot samassa järjestyksessä.

Todistus. " $\Rightarrow$ ": Oletetaan, että $P^{-1}AP = D$, missä $P \in \mathbb{R}^{n \times n}$ on jokin kääntyvä matriisi ja $D \in \mathbb{R}^{n \times n}$ lävistäjämatriisi. Nyt $AP = PD$. Olkoot matriisin P sarakkeet $\mathbf{p}_1, \dots, \mathbf{p}_n$ ja matriisin D lävistäjäalkiot $\lambda_1, \dots, \lambda_n$. Nyt siis

$$P = \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix} \quad \text{ja} \quad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

Matriisituloa laskettaessa tulon AP jokainen sarake saadaan kertomalla matriisilla A vas-

taava sarake matriisista P :

$$AP = A \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{p}_1 & \cdots & A\mathbf{p}_n \end{bmatrix}.$$

Toisaalta lävistäjämatriisia D kerrotaessa tullaan kertoneeksi matriisin P jokainen sarake vastaavalla lävistäjäalkiolla:

$$PD = \begin{bmatrix} \lambda_1 \mathbf{p}_1 & \cdots & \lambda_n \mathbf{p}_n \end{bmatrix}.$$

Koska $AP = PD$, nähdään nyt, että $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$ kaikilla $i \in \{1, \dots, n\}$. Siis jokainen λ_i on ominaisarvo ja $\mathbf{p}_i$ sitä vastaava ominaisvektori.

On vielä osoitettava, että ominaisvektorit $\mathbf{p}_1, \dots, \mathbf{p}_n$ ovat lineaarisesti riippumattomia. Koska P on kääntyvä, yhtälöllä $P\mathbf{x} = \mathbf{0}$ on [lauseen 4.5.1](#) mukaan täsmälleen yksi ratkaisu $\mathbf{x} = \mathbf{0}$. Yhtälö $P\mathbf{x} = \mathbf{0}$ voidaan kirjoittaa myös muotoon

$$x_1 \mathbf{p}_1 + x_2 \mathbf{p}_2 + \cdots + x_n \mathbf{p}_n = \mathbf{0}.$$

Tämän yhtälön ainoa ratkaisu on siis $x_1 = 0, \dots, x_n = 0$. Näin ollen matriisin A ominaisvektoreiden jono $(\mathbf{p}_1, \dots, \mathbf{p}_n)$ on lineaarisesti riippumaton.

” $\Leftarrow$ ”: Oletetaan, että $\mathbf{p}_1, \dots, \mathbf{p}_n$ ovat jotkin matriisin A lineaarisesti riippumattomat ominaisvektorit. Olkoot niitä vastaavat ominaisarvot $\lambda_1, \dots, \lambda_n$. Nyt $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$ kaikilla $i \in \{1, \dots, n\}$. Olkoon P matriisi, jonka sarakkeet ovat ominaisvektorit: $P = [\mathbf{p}_1 \cdots \mathbf{p}_n]$. Olkoon D puolestaan lävistäjämatriisi, jonka lävistäjäalkiot ovat $\lambda_1, \dots, \lambda_n$. Tällöin nähdään samaan tapaan kuin edellä, että $AP = PD$.

Koska matriisin P sarakkeet ovat lineaarisesti riippumattomat, on yhtälöllä $P\mathbf{x} = \mathbf{0}$ täsmälleen yksi ratkaisu $\mathbf{x} = \mathbf{0}$. (Tämä nähdään samalla tavalla kuin todistuksen ensimmäisessä osassa.) [Lauseen 4.6.6](#) nojalla matriisi P on nyt kääntyvä. Yhtälö $AP = PD$ saadaan siis muotoon

$$P^{-1}AP = D.$$

□

Esimerkki 6.6.4

Tutkitaan, onko [esimerkin 6.5.2](#) matriisi

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

diagonalisoituva. Esimerkissä 6.5.2 todettiin, että matriisin ominaisarvot ovat 4 ja -1 . Eräät näitä ominaisarvoja vastaavat ominaisvektorit ovat $(2, 3)$ ja $(-1, 1)$. Nämä ominaisvektorit ovat lineaarisesti riippumattomat joten [lauseen 6.6.3](#) perusteella A

on diagonalisoituva. Muodostetaan ominaisvektoreista matriisi

$$P = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$$

ja ominaisarvoista matriisi

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

Nyt lauseen 6.6.3 nojalla pätee $P^{-1}AP = D$. Tämän voi vielä tarkistaa laskemalla.

Esimerkki 6.6.5

Diagonalisoidaan matriisi

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

jos mahdollista. Selvitetään aluksi matriisin ominaisarvot. Koska matriisi A on kolmiomatriisi, sen ominaisarvot ovat sen lävistäjän alkiot. Näin matriisin A ainoa ominaisarvo on 2. Ominaisarvoa vastaavat ominaisvektorit saadaan yhtälöstä $A\mathbf{x} = 2\mathbf{x}$. Kun yhtälö ratkaistaan, nähdään sen ratkaisujen olevan muotoa $\mathbf{x} = (t, 0)$, missä $t \in \mathbb{R} \setminus \{0\}$. Matriisilla A ei siis ole kahta lineaarisesti riippumatonta ominaisvektoria, joten A ei ole diagonalisoituva.

Esimerkki 6.6.6 (Diagonalisoituvan matriisin potenssit)

Lasketaan [esimerkissä 6.6.2](#) esiintyneen matriisin

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

seitsemäs potenssi. Suora matriisikertolasku olisi työläs suorittaa, mutta koska matriisi A on diagonalisoituva, voidaan käyttää hyväksi sen ominaisarvoja.

Esimerkissä 6.6.2 todettiin, että $P^{-1}AP = D$, missä

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{ja} \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}.$$

Jos kerrotaan yhtälöä $P^{-1}AP = D$ vasemmalta matriisilla P ja oikealta matriisilla

P^{-1} , saadaan $A = PDP^{-1}$. Nyt voidaan laskea

$$\begin{aligned}
 A^7 &= (PDP^{-1})^7 \\
 &= \underbrace{(PDP^{-1})(PDP^{-1}) \dots (PDP^{-1})(PDP^{-1})}_{7 \text{ kpl}} \\
 &= PD(P^{-1}P)D \dots (P^{-1}P)DP^{-1} \\
 &= P \underbrace{D \dots D}_{7 \text{ kpl}} P^{-1} \\
 &= PD^7P^{-1}.
 \end{aligned}$$

Matriisin A potenssin määrittäminen on siis muuttunut matriisin D potenssin määrittämiseksi. Se osoittautuu helposksi. Huomataan nimittäin, että

$$D^7 = \begin{bmatrix} 4^7 & 0 \\ 0 & 2^7 \end{bmatrix} = \begin{bmatrix} 16384 & 0 \\ 0 & 128 \end{bmatrix}.$$

Lävistäjämatrisiin potenssi saadaan itse asiassa aina selville laskemalla pelkät lävistäjäalkioiden potenssit.

Nyt

$$\begin{aligned}
 A^7 &= (PDP^{-1})^7 = PD^7P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 16384 & 0 \\ 0 & 128 \end{bmatrix} \left(\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \\
 &= \frac{1}{2} \begin{bmatrix} 16384 & 128 \\ 16384 & -128 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 16512 & 16256 \\ 16256 & 16512 \end{bmatrix} = \begin{bmatrix} 8256 & 8128 \\ 8128 & 8256 \end{bmatrix}.
 \end{aligned}$$

Matriisipotenssin laskeminen saatiin siis muutettua pariksi matriisikertolaskuksi sekä tavallisten kokonaislukujen potenssiksi. Samalla vaivalla voitaisiin laskea paljon suurempiakin potensseja. Tämä temppu onnistuu kuitenkin vain, jos alkuperäinen matriisi on diagonalisoituva.

Palataan vielä tutkimaan matriisin ominaisvektoreita. Seuraava lause osoittaa, että eri ominaisarvoja vastaavat ominaisvektorit ovat lineaarisesti riippumattomia. Tästä tuloksesta on toisinaan hyötyä, kun tutkitaan, onko matriisi diagonalisoituva.

Lause 6.6.7

Oletetaan, että A on $n \times n$ -matriisi. Oletetaan, että $\lambda_1, \dots, \lambda_m$ ovat matriisin A eri ominaisarvoja ja $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^n$ jotkin niitä vastaavat ominaisvektorit. Tällöin vektorit $\mathbf{v}_1, \dots, \mathbf{v}_m$ ovat lineaarisesti riippumattomia.

Todistus. Oletetaan vastoin väitettä, että vektorit $\mathbf{v}_1, \dots, \mathbf{v}_m$ ovat lineaarisesti riippuvia. Nyt lauseen 5.4.9 nojalla jokin jonon vektoreista on muiden lineaarikombinaatio. Tästä seuraa, että jokin jonon vektoreista on sitä edeltävien jonon vektoreiden lineaarikombinaatio. Olkoon $\mathbf{v}_{k+1}$ jonon ensimmäinen vektori, joka on sitä edeltävien vektoreiden lineaarikombinaatio. Tällöin on olemassa reaalityyppiset $c_1, \dots, c_k$, joille pätee

$$c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k = \mathbf{v}_{k+1}. \quad (6.6.1)$$

Lisäksi vektorit $\mathbf{v}_1, \dots, \mathbf{v}_k$ ovat lineaarisesti riippumattomia. Jos ne nimittäin olisivat lineaarisesti riippuvia, $\mathbf{v}_{k+1}$ ei olisikaan ensimmäinen vektori, joka on sitä edeltävien vektoreiden lineaarikombinaatio.

Kertomalla yhtälön (6.6.1) molemmat puolet vasemmalta matriisilla A saadaan yhtälö

$$A(c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k) = A \mathbf{v}_{k+1}.$$

Matriisien laskusääntöjen avulla yhtälö saa muodon $c_1 A \mathbf{v}_1 + \dots + c_k A \mathbf{v}_k = A \mathbf{v}_{k+1}$. Kun vielä muistetaan, että vektorit $\mathbf{v}_1, \dots, \mathbf{v}_k$ ovat matriisin A ominaisvektoreita, saadaan lopulta yhtälö

$$c_1 \lambda_1 \mathbf{v}_1 + \dots + c_k \lambda_k \mathbf{v}_k = \lambda_{k+1} \mathbf{v}_{k+1}. \quad (6.6.2)$$

Toisaalta voidaan kertoa yhtälön (6.6.1) molemmat puolet luvulla λ_{k+1} päätyen yhtälöön

$$c_1 \lambda_{k+1} \mathbf{v}_1 + \dots + c_k \lambda_{k+1} \mathbf{v}_k = \lambda_{k+1} \mathbf{v}_{k+1}. \quad (6.6.3)$$

Vähennetään yhtälöstä (6.6.2) puolittain yhtälö (6.6.3), jolloin saadaan

$$c_1 (\lambda_1 - \lambda_{k+1}) \mathbf{v}_1 + \dots + c_k (\lambda_k - \lambda_{k+1}) \mathbf{v}_k = \mathbf{0}.$$

Jono $(\mathbf{v}_1, \dots, \mathbf{v}_k)$ on lineaarisesti riippumaton, joten kaikkien yhtälössä olevien kertoimien on oltava nollia:

$$c_1 (\lambda_1 - \lambda_{k+1}) = 0, \quad c_2 (\lambda_2 - \lambda_{k+1}) = 0, \quad \dots, \quad c_k (\lambda_k - \lambda_{k+1}) = 0.$$

Koska $\lambda_1, \dots, \lambda_m$ ovat kaikki eri ominaisarvoja, niin tiedetään, että $(\lambda_i - \lambda_{k+1}) \neq 0$ kaikilla $i \in \{1, \dots, k\}$. Tulon nollasäännön nojalla

$$c_1 = 0, \quad c_2 = 0, \quad \dots, \quad c_k = 0.$$

Näin ollen

$$\mathbf{v}_{k+1} = c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k = 0 \mathbf{v}_1 + \dots + 0 \mathbf{v}_k = \mathbf{0}.$$

Toisaalta oletuksen mukaan $\mathbf{v}_{k+1}$ on matriisin A ominaisvektori, joten $\mathbf{v}_{k+1} \neq \mathbf{0}$. Koska päädyttiin ristiriitaan, vastaoletus ei voi olla tosi. Siis alkuperäinen väite pätee, eli vektorit $\mathbf{v}_1, \dots, \mathbf{v}_m$ ovat lineaarisesti riippumattomia. $\square$

Edellisestä lauseesta seuraa, että toisinaan matriisin diagonalisoituvuus on helppo todeta.

Seuraus 6.6.8

Oletetaan, että $n \times n$ -matriisilla on n eri ominaisarvoa. Tällöin A on diagonalisoituva.

Todistus. Olkoot $\mathbf{v}_1, \dots, \mathbf{v}_n$ jotkin eri ominaisarvoihin liittyvät ominaisvektorit. Ne ovat lineaarisesti riippumattomia lauseen 6.6.7 nojalla. Koska matriisilla A on n lineaarisesti riippumatonta ominaisvektoria, on A diagonalisoituva lauseen 6.6.3 nojalla. $\square$

Huomaa, että diagonalisoituvan $n \times n$ -matriisin ominaisarvojen lukumäärän ei tarvitse olla n . Esimerkiksi lävistäjämatriisi

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

on diagonalisoituva, sillä $I^{-1}AI = A$. Lävistäjä-matriisi on kolmiomatriisi, joten sen ominaisarvot voidaan lukea suoraan lävistäjältä. Havaitaan, että matriisilla A on vain yksi ominaisarvo, -3 .

Tiivistelmä

- Jotkin matriisit on mahdollista muuttaa matriisikertolaskun avulla lävistäjä- eli diagonaalimatriiseiksi.
- Tällaisia matriiseja kutsutaan diagonalisoituviksi ja prosessia kutsutaan diagonalisoinniksi.
- Koska lävistäjämatriiseja on helppo käsitellä, voidaan diagonalisoinnin avulla helpottaa laskuja kuten potenssiin korotusta.

7 SOVELLUS: PIENIMMÄN NELIÖSUMMAN MENETELMÄ

7.1 Kun ratkaisut puuttuvat

Palataan hetkeksi ratkaisemaan yhtälöryhmää $A\mathbf{x} = \mathbf{b}$, missä A on matriisi ja $\mathbf{b}$ tunnettu vektori. Sovelluksissa tulee usein vastaan ongelmia, missä tällä yhtälöryhmällä ei yksinkertaisesti ole ratkaisua. Yksi yleinen syy on se, että reaali maailman mittaukset eivät ole virheettömiä ja tämän vuoksi tuottavat ristiriitoja käytössä olevaan malliin. Tällöin yhtälöryhmälle $A\mathbf{x} = \mathbf{b}$ voidaan tarkan ratkaisun sijaan etsiä ”paras mahdollinen” ratkaisu, jota kutsutaan *pienimmän neliösumman ratkaisuksi*. Tässä osiossa johdetaan keino määrittää pienimmän neliösumman ratkaisu annetulle matriisiyhtälölle, sekä sovelletaan menetelmää polynomisovituksen etsimiseen.

Ydinsisältö

- **Pienimmän neliösumman ratkaisu** ja sen etsiminen **normaaliryhmän** avulla
- Moore-Penrosen **pseudoinverssi** ja pienimmän neliösumman ratkaisun **olemassaolo**
- Polynomifunktion **sovittaminen** pisteistöön

7.2 Pienimmän neliösumman menetelmä

Määritelmä 7.2.1

Olkoon A jokin $m \times n$ -matriisi ja $\mathbf{b}$ avaruuden $\mathbb{R}^m$ vektori. Tällöin yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ *pienimmän neliösumman ratkaisu* (*least squares solution*) on avaruuden $\mathbb{R}^n$ vektori $\bar{\mathbf{x}}$, joka toteuttaa epäyhtälön

$$\|\mathbf{b} - A\bar{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

aina, kun $\mathbf{x} \in \mathbb{R}^n$.

Jos yhtälöryhmällä $A\mathbf{x} = \mathbf{b}$ ei ole ratkaisua, niin $\mathbf{b} \notin \mathcal{R}(A)$. Pienimmän neliösumman ratkaisu on sellainen vektori $\bar{\mathbf{x}}$, että vektori $A\bar{\mathbf{x}} \in \mathcal{R}(A)$ on ikään kuin lähimpänä vektoria $\mathbf{b}$. Nyt kun periaate on selvä, tarvitsee enää tietää, miten $\bar{\mathbf{x}}$ lasketaan.

Aloitetaan pohtimalla, miten voitaisiin laskea pisteen $\mathbf{b}$ kohtisuora etäisyys tasosta S . Aiemmassa [esimerkissä 2.2.9](#) etsittiin pisteen $\mathbf{b}$ kohtisuoraa etäisyyttä tasosta S .

Ongelman ratkaisemiseksi tarvitaan jokin vektori $\mathbf{v}$ tasolta S pisteeseen $\mathbf{b}$, joka tämän jälkeen projisoidaan tason normaalille $\mathbf{n}$. Haluttu etäisyys on vektorin $\text{proj}_{\mathbf{n}}(\mathbf{v})$ normi. Toisin muotoiltuna tehtävän ratkaisu on tasolta S kohtisuorasti pisteeseen $\mathbf{b}$ kulkevan vektorin $\mathbf{b} - \text{proj}_S(\mathbf{b})$ pituus. Tämä tarkoittaa edelleen, että $\mathbf{b} - (\mathbf{b} - \text{proj}_S(\mathbf{b})) = \text{proj}_S(\mathbf{b})$ on vektoria $\mathbf{b}$ lähimmäksi sijoittuva tason S vektori.

Huomautus 7.2.2

Merkinnällä $\text{proj}_S(\mathbf{b})$ tarkoitetaan vektorin $\mathbf{b}$ projektiota aliavaruudelle S . Se voitaisiin laskea kirjoittamalla aliavaruuden yleinen vektori kantavektorien lineaarikombinaationa, ja käyttämällä sitten tavallisen vektoriprojektion kaavaa.

Korvaamalla edellisessä pohdinnassa tason S matriisin A sarakeavaruudella $\mathcal{R}(A)$ nähdään, että $\text{proj}_{\mathcal{R}(A)}(\mathbf{b})$ on lähimmäksi vektoria $\mathbf{b}$ sijoittuva aliavaruuden $\mathcal{R}(A)$ vektori. Normi $\|\mathbf{b} - \text{proj}_{\mathcal{R}(A)}(\mathbf{b})\|$ on siis pienin mahdollinen, eli $A\bar{\mathbf{x}} = \text{proj}_{\mathcal{R}(A)}(\mathbf{b})$, kun $\bar{\mathbf{x}}$ on yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ pienimmän neliösumman ratkaisu. Nyt projektion laskemiseksi havaitaan, että $\mathbf{b} - A\bar{\mathbf{x}}$ on ortogonaalinen mielivaltaista aliavaruuden $\mathcal{R}(A)$ vektoria vasten. Erityisesti siis $\mathbf{a}_j \cdot (\mathbf{b} - A\bar{\mathbf{x}}) = \mathbf{a}_j^\top (\mathbf{b} - A\bar{\mathbf{x}}) = 0$, kun $\mathbf{a}_j$ on $m \times n$ -matriisin A sarake. Keräämällä kaikki n tällaista yhtälöä yhteen nähdään, että

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^\top (\mathbf{b} - A\bar{\mathbf{x}}) \\ \mathbf{a}_2^\top (\mathbf{b} - A\bar{\mathbf{x}}) \\ \vdots \\ \mathbf{a}_n^\top (\mathbf{b} - A\bar{\mathbf{x}}) \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \\ \vdots \\ \mathbf{a}_n^\top \end{bmatrix} (\mathbf{b} - A\bar{\mathbf{x}}) = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}^\top (\mathbf{b} - A\bar{\mathbf{x}}) = A^\top (\mathbf{b} - A\bar{\mathbf{x}}).$$

Pienimmän neliösumman ratkaisu $\bar{\mathbf{x}}$ siis toteuttaa aina yhtälöryhmän $A^\top A\mathbf{x} = A^\top \mathbf{b}$.

Määritelmä 7.2.3

Yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ *normaaliryhmä* on $A^\top A\mathbf{x} = A^\top \mathbf{b}$.

Käy ilmi, että normaaliryhmällä on aina ratkaisuja, ja jokainen näistä on myös pienimmän neliösumman ratkaisu. Tämän osoittamiseksi lähestytään ongelmaa teoreettisemmin seuraavasta näkökulmasta. Jos matriisiyhtälössä $A\mathbf{x} = \mathbf{b}$ matriisi A ei ole kääntyvä, voidaan silti yrittää löytää mahdollisimman hyvin käänteismatriisin käyttäytymistä imitoiva matriisi. Tällaista matriisia kutsutaan *pseudoinverssiksi*. Seuraava tulos otetaan käyttöön todistamatta.

Lause 7.2.4 (Moore-Penrosen pseudoinverssi)

Jos A on $m \times n$ -matriisi, niin on olemassa yksikäsitteinen $n \times m$ -matriisi A^+ , joka toteuttaa ehdot

1. $AA^+A = A$,
2. $A^+AA^+ = A^+$,
3. $(AA^+)^\top = AA^+$
4. $(A^+A)^\top = A^+A$.

Matriisia A^+ sanotaan matriisin A *Moore-Penrosen pseudoinverssiksi*.

Lemma 7.2.5

Jos A on $m \times n$ -matriisi, niin $\mathcal{R}(A^\top A) = \mathcal{R}(A^\top)$.

Todistus. Oletetaan, että $\mathbf{y} \in \mathcal{R}(A^\top A)$. Tällöin $\mathbf{y} = A^\top A\mathbf{x}$ jollekin avaruuden $\mathbb{R}^n$ vektorille $\mathbf{x}$ ja $A\mathbf{x} = \mathbf{z}$ on avaruuden $\mathbb{R}^m$ vektori. Täten $\mathbf{y} = A^\top \mathbf{z}$ jollekin avaruuden $\mathbb{R}^m$ vektorille $\mathbf{z}$, eli $\mathbf{y} \in \mathcal{R}(A^\top)$.

Oletetaan sitten, että $\mathbf{y} \in \mathcal{R}(A^\top)$, eli $\mathbf{y} = A^\top \mathbf{x}$ jollekin avaruuden $\mathbb{R}^m$ vektorille $\mathbf{x}$. Si-
joitetaan matriisin A paikalle Moore-Penrosen pseudoinverssin A^+ sisältävä matriisitulo
edellisen lauseen ensimmäisen ehdon mukaisesti ja sovelletaan muitakin pseudoinverssin
ominaisuuksia.

$$\mathbf{y} = (AA^+A)^\top \mathbf{x} = A^\top (AA^+)^\top \mathbf{x} = A^\top (AA^+) \mathbf{x} = (A^\top A)A^+ \mathbf{x}$$

Tässä $A^+ \mathbf{x} = \mathbf{z}$ on avaruuden $\mathbb{R}^n$ vektori, jolle $\mathbf{y} = A^\top A\mathbf{z}$, ja täten $\mathbf{y} \in \mathcal{R}(A^\top A)$.

Edellä olevista päättelyistä seuraa, että $\mathcal{R}(A^\top A) \subseteq \mathcal{R}(A^\top)$ ja $\mathcal{R}(A^\top) \subseteq \mathcal{R}(A^\top A)$, joten
 $\mathcal{R}(A^\top A) = \mathcal{R}(A^\top)$. □

Nyt voidaan lopulta osoittaa, että yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ normaaliryhmällä $A^\top A\mathbf{x} = A^\top \mathbf{b}$
on aina vähintään yksi ratkaisu.

Lause 7.2.6

Olkoon A $m \times n$ -matriisi ja $\mathbf{b} \in \mathbb{R}^m$. Tällöin normaaliryhmällä $A^\top A\mathbf{x} = A^\top \mathbf{b}$ on
ratkaisu $\bar{\mathbf{x}}$, joka on myös yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ pienimmän neliösumman ratkaisu.

Todistus. Vektori $A^\top \mathbf{b}$ kuuluu luonnollisesti matriisin $A^\top$ sarakeavaruuteen $\mathcal{R}(A^\top)$. Edel-
lisen lemmän nojalla siis $A^\top \mathbf{b} \in \mathcal{R}(A^\top A)$, ja täten normaaliryhmällä $A^\top A\mathbf{x} = A^\top \mathbf{b}$ on
ratkaisu $\bar{\mathbf{x}}$. Olkoon nyt $\mathbf{x}$ avaruuden $\mathbb{R}^n$ vektori ja tarkastellaan normin neliötä

$$\|\mathbf{b} - A\mathbf{x}\|^2 = \|\mathbf{b} - A\bar{\mathbf{x}} + A\bar{\mathbf{x}} - A\mathbf{x}\|^2 = \|\mathbf{b} - A\bar{\mathbf{x}}\|^2 + \|A(\bar{\mathbf{x}} - \mathbf{x})\|^2 + 2A(\bar{\mathbf{x}} - \mathbf{x}) \cdot (\mathbf{b} - A\bar{\mathbf{x}}).$$

Pistetulo voidaan kirjoittaa matriisitulona, jolloin

$$\|\mathbf{b} - \mathbf{Ax}\|^2 = \|\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}\|^2 + \|\mathbf{A}(\bar{\mathbf{x}} - \mathbf{x})\|^2 + 2(\bar{\mathbf{x}} - \mathbf{x})^\top (\mathbf{A}^\top \mathbf{b} - \mathbf{A}^\top \mathbf{A}\bar{\mathbf{x}}) = \|\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}\|^2 + \|\mathbf{A}(\bar{\mathbf{x}} - \mathbf{x})\|^2,$$

sillä $\bar{\mathbf{x}}$ on normaaliryhmän ratkaisu. Täten

$$\|\mathbf{b} - \mathbf{Ax}\|^2 \geq \|\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}\|^2,$$

eli $\|\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}\| \leq \|\mathbf{b} - \mathbf{Ax}\|$, mielivaltaista vektoria $\mathbf{x}$ kohti, ja näin $\bar{\mathbf{x}}$ on yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$ pienimmän neliösumman ratkaisu. $\square$

Pienimmän neliösumman ratkaisu on siis aina olemassa, ja edellinen lause tarjoaa keinon löytää sen. Ratkaisun *residuaaliksi*, eli *virhevektoriksi* kutsutaan vektoria $\mathbf{r} = \mathbf{b} - \mathbf{A}\bar{\mathbf{x}}$, ja ratkaisun *virhe* on tämän normi $\|\mathbf{r}\| = \|\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}\|$. Jos alkuperäisellä yhtälöryhmällä on ratkaisuja, niin pienimmän neliösumman ratkaisun virhe on 0.

Huomautus 7.2.7

Pienimmän neliösumman ratkaisun olemassaolo ei tarkoita sitä, että se olisi yksikäsitteinen. On hyvinkin mahdollista, että matriisi normaaliryhmän kerroinmatriisi $\mathbf{A}^\top \mathbf{A}$ ei ole kääntyvä, ja tällöin normaaliryhmällä on ääretön määrä ratkaisuja.

Varmistetaan vielä, että pienimmän neliösumman ratkaisu toimii kuten yhtälön $\mathbf{Ax} = \mathbf{b}$ tarkkakin ratkaisu ainakin seuraavan lauseen mielessä.

Lause 7.2.8

Olkoon $\bar{\mathbf{x}}$ yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$ pienimmän neliösumman ratkaisu. Jos yhtälöryhmällä $\mathbf{Ax} = \mathbf{b}$ on ratkaisuja, niin tällöin myös $\bar{\mathbf{x}}$ on sen ratkaisu.

Todistus. Koska yhtälöryhmällä on ratkaisuja, löydetään vektori $\mathbf{x}$, jolle $\|\mathbf{b} - \mathbf{Ax}\| = 0$. Koska pienimmän neliösumman ratkaisu minimoi yhtälön vasemman puolen lausekkeen, on oltava $\|\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}\| = 0$. Ainoastaan nollavektorin normi on nolla, joten $\mathbf{b} - \mathbf{A}\bar{\mathbf{x}} = \mathbf{0}$ ja edelleen $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$. $\square$

Normaaliryhmä $\mathbf{A}^\top \mathbf{Ax} = \mathbf{A}^\top \mathbf{b}$ voidaan ratkaista **Gaussin–Jordanin menetelmällä**, kuten mikä tahansa muukin yhtälöryhmä. Erityisen kiintoisa tapaus syntyy silloin, kun ratkaisu on yksikäsitteinen, eli kerroinmatriisi $\mathbf{A}^\top \mathbf{A}$ on kääntyvä. Tällöin pienimmän neliösumman ratkaisuksi tulee

$$\bar{\mathbf{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}.$$

Moore–Penrosen pseudoinverssi sattuu tässä tapauksessa olemaan täsmälleen se matriisi, joka ”kääntää” matriisin $\mathbf{A}$: jos $\mathbf{A}^\top \mathbf{A}$ on kääntyvä, niin $\mathbf{A}^+ = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top$ ja pienimmän neliösumman ratkaisu on $\bar{\mathbf{x}} = \mathbf{A}^+ \mathbf{b}$.

Lause 7.2.9

Olkoon A $m \times n$ -matriisi. Tällöin seuraavat väitteet ovat voimassa.

1. Jos $A^T A$ on kääntyvä, niin $A^+ A = I_n$.
2. Jos A on kääntyvä, niin $A^+ = A^{-1}$.

Todistus. Jätetään harjoitustehtäväksi. □

Käänteismatriisia $(A^T A)^{-1}$ kutsutaan yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ *normalisoivaksi kertoimeksi*, ja sen olemassaololle löydetään riittävä ja välttämätön ehto.

Lause 7.2.10

Normalisoiva kerroin $(A^T A)^{-1}$ on olemassa täsmälleen silloin, kun matriisin A sarakkeet ovat lineaarisesti riippumattomat.

Todistus. Tiedetään, että $m \times n$ -matriisin A sarakkeet ovat lineaarisesti riippumattomat täsmälleen silloin, kun $\text{rank}(A) = n$. Vastaavasti $n \times n$ -matriisi $A^T A$ on kääntyvä jos ja vain jos $\text{rank}(A^T A) = n$. Mutta [lemman 7.2.5](#) nojalla $\mathcal{R}(A^T A) = \mathcal{R}(A^T)$, joten

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

ja väite on todistettu. □

Esimerkki 7.2.11

Tutki yhtälöryhmän

$$\begin{cases} x + 5y = 3 \\ 2x - 2y = 2 \\ -x + y = 5 \end{cases}$$

ratkaisujen olemassaoloa. Jos ratkaisuja ei ole, etsi residuaalin normin minimoiva approksimaatio ryhmän ratkaisulle. Jos tämä arvio on yksikäsitteinen, etsi yhtälöryhmän kerroinmatriisille Moore-Penrosen pseudoinverssi.

Ratkaisu. Viedään yhtälöryhmän kokonaismatriisi redusoituun riviporrasmuotoon.

$$[A \mid \mathbf{b}] = \left[\begin{array}{cc|c} 1 & 5 & 3 \\ 2 & -2 & 2 \\ -1 & 1 & 5 \end{array} \right] \longrightarrow \text{rref}[A \mid \mathbf{b}] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]$$

Tästä nähdään, että yhtälöryhmällä ei ole tarkkaa ratkaisua, joten etsitään pienimmän neliösumman ratkaisu. Normaaliryhmän kerroinmatriisi ja vakiovektori

$$A^T A = \begin{bmatrix} 6 & 0 \\ 0 & 30 \end{bmatrix} \quad \text{ja} \quad A^T \mathbf{b} = \begin{bmatrix} 2 \\ 16 \end{bmatrix}.$$

Matriisi $A^T A$ on selvästi kääntyvä, sillä se on kolmiomatriisi, jolla ei ole nolla-alkiota diagonaalillaan. Täten normaaliryhmällä $A^T A \mathbf{x} = A^T \mathbf{b}$ on yksikäsitteinen pienimmän neliösumman ratkaisu

$$\bar{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} = \frac{1}{180} \begin{bmatrix} 30 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 16 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{8}{15} \end{bmatrix}.$$

Ratkaisun virhe on $\|\mathbf{b} - A\bar{\mathbf{x}}\| = \sqrt{(3-3)^2 + \left(2 + \frac{2}{5}\right)^2 + \left(5 - \frac{1}{5}\right)^2} = \frac{12}{\sqrt{5}}$ ja kerroinmatriisiin A pseudoinverssiksi saadaan

$$A^+ = (A^T A)^{-1} A^T = \frac{1}{180} \begin{bmatrix} 30 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{15} & \frac{1}{30} \end{bmatrix}. \quad \diamond$$

7.3 Polynomin sovittaminen pisteistöön

Taulukkolaskentaohjelmat tarjoavat usein mahdollisuuden sovittaa lineaarinen tai korkeampiasteinen polynomimalli annettuun mittauspisteistöön. Sen tarkoituksena on antaa yksinkertainen, ikään kuin paras mahdollinen kuva pisteiden sijoittumisesta tasoon. Tarkastellaan seuraavaksi sitä, miten ohjelma tämän oikein tekee.

Oletetaan tason pisteet $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$ tunnetuiksi ja yritetään löytää polynomifunktio

$$f(x) = c_0 + c_1 x + \dots + c_{n-1} x^{n-1} + c_n x^n,$$

joka kuvaa niiden sijoittumista tasoon. Tuntemattomina ovat siis polynomin kertoimet $c_0, c_1, \dots, c_{n-1}, c_n$. Niiden määrittämiseksi kirjoitetaan yhtälöryhmä

$$\begin{bmatrix} c_0 + c_1 x_1 + \dots + c_{n-1} x_1^{n-1} + c_n x_1^n \\ c_0 + c_1 x_2 + \dots + c_{n-1} x_2^{n-1} + c_n x_2^n \\ \vdots \\ c_0 + c_1 x_m + \dots + c_{n-1} x_m^{n-1} + c_n x_m^n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} & x_1^n \\ 1 & x_2 & \dots & x_2^{n-1} & x_2^n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_m & \dots & x_m^{n-1} & x_m^n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix},$$

jolle etsitään pienimmän neliösumman ratkaisu. Yhtälöryhmässä esiintyvää $m \times (n+1)$ -

kerroinmatriisia

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} & x_m^n \end{bmatrix}$$

kutsutaan *Vandermonden matriisiksi*, ja sillä on monia mielenkiintoisia ominaisuuksia. Menemättä niihin syvemmälle todetaan, että pienimmän neliösumman menetelmä toimii käytännössä silloin, kun mittauspisteitä on enemmän kuin tuntemattomia kertoimia, eli kun $m > n + 1$, missä m on mittapisteiden lukumäärä ja n kysytyn polynomifunktion aste.

Esimerkki 7.3.1

Mittaustuloksina saatiin $(-1, 2)$, $(1, 2)$, $(3, 4)$ ja $(5, 6)$. Sovita tähän pisteistöön

1. suora $y = ax + b$,
2. paraabeli $y = ax^2 + bx + c$.

Ratkaisu. 1. Kirjoitetaan mittauspisteiden avulla yhtälöryhmä

$$\begin{cases} y(-1) = -a + b = 2 \\ y(1) = a + b = 2 \\ y(3) = 3a + b = 4 \\ y(5) = 5a + b = 6, \end{cases} \quad \text{eli} \quad V\mathbf{c} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 6 \end{bmatrix} = \mathbf{y}.$$

Normaaliryhmään liittyviksi matriiseiksi lasketaan

$$V^T V = \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix}^T \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 36 & 8 \\ 8 & 4 \end{bmatrix} \quad \text{ja} \quad V^T \mathbf{y} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix}^T \begin{bmatrix} 2 \\ 2 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 42 \\ 14 \end{bmatrix}.$$

Normaaliryhmän $V^T V \mathbf{c} = V^T \mathbf{y}$ kokonaimatriisin redusoiduksi riviporrasmuodoksi saadaan

$$\left[\begin{array}{cc|c} 1 & 0 & 0,7 \\ 0 & 1 & 2,1 \end{array} \right],$$

eli $a = 0,7$ ja $b = 2,1$. Tämän pienimmän neliösumman ratkaisun antama sovite on siis $y = 0,7x + 2,1$.

2. Kirjoitetaan mittauspisteiden avulla yhtälöryhmä

$$\begin{cases} y(-1) = a - b + c = 2 \\ y(1) = a + b + c = 2 \\ y(3) = 9a + 3b + c = 4 \\ y(5) = 25a + 5b + c = 6, \end{cases} \quad \text{eli} \quad V\mathbf{c} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 9 & 3 & 1 \\ 25 & 5 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 6 \end{bmatrix} = \mathbf{y}.$$

Normaaliryhmään liittyviksi matriiseiksi lasketaan

$$V^T V = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 9 & 3 & 1 \\ 25 & 5 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 9 & 3 & 1 \\ 25 & 5 & 1 \end{bmatrix} = \begin{bmatrix} 708 & 152 & 36 \\ 152 & 36 & 8 \\ 36 & 8 & 3 \end{bmatrix} \quad \text{ja}$$

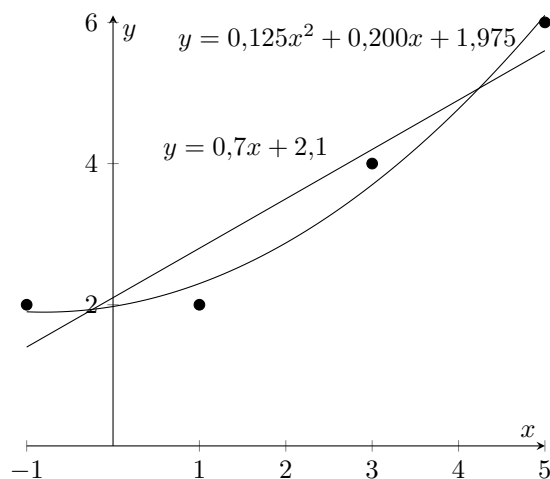
$$V^T \mathbf{y} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 9 & 3 & 1 \\ 25 & 5 & 1 \end{bmatrix}^T \begin{bmatrix} 2 \\ 2 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 190 \\ 42 \\ 14 \end{bmatrix}.$$

Normaaliryhmän $V^T V \mathbf{c} = V^T \mathbf{y}$ kokonaismatriisin redusoiduksi riviporrasmuodoksi saadaan

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0,125 \\ 0 & 1 & 0 & 0,200 \\ 0 & 0 & 1 & 1,975 \end{array} \right],$$

eli $a = 0,125$, $b = 0,200$ ja $c = 1,975$. Tämän pienimmän neliösumman ratkaisun antama sovite on siis $y = 0,125x^2 + 0,200x + 1,975$.

Piirretään lopuksi vielä kuvat näistä ensimmäisen ja toisen asteen soviteista.



Kuva 7.3.1. Ensimmäisen ja toisen asteen polynomisovitteet pisteistöön.

◇

HAKEMISTO

- A^{-1} , 68
- I , 62
- I_n , 62
- O , 62
- $O_{m \times n}$, 62
- $\mathbb{R}^n$, 3
- $\mathcal{E}_n$, 131
- aliavaruus, 116
- alkeismatriisi, 99
- alkeisrivimuunnos, 80
- alkio, matriisin, 48
- avaruus, 3
- determinantti
 - kehityskaava, 147
- diagonaali, 62
- diagonaalimatriisi, 62
- diagonalisointi, 166
- diagonalisoituva matriisi, 166
- dimensio, 136
- ekvivalentit yhtälöryhmät, 80
- eliminointimentelmä
 - Gaussin, 91
 - Gaussin-Jordanin, 84
- erotus
 - matriisien, 49
 - vektorien, 7
- etäisyys
 - pisteen suorasta, 37
 - pisteen tasosta, 45
 - vektorien välinen, 20
- Gaussin-Jordanin eliminointimenetelmä, 84
- homogeeninen yhtälöryhmä, 96
- identiteettimatriiseiksi, 62
- johtava alkio, 83
- kanta
 - aliavaruuden, 133
 - luonnollinen, 131
 - vektoriavaruuden $\mathbb{R}^n$, 131
- karakteristinen polynomi, 163
- kerroinmatriisi, 98
- kertolasku
 - matriisien, 50
- kohtisuora komponentti, 24
- kohtisuoruus, 16
- koko, matriisin, 48
- komponentti, 4
- koordinaatit, 137
- kulma, vektorien välinen, 17
- kuva-avaruus, 71
- käänteismatriisi, 66
- kääntyvä matriisi, 66
- liitännäisyys
 - vektoreille, 11
- linearikombinaatio, 8
- lineaarinen yhtälöryhmä, 75
- lineaariyhdistelmä, 8
- luonnollinen kanta, 131
- lävistäjä, 62
- lävistäjämatriisi, 62
- matriisi, 48
- matriisikertolasku, 50
- matriisipotenssi, 57
- neliömatrissi, 62
- nolla-avaruus, 72
- nollamatriisi, 62

- nollavektori, 8
- normaalivektori
 - suoran, 38
 - tason, 42
- normi, 14
- ominaisarvo, 158
- ominaisavaruus, 160
- ominaisvektori, 158
- ortogonaalinen, 16
- osittelulaki
 - vektoreille, 11
- paikkavektori
 - suoran, 33
 - tason, 40
- parametrimuoto, suoran, 34
- pistetulo, 12
- pituus, vektorin, 14
- porrasmatriisi, 83
 - redusoitu, 83
- potenssi
 - matriisien, 57
- projektio, 21
- redusoitu porrasmatriisi, 83
- ristitulo, 26
- riviekvivalentti, 82
- sarake-avaruus, 71
- sidottu, 124
- singulaarinen matriisi, 66
- skalaari, 5
- skalaarikertolasku, 5
 - matriisien, 49
- skalaarimatriisi, 62
- skalaarimonikerta
 - matriisin, 49
 - vektorin, 5
- span, 116
- summa
 - matriisien, 49
 - vektoreiden, 4
- suora, 33
- suuntavektori
 - suoran, 33
 - tason, 40
- symmetrinen, 64
- säännöllinen matriisi, 66
- taso, 40
- transpoosi, 64
- tuntematon, 76
- ulotteinen
 - vektoriavaruus, 136
- vaihdannaisuus
 - vektoreille, 11
- vapaa, 124
- vapaa muuttuja, 89
- vastavektori, 7
- vektoreiden virittämä aliavaruus, 116
- vektori, 3
- vektoriavaruus, 3
- virittäjä
 - aliavaruuden, 117
 - avaruuden $\mathbb{R}^n$, 112
- virittäminen
 - aliavaruuden, 116
 - vektoriavaruuden $\mathbb{R}^n$, 112
- ydin, 72
- yhdensuuntaisuus, 6
- yhteenlasku
 - matriisien, 49
 - vektoreiden, 4
- yhtälöryhmä, 75
- yhtäpitävät yhtälöryhmät, 80
- ykkösmatriisi, 62
- yksikkömatrisi, 62
- yksikkövektori, 15