Johdatus lineaarialgebraan
Osa I

Jokke Häsä, Lotta Oinonen, Johanna Rämö

4. syyskuuta 2017

Helsingin yliopisto
Matematiikan ja tilastotieteen laitos

Sisältö

1 Vektoriavaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ vektorit 4
1.1 Kaksiulotteisen avaruuden vektorit 4
1.2 Vektorien laskutoimitukset 6
1.3 Kolmiulotteinen vektoriavaruus 8
2 Vektoriavaruus $\mathbb{R}^n$ 9
3 Lineaariset yhtälöryhmät 13
3.1 Lineaarisen yhtälöryhmän määritelmä 13
3.2 Alkeisrivitoimitukset ja porrasmatriisit 14
3.3 Gaussin–Jordanin eliminointimenetelmä 17
3.4 Yhtälöryhmien ekvivalenssin todistus 25
4 Virittäminen 27
5 Vapaus 31
5.1 Vapauden määritelmä 31
5.2 Homogeeniset yhtälöryhmät ja vapaus 35
6 Kanta 37
6.1 Koordinaatit 40
6.2 Vektoriavaruuden $\mathbb{R}^n$ dimensio 41
7 Aliavaruudet 43
7.1 Aliavaruuden kanta 48
7.2 Aliavaruuden dimensio 50
8 Suorat ja tasot 51
8.1 Suora 51
8.2 Taso 55
9 Matriisit 58
9.1 Matriisien laskutoimituksia 58
9.2 Erityisiä matriiseja 60
9.3 Matriisien laskusääntöjä 61
9.4 Matriisin transpoosi 62
9.5 Käänteismatriisi 63
9.6 Sarakevektorit 68
10 Matriisit ja yhtälöryhmät 70
10.1 Alkeismatriisit 71
10.2 Käänteismatriisin määrittäminen - Miksi menetelmä toimii? 75
10.3 Lisää kääntyvyyteen liittyviä tuloksia 76
11 Determinantti 77
11.1 Pienten matriisien determinantit 77
11.2 Determinantin kehityskaavat 79

11.3 Determinantin ominaisuuksia . . . 81
12 Ominaisarvot ja diagonalisointi . . . 85
12.1 Ominaisarvon määritelmä . . . 85
12.2 Ominaisarvojen löytäminen . . . 89
12.3 Diagonalisointi . . . 91
13 Pistetulo . . . 96
13.1 Vektorin normi . . . 97
13.2 Vektorien kohtisuoruus ja projektio . . . 99
13.3 Vektorien välinen kulma . . . 102
13.4 Pistetulon sovelluksia . . . 104
14 Ristitulo . . . 110

Hakemisto **116**

1 Vektoriavaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ vektorit

Lukiomatematiikassa vektorit esitetään olioina, joilla on suunta ja pituus. Tason ja kolmiulotteisen avaruuden vektorit kirjoitetaan yleensä yksikkövektorien $\bar{\imath}$, $\bar{\jmath}$ ja $\bar{k}$ avulla. Eräs tason vektori voisi olla vaikkapa $\bar{v} = 3\bar{\imath} - 2\bar{\jmath}$. Jokaisella koordinaatiston pisteellä on paikkavektori, jonka komponentit ovat pisteen koordinaatit. Esimerkiksi pisteen $(3, -2)$ paikkavektori on edellä mainittu vektori $\bar{v}$.

Kun vektorin käsitettä yleistetään korkeampiin ulottuvuuksiin, osoittautuu helpoimmaksi käsitellä vektoreita ja avaruuden pisteitä samalla tavalla. Esimerkiksi merkintä $(3, -2)$ tarkoittaa sekä pistettä $(3, -2)$ että yllä määriteltyä vektoria $\bar{v}$. Merkinnät myös yksinkertaistuvat, kun voidaan luopua yksikkövektoreiden $\bar{\imath}$, $\bar{\jmath}$ ja $\bar{k}$ käytöstä.

Tässä luvussa käsitellään tason ja kolmiulotteisen avaruuden vektoreita. Tasoa ja kolmiulotteista avaruutta kutsutaan nimillä $\mathbb{R}^2$ ja $\mathbb{R}^3$.

1.1 Kaksiulotteisen avaruuden vektorit

Määritelmä 1.1. Vektoriavaruus $\mathbb{R}^2$ koostuu reaalilukupareista. Toisin sanoen

$$\mathbb{R}^2 = \{(a, b) \mid a \in \mathbb{R} \text{ ja } b \in \mathbb{R}\}.$$

Vektoriavaruuden $\mathbb{R}^2$ alkioita kutsutaan *vektoreiksi.*

Esimerkki 1.2. Esimerkiksi $(-3, -1)$ ja $\left(\frac{1}{2}, -\sqrt{5}\right)$ ovat vektoriavaruuden $\mathbb{R}^2$ vektoreita.

Huom. 1. Määritelmä tarkoittaa sopimusta. Tässä siis sovitaan, mitä vektoriavaruudella $\mathbb{R}^2$ ja vektoreilla tarkoitetaan. Määritelmää ei tarvitse perustella millään tavalla.

Huom. 2. Tarkalleen ottaen vektori (a, b) on niin kutsuttu *järjestetty pari.* Tämä tarkoittaa sitä, että lukujen a ja b järjestyksellä on väliä. Esimerkiksi järjestetty pari $(1, 2)$ ei ole sama kuin järjestetty pari $(2, 1)$.

Huom. 3. Usein termin vektoriavaruus sijasta käytetään lyhyesti vain ilmaisua *avaruus.*

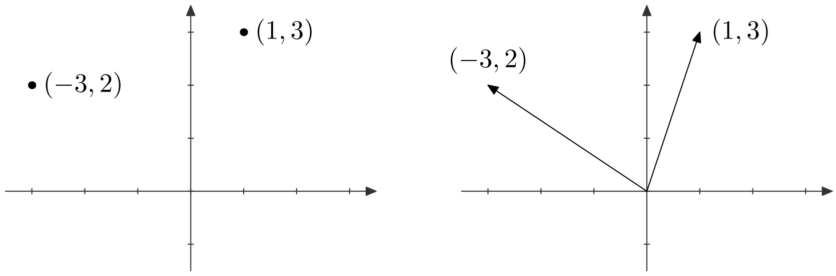
Vektoreita merkitään tässä tekstissä yleensä pienellä kirjaimella, jonka päällä on viiva. Voidaan esimerkiksi kirjoittaa $\bar{v} = (a, b)$. Luvut a ja b ovat tällöin vektorin $\bar{v}$ *komponentteja.* Vektoriavaruuden $\mathbb{R}^2$ vektoreissa on aina kaksi komponenttia, ja vektoriavaruutta $\mathbb{R}^2$ kutsutaan *kaksiulotteiseksi vektoriavaruudeksi.*

Esimerkki 1.3. Merkitään $\bar{v} = (4, -1)$ ja $\bar{u} = \left(\frac{1}{2}, -\sqrt{5}\right)$. Vektorin $\bar{v}$ komponentit ovat 4 ja -1. Vektorin $\bar{u}$ komponentit ovat puolestaan $\frac{1}{2}$ ja $-\sqrt{5}$.

Voi tuntua hieman kummalliselta kutsua tasoa avaruudeksi, sillä avaruuden ajatellaan yleensä olevan jotain kolmiulotteista. Matematiikassa sanaa avaruus käytetään kuitenkin paljon laajemmin kuin arkikielessä. Voimmekin puhua 2-ulotteisista avaruuksista ja yhtä hyvin vaikkapa 4-, 1- tai 100-ulotteisista avaruuksista.

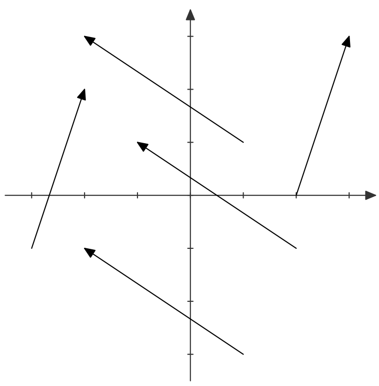
Kaksiulotteisen vektoriavaruuden vektoreita voidaan havainnollistaa eri tavoin. Eräs tapa on ajatella vektorit koordinaatiston pisteinä. Vektoria (a, b) vastaa piste, jonka vaakakoordinaatti on a ja pystykoordinaatti b. Kuvassa 1.1 on esitetty vektoreita $(1, 3)$ ja $(-3, 2)$ vastaavat tason pisteet.

Vektoria (a, b) voi kuvata myös pisteen (a, b) paikkavektorina eli suuntajanana, jonka lähtöpiste on origo ja päätepiste (a, b). Vektoreita $(1, 3)$ ja $(-3, 2)$ vastaavat paikkavektorit on myös esitetty kuvassa 1.1.



Kuva 1.1: Vektoreita $(1, 3)$ ja $(-3, 2)$ vastaavat tason pisteet sekä paikkavektorit.

Pisteen ja paikkavektorin lisäksi vektoriavaruuden $\mathbb{R}^2$ vektoria voi havainnollistaa mistä tahansa pisteestä lähtevällä suuntajanana. Suuntajanan paikalla ei ole väliä, ainoastaan sen suunta ja pituus merkitsevät. Vektoria (a, b) vastaavalla suuntajanalla on sama suunta ja pituus kuin pisteen (a, b) paikkavektorilla. Kuvassa 1.2 on esitetty vektoria $(1, 3)$ vastaavia suuntajanoja sekä vektoria $(-3, 2)$ vastaavia suuntajanoja.



Kuva 1.2: Vektoreita $(1, 3)$ ja $(-3, 2)$ vastaavia suuntajanoja.

Vektoriavaruuden $\mathbb{R}^2$ vektori on siis tällä kurssilla määritelmänsä mukaan kahdesta reaalilukusta koostuva järjestetty pari. Vektoreita voidaan havainnollistaa pisteinä, paikkavektoreina tai suuntajanoina. Se, millainen havainnollistamistapa on paras, riippuu siitä, mitä ollaan tekemässä. Usein vektoreita käsiteltäessä on pystyttävä vaihtamaan sulavasti yhdestä esitystavasta toiseen.

Koulusta tutut tason yksikkövektorit ovat määritelmän mukaan $\bar{\imath} = (1, 0)$ ja $\bar{\jmath} = (0, 1)$. Merkintöjä $\bar{\imath}$ ja $\bar{\jmath}$ ei käytetä tällä kurssilla.

1.2 Vektorien laskutoimitukset

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

Määritelmä 1.4. Oletetaan, että $\bar{v} = (v_1, v_2) \in \mathbb{R}^2$ ja $\bar{w} = (w_1, w_2) \in \mathbb{R}^2$. Vektoreiden $\bar{v}$ ja $\bar{w}$ *summa* on vektori

$$\bar{v} + \bar{w} = (v_1 + w_1, v_2 + w_2).$$

Lisäksi vektoreita voidaan kertoa reaaliluvuilla. Tätä operaatiota kutsutaan *skalaarikertolaskuksi*. Skalaarikertolaskussa vektorin molemmat komponentit kerrotaan samalla luvulla.

Määritelmä 1.5. Jos $\bar{v} = (v_1, v_2) \in \mathbb{R}^2$ ja $c \in \mathbb{R}$, määritellään

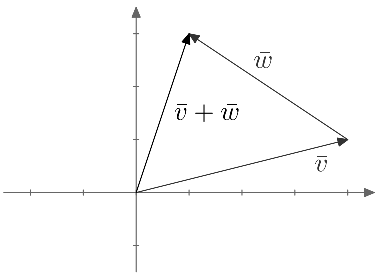
$$c\bar{v} = (cv_1, cv_2).$$

Vektorien yhteydessä reaalilukuja kutsutaan usein *skalaareiksi*, ja siitä johtuu myös skalaarikertolaskun nimitys.

Esimerkki 1.6. Tarkastellaan vektoreita $\bar{v} = (4, 1)$ ja $\bar{w} = (-3, 2)$. Niiden summa on

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

Yhteenlaskua voidaan havainnollistaa geometrisesti (kuva 1.3). Nyt on hyödyllistä ajatella vektoreita suuntajanoina, joiden paikalla ei ole merkitystä. 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.

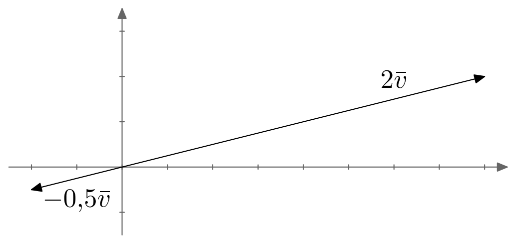


Kuva 1.3: Vektorit $\bar{v}$ ja $\bar{w}$ sekä niiden summa $\bar{v} + \bar{w}$.

Tutkitaan sitten skalaarikertolaskua. Määritelmän mukaan $2\bar{v} = (2 \cdot 4, 2 \cdot 1) = (8, 2)$ ja

$$-\frac{1}{2}\bar{v} = \left(-\frac{1}{2} \cdot 4, -\frac{1}{2} \cdot 1\right) = \left(-2, -\frac{1}{2}\right).$$

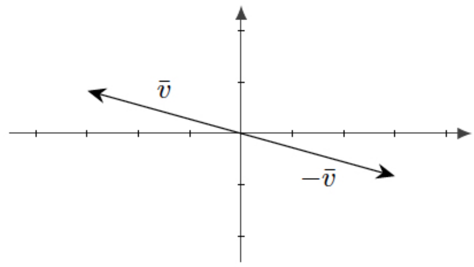
Vektorit $2\bar{v}$ ja $-\frac{1}{2}\bar{v}$ on piirretty kuvaan 1.4. Huomataan, että skalaarilla kertominen venyttää (eli skaalaa) vektoria vastaavan suuntajanan pituutta, mutta säilyttää suuntajanan suunnan, kuitenkin niin, että negatiivisella skalaarilla kertominen kääntää suunnan vastakkaiseksi.



Kuva 1.4: Skalaarimonikerrat $2\bar{v}$ ja $-\frac{1}{2}\bar{v}$.

Määritelmä 1.7. Vektorille $(-1)\bar{v}$ käytetään merkintää $-\bar{v}$. Sitä kutsutaan vektorin $\bar{v}$ *vastavektoriksi.*

Esimerkiksi vektorin $\bar{v} = (-3, 5/6)$ vastavektori on $-\bar{v} = (3, -5/6)$. Näitä on havainnollistettu kuvassa 1.5.



Kuva 1.5: Vektori $\bar{v}$ ja sen vastavektori $-\bar{v}$.

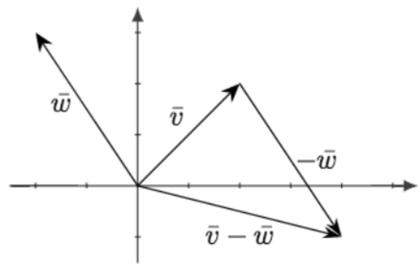
Määritelmä 1.8. Summalle $\bar{v} + (-\bar{w})$ käytetään merkintää $\bar{v} - \bar{w}$. Tätä kutsutaan vektorien $\bar{v}$ ja $\bar{w}$ *erotukseksi.*

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

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

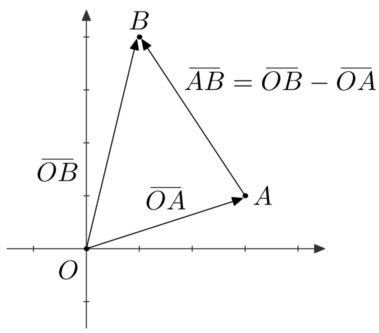
Vektoreiden erotus on erikoistapaus vektorien summasta, ja erotuksen voikin määrittää kuvasta samaan tapaan kuin summan (kuva 1.6). Määritelmän mukaan vektorien $\bar{v}$ ja $\bar{w}$ erotus $\bar{v} - \bar{w}$ saadaan laskemalla yhteen vektori $\bar{v}$ ja vastavektori $-\bar{w}$. Piirroksessa tämä tarkoittaa sitä, että jälkimmäisen vektorin suunta on käännettävä ennen yhteenlaskua.

Kuten edellisissä esimerkeissä nähtiin, vektoreiden summia ja erotuksia havainnollistaessa on erityisen kätevää ajatella vektoria suuntajanana, jonka voi kuvassa siirtää alkamaan mistä pisteestä tahansa. Silloin kun ei tutkita vektoreiden summia tai erotuksia, kannattaa vektoreita havainnollistaa joko koordinaatiston pisteinä tai origosta lähtevinä paikkavektoreina. Jos vektoreita ryhtyy turhaan siirtelemään koordinaatistossa, voi aiheuttaa itselleen ylimääräistä hämmennystä.



Kuva 1.6: Vektorit $\bar{v}$ ja $\bar{w}$ sekä niiden erotus $\bar{v}-\bar{w}$.

Otetaan vielä käyttöön merkintä, joka voi olla hyödyllinen silloin, kun vektoreita ajatellaan suuntajanoina. Oletetaan, että A ja B ovat tason pisteitä. Vektori $\overline{AB}$ on vektori, jota vastaavan suuntajanan alkupiste on A ja päätepiste on B (kuva 1.7). Origoa on tapana merkitä kirjaimella O. Siten pisteen A paikkavektorille saadaan merkintä $\overline{OA}$. Tällöin $\overline{AB} = \overline{OB} - \overline{OA}$.



Kuva 1.7: Vektorit $\overline{OA}$, $\overline{OB}$ ja $\overline{AB}$.

1.3 Kolmiulotteinen vektoriavaruus

Kaikki edellä esitellyt käsitteet voidaan määritellä myös kolmiulotteisessa avaruudessa. Vektoriavaruus $\mathbb{R}^3$ on joukko

$$\{(a,b,c) \mid a,b,c \in \mathbb{R}\}.$$

Myös vektoriavaruuden $\mathbb{R}^3$ alkioita nimitetään vektoreiksi, ja vektoriavaruutta $\mathbb{R}^3$ kutsutaan kolmiulotteiseksi vektoriavaruudeksi. Vektoriavaruuden $\mathbb{R}^3$ vektoreita voidaan ajatella avaruuskoordinaatiston pisteinä, paikkavektoreina tai suuntajanoina. Yhteenlasku ja skalaarikertolasku määritellään komponenteittain samalla tavalla kuin vektoriavaruudessa $\mathbb{R}^2$.

Kolmiulotteisen vektoriavaruuden koordinaattiakselien suuntaiset yksikkövektorit ovat $\bar{\imath} = (1,0,0)$, $\bar{\jmath} = (0,1,0)$ ja $\bar{k} = (0,0,1)$. Näitä merkintöjä ei jatkossa tulla tarvitsemaan.

2 Vektoriavaruus $\mathbb{R}^n$

Edellisessä luvussa käsiteltiin vektoriavaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ vektoreita eli reaalilukupareja ja reaalilukukolmikoita. Näitä vektoriavaruuksia voidaan yleistää määrittelemällä n-ulotteinen vektoriavaruus $\mathbb{R}^n$.

Määritelmä 2.1. Oletetaan, että $n \in \{1, 2, 3, \dots\}$. Vektoriavaruuden $\mathbb{R}^n$ alkiot ovat reaaliluvuista koostuvia n-jonoja. Toisin sanoen

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

Vektoriavaruuden $\mathbb{R}^n$ alkioita kutsutaan *vektoreiksi*.

Usein termin vektoriavaruus sijasta käytetäään lyhyesti ilmaisua avaruus.

Jos $\bar{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$, niin lukuja $v_1, v_2, \dots, v_n$ kutsutaan vektorin $\bar{v}$ *komponenteiksi*. Sovimme, että ellei toisin mainita, vektorin $\bar{v}$ komponentteja merkitään symboleilla $v_1, v_2, \dots, v_n$.

Yleisille n-ulotteisen vektoriavaruuden vektoreille määritellään laskutoimitukset samoin kuin kaksi- ja kolmiulotteisessa tapauksessa.

Määritelmä 2.2. Oletetaan, että $\bar{v} \in \mathbb{R}^n$, $\bar{w} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$. Tällöin

$$\begin{aligned} \bar{v} + \bar{w} &= (v_1 + w_1,\, v_2 + w_2,\, \dots,\, v_n + w_n) \quad \text{ja} \\ c\bar{v} &= (cv_1,\, cv_2,\, \dots,\, cv_n). \end{aligned}$$

Ensimmäistä laskutoimitusta nimitetään vektorien *yhteenlaskuksi* ja toista *skalaarikertolaskuksi*.

Vektorien yhteydessä reaalilukuja kutsutaan *skalaareiksi*. Jos $\bar{v} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$, vektoria $c\bar{v}$ nimitetään vektorin $\bar{v}$ *skalaarimonikerraksi*.

Huom. 1. Ainoastaan samanulotteisen vektoriavaruuden vektoreita voi laskea yhteen.

Huom. 2. Skalaarikertolasku on aina kirjoitettava järjestyksessä, jossa reaaliluku edeltää vektoria. Esimerkiksi merkintä $\bar{v}3$ ei tarkoita mitään, vaan oikea kirjoitusasu on $3\bar{v}$.

Määritelmä 2.3. Vektorin $\bar{v}$ *vastavektori* on skalaarimonikerta $(-1)\bar{v}$. Sitä merkitään $-\bar{v}$. Vektoreiden $\bar{v}$ ja $\bar{w}$ *erotus* on summa $\bar{v} + (-\bar{w})$. Sitä merkitään $\bar{v} - \bar{w}$. Vektoria $(0, 0, \dots, 0)$ kutsutaan *nollavektoriksi*. Sille käytetään merkintää $\bar{0}$.

Esimerkki 2.4. Merkitään $\bar{v} = (-5, 3, 0, 1, -1)$ ja $\bar{w} = (-2, -4, 2, 3, 5)$. Tällöin $\bar{v}$ ja $\bar{w}$ ovat vektoriavaruuden $\mathbb{R}^5$ vektoreita. Lasketaan vektorit $2\bar{v} - 3\bar{w}$ ja $-5\bar{v} - \bar{w}$:

$$\begin{aligned} 2\bar{v} - 3\bar{w} &= (-10, 6, 0, 2, -2) - (-6, -12, 6, 9, 15) = (-4, 18, -6, -7, -17) \\ -5\bar{v} - \bar{w} &= (25, -15, 0, -5, 5) - (-2, -4, 2, 3, 5) = (27, -11, -2, -8, 0). \end{aligned}$$

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

Lause 2.5. *Oletetaan, että $\bar{v}, \bar{w}, \bar{u} \in \mathbb{R}^n$ ja $a, b \in \mathbb{R}$. Tällöin pätee:*

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

Huom. 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 $\bar{v} \in \mathbb{R}^n$ ja $\bar{w} \in \mathbb{R}^n$. Tällöin $\bar{v} = (v_1, \ldots, v_n)$ ja $\bar{w} = (w_1, \ldots, w_n)$, ja luvut $v_1, \ldots, v_n$ ja $w_1, \ldots, w_n$ ovat reaalilukuja. Koska reaalilukujen yhteenlasku on vaihdannainen, jokaisella $i \in \{1, \ldots, n\}$ pätee $v_i + w_i = w_i + v_i$. Nyt nähdään, että

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

Väite on todistettu. □

Tasovektorien yhteydessä todettiin, että skalaarikertolasku säilyttää (tai kääntää vastakkaiseksi) vektorin suunnan. Otetaan tämä havainto yleisten vektorien yhdensuuntaisuuden määritelmäksi.

Määritelmä 2.6. Vektoriavaruuden $\mathbb{R}^n$ vektorit $\bar{v}$ ja $\bar{w}$ ovat *yhdensuuntaiset*, jos $\bar{v} = r\bar{w}$ jollakin $r \in \mathbb{R} \setminus \{0\}$. Tällöin merkitään $\bar{v} \parallel \bar{w}$.

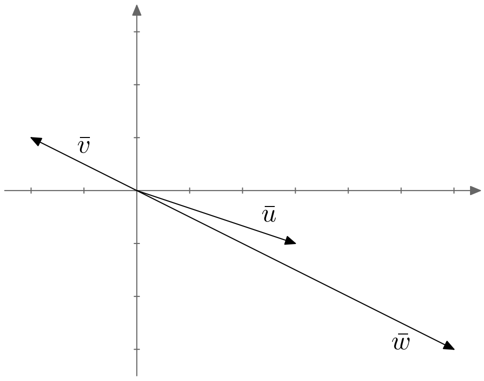
Esimerkki 2.7. Tutkitaan vektoreita $\bar{v} = (-2, 1)$, $\bar{w} = (6, -3)$ ja $\bar{u} = (3, -1)$. Ne on esitetty kuvassa 2.8.

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

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

Osoitetaan sitten, että vektorit $\bar{v} = (-2, 1)$ ja $\bar{u} = (3, -1)$ 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 $\bar{v} = r\bar{u}$. Tästä seuraa, että

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



Kuva 2.8: Esimerkin 2.7 vektorit $\bar{v}$, $\bar{w}$ ja $\bar{u}$.

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

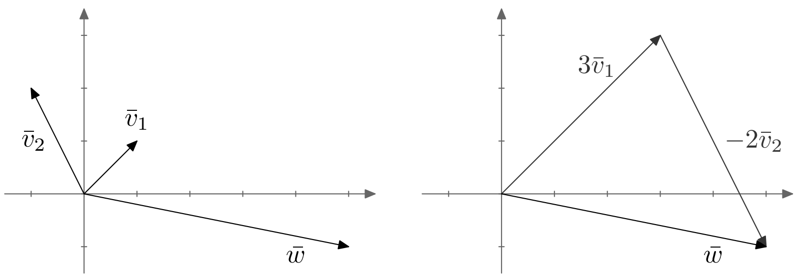
Määritelmä 2.8. Vektori $\bar{w} \in \mathbb{R}^n$ on vektoreiden $\bar{v}_1, \bar{v}_2, \dots \bar{v}_k \in \mathbb{R}^n$ *lineaarikombinaatio* eli *lineaariyhdistelmä*, jos on olemassa sellaiset reaaliluvut $a_1, a_2, \dots, a_k$, että

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

Esimerkki 2.9. Merkitään $\bar{v}_1 = (1,1)$, $\bar{v}_2 = (-1,2)$ ja $\bar{w} = (5,-1)$. Tutkitaan, onko $\bar{w}$ vektoreiden $\bar{v}_1$ ja $\bar{v}_2$ lineaarikombinaatio. Hetken pohdinnan jälkeen huomataan, että

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

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



Kuva 2.9: Vektori $\bar{w}$ on vektoreiden $\bar{v}_1$ ja $\bar{v}_2$ lineaarikombinaatio.

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

Esimerkki 2.10. Tutkitaan, onko vektori $\bar{w} = (-2, 3, 2, -1)$ vektoreiden

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

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

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

Toisin sanoen on pääteltävä, onko $\bar{w}$ vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio.

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

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

Kun tarkastellaan jokaista komponenttia erikseen, saatua vektoriyhtälöä vastaa yhtälöryhmä

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

Miten tällainen yhtälöryhmä ratkaistaan? Ennen kuin syvennymme enemmän lineaarikombinaatioihin, on syytä perehtyä yhtälöryhmien ratkaisemiseen.

3 Lineaariset yhtälöryhmät

Edellisen luvun lopun esimerkissä päädyttiin yhtälöryhmään, jonka ratkaisusta riippui, onko vektori eräiden toisten vektoreiden lineaarikombinaatio (esim. 2.10). Tämäntyyppisiä tilanteita esiintyy lineaarialgebrassa jatkuvasti, ja kysymykset voivat olla hyvin monimuotoisia. Esimerkiksi mainitussa esimerkissä ei itse asiassa tarvittu yhtälöryhmän varsinaista ratkaisua, vaan oli ainoastaan osoitettava sen *olemassaolo*. Toisissa kysymyksissä olennaista saattaa olla, onko mahdollisia ratkaisuja yksi vai useampia. Joidenkin yhtälöryhmien kohdalla haluamme selvittää, minkälaisen joukon ratkaisut muodostavat.

Esimerkki 3.1. Tarkastellaan yhtälöryhmää

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

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 monimutkaistuessa 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 kaikkiin (ainakin tällä kurssilla tarvittaviin) kysymyksiin voidaan helposti antaa vastaus.

3.1 Lineaarisen yhtälöryhmän määritelmä

Lineaarinen yhtälöryhmä on yhtälöryhmä, joka on muotoa

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

missä $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \in \mathbb{R}$. Symbolit $x_1, x_2, \ldots, x_n$ ovat yhtälöiden *tuntemattomia*. Lukuja $a_{11}, \ldots, a_{mn}$ nimitetään yhtälöryhmän *kertoimiksi* ja lukuja $b_1, b_2, \ldots, b_m$ *vakioiksi*. Jos tuntemattomia on vähän, niitä voidaan merkitä myös symboleilla x, y, z ja niin edelleen.

Esimerkiksi

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

on lineaarinen yhtälöryhmä.

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

Lineaarisen yhtälöryhmän ratkaisemisen kannalta oleellista ovat vain kertoimien ja vakioiden arvot, esimerkiksi tuntemattomien nimityksellä ei ole merkitystä. Kaikki tieto yhtälöryhmästä voidaankin tiivistää lukutaulukkoon eli *matriisiin*, jossa luetellaan kaikki kertoimet sekä vakiot. Kun käsitellään yhtälöryhmien sijasta matriiseja, päästään helpommalla, sillä tuntemattomia ei tarvitse kirjata ylös.

Esimerkiksi edellä esitellyn yhtälöryhmän matriisi on

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

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.

Kappaleessa 9 käsitellään matriisien teoriaan yleisemmin. Tässä luvussa käsittelemme vain yhtälöryhmistä saatuja matriiseja.

3.2 Alkeisrivitoimitukset ja porrasmatriisit

Seuraavaksi tutustutaan menetelmään, jolla voidaan ratkaista mikä tahansa lineaarinen yhtälöryhmä. 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 3.2. Yhtälöryhmiä kutsutaan *ekvivalenteiksi*, jos niillä on täsmälleen samat ratkaisut.

Ryhdymme muokkaamaan yhtälöryhmiä niin kutsutuilla alkeisrivitoimituksilla. Niiden avulla tuotetaan uusia yhtälöryhmiä, jotka ovat ekvivalentteja alkuperäisen yhtälöryhmän kanssa. Koska matriisien käsitteleminen on helpompaa kuin yhtälöryhmien, tehdään alkeisrivitoimitukset suoraan matriiseille.

Määritelmä 3.3. Seuraavat kolme operaatiota ovat *alkeisrivitoimituksia*:

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

Alkeisrivitoimituksille 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 ($i \neq j$).

Seuraavassa on annettu esimerkit erilaisista alkeisrivitoimituksista:

$$\left[\begin{array}{cc|c} -4 & 3 & 4 \\ 1 & 2 & -1 \\ 5 & 3 & 2 \\ 0 & 6 & 4 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|c} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 5 & 3 & 2 \\ 0 & 6 & 4 \end{array}\right] \xrightarrow{R_3 - 5R_1} \left[\begin{array}{cc|c} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 0 & -7 & 7 \\ 0 & 6 & 4 \end{array}\right] \xrightarrow{-\frac{1}{7}R_3} \left[\begin{array}{cc|c} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 0 & 1 & -1 \\ 0 & 6 & 4 \end{array}\right]$$

Määritelmä 3.4. Matriisi A on *riviekvivalentti* matriisin B kanssa, jos B saadaan matriisista A alkeisrivitoimituksilla.

Esimerkiksi edellisen esimerkin matriisit

$$\left[\begin{array}{cc|c} -4 & 3 & 4 \\ 1 & 2 & -1 \\ 5 & 3 & 2 \\ 0 & 6 & 4 \end{array}\right] \quad \text{ja} \quad \left[\begin{array}{cc|c} 1 & 2 & -1 \\ -4 & 3 & 4 \\ 0 & 1 & -1 \\ 0 & 6 & 4 \end{array}\right]$$

ovat riviekvivalentit. Alkeisrivitoimituksia voidaan ajatella tehtävän myös nolla kappaletta. Siten jokainen matriisi on itsensä kanssa riviekvivalentti.

Lause 3.5. *Jos yhtälöryhmiä vastaavat matriisit ovat riviekvivalentit, yhtälöryhmät ovat ekvivalentit.*

Lause voidaan muotoilla myös toisin: jos yhtälöryhmiä vastaavat matriisit ovat riviekvivalentit, yhtälöryhmillä on täsmälleen samat ratkaisut. Alkeisrivitoimituksen tekeminen ei siis muuta yhtälöryhmän ratkaisuja. Lauseen todistus on esitetty luvun lopussa.

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

Määritelmä 3.6. 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.

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right] \rightsquigarrow \underset{\text{toimituksia}}{\overset{\text{alkeisrivi-}}{\cdots}} \rightsquigarrow \left[\begin{array}{cccc|c} c_{11} & c_{12} & \dots & c_{1n} & d_1 \\ c_{21} & c_{22} & \dots & c_{2n} & d_2 \\ & & \vdots & & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mn} & d_m \end{array}\right]$$

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \qquad\vdots \qquad\qquad = \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \quad \leftarrow \underset{\text{ratkaisut}}{\text{samat}} \rightarrow \quad \begin{cases} c_{11}x_1 + \cdots + c_{1n}x_n = d_1 \\ c_{21}x_1 + \cdots + c_{2n}x_n = d_2 \\ \qquad\vdots \qquad\qquad = \vdots \\ c_{m1}x_1 + \cdots + c_{mn}x_n = d_m \end{cases}$$

Kuva 3.10: Gaussin–Jordanin eliminointimenetelmän perusta.

Esimerkiksi seuraavat matriisit ovat porrasmatriiseja. Niiden johtavat alkiot on lihavoitu.

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

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

Määritelmä 3.7. Matriisi on *redusoitu porrasmatriisi*, jos seuraavat ehdot toteutuvat:

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

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

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

Esimerkki 3.8. 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.

3.3 Gaussin–Jordanin eliminointimenetelmä

Tavoitteena on muuttaa yhtälöryhmän matriisi alkeisrivitoimitusten avulla redusoiduksi porrasmatriisiksi, josta ratkaisut näkyvät suoraan. Voidaan osoittaa, että mikä tahansa matriisi voidaan muuttaa tällä tavoin redusoiduksi porrasmatriisiksi ja että alkeisrivitoimitusten käyttämisjärjestys ei vaikuta tulokseen. Seuraava esimerkki näyttää, kuinka tämä tehdään.

Esimerkki 3.9. Muutetaan matriisi

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

redusoiduksi 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 alkiot 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

$$\xrightarrow{-1 \cdot R_2} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 2 & 4 \end{bmatrix}.$$

Toisen rivin johtavan alkion avulla voidaan muuttaa sen alla oleva alkio nollaksi. 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 3.5 nojalla.

Ohjeita redusoidun porrasmatriisin aikaansaamiseksi:

- Porrasmatriisia muodostetaan vasemmalta oikealle ja ylhäältä alaspäin.
- Johtavat alkiot kannattaa useimmiten muuttaa ykkösiksi.
- Johtavien alkioiden avulla muutetaan niiden alapuolella olevat alkiot nolliksi. Näin saadaan aikaan porrasmatriisi.
- Redusoitua porrasmatriisia muodostetaan oikealta vasemmalle ja alhaalta ylöspäin.
- Johtavien alkioiden avulla muutetaan niiden yläpuolella olevat alkiot nolliksi.
- Tee vain yksi alkeisrivitoimitus kerrallaan!

Toisinaan redusoituun porrasmatriisin voi päätyä nopeammin käyttämällä jotakin toista reittiä. Edellä kuvattujen välivaiheiden seuraaminen on kuitenkin turvallista, sillä ne tuottavat aina redusoidun porrasmatriisin.

Nyt olemme valmiita ratkaisemaan yhtälöryhmiä. Yhtälöryhmän ratkaiseminen Gaussin–Jordanin menetelmää käyttäen sisältää seuraavat vaiheet:

1. Kirjoita yhtälöryhmän matriisi.
2. Muuta matriisi alkeisrivitoimituksilla porrasmatriisiksi.

3. Muuta porrasmatriisi redusoiduksi porrasmatriisiksi.
4. Lue ratkaisut redusoidusta porrasmatriisista.

Esimerkki 3.10. Ratkaistaan yhtälöryhmä

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

Yhtälöryhmän matriisi on

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

Tämä matriisi muutettiin redusoiduksi porrasmatriisiksi esimerkissä 3.9:

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

Redusoitua porrasmatriisia vastaava yhtälöryhmä on

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

Koska alkuperäisen yhtälöryhmän ratkaisut ovat lauseen 3.5 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 3.11. Ratkaistaan lineaarinen yhtälöryhmä

$$\begin{cases} x + 2y + z & = & 8 \\ -3x - 6y - 3z & = & -21 \end{cases}$$

Muutetaan yhtälöryhmän matriisi redusoiduksi porrasmatriisiksi:

$$\left[\begin{array}{rrr|r} 1 & 2 & 1 & 8 \\ -3 & -6 & -3 & -21 \end{array}\right] \xrightarrow{R_2+3R_1} \left[\begin{array}{rrr|r} 1 & 2 & 1 & 8 \\ 0 & 0 & 0 & 3 \end{array}\right].$$

Vastaava yhtälöryhmä on

$$\begin{cases} x + 2y + z = 8 \\ \qquad\quad 0 = 3. \end{cases}$$

Alin yhtälö on aina epätosi, joten yhtälöryhmällä ei ole ratkaisuja.

Esimerkki 3.12. Ratkaistaan lineaarinen yhtälöryhmä

$$\begin{cases} 3x_1 + 3x_2 - 15x_3 &= 9 \\ x_1 \qquad\; - 2x_3 &= 1 \\ 2x_1 - x_2 - x_3 &= 0 \end{cases}$$

Muutetaan yhtälöryhmän matriisi redusoiduksi porrasmatriisiksi:

$$\left[\begin{array}{rrr|r} 3 & 3 & -15 & 9 \\ 1 & 0 & -2 & 1 \\ 2 & -1 & -1 & 0 \end{array}\right] \xrightarrow{\frac{1}{3}R_1} \left[\begin{array}{rrr|r} 1 & 1 & -5 & 3 \\ 1 & 0 & -2 & 1 \\ 2 & -1 & -1 & 0 \end{array}\right] \xrightarrow{R_2-R_1} \left[\begin{array}{rrr|r} 1 & 1 & -5 & 3 \\ 0 & -1 & 3 & -2 \\ 2 & -1 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{R_3-2R_1} \left[\begin{array}{rrr|r} 1 & 1 & -5 & 3 \\ 0 & -1 & 3 & -2 \\ 0 & -3 & 9 & -6 \end{array}\right] \xrightarrow{-1\cdot R_2} \left[\begin{array}{rrr|r} 1 & 1 & -5 & 3 \\ 0 & 1 & -3 & 2 \\ 0 & -3 & 9 & -6 \end{array}\right] \xrightarrow{R_3+3R_2} \left[\begin{array}{rrr|r} 1 & 1 & -5 & 3 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_1-R_2} \left[\begin{array}{rrr|r} 1 & 0 & -2 & 1 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Saatua matriisia vastaa yhtälöryhmä

$$\begin{cases} x_1 - 2x_3 = 1 \\ x_2 - 3x_3 = 2 \\ \qquad\quad 0 = 0 \end{cases}$$

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

$$\begin{cases} x_1 = 1 + 2t \\ x_2 = 2 + 3t \qquad \text{missä } t \in \mathbb{R}. \\ x_3 = t, \end{cases}$$

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.

Yhtälöryhmässä saattaa olla useitakin vapaita muuttujia. Nämä löytyvät redusoidussa porrasmatriisissa niistä sarakkeista, joissa ei ole lainkaan johtavaa alkiota.

Esimerkki 3.13. Lineaarisen yhtälöryhmän matriisi muutettiin alkeisrivitoimituksilla 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 alkiot 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

$$\begin{cases} x_1 + 3r + 4s = 7 \\ \qquad\quad x_3 + 2s = 0 \\ \qquad\qquad\quad x_6 = 3. \end{cases}$$

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

$$\begin{cases} x_1 = 7 - 3r - 4s \\ x_2 = r \\ x_3 = -2s \\ x_4 = s \\ x_5 = t \\ x_6 = 3, \end{cases} \qquad \text{missä } r, s, t \in \mathbb{R}.$$

Luvut r, s ja t voidaan valita täysin vapaasti, ja jokainen valinta tuottaa yhtälöryhmän erään ratkaisun.

Porrasmatriisien tulkinta

Edelliset esimerkit kuvaavat tilanteita, joihin Gaussin–Jordanin menetelmää käyttäen voidaan päätyä. Kootaan vielä yhteen redusoidun porrasmatriisin M merkitys sitä vastaavan yhtälöryhmän kannalta eri tapauksissa.

- Jos jokin matriisin M viimeisistä riveistä on muotoa $\left[\begin{array}{ccc|c} 0 & \cdots & 0 & 1 \end{array}\right]$ (eli rivin johtava ykkönen on pystyviivan oikealla puolella), kyseistä riviä vastaa epätosi yhtälö $0 = 1$. Yhtälöryhmällä ei ole ratkaisua.
- 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.

- Oletetaan, että edelliset tapaukset eivät toteudu. Tällöin matriisin M jokaisessa sarakkeessa pystyviivan vasemmalla puolella on johtava alkio, ja yhtälöryhmällä on täsmälleen yksi ratkaisu.

Esimerkki 3.14. 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 esimerkissä esitettyjä väitteitä ei perustella tarkasti, sillä pitkien ja teknisten todistusten kirjoittaminen ei olisi mielekästä.

Oletetaan, että yhtälöryhmän matriisi on saatu alkeisrivitoimituksilla muotoon

$$\left[\begin{array}{rrr|r} 4 & -3 & 1 & 0 \\ 0 & -1 & 2 & 2 \\ 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

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 alkeisrivitoimituksilla porrasmatriisiksi

$$\left[\begin{array}{rrrr|r} 4 & -3 & 0 & 1 & 0 \\ 0 & 6 & 2 & 3 & 2 \\ 0 & 0 & 0 & -1 & -4 \end{array}\right].$$

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 alkiota. Jos matriisi muutettaisiin redusoiduksi porrasmatriisiksi, ei siinäkään olisi johtavaa alkiota kolmannessa sarakkeessa. Siten yhtälöryhmällä on äärettömän monta ratkaisua, ja sen näkee suoraan porrasmatriisista.

Tarkastellaan vielä lopuksi yhtälöryhmää, jonka matriisi on saatu alkeisrivitoimituksilla porrasmatriisiksi

$$\left[\begin{array}{rrr|r} 4 & -3 & 1 & 0 \\ 0 & 6 & 2 & 2 \\ 0 & 0 & 2 & -4 \end{array}\right].$$

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 3.15. 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:

$$\left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 1 & k & 1 & 1 \\ k & 1 & 1 & -2 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & k-1 & 1-k & 0 \\ k & 1 & 1 & -2 \end{array}\right]$$
$$\xrightarrow{R_3 - kR_1} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & k-1 & 1-k & 0 \\ 0 & 1-k & 1-k^2 & -2-k \end{array}\right] \xrightarrow{R_3 + R_2} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & k-1 & 1-k & 0 \\ 0 & 0 & 2-k-k^2 & -2-k \end{array}\right]$$

Kaikki alkeisrivitoimitukset 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 alkiota, 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 nollasta poikkeavia. Tällöin voidaan jatkaa alkeisrivitoimitusten 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 $k \neq -2$.

Esimerkki 3.16. Redusoidut porrasmatriisit ovat teoreettisesti mielenkiintoisia, koska jokaista matriisia vastaa täsmälleen yksi redusoitu porrasmatriisi. Edellä kuitenkin nähtiin, että yhtälöryhmän ratkaisujen lukumärä on luettavissa jo porrasmatriisista. Itse asiassa yhtälöryhmä voidaan jopa ratkaista porrasmatriisivaiheen avulla.

Esimerkin 3.12 yhtälöryhmää vastaava porrasmatriisi oli

$$\left[\begin{array}{ccr|c} 1 & 1 & -5 & 3 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Koska kolmannessa sarakkeessa ei ole johtavaa alkiota, 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.

Historiallinen huomautus. 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.

Geometrinen tulkinta

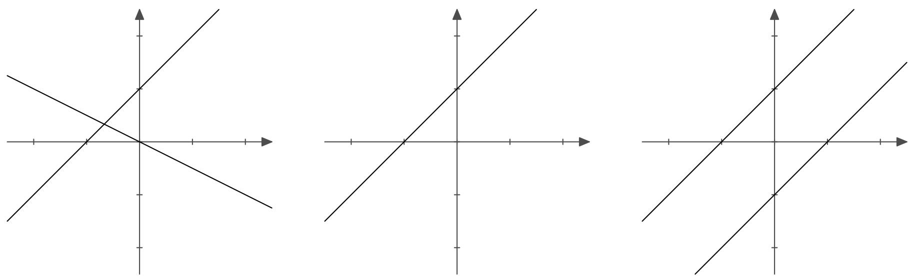
Yhtälöryhmällä voi siis olla täsmälleen yksi ratkaisu, äärettömän monta ratkaisua tai ei yhtään ratkaisua. Kun muuttujia on kaksi tai kolme, tilannetta voi havainnollistaa analyyttisen geometrian avulla. Tutkitaan yhtälöparia

$$\begin{cases} ax + by = c \\ dx + ey = f. \end{cases}$$

Oletetaan, että yhtälöllä on ratkaisu $x = r$, $y = s$. Sitä voidaan ajatella tason pisteenä (r, s). Koska ratkaisu toteuttaa ylemmän yhtälön, piste (r, s) on suoralla, jonka yhtälö on $ax+by = c$.

Vastaavasti piste (r, s) on suoralla, jonka yhtälö on $dx + ey = f$. Piste (r, s) on siis molemmilla suorilla, eli se on suorien leikkauspiste.

Jos yhtälöt määrittävät kaksi erisuuntaista suoraa, niillä on täsmälleen yksi leikkauspiste. Tällöin yhtälöparilla on täsmälleen yksi ratkaisu. Jos yhtälöt määrittävät saman suoran, on leikkauspisteitä äärettömän monta. Silloin ratkaisujakin on äärettömän monta. Jos yhtälöiden määrittämät suorat eivät ole samat mutta ovat kuitenkin yhdensuuntaiset, ei leikkauspisteitä ole. Silloin ei myöskään yhtälöparilla ole ratkaisuja.



Kuva 3.11: Yhtälöryhmällä on tasan yksi ratkaisu, äärettömän monta ratkaisua tai ei yhtään ratkaisua.

Kun muuttujia on kolme, yhtälöt kuvaavat tasoja. Esimerkiksi kolmen tason leikkausjoukko voi olla piste, suora tai taso. Nämä vastaavat tilanteista, joissa ratkaisuun tulee nolla, yksi tai kaksi vapaata muuttujaa. Leikkausjoukko voi olla myös tyhjä, jolloin yhtälöryhmällä ei ole ratkaisuja.

3.4 Yhtälöryhmien ekvivalenssin todistus

Käydään vielä lopuksi todistus sille, että yhtälöryhmillä on samat ratkaisut, jos niiden matriisit ovat riviekvivalentit (lause 3.5).

Lauseen 3.5 todistus. Osoitetaan, että jos yhtälöryhmiä vastaavat matriisit ovat riviekvivalentit, yhtälöryhmät ovat ekvivalentit. Tätä varten riittää näyttää, että alkeisrivitoimituksen tekeminen ei vaikuta yhtälöryhmän ratkaisuihin. Tutkitaan yhtälöryhmää

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & & & & & & \vdots & \\ a_{i1}x_1 & + & a_{i2}x_2 & + & \dots & + & a_{in}x_n & = & b_i \\ & & & & & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m, \end{array}\right. \tag{1}$$

missä $a_{11}, \dots, a_{mn}, b_1, \dots, b_m \in \mathbb{R}$.

1. Ensinnäkin huomataan, että yhtälöryhmän rivien järjestyksellä ei ole väliä. Siten kahden rivin paikkojen vaihtaminen ei muuta yhtälöryhmän ratkaisuja.

2. Tutkitaan sitten alkeisrivitoimitusta, joka kertoo rivin i luvulla $c \in \mathbb{R} \setminus \{0\}$. Tuloksena on yhtälöryhmä

$$
\begin{cases}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
 & & & & & & & \vdots & \\
ca_{i1}x_1 & + & ca_{i2}x_2 & + & \cdots & + & ca_{in}x_n & = & cb_i \\
 & & & & & & & \vdots & \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m.
\end{cases} \tag{2}
$$

On osoitettava, yhtälöryhmillä (1) ja (2) on samat ratkaisut. Tämä tehdään kahdessa osassa. Ensin näytetään, että jokainen yhtälöryhmän (1) ratkaisu on myös yhtälöryhmän (2) ratkaisu. Sitten näytetään, että jokainen yhtälöryhmän (2) ratkaisu on myös yhtälöryhmän (1) ratkaisu.

Oletetaan ensin, että $x_1 = r_1, \ldots, x_n = r_n$ on yhtälöryhmän (1) ratkaisu ja osoitetaan, että se on myös ryhmän (2) ratkaisu. Oletuksen perusteella pätee

$$
\begin{cases}
a_{11}r_1 & + & a_{12}r_2 & + & \cdots & + & a_{1n}r_n & = & b_1 \\
 & & & & & & & \vdots & \\
a_{i1}r_1 & + & a_{i2}r_2 & + & \cdots & + & a_{in}r_n & = & b_i \\
 & & & & & & & \vdots & \\
a_{m1}r_1 & + & a_{m2}r_2 & + & \cdots & + & a_{mn}r_n & = & b_m.
\end{cases}
$$

Kun i:nnen yhtälön molemmat puolet kerrotaan luvulla c, saadaan yhtälö

$$
ca_{i1}r_1 + \cdots + ca_{in}r_n = cb_i.
$$

Nyt siis $x_1 = r_1, \ldots x_n = r_n$ toteuttaa yhtälöryhmän (2), ja siten se on myös yhtälöryhmän (2) ratkaisu.

Oletetaan sitten, että $x_1 = s_1, \ldots, x_n = s_n$ on yhtälöryhmän (2) ratkaisu ja osoitetaan, että se on myös ryhmän (1) ratkaisu. Nyt siis

$$
\begin{cases}
a_{11}s_1 & + & a_{12}s_2 & + & \cdots & + & a_{1n}s_n & = & b_1 \\
 & & & & & & & \vdots & \\
ca_{i1}s_1 & + & ca_{i2}s_2 & + & \cdots & + & ca_{in}s_n & = & cb_i \\
 & & & & & & & \vdots & \\
a_{m1}s_1 & + & a_{m2}s_2 & + & \cdots & + & a_{mn}s_n & = & b_m.
\end{cases}
$$

Koska $c \neq 0$, voidaan i:nnen yhtälön molemmat puolet jakaa luvulla c. Tällöin saadaan yhtälö $a_{i1}s_1 + \cdots + a_{in}s_n = b_i$. Nyt nähdään, että $x_1 = s_1, \ldots, x_n = s_n$ toteuttaa myös yhtälöryhmän (1), joten se on myös yhtälöryhmän (1) ratkaisu. Siten yhtälöryhmillä on samat ratkaisut.

3. Kolmannen alkeisrivitoimituksen tarkastelu jätetään lukijalle. □

4 Virittäminen

Edellisessä luvussa opittiin vastaaamaan erilaisiin lineaaristen yhtälöryhmien ratkaisuja koskeviin kysymyksiin. Hyödynnetään näitä tietoja nyt vektoriavaruuden $\mathbb{R}^n$ vektoreiden tutkimiseen.

Osaamme muun muassa vastata esimerkissä 2.10 esitettyyn kysymykseen. Esimerkissä haluttiin tietää, onko vektori $\bar{w} = (-2, 3, 2, -1)$ vektoreiden

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

lineaarikombinaatio. Toisin sanoen oli selvitettävä, onko yhtälöllä

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

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

$$\left[\begin{array}{rrr|r} 1 & -1 & 0 & -1 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & 8 & 6 \\ 0 & 0 & 0 & -2 \end{array}\right].$$

Porrasmatriisissa on epätosi yhtälö, jonka perusteella tiedetään, että ratkaisuja ei ole. Siten $\bar{w}$ ei ole vektorien $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio.

Olemme siis oppineet selvittämään, onko jokin vektori toisten vektoreiden lineaarikombinaatio. Seuraavaksi tutkimme, milloin *jokainen* vektoriavaruuden vektori voidaan kirjoittaa joidenkin tiettyjen vektoreiden lineaarikombinaationa. Esimerkiksi varuuden $\mathbb{R}^2$ 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 $\mathbb{R}^2$.

Määritelmä 4.1. Vektorit $\bar{v}_1, \ldots, \bar{v}_k \in \mathbb{R}^n$ *virittävät* vektoriavaruuden $\mathbb{R}^n$, jos jokainen vektoriavaruudenavaruuden $\mathbb{R}^n$ alkio voidaan kirjoittaa vektoreiden $\bar{v}_1, \ldots, \bar{v}_k$ lineaarikombinaationa.

Virittämisen määritelmä voidaan ilmaista myös toisin sanoin: vektorit $\bar{v}_1, \ldots, \bar{v}_k \in \mathbb{R}^n$ virittävät vektoriavaruuden $\mathbb{R}^n$, jos

$$\mathbb{R}^n = \{a_1\bar{v}_1 + a_2\bar{v}_2 + \cdots + a_k\bar{v}_k \mid a_1, a_2, \ldots, a_k \in \mathbb{R}\}.$$

Vektoreita $\bar{v}_1, \ldots, \bar{v}_k$ 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 4.2. Osoitetaan, että vektorit $\bar{v}_1 = (1,1,1)$, $\bar{v}_2 = (1,1,0)$ ja $\bar{v}_3 = (1,0,0)$ virittävät avaruuden $\mathbb{R}^3$.

On siis osoitettava, että jokainen avaruuden $\mathbb{R}^3$ vektori voidaan esittää vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaationa. Oletetaan tätä varten, että $\bar{w} \in \mathbb{R}^3$. Nyt $\bar{w} = (w_1, w_2, w_3)$ joillakin $w_1, w_2, w_3 \in \mathbb{R}$.

Jotta vektori $\bar{w}$ olisi vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio, täytyy löytyä luvut x_1, x_2, $x_3 \in \mathbb{R}$, joille pätee

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

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 $\bar{w}$ on vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio.

Olemme nyt osoittaneet, että jokainen vektoriavaruuden $\mathbb{R}^3$ vektori on vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio. Näin ollen kyseiset kolme vektoria virittävät vektoriavaruuden $\mathbb{R}^3$.

Esimerkki 4.3. Tutkitaan, virittävätkö vektorit

$$\bar{v}_1 = (3,2,-1), \quad \bar{v}_2 = (2,-2,6) \quad \text{ja} \quad \bar{v}_3 = (3,4,-5)$$

vektoriavaruuden $\mathbb{R}^3$.

Oletetaan, että $\bar{w} \in \mathbb{R}^3$. Nyt $\bar{w} = (w_1, w_2, w_3)$ joillakin $w_1, w_2, w_3 \in \mathbb{R}^3$. Tutkitaan, onko $\bar{w}$ vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio. Täytyy siis selvittää, onko olemassa lukuja x_1, x_2, $x_3 \in \mathbb{R}$, joille pätee

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

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

$$\left[\begin{array}{rrr|c} 1 & -6 & 5 & -w_3 \\ 0 & 10 & -6 & w_2 + 2w_3 \\ 0 & 0 & 0 & w_1 - 2w_2 - w_3 \end{array}\right].$$

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 $\bar{w} = (w_1, w_2, w_3)$ on vektoren $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio, jos ja vain jos $w_1 - 2w_2 - w_3 = 0$.

Näin ollen esimerkiksi vektori $(1, 0, 0)$ ei ole vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio, sillä se ei toteuta edellä saatua yhtälöä. On siis olemassa vektoriavaruuden $\mathbb{R}^3$ vektori, joka ei ole vektoreiden $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ lineaarikombinaatio. Siten vektorit $\bar{v}_1$, $\bar{v}_2$ ja $\bar{v}_3$ eivät viritä avaruutta $\mathbb{R}^3$.

Esimerkki 4.4. Tutkitaan, virittävätkö vektorit

$$\bar{u}_1 = (1, 1, 0), \quad \bar{u}_2 = (1, 0, 1), \quad \bar{u}_3 = (0, 1, 1) \quad \text{ja} \quad \bar{u}_4 = (-2, 1, 1)$$

avaruuden $\mathbb{R}^3$. Oletetaan, että $\bar{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$. On selvitettävä, onko olemassa lukuja $x_1, x_2, x_3, x_4 \in \mathbb{R}$, joille pätee

$$x_1\bar{u}_1 + x_2\bar{u}_2 + x_3\bar{u}_3 + x_4\bar{u}_4 = \bar{w}.$$

Saadaan yhtälöryhmä, jonka matriisi on

$$\left[\begin{array}{rrrr|c} 1 & 1 & 0 & -2 & w_1 \\ 1 & 0 & 1 & 1 & w_2 \\ 0 & 1 & 1 & 1 & w_3 \end{array}\right].$$

Tästä saadaan alkeisrivitoimituksilla porrasmatriisi

$$\left[\begin{array}{rrrr|c} 1 & 1 & 0 & -2 & w_1 \\ 0 & 1 & -1 & -3 & w_1 - w_2 \\ 0 & 0 & 1 & 2 & \frac{1}{2}(w_3 + w_2 - w_1) \end{array}\right].$$

Yhtälöryhmällä on ratkaisuja, sillä porrasmatriisiss ei näy epätosia yhtälöitä. Tämä ei riipu mitenkään luvuista w_1, w_2 ja w_3. Siten $\bar{w}$ on vektoreiden $\bar{u}_1$, $\bar{u}_2$, $\bar{u}_3$ ja $\bar{u}_4$ lineaarikombinaatio. Näin ollen kyseiset vektorit virittävät avaruuden $\mathbb{R}^3$.

Edellisen esimerkin virittäjät $\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4$ eivät ole parhaat mahdolliset. Koska yhtälöryhmässä on vapaita muuttujia, on yhtälöryhmällä äärettömän monta ratkaisua. Vektoriavaruuden $\mathbb{R}^3$ alkiot voidaan siis kirjoittaa *usealla* eri tavalla virittäjävektorien lineaarikombinaatioina. Esimerkiksi jos $\bar{w} = (1, 2, 3)$, niin ratkaisut ovat

$$\begin{cases} x_1 = t \\ x_2 = 1 + t \\ x_3 = 2 - 2t \\ x_4 = t \end{cases} \qquad \text{missä } t \in \mathbb{R}.$$

Valitsemalla $t = 3$ saadaan

$$(1, 2, 3) = 3\bar{u}_1 + 4\bar{u}_2 - 4\bar{u}_3 + 3\bar{u}_4$$

ja toisaalta valitsemalla $t = 1$ saadaan

$$(1, 2, 3) = \bar{u}_1 + 2\bar{u}_2 + 0\bar{u}_3 + \bar{u}_4.$$

Tämä ei ole aina toivottavaa, vaan tavoitteena on löytää sellainen virittäjäjoukko, että aliavaruuden vektorit voidaan ilmaista virittäjävektorien lineaarikombinaationa täsmälleen yhdellä tavalla. Tällaisia virittäjäjoukkoja tutkitaan seuraavassa luvussa.

5 Vapaus

Edellisen luvun lopussa tutkittiin virittävätkö vektorit $\bar{u}_1 = (1,1,0)$, $\bar{u}_2 = (1,0,1)$, $\bar{u}_3 = (0,1,1)$ ja $\bar{u}_4 = (-2,1,1)$ avaruuden $\mathbb{R}^3$. Silloin huomattiin, että avaruuden $\mathbb{R}^3$ vektorin voi kirjoittaa monella eri tavalla virittäjävektoreiden lineaarikombinaationa. Esimerkiksi nollavektorin voi kirjoittaa muodossa $\bar{0} = 0\bar{u}_1+0\bar{u}_2+0\bar{u}_3+0\bar{u}_4$, muodossa $\bar{0} = 1\bar{u}_1+1\bar{u}_2+(-2)\bar{u}_3+1\bar{u}_4$ tai muodossa $\bar{0} = 2\bar{u}_1 + 2\bar{u}_2 + (-4)\bar{u}_3 + 2\bar{u}_4$. Tämä ei ole useinkaan toivottavaa, sillä halutaan, että vektorit on mahdollista kirjoittaa virittäjävektorien lineaarikombinaationa täsmälleen yhdellä tavalla. Tällöin vaikkapa nollavektori saadaan aikaiseksi *ainoastaan* niin, että jokaisen virittäjävektorin kertoimena on nolla. Tavoitteen täyttävää virittäjäjoukkoa kutsutaan vapaaksi.

5.1 Vapauden määritelmä

Määritelmä 5.1. Oletetaan, että $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k \in \mathbb{R}^n$. Vektorijono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on *vapaa* eli *lineaarisesti riippumaton*, jos yhtälöllä

$$x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_k\bar{v}_k = \bar{0}$$

on täsmälleen yksi ratkaisu $x_1 = 0$, $x_2 = 0$, $\ldots$, $x_k = 0$. (Tässä tuntemattomat $x_1, \ldots, x_k$ ovat reaalilukuja.)

Jos jono $(\bar{v}_1, \ldots, \bar{v}_k)$ on vapaa eli lineaarisesti riippumaton, voidaan myös sanoa, että vektorit $\bar{v}_1, \ldots, \bar{v}_k$ ovat lineaarisesti riippumattomia toisistaan. Jos jono ei ole vapaa, sanotaan, että se on *sidottu*.

Huom. 1. Vektorijonolla $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ tarkoitetaan yksinkertaisesti tiettyjen vektorien muodostamaa kokoelmaa. Sitä ei saa sekoittaa vektorimerkintään. Kyseessä ei siis ole jonkinlainen ”vektorien vektori”.

Huom. 2. Määritelmässä mainitulla yhtälöllä $x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_k\bar{v}_k = \bar{0}$ on aina ratkaisu $x_1 = 0$, $x_2 = 0$, $\ldots$, $x_k = 0$, oli jono vapaa tai ei. Tämä on yhtälön niin kutsuttu *triviaaliratkaisu*, joka on aina olemassa. Vapaiden vektorijonojen erityisominaisuus on siis se, että yhtälöllä ei ole mitään muita ratkaisuja kuin triviaaliratkaisu.

Luvussa 6 tullaan näkemään, että jos jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa ja virittää vektoriavaruuden $\mathbb{R}^n$, avaruuden $\mathbb{R}^n$ alkiot voidaan kirjoittaa täsmälleen yhdellä tavalla virittäjävektorien lineaarikombinaatioina. Virittäjien joukossa ei siis tällöin ole tarpeettomia vektoreita.

Aloitetaan vapaisiin ja sidottuihin jonoihin tutustuminen esimerkeillä.

Esimerkki 5.2. Merkitään $\bar{v}_1 = (1,2)$ ja $\bar{v}_2 = (-3,-1)$. Tutkitaan, onko jono $(\bar{v}_1, \bar{v}_2)$ vapaa vai sidottu.

Tarkastellaan yhtälöä $x_1\bar{v}_1 + x_2\bar{v}_2 = \bar{0}$, missä $x_1, x_2 \in \mathbb{R}$. Toisin sanoen tutkittava yhtälö on

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

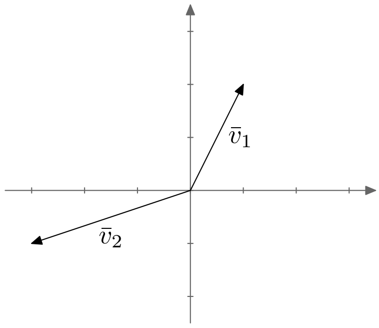
eli

$$\begin{cases} x_1 - 3x_2 = 0 \\ 2x_1 - x_2 = 0. \end{cases}$$

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 $x_2 = 0$. Jono $(\bar{v}_1, \bar{v}_2)$ on siis vapaa.

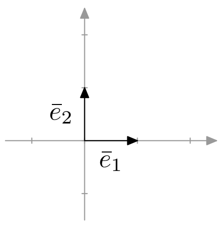


Kuva 5.12: Jono $(\bar{v}_1, \bar{v}_2)$ on vapaa.

Esimerkki 5.3. Merkitään $\bar{e}_1 = (1, 0)$ ja $\bar{e}_2 = (0, 1)$. Tutkitana, onko avaruuden $\mathbb{R}^2$ jono $(\bar{e}_1, \bar{e}_2)$ vapaa. Tarkastellaan siis yhtälöä

$$x_1\bar{e}_1 + x_2\bar{e}_2 = \bar{0},$$

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 jono $(\bar{e}_1, \bar{e}_2)$ on vapaa.



Kuva 5.13: Jono $(\bar{e}_1, \bar{e}_2)$ on vapaa.

Esimerkki 5.4. Kun osoitetaan jono sidotuksi, ei välttämättä tarvitse ratkaista yhtälöryhmää. Toisinaan on nimittäin helppo nähdä, minkälaisten kertoimien avulla lineaarikombinaatiosta muodostuu nollavektori.

Merkitään $\bar{w}_1 = (2, 1)$ ja $\bar{w}_2 = (-4, -2)$. Huomataan, että $2\bar{w}_1 + \bar{w}_2 = (4, 2) + (-4, -2) = \bar{0}$. Koska vektorien $\bar{w}_1$ ja $\bar{w}_2$ lineaarikombinaatio on nollavektori, vaikka kertoimet eivät ole nollia, jono $(\bar{w}_1, \bar{w}_2)$ on määritelmän nojalla sidottu.

Esimerkki 5.5. Merkitään $\bar{v}_1 = (1, 2)$, $\bar{v}_2 = (-3, -1)$ ja $\bar{v}_3 = (-1, 1)$. Tutkitaan, onko jono $(\bar{v}_1, \bar{v}_2, \bar{v}_3)$ vapaa vai sidottu. Tarkastellaan yhtälöä $c_1\bar{v}_1 + c_2\bar{v}_2 + c_3\bar{v}_3 = \bar{0}$, 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

$$\begin{cases} x_1 - 3x_2 - c_3 = 0 \\ 2x_1 - x_2 + c_3 = 0. \end{cases}$$

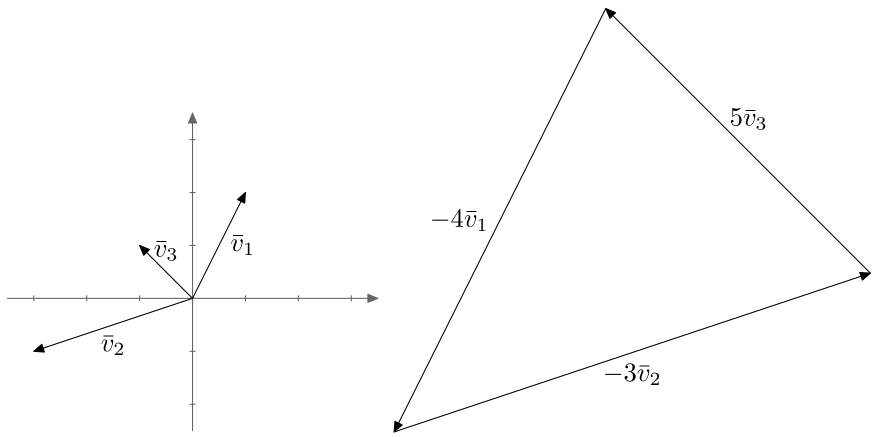
Ratkaistaan tästä x_1 ja x_2:

$$\left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 2 & -1 & 1 & 0 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 0 & 5 & 3 & 0 \end{array}\right] \xrightarrow{\frac{1}{5}R_2} \left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 0 & 1 & 3/5 & 0 \end{array}\right] \xrightarrow{R_1 + 3R_2} \left[\begin{array}{ccc|c} 1 & 0 & 4/5 & 0 \\ 0 & 1 & 3/5 & 0 \end{array}\right].$$

Huomataan, että yhtälöryhmällä on äärettömän monta ratkaisua:

$$\begin{cases} x_1 = -(4/5)t \\ x_2 = -(3/5)t \qquad \text{missä } t \in \mathbb{R}. \\ x_3 = t \end{cases}$$

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 $-4\bar{v}_1 - 3\bar{v}_2 + 5\bar{v}_3 = \bar{0}$. Jono $(\bar{v}_1, \bar{v}_2, \bar{v}_3)$ on siis sidottu. Tilannetta on havainnollistettu kuvassa 5.14.



Kuva 5.14: Jono $(\bar{v}_1, \bar{v}_2, \bar{v}_3)$ on sidottu.

Määritelmän mukaan jono $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ on vapaa, jos yhtälöllä

$$x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_k\bar{v}_k = \bar{0}$$

on täsmälleen yksi ratkaisu $x_1 = 0$, $x_2 = 0$, ..., $x_k = 0$. Näin ollen vapauden ehdon voi kirjoittaa myös muodossa

$$x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_k\bar{v}_k = \bar{0}, \quad \text{jos ja vain jos} \quad x_1 = 0, x_2 = 0, \dots, x_k = 0.$$

Tiedetään kuitenkin, että jos $x_1 = 0, x_2 = 0, \dots, x_k = 0$, niin $x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_k\bar{v}_k = \bar{0}$. Ekvialenssin toinen suunta on siis aina totta. Siksi vapauden määritelmä voidaan lyhentää seuraavanlaiseen muotoon:

$$\text{jos} \quad x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_k\bar{v}_k = \bar{0}, \quad \text{niin} \quad x_1 = 0, x_2 = 0, \dots, x_k = 0.$$

Tällaista muotoilua vapauden määritelmälle on kätevä käyttää esimerkiksi todistuksissa, joissa ei käsitellä konkreettisia vektoreita.

Kahden vektorin tapauksessa lineaarinen riippumattomuus on helppo tarkistaa. Rittää tutkia, ovatko vektorit yhdensuuntaisia.

Lause 5.6. *Oletetaan, että $\bar{v}, \bar{w} \in \mathbb{R}^n$ ja kumpikaan vektoreista ei ole nollavektori. Tällöin $(\bar{v}, \bar{w})$ on vapaa, jos ja vain jos vektorit $\bar{v}$ ja $\bar{w}$ eivät ole yhdensuuntaisia.*

Todistus. Todistus jätetään harjoitustehtäväksi. □

Seuraava lause osoittaa, että vektorijono on sidottu, jos ja vain jos jokin sen vektoreista voidaan ilmaista toisten lineaarikombinaationa.

Lause 5.7. *Oletetaan, että $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k \in \mathbb{R}^n$ ja $k \geq 2$. Jono $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ on sidottu, jos ja vain jos jollakin $j \in \{1, 2, \dots, k\}$ vektori $\bar{v}_j$ on vektoreiden $\bar{v}_1, \dots, \bar{v}_{j-1}, \bar{v}_{j+1}, \dots, \bar{v}_k$ lineaarikombinaatio.*

Todistus. "$\Rightarrow$": Oletetaan, että jono $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ on sidottu. On siis olemassa reaaliluvut $c_1, \dots, c_k$, joilla pätee

$$c_1\bar{v}_1 + c_2\bar{v}_2 + \cdots + c_k\bar{v}_k = \bar{0},$$

ja lisäksi $c_j \neq 0$ jollakin $j \in \{1, 2, \dots, k\}$. Nyt

$$c_j\bar{v}_j = -c_1\bar{v}_1 - \cdots - c_{j-1}\bar{v}_{j-1} - c_{j+1}\bar{v}_{j+1} - \cdots - c_k\bar{v}_k$$

ja edelleen

$$\bar{v}_j = -\frac{c_1}{c_j}\bar{v}_1 - \cdots - \frac{c_{j-1}}{c_j}\bar{v}_{j-1} - \frac{c_{j+1}}{c_j}\bar{v}_{j+1} - \cdots - \frac{c_k}{c_j}\bar{v}_k.$$

Siis $\bar{v}_j$ on vektoreiden $\bar{v}_1, \dots, \bar{v}_{j-1}, \bar{v}_{j+1}, \dots, \bar{v}_k$ lineaarikombinaatio.

"$\Leftarrow$": Oletetaan sitten, että $\bar{v}_j$ on vektoreiden $\bar{v}_1, \dots, \bar{v}_{j-1}, \bar{v}_{j+1}, \dots, \bar{v}_k$ 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

$$\bar{v}_j = c_1\bar{v}_1 + \cdots + c_{j-1}\bar{v}_{j-1} + c_{j+1}\bar{v}_{j+1} + \cdots + c_k\bar{v}_k.$$

Tästä seuraa, että

$$\bar{0} = c_1\bar{v}_1 + \cdots + c_{j-1}\bar{v}_{j-1} + (-1)\bar{v}_j + c_{j+1}\bar{v}_{j+1} + \cdots + c_k\bar{v}_k.$$

Koska kerroin -1 ei ole nolla, on jono $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ sidottu. □

Esimerkki 5.8. Tarkastellaan vektoriavaruuden $\mathbb{R}^3$ vektoreita $\bar{v}_1 = (1, -1, 0)$, $\bar{v}_2 = (1, 1, 0)$, $\bar{v}_3 = (0, 0, 2)$ ja $\bar{v}_4 = (3, -1, 0)$. Näillä pätee muun muassa

$$2\bar{v}_1 + \bar{v}_2 + 0\bar{v}_3 - \bar{v}_4 = (2, -2, 0) + (1, 1, 0) + (0, 0, 0) - (3, -1, 0) = (0, 0, 0),$$

joten jono $(\bar{v}_1, \bar{v}_2, \bar{v}_3, \bar{v}_4)$ on sidottu. Edellisen lauseen perusteella jokin vektoreista voidaan kirjoittaa toisten lineaarikombinaationa. Yllä olevasta yhtälöstä nähdäänkin, että

$$\bar{v}_2 = -2\bar{v}_1 + \bar{v}_4.$$

Kaikkia vektoreita ei kuitenkaan välttämättä voida kirjoittaa toisten lineaarikombinaationa. Esimerkiksi ei ole olemassa sellaisia lukuja a, b ja c, että pätisi

$$\bar{v}_3 = a\bar{v}_1 + b\bar{v}_2 + c\bar{v}_4.$$

(Tämän täsmällinen todistaminen jätetään lukijalle.)

5.2 Homogeeniset yhtälöryhmät ja vapaus

Vektorijonon vapautta tutkittaessa päädytään ratkaisemaan yhtälöryhmiä, joissa vakiot ovat nollia. Tällaista yhtälöryhmää kutsutaan homogeeniseksi.

Määritelmä 5.9. Lineaarinen yhtälöryhmä, jonka kaikki vakiot ovat nollia, on nimeltään *homogeeninen yhtälöryhmä.*

Homogeeninen yhtälöryhmä on siis muotoa

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & 0 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & 0, \end{cases}$$

missä $a_{11}, \ldots, a_{mn} \in \mathbb{R}$. Homogeenisella yhtälöryhmällä on aina ainakin yksi ratkaisu:

$$x_1 = 0, \quad x_2 = 0, \quad \ldots \quad x_n = 0.$$

Tätä kutsutaan yhtälöryhmän triviaaliksi ratkaisuksi.

Lause 5.10. *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 alkiota. Näin ollen löytyy vapaa muuttuja, mistä seuraa, että yhtälöryhmällä on äärettömän monta ratkaisua. □

Korollaari 5.11. *Oletetaan, että* $\bar{v}_1$, $\bar{v}_2, \ldots, \bar{v}_n \in \mathbb{R}^m$, *missä* $n \in \{1, 2, \ldots\}$. *Jos* $n > m$, *niin jono* $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n)$ *on sidottu.*

Huom. Korollaari on lause joka seuraa suoraan tai lähes suoraan toisesta lauseesta. Tämä korollaari on lauseen 5.10 seuraus.

Todistus. Merkitään $\bar{v}_k = (v_{1k}, v_{2k}, \ldots, v_{mk})$ kaikilla $k \in \{1, \ldots, n\}$. Nyt yhtälöä

$$x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_n\bar{v}_n = \bar{0}$$

vastaavaksi yhtälöryhmäksi saadaan

$$\begin{cases} v_{11}x_1 + v_{12}x_2 + \cdots + v_{1n}x_n & = & 0 \\ v_{21}x_1 + v_{22}x_2 + \cdots + v_{2n}x_n & = & 0 \\ \vdots & & \vdots \\ v_{m1}x_1 + v_{m2}x_2 + \cdots + v_{mn}x_n & = & 0. \end{cases}$$

Tässä homogeenisessa yhtälöryhmässä on enemmän tuntemattomia kuin yhtälöitä. Siten yhtälöryhmällä on äärettömän monta ratkaisua. Koska löytyy muitakin ratkaisuja kuin triviaaliratkaisu, ei jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n)$ ole vapaa. □

Jos homogeenisessa yhtälöryhmässä tuntemattomien määrä n on pienempi tai yhtä suuri kuin yhtälöiden määrä m, ei ratkaisujen määrästä voi äkkiseltään sanoa mitään varmaa. Ratkaisuja voi olla täsmälleen yksi (triviaali ratkaisu) tai äärettömän monta. Jos siis avaruuden $\mathbb{R}^n$ jonon $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m)$ pituus m on *pienempi* kuin n, ei jonon lineaarisesta riippumattomuudesta voida sanoa sen perusteella mitään.

6 Kanta

Tässä luvussa tarkastellaan vektoriavaruuden virittäjävektoreita, jotka muodostavat vapaan jonon.

Määritelmä 6.1. Olkoot $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k \in \mathbb{R}^n$. Vektorijono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vektoriavaruuden $\mathbb{R}^n$ *kanta*, jos seuraavat ehdot pätevät:

a) vektorit $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ virittävät avaruuden $\mathbb{R}^n$.

b) jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa.

Esimerkki 6.2. Luvun 4 alussa osoitettiin, että vektorit $\bar{e}_1 = (1, 0)$ ja $\bar{e}_2 = (0, 1)$ virittävät avaruuden $\mathbb{R}^2$. Jono $(\bar{e}_1, \bar{e}_2)$ on lisäksi vapaa esimerkin 5.3 perusteella. Siten $(\bar{e}_1, \bar{e}_2)$ on avaruuden $\mathbb{R}^2$ kanta.

Vastaavasti avaruudella $\mathbb{R}^n$ on kanta

$$(\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n).$$

Tässä $\bar{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, 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 $\mathcal{E}_n$.

Määritelmän mukaan vektoriavaruuden $\mathbb{R}^n$ kanta on vapaa vektorijono, joka virittää avaruuden $\mathbb{R}^n$. Seuraava lause osoittaa, että avaruuden $\mathbb{R}^n$ vektorit voidaan kirjoittaa kannan vektoreiden lineaarikombinaatioina täsmälleen yhdellä tavalla. Myös kääntäinen väite pätee: jos avaruuden $\mathbb{R}^n$ vektorit voidaan kirjoittaa joidenkin vektoreiden lineaarikombinaatioina täsmälleen yhdellä tavalla, kyseessä on kanta.

Lause 6.3. *Jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vektoriavaruuden $\mathbb{R}^n$ kanta, jos ja vain jos jokainen avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa täsmälleen yhdellä tavalla vektoreiden $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ lineaarikombinaationa.*

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 ensin, että jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vektoriavaruuden $\mathbb{R}^n$ kanta. Osoitetaan, että jokainen avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa täsmälleen yhdellä tavalla virittäjävektorien $\bar{v}_1, \ldots, \bar{v}_k$ lineaarikombinaationa.

Kannan määritelmän nojalla vektorit $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ virittävät avaruuden $\mathbb{R}^n$. Jokainen avaruuden vektori voidaan siis kirjoittaa vektorien $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ lineaarikombinaationa. On vielä osoitettava, että vektorit voidaan kirjoittaa lineaarikombinaationa *vain yhdellä tavalla.*

Oletetaan, että alkio $w \in \mathbb{R}^n$ voidaan kirjoittaa lineaarikombinaationa

$$\bar{w} = a_1\bar{v}_1 + \cdots + a_k\bar{v}_k \tag{3}$$

ja lineaarikombinaationa

$$\bar{w} = b_1\bar{v}_1 + \cdots + b_k\bar{v}_k \tag{4}$$

joillakin $a_1, \ldots, a_k, b_1, \ldots, b_k \in \mathbb{R}$. Nyt $a_1\bar{v}_1 + \cdots + a_k\bar{v}_k = b_1\bar{v}_1 + \cdots + b_k\bar{v}_k$, joten

$$a_1\bar{v}_1 + \cdots + a_k\bar{v}_k - (b_1\bar{v}_1 + \cdots + b_k\bar{v}_k) = \bar{0}.$$

Vektorien yhteenlaskun ja skalaarikertolaskun ominaisuuksien perusteella pätee

$$(a_1 - b_1)\bar{v}_1 + \cdots + (a_k - b_k)\bar{v}_k = \bar{0}.$$

Jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on oletuksen nojalla vapaa, joten yllä olevasta yhtälöstä seuraa, että kaikki kertoimet ovat nollia: $a_1 - b_1 = 0$, $\ldots$, $a_k - b_k = 0$. Siten $a_1 = b_1$, $\ldots$, $a_k = b_k$. Näin ollen tutkitut lineaarikombinaatiot (3) ja (4) ovatkin itse asiassa samanlaiset (niissä on samat kertoimet). Siksi vektoria $\bar{w}$ ei voida kirjoittaa usealla eri tavalla virittäjävektoreiden lineaarikombinaationa.

"$\Leftarrow$": Oletetaan sitten, että jokainen avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa täsmälleen yhdellä tavalla vektorien $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ lineaarikombinaationa. Osoitetaan, että jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on avaruuden $\mathbb{R}^n$ kanta.

Koska avaruuden $\mathbb{R}^n$ vektorit voidaan kirjoittaa vektorien $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ lineaarikombinaationa, virittävät vektorit $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ avaruuden $\mathbb{R}^n$.

On vielä osoitettava, että jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa. Sitä varten tutkitaan yhtälöä

$$x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_k\bar{v}_k = \bar{0},$$

missä $x_1 = 0, \ldots, x_k \in \mathbb{R}$. Tiedetään, että ainakin

$$0\bar{v}_1 + 0\bar{v}_2 + \cdots + 0\bar{v}_k = \bar{0}.$$

Koska nollavektori $\bar{0}$ on vektoriavaruuden $\mathbb{R}^n$ alkio, se voidaan oletuksen mukaan kirjoittaa virittäjävektorien lineaarikombinaationa täsmälleen yhdellä tavalla. Siksi pätee $x_1 = 0$, $\ldots$, $x_k = 0$. Siis jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa. $\square$

Kannan määritelmässä on kaksi ehtoa, joista ensimmäinen koskee virittämistä ja toinen vapautta. Toisaalta edellisen lauseen nojalla annetut vektorit muodostavat kannan, jos ja vain jos avaruuden vektorit voidaan kirjoittaa niiden lineaarikombinaatioina täsmälleen yhdellä tavalla. Virittäminen on yhtäpitävää sen kanssa, että avaruuden vektorit *voidaan* kirjoittaa annettujen vektoreiden lineaarikombinaatioina. Lauseen todistuksesta nähdään, että vapaus puolestaan on yhtäpitävää sen kanssa, että avaruuden vektorit voidaan kirjoittaa annettujen vektorien lineaarikombinaatioina *täsmälleen yhdellä tavalla.*

Esimerkki 6.4. Merkitään $\bar{w}_1 = (2, -1)$, $\bar{w}_2 = (1, 3)$. Osoitetaan lauseen 6.3 avulla, että $(\bar{w}_1, \bar{w}_2)$ on avaruuden $\mathbb{R}^2$ kanta. Oletetaan, että $\bar{v} \in \mathbb{R}^2$. Ratkaistaan yhtälö $x_1\bar{w}_1 + x_2\bar{w}_2 = \bar{v}$ 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 $\bar{v} \in \mathbb{R}^2$.

Siten jokainen avaruuden $\mathbb{R}^2$ vektori voidaan kirjoittaa vektorien $\bar{w}_1$ ja $\bar{w}_2$ lineaarikombinaationa täsmälleen yhdellä tavalla. Näin ollen jono $(\bar{w}_1, \bar{w}_2)$ on avaruuden $\mathbb{R}^2$ 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 maäärä vektoreita. Seuraava lause osoittaa, että jokaisessa vektoriavaruuden kannassa on aina sama määrä vektoreita.

Lause 6.5. *Aliavaruuden $\mathbb{R}^n$ jokaisessa kannassa on yhtä monta vektoria.*

Todistus. Oletetaan, että $\mathcal{B} = (\bar{v}_1, \ldots, \bar{v}_j)$ ja $\mathcal{C} = (\bar{w}_1, \ldots, \bar{w}_k)$ ovat molemmat vektoriavaruuden $\mathbb{R}^n$ 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

$$x_1\bar{w}_1 + \cdots + x_k\bar{w}_k = \bar{0}\,. \tag{5}$$

Koska $\mathcal{B}$ on aliavaruuden W kanta, voidaan kaikki kannan $\mathcal{C}$ vektorit kirjoittaa kannan $\mathcal{B}$ vektorien lineaarikombinaatioina:

$$\begin{aligned} \bar{w}_1 &= a_{11}\bar{v}_1 + a_{12}\bar{v}_2 + \cdots + a_{1j}\bar{v}_j \\ \bar{w}_2 &= a_{21}\bar{v}_1 + a_{22}\bar{v}_2 + \cdots + a_{2j}\bar{v}_j \\ &\vdots \\ \bar{w}_k &= a_{k1}\bar{v}_1 + a_{k2}\bar{v}_2 + \cdots + a_{kj}\bar{v}_j \end{aligned}$$

joillakin $a_{11}, \ldots, a_{kj} \in \mathbb{R}$. Sijoittamalla nämä yhtälöön (5) muodostuu yhtäpitävä yhtälö:

$$\begin{multline} x_1(a_{11}\bar{v}_1 + a_{12}\bar{v}_2 + \cdots + a_{1j}\bar{v}_j) \\ + x_2(a_{21}\bar{v}_1 + a_{22}\bar{v}_2 + \cdots + a_{2j}\bar{v}_j) \\ + \ldots \\ + x_k(a_{k1}\bar{v}_1 + a_{k2}\bar{v}_2 + \cdots + a_{kj}\bar{v}_j) = \bar{0}, \end{multline}$$

josta saadaan edelleen ryhmittelemällä

$$\begin{multline} (x_1a_{11} + x_2a_{21} + \cdots + x_ka_{k1})\bar{v}_1 \\ + (x_1a_{12} + x_2a_{22} + \cdots + x_ka_{k2})\bar{v}_2 \\ + \ldots \\ + (x_1a_{1j} + x_2a_{2j} + \cdots + x_ka_{kj})\bar{v}_j = \bar{0}. \end{multline}$$

Jono $\mathcal{B} = (\bar{v}_1, \ldots, \bar{v}_j)$ on kanta, joten se on vapaa. Siten edellinen yhtälö toteutuu, jos ja vain jos kaikki kertoimet ovat nollia:

$$\begin{cases} x_1 a_{11} + x_2 a_{21} + \cdots + x_k a_{k1} &= 0 \\ x_1 a_{12} + x_2 a_{22} + \cdots + x_k a_{k2} &= 0 \\ \qquad\qquad \vdots & \ \ \vdots \\ x_1 a_{1j} + x_2 a_{2j} + \cdots + x_k a_{kj} &= 0 \end{cases}$$

Kyseessä on homogeeninen yhtälöryhmä, jossa tuntemattomien määrä k on suurempi kuin yhtälöiden määrä j. Lauseen 5.10 mukaan yhtälöryhmällä on muitakin ratkaisuja kuin triviaaliratkaisu $x_1 = 0, \ldots, x_k = 0$. Siis jono $\mathcal{C} = (\bar{w}_1, \ldots, \bar{w}_k)$ on sidottu. Tämä on ristiriita, sillä $\mathcal{C}$ on aliavaruuden W kanta.

Tapaus $j > k$ käsitellään vastaavasti. Tällöinkin päädytään ristiriitaan. Täytyy siis päteä $j = k$. □

6.1 Koordinaatit

Kun vektoriavaruuden vektori kirjoitetaan kannan vektorien lineaarikombinaationa, kertoimia kutsutaan vektorin koordinaateiksi.

Määritelmä 6.6. Oletetaan, että $\mathcal{B} = (\bar{v}_1, \ldots, \bar{v}_k)$ on vektoriavaruuden $\mathbb{R}^n$ kanta. Oletetaan, että $\bar{u} \in \mathbb{R}^n$. Vektorin $\bar{u}$ *koordinaateiksi kannan $\mathcal{B}$ suhteen* kutsutaan reaalilukuja $b_1, \ldots, b_k$, joilla

$$\bar{u} = b_1\bar{v}_1 + \cdots + b_k\bar{v}_k\,.$$

Huom. Vektorin koordinaatit jonkin tietyn kannan suhteen ovat yksikäsitteiset, sillä vektori voidaan lauseen 6.3 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.

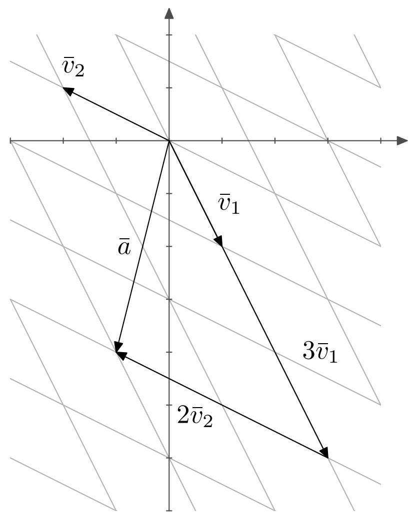
Esimerkki 6.7. Luonnollisen kannan suhteen koordinaatit on helppo määrittää. Esimerkiksi vektorin $\bar{a} = (-1, -4)$ koordinaatit avaruuden $\mathbb{R}^2$ luonnollisen kannan $\mathcal{E}_2 = (\bar{e}_1, \bar{e}_2)$ suhteen ovat -1 ja -4, sillä $\bar{a} = (-1, -4) = (-1)(1, 0) + (-4)(0, 1) = (-1)\bar{e}_1 + (-4)\bar{e}_2$.

Tutkitaan sitten erästä toista avaruuden $\mathbb{R}^2$ kantaa. Merkitään $\bar{v}_1 = (1, -2)$, $\bar{v}_2 = (-2, 1)$. Nyt $\mathcal{B} = (\bar{v}_1, \bar{v}_2)$ on avaruuden $\mathbb{R}^2$ kanta, minkä osoittaminen jätetään lukijalle. Määritetään vektorin $\bar{a} = (-1, -4)$ koordinaatit kannan $\mathcal{B}$ suhteen. On ratkaistava yhtälö $\bar{a} = x_1\bar{v}_1 + x_2\bar{v}_2$ eli yhtälö

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

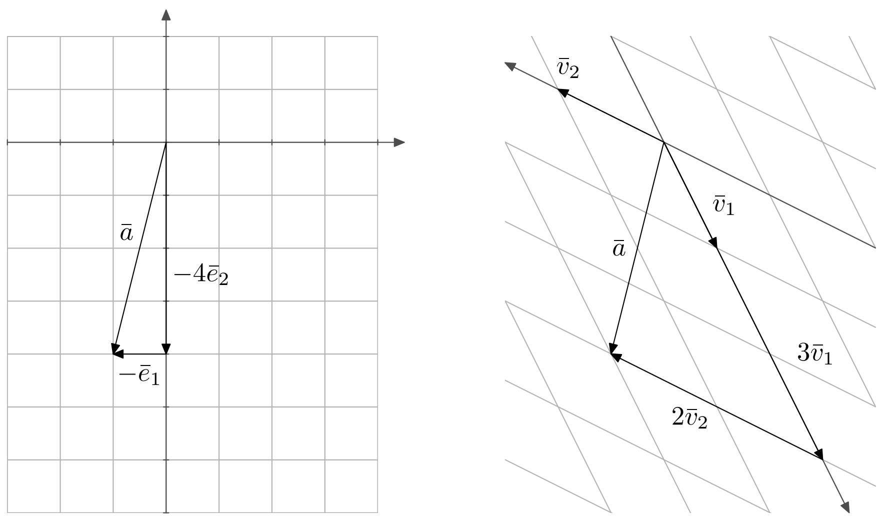
Tästä yhtälöstä saadaan tuttuun tapaan yhtälöryhmä, jonka ratkaisuksi paljastuu $x_1 = 3$, $x_2 = 2$. Näin ollen $\bar{a} = 3\bar{v}_1 + 2\bar{v}_2$ eli vektorin $\bar{a}$ koordinaatit kannan $\mathcal{B}$ suhteen ovat 3 ja 2.

Vektorin $\bar{a} = (-1, -4)$ koordinaatteja kannan $\mathcal{B} = (\bar{v}_1, \bar{v}_2)$ suhteen on havainnollistettu kuvassa 6.15. Jotta päästään pisteeseen $(-1, -4)$ on mentävä 3 ruutua vektorin $\bar{v}_1$ suuntaan ja 2 ruutua vektorin $\bar{v}_2$ suuntaan. Siten vektorin $\bar{a} = (-1, -4)$ koordinaatit kannan $\mathcal{B}$ suhteen ovat 3 ja 2.



Kuva 6.15: Vektori $\bar{a} = (-1, -4)$ ilmaistuna kannan $\mathcal{B} = (\bar{v}_1, \bar{v}_2)$ vektoreiden lineaarikombinaationa. Vektorin koordinaatit kannan $\mathcal{B}$ suhteen ovat 3 ja 2.

Kun käytetään jotakin muuta kuin luonnollista kantaa, vääntyy koordinaatisto vinoon. Kuvassa 6.16 vasemmalla vektori $\bar{a}$ on piirretty luonnollista kantaa $\mathcal{E}_2 = (\bar{e}_1, \bar{e}_2)$ vastaavaan tavalliseen koordinaatistoon ja oikealla kantaa $\mathcal{B} = (\bar{v}_1, \bar{v}_2)$ vastaavaan koordinaatistoon.



Kuva 6.16: Vektori $\bar{a}$ luonnollisen kannan $\mathcal{E}_2$ määräämässä koordinaatistossa ja kannan $\mathcal{B}$ määräämässä koordinaatistossa. Kun käytetään jotakin muuta kuin luonnollista kantaa, vääntyy koordinaatisto vinoon.

6.2 Vektoriavaruuden $\mathbb{R}^n$ dimensio

Lukijalla on todennäköisesti jonkinlainen käsitys siitä, mitä ulottuvuudella eli dimensiolla tarkoitetaan. Esimerkiksi avaruus $\mathbb{R}^2$ on kaksiulotteinen ja avaruus $\mathbb{R}^3$ kolmiulotteinen.

Dimensio määritellään täsmällisesti kannan avulla: avaruuden dimensio on sen kannassa olevien vektorien lukumäärä. Dimension määritteleminen tällä tavalla ei ole aivan suoraviivaista, sillä yhdellä vektoriavaruudella voi olla monta erilaista kantaa. Periaatteessa eri kannoissa voisi olla eri määrä vektoreita, eikä silloin olisi selvää, mikä niistä määräisi aliavaruuden dimension. Lauseen 6.5 perusteella kuitenkin tiedetään, että vektorivaruuden jokaisessa kannassa on aina yhtä monta vektoria. Siten dimension määritelmä on järkevä.

Määritelmä 6.8. Vektoriavaruuden $\mathbb{R}^n$ dimensio on sen kannassa olevien vektorien lukumäärä.

Jos avaruuden dimensio on n, sanotaan, että avaruus on *n-ulotteinen*.

Esimerkin 6.2 perusteella vektoriavaruudella $\mathbb{R}^n$ on aina luonnollinen kanta

$$(\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n).$$

Tässä kannassa on n vektoria. Siten avaruuden $\mathbb{R}^n$ dimensio on n, eli $\mathbb{R}^n$ on n-ulotteinen.

7 Aliavaruudet

Luvussa 4 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 $\mathbb{R}^2$ vektoreita ei voida kirjoittaa näiden kahden vektorin lineaarikombinaatioina. (Näiden väitteiden todistaminen jätetään lukijalle.)

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 vektoriavaruusta $\mathbb{R}^3$. Tällaisia avaruuksia kutsutaan aliavaruuksiksi. Annetaan seuraavaksi täsmällinen määritelmä vektorien virittämälle aliavaruudelle, ja tutkitaan, miltä tällaiset aliavaruudet näyttävät.

Vektorien virittämä aliavaruus koostuu kaikista kyseisten vektorien lineaarikombinaatioista.

Määritelmä 7.1. Vektoreiden $\bar{v}_1, \ldots, \bar{v}_k \in \mathbb{R}^n$ virittämä *aliavaruus* on joukko

$$\{a_1\bar{v}_1 + a_2\bar{v}_2 + \cdots + a_k\bar{v}_k \mid a_1, a_2, \ldots, a_k \in \mathbb{R}\}.$$

Tätä joukkoa merkitään $\operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$.

Jos $W = \operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$, sanotaan, että vektorit $\bar{v}_1, \ldots, \bar{v}_k \in \mathbb{R}^n$ virittävät aliavaruuden W. Vektoreita $\bar{v}_1, \ldots, \bar{v}_k$ kutsutaan aliavaruuden W *virittäjiksi*.

Vektorien virittämää aliavaruutta kutsutaan toisinaan lyhyesti aliavaruudeksi. Huomaa, että merkinnässä $\operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$ 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.

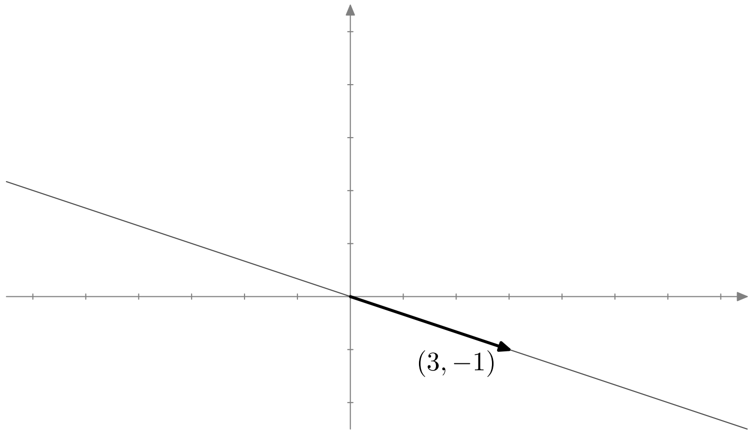
Esimerkki 7.2. Tarkastellaan vektorin $(3, -1)$ virittämää aliavaruutta $\operatorname{span}((-3, 1))$. Se koostuu määritelmän mukaan kaikista vektorin $(3, -1)$ lineaarikombinaatioista. Koska vektoreita on vain yksi, ovat lineaarikombinaatiot itse asiassa skalaarimonikertoja. Vektorin $(-3, 1)$ virittämä aliavaruus on

$$\operatorname{span}((-3, 1)) = \{a(3, -1) \mid a \in \mathbb{R}\}.$$

Tutkitaan, miltä aliavaruus $\operatorname{span}((-3, 1))$ 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 $\operatorname{span}((-3, 1))$ alkioita ajatellaan koordinaatiston pisteinä, huomataan pisteiden sijaitsevat samalla suoralla (kuva 7.17). Aliavaruus on $\operatorname{span}((-3, 1))$ siis suora. Koska $(0, 0)$ on tämän suoran alkio, kulkee suora origon kautta.

Esimerkki 7.3. Tarkastellaan seuraavaksi, miltä näyttää vektorien $(-3, 1, 1)$ ja $(2, -1, 2)$ virittämä aliavaruus $\operatorname{span}((-3, 1, 1), (2, -1, 2))$. Se on vektoriavaruuden $\mathbb{R}^3$ osajoukko. Vektoreiden lineaarikombinaatiot muodostavat joukon

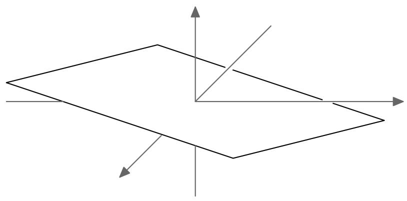
$$\operatorname{span}((-3, 1, 1), (2, -1, 2)) = \{a_1(-3, 1, 1) + a_2(2, -1, 2) \mid a_1, a_2 \in \mathbb{R}\}.$$



Kuva 7.17: Aliavaruus $\mathrm{span}((3,-1))$ on origon kautta kulkeva suora.

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 $\mathrm{span}((-3,1,1),(2,-1,2))$ on hieman vaikeampi hahmotella kuvaa kuin aliavaruudesta $\mathrm{span}((-3,1))$. Kaikki vektorien $(-3,1,1)$ ja $(2,-1,2)$ lineaarikombinaatiot ovat kuitenkin samassa tasossa kuin $(-3,1,1)$ ja $(2,-1,2)$. Joukko $\mathrm{span}((-3,1,1),(2,-1,2))$ 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 7.18: Aliavaruus $\mathrm{span}((-3,1,1),(2,-1,2))$ 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 $\mathrm{span}(\bar{0}) = \{a\bar{0} \mid a \in \mathbb{R}\} = \{\bar{0}\}$. Tässä aliavaruudessa on siis ainoastaan nollavektori. Myös koko avaruus $\mathbb{R}^3$ on eräiden vektoreiden virittämä aliavaruus:

$$\begin{aligned}\mathrm{span}((1,0,0),(0,1,0),(0,0,1)) &= \{a_1(1,0,0) + a_2(0,1,0), a_3(0,0,1) \mid a_1,a_2,a_3 \in \mathbb{R}\} \\ &= \{(a_1,a_2,a_3) \mid a_1,a_2,a_3 \in \mathbb{R}\} = \mathbb{R}^3.\end{aligned}$$

Esimerkki 7.4. Tutkitaan, miltä näyttävät avaruuden $\mathbb{R}^4$ aliavaruuden

$$\mathrm{span}((1,0,-2,5),(0,-1,4,0),(0,0,0,1))$$

alkiot. Määritelmän mukaan

$$\begin{aligned}
&\operatorname{span}((1,0,-2,5),(0,-1,4,0),(0,0,0,1)) \\
&\qquad = \{a_1(1,0,-2,5)+a_2(0,-1,4,0)+a_3(0,0,0,1) \mid a_1,a_2,a_3 \in \mathbb{R}\} \\
&\qquad = \{(a_1,\ 0,\ -2a_1,\ 5a_1)+(0,\ -a_2,\ 4a_2,\ 0)+(0,\ 0,\ 0,\ a_3) \mid a_1,a_2,a_3 \in \mathbb{R}\} \\
&\qquad = \{(a_1,\ -a_2,\ -2a_1+4a_2,\ 5a_1+a_3) \mid a_1,a_2,a_3 \in \mathbb{R}\}.
\end{aligned}$$

Aliavaruuden $\operatorname{span}((1,0,-2,2),(0,-1,4,5),(0,0,0,1))$ alkiot ovat siis muotoa

$$(a_1,\ -a_2,\ -2a_1+4a_2,\ 5a_1+a_3),$$

missä $a_1,a_2,a_3 \in \mathbb{R}$.

Esimerkki 7.5. 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 7.4, mutta käännetään päättelyn suunta:

$$\begin{aligned}
W &= \{(4a_1-2a_2,\ 3a_1-a_2,\ 5a_1+a_2) \mid a_1,a_2 \in \mathbb{R}\} \\
&= \{(4a_1,3a_1,5a_1)+(-2a_2,-a_2,a_2) \mid a_1,a_2 \in \mathbb{R}\} \\
&= \{a_1(4,3,5)+a_2(-2,-1,1) \mid a_1,a_2 \in \mathbb{R}\} \\
&= \operatorname{span}((4,3,5),(-2,-1,1)).
\end{aligned}$$

Kyseessä on siis vektorien $(4,3,5)$ ja $(-2,-1,1)$ virittämä aliavaruus. Se on origon kautta kulkeva taso.

Eri vektorit saattavat virittää saman aliavaruuden. Tätä tutkitaan seuraavassa esimerkissä.

Esimerkki 7.6. Aliavaruudella $W = \operatorname{span}((3,9,-3),(1,5,0))$ on virittäjävektorit $(3,9,-3)$ ja $(1,5,0)$. Aliavaruudelle W on kuitenkin mahdollista löytää paljon muitakin virittäjiä. Tässä esimerkissä osoitetaan, että myös vektorit $(1,3,-1)$ ja $(0,2,1)$ virittävät aliavaruuden W. Tämä tarkoittaa, että vektoreiden $(3,9,-3)$ ja $(1,5,0)$ lineaarikombinaatioiden muodostama joukko on täsmälleen sama vektorit kuin vektoreiden $(1,3,-1)$ ja $(0,2,1)$ lineaarikombinaatioiden muodostama joukko.

On siis näytettävät, että $\operatorname{span}((3,9,-3),(1,5,0)) = \operatorname{span}((1,3,-1),(0,2,1))$. Kaksi joukkoa osoitetaan samoiksi näyttämällä, että kumpikin on toisen osajoukko.

"$\subset$": Osoitetaan ensin, että $\operatorname{span}((3,9,-3),(1,5,0)) \subset \operatorname{span}((1,3,-1),(0,2,1))$. Tämä tehdään näyttämällä, että jokainen joukon $\operatorname{span}((3,9,-3),(1,5,0))$ alkio on myös alkiona joukossa $\operatorname{span}((1,3,-1),(0,2,1))$. Oletetaan siis, että $\bar{b} \in \operatorname{span}((3,9,-3),(1,5,0))$. Tällöin olemassa reaaliluvut a_1 ja a_2, joille pätee

$$\bar{b} = a_1(3,9,-3)+a_2(1,5,0). \tag{6}$$

Tavoitteena on osoittaa, että $\bar{b} \in \operatorname{span}((1,3,-1),(0,2,1))$. On siis näytettävä, että $\bar{b}$ on vektoreiden $(1,3,-1)$ ja $(0,2,1)$ lineaarikombinaatio. Virittäjävektorit voidaan kirjoittaa näiden vektoreiden lineaarkombinaatioina:

$$(3,9,-3) = 3(1,3,-1) \quad \text{ja} \quad (1,5,0) = (1,3,-1)+(0,2,1).$$

Sijoitetaan lineaarikombinaatiot yhtälöön (6) ja sievennetään:

$$\begin{aligned}\bar{b} &= a_1(3,9,-3) + a_2(1,5,0) = a_1\big((3(1,3,-1))\big) + a_2\big((1,3,-1) + (0,2,1)\big) \\ &= (3a_1 + a_2)(1,3,-1) + a_2(0,2,1).\end{aligned}$$

Näin ollen $\bar{b}$ on vektoreiden $(1,3,-1)$ ja $(0,2,1)$ lineaarikombinaatio, eli $\bar{b} \in \operatorname{span}((1,3,-1),(0,2,1))$.

"$\supset$": Oletetaan, että $\bar{b} \in \operatorname{span}((1,3,-1),(0,2,1))$. Nyt olemassa reaaliluvut a_1 ja a_2, joille pätee

$$\bar{b} = a_1(1,3,-1) + a_2(0,2,1).$$

Osoitetaan, että $\bar{b} \in \operatorname{span}((3,9,-3),(1,5,0))$. On siis näytettävät, että $\bar{b}$ on vektoreiden $(3,9,-3)$ ja $(1,5,0)$ lineaarikombinaatio. Huomataan, että

$$(1,3,-1) = \frac{1}{3}(3,9,-3) \quad \text{ja} \quad (0,2,1) = -\frac{1}{3}(3,9,-3) + (1,5,0).$$

Siten

$$\begin{aligned}\bar{b} &= a_1(1,3,-1) + a_2(0,2,1) = a_1\left(\frac{1}{3}(3,9,-3)\right) + a_2\left(-\frac{1}{3}(3,9,-3) + (1,5,0)\right) \\ &= \left(\frac{1}{3}a_1 - \frac{1}{3}a_2\right)(3,9,-3) - a_2(1,5,0).\end{aligned}$$

Näin ollen $\bar{b} \in \operatorname{span}((3,9,-3),(1,5,0))$.

Toisinaan kaikkia virittäjävektoreita ei tarvita aliavaruuden virittämiseen. Seuraava lause osoittaa, että jos jokin virittäjävektori on toisten virittäjävektoreiden lineaarikombinaatio, se voidaan pudottaa pois virittäjävektoreiden joukosta.

Lause 7.7. *Oletetaan, että $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k \in \mathbb{R}^n$. Oletetaan lisäksi, että $\bar{w}$ on vektoreiden $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k$ lineaarikombinaatio. Tällöin*

$$\operatorname{span}(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k, \bar{w}) = \operatorname{span}(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k).$$

Todistus. "$\subset$": Oletetaan, että $\bar{b} \in \operatorname{span}(\bar{v}_1, \dots, \bar{v}_k, \bar{w})$. Nyt on olemassa reaalilukukertoimet $a_1, \dots, a_k, a_w \in \mathbb{R}$, joille pätee

$$\bar{b} = a_1\bar{v}_1 + \dots + a_k\bar{v}_k + a_w\bar{w}.$$

Toisaalta koska $\bar{w}$ on vektoreiden $\bar{v}_1, \dots, \bar{v}_k$ lineaarikombinaatio, on olemassa $c_1, \dots, c_k \in \mathbb{R}$, joille pätee

$$\bar{w} = c_1\bar{v}_1 + \dots + c_k\bar{v}_k.$$

Sijoitetaan tämä ensimmäiseen yhtälöön, jolloin saadaan

$$\begin{aligned}\bar{b} &= a_1\bar{v}_1 + \dots + a_k\bar{v}_k + a_w(c_1\bar{v}_1 + \dots + c_k\bar{v}_k) \\ &= a_1\bar{v}_1 + \dots + a_k\bar{v}_k + a_wc_1\bar{v}_1 + \dots + a_wc_k\bar{v}_k \\ &= (a_1 + a_wc_1)\bar{v}_1 + \dots + (a_k + a_wc_k)\bar{v}_k.\end{aligned}$$

Tästä nähdään, että $\bar{b} \in \text{span}(\bar{v}_1, \ldots, \bar{v}_k)$, joten

$$\text{span}(\bar{v}_1, \ldots, \bar{v}_k, \bar{w}) \subset \text{span}(\bar{v}_1, \ldots, \bar{v}_k).$$

"$\supset$": Todistuksen toinen osa jätetään harjoitustehtäväksi. $\square$

Tarkastellaan vektoreita $(-1, 0, -3)$, $(-6, 1, -1)$ ja $(-5, 1, 2)$. Vektoreista keskimmäinen on kahden muun summa. Siten edellisen lauseen nojalla

$$\text{span}((-1, 0, -3), (-6, 1, -1), (-5, 1, 2)) = \text{span}((-1, 0, -3), (-5, 1, 2)).$$

Pohditaan vielä vapaamuotoisesti, mistä tämä johtuu. Vektori $(-6, 1, -1)$ on muiden virittäjävektoreiden lineaarikombinaatio. Siksi kaikki, mikä saadaan aikaiseksi vektoria $(-6, 1, -1)$ käyttäen, voidaan saada aikaiseksi jo muilla virittäjävektoreilla. Vektori $(-6, 1, -1)$ ei siis tuota aliavaruuteen mitään uusia vektoreita, joten sen voi jättää pois virittäjävektorien joukosta ilman, että aliavaruus muuttuu miksikään.

Seuraavassa esimerkissä tutkitaan, kuinka aliavaruuden ovat pienenmpiä vektoriavaruuksia toisten vektoriavaruuksien sisässä.

Esimerkki 7.8. 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 alkiota lasketaan yhteen. Oletetaan, että $\bar{v}, \bar{w} \in W$. Tällöin on olemassa sellaiset reaaliluvut a ja b, että $\bar{v} = a(-2, -1)$ ja $\bar{w} = b(-2, -1)$. Nähdään, että

$$\bar{v} + \bar{w} = a(-2, -1) + b(-2, -1) = (a + b)(-2, -1),$$

missä $a + b \in \mathbb{R}$. Havaitaan, että summa $\bar{v} + \bar{w}$ on vektorin $(-2, -1)$ skalaarimonikerta, joten se on aliavaruuden W alkio. Jos lasketaan yhteen mitkä tahansa kaksi aliavaruuden $W = \text{span}((-2, -1))$ alkiota, on tuloksena siis edelleen aliavaruuden W alkio.

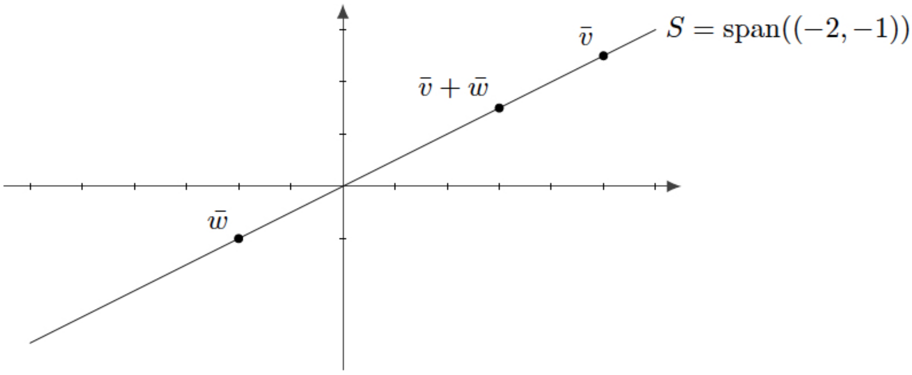
Tarkastellaan sitten aliavaruuden alkioiden skalaarimonikertoja. Oletetaan, että $\bar{u} \in W$ ja $k \in \mathbb{R}$. Tällöin on olemassa $c \in \mathbb{R}$, jolle pätee $\bar{u} = c(-2, -1)$. Huomataan, että

$$k\bar{u} = k(c(-2, -1)) = (kc)(-2, -1),$$

missä $kc \in \mathbb{R}$. Havaitaan, että vektori $k\bar{u}$ voidaan kirjoittaa vektorin $(-2, -1)$ skalaarimonikertana, joten $k\bar{u} \in W$. 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. Sama pätee origon kautta kulkeviin tasoihin.

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.



Kuva 7.19: Suoran $W = \mathrm{span}((-2,-1))$ alkioiden $\bar{v}$ ja $\bar{w}$ summa $\bar{v}+\bar{w}$ on aliavaruuden W alkio.

Lause 7.9. *Oletetaan, että* $\bar{v}_1,\dots,\bar{v}_k \in \mathbb{R}^n$. *Olkoon* $W = \mathrm{span}(\bar{v}_1,\dots,\bar{v}_k)$. *Tällöin seuraavat väitteet pätevät:*

a) Jos $\bar{u}$, $\bar{w} \in W$, *niin* $\bar{u}+\bar{w} \in W$.

b) Jos $\bar{w} \in W$ *ja* $c \in \mathbb{R}$, *niin* $c\bar{w} \in W$.

c) $\bar{0} \in W$.

Todistus. Osoitetaan kohta a) ja jätetään loput kohdat harjoitustehtäviksi. Oletetaan, että $\bar{u},\bar{w} \in W$. Nyt $\bar{u} = a_1\bar{v}_1 + \cdots + a_k\bar{v}_k$ joillakin $a_1,\dots,a_k \in \mathbb{R}$ ja $\bar{w} = b_1\bar{v}_1 + \cdots + b_k\bar{v}_k$ joillakin $b_1,\dots,b_k \in \mathbb{R}$. Osoitetaan, että summa $\bar{u}+\bar{w}$ on aliavaruuden W alkio. Huomataan, että

$$\begin{aligned}\bar{u}+\bar{w} &= (a_1\bar{v}_1 + \cdots + a_k\bar{v}_k) + (b_1\bar{v}_1 + \cdots + b_k\bar{v}_k)\\ &= (a_1+b_1)\bar{v}_1 + \cdots + (a_k+b_k)\bar{v}_k.\end{aligned}$$

Koska $\bar{u}+\bar{w}$ on vektoreiden $\bar{v}_1,\bar{v}_2,\dots,\bar{v}_k$ lineaarikombinaatio, pätee $\bar{u}+\bar{w} \in W$. □

7.1 Aliavaruuden kanta

Samalla tavalla kuin vektoriavaruudelle määriteltiin kanta voidaan määritellä kanta aliavaruudelle.

Määritelmä 7.10. Oletetaan, että W on vektoriavaruuden $\mathbb{R}^n$ aliavaruus, johon kuuluvat vektorit $\bar{w}_1,\bar{w}_2,\dots,\bar{w}_k$. Vektorijono $(\bar{w}_1,\bar{w}_2,\dots,\bar{w}_k)$ on aliavaruuden W *kanta*, jos seuraavat ehdot pätevät:

a) vektorit $\bar{w}_1,\bar{w}_2,\dots,\bar{w}_k$ virittävät aliavaruuden W

b) jono $(\bar{w}_1,\bar{w}_2,\dots,\bar{w}_k)$ on vapaa.

Esimerkki 7.11. 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 5.6). Siten vektorit $(3,9,-3)$ ja $(1,5,1)$ muodostavat aliavaruuden W kannan.

Esimerkki 7.12. 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

$$x_1(3,-1,5) + x_2(2,1,3) + x_3(0,-5,1) = (0,0,0), \tag{7}$$

missä $x_1, x_2, x_3 \in \mathbb{R}$. Yhtälöstä saadaan yhtälöryhmä, jota vastaa matriisi

$$\left[\begin{array}{rrr|r} 3 & 2 & 0 & 0 \\ -1 & 1 & -5 & 0 \\ 5 & 3 & 1 & 0 \end{array}\right].$$

Tästä saadaan alkeisrivitoimituksilla matriisi

$$\left[\begin{array}{lrr|r} 1 & -1 & 5 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Porrasmatriisista nähdään, että yhtälöryhmällä on äärettömän monta ratkaisua, joten jono $((3,-1,5),(2,1,3),(0,-5,1))$ ei ole vapaa eikä siis myöskään aliavaruuden kanta.

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 (7) 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 7.7 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 5.6 nojalla. Siten $((3,-1,5),(2,1,3))$ on aliavaruuden W kanta.

Samaa menetelmää käyttäen voidaan aina löytää vektorien virittämälle aliavaruudelle kanta.

Aliavaruudella voi olla monta erilaista kantaa. Esimerkissä 7.11 osoitettiin, että aliavaruudella $W = \text{span}((3,9,-3),(1,5,0))$ on kanta $((3,9,-3),(1,5,0))$. Toisaalta esimerkin 7.6 nojalla aliavaruuden W virittävät myös vektorit $(1,3,-1)$ ja $(0,2,1)$. Nämä vektorit ovat lineaarisesti riippumattomia, joten ne muodostavat aliavaruuden kannan. Siten aliavaruudella W on myös kanta $((1,3,-1),(0,2,1))$. Vaikka kannoissa on eri vektorit, molemmissa kannoissa on kuitenkin kaksi vektoria. Kantavektoreita on siis yhtä paljon.

Lause 7.13. *Aliavaruuden W jokaisessa kannassa on yhtä monta vektoria.*

Todistus. Todistus on samankaltainen kuin lauseen 6.5 todistus. □

7.2 Aliavaruuden dimensio

Alaluvussa 6.2 käsiteltiin vektoriavaruuden $\mathbb{R}^n$ dimensiota eli ulottuvuutta ja määriteltiin sen olevan avaruuden kannan vektorien lukumäärä. Samalla tavalla voidaan määritellä aliavaruuden dimensio.

Esimerkissä 7.11 tutkittiin aliavaruutta $W = \text{span}((3, 9, -3), (1, 5, 1))$, joka on taso. Tuntuisi luonnolliselta ajatella, että sen dimensio olisi kaksi. Esimerkissä nähtiin, että aliavaruudella W on kanta $((3, 9, -3), (1, 5, 1))$, jossa on kuin onkin kaksi vektoria. Dimensio on siis kaksi.

Määritelmä 7.14. Aliavaruuden W *dimensio* $\dim(W)$ on aliavaruuden W kannan vektoreiden lukumäärä. Jos aliavaruuden dimensio on n, sanotaan, että aliavaruus on *n-ulotteinen*.

Lauseen 7.13 nojalla aliavaruuden jokaisessa kannassa on aina yhtä monta vektoria, joten dimension voi määritellä edellä mainitulla tavalla.

Esimerkki 7.15. 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 5.6 mukaan ole lineaarisesti riippumattomia kuten kantavektorien pitäisi olla.

Koska toinen virittäjävektori on ensimmäisen skalaarimonikerta, lauseen 7.7 nojalla

$$W = \text{span}((1, 3, -1), (-3, -9, 3)) = \text{span}((1, 3, -1)).$$

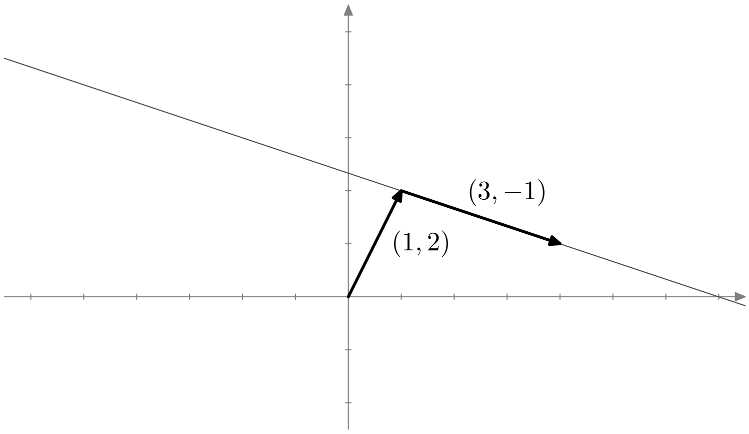
Yhden nollavektorista poikkeavan vektorin muodostama jono on aina vapaa. Näin ollen aliavaruuden W kannan muodostaa vektori $(1, 3, -1)$, joten $\dim(W) = 1$.

8 Suorat ja tasot

8.1 Suora

Edellisessä luvusssa tutkittiin vektorien virittämiä aliavaruuksia ja huomattiin, että ne voivat olla esimerkiksi origon kautta kulkevia suoria ja tasoja. Tässä luvussa tutustutaan hieman yleisemmin suoriin ja tasoihin. Tarkastelun kohteena ovat myös sellaiset suorat ja tasot, jotka eivät kulje origon kautta eivätkä siten ole aliavaruuksia.

Luvussa 7 huomattiin, että joukko $\{t(3,-1) \mid t \in \mathbb{R}\}$ muodostaa origon kautta kulkevan suoran. Jos halutaan muodostaa vektorimerkintöjen avulla suora, joka ei kulje origon kautta, on suora ensin ”siirrettävä” haluttuun paikkaan. Esimerkiksi suora $\{(1,2)+t(3,-1) \mid t \in \mathbb{R}\}$ ei kulje origon kautta. Tämän suoran pisteet saadaan lisäämällä vektoriin $(1,2)$ vektoreita, jotka ovat yhdensuuntaisia vektorin $(3,-1)$ kanssa. Tilannetta on havainnollistettu kuvassa 8.20.



Kuva 8.20: Suora $\{(1,2) + t(3,-1) \mid t \in \mathbb{R}\}$.

Määritelmä 8.1. Vektoriavaruuden $\mathbb{R}^n$ *suora* on joukko

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

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

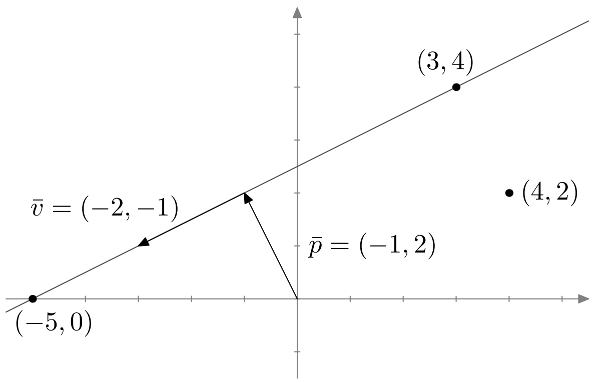
Huomaa, että yllä ei ole annettu mitä tahansa suoran kuvailua, vaan suoran *määritelmä*. Emme siis ole lähteneet esimerkiksi jostakin suoran geometrisestä muotoilusta ja ilmaisseet saman asian vektoreilla, vaan määritelleet suoran käsitteen tämän kurssin tarpeita varten.

Määritelmän mukaan suoran alkiot ovat vektoreita. Vektoreita voidaan kuitenkin havainnollistaa pisteinä, joten voimme kutsua suoran alkoita myös pisteiksi. Olkoon S vektoriavaruuden $\mathbb{R}^2$ suora. 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 8.2. Esimerkiksi joukko

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

on suora vektoriavaruudessa $\mathbb{R}^2$. Se on piirretty kuvaan 8.21. Määritelmän mukaan mikä tahansa suoran S piste voidaan kirjoittaa summana vektorista $\bar{p} = (-1, 2)$ ja jostakin vektorin $\bar{v} = (-2, -1)$ skalaarimonikerrasta. Esimerkiksi $(-5, 0) = \bar{p} + 2\bar{v}$ ja $(3, 4) = \bar{p} - 2\bar{v}$, joten $(-5, 0) \in S$ ja $(3, 4) \in S$. Siis suora S kulkee pisteiden $(-5, 0)$ ja $(3, 4)$ kautta.



Kuva 8.21: Suoran S paikkavektori on $\bar{p} = (-1, 2)$ ja suuntavektori $\bar{v} = (-2, -1)$.

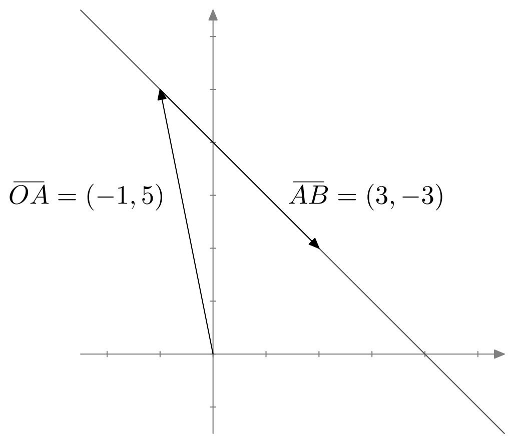
Toisaalta piste $(4, 2)$ ei ole suoralla S. Jos nimittäin $(4, 2) = (-1, 2) + t(-2, -1)$ jollakin $t \in \mathbb{R}$, niin $4 = -1 - 2t$ ja $2 = 2 - t$. Ensimmäisen yhtälön perusteella $t = -5/2$ ja toisen perusteella $t = 0$. Tämä on mahdotonta, joten ei ole olemassa sellaista lukua $t \in \mathbb{R}$,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

Ryhdytään seuraavaksi määrittämään suoraa, joka kulkee annettujen pisteiden kautta.

Esimerkki 8.3. 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 $\overline{OA} = (-1, 5)$. (Merkintä $\overline{OA}$ on esitelty alaluvussa 1.2.) Suuntavektoriksi käy mikä tahansa suoran kanssa yhdensuuntainen vektori, esimerkiksi vektori

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

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



Kuva 8.22: 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

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

paikkavektoriksi $\bar{p}$ voidaan valita suoran minkä tahansa pisteen paikkavektori, ja suuntavektoriksi $\bar{v}$ voidaan valita mikä tahansa suoran suuntainen vektori. Väitteen todistusta ei esitetä tässä, mutta asiaan palataan lauseessa 8.5. Seuraava esimerkki havainnollistaa asiaa.

Esimerkki 8.4. Tutkitaan esimerkin 8.2 suoraa $S = \{(-1, 2) + t(-2, -1) \mid t \in \mathbb{R}\}$. Tulemme näkemään, että 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 $(-4, -2) = 2(-2, -1)$.

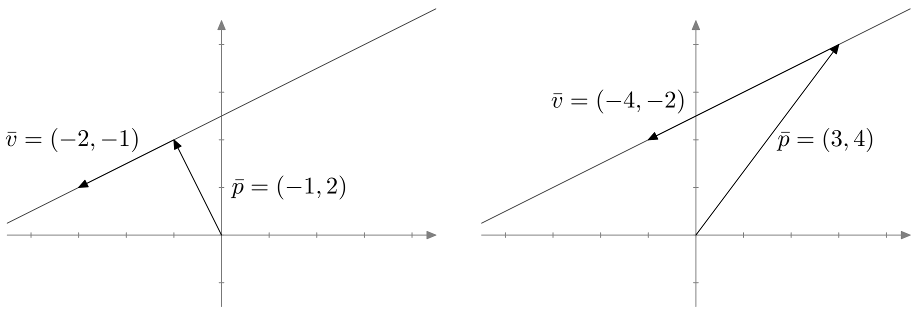
Suora S voidaan siis kirjoittaa muodossa

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

Vaikka tämä joukko onkin äkkiseltään katsottuna erilainen kuin suoran S alkuperäinen määritelmä, on joukoissa täsmälleen samat alkiot. Asiaa on havainnollistettu kuvassa 8.23.

Kuvasta katsominen ei vielä todista suoria samoiksi. Osoitetaan huolellisesti, että joukot $S = \{(-1, 2) + t(-2, -1) \mid t \in \mathbb{R}\}$ ja $S' = \{(3, 4) + s(-4, -2) \mid s \in \mathbb{R}\}$ ovat samat. Kaksi joukkoa osoitetaan samoiksi näyttämällä, että kumpikin on toisen osajoukko.

”$\subset$”: Osoitetaan ensin, että $S \subset S'$. Tämä tehdään näyttämällä, että jokainen joukon S alkio on joukossa S'. Oletetaan, että $\bar{a} \in S$. Tällöin joukon S määritelmän perusteella pätee $\bar{a} = (-1, 2) + t(-2, -1)$ jollakin $t \in \mathbb{R}$. On osoitettava, että $\bar{a} \in S'$. Tavoitteena on siis kirjoittaa



Kuva 8.23: Suoran paikkavektori ja suuntavektori eivät ole yksikäsitteisiä.

vektori $\bar{a} = (-1, 2) + t(-2, -1)$ muodossa $(3, 4) + s(-4, -2)$, missä s on jokin reaaliluku. Ryhdytään muokkaamaan vektoria $\bar{a}$ haluttuun muotoon. Huomataan, että

$$\begin{aligned}\bar{a} &= (-1, 2) + t(-2, -1) = (3, 4) + (-4, -2) + t(-2, -1) \\ &= (3, 4) + 2(-2, -1) + t(-2, -1) = (3, 4) + (2 + t)(-2, -1) \\ &= (3, 4) + \frac{2 + t}{2}(-4, -2).\end{aligned}$$

Koska $(2 + t)/2 \in \mathbb{R}$, viimeisestä muodosta nähdään, että $\bar{a} \in S'$. (Joukon S' määritelmässä voidaan valita $s = (2 + t)/2$.) On siis osoitettu, että $S \subset S'$.

"$\supset$": Osoitetaan sitten, että $S' \subset S$, eli näytetään, että jokainen joukon S' alkio on joukossa S. Oletetaan, että $\bar{a} \in S'$. Nyt $\bar{a} = (3, 4) + s(-4, -2)$ jollakin $s \in \mathbb{R}$. On osoitettava, että $\bar{a} \in S$. Tavoitteena on siis kirjoittaa vektori $\bar{a}$ muodossa $(-1, 2) + t(-2, -1)$, missä t on jokin reaaliluku. Huomataan, että

$$\begin{aligned}\bar{a} &= (3, 4) + s(-4, -2) = (-1, 2) + (4, 2) + s(-4, -2) \\ &= (-1, 2) + (-1)(-4, -2) + s(-4, -2) = (-1, 2) + (-1 + s)(-4, -2) \\ &= (-1, 2) + 2(-1 + s)(-2, -1) = (-1, 2) + (-2 + 2s)(-2, -1).\end{aligned}$$

Koska $-2 + 2s \in \mathbb{R}$, tiedetään, että $\bar{a} \in S$. Näin on osoitettu, että $S' \subset S$.

Osoitetaan sitten yleisessä tapauksessa, että paikkavektoriksi voidaan valita suoran mikä tahansa alkio ja ja suuntavektoriksi mikä tahansa vektori, joka on yhdensuuntainen suoran suuntavektorin kanssa. Todistus mukailee edellistä esimerkkiä.

Lause 8.5. *Olkoon $S = \{\bar{p} + t\bar{v} \mid t \in \mathbb{R}\}$ vektoriavaruuden $\mathbb{R}^n$ suora. Oletetaan, että $\bar{q} \in S$ ja $\bar{w}$ on yhdensuuntainen vektorin $\bar{v}$ kanssa. Tällöin $S = \{\bar{q} + t\bar{w} \mid t \in \mathbb{R}\}$*

Todistus. Merkitään $S' = \{\bar{q} + t\bar{w} \mid t \in \mathbb{R}\}$ ja osoitetaan, että $S = S'$. Koska vektorit $\bar{v}$ ja $\bar{w}$ ovat yhdensuuntaisia, on olemassa $k \in \mathbb{R} \setminus \{0\}$, jolle pätee $\bar{v} = k\bar{w}$. Tästä seuraa, että $\bar{w} = (1/k)\bar{v}$. Toisaalta $\bar{q} \in S$, joten $\bar{q} = \bar{p} + b\bar{v}$ jollakin $b \in \mathbb{R}$. Tästä seuraa, että $\bar{p} = \bar{q} - b\bar{v} = \bar{q} - bk\bar{w}$

"$\subset$": Oletetaan, että $\bar{a} \in S$. Joukon S määritelmän perusteella $\bar{a} = \bar{p} + t\bar{v}$ jollakin $t \in \mathbb{R}$. Huomataan, että

$$\bar{a} = \bar{p} + t\bar{v} = \bar{q} - bk\bar{w} + t(k\bar{w}) = \bar{q} - bk\bar{w} + tk\bar{w} = \bar{q} + (-bk + tk)\bar{w} \in S'$$

On siis osoitettu, jokainen joukon S alkio on joukossa S'. Toisin sanoen $S \subset S'$.

"$\supset$": Oletetaan, että $\bar{a} \in S'$. Nyt $\bar{a} = \bar{q} + t\bar{w}$ jollakin $t \in \mathbb{R}$. Huomataan, että

$$\bar{a} = \bar{q} + t\bar{w} = \bar{p} + b\bar{v} + t((1/k)\bar{v}) = \bar{p} + b\bar{v} + (t/k)\bar{v} = \bar{p} + (b + t/k)\bar{v} \in S.$$

Näin on osoitettu, että $S' \subset S$.

Siten $S = S'$. □

8.2 Taso

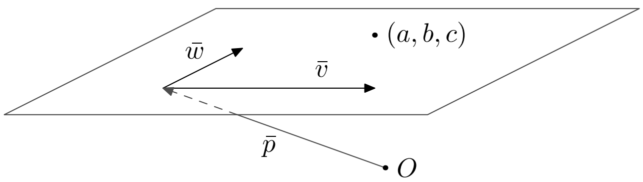
Kolmiulotteisessa vektoriavaruudessa voidaan määritellä tasot samaan tapaan kuin suorat. Nyt tarvitaan kuitenkin kaksi suuntavektoria, eivätkä ne saa olla yhdensuuntaisia.

Määritelmä 8.6. Vektoriavaruuden $\mathbb{R}^n$ *taso* on joukko

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

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

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 8.24. Vastaavia ilmauksia käytetään vektoriavaruudessa $\mathbb{R}^n$.



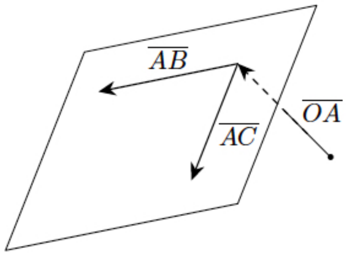
Kuva 8.24: Piste (a, b, c) on tasossa T.

Esimerkki 8.7. 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 $\overline{OA} = (0, 1, 0)$ käy tähän tarkoitukseen. Lisäksi tarvitaan tason suuntaiset suuntavektorit:

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

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

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



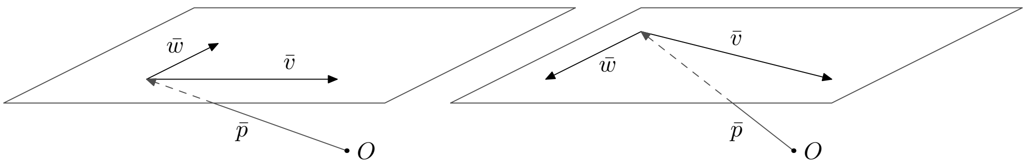
Kuva 8.25: Taso T kulkee pisteiden A, B ja C kautta.

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

$$\begin{aligned} A &= (0,1,0) + 0(-1,2,2) + 0(-2,-1,1), \\ B &= (0,1,0) + 1(-1,2,2) + 0(-2,-1,1) \quad \text{ja} \\ C &= (0,1,0) + 0(-1,2,2) + 1(-2,-1,1). \end{aligned}$$

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

Edellisessä käytetty menetelmä toimii yleisemminkin. Jos taso halutaan kirjoittaa muodossa $\{\bar{p} + s\bar{w} + t\bar{v} \mid s,t \in \mathbb{R}\}$, paikkavektoriksi $\bar{p}$ voidaan valita tason minkä tahansa pisteen paikkavektori. Suuntavektoreiksi $\bar{w}$ ja $\bar{v}$ voidaan valita mitkä tahansa tason suuntaiset vektorit, kunhan $\bar{w}$ ja $\bar{v}$ eivät ole yhdensuuntaiset. Taso on siis mahdollista kirjoittaa usealla eri tavalla joukkona $\{\bar{p} + s\bar{w} + t\bar{v} \mid s,t \in \mathbb{R}\}$. Tilannetta on havainnollistettu kuvassa 8.26.



Kuva 8.26: Taso voidaan kirjoittaa eri tavoin joukkona $\{\bar{p} + s\bar{w} + t\bar{v} \mid s,t \in \mathbb{R}\}$.

Esimerkki 8.8. Tutkitaan tasoa $T = \{a_1(3,-1,5) + a_2(2,1,3) \mid a_1, a_2 \in \mathbb{R}\}$ ja kirjoitetaan se hieman toisenlaisessa muodossa.

Oletetaan, että $\bar{u} = (u_1, u_2, u_3) \in \mathbb{R}^3$. Selvitetään, mikä ehto vektorin $\bar{u}$ komponenttien pitää toteuttaa, jotta $\bar{u}$ on tason T vektori. Ratkaistaan yhtälö

$$a_1(3,-1,5) + a_2(2,1,3) = \bar{u}.$$

Sitä vastaa yhtälöryhmä

$$\begin{cases} 3a_1 + 2a_2 = u_1 \\ -a_1 + a_2 = u_2 \\ 5a_1 + 3a_2 = u_3. \end{cases}$$

Yhtälöryhmän matriisi on

$$\left[\begin{array}{cc|c} 3 & 2 & u_1 \\ -1 & 1 & u_2 \\ 5 & 3 & u_3 \end{array}\right],$$

ja se saadaan alkeisrivitoimituksilla muutettua matriisiksi

$$\left[\begin{array}{cc|c} 1 & -1 & -u_2 \\ 0 & 1 & \frac{1}{5}u_1 + \frac{3}{5}u_2 \\ 0 & 0 & -\frac{8}{5}u_1 + \frac{1}{5}u_2 + u_3 \end{array}\right].$$

Yhtälöryhmällä ratkaisu, jos ja vain jos matriisin alinta riviä vastaava yhtälö

$$0 = -\frac{8}{5}u_1 + \frac{1}{5}u_2 + u_3$$

on tosi eli $-8u_1 + u_2 + 5u_3 = 0$. Toisin sanoen, vektori $\bar{u}$ on tasossa T, jos ja vain jos $-8u_1 + u_2 + 5u_3 = 0$. Siten voidaan kirjoittaa

$$T = \{(u_1, u_2, u_3) \in \mathbb{R}^3 \mid -8u_1 + u_2 + 5u_3 = 0\}$$

Yhtälöä $-8u_1 + u_2 + 5u_3 = 0$ kutsutaan tason T *normaalimuotoiseksi yhtälöksi*. Yhtälö määrittää täysin tason T, sillä tasossa olevat pisteet ovat täsmälleen ne pisteet, joiden komponentit toteuttavat yhtälön. Aiheeseen palataan alaluvussa 13.4.

9 Matriisit

Aiemmissa luvuissa matriiseja on käsitelty siinä määrin kuin on ollut tarpeellista yhtälönratkaisun kannalta. Matriiseja käytetään kuitenkin myös muihin tarkoituksiin, ja siksi on hyödyllistä selvittää, minkälaisia ominaisuuksia matriiseilla itsellään on. Samalla saadaan joitakin uusia tapoja myös yhtälöryhmien ratkaisujen tutkimiseen.

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

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

on $m \times n$ -matriisi. Voidaan myös sanoa, että matriisin A *tyyppi* on $m \times n$. Kaikkien reaalikertoimisten $m \times n$ -matriisien joukkoa merkitään $\mathbb{R}^{m \times n}$. Matriisissa olevia lukuja kutsutaan matriisin *alkioiksi*, ja rivillä i sarakkeessa j olevaa alkiota merkitään $A(i,j)$. Esimerkiksi

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

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

9.1 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 samoissa kohdissa olevat alkiot. Tuloksena on $m \times n$ -matriisi $A + B$, jolle pätee

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

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 matriiseja, joilla on sama tyyppi, voidaan laskea yhteen.

Skalaarikertolasku

Minkä tahansa matriisin $A \in \mathbb{R}^{m\times n}$ voi kertoa reaaliluvulla c, ja tätä toimitusta kutsutaan *skalaarikertolaskuksi*. Saatava tulos on $m \times n$ -matriisi cA, jota nimitetään matriisin A *skalaarimonikerraksi* ja jolle pätee

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

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*.

Matriisikertolasku

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

Kaksi matriisia voidaan kertoa keskenään vain, jos ensimmäisessä on yhtä paljon sarakkeita kuin toisessa on rivejä. Olkoot siis $A \in \mathbb{R}^{m\times n}$ ja $B \in \mathbb{R}^{n\times p}$. Tällöin tulo AB on $m \times p$ -matriisi, jolle pätee

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

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

Esimerkki 9.1. Lasketaan matriisien

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 8 & 9 & 0 \end{bmatrix} \quad \text{ja} \quad B = \begin{bmatrix} 4 & 5 \\ 6 & 7 \\ -1 & -2 \end{bmatrix}$$

tulo. Koska matriisissa A on kolme saraketta (tyyppi 2×3) ja matriisissa B on vastaavasti kolme riviä (tyyppi 3×2), matriisit voidaan kertoa keskenään. Tulosmatriisi on tyyppiä 2×2. Määritelmän mukaan

$$AB = \begin{bmatrix} 1 & 2 & 3 \\ 8 & 9 & 0 \end{bmatrix}\begin{bmatrix} 4 & 5 \\ 6 & 7 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 1\cdot 4 + 2\cdot 6 + 3\cdot(-1) & 1\cdot 5 + 2\cdot 7 + 3\cdot(-2) \\ 8\cdot 4 + 9\cdot 6 + 0\cdot(-1) & 8\cdot 5 + 9\cdot 7 + 0\cdot(-2) \end{bmatrix} = \begin{bmatrix} 13 & 13 \\ 86 & 103 \end{bmatrix}.$$

---

[1] Merkintä $\sum_{k=1}^{n} c_k$ tarkoittaa summaa $c_1 + c_2 + \cdots + c_n$.

Esimerkki 9.2. 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 $AB \neq BA$.

Oletetaan, että A on $n \times n$-matriisi ja $k \in \{1, 2, \dots\}$. Tällöin voidaan matriisitulon avulla määritellä *matriisipotenssi*

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

9.2 Erityisiä matriiseja

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 usein myös *yksikkömatriiseiksi*.

Huom. 1. Nollamatriisi voi olla mitä tahansa tyyppiä. Sen sijaan ykkösmatriisissa on aina yhtä paljon rivejä ja sarakkeita.

Huom. 2. Jos matriisien tyypeistä ei ole epäselvyyttä, saatetaan merkitä yksinkertaisemmin $O_{m \times n} = O$ ja $I_n = I$.

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

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

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

Neliömatriisi on matriisi, jossa on yhtä monta riviä ja saraketta. Esimerkiksi ykkösmatriisit ovat neliömatriiseja.

Neliömatriisin alkio on *lävistäjällä* eli *diagonaalilla*, jos alkion rivin ja sarakkeen numerot ovat samat. Matriisi, jonka kaikki nollasta poikkeavat alkiot ovat lävistäjällä, on *lävistäjämatriisi*. Lävistäjämatriisi, jonka kaikki lävistäjäalkiot ovat samoja, on puolestaan *skalaarimatriisi*. Skalaarimatriisit ovat ykkösmatriisin 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.

9.3 Matriisien laskusääntöjä

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

Lause 9.3. *Seuraavat säännöt pätevät matriiseille A, B ja C sekä reaaliluvuille a ja b, jos laskutoimitukset on määritelty:*

a) $A + B = B + A$

b) $A + (B + C) = (A + B) + C$

c) $A(BC) = (AB)C$

d) $A(B + C) = AB + AC$

e) $(A + B)C = AC + BC$

f) $(ab)A = a(bA)$

g) $a(AB) = (aA)B = A(aB)$.

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

Todistus. Osoitetaan esimerkin vuoksi kohta d). Muiden kohtien tarkistaminen jätetään lukijalle.

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 $i \in \{1, 2, \ldots , m\}$ ja $j \in \{1, 2, \ldots , p\}$. Nähdään, että

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

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

9.4 Matriisin transpoosi

Määritelmä 9.4. Oletetaan, että A on $m \times n$ -matriisi. Sen *transpoosi* $A^\top$ 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

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

Määritelmä 9.5. Neliömatriisin A sanotaan olevan *symmetrinen*, jos $A^\top = A$. Neliömatriisin A sanotaan olevan *antisymmetrinen*, jos $A^\top = -A$.

Esimerkki 9.6. Merkitään

$$B = \begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 0 \end{bmatrix} \quad \text{ja} \quad C = \begin{bmatrix} 0 & 4 & -5 \\ -4 & 0 & -6 \\ 5 & 6 & 0 \end{bmatrix}.$$

Tällöin

$$B^\top = \begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 0 \end{bmatrix} = B \quad \text{ja} \quad C^\top = \begin{bmatrix} 0 & -4 & 5 \\ 4 & 0 & 6 \\ -5 & -6 & 0 \end{bmatrix} = -C.$$

Siis B on symmetrinen ja C on antisymmetrinen.

Transpoosioperaation käyttäytymistä matriisien laskutoimitusten kanssa valottaa seuraava lause.

Lause 9.7. *Seuraavat säännöt pätevät matriiseille A ja B sekä reaaliluvulle t, jos laskutoimitukset on määritelty (ts. matriisit ovat sopivaa tyyppiä):*

a) $(A^\top)^\top = A$

b) $(A + B)^\top = A^\top + B^\top$

c) $(AB)^\top = B^\top A^\top$

d) $(tA)^\top = t(A^\top)$.

Erityisesti kannattaa huomata tulon tekijöiden järjestyksen vaihtuminen kohdassa (c).

Todistus. Osoitetaan todeksi kohta (c) ja jätetään loput kohdat lukijalle. Olkoot $A \in \mathbb{R}^{m\times n}$ ja $B \in \mathbb{R}^{n\times p}$. Nyt sekä $(AB)^\top$ että $B^\top A^\top$ 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

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

Siten $(AB)^\top = B^\top A^\top$. □

9.5 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. 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 9.8. Olkoon A neliömatriisi. Jos on olemassa saman tyyppinen neliömatriisi B, 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.

Joissakin matematiikan kirjoissa kääntyviä matriiseja kutsutaan säännöllisiksi matriiseiksi. Sellaisia matriiseja, joilla ei ole käänteismatriisia, kutsutaan toisinaan singulaarisiksi.

Esimerkki 9.9. 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 9.10. Läheskään kaikilla matriiseilla ei ole käänteismatrisia. 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 matrisiitulo, 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 9.11. *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. □

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 9.12. *Oletetaan, että matriisit A ja B ovat kääntyviä. Tällöin myös matriisit A^{-1}, AB ja $A^\top$ ovat kääntyviä, ja niiden käänteismatriisit ovat seuraavat:*

a) $(A^{-1})^{-1} = A$

b) $(A^\top)^{-1} = (A^{-1})^\top$

c) $(AB)^{-1} = B^{-1}A^{-1}$.

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

a) 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 $(A^{-1})^{-1} = A$.

b) Osoitetaan, että matriisin $A^\top$ käänteismatriisi on $(A^{-1})^\top$. Lauseen 9.7 nojalla pätee

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

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

c) Matriisia AB koskevan väitteen todistaminen jätetään harjoitustehtäväksi. □

2×2 -matriisin käänteismatriisi

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

Lause 9.13. *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 tuloon AB tulee nimittäin välttämättä nollarivi ja jälkimmäisessä tapauksessa tuloon BA tulee välttämättä nollasarake. 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

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

eli

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

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. □

Käänteismatriisin määrittäminen yleisessä tapauksessa

Suurempien kuin 2×2 -matriisien käänteismatriisien laskemiseksi ei helppoa kaavaa. On kuitenkin olemassa menetelmä, jolla voidaan aina selvittää, onko matriisi kääntyvä. Jos matriisin on kääntyvä, voidaan menetelmän avulla lisäksi selvittää sen käänteismatriisi. Tässä luvussa kerrotaan, kuinka matrisiin kääntyvyyden ja mahdollisen käänteismatriisin voi selvittää. Luvussa 10 puolestaan paneudutaan siihen, miksi menetelmä toimii.

Matriisin kääntyvyyden selvittäminen ja käänteismatriisin etsiminen voidaan tehdä yhtä aikaa, ja se tapahtuu seuraavasti. Oletetaan, että halutaan selvittää, onko matriisi A kääntyvä. Yhdistetään matriisi A ja ykkösmatriisi I matriisiksi $[A \mid I]$. Tehdään tälle matriisille alkeisrivitoimituksia, joilla yritetään muuttaa A redusoiduksi porrasmatriisiksi. Jos A saadaan muutettua alkeisrivitoimitusten avulla ykkösmatriisiksi, on A kääntyvä. Lisäksi matriisin I paikalle muodostuu matriisin A käänteismatriisi.

Esimerkki 9.14. Tutkitaan, onko matriisilla

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

käänteismatriisi. Muokataan yhdistettyä matriisia

$$\left[\begin{array}{ccc|ccc} 2 & 4 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array}\right]$$

alkeisrivitoimituksilla samaan tapaan kuin Gaussin–Jordanin menetelmässä. Tavoitteena on saada vasemmalle puolelle ykkösmatriisi. Muokkaus voi tapahtua esimerkiksi seuraavasti:

$$\left[\begin{array}{ccc|ccc} 2 & 4 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 2 & 4 & -2 & 1 & 0 & 0 \end{array}\right] \xrightarrow{R_3 - 2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 4 & -10 & 1 & 0 & -2 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 4 & -10 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right] \xrightarrow{\frac{1}{4}R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & -\frac{5}{2} & \frac{1}{4} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right] \xrightarrow{R_2 + \frac{5}{2}R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 0 & 0 & 1 \\ 0 & 1 & 0 & \frac{1}{4} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right] \xrightarrow{R_1 - 4R_3}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -4 & 1 \\ 0 & 1 & 0 & \frac{1}{4} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right].$$

Koska matriisi A saatiin muutettua alkeisrivitoimituksilla ykkösmatriisiksi, on A kääntyvä. 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 kääntyvä, tulee myös se ilmi matriisia redusoitaessa. Jos matriisia voidaan muokata alkeisrivitoimituksilla niin, että saadaan aikaiseksi nollarivi, ei matriisi ole kääntyvä. Toisin sanoen matriisi ei voi olla kääntyvä, jos se on riviekvivalentti sellaisen matriisin kanssa, jossa on nollarivi.

Esimerkki 9.15. Tutkitaan, onko matriisilla

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

käänteismatriisi. Ryhdytään muokkaamaan yhdistettyä matriisia alkeisrivitoimituksilla:

$$\left[\begin{array}{ccc|ccc} -1 & 0 & 2 & 1 & 0 & 0 \\ 4 & 1 & -3 & 0 & 1 & 0 \\ 3 & 1 & -1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{(-1)R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 4 & 1 & -3 & 0 & 1 & 0 \\ 3 & 1 & -1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_2-4R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 5 & 4 & 1 & 0 \\ 3 & 1 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-3R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 5 & 4 & 1 & 0 \\ 0 & 1 & 5 & 3 & 0 & 1 \end{array}\right] \xrightarrow{R_3-R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 5 & 4 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right].$$

Koska matriisin B viimeisen rivin paikalle tuli nollarivi, matriisi B ei ole kääntyvä.

9.6 Sarakevektorit

Usein avaruuden $\mathbb{R}^n$ vektori $(v_1, v_2, \ldots, v_n)$ samastetaan matriisin

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$$

kanssa. Joukon $\mathbb{R}^{n\times 1}$ matriiseja kutsutaan *sarakevektoreiksi*. Kun avaruuden $\mathbb{R}^n$ vektoreita ajatellaan sarakevektoreina, voi niitä kertoa matriiseilla: jos $A \in \mathbb{R}^{m\times n}$ ja $\bar{v} \in \mathbb{R}^n$, on tulo $A\bar{v}$ määritelty, kun $\bar{v}$ tulkitaan joukon $\mathbb{R}^{n\times 1}$ alkioksi.

Esimerkki 9.16. Matriisin

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

ja vektorin $\bar{v} = (-5, 3)$ tulo on

$$A\bar{v} = \begin{bmatrix} 2 & 4 \\ -1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} -5 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 13 \end{bmatrix}.$$

Saatu tulos voidaan puolestaan samastaa vektorin $(2, 5, 13)$ kanssa.

Tilan säästämiseksi sarakevektorit kirjoitetaan usein transpoosimerkinnän avulla. Esimerkiksi sarakevektori

$$\begin{bmatrix} 2 \\ 5 \\ 13 \end{bmatrix}$$

saa tällöin muodon $[2\ 5\ 13]^\top$.

10 Matriisit ja yhtälöryhmät

Tässä luvussa esitellään uusi tapa kirjoittaa lineaarinen yhtälöryhmä matriisien avulla käyttäen hyväksi matriisikertolaskua sekä sarakevektoreita. Pilkotaan sitä varten yhtälöryhmän

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

eri osat omiksi matriiseikseen. Ensinnäkin yhtälöryhmän *kerroinmatriisiksi* kutsutaan matriisia

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

Kerroinmatriisi sisältää siis yhtälöryhmän kertoimet järjestyksessä. Kerätään vielä muuttujat ja vakiot omiksi matriiseikseen:

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{ja} \qquad \bar{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Kaikki tuntemattomat ovat sarakevektorissa $\bar{x}$ ja kaikki vakiot sarakevektorissa $\bar{b}$.

Nyt yhtälöryhmä voidaan kirjoittaa matriisien avulla. Huomataan nimittäin, että

$$A\bar{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

Sarakevektorin $A\bar{x}$ alkiot vastaavat yhtälöryhmän yhtälöiden vasempia puolia. Koska sarakevektori $\bar{b}$ sisältää yhtälöiden oikeat puolet samassa järjestyksessä, yhtälöryhmä voidaan kirjoittaa muodossa $A\bar{x} = \bar{b}$ eli

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

Kerroinmatriisin kääntyvyys vaikuttaa nyt merkittävästi yhtälöryhmän ratkaisuihin.

Lause 10.1. *Jos matriisi $A \in \mathbb{R}^{n \times n}$ on kääntyvä ja $\bar{b} \in \mathbb{R}^n$, yhtälöllä $A\bar{x} = \bar{b}$ on täsmälleen yksi ratkaisu, ja se on $\bar{x} = A^{-1}\bar{b}$.*

Todistus. Oletetaan, että matriisi A on kääntyvä. Todistuksessa on kaksi osaa. On osoitettava, että $A^{-1}\bar{b}$ on yhtälön ratkaisu ja että muita ratkaisuja ei ole.

Osoitetaan ensin, että $A^{-1}\bar{b}$ on yhtälön ratkaisu sijoittamalla se yhtälön vasemmalle puolelle vektorin $\bar{x}$ paikalle:

$$A(A^{-1}\bar{b}) = (AA^{-1})\bar{b} = I\bar{b} = \bar{b}.$$

Koska tuloksena oli yhtälön oikea puoli, esitetty ratkaisu toteuttaa yhtälön.

Osoitetaan sitten, ettei muita ratkaisuja ole. Oletetaan, että $\bar{y}$ on jokin (toinen) ratkaisu. Tällöin $A\bar{y} = \bar{b}$. Kerrotaan yhtälön molemmat puolet vasemmalta matriisilla A^{-1}, jolloin saadaan

$$A^{-1}A\bar{y} = A^{-1}\bar{b}.$$

Koska $A^{-1}A = I$, edellinen yhtälö sievenee muotoon $\bar{y} = A^{-1}\bar{b}$. Kysymyksessä on siis sama ratkaisu kuin edellä. Siten ratkaisuja on vain yksi ja se on $A^{-1}\bar{b}$. □

Huomaa, että todistuksessa tarvittiin käänteismatriisin kumpaakin ominaisuutta: $AA^{-1} = I$ ja $A^{-1}A = I$.

10.1 Alkeismatriisit

Myös alkeisrivitoimitukset voi ilmaista matriisikertolaskun avulla. Osoittautuu, että jos matriisia kerrotaan niin kutsutulla alkeismatriisilla, tullaan matriisille tehneeksi jokin alkeisrivitoimitus. Tästä havainnosta tulee olemaan hyötyä kääntyvien matriisien käsittelyssä.

Määritelmä 10.2. Matriisi on *alkeismatriisi*, jos se on saatu ykkösmatriisista yhdellä alkeisrivitoimituksella.

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 alkeisrivitoimitukset $-\frac{1}{2}R_3$, $R_2 \leftrightarrow R_4$ ja $R_3 + 3R_1$.

Esimerkki 10.3. Osoittautuu, että alkeismatriiseilla kertominen vastaa alkeisrivitoimitusten 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

$$E_1A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\mathbf{\frac{1}{2}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -\mathbf{\frac{1}{2}}a_{31} & -\mathbf{\frac{1}{2}}a_{32} & -\mathbf{\frac{1}{2}}a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix},$$

$$E_2A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{1} & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} & \boldsymbol{a_{23}} \\ a_{31} & a_{32} & a_{33} \\ \boldsymbol{a_{41}} & \boldsymbol{a_{42}} & \boldsymbol{a_{43}} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \boldsymbol{a_{41}} & \boldsymbol{a_{42}} & \boldsymbol{a_{43}} \\ a_{31} & a_{32} & a_{33} \\ \boldsymbol{a_{21}} & \boldsymbol{a_{22}} & \boldsymbol{a_{23}} \end{bmatrix} \quad \text{ja}$$

$$E_3A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \mathbf{3} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \boldsymbol{3a_{11}} + a_{31} & \boldsymbol{3a_{12}} + a_{32} & \boldsymbol{3a_{13}} + a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}.$$

Huomataan, että jokaisella alkeismatriisilla kerrottaessa matriisille A tullaan tehneeksi sama alkeisrivitoimitus, jonka avulla alkeismatriisi muodostettiin.

Yksittäinen esimerkki ei takaa, että alkeismatriisilla kertominen vastaa aina alkeisrivitoimituksen 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 10.4. *Oletetaan, että $A \in \mathbb{R}^{n \times m}$. Olkoon E alkeismatriisi, joka saadaan tekemällä jokin alkeisrivitoimitus ykkösmatriisille I_n. Jos matriisille A tehdään sama alkeisrivitoimitus, tuloksena on matriisi EA.*

Huom. Lemma tarkoittaa apulausetta. Se on siis pieni tulos, jota voidaan käyttää hyväksi tärkeämpien lauseiden todistamisessa.

Lause 10.5. *Alkeismatriisit ovat kääntyviä, ja alkeismatriisin käänteismatriisi on myös alkeismatriisi.*

Todistus. Myöskään tämän tuloksen tarkkaa todistusta ei esitetä tässä. Käydään kuitenkin läpi todistuksen idea.

Jokainen alkeisrivitoimitus voidaan peruuttaa toisella alkeisrivitoimituksella kuten kohta nähdään. Kutsutaan tätä alkeisrivitoimitusta alkuperäisen alkeisrivitoimituksen *käänteistoimitukseksi.*

Oletetaan, että $a, b \in \mathbb{R}$ ja $a \neq 0$. Jos matriisille tehdään alkeisrivitoimitus $R_i \leftrightarrow R_j$, päästään takaisin alkutilanteeseen tekemällä sama alkeisrivitoimitus uudelleen. Alkeisrivitoimitus $R_i \leftrightarrow R_j$ on siis itsensä käänteistoimitus. Alkeisrivitoimituksen aR_i käänteistoimitus on puolestaan $\frac{1}{a}R_i$, ja alkeisrivitoimituksen $R_i + bR_j$ käänteistoimitus on $R_i - bR_j$.

Alkeismatriisin käänteismatriisi saadaan aina käänteistoimitusta vastaavasta alkeismatriisista. Alkeisrivitoimitusta $R_i \leftrightarrow R_j$ vastaava alkeismatriisi on oma käänteismatriisinsa, alkeisrivitoimitusta aR_i vastaavan alkeismatriisin käänteismatriisi on alkeisrivitoimitusta $\frac{1}{a}R_i$

vastaava alkeismatriisi ja niin edelleen. Alkeisrivitoimituksen tekeminen vastaa nimittäin edellisen lemman nojalla alkeismatriisilla kertomista. Esimerkiksi alkeisrivitoimitukset 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 10.6. 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 alkeisrivitoimitusta $R_3 + 3R_1$. Tämän alkeisrivitoimituksen voi kumota tekemällä alkeisrivitoimituksen $R_3 - 3R_1$. 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 10.1 todettiin jo kääntyvien matriisien merkitys yhtälöryhmän ratkaisun kannalta. Nyt tuota tulosta voidaan täydentää tarkastelemalla lisäksi alkeisrivitoimituksia ja niitä vastaavia alkeismatriiseja.

Lause 10.7 (Kääntyvien matriisien lause)**.** *Oletetaan, että A on $n \times n$ -neliömatriisi. Seuraavat ehdot ovat yhtäpitäviä:*

a) Matriisi A on kääntyvä.

b) Yhtälöllä $A\bar{x} = \bar{b}$ on täsmälleen yksi ratkaisu kaikilla $\bar{b} \in \mathbb{R}^n$.

c) Yhtälöllä $A\bar{x} = \bar{0}$ on vain triviaali ratkaisu $\bar{x} = \bar{0}$.

d) Matriisi A on riviekvivalentti ykkösmatriisin kanssa.

e) Matriisi A on alkeismatriisien tulo.

f) Matriisi A ei ole riviekvivalentti minkään nollarivin sisältävän matriisin kanssa.

Todistus. Osoitetaan väite todistamalla seuraavat päättelyketjut:

$$\text{a)} \Rightarrow \text{b)} \Rightarrow \text{c)} \Rightarrow \text{d)} \Rightarrow \text{e)} \Rightarrow \text{a)} \qquad \text{ja} \qquad \text{d)} \Rightarrow \text{f)} \Rightarrow \text{d)}.$$

Tämän jälkeen tiedetään, että jokainen lauseen kohta on yhtäpitävä toisten kohtien kanssa.

a) $\Rightarrow$ b): Väite on osoitettu lauseessa 10.1.

b) $\Rightarrow$ c): Oletetaan, että yhtälöllä $A\bar{x} = \bar{b}$ on täsmälleen yksi ratkaisu kaikilla $\bar{b} \in \mathbb{R}^n$. Tämä pätee myös, jos $\bar{b} = \bar{0}$. Toisaalta yhtälöllä $A\bar{x} = \bar{0}$ on aina ratkaisu $\bar{x} = \bar{0}$. Siten $\bar{x} = \bar{0}$ on ainoa ratkaisu.

c) $\Rightarrow$ d): Oletetaan, että yhtälöllä $A\bar{x} = \bar{0}$ on vain ratkaisu $\bar{x} = \bar{0}$. Merkitään $A(i,j) = a_{ij}$ kaikilla $i, j \in \{1, 2, \ldots, n\}$. Yhtälöä $A\bar{x} = \bar{0}$ vastaava lineaarinen yhtälöryhmä on

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & 0 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & 0 \end{cases}$$

Yhtälöryhmässä on sama määrä yhtälöitä ja tuntemattomia. Koska yhtälöryhmällä on täsmälleen yksi ratkaisu $\bar{x} = \bar{0}$, se on ekvivalentti yhtälöryhmän

$$\begin{cases} x_1 = 0 \\ x_2 = 0 \\ \quad\vdots \\ x_n = 0 \end{cases}$$

kanssa. Tämä tarkoittaa, että matriisi A saadaan alkeisrivitoimituksilla muutettua ykkösmatriisiksi. Toisin sanottuna A on riviekvivalentti ykkösmatriisin kanssa.

d) $\Rightarrow$ e): Oletetaan, että matriisi A on riviekvivalentti ykkösmatriisin kanssa. Toisin sanoen A voidaan muuttaa alkeisrivitoimituksilla ykkösmatriisiksi I. Tällöin lemman 1.4 perusteella on olemassa alkeismatriisit $E_1, \ldots, E_k$, joilla kertomalla matriisista A saadaan ykkösmatriisi. Pätee 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.

e) $\Rightarrow$ a): Oletetaan, että $A = E_1 \cdots E_k$, missä $E_1, \ldots, E_k$ ovat alkeismatriiseja. Merkitään $B = E_k^{-1} \cdots E_1^{-1}$. Nyt

$$\begin{aligned} AB &= (E_1 \cdots E_k)(E_k^{-1} \cdots E_1^{-1}) \\ &= E_1 \cdots (E_k E_k^{-1}) \cdots E_1^{-1} \\ &= E_1 \cdots E_{k-1} I E_{k-1}^{-1} \cdots E_1^{-1} \\ &\;\;\vdots \\ &= E_1 E_1^{-1} = I. \end{aligned}$$

Siispä $AB = I$. Samalla tavalla nähdään, että $BA = I$. Siten B on matriisin A käänteismatriisi ja A on kääntyvä.

d) $\Rightarrow$ f): Oletetaan, että A on riviekvivalentti ykkösmatriisin kanssa. Tehdään lisäksi vastaoletus, että A on riviekvivalentti jonkin matriisin B kanssa, joka sisältää nollarivin. Tällöin matriisiyhtälöitä $A\bar{x} = \bar{0}$, $I\bar{x} = \bar{0}$ ja $B\bar{x} = \bar{0}$ 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 $B\bar{x} = \bar{0}$ on vapaa muuttuja ja ratkaisuja on ääretön määrä. Kuitenkin yhtälöllä $I\bar{x} = \bar{0}$ on vain yksi ratkaisu: $\bar{x} = \bar{0}$. Tämä on ristiriita, joten vastaoletus on väärä ja A ei ole riviekvivalentti nollarivin sisältävän matriisin kanssa.

f) $\Rightarrow$ d): 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. □

10.2 Käänteismatriisin määrittäminen - Miksi menetelmä toimii?

Luvussa 9.5 esitettiin menetelmä, jonka avulla voidaan tutkia, onko matriisi kääntyvä. Nyt voidaan vihdoin perustella, miksi menetelmä toimii.

Menetelmässä matriisia muokataan alkeisrivitoimituksilla. Jos näin saadaan aikaiseksi ykkösmatriisi, on alkuperäinen matriisi kääntyvä. Tämä perustuu siihen, että jos matriisi A onnistutaan muuttamaan alkeisrivitoimituksilla ykkösmatriisiksi, niin A on lauseen 10.7 nojalla kääntyvä eli sillä on käänteismatriisi A^{-1}.

Jos matriisi on kääntyvä, käytetyistä alkeisrivitoimituksista saadaan myös selville, mikä käänteismatriisi on. Oletetaan, että matriisi A on muutettu ykkösmatriisiksi alkeisrivitoimituksilla, joita vastaavat alkeismatriisit $E_1, \ldots, E_k$. 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 alkeisrivitoimitukset kuin tehtiin alunperin matriisille A päädytään käänteismatriisiin A^{-1}. Siis

$$[A \mid I] \quad \rightsquigarrow \quad [I \mid A^{-1}].$$

Siksi käänteismatriisi A^{-1} ilmestyy ykkösmatriisin I paikalle.

Lause 10.7 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 alkeisrivitoimituksilla muodostettua nollarivi. Näin olemme perustelleet loputkin luvussa 9.5 esitetystä menetelmästä. Matriisin kääntyvyyden selvittämisessä käytettävä menetelmä perustuu siten täysin lauseeseen 10.7.

10.3 Lisää kääntyvyyteen liittyviä tuloksia

Tässä alaluvussa esitetään kaksi tulosta, joista on toisinaan hyötyä. Molempien todistukset perustuvat matriisien ja yhtälöryhmien väliseen suhteeseen. Ensimmäisessä todetaan, että matriisin kääntyvyyden toteamiseksi ei tarvitse tarkastella matriisituloa molemmin päin.

Lause 10.8. *Olkoot $A, B \in \mathbb{R}^{n\times n}$. Jos $BA = I$, niin $AB = I$ eli matriisit A ja B ovat toistensa käänteismatriiseja.*

Todistus. Oletetaan, että $BA = I$. Osoitetaan ensin, että matriisi A on kääntyvä. Tehdään tämä osoittamalla, että homogeenisella yhtälöryhmällä $A\bar{x} = \bar{0}$ on vain triviaaliratkaisu. Kertomalla yhtälö puolittain matriisilla B saadaan $BA\bar{x} = B\bar{0}$. Oletuksen mukaan $BA = I$ ja toisaalta $B\bar{0} = \bar{0}$, joten yhtälö tulee muotoon $\bar{x} = \bar{0}$. Voidaan siis todeta, että yhtälöryhmän $A\bar{x} = \bar{0}$ ainoa ratkaisu on $\bar{x} = \bar{0}$. Lauseen 10.7 nojalla tästä seuraa, että A on kääntyvä.

Koska A on kääntyvä, käänteismatriisi A^{-1} on olemassa. Kerrotaan sillä yhtälön $BA = I$ molemmat puolet oikealta. Näin saadaan yhtälö $(BA)A^{-1} = IA^{-1}$, joka sievenee muotoon $B = A^{-1}$. Siten B on matriisin A käänteismatriisi. Muut väitteet seuraavat tästä. Nyt tiedetään, että $AB = I$. Lisäksi matriisin B käänteismatriisi on lauseen 9.12 nojalla A. □

Toinen tulos liittyy siihen, millä ehdoilla kahden matriisin tulo on tai ei ole kääntyvä.

Lause 10.9. *Olkoot A ja B neliömatriiseja. Tällöin seuraavat väitteet pätevät:*

a) Jos A ja B ovat molemmat kääntyviä, myös tulo AB on kääntyvä.

b) Jos vain toinen matriiseista A ja B on kääntyvä, tulo AB ei ole kääntyvä.

c) Jos kumpikaan matriiseista A ja B ei ole kääntyvä, tulo AB ei ole kääntyvä.

Todistus. a) Tämä on todistettu lauseessa 9.12.

b) Oletetaan esimerkiksi, että A on kääntyvä ja B ei. Tehdään vastaoletus, että tulo AB on kääntyvä ja sillä on käänteismatriisi C. Tällöin voidaan laskea

$$(CA)B = C(AB) = I.$$

Lauseen 10.8 nojalla B on kääntyvä. Tämä on ristiriita, joten vastaoletus on väärä, eli tulo AB ei ole kääntyvä.

Tapaus, jossa A ei ole kääntyvä ja B on kääntyvä, todistetaan samaan tapaan.

c) Tämän kohdan todistuksessa tarvitaan matriisien ja yhtälöryhmien välistä yhteyttä. Oletetaan, että A ja B eivät ole kääntyviä. Tarkastellaan homogeenista yhtälöryhmää $(AB)\bar{x} = \bar{0}$ ja yritetään osoittaa, että sillä on muitakin kuin triviaaliratkaisu.

Koska B ei ole kääntyvä, yhtälöryhmällä $B\bar{x} = \bar{0}$ on jokin epätriviaali ratkaisu $\bar{x} = \bar{v} \neq \bar{0}$. Tällöin

$$(AB)\bar{v} = A(B\bar{v}) = A\bar{0} = \bar{0}.$$

Siispä $\bar{x} = \bar{v}$ on myös yhtälöryhmän $(AB)\bar{x} = \bar{0}$ ratkaisu. Koska kyseisellä yhtälöryhmällä on epätriviaali ratkaisu, lauseesta 10.7 seuraa, että matriisi AB ei ole kääntyvä. □

11 Determinantti

Neliömatriisille voidaan laskea luku, joka kertoo muun muassa, onko matriisi kääntyvä vai ei. Tätä lukua kutsutaan matriisin determinantiksi. Determinantilla on muitakin sovelluksia, mutta tässä yhteydessä tarkastelemme vain sen yhteyttä matriisin kääntyvyyteen.

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.

11.1 Pienten matriisien determinantit

Tarkastellaan ensin korkeintaan 3×3 -matriisien determinantteja. Determinantti voidaan laskea vain neliömatriisille.

Määritelmä 11.1.

a) Matriisin
$$A = \begin{bmatrix} a \end{bmatrix}$$
determinantti on $\det(A) = a$.

b) Matriisin
$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
determinantti on $\det(B) = ad - bc$.

c) Matriisin
$$C = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$
determinantti on
$$\begin{aligned} \det(C) &= a_1(b_2c_3 - c_2b_3) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1) \\ &= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 - a_2b_1c_3 - a_1b_3c_2. \end{aligned}$$

Matriisin determinanttia voi merkitä myös pystyviivojen avulla:
$$\det(A) = \begin{vmatrix} a \end{vmatrix}, \quad \det(B) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{ja} \quad \det(C) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Esimerkki 11.2. Matriisin $A = [4]$ determinantti on $\det(A) = 4$. Matriisin
$$B = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

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

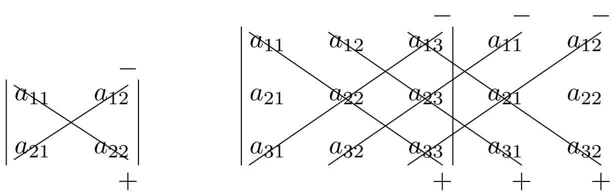
$$C = \begin{bmatrix} -2 & 3 & 2 \\ 0 & 1 & -1 \\ 1 & 2 & 4 \end{bmatrix}$$

determinantti on

$$\begin{aligned} \det(C) &= -2(1 \cdot 4 - (-1) \cdot 2) - 3(0 \cdot 4 - (-1) \cdot 1) + 2(0 \cdot 2 - 1 \cdot 1) \\ &= -2 \cdot 6 - 3 \cdot 1 + 2 \cdot (-1) = -17. \end{aligned}$$

Matriisille, jonka tyyppi on 2×2, voi käyttää determinantin laskemiseen kuvassa 11.27 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.

Kuvassa 11.27 on esitetty laskemista helpottava muistisääntö myös suuremman, tyyppiä 3×3 olevan matriisin determinantille. Kirjoitetaan matriisin vierelle matriisin ensimmäinen ja toinen sarake. Piirretään kuvion päälle matriisin lävistäjän suuntaisia viivoja sekä vastakkaissuuntaisia viivoja. Samalla viivalla olevat alkiot kerrotaan keskenään. Jos viiva on lävistäjän suuntainen, tulee tulon eteen plusmerkki. Jos viiva on vastakkaissuuntainen, tulee tulon eteen miinusmerkki. Lopuksi tulot lasketaan yhteen.



Kuva 11.27: Muistisäännöt 2×2 -determinantin ja 3×3 -determinantin laskemiseksi.

Determinantin merkitys näkyy siinä, että se kertoo matriisin kääntyvyydestä. Lauseen 9.13 nojalla 2×2 -matriisi

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

on kääntyvä, jos ja vain jos $ad - bc \neq 0$. Toisaalta matriisin A determinantti on $ad - bc$. Matriisi A on siis kääntyvä, jos ja vain jos $\det(A) \neq 0$. Samanlainen tulos pätee myös 3×3 -matriisien determinanteille sekä myöhemmin määriteltäville suurempien matriisien determinanteille. Todistus on esitetty tämän luvun loppupuolella.

Lause 11.3. *Oletetaan, että A on $n \times n$ -matriisi. Matriisi A on kääntyvä, jos ja vain jos $\det(A) \neq 0$.*

Esimerkki 11.4. Esimerkissä 11.2 saatiin matriisin

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

determinantiksi -17. Koska determinantti ei ole nolla, matriisi C on kääntyvä.

Esimerkki 11.5. Tutkitaan, onko vektorijono $((2,1,-1),(0,1,-3),(-2,1,-5))$ avaruuden $\mathbb{R}^3$ kanta. Kyseessä on kanta, mikäli jokainen vektori $\bar{w} \in \mathbb{R}^3$ voidaan ilmaista yksikäsitteisesti annettujen vektorien lineaarikombinaationa. Tutkittava yhtälöryhmä voidaan ilmaista matriisimuodossa yhtälönä $A\bar{x} = \bar{w}$, 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 kääntyvä (lause 10.7). Toisaalta lauseen 11.3 nojalla A on kääntyvä, jos ja vain jos sen determinantti on nollasta poikkeava. Lasketaan determinantti:

$$\begin{aligned} \det(A) = \begin{vmatrix} 2 & 0 & -2 \\ 1 & 1 & 1 \\ -1 & -3 & -5 \end{vmatrix} &= 2\big(1 \cdot (-5) - 1 \cdot (-3)\big) - 0 - 2\big(1 \cdot (-3) - 1 \cdot (-1)\big) \\ &= 2 \cdot (-2) - 2 \cdot (-2) = 0. \end{aligned}$$

Koska determinantti on 0, matriisi A ei ole kääntyvä. Tästä syystä tutkittavan yhtälöryhmän ratkaisua ei ole olemassa tai se ei ole yksikäsitteinen. Annettu vektorijono ei siis muodosta kantaa.

11.2 Determinantin kehityskaavat

Suurempien matriisien determinantit voidaan laskea pienempien matriisien determinanttien avulla.

Määritelmä 11.6. Olkoon A jokin $n \times n$ -matriisi. Merkitään

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

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.

Yleisen determinantin määritelmä ei ole ristiriidassa aiempien determinantin määritelmien kanssa. Esimerkiksi 3×3 -matriisin determinantti on uuden määritelmän mukaan

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$
$$= a_1(b_2c_3 - c_2b_3) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1).$$

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 11.7. *Oletetaan, että $A \in \mathbb{R}^{n\times n}$, ja merkitään $A(i,j) = a_{ij}$ kaikilla $i, j \in \{1, \ldots, n\}$.*

a) Olkoon $i \in \{1, \ldots, n\}$. 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.

b) Olkoon $j \in \{1, \ldots, n\}$. 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.

Todistuksen idea. Lause voidaan todistaa tarkastelemalla, millaiseen lausekkeeseen määritelmän 11.6 kehityskaava lopulta johtaa. Syntyvä lauseke on summa tuloista

$$\pm a_{1k_1} a_{2k_2} \cdots a_{nk_n},$$

missä $k_1, \ldots, k_n$ 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 $k_1, \ldots, k_n$ alkuperäisen järjestyksen $1, \ldots, n$ niin sanottu parillinen vai pariton permutaatio. Tämän havainnon 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 11.8. 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:

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 0 \\ 0 & 0 & -1 & 0 \\ 8 & 9 & -4 & -2 \end{vmatrix} &= 0 \cdot \begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 0 \\ 9 & -4 & -2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 5 & 7 & 0 \\ 8 & -4 & -2 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 0 \\ 8 & 9 & -2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 8 & 9 & -4 \end{vmatrix} \\ &= - \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 0 \\ 8 & 9 & -2 \end{vmatrix} \end{aligned}$$

Saatu 3×3-determinantti voidaan kehittää kolmannen sarakkeen suhteen:

$$\begin{aligned} - \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 0 \\ 8 & 9 & -2 \end{vmatrix} &= - \left(4 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 2 \\ 8 & 9 \end{vmatrix} + (-2) \cdot \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} \right) \\ &= -\big(4 \cdot (45 - 48) - 0 - 2 \cdot (6 - 10)\big) = -(-12 - 0 + 8) = -(-4) = 4. \end{aligned}$$

Siis $\det(D) = 4$.

11.3 Determinantin ominaisuuksia

Tarkastellaan vielä, miten determinantti suhtautuu matriisien alkeisrivitoimituksiin sekä laskutoimituksiin. Seuraavan tuloksen voi todistaa pienille matriiseille suoraan laskemalla, ja yleisen tapauksen voi johtaa determinanttien kehityskaavoista.

Lause 11.9. *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 11.10. *Oletetaan, että* A *on neliömatriisi. Tällöin* $\det(A^\top) = \det(A)$.

Lauseesta 11.10 seuraa, että determinantin sarakkeet käyttäytyvät täsmälleen samalla tavalla kuin sen rivit. Lauseesta 11.9 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 11.9 seuraa myös, että joidenkin matriisien determinantti on helppo määrittää.

Lause 11.11. *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 alkiot 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 11.9 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 11.9 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. $\square$

Edeltäviä lauseita voidaan käyttää hyväksi determinantin laskemisessa. Kun matriisi muutetaan porrasmatriisiksi alkeisrivitoimitusten avulla, determinantti muuttuu lauseessa 11.9 kuvatulla tavalla. Porrasmatriisin determinantti voidaan puolestaan määrittää lauseen 11.11 avulla.

Esimerkki 11.12. 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 alkeisrivitoimitukset ovat seuraavat:

$$\begin{bmatrix} 1 & -2 & 0 & 1 \\ -2 & 4 & 1 & 0 \\ 2 & 0 & -2 & 1 \\ 3 & 2 & -1 & -5 \end{bmatrix} \xrightarrow{R_2+2R_1} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 2 & 0 & -2 & 1 \\ 3 & 2 & -1 & -5 \end{bmatrix} \xrightarrow{R_3-2R_1} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 3 & 2 & -1 & -5 \end{bmatrix} \xrightarrow{R_4-3R_1}$$
$$\begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 0 & 8 & -1 & -8 \end{bmatrix} \xrightarrow{R_4-2R_3} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 0 & 0 & 3 & -6 \end{bmatrix} \xrightarrow{R_4-3R_2} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 4 & -2 & -1 \\ 0 & 0 & 0 & -12 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3}$$
$$\begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 4 & -2 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -12 \end{bmatrix}.$$

Tuloksena on yläkolmiomatriisi, jonka determinantti on lävistäjäalkioiden tulo eli tässä tapauksessa -48. Lauseen 11.9 perusteella ainoastaan viimeinen alkeisrivitoimitus 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 11.3. Tehdään se tarkastelemalla alkeisrivitoimituksia.

Lauseen 11.3 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 alkeisrivitoimituksilla ykkösmatriisiksi, muuten porrasmatriisiin tulee nollarivi (lause 10.7). Lauseen 11.11 nojalla ykkösmatriisin determinantti on 1 ja nollarivin omaavan matriisin determinantti on 0. Toisaalta lauseesta 11.9 seuraa, että jokainen alkeisrivitoimitus 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 determinanttiin.

Lause 11.13. *Oletetaan, että A ja B ovat neliömatriiseja. Tällöin* $\det(AB) = \det(A)\det(B)$.

Todistuksen idea. Jos jompikumpi tai molemmat matriiseista A ja B eivät ole kääntyviä, lauseen 10.9 perusteella myöskään niiden tulo ei ole kääntyvä. Tällöin väite pätee lauseen 11.3 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 11.9 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 11.14. *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 b)-kohdan nojalla pätee

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1.$$

Toisaalta lauseen 11.3 mukaan $\det(A) \neq 0$, sillä A on kääntyvä. Jakamalla puolittain saadaan $\det(A^{-1}) = 1/\det(A)$. □

12 Ominaisarvot ja diagonalisointi

Tässä luvussa tarkastellaan neliömatriisin ominaisarvoja ja diagonalisointia. Sovelluksena esitetään laskumenetelmä diagonalisoituvan matriisin potensseille. Ominaisarvojen tarkastelua jatketaan kurssin toisessa osassa.

12.1 Ominaisarvon määritelmä

Palautetaan mieleen, että $m \times n$ -matriisilla voi kertoa avaruuden $\mathbb{R}^n$ vektoria, kun vektori kirjoitetaan sarakevektorina (ks. luku 9.6). Esimerkiksi matriisin

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

ja vektorin $(1, 2)$ tulo on

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

eli vektori $(5, 10)$.

Huomataan, että matriisilla A kertominen vastaa vektorin $(1, 2)$ tapauksessa skalaarilla viisi kertomista. Sanotaan, että matriisilla A on ominaisarvo 5, johon liittyy ominaisvektori $(1, 2)$. 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 12.1. Oletetaan, että A on $n \times n$ -neliömatriisi. Luku $\lambda \in \mathbb{R}$ on matriisin A *ominaisarvo*, jos on olemassa sellainen vektori $\bar{v} \in \mathbb{R}^n$, että

$$\bar{v} \neq \bar{0} \quad \text{ja} \quad A\bar{v} = \lambda\bar{v}.$$

Vektoria $\bar{v}$, joka toteuttaa yllä mainitut ehdot kutsutaan ominaisarvoon λ liittyväksi *ominaisvektoriksi*.

Huom. 1. Edellinen määritelmä on sekä ominaisarvon että ominaisvektorin määritelmä. Ominaisarvoa ei voida määritellä ilman ominaisvektoreita eikä ominaisvektoreista voida puhua mainitsematta, mihin ominaisarvoon ne liittyvät.

Huom. 2. Nollavektorin ei haluta olevan ominaisvektori, sillä jos niin olisi, kaikki reaaliluvut olisivat kaikkien matriisien ominaisarvoja, koska $A\bar{0} = \lambda\bar{0}$ kaikilla $\lambda \in \mathbb{R}$.

Esimerkki 12.2. Matriisilla

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

on ominaisarvo 4, johon liitty ominaisvektori $\bar{v}_1 = (1, 1)$. Tämä nähdään laskemalla matriisin A ja vektorin $\bar{v}_1$ tulo:

$$A\bar{v}_1 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4\bar{v}_1.$$

Samaa ominaisarvoa voi vastata useampi eri ominaisvektori. Esimerkiksi $3\bar{v}_1 = (3,3)$ on myös matriisin A ominaisarvoa 4 vastaava ominaisvektori, sillä

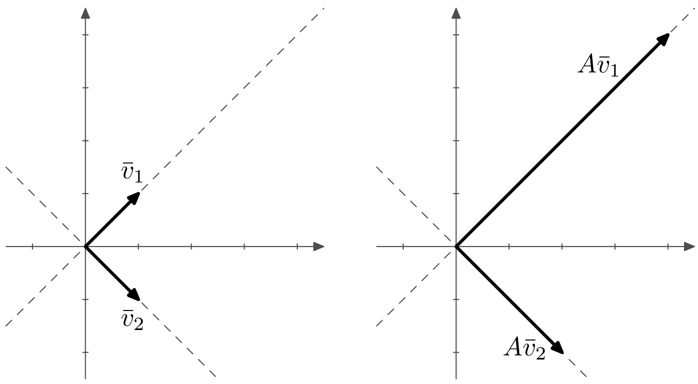
$$A(3\bar{v}_1) = 3(A\bar{v}_1) = 3(4\bar{v}_1) = 12\bar{v}_1 = 4(3\bar{v}_1).$$

Valitsemalla $\bar{v}_2 = (1,-1)$ huomataan, että matriisilla A on ominaisarvon 4 lisäksi toinenkin ominaisarvo. Nimittäin

$$A\bar{v}_2 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2\bar{v}_2.$$

Siten myös luku 2 on matriisin A ominaisarvo ja $\bar{v}_2 = (1,-1)$ on yksi siihen liittyvä ominaisvektori. (Matriisin ominaisarvot opetellaan etsimään kappaleessa 12.2).

Kun matriisilla kertoo ominaisvektoria $\bar{v}$, tuloksena on vektorin $\bar{v}$ skalaarimonikerta. Toisin sanoen tulos on vektorin $\bar{v}$ virittämällä suoralla (ks. kuva 12.28).



Kuva 12.28: Vektori $\bar{v}_1 = (1,1)$ on matriisin A ominaisvektori, sillä $A\bar{v}_1 = 4\bar{v}_1$ on vektorin $\bar{v}_1$ virittämällä suoralla. Samoin vektori $\bar{v}_2 = (1,-1)$ on matriisin A ominaisvektori, sillä $A\bar{v}_2 = 2\bar{v}_2$ on vektorin $\bar{v}_2$ virittämällä suoralla.

Tutkitaan vielä lopuksi, onko vektori $\bar{w} = (2,1)$ matriisin A ominaisvektori.

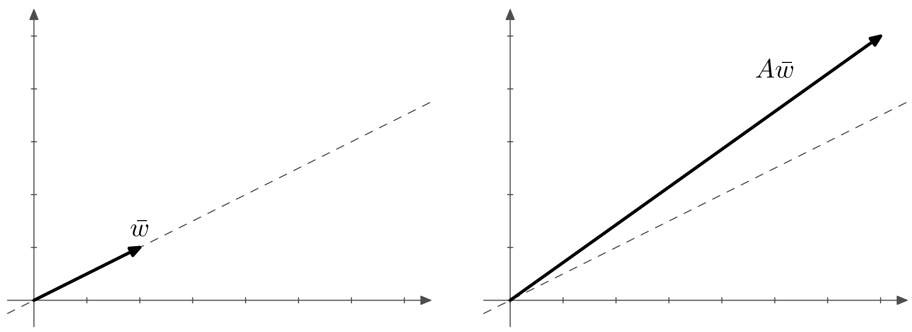
$$A\bar{w} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}.$$

Nähdään, että $A\bar{w}$ ei ole vektorin $\bar{w}$ skalaarimonikerta, joten $\bar{w}$ ei ole matriisin A ominaisvektori. Tätä on havainnollistettu kuvassa 12.29.

Kuten edellinen esimerkki osoittaa, matriisilla voi olla useampi kuin yksi ominaisarvo. Kuhunkin ominaisarvoon liittyy useita ominaisvektoreita. Seuraava esimerkki näyttää, miten tiettyyn ominaisarvoon liittyvät ominaisvektorit löydetään.

Esimerkki 12.3. Jatketaan edellistä esimerkkiä ja etsitään kaikki matriisin A ominaisarvoa 4 vastaavat ominaisvektorit. On siis ratkaistava yhtälöstä $A\bar{v} = 4\bar{v}$ tuntematon $\bar{v}$. Yhtälö saadaan muotoon

$$A\bar{v} - 4\bar{v} = \bar{0}.$$



Kuva 12.29: Vektori $\bar{w} = (2, 1)$ ei ole matriisin A ominaisvektori, sillä $A\bar{w}$ ei ole vektorin $\bar{w}$ virittämällä suoralla.

Tästä yhtälöstä haluttaisiin nyt ottaa yhteiseksi tekijäksi $\bar{v}$, 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 $\bar{v}$ samalla tavalla kuin luvulla 4 kertominen:

$$4I\bar{v} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 4v_1 + 0 \\ 0 + 4v_2 \end{bmatrix} = \begin{bmatrix} 4v_1 \\ 4v_2 \end{bmatrix} = 4\bar{v}$$

Nyt yhtälö saadaan muotoon $A\bar{v} - 4I\bar{v} = \bar{0}$, josta seuraa

$$(A - 4I)\bar{v} = \bar{0}.$$

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

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Päädytään siis ratkaisemaan yhtälöryhmä

$$\begin{cases} -v_1 + v_2 = 0 \\ v_1 - v_2 = 0 \end{cases}$$

Muutetaan yhtälöryhmän matriisi porrasmuotoon:

$$\left[\begin{array}{rr|r} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right] \xrightarrow{R_2 + R_1} \left[\begin{array}{rr|r} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \xrightarrow{(-1)\cdot R_1} \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Merkitään $v_2 = t$. Tällöin $v_1 = v_2 = t$. Siten yhtälön ratkaisu on

$$\begin{cases} v_1 = t \\ v_2 = t, \end{cases} \quad \text{missä } t \in \mathbb{R}.$$

Nollavektoria ei kuitenkaan määritelmän mukaan kelpuuteta ominaisvektoriksi. Siten ominaisarvoa 4 vastaavat ominaisvektorit ovat muotoa (t, t), missä $t \in \mathbb{R} \setminus \{0\}$.

Ominaisvektorien muodostama joukko on

$$\{(t,t) \mid t \in \mathbb{R} \setminus \{0\}\} = \{t(1,1) \mid t \in \mathbb{R} \setminus \{0\}\}.$$

Kyseessä on suora, josta on poistettu origo.

Jos matriisille A löytyy yksikin ominaisvektori, sillä on välttämättä äärettömän monta ominaisvektoria. Jokainen ominaisvektorin $\bar{v}$ skalaarimonikerta nollavektoria lukuunottamatta on nimittäin myös ominaisvektori, sillä $A(c\bar{v}) = c(A\bar{v}) = c(\lambda\bar{v}) = \lambda(c\bar{v})$ kaikilla $c \in \mathbb{R}$.

Ominaisarvoa vastaavien ominaisvektorien ei kuitenkaan tarvitse kaikkien olla toistensa skalaarimonikertoja kuten seuraava esimerkki osoittaa.

Esimerkki 12.4. 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 vastaavat ominaisvektorit näyttävät. Laskemalla samaan tapaan kuin esimerkissä 12.3 nähdään, että ominaisarvoa 1 vastaavat vektorit ovat täsmälleen muotoa

$$(s - 4t,\ 3t,\ 3s), \quad \text{missä } s, t \in \mathbb{R}.$$

Ominaisvektorien muodostama joukko on

$$\begin{aligned} \{(s-4t,\ 3t,\ 3s) \mid s,t \in \mathbb{R} \setminus \{0\}\} &= \{(s,0,3s) + (-4t,3t,0) \mid s,t \in \mathbb{R} \setminus \{0\}\} \\ &= \{s(1,0,3) + t(-4,3,0), \mid s,t \in \mathbb{R} \setminus \{0\}\}. \end{aligned}$$

Ominaisvektorit muodostavat siis tason, josta on poistettu origo.

12.2 Ominaisarvojen löytäminen

Ominaisarvojen ja ominaisvektoreiden löytäminen perustuu yhtälön

$$A\bar{v} - \lambda\bar{v} = \bar{0} \tag{8}$$

ratkaisemiseen. Ominaisvektoria $\bar{v}$ ei kuitenkaan voida ratkaista ennen kuin tunnetaan ominaisarvo λ. Sen löytämiseksi muutetaan yhtälö hieman toiseen muotoon samaan tapaan kuin esimerkissä 12.3. Ensinnäkin huomataan, että $\lambda\bar{v} = \lambda I\bar{v}$, missä I on ykkösmatriisi. Näin ollen

$$A\bar{v} - \lambda\bar{v} = A\bar{v} - \lambda I\bar{v} = (A - \lambda I)\bar{v}.$$

Nyt yhtälö (8) tulee muotoon

$$(A - \lambda I)\bar{v} = \bar{0}. \tag{9}$$

Yhtälöä (9) vastaa homogeeninen yhtälöryhmä, joten sillä on aina triviaaliratkaisu $\bar{v} = \bar{0}$. Tämä ei kuitenkaan kelpaa ominaisvektoriksi, joten tavoitteena on löytää jokin epätriviaali ratkaisu. Lauseen 10.7 nojalla yhtälöllä on epätriviaaleja ratkaisuja täsmälleen silloin, kun kerroinmatriisi $A - \lambda I$ ei ole kääntyvä. Toisaalta lauseen 11.3 mukaan neliömatriisi ei ole kääntyvä täsmälleen silloin, kun sen determinantti on 0. Näin saadaan seuraava lause.

Lause 12.5. *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 *karakteristiseksi polynomiksi*. Edellinen lause voidaan siis muotoilla myös niin, että matriisin A ominaisarvot ovat sen karakteristisen polynomin nollakohdat.

Esimerkki 12.6. Määritetään matriisin

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

ominaisarvot ja niitä vastaavat ominaisvektorit. Lähdetään liikkeelle laskemalla lauseessa 12.5 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 12.5 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 ominaisvektorit. Kumpaakin ominaisarvoa vastaavat omat ominaisvektorinsa. Tarkastellaan ensin ominaisarvoa $\lambda_1 = 4$. Tällöin ratkaistavana on yhtälö $(A - 4I)\bar{v} = \bar{0}$. Ratkaistavaksi saadaan siis yhtälöryhmä, jota vastaa matriisi

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

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 vastaavat ominaisvektorit ovat siis muotoa $((2/3)t, t)$, missä $t \in \mathbb{R} \setminus \{0\}$.

Tarkastellaan sitten ominaisarvoa $\lambda_2 = -1$. Nyt ratkaistavana oleva yhtälö on $(A+I)\bar{v} = \bar{0}$. Sen ratkaisuksi saadaan

$$\begin{cases} x_1 = -t \\ x_2 = t, \end{cases} \quad \text{missä } t \in \mathbb{R}.$$

Ominaisarvoa -1 vastaavat ominaisvektorit ovat siis muotoa $(-t, t)$, missä $t \in \mathbb{R} \setminus \{0\}$.

Matriisilla voi olla vain äärellisen monta ominaisarvoa.

Lause 12.7. *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 $c_0 + c_1\lambda + \cdots + c_n\lambda^n$, missä $c_0, \ldots, c_n \in \mathbb{R}$. Voidaan osoittaa, että yhtälöllä

$$c_0 + c_1\lambda + \cdots + c_n\lambda^n = 0$$

on enintään n eri ratkaisua. (Tämä tehdään esimerkiksi kurssilla Algebra I.) Näin ollen matriisilla A on enintään n eri ominaisarvoa. □

Joidenkin matriisien ominaisarvojen löytäminen onnistuu helposti. Jos matriisi A on kolmiomatriisi eli kaikki sen lävistäjän alapuoliset tai yläpuoliset alkiot ovat nollia, niin myös $A - \lambda I$ on kolmiomatriisi. Tällöin sen determinantti $\det(A - \lambda I)$ on lävistäjäalkioiden tulo lauseen 11.11 nojalla. Näin ollen kolmiomatriisin A karakteristinen polynomin nollakohdat saadaan yhtälöstä

$$(a_{11} - \lambda)(a_{22} - \lambda) \ldots (a_{nn} - \lambda) = 0.$$

Lauseen 12.5 nojalla saadaan tästä seuraava tulos:

Lause 12.8. *Oletetaan, että neliömatriisi A on kolmiomatriisi eli kaikki sen lävistäjän alapuoliset tai yläpuoliset alkiot ovat nollia. Tällöin matriisin A ominaisarvot ovat sen lävistäjän alkiot.*

Esimerkki 12.9. 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 alkiot. Siis matriisin A ominaisarvot ovat $\lambda_1 = 1$, $\lambda_2 = 4$ ja $\lambda_3 = -12$.

12.3 Diagonalisointi

Edellä nähtiin, että kolmiomatriisien ja siten myös lävistäjämatriisien ominaisarvot voidaan lukea suoraan matriisista. Tässä kappaleessa tutustutaan menetelmään, jonka avulla tietynlaiset neliömatriisit saadaan muutettua lävistäjämatriiseksi, joilla on samat ominaisarvot kuin alkuperäisillä matriiseilla. Matriiseja, joilla tämä menetelmä toimii, kutsutaan diagonalisoituviksi.

Määritelmä 12.10. 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 12.11. Esimerkin 12.2 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 9.13 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 12.10 vaadittu lävistäjämatriisi on siis

$$D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}.$$

Esimerkissä 12.11 matriisi P vain tupsahti jostakin. Vertaamalla esimerkkiin 12.2 huomataan kuitenkin, että matriisin D lävistäjäalkiot ovat matriisin A ominaisarvot, ja matriisin P sarakkeet ovat jotkin niitä vastaavat ominaisvektorit. Seuraava lause osoittaa, että näin on aina, jos matriisi on diagonalisoituva.

Lause 12.12. *Neliömatriisi $A \in \mathbb{R}^{n\times n}$ on diagonalisoituva, jos ja vain jos sillä on n lineaarisesti riippumatonta 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. "⇒": 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 $\bar{p}_1, \dots, \bar{p}_n$ ja matriisin D lävistäjäalkiot $\lambda_1, \dots, \lambda_n$. Nyt siis

$$P = \begin{bmatrix} \bar{p}_1 & \dots & \bar{p}_n \end{bmatrix} \quad \text{ja} \quad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

Matriisituloa laskettaessa tulon AP jokainen sarake saadaan kertomalla matriisilla A vastaava sarake matriisista P:

$$AP = A[\bar{p}_1 \cdots \bar{p}_n] = [A\bar{p}_1 \cdots A\bar{p}_n].$$

Toisaalta lävistäjämatriisia D kerrottaessa tullaan kertoneeksi matriisin P jokainen sarake vastaavalla lävistäjäalkiolla:

$$PD = [\lambda_1\bar{p}_1 \cdots \lambda_n\bar{p}_n].$$

Koska $AP = PD$, nähdään nyt, että $A\bar{p}_i = \lambda_i\bar{p}_i$ kaikilla $i \in \{1, \dots, n\}$. Siis jokainen λ_i on ominaisarvo ja $\bar{p}_i$ sitä vastaava ominaisvektori.

On vielä osoitettava, että ominaisvektorien jono $(\bar{p}_1, \dots, \bar{p}_n)$ on vapaa. Koska P on kääntyvä, yhtälöllä $P\bar{x} = \bar{0}$ on lauseen 10.1 mukaan täsmälleen yksi ratkaisu $\bar{x} = \bar{0}$. Yhtälö $P\bar{x} = \bar{0}$ voidaan kirjoittaa myös muotoon

$$x_1\bar{p}_1 + x_2\bar{p}_2 + \cdots + x_n\bar{p}_n = \bar{0}.$$

Tämän yhtälön ainoa ratkaisu on siis $x_1 = 0, \dots, x_n = 0$. Näin ollen matriisin A ominaisvektoreiden jono $(\bar{p}_1, \dots, \bar{p}_n)$ on vapaa.

"⇐": Oletetaan, että $\bar{p}_1, \dots, \bar{p}_n$ ovat jotkin matriisin A lineaarisesti riippumattomat ominaisvektorit. Olkoot niitä vastaavat ominaisarvot $\lambda_1, \dots, \lambda_n$. Nyt $A\bar{p}_i = \lambda_i\bar{p}_i$ kaikilla $i \in \{1, \dots, n\}$. Olkoon P matriisi, jonka sarakkeet ovat ominaisvektorit: $P = [\bar{p}_1 \cdots \bar{p}_n]$. 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 $P\bar{x} = \bar{0}$ täsmälleen yksi ratkaisu $\bar{x} = \bar{0}$. (Tämä nähdään samalla tavalla kuin todistuksen ensimmäisessä osassa.) Lauseen 10.7 nojalla matriisi P on nyt kääntyvä. Yhtälö $AP = PD$ saadaan siis muotoon

$$P^{-1}AP = D.$$

□

Esimerkki 12.13. Tutkitaan, onko esimerkin 12.6 matriisi

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

diagonalisoituva. Esimerkissä 12.6 todettiin, että matriisin ominaisarvot ovat 4 ja -1. Eräät näitä ominaisarvoja vastaavat ominaisvektorit ovat $(2,3)$ ja $(-1,1)$. Nämä ominaisvektorit ovat lineaarisesti riippumattomatm joten lauseen 12.12 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 12.12 nojalla pätee $P^{-1}AP = D$. Tämän voi vielä tarkistaa laskemalla.

Esimerkki 12.14. 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 $A\bar{x} = 2\bar{x}$. Kun yhtälö ratkaistaan, nähdään sen ratkaisujen olevan muotoa $\bar{x} = (t, 0)$, missä $t \in \mathbb{R} \setminus \{0\}$. Matriisilla A ei siis ole kahta lineaarisesti riippumatonta ominaisarvoa, joten A ei ole diagonalisoituva.

Esimerkki 12.15 (Diagonalisoituvan matriisin potenssit)**.** Lasketaan esimerkissä 12.11 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 12.11 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

$$\begin{aligned} A^7 &= (PDP^{-1})^7 \\ &= \underbrace{(PDP^{-1})(PDP^{-1})\ldots(PDP^{-1})(PDP^{-1})}_{\text{7 kpl}} \\ &= PD(P^{-1}P)D\ldots(P^{-1}P)DP^{-1} \\ &= P\underbrace{D\ldots D}_{\text{7 kpl}}P^{-1} \\ &= PD^7P^{-1}. \end{aligned}$$

Matriisin A potenssin määritäminen on siis muuttunut matriisin D potenssin määrittämiseksi. Se osoittautuu helposksi. Huomataan nimittäin, että

$$D^7 = \begin{bmatrix} 4^7 & 0 \\ 0 & 2^7 \end{bmatrix} = \begin{bmatrix} 16384 & 0 \\ 0 & 128 \end{bmatrix}.$$

Lävistäjämatriisin potenssi saadaan itse asiassa aina selville laskemalla pelkät lävistäjäalkioiden potenssit.

Nyt

$$\begin{aligned} A^7 = (PDP^{-1})^7 = PD^7P^{-1} &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 16384 & 0 \\ 0 & 128 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 16384 & 128 \\ 16384 & -128 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 16512 & 16256 \\ 16256 & 16512 \end{bmatrix} = \begin{bmatrix} 8256 & 8128 \\ 8128 & 8256 \end{bmatrix}. \end{aligned}$$

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 12.16. *Oletetaan, että A on $n \times n$-matriisi. Oletetaan, että $\lambda_1, \ldots, \lambda_m$ ovat matriisin A eri ominaisarvoja ja $\bar{v}_1, \ldots, \bar{v}_m \in \mathbb{R}^n$ jotkin niitä vastaavat ominaisvektorit. Tällöin jono $(\bar{v}_1, \ldots, \bar{v}_m)$ on vapaa.*

Todistus. Oletetaan vastoin väitettä, että jono $(\bar{v}_1, \ldots, \bar{v}_m)$ on sidottu. Nyt lauseen 5.7 nojalla jokin jonon vektoreista on muiden lineaarikombinaatio. Tästä seuraa, että jokin jonon vektoreista on sitä edeltävien jonon vektoreiden lineaarikombinaatio. Olkoon $\bar{v}_{k+1}$ jonon ensimmäinen vektori, joka on sitä edeltävien vektoreiden lineaarikombinaatio. Tällöin on olemassa reaaliluvut $c_1, \ldots, c_k$, joille pätee

$$c_1\bar{v}_1 + \cdots + c_k\bar{v}_k = \bar{v}_{k+1}. \tag{10}$$

Lisäksi jono $(\bar{v}_1, \ldots, \bar{v}_k)$ on vapaa. Jos se nimittäin ei olisi vapaa, $\bar{v}_{k+1}$ ei olisikaan ensimmäinen vektori, joka on sitä edeltävien vektoreiden lineaarikombinaatio.

Kertomalla yhtälön (10) molemmat puolet vasemmalta matriisilla A saadaan yhtälö

$$A(c_1\bar{v}_1 + \cdots + c_k\bar{v}_k) = A\bar{v}_{k+1}.$$

Matriisien laskusääntöjen avulla yhtälö saa muodon $c_1A\bar{v}_1 + \cdots + c_kA\bar{v}_k = A\bar{v}_{k+1}$. Kun vielä muistetaan, että vektorit $\bar{v}_1, \ldots, \bar{v}_k$ ovat matriisin A ominaisvektoreita, saadaan lopulta yhtälö

$$c_1\lambda_1\bar{v}_1 + \cdots + c_k\lambda_k\bar{v}_k = \lambda_{k+1}\bar{v}_{k+1}. \tag{11}$$

Toisaalta voidaan kertoa yhtälön (10) molemmat puolet luvulla λ_{k+1} päätyen yhtälöön

$$c_1\lambda_{k+1}\bar{v}_1 + \cdots + c_k\lambda_{k+1}\bar{v}_k = \lambda_{k+1}\bar{v}_{k+1}. \tag{12}$$

Vähennetään yhtälöstä (11) puolittain yhtälö (12), jolloin saadaan

$$c_1(\lambda_1 - \lambda_{k+1})\bar{v}_1 + \cdots + c_k(\lambda_k - \lambda_{k+1})\bar{v}_k = \bar{0}.$$

Jono $(\bar{v}_1, \ldots, \bar{v}_k)$ on vapaa, joten kaikkien yhtälössä olevien kertoimien on oltava nollia:

$$c_1(\lambda_1 - \lambda_{k+1}) = 0,\ c_2(\lambda_2 - \lambda_{k+1}) = 0, \ldots,\ c_k(\lambda_k - \lambda_{k+1}) = 0.$$

Koska $\lambda_1, \ldots, \lambda_m$ ovat kaikki eri ominaisarvoja, niin tiedetään, että $(\lambda_i - \lambda_{k+1}) \neq 0$ kaikilla $i \in \{1, \ldots, k\}$. Tulon nollasäännön nojalla

$$c_1 = 0,\ c_2 = 0,\ \ldots,\ c_k = 0.$$

Näin ollen

$$\bar{v}_{k+1} = c_1\bar{v}_1 + \cdots + c_k\bar{v}_k = 0\bar{v}_1 + \cdots + 0\bar{v}_k = \bar{0}.$$

Toisaalta oletuksen mukaan $\bar{v}_{k+1}$ on matriisin A ominaisvektori, joten $\bar{v}_{k+1} \neq \bar{0}$. Koska päädyttiin ristiriitaan, vastaoletus ei voi olla tosi. Siis alkuperäinen väite pätee, eli jono $(\bar{v}_1, \ldots, \bar{v}_m)$ on vapaa. □

Edellisestä lauseesta seuraa, että toisinaan matriisin diagonalisoituvuus on helppo todeta.

Korollaari 12.17. *Oletetaan, että $n \times n$-matriisilla on n eri ominaisarvoa. Tällöin A on diagonalisoituva.*

Todistus. Olkoot $\bar{v}_1, \ldots, \bar{v}_n$ jotkin eri ominaisarvoihin liittyvät ominaisvektorit. Ne ovat lineaarisesti riippumattomia lauseen 12.16 nojalla. Koska matriisilla A on n lineaarisesti riippumatonta ominaisvektoria, on A diagonalisoituva lauseen 12.12 nojalla. □

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 $I^{-1}AI = A$. Lävistäjämatriisi on kolmiomatriisi, joten sen ominaisarvot voidaan lukea suoraan lävistäjältä. Havaitaan, että matriisilla A on vain yksi ominaisarvo, -3.

13 Pistetulo

Avaruuksissa $\mathbb{R}^2$ ja $\mathbb{R}^3$ on totuttu puhumaan vektorien pituuksista ja vektoreiden välisistä kulmista. Kuten tavallista, 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 13.1. Vektoreiden $\bar{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ ja $\bar{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n$ *pistetulo* on

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

Vektorien pistetulo on aina reaaliluku. Jos esimerkiksi $\bar{v} = (3, -2, 0)$ ja $\bar{w} = (1, -2, \sqrt{3})$, vektorien $\bar{v}$ ja $\bar{w}$ pistetulo on

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

Pistetulolle voidaan johtaa laskusääntöjä.

Lause 13.2. *Oletetaan, että $\bar{v}, \bar{w}, \bar{u} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$. Tällöin*

a) $\bar{v} \cdot \bar{w} = \bar{w} \cdot \bar{v}$

b) $\bar{v} \cdot (\bar{w} + \bar{u}) = \bar{v} \cdot \bar{w} + \bar{v} \cdot \bar{u}$

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

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

$$\begin{aligned}
\bar{v} \cdot (\bar{w} + \bar{u}) &= (v_1, \ldots, v_n) \cdot (w_1 + u_1, w_2 + u_2, \ldots, w_n + u_n) \\
&= v_1(w_1 + u_1) + v_2(w_2 + u_2) + \cdots + v_n(w_n + u_n) \\
&= v_1 w_1 + v_1 u_1 + v_2 w_2 + v_2 u_2 + \cdots + v_n w_n + v_n u_n \\
&= (v_1 w_1 + v_2 w_2 + \cdots + v_n w_n) + (v_1 u_1 + v_2 u_2 + \cdots + v_n u_n) \\
&= \bar{v} \cdot \bar{w} + \bar{v} \cdot \bar{u}.
\end{aligned}$$

Tässä käytettiin reaalilukujen yhteenlaskun ja kertolaskun osittelulakia. □

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

$$(\bar{v} + \bar{w}) \cdot \bar{u} \overset{a)}{=} \bar{u} \cdot (\bar{v} + \bar{w}) \overset{b)}{=} \bar{u} \cdot \bar{v} + \bar{u} \cdot \bar{w} \overset{a)}{=} \bar{v} \cdot \bar{u} + \bar{w} \cdot \bar{u}.$$

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

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

Lause 13.3. *Oletetaan, että $\bar{v} \in \mathbb{R}^n$. Tällöin*

a) $\bar{v} \cdot \bar{v} \geq 0$

b) $\bar{v} \cdot \bar{v} = 0$, *jos ja vain jos* $\bar{v} = \bar{0}$.

Todistus.

a) Nähdään, että $\bar{v} \cdot \bar{v} = v_1^2 + v_2^2 + \cdots + v_n^2 \geq 0 + 0 + \cdots + 0 = 0$, sillä reaaliluvun neliö on aina epänegatiivinen. Tämä todistaa väitteen.

b) "$\Rightarrow$": Oletetaan, että $\bar{v} \cdot \bar{v} = 0$. Tällöin $v_1^2 + v_2^2 + \cdots + v_n^2 = 0$. Koska jokainen yhteenlaskettava on epänegatiivinen, täytyy yhteenlaskettavien olla nollia. Toisin sanoen $v_i^2 = 0$ kaikilla $i \in \{1, \ldots, n\}$. Tästä seuraa, että $v_i = 0$ kaikilla $i \in \{1, \ldots, n\}$. Siten $\bar{v} = (0, 0, \ldots, 0) = \bar{0}$.

"$\Leftarrow$": Oletetaan, että $\bar{v} = \bar{0}$. Nyt $\bar{v} \cdot \bar{v} = 0^2 + 0^2 + \cdots + 0^2 = 0$. Väite on todistettu. $\square$

13.1 Vektorin normi

Pistetulon avulla voidaan määritellä avaruuden $\mathbb{R}^n$ vektorin normi eli pituus. Lauseen 13.3 nojalla pätee $\bar{v} \cdot \bar{v} \geq 0$, joten seuraavassa määritelmässä juurrettava on epänegatiivinen, kuten kuuluu olla.

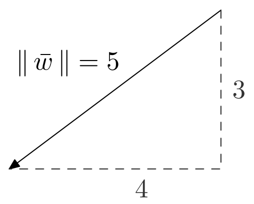
Määritelmä 13.4. Vektorin $\bar{v} \in \mathbb{R}^n$ *normi* eli *pituus* on

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

Määritelmästä seuraa, että $\|\bar{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$, kun pistetulo lasketaan auki. Esimerkiksi vektorin $\bar{v} = (1/2, 3, -2, 0)$ normi on

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

Tasossa normia voi havainnollistaa Pythagoraan lauseen avulla. Kuvaan 13.30 on piirretty vektori $\bar{w} = (-4, -3)$. 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 13.30: Vektorin $\bar{w} = (-4, -3)$ 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 13.5. *Oletetaan, että* $\bar{v} \in \mathbb{R}^n$. *Tällöin*

a) $\|\bar{v}\| \geq 0$

b) $\|\bar{v}\| = 0$, *jos ja vain jos* $\bar{v} = \bar{0}$.

Todistus. Tulokset seuraavat suoraan normin määritelmästä, neliöjuuren ominaisuuksista ja lauseesta 13.3.

a) Määritelmän mukaan $\|\bar{v}\| = \sqrt{\bar{v} \cdot \bar{v}}$. Neliöjuuren arvo on aina epänegatiivinen, joten $\|\bar{v}\| \geq 0$.

b) Huomataan, että $\|\bar{v}\| = \sqrt{\bar{v} \cdot \bar{v}} = 0$, jos ja vain jos juurrettava $\bar{v} \cdot \bar{v}$ on nolla. Lauseen 13.3 nojalla taas $\bar{v} \cdot \bar{v} = 0$, jos ja vain jos $\bar{v} = \bar{0}$. Tämä todistaa väitteen. □

Kun vektoria kerrotaan skalaarilla, tavallisinta on sanoa, että vektorin suunta pysyy samana, mutta sen pituutta ”skaalataan”. Tähän asti kyse on ollut vain intuitiivisesta kuvailusta, mutta normin käsitteen avulla asia voidaan ilmaista tarkasti. On vain otettava huomioon, että pituus on aina epänegatiivinen, vaikka skalaarikerroin olisikin negatiivinen.

Lause 13.6. *Oletetaan, että* $\bar{v} \in \mathbb{R}^n$ *ja* $c \in \mathbb{R}$. *Tällöin* $\|c\bar{v}\| = |c|\|\bar{v}\|$.

Todistus. Todistus jätetään harjoitustehtäväksi. □

Jos vain vektorin suunnalla on merkitystä, pyritään usein yksinkertaisuuden vuoksi rajoittumaan vektoreihin, joiden pituus on yksi. Tällaisilla vektoreilla on oma nimityksensä.

Määritelmä 13.7. Vektori $\bar{u} \in \mathbb{R}^n$ on *yksikkövektori*, jos sen normi on yksi eli

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

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

Esimerkki 13.8. Etsitään jokin yksikkövektori, joka on yhdensuuntainen vektorin $\bar{v} = (2, -1, 0)$ kanssa. Vektorin $\bar{v}$ normi on $\|\bar{v}\| = \sqrt{4 + 1} = \sqrt{5}$. Jos vektori $\bar{v}$ kerrotaan skalaarilla $1/\sqrt{5}$, saadaan vektori $(1/\sqrt{5})\bar{v}$, jonka pituus on lauseen 13.6 nojalla

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

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

Edellistä esimerkkiä mukaillen saadaan seuraava yleinen tulos.

Lause 13.9. *Oletetaan, että* $\bar{v} \in \mathbb{R}^n \setminus \{\bar{0}\}$. *Tällöin vektori* $\frac{1}{\|\bar{v}\|}\bar{v}$ *on yksikkövektori, joka on samansuuntainen vektorin* $\bar{v}$ *kanssa.*

Todistus. Todistus jätetään harjoitustehtäväksi. □

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

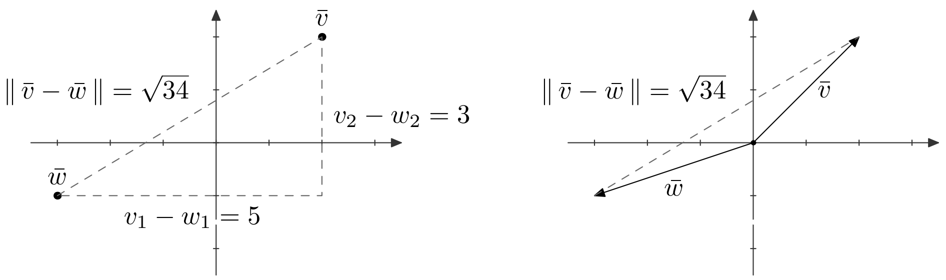
Määritelmä 13.10. Oletetaan, että $\bar{v}$, $\bar{w} \in \mathbb{R}^n$. Vektorien $\bar{v}$ ja $\bar{w}$ välinen *etäisyys* on

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

Esimerkki 13.11. Vektoreiden $\bar{v} = (2, 2)$ ja $\bar{w} = (-3, -1)$ välinen etäisyys on

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

Etäisyyttä on havainnollistettu kahdella eri tavalla kuvassa 13.31.



Kuva 13.31: Vektoreiden $\bar{v}$ ja $\bar{w}$ välinen etäisyys. Ensimmäisessä kuvassa vektorit on havainnollistettu tason pisteinä, jälkimmäisessä kuvassa paikkavektoreina.

13.2 Vektorien kohtisuoruus ja projektio

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

Määritelmä 13.12. Vektorit $\bar{v} \in \mathbb{R}^n$ ja $\bar{w} \in \mathbb{R}^n$ ovat *kohtisuorassa toisiaan vastaan* eli *ortogonaaliset*, jos $\bar{v} \cdot \bar{w} = 0$. Tällöin merkitään $\bar{v} \perp \bar{w}$.

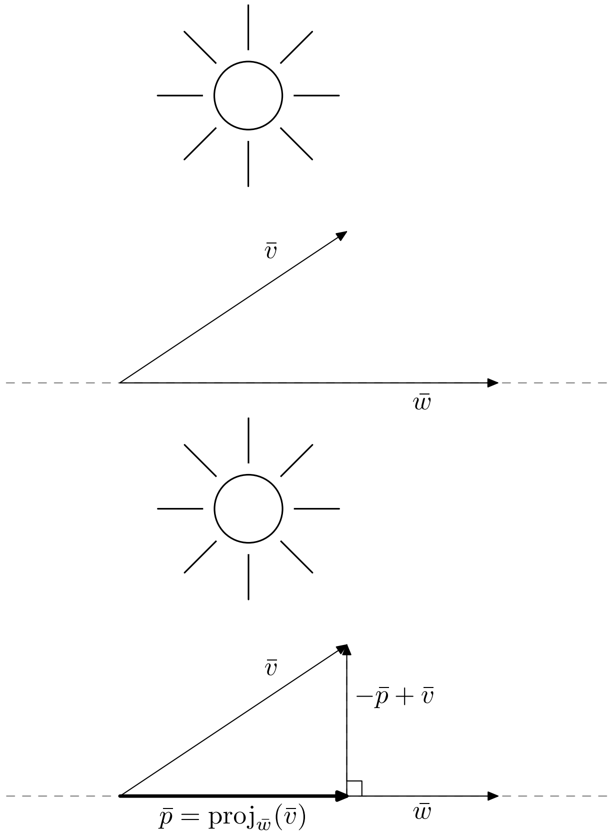
Esimerkiksi tason vektorit $(2, 1)$ ja $(-2, 4)$ ovat kohtisuorassa toisiaan vastaan, sillä

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

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

Kohtisuoruuden käsite tarjoaa mahdollisuuden määritellä vektorin *projektio*. Tätä voidaan ajatella vektorin heittämänä varjona kuvan 13.32 mukaan. Projektio on hyödyllinen työkalu esimerkiksi tietokonegrafiikassa ja geometriassa (ks. esim. 13.22), mutta sillä on myös teoreettiset sovelluksensa. Projektiota tarkastellaan lähemmin kurssin toisessa osassa. Tässä rajoitumme vektorin projektioon jonkin toisen vektorin virittämälle aliavaruudelle (eli suoralle).

Palataan vielä varjovertaukseen. Voidaan ajatella, että vektorin $\bar{v}$ projektio vektorin $\bar{w}$ suuntaiselle suoralle on vektorin $\bar{v}$ heittämä varjo, kun aurinko paistaa kohtisuoraan suoraa vastaan kuten kuvassa 13.32. Matemaattisesti tämä voidaan ilmaista sanomalla, että projektio $\bar{p}$ on yhdensuuntainen vektorin $\bar{w}$ kanssa ja vektorit $\bar{w}$ sekä $-\bar{p}+\bar{v}$ ovat kohtisuorassa toisiaan vastaan.



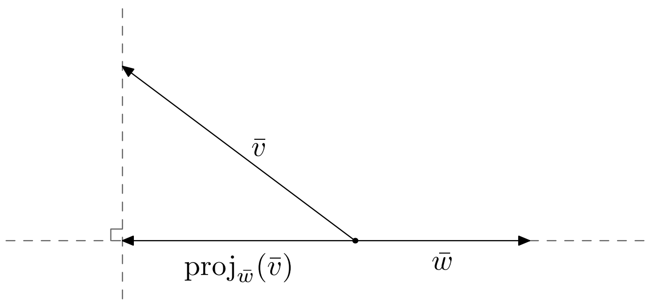
Kuva 13.32: Projektiota voi ajatella vektorin heittämänä varjona.

Määritelmä 13.13. Olkoot $\bar{v}, \bar{w} \in \mathbb{R}^n$ ja $\bar{w} \neq \bar{0}$. Tällöin vektorin $\bar{v}$ *projektio* vektorin $\bar{w}$ virittämälle aliavaruudelle on sellainen vektori $\bar{p} \in \mathbb{R}^n$, että

a) vektori $\bar{p}$ on yhdensuuntainen vektorin $\bar{w}$ kanssa
b) vektori $\bar{v} - \bar{p}$ on kohtisuorassa vektoria $\bar{w}$ vastaan.

Projektiota $\bar{p}$ merkitään $\mathrm{proj}_{\bar{w}}(\bar{v})$.

Vektorin projektion voi määrittää kuvan avulla seuraavasti. Piirretään vektorit $\bar{v}$ ja $\bar{w}$ alkamaan samasta pisteestä ja piirretään vektorin $\bar{w}$ suuntainen suora (ks. kuva 13.33). Projektio $\mathrm{proj}_{\bar{w}}(v)$ löydetään piirtämällä suora, joka on kohtisuorassa vektorin $\bar{w}$ suuntaista suoraa vastaan ja kulkee vektorin $\bar{v}$ kärjen kautta.



Kuva 13.33: Vektorin $\bar{v}$ projektion määrittäminen piirtämällä.

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

Lause 13.14. *Olkoot $\bar{v}, \bar{w} \in \mathbb{R}^n$ ja $\bar{w} \neq \bar{0}$. Tällöin on olemassa täsmälleen yksi projektio $\text{proj}_{\bar{w}}(\bar{v})$ ja se saadaan kaavasta*

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

Todistus. Aloitetaan osoittamalla, että vektori

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

täyttää projektion määritelmässä mainitut ehdot.

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

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

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

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

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

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

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

Pistetulosta tulee aina tulokseksi reaaliluku, joten seuraavaksi voidaan käyttää reaalilukujen laskusääntöjä:

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

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

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

on vektorin $\bar{v}$ projektio vektorin $\bar{w}$ virittämälle aliavaruudelle eli

$$\mathrm{proj}_{\bar{w}}(\bar{v}) = \frac{\bar{v}\cdot\bar{w}}{\bar{w}\cdot\bar{w}}\bar{w}.$$

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

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

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

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

Väite on todistettu. $\square$

Esimerkki 13.15. Esimerkiksi vektorin $\bar{v} = (1,2)$ projektio vektorin $\bar{w} = (-1,3)$ virittämälle aliavaruudelle on

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

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

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

Projektiot on esitetty kuvassa 13.34.

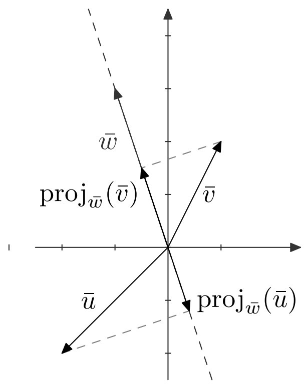
13.3 Vektorien välinen kulma

Suoran kulman lisäksi pistetulo soveltuu myös muiden kulmien määrittämiseen. Ennen määritelmää tarvitaan kuitenkin pari aputulosta. Ensimmäinen on eräs yleisen teorian kannalta tärkeä tulos, jota tässä vaiheessa kuitenkin tarvitaan lähinnä sitä seuraavan lemman todistamiseen.

Lause 13.16 (Schwarzin epäyhtälö)**.** *Oletetaan, että $\bar{v} \in \mathbb{R}^n$ ja $\bar{w} \in \mathbb{R}^n$. Tällöin*

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

Schwarzin epäyhtälön todistus on melko tekninen, ja siksi sen esitystä lykätään kurssin toiseen osaan.



Kuva 13.34: Vektoreiden $\bar{v}$ ja $\bar{u}$ projektiot vektorin $\bar{w}$ virittämälle aliavaruudelle.

Lemma 13.17. *Oletetaan, että* $\bar{v}, \bar{w} \in \mathbb{R}^n \setminus \{\bar{0}\}$. *Tällöin*

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

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

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

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

$$-1 \leq \frac{\bar{v} \cdot \bar{w}}{\|\bar{v}\|\|\bar{w}\|} \leq 1. \qquad \square$$

Nyt voidaan määritellä kahden vektorin välinen kulma. Kosinifunktio on määritelty niin, että jokaista lukua $x \in [-1, 1]$ vastaa täsmälleen yksi sellainen kulma α, että $0° \leq \alpha \leq 180°$ ja $\cos \alpha = x$. Edellisen lemman nojalla voidaan siis asettaa seuraava määritelmä.

Määritelmä 13.18. Vektorien $\bar{v} \in \mathbb{R}^n \setminus \{\bar{0}\}$ ja $\bar{w} \in \mathbb{R}^n \setminus \{\bar{0}\}$ välinen *kulma* on se kulma α, jolle pätee $0° \leq \alpha \leq 180°$ ja

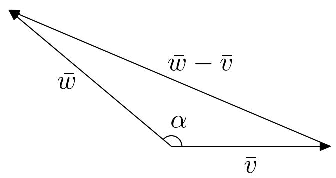
$$\cos \alpha = \frac{\bar{v} \cdot \bar{w}}{\|\bar{v}\|\|\bar{w}\|}.$$

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

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

Lisäksi täytyy päteä $0° \leq \alpha \leq 180°$. Laskimella saadaan vektorien välisen kulman likiarvoksi $\alpha \approx 46{,}65°$.

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



Kuva 13.35: Vektoreiden $\bar{v}$ ja $\bar{w}$ välinen kulma kosinilauseen näkökulmasta.

Kosinilauseen mukaan kuvan 13.35 kolmiossa pätee

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

Toisaalta normin määritelmän ja pistetulon ominaisuuksien nojalla

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

Saadaan siis yhtälö

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

josta edelleen

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

Aiemmin määriteltiin, että vektorit ovat kohtisuorassa toisiaan vastaan, jos niiden pistetulo on nolla. Olisi luonnollista, että toisiaan vastaan kohtisuorien vektorien välinen kulma olisi 90°. Tämä pitääkin paikkansa. Oletetaan nimittäin, että $\bar{v}, \bar{w} \in \mathbb{R}^n \setminus \{\bar{0}\}$. Määritelmän 13.3 mukaan vektoreiden $\bar{v}$ ja $\bar{w}$ välinen kulma on 90°, jos ja vain jos

$$\frac{\bar{v}\cdot\bar{w}}{\|\bar{v}\|\|\bar{w}\|} = \cos 90^\circ.$$

Koska $\cos 90^\circ = 0$, yhtälö pätee, jos ja vain jos $\bar{v}\cdot\bar{w} = 0$. Siten vektoreiden $\bar{v}$ ja $\bar{w}$ välinen kulma on 90°, jos ja vain jos $\bar{v}\cdot\bar{w} = 0$.

13.4 Pistetulon sovelluksia

Tässä alaluvussa esitellään muutamia pistetulon sovelluksia. Kaikille luvun väitteille ei anneta tarkkoja todistuksia.

Normaalimuotoiset yhtälöt

Esimerkissä 8.8 tutkittiin origon kautta kulkevaa tasoa

$$T = \{a_1(3,-1,5) + a_2(2,1,3) \mid a_1, a_2 \in \mathbb{R}\}$$

ja osoitettiin, että sen voi kirjoitta myös muodossa

$$T = \{(u_1, u_2, u_3) \in \mathbb{R}^3 \mid -8u_1 + u_2 + 5u_3 = 0\}.$$

Nyt taso voidaan vielä kirjoittaa hieman eri muodossa pistetulon avulla. Huomataan nimittäin, että

$$-8u_1 + u_2 + 5u_3 = (-8, 1, 5) \cdot (u_1, u_2, u_3),$$

joten

$$T = \{(u_1, u_2, u_3) \in \mathbb{R}^3 \mid (-8, 1, 5) \cdot (u_1, u_2, u_3) = 0\}.$$

Toisin sanoen taso T koostuu täsmälleen niistä vektoreista, jotka ovat kohtisuorassa vektoria $(-8, 1, 5)$ vastaan.

Lähestytään samaa asiaa sitten toisesta näkökulmasta. Tarkastellaan kaikkia avaruuden $\mathbb{R}^3$ vektoreita $\bar{v}$, jotka ovat kohtisuorassa vektoria $\bar{n} = (1, 2, 3)$ vastaan. Millaisen joukon ne muodostavat? Vektorit toteuttavat yhtälön $\bar{n} \cdot \bar{v} = 0$ eli

$$v_1 + 2v_2 + 3v_3 = 0.$$

Vektorien muodostama joukko A on

$$\begin{aligned}
A &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid (1, 2, 3) \cdot (v_1, v_2, v_3) = 0\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 + 2v_2 + 3v_3 = 0\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = -2v_2 - 3v_3\} \\
&= \{(-2v_2 - 3v_3,\ v_2,\ v_3) \mid v_2, v_3 \in \mathbb{R}\} \\
&= \{v_2(-2, 1, 0) + v_3(-3, 0, 1) \mid v_2, v_3 \in \mathbb{R}\}.
\end{aligned}$$

Viimeisestä muodosta nähdään, että kyseessä on origon kautta kulkeva taso.

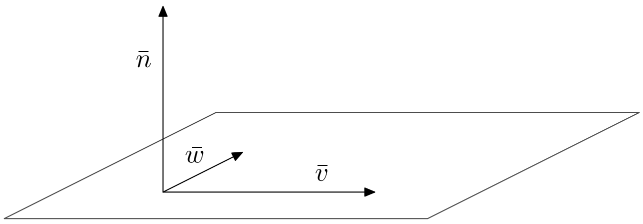
Esimerkkien perusteella origon kautta kulkeva taso muodostuu tämälleen niistä vektoreista, jotka ovat kohtisuorassa jotakin annettua vektoria vastaan. Toisaalta jos valitaan jokin vektori $\bar{n} \in \mathbb{R}^3 \setminus \{\bar{0}\}$ ja muodostetaan joukko, jossa ovat kaikki vektoria $\bar{n}$ vastaan kohtisuorassa olevat vektorit, saadaan aikaiseksi taso. Tasoja voidaan siis luonnehtia niitä vastaan kohtisuorassa olevien vektorien avulla.

Vektoria, joka on kohtisuorassa kaikkia origon kautta kulkevan kyseisen tason vektoreita kohtaan, kutsutaan tason *normaaliksi.* Edellä esiintynyt vektori $(-8, 1, 5)$ on tason T normaali, ja vektori $(1, 2, 3)$ on tason A normaali. Jotta voidaan ottaa huomioon myös sellaiset tasot, jotka eivät kulje origon kautta, täytyy normaali määritellä hieman yleisemmällä tavalla.

Määritelmä 13.19. Olkoon T avaruuden $\mathbb{R}^3$ taso. Vektoria, joka on kohtisuorassa tason T suuntavektoreita vastaan, kutsutaan tason *normaaliksi*.

Esimerkki 13.20. Vektori $(2, 20, 12)$ on tason

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



Kuva 13.36: Tason normaali $\bar{n}$.

normaali, sillä $(2,20,12)\cdot(4,0,-1)=0$ ja $(2,20,12)\cdot(0,3,-5)=0$. Vektori $(2,20,12)$ on itse asiassa kohtisuorassa kaikkia tason T vektoreita vastaan, sillä

$$\begin{aligned}(2,20,12)\cdot(s(4,0,-1)+t(0,3,-5)) &= s(2,20,12)\cdot(s(4,0,-1))+(2,20,12)\cdot(t(0,3,-5))\\ &= s((2,20,12)\cdot(4,0,-1))+t((2,20,12)\cdot(0,3,-5))\\ &= s\cdot 0+t\cdot 0=0.\end{aligned}$$

Vektori $(2,20,12)$ 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:

$$(2,20,12)\cdot(4,0,-1)=0 \quad \text{ja} \quad (2,20,12)\cdot(0,3,-5)=0.$$

Vektori $(2,20,12)$ ei kuitenkaan ole kohtisuorassa mitään tason U vektoria vastaan. (Tämän tarkistaminen jätetään lukijalle.) Erona edelliseen tapaukeen on se, että taso U ei kulje origon kautta.

Määritelmänsä mukaan normaali on kohtisuorassa tason suuntavektoreita vastaan. Jos taso kulkee origon kautta, tason normaali on kohtisuorassa *kaikkia* tason vektoreita vastaan. Toisaalta mitkään muut kuin kyseisen tason vektorit eivät ole kohtisuorassa normaalia vastaan. (Väitteiden osoittaminen jätetään harjoitustehtäväksi.) Toisin sanoen, jos $\bar{n}$ on origon kautta kulkevan tason T normaali, piste $\bar{x}$ on tasossa T, jos ja vain jos

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

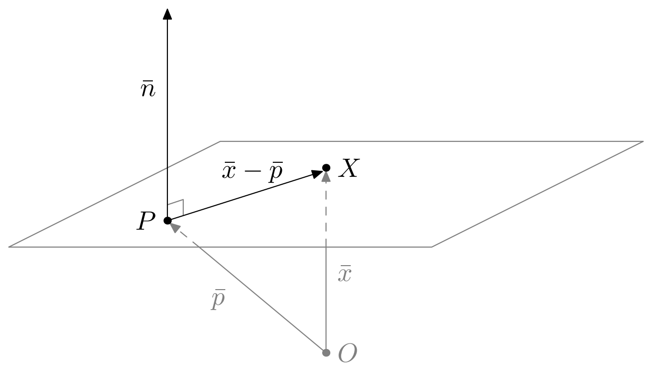
Jos taso ei kulje origon kautta, saa tason yhtälö hiukan toisen muodon. Oletetaan, että T on avaruuden $\mathbb{R}^3$ taso, jonka paikkavektori on $\bar{p}$ ja jolla on normaali $\bar{n}$. Tällöin $\bar{x}$ on tasossa T, jos ja vain jos

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

Tilannetta on havainnollistettu kuvassa 13.37. Yhtälöä

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

kutsutaan tason T *normaalimuotoiseksi yhtälöksi*.



Kuva 13.37: Tason T normaalimuotoisen yhtälön havainnollistus.

Esimerkki 13.21. Oletetaan, että avaruuden $\mathbb{R}^3$ taso T kulkee pisteen $P = (6, 0, 1)$ kautta ja sillä on normaali $\bar{n} = (1, 2, 3)$. Tason T normaalimuotoinen yhtälö on tällöin

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

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

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

Kirjoitetaan yhtälö vielä hiukan toisenlaisessa muodossa. Merkitään $\bar{x} = (x, y, z)$, missä $x, y, z \in \mathbb{R}$. Nyt

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

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

Edellisestä esimerkkejä mukaillen voidaan osoittaa, että tason normaalimuotoinen yhtälö on aina muotoa

$$ax + by + cz + d = 0,$$

missä $a, b, c, d \in \mathbb{R}$. Avaruuden $\mathbb{R}^2$ suorille on mahdollista johtaa normaalimuotoinen yhtälö samalla tavalla kuin avaruuden $\mathbb{R}^3$ tasoille. Suorien normaalimuotoiset yhtälöt ovat muotoa

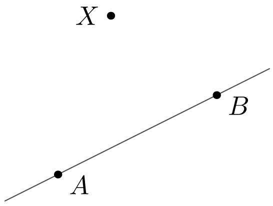
$$ax + bx + c = 0,$$

missä $a, b, c \in \mathbb{R}$.

Pisteen etäisyys suorasta

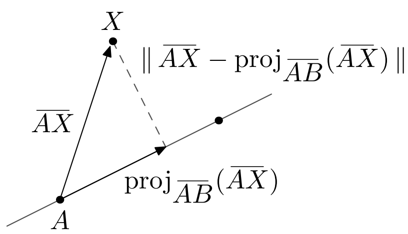
Pisteen X etäisyys suorasta $S = \{\bar{p} + t\bar{v} \mid t \in \mathbb{R}\}$ on kaikkein lyhin välimatka, joka voi olla pisteen X ja suoran S jonkin pisteen välillä. Täsmällisesti ilmaistuna pisteen X etäisyys suorasta S on $\min\{d(\bar{x}, \bar{a}) \mid \bar{a} \in S\}$, missä $\bar{x}$ on pisteen X paikkavektori. 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. Tutkitaan tätä esimerkin avulla.

Esimerkki 13.22. Määritetään pisteen $X = (4, -1, 9)$ etäisyys suorasta S, joka kulkee pisteiden $A = (2, -3, 5)$ ja $B = (4, 1, 7)$ kautta (ks. kuva 13.38).



Kuva 13.38: Pisteiden A ja B kautta kulkeva suora S.

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



Kuva 13.39: Pisteen X etäisyys suorasta S.

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

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

Erotus $\overline{AX} - \text{proj}_{\overline{AB}}(\overline{AX})$ 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 $\frac{1}{3}\sqrt{66}$.

14 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 nimenomaan kolmiulotteisen avaruuden laskutoimitus.

Määritelmä 14.1. Vektorien $\bar{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$ ja $\bar{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$ *ristitulo* on vektori

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

Ristitulon $\bar{v} \times \bar{w}$ laskemiseen voi käyttää kuvassa 14.40 esitettyä laskusääntöä. Yhtenäisellä viivalla yhdistettyjen komponenttien tulosta vähennetään katkoviivalla yhdistettyjen komponenttien tulo.

$$\begin{array}{ccccc} v_1 & v_2 & v_3 & v_1 & v_2 \\ & & & & \\ w_1 & w_2 & w_3 & w_1 & w_2 \end{array}$$

Kuva 14.40: Ristitulon $\bar{v} \times \bar{w}$ laskeminen.

Esimerkki 14.2. Merkitään $\bar{a} = (2, 1, 4)$ ja $\bar{b} = (3, -1, -3)$. Kuvan 14.41 perusteella voidaan laskea

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

$$\begin{array}{ccccc} 2 & 1 & 4 & 2 & 1 \\ & & & & \\ 3 & -1 & -3 & 3 & -1 \end{array}$$

Kuva 14.41: Ristitulon $\bar{a} \times \bar{b}$ laskeminen.

Ristitulolle saadaan toinen muistisääntö determinantin avulla. Vektoreiden $\bar{v}$ ja $\bar{w}$ ristitulo saadaan laskemalla determinantti

$$\begin{vmatrix} \bar{e}_1 & \bar{e}_2 & \bar{e}_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

Tässä $\bar{e}_1 = (1, 0, 0)$, $\bar{e}_2 = (0, 1, 0)$ ja $\bar{e}_3 = (0, 0, 1)$. Tarkalleen ottaen oikean determinantin alkiot eivät voisi olla vektoreita, mutta kyseessä on vain muistisääntö: vektoreiden $\bar{e}_1$, $\bar{e}_2$ ja $\bar{e}_3$ ajatellaan käyttäytyvän determinanttia laskettaessa reaalilukujen tavoin.

Esimerkki 14.3. Esimerkiksi vektoreiden $\bar{a} = (2, 1, 4)$ ja $\bar{b} = (3, -1, -3)$ ristitulo on

$$\begin{aligned}\begin{vmatrix} \bar{e}_1 & \bar{e}_2 & \bar{e}_3 \\ 2 & 1 & 4 \\ 3 & -1 & -3 \end{vmatrix} &= \bar{e}_1(1 \cdot (-3) - 4 \cdot (-1)) - \bar{e}_2(2 \cdot (-3) - 4 \cdot 3) + \bar{e}_3(2 \cdot (-1) - 1 \cdot 3) \\ &= \bar{e}_1 + 18\bar{e}_2 - 5\bar{e}_3 = (1, 18, -5).\end{aligned}$$

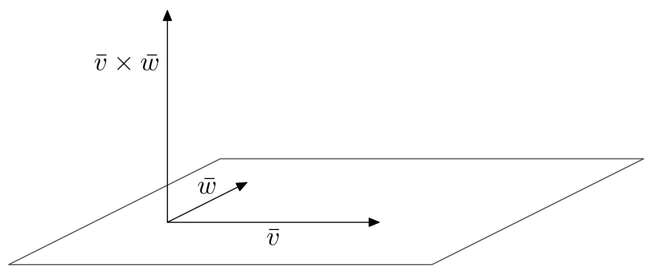
Eräs ristitulon sovelluksista on, että sen avulla voidaan löytää vektori, joka on kohtisuorassa yhtä aikaa kahta vektoria vastaan.

Lause 14.4. *Oletetaan, että $\bar{v}, \bar{w} \in \mathbb{R}^3$. Tällöin $(\bar{v} \times \bar{w}) \perp \bar{v}$ ja $(\bar{v} \times \bar{w}) \perp \bar{w}$.*

Todistus. Laskemalla huomataan, että

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

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



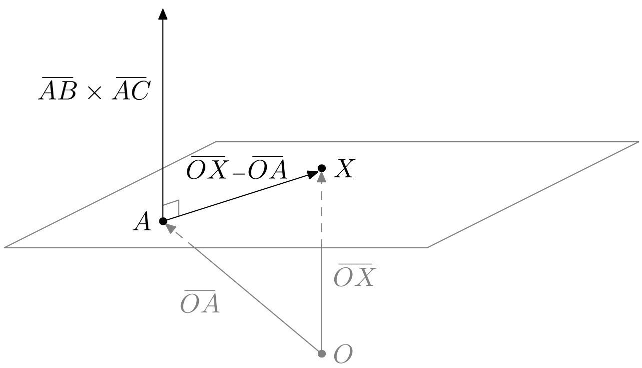
Kuva 14.42: Ristitulo $\bar{v} \times \bar{w}$ on kohtisuorassa vektoria $\bar{v}$ ja vektoria $\bar{w}$ vastaan.

Edellisestä lauseesta seuraa, että ristitulon avulla voidaan löytää tason normaali (eli vektori, joka on kohtisuorassa tason suuntavektoreita vastaan). Tästä on hyötyä tason normaalimuotoisen yhtälön määrittämisessä.

Esimerkki 14.5. Määritetään esimerkkiä 13.21 mukaillen 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 normaali. Normaali on vektori, joka on kohtisuorassa tason suuntavektoreita vastaan. Valitaan suuntavektoreiksi suuntajanat $\overline{AB} = (-1, 2, 2)$ ja $\overline{AC} = (-2, -1, 1)$, jolloin normaaliksi käy edellisen lauseen nojalla 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 $\bar{x} = \overline{OX} = (x, y, z)$, tason normaalimuotoiseksi yhtälöksi saadaan esimerkin 13.21 mukaisesti

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



Kuva 14.43: Tason T normaalimuotoisen yhtälön määrittäminen.

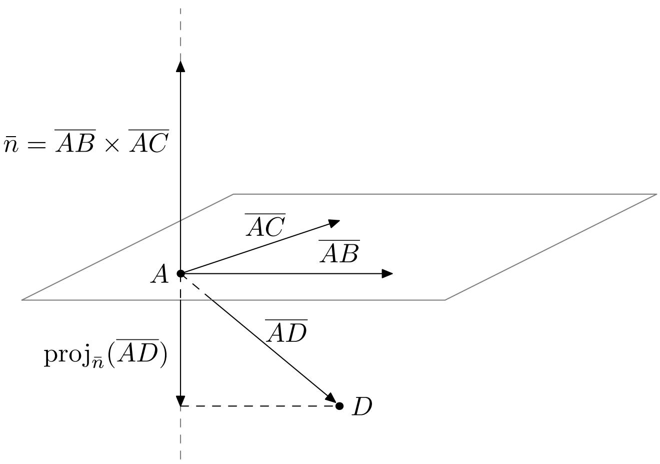
Kun yhtälöön sijoitetaan luvut, se tulee muotoon

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

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

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

Esimerkki 14.6. Pisteen etäisyys tasosta voidaan määrittää ristitulon ja projektion avulla. Merkitään $A = (0, 1, 0)$, $B = (-1, 3, 2)$ ja $C = (-2, 0, 1)$. Oletetaan, että taso T kulkee pisteiden A, B ja C kautta. Määritetään pisteen $D = (1, 2, 3)$ etäisyys tasosta T (ks. kuva 14.44).



Kuva 14.44: Pisteen D etäisyys tasosta T.

Tason suuntaisten vektoreiden $\overline{AB} = (-1, 2, 2)$ ja $\overline{AC} = (-2, -1, 1)$ ristitulo $\overline{AB} \times \overline{AC} = (4, -3, 5)$ on tason normaali. Lisäksi tarvitaan vektori jostakin tason pisteestä pisteeseen $D =$

$(1, 2, 3)$. Valitaan vektori $\overline{AD} = \overline{OD} - \overline{OA} = (1, 1, 3)$. Vektorin $\overline{AD}$ projektio normaalin $\bar{n} = (4, -3, 5)$ virittämälle aliavaruudelle on

$$\operatorname{proj}_{\bar{n}}(\overline{AD}) = \frac{\overline{AD} \cdot \bar{n}}{\bar{n} \cdot \bar{n}}\, \bar{n} = \frac{16}{50}(4, -3, 5) = \frac{8}{25}(4, -3, 5).$$

Tämän projektion normi (eli pituus) on pisteen D etäisyys tasosta T:

$$\|\operatorname{proj}_{\bar{n}}(\overline{AD})\| = \frac{8}{25}\,\|(4, -3, 5)\| = \frac{8}{25}\sqrt{16 + 9 + 25} = \frac{8}{25}\sqrt{50} = \frac{8}{5}\sqrt{2}.$$

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

Lause 14.7. *Oletetaan, että $\bar{u}$, $\bar{v}$, $\bar{w} \in \mathbb{R}^3$ ja $c \in \mathbb{R}$. Tällöin*

a) $\bar{v} \times \bar{w} = -(\bar{w} \times \bar{v})$ *(antikommutointi)*

b) $\bar{u} \times (\bar{v} + \bar{w}) = \bar{u} \times \bar{v} + \bar{u} \times \bar{w}$ *(osittelulaki)*

c) $(\bar{v} + \bar{w}) \times \bar{u} = \bar{v} \times \bar{u} + \bar{w} \times \bar{u}$ *(osittelulaki)*

d) $c(\bar{v} \times \bar{w}) = (c\bar{v}) \times \bar{w} = \bar{v} \times (c\bar{w})$

e) $\bar{v} \times \bar{v} = \bar{0}$

f) $\bar{0} \times \bar{v} = \bar{0}$ *ja* $\bar{v} \times \bar{0} = \bar{0}$

g) $\bar{u} \cdot (\bar{v} \times \bar{w}) = (\bar{u} \times \bar{v}) \cdot \bar{w}$

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

Ristitulolla ja pistetulolla on yhteyksiä, jotka eivät ole aivan itsestään selviä. Niitä on koottu seuraavaan lauseeseen.

Lause 14.8. *Oletetaan, että $\bar{u}$, $\bar{v}$, $\bar{w} \in \mathbb{R}^3$. Tällöin*

a) $(\bar{u} \times \bar{v}) \times \bar{w} = (\bar{u} \cdot \bar{w})\bar{v} - (\bar{v} \cdot \bar{w})\bar{u}$

b) $\bar{u} \times (\bar{v} \times \bar{w}) = (\bar{u} \cdot \bar{w})\bar{v} - (\bar{u} \cdot \bar{v})\bar{w}$

c) $\|\bar{v} \times \bar{w}\|^2 = \|\bar{v}\|^2\|\bar{w}\|^2 - (\bar{v} \cdot \bar{w})^2$ *(Lagrangen identiteetti)*

Todistus. Osoitetaan kohta c) (eli Lagrangen identiteetti) ja jätetään muut kohdat harjoitustehtäviksi. Käyttämällä lauseen 14.7 kohtaa g) ja lauseen 14.8 kohtaa a) saadaan

$$\begin{aligned}
\|\bar{v} \times \bar{w}\|^2 &= (\bar{v} \times \bar{w}) \cdot (\bar{v} \times \bar{w}) = \big((\bar{v} \times \bar{w}) \times \bar{v}\big) \cdot \bar{w} \\
&= \big((\bar{v} \cdot \bar{v})\bar{w} - (\bar{v} \cdot \bar{w})\bar{v}\big) \cdot \bar{w} = \big(\|\bar{v}\|^2\bar{w} - (\bar{v} \cdot \bar{w})\bar{v}\big) \cdot \bar{w} \\
&= \|\bar{v}\|^2(\bar{w} \cdot \bar{w}) - (\bar{v} \cdot \bar{w})(\bar{v} \cdot \bar{w}) = \|\bar{v}\|^2\|\bar{w}\|^2 - (\bar{v} \cdot \bar{w})^2.
\end{aligned}$$

Siten Lagrangen identiteetti pätee. □

Lauseen 14.4 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 14.9. *Oletetaan, että $\bar{v}, \bar{w} \in \mathbb{R}^3$. Jos $\bar{v} \neq \bar{0}$ ja $\bar{w} \neq \bar{0}$, niin*

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

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

Todistus. Todistuksessa käytetään Lagrangen identiteettiä (lause 14.8). Vektorien välisen kulman määritelmän mukaan $\cos \alpha = (\bar{v} \cdot \bar{w})/(\|\bar{v}\|\|\bar{w}\|)$, ja lisäksi pätee $\sin^2 \alpha + \cos^2 \alpha = 1$. Nyt Lagrangen identiteetistä saadaan

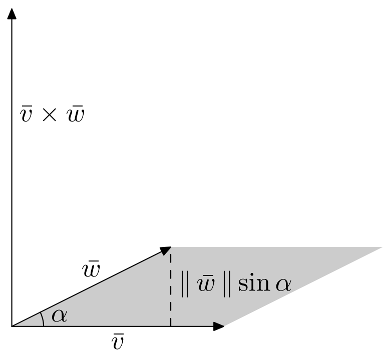
$$\begin{aligned}
\|\bar{v} \times \bar{w}\|^2 &= \|\bar{v}\|^2\|\bar{w}\|^2 - (\bar{v} \cdot \bar{w})^2 = \|\bar{v}\|^2\|\bar{w}\|^2 - (\cos \alpha \cdot \|\bar{v}\|\|\bar{w}\|)^2 \\
&= \|\bar{v}\|^2\|\bar{w}\|^2 - \cos^2 \alpha \cdot \|\bar{v}\|^2\|\bar{w}\|^2 = \|\bar{v}\|^2\|\bar{w}\|^2(1 - \cos^2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

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

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

Tämä todistaa väitteen. □

Edellisestä lauseesta seuraa, että ristitulovektorin $\bar{v} \times \bar{w}$ pituus on yhtä suuri kuin vektorien $\bar{v}$ ja $\bar{w}$ määräämän suunnikkaan ala (kuva 14.45). Oletetaan nimittäin, että vektorien $\bar{v}$ ja $\bar{w}$ välinen kulma on α. Tällöin suunnikkaan korkeus on $\|\bar{w}\| \sin \alpha$. Näin suunnikkaan pinta-alaksi saadaan $\|\bar{w}\| \sin \alpha \cdot \|\bar{v}\| = \|\bar{v} \times \bar{w}\|$.



Kuva 14.45: Ristitulovektorin $\bar{v} \times \bar{w}$ pituus on yhtä suuri kuin vektorien $\bar{v}$ ja $\bar{w}$ määräämän suunnikkaan ala.

Ristitulon avulla voidaan määrittää myös suuntaissärmiön tilavuus. Vektoreiden $\bar{v}$, $\bar{w}$ ja $\bar{u}$ määräämän suuntaissärmiön tilavuus on pohjan pinta-alan $\|\bar{v} \times \bar{w}\|$ ja korkeuden h tulo (kuva 14.46). Korkeuden h selvittämiseksi lasketaan vektorin $\bar{u}$ projektio ristitulovektorin $\bar{v}\times\bar{w}$ virittämälle aliavaruudelle:

$$\operatorname{proj}_{\bar{v}\times\bar{w}}(\bar{u}) = \frac{(\bar{v}\times\bar{w})\cdot\bar{u}}{(\bar{v}\times\bar{w})\cdot(\bar{v}\times\bar{w})}\,(\bar{v}\times\bar{w}).$$

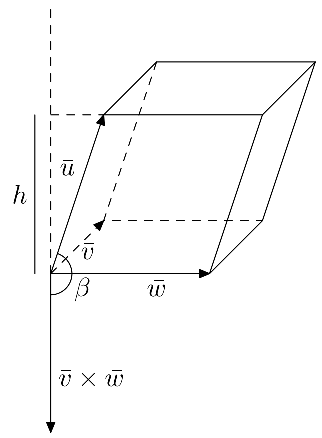
Korkeus h on tämän vektorin pituus eli normi:

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

Tilavuudeksi saadaan pohjan pinta-ala kertaa korkeus:

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

Suuntaissärmiön tilavuus on siis niin kutsutun *skalaarikolmitulon* $(\bar{v}\times\bar{w})\cdot\bar{u}$ itseisarvo. Lauseesta 14.7 seuraa, että vektorien $\bar{v}$, $\bar{w}$ ja $\bar{u}$ järjestyksellä tässä kaavassa ei ole väliä.



Kuva 14.46: Vektoreiden $\bar{v}$, $\bar{w}$ ja $\bar{u}$ määräämän suuntaissärmiön tilavuus.

Hakemisto

A^{-1}, 65
I, 60
I_n, 60
O, 60
O_n, 60
$\mathcal{E}_n$, 37
$\mathbb{R}^n$, 9

aliavaruus, 43
alkeismatriisi, 71
alkeisrivitoimitus, 15
alkio, matriisin, 58
antisymmetrinen, 62
avaruus, 9

determinantti, 77
    kehityskaava, 79
diagonaali, 61
diagonalisointi, 91
diagonalisoituva matriisi, 91
dimensio
    aliavaruuden, 50
    vektoriavaruuden $\mathbb{R}^n$, 42

ekvivalentit yhtälöryhmät, 14
eliminointimentelmä
    Gaussin, 24
    Gaussin-Jordanin, 17
erotus
    matriisien, 59
    vektorien, 9
etäisyys
    pisteen suorasta, 108
    vektorien välinen, 99

Gaussin-Jordanin eliminointimenetelmä, 17

homogeeninen yhtälöryhmä, 35

johtava alkio, 15

käänteismatriisi, 64
kääntyvä matriisi, 64
kanta
    aliavaruuden, 48
    luonnollinen, 37
    vektoriavaruuden $\mathbb{R}^n$, 37
karakteristinen polynomi, 89
kerroinmatriisi, 70
kertolasku
    matriisien, 59
kohtisuoruus, 99
komponentti, 9
koordinaatit, 40
kulma, vektorien välinen, 103

lävistäjä, 61
lävistäjämatriisi, 61
liitännäisyys
    vektoreille, 10
lineaarikombinaatio, 11
lineaarinen yhtälöryhmä, 13
lineaarisesti riippumaton, 31
lineaariyhdistelmä, 11
luonnollinen kanta, 37

matriisi, 58
matriisikertolasku, 59
matriisipotenssi, 60

neliömatriisi, 61
nollamatriisi, 60
nollavektori, 9
normaali

tason, 105
normaalimuotoinen yhtälö
suoran, 107
tason, 57, 106, 107
normi, 97

ominaisarvo, 85
ominaisvektori, 85
ortogonaalinen, 99
osittelulaki
vektoreille, 10

paikkavektori
suoran, 51
tason, 55
pistetulo, 96
pituus, vektorin, 97
porrasmatriisi, 15
redusoitu, 16
potenssi
matriisien, 60
projektio, 100

redusoitu porrasmatriisi, 16
ristitulo, 110
riviekvivalentti, 15

säännöllinen matriisi, 64
sarakevektori, 68
sidottu, 31
singulaarinen matriisi, 64
skalaari, 9
skalaarikertolasku
matriisien, 59
vektoriavaruuden $\mathbb{R}^n$ vektorien, 9
skalaarimatriisi, 61
skalaarimonikerta
matriisin, 59
vektorin, 9
span, 43
summa
matriisien, 58
vektoreiden, 9
suora, 51
suuntavektori
suoran, 51
tason, 55
symmetrinen, 62

taso, 55
transpoosi, 62
tuntematon, 14
tyyppi, matriisin, 58

ulotteinen
aliavaruus, 50
vektoriavaruus, 42

vaidannaisuus
vektoreille, 10
vapaa, 31
vapaa muuttuja, 20
vastavektori, 9
vektoreiden virittämä aliavaruus, 43
vektori, 9
vektoriavaruus, 9
vektoriavaruus $\mathbb{R}^n$, 9
vektorijono, 31
virittäjä
aliavaruuden, 43
avaruuden $\mathbb{R}^n$, 28
virittäminen
aliavaruuden, 43
vektoriavaruuden $\mathbb{R}^n$, 27

yhdensuuntaisuus, 10
yhtälöryhmä, 13
yhteenlasku
matriisien, 58
vektoreiden, 9
ykkösmatriisi, 60
yksikkömatriisi, 60
yksikkövektori, 98