Johdatus lineaarialgebraan

Lotta Oinonen ja Johanna Rämö

6. joulukuuta 2012

Helsingin yliopisto
Matematiikan ja tilastotieteen laitos
2012

Sisältö

1 Avaruus $\mathbb{R}^n$ **4**

1 Avaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ vektorit 4
2 Avaruus $\mathbb{R}^n$ 8
3 Suorat ja tasot 11
3.1 Suora 11
3.2 Taso 14
4 Avaruuden $\mathbb{R}^n$ aliavaruudet 16
5 Lineaariset yhtälöryhmät 19
5.1 Gaussin–Jordanin eliminointimenetelmä 19
6 Virittäminen 31
7 Vapaus 34
7.1 Homogeeniset yhtälöryhmät ja vapaus 39
8 Kanta 41
8.1 Koordinaatit 42
8.2 Dimensio 43
9 Matriisit 46
9.1 Matriisien laskutoimituksia 46
9.2 Erityisiä matriiseja 48
9.3 Matriisien laskusääntöjä 49
9.4 Matriisin transpoosi 49
9.5 Käänteismatriisi 50
9.6 Sarakevektorit 53
10 Matriisit ja yhtälöryhmät 54
10.1 Alkeismatriisit 55
10.2 Käänteismatriisin määrittäminen 58
11 Determinantti 61
11.1 Determinantin kehityskaavat 63
11.2 Determinantin ominaisuuksia 64
12 Pistetulo 66
12.1 Vektorin normi 67
12.2 Vektorien välinen kulma ja kohtisuoruus 69
12.3 Projektio 72
12.4 Ortogonaalinen ja ortonormaali kanta 75

13 Ristitulo 79

2 Vektoriavaruudet **85**
14 Vektoriavaruus 85
15 Aliavaruus 91
15.1 Vektoreiden virittämä aliavaruus 93
16 Vapaus 97
17 Kanta 99
17.1 Dimensio 100
18 Lineaarikuvaus 104
18.1 Lineaarikuvausten yhdistetyt kuvaukset 110
18.2 Aliavaruuden kuva lineaarikuvauksessa 112
19 Ydin ja kuva 113
19.1 Lineaarikuvauksen ydin 113
19.2 Lineaarikuvauksen kuva 116
20 Isomorfismi 118
21 Kanta ja lineaarikuvaukset 120
21.1 Lineaarikuvauksen matriisi 124
22 Ominaisarvot 129
22.1 Karakteristinen polynomi 134
22.2 Diagonalisointi 137
23 Sisätulo 146
23.1 Normi ja kohtisuoruus 147
23.2 Kohtisuora komplementti 148
23.3 Kohtisuora projektio 153
23.4 Ortogonaaliset ja ortonormaalit kannat 158

Luku 1

Avaruus $\mathbb{R}^n$

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

Tässä luvussa käsitellään avaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ vektoreita. Käsiteltävä asia on tuttua koulusta, mutta käytettävät merkinnät ja nimitykset saattavat olla uusia. Jos joukko-opin merkinnät (esim. $\in$, $\subset$, $\setminus$) eivät ole tuttuja, voit katsoa apua kurssisivulla olevasta tiedostosta ”Joukko-opin merkintöjä”.

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

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

Avaruuden $\mathbb{R}^2$ alkoita kutsutaan *vektoreiksi*.

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

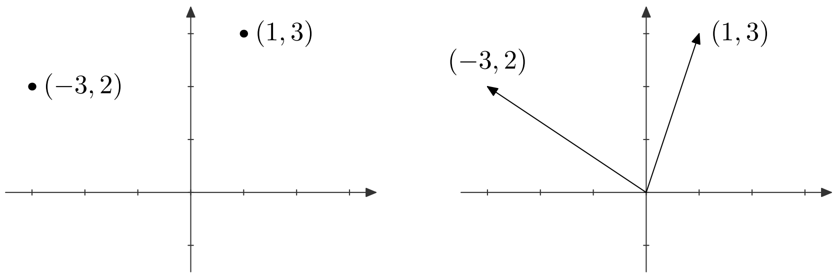
Vektoreita merkitään tässä tekstissä yleensä kirjaimella, jonka päällä on viiva. Voidaan esimerkiksi kirjoittaa $\bar{v} = (a,b)$. Luvut a ja b ovat vektorin $\bar{v}$ *komponentteja*.

Esimerkki 1.2. Esimerkiksi $\bar{v} = (4,-1)$ ja $\bar{u} = (\frac{1}{2}, -\sqrt{5})$ ovat avaruuden $\mathbb{R}^2$ vektoreita. Vektorin $\bar{v}$ komponentit ovat 4 ja -1. Vektorin $\bar{u}$ komponentit ovat puolestaan $\frac{1}{2}$ ja $-\sqrt{5}$.

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 $(2,1)$.

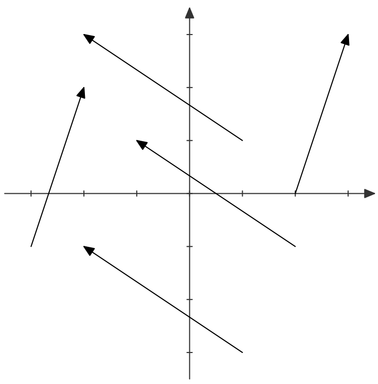
Vektoreita voidaan havainnollistaa koordinaatiston pisteinä. Vektoria (a,b) vastaa piste, jonka vaakakoordinaatti on a ja pystykoordinaatti b. Kuvassa 1.1 on esitetty vektoreita $(1,3)$ ja $(-3,2)$ vastaavat tason pisteet.

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 esitetty kuvassa 1.1.



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

Pisteen ja paikkavektorin lisäksi avaruuden $\mathbb{R}^2$ vektoria voi havainnollistaa mistä tahansa pisteestä lähtevällä suuntajanana. 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.

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.

Tällä kurssilla avaruuden $\mathbb{R}^2$ vektori on määritelmänsä mukaan kahdesta reaaliluvusta koostuva järjestetty pari. Vektoreita voidaan kuitenkin havainnollistaa pisteinä, paikkavektoreina ja suuntajanoina. Se, millainen havainnollistamistapa on paras, riippuu siitä, mitä ollaan tekemässä. Usein vektoreita käsitellessä on pystyttävä vaihtamaan sulavasti yhdestä esitystavasta toiseen.

Vektoreille voidaan määritellä erilaisia laskutoimituksia.

Määritelmä 1.3. 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 voidaa kertoa reaaliluvuilla. Tätä operaatiota kutsutaan *skalaarikertolaskuksi.* Jos $\bar{v} = (v_1, v_2) \in \mathbb{R}^2$ ja $c \in \mathbb{R}$, niin määritellään

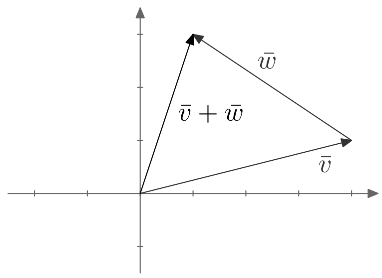
$$c\bar{v} = (cv_1, cv_2).$$

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

Esimerkki 1.4. Tarkastellaan vektoreita $\bar{v} = (4, 1)$ ja $\bar{w} = (-3, 2)$. Niiden summa on

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

Yhteenlaskussa vektorien komponentit lasketaan yhteen, ja siksi yhteenlaskua voidaan havainnollistaa geometrisesti (ks. kuva 1.3). Vektorien summa nähdään asettamalla vektorit peräkkäin niin, että jälkimmäinen vektori alkaa siitä, mihin ensimmäinen päättyi. Summavektorin alkupiste on ensimmäisen vektorin alkupiste ja päätepiste jälkimmäisen vektorin päätepiste.



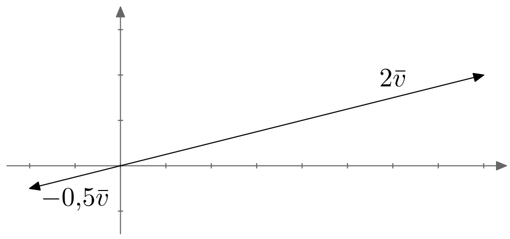
Kuva 1.3: Vektorit $\bar{v}$ ja $\bar{w}$ sekä niiden summa $\bar{v} + \bar{w}$.

Tutkitaan sitten skalaarikertolaskua. Määritelmän mukaan

$$2\bar{v} = (2 \cdot 4, 2 \cdot 1) = (8, 2) \quad \text{ja}$$
$$-\frac{1}{2}\bar{v} = (-\frac{1}{2} \cdot 4; -\frac{1}{2} \cdot 1) = (-2;\ -\frac{1}{2}).$$

Vektorit $2\bar{v}$ ja $-\frac{1}{2}\bar{v}$ on piirretty kuvaan 1.4.

Huomataan, että positiivisella skalaarilla kertominen säilyttää vektorin suunnan ja negatiivisella skalaarilla kertominen kääntää vektorin suunnan vastakkaiseksi.



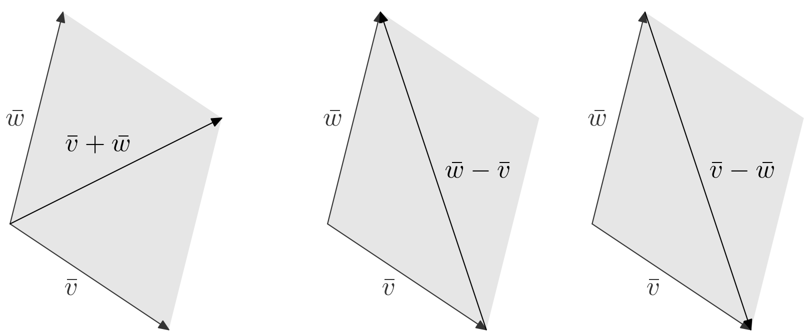
Kuva 1.4: Skalaarimonikerrat $2\bar{v}$ ja $-\frac{1}{2}\bar{v}$.

Määritelmä 1.5. Vektorille $(-1)\bar{v}$ käytetään merkintää $-\bar{v}$. Summalle $\bar{v} + (-\bar{w})$ puolestaan käytetään merkintää $\bar{v} - \bar{w}$. Tätä kutsutaan vektorien $\bar{v}$ ja $\bar{w}$ *erotukseksi*.

Esimerkiksi vektorien $(-1, 6)$ ja $(4, 2)$ erotus on

$$(-1, 6) - (4, 2) = (-1, 6) + (-1)(4, 2) = (-1, 6) + (-4, -2) = (-5, 4).$$

Vektorien erotuksen voi määrittää kuvan perusteella samaan tapaan kuin summan. Nyt vain jälkimmäisen vektorin suunta on käännettävä. Vektorien summaa ja erotusta on havainnollistettu kuvassa 1.5.



Kuva 1.5: Summa $\bar{v} + \bar{w}$ sekä erotukset $\bar{v} - \bar{w}$ ja $\bar{w} - \bar{v}$.

Kaikki edellä esitellyt käsitteet voidaan määritellä myös kolmiulotteisessa avaruudessa. Avaruus $\mathbb{R}^3$ on joukko $\{(a, b, c) \mid a, b, c \in \mathbb{R}\}$. Sen alkioita voidaan ajatella avaruuskoordinaatiston pisteinä. Yhteenlasku ja skalaarikertolasku määritellään komponenteittain samalla tavalla kuin avaruudessa $\mathbb{R}^2$.

Toisinaan merkitään $\bar{i} = (1, 0)$ ja $\bar{j} = (0, 1)$ tai $\bar{i} = (1, 0, 0)$, $\bar{j} = (0, 1, 0)$ ja $\bar{k} = (0, 0, 1)$. Kyseisiä merkintöjä ei juurikaan käytetä tällä kurssilla.

2 Avaruus $\mathbb{R}^n$

Edellisessä luvussa käsiteltiin avaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ vektoreita eli reaalilukupareja ja reaalilukukolmikoita. Näitä avaruuksia voidaan yleistää määrittelemällä avaruus $\mathbb{R}^n$.

Oletetaan, että $n \in \{1, 2, 3, \dots\}$.

Määritelmä 2.1. Avaruuden $\mathbb{R}^n$ alkiot ovat reaaliluvuista koostuvia n-jonoja. Toisin sanoen

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

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

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 $\bar{v}$ komponetteja merkitään symboleilla $v_1, v_2, \dots, v_n$.

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

$$\bar{v} + \bar{w} = (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n) \quad \text{ja}$$

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

Ensimmäistä laskutoimitusta nimitetään vektorien *yhteenlaskuksi* ja toista *skalaarikertolaskuksi*. Jos $\bar{v} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$, vektoria $c\bar{v}$ nimitetään vektorin $\bar{v}$ *skalaarimonikerraksi*. Vektorien yhteydessä reaalilukuja kutsutaan usein *skalaareiksi*.

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 merkitää $\bar{0}$.

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 avaruuden $\mathbb{R}^5$ 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 avaruuden $\mathbb{R}^n$ vektoreille pätevät tutut laskusäännöt.

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

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. $c(\bar{v} + \bar{w}) = c\bar{v} + c\bar{w}$ *(osittelulaki)*

6. $(a+c)\bar{v} = a\bar{v} + c\bar{v}$ *(osittelulaki)*

7. $a(c\bar{v}) = (ac)\bar{v}$

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. Kirjoitetaan $\bar{v} = (v_1, \ldots, v_n)$ ja $\bar{w} = (w_1, \ldots, w_n)$, missä $v_1, \ldots, v_n \in \mathbb{R}$ ja $w_1, \ldots, w_n \in \mathbb{R}$. Nyt nähdään, että

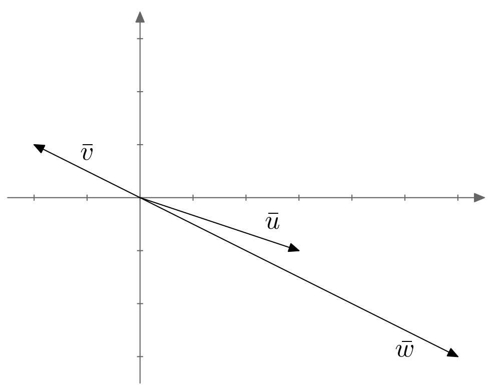
$$\begin{aligned}\bar{v} + \bar{w} &= (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n) \\ &= (w_1 + v_1, w_2 + v_2, \ldots, w_n + v_n) = \bar{w} + \bar{v}.\end{aligned}$$

Tässä käytettiin hyväksi reaalilukujen yhteenlaskun vaihdannaisuutta. Siten väite on todistettu. □

Skalaarikertolaskun avulla voidaan määritellä vektorien yhdensuuntaisuus.

Määritelmä 2.6. Vektorit $\bar{v} \in \mathbb{R}^n$ ja $\bar{w} \in \mathbb{R}^n$ 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. Vektorit $\bar{v} = (-2, 1)$ ja $\bar{w} = (6, -3)$ ovat yhdensuuntaiset, sillä $\bar{v} = -(1/3)\bar{w}$. Vektorit $\bar{v}$ ja $\bar{u} = (3, -1)$ eivät puolestaan ole yhdensuuntaiset. Jos nimittäin olisi olemassa $r \in \mathbb{R}$, jolle pätisi $\bar{v} = r\bar{u}$, niin $-2 = 3r$ ja $1 = -r$. Ensimmäisen yhtälön mukaan $r = -2/3$, mutta toisen yhtälön mukaan $r = -1$. Tämä on mahdotonta, joten ei ole olemassa sellaista lukua r, jolle pätee $\bar{v} = r\bar{u}$. Siten vektorit $\bar{v}$ ja $\bar{u}$ eivät ole yhdensuutaiset.



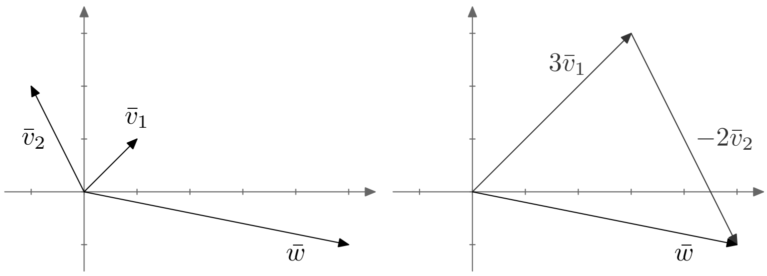
Kuva 2.6: Vektorit $\bar{v}$, $\bar{w}$ ja $\bar{u}$.

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

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

Esimerkki 2.9. Merkitään $\bar{v}_1 = (1,1)$, $\bar{v}_2 = (-1,2)$ ja $\bar{w} = (5,-1)$. Vektori $\bar{w}$ on vektoreiden $\bar{v}_1$ ja $\bar{v}_2$ lineaarikombinaatio, sillä

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



Kuva 2.7: Vektori $\bar{w}$ on vektoreiden $\bar{v}_1$ ja $\bar{v}_2$ lineaarikombinaatio.

3 Suorat ja tasot

Tässä luvussa käsitellään avaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ suoria ja tasoja vektoreiden näkökulmasta.

3.1 Suora

Määritelmä 3.1. Oletetaan, että $n = 2$ tai $n = 3$. Avaruuden $\mathbb{R}^n$ *suora* on joukko

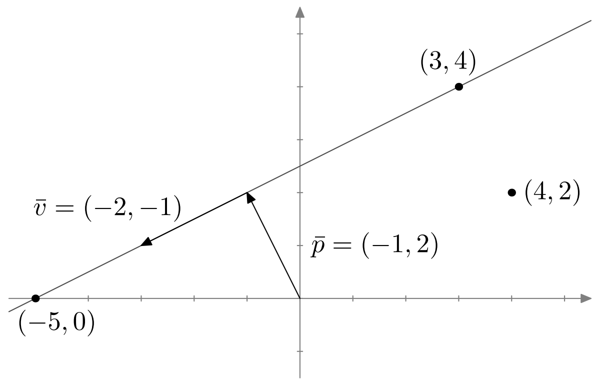
$$\{\bar{p} + k\bar{v} \mid k \in \mathbb{R}\},$$

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.*

Olkoon S avaruuden $\mathbb{R}^2$ suora. Jos $(a, b) \in S$, niin sanotaan, että piste (a, b) on suoralla S tai että suora S kulkee pisteen (a, b) kautta. Vastaavia ilmauksia käytetään avaruudessa $\mathbb{R}^3$.

Esimerkki 3.2. Esimerkiksi joukko $S = \{(-1, 2) + k(-2, -1) \mid k \in \mathbb{R}\}$ on suora. Se on piirretty kuvaan 3.8.

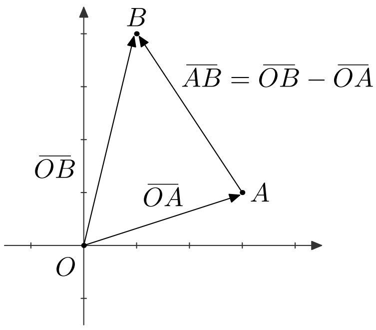


Kuva 3.8: Suora S avaruudessa $\mathbb{R}^2$.

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 kulkeen pisteiden $(-5, 0)$ ja $(3, 4)$ kautta.

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

Ryhdytään seuraavaksi määrittämään suoraa, joka kulkee annettujen pisteiden kautta. Sitä ennen otetaan käyttöön muutama merkintä. Oletetaan, että A ja B ovat avaruuden $\mathbb{R}^2$ tai $\mathbb{R}^3$ pisteitä. Vektori $\overline{AB}$ on vektori, jota vastaavan suuntajanan alkupiste on A ja päätepiste on B. Origoa on tapana merkitä symbolilla O. Siten pisteen A paikkavektorille saadaan merkintä $\overline{OA}$.

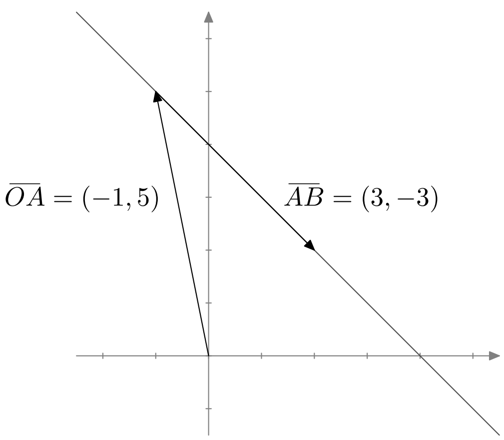


Kuva 3.9: Vektorit $\overline{OA}$, $\overline{OB}$ ja $\overline{AB}$.

Esimerkki 3.3. Tutkitaan, millainen on pisteiden $A = (-1, 5)$ ja $B = (2, 2)$ kautta kulkeva suora. Tätä varten tarvitaan suoralle paikkavektori. Paikkavektoriksi käy minkä tahansa suoran pisteen paikkavektori, esimerkiksi vektori $\overline{OA} = (-1, 5)$. Suuntavektoriksi käy mikä tahansa suoran suuntainen vektori, esimerkiksi vektori

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

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



Kuva 3.10: Suora $\{(-1, 5) + t(3, -3) \mid t \in \mathbb{R}\}$.

Varmistutaan vielä siitä, että annetut pisteet A ja B todellakin ovat suoralla. Huomataan, että $(-1,5) = (-1,5) + 0 \cdot (3,-3)$ ja $(2,2) = (-1,5) + (3,-3)$. Siten pisteet A ja B ovat suoralla.

Vastaavalla menetelmällä on aina mahdollista määrittää kahden pisteen kautta kulkeva suora, vaikka asiaa ei sen tarkemmin tässä osoitetakaan.

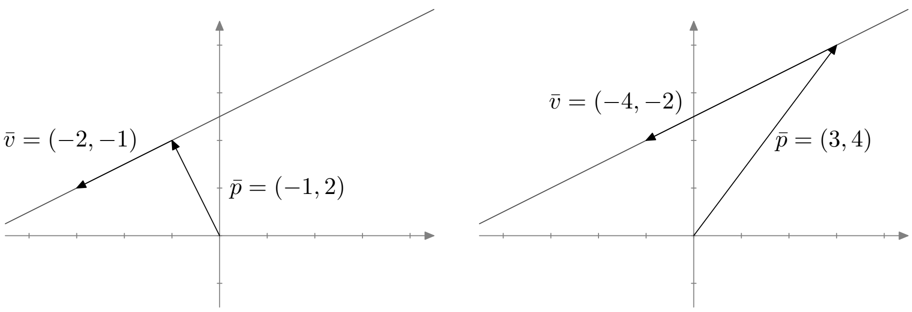
Suoran paikka- ja suuntavektorit eivät ole yksikäsitteisiä, sillä sama suora voidaan kirjoittaa joukkona $\{\bar{p} + t\bar{v} \mid t \in \mathbb{R}\}$ usealla eri tavalla. Voidaan osoittaa, että

- vektoriksi $\bar{p}$ voidaan valita suoran minkä tahansa pisteen paikkavektori.
- vektoriksi $\bar{v}$ voidaan valita mikä tahansa suoran suuntainen vektori.

Esimerkki 3.4. Esimerkin 3.2 suoralle $S = \{(-1,2) + k(-2,-1) \mid k \in \mathbb{R}\}$ on mahdollista valita paikkavektoriksi piste $(3,4)$ ja suuntavektoriksi vektori $(-4,-2)$. Tällöin S voidaan kirjoittaa muodossa

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

Vaikka tämä joukko onkin äkkiseltään katsottuna erilainen kuin suoran S alkuperäinen määritelmä, on joukoissa täsmälleen samat alkiot. Asiaa on havainnollistettu kuvassa 3.11.



Kuva 3.11: Suoran paikkavektori ja suuntavektori eivät ole yksikäsitteisiä.

Osoitetaan vielä huolellisesti, että joukot $S = \{(-1,2) + k(-2,-1) \mid k \in \mathbb{R}\}$ ja $S' = \{(3,4) + t(-4,-2) \mid t \in \mathbb{R}\}$ 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 $\bar{a} \in S$. Nyt $\bar{a} = (-1,2) + k(-2,-1)$ jollakin $k \in \mathbb{R}$. Huomataan, että

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

missä $(k+2)/2 \in \mathbb{R}$. Siten $\bar{a} \in S'$. Näin on 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 $\bar{a} = (3,4) + t(-4,-2)$ jollakin $t \in \mathbb{R}$. Huomataan, että

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

missä $2(t-1) \in \mathbb{R}$. Siten $\bar{a} \in S$. Näin on osoitettu, että $S' \subset S$.

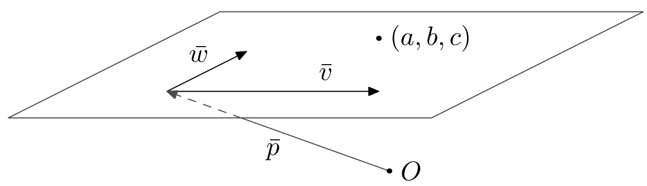
3.2 Taso

Määritelmä 3.5. Avaruuden $\mathbb{R}^3$ *taso* on joukko

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

missä $\bar{p} \in \mathbb{R}^3$, $\bar{v}, \bar{w} \in \mathbb{R}^3 \setminus \{\bar{0}\}$ 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*. Olkoon T avaruuden $\mathbb{R}^3$ taso. Jos $(a,b,c) \in T$, niin sanotaan, että piste (a,b,c) on tasossa T tai että taso T kulkee pisteen (a,b,c) kautta.



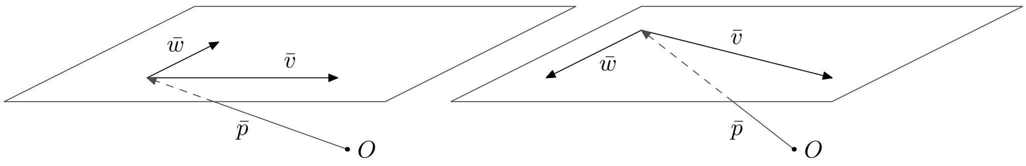
Kuva 3.12: Piste (a,b,c) on tasossa T.

Voidaan osoittaa, että sama taso on mahdollista kirjoittaa usealla eri tavalla joukkona $\{\bar{p} + s\bar{w} + t\bar{v} \mid s,t \in \mathbb{R}\}$:

- Vektoriksi $\bar{p}$ voidaan valita tason minkä tahansa pisteen paikkavektori.
- Vektoreiksi $\bar{w}$ ja $\bar{v}$ voidaan valita mitkä tahansa tason suuntaisen vektorit, kunhan $\bar{w}$ ja $\bar{v}$ eivät ole yhdensuuntaiset.

Esimerkki 3.6. 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:

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



Kuva 3.13: Taso voidaan kirjoittaa eri tavoin joukkona $\{\bar{p} + s\bar{w} + t\bar{v} \mid s, t \in \mathbb{R}\}$.

Huomaa, että vektorit eivät ole yhdensuuntaiset, sillä $\overline{AB} \neq t\overline{AC}$ kaikilla $t \in \mathbb{R} \setminus \{0\}$. Näin saadaan taso

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

4 Avaruuden $\mathbb{R}^n$ aliavaruudet

Edellisessä luvussa käsittelimme avaruuksien $\mathbb{R}^2$ ja $\mathbb{R}^3$ suoria ja tasoja. Osoittautuu, että erityisesti origon kautta kulkevat suorat ja tasot ovat mielenkiintoisia. Tässä luvussa yleistämme tällaiset suorat ja tasot avaruuteen $\mathbb{R}^n$ ja tutkimme niin kutsuttuja vektoreiden virittämiä aliavaruuksia.

Määritelmä 4.1. Vektoreiden $\bar{v}_1, \dots, \bar{v}_k \in \mathbb{R}^n$ *virittämä aliavaruus* on joukko

$$\mathrm{span}(\bar{v}_1, \dots, \bar{v}_k) = \{a_1\bar{v}_1 + a_2\bar{v}_2 + \dots + a_k\bar{v}_k \mid a_1, a_2, \dots, a_k \in \mathbb{R}\}.$$

Vektoreiden virittämä aliavaruus koostuu siis kaikista vektoreiden lineaarikombinaatioista. Englannin kielen verbi ”span” tarkoittaa virittämistä tai ulottumista.

Esimerkki 4.2. Esimerkiksi avaruuden $\mathbb{R}^2$ suora

$$S = \{\bar{0} + t(-3, 1) \mid t \in \mathbb{R}\} = \{t(-3, 1) \mid t \in \mathbb{R}\}$$

on vektorin $(-3, 1)$ virittämä aliavaruus. Toisin sanoen $S = \mathrm{span}\big((-3, 1)\big)$. Avaruuden $\mathbb{R}^3$ taso

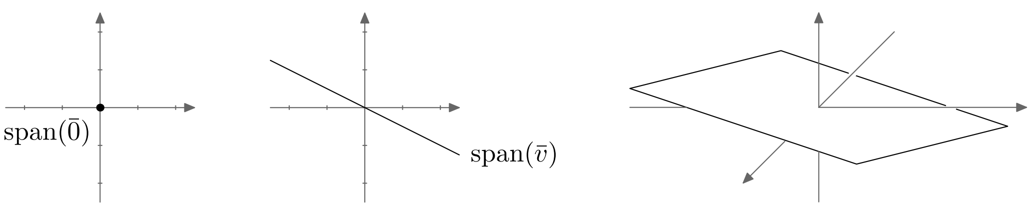
$$T = \{\bar{0} + t(-3, 1) + s(2, 2) \mid t, s \in \mathbb{R}\} = \{t(-3, 1) + s(2, 2) \mid t, s \in \mathbb{R}\}$$

taas on vektorien $(-3, 1)$ ja $(2, 2,)$ virittämä aliavaruus, eli $T = \mathrm{span}((-3, 1), (2, 2))$.

Tutkitaan millaisia vektorien virittämät aliavaruudet voivat olla avaruuksissa $\mathbb{R}^2$ ja $\mathbb{R}^3$. Nollavektorin virittämä aliavaruus on $\mathrm{span}(\bar{0}) = \{a \cdot \bar{0} \mid a \in \mathbb{R}\} = \{\bar{0}\}$. Aliavaruudessa on siis ainoastaan nollavektori.

Oletetaan sitten, että $\bar{v} \in \mathbb{R}^2 \setminus \{\bar{0}\}$ tai $\bar{v} \in \mathbb{R}^3 \setminus \{\bar{0}\}$. Tällöin vektorin $\bar{v}$ virittämä aliavaruus $\mathrm{span}(\bar{v}) = \{a\bar{v} \mid a \in \mathbb{R}\}$ on vektorin $\bar{v}$ suuntainen suora. Tämä suora kulkee origon kautta.

Jos taas $\bar{v}$, $\bar{w} \in \mathbb{R}^3 \setminus \{\bar{0}\}$ ja $\bar{v}$ ja $\bar{w}$ eivät ole yhdensuuntaisia, vektorien $\bar{v}$ ja $\bar{w}$ virittämä aliavaruus $\mathrm{span}(\bar{v}, \bar{w}) = \{s\bar{v} + t\bar{w} \mid s, t \in \mathbb{R}\}$ on taso, joka kulkee origon kautta.



Kuva 4.14: Nollavektorin virittämä aliavaruus on $\{\bar{0}\}$. Vektorin $\bar{v} \neq \bar{0}$ virittämä aliavaruus on origon kautta kulkeva suora. Kahden vektorin virittämä aliavaruus voi olla origon kautta kulkeva taso.

Vektoreiden virittämän aliavaruus yleistää siis origon kautta kulkevan suoran ja tason käsitteitä. Seuraava esimerkki osoittaa, miksi juuri origon kautta kulkevat suorat ja tasot ovat erityisen kiinnostavia.

Esimerkki 4.3. Tarkastellaan origon kautta kulkevaa suoraa

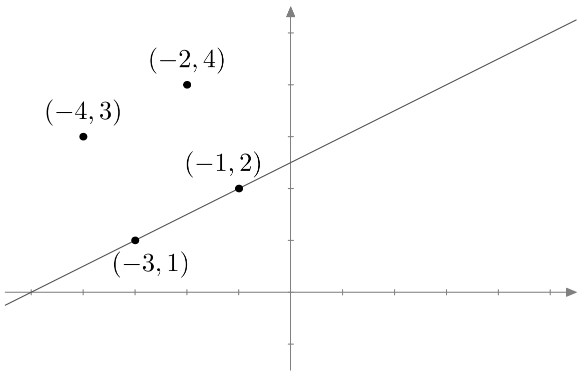
$$S = \operatorname{span}(\bar{v}) = \{k\bar{v} \mid k \in \mathbb{R}\},$$

missä $\bar{v} \in \mathbb{R}^2 \setminus \{\bar{0}\}$. Jos $\bar{w}$, $\bar{u} \in S$, niin $\bar{w} = a\bar{v}$ ja $\bar{u} = b\bar{v}$ joillakin a, $b \in \mathbb{R}$. Nyt $\bar{w} + \bar{u} = (a + b)\bar{v}$, joten summa $\bar{w} + \bar{u}$ on suoran S alkio. Lisäksi jos $c \in \mathbb{R}$, niin $c\bar{w} = c(a\bar{v}) = (ca)\bar{v}$. Siten kaikkien suoran S alkioiden skalaarimonikerrat ovat edelleen suoran S alkioita. Tavallaan suora S on oma pieni avaruutensa avaruuden $\mathbb{R}^2$ sisässä, ja siellä voidaan laskea vektoreita yhteen ja kertoa niitä reaaliluvuilla. Sama pätee origon kautta kulkeviin tasoihin.

Tilanne on aivan toinen, jos suora tai taso ei kulje origon kautta. Tutkitaan vaikkapa esimerkin 3.2 suoraa

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

Nyt esimerkiksi $(-1, 2)$ ja $(-3, 1)$ ovat suoralla S. Summa $(-1, 2) + (-3, 1) = (-4, 3)$ ei kuitenkaan ole suoralla S (ks. kuva 4.15). Myöskään skalaarimonikerta $2 \cdot (-1, 2) = (-2, 4)$ ei ole suoralla S.



Kuva 4.15: Esimerkin 3.2 suora S.

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 vektorien virittämään aliavaruuteen.

Lause 4.4. *Oletetaan, että* $\bar{v}_1$, $\bar{v}_2$, $\ldots$, $\bar{v}_k \in \mathbb{R}^n$. *Merkitään* $W = \operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$. *Tällöin* *a) jos $\bar{u}$, $\bar{w} \in W$, niin $\bar{u} + \bar{w} \in W$.*

b) jos $\bar{w} \in W$ ja $a \in \mathbb{R}$, niin $a\bar{w} \in W$.

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 $a_1, \ldots, a_k \in \mathbb{R}$ ja $\bar{w} = b_1\bar{v}_1 + \cdots + b_k\bar{v}_k$ joillakin $b_1, \ldots, b_k \in \mathbb{R}$. Huomataan, että

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

Siten $\bar{u} + \bar{w} \in W$. □

Esimerkki 4.5. Tutkitaan, kuuluuko vektori $\bar{w} = (-2, 3, 2, -1)$ vektoreiden

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

virittämään aliavaruuteen $\mathrm{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3)$. On siis selvitettävä, onko olemassa reaalilukuja x_1, x_2, x_3, joille pätee

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

eli

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

Tämä yhtälö voidaan vielä muuttaa muotoon

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

ja edelleen yhtälöksi

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

Toisin sanoen on ratkaistava yhtälöryhmä

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

Miten tällainen yhtälöryhmä ratkaistaan? Ennen kuin syvennymme vektoreiden virittämiin aliavaruuksiin lisää, on syytä perehtyä yhtälöryhmien ratkaisemiseen.

5 Lineaariset yhtälöryhmät

Lineaarinen yhtälöryhmä on yhtälöryhmä, joka on muotoa

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

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

Esimerkiksi

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

on lineaarinen yhtälöryhmä.

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

Lineaarisessa yhtälöryhmässä oleellisia ovat vain kertoimet ja vakiot. Kaikki tieto yhtälöryhmästä voidaankin tiivistää lukutaulukkoon eli *matriisiin*, jossa luetellaan kaikki kertoimet sekä vakiot. Kun käsitellään yhtälöryhmien sijasta matriiseja, on kirjoittevaa paljon vähemmän, 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. Huomaa, että matriisiin on kirjoitettava nolla niiden termien kohdalle, jotka puuttuvat yhtälöryhmästä. Kyseisten termien kertoimena on nimittäin nolla.

Kappaleessa 8 tutustutaan matriisien teoriaan yleisemmin. Tässä luvussa käsittelemme vain yhtälöryhmistä saatuja matriiseja.

5.1 Gaussin–Jordanin eliminointimenetelmä

Seuraavaksi käydään läpi menetelmä, jolla voidaan ratkaista mikä tahansa lineaarinen yhtälöryhmä. Ideana on muokata yhtälöryhmästä uusia yhtälöryhmiä, joilla on samat ratkaisut kuin alkuperäisellä yhtälöryhmällä. Viimeisenä saatu yhtälöryhmä on helppo ratkaista. Koska viimeisen yhtälöryhmän ratkaisut ovat samat kuin alkuperäisen yhtälöryhmän, on alkuperäisen yhtälöryhmän ratkaisut löydetty.

Määritelmä 5.1. 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 matriiseille.

Määritelmä 5.2. 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 ($i \neq j$).

Esimerkki 5.3. Seuraavassa on annettu esimerkit erilaisista alkeisrivitoimituksista:

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

Määritelmä 5.4. Matriisi A on *riviekvivalentti* matriisin B kanssa, jos B saadaan matriisista A alkeisrivitoimituksilla.

Esimerkiksi edellisen esimerkin matriisit

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

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

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

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

Kuva 5.16: Gaussin-Jordanin eliminointimenetelmän perusta.

Lause 5.5. *Jos yhtälöryhmiä vastaavat matriisit ovat riviekvivalentit, yhtälöryhmät ovat ekvivalentit.*

Lause voidaan muotoilla myös toisin: jos yhtälöryhmiä vastaavat matriisit ovat riviekvivalenit, 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 nähdä. Tutustutaan aluksi porrasmatriiseihin ja siirrytään sitten tutkimaan redusoituja porrasmatriiseja.

Määritelmä 5.6. Matriisi on *porrasmatriisi*, jos

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

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

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

Määritelmä 5.7. Matriisi on *redusoitu porrasmatriisi*, jos

1. se 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:

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

Esimerkki 5.8. 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

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

Huomataan, että yhtälöryhmästä näkyy suoraan sen ratkaisu, joka on

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

Ratkaisut on helppo nähdä juuri sen vuoksi, että yhtälöryhmän matriisi on redusoitu porrasmatriisi.

Tavoitteena on siis muuttaa yhtälöryhmän matriisi redusoiduksi porrasmatriisiksi, josta ratkaisut näkyvät suoraan. Voidaan osoittaa, että mikä tahansa matriisi voidaan muuttaa alkeisrivitoimituksilla redusoiduksi porrasmatriisiksi. Seuraava esimerkki näyttää, kuinka tämä tehdään.

Esimerkki 5.9. Tutkitaan matriisia

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

ja muutetaan se 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 muokkaaman 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 kaikki johtavat alkiot ykkösiksi:

$$\xrightarrow{\frac{1}{4}R_3}\begin{bmatrix}1 & 0 & 2 & 1\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1\end{bmatrix}.$$

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.

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!

Gaussin-Jordanin menetelmä

Nyt olemme valmiita ottamaan käyttöön niin kutsutun Gaussin-Jordanin menetelmän, jonka avulla lineaariset yhtälöryhmät voidaan aina ratkaista.

1. Kirjoita yhtälöryhmän matriisi.
2. Muuta matriisi alkeisrivitoimituksilla porrasmatriisiksi.
3. Muuta porrasmatriisi redusoiduksi porrasmatriisiksi.
4. Lue ratkaisut redusoidusta porrasmatriisista.

Esimerkki 5.10. Ratkaistaan yhtälöryhmä

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

Sen 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 redusoituksi porrasmatriisiksi esimerkissä **??**:

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

Redusoitua porrasmatriisia vastaava yhtälöryhmä on

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

Koska alkuperäisen yhtälöryhmän ratkasut ovat lauseen 5.5 nojalla samat kuin lopuksi saadun yhtälöryhmän, on yhtälöryhmä ratkaistu. Sen ratkaisu on

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

Esimerkki 5.11. Ratkaistaan lineaarinen yhtälöryhmä

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

Muutetaan sen 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 ratkaisua.

Esimerkki 5.12. Ratkaistaan lineaarinen yhtälöryhmä

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

Muutetaan sen 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

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

Alin yhtälö $0 = 0$ on aina tosi. 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. Ylin yhtälö saadaan muotoon $x_1 - 2t = 1$, joten $x_1 = 1 + 2t$. Toinen yhtälö puolestaan on $x_2 - t = 2$ eli $x_2 = 2 + t$. Siten yhtälön ratkaisu on

$$\begin{cases} x_1 = 1 + 2t \\ x_2 = 2 + t \\ x_3 = t, \end{cases}$$

missä $t \in \mathbb{R}$. Ratkaisuja on siis äärettömän monta. Esimerkiksi $x_1 = 3$, $x_2 = 3$ ja $x_3 = 1$ sekä $x_1 = -1$, $x_2 = 1$ ja $x_3 = -1$ ovat yhtälöryhmän ratkaisuja. Jokaisella reaaliluvulla t yhtälöryhmälle saadaan eri ratkaisu.

Huomaa, että vapaat muuttujat näkyvät redusoidussa porrasmatriisissa sarakkeina, joissa ei ole lainkaan johtavaa alkiota.

Esimerkki 5.13. 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?

Havaitaan, että 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

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

ja edelleen

$$\begin{cases} x_1 = 7 - 3r - 4s \\ x_3 = -2s \\ x_6 = 3. \end{cases}$$

Yhtälöryhmän ratkaisu on siis

$$\begin{cases} x_1 = 7 - 3r - 4s \\ x_2 = r \\ x_3 = -2s \\ x_4 = s \\ x_5 = t \\ x_6 = 3, \end{cases}$$

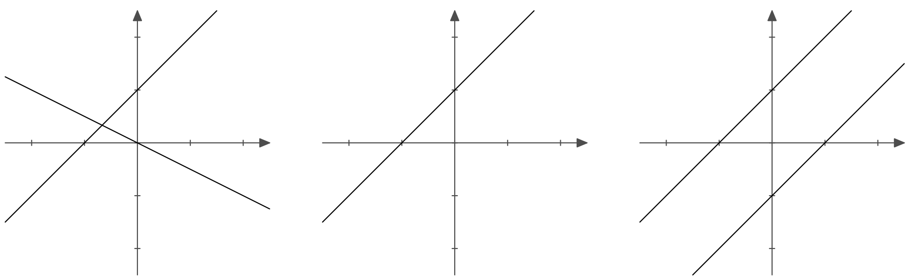
missä $r, s, t \in \mathbb{R}$.

Esimerkeistä huomataan, että yhtälöryhmällä voi olla täsmälleen yksi ratkaisu, äärettömän monta ratkaisua tai ei yhtään ratkaisua. Kun muuttujia on kaksi, tilannetta voi havainnollistaa kuvan 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 ensimmäisen yhtälön, piste (r, s) on suoralla, jonka yhtälö on $ax + by = c$. Samoin 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, on 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 5.17: Yhtälöryhmällä on tasan yksi ratkaisu, äärettömän monta ratkaisua tai ei yhtään ratkaisua.

Esimerkki 5.14. 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

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

  Havaitaan, että alinta riviä vastaava yhtälö $0=0$ on aina tosi. Lisäksi 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

$$\xrightarrow{\frac{1}{k-1}R_2} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 2-k-k^2 & -2-k \end{array}\right] \xrightarrow{\frac{1}{2-k-k^2}R_3} \left[\begin{array}{ccc|c} 1 & 1 & k & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & (-2-k)/(2-k-k^2) \end{array}\right].$$

Tällöin 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 $k\neq -2$.

Käydään vielä lopuksi läpi lauseen 5.5 todistus, joka sivuutettiin luvun alussa.

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

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

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 $x_1 = r_1, \dots, x_n = r_n$ 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

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

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

Oletetaan sitten, että $x_1 = s_1, \dots, x_n = s_n$ 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 $x_1 = s_1, \dots, x_n = s_n$ toteuttaa myös yhtälöryhmän (1), joten se on myös yhtälöryhmän (1) ratkaisu. Siten yhtälöryhmillä on samat ratkaisut.

3. Kolmannen alkeisrivitoimituksen tarkastelu jätetään lukijalle.

□

6 Virittäminen

Palataan sitten avaruuden $\mathbb{R}^n$ aliavaruuksiin. Palautetaan mieleen, että vektoreiden $\bar{v}_1, \dots, \bar{v}_k \in \mathbb{R}^n$ virittämä aliavaruus on joukko

$$\operatorname{span}(\bar{v}_1, \dots, \bar{v}_k) = \{a_1\bar{v}_1 + a_2\bar{v}_2 + \dots + a_k\bar{v}_k \mid a_1, a_2, \dots, a_k \in \mathbb{R}\}.$$

Nyt osaamme vastata esimerkissä 4.5 esitettyyn kysymykseen. Esimerkissä haluttiin tietää, kuuluuko vektori $\bar{w} = (-2, 3, 2, -1)$ vektoreiden

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

virittämään aliavaruuteen $\operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3)$. 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 yhtälöryhmä ratkaistaan Gaussin-Jordanin eliminointimenetelmällä, nähdään, että ratkaisuja ei ole. Siten $\bar{w} \notin \operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3)$.

Esimerkki 6.1. Merkitään $\bar{e}_1 = (1, 0)$ ja $\bar{e}_2 = (0, 1)$. Osoitetaan, että

$$\operatorname{span}(\bar{e}_1, \bar{e}_2) = \mathbb{R}^2.$$

Kaksi joukkoa osoitetaan samoiksi näyttämällä, että kumpikin on toisen osajoukko. Tiedetään, että jokainen joukon $\operatorname{span}(\bar{e}_1, \bar{e}_2)$ vektori on avaruuden $\mathbb{R}^2$ vektori, joten on selvää, että $\operatorname{span}(\bar{e}_1, \bar{e}_2) \subset \mathbb{R}^2$. Näin ollen riittää näyttää, että $\mathbb{R}^2 \subset \operatorname{span}(\bar{e}_1, \bar{e}_2)$. On siis osoitettava, että jokainen avaruuden $\mathbb{R}^2$ vektori voidaan esittää vektoreiden $\bar{e}_1$ ja $\bar{e}_2$ lineaarikombinaationa.

Oletetaan, että $\bar{v} = (v_1, v_2) \in \mathbb{R}^2$. Nyt

$$\bar{v} = (v_1, v_2) = v_1(1, 0) + v_2(0, 1) = v_1\bar{e}_1 + v_2\bar{e}_2,$$

joten $\bar{v} \in \operatorname{span}(\bar{e}_1, \bar{e}_2)$. Siten $\mathbb{R}^2 \subset \operatorname{span}(\bar{e}_1, \bar{e}_2)$. On siis osoitettu, että $\operatorname{span}(\bar{e}_1, \bar{e}_2) = \mathbb{R}^2$.

Esimerkki 6.2. Tutkitaan, millä ehdolla vektori $\bar{w} = (w_1, w_2, w_3)$ kuuluu vektoreiden

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

virittämään aliavaruuteen $\operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3)$.

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

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

Tästä saadaan 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

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

Matriisista 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 aliavaruudessa $\operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3)$, jos ja vain jos $w_1 - 2w_2 - w_3 = 0$.

Nyt aliavaruus $\operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3)$ voidaan kirjoittaa uudessa muodossa:

$$\begin{aligned} \operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3) &= \{(w_1, w_2, w_3) \in \mathbb{R}^3 \mid w_1 - 2w_2 - w_3 = 0\} \\ &= \{(w_1, w_2, w_3) \in \mathbb{R}^3 \mid w_3 = w_1 - 2w_2\} \\ &= \{(w_1, w_2, w_1 - 2w_2) \mid w_1, w_2 \in \mathbb{R}^2\} \\ &= \{(w_1, 0, w_1) + (0, w_2, -2w_2) \mid w_1, w_2 \in \mathbb{R}^2\} \\ &= \{w_1(1, 0, 1) + w_2(0, 1, -2) \mid w_1, w_2 \in \mathbb{R}^2\}. \end{aligned}$$

Kyseessä on siis vektorien $(1, 0, 1)$ ja $(0, 1, -2)$ virittämä aliavaruus. Toisin sanottuna

$$\operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3) = \operatorname{span}((1, 0, 1), (0, 1, -2)).$$

Eri vektorit voivat siis virittää saman aliavaruuden. Edes virittäjävektorien määrän ei tarvitse olla sama.

Esimerkki 6.3. 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

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

Matriisista nähdään, että yhtälöryhmällä on ratkaisuja olivat w_1, w_2 ja w_3 mitä lukuja tahansa. Siten $\bar{w} \in \operatorname{span}(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$. Näin ollen $\mathbb{R}^3 = \operatorname{span}(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$.

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

$$\begin{cases} x_1 = t \\ x_2 = 1 + t \\ x_3 = 2 - 2t \\ x_4 = t\,, \end{cases}$$

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

7 Vapaus

Määritelmä 7.1. 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 seuraava ehto pätee:

$$\text{jos} \quad c_1\bar{v}_1 + c_2\bar{v}_2 + \cdots + c_k\bar{v}_k = \bar{0} \quad \text{joillakin } c_1, \ldots, c_k \in \mathbb{R},$$

niin $c_1 = 0, c_2 = 0, \ldots, c_k = 0$.

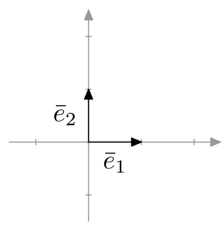
Jos jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa eli lineaarisesti riippumaton, voidaan myös sanoa, että vektorit $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ ovat lineaarisesti riippumattomia. Jos jono ei ole vapaa, sanotaan, että se on *sidottu*.

Tulemme näkemään, että jos jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa, voidaan aliavaruuden $\text{span}(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ alkiot kirjoittaa täsmälleen yhdellä tavalla virittäjävektorien lineaarikombinaatioina. Virittäjien joukossa ei siis ole tarpeettomia vektoreita.

Esimerkki 7.2. Merkitään $\bar{e}_1 = (1, 0)$ ja $\bar{e}_2 = (0, 1)$. Osoitetaan, että avaruuden $\mathbb{R}^2$ jono $(\bar{e}_1, \bar{e}_2)$ on lineaarisesti riippumaton. Oletetaan, että reaaliluvut c_1 ja c_2 ovat sellaisia, että

$$c_1\bar{e}_1 + c_2\bar{e}_2 = \bar{0}.$$

Nyt $c_1(1, 0) + c_2(0, 1) = (0, 0)$, joten $(c_1, c_2) = (0, 0)$. Siten pätee $c_1 = 0$ ja $c_2 = 0$. Näin on osoitettu, että jono $(\bar{e}_1, \bar{e}_2)$ on vapaa.



Kuva 7.18: Jono $(\bar{e}_1, \bar{e}_2)$ on vapaa.

Esimerkki 7.3. Merkitään $\bar{v}_1 = (1, 2)$ ja $\bar{v}_2 = (-3, -1)$. Tutkitaan, onko jono $(\bar{v}_1, \bar{v}_2)$ vapaa vai sidottu.

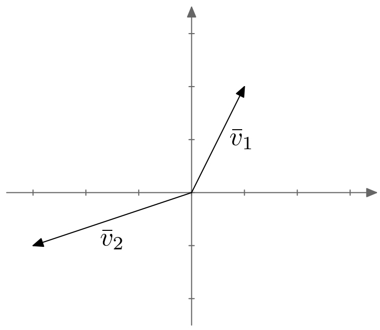
Oletetaan, että $c_1\bar{v}_1 + c_2\bar{v}_2 = \bar{0}$ joillakin c_1, $c_2 \in \mathbb{R}$. Nyt $c_1(1, 2) + c_2(-3, -1) = (0, 0)$ eli

$$\begin{cases} c_1 - 3c_2 = 0 \\ 2c_1 - c_2 = 0. \end{cases}$$

Ratkaistaan tästä c_1 ja c_2:

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

Ainoa ratkaisu on $c_1 = 0$ ja $c_2 = 0$. Jono $(\bar{v}_1, \bar{v}_2)$ on siis vapaa.



Kuva 7.19: Jono $(\bar{v}_1, \bar{v}_2)$ on vapaa.

Esimerkki 7.4. Kun osoitetaan jono sidotuksi, ei välttämättä tarvitse ratkaista yhtälöryhmää. Toisinaan on nimittäin helppo nähdä millaisten kertoimien avulla lineaarikombinaatiosta muodostuu nollavektori.

Merkitään $\bar{w}_1 = (2, 1)$ ja $\bar{w}_2 = (-4, -2)$. Huomataan, että

$$2\bar{w}_1 + \bar{w}_2 = \bar{0}.$$

Siten jono $(\bar{w}_1, \bar{w}_2)$ on sidottu.

Esimerkki 7.5. 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. Oletetaan, että $c_1\bar{v}_1 + c_2\bar{v}_2 + c_3\bar{v}_3 = \bar{0}$ joillakin c_1, c_2, $c_3 \in \mathbb{R}$. Tällöin $c_1(1, 2) + c_2(-3, -1) + c_3(-1, 1) = (0, 0)$ eli

$$\begin{cases} c_1 - 3c_2 - c_3 = 0 \\ 2c_1 - c_2 + c_3 = 0. \end{cases}$$

Ratkaistaan tästä c_1 ja c_2:

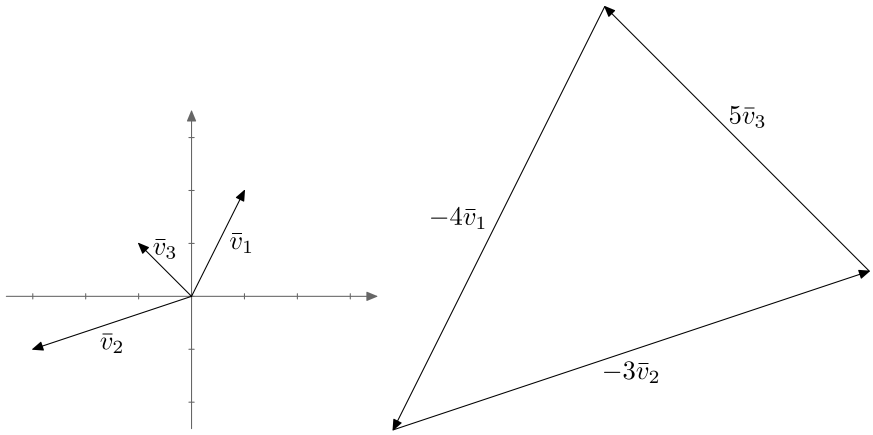
$$\left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 2 & -1 & 1 & 0 \end{array}\right] \xrightarrow{R_2-2R_1} \left[\begin{array}{ccc|c} 1 & -3 & -1 & 0 \\ 0 & 5 & 3 & 0 \end{array}\right]$$

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

Huomataan, että yhtälöryhmällä on äärettömän monta ratkaisua:

$$\begin{cases} x_1 = -(4/5)t \\ x_2 = -(3/5)t \\ x_3 = t, \end{cases}$$

missä $t \in \mathbb{R}$. Siten $c_1 = 0$, $c_2 = 0$, $c_3 = 0$ 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 7.20.



Kuva 7.20: Jono $(\bar{v}_1, \bar{v}_2, \bar{v}_3)$ on sidottu.

Määritelmän mukaan jonon vapaus kertoo siitä, että nollavektori voidaan kirjoittaa vain yhdellä tavalla jonon vektorien lineaarikombinaationa. Yhtälö

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

toteutuu ainakin, jos kertoimiksi $c_1, \dots, c_k$ valitaan nollat. Toisinaan yhtälö kuitenkin toteutuu myös joillakin muilla kertoimilla. Jono on vapaa, jos nollavektori voidaan kirjoittaa jonon vektorien lineaarikombinaationa *täsmälleen yhdellä tavalla.*

Seuraava lause osoittaa, että vektorijono $(\bar{v}_1, \dots, \bar{v}_k)$ on vapaa, jos ja vain jos aliavaruuden $\mathrm{span}(\bar{v}_1, \dots, \bar{v}_k)$ kaikki vektorit voidaan ilmaista täsmälleen yhdellä tavalla virittäjävektorien $\bar{v}_1, \dots, \bar{v}_k$ lineaarikombinaatioina. Siis jos nollavektori voidaan kirjoittaa vain yhdellä tavalla virittäjävektoreiden lineaarikombinaationa, myös kaikki muut aliavaruuden $\mathrm{span}(\bar{v}_1, \dots, \bar{v}_k)$ vektorit voidaan kirjoittaa vain yhdellä tavalla virittäjävektoreiden lineaarikombinaationa. Vapaat vektorijonot ovat kiinnostavia nimen omaan sen vuoksi, että niistä saadaan virittäjäjoukko, jossa ei ole turhia vektoreita.

Lause 7.6. *Oletetaan, että $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k \in \mathbb{R}^n$. Jono $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ on vapaa, jos ja vain jos jokainen aliavaruuden $\mathrm{span}(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ alkio voidaan kirjoittaa täsmälleen yhdellä tavalla vektorien $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k$ 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 tällöin 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 vapaa. Osoitetaan, että jokainen aliavaruuden $W = \text{span}(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ alkio voidaan kirjoittaa täsmälleen yhdellä tavalla virittäjävektorien $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ lineaarikombinaationa. Oletetaan, että alkio $w \in W$ voidaan kirjoittaa lineaarikombinaationa

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

ja lineaarikombinaationa

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

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 sen 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 ovatkin välttämättä samat. Siksi vektoria $\bar{w}$ ei voida kirjoittaa usealla eri tavalla virittäjävektoreiden lineaarikombinaationa.

"$\Leftarrow$": Oletetaan sitten, että jokainen aliavaruuden $W = \text{span}(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ alkio voidaan kirjoittaa täsmälleen yhdellä tavalla virittäjävektorien $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ lineaarikombinaationa. Osoitetaan, että jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa. Sitä varten oletetaan, että luvut $c_1, \ldots, c_k \in \mathbb{R}$ ovat sellaisia, että

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

Tiedetään, että ainakin

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

Koska vektori $\bar{0}$ on aliavaruuden W alkio, se voidaan kirjoittaa virittäjävektorien lineaarikombinaationa täsmälleen yhdellä tavalla. Siten täytyy päteä $c_1 = 0, \ldots, c_k = 0$. Siis jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa. $\square$

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

Lause 7.7. *Oletetaan, että $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k \in \mathbb{R}^n$ ja $k \geq 2$. Jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on sidottu, jos ja vain jos*

$$\bar{v}_j \in \text{span}(\bar{v}_1, \ldots, \bar{v}_{j-1}, \bar{v}_{j+1}, \ldots, \bar{v}_k)$$

jollakin $j \in \{1, 2, \ldots, k\}$.

Todistus. "$\Rightarrow$": Oletetaan, että jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on sidottu. On siis olemassa luvut $c_1, \ldots, c_k \in \mathbb{R}$, 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 $j \in \{1, 2, \ldots, k\}$. 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 $\bar{v}_j \in \operatorname{span}(\bar{v}_1, \ldots, \bar{v}_{j-1}, \bar{v}_{j+1}, \ldots, \bar{v}_k)$.

"$\Leftarrow$": Oletetaan sitten, että

$$\bar{v}_j \in \operatorname{span}(\bar{v}_1, \ldots, \bar{v}_{j-1}, \bar{v}_{j+1}, \ldots, \bar{v}_k)$$

jollakin $j \in \{1, 2, \ldots, k\}$. Nyt on olemassa sellaiset $c_1, \ldots, c_{j-1}, c_{j+1}, \ldots, c_k \in \mathbb{R}$, 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 $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ sidottu. □

Esimerkki 7.8. Merkitään $\bar{v}_1 = (1, -1, 0)$, $\bar{v}_2 = (1, 1, 0)$, $\bar{v}_3 = (0, 0, 2)$ ja $\bar{v}_4 = (3, -1, 0)$. Tällöin esimerkiksi

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

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. Nähdään, että esimerkiksi

$$\bar{v}_2 = -2\bar{v}_1 + 0\bar{v}_3 + \bar{v}_4.$$

Kaikkia vektoreita ei kuitenkaan välttämättä voida kirjoittaa toisten lineaarikombinaationa. Esimerkiksi

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

kaikilla a, b, $c \in \mathbb{R}$. (Tämän todistuksen yksityiskohdat jätetään lukijalle.)

7.1 Homogeeniset yhtälöryhmät ja vapaus

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

Määritelmä 7.9. Lineaarinen yhtälöryhmä, jonka kaikki vakiot ovat nollia, on nimeltään *homogeeninen yhtälöryhmä.*

Homogeeninen yhtälöryhmä on siis muotoa

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

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

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

Tätä kutsutaan yhtälöryhmän triviaaliksi ratkaisuksi.

Lause 7.10. *Jos homogeenisessa yhtälöryhmässä tuntemattomien määrä n on suurempi kuin yhtälöiden määrä m, niin homogeenisella 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 enemmän sarakkeita kuin rivejä. Johtavia alkioita enintään yksi joka rivillä. Koska sarakkeita on enemmän kuin rivejä, on matriisissa ainakin yksi sarake, jossa ei ole johtavaa alkiota. Siten vapaita muuttujia on ainakin yksi, ja yhtälöryhmällä on äärettömän monta ratkaisua. □

Korollaari 7.11. *Oletetaan, että $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m \in \mathbb{R}^n$, missä $n \in \{1, 2, \ldots\}$. Jos $m > n$, niin jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m)$ on sidottu.*

Huom. Korollaari on lause joka seuraa suoraan tai lähes suoraan toisesta lauseesta. Tämä korollaari on lauseen 7.10 seuraus.

Todistus. Merkitään $\bar{v}_k = (v_{1k}, v_{2k}, \ldots, v_{nk})$ kaikilla $k \in \{1, \ldots, m\}$. Nyt yhtälöä $x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_m\bar{v}_m = \bar{0}$ vastaavaksi yhtälöryhmäksi saadaan

$$\begin{cases} v_{11}x_1 + v_{12}x_2 + \cdots + v_{1m}x_m = 0 \\ v_{21}x_1 + v_{22}x_2 + \cdots + v_{2m}x_m = 0 \\ \qquad\qquad\qquad \vdots \qquad\quad = \vdots \\ v_{n1}x_1 + v_{n2}x_2 + \cdots + v_{nm}x_m = 0. \end{cases}$$

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

8 Kanta

Tässä luvussa W on vektoreiden $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m \in \mathbb{R}^n$ virittämä aliavaruus eli

$$W = \operatorname{span}(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m).$$

Määritelmä 8.1. Oletetaan, että $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k \in W$. Vektorijono $(\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k)$ on aliavaruuden W kanta, jos

a) $W = \operatorname{span}(\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k)$ ja

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

Esimerkki 8.2. Esimerkissä 6.1 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 7.2 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*. Lukijan tehtäväksi jätetään osoittaa, että kyseessä on todellakin kanta.

Kanta siis virittää aliavaruuden ja on lisäksi vapaa. Lauseesta 7.6 saadaan seuraava hyvin käyttökelpoinen tulos:

Lause 8.3. *Jono $(\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k)$ on aliavaruuden W kanta, jos ja vain jos jokainen aliavaruuden W vektori voidaan kirjoittaa täsmälleen yhdellä tavalla vektoreiden $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k$ lineaarikombinaationa.*

Todistus. "$\Rightarrow$": Oletetaan, että jono $(\bar{w}_1, \ldots, \bar{w}_k)$ on aliavaruuden W kanta. Tällöin kannan määritelmän nojalla $W = \operatorname{span}(\bar{w}_1, \ldots, \bar{w}_k)$ ja jono $(\bar{w}_1, \ldots, \bar{w}_k)$ on vapaa. Lauseesta 7.6 seuraa, että jokainen aliavaruuden $W = \operatorname{span}(\bar{w}_1, \ldots, \bar{w}_k)$ vektori voidaan kirjoittaa tasan yhdellä tavalla vektoreiden $\bar{w}_1, \ldots, \bar{w}_k$ lineaarikombinaationa.

"$\Leftarrow$": Oletetaan, että jokainen aliavaruuden W vektori voidaan kirjoittaa täsmälleen yhdellä tavalla vektoreiden $\bar{w}_1, \ldots, \bar{w}_k$ lineaarikombinaationa. Tästä seuraa ensinnäkin, että $W = \operatorname{span}(\bar{w}_1, \ldots, \bar{w}_k)$. Nyt lauseen 7.6 mukaan jono $(\bar{w}_1, \ldots, \bar{w}_k)$ on vapaa. Näin kannan määritelmän molemmat ehdot täyttyvät. Siis $(\bar{w}_1, \ldots, \bar{w}_k)$ on aliavaruuden W kanta. $\square$

Esimerkki 8.4. Merkitään $\bar{w}_1 = (2, -1)$, $\bar{w}_2 = (1, 3)$. Osoitetaan lauseen 8.3 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 alkeisriviotoimituksilla, saadaan matriisi

$$\left[\begin{array}{cc|c} 1 & 0 & (3v_1 - v_2)/7 \\ 0 & 1 & (v_1 + 2v_2)/7 \end{array}\right].$$

Matriisista nähdään, että yhtälöryhmällä on täsmälleen yksi ratkaisu riippumatta vektorista $\bar{v} \in \mathbb{R}^2$. Siis jono $(\bar{w}_1, \bar{w}_2)$ on avaruuden $\mathbb{R}^2$ kanta.

8.1 Koordinaatit

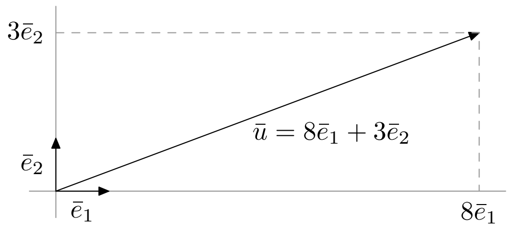
Määritelmä 8.5. Oletetaan, että $\mathcal{B} = (\bar{w}_1, \ldots, \bar{w}_k)$ on aliavaruuden W kanta. Oletetaan, että $\bar{u} \in W$. Vektorin $\bar{u}$ *koordinaateiksi kannan* $\mathcal{B}$ *suhteen* kutsutaan reaalilukuja $a_1, \ldots, a_k$, joilla

$$\bar{u} = a_1\bar{w}_1 + \cdots + a_k\bar{w}_k\,.$$

Huomaa, että vektorin koordinaatit jokin tietyn kannan suhteen ovat yksikäsitteiset, sillä vektori voidaan lauseen 8.3 mukaan kirjoitaa vain yhdellä tavalla kannan alkioiden lineaarikombinaationa. Vektorilla on siis vain yhdet koordinaatit jonkin tietyn kannan suhteen. Eri kantojen suhteen saman vektorin koordinaatit voivat tietenkin olla erilaisia.

Esimerkki 8.6. Määritetään vektorin $\bar{u} = (8, 3)$ koordinaatit avaruuden $\mathbb{R}^2$ luonnollisen kannan $\mathcal{E}_2 = (\bar{e}_1, \bar{e}_2)$ suhteen. Koordinaatit ovat 8 ja 3, sillä

$$\bar{u} = 8(1, 0) + 3(0, 1) = 8\bar{e}_1 + 3\bar{e}_2.$$



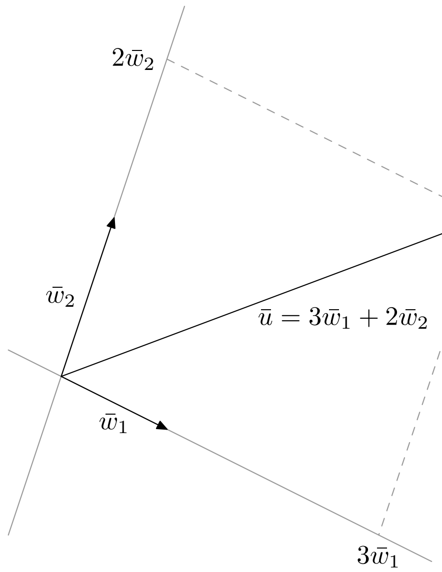
Kuva 8.21: Vektorin $\bar{u}$ koordinaatit kannan $\mathcal{E}_2 = (\bar{e}_1, \bar{e}_2)$ suhteen ovat 8 ja 3.

Tutkitaan sitten vektorin $\bar{u}$ koordinaatteja jonkin toisen kannan suhteen. Merkitään $\bar{w}_1 = (2, -1)$, $\bar{w}_2 = (1, 3)$. Esimerkin 8.4 perusteella $(\bar{w}_1, \bar{w}_2)$ on avaruuden $\mathbb{R}^2$ kanta.

Määritetään vektorin $\bar{u}$ koordinaatit kannan $\mathcal{B} = (\bar{w}_1, \bar{w}_2)$ suhteen. On siis ratkaistava yhtälö $x_1\bar{w}_1 + x_2\bar{w}_2 = \bar{u}$ eli yhtälö $x_1(2, -1) + x_2(1, 3) = (8, 3)$. Esimerkistä 8.4 nähdään, että yhtälön ratkaisu on

$$\begin{cases} x_1 = (3u_1 - u_2)/7 = (24 - 3)/7 = 3 \\ x_2 = (v_1 + 2v_2)/7 = (8 + 6)/8 = 2\,. \end{cases}$$

Siis $\bar{u} = 3\bar{w}_1 + 2\bar{w}_2$, eli kysytyt koordinaatit ovat 3 ja 2. Tilannetta on havainnollistettu kuvassa 8.22.



Kuva 8.22: Vektorin $\bar{u}$ koordinaatit kannan $\mathcal{B} = (\bar{w}_1, \bar{w}_2)$ suhteen ovat 3 ja 2.

8.2 Dimensio

Aliavaruudella voi olla useita eri kantoja, mutta jokaisessa niistä on yhtä monta vektoria.

Lause 8.7. *Aliavaruuden W 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 aliavaruuden W kantoja. Pyritään osoittamaan, että $j = k$. Tehdään se osoittamalla, että muut vaihtoehdot $j < k$ ja $k < j$ johtavat ristiriitaan.

Oletetaan, että $j < k$. Tarkastellaan yhtälöä

$$x_1\bar{w}_1 + \cdots + x_k\bar{w}_k = \bar{0}\,. \tag{3}$$

Koska $\mathcal{B}$ on W:n 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 $a_{11}, \dots, a_{kj} \in \mathbb{R}$.

Sijoittamalla nämä yhtälöön (3) saadaan yhtäpitävä yhtälö:

$$\begin{aligned}
&x_1(a_{11}\bar{v}_1 + a_{12}\bar{v}_2 + \dots + a_{1j}\bar{v}_j) \\
&\qquad\qquad + x_2(a_{21}\bar{v}_1 + a_{22}\bar{v}_2 + \dots + a_{2j}\bar{v}_j) \\
&\qquad\qquad\qquad\qquad + \dots + x_k(a_{k1}\bar{v}_1 + a_{k2}\bar{v}_2 + \dots + a_{kj}\bar{v}_j) = \bar{0}
\end{aligned}$$

ja edelleen ryhmittelemällä:

$$\begin{aligned}
&(x_1a_{11} + x_2a_{21} + \dots + x_ka_{k1})\bar{v}_1 \\
&\qquad\qquad + (x_1a_{12} + x_2a_{22} + \dots + x_ka_{k2})\bar{v}_2 \\
&\qquad\qquad\qquad\qquad + \dots + (x_1a_{1j} + x_2a_{2j} + \dots + x_ka_{kj})\bar{v}_j = \bar{0}
\end{aligned}$$

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

$$\begin{cases}
x_1a_{11} + x_2a_{21} + \dots + x_ka_{k1} = 0 \\
x_1a_{12} + x_2a_{22} + \dots + x_ka_{k2} = 0 \\
\qquad\qquad\qquad \vdots \qquad\qquad = \vdots \\
x_1a_{1j} + x_2a_{2j} + \dots + x_ka_{kj} = 0
\end{cases}$$

Kyseessä on homogeeninen yhtälöryhmä, jossa tuntemattomien määrä k on suurempi kuin yhtälöiden määrä j. Lauseen 7.10 mukaan yhtälöryhmällä on muitakin ratkaisuja kuin $x_1 = 0, \dots, x_k = 0$. Siis jono $\mathcal{C} = (\bar{w}_1, \dots, \bar{w}_k)$ on sidottu. Tämä on ristiriita, sillä $\mathcal{C}$ on aliavaruuden W kanta.

Tapaus $j > k$ käsitellään vastaavasti. Tällöinkin päädytään ristiriitaan. Siten täytyy päteä $j = k$. □

Voidaan osoittaa, että jokaisella avaruuden $\mathbb{R}^n$ aliavaruudella on kanta. (Tämä todistetaan kurssin toisessa osassa.) Lisäksi edellä nähtiin, että jokaisessa kannassa on yhtä monta vektoria. Kantavektorien lukumäärää kutsutaan avaruuden dimensioksi.

Määritelmä 8.8. Aliavaruuden W *dimensio* $\dim(W)$ on aliavaruuden W kannan vektoreiden lukumäärä.

Jos aliavaruuden dimensio on n, sanotaan, että aliavaruus on n*-ulotteinen*.

Esimerkki 8.9. Avaruuden $\mathbb{R}^2$ dimensio on 2, sillä avaruudella on kanta $(\bar{e}_1, \bar{e}_2)$. Vastaavasti avaruuden $\mathbb{R}^n$ dimensio on n, sillä avaruuden luonnollisen kannan vektorien lukumäärä on n.

Esimerkki 8.10. Merkitään $\bar{v}_1 = (3, -1, 5)$, $\bar{v}_2 = (2, 1, 3)$ ja $\bar{v}_3 = (0, -5, 1)$. Olkoon $W = \operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3)$. Määritetään aliavaruuden W dimensio.

Oletetaan, että $\bar{u} \in \mathbb{R}^3$. Selvitetään, mikä ehto vektorin $\bar{u}$ komponenttien pitää toteuttaa, jotta $\bar{u}$ on aliavaruudessa W. Ratkaistaan yhtälö $x_1\bar{v}_1 + x_2\bar{v}_2 + x_3\bar{v}_3 = \bar{u}$ eli yhtälö

$$x_1(3, -1, 5) + x_2(2, 1, 3) + x_3(0, -5, 1) = (u_1, u_2, u_3).$$

Sitä vastaa yhtälöryhmä

$$\begin{cases} 3x_1 + 2x_2 \phantom{{}-5x_3} = u_1 \\ -x_1 + x_2 - 5x_3 = u_2 \\ 5x_1 + 3x_2 + x_3 = u_3 \,. \end{cases}$$

Yhtälöryhmän matriisi on

$$\left[\begin{array}{rrr|c} 3 & 2 & 0 & u_1 \\ -1 & 1 & -5 & u_2 \\ 5 & 3 & 1 & u_3 \end{array}\right]$$

ja se saadaan alkeisrivitoimituksilla muutettua matriisiksi

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

Havaitaan, että yhtälöryhmällä on ratkaisu, jos ja vain jos alinta riviä vastaava yhtälö $0x_1 + 0x_2 + 0x_3 = (5u_3 + u_2 - 8u_1)/5$ on tosi eli $5u_3 + u_2 - 8u_1 = 0$. Siten

$$\begin{aligned} W &= \operatorname{span}(\bar{v}_1, \bar{v}_2, \bar{v}_3) \\ &= \{(u_1, u_2, u_3) \in \mathbb{R}^3 \mid 5u_3 + u_2 - 8u_1 = 0\} \\ &= \{(u_1, 8u_1 - 5u_3, u_3) \mid u_1, u_3 \in \mathbb{R}\} \\ &= \{u_1(1, 8, 0) + u_3(0, -5, 1) \mid u_1, u_3 \in \mathbb{R}\} \\ &= \operatorname{span}((1, 8, 0), (0, -5, 1)). \end{aligned}$$

Nyt tiedetään, että vektorit $(1, 8, 0)$ ja $(0, -5, 1)$ virittävät aliavaruuden. Lisäksi nämä kaksi vektoria ovat lineaarisesti riippumattomia. (Tämän osoittaminen jätetään lukijalle.) Siten jono $((1, 8, 0), (0, -5, 1))$ on avaruuden W kanta, ja $\dim(W) = 2$.

9 Matriisit

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 $B(2, 2) = 11$.

Samalla tavalla voidaan määritellä kompleksialkioisten matriisien joukot $\mathbb{C}^{m \times n}$. Kaikki jatkossa esiteltävät ominaisuudet pätevät myös kompleksimatriiseille, vaikka tekstissä puhutaan vain reaalialkioisista matriiseista.

9.1 Matriisien laskutoimituksia

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.

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 $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, n\}$. Esimerkiksi

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

Matriisia $(-1)A$ on tapana merkitä $-A$ ja matriisisummaa $A + (-B)$ on tapana merkitä $A - B$.

Matriiseille voidaan määritellä myös *matriisikertolasku*. Tämä laskutoimitus on hieman monimutkaisempi kuin edellä määritellyt. 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 $i \in \{1, \ldots, m\}$ ja $j \in \{1, \ldots, p\}$.[1]

Esimerkki 9.1. Lasketaan matriisien

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

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

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

Esimerkki 9.2. Matriisikertolasku ei ole vaihdannainen operaatio eli tulon tekijöiden järjestystä ei voi vaihtaa. Tarkastellaan vaikkapa matriiseja

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

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

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 määritellä *matriisipotenssi*

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

9.2 Erityisiä matriiseja

Matriisia

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

jonka kaikki alkiot ovat nollia, kutsutaan *nollamatriisiksi*. *Ykkösmatriisi* puolestaan on

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

Ei ole vaikea nähdä, että matriisikertolaskussa ykkösmatriisit käyttäytyvät kuin reaaliluku 1 tavallisessa kertolaskussa: kaikilla $A \in \mathbb{R}^{m \times n}$ pätee

$$I_m A = A I_n = A.$$

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

Neliömatriisi on matriisi, jossa on yhtä monta riviä ja saraketta. Neliömatriisin alkio on *lävistäjällä*, 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

Lause 9.3. *Seuraavat säännöt pätevät matriiseille A, B ja C sekä reaaliluvulle a, 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) $a(AB) = (aA)B = A(aB)$

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

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

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

Esimerkiksi matriisin

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

transpoosi on

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

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

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

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

Siis B on symmetrinen ja C on antisymmetrinen.

Lause 9.6. *Seuraavat säännöt pätevät matriiseille A ja B sekä reaaliluvulle t, jos laskutoimitukset on määritelty (ts. matriisit ovat sopivaa tyyppiä):*

a) $(A^T)^T = A$

b) $(A + B)^T = A^T + B^T$

c) $(AB)^T = B^T A^T$

d) $(tA)^T = t(A^T)$.

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

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

Siten $(AB)^T = B^T A^T$. □

9.5 Käänteismatriisi

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*.

Esimerkki 9.7. Matriisin

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

käänteismatriisi on

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

sillä

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

ja

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

Lause 9.8. *Matriisilla on korkeintaan yksi käänteismatriisi.*

Todistus. Oletetaan, että matriisilla A on käänteismatriisit B ja B'. Nyt

$$B = B \cdot I = B \cdot (A \cdot B') = (B \cdot A) \cdot B' = I \cdot B' = B'.$$

Siten B ja B' ovat välttämättä sama matriisi. Näin ollen 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 perustellut, että matriisi A todella on kääntyvä.

Lause 9.9. *Oletetaan, että matriisit A ja B ovat kääntyviä. Tällöin myös matriisit A^{-1}, AB ja A^T ovat kääntyviä. Niiden käänteismatriisit ovat seuraavat:*

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

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

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

Todistus. Lauseen matriisit osoitetaan kääntyviksi näyttämällä, että niillä on käänteismatriisi.

Osoitetaan matriisi A^{-1} on kääntyväksi näyttämällä, että sen käänteismatriisi on A. Koska $A^{-1}A = I$ ja $AA^{-1} = I$, on A matriisin A^{-1} käänteismatriisi eli $(A^{-1})^{-1} = A$. Matriisia AB koskevien väitteiden todistaminen jätetään harjoitustehtäväksi.

Osoitetaan lopuksi, että $(A^T)^{-1} = (A^{-1})^T$. Lauseen 9.6 nojalla

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

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

Lause 9.10. *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 $ad - bc = 0$. Nyt on tutkittava kaksi eri tapausta: joko $a = 0$ tai $a \neq 0$. Jos $a = 0$, niin $bc = 0$. 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 $AB = I$. Tuloon AB tulee nimittäin välttämättä nollarivi tai nollasarake. Siten A ei ole kääntyvä.

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

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

Oletetaan, että on olemassa sellainen matriisi

$$B = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

että $AB = I$. Tällöin

$$\begin{cases} ax + bz &= 1 \\ ay + bw &= 0 \\ cx + (bc/a)z &= 0 \\ cy + (bc/a)w &= 1. \end{cases}$$

Kolmannen yhtälön perusteella $c(x+bz/a) = 0$. Jos $c = 0$, päädytään samanlaiseen tilanteeseen kuin silloin, jos $a = 0$. Siten voidaa olettaa, että $c \neq 0$. Tällöin täytyy pätea $x + bz/a = 0$ 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. □

Suurempien matriisien käänteismatriiseille ei ole yhtä helppoa kaavaa. Käänteismatriisin määrittämistä käsitellään lisää luvussa 10.2.

9.6 Sarakevektorit

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

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

kanssa. Joukon $\mathbb{R}^{n\times 1}$ alkioita kutsutaan *sarakevektoreiksi*. Jos 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 9.11. Matriisin

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

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

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

Tämä sarakevektori voidaan samastaa vektorin $(2, 5, 13)$ kanssa.

10 Matriisit ja yhtälöryhmät

Tulemme näkemään, että lineaariset yhtälöryhmät voidaan esittää käyttäen matriisikertolaskua. Yhtälöryhmän

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ & \vdots & & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m, \end{cases} \tag{4}$$

kerroinmatriisiksi kutsutaan matriisia

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

Kerätään vielä muuttujat ja vakiot omiksi matriiseikseen:

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

Nyt yhtälöryhmä voidaan kirjoittaa matriisien avulla. Huomataan nimittäin, että

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

Matriisissa $A\bar{x}$ näkyy siis yhtälöryhmän vasen puoli. Yhtälöryhmä voidaan näin ollen kirjoittaa muodossa $A\bar{x} = \bar{b}$.

Lause 10.1. *Jos matriisi A on kääntyvä, yhtälöllä on $A\bar{x} = \bar{b}$ täsmälleen yksi ratkaisu.*

Todistus. Oletetaan, että matriisi A on kääntyvä. Todistuksessa on kaksi osaa. On osoitettava, että yhtälöllä on jokin ratkaisu ja että ratkaisuja ei ole enempää kuin yksi.

Osoitetaan ensin, että yhtälölle löytyy jokin ratkaisu. Koska A on kääntyvä, on olemassa käänteismatriisi A^{-1}. Nähdään, että $A^{-1}\bar{b}$ on yhtälön ratkaisu, sillä

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

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

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

ja edelleen $\bar{y} = A^{-1}\bar{b}$. Kysymyksessä onkin sama ratkaisu, joka löydettiin jo aikaisemmin. Siten ratkaisuja on vain yksi ja se on $A^{-1}\bar{b}$. $\square$

10.1 Alkeismatriisit

Tässä luvussa nähdään, miten alkeisrivitoimitukset voi ilmaista matriisikertolaskun avulla. Osoittautuu, että jos matriisia kerrotaan niin kutsutulla alkeismatriisilla, tullaan matriisille tehneeksi alkeisrivitoimitus. Tästä tulee olemaan hyötyä kääntyvien matriisien käsittelyssä.

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

Esimerkiksi seuraavat matriisit ovat alkeismatriiseja:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Nämä alkeismatriisit on saatu ykkösmatriisista tekemällä 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. Nähdään, että

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

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

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 tässä tapauksessa alkeismatriisilla kerrottaessa matriisille A tullaan tehneeksi sama alkeisrivioperaatio, jonka avulla alkeismatriisi muodostettiin.

Yksittäinen esimerkki ei takaa, että alkeismatriisilla kertominen vastaa aina alkeisrivitoimituksen tekemistä. Esimerkin perusteella voi kuitenkin ymmärtää, miksi näin on. Väitteen todistaminen on melko työlästä, joten se jätetään väliin.

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 vaikkapa suurempien lauseiden todistamisessa.

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

Todistus. 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 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.

$\square$

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 $E^{-1} = F$.

Lause 10.7. *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.

Todistus. Osoitetaan väite todistamalla seuraava päättelyketju:

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

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 vastaava lineaarinen yhtälöryhmä on

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

Koska yhtälöryhmällä on täsmälleen yksi ratkaisu ja muuttujia on yhtä monta kuin yhtälöitä, täytyy yhtälöryhmän olla ekvivalentti yhtälöryhmän

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

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

d) $\Rightarrow$ e): Oletetaan, että matriisi A riviekvivalentti ykkösmatrisiin kanssa. Olkoot $E_1, \dots, E_k$ ne alkeismatriisit, joilla kertomalla matriisista A saadaan redusoitu porrasmatriisi. Nyt siis pätee

$$E_k \cdots E_1 A = I.$$

Kun yhtälön vasemmat puolet kerrotaan matriisilla E_k^{-1}, saadaan $E_{k-1} \cdots E_1 A = E_k^{-1}$. Kun tämän yhtälön vasemmat puolet kerrotaan 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 $E_1, \dots, E_k$ ovat alkeismatriiseja. Merkitään

$$B = E_k^{-1} \cdots E_1^{-1}.$$

Nyt

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

Samalla tavalla nähdään, että $BA = I$. Siten B on matriisin A käänteismatriisi. $\square$

10.2 Käänteismatriisin määrittäminen

Muuttamalla matriisi redusoiduksi porrasmatriisiksi voidaan nähdä, onko matriisi kääntyvä. Jos matriisi A onnistutaan muuttamaan alkeisrivitoimituksilla ykkösmatriisiksi, niin A on kääntyvä eli sillä on käänteismatriisi A^{-1}. Muussa tapauksessa A ei ole kääntyvä.

Jos matriisi on käntyvä, käytetyistä alkeisrivitoimituksista saadaan myös selville, mikä käänteismatriisi on. Oletetaan, että matriisi A on muutettu ykkösmatriisiksi alkeisrivitoimituksilla, joita vastaavat alkeismatriisit $E_1, \dots, E_k$. Nyt

$$E_k \dots E_1 A = I.$$

Tällöin käänteismatriisille pätee

$$\begin{aligned} A^{-1} = IA^{-1} &= (E_k \dots E_1 A) A^{-1} = E_k \dots E_1 (AA^{-1}) \\ &= E_k \dots E_1 I. \end{aligned}$$

Tämä tarkoittaa, että tekemällä samat alkeisrivitoimitukset ykkösmatriisille I päädytään matriisiin A^{-1}.

Matriisin A kääntyvyyden selvittäminen ja käänteismatriisin etsiminen voidaan tehdä yhtä aikaa. Yhdistetään matriisit A ja I matrisiiksi $[A \mid I]$. Tehdään tälle matriisille alkeisrivitoimituksia, joilla A muutetaan redusoiduksi porrasmatriisiksi. Jos matriisi A saadaan muutettua alkeisrivitoimitusten avulla ykkösmatriisiksi, on A kääntyvä. Kuten edellä todettiin, samat alkeisrivitoimitukset muuttavat ykkösmatriisin I matriisin A käänteismatriisiksi A^{-1}. Siis

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

Esimerkki 10.8. 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.

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

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

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

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

Koska matriisi A saatiin muutettua alkeisrivitoimituksilla ykkösmatriisiksi, on A kääntyvä. Lisäksi sen käänteismatriisi on

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

Entä sitten, jos matriisi ei ole kääntyvä? Kuinka voidaan osoittaa, että matriisista ei saada alkeisrivitoimituksilla ykkösmatriisia? Voidaan osoittaa, että jos alkeisrivitoimitusten avulla saadaan aikaan nollarivi, ei matriisi voi olla riviekvivalentti ykkösmatriisin kanssa. (Todistusta ei esitetä tässä.) Nollarivi on siis merkki siitä, ettei matriisi ei ole kääntyvä.

Esimerkki 10.9. Tutkitaan, onko matriisilla

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

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 paikalle tuli nollarivi, matriisista B ei saada alkeisrivitoimituksilla ykkösmatriisia. Siten B ei ole kääntyvä.

11 Determinantti

Matriisin determinantti on reaaliluku, joka kertoo matriisin ominaisuuksista. Determinantti on kätevä työkalu esimerkiksi silloin, kun halutaan tietää, onko matriisi kääntyvä. Monet determinanttiin liittyvistä todistuksista ovat kuitenkin niin työläitä, että ne sivuutetaan tässä.

Määritelmä 11.1. Oletetaan, että A on $n \times n$-matriisi. Merkitään

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

a) Jos $n = 1$, niin $\det(A) = a_{11}$.

b) Muussa tapauksessa

$$\det(A) = \sum_{i=1}^{n} (-1)^{1+j} a_{1j} \det(A_{1j}),$$

missä A_{ij} on matriisi, joka on saatu matriisista A poistamalla i:s rivi ja j:s sarake.

Matriisin A determinantille voidaan käyttää merkintää

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

Esimerkki 11.2. Matriisin $A = [4]$ determinantti on $\det(A) = 4$. Matriisin

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

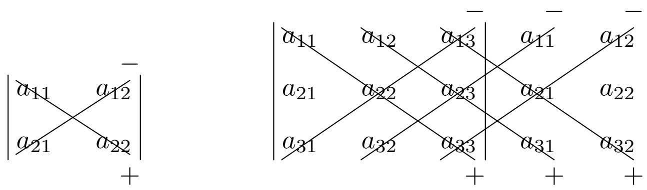
determinantti on puolestaan

$$\det(B) = 1 \cdot \det([4]) - (-1) \cdot \det([2]) = \det([4]) + \det([2]) = 4 + 2 = 6.$$

Määritelmän mukaan 2×2-matriisin determinantti lasketaan seuraavasti:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Sen laskemiseen voi käyttää kuvassa 11.23 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.



Kuva 11.23: Muistisäännöt 2×2-determinantin ja 3×3-determinantin laskemiseksi.

Esimerkki 11.3. Matriisin

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

determinantti on

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

Määritelmän mukaan 3×3-matriisin determinantti lasketaan seuraavasti:

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ &\quad + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned}$$

Tällekin determinantille on olemassa laskemista helpottava muistisääntö (ks. kuva 11.23). 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.

Lauseen 9.10 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 kaikille matriiseille. Sen osoittaminen on työlästä ja jätetään siksi väliin.

Lause 11.4. *Oletetaan, että A on $n \times n$-matriisi. Matriisi A on kääntyvä, jos ja vain jos* $\det(A) \neq 0$.

11.1 Determinantin kehityskaavat

Determinantin määritelmässä olevat kertoimet otetaan matriisin ensimmäiseltä riviltä. Sanotaan, että determinantti on tällöin kehitetty ensimmäisen rivin suhteen. Yhtä hyvin voidaan käyttää muita rivejä tai jopa muita sarakkeita.

Lause 11.5. *Oletetaan, että A on $n \times n$-matriisi. Merkitään $A(i,j) = a_{ij}$ kaikilla $i, j \in \{1, \dots, n\}$*

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

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}),$$

missä A_{ij} on matriisi, joka on saatu matriisista A poistamalla i:s rivi ja j:s sarake. Kyseessä on kehitys rivin i suhteen.

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

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}),$$

missä A_{ij} on matriisi, joka on saatu matriisista A poistamalla i:s rivi ja j:s sarake. Kyseessä on kehitys sarakkeen j suhteen.

Toisinaan voi säästää vaivaa, jos valitsee viisaasti rivin tai sarakkeen, jonka suhteen determinantin kehittää. Lasketaan matriisin

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

determinantti kehittämällä se aluksi kolmannen rivin ja sitten kolmannen sarakkeen suhteen:

$$\begin{aligned}
\det(D) &= \begin{vmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 2 & 1 & -4 & -1 \end{vmatrix} \\
&= 0 \cdot \begin{vmatrix} 2 & 3 & 0 \\ 1 & 0 & -1 \\ 1 & -4 & -1 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 3 & 0 \\ 0 & 0 & -1 \\ 2 & -4 & -1 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & -1 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 2 & 1 & -4 \end{vmatrix} \\
&= - \begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & -1 \end{vmatrix} = - \left(0 \cdot \begin{vmatrix} 0 & 1 \\ 2 & 1 \end{vmatrix} - (-1) \cdot \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} \right) \\
&= -\Big(0 + (1 \cdot 1 - 2 \cdot 2) - (1 \cdot 1 - 0 \cdot 2) \Big) = -(0 - 3 - 1) = 4
\end{aligned}$$

Kehityskaavojen plus- ja miinusmerkkien vaihtelu (eli kaavoissa oleva kerroin $(-1)^{i+j}$) 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. Huomaa, että alkion oma etumerkki säilyy joka tapauksessa.

11.2 Determinantin ominaisuuksia

Lause 11.6. *Oletetaan, että A ja B ovat $n \times n$-matriiseja. Tällöin*

a) $\det(A^T) = \det(A)$

b) $\det(AB) = \det(A)\det(B)$.

Lause 11.7. *Oletetaan, että matriisi A on kääntyvä. Tällöin*

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

Todistus. Oletuksen mukaan matriisi A on kääntyvä, joten sillä on käänteismatriisi A^{-1}. Lauseen 11.6 kohdan b) nojalla

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

Toisaalta lauseen 11.4 mukaan $\det(A) \neq 0$, sillä A on kääntyvä. Siten $\det(A^{-1}) = 1/\det(A)$. □

Seuraava lause kertoo, miten alkeisrivitoimitusten tekeminen vaikuttaa matriisin determinanttiin.

Lause 11.8. *Oletetaan, että A on neliömatriisi.*

1) Jos matriisi B saadaan matriisista A vaihtamalla kaksi riviä keskenään, niin $\det(B) = -\det(A)$.

2) Jos matriisi B saadaan matriisista A kertomalla jokin rivi reaaliluvulla $t \neq 0$, niin $\det(B) = t\det(A)$.

3) Jos matriisi B saadaan matriisista A lisäämällä johonkin riviin jokin toinen rivi reaaliluvulla k kerrottuna, niin $\det(B) = \det(A)$.

Lauseen 11.6 kohdasta a) seuraa, että determinantin sarakkeet käyttäytyvät täsmälleen samalla tavalla kuin sen rivit. Lauseesta 11.8 saadaan siis seuraavat muistisäännöt:

1) Jos matriisin kaksi riviä (saraketta) vaihtaa keskenään, niin 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 (sarakkeessa) kaikilla alkioilla on yhteinen tekijä, niin 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 (sarakkeeseen) lisätään jokin toinen rivi (sarake) vakiolla kerrottuna, niin matriisin determinantti ei muutu:
$$\begin{vmatrix} 1 & 3 & 0 \\ 3 & 1 & 7 \\ 5 & 4 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 0 \\ 0 & -8 & 7 \\ 5 & 4 & 4 \end{vmatrix}.$$

Joidenkin matriisien determinantti on helppo määrittää.

Lause 11.9. *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. Kohdat 1 ja 3 voidaan todistaa käyttämällä kehityskaavoja. Kohta 2 palautuu kohtaan 1 käyttämällä lauseen 11.8 kohtaa 3. □

Esimerkiksi
$$\begin{vmatrix} 1 & 7 & 0 & 6 \\ 0 & 3 & 4 & 5 \\ 0 & 0 & 8 & 2 \\ 0 & 0 & 0 & 9 \end{vmatrix} = 1 \cdot 3 \cdot 4 \cdot 9 = 108 \quad \text{ja} \quad \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 5 & 8 & 0 \\ 4 & 6 & 7 & 9 \end{vmatrix} = 1 \cdot 0 \cdot 8 \cdot 9 = 0.$$

12 Pistetulo

Avaruuksissa $\mathbb{R}^2$ ja $\mathbb{R}^3$ voidaan puhua vektorien pituuksista ja vektoreiden välisistä kulmista. Nämä käsitteet yleistetään avaruuteen $\mathbb{R}^n$ pistetulon avulla.

Määritelmä 12.1. 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

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

Esimerkiksi vektorien $\bar{v} = (3, -2, 0)$ ja $\bar{w} = (1, -2, \sqrt{3})$ pistetulo on

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

Huomaa, että pistetulosta tulee aina tulokseksi reaaliluku.

Pistetulolle voidaan todistaa laskusääntöjä.

Lause 12.2. *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 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. □

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

Lause 12.3. *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 $\bar{v} \cdot \bar{v} = 0^2 + 0^2 + \cdots + 0^2 = 0$. Siten väite on todistettu.

□

12.1 Vektorin normi

Pistetulon avulla voidaan määritellä avaruuden $\mathbb{R}^n$ vektorin normi eli pituus. Lauseen 12.3 nojalla $\bar{v} \cdot \bar{v} \geq 0$, joten seuraavassa määritelmässä juurrettava on epänegatiivinen, kuten kuuluu olla.

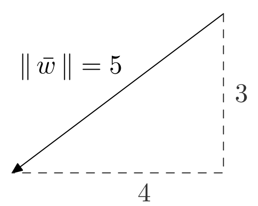
Määritelmä 12.4. Vektorin $\bar{v} \in \mathbb{R}^n$ *normi* eli pituus on

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

Määritelmästä seuraa, että $\|\bar{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$. Esimerkiksi vektorin $\bar{v} = (1/2, 3, -2, 0)$ normi on

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

Tason vektoreiden normia voidaan havainnollistaa Pythagoraan lauseen avulla. Kuvassa 12.24 on esitetty vektori $\bar{w} = (-4, -3)$. Sen pituus on Pythagoraan lauseen nojalla $\sqrt{3^2 + 4^2} = \sqrt{25} = 5$. Pituuden geometrinen tulkinta antaa siis saman tuloksen kuin määritelmä 12.4.



Kuva 12.24: Vektorin $\bar{w}$ normi eli pituus.

Vektorin normi on aina epänegatiivinen ja nollavektori on ainoa vektori, jonka normi on nolla.

Lause 12.5. *Oletetaan, että $\bar{v} \in \mathbb{R}^n$. Tällöin*

a) $\|\bar{v}\| \geq 0$

b) $\|\bar{v}\| = 0$, *jos ja vain jos* $\bar{v} = 0$.

Todistus. Tulokset seuraavat suoraan neliöjuuren ominaisuuksista ja lauseesta 12.3.

a) Määritelmän mukaan $\|\bar{v}\| = \sqrt{\bar{v} \cdot \bar{v}}$. Neliöjuuren arvo on aina epänegatiivinen, joten $\|\bar{v}\| \geq 0$.

b) Huomataan, että $\|\bar{v}\| = 0$, jos ja vain jos juurrettava $\bar{v} \cdot \bar{v}$ on nolla. Lauseen 12.3 nojalla taas $\bar{v} \cdot \bar{v} = 0$, jos ja vain jos $\bar{v} = \bar{0}$. Tämä todistaa väitteen.

$\square$

Lause 12.6. *Oletetaan, että* $\bar{v} \in \mathbb{R}^n$ *ja* $c \in \mathbb{R}$. *Tällöin* $\|c\bar{v}\| = |c|\|\bar{v}\|$.

Todistus. Pistetulon ominaisuuksien perusteella

$$\|c\bar{v}\| = \sqrt{c\bar{v} \cdot c\bar{v}} = \sqrt{c(\bar{v} \cdot c\bar{v})} = \sqrt{c^2(\bar{v} \cdot \bar{v})} = |c|\sqrt{(\bar{v} \cdot \bar{v})} = |c|\|\bar{v}\|.$$

$\square$

Määritelmä 12.7. Vektori $\bar{u} \in \mathbb{R}^n$ on *yksikkövektori*, jos sen normi on yksi eli

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

Esimerkiksi vektorit $(1, 0)$ ja $(0, 1)$ sekä vektorit $(1, 0, 0)$, $(0, 1, 0)$ ja $(0, 0, 1)$ ovat yksikkövektoreita.

Esimerkki 12.8. Etsitään 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 12.6 nojalla

$$(1/\sqrt{5}) \cdot \|\bar{v}\| = 1/\sqrt{5} \cdot \sqrt{5} = 1.$$

Lisäksi vektorit $\bar{v}$ ja $(1/\sqrt{5})\bar{v}$ ovat yhdensuuntaiset.

Lause 12.9. *Oletetaan, että* $\bar{v} \in \mathbb{R}^n \setminus \{\bar{0}\}$. *Tällöin vektori* $\dfrac{1}{\|\bar{v}\|}\bar{v}$ *on yksikkövektori, joka on yhdensuuntainen* $\bar{v}$*:n kanssa.*

Todistus. Väite seuraa lauseesta 12.6 samalla tavalla kuin esimerkissä 12.8. $\square$

Normin avulla voidaan määritellä vektorien välinen etäisyys.

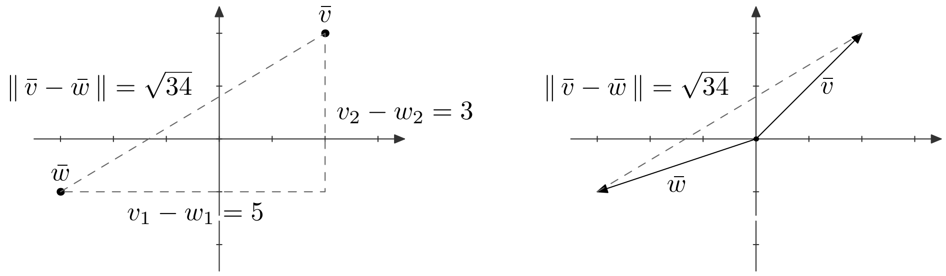
Määritelmä 12.10. Oletetaan, että $\bar{v}$, $\bar{w} \in \mathbb{R}^n$. Vektorien $\bar{v}$ ja $\bar{w}$ välinen etäisyys on

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

Esimerkki 12.11. Vektoreiden $\bar{v} = (2, 2)$ ja $\bar{w} = (-3, -1)$ välinen etäisyys on

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

Sitä on havainnollistettu kahdella eri tavalla kuvassa 12.25.



Kuva 12.25: Vektoreiden $\bar{v}$ ja $\bar{w}$ välinen etäisyys.

Lause 12.12 (Schwarzin epäyhtälö)**.** *Oletetaan, että* $\bar{v} \in \mathbb{R}^n$ *ja* $\bar{w} \in \mathbb{R}^n$*. Tällöin*

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

Todistus. Schwarzin epäyhtälöä tarvitaan tässä vaiheessa lähinnä lemman 12.13 todistamiseen. Lauseen todistus on kuitenkin melko tekninen, joten sitä lykätään kurssin toiseen osaan. □

12.2 Vektorien välinen kulma ja kohtisuoruus

Avaruudessa $\mathbb{R}^n$ kahden vektorin välinen kulma määritetään pistetulon avulla. Vektorien välisen kulman määrittelyyn tarvitaan seuraavaa lemmaa.

Lemma 12.13. *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 12.12 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

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

□

Kosinifunktio on määritelty niin, että jokaista lukua $a \in [-1, 1]$ vastaa täsmälleen yksi sellainen kulma α, että $0^\circ \leq \alpha \leq 180^\circ$ ja $\cos \alpha = a$. Edellisen lemman nojalla voidaan siis asettaa seuraava määritelmä.

Määritelmä 12.14. 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 $0^\circ \leq \alpha \leq 180^\circ$ 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 $0° \leq \alpha \leq 180°$. Näin vektorien väliseksi kulmaksi saadaan $\alpha \approx 46{,}65°$.

Tason vektorien tapauksessa vektorien välisen kulman määritelmä vastaa geometrista käsitystämme vektorien välisestä kulmasta. Kosinilauseen mukaan kuvan 12.26 kolmiossa

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

Toisaalta normin määritelmän nojalla

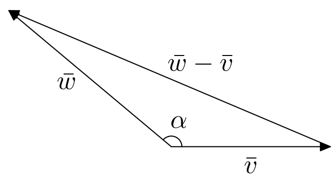
$$\begin{aligned}\|\bar{w} - \bar{v}\|^2 &= (\bar{w} - \bar{v}) \cdot (\bar{w} - \bar{v}) = \bar{w} \cdot \bar{w} - \bar{w} \cdot \bar{v} - \bar{v} \cdot \bar{w} + \bar{v} \cdot \bar{v} \\ &= \|\bar{v}\|^2 - 2(\bar{v} \cdot \bar{w}) + \|\bar{w}\|^2.\end{aligned}$$

Siten

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

ja edelleen

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



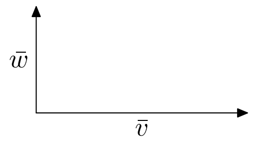
Kuva 12.26: Vektoreiden $\bar{v}$ ja $\bar{w}$ välinen kulma kosinilauseen näkökulmasta.

Määritelmä 12.15. Vektorit $\bar{v} \in \mathbb{R}^n$ ja $\bar{w} \in \mathbb{R}^n$ ovat *ortogonaaliset* eli *kohtisuorassa toisiaan vastaan*, jos $\bar{v} \cdot \bar{w} = 0$. Tällöin merkitään $\bar{v} \perp \bar{w}$.

Yleensä kahden olion ajatellaan olevan kohtisuorassa toisiaan vastaan, jos niiden välinen kulma on $90°$. Tämä pitää paikkansa myös vektoreiden tapauksessa. Oletetaan, että $\bar{v}, \bar{w} \in \mathbb{R}^n \setminus \{\bar{0}\}$. Määritelmän mukaan vektoreiden $\bar{v}$ ja $\bar{w}$ välinen kulma on $90°$, jos ja vain jos

$$\frac{\bar{v} \cdot \bar{w}}{\|\bar{v}\|\ \|\bar{w}\|} = \cos 90° = 0.$$

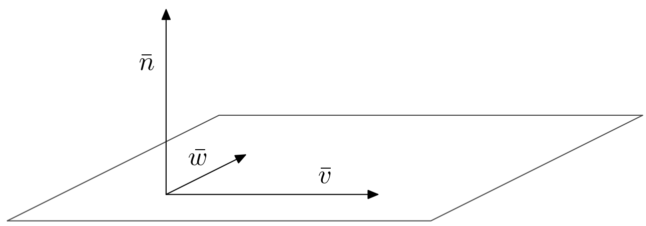
Tämä puolestaan pätee, jos ja vain jos $\bar{v} \cdot \bar{w} = 0$. Siten vektoreiden $\bar{v}$ ja $\bar{w}$ välinen kulma on $90°$, jos ja vain jos $\bar{v} \cdot \bar{w} = 0$.



Kuva 12.27: Vektorit $\bar{v}$ ja $\bar{w}$ ovat ortogonaaliset eli kohtisuorassa toisiaan vastaan.

Pistetulon sovellus: Tason normaalimuotoinen yhtälö

Vektorin sanotaan olevan kohtisuorassa tasoa vastaan, jos se on kohtisuorassa tason suuntavektoreita vastaan. Tällaista vektoria kutsutaan tason *normaaliksi* (ks. kuva 12.28).



Kuva 12.28: Tason normaali $\bar{n}$.

Oletetaan, että T on avaruuden $\mathbb{R}^3$ taso, joka kulkee pisteen P kautta ja jolla on normaali $\bar{n}$. Voidaan osoittaa, että piste $Q = (x, y, z)$ on tasossa T, jos ja vain jos

$$\bar{n} \cdot (\bar{q} - \bar{p}) = 0,$$

missä $\bar{q} = \overline{OQ}$ ja $\bar{p} = \overline{OP}$. Tilannetta on havainnollistettu kuvassa 12.29.

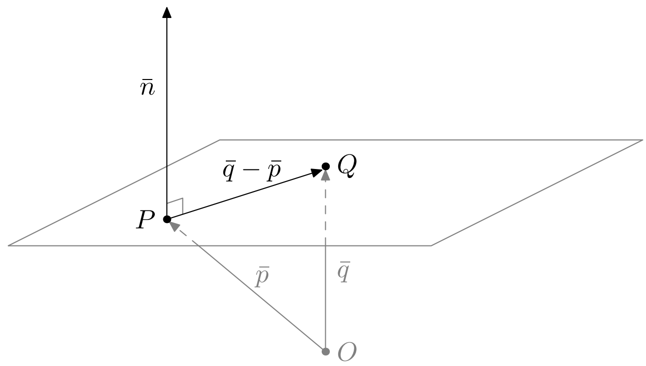
Edellä esitettyä yhtälöä kutsutaan tason T *normaalimuotoiseksi yhtälöksi*. Piste Q on tasossa T, jos ja vain jos pisteen paikkavektori $\bar{q}$ toteuttaa yhtälön.

Esimerkki 12.16. Oletetaan, että taso T kulkee pisteen $P = (6, 0, 1)$ kautta ja sillä on normaali $\bar{n} = (1, 2, 3)$. Tason T normaalimuotoinen yhtälö on tällöin

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

Tasossa T ovat siis ne pisteet Q, joiden paikkavektori $\bar{q}$ toteuttaa edellä esitetyn yhtälön. Toisin sanoen

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



Kuva 12.29: Tason T normaalimuotoisen yhtälön havainnollistus.

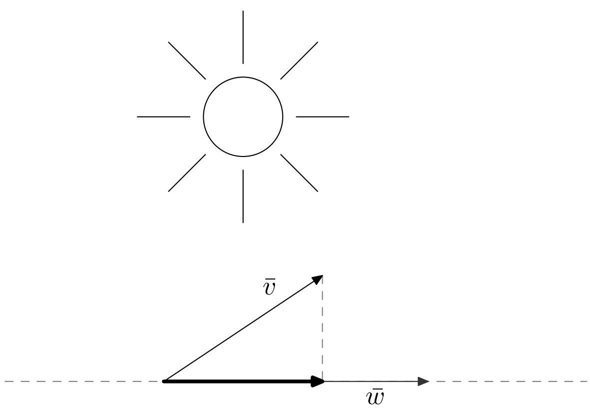
Kirjoitetaan taso vielä hiukan toisenlaisessa muodossa. Merkitään $\bar{q} = (x, y, z)$, missä $x, y, z \in \mathbb{R}$. Nyt

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

joten voidaan kirjoittaa $T = \{\bar{q} \in \mathbb{R}^3 \mid x + 2y + 3z - 9 = 0\}$.

12.3 Projektio

Ryhdymme määrittelemään vektorin $\bar{v}$ projektiota vektorin $\bar{w}$ virittämälle aliavaruudelle $\operatorname{span}(\bar{w})$ (eli vektorin $\bar{w}$ suuntaiselle suoralle). Voidaan ajatella, että projektio on vektorin $\bar{v}$ heittämä varjo, kun aurinko paistaa kohtisuoraan vektoria $\bar{w}$ vastaan kuten kuvassa 12.30.



Kuva 12.30: Projektion havainnollistus.

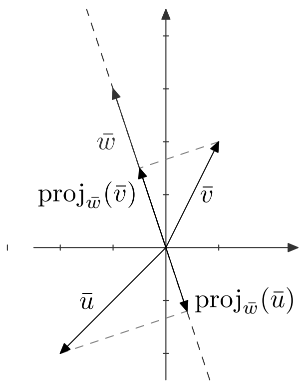
Määritelmä 12.17. Oletetaan, että $n \in \{1, 2, \dots\}$. Olkoot $\bar{v}, \bar{w} \in \mathbb{R}^n$ ja $\bar{w} \neq \bar{0}$. Tällöin vektorin $\bar{v}$ *projektio* vektorin $\bar{w}$ virittämälle aliavaruudelle on

$$\mathrm{proj}_{\bar{w}}(\bar{v}) = \frac{\bar{v} \cdot \bar{w}}{\bar{w} \cdot \bar{w}} \bar{w}.$$

Esimerkki 12.18. Esimerkiksi vektorin $\bar{v} = (1, 2)$ projektio vektorin $\bar{w} = (-1, 3)$ virittämälle aliavaruudelle on

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

Projektio on esitety kuvassa 12.31.



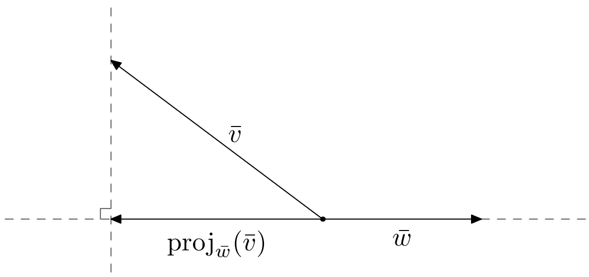
Kuva 12.31: Vektoreiden $\bar{v}$ ja $\bar{u}$ projektiot vektorin $\bar{w}$ virittämälle aliavaruudelle.

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

$$\mathrm{proj}_{\bar{w}}(v) = \frac{-4}{10}(-1, 3) = -\frac{2}{5}(-1, 3) = \left(\frac{2}{5}, -\frac{6}{5}\right).$$

Määritelmästä nähdään, että $\mathrm{proj}_{\bar{w}}(v)$ on aina yhdensuuntainen vektorin $\bar{w}$ kanssa. (Vektoria $\bar{w}$ kerrotaan nimittäin skalaarilla $(\bar{v} \cdot \bar{w})/(\bar{w} \cdot \bar{w})$.) Ei ole myöskään vaikea osoittaa, että vektorit $\bar{v} - \mathrm{proj}_{\bar{w}}(\bar{v})$ ja $\bar{w}$ ovat ortogonaaliset eli kohtisuorassa toisiaan vastaan.

Vektorin projektion voi määrittää myös geometrisesti (ks. kuva 12.32). Piirretään vektorit $\bar{v}$ ja $\bar{w}$ alkamaan samasta pisteestä ja piirretään vektorin $\bar{w}$ suuntainen suora. 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.

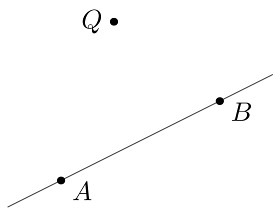


Kuva 12.32: Vektorin $\bar{v}$ projektio vektorin $\bar{w}$ virittämälle aliavaruudelle.

Projektion sovellus: Pisteen etäisyys suorasta

Pisteen etäisyys suorasta voidaan määrittää projektion avulla. Pisteen Q etäisyys suorasta $S = \{\bar{p} + t\bar{v} \mid t \in \mathbb{R}\}$ on kaikkein lyhin välimatka, joka voi olla pisteen Q ja suoran S pisteen välillä. Täsmällisesti ilmaistuna pisteen Q etäisyys suorasta S on $\min\{d(\bar{q}, \bar{a}) \mid \bar{a} \in S\}$, missä $\bar{q}$ on pisteen Q paikkavektori.

Tutkitaan esimerkin avulla, kuinka projektiota voidaan käyttää etäisyyden määrittämisessä. Tarkkoja todistuksia ei esitetä. Määritetään pisteen $Q = (4, -1, 9)$ etäisyys suorasta S, joka kulkee pisteiden $A = (2, -3, 5)$ ja $B = (4, 1, 7)$ kautta (ks. kuva 12.33).



Kuva 12.33: Pisteiden A ja B kautta kulkeva suora S.

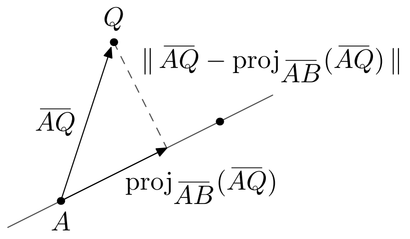
Määritetään ensin vektori jostakin suoran pisteestä tutkittavaan pisteeseen. Esimerkiksi vektori

$$\overline{AQ} = \overline{OQ} - \overline{OA} = (2, 2, 4)$$

käy tähän tarkoitukseen. Lisäksi tarvitaan suoran suuntainen vektori, kuten vaikkapa vektori $\overline{AB} = (2, 4, 2)$.

Vektorin $\overline{AQ}$ projektio suoralle S on

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



Kuva 12.34: Pisteen Q etäisyys suorasta S.

Erotus $\overline{AQ} - \mathrm{proj}_{\overline{AB}}(\overline{AQ})$ on kohtisuorassa suoraa S vastaan. Lasketaan erotus:

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

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

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

Siten pisteen Q etäisyys suorasta S on $\frac{1}{3}\sqrt{66}$.

12.4 Ortogonaalinen ja ortonormaali kanta

Määritelmä 12.19. Avaruuden $\mathbb{R}^n$ aliavaruuden W kanta $(\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k)$ on *ortogonaalinen*, jos

$$\bar{w}_i \cdot \bar{w}_j = 0 \quad \text{kaikilla } i, j \in \{1, 2, \ldots, k\}, \text{ missä } i \neq j.$$

Toisin sanoen kantavektorit ovat kohtisuorassa toisiaan vastaan.

Kanta $(\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k)$ on *ortonormaali*, jos se on ortogonaalinen ja lisäksi

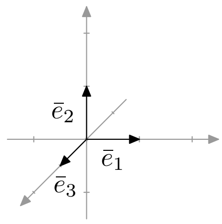
$$\|\bar{w}_i\| = 1 \quad \text{kaikilla } i \in \{1, 2, \ldots, k\}.$$

Toisin sanoen kantavektorit ovat kohtisuorassa toisiaan vastaan ja niiden normi on yksi.

Esimerkki 12.20. Avaruuden $\mathbb{R}^n$ luonnollinen kanta $\mathcal{E}_n = (\bar{e}_1, \ldots, \bar{e}_n)$ on ortonormaali. Huomataan nimittäin, että

$$\bar{e}_i \cdot \bar{e}_j = 0, \quad \text{jos } i \neq j.$$

Lisäksi $\|\bar{e}_i\| = 1$ kaikilla $i \in \{1, \ldots, n\}$.

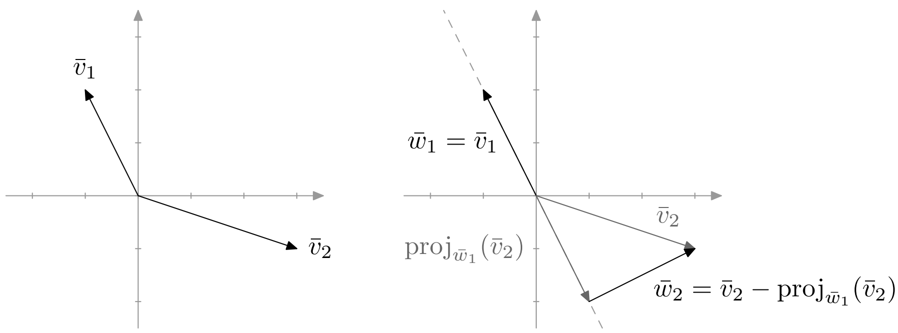


Kuva 12.35: Avaruuden $\mathbb{R}^3$ luonnollinen kanta $\mathcal{E}_3 = (\bar{e}_1, \bar{e}_2, \bar{e}_3)$ on ortonormaali.

Ortogonaaliset kannat ovat monissa tilanteissa hyvin käyttökelpoisia. Mistä tahansa kannasta voidaan muodostaan ortogonaalinen kanta projektiota apuna käyttäen. Seuraavassa esimerkissä näytetään, miten tämä tapahtu avaruudessa $\mathbb{R}^2$. Asiaan palataan tarkemmin kurssin toisessa osassa.

Esimerkki 12.21. Merkitään $\bar{v}_1 = (-1, 2)$ ja $\bar{v}_2 = (3, -1)$. Jono $(\bar{v}_1, \bar{v}_2)$ on avaruuden $\mathbb{R}^2$ kanta. (Tämän todistaminen jätetään lukijalle.)

Etsitään ortogonaalinen kanta muodostamalla uusi jono $(\bar{w}_1, \bar{w}_2)$ valitsemalla $\bar{w}_1 = \bar{v}_1$ ja $\bar{w}_2 = \bar{v}_2 - \text{proj}_{\bar{w}_1}(\bar{v}_2)$. Nyt vektorit $\bar{w}_1$ ja $\bar{w}_2$ ovat ortogonaaliset kuten luvussa 12.3 todettiin. Lisäksi $(\bar{w}_1, \bar{w}_2)$ on avaruuden $\mathbb{R}^2$ kanta, minkä todistaminen jätetään jälleen lukijalle.

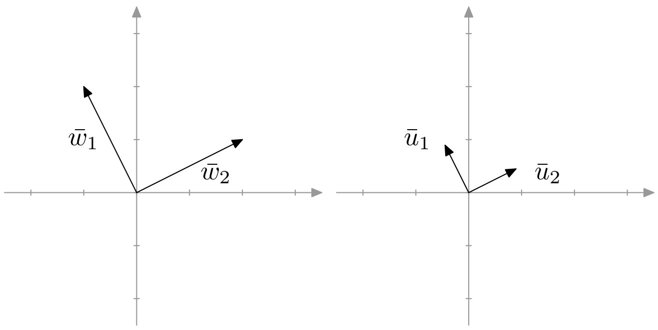


Kuva 12.36: Kannan $(\bar{v}_1, \bar{v}_2)$ muuttaminen ortogonaaliseksi kannaksi $(\bar{w}_1, \bar{w}_2)$.

Näin saadusta ortogonaalisesta kannasta voidaan vielä muodostaa ortonormaali kanta $(\bar{u}_1, \bar{u}_2)$ valitsemalla

$$\bar{u}_1 = \frac{1}{\|\bar{w}_1\|}\bar{w}_1 = \frac{1}{\sqrt{5}}(-1, 2) \quad \text{ja} \quad \bar{u}_2 = \frac{1}{\|\bar{w}_2\|}\bar{w}_2 = \frac{1}{\sqrt{5}}(2, 1).$$

Jono $(\bar{u}_1, \bar{u}_2)$ on avaruuden $\mathbb{R}^2$ ortonormaali kanta.



Kuva 12.37: Ortogonaalinen kanta $(\bar{w}_1, \bar{w}_2)$ ja ortonormaali kanta $(\bar{u}_1, \bar{u}_2)$.

Vektorin koordinaatit ortonormaalin kannan suhteen saadaan pistetulon avulla.

Lause 12.22. *Oletetaan, että $\mathcal{B} = (\bar{u}_1, \dots, \bar{u}_k)$ on aliavaruuden W ortonormaali kanta. Oletetaan, että $\bar{w} \in W$. Tällöin vektorin $\bar{w}$ koordinaatit kannan $\mathcal{B}$ suhteen ovat $\bar{w} \cdot \bar{u}_1$, $\bar{w} \cdot \bar{u}_2, \dots, \bar{w} \cdot \bar{u}_k$ eli*

$$\bar{w} = (\bar{w} \cdot \bar{u}_1)\bar{u}_1 + (\bar{w} \cdot \bar{u}_2)\bar{u}_2 + \cdots + (\bar{w} \cdot \bar{u}_k)\bar{u}_k\,.$$

Todistus. Tutkitaan vektorin $\bar{w} \in W$ koordinaatteja kannan S suhteen. Olkoot koordinaatit $a_1, \dots, a_k$ eli $\bar{w} = a_1\bar{w}_1 + a_2\bar{w}_2 + \cdots + a_k\bar{w}_k$. Huomataan, että

$$\begin{aligned}
\bar{w} \cdot \bar{w}_1 &= (a_1\bar{w}_1 + a_2\bar{w}_2 + \cdots + a_k\bar{w}_k) \cdot \bar{w}_1 \\
&= a_1(\bar{w}_1 \cdot \bar{w}_1) + a_2(\bar{w}_2 \cdot \bar{w}_1) + \cdots + a_k(\bar{w}_k \cdot \bar{w}_1) \\
&= a_1 \cdot 1 + a_2 \cdot 0 + \cdots + a_k \cdot 0 = a_1.
\end{aligned}$$

Vastaavalla tavalla nähdään, että $\bar{w} \cdot \bar{w}_i = a_i$ kaikilla $i \in \{1, 2, \dots, k\}$. Vektorin $\bar{w}$ koordinaatit kannan $\mathcal{B}$ suhteen saadaan siis laskemalla $\bar{w}$:n pistetulo kantavektorien kanssa. □

Esimerkki 12.23. Määritetään vektorin $\bar{w} = (2, 9, -7)$ koordinaantit ortonormaalin kannan $\mathcal{E}_3 = (\bar{e}_1, \bar{e}_2, \bar{e}_3)$ suhteen käyttäen edellä osoitettua tulosta. Koska

$$\begin{aligned}
\bar{w} \cdot \bar{e}_1 &= (2, 9, -7) \cdot (1, 0, 0) = 2, \\
\bar{w} \cdot \bar{e}_2 &= (2, 9, -7) \cdot (0, 1, 0) = 9, \\
\bar{w} \cdot \bar{e}_3 &= (2, 9, -7) \cdot (0, 0, 1) = -7,
\end{aligned}$$

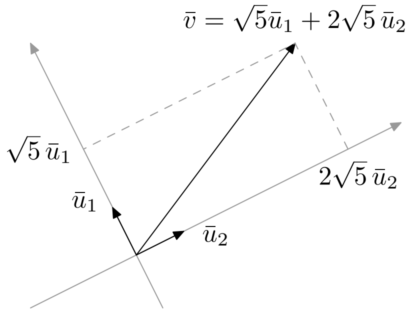
saadaan $\bar{w} = 2\bar{e}_1 + 9\bar{e}_2 - 7\bar{e}_3$. (Tämän olisimme toki voineet päätellä suoraankin.)

Esimerkki 12.24. Tarkastellaan esimerkissä 12.20 muodostettua avaruuden $\mathbb{R}^2$ ortonormaalia kantaa $(\bar{u}_1, \bar{u}_2)$, jossa

$$\bar{u}_1 = \frac{1}{\sqrt{5}}(-1, 2), \qquad \bar{u}_2 = \frac{1}{\sqrt{5}}(2, 1).$$

Vektorin $\bar{v} = (3, 4)$ koordinaatit tämän kannan suhteen ovat

$$\begin{aligned}
\bar{v} \cdot \bar{u}_1 &= \frac{1}{\sqrt{5}}\Big((3,4) \cdot (-1,2)\Big) = \frac{5}{\sqrt{5}} = \sqrt{5}, \\
\bar{v} \cdot \bar{u}_2 &= \frac{1}{\sqrt{5}}\Big((3,4) \cdot (2,1)\Big) = \frac{10}{\sqrt{5}} = 2\sqrt{5}.
\end{aligned}$$



Kuva 12.38: Vektorin $\bar{v}$ koordinaantit ortonormaalin kannan $(\bar{u}_1, \bar{u}_2)$ suhteen.

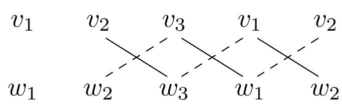
13 Ristitulo

Avaruuden $\mathbb{R}^3$ vektoreille voidaan määritellä ristitulo. Ristitulon tulos on avaruuden $\mathbb{R}^3$ vektori. Ristitulosta on hyötyä esimerkiksi silloin, kun tarvitaan vektori, joka on kohtisuorassa tasoa vastaan.

Määritelmä 13.1. 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

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

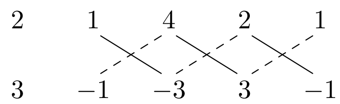
Ristitulon $\bar{v} \times \bar{w}$ laskemiseen voi käyttää kuvassa 13.39 esitettyä laskusääntöä. Yhtenäisellä viivalla yhdistettyjen komponenttien tulosta vähennetään katkoviivalla yhdistettyjen komponenttien tulo.



Kuva 13.39: Ristitulon $\bar{v} \times \bar{w}$ laskeminen.

Esimerkki 13.2. Merkitään $\bar{a} = (2, 1, 4)$ ja $\bar{b} = (3, -1, -3)$. Tällöin

$$\begin{aligned} \bar{a} \times \bar{b} &= (1 \cdot (-3) - 4 \cdot (-1),\ 4 \cdot 3 - 2 \cdot (-3),\ 2 \cdot (-1) - 1 \cdot 3) \\ &= (1, 18, -5). \end{aligned}$$



Kuva 13.40: Ristitulon $\bar{a} \times \bar{b}$ laskeminen.

Ristitulolle saadaan toinen muistisääntö determinantin avulla. Vektoreiden $\bar{v}$ ja $\bar{w}$ ristitulo saadaan laskemalla determinantti

$$\begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

Tässä $\bar{i} = (1, 0, 0)$, $\bar{j} = (0, 1, 0)$ ja $\bar{k} = (0, 0, 1)$. Tarkalleen ottaen determinantin alkiot eivät voi olla vektoreita. Kyseessä on kuitenkin vain muistisääntö, ja vektoreiden $\bar{i}$, $\bar{j}$ ja $\bar{k}$ ajatellaan käyttäytyvän determinanttia laskettaessa reaalilukujen tavoin.

Esimerkiksi vektoreiden $\bar{a} = (2, 1, 4)$ ja $\bar{b} = (3, -1, -3)$ ristitulo on

$$\begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ 2 & 1 & 4 \\ 3 & -1 & -3 \end{vmatrix} = -3\bar{i} + 12\bar{j} - 2\bar{k} - 3\bar{k} + 4\bar{i} + 6\bar{j} = \bar{i} + 18\bar{j} - 5\bar{k} = (1, 18, -5).$$

Ristitulon avulla voidaan löytää vektori, joka on kohtisuorassa kahta vektoria vastaan.

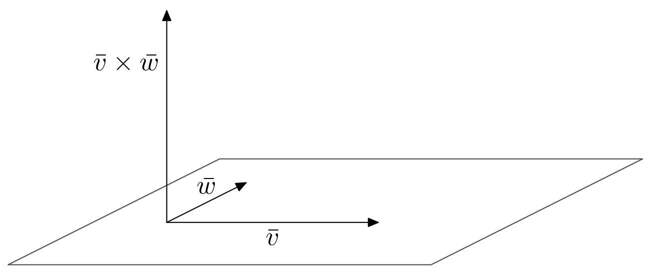
Lause 13.3. *Oletetaan, että $\bar{v}$, $\bar{w} \in \mathbb{R}^3$. Tällöin*

$$(\bar{v} \times \bar{w}) \perp \bar{v} \quad \textit{ja} \quad (\bar{v} \times \bar{w}) \perp \bar{w}.$$

Todistus. 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. □



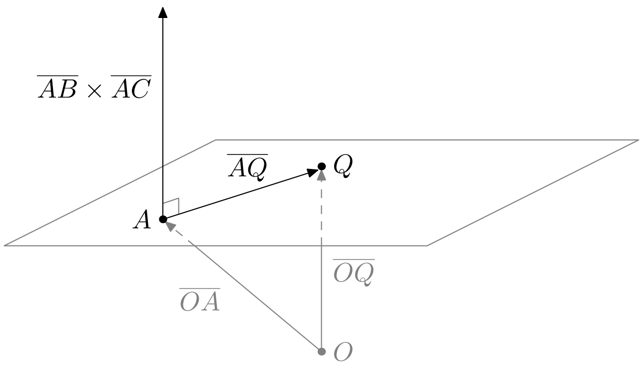
Kuva 13.41: Ristitulo $\bar{v} \times \bar{w}$ on kohtisuorassa vektoria $\bar{v}$ ja vektoria $\bar{w}$ vastaan.

Ristitulon avulla voidaan löytää tason normaali (eli vektori, joka on kohtisuorassa tason suuntavektoreita vastaan). Tästä on hyötyä tason normaalimuotoisen yhtälön määrittämisessä.

Esimerkki 13.4. Määritetään normaalimuotoinen yhtälö tasolle T, joka kulkee pisteiden $A = (0, 1, 0)$, $B = (-1, 3, 2)$ ja $C = (-2, 0, 1)$ kautta. Tätä varten tarvitaan tason T normaali. Vektorien $\overline{AB} = (-1, 2, 2)$ ja $\overline{AC} = (-2, -1, 1)$ ristitulo

$$\overline{AB} \times \overline{AC} = (4, -3, 5)$$

käy edellisen lauseen nojalla tähän tarkoitukseen.



Kuva 13.42: Tason T normaalimuotoisen yhtälön määrittäminen.

Lisäksi tarvitaan vektori jostakin tason pisteestä pisteeseen $Q = (x, y, z)$. Valitaan vektori $\overline{AQ} = \overline{OQ} - \overline{OA} = (x, y-1, z)$. Tason T normaalimuotoiseksi yhtälöksi saadaan $(\overline{AB} \times \overline{AC}) \cdot \overline{AQ} = 0$ eli

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

Laskemalla pistetulo saadaan yhtälö muotoon

$$4x - 3y + 5z + 3 = 0.$$

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

Esimerkki 13.5. 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 13.43).

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

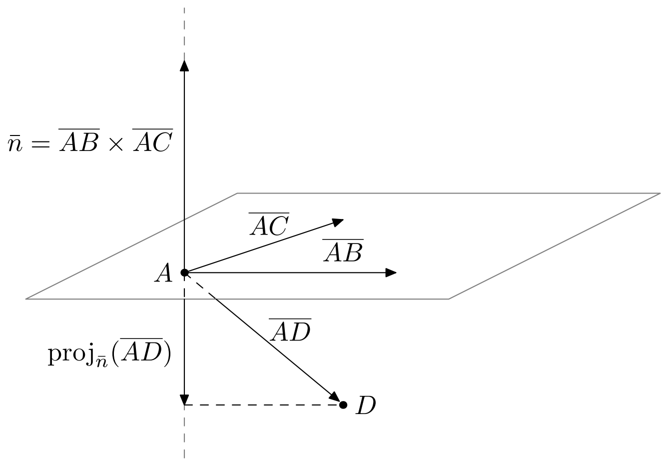
$$\overline{AD} = \overline{OD} - \overline{OA} = (1, 1, 3).$$

Vektorin $\overline{AD}$ projektio normaalin $\bar{n} = \overline{AB} \times \overline{AC}$ 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 P etäisyys tasosta T:

$$\|\operatorname{proj}_{\bar{n}}(\overline{AD})\| = \frac{8}{25}\|(4, -3, 5)\| = \frac{8}{25}\sqrt{16 + 9 + 25} = \frac{8}{25}\sqrt{50} = \frac{8}{5}\sqrt{2}.$$



Kuva 13.43: Pisteen D etäisyys tasosta T.

Käydään vielä läpi muutamia ristituloon liittyviä laskusääntöjä.

Lause 13.6. *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 on myös pistetuloon liittyviä laskusääntöjä.

Lause 13.7. *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 13.6 kohtaa g) ja lauseen 13.7 kohtaa a) saadaan

$$\begin{aligned}\|\bar{v} \times \bar{w}\|^2 &= (\bar{v} \times \bar{w}) \cdot (\bar{v} \times \bar{w}) = ((\bar{v} \times \bar{w}) \times \bar{v}) \cdot \bar{w} \\ &= ((\bar{v} \cdot \bar{v})\bar{w} - (\bar{v} \cdot \bar{w})\bar{v}) \cdot \bar{w} = (\|\bar{v}\|^2\bar{w} - (\bar{v} \cdot \bar{w})\bar{v}) \cdot \bar{w} \\ &= \|\bar{v}\|^2(\bar{w} \cdot \bar{w}) - (\bar{v} \cdot \bar{w})(\bar{v} \cdot \bar{w}) = \|\bar{v}\|\|\bar{w}\|^2 - (\bar{v} \cdot \bar{w})^2.\end{aligned}$$

Siten Lagrangen identiteetti pätee. $\square$

Lause 13.8. *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 13.7). 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

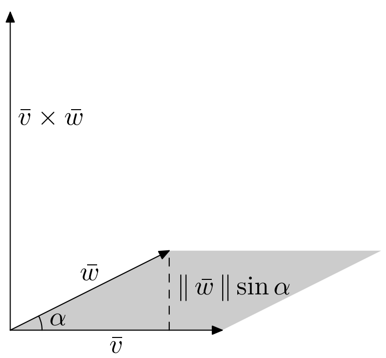
$$\begin{aligned}
\|\bar{v} \times \bar{w}\|^2 &= \|\bar{v}\|^2\|\bar{w}\|^2 - (\bar{v} \cdot \bar{w})^2 = \|\bar{v}\|^2\|\bar{w}\|^2 - (\cos \alpha \|\bar{v}\|\|\bar{w}\|)^2 \\
&= \|\bar{v}\|^2\|\bar{w}\|^2 - \cos^2 \alpha \|\bar{v}\|^2\|\bar{w}\|^2 = \|\bar{v}\|^2\|\bar{w}\|^2(1 - \cos^2 \alpha) \\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

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. $\square$

Edellisestä lauseesta seuraa, että ristitulovektorin $\bar{v} \times \bar{w}$ pituus on yhtä suuri kuin vektorien $\bar{v}$ ja $\bar{w}$ määräämän suunnikkaan ala (ks. kuva 13.44). Oletetaan, 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 13.44: Ristitulovektorin $\bar{v} \times \bar{w}$ pituus on yhtä suuri kuin vektorien $\bar{v}$ ja $\bar{w}$ määräämän suunnikkaan ala.

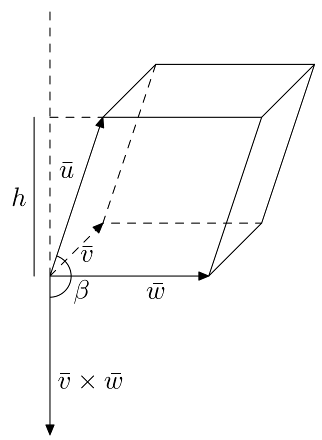
Ristitulon avulla voidaan määrittää myös suuntaissärmiön tilavuus. Vektoreiden $\bar{v}$, $\bar{w}$ ja $\bar{u}$ määräämään suuntaissärmiön tilavuus on pohjan pinta-alan $\|\bar{v} \times \bar{w}\|$ ja korkeuden h tulo (ks. kuva 13.45). Pohjan pinta-alan tiedetään edellisen kappaleen perusteella olevan $\|\bar{v} \times \bar{w}\|$. Määritetään vielä korkeus h. Olkoon β vektoreiden $\bar{u}$ ja $\bar{v} \times \bar{w}$ välinen kulma. Nyt

$$h = \|\bar{u}\| \left|\cos(180^\circ - \beta)\right| = \|\bar{u}\| \left|\cos \beta\right|.$$

Siten tilavuus on

$$\|\bar{v} \times \bar{w}\| \|\bar{u}\| \left|\cos \beta\right| = \Big| \|\bar{v} \times \bar{w}\| \|\bar{u}\| \cos \beta \Big| = \Big| (\bar{v} \times \bar{w}) \cdot \bar{u} \Big|.$$

Viimeisessä välivaiheessa käytettiin vektorien $\bar{u}$ ja $\bar{v} \times \bar{w}$ välisen kulman määritelmää. Suunnikkaan tilavuus on siis niin kutsutun *skalaarikolmitulon* itseisarvo.



Kuva 13.45: Vektoreiden $\bar{v}$, $\bar{w}$ ja $\bar{u}$ määräämään suuntaissärmiön tilavuus.

Luku 2

Vektoriavaruudet

14 Vektoriavaruus

Kurssin ensimmäisessä osassa käsiteltiin avaruuden $\mathbb{R}^n$ vektoreita. Nyt määrittelemme abstraktimmat avaruuden ja vektorin käsitteet, jotka ovat avaruuden $\mathbb{R}^n$ ja sen vektoreiden yleistyksiä. Lähtökohtana ovat lauseessa 2.5 esitetyt vektorien laskusäännöt. Tulemme yleistämään myös muita avaruuteen $\mathbb{R}^n$ liittyvia käsitteitä.

Määritelmä 14.1. Oletetaan, että joukossa V on määritelty yhteenlasku ja skalaarikertolasku jollakin tavalla. Jos alla listatut ehdot pätevät kaikilla $\bar{v}, \bar{w}, \bar{u} \in V$ ja $a, b \in \mathbb{R}$, niin joukkoa V kutsutaan *vektoriavaruudeksi* ja sen alkioita *vektoreiksi*.

1) $\bar{v} + \bar{w} = \bar{w} + \bar{v}$ kaikilla $\bar{v}, \bar{w} \in V$.
2) $(\bar{v} + \bar{w}) + \bar{u} = \bar{v} + (\bar{w} + \bar{u})$ kaikilla $\bar{v}, \bar{w}, \bar{u} \in V$.
3) On olemassa niin kutsuttu *nollavektori* $\bar{0}$, jolle pätee $\bar{v} + \bar{0} = \bar{v}$ kaikilla $\bar{v} \in V$.
4) Jokaisella vektorilla $\bar{v} \in V$ on niin kutsuttu *vastavektori* $-\bar{v}$, jolle pätee $\bar{v} + (-\bar{v}) = \bar{0}$.
5) $a(\bar{v} + \bar{w}) = a\bar{v} + a\bar{w}$ kaikilla $\bar{v}, \bar{w} \in V$ ja $a \in \mathbb{R}$.
6) $(a + b)\bar{v} = a\bar{v} + b\bar{v}$ kaikilla $\bar{v} \in V$ ja $a, b \in \mathbb{R}$.
7) $(ab)\bar{v} = a(b\bar{v})$ kaikilla $\bar{v} \in V$ ja $a, b \in \mathbb{R}$.
8) $1\bar{v} = \bar{v}$ kaikilla $\bar{v} \in V$.

Vektoriavaruuden määritelmässä vaaditaan, että yhteenlasku ja skalaarikertolasku on määritelty joukossa V: jos $\bar{v}$, $\bar{w} \in V$ ja $a \in \mathbb{R}$, niin täytyy päteä $\bar{v} + \bar{w} \in V$ ja $a\bar{v} \in V$.

Skalaari tarkoittaa tällä kurssilla reaalilukua, sillä käsittelemme reaalikertoimisia vektoriavaruuksia. Kompleksikertoimisilla vektoriavaruuksilla skalaarit ovat kompleksilukuja. Periaatteessa skalaarit voivat olla minkä tahansa *kunnan* alkioita. (Kunnista kerrotaan lisää kurssilla Algebra I.)

Tarpeen tullen vektoriavaruuden V nollavektoria voidaan merkitä $\bar{0}_V$. Tällöin ei tule sekaannusta siitä, minkä vektoriavaruuden nollavektorista on kyse.

Esimerkki 14.2. Kaikki vektoriavaruuden määritelmän ehdot pätevät lauseen 2.5 perusteella avaruuden $\mathbb{R}^n$ yhteenlaskulle ja skalaarikertolaskulle. Siten $\mathbb{R}^n$ on vektoriavaruus. Vektoriavaruuden käsite siis tosiaan yleistää avaruutta R^n.

Myös reaalilukujen joukko $\mathbb{R}$ on vektoriavaruus, kun yhteenlaskuna on reaalilukujen yhteenlasku ja skalaarikertolaskuna reaalilukujen kertolasku.

Esimerkki 14.3. Kokonaislukujen joukko $\mathbb{Z}$ varustettuna tavallisella yhteenlaskulla ja skalaarikertolaskulla (reaaliluvulla kertominen) ei ole vektoriavaruus. Tämä johtuu siitä, että esimerkiksi $0{,}5 \in \mathbb{R}$ ja $3 \in \mathbb{Z}$, mutta $0{,}5 \cdot 3 = 1{,}5 \notin \mathbb{Z}$. Skalaarikertolaskun tulos ei siis välttämättä ole joukossa $\mathbb{Z}$.

Vektoriavaruuden määritelmä ei kerro, mitä otuksia vektoriavaruuden alkiot ovat. Se ei myöskään sano, miltä yhteenlasku ja skalaarikertolasku näyttävät. Ne saattavat olla tuttuja laskutoimituksia mutta myös jotain aivan muuta.

Esimerkki 14.4. Vektoriavaruuden alkiot voivat olla vaikkapa matriiseja. Matriiseja voidaan nimittäin laskea yhteen ja niitä voidaan kertoa reaaliluvuilla. Lisäksi seuraavat säännöt pätevät $m \times n$-matriiseille A, B ja C sekä reaaliluvuille a ja b:

1) $A + B = B + A$
2) $A + (B + C) = (A + B) + C$
3) $A + O = A$
4) $A + (-A) = O$
5) $a(A + B) = aA + aB$
6) $(a + b)A = aA + bA$
7) $(ab)A = a(bA)$
8) $1A = A$.

(Osa näistä ehdoista on todettu lauseessa 9.3 ja loppujen todistaminen on suoraviivaista.) Kaikkien $m \times n$-matriisien joukko $\mathbb{R}^{m\times n}$ on siis vektoriavaruus. Nollavektori on nollamatriisi O, ja matriisin vastavektori saadaan muuttamalla kaikkien matriisin alkioiden merkki.

Esimerkki 14.5. Vektoriavaruuden alkiot voivat olla myös kuvauksia. Olkoon $\mathcal{F}$ kaikkien kuvausten $\mathbb{R} \to \mathbb{R}$ joukko. Jos $f \in \mathcal{F}$, $g \in \mathcal{F}$ ja $a \in \mathbb{R}$, niin kuvaukset $f + g$ ja af määritellään seuraavasti:

$$\begin{aligned} f + g\colon \mathbb{R} \to \mathbb{R}, &\quad x \mapsto f(x) + g(x) \quad \text{ja} \\ af\colon \mathbb{R} \to \mathbb{R}, &\quad x \mapsto af(x). \end{aligned}$$

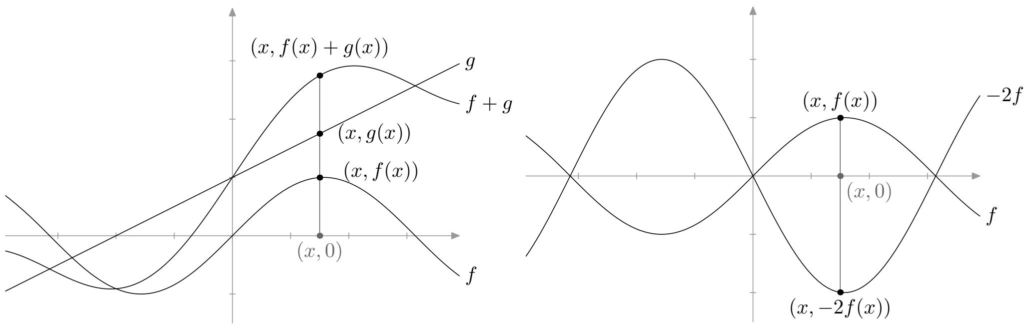
Sanotaan, että funktioiden laskutoimitukset on tällöin määritelty *pisteittäin*.

Tarkastellaan esimerkiksi funktioita

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) = \sin x \quad \text{ja} \quad g\colon \mathbb{R} \to \mathbb{R},\ g(x) = 0{,}5x + 1.$$

Nyt funktiot $f + g$ ja $(-2)f$ näyttävät seuraavilta:

$$f + g\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \sin x + 0{,}5x + 1 \quad \text{ja} \quad (-2)f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto -2\sin x.$$



Kuva 14.1: Funktiot f ja g sekä niiden summa $f+g$ ja skalaarimonikerta $(-2)f$.

Joukko $\mathcal{F}$, jossa yhteenlasku ja skalaarikertolasku määritellään pisteittäin, on vektoriavaruus. Tämä osoitetaan käymällä läpi vektoriavaruuden määritelmän ehdot. Seuraavassa osoitetaan osa ehdoista. Loppujen ehtojen tarkistaminen jätetään harjoitustehtäväksi.

1) Oletetaan, että $f, g \in \mathcal{F}$, ja osoitetaan, että $f+g = g+f$. Olkoon $x \in \mathbb{R}$. Kuvausten yhteenlaskun määritelmän mukaan

$$(f+g)(x) = f(x) + g(x) \quad \text{ja}$$
$$(g+f)(x) = g(x) + f(x).$$

Kuvausten f ja g arvot $f(x)$ ja $g(x)$ ovat reaalilukuja, joten $f(x)+g(x) = g(x)+f(x)$. Näin ollen

$$(f+g)(x) = (g+f)(x).$$

Kuvauksilla $f+g$ ja $g+f$ on siis samat arvot, joten ne ovat sama kuvaus. Toisin sanoen

$$f+g = g+f.$$

3) Osoitetaan, että nollavektoriksi kelpaa kuvaus $f_0 \colon \mathbb{R} \to \mathbb{R}$, joka määritellään kaavalla $f_0(x) = 0$ kaikilla $x \in \mathbb{R}$. Oletetaan, että $g \in \mathcal{F}$, ja osoitetaan, että $g + f_0 = g$. Olkoon $x \in \mathbb{R}$. Kuvausten yhteenlaskun määritelmän mukaan

$$(g+f_0)(x) = g(x) + f_0(x) = g(x) + 0 = g(x)\,.$$

Kuvauksilla $g + f_0$ ja g on siis samat arvot, joten

$$g + f_0 = g.$$

4) Osoitetaan, että kuvauksen $g \in \mathcal{F}$ vastavektoriksi kelpaa kuvaus $g' \colon \mathbb{R} \to \mathbb{R}$, joka määritellään kaavalla $x \mapsto -g(x)$ kaikilla $x \in \mathbb{R}$. Osoitetaan siis, että $g + g' = f_0$. Olkoon $x \in \mathbb{R}$. Kuvausten yhteenlaskun määritelmän mukaan

$$\begin{aligned}(g+g')(x) &= g(x) + g'(x) \\ &= g(x) + (-g(x)) = 0 = f_0(x).\end{aligned}$$

Kuvauksilla $g + g'$ ja f_0 on siis samat arvot, joten

$$g + g' = f_0.$$

6) Oletetaan, että $f \in \mathcal{F}$ ja $a,\ b \in \mathbb{R}$, ja osoitetaan, että $(a+b)f = af + bf$. Olkoon $x \in \mathbb{R}$. Kuvausten skalaarikertolaskun ja yhteenlaskun määritelmien mukaan

$$\begin{aligned}\big((a+b)f\big)(x) &= (a+b)f(x) \quad \text{ja} \\ (af+bf)(x) &= (af)(x) + (bf)(x) = af(x) + bf(x).\end{aligned}$$

Kuvauksen f arvo $f(x)$ on reaaliluku, joten $(a+b)f(x) = af(x) + bf(x)$. Näin ollen

$$\big((a+b)f\big)(x) = (af+bf)(x).$$

Kuvauksilla $(a+b)f$ ja $af + bf$ on siis samat arvot, joten

$$(a+b)f = af + bf.$$

Esimerkki 14.6. Myös polynomit muodostavat vektoriavaruuksia. Reaalikertoiminen *polynomi* on muotoa

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

oleva summa, missä $n \in \mathbb{N}$ ja $a_n, \ldots, a_0 \in \mathbb{R}$. Lukuja $a_n, \ldots, a_0$ kutsutaan polynomin *kertoimiksi* ja symbolia x polynomin *tuntemattomaksi*. Summattavat $a_i x^i$ ovat polynomin *termejä*.

Polynomeille voidaan määritellä yhteenlasku, jossa toisiaan vastaavien termien kertoimet lasketaan yhteen. Esimerkiksi polynomien

$$p = 3x^2 - 4x + 10 \qquad \text{ja} \qquad q = -2x^5 - x^3 + 5x^2 + 4x$$

summa on polynomi $p + q = -2x^5 - x^3 + 8x^2 + 10$. Skalaarikertolaskussa puolestaan kukin polynomin kerroin kerrotaan reaaliluvulla. Esimerkiksi polynomi $(-3)p$ saadaan kertomalla kaikki polynomin p kertoimet luvulla -3:

$$(-3)p = -9x^2 + 12x - 30.$$

Voidaan osoittaa, että reaalikertoimisten polynomien joukko muodostaa vektoriavaruuden. Tätä vektoriavaruutta merkitään symbolilla $\mathcal{P}$.

Esimerkki 14.7. Määritellään joukossa $\mathbb{R}^2$ skalaarikertolasku $*$ seuraavasti: jos $(v_1, v_2) \in \mathbb{R}^2$ ja $a \in \mathbb{R}$, niin

$$a * (v_1, v_2) = (av_1, 0)\,.$$

Osoitetaan, että joukko $\mathbb{R}^2$ varustettuna tavallisella yhteenlaskulla $+$ ja skalaarikertolaskulla $*$ ei ole vektoriavaruus.

Havaitaan, että esimerkiksi

$$1 * (5, 9) = (5, 0).$$

Näin ollen

$$1 * (5, 9) \neq (5, 9),$$

joten vektoriavaruuden määritelmän ehto (8) ei täyty.

Esimerkki 14.8. Määritellään positiivisten reaalilukujen joukossa

$$\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$$

yhteenlasku $\oplus$ ja skalaarikertolasku $\odot$ seuraavasti: jos x, $y \in \mathbb{R}_+$ ja $c \in \mathbb{R}$, niin

$$x \oplus y = x \cdot y \quad \text{ja} \quad c \odot x = x^c.$$

Voidaan osoittaa, että tällöin $\mathbb{R}_+$ on vektoriavaruus. Tätä vektoriavaruutta tutkitaan lisää harjoitustehtävissä.

Lause 14.9. *Oletetaan, että V on vektoriavaruus. Tällöin*

a) nollavektoreita on täsmälleen yksi

b) jokaisella vektorilla $\bar{v} \in V$ on täsmälleen yksi vastavektori.

Todistus. a) Oletetaan, että vektoriavaruudessa V on kaksi nollavektoria, $\bar{0}$ ja $\bar{0}'$. Nyt siis pätee $\bar{v} + \bar{0} = \bar{v}$ ja $\bar{v} + \bar{0}' = \bar{v}$ kaikilla $\bar{v} \in V$. Tarkastellaan nyt vektoria $\bar{a} = \bar{0} + \bar{0}'$. Ensinnäkin pätee $\bar{a} = \bar{0}' + \bar{0} = \bar{0}'$, sillä $\bar{0}$ on nollavektori. Toisaalta pätee myös $\bar{a} = \bar{0} + \bar{0}' = \bar{0}$, sillä $\bar{0}'$ on nollavektori. Näin on osoitettu, että $\bar{0} = \bar{0}'$.

b) Oletetaan, että $\bar{v} \in V$. Oletetaan lisäksi, että $\bar{u}$ ja $\bar{w}$ ovat kumpikin vektorin $\bar{v}$ vastavektoreita eli $\bar{v} + \bar{u} = \bar{0}$ ja $\bar{v} + \bar{w} = \bar{0}$. Tällöin

$$\bar{u} = \bar{u} + \bar{0} = \bar{u} + (\bar{v} + \bar{w}) = (\bar{u} + \bar{v}) + \bar{w} = (\bar{v} + \bar{u}) + \bar{w} = \bar{0} + \bar{w} = \bar{w}. \qquad \square$$

Lause 14.10. *Oletetaan, että V on vektoriavaruus ja $\bar{v} \in V$, $a \in \mathbb{R}$. Tällöin*

a) $0\bar{v} = \bar{0}$

b) $a\bar{0} = \bar{0}$

c) $(-1)\bar{v} = -\bar{v}$

d) jos $a\bar{v} = \bar{0}$, niin $a = 0$ tai $\bar{v} = \bar{0}$ (tulon nollasääntö).

Todistus. Osoitetaan kohdat b) ja d) ja jätetään loput kohdat harjoitustehtäviksi.

b) Nähdään, että
$$a\bar{0} = a(\bar{0} + \bar{0}) = a\bar{0} + a\bar{0}.$$
Lisäämällä tämän yhtälön molemmille puolille vektori $-(a\bar{0})$ saadaan
$$\bar{0} = a\bar{0}.$$

d) Oletetaan, että $a\bar{v} = \bar{0}$. On osoitettava, että tästä seuraa $a = 0$ tai $\bar{v} = \bar{0}$. Tutkitaan kahta tapausta. Oletetaan ensin, että $a \neq 0$. Tällöin on olemassa käänteisluku $1/a$, ja voimme kertoa yhtälön $a\bar{v} = \bar{0}$ molemmat puolet tällä käänteisluvulla. Näin saadaan yhtälö $(1/a)(a\bar{v}) = (1/a)\bar{0}$. Tämän yhtälön vasen puoli on
$$\frac{1}{a}(a\bar{v}) = \left(\frac{1}{a}a\right)\bar{v} = 1\bar{v} = \bar{v}.$$
Oikea puoli on puolestaan kohdan a) perusteella $(1/a)\bar{0} = \bar{0}$. Näin ollen $\bar{v} = \bar{0}$, joten väite pätee silloin, kun $a \neq 0$. Jos taas $a = 0$, on selvää, että väite pätee. □

Edellinen lause osoittaa, että avaruudesta $\mathbb{R}^n$ tutut laskusäännöt pätevät myös yleisemmissä vektoriavaruuksissa. Myös erotuksen ja lineaarikombinaation käsitteet voidaan määritellä tutulla tavalla.

Määritelmä 14.11. Oletetaan, että V on vektoriavaruus ja $\bar{v}, \bar{w} \in V$. Vektoreiden $\bar{v}$ ja $\bar{w}$ *erotus* $\bar{v} - \bar{w}$ tarkoittaa summaa $\bar{v} + (-\bar{w})$.

Määritelmä 14.12. Oletetaan, että V on vektoriavaruus ja $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k \in V$. Vektoreiden $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k$ *lineaarikombinaatio* on vektori
$$a_1\bar{v}_1 + a_2\bar{v}_2 + \cdots + a_k\bar{v}_k,$$
missä $a_1, a_2, \dots, a_k \in \mathbb{R}$.

15 Aliavaruus

Kurssin ensimmäisessä osassa määrittelimme avaruuden $\mathbb{R}^n$ vektorien virittämän aliavaruuden. Nyt esittelemme yleisemmän aliavaruuden käsitteen.

Määritelmä 15.1. Olkoon V vektoriavaruus. Sen osajoukko W on vektoriavaruuden V *aliavaruus*, jos seuraavat ehdot pätevät:

a) $\bar{w} + \bar{u} \in W$ kaikilla $\bar{w}, \bar{u} \in W$

b) $r\bar{w} \in W$ kaikilla $r \in \mathbb{R}$ ja $\bar{w} \in W$

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

Lauseen 4.4 nojalla avaruuden $\mathbb{R}^n$ vektoreiden $\bar{v}_1, \dots, \bar{v}_k$ virittämä aliavaruus $\text{span}(\bar{v}_1, \dots, \bar{v}_k)$ on vektoriavaruuden $\mathbb{R}^n$ aliavaruus.

Esimerkki 15.2. Osoitetaan, että joukko

$$W = \{(a, b, a) \mid a, b \in \mathbb{R}\}$$

on vektoriavaruuden $\mathbb{R}^3$ aliavaruus. Joukko W muodostuu siis sellaisista vektoreista, joiden ensimmäinen ja viimeinen komponentti ovat samat.

Joukko W on määritelmänsä mukaan vektoriavaruuden $\mathbb{R}^3$ osajoukko.

a) Oletetaan, että $\bar{w}, \bar{u} \in W$. Nyt $\bar{w} = (a, b, a)$ ja $\bar{u} = (c, d, c)$ joillakin reaaliluvuilla $a, b, c, d \in \mathbb{R}$. Huomataan, että

$$\bar{w} + \bar{u} = (a + c, b + d, a + c).$$

Koska summavektorin ensimmäinen ja viimeinen komponentti ovat samat, on se joukon W alkio. Siten $\bar{w} + \bar{u} \in W$.

b) Oletetaan, että $\bar{w} \in W$ ja $r \in \mathbb{R}$. Nyt $\bar{w} = (a, b, a)$ joillakin $a, b \in \mathbb{R}$. Nähdään, että $r\bar{w} = (ra, rb, ra)$. Vektorin $r\bar{w}$ ensimmäinen ja viimeinen komponentti ovat samat, joten pätee $r\bar{w} \in W$.

c) Nollavektori $(0, 0, 0)$ on joukon W alkio, sillä sen ensimmäinen ja viimeinen komponentti ovat samat.

Siten W on vektoriavaruuden $\mathbb{R}^3$ aliavaruus.

Esimerkki 15.3. Tutkitaan sitten erästä polynomiavaruuden $\mathcal{P}$ aliavaruutta. Sitä varten on määriteltävä polynomin aste. Olkoon $p = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ polynomi, jolle pätee $a_n \neq 0$. Lukua n kutsutaan polynomin *asteeksi* ja merkitään $\deg(p)$.

Osoitetaan, että polynomiavaruudella $\mathcal{P}$ on aliavaruus

$$\mathcal{P}_2 = \{p \in \mathcal{P} \mid p = 0 \text{ tai } \deg(p) \leq 2\}.$$

Oletetaan, että $p, q \in \mathcal{P}_2$ ja $r \in \mathbb{R}$. Nyt $p = a_2x^2 + a_1x + a_0$ ja $q = b_2x^2 + b_1x + b_0$ joillakin $a_2, a_1, a_0, b_2, b_1, b_0 \in \mathbb{R}$. Huomataan, että

$$p + q = (a_2 + b_2)x^2 + (a_1 + b_1)x + a_0 + b_0,$$

joten polynomin $p + q$ aste on korkeintaan kaksi. Siis $p + q \in \mathcal{P}_2$. Lisäksi

$$rp = (ra_2)x^2 + (ra_1)x + ra_0,$$

joten myös polynomin rp aste on korkeintaan kaksi. Näin ollen $rp \in \mathcal{P}_2$. Vektoriavaruuden $\mathcal{P}$ nollavektori on nollapolynomi 0, joka kuuluu määritelmän mukaan joukkoon $\mathcal{P}_2$. Siten $\mathcal{P}_2$ on vektoriavaruuden $\mathcal{P}$ aliavaruus.

Samalla tavoin voidaan osoittaa, että joukko

$$\mathcal{P}_n = \{p \in \mathcal{P} \mid p = 0 \text{ tai } \deg(p) \leq n\}$$

on vektoriavaruuden $\mathcal{P}$ aliavaruus kaikilla $n \in \mathbb{N}$.

Esimerkki 15.4. Tarkastellaan $n \times n$-matriisien muodostamaa vektoriavaruutta $\mathbb{R}^{n\times n}$. Olkoon W symmetristen $n \times n$-matriisien joukko. Toisin sanoen

$$W = \{C \in \mathbb{R}^{n\times n} \mid C^T = C\}.$$

Osoitetaan, että W on vektoriavaruuden $\mathbb{R}^{n\times n}$ aliavaruus.

Ensinnäkin W on määritelmänsä mukaan joukon $\mathbb{R}^{n\times n}$ osajoukko. Oletetaan, että A, $B \in W$ ja $c \in \mathbb{R}$. Tällöin $A^T = A$ ja $B^T = B$.

Transpoosin laskusääntöjen nojalla

$$(A + B)^T = A^T + B^T = A + B,$$

joten $A + B \in W$. Lisäksi

$$(cA)^T = cA^T = cA,$$

joten $cA \in W$. Nollavektori on $n \times n$-nollamatriisi O. Sille pätee

$$O^T = O,$$

joten $O \in W$.

Näin ollen W on vektoriavaruuden $\mathbb{R}^{n\times n}$ aliavaruus.

Esimerkki 15.5. Tutkitaan, onko joukko

$$W = \left\{ \begin{bmatrix} a & a+1 \\ 0 & b \end{bmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}$$

vektoriavaruuden $\mathbb{R}^{2\times 2}$ aliavaruus.

Havaitaan, että nollavektori eli nollamatriisi

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

ei ole joukossa W, joten aliavaruuden määritelmän ehto c) ei täyty. Siis W ei ole vektoriavaruuden $\mathbb{R}^{2\times 2}$ aliavaruus.

Esimerkki 15.6. Tutkitaan, onko joukko $W = \{A \in \mathbb{R}^{2\times 2} \mid \det(A) = 0\}$ vektoriavaruuden $\mathbb{R}^{2\times 2}$ aliavaruus.

Valitaan esimerkiksi

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

Tällöin $\det(A) = 0$ ja $\det(B) = 0$, joten $A, B \in W$.

Kuitenkin

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

ja siten $\det(A+B) = 2 \neq 0$. Näin ollen $A+B \notin W$, joten W ei ole vektoriavaruuden $\mathbb{R}^{2\times 2}$ aliavaruus.

Seuraava lause osoittaa, että jokainen aliavaruus on itsekin pieni vektoriavaruus.

Lause 15.7. *Oletetaan, että V on vektoriavaruus, jolla on aliavaruus W. Tällöin myös aliavaruus W on vektoriavaruus.*

Todistus. Vektoriavaruuden yhteenlaskua ja skalaarikertolaskua koskevat ehdot 1)–2) ja 5)–8) pysyvät voimassa, vaikka rajoitutaan tarkastelemaan alkuperäisen vektoriavaruuden V osajoukkoa W.

Nollavektoria käsittelevä ehto 3) seuraa aliavaruuden määritelmästä, sillä nollavektori kuuluu aina aliavaruuteen. Vastavektoriin liittyvä ehto 4) puolestaan seuraa aliavaruuden määritelmän ehdosta b) sekä lauseen 14.10 kohdasta c). Jos nimittäin $\bar{v} \in W$, niin $-\bar{v} = (-1)\bar{v} \in W$. Siten jokaisella W:n vektorilla on vastavektori joukossa W.

Aliavaruuden määritelmän ehdot a) ja b) takaavat, että yhteenlasku ja skalaarikertolasku ovat joukon W laskutoimituksia. □

15.1 Vektoreiden virittämä aliavaruus

Määritelmä 15.8. Olkoon V jokin vektoriavaruus. *Vektoreiden $\bar{v}_1, \ldots, \bar{v}_k \in V$ virittämä aliavaruus* on joukko

$$\operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k) = \{a_1\bar{v}_1 + \cdots + a_k\bar{v}_k \mid a_1, \ldots, a_k \in \mathbb{R}\}.$$

Esimerkki 15.9. Esimerkissä 15.2 osoitettiin, että $W = \{(r, s, r) \mid r, s \in \mathbb{R}\}$ on vektoriavaruuden $\mathbb{R}^3$ aliavaruus. Etsitään sille virittäjävektorit.

Havaitaan, että

$$\begin{aligned} W &= \{(a, b, a) \mid a, b \in \mathbb{R}\} \\ &= \{a(1, 0, 1) + b(0, 1, 0) \mid a, b \in \mathbb{R}\} \\ &= \operatorname{span}\big((1, 0, 1), (0, 1, 0)\big). \end{aligned}$$

Siis W on vektoreiden $(1, 0, 1)$ ja $(0, 1, 0)$ virittämä vektoriavaruuden $\mathbb{R}^3$ aliavaruus.

Esimerkki 15.10. Merkitään

$$W = \left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \,\middle|\, a, b, c \in \mathbb{R} \right\}.$$

Osoitetaan, että W on 2×2-matriisien muodostaman vektoriavaruuden $\mathbb{R}^{2\times 2}$ aliavaruus. Tehdään tämä etsimällä W:lle virittäjävektorit.

Havaitaan, että

$$\begin{aligned} W &= \left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \,\middle|\, a, b, c \in \mathbb{R} \right\} \\ &= \left\{ a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \,\middle|\, a, b, c \in \mathbb{R} \right\} \\ &= \operatorname{span}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right). \end{aligned}$$

Siis W on vektoreiden (eli tässä tapauksessa matriisien)

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

virittämä vektoriavaruuden $\mathbb{R}^{2\times 2}$ aliavaruus.

Vektorien virittämä aliavaruus on aliavaruus myös määritelmän 15.1 mielessä. Tämä osoitetaan täsmällisesti seuraavassa lauseessa.

Lause 15.11. *Jos* $\bar{v}_1, \ldots, \bar{v}_k \in V$, *niin* $\operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$ *on vektoriavaruuden* V *aliavaruus. Lisäksi* $\operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$ *on pienin aliavaruus, joka sisältää vektorit* $\bar{v}_1, \ldots, \bar{v}_k$.

Todistus. Ensinnäkin $\operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$ on vektoriavaruuden V osajoukko, sillä se koostuu V:n vektorien lineaarikombinaatioista.

Oletetaan, että $\bar{u}$, $\bar{w} \in \operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$ ja $c \in \mathbb{R}$. Tällöin

$$\bar{u} = a_1\bar{v}_1 + \cdots + a_k\bar{v}_k \quad \text{ja} \quad \bar{w} = b_1\bar{v}_1 + \cdots + b_k\bar{v}_k$$

joillakin $a_1, \ldots, a_k$, $b_1, \ldots, b_k \in \mathbb{R}$. Nähdään, että

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

joten $\bar{u} + \bar{w} \in \operatorname{span}(\bar{v}_1, \ldots, \bar{v}_k)$. Lisäksi

$$c\bar{u} = c(a_1\bar{v}_1 + \cdots + a_k\bar{v}_k) = (ca_1)\bar{v}_1 + \cdots + (ca_k)\bar{v}_k,$$

joten $c\bar{u} \in \mathrm{span}(\bar{v}_1, \ldots, \bar{v}_k)$. Nollavektori voidaan lauseen 14.10 a)-kohdan nojalla kirjoittaa muodossa

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

joten $\bar{0} \in \mathrm{span}(\bar{v}_1, \ldots, \bar{v}_k)$. Siten $\mathrm{span}(\bar{v}_1, \ldots, \bar{v}_k)$ on vektoriavaruuden V aliavaruus.

Vektorit $\bar{v}_1, \ldots, \bar{v}_k$ kuuluvat aliavaruuteen V, sillä

$$\begin{aligned}
\bar{v}_1 &= 1\bar{v}_1 + 0\bar{v}_2 + \cdots + 0\bar{v}_k \\
\bar{v}_2 &= 0\bar{v}_1 + 1\bar{v}_2 + \cdots + 0\bar{v}_k \\
&\vdots \\
\bar{v}_k &= 0\bar{v}_1 + 0\bar{v}_2 + \cdots + 1\bar{v}_k.
\end{aligned}$$

Tarvitsee enää osoittaa, että $\mathrm{span}(\bar{v}_1, \ldots, \bar{v}_k)$ on pienin aliavaruus, joka sisältää vektorit $\bar{v}_1, \ldots, \bar{v}_k$. Oletetaan, että W on vektoriavaruuden V jokin sellainen aliavaruus, että $\bar{v}_1, \ldots, \bar{v}_k \in W$. Koska W on aliavaruus, se sisältää kaikkien vektoriensa summat ja skalaarimonikerrat. Siis $a_1\bar{v}_1 + \cdots + a_k\bar{v}_k \in W$ kaikilla $a_1, \ldots, a_k \in \mathbb{R}$. Näin ollen $\mathrm{span}(\bar{v}_1, \ldots, \bar{v}_k) \subset W$. □

Esimerkki 15.12. Osoitetaan, että 2×2-matriiseista muodostuva vektoriavaruus $\mathbb{R}^{2\times 2}$ on seuraavien vektoreiden virittämä:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{ja} \quad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Oletetaan, että

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Tällöin

$$A = a_{11}E_{11} + a_{12}E_{12} + a_{21}E_{21} + a_{22}E_{22},$$

joten A on vektoreiden E_{11}, E_{12}, E_{21} ja E_{22} lineaarikombinaatio. Siten on osoitettu, että $\mathbb{R}^{2\times 2} = \mathrm{span}(E_{11}, E_{12}, E_{21}, E_{22})$.

Esimerkki 15.13. Polynomit 1 ja x virittävät polynomiavaruuden

$$\mathcal{P}_1 = \{p \in \mathcal{P} \mid p = 0 \text{ tai } \deg(p) \leq 1\}.$$

Jos nimittäin $p \in \mathcal{P}_1$, niin $p = ax + b = ax + b \cdot 1$ joillakin $a, b \in \mathbb{R}$. Siten p on polynomien x ja 1 lineaarikombinaatio.

Samalla tavoin voidaan osoittaa, että $\mathcal{P}_n = \mathrm{span}(1, x, x^2, \ldots, x^n)$.

Esimerkki 15.14. Merkitään

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{ja} \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Määritetään $\operatorname{span}(A, B, I)$.

Jokainen vektoreiden (matriisien) A, B ja I lineaarikombinaatio on muotoa

$$xA + yB + zI = \begin{bmatrix} x+z & x+y \\ x+y & z \end{bmatrix},$$

missä x, y, $z \in \mathbb{R}$. Havaitaan, että tällainen lineaarikombinaatio on symmetrinen matriisi. Siten

$$\operatorname{span}(A, B, I) \subset \{\, C \in \mathbb{R}^{2\times 2} \mid C^T = C \,\}.$$

Osoitetaan sitten, että jokainen symmetrinen matriisi voidaan kirjoittaa vektoreiden A, B ja I lineaarikombinaationa. Oletetaan, että C on symmetrinen matriisi. Tällöin

$$C = \begin{bmatrix} d & e \\ e & f \end{bmatrix},$$

missä d, e, $f \in \mathbb{R}$. Pienen pohdiskelun jälkeen havaitaan, että

$$\begin{aligned} C = \begin{bmatrix} d & e \\ e & f \end{bmatrix} &= (d-f)\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + (e-d+f)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + f\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= (d-f)A + (e-d+f)B + fI. \end{aligned}$$

Siis jokainen symmetrinen matriisi on vektoreiden A, B ja I lineaarikombinaatio. Tämä tarkoittaa sitä, että $\{C \in \mathbb{R}^{2\times 2} \mid C^T = C\} \subset \operatorname{span}(A, B, I)$.

Näin ollen $\operatorname{span}(A, B, I) = \{C \in \mathbb{R}^{2\times 2} \mid C^T = C\}$.

Laajennetaan lopuksi virittämisen määritelmää hieman. Määritelmässä 15.8 puhutaan yhden tai useamman vektorin virittämistä aliavaruuksista. Toisinaan halutaan ottaa huomioon myös tapaus, jossa virittäjävektoreita ei ole yhtään. Sovimme, että nollan vektorin virittämä aliavaruus on $\{\bar{0}\}$.

Lisäksi virittämisen määritelmää voidaan laajentaa koskemaan myös äärettömiä vektorijoukkoja. Aliavaruus $\operatorname{span}(S)$, missä S on äärettömän monen vektorin muodostama joukko, koostuu kaikista (äärellisistä) lineaarikombinaatioista, jotka voidaan muodostaa joukon S vektoreista. Esimerkiksi kaikkien polynomien muodostama vektoriavaruus $\mathcal{P}$ on vektoreiden $1, x, x^2, \ldots$ virittämä.

16 Vapaus

Kurssin ensimmäisessä osassa käsiteltiin avaruuden $\mathbb{R}^n$ vapaita vektorijonoja. Tämä käsite voidaan yleistää mihin tahansa vektoriavaruuteen.

Määritelmä 16.1. Vektoriavaruuden V vektoreista muodostuva jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on *vapaa*, jos seuraava ehto pätee:

jos $c_1\bar{v}_1 + c_2\bar{v}_2 + \cdots + c_k\bar{v}_k = \bar{0}$ joillakin $c_1, \ldots, c_k \in \mathbb{R}$,
niin $c_1 = 0, c_2 = 0, \ldots, c_k = 0$.

Jos jono ei ole vapaa, sanotaan, että se on *sidottu*. Vapaata jonoa voidaan kutsua myös *lineaarisesti riippumattomaksi* ja sidottua *lineaarisesti riippuvaksi*.

Tyhjä jono on jono, jossa on ei ole yhtään vektoria. Sovimme, että tyhjä jono on vapaa.

Esimerkki 16.2. Esimerkissä 15.12 määriteltiin avaruuden $\mathbb{R}^2$ matriisit E_{11}, E_{12}, E_{21} ja E_{22}. Osoitetaan, että jono $(E_{11}, E_{12}, E_{21}, E_{22})$ on vapaa. Oletetaan, että luvut $c_1, c_2, c_3, c_4 \in \mathbb{R}$ ovat sellaisia, että

$$c_1E_{11} + c_2E_{12} + c_3E_{21} + c_4E_{22} = O.$$

(Tässä tapauksessa nollavektori on nollamatriisi O.) Yhtälön vasen puoli saadaan muotoon

$$\begin{bmatrix} c_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & c_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c_3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c_4 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}.$$

Nyt siis

$$\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

mistä seuraa, että $c_1 = 0$, $c_2 = 0$, $c_3 = 0$ ja $c_4 = 0$. Siten jono $(E_{11}, E_{12}, E_{21}, E_{22})$ on vapaa.

Esimerkki 16.3. Osoitetaan, että vektoriavaruuden $\mathcal{P}_n$ jono $(1, x, x^2, \ldots, x^n)$ on vapaa. Oletetaan, että $c_0, c_1, \ldots, c_n \in \mathbb{R}$ ovat sellaisia, että

$$c_0 1 + c_1x + c_2x^2 + \cdots + c_nx^n = 0.$$

(Yhtälön oikealla puolella on avaruuden $\mathcal{P}_n$ nollavektori eli nollapolynomi.) Kaksi polynomia ovat samat, jos ja vain jos niiden kertoimet ovat samat. Täytyy siis päteä $c_0 = 0, c_1 = 0, \ldots, c_n = 0$. Siten jono $(1, x, x^2, \ldots, x^n)$ on vapaa.

Vapauden määritelmän mukaan jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on sidottu, jos ja vain jos on olemassa sellaiset kertoimet $c_1, \ldots, c_k \in \mathbb{R}$, että $c_1\bar{v}_1 + c_2\bar{v}_2 + \cdots + c_k\bar{v}_k = \bar{0}$ ja jokin kertoimista $c_1, \ldots, c_k$ ei ole nolla.

Toisinaan jono on helppo osoittaa sidotuksi keksimällä sopivat kertoimet. Esimerkiksi vektoriavaruuden V jono $(\bar{v}_1, \bar{v}_1, \bar{v}_2)$ on sidottu, sillä

$$1\bar{v}_1 + (-1)\bar{v}_1 + 0\bar{v}_2 = \bar{0}.$$

Seuraavaksi osoitamme vapauteen liittyviä lauseita. Monet tuloksista olivat esillä jo luvussa 7 avaruuden $\mathbb{R}^n$ vektoreille. Yleisessä tapauksessa todistukset ovat hyvin samanlaisia, joten niitä ei esitetä tässä.

Seuraava lause osoittaa, että vektorien vapaus takaa yksikäsitteisen esityksen.

Lause 16.4. *Oletetaan, että V on vektoriavaruus ja $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k \in V$. Jono $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ on vapaa, jos ja vain jos jokainen aliavaruuden* $\text{span}(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ *alkio voidaan kirjoittaa täsmälleen yhdellä tavalla vektorien $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k$ lineaarikombinaationa.*

Todistus. Todistus on samanlainen kuin lauseen 7.6 todistus. □

Vektorijono on sidottu, jos ja vain jos jokin sen vektoreista voidaan ilmaista toisten lineaarikombinaationa.

Lause 16.5. *Oletetaan, että V on vektoriavaruus, $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k \in V$ ja $n \geq 2$. Jono $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_k)$ on sidottu, jos ja vain jos vain jos*

$$\bar{v}_j \in \text{span}(\bar{v}_1, \dots, \bar{v}_{j-1}, \bar{v}_{k+1}, \dots, \bar{v}_k)$$

jollakin $j \in \{1, 2, \dots, k\}$.

Todistus. Todistus on samanlainen kuin lauseen 7.7 todistus. □

Vapaasta jonosta voidaan tietyin ehdoin muodostaa vielä pidempi vapaa jono.

Lause 16.6. *Oletetaan, että vektoriavaruuden V jono $(\bar{v}_1, \dots, \bar{v}_k)$ on vapaa. Oletetaan lisäksi, että $\bar{w} \in V$. Tällöin jono $(\bar{v}_1, \dots, \bar{v}_k, \bar{w})$ on vapaa, jos ja vain jos*

$$\bar{w} \notin \text{span}(\bar{v}_1, \dots, \bar{v}_k).$$

Todistus. "$\Rightarrow$" Oletetaan, että jono $(\bar{v}_1, \dots, \bar{v}_k, \bar{w})$ on vapaa. On osoitettava, että $\bar{w} \notin \text{span}(\bar{v}_1, \dots, \bar{v}_k)$. Oletetaan vastoin väitettä, että $\bar{w} \in \text{span}(\bar{v}_1, \dots, \bar{v}_k)$. Tällöin

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

joillakin $a_1, \dots, a_k \in \mathbb{R}$. Nyt $a_1\bar{v}_1 + \cdots + a_k\bar{v}_k + (-1)\bar{w} = \bar{0}$, joten jono $(\bar{v}_1, \dots, \bar{v}_k, \bar{w})$ ei ole vapaa. Tämä on ristiriita. Siten $\bar{w} \notin \text{span}(\bar{v}_1, \dots, \bar{v}_k)$.

"$\Leftarrow$" Oletetaan, että $\bar{w} \notin \text{span}(\bar{v}_1, \dots, \bar{v}_k)$, ja osoitetaan, että jono $(\bar{v}_1, \dots, \bar{v}_k, \bar{w})$ on vapaa. Oletetaan, että

$$c_1\bar{v}_1 + \cdots + c_k\bar{v}_k + c_{k+1}\bar{w} = \bar{0}$$

joillakin $c_1, \dots, c_{k+1} \in \mathbb{R}$. Jos $c_{k+1} \neq 0$, niin

$$\bar{w} = \frac{c_1}{c_{k+1}}\bar{v}_1 + \cdots + \frac{c_k}{c_{k+1}}\bar{v}_k.$$

Nyt siis $\bar{w} \in \text{span}(\bar{v}_1, \dots, \bar{v}_k)$. Tämä on kuitenkin vastoin oletusta, joten täytyy päteä $c_{k+1} = 0$. Tällöin

$$c_1\bar{v}_1 + \cdots + c_k\bar{v}_k = \bar{0}.$$

Koska jono $(\bar{v}_1, \dots, \bar{v}_k)$ on vapaa, tiedetään, että $c_1 = 0, \dots, c_k = 0$. Koska myös kerroin c_{k+1} on nolla, on jono $(\bar{v}_1, \dots, \bar{v}_k, \bar{w})$ vapaa. □

Vapaan jonon jokainen osajono on vapaa.

Lause 16.7. *Oletetaan, että vektoriavaruuden V jono jono $\mathcal{S} = (\bar{v}_1, \ldots, \bar{v}_k)$ on vapaa. Tällöin jokainen jonon $\mathcal{S}$ osajono on myöskin vapaa.*

Todistus. Osajono tarkoittaa jonoa, joka saadaan poistamalla alkuperäisestä jonosta vektoreita. Myös jono itse on yksi osajonoista.

Oletetaan, että vektorijono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa. Osoitetaan, että mikä tahansa sen osajono on vapaa. Jos osajono on tyhjä jono, se on sopimuksemme mukaan vapaa. Tutkitaan sitten epätyhjiä jonoja. Koska vapautta tutkittaessa vektoreiden järjestyksellä ei ole väliä, riittää osoittaa, että jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m)$ on vapaa kaikilla $m \in \{1, \ldots, k\}$.

Oletetaan siis, että $m \in \{1, \ldots, k\}$. Olkoot luvut $c_1, \ldots, c_m \in \mathbb{R}$ sellaisia, että

$$c_1\bar{v}_1 + c_2\bar{v}_2 + \cdots + c_m\bar{v}_m = \bar{0}.$$

Tästä seuraa, että

$$c_1\bar{v}_1 + c_2\bar{v}_2 + \cdots + c_m\bar{v}_m + 0\bar{v}_{m+1} + \cdots + 0\bar{v}_k = \bar{0}.$$

Koska jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vapaa, täytyy yllä olevassa lineaarikombinaatiossa kaikkien kertoimien olla nollia. Siis $c_1 = 0, \ldots, c_m = 0$. Siten jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m)$ on vapaa. □

Vapauden määritelmässä käsitellään vain äärellisiä vektorijonoja. Määritelmää voidaan kuitenkin laajentaa koskemaan myös äärettömän monen vektorin muodostamia jonoja samalla tavalla kuin virittämisen tapauksessa. Vektoriavaruuden V jono $(\bar{v}_1, \bar{v}_2, \ldots)$ on vapaa, jos sen kaikki äärelliset osajonot ovat vapaita. Esimerkiksi polynomiavaruuden $\mathcal{P}$ jono $(1, x, x^2, \ldots)$ on vapaa.

17 Kanta

Määritelmä 17.1. Oletetaan, että $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k \in V$. Jono $(\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k)$ on vektoriavaruuden V kanta, jos

a) $V = \operatorname{span}(\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_k)$

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

Esimerkki 17.2. Esimerkeissä 15.12 ja 16.2 osoitettiin, että matriisiavaruuden $\mathbb{R}^{2\times2}$ jono $(E_{11}, E_{12}, E_{21}, E_{22})$ virittää avaruuden $\mathbb{R}^{2\times2}$ ja on lisäksi vapaa. Siten jono $(E_{11}, E_{12}, E_{21}, E_{22})$ on avaruuden $\mathbb{R}^{2\times2}$ kanta.

Esimerkki 17.3. Polynomiavaruuden $\mathcal{P}_n$ jono $(1, x, x^2, \ldots, x^n)$ virittää avaruuden $\mathcal{P}_n$ ja on lisäksi vapaa esimerkkien 15.13 ja 16.3 perusteella. Jono $(1, x, x^2, \ldots, x^n)$ on siis avaruuden $\mathcal{P}_n$ kanta.

Vektoriavaruuden vektorit voidaan kirjoittaa täsmälleen yhdellä tavalla kannan vektoreiden lineaarikombinaatioina.

Lause 17.4. *Jono $\mathcal{B} = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on vektoriavaruuden V kanta, jos ja vain jos jokainen vektorivaruuden V alkio voidaan kirjoittaa täsmälleen yhdellä tavalla vektorien $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ lineaarikombinaationa.*

Todistus. Väite seuraa lähes suoraan lauseesta 16.4 samalla tavalla kuin vastaava avaruutta $\mathbb{R}^n$ koskeva lause 8.3. □

Määritelmä 17.5. Olkoon $\mathcal{B} = (\bar{v}_1, \ldots, \bar{v}_n)$ on vektoriavaruuden V kanta. Oletetaan, että $\bar{u} \in V$. Vektorin $\bar{u}$ *koordinaateiksi kannan $\mathcal{B}$ suhteen* kutsutaan reaalilukuja $a_1, \ldots, a_n$, joilla

$$\bar{u} = a_1\bar{v}_1 + \cdots + a_n\bar{v}_n.$$

Esimerkki 17.6. Määritetään avaruuden $\mathbb{R}^{2\times 2}$ matriisin

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

koordinaatit kannan $(E_{11}, E_{12}, E_{21}, E_{22})$ suhteen. Huomataan, että

$$A = 1E_{11} + 2E_{12} + (-1)E_{21} + 0E_{22},$$

joten kooridinaatit ovat 1, 2, -1 ja 0.

Polynomiavaruuden $\mathcal{P}_3$ alkion $x^3 - 4x + 2$ koordinaatit kannan $(1, x, x^2, x^3)$ suhteen taas ovat 2, -4, 0 ja 1.

Lause 17.7. *Oletetaan, että V on vektoriavaruus, jolla on kanta $(\bar{v}_1, \ldots, \bar{v}_n)$. Jos vektoriavaruuden V vektorijonon pituus on suurempi kuin n, kyseinen jono ei voi olla vapaa.*

Todistus. Väite on aikaisemmin todistettu avaruudelle $\mathbb{R}^n$. (Ks. korollaari 7.11.) Yleisessä tapauksessa todistus on samanlainen. □

Aiemmin mainittiin, että voidaan puhua myös äärettömän monen vektorin virittämistä aliavaruuksista sekä äärettömän monen vektorin muodostamista vapaista jonoista. Siksi myös kannan määritelmä voidaan yleistää koskemaan äärettömiä jonoja. Esimerkiksi jono $(1, x, x^2, \ldots)$ on polynomiavaruuden $\mathcal{P}$ kanta.

17.1 Dimensio

Vektoriavaruudella voi olla useita eri kantoja. Saman avaruuden eri kannoissa on kuitenkin aina yhtä monta vektoria.

Lause 17.8. *Vektoriavaruuden jokaisessa kannassa on yhtä monta vektoria.*

Todistus. Oletetaan vastoin väitettä, että vektoriavaruudella V on kaksi eripituista kantaa $(\bar{v}_1, \dots, \bar{v}_k)$ ja $(\bar{w}_1, \dots, \bar{w}_m)$, missä $k > m$. Koska jono $(\bar{v}_1, \dots, \bar{v}_k)$ on kanta, se on vapaa. Nyt päädytään ristiriitaan, sillä lauseen 17.7 nojalla vektoriavaruudessa ei voi olla vapaata jonoa, joka on pitempi kuin jokin vektoriavaruuden kanta. Siten väite on todistettu. □

Edellisen lauseen perusteella voidaan määritellä vektoriavaruuden dimensio.

Määritelmä 17.9. Vektoriavaruus V on *äärellisulotteinen*, jos sillä on äärellisen monesta vektorista koostuva kanta. Tällöin avaruuden *dimensio* $\dim(V)$ on kannan vektoreiden lukumäärä. Jos vektoriavaruudella ei ole äärellistä kantaa, avaruus on *ääretönulotteinen* ja sen dimensio on ääretön.

Jos vektoriavaruuden dimensio on n, on tapana sanoa, että vektoriavaruus on *n-ulotteinen.*

Esimerkki 17.10. Matriisiavaruuden $\mathbb{R}^{2\times 2}$ dimensio on neljä, sillä avaruudella on kanta $(E_{11}, E_{12}, E_{21}, E_{22})$. Polynomiavaruuden $\mathcal{P}_n$ dimensio on $n+1$, sillä avaruudella on kanta $(1, x, x^2, \dots, x^n)$.

Esimerkki 17.11. Vektoriavaruuden $\{\bar{0}\}$ dimensio on nolla, sillä sen kanta on tyhjä jono, jossa on nolla vektoria. Olemme nimittäin sopineet, että tyhjän jonon virittämä avaruus on $\{\bar{0}\}$ ja että tyhjä jono on vapaa.

Vapaa jono voidaan jatkaa vektoriavaruuden kannaksi.

Lause 17.12. *Oletetaan, että V äärellisulotteinen vektoriavaruus. Oletetaan lisäksi, että $\mathcal{S} = (\bar{w}_1, \dots, \bar{w}_k)$ on avaruuden V vapaa jono. Tällöin jonoon $\mathcal{S}$ voidaan lisätä vektoreita niin, että tuloksena on avaruuden V kanta.*

Todistus. Tarkastellaan kaikkia sellaisia avaruuden V jonoja $(\bar{u}_1, \dots, \bar{u}_m)$, joilla jono $(\bar{w}_1, \dots, \bar{w}_k, \bar{u}_1, \dots, \bar{u}_m)$ on vapaa. Valitaan näistä jonoista jokin sellainen, jonka pituus on mahdollisimman suuri. (Tällainen on olemassa, sillä lauseen 17.7 nojalla minkään vapaan jonon pituus ei voi olla suurempi kuin avaruuden V ulottuvuus.) Olkoon tuo mahdollisimman pitkä jono $\mathcal{B} = (\bar{w}_1, \dots, \bar{w}_k, \bar{a}_1, \dots, \bar{a}_r)$.

Osoitetaan, että jono $\mathcal{B}$ on avaruuden V kanta. Koska $\mathcal{B}$ on vapaa, riittää osoittaa, että $\mathcal{B}$ virittää avaruuden V. Oletetaan sitä varten, että $\bar{v} \in V$. Jos $\bar{v} \notin \operatorname{span}(\mathcal{B})$, niin lauseen 16.6 nojalla jono $(\bar{w}_1, \dots, \bar{w}_k, \bar{a}_1, \dots, \bar{a}_r, \bar{v})$ on vapaa. Tämä on ristiriita, sillä näin saatu jono on pidempi kuin $\mathcal{B}$, jonka pituus oli suurin mahdollinen. Siten $\bar{v} \in \operatorname{span}(\mathcal{B})$, mistä seuraa, että $V = \operatorname{span}(\mathcal{B})$. Näin on osoitettu, että $\mathcal{B} = (\bar{w}_1, \dots, \bar{w}_k, \bar{a}_1, \dots, \bar{a}_r)$ on avaruuden V kanta. □

Virittäjäjono voidaan lyhentää vektoriavaruuden kannaksi.

Lause 17.13. *Oletetaan, että jono $\mathcal{S} = (\bar{w}_1, \dots, \bar{w}_k)$ virittää avaruuden V. Tällöin jonosta $\mathcal{S}$ voidaan poistaa vektoreita niin, että tuloksena on avaruuden V kanta.*

Todistus. Todistus on hyvin samankaltainen kuin edellisen lauseen todistus. Tarkastellaan kaikkia sellaisia jonon $(\bar{w}_1, \dots, \bar{w}_k)$ osajonoja, jotka ovat vapaita. (Tällaisia jonoja on olemassa, sillä esimerkiksi tyhjä jono on vapaa.) Valitaan näistä jonoista jokin sellainen, jonka pituus on mahdollisimman suuri. Olkoon tuo mahdollisimman pitkä vapaa osajono $\mathcal{B} = (\bar{b}_1, \dots, \bar{b}_r)$.

Osoitetaan, että $\mathcal{B}$ on avaruuden V kanta. Koska $\mathcal{B}$ on vapaa, riittää osoittaa, että $\mathcal{B}$ virittää avaruuden V. Tämä puolestaan on todistettu, jos voidaan osoittaa, että $\bar{w} \in \operatorname{span}(\mathcal{B})$ kaikilla $\bar{w} \in \mathcal{S}$. Oletetaan vastoin väitettä, että $\bar{w} \notin \operatorname{span}(\mathcal{B})$ jollakin $\bar{w} \in \mathcal{S}$. Nyt lauseen 16.6 nojalla jono $(\bar{b}_1, \dots, \bar{b}_r, \bar{w})$ on vapaa. Tämä on ristiriita, sillä näin saatu osajono on pidempi kuin $\mathcal{B}$, jonka pituus oli suurin mahdollinen. Siten $\bar{w} \in \operatorname{span}(\mathcal{B})$, mistä seuraa, että $V = \operatorname{span}(\mathcal{B})$. Näin on osoitettu, että $\mathcal{B} = (\bar{b}_1, \dots, \bar{b}_r)$ on avaruuden V kanta. □

Lause 17.14. *Oletetaan, että V on vektoriavaruus, jonka dimensio on n.*

a) Jokainen vektoriavaruuden V vapaa jono, jonka pituus on n, on V:n kanta.

b) Jokainen vektoriavaruuden V virittäjäjono, jonka pituus on n, on V:n kanta.

Todistus. a) Oletetaan, että avaruuden V jono $(\bar{v}_1, \dots, \bar{v}_n)$ on vapaa. Nyt lauseen 17.12 nojalla jonoon voidaan lisätä vektoreita niin, että saadaan aikaan kanta. Kuitenkin vektoriavaruuden V dimensio on n, joten kannassa on oltava n vektoria. Jonon $(\bar{v}_1, \dots, \bar{v}_n)$ täytyy siis jo olla kanta.

b) Oletetaan, että jono $(\bar{v}_1, \dots, \bar{v}_n)$ virittää avaruuden V. Nyt lauseen 17.13 nojalla jonosta voidaan ottaa pois vektoreita niin, että saadaan aikaan kanta. Kuitenkin vektoriavaruuden kannassa on oltava n vektoria. Jonon $(\bar{v}_1, \dots, \bar{v}_n)$ täytyy siis jo olla kanta. □

Edellä osoitetuista lauseista on hyötyä, kun tutkitaan, onko annettu jono vektoriavaruuden kanta. Kootaan vielä kaikki tiedot yhteen. Oletetaan, että V on vektoriavaruus, jonka dimensio on n.

- Jos vektoriavaruuden V vektorijonon pituus on suurempi kuin n, kyseinen jono ei voi olla vapaa.
- Jos vektoriavaruuden V vektorijonon pituus on pienempi kuin n, kyseinen jono ei voi virittää avaruutta V.
- Jos vektoriavaruuden V vektorijono virittää avaruuden ja sen pituus on n, kyseinen jono on kanta.
- Jos vektoriavaruuden V vektorijono on vapaa ja sen pituus on n, kyseinen jono on kanta.

Aliavaruuden dimensio ei voi olla suurempi kuin itse vektoriavaruuden dimensio.

Lause 17.15. *Oletetaan, että V on äärellisulotteinen vektoriavaruus, jolla on aliavaruus W. Tällöin myös W on äärellisulotteinen ja* $\dim(W) \le \dim(V)$. *Lisäksi* $\dim(W) = \dim(V)$, *jos ja vain jos $W = V$.*

Todistus. Osoitetaan ensin, että aliavaruudella W on kanta. Tarkastellaan kaikkia aliavaruuden W vapaita jonoja. Koska ne ovat myös vektoriavaruuden V vapaita jonoja, niiden pituus on lauseen 17.7 nojalla pienempi kuin $\dim(V)$. Valitaan jono, jonka pituus on suurin mahdollinen. Samaan tapaan kuin lauseen 17.12 todistuksessa voidaan osoittaa, että näin valittu jono on aliavaruuden W kanta. Kanta on siis olemassa. Tästä nähdään myös, että aliavaruudella W on äärellinen kanta ja tuon kannan pituus on väistämättä lyhyempi kuin $\dim(V)$. Siten W on äärellisulotteinen, ja $\dim(W) \leq \dim(V)$.

Osoitetaan vielä väitteen toinen osa. Jos $V = W$, niin $\dim(W) = \dim(V)$. Oletetaan sitten, että $\dim(W) = \dim(V)$, ja osoitetaan, että $W = V$. Olkoon $\mathcal{B}$ aliavaruuden W kanta. Nyt $\mathcal{B}$ on avaruuden V vapaa jono, jonka pituus on sama kuin avaruuden V dimensio. Lauseen 17.14 perusteella $\mathcal{B}$ on avaruuden V kanta. Tästä seuraa, että $V = \text{span}(\mathcal{B}) = W$.

□

18 Lineaarikuvaus

Määritelmä 18.1. Olkoot V ja U vektoriavaruuksia. Kuvaus $L\colon V \to U$ on *lineaarikuvaus*, jos seuraavat ehdot pätevät:

a) $L(\bar{v} + \bar{w}) = L(\bar{v}) + L(\bar{w})$ kaikilla $\bar{v}, \bar{w} \in V$

b) $L(c\bar{v}) = cL(\bar{v})$ kaikilla $c \in \mathbb{R}$ ja $\bar{v} \in V$.

Jos kuvaus L on lineaarikuvaus, voidaan myös sanoa, että L on *lineaarinen*. Merkintä $L\colon V \to U$ tarkoittaa, että vektoriavaruus V on kuvauksen L lähtöavaruus. Vektoriavaruus U on puolestaan kuvauksen L maaliavaruus.

Esimerkki 18.2. Tarkastellaan kuvausta $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = 3x$. Osoitetaan, että f on lineaarikuvaus.

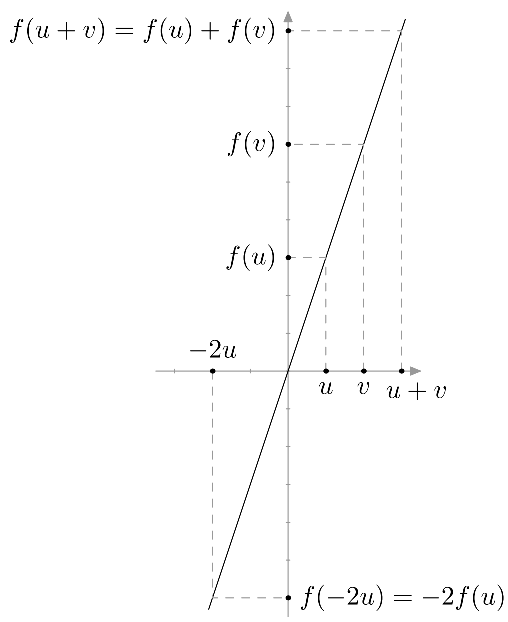
Oletetaan, että $v, u \in \mathbb{R}$ ja $c \in \mathbb{R}$. Tällöin

$$f(v + u) = 3(v + u) = 3v + 3u = f(v) + f(u)$$

ja

$$f(cv) = 3(cv) = c(3v) = cf(v).$$

Kuvaus f täyttää siis lineaarikuvauksen määritelmän ehdot.



Kuva 18.2: Kuvaus f täyttää lineaarikuvauksen määritelmän ehdot.

Esimerkki 18.3. Tarkastellaan kuvausta $g\colon \mathbb{R} \to \mathbb{R}$, $g(x) = x^3 - 2x + 1$. Osoitetaan, että g ei ole lineaarikuvaus. Tämän osoittamiseen riittää yksi tapaus, jossa lineaarikuvauksen ehto ei toteudu.

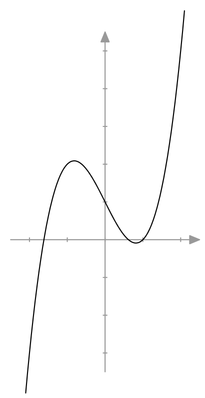
Valitaan esimerkiksi $v = -1$ ja $w = 2$. Tällöin

$$g(v + w) = g(1) = 0,$$

mutta

$$g(v) + g(w) = g(-1) + g(2) = 2 + 5 = 7.$$

Siis $g(-1 + 2) \neq g(-1) + g(2)$, joten g ei ole lineaarikuvaus.



Kuva 18.3: Kuvaus g ei ole lineaarikuvaus.

Tällä kurssilla on tärkeää tehdä ero kuvauksen ja kuvauksen arvon välillä. Edellisessä esimerkissä määriteltiin kuvaus g kirjoittamalla $g\colon \mathbb{R} \to \mathbb{R}$, $g(x) = x^3 - 2x + 1$. Merkintä $g\colon \mathbb{R} \to \mathbb{R}$ on tässä oleellinen. Jos kirjoitetaan pelkästään $g(x) = x^3 - 2x + 1$, tarkoitetaan kuvauksen g arvoa jossakin yksittäisessä pisteessä x. Samalla tavalla on tehtävä ero merkintöjen g ja $g(x)$ välillä. Ensimmäinen tarkoittaa kuvausta ja toinen kuvauksen arvoa pisteessä x.

Esimerkki 18.4. Osoitetaan, että kuvaus

$$L\colon \mathbb{R}^3 \to \mathbb{R}^2, \quad L(x_1, x_2, x_3) = (7x_2,\ x_1 - 3x_3)$$

on lineaarikuvaus. (Tässä vektorin (x_1, x_2, x_3) kuva pitäisi tarkalleen ottaen kirjoittaa muodossa $L((x_1, x_2, x_3))$ mutta on yleisesti sovittu, että toiset sulut saa tässä jättää pois.)

Oletetaan, että $\bar{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$, $\bar{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$ ja $c \in \mathbb{R}$.

a) Huomataan, että

$$\begin{aligned} L(\bar{v} + \bar{w}) &= L(v_1 + w_1, v_2 + w_2, v_3 + w_3) = (7(v_2 + w_2),\ (v_1 + w_1) - 3(v_3 + w_3)) \\ &= (7v_2 + 7w_2,\ v_1 + w_1 - 3v_3 - 3w_3). \end{aligned}$$

Toisaalta

$$\begin{aligned} L(\bar{v}) + L(\bar{w}) &= (7v_2,\ v_1 - 3v_3) + (7w_2,\ w_1 - 3w_3) \\ &= (7v_2 + 7w_2,\ v_1 + w_1 - 3v_3 - 3w_3), \end{aligned}$$

joten $L(\bar{v} + \bar{w}) = L(\bar{v}) + L(\bar{w})$.

b) Nähdään, että

$$\begin{aligned} L(c\bar{v}) &= L(cv_1, cv_2, cv_3) = (7cv_2,\ cv_1 - 3cv_3) \\ &= c(7v_2,\ v_1 - 3v_2) = cL(\bar{v}). \end{aligned}$$

Siten L on lineaarikuvaus.

Esimerkki 18.5. Kuvaus $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L(x_1, x_2) = (x_1x_2,\ x_1)$ ei ole lineaarikuvaus. Huomataan nimittäin, että

$$L((1,1) + (1,0)) = L(2,1) = (0,2)$$

ja

$$L(1,1) + L(1,0) = (1,1) + (0,1) = (1,2) \neq (0,2).$$

Siis $L((1,1) + (1,0)) \neq L(1,1) + L(1,0)$, eikä kuvaus siten ole lineaarikuvaus.

Vaihtoehtoisesti voidaan myös tarkastella lineaarikuvauksen jälkimmäistä ehtoa ja osoittaa, että se ei päde. Huomataan, että $L(2(1,1)) = L(2,2) = (4,2)$ ja $2L(1,1) = 2(1,1) = (2,2)$. Siten $L(2(1,1)) \neq 2L(1,1)$, eikä kuvaus ole lineaarikuvaus.

Esimerkki 18.6. Osoitetaan, että kuvaus $L\colon \mathbb{R}^2 \to \mathcal{P}_1$, $L(a,b) = ax + b$ on lineaarikuvaus. Oletetaan, että $(a,b), (c,d) \in \mathbb{R}^2$ ja $r \in \mathbb{R}$. Nyt

$$\begin{aligned} L((a,b) + (c,d)) &= L(a+c, b+d) = (a+c)x + (b+d) \\ &= ax + b + cx + d = L(a,b) + L(c,d) \end{aligned}$$

ja

$$L(r(a,b)) = L(ra, rb) = rax + rb = r(ax+b) = rL(a,b).$$

Siten L on lineaarikuvaus.

Esimerkki 18.7. Jos tiedetään virittäjävektorien arvot lineaarikuvauksessa, voidaan tämän avulla päätellä kaikkien muidenkin vektoreiden arvot.

Oletetaan, että $L\colon \mathbb{R}^2 \to \mathbb{R}^3$ on lineaarikuvaus, jolle pätee $L(1,0) = (1,4,5)$ ja $L(0,1) = (0,-1,-1)$. Määritetään tämän tiedon avulla vaikkapa vektorin $(-3,4)$ kuva:

$$\begin{aligned} L(-3,4) &= L(-3(1,0) + 4(0,1)) \\ &= L(-3(1,0)) + L(4(0,1)) \\ &= -3L(1,0) + 4L(0,1) \\ &= -3(1,4,5) + 4(0,-1,-1) \\ &= (-3,-16,-19). \end{aligned}$$

Mistä tahansa matriisista saadaan lineaarikuvaus.

Lause 18.8. *Oletetaan, että A on $m \times n$-matriisi. Matriisin A määräämä kuvaus $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$, $L_A(\bar{v}) = A\bar{v}$ on lineaarikuvaus.*

Tässä avaruuden $\mathbb{R}^n$ alkiot tulkitaan $n \times 1$-matriiseiksi.

Todistus. Oletetaan, että $\bar{v}, \bar{w} \in \mathbb{R}^n$ ja $c \in \mathbb{R}$. Nyt matriisien kertolaskun ominaisuuksien perusteella pätee

$$L_A(\bar{v} + \bar{w}) = A(\bar{v} + \bar{w}) = A\bar{v} + A\bar{w} = L_A(\bar{v}) + L_A(\bar{w})$$

ja

$$L_A(c\bar{v}) = A(c\bar{v}) = c(A\bar{v}) = cL_A(\bar{v}).$$

Siten L_A on lineaarinen. □

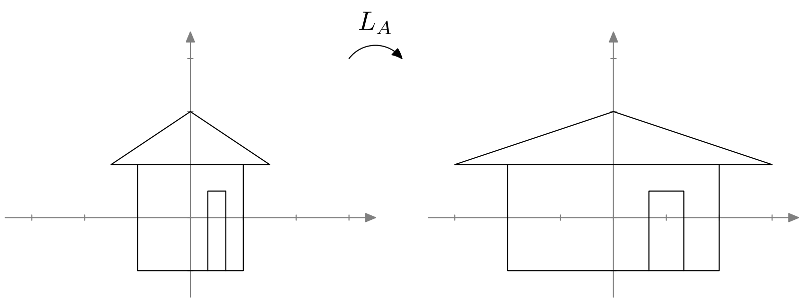
Esimerkki 18.9. Tutkitaan, millaisen lineaarikuvauksen antavat matriisit

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

Matriisista A saadaan kuvaus $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L_A(\bar{v}) = A\bar{v}$. Avaruuden $\mathbb{R}^2$ vektori (x_1, x_2) kuvautuu siis vektoriksi

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

Toisin sanoen $L_A(x_1, x_2) = (2x_1, x_2)$. Tästä nähdään, että kuvaus L_A *venyttää* vektoreita x_1-akselin suunnassa (ks. kuva 18.4).

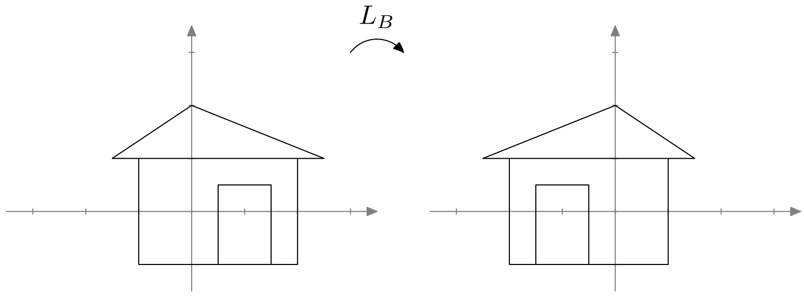


Kuva 18.4: Lineaarikuvaus L_A venyttää vektoreita x_1-akselin suunnassa.

Matriisista B saadaan puolestaan kuvaus $L_B\colon \mathbb{R}^2 \to \mathbb{R}^2$, jolle pätee

$$L_B\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}.$$

Kuvaus L_B *peilaa* vektorit x_2-akselin suhteen (ks. kuva 18.5).



Kuva 18.5: Lineaarikuvaus L_B peilaa vektorit x_2-akselin suhteen.

Matriisi C määrää kuvauksen $L_C\colon \mathbb{R}^2 \to \mathbb{R}^2$, jolle pätee

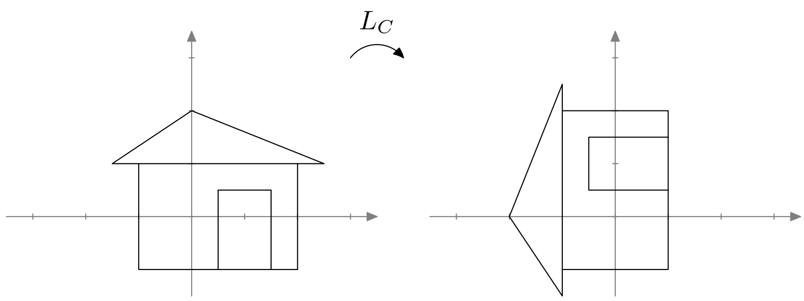
$$L_C\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$

Kuvaus L_C *kiertää* vektoreita 90° vastapäivään eli positiiviseen kiertosuuntaan (ks. kuva 18.6).

Voidaan osoittaa, että matriisin

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

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



Kuva 18.6: Lineaarikuvaus L_C kiertää vektoreita 90° positiiviseen kiertosuuntaan.

Tulemme myöhemmin osoittamaan, että jokainen lineaarikuvaus $\mathbb{R}^m \to \mathbb{R}^n$ on jonkin matriisin määräämä kuvaus. Seuraavat esimerkit antavat tästä esimakua.

Esimerkki 18.10. Tutkitaan kuvausta $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L(x_1, x_2) = (x_1, -x_2)$. Se peilaa vektorit x_1-akselin suhteen (ks. kuva 18.7). Kuvaus L on itse asiassa matriisin määräämä lineaarikuvaus eli kuvauksen L arvot saadaan kertomalla vektoreita jollakin matriisilla. Selvitetään tämä matriisi.

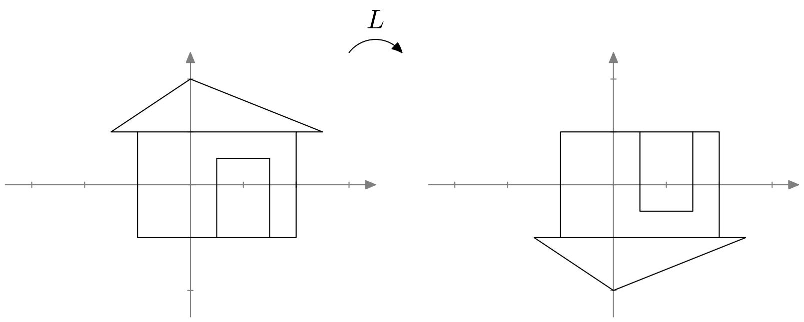
Kun tulkitaan avaruuden $\mathbb{R}^2$ vektori (x_1, x_2) matriisiksi, nähdään, että

$$L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix} = \begin{bmatrix} 1x_1 + 0x_2 \\ 0x_1 - x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Siten kuvaus L on matriisin

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

määräämä kuvaus, jolle pätee $L(\bar{x}) = D\bar{x}$ kaikilla $\bar{x} \in \mathbb{R}^2$. Koska L on matriisin määräämä kuvaus, se on lauseen 18.8 nojalla lineaarinen.



Kuva 18.7: Lineaarikuvaus L peilaa vektorit vaaka-akselin suhteen.

Esimerkki 18.11. Tarkastellaan kuvausta $P\colon \mathbb{R}^2 \to \mathbb{R}^2$, $P(\bar{v}) = \operatorname{proj}_{\bar{e}_1}(\bar{v})$, joka projisoi tason $\mathbb{R}^2$ vektorit vektorin $\bar{e}_1 = (1, 0)$ virittämälle aliavaruudelle (ks. kuva 18.8). Jos $\bar{v} = (v_1, v_2) \in \mathbb{R}^2$, niin

$$P(\bar{v}) = \frac{\bar{v} \cdot \bar{e}_1}{\bar{e}_1 \cdot \bar{e}_1}\bar{e}_1 = \frac{(v_1, v_2) \cdot (1, 0)}{(1, 0) \cdot (1, 0)}(1, 0) = v_1(1, 0) = (v_1, 0).$$

Toisin sanoen $P(v_1, v_2) = (v_1, 0)$.

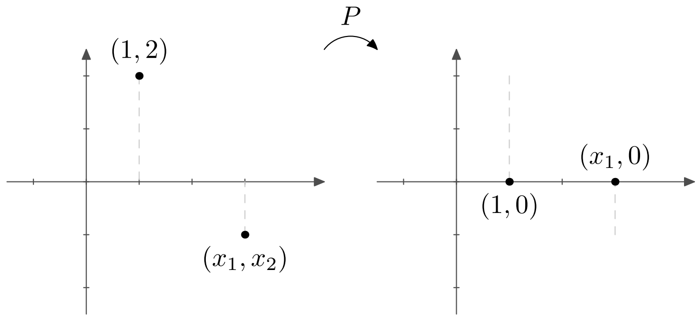
Osoitetaan, että kuvaus P on lineaarinen etsimällä matriisi, jonka määräämä lineaarikuvaus P on. Oletetaan, että $(x_1, x_2) \in \mathbb{R}^2$. Nyt

$$P\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1x_1 + 0x_2 \\ 0x_1 + 0x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

joten P on matriisin

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

määräämä lineaarikuvaus. Siten P on lineaarinen.



Kuva 18.8: Kuvaus P projisoi vektorit vaaka-akselille.

Lineaarikuvaus kuvaa nollavektorin aina nollavektoriksi.

Lause 18.12. *Jos $L\colon V \to U$ on lineaarikuvaus, niin $L(\bar{0}_V) = \bar{0}_U$.*

Todistus. Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Nyt

$$L(\bar{0}_V) = L(\bar{0}_V + \bar{0}_V) = L(\bar{0}_V) + L(\bar{0}_V).$$

Lisätään tämän yhtälön molemmille puolille avaruuden U vektori $-L(\bar{0}_V)$, jolloin saadaan

$$L(\bar{0}_V) - L(\bar{0}_V) = L(\bar{0}_V) + L(\bar{0}_V) - L(\bar{0}_V).$$

Nyt nähdään, että $\bar{0}_U = L(\bar{0}_V)$. Siten väite on todistettu. □

Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Jos $L(\bar{v}) = \bar{w}$ joillakin $\bar{v}, \bar{w} \in V$, voidaan merkitä $\bar{v} \mapsto \bar{w}$. Voidaan esimerkiksi kirjoittaa $0_V \mapsto 0_U$.

18.1 Lineaarikuvausten yhdistetyt kuvaukset

Oletetaan, että $L\colon V \to U$ ja $T\colon U \to W$ ovat lineaarikuvauksia. *Yhdistetty kuvaus*[1] $T \circ L$ tarkoittaa kuvausta $V \to W$, jolle pätee

$$(T \circ L)(\bar{v}) = T(L(\bar{v})) \quad \text{eli} \quad \bar{v} \mapsto T(L(\bar{v}))$$

kaikilla $\bar{v} \in V$.

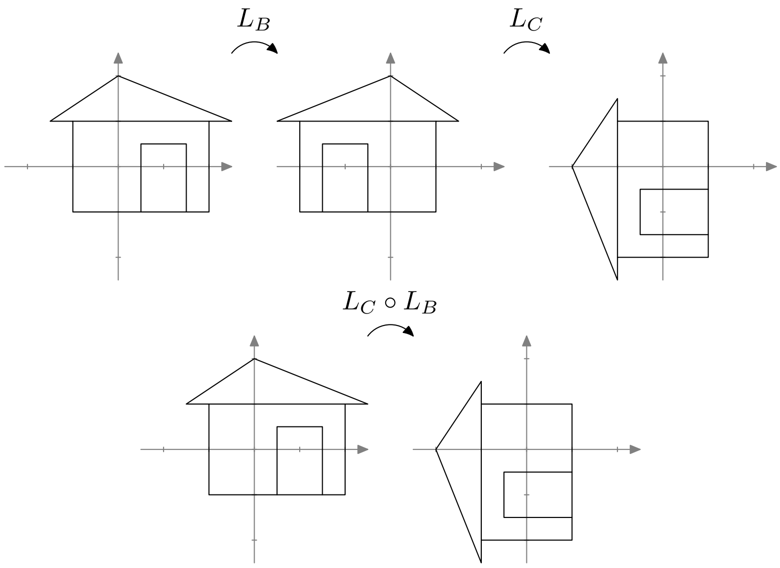
Esimerkki 18.13. Esimerkissä 18.9 esiteltiin peilaus $L_B\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L_B(x_1, x_2) = (-x_1, x_2)$ ja kierto $L_C\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L_C(x_1, x_2) = (-x_2, x_1)$. Koska kuvauksen L_B maalijoukko on sama kuin kuvauksen L_C lähtöjoukko, voidaan määritellä yhdistetty kuvaus $L_C \circ L_B$. Tämä kuvaus ensin peilaa vektorit x_2-akselin suhteen ja

[1] Yhdistetyistä kuvauksista voidaan puhua muidenkin kuvausten kuin lineaarikuvausten yhteydessä. Tällä kurssilla keskitytään kuitenkin lineaarikuvauksiin.

sen jälkeen kiertää niitä 90° vastapäivään (ks. kuva 18.9). Vektorin $(x_1, x_2) \in \mathbb{R}^2$ kuvavektori on

$$(L_C \circ L_B)(x_1, x_2) = L_C(L_B(x_1, x_2)) = L_C(-x_1, x_2) = (-x_2, -x_1).$$

Saadaan siis kuvaus $L_C \circ L_B \colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x_1, x_2) \mapsto (-x_2, -x_1)$.



Kuva 18.9: Kuvaukset L_B ja L_C voidaan yhdistää.

Lause 18.14. *Oletetaan, että* $L \colon V \to U$ *ja* $T \colon U \to W$ *ovat lineaarikuvauksia. Tällöin yhdistetty kuvaus* $T \circ L \colon V \to W$ *on lineaarinen.*

Todistus. Oletetaan, että $\bar{v}_1$, $\bar{v}_2 \in V$ ja $a \in \mathbb{R}$. Tarkistetaan lineaarikuvauksen määritelmän ehdot:

a) Yhdistetyn kuvauksen määritelmän ja kuvausten L ja T lineaarisuuden avulla saadaan

$$\begin{aligned}(T \circ L)(\bar{v}_1 + \bar{v}_2) &= T(L(\bar{v}_1 + \bar{v}_2)) \\ &= T(L(\bar{v}_1) + L(\bar{v}_2)) \\ &= T(L(\bar{v}_1)) + T(L(\bar{v}_2)) \\ &= (T \circ L)(\bar{v}_1) + (T \circ L)(\bar{v}_2)\end{aligned}$$

b) Nähdään, että

$$(T \circ L)(a\bar{v}_1) = T(L(a\bar{v}_1)) = T(aL(\bar{v}_1)) = aT(L(\bar{v}_1))) = a(T \circ L)(\bar{v}_1). \qquad \square$$

Jos kuvaukset ovat matriisien määräämiä lineaarikuvauksia, niiden yhdistäminen vastaa matriisien kertomista keskenään.

Lause 18.15. *Oletetaan, että A on $m \times n$-matriisi ja B on $p \times m$-matriisi. Tällöin*

$$L_B \circ L_A = L_{BA},$$

eli tulomatriisin BA määräämä kuvaus $L_{BA}\colon \mathbb{R}^n \to \mathbb{R}^p$ on sama kuvaus kuin yhdistetty kuvaus $L_B \circ L_A\colon \mathbb{R}^n \to \mathbb{R}^p$.

Todistus. Oletetaan, että $\bar{v} \in \mathbb{R}^n$. Tällöin matriisien laskusääntöjen mukaan

$$\begin{aligned} L_{BA}(\bar{v}) &= (BA)\bar{v} = B(A\bar{v}) = L_B(A\bar{v}) = L_B(L_A(\bar{v})) \\ &= (L_B \circ L_A)(\bar{v}). \end{aligned}$$

Siis $L_{BA}\colon \mathbb{R}^n \to \mathbb{R}^p$ ja $L_B \circ L_A\colon \mathbb{R}^n \to \mathbb{R}^p$ ovat sama kuvaus. □

Esimerkki 18.16. Esimerkissä 18.9 kuvaus L_B saatiin matriisista

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

ja kuvaus L_C matriisista

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

Näiden matriisien tulo on

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

Tämä tulomatriisi määrittää kuvauksen $\mathbb{R}^2 \to \mathbb{R}^2$, jolle pätee $(x_1, x_2) \mapsto (-x_2, -x_1)$. Toisaalta esimerkissä 18.13 nähtiin, että yhdistetylle kuvaukselle $L_C \circ L_B$ pätee $(x_1, x_2) \mapsto (-x_2, -x_1)$. Kyseessä on siis sama kuvaus aivan kuten lauseen 18.15 perusteella pitääkin olla.

18.2 Aliavaruuden kuva lineaarikuvauksessa

Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Avaruuden V aliavaruuden W *kuva*[2] kuvauksessa L on joukko

$$LW = \{L(\bar{w}) \mid \bar{w} \in W\}.$$

Aliavaruuden kuva koostuu siis kaikista niistä maaliavaruuden vektoreista, joille kuvautuu jokin aliavaruuden W vektori.

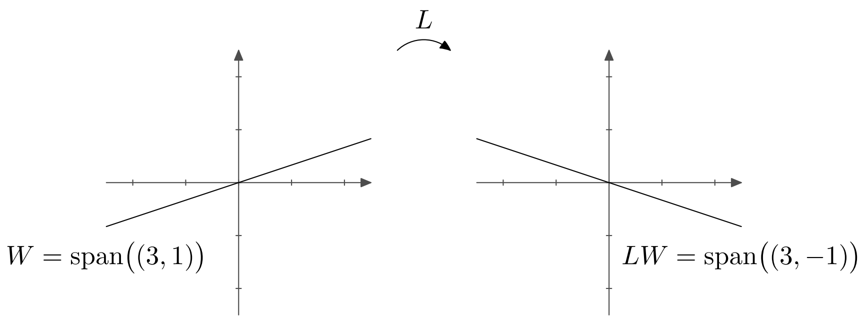
[2] Yleisesti kuvauksien yhteydessä voidaan puhua osajoukon kuvasta. Tällä kurssilla keskitytään kuitenkin lineaarikuvauksiin ja aliavaruuksien kuviin lineaarikuvauksissa.

Esimerkki 18.17. Tarkastellaan esimerkin 18.10 lineaarikuvausta $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L(x_1, x_2) = (x_1, -x_2)$, joka peilaa jokaisen pisteen x_1-akselin suhteen.

Olkoon $W = \operatorname{span}\big((3,1)\big)$. Tällöin W on vektorin $(3,1)$ virittämä aliavaruus (eli vektorin $\bar{w}$ suuntainen origon kautta kulkeva suora). Aliavaruuden W kuva on

$$\begin{aligned}
LW &= \{L(\bar{w}) \mid \bar{w} \in W\} \\
&= \{L(\bar{w}) \mid \bar{w} \in \operatorname{span}\big((3,1)\big)\} \\
&= \{L(\bar{w}) \mid \bar{w} = t(3,1) \text{ jollakin } t \in \mathbb{R}\} \\
&= \{L(t(3,1)) \mid t \in \mathbb{R}\} \\
&= \{tL(3,1) \mid t \in \mathbb{R}\} \\
&= \{t(3,-1) \mid t \in \mathbb{R}\} \\
&= \operatorname{span}\big((3,-1)\big).
\end{aligned}$$

Nähdään, että aliavaruuden W kuva on myöskin aliavaruus.



Kuva 18.10: Aliavaruus W ja sen kuva LW.

Lineaarikuvauksesssa aliavaruudet kuvautuvat aliavaruuksiksi.

Lause 18.18. *Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Jos W on avaruuden V aliavaruus, niin kuva*

$$LW = \{L(\bar{w}) \mid \bar{w} \in W\}$$

on avaruuden U aliavaruus.

Todistus. Lauseen todistus jätetään harjoitustehtäväksi. □

19 Ydin ja kuva

19.1 Lineaarikuvauksen ydin

Lineaarikuvauksen ydin koostuu kaikista niistä vektoreista, jotka kuvautuvat nollavektorille.

Määritelmä 19.1. Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Sen *ydin* on joukko

$$\operatorname{Ker} L = \{\bar{v} \in V \mid L(\bar{v}) = \bar{0}\}.$$

Huomaa, että ydin on aina joukko eikä jokin yksittäinen vektori. Monesti on olemassa useita vektoreita, jotka kuvautuvat nollavektorille. Ytimessä saattaa olla siis yksi alkio tai useampia alkioita. Lineaarikuvauksen ydin ei ole koskaan tyhjä joukko, sillä lauseen 18.12 mukaan nollavektori on aina kuvauksen ytimessä.

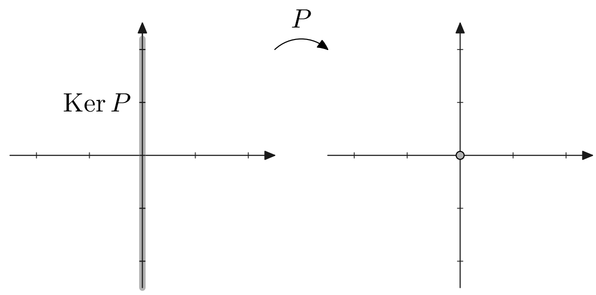
Merkitä Ker tulee englannin kielen sanasta ”kernel”, joka tarkoittaa ydintä.

Esimerkki 19.2. Tarkastellaan esimerkin 18.11 projektiokuvausta $P\colon \mathbb{R}^2 \to \mathbb{R}^2$, $P(x_1, x_2) = (x_1, 0)$. Esimerkiksi vektori $(0, \sqrt{5})$ on kuvauksen P ytimessä, sillä $P(0, \sqrt{5}) = (0, 0)$. Toisin sanoen $(0, \sqrt{5}) \in \operatorname{Ker} P$.

Määritetään sitten kaikki ytimen alkiot. Huomataan, että

$$\begin{aligned}
\operatorname{Ker} P &= \{\bar{v} \in \mathbb{R}^2 \mid P(\bar{v}) = \bar{0}\} \\
&= \{(v_1, v_2) \in \mathbb{R}^2 \mid (v_1, 0) = (0, 0)\} \\
&= \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 = 0\} \\
&= \{(0, v_2) \mid v_2 \in \mathbb{R}\} \\
&= \{v_2(0, 1) \mid v_2 \in \mathbb{R}\} \\
&= \operatorname{span}\big((0, 1)\big).
\end{aligned}$$

Lineaarikuvauksen P ydin on siis vektorin $\bar{e}_2 = (0, 1)$ virittämä aliavaruus eli origon kautta kulkeva, vektorin $\bar{e}_2$ suuntainen suora (ks. kuva 19.11).



Kuva 19.11: Lineaarikuvauksen P ytimen vektorit kuvautuvat nollavektoriksi.

Esimerkki 19.3. Määritetään esimerkin 18.4 lineaarikuvauksen $L\colon \mathbb{R}^3 \to \mathbb{R}^2$, $L(x_1, x_2, x_3) = (7x_2,\ x_1 - 3x_3)$ ydin. Vektori $(x_1, x_2, x_3) \in \mathbb{R}^2$ on ytimessä, jos ja vain jos $(7x_2,\ x_1 - 3x_3) = (0, 0)$. Tästä saadaan yhtälöryhmä

$$\begin{cases} 7x_2 = 0 \\ x_1 - 3x_3 = 0. \end{cases}$$

Yhtälöryhmän ratkaisuksi saadaan Gaussin-Jordanin menetelmällä

$$\begin{cases} x_1 &= 3s \\ x_2 &= 0 \\ x_3 &= s, \qquad \text{missä } s \in \mathbb{R}. \end{cases}$$

Siten $\operatorname{Ker} L = \{(3s, 0, s) \mid s \in \mathbb{R}\}$.

Huomataan, että $\{(3s, 0, s) \mid s \in \mathbb{R}\} = \{s(3, 0, 1) \mid s \in \mathbb{R}\}$ eli ydin $\operatorname{Ker} L$ on vektorin $(3, 0, 1)$ virittämä aliavaruus $\operatorname{span}\big((3, 0, 1)\big)$.

Esimerkki 19.4. Määritetään esimerkin 18.6 lineaarikuvauksen

$$L\colon \mathbb{R}^2 \to \mathcal{P}_1, \quad L(a, b) = ax + b$$

ydin. Vektori $(a, b) \in \mathbb{R}^2$ on ytimessä, jos ja vain jos $ax + b$ on nollapolynomi 0. Tämä pätee, jos ja vain jos $a = 0$ ja $b = 0$. Siten $\operatorname{Ker} L = \{(0, 0)\}$.

Tässäkin tapauksessa ydin on siis aliavaruus.

Lause 19.5. *Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Tällöin ydin* $\operatorname{Ker} L$ *on avaruuden V aliavaruus.*

Todistus. Ensinnäkin $\operatorname{Ker} L$ on määritelmänsä mukaan vektoriavaruuden V osajoukko.

Oletetaan, että $\bar{w}, \bar{u} \in \operatorname{Ker} L$ ja $c \in \mathbb{R}$. Nyt siis $L(\bar{w}) = \bar{0}$ ja $L(\bar{u}) = \bar{0}$. Nähdään, että

$$L(\bar{w} + \bar{u}) = L(\bar{w}) + L(\bar{u}) = \bar{0} + \bar{0} = \bar{0}.$$

Koska $L(\bar{w} + \bar{u}) = \bar{0}$ ja ytimeen kuuluvat kaikki nollavektorille kuvautuvat vektorit, pätee $\bar{w} + \bar{u} \in \operatorname{Ker} L$. Lisäksi

$$L(c\bar{w}) = cL(\bar{w}) = c\bar{0} = \bar{0},$$

ja näin ollen $c\bar{w} \in \operatorname{Ker} L$. Lopuksi vielä todetaan, että lauseen 18.12 perusteella $L(\bar{0}) = \bar{0}$, joten $\bar{0} \in \operatorname{Ker} L$.

Siten $\operatorname{Ker} L$ on avaruuden V aliavaruus. □

Lineaarikuvaus $L\colon V \to U$ on *injektiivinen*[3] tai *injektio*, jos eri vektoreilla on aina eri kuvat. Toisin sanoen kaikilla $\bar{v}, \bar{w} \in V$ ehdosta $L(\bar{v}) = L(\bar{w})$ seuraa $\bar{v} = \bar{w}$. Kullekin maalijoukon alkiolle kuvautuu siis korkeintaan yksi lähtöjoukon alkio.

Vaikkapa esimerkin 18.11 projektiokuvaus $P\colon \mathbb{R}^2 \to \mathbb{R}^2$, $P(x_1, x_2) = (x_1, 0)$ ei ole injektio, sillä $P(1, 1) = (1, 0)$ ja $P(1, 2) = (1, 0)$.

Lause 19.6. *Lineaarikuvaus $L\colon V \to U$ on injektio, jos ja vain jos* $\operatorname{Ker} L = \{\bar{0}\}$.

[3] Injektiivisyydestä voidaan puhua muidenkin kuvausten kuin lineaarikuvausten yhteydessä.

Todistus. "$\Rightarrow$": Oletetaan ensin, että L on injektio. Osoitetaan, että $\operatorname{Ker} L = \{\bar{0}\}$. Tiedetään, että $L(\bar{0}) = \bar{0}$, joten $\bar{0} \in \operatorname{Ker} L$. Injektiivisyyden nojalla mikään muu vektori ei voi kuvautua nollavektorille, joten ytimessä on vain yksi vektori, $\bar{0}$.

"$\Leftarrow$": Oletetaan sitten, että $\operatorname{Ker} L = \{\bar{0}\}$. Osoitetaan, että L on injektio. Oletetaan, että vektoreille $\bar{v}, \bar{w} \in V$ pätee $L(\bar{v}) = L(\bar{w})$. Lisäämällä yhtälön molemmille puolille vektori $-L(\bar{w})$, saadaan $L(\bar{v}) - L(\bar{w}) = \bar{0}$. Koska L on lineaarikuvaus, seuraa tästä, että $L(\bar{v} - \bar{w}) = \bar{0}$. Siis $\bar{v} - \bar{w}$ on ytimen alkio. Koska $\operatorname{Ker}(L) = \{\bar{0}\}$, täytyy päteä $\bar{v} - \bar{w} = \bar{0}$. Kun tämän yhtälön molemmille puolille lisätään vektori $\bar{w}$, saadaan $\bar{v} = \bar{w}$. On siis osoitettu, että L on injektio. □

Esimerkki 19.7. Esimerkissä 19.3 nähtiin, että lineaarikuvauksen $L\colon \mathbb{R}^3 \to \mathbb{R}^2$, $L(x_1, x_2, x_3) = (7x_2,\ x_1 - 3x_3)$ ydin on $\{(3s, 0, s) \mid s \in \mathbb{R}\}$. Koska ytimessä on muitakin vektoreita kuin nollavektori, ei kuvaus ole injektio.

Esimerkissä 19.4 puolestaan pääteltiin, että lineaarikuvauksen $L\colon \mathbb{R}^2 \to \mathcal{P}_1$, $L(a, b) = ax + b$ ydin on $\{(0, 0)\}$. Koska ytimen ainoa vektori on nollavektori, on kuvaus injektio.

19.2 Lineaarikuvauksen kuva

Määritelmä 19.8. Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Sen *kuva* on joukko

$$\operatorname{Im} L = LV = \{L(\bar{v}) \mid \bar{v} \in V\}.$$

Lineaarikuvauksen kuva on erikoistapaus aiemmin määritellystä aliavaruuden kuvasta. Nyt aliavaruutena on koko vektoriavaruus V.

Merkintä Im tulee englannin kielen sanasta "image", joka tarkoittaa kuvaa. Huomaa, että kuva $\operatorname{Im} L$ on sama asia kuin kuvauksen arvojoukko. Lineaarikuvausten yhteydessä vain on tapana käyttää termiä kuva.

Esimerkki 19.9. Tarkastellaan jälleen esimerkin 18.11 kuvausta $P\colon \mathbb{R}^2 \to \mathbb{R}^2$, $P(x_1, x_2) \mapsto (x_1, 0)$, joka projisoi vektorit x_1-akselille. Lineaarikuvauksen P kuva on

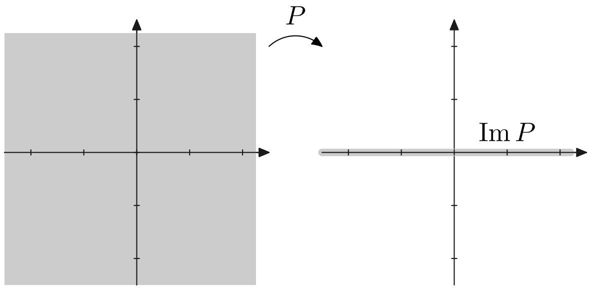
$$\begin{aligned}
\operatorname{Im} P &= \{L(\bar{v}) \mid \bar{v} \in \mathbb{R}^2\} \\
&= \{(v_1, 0) \mid (v_1, v_2) \in \mathbb{R}^2\} \\
&= \{(v_1, 0) \mid v_1 \in \mathbb{R}\} \\
&= \{v_1(1, 0) \mid v_1 \in \mathbb{R}\} \\
&= \operatorname{span}\big((1, 0)\big).
\end{aligned}$$

Lineaarikuvauksen P kuva on siis vektorin $\bar{e}_1 = (1, 0)$ virittämä aliavaruus eli origon kautta kulkeva, vektorin $\bar{e}_1$ suuntainen suora.

Esimerkki 19.10. Määritetään esimerkin 18.6 lineaarikuvauksen $L\colon \mathbb{R}^2 \to \mathcal{P}_1$, $(a, b) \mapsto ax + b$ kuva. Nähdään, että

$$\begin{aligned}
\operatorname{Im} L &= \{L(\bar{v}) \mid \bar{v} \in \mathbb{R}^2\} = \{v_1 x + v_2 \mid (v_1, v_2) \in \mathbb{R}^2\} \\
&= \{v_1 x + v_2 \mid v_1, v_2 \in \mathbb{R}\} = \mathcal{P}_1.
\end{aligned}$$

Siten kuva on koko avaruus $\mathcal{P}_1$.



Kuva 19.12: Lineaarikuvauksen P kuva on vektorin $\bar{e}_1$ virittämä aliavaruus.

Lause 19.11. *Oletetaan, että $L\colon V \to U$ on lineaarikuvaus. Tällöin kuva $\operatorname{Im} L$ on avaruuden U aliavaruus.*

Todistus. Vektoriavaruus V on itsensä aliavaruus. Nyt lauseen 18.18 nojalla kuva $LV = \operatorname{Im} L$ on avaruuden U aliavaruus. □

Lineaarikuvaus $f\colon V \to U$ on *surjektiivinen*[4] eli *surjektio*, jos jokaiselle avaruuden U alkiolle kuvautuu jokin V:n alkio. Nähdään, että lineaarikuvaus $L\colon V \to U$ on surjektio, jos ja vain jos $\operatorname{Im}(L) = U$.

Lineaarikuvausta $L\colon V \to U$ voidaan nyt luonnehtia sen ytimen ja kuvan avulla.

- Kuvaus L on injektio, jos ja vain jos $\operatorname{Ker} L = \{\bar{0}\}$.
- Kuvaus L on surjektio, jos ja vain jos $\operatorname{Im} L = U$.

[4] Surjektiivisuudesta voidaa puhua muidenkin kuvausten kuin lineaarikuvausten yhteydessä.

20 Isomorfismi

Kuvaus on *bijektiivinen* eli *bijektio*, jos se on sekä injektio että surjektio. Tällöin kullekin maalijoukon alkiolle kuvautuu täsmälleen yksi lähtöjoukon alkio.

Määritelmä 20.1. *Isomorfismi* on bijektiivinen lineaarikuvaus.

Jos vektoriavaruuksien V ja U välillä on isomorfismi, sanotaan, että V ja U ovat *isomorfiset*. Tällöin merkitään $V \cong U$. Isomorfiset vektoriavaruudet ovat ominaisuuksiltaan samankaltaiset.

Esimerkki 20.2. Osoitetaan, että kuvaus $L\colon \mathbb{R}^2 \to \mathcal{P}_1$, $L(a,b) = ax + b$ on isomorfismi. Esimerkin 18.6 perusteella kuvauksen tiedetään olevan lineaarikuvaus. Lisäksi on osoitettava, että kuvaus on bijektio eli että kuvaus on sekä injektio että surjektio. Kuvauksen ydin $\operatorname{Ker} L$ on esimerkin 19.4 mukaan $\{(0,0)\}$, joten kuvaus on injektio. Kuva $\operatorname{Im} L$ puolestaan on esimerkin 19.10 nojalla koko maalijoukko $\mathcal{P}_1$. Siten L on surjektio. Näin ollen L on bijektio ja edelleen isomorfismi.

Koska L on isomorfismi, vektorivaruudet $\mathbb{R}^2$ ja $\mathcal{P}_1$ ovat isomorfisia. Huomataan, että avaruudet muistuttavat toisiaan. Sekä alkiossa (a,b) että alkiossa $ax+b$ näkyvät reaaliluvut a ja b. Kaikki oleellinen tieto alkiosta sisältyy näihin reaalilukuihin. Lisäksi nämä reaaliluvut käyttäytyvät samalla tavoin yhteenlaskussa ja skalaarikertolaskussa. Tätä on havainnollistettu taulukossa 2.1.

Vektoriavaruus	summa
$\mathbb{R}^2$	$(a,b)+(c,d)=(a+c,b+d)$
$\mathcal{P}_1$	$(ax+b)+(cx+d)=(a+c)x+(b+d)$

Vektoriavaruus	skalaarimonikerta
$\mathbb{R}^2$	$r(a,b)=(ra,rb)$
$\mathcal{P}_1$	$r(ax+b)=rax+rb$

Taulukko 2.1: Vektorien yhteenlasku ja skalaarikertolasku avaruuksissa $\mathbb{R}^2$ ja $\mathcal{P}_1$.

Oletetaan, että $L\colon V \to U$ on kuvaus. Jos on olemassa kuvaus $T\colon U \to V$, jolle pätee

$$T \circ L = \mathrm{id}_V \quad \text{ja} \quad L \circ T = \mathrm{id}_U,$$

sanotaan, että kuvaus T on kuvauksen L *käänteiskuvaus*[5].

Tässä merkintä id_V tarkoittaa avaruuden V *identtistä kuvausta*: $\mathrm{id}_V\colon V \to V$, jolla $\bar{v} \mapsto \bar{v}$ kaikilla $\bar{v} \in V$. Vastaavasti id_U tarkoittaa avaruuden U identtistä kuvausta. Käänteiskuvauksen ehto voidaan siis kirjoittaa myös muodossa

$$T(L(\bar{v})) = \bar{v} \quad \text{kaikilla } \bar{v} \in V \qquad \text{ja} \qquad L(T(\bar{u})) = \bar{u} \quad \text{kaikilla } \bar{u} \in V.$$

Kuvauksen L käänteiskuvausta merkitään L^{-1}. Kaikilla kuvauksilla ei ole käänteiskuvausta.

[5]Käänteiskuvauksista voidaan puhua muidenkin kuvausten kuin lineaarikuvausten yhteydessä.

Esimerkki 20.3. Osoitetaan, että kuvauksen $L\colon \mathbb{R}^2 \to \mathcal{P}_1$, $L(a,b) = ax + b$ käänteiskuvaus on kuvaus $T\colon \mathcal{P}_1 \to \mathbb{R}^2$, $T(ax+b) = (a,b)$. Olkoon $(a,b) \in \mathbb{R}^2$. Nyt

$$(T \circ L)(a,b) = T(L(a,b)) = T(a+bx) = (a,b),$$

joten $T \circ L = \mathrm{id}_{\mathbb{R}^2}$. Toisaalta

$$(L \circ T)(ax+b) = L(T(ax+b)) = L(a,b) = ax+b,$$

joten $L \circ T = \mathrm{id}_{\mathcal{P}_1}$. Siten T on kuvauksen L käänteiskuvaus eli $T = L^{-1}$.

Voidaan osoittaa, että kuvauksella L on käänteiskuvaus, jos ja vain jos L on bijektio. Tästä saadaan seuraava tulos:

Lause 20.4. *Lineaarikuvaus on isomorfismi, jos ja vain jos sillä on käänteiskuvaus.*

Lause 20.5. *Oletetaan, että V, U ja W ovat vektoriavaruuksia. Tällöin*

a) $V \cong V$

b) jos $V \cong U$, *niin* $U \cong V$

c) jos $V \cong U$ *ja* $U \cong W$, *niin* $V \cong W$.

Todistus.

a) Ei ole vaikea osoittaa, että identtinen kuvaus $\mathrm{id}\colon V \to V$, $\mathrm{id}(\bar{v}) = \bar{v}$ on bijektiivinen lineaarikuvaus. Siten se on isomorfismi vektoriavaruudelta V itselleen.

b) Oletetaan, että $L\colon V \to U$ on isomorfismi. Koska L on bijektio, on olemassa käänteiskuvaus $L^{-1}\colon U \to V$. Osoitetaan, että L^{-1} on isomorfismi. Ensinnäkin tiedetään, että bijektion käänteiskuvaus on myöskin bijektio, joten L^{-1} on bijektio. Tarkistetaan vielä lineaarikuvauksen ehdot.

   Oletetaan, että $\bar{u}_1, \bar{u}_2 \in U$ ja $c \in \mathbb{R}$. Koska L on bijektio ja siten surjektiivinen, on olemassa vektorit $\bar{v}_1, \bar{v}_2 \in V$, joille pätee $L(\bar{v}_1) = \bar{u}_1$ ja $L(\bar{v}_2) = \bar{u}_2$. Huomaa, että tällöin $\bar{v}_1 = L^{-1}(\bar{u}_1)$ ja $\bar{v}_2 = L^{-1}(\bar{u}_2)$. Nyt

$$\begin{aligned} L^{-1}(\bar{u}_1 + \bar{u}_2) &= L^{-1}(L(\bar{v}_1) + L(\bar{v}_2)) = L^{-1}(L(\bar{v}_1 + \bar{v}_2)) \\ &= \bar{v}_1 + \bar{v}_2 = L^{-1}(\bar{u}_1) + L^{-1}(\bar{u}_2) \end{aligned}$$

   ja

$$L^{-1}(c\bar{u}_1) = L^{-1}(cL(\bar{v}_1)) = L^{-1}(L(c\bar{v}_1)) = c\bar{v}_1 = cL^{-1}(\bar{u}_1).$$

   Siten L^{-1} on lineaarikuvaus. Näin on osoitettu, että L^{-1} on isomorfismi.

c) Todistus jätetään harjoitustehtäväksi. □

21 Kanta ja lineaarikuvaukset

Esimerkissä 18.7 näytettiin, miten kantavektorien arvoista voidaan johtaa kaikkien muidenkin vektorien arvot, jos kuvauksen tiedetään olevan lineaarikuvaus. Vieläkin vahvempi tulos pitää paikkansa. Lineaarikuvaus voidaan jopa *määäritellä* antamalla pelkkien kantavektorien arvot.

Lause 21.1. *Olkoot V ja U vektoriavaruuksia. Oletetaan, että $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n)$ on avaruuden V kanta ja $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n \in U$. Tällöin on olemassa täsmälleen yksi lineaarikuvaus $L\colon V \to U$, jolle pätee $L(\bar{v}_i) = \bar{u}_i$ kaikilla $i \in \{1, \dots, n\}$.*

Todistus. Jos $\bar{w} \in V$, on olemassa yksikäsitteiset reaaliluvut $w_1 \dots, w_n$, joille pätee $\bar{w} = w_1\bar{v}_1 + w_2\bar{v}_2 + \cdots + w_n\bar{v}_n$. Määritellään kuvaus

$$L\colon V \to U, \quad L(\bar{w}) = w_1\bar{u}_1 + w_2\bar{u}_2 + \cdots + w_n\bar{u}_n,$$

missä luvut $w_1 \dots, w_n$ ovat kuten yllä. (Kyseessä ovat siis vektorin $\bar{w}$ koordinaatit kannan $(\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n)$ suhteen.)

Osoitetaan, että L on etsitty kuvaus. Jos $i \in \{1, \dots, n\}$, niin

$$\bar{v}_i = 0\bar{v}_1 + \cdots + 0\bar{v}_{i-1} + 1\bar{v}_i + 0\bar{v}_{i+1} + 0\bar{v}_n.$$

Siis

$$L(\bar{v}_i) = 0\bar{u}_1 + \cdots + 0\bar{u}_{i-1} + 1\bar{u}_i + 0\bar{u}_{i+1} + 0\bar{u}_n = \bar{u}_i,$$

joten kantavektorit kuvautuvat halutulla tavalla.

Osoitetaan vielä, että L on lineaarikuvaus. Oletetaan, että $\bar{x}, \bar{y} \in V$. Nyt

$$\bar{x} = x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_n\bar{v}_n \quad \text{ja} \quad \bar{y} = y_1\bar{v}_1 + y_2\bar{v}_2 + \cdots + y_n\bar{v}_n$$

joillakin $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$. Nähdään, että

$$\begin{aligned}
L(\bar{x} + \bar{y}) &= L\big((x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_n\bar{v}_n) + (y_1\bar{v}_1 + y_2\bar{v}_2 + \cdots + y_n\bar{v}_n)\big) \\
&= L\big((x_1 + y_1)\bar{v}_1 + (x_2 + y_2)\bar{v}_2 + \cdots + (x_n + y_n)\bar{v}_n\big) \\
&= (x_1 + y_1)\bar{u}_1 + (x_2 + y_2)\bar{u}_2 + \cdots + (x_n + y_n)\bar{u}_n \\
&= (x_1\bar{u}_1 + x_2\bar{u}_2 + \cdots + x_n\bar{u}_n) + (y_1\bar{u}_1 + y_2\bar{u}_2 + \cdots + y_n\bar{u}_n) \\
&= L(x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_n\bar{v}_n) + L(y_1\bar{v}_1 + y_2\bar{v}_2 + \cdots + y_n\bar{v}_n) \\
&= L(\bar{x}) + L(\bar{y}).
\end{aligned}$$

Oletetaan vielä, että $c \in \mathbb{R}$. Nyt

$$\begin{aligned}
L(c\bar{x}) &= L\big(c(x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_n\bar{v}_n)\big) \\
&= L(cx_1\bar{v}_1 + cx_2\bar{v}_2 + \cdots + cx_n\bar{v}_n) \\
&= cx_1\bar{u}_1 + cx_2\bar{u}_2 + \cdots + cx_n\bar{u}_n \\
&= c(x_1\bar{u}_1 + x_2\bar{u}_2 + \cdots + x_n\bar{u}_n) \\
&= cL(x_1\bar{v}_1 + x_2\bar{v}_2 + \cdots + x_n\bar{v}_n) \\
&= cL(\bar{x}).
\end{aligned}$$

Siten L on lineaarikuaus. On siis olemassa vähintään yksi kuvaus, joka täyttää annetut ehdot.

Osoitetaan sitten, että lauseen vaatimukset täyttäviä lineaarikuvauksia on enintään yksi. Oletetaan, että $L\colon V \to W$ ja $T\colon V \to W$ ovat molemmat lineaarikuvauksia, joille pätee

$$\begin{aligned}
&L(\bar{v}_1) = \bar{w}_1, L(\bar{v}_2) = \bar{w}_2, \ldots, L(\bar{v}_n) = \bar{w}_n \quad \text{ja} \\
&T(\bar{v}_1) = \bar{w}_1, T(\bar{v}_2) = \bar{w}_2, \ldots, T(\bar{v}_n) = \bar{w}_n.
\end{aligned}$$

Osoitetaan, että kuvaukset L ja T ovat samat.

Oletetaan, että $\bar{v} \in V$. Tällöin $\bar{v} = a_1\bar{v}_1 + \cdots + a_n\bar{v}_n$ joillakin $a_1, \ldots, a_n \in \mathbb{R}$, sillä $(\bar{v}_1, \ldots, \bar{v}_n)$ on avaruuden V kanta. Kuvausten L ja T lineaarisuutta käyttäen saadaan

$$\begin{aligned}
L(\bar{v}) &= L(a_1\bar{v}_1 + \cdots + a_n\bar{v}_n) = a_1L(\bar{v}_1) + \cdots + a_nL(\bar{v}_n) \\
&= a_1\bar{w}_1 + \cdots + a_n\bar{w}_n = a_1T(\bar{v}_1) + \cdots + a_nT(\bar{v}_n) \\
&= T(a_1\bar{v}_1 + \cdots + a_n\bar{v}_n) = T(\bar{v}).
\end{aligned}$$

Kuvauksilla $L\colon V \to W$ ja $T\colon V \to W$ on samat arvot, joten ne ovat sama kuvaus.

Siten lauseen ehdot täyttäviä lineaarikuvauksia on täsmälleen yksi. □

Nyt voidaan määritellä vaikkapa lineaarikuvaus $L\colon \mathbb{R}^5 \to \mathbb{R}^3$ asettamalla

$$\begin{aligned}
&L(\bar{e}_1) = (1, 5, -2),\ L(\bar{e}_2) = (0, 0, 0),\ L(\bar{e}_3) = (-1, 2, 6), \\
&L(\bar{e}_4) = (7, 4, 4),\ L(\bar{e}_5) = (0, 1, 0).
\end{aligned}$$

Lauseen 21.1 perusteella on olemassa täsmälleen yksi lineaarikuvaus, joka täyttää nämä viisi ehtoa. Niiden antaminen riittää siis määrittelemään kuvauksen L. Kuvauksen muut arvot voidaan laskea samaan tapaan kuin esimerkissä 18.7.

Äärellisulotteisen vektoriavaruuden tapauksessa lineaarikuvauksen ytimen ja kuvan dimensiot riippuvat toisistaan.

Lause 21.2 (Dimensiolause)**.** *Olkoot V ja U vektoriavaruuksia. Oletetaan, että V on äärellisulotteinen. Olkoon $L\colon V \to U$ lineaarikuvaus. Tällöin*

$$\dim(V) = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L).$$

Todistus. Olkoon $(\bar{v}_1, \ldots, \bar{v}_k)$ aliavaruuden $\operatorname{Ker} L$ kanta. Koska jono $(\bar{v}_1, \ldots, \bar{v}_k)$ on vapaa, se voidaan lauseen 17.12 nojalla täydentää vektoriavaruuden V kannaksi $(\bar{v}_1, \ldots, \bar{v}_k, \bar{u}_{k+1}, \ldots \bar{u}_n)$. Nyt siis $\dim(\operatorname{Ker} L) = k$ ja $\dim(V) = n$. Osoitetaan, että $(L(\bar{u}_{k+1}), \ldots L(\bar{u}_n))$ on aliavaruuden $\operatorname{Im} L$ kanta, jolloin $\dim(\operatorname{Im} L) = n - k$. Tämä todistaa väitteen.

Osoitetaan ensin, että $(L(\bar{u}_{k+1}), \ldots, L(\bar{u}_n))$ virittää aliavaruuden $\operatorname{Im} L$. Oletetaan, että $\bar{w} \in \operatorname{Im} L$. Koska kuvaus L on bijektio, on olemassa $\bar{v} \in V$, jolle pätee $\bar{w} = L(\bar{v})$. Toisaalta tiedetään, että $(\bar{v}_1, \ldots, \bar{v}_k, \bar{u}_{k+1}, \ldots \bar{u}_n)$ on avaruuden V kanta, joten

$$\bar{v} = a_1\bar{v}_1 + \cdots + a_k\bar{v}_k + a_{k+1}\bar{u}_{k+1} + \cdots + a_n\bar{u}_n$$

joillakin $a_1, \dots, a_n \in \mathbb{R}$. Nyt

$$\begin{aligned}\bar{w} &= L(\bar{v}) = L(a_1\bar{v}_1 + \cdots + a_k\bar{v}_k + a_{k+1}\bar{u}_{k+1} + \cdots + a_n\bar{u}_n) \\ &= a_1L(\bar{v}_1) + \cdots + a_kL(\bar{v}_k) + a_{k+1}L(\bar{u}_{k+1}) + \cdots + a_nL(\bar{u}_n) \\ &= \bar{0} + \cdots + \bar{0} + a_{k+1}L(\bar{u}_{k+1}) + \cdots + a_nL(\bar{u}_n),\end{aligned}$$

sillä $L(\bar{v}_1), \dots, L(\bar{v}_k) \in \operatorname{Ker} L$. Siten jokainen avaruuden U alkio voidaan kirjoittaa vektoreiden $L(\bar{u}_{k+1}), \dots, L(\bar{u}_n)$ lineaarikombinaationa eli

$$U = \operatorname{span}(L(\bar{u}_{k+1}), \dots L(\bar{u}_n)).$$

Osoitetaan sitten, että jono $(L(\bar{u}_{k+1}), \dots L(\bar{u}_n))$ on vapaa. Oletetaan, että

$$c_{k+1}L(\bar{u}_{k+1}) + \cdots + c_nL(\bar{u}_n) = \bar{0}$$

joillakin $c_{k+1}, \dots, c_n \in \mathbb{R}$. Nyt

$$L(c_{k+1}\bar{u}_{k+1} + \cdots + c_n\bar{u}_n) = \bar{0},$$

joten $c_{k+1}\bar{u}_{k+1} + \cdots + c_n\bar{u}_n \in \operatorname{Ker} L$. Näin ollen vektori $c_{k+1}\bar{u}_{k+1} + \cdots + c_n\bar{u}_n$ voidaan kirjoittaa aliavaruuden $\operatorname{Ker} L$ kannan alkioiden lineaarikombinaationa. On siis olemassa luvut $b_1, \dots, b_k \in \mathbb{R}$, joille pätee

$$c_{k+1}\bar{u}_{k+1} + \cdots + c_n\bar{u}_n = b_1\bar{v}_1 + \cdots + b_k\bar{v}_k.$$

Tästä saadaan yhtälö

$$-b_1\bar{v}_1 - \cdots - b_k\bar{v}_k + c_{k+1}\bar{u}_{k+1} + \cdots + c_n\bar{u}_n = \bar{0}.$$

Jono $(\bar{v}_1, \dots, \bar{v}_k, \bar{u}_{k+1}, \dots \bar{u}_n)$ on kuitenkin vektoriavaruuden V kanta, joten se on vapaa. Kaikkien lineaarikombinaation kertoimien pitää siis olla nollia:

$$b_1 = 0, \dots, b_k = 0, c_{k+1} = 0, \dots, c_n = 0.$$

Näin ollen tiedetään, että $c_{k+1} = 0, \dots, c_n = 0$, mistä seuraa, että alunperin tutkittu jono $(L(\bar{u}_{k+1}), \dots L(\bar{u}_n))$ on vapaa. Siten $(L(\bar{u}_{k+1}), \dots L(\bar{u}_n))$ on aliavaruuden $\operatorname{Im} L$ kanta.

Nyt $\dim(\operatorname{Im} L) = n - k$, $\dim(\operatorname{Ker} L) = k$ ja $\dim V = n$. Tästä seuraa, että $\dim(V) = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L)$, joten väite on todistettu. □

Esimerkki 21.3. Tarkastellaan esimerkin 18.11 projektiokuvausta $P\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L(x_1, x_2) = (x_1, 0)$. Esimerkissä 19.2 osoitettiin, että $\ker P = \operatorname{span}((1, 0))$, ja esimerkissä 19.9 näytettiin, että $\operatorname{Im} P = \operatorname{span}((1, 0))$. Nyt nähdään, että

$$\dim(\mathbb{R}^2) = 2 = 1 + 1 = \dim(\ker P) + \dim(\operatorname{Im} P)$$

aivan kuten Dimensiolauseen mukaan kuuluukin olla.

Jos lähtö- ja maaliavaruuden dimensiot ovat samat (ja äärelliset), ovat kaikki injektiot surjektioita ja surjektiot injektioita.

Lause 21.4. *Oletetaan, että V ja U ovat äärellisulotteisia vektoriavaruuksia ja* $\dim(V) = \dim(U)$*. Olkoon $L\colon V \to U$ lineaarikuvaus. Tällöin seuraavat väitteet ovat yhtäpitäviä:*

a) L on injektio

b) L on surjektio.

Huomaa, että lauseessa oletetaan, että lineaarikuvauksen lähtö- ja maaliavaruudella on sama dimensio. Jos tämä ehto ei ole voimassa, ei lauseen tuloskaan välttämättä päde.

Todistus. On osoitettava, että väite a) pätee, jos ja vain jos väite b) pätee. Perustana on Dimensiolause 21.2, jonka mukaan $\dim(V) = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L)$.

"a) $\Rightarrow$ b)": Oletetaan, että L on injektio. Nyt lauseen 19.6 nojalla $\operatorname{Ker} L = \{\bar{0}\}$, joten $\dim(\operatorname{Ker} L) = 0$. Tästä seuraa, että

$$\dim(\operatorname{Im} L) = \dim(V) - \dim(\operatorname{Ker} L) = \dim(V) = \dim(U),$$

koska oletimme, että avaruuksien V ja U dimensiot ovat samat. Nyt tiedetään, että kuva $\operatorname{Im} L$ on vektoriavaruuden U aliavaruus, jolla on sama dimensio kuin avaruudella U. Lauseen 17.15 perusteella $\operatorname{Im} L = U$. Siten L on surjektio.

"b) $\Rightarrow$ c)": Oletetaan sitten, että L on surjektio. Nyt $\operatorname{Im} L = U$, joten pätee $\dim(\operatorname{Im} L) = \dim(U) = \dim(V)$. Tästä seuraa, että

$$\dim(\operatorname{Ker} L) = \dim(V) - \dim(\operatorname{Im} L) = 0.$$

Siten $\operatorname{Ker} L = \{\bar{0}\}$, ja L on injektio. □

Jos äärellisulotteiset vektoriavaruudet ovat isomorfiset, niiden dimensiot ovat samat. Myös käänteinen väite pätee. Jos vektoriavaruuksien dimensiot ovat samat, ne ovat isomorfiset. Äärellisulotteisen vektoriavaruuden dimensio riittää siis kertomaan vektoriavaruudesta kaiken oleellisen.

Lause 21.5. *Oletetaan, että V ja W ovat äärellisulotteisia vektoriavaruuksia. Vektoriavaruudet V ja W ovat isomorfiset, jos ja vain jos* $\dim(V) = \dim(W)$*.*

Todistus. "$\Rightarrow$": Oletetaan, että V ja W ovat isomorfiset. Tällöin on olemassa bijektiivinen lineaarikuvaus $L\colon V \to W$. Koska L on injektio, niin $\operatorname{Ker} L = \{\bar{0}\}$. Dimensiolauseen 21.2 nojalla pätee

$$\dim(\operatorname{Im} L) = \dim(V) - \dim(\operatorname{Ker} L) = \dim(V) - 0 = \dim(V).$$

Toisaalta L on myös surjektio, joten $\operatorname{Im} L = W$. Siis

$$\dim(V) = \dim(\operatorname{Im} L) = \dim(W).$$

"$\Leftarrow$": Oletetaan, että $\dim(V) = \dim(W) = n$. Vektoriavaruuksien välinen isomorfismi saadaan kuvaamalla avaruuden V kantavektorit avaruuden W kantavektoreille. Olkoon siis $(\bar{v}_1, \dots, \bar{v}_n)$ vektoriavaruuden V kanta ja $(\bar{w}_1, \dots, \bar{w}_n)$ vektoriavaruuden W kanta. Olkoon $L\colon V \to W$ se lineaarikuvaus, jolle pätee

$$L(\bar{v}_1) = \bar{w}_1, L(\bar{v}_2) = \bar{w}_2, \dots, L(\bar{v}_n) = \bar{w}_n.$$

Lauseen 21.1 mukaan tällaisia lineaarikuvauksia on täsmälleen yksi. Osoitetaan, että L on bijektio.

Näytetään ensin, että L on injektio osoittamalla, että $\operatorname{Ker} L = \{\bar{0}\}$. Oletetaan, että $\bar{v} \in \operatorname{Ker} L$. Tällöin $L(\bar{v}) = \bar{0}$. Kirjoitetaan $\bar{v}$ kantavektorien lineaarikombinaationa: $\bar{v} = a_1\bar{v}_1 + \cdots + a_n\bar{v}_n$ joillakin $a_1, \dots, a_n \in \mathbb{R}$. Nyt saadaan

$$\begin{aligned}\bar{0} &= L(\bar{v}) = L(a_1\bar{v}_1 + \cdots + a_n\bar{v}_n) = a_1L(\bar{v}_1) + \cdots + a_nL(\bar{v}_n) \\ &= a_1\bar{w}_1 + \cdots + a_n\bar{w}_n.\end{aligned}$$

Jono $(\bar{w}_1, \dots, \bar{w}_n)$ on kanta ja siten vapaa, joten tästä yhtälöstä seuraa, että

$$a_1 = 0, a_2 = 0, \dots, a_n = 0.$$

Siis

$$\bar{v} = a_1\bar{v}_1 + \cdots + a_n\bar{v}_n = 0\bar{v}_1 + \cdots + 0\bar{v}_n = \bar{0}.$$

Näin on osoitettu, että ytimessä ei ole muita vektoreita kuin nollavektori, eli $\operatorname{Ker} L = \{\bar{0}\}$. Siis L on injektio.

Osoitetaan sitten surjektiivisuus. Oletuksen mukaan $\dim(V) = \dim(W)$. Koska lineaarikuvaus $L\colon V \to W$ on injektio, se on lauseen 21.4 mukaan myös surjektio. Siis L on bijektiivinen lineaarikuvaus eli isomorfismi. Näin ollen $V \cong W$. □

21.1 Lineaarikuvauksen matriisi

Tässä luvussa osoitamme, että jokainen lineaarikuvaus avaruudelta $\mathbb{R}^m$ avaruudelle $\mathbb{R}^n$ on jonkin matriisin määräämä lineaarikuvaus. Tutkitaan kuitenkin sitä ennen, miten matriisien määräämät lineaarikuvaukset käyttäytyvät.

Esimerkki 21.6. Tarkastellaan matriisin

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

määräämää lineaarikuvausta $L_A\colon \mathbb{R}^3 \to \mathbb{R}^2$, $L_A(\bar{x}) = A\bar{x}$. Avaruuden $\mathbb{R}^3$ vektorin (x_1, x_2, x_3) kuva tässä kuvauksessa on

$$\begin{aligned}\begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 1x_1 + 2x_2 + 0x_3 \\ 0x_1 - 1x_2 - 3x_3 \end{bmatrix} \\ &= \begin{bmatrix} 1x_1 \\ 0x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ -1x_2 \end{bmatrix} + \begin{bmatrix} 0x_3 \\ -3x_3 \end{bmatrix} \\ &= x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ -1 \end{bmatrix} + x_3\begin{bmatrix} 0 \\ -3 \end{bmatrix}.\end{aligned}$$

Vektorin (x_1, x_2, x_3) kuva kuvauksessa L_A on siis matriisin A sarakkeiden lineaarikombinaatio, jossa kertoimina ovat vektorin komponentit.

Nähdään, että kantavektorin $\bar{e}_1 = (1, 0, 0)$ kuva on

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

Samalla tavalla kantavektorin $\bar{e}_2 = (0, 1, 0)$ kuvaksi saadaan $(2, -1)$ ja kantavektorin $\bar{e}_3 = (0, 0, 1)$ kuvaksi $(0, -3)$. Luonnollisen kannan vektorien kuvat ovat siis matriisin A sarakkeet.

Lauseessa 18.8 osoitettiin, että matriisista saadaan aina lineaarikuvaus. Nyt osoitamme, että jokainen lineaarikuvaus $\mathbb{R}^n \to \mathbb{R}^m$ on matriisin määräämä. Lineaarikuvaukset $\mathbb{R}^n \to \mathbb{R}^m$ ja $m \times n$-matriisit siis vastaavat toisiaan.

Lause 21.7. *Oletetaan, että $T\colon \mathbb{R}^n \to \mathbb{R}^m$ on lineaarikuvaus. Tällöin on olemassa täsmälleen yksi matriisi $A \in \mathbb{R}^{m \times n}$, jolla $T(\bar{v}) = A\bar{v}$ kaikilla $\bar{v} \in \mathbb{R}^n$.*

Todistus. Aloitetaan tutkimalla, mitä tapahtuu, kun matriisilla kerrotaan vektoria. Samaan tapaan kuin esimerkissä 21.6 nähdään, että

$$\begin{aligned}
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} &= \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix} \\
&= \begin{bmatrix} v_1a_{11} \\ v_1a_{21} \\ \vdots \\ v_1a_{m1} \end{bmatrix} + \begin{bmatrix} v_2a_{12} \\ v_2a_{22} \\ \vdots \\ v_2a_{m2} \end{bmatrix} + \cdots + \begin{bmatrix} v_na_{1n} \\ v_na_{2n} \\ \vdots \\ v_na_{mn} \end{bmatrix} \\
&= v_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.
\end{aligned}$$

Tulo on siis matriisin sarakkeiden lineaarikombinaatio, jossa kertoimina ovat vektorin komponentit.

Muodostetaan etsitty matriisi A seuraavasti: Katsotaan, miten avaruuden $\mathbb{R}^n$ luonnollisen kannan $(\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n)$ vektorit kuvautuvat lineaarikuvauksessa T eli määritetään $T(\bar{e}_1)$, $T(\bar{e}_2), \dots, T(\bar{e}_n)$. Asetetaan vektorit $T(\bar{e}_1)$, $T(\bar{e}_2), \dots, T(\bar{e}_n)$ matriisin A sarakkeiksi tässä järjestyksessä. Voidaan merkitä lyhyesti

$$A = \begin{bmatrix} T(\bar{e}_1) & T(\bar{e}_2) & \dots & T(\bar{e}_n) \end{bmatrix}.$$

Huomaa, että matriisin jokaisessa sarakkeessa on m alkiota ja sarakkeita on n kappaletta, joten A todella on $m \times n$-matriisi.

Osoitetaan, että matriisin A määräämä lineaarikuvaus $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ on sama kuvaus kuin T. Koska kantavektorien kuvavektorit määräävät lineaarikuvauksen (lause 21.1), riittää osoittaa, että kantavektorit $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n$ kuvautuvat samalla tavalla kuvauksissa L_A ja T.

Määritetään luonnollisen kannan vektorien kuvat samaan tapaan kuin esimerkissä 21.6. Edellä osoitettiin, että vektorin kuva lineaarikuvauksessa L_A on matriisin A sarakkeiden lineaarikombinaatio, jossa kertoimina ovat vektorin komponentit. Tästä seuraa, että luonnollisen kannan vektorien $\bar{e}_1, \dots, \bar{e}_n$ kuvavektorit ovat matriisin A sarakkeet. Esimerkiksi

$$L_A(\bar{e}_1) = A\bar{e}_1 = 1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + 0\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + 0\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}.$$

Matriisin A sarakkeet taas valittiin niin, että ne ovat kantavektorien kuvavektorit kuvauksessa T. Siten $L_A(\bar{e}_i) = T(\bar{e}_i)$ kaikilla $i \in \{1, \dots, n\}$. Lineaarikuvaukset L_A ja T ovat siten lauseen 21.1 nojalla sama kuvaus, eli $T(\bar{v}) = A\bar{v}$ kaikilla $\bar{v} \in \mathbb{R}^n$.

Osoitetaan vielä, että ehdon toteuttavia matriiseja ei ole enempää kuin yksi. Oletetaan, että sekä matriisin A että matriisin B määräämä kuvaus on T eli $L_A = T$ ja $L_B = T$. Edellä nähtiin, että matriisin määräämässä kuvauksessa luonnollisen kannan vektorien kuvavektorit ovat matriisin sarakkeet. Koska kuvaukset L_A ja L_B ovat samat, on luonnollisen kannan vektoreilla näissä kuvauksissa samat kuvavektorit. Siten matriisien A ja B sarakkeet ovat samat, eli kyseessä on sama matriisi. Näin ollen on olemassa vain yksi matriisi, jonka määrämä lineaarikuvaus on T. □

Määritelmä 21.8. Oletetaan, että $T\colon \mathbb{R}^n \to \mathbb{R}^m$ on lineaarikuvaus. Edellisessä lauseessa määriteltyä matriisia

$$A = \begin{bmatrix} T(\bar{e}_1) & T(\bar{e}_2) & \dots & T(\bar{e}_n) \end{bmatrix}$$

kutsutaan lineaarikuvauksen T *(standardi)matriisiksi.*

Jos A on lineaarikuvauksen $T\colon \mathbb{R}^n \to \mathbb{R}^m$ matriisi, niin $T(\bar{v}) = A\bar{v}$ kaikilla $\bar{v} \in \mathbb{R}^n$.

Esimerkki 21.9. Määritetään esimerkin 18.4 lineaarikuvauksen

$$L\colon \mathbb{R}^3 \to \mathbb{R}^2, \quad L(x_1, x_2, x_3) = (7x_2,\ x_1 - 3x_3)$$

matriisi lauseen 21.7 avulla. Tätä varten tarvitaan luonnollisen kannan vektorien kuvavektorit kuvauksessa L:

$$L(1,0,0) = (0,1), \quad L(0,1,0) = (7,0), \quad L(0,0,1) = (0,-3).$$

Nyt kuvauksen L matriisi saadaan asettamalla nämä kuvavektorit matriisin sarakkeiksi. Siten matriisi on

$$\begin{bmatrix} 0 & 7 & 0 \\ 1 & 0 & -3 \end{bmatrix}.$$

Lopuksi voi vielä halutessaan tarkistaa, että matriisilla kertominen tosiaankin tuottaa lineaarikuvauksen arvot. Oletetaan, että $(x_1, x_2, x_3) \in \mathbb{R}^2$. Nyt

$$\begin{bmatrix} 0 & 7 & 0 \\ 1 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7x_2 \\ x_1 - 3x_3 \end{bmatrix},$$

ja nähdään, että tuloksena on vektorin (x_1, x_2) kuva kuten pitikin.

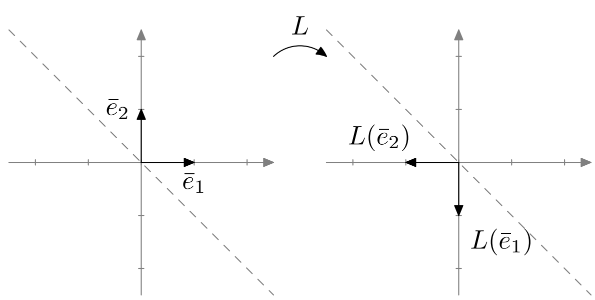
Vaihtoehtoisesti kuvauksen matriisin voi etsiä samalla tavalla kuin esimerkissä 18.10.

Esimerkki 21.10. Esimerkissä 18.9 näytetiin, miten joitakin lineaarikuvauksia voidaan ajatella venytyksinä, peilauksina tai kiertoina. Nyt kun tiedetään, kuinka lineaarikuvauksen matriisi muodostetaan, on geometrisen tulkinnan avulla helppo johtaa tällaisen lineaarikuvauksen matriisi.

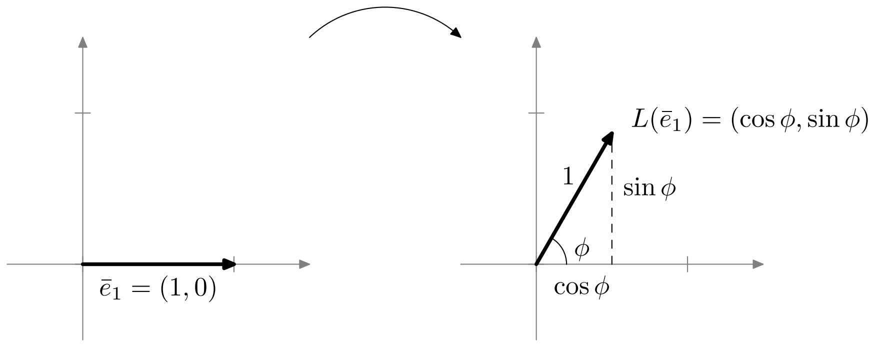
Tutkitaan vaikkapa, millainen matriisi on lineaarikuvauksella L, joka peilaa tason vektorit suoran $\text{span}\big((-1,1)\big)$ suhteen. (Jos ollaan tarkkoja, pitäisi ensin osoittaa, että kyseinen kuvaus todellakin on lineaarikuvaus. Se jätetään tällä kertaa väliin.) Matriisin sarakkeiksi tulevat luonnollisen kannan vektorien kuvavektorit. Nähdään, että kantavektori $(1,0)$ kuvautuu vektorille $(0,-1)$ ja kantavektori $(0,1)$ kuvautuu vektorille $(-1,0)$. Nämä vektorit ovat kuvauksen L matriisin B sarakkeet:

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

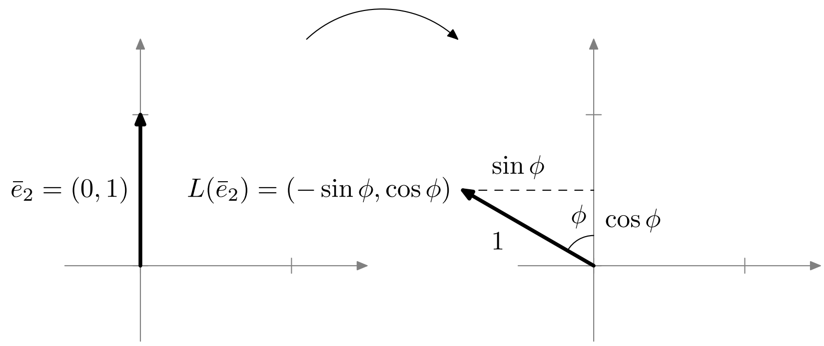
Nyt siis $L(\bar{v}) = B\bar{v}$ kaikilla $\bar{v} \in \mathbb{R}^2$.



Kuva 21.13: Kantavektorien $\bar{e}_1$ ja $\bar{e}_2$ kuvautuminen lineaarikuvauksessa L.



Kuva 21.14: Kantavektorin $\bar{e}_1$ kuvautuminen kiertokuvauksessa.



Kuva 21.15: Kantavektorin $\bar{e}_2$ kuvautuminen kiertokuvauksessa.

Esimerkki 21.11. Tarkastellaan lineaarikuvausta, joka kiertää tason vektoreita ϕ astetta vastapäivään. (Jätämme jälleen todistamatta, että kuvaus on todellakin lineaarikuvaus.) Lineaarikuvauksen matriisia varten tarvitsemme luonnollisen kannan vektorien kuvavektorit. Kuvista 21.14 ja 21.15 näkyy, että vektorin $(1,0)$ kuva on $(\cos\phi, \sin\phi)$ ja vektorin $(0,1)$ kuva on $(-\sin\phi, \cos\phi)$. Siten kiertokuvauksen matriisi on

$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

22 Ominaisarvot

Määritelmä 22.1. 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 $\bar{v} \neq \bar{0}$ ja

$$A\bar{v} = \lambda\bar{v}.$$

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

Matriisin A ominaisvektori on siis vektori, jolle matriisilla A kertominen vastaa reaaliluvulla λ kertomista. Huomaa, että edellinen määritelmä on sekä ominaisarvon että ominaisvektorin määritelmä. Ominaisarvoa ei voida määritellä ilman ominaisvektoreita eikä ominaisvektoreista voida puhua mainitsematta ominaisarvoa.

Nollavektorin ei haluta olevan ominaisvektori, sillä jos niin olisi, kaikki reaaliluvut olisivat kaikkien matriisien ominaisarvoja.

Esimerkki 22.2. Matriisilla

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

on ominaisarvo 4, sillä

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Eräs ominaisarvoa 4 vastaava ominaisvektori on siis $(1, 1)$.

Samaa ominaisarvoa voi vastata useampi eri ominaisvektori. Esimerkiksi $(2, 2)$ on myös matriisin A ominaisarvoa 4 vastaava ominaisvektori, sillä

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 8 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Matriisilla A on toinenkin ominaisarvo. Huomataan nimittäin, että

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

Siten myös luku 2 on matriisin A ominaisarvo.

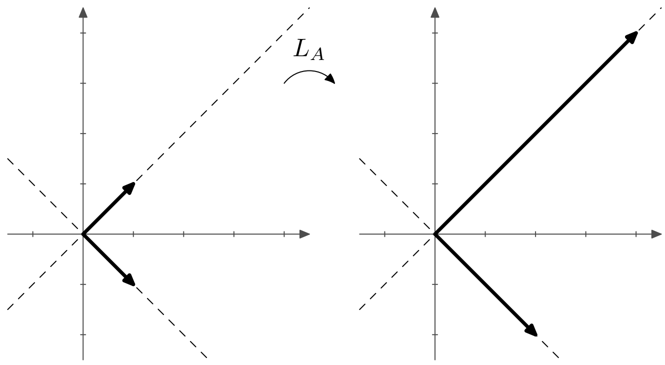
Esimerkki 22.3. Jokainen matriisi määrä lineaarikuvauksen. Kun matriisia ajatellaan lineaarikuvauksena, saavat ominaisarvot ja ominaisvektorit geometrisen tulkinnan.

Edellisen esimerkin matriisi

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

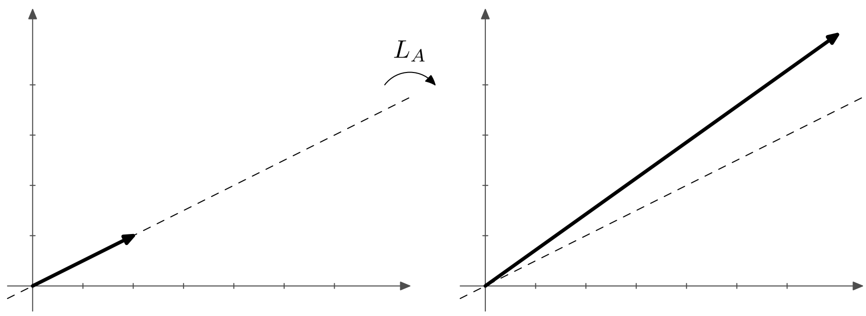
määrää lineaarikuvauksen $L_A \colon \mathbb{R}^2 \to \mathbb{R}^2$, $L_A(\bar{x}) = A\bar{x}$. Tässä kuvauksessa ominaisvektorin $(1, 1)$ kuvavektori on $(4, 4)$ ja ominaisvektorin $(2, 2)$ kuvavektori on $(8, 8)$.

Kuvaus siis venyttää ominaisarvoa 4 vastaavien ominaisvektoreiden pituuden nelinkertaiseksi. Ominaisarvoa 2 vastaava ominaisvektorin $(1, -1)$ pituus puolestaan kaksinkertaistuu.



Kuva 22.16: Vektorit $(1, 1)$ ja $(1, -1)$ ovat matriisin A ominaisvektoreita.

Jos sen sijaan tutkitaan vektoria $(2, 1)$ nähdään, että sen kuvavektori on $(7, 5)$. Vektori $(2, 1)$ ei ole ominaisvektori, koska sen kuvavektori ei ole vektorin $(2, 1)$ virittämällä suoralla.



Kuva 22.17: Vektori $(2, 1)$ ei ole matriisin A ominaisvektori.

Kun kaikki matriisin A ominaisarvoa λ vastaavat vektorit (sekä nollavektori) kerätään yhteen, saadaan ominaisarvoa vastaava ominaisavaruus.

Määritelmä 22.4. Oletetaan, että matriisilla $A \in \mathbb{R}^{n \times n}$ on ominaisarvo $\lambda \in \mathbb{R}$. Ominaisarvoa λ vastaava *ominaisavaruus* on joukko

$$V_\lambda = \{\bar{v} \in \mathbb{R}^n \mid A\bar{v} = \lambda\bar{v}.\}$$

Esimerkki 22.5. Määritetään esimerkin 22.2 matriisin

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

ominaisarvoa 4 vastaava ominaisavaruus eli kaikki ominaisarvoa 4 vastaavat ominaisvektorit.

Vektori $\bar{v} = (v_1, v_2) \in \mathbb{R}^2$ on ominaisarvoa 4 vastaava ominaisvektori, jos ja vain jos

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 4v_1 \\ 4v_2 \end{bmatrix}.$$

Tämä yhtälö toteutuu täsmälleen silloin, kun

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} - \begin{bmatrix} 4v_1 \\ 4v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

eli

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

Yhtälöä vastaa lineaarinen yhtälöryhmä

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

jossa muuttujina ovat v_1 ja v_2. Gaussin–Jordanin menetelmällä yhtälöryhmän ratkaisuiksi saadaan

$$\begin{cases} v_1 & = & s \\ v_2 & = & s, \end{cases} \qquad \text{missä } s \in \mathbb{R}.$$

Koska nollavektori ei ole ominaisvektori, ovat ominaisvektorit muotoa (s, s), missä $s \in \mathbb{R} \setminus \{0\}$.

Ominaisavaruuteen myös nollavektori otetaan mukaan. Siten ominaisarvoa 4 vastaava ominaisavaruus on

$$V_4 = \{(s, s) \mid s \in \mathbb{R}\}.$$

Huomataan, että ominaisavaruus on vektorin $(1, 1)$ virittämä aliavaruus:

$$V_4 = \{(s, s) \mid s \in \mathbb{R}\} = \{s(1, 1) \mid s \in \mathbb{R}\} = \operatorname{span}((1, 1)).$$

Ominaisavaruudet ovat aliavaruuksia aivan kuten niiden nimikin antaa olettaa.

Lause 22.6. *Matriisin $A \in \mathbb{R}^{n \times n}$ ominaisarvoa λ vastaava ominaisavaruus V_λ on vektoriavaruuden $\mathbb{R}^n$ aliavaruus.*

Todistus. Ensinnäkin V_λ on määritelmänsä perusteella joukon $\mathbb{R}^n$ osajoukko. Oletetaan, että $\bar{v}, \bar{w} \in V_\lambda$ ja $c \in \mathbb{R}$. Nyt $A\bar{v} = \lambda\bar{v}$ ja $A\bar{w} = \lambda\bar{w}$.

a) Ensinnäkin on osoitettava, että $\bar{v} + \bar{w} \in V_\lambda$. Matriisien laskusääntöjen perusteella

$$A(\bar{v} + \bar{w}) = A\bar{v} + A\bar{w} = \lambda\bar{v} + \lambda\bar{w} = \lambda(\bar{v} + \bar{w}).$$

Siis $\bar{v} + \bar{w} \in V_\lambda$.

b) Näytetään sitten, että $c\bar{v} \in V_\lambda$. Koska

$$A(c\bar{v}) = c(A\bar{v}) = c(\lambda\bar{v}) = c\lambda(\bar{v}),$$

pätee $\bar{v} \in V_\lambda$.

c) Lopuksi huomataan, että $A\bar{0} = \bar{0} = \lambda\bar{0}$, joten $\bar{0} \in V_\lambda$.

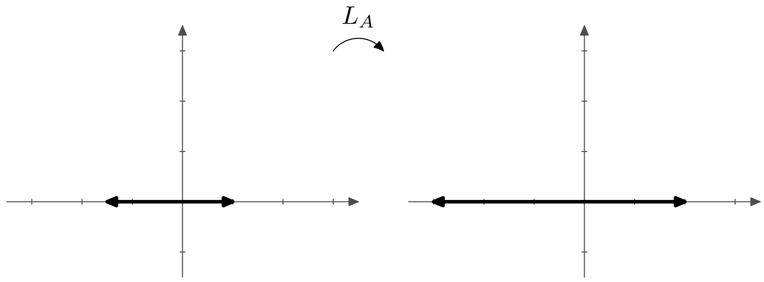
□

Esimerkki 22.7. Toisinaan matriisin ominaisarvot ja -vektorit voi päätellä tarkastelemalla lineaarikuvausta, joka matriisista saadaan. Tutkitaan esimerkin 18.9 matriisien

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{ja} \quad C = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

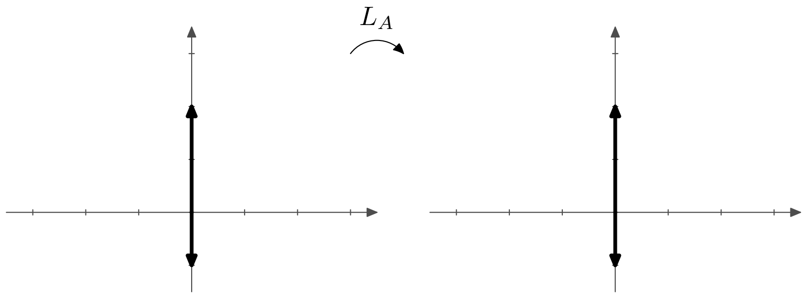
ominaisarvoja ja ominaisvektoreita.

Matriisia A vastaava lineaarikuvaus L_A venyttää vektoreita vaaka-akselin suunnassa kaksinkertaisiksi. Tällaisessa kuvauksessa vektorin $(1, 0)$ suuntaiset vektorit venyvät kaksinkertaisiksi. Tästä voidaan päätellä, että vektorit $t(1, 0)$, missä $t \in \mathbb{R} \setminus \{0\}$, ovat matriisin A ominaisvektoreita. Vastaava ominaisarvo on 2.



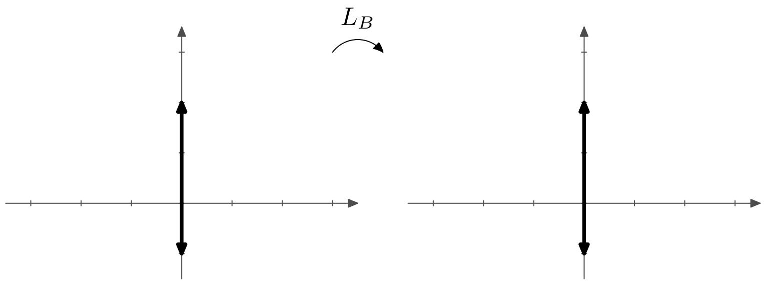
Kuva 22.18: Vaaka-akselin suuntaiset vektorit ovat matriisin A ominaisvektoreita.

Toisaalta vektorin $(0, 1)$ suuntaisille vektoreille ei tapahdu mitään. Siten myös vektorit $t(0, 1)$, missä $t \in \mathbb{R}\setminus\{0\}$, ovat matriisin A ominaisvektoreita. Niitä vastaava ominaisarvo on 1.



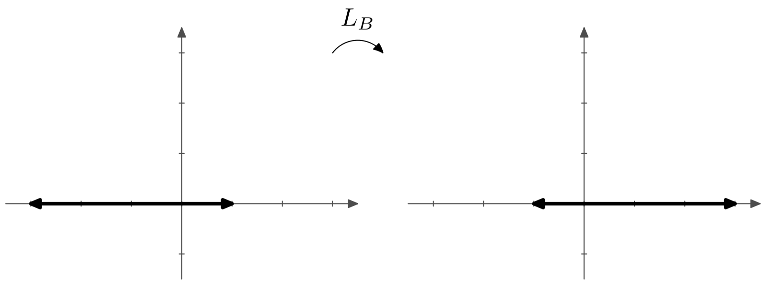
Kuva 22.19: Pystyakselin suuntaiset vektorit ovat matriisin A ominaisvektoreita.

Matriisia B vastaava lineaarikuvaus L_B peilaa vektorit pystyakselin suhteen. Tällaisessa kuvauksessa vektorin $(1, 0)$ suuntaisille vektoreille ei tapahdu mitään. Näin ollen vektorit $t(1, 0)$, missä $t \in \mathbb{R} \setminus \{0\}$, ovat matriisin B ominaisvektoreita, ja vastaava ominaisarvo on 1.



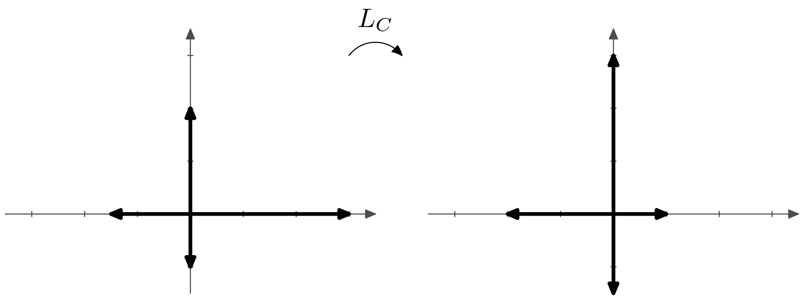
Kuva 22.20: Pystyakselin suuntaiset vektorit ovat matriisin B ominaisvektoreita.

Vektorin $(0, 1)$ suuntaiset vektorit kuvautuvat vastavektoreikseen. Siis vektorit $t(0, 1)$, missä $t \in \mathbb{R} \setminus \{0\}$, ovat ominaisvektoreita. Niitä vastaava ominaisarvo on -1.



Kuva 22.21: Vaaka-akselin suuntaiset vektorit ovat matriisin B ominaisvektoreita.

Matriisia C vastaava lineaarikuvaus L_C kiertää vektoreita origon ympäri $90°$ vastapäivään eli positiiviseen kiertosuuntaan. Kuvauksessa L_C kaikkien nollasta poikkeavien vektorien suunta muuttuu $90°$. Tästä voidaan päätellä, että matriisilla C ei ole ominaisvektoreita eikä ominaisarvoja.



Kuva 22.22: Matriisilla C kertominen muuttaa vektoreiden suuntaa $90°$.

22.1 Karakteristinen polynomi

Seuraava lause antaa suoraviivaisen tavan löytää matriisin ominaisarvot.

Lause 22.8. *Oletetaan, että* $A \in \mathbb{R}^{n\times n}$*. Reaaliluku* λ *on matriisin* A *ominaisarvo, jos ja vain jos* $\det(A - \lambda I) = 0$.

Todistus. "$\Rightarrow$": Oletetaan, että $\lambda \in \mathbb{R}$ on matriisin A ominaisarvo. Nyt on olemassa $\bar{v} \in \mathbb{R}^n \setminus \{\bar{0}\}$, jolle pätee $A\bar{v} = \lambda\bar{v}$. Vektori $\lambda\bar{v}$ voidaan yhtä hyvin kirjoittaa muodossa $\lambda I\bar{v}$, jolloin kyseessä onkin matriisin λI ja vektorin $\bar{v}$ matriisitulo. Nyt tutkittavana on yhtälö $A\bar{v} = \lambda I\bar{v}$. Matriisien laskusääntöjen nojalla tästä saadaan $A\bar{v} - \lambda I\bar{v} = \bar{0}$ ja edelleen $(A - \lambda I)\bar{v} = \bar{0}$.

Koska $(A - \lambda I)\bar{v} = \bar{0}$, yhtälöllä $(A - \lambda I)\bar{x} = \bar{0}$ on epätriviaali (eli nollasta poikkeava) ratkaisu $\bar{x} = \bar{v}$. Nyt nähdään, että lauseen 10.7 nojalla $A - \lambda I$ ei ole kääntyvä. Lauseen 11.4 perusteella puolestaan $\det(A - \lambda I) = 0$.

"$\Leftarrow$": Oletetaan sitten, että $\det(A-\lambda I) = 0$ jollakin reaaliluvulla λ. Nyt matriisi $A - \lambda I$ ei ole kääntyvä. Tästä seuraa, että yhtälöllä $(A - \lambda I)\bar{x} = \bar{0}$ on epätriviaali ratkaisu. Olkoon tuo ratkaisu $\bar{v}$. Nyt siis $\bar{v} \neq \bar{0}$. Koska $(A - \lambda I)\bar{v} = \bar{0}$, saadaan matriisien laskusääntöjen avulla yhtälö $A\bar{v}-\lambda I\bar{v} = \bar{0}$, mistä seuraa samaan tapaan kuin yllä, että $A\bar{v} = \lambda\bar{v}$. Siten λ on matriisin A ominaisarvo. □

Määritelmä 22.9. Oletetaan, että A on $n \times n$-matriisi. Muuttujan λ polynomi, joka saadaan kirjoittamalla auki determinantti $\det(A-\lambda I)$, on nimeltään matriisin A *karakteristinen poynomi.*

Lauseen 22.8 perusteella matriisin ominaisarvot ovat sen karakteristisen polynomin nollakohdat.

Esimerkki 22.10. Määritetään matriisin $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ ominaisarvot karakteristisen polynomin avulla. Nähdään, että

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

Nyt

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda) - 6 \\ &= 2 - \lambda - 2\lambda + \lambda^2 - 6 = \lambda^2 - 3\lambda - 4, \end{aligned}$$

joten A:n karakteristinen polynomi on $\lambda^2 - 3\lambda - 4$.

Matriisin A ominaisarvot ovat nyt yhtälön $\lambda^2 - 3\lambda - 4 = 0$ ratkaisut. Toisen asteen ratkaisukaavan avulla saadaan yhtälön ratkaisuiksi $\lambda = 4$ ja $\lambda - 1$. Siten matriisin ominaisarvot ovat 4 ja -1.

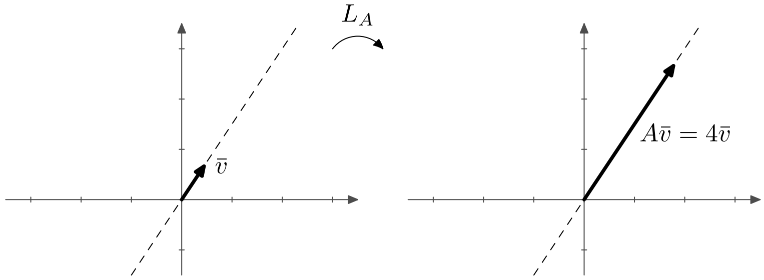
Selvitetään vielä ominaisarvoja vastaavat ominaisavaruudet. Ominaisarvoa 4 vastaava ominaisavaruus on joukko $V_4 = \{\bar{v} \in \mathbb{R}^2 \mid A\bar{v} = 4\bar{v}\}$. Sen alkiot löydetään ratkaisemalla yhtälö $A\bar{x} = 4\bar{x}$ eli yhtälö $(A - 4I)\bar{x} = \bar{0}$. Tätä yhtälöä vastaavan yhtälöryhmän ratkaisut ovat

$$\begin{cases} x_1 = (2/3)t \\ x_2 = t \end{cases}, \quad \text{missä } t \in \mathbb{R}.$$

Näin ollen

$$\begin{aligned} V_4 &= \{\bar{v} \in \mathbb{R}^2 \mid A\bar{v} = 4\bar{v}\} = \{((2/3)t, t) \mid t \in \mathbb{R}\} = \{t(2/3, 1) \mid t \in \mathbb{R}\} \\ &= \operatorname{span}\big((2/3, 1)\big) = \operatorname{span}\big((2, 3)\big). \end{aligned}$$

Matriisin A määräämä kuvaus L_A venyttää ominaisavaruuden $V_4 = \operatorname{span}\big((2,3)\big)$ vektorit nelinkertaisiksi, kuten kuvasta 22.23 nähdään.



Kuva 22.23: Ominaisavaruuden V_4 vektorit nelinkertaistuvat.

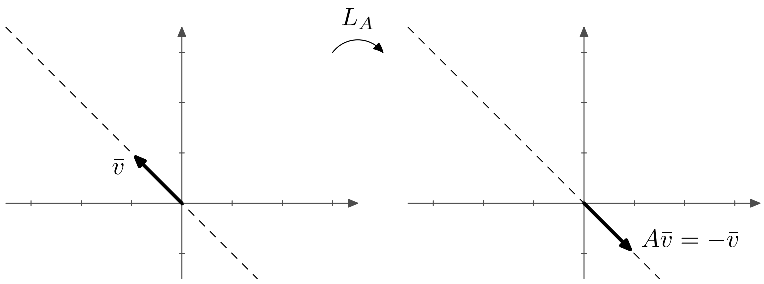
Ominaisarvoa -1 vastaava ominaisavaruus on $V_{-1} = \{\bar{v} \in \mathbb{R}^2 \mid A\bar{v} = -\bar{v}\}$. Sen alkiot löydetään vastaavasti ratkaisemalla yhtälö $A\bar{x} = -\bar{x}$ eli yhtälö $(A+I)\bar{x} = \bar{0}$. Tätä yhtälöä vastaavan yhtälöryhmän ratkaisut ovat

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

Näin ollen

$$\begin{aligned} V_{-1} &= \{\bar{v} \in \mathbb{R}^2 \mid A\bar{v} = -\bar{v}\} = \{(-t, t) \mid t \in \mathbb{R}\} = \{t(-1, 1) \mid t \in \mathbb{R}\} \\ &= \operatorname{span}\big((-1, 1)\big). \end{aligned}$$

Matriisin A määräämä kuvaus L_A kääntää ominaisavaruuden $V_{-1} = \operatorname{span}\big((-1,1)\big)$ vektorien suunnan päinvastaiseksi (ks. kuva 22.24).



Kuva 22.24: Ominaisavaruuden V_{-1} vektorien suunta vaihtuu päinvastaiseksi.

Lause 22.11. *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. □

Eri ominaisarvoja vastaavat ominaisvektorit ovat lineaarisesti riippumattomia.

Lause 22.12. *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 16.5 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

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

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 (1) 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

$$c_1\lambda_1\bar{v}_1 + \cdots + c_k\lambda_k\bar{v}_k = \lambda_{k+1}\bar{v}_{k+1}. \tag{2}$$

Toisaalta voidaan kertoa yhtälön (1) molemmat puolet luvulla λ_{k+1} päätyen yhtälöön

$$c_1\lambda_{k+1}\bar{v}_1 + \cdots + c_k\lambda_{k+1}\bar{v}_k = \lambda_{k+1}\bar{v}_{k+1}. \tag{3}$$

Vähennetään yhtälöstä (2) puolittain yhtälö (3), 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. □

22.2 Diagonalisointi

Joistakin matriiseista voidaan muodostaa lävistäjämatriiseja, joilla on samoja ominaisuuksia kuin alkuperäisellä matriisilla. Lävistäjämatriisit ovat hyvin yksinkertaisia matriiseja, joita on helppo käsitellä. Siksi monissa tapauksissa laskut helpottuvat suuresti, kun siirrytään käyttämään lävistäjämatriisia.

Tässä luvussa käsitellään similaarisia matriiseja, joiden ominaisuudet muistuttavat toisiaan. Lisäksi esitellään diagonalisoinnin käsite. Diagonalisoinnissa lähdetään liikkeelle neliömatriisista A ja etsitään lävistäjämatriisi, joka on similaarinen matriisin A kanssa.

Määritelmä 22.13. Matriisi $A \in \mathbb{R}^{n\times n}$ on *similaarinen* matriisin $B \in \mathbb{R}^{n\times n}$ kanssa, jos on olemassa kääntyvä matriisi $P \in \mathbb{R}^{n\times n}$, jolle pätee

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

Tällöin merkitään $A \sim B$.

Similaarisuus on määritelty vain neliömatriiseille. Huomaa, että $P^{-1}AP = B$, jos ja vain jos $AP = PB$.

Esimerkki 22.14. Merkitään

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

Matriisi A on similaarinen matriisin B kanssa. Tämän osoittamiseen voidaan käyttää matriisia

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

Ensinnäkin P on kääntyvä, sillä $\det(P) = 1 - (-1) = 2 \neq 0$. Lisäksi

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

joten $AP = PB$. Tästä seuraa, että

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

Lause 22.15. *Oletetaan, että A, B ja C ovat $n \times n$-matriiseja. Tällöin*

a) $A \sim A$

b) jos $A \sim B$, niin $B \sim A$

c) jos $A \sim B$ ja $B \sim C$, niin $A \sim C$.

Todistus. Osoitetaan kohta b) ja jätetään loput kohdat harjoitustehtäviksi. Oletetaan, että $A \sim B$. Tällöin on olemassa kääntyvä matriisi P, jolla $P^{-1}AP = B$. Kertomalla tätä yhtälöä vasemmalta matriisilla P ja oikealta matriisilla P^{-1} saadaan $A = PBP^{-1}$.

Merkitään $Q = P^{-1}$. Tällöin Q on kääntyvä ja

$$Q^{-1}BQ = (P^{-1})^{-1}BP^{-1} = PBP^{-1} = A.$$

Siis $B \sim A$. □

Similaariset matriisit muistuttavat toisiaan monessa suhteessa. Niillä on esimerkiksi sama determinantti ja samat ominaisarvot.

Lause 22.16. *Oletetaan, että A ja B ovat $n \times n$ -neliömatriiseja. Oletetaan lisäksi, että $A \sim B$. Tällöin*

a) $\det(A) = \det(B)$

b) A on kääntyvä, jos ja vain jos B on kääntyvä

c) matriiseilla A ja B on sama karakteristinen polynomi

d) matriiseilla A ja B on samat ominaisarvot.

Todistus. Todistetaan kohta c) ja jätetään loput kohdat harjoitustehtäviksi. Matriisin B karakteristinen polynomi on $\det(B - \lambda I)$. Koska $A \sim B$, on olemassa kääntyvä matriisi P, jolle pätee $P^{-1}AP = B$. Näin ollen

$$\begin{aligned}
\det(B - \lambda I) &= \det(P^{-1}AP - \lambda I) \\
&= \det(P^{-1}AP - \lambda P^{-1}IP) \\
&= \det(P^{-1}AP - P^{-1}\lambda IP) \\
&= \det(P^{-1}(AP - \lambda IP)) \\
&= \det(P^{-1}(A - \lambda I)P) \\
&= \det(P^{-1})\det(A - \lambda I)\det(P) \\
&= \left(\frac{1}{\det(P)}\right)\det(A - \lambda I)\det(P) \\
&= \det(A - \lambda I).
\end{aligned}$$

Siis matriisien B ja A karakteristiset polynomit $\det(B - \lambda I)$ ja $\det(A - \lambda I)$ ovat samat. $\square$

Esimerkki 22.17. Edellinen tulos on käyttökelpoinen, jos pitää osoittaa, että jotkin matriisit *eivät* ole similaariset.

Merkitään

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

Huomataan, että $\det(A) = -3$ ja $\det(B) = 3$. Siis

$$\det(A) \neq \det(B),$$

joten matriisit A ja B eivät ole similaariset.

Esimerkki 22.18. Huomaa, että vaikka jokin lauseen 22.16 ehdoista pätisikin, se ei takaa, että matriisit ovat similaariset. Tutkitaan vaikkapa matriiseja

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

Huomataan, että $\det(A) = -4 = \det(B)$. Kuitenkin matriisin A karakteristinen polynomi on

$$\det(A - \lambda I) = (1 - \lambda)(2 - \lambda) - 6 = \lambda^2 - 3\lambda - 4$$

ja matriisin B karakteristinen polynomi on

$$\det(B - \lambda I) = (1 - \lambda)(-1 - \lambda) - 3 = \lambda^2 - 4.$$

Siten matriisit A ja B eivät ole similaariset lauseen 22.16 kohdan c) nojalla.

Jopa kaikki lauseen ehdot voivat päteä ilman, että matriisit ovat similaarisia. Tarkastellaan matriiseja

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{ja} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Huomataan, että $\det(I) = 1 = \det(B)$. Tästä seuraa, että molemmat matriisit ovat kääntyviä. Lisäksi matriiseilla I ja B on sama karakteristinen polynomi $(1 - \lambda)^2$, ja kummankin matriisin ainoa ominaisarvo on 1. Tästä huolimatta matriisit I ja B eivät ole similaariset: jos P on mikä tahansa kääntyvä matriisi, niin

$$P^{-1} I P = I \neq B.$$

Määritelmä 22.19. Neliömatriisi A on *diagonalisoituva*, jos A on similaarinen jonkin lävistäjämatriisin kanssa.

Toisin sanoen A on diagonalisoituva, jos ja vain jos on olemassa kääntyvä matriisi P ja lävistäjämatriisi D, joille pätee $P^{-1}AP = D$. Matriisien P ja D etsimistä kutsutaan matriisin A *diagonalisoimiseksi.*

Esimerkki 22.20. Osoitetaan, että esimerkin 22.10 matriisi

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

on diagonalisoituva. Valitaan

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

Nyt matriisi P on kääntyvä, sillä $\det(P) = 2 - (-3) = 5 \neq 0$. Havaitaan, että

$$\begin{aligned} AP &= \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 1 \\ 12 & -1 \end{bmatrix} \quad \text{ja} \\ PD &= \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 8 & 1 \\ 12 & -1 \end{bmatrix}. \end{aligned}$$

Siis $AP = PD$, mistä seuraa, että

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

Huomaa, että oli ehdottoman tärkeää tarkistaa, että P on kääntyvä. Muuten ei voida puhua käänteismatriisista P^{-1}.

Määritelmän perusteella on melko hankalaa tutkia, onko matriisi diagonalisoituva. Jostakin on löydettävä matriisit P ja D, jotka toteuttavat määritelmän ehdon, tai todistettava, ettei tällaisia matriiseja ole. Tulemmekin myöhemmin esittelemään tuloksen, joka helpottaa diagonalisointia. Tutkitaan kuitenkin ensin esimerkkiä, joka näyttää yhden syyn sille, miksi matriisin diagonalisoiminen on hyödyllistä.

Esimerkki: Matriisin potenssien laskeminen

Lävistäjämatriisin potenssien laskeminen on vaivattomampaa kuin muiden matriisien. Induktiolla voidaan osoittaa, että mille tahansa kokonaisluvulle $k \geq 1$ ja lävistäjämatriisille

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 & 0 \\ 0 & d_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & d_{n-1} & 0 \\ 0 & 0 & \dots & 0 & d_n \end{bmatrix}$$

pätee

$$D^k = \begin{bmatrix} d_1^k & 0 & \dots & 0 & 0 \\ 0 & d_2^k & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & d_{n-1}^k & 0 \\ 0 & 0 & \dots & 0 & d_n^k \end{bmatrix}.$$

Myös diagonalisoituvan matriisin potenssien laskeminen on helppoa, sillä sen potenssit voidaan esittää lävistäjämatriisin potenssien avulla.

Esimerkki 22.21. Tutkitaan esimerkin 22.20 matriiseja

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

Esimerkissä osoitettiin, että matriisi A on diagonalisoituva ja $P^{-1}AP = D$. Kertomalla tätä yhtälöä vasemmalta matriisilla P ja oikealta matriisilla P^{-1} saadaan $A = PDP^{-1}$. Jos nyt $k \in \{1, 2, \dots\}$, niin

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

Matsiisin A potenssit voidaan siis laskea lävistäjämatriisin D avulla.

Potenssin laskemista varten on määritettävä matriisin P käänteismatriisi. Tämä voidaan tehdä luvussa 10.2 esitetyn menetelmän avulla. Käänteismatriisiksi saadaan

$$\frac{1}{5}\begin{bmatrix} 1 & 1 \\ -3 & 2 \end{bmatrix}.$$

Nyt voidaan laskea vaikkapa 10:s potenssi:

$$\begin{aligned} A^{10} = PD^{10}P^{-1} &= \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 4^{10} & 0 \\ 0 & (-1)^{10} \end{bmatrix}\frac{1}{5}\begin{bmatrix} 1 & 1 \\ -3 & 2 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} 2 \cdot 4^{10} + 3 & 2 \cdot 4^{10} - 2 \\ 3 \cdot 4^{10} - 3 & 3 \cdot 4^{10} + 2 \end{bmatrix}. \end{aligned}$$

Diagonalisoituvuuden selvittäminen

Palataan vielä esimerkin 22.20 matriisiin

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

Esimerkissä osoitettiin, että A on diagonalisoituva näyttämällä, että $P^{-1}AP = D$, missä

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

Tutkitaan, mistä matriisit P ja D tulevat.

Esimerkin 22.10 nojalla matriisin A ominaisarvot ovat 4 ja -1. Näitä vastaavat ominaisvektorit ovat muotoa $t(2,3)$ ja $s(1,-1)$, missä $t,s \in \mathbb{R} \setminus \{0\}$. Nähdään, että matriisin P sarakkeina ovat ominaisvektorit $(2,3)$ ja $(1,-1)$. Matriisin D lävistäjällä puolestaan ovat vastavat ominaisarvot eli 4 ja -1.

Osoittautuu, että jos matriisi on diagonalisoituva, tarvittavat matriisit P ja D voidaan aina löytää vastaavalla tavalla ominaisvektoreiden ja ominaisarvojen avulla. Oleellista on kuitenkin se, että käytettävät ominaisvektorit ovat lineaarisesti riippumattomia. (Edellisessä esimerkissä $(2,3)$ ja $(1,-1)$ ovat lineaarisesti riippumattomia.) Muutoin P ei ole kääntyvä.

Lause 22.22. *Matriisi $A \in \mathbb{R}^{n\times n}$ on diagonalisoituva, jos ja vain jos A:lla on n lineaarisesti riippumatonta ominaisvektoria.*

Todistus. "$\Rightarrow$": Oletetaan, että $P^{-1}AP = D$, missä $P \in \mathbb{R}^{n\times n}$ on jokin kääntyvä matriisi ja $D \in \mathbb{R}^{n\times n}$ lävistäjämatriisi. Nyt $AP = PD$. Olkoot P:n sarakkeet $\bar{p}_1, \dots, \bar{p}_n$ ja D:n 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 lävistäjämatriisilla D kerrottaessa tullaan kertoneeksi 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.

"$\Leftarrow$": Oletetaan, että $\bar{p}_1, \dots, \bar{p}_n$ ovat jotkin A:n lineaarisesti riippumattomat ominaisvektorit. Olkoot niitä vastaavat ominaisarvot $\lambda_1, \dots, \lambda_n$. Nyt $A\bar{p}_i = \lambda_i\bar{p}_i$

kaikilla $i \in \{1, \ldots, n\}$. Olkoon P matriisi, jonka sarakkeet ovat ominaisvektorit: $P = [\bar{p}_1 \cdots \bar{p}_n]$. Olkoon D puolestaan lävistäjämatriisi, jonka lävistäjäalkiot ovat $\lambda_1, \ldots, \lambda_n$. 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

$$P^{-1}AP = D. \qquad \square$$

Edellinen lause antaa suoraviivaisen keinon tarkistaa, onko matriisi diagonalisoituva. Lisäksi lauseen todistuksesta nähdään, miten diagonalisoituvuuden määritelmässä mainitut matrisiit P ja D löydetään. Kerätään nämä kaikki tiedot vielä yhteen.

Näin selvitetään, onko matriisi $A \in \mathbb{R}^{n \times n}$ diagonalisoituva:

1) Etsitään matriisin A ominaisarvot.
2) Määritetään jokaista ominaisarvoa vastaava ominaisavaruus.
3) Tutkitaan, löytyykö lineaarisesti riippumattomia ominaisvektoreita n kappaletta. Jos niitä on vähemmän kuin n kappaletta, matriisi A ei ole diagonalisoituva. Muussa tapauksessa A on diagonalisoituva, ja voidaan jatkaa seuraavaan kohtaan.
4) Muodostetaan matriisi P laittamalla löydetyt lineaarisesti riippumattomat ominaisvektorit sen sarakkeiksi. Tällöin P on lauseen 22.22 todistuksen nojalla kääntyvä. (Tämän voi tarkistaa esimerkiksi determinantin avulla.)
5) Muodostetaan lävistäjämatriisi D laittamalla sen sarakkeisiin matriisin P sarakkeita vastaavat ominaisarvot. Tällöin $P^{-1}AP = D$ lauseen 22.22 todistuksen nojalla. (Tämän voi tarkistaa laskemalla tulot AP ja PD.)

Esimerkki 22.23. Merkitään

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

Diagonalisoidaan matriisi A, jos mahdollista.

Aloitetaan määrittämällä matriisin A ominaisarvot. Matriisin karakteristinen polynomi on

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

Ominaisarvot löydetään siis ratkaisemalla yhtälö $-\lambda^2(\lambda+2)=0$. Näin ollen ominaisarvot ovat 0 ja -2.

Määritetään seuraavaksi ominaisarvoja vastaavat ominaisavaruudet. Ominaisarvoa 0 vastaava ominaisavaruus on

$$V_0 = \{\bar{v} \in \mathbb{R}^3 \mid A\bar{v} = 0\bar{v}\}.$$

Sen alkioita ovat yhtälön $A\bar{x} = 0\bar{x}$ ratkaisut. Muodostamalla yhtälöä vastaava yhtälöpari ja ratkaisemalla se nähdään, että ratkaisut ovat muotoa $\bar{x} = (t, s, t)$, missä s, $t \in \mathbb{R}$. Siis

$$\begin{aligned} V_0 &= \{(t,s,t) \mid s,t \in \mathbb{R}\} = \{t(1,0,1) + s(0,1,0) \mid s,t \in \mathbb{R}\} \\ &= \operatorname{span}\big((1,0,1),\ (0,1,0)\big). \end{aligned}$$

Ominaisarvoa -2 vastaava ominaisavaruus on

$$V_{-2} = \{\bar{v} \in \mathbb{R}^3 \mid A\bar{v} = -2\bar{v}\}.$$

Kun ratkaistaan yhtälö $A\bar{x} = -2\bar{x}$ Gaussin–Jordanin menetelmällä, nähdään, että ratkaisut ovat muotoa $\bar{x} = (-t, 3t, t)$, missä $t \in \mathbb{R}$. Siis

$$V_{-2} = \{(-t,3t,t) \mid t \in \mathbb{R}\} = \{t(-1,3,1) \mid t \in \mathbb{R}\} = \operatorname{span}\big((-1,3,1)\big).$$

Matriisi A on diagonalisoituva, jos ja vain jos on olemassa kolme lineaarisesti riippumatonta ominaisvektoria. Osoitetaan, että ominaisvektorit $\bar{p}_1 = (1,0,1)$, $\bar{p}_2 = (0,1,0)$ ja $\bar{p}_3 = (-1,3,1)$ ovat lineaarisesti riippumattomat. On siis tutkittava yhtälöä

$$c_1(1,0,1) + c_2(0,1,0) + c_3(0,1,0) = (0,0,0),$$

missä $c_1, c_2, c_3 \in \mathbb{R}$. Muodostamalla yhtälöä vastaava yhtälöryhmä ja ratkaisemalla se Gaussin–Jordanin eliminointimenetelmällä nähdään, että yhtälöryhmän ainoa ratkaisu on $c_1 = 0$, $c_2 = 0$, $c_3 = 0$. Siten tutkittavat vektorit ovat lineaarisesti riippumattomia. Näin ollen A on diagonalisoituva lauseen 22.22 nojalla.

Merkitään

$$P = \begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{ja} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Tällöin lauseen 22.22 todistuksen mukaan P on kääntyvä ja $P^{-1}AP = D$.

Jos on epävarma ratkaisustaan, voi vielä lopukasi tarkistaa, että $\det(P) \neq 0$ ja $AP = PD$.

Esimerkki 22.24. Diagonalisoidaan matriisi

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

jos mahdollista. Selvitetään aluksi matriisin ominaisarvot. Matriisin karakteristinen polynomi on $(\lambda - 2)^2$, ja tämän polynomin ainoa nollakohta on 2. Siten ominaisarvoja on yksi, ja se 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.

Toisinaan matriisin diagonalisoituvuus on helppo todeta.

Lause 22.25. *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 22.12 nojalla. Koska matriisilla A on n lineaarisesti riippumatonta ominaisvektoria, on A diagonalisoituva lauseen 22.22 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 A on similaarinen itsensä kanssa. Kuitenkin karakteristisen polynomin avulla nähdään, että matriisilla on vain yksi ominaisarvo, -3.

23 Sisätulo

Ottamalla lähtökohdaksi avaruuden $\mathbb{R}^n$ vektorien pistetulon ominaisuudet, voidaan määritellä vektoriavaruuteen V yleisempi sisätulon käsite.

Määritelmä 23.1. Vektoriavaruuden V *sisätulo* on sääntö, joka liittää jokaiseen vektoriavaruuden V alkiopariin $(\bar{v}, \bar{w})$ yksikäsitteisen reaaliluvun $\langle \bar{v}, \bar{w} \rangle$. Lisäksi sisätulon on toteutettava seuraavat ehdot kaikilla $\bar{v}, \bar{w}, \bar{u} \in V$ ja $c \in \mathbb{R}$:

1) $\langle \bar{v}, \bar{w} \rangle = \langle \bar{w}, \bar{v} \rangle$
2) $\langle \bar{v}, \bar{w} + \bar{u} \rangle = \langle \bar{v}, \bar{w} \rangle + \langle \bar{v}, \bar{u} \rangle$
3) $\langle c\bar{v}, \bar{w} \rangle = c\langle \bar{v}, \bar{w} \rangle$
4) $\langle \bar{v}, \bar{v} \rangle \geq 0$; lisäksi $\langle \bar{v}, \bar{v} \rangle = 0$ jos ja vain jos $\bar{v} = \bar{0}$.

Vektoriavaruutta, jossa on määritelty sisätulo, kutsutaan *sisätuloavaruudeksi*. Huomaa, että ehdoista 1 ja 2 seuraa, että $\langle \bar{v} + \bar{u}, \bar{w} \rangle = \langle \bar{v}, \bar{w} \rangle + \langle \bar{u}, \bar{w} \rangle$ kaikilla $\bar{v}, \bar{w}, \bar{u} \in V$. Samalla tavalla $\langle \bar{v}, c\bar{w} \rangle = c\langle \bar{v}, \bar{w} \rangle$ kaikilla $\bar{v}, \bar{w} \in V$ ja $c \in \mathbb{R}$.

Esimerkki 23.2. Vektoriavaruuden $\mathbb{R}^n$ pistetulo on sisätulo. Tässä tapauksessa $\langle \bar{v}, \bar{w} \rangle = \bar{v} \cdot \bar{w}$, kun $\bar{v}, \bar{w} \in V$.

Esimerkki 23.3. Vektoriavaruudessa $\mathbb{R}^n$ voidaan määritellä myös muita sisätuloja. Osoitetaan, että kaava $\langle \bar{v}, \bar{w} \rangle = v_1w_1 + 2v_2w_2$ määrittää avaruuden $\mathbb{R}^2$ sisätulon. Oletetaan, että $\bar{v}, \bar{w}, \bar{u} \in V$ ja $c \in \mathbb{R}$.

1) Reaalilukujen kertolaskun vaihdannaisuudesta seuraa, että
$$\langle \bar{v}, \bar{w} \rangle = v_1w_1 + 2v_2w_2 = w_1v_1 + 2w_2v_2 = \langle \bar{w}, \bar{v} \rangle.$$
Siis ehto 1 pätee.

2) Ehto 2 saadaan puolestaan reaalilukujen osittelulaista sekä yhteenlaskun liitännäisyydestä ja vaihdannaisuudesta:
$$\begin{aligned}\langle \bar{v}, \bar{w} + \bar{u} \rangle &= v_1(w_1 + u_1) + 2v_2(w_2 + u_2)\\ &= v_1w_1 + v_1u_1 + 2v_2w_2 + 2v_2u_2\\ &= v_1w_1 + 2v_2w_2 + v_1u_1 + 2v_2u_2\\ &= \langle \bar{v}, \bar{w} \rangle + \langle \bar{v}, \bar{u} \rangle.\end{aligned}$$

3) Myös ehto 3 pätee, sillä
$$\langle c\bar{v}, \bar{w} \rangle = cv_1w_1 + 2cv_2w_2 = c(v_1w_1 + 2v_2w_2) = c\langle \bar{v}, \bar{w} \rangle.$$

4) Nähdään, että $\langle \bar{v}, \bar{v} \rangle = v_1v_1 + 2v_2v_2 = v_1^2 + v_2^2 \geq 0$. Osoitetaan vielä, että $\langle \bar{v}, \bar{v} \rangle = 0$, jos ja vain jos $\bar{v} = \bar{0}$. Tämä täytyy tehdä kahdessa osassa.

"$\Rightarrow$": Oletetaan ensin, että $\langle \bar{v}, \bar{v} \rangle = 0$. Nyt $v_1^2 + 2v_2^2 = 0$. Koska kumpikaan yhtälön vasemman puolen summattavista ei ole negatiivinen, täytyy päteä $v_1^2 = 0$ ja $2v_2^2 = 0$. Tästä seuraa, että $v_1 = 0$ ja $v_2 = 0$. Siis $\bar{v} = \bar{0}$.

"$\Leftarrow$": Oletetaan sitten, että $\bar{v} = \bar{0}$. Nyt

$$\langle \bar{v}, \bar{v} \rangle = \langle \bar{0}, \bar{0} \rangle = 0^2 + 2 \cdot 0^2 = 0.$$

Siten viimeinenkin ehto pätee, ja kyseessä on sisätulo.

Esimerkki 23.4. Joukko

$$C([0,1]) = \{f\colon [0,1] \to \mathbb{R} \mid f \text{ on jatkuva}\}$$

on vektoriavaruus, kun yhteenlasku ja skalaarikertolasku määritellään pisteittäin samaan tapaan kuin esimerkissä 14.5. Vektoriavaruudessa $C([0,1])$ voidaan määritellä sisätulo kaavalla

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx.$$

Lemma 23.5. *Sisätuloavaruudessa V pätee $\langle \bar{v}, \bar{0} \rangle = 0$ ja $\langle \bar{0}, \bar{v} \rangle = 0$ kaikilla $\bar{v} \in V$.*

Todistus. Oletetaan, että V on sisätuloavaruus ja $\bar{v} \in V$. Sisätulon määritelmän mukaan

$$\langle \bar{v}, \bar{0} \rangle = \langle \bar{v}, \bar{0} + \bar{0} \rangle = \langle \bar{v}, \bar{0} \rangle + \langle \bar{v}, \bar{0} \rangle.$$

Vähentämällä yhtälön molemmilta puolilta luku $\langle \bar{v}, \bar{0} \rangle$, saadaan $\langle \bar{v}, \bar{0} \rangle = 0$. Lisäksi sisätulon määritelmän perusteella $\langle \bar{0}, \bar{v} \rangle = \langle \bar{v}, \bar{0} \rangle = 0$. □

23.1 Normi ja kohtisuoruus

Sisätuloavaruudessa voidaan määritellä normi samaan tapaan kuin avaruuden $\mathbb{R}^n$ pistetulon tapauksessa.

Määritelmä 23.6. Olkoon V sisätuloavaruus. Tällöin vektorin $\bar{v} \in V$ *normi* on

$$\|\bar{v}\| = \sqrt{\langle \bar{v}, \bar{v} \rangle}.$$

Sisätulon määritelmän mukaan $\langle \bar{v}, \bar{v} \rangle \geq 0$ kaikilla $\bar{v} \in V$, joten normi on aina määritelty. Sisätuloavaruuden normille pätevät tutut tulokset.

Lause 23.7. *Oletetaan, että V on sisätuloavaruus, $\bar{v} \in V$ ja $c \in \mathbb{R}$. Tällöin*

a) $\|\bar{v}\| \geq 0$

b) $\|\bar{v}\| = 0$*, jos ja vain jos* $\bar{v} = 0$

c) $\|c\bar{v}\| = c\|\bar{v}\|$.

Todistus. Lauseen todistus on samanlainen kuin pistetulon tapauksessa (lauseet 12.5 ja 12.6). □

Esimerkki 23.8. Sisätuloavaruuden *yksikköympyrä* koostuu kaikista niistä vektoreista, joiden pituus on yksi. Esimerkiksi avaruuden $\mathbb{R}^2$ ja pistetulon tapauksessa yksikköympyrä on joukko

$$\{\bar{v} \in \mathbb{R}^2 \mid \|\bar{v}\| = 1\} = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1^2 + v_2^2 = 1\}.$$

Kun vektorit tulkitaan tason pisteiksi, voi yksikköympyrästä piirtää kuvan koordinaatistoon. Sen alkiot muodostavat ympyrän, jonka säde on yksi ja keskipiste $(0, 0)$.

Tutkitaan sitten esimerkissä 23.3 esiteltyä avaruuden $\mathbb{R}^2$ sisätuloa, joka määriteltiin kaavalla $\langle \bar{v}, \bar{w} \rangle = v_1 w_1 + 2v_2 w_2$. Tällöin yksikköympyrä on joukko

$$\{\bar{v} \in \mathbb{R}^2 \mid \|\bar{v}\| = 1\} = \{\bar{v} \in \mathbb{R}^2 \mid \langle \bar{v}, \bar{v} \rangle = 1\} = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1^2 + 2v_2^2 = 1\}.$$

Huomataan, että joukko $\{(v_1, v_2) \in \mathbb{R}^2 \mid v_1^2 + 2v_2^2 = 1\}$ on ellipsi, jonka keskipiste on origo ja akselien pituudet 2 ja $\sqrt{2}$. Tämän sisätulon tapauksessa yksikköympyrä ei siis näytäkään ympyrältä.

Sisätulon avulla voidaan määritellä vektorien kohtisuoruus.

Määritelmä 23.9. Olkoon V sisätuloavaruus. Oletetaan, että $\bar{v}, \bar{w} \in V$. Tällöin vektorit $\bar{v} \in V$ ja $\bar{w} \in V$ ovat *ortogonaaliset* eli *kohtisuorassa toisiaan vastaan*, jos

$$\langle \bar{v}, \bar{w} \rangle = 0.$$

Koulusta tuttu Pythagoraan lause voidaan yleistää mihin tahansa sisätuloavaruuteen.

Lause 23.10 (Pythagoraan lause). *Olkoon V sisätuloavaruus. Vektorit $\bar{v} \in V$ ja $\bar{w} \in V$ ovat ortogonaaliset, jos ja vain jos*

$$\|\bar{v}\|^2 + \|\bar{w}\|^2 = \|\bar{v} + \bar{w}\|^2.$$

Todistus. Todistus jätetään harjoitustehtäväksi. □

23.2 Kohtisuora komplementti

Aliavaruuden kohtisuora komplementti koostuu niistä vektoreista, jotka ovat kohtisuorassa kaikkia aliavaruuden vektoreita vastaan.

Määritelmä 23.11. Olkoon W sisätuloavaruuden V aliavaruus. Sen *kohtisuora komplementti* on joukko

$$W^\perp = \{\bar{v} \in V \mid \langle \bar{v}, \bar{w} \rangle = 0 \text{ kaikilla } \bar{w} \in W\}.$$

Esimerkki 23.12. Tarkastellaan avaruuden $\mathbb{R}^2$ aliavaruutta

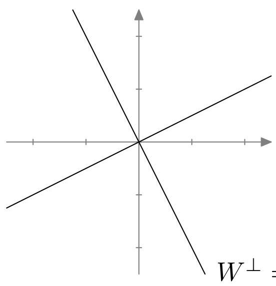
$$W = \operatorname{span}\big((2, 1)\big) = \{t(2, 1) \mid t \in \mathbb{R}\},$$

joka on vektorin $(2,1)$ suuntainen origon kautta kulkeva suora. Määritetään kohtisuora komplementti $W^\perp$, kun sisätulona on tavallinen pistetulo. Kohtisuora komplementti koostuu kaikista niistä vektoreista, jotka ovat kohtisuorassa suoraa vastaan. Kuvan perusteella voidaan päätellä, että kohtisuora komplementti on suora, joka on kohtisuorassa W:tä vastaan (ks. kuva 23.25).

Määritetään aliavaruuden kohtisuora komplementti $W^\perp$ vielä täsmällisesti. Nähdään, että

$$\begin{aligned}
W^\perp &= \{\bar{v} \in \mathbb{R}^2 \mid \bar{v} \cdot \bar{w} = 0 \text{ kaikilla } \bar{w} \in W\} \\
&= \{(v_1, v_2) \in \mathbb{R}^2 \mid (v_1, v_2) \cdot t(2,1) = 0 \text{ kaikilla } t \in \mathbb{R}\} \\
&= \{(v_1, v_2) \in \mathbb{R}^2 \mid t(2v_1 + v_2) = 0 \text{ kaikilla } t \in \mathbb{R}\} \\
&= \{(v_1, v_2) \in \mathbb{R}^2 \mid 2v_1 + v_2 = 0\} \\
&= \{(v_1, v_2) \in \mathbb{R}^2 \mid v_2 = -2v_1\} \\
&= \{(v_1, -2v_1) \mid v_1 \in \mathbb{R}\} \\
&= \{v_1(1,-2) \mid v_1 \in \mathbb{R}\} \\
&= \operatorname{span}\big((1,-2)\big).
\end{aligned}$$

Ortogonaalinen komplementti on siis origon kautta kulkeva vektorin $(1,-2)$ suuntainen suora.



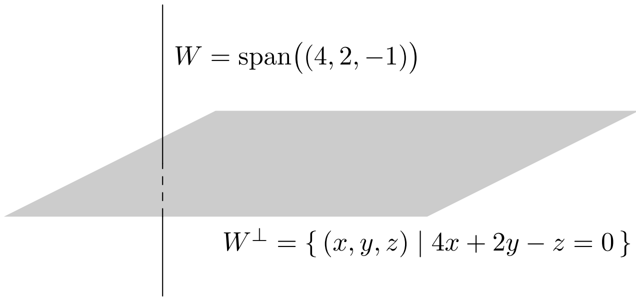
Kuva 23.25: Avaruuden $\mathbb{R}^2$ suoran $\operatorname{span}\big((2,1)\big)$ kohtisuora komplementti on suora $\operatorname{span}\big((1,-2)\big)$.

Esimerkki 23.13. Tarkastellaan avaruuden $\mathbb{R}^3$ aliavaruutta $W = \operatorname{span}\big((4,2,-1)\big)$, joka on vektorin $(4,2,-1)$ suuntainen, origon kautta kulkeva suora. Määritetään aliavaruuden kohtisuora komplementti $W^\perp$, kun sisätulona on tavallinen pistetulo. Kuvan perusteella vaikuttaisi siltä, että tämän suoran kohtisuora komplementti olisi taso, joka on kohtisuorassa suoraa vastaan (ks. kuva 23.26).

Määritetään kohtisuora komplementti vielä täsmällisesti. Nähdään, että

$$\begin{aligned}
W^\perp &= \{\bar{v} \in \mathbb{R}^3 \mid \bar{v} \cdot \bar{w} = 0 \text{ kaikilla } \bar{w} \in W\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid (v_1, v_2, v_3) \cdot t(4,2,-1) = 0 \text{ kaikilla } t \in \mathbb{R}\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid t(4v_1 + 2v_2 - v_3) = 0 \text{ kaikilla } t \in \mathbb{R}\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid 4v_1 + 2v_2 - v_3 = 0\}.
\end{aligned}$$

Ortogonaalinen komplementti on siis origon kautta kulkeva avaruuden $\mathbb{R}^3$ taso, jonka eräs normaali on vektori $(4, 2, -1)$.



Kuva 23.26: Avaruuden $\mathbb{R}^3$ suoran $\operatorname{span}\big((4,2,-1)\big)$ kohtisuora komplementti on taso $\{(v_1, v_2, v_3) \mid 4v_1 + 2v_2 - v_3 = 0\}$.

Aliavaruuden kohtisuora komplementti on aina aliavaruus.

Lause 23.14. *Olkoon W sisätuloavaruuden V aliavaruus. Sen kohtisuora komplementti $W^\perp$ on myös V:n aliavaruus.*

Todistus. Oletetaan, että $\bar{v}, \bar{u} \in W^\perp$ ja $c \in \mathbb{R}$. Nyt kaikilla $\bar{w} \in W$ pätee $\langle \bar{v}, \bar{w}\rangle = 0$ ja $\langle \bar{u}, \bar{w}\rangle = 0$. On osoitettava, että $\bar{v}+\bar{u} \in W^\perp$, $c\bar{v} \in W^\perp$ ja $\bar{0} \in W^\perp$. Toisin sanoen täytyy näyttää, että

$$\langle \bar{v} + \bar{u}, \bar{w}\rangle = 0, \quad \langle c\bar{v}, \bar{w}\rangle = 0 \quad \text{ja} \quad \langle \bar{0}, \bar{w}\rangle = 0$$

kaikilla $\bar{w} \in W$.

Oletetaan tätä varten, että $\bar{w} \in W$. Sisätulon määritelmän nojalla

$$\langle \bar{v} + \bar{u}, \bar{w}\rangle = \langle \bar{v}, \bar{w}\rangle + \langle \bar{u}, \bar{w}\rangle = 0 + 0 = 0.$$

Samalla tavalla nähdään, että

$$\langle c\bar{v}, \bar{w}\rangle = c\langle \bar{v}, \bar{w}\rangle = c \cdot 0 = 0.$$

Lisäksi lauseen 23.5 nojalla pätee $\langle \bar{0}, \bar{w}\rangle = 0$.

Siten $\langle \bar{v}+\bar{u}, \bar{w}\rangle = 0$, $\langle c\bar{v}, \bar{w}\rangle = 0$ ja $\langle \bar{v}, \bar{w}\rangle = 0$ kaikilla $\bar{w} \in W$, joten $\bar{v}+\bar{u} \in W^\perp$, $c\bar{v} \in W^\perp$ ja $\bar{0} \in W^\perp$. Näin on osoitettu, että $W^\perp$ on aliavaruus. □

Jos vektori on kohtisuorassa aliavaruuden virittäjävektoreita vastaan, se on kohtisuorassa kaikkia aliavaruuden vektoreita vastaan. Pelkkien virittäjävektorien tarkasteleminen siis riittää.

Lause 23.15. *Oletetaan, että W on sisätuloavaruuden V aliavaruus. Oletetaan lisäksi, että $W = \mathrm{span}(\bar{w}_1, \dots, \bar{w}_k)$ ja $\bar{v} \in V$. Jos $\langle \bar{v}, \bar{w}_i \rangle = 0$ kaikilla $i \in \{1, \dots, k\}$, niin $\bar{v} \in W^\perp$.*

Todistus. On osoitettava, että $\langle \bar{v}, \bar{w} \rangle = 0$ kaikilla $\bar{w} \in W$. Oletetaan siis, että $\bar{w} \in W$. Tällöin $\bar{w} = a_1\bar{w}_1 + a_2\bar{w}_2 + \dots + a_k\bar{w}_k$ joillakin $a_1, \dots, a_k \in \mathbb{R}$. Sisätulon määritelmän ehtojen nojalla

$$\begin{aligned}
\langle \bar{v}, \bar{w} \rangle &= \langle \bar{v}, a_1\bar{w}_1 + a_2\bar{w}_2 + \dots + a_k\bar{w}_k \rangle \\
&= \langle \bar{v}, a_1\bar{w}_1 \rangle + \langle \bar{v}, a_2\bar{w}_2 \rangle + \dots + \langle \bar{v}, a_k\bar{w}_k \rangle \\
&= a_1\langle \bar{v}, \bar{w}_1 \rangle + a_2\langle \bar{v}, \bar{w}_2 \rangle + \dots + a_k\langle \bar{v}, \bar{w}_k \rangle \\
&= a_1 \cdot 0 + a_2 \cdot 0 + \dots + a_k \cdot 0 = 0\,.
\end{aligned}$$

Siten $\langle \bar{v}, \bar{w} \rangle = 0$, olipa $\bar{w}$ mikä tahansa aliavaruuden W vektori. Näin ollen $\bar{v} \in W^\perp$. □

Esimerkki 23.16. Tarkastellaan vektoriavaruuden $\mathbb{R}^3$ aliavaruutta

$$W = \mathrm{span}\big((1,0,3)\,(2,1,5)\big),$$

joka on vektoreiden $\bar{w}_1 = (1,0,3)$ ja $\bar{w}_2 = (2,1,5)$ virittämä, origon kautta kulkeva taso. Kun määritetään tämän aliavaruuden kohtisuoraa komplementtia, riittää tarkastella pelkkiä virittäjävektoreita. Nähdään, että

$$\begin{aligned}
W^\perp &= \{\bar{v} \in \mathbb{R}^3 \mid \langle \bar{v}, \bar{w} \rangle = 0 \text{ kaikilla } \bar{w} \in W\} \\
&= \{\bar{v} \in \mathbb{R}^3 \mid \langle \bar{v}, \bar{w}_1 \rangle = 0 \text{ ja } \langle \bar{v}, \bar{w}_2 \rangle = 0\} \\
&= \{\bar{v} \in \mathbb{R}^3 \mid \bar{v} \cdot \bar{w}_1 = 0 \text{ ja } \bar{v} \cdot \bar{w}_2 = 0\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 + 3v_3 = 0 \text{ ja } 2v_1 + v_2 + 5v_3 = 0\}.
\end{aligned}$$

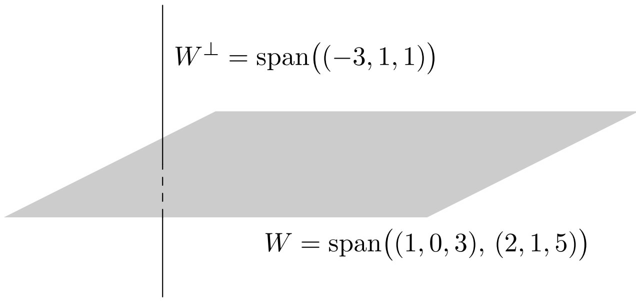
On siis ratkaistava yhtälöpari

$$\begin{cases} v_1 \phantom{{}+v_2} + 3v_3 = 0 \\ 2v_1 + v_2 + 5v_3 = 0. \end{cases}$$

Sen ratkaisut ovat $\bar{v} = (-3t, t, t)$, missä $t \in \mathbb{R}$. Näin ollen

$$\begin{aligned}
W^\perp &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 + 3v_3 = 0 \text{ ja } 2v_1 + v_2 + 5v_3 = 0\} \\
&= \{t(-3,1,1) \mid t \in \mathbb{R}\} \\
&= \mathrm{span}\big((-3,1,1)\big).
\end{aligned}$$

Ortogonaalinen komplementti on siis vektorin $(-3,1,1)$ virittämä suora (ks. kuva 23.27).



Kuva 23.27: Avaruuden $\mathbb{R}^3$ tason $W = \operatorname{span}\big((1,0,3)\,(2,1,5)\big)$ kohtisuora komplementti on suora $\operatorname{span}\big((-3,1,1)\big)$.

Ainoa vektori, joka on sekä aliavaruudessa että sen kohtisuorassa komplementissa, on nollavektori.

Lause 23.17. *Olkoon W on sisätuloavaruuden V aliavaruus. Tällöin*

$$W \cap W^{\perp} = \{\bar{0}\}.$$

Todistus. "$\subset$": Oletetaan, että $\bar{u} \in W \cap W^{\perp}$. Nyt $\bar{u} \in W$ ja $\bar{u} \in W^{\perp}$. Kohtisuoran komplementin määritelmän mukaan tällöin pätee $\langle \bar{u}, \bar{w} \rangle = 0$ kaikilla $\bar{w} \in W$. Erityisesti $\langle \bar{u}, \bar{u} \rangle = 0$. Sisätulon määritelmästä seuraa, että $\bar{u} = \bar{0}$. Siis $W \cap W^{\perp} \subset \{\bar{0}\}$.

"$\supset$": Koska W ja $W^{\perp}$ ovat kumpikin aliavaruuksia, niin $\bar{0} \in W$ ja $\bar{0} \in W^{\perp}$. Siten $\bar{0} \in W \cap W^{\perp}$. Siis $\{\bar{0}\} \subset W \cap W^{\perp}$. $\square$

23.3 Kohtisuora projektio

Avaruudessa $\mathbb{R}^n$ vektorin $\bar{v}$ projektio vektorin $\bar{w}$ virittämälle alivaruudelle määriteltiin kaavalla

$$\mathrm{proj}_{\bar{w}}(\bar{v}) = \frac{\bar{v} \cdot \bar{w}}{\bar{w} \cdot \bar{w}} \bar{w}.$$

Ryhdymme nyt yleistämään projektion käsitettä. Uudessa määritelmässä aliavaruus, jolle projisoidaan, voi olla useamman kuin yhden vektorin virittämä. Lisäksi sisätuloavaruuden ei tarvitse olla avaruus $\mathbb{R}^n$ pistetulolla varustettuna, vaan voidaan käsitellä mitä tahansa sisätuloavaruutta.

Ennen kuin voimme antaa projektion yleisen määritelmän, on käsiteltävä ortogonaalisia jonoja. Sisätuloavaruuden V jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on *ortogonaalinen*, jos seuraavat ehdot pätevät:

a) $\langle \bar{v}_i, \bar{v}_j \rangle = 0$, kun $i \neq j$

b) $\bar{v}_i \neq \bar{0}$ kaikilla $i \in \{1, 2, \ldots, k\}$.

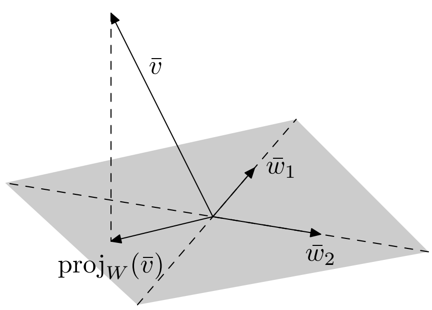
Toisin sanoen vektorit ovat kohtisuorassa toisiaan vastaan eikä mikään vektoreista ole nollavektori. Sisätuloavaruuden V jono $(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k)$ on *ortonormaali*, jos se on ortogonaalinen ja lisäksi

$$\|\bar{v}_i\| = 1 \quad \text{kaikilla } i \in \{1, 2, \ldots, k\}.$$

Toisin sanoen vektorit ovat kohtisuorassa toisiaan vastaan ja niiden normi on yksi.

Määritelmä 23.18. Oletetaan, että W on sisätuloavaruuden V aliavaruus, jolla on ortogonaalinen kanta $(\bar{w}_1, \ldots, \bar{w}_k)$. Vektorin $\bar{v} \in V$ *kohtisuora projektio aliavaruudelle* W on

$$\mathrm{proj}_W(\bar{v}) = \frac{\langle \bar{v}, \bar{w}_1 \rangle}{\langle \bar{w}_1, \bar{w}_1 \rangle} \bar{w}_1 + \frac{\langle \bar{v}, \bar{w}_2 \rangle}{\langle \bar{w}_2, \bar{w}_2 \rangle} \bar{w}_2 + \cdots + \frac{\langle \bar{v}, \bar{w}_k \rangle}{\langle \bar{w}_k, \bar{w}_k \rangle} \bar{w}_k.$$



Kuva 23.28: Vektorin $\bar{v}$ kohtisuora projektio aliavaruudelle $W = \mathrm{span}(\bar{w}_1, \bar{w}_2)$.

Usein kohtisuoraa projektiota kutsutaan lyhyesti vain projektioksi. Jos aliavaruus W on vain yhden vektorin virittämä eli $W = \operatorname{span}(\bar{w})$ jollakin $\bar{w} \in V$, voidaan kirjoittaa $\operatorname{proj}_W(\bar{v}) = \operatorname{proj}_{\bar{w}}(\bar{v})$. Tätä merkintää käyttäen

$$\operatorname{proj}_W(\bar{v}) = \operatorname{proj}_{\bar{w}_1}(\bar{v}) + \operatorname{proj}_{\bar{w}_2}(\bar{v}) + \cdots + \operatorname{proj}_{\bar{w}_k}(\bar{v}).$$

Esimerkki 23.19. Merkitään $\bar{w}_1 = (1,1,0)$ ja $\bar{w}_2 = (-1,1,1)$. Tarkastellaan avaruuden $\mathbb{R}^3$ aliavaruutta $W = \operatorname{span}(\bar{w}_1, \bar{w}_2)$, joka on vektoreiden $\bar{w}_1$ ja $\bar{w}_2$ virittämä, origon kautta kulkeva taso. Sisätulona on tavallinen pistetulo. Havaitaan, että $\bar{w}_1 \nparallel \bar{w}_2$, joten jono $(\bar{w}_1, \bar{w}_2)$ on vapaa. Siten se on virittämänsä aliavaruden $W = \operatorname{span}(\bar{w}_1, \bar{w}_2)$ kanta. Lisäksi $\bar{w}_1 \cdot \bar{w}_2 = -1 + 1 = 0$, joten kanta $(\bar{w}_1, \bar{w}_2)$ on ortogonaalinen.

Nyt voidaan määrittää vektorin $\bar{v} = (3,-1,2)$ projektio aliavaruudelle W:

$$\begin{aligned}
\operatorname{proj}_W(\bar{v}) &= \frac{\langle \bar{v}, \bar{w}_1 \rangle}{\langle \bar{w}_1, \bar{w}_1 \rangle}\bar{w}_1 + \frac{\langle \bar{v}, \bar{w}_2 \rangle}{\langle \bar{w}_2, \bar{w}_2 \rangle}\bar{w}_2 \\
&= \frac{\bar{v} \cdot \bar{w}_1}{\bar{w}_1 \cdot \bar{w}_1}\bar{w}_1 + \frac{\bar{v} \cdot \bar{w}_2}{\bar{w}_2 \cdot \bar{w}_2}\bar{w}_2 \\
&= \frac{3-1+0}{1+1+0}\bar{w}_1 + \frac{-3-1+2}{1+1+1}\bar{w}_2 \\
&= \bar{w}_1 - \frac{2}{3}\bar{w}_2 \\
&= \frac{1}{3}(5,1,-2).
\end{aligned}$$

Kohtisuoran projektion määritelmässä on käytettävä avaruuden W *ortogonaalista kantaa*. Jos projektion kaavassa käyttää kantaa, joka ei ole ortogonaalinen, ei tulos ole toivottu. Luvussa 23.4 tulemme osoittamaan että jokaiselle äärellisulotteiselle avaruudelle löytyy ortogonaalinen kanta. Lisäksi osoitamme lauseessa 23.24, että valitulla kannalla ei ole vaikutusta siihen, mikä projektiovektori on. Oleellista on vain, että kanta on ortogonaalinen.

Esimerkki 23.20. Tarkastellaan vektorin $\bar{v} = (3,2,1)$ projektiota aliavaruudelle $W = \operatorname{span}\big(\bar{e}_1, \bar{e}_2\big)$, missä $\bar{e}_1 = (1,0,0)$ ja $\bar{e}_2 = (0,1,0)$. Jos aliavaruus W piirrettäisiin tavalliseen tapaan koordinaatistoon, se olisi origon kautta kulkeva vaakataso (eli xy-taso). Koska jono $(\bar{e}_1, \bar{e}_2)$ on vapaa, se on aliavaruuden W kanta, ja lisäksi vektorit $\bar{e}_1$ ja $\bar{e}_2$ ovat ortogonaaliset. Projektion määritelmän mukaan $\operatorname{proj}_W(\bar{v})$ määritetään laskemalla vektorin $\bar{v}$ projektiot vektoreiden $(1,0,0)$ ja $(0,1,0)$ virittämille aliavaruuksille:

$$\begin{aligned}
\operatorname{proj}_W(\bar{v}) &= \frac{\langle \bar{v}, \bar{e}_1 \rangle}{\langle \bar{e}_1, \bar{e}_1 \rangle}\bar{e}_1 + \frac{\langle \bar{v}, \bar{e}_2 \rangle}{\langle \bar{e}_2, \bar{e}_2 \rangle}\bar{e}_2 = \frac{\bar{v} \cdot \bar{e}_1}{\bar{e}_1 \cdot \bar{e}_1}\bar{e}_1 + \frac{\bar{v} \cdot \bar{e}_2}{\bar{e}_2 \cdot \bar{e}_2}\bar{e}_2 \\
&= \frac{3}{1}\bar{e}_1 + \frac{2}{1}\bar{e}_2 = 3\bar{e}_1 + 2\bar{e}_2 = (3,2,0).
\end{aligned}$$

Voidaan ajatella, että vektorin projektio tasolle on varjo, jonka projektio heittää, kun aurinko paistaa kohtisuoraan tasoa vastaan. Nähdään, että tämän esimerkin tapauksessa näin todellakin on. Projisoitaessa vektorista katoaa kolmas komponentti.

Tutkitaan sitten, mitä tapahtuu, jos aliavaruudelle valitaan kanta, joka ei olekaan ortogonaalinen. Aliavaruudella W on myös kanta $(\bar{u}_1, \bar{u}_2)$, missä $\bar{u}_1 = (1,1,0)$ ja $\bar{u}_2 = (0,1,0)$, eikä tämä kanta ole ortogonaalinen. Nyt vektorin $\bar{v} = (3,2,1)$ projektio vektorin $\bar{u}_1$ virittämälle aliavaruudelle on

$$\text{proj}_{\bar{u}_1}(\bar{v}) = \frac{5}{2}\bar{u}_1$$

ja vektorin $\bar{u}_2$ virittämälle aliavaruudelle puolestaan

$$\text{proj}_{\bar{u}_2}(\bar{v}) = \frac{2}{1}\bar{u}_2 = 2\bar{u}_2.$$

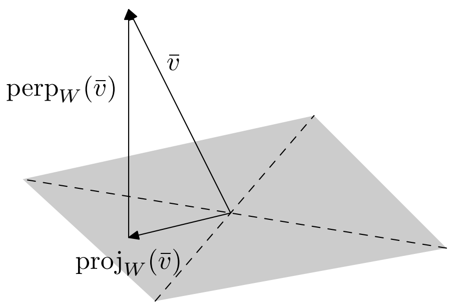
Näiden projektiovektoreiden summa on $(5/2, 9/2, 0)$, joten tulos on aivan erilainen kuin edellisissä laskuissa. Se ei vastaa käsitystämme siitä, miltä projektion pitäisi näyttää. Tämä johtui siitä, ettei käytetty kanta ollut ortogonaalinen.

Edellä mainittiin, että vektorin $\bar{v}$ projektio tasolle W on varjo, jonka projektio heittää, kun aurinko paistaa kohtisuoraan tasoa vastaan. Matemaattisesti tämä tarkoittaa sitä, että vektorit $\bar{v}$ ja $\bar{v} - \text{proj}_W(\bar{v})$ ovat kohtisuorassa toisiaan vastaan.

Määritelmä 23.21. Oletetaan, että W on sisätuloavaruuden V äärellisulotteinen aliavaruus, jolla on ortogonaalinen kanta. Vektorin $\bar{v} \in V$ *kohtisuora komponentti aliavaruutta W vastaan* on

$$\text{perp}_W(\bar{v}) = \bar{v} - \text{proj}_W(\bar{v}) = \frac{\langle \bar{v}, \bar{e}_1 \rangle}{\langle \bar{e}_1, \bar{e}_1 \rangle}\bar{w}_1.$$

Merkintä perp tulee englannin kielen sanasta ”perpendicular”, joka tarkoittaa kohtisuoraa.



Kuva 23.29: Vektorin $\bar{v}$ kohtisuora kohtisuora komponentti aliavaruutta W vastaan.

Esimerkki 23.22. Esimerkissä 23.19 nähtiin, että vektorin $\bar{v} = (3,-1,2)$ kohtisuora projektio vektorien $\bar{w}_1 = (1,1,0)$ ja $\bar{w}_2 = (-1,1,1)$ virittämälle aliavaruudelle W on $\frac{1}{3}(5,1,-2)$. Vektorin $\bar{v}$ kohtisuora komponentti aliavaruutta W vastaan on siten

$$\text{perp}_W(\bar{v}) = \bar{v} - \text{proj}_W(\bar{v}) = (3,-1,2) - \frac{1}{3}(5,1,-2) = \frac{1}{3}(4,-4,4).$$

Vektorin $\bar{v}$ kohtisuora komponentti $\mathrm{perp}_W(\bar{v})$ on kohtisuorassa aliavaruutta W vastaan. Se kuuluu siis kohtisuoraan komplementtiin $W^{\perp}$.

Lause 23.23. *Oletetaan, että W on sisätuloavaruuden V äärellisulotteinen aliavaruus ja $\bar{v} \in V$. Tällöin* $\mathrm{perp}_W(\bar{v}) \in W^{\perp}$.

Todistus. Olkoon $(\bar{w}_1, \ldots, \bar{w}_k)$ aliavaruuden W ortogonaalinen kanta. Nyt siis

$$\langle \bar{w}_i, \bar{w}_j \rangle = 0, \text{kun } i \neq j.$$

Lauseen 23.15 nojalla riittää osoittaa, että $\langle \mathrm{perp}_W(\bar{v}), \bar{w}_i \rangle = 0$ kaikilla $i \in \{1, \ldots, n\}$. Oletetaan, että $i \in \{1, \ldots, n\}$. Tällöin

$$\begin{aligned}
\langle \mathrm{perp}_W(\bar{v}), \bar{w}_i \rangle &= \langle \bar{v} - \mathrm{proj}_W(\bar{v}), \bar{w}_i \rangle \\
&= \langle \bar{v}, \bar{w}_i \rangle - \langle \mathrm{proj}_W(\bar{v}), \bar{w}_i \rangle \\
&= \langle \bar{v}, \bar{w}_i \rangle - \left\langle \frac{\langle \bar{v}, \bar{w}_1 \rangle}{\langle \bar{w}_1, \bar{w}_1 \rangle} \bar{w}_1 + \cdots + \frac{\langle \bar{v}, \bar{w}_k \rangle}{\langle \bar{w}_k, \bar{w}_k \rangle} \bar{w}_k, \bar{w}_i \right\rangle \\
&= \langle \bar{v}, \bar{w}_i \rangle - \left(\frac{\langle \bar{v}, \bar{w}_1 \rangle}{\langle \bar{w}_1, \bar{w}_1 \rangle} \langle \bar{w}_1, \bar{w}_i \rangle + \cdots + \frac{\langle \bar{v}, \bar{w}_k \rangle}{\langle \bar{w}_k, \bar{w}_k \rangle} \langle \bar{w}_k, \bar{w}_i \rangle \right) \\
&= \langle \bar{v}, \bar{w}_i \rangle - \frac{\langle \bar{v}, \bar{w}_i \rangle}{\langle \bar{w}_i, \bar{w}_i \rangle} \langle \bar{w}_i, \bar{w}_i \rangle \\
&= \langle \bar{v}, \bar{w}_i \rangle - \langle \bar{v}, \bar{w}_i \rangle = 0,
\end{aligned}$$

sillä $\langle \bar{w}_i, \bar{w}_j \rangle = 0$, kun $i \neq j$.

Vektori $\mathrm{perp}_W(\bar{v})$ on siis kohtisuorassa kaikkia kantavektoreita vastaan. Tästä seuraa, että $\mathrm{perp}_W(\bar{v}) \in W^{\perp}$. Huomaa, että todistuksessa käytettiin hyväksi sitä, että aliavaruudella W on ortogonaalinen kanta. □

Lause 23.24. *Oletetaan, että W on sisätuloavaruuden V äärellisulotteinen aliavaruus ja $v \in V$. Tällöin on olemassa täsmälleen yhdet vektorit $\bar{w} \in W$ ja $\bar{w}^{\perp} \in W^{\perp}$, joille pätee*

$$\bar{v} = \bar{w} + \bar{w}^{\perp}.$$

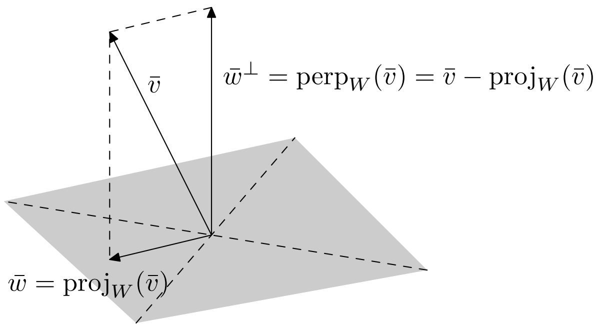
Osoittautuu, että lauseessa mainitut vektorit ovat $\mathrm{proj}_W(\bar{v})$ ja $\mathrm{perp}_W(\bar{v})$.

Todistus. Valitaan $\bar{w} = \mathrm{proj}_W(\bar{v})$ ja $\bar{w}^{\perp} = \mathrm{perp}_W(\bar{v}) = \bar{v} - \mathrm{proj}_W(\bar{v})$. Projektion määritelmästä nähdään, että $\bar{w} \in W$, ja lauseen 23.23 nojalla $\bar{w}^{\perp} \in W^{\perp}$. Lisäksi $\bar{v} = \bar{w} + \bar{w}^{\perp}$.

Osoitetaan vielä, että mitkään muut vektorit eivät toteuta annettuja ehtoja. Oletetaan, että $\bar{w}, \bar{u} \in W$, $\bar{w}^{\perp}, \bar{u}^{\perp} \in W^{\perp}$ ovat ehdot toteuttavia vektoreita. Nyt siis $\bar{v} = \bar{w} + \bar{w}^{\perp} = \bar{u} + \bar{u}^{\perp}$. Tästä seuraa, että

$$\bar{w} - \bar{u} = \bar{u}^{\perp} - \bar{w}^{\perp}.$$

Toisaalta W ja $W^{\perp}$ ovat aliavaruuksia, joten $\bar{w} - \bar{u} \in W$ ja $\bar{u}^{\perp} - \bar{w}^{\perp} \in W^{\perp}$. Kuitenkin lauseen 23.17 nojalla $W \cap W^{\perp} = \{\bar{0}\}$, joten $\bar{w} - \bar{u} = \bar{0}$ ja $\bar{w}^{\perp} - \bar{u}^{\perp} = \bar{0}$. Tästä seuraa, että $\bar{w} = \bar{u}$ ja $\bar{w}^{\perp} = \bar{u}^{\perp}$. Siten ehdot toteuttavia vektoreita on vain yhdet. □



Kuva 23.30: Vektori $\bar{v}$ voidaan kirjoittaa summana vektoreista, jotka ovat aliavaruuksien W ja $W^{\perp}$ alkioita.

Todistetaan vielä lopuksi projektion avulla muutama hyödyllinen sisätuloon ja normiin liittyvä tulos. Esimerkiksi Schwarzin epäyhtälöä käytettiin jo luvussa 12.

Lause 23.25 (Schwarzin epäyhtälö)**.** *Olkoon V sisätuloavaruus. Oletetaan, että $\bar{v}, \bar{w} \in V$. Tällöin*

$$|\langle \bar{v}, \bar{w} \rangle| \leq \|\bar{v}\| \|\bar{w}\|.$$

Todistus. Oletetaan ensin, että $\bar{w} = \bar{0}$. Nyt $\langle \bar{v}, \bar{w} \rangle = 0$ ja $\|\bar{w}\| = 0$, joten väite pätee.

Oletetaan sitten, että $\bar{w} \neq \bar{0}$. Nyt

$$\begin{aligned}
0 &\leq \|\bar{v} - \mathrm{proj}_w(\bar{v})\|^2 \\
&= \langle \bar{v} - \mathrm{proj}_w(\bar{v}), \bar{v} - \mathrm{proj}_w(\bar{v}) \rangle \\
&= \left\langle \bar{v} - \frac{\langle \bar{w}, \bar{v} \rangle}{\langle \bar{w}, \bar{w} \rangle} \bar{w},\ \bar{v} - \frac{\langle \bar{w}, \bar{v} \rangle}{\langle \bar{w}, \bar{w} \rangle} \bar{w} \right\rangle \\
&= \langle \bar{v}, \bar{v} \rangle - \frac{\langle \bar{v}, \bar{w} \rangle}{\langle \bar{w}, \bar{w} \rangle} \langle \bar{v}, \bar{w} \rangle - \frac{\langle \bar{v}, \bar{w} \rangle}{\langle \bar{w}, \bar{w} \rangle} \langle \bar{w}, \bar{v} \rangle + \frac{\langle \bar{v}, \bar{w} \rangle^2}{\langle \bar{w}, \bar{w} \rangle^2} \langle \bar{w}, \bar{w} \rangle \\
&= \|\bar{v}\|^2 - 2 \frac{\langle v, w \rangle^2}{\|\bar{w}\|^2} + \frac{\langle v, w \rangle^2}{\|\bar{w}\|^2} \\
&= \|\bar{v}\|^2 - \frac{\langle v, w \rangle^2}{\|\bar{w}\|^2}.
\end{aligned}$$

Tästä seuraa, että

$$\langle \bar{v}, \bar{w} \rangle^2 \leq \|\bar{v}\|^2 \|\bar{w}\|^2.$$

Nyt voidaan päätellä, että $|\langle \bar{v}, \bar{w} \rangle| = \sqrt{\langle \bar{v}, \bar{w} \rangle^2} \leq \sqrt{\|\bar{v}\|^2 \|\bar{w}\|^2} = \|\bar{v}\| \|\bar{w}\|.$ □

Lause 23.26 (Kolmioepäyhtälö)**.** *Olkoon V sisätuloavaruus. Tällöin*

$$\|\bar{v} + \bar{w}\| \leq \|\bar{v}\| + \|\bar{w}\|$$

kaikilla $\bar{v}, \bar{w} \in V$.

Todistus. Ensinnäkin huomataan, että

$$\begin{aligned}\|\bar v+\bar w\|^2 &= \langle \bar v+\bar w, \bar v+\bar w\rangle \\ &= \langle \bar v,\bar v\rangle + \langle \bar v,\bar w\rangle + \langle \bar w,\bar v\rangle + \langle \bar w,\bar w\rangle \\ &= \|\bar v\|^2 + 2\langle \bar v,\bar w\rangle + \|\bar w\|^2 \\ &\le \|\bar v\|^2 + 2|\langle \bar v,\bar w\rangle| + \|\bar w\|^2.\end{aligned}$$

Nyt voidaan käyttää Schwarzin epäyhtälöä:

$$\begin{aligned}\|\bar v\|^2 + 2|\langle \bar v,\bar w\rangle| + \|\bar w\|^2 &\le \|\bar v\|^2 + 2\|\bar v\|\|\bar w\| + \|\bar w\|^2 \\ &= (\|\bar v\| + \|\bar w\|)^2.\end{aligned}$$

Näin saadaan $\|\bar v+\bar w\|^2 \le (\|\bar v\| + \|\bar w\|)^2$. Koska normit ovat positiivisia, tästä voidaan päätellä, että

$$\|\bar v+\bar w\| \le \|\bar v\| + \|\bar w\|.$$

□

23.4 Ortogonaaliset ja ortonormaalit kannat

Luvussa 12.4 tutkittiin avaruuden $\mathbb{R}^n$ ortogonaalisia ja ortonormaaleja kantoja. Monissa tilanteissa nämä kannat ovat kätevämpiä kuin muut kannat. Nyt yleistämme tulokset sisätuloavaruuksille sekä esittelemme joitakin uusia tuloksia.

Ensinnäkin ortogonaaliset jonot ovat vapaita.

Lause 23.27. *Oletetaan, että $(\bar v_1, \bar v_2, \dots, \bar v_k)$ on sisätuloavaruuden V ortogonaalinen jono. Tällöin $(\bar v_1, \bar v_2, \dots, \bar v_k)$ on vapaa.*

Todistus. Olkoot $c_1, \dots, c_k \in \mathbb{R}$ sellaisia, että

$$c_1\bar v_1 + \cdots + c_k\bar v_k = \bar 0.$$

Oletetaan, että $i \in \{1, \dots, k\}$. Nyt

$$\langle \bar v_i, c_1\bar v_1 + \cdots + c_k\bar v_k\rangle = \langle \bar v_i, \bar 0\rangle.$$

Yhtälön vasen puoli saadaan muotoon

$$c_1\langle \bar v_i, \bar v_1\rangle + \cdots + c_k\langle \bar v_i, \bar v_k\rangle = c_i\langle \bar v_i, \bar v_i\rangle,$$

sillä jono $(\bar v_1, \bar v_2, \dots, \bar v_k)$ on ortogonaalinen. Toisaalta $\langle \bar v_i, \bar 0\rangle = 0$ lemman 23.5 nojalla. Siis $c_i\langle \bar v_i, \bar v_i\rangle = 0$.

Koska ortogonaalisessa jonossa ei määritelmän mukaan ole nollavektoreita, ei $\bar v_i$ ole nollavektori. Näin ollen $\langle \bar v_i, \bar v_i\rangle \neq 0$, mistä seuraa, että $c_i = 0$. Siten $c_i = 0$ kaikilla $i \in \{1, \dots, k\}$, ja jono $(\bar v_1, \bar v_2, \dots, \bar v_k)$ on vapaa. □

Ortonormaalin kannan suhteen vektorin koordinaatit on helppo määrittää.

Lause 23.28. *Oletetaan, että $\mathcal{B} = (\bar{u}_1, \ldots, \bar{u}_n)$ on sisätuloavaruuden V ortonormaali kanta. Olkoon $\bar{v} \in V$. Tällöin*

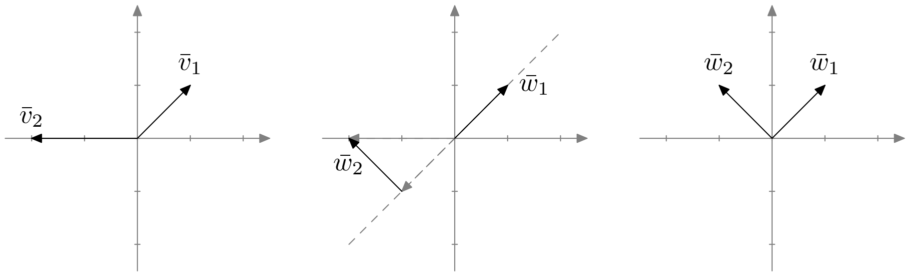
$$\bar{v} = \langle \bar{v}, \bar{u}_1 \rangle \bar{u}_1 + \langle \bar{v}, \bar{u}_2 \rangle \bar{u}_2 + \cdots + \langle \bar{v}, \bar{u}_n \rangle \bar{u}_n.$$

Toisin sanoen vektorin $\bar{v}$ koordinaatit kannan $\mathcal{B}$ suhteen saadaan sisätulon avulla.

Todistus. Lauseen todistus on samanlainen kuin lauseen 12.22 todistus. □

Ortonormaalin kannan etsiminen

Kohta esiteltävän Gramin–Schmidtin menetelmän avulla kannasta voidaan muokata ortogonaalinen ja edelleen ortonormaali. Luvussa 12.4 näytettiin, kuinka projektion (tai oikeammin kohtisuoran komplementin) avulla voidaan kahden vektorin muodostamasta kannasta muokata ortogonaalinen (ks. kuva 23.31). Nyt yleistämme tämän menetelmän mille tahansa äärelliselle kannalle.



Kuva 23.31: Ortogonaalisen kannan etsiminen kahden kantavektorin tapauksessa.

Esimerkki 23.29. Vektorit $\bar{v}_1 = (1, 1, 0)$, $\bar{v}_2 = (-2, 0, 1)$ ja $\bar{v}_3 = (0, 1, 1)$ muodostavat avaruuden $\mathbb{R}^3$ kannan. (Tämän osoittamiseksi riittää näyttää, että vektorit ovat lineaarisesti riippumattomia. Se jätetään lukijalle.) Ryhdytään muodostamaan näistä kolmesta vektorista kantaa $(\bar{v}_1, \bar{v}_2, \bar{v}_3)$, joka on ortogonaalinen tavallisen pistetulon suhteen.

Uuden kannan ensimmäiseksi vektoriksi voidaan ottaa mikä tahansa vektoreista $\bar{w}_1, \bar{w}_2, \bar{w}_3$. Valitaan $\bar{w}_1 = \bar{v}_1$. Toiseksi vektoriksi valitaan vektorin $\bar{v}_2$ kohtisuora komponentti aliavaruutta $\mathrm{span}(\bar{w}_1)$ vastaan:

$$\bar{w}_2 = \mathrm{perp}_{\mathrm{span}(\bar{w}_1)}(\bar{v}_2) = \bar{v}_2 - \mathrm{proj}_{\mathrm{span}(\bar{w}_1)}(\bar{v}_2) = \bar{v}_2 - \frac{\bar{v}_2 \cdot \bar{w}_1}{\bar{w}_1 \cdot \bar{w}_1}\bar{w}_1 = (-1, 1, 1).$$

Nyt vektorit $\bar{w}_1$ ja $\bar{w}_2$ ovat lemman 23.23 nojalla kohtisuorassa toisiaan vastaan. Lauseen 23.27 perusteella jono $(\bar{v}_1, \bar{v}_2)$ on vapaa, joten se muodostaa aliavaruuden $\mathrm{span}(\bar{w}_1, \bar{w}_2)$ ortogonaalisen kannan (ks. lause 17.14).

Kolmanneksi vektoriksi valitaan vektorin $\bar{v}_3$ kohtisuora komponentti aliavaruutta $\mathrm{span}(\bar{w}_1, \bar{w}_2)$ vastaan:

$$
\begin{aligned}
\bar{w}_3 &= \operatorname{perp}_{\operatorname{span}(\bar{w}_1,\bar{w}_2)}(\bar{v}_3) \\
&= \bar{v}_3 - \operatorname{proj}_{\operatorname{span}(\bar{w}_1,\bar{w}_2)}(\bar{v}_3) \\
&= \bar{v}_3 - \frac{\bar{v}_3 \cdot \bar{w}_1}{\bar{w}_1 \cdot \bar{w}_1}\bar{w}_1 - \frac{\bar{v}_3 \cdot \bar{w}_2}{\bar{w}_2 \cdot \bar{w}_2}\bar{w}_2 \\
&= \frac{1}{6}(1,-1,2).
\end{aligned}
$$

Huomaa, että projektion laskemiseen tarvitaan tietoa siitä, että $(\bar{w}_1, \bar{w}_2)$ on aliavaruuden $\operatorname{span}(\bar{w}_1, \bar{w}_2)$ ortogonaalinen kanta. Vektori $\bar{w}_3$ on kohtisuorassa vektoreita $\bar{w}_1$ ja $\bar{w}_2$ vastaan lemman 23.23 perusteella. Siten jono $(\bar{w}_1, \bar{w}_2, \bar{w}_3)$ on ortogonaalinen.

Lauseen 23.27 nojalla jono $(\bar{w}_1, \bar{w}_2, \bar{w}_3)$ on vapaa. Koska avaruuden $\mathbb{R}^3$ dimensio on kolme, on jono $(\bar{w}_1, \bar{w}_2, \bar{w}_3)$ avaruuden kanta. Näin saatiin siis aikaan ortogonaalinen kanta.

Lause 23.30 (Gramin–Schmidtin menetelmä)**.** *Oletetaan, että $\mathcal{B} = (\bar{v}_1, \dots, \bar{v}_n)$ on sisätuloavaruuden V kanta. Tällöin avaruudella V on ortogonaalinen kanta $(\bar{w}_1, \dots, \bar{w}_n)$, joka saadaan seuraavasti:*

$$
\begin{aligned}
\bar{w}_1 &= \bar{v}_1 \\
\bar{w}_2 &= \operatorname{perp}_{\operatorname{span}(\bar{w}_1)}(\bar{v}_2) \\
\bar{w}_3 &= \operatorname{perp}_{\operatorname{span}(\bar{w}_1,\bar{w}_2)}(\bar{v}_3) \\
&\vdots \\
\bar{w}_n &= \operatorname{perp}_{\operatorname{span}(\bar{w}_1,\dots,\bar{w}_{n-1})}(\bar{v}_n).
\end{aligned}
$$

Tästä ortogonaalisesta kannasta saadaan ortonormaali kanta $(\bar{w}_1, \dots, \bar{w}_n)$ asettamalla

$$\bar{u}_i = \frac{1}{\|\bar{w}_i\|}\bar{w}_i$$

kaikilla $i \in \{1, 2, \dots, n\}$.

Jos kantavektorit kirjoitetaan auki, nähdään, että

$$
\begin{aligned}
\bar{w}_1 &= \bar{v}_1 \\
\bar{w}_2 &= \bar{v}_2 - \frac{\langle \bar{v}_2, \bar{w}_1 \rangle}{\langle \bar{w}_1, \bar{w}_1 \rangle}\bar{w}_1 \\
\bar{w}_3 &= \bar{v}_3 - \frac{\langle \bar{v}_3, \bar{w}_1 \rangle}{\langle \bar{w}_1, \bar{w}_1 \rangle}\bar{w}_1 - \frac{\langle \bar{v}_3, \bar{w}_2 \rangle}{\langle \bar{w}_2, \bar{w}_2 \rangle}\bar{w}_2 \\
&\vdots \\
\bar{w}_n &= \bar{v}_2 - \frac{\langle \bar{v}_2, \bar{w}_1 \rangle}{\langle \bar{w}_1, \bar{w}_1 \rangle}\bar{w}_{n-1} - \frac{\langle \bar{v}_2, \bar{w}_2 \rangle}{\langle \bar{w}_2, \bar{w}_2 \rangle}\bar{w}_{n-1} - \cdots - \frac{\langle \bar{v}_n, \bar{w}_{n-1} \rangle}{\langle \bar{w}_{n-1}, \bar{w}_{n-1} \rangle}\bar{w}_{n-1}\bar{w}_1.
\end{aligned}
$$

Todistus. Aloitetaan osoittamalla, että jono $(\bar{w}_1, \dots, \bar{w}_n)$ on ortogonaalinen. Koska $\bar{w}_2$ on vektorin $\bar{v}_2$ kohtisuora komponentti aliavaruutta $\mathrm{span}(\bar{w}_1)$ vastaan, vektorit $\bar{w}_1$ ja $\bar{w}_2$ ovat ortogonaaliset lemman 23.23 nojalla. Samalla tavalla $\bar{w}_3$ on vektorin $\bar{v}_3$ kohtisuora komponentti aliavaruutta $\mathrm{span}(\bar{w}_1, \bar{w}_2)$ vastaan, joten vektorit $\bar{w}_3$ ja $\bar{w}_2$ sekä $\bar{w}_3$ ja $\bar{w}_1$ ovat ortogonaaliset. Jatkaen samaan tapaan nähdään, että kaikki vektorit $\bar{w}_1, \dots, \bar{w}_n$ ovat kohtisuorassa toisiaan vastaan.

On vielä osoitettava, että mikään vektoreista ei ole nollavektori. Oletetaan vastoin väitettä, että $\bar{w}_i = \bar{0}$ jollakin $i \in \{1, \dots, n\}$. Nyt

$$\bar{v}_i = \bar{w}_i + \mathrm{proj}_{\mathrm{span}(\bar{w}_1,\dots,\bar{w}_{i-1})}(\bar{v}_i) = \mathrm{proj}_{\mathrm{span}(\bar{w}_1,\dots,\bar{w}_{i-1})}(\bar{v}_i) \in \mathrm{span}(\bar{v}_1, \dots, \bar{v}_{i-1}).$$

Lauseen 16.5 perusteella tästä seuraa, että jono $(\bar{v}_1, \dots, \bar{v}_i)$ on sidottu. Tämä on ristiriita, sillä $(\bar{v}_1, \dots, \bar{v}_n)$ on kanta ja siksi vapaa, eikä vapaan jonon osajono voi olla sidottu (ks. lause 16.7). Siten vastaoletus on väärä. Näin ollen $\bar{w}_i \neq \bar{0}$ kaikilla $i \in \{1, \dots, n\}$. Siis jono $(\bar{w}_1, \dots, \bar{w}_n)$ on ortogonaalinen.

Lauseen 23.27 nojalla jono $(\bar{w}_1, \dots, \bar{w}_n)$ on vapaa ja lisäksi sen pituus on sama kuin avaruuden V dimensio. Lauseen 17.14 nojalla se on avaruuden V kanta.

Skalaarilla kertominen ei vaikuta vektorien ortogonaalisuuteen, joten myös jono $(\bar{u}_1, \dots, \bar{u}_n)$ on ortogonaalinen. Lisäksi sen jokaisen vektorin normi on yksi, joten jono on ortonormaali. □

Lauseesta seuraa, että jokaisella äärellisulotteisella vektoriavaruudella on ortogonaalinen kanta.

Esimerkki 23.31. Etsitään avaruudelle $\mathbb{R}^3$ ortogonaalinen kanta, jonka yksi vektori on $\bar{w}_1 = (1, 2, 3)$. Valitaan aluksi vaikkapa $\bar{v}_2 = (0, 1, 0)$ ja $\bar{v}_3 = (0, 0, 1)$. Tällöin jono $(\bar{w}_1, \bar{v}_2, \bar{v}_3)$ avaruuden $\mathbb{R}^3$ kanta. Tämän osoittamiseen riittää lauseen 17.14 perusteella näyttää, että jono on vapaa tai että se virittää avaruuden $\mathbb{R}^3$. Tämä jätetään lukijalle.

Ortogonalisoidaan kanta $(\bar{w}_1, \bar{v}_2, \bar{v}_3)$. Valitaan toiseksi kantavektoriksi

$$\bar{w}_2' = \bar{v}_2 - \frac{\bar{v}_2 \cdot \bar{w}_1}{\bar{w}_1 \cdot \bar{w}_1}\bar{w}_1 = \bar{v}_2 - \frac{2}{14}\bar{w}_1 = (-1/7, 5/7, -3/7).$$

Tässä vaiheessa vektoria $\bar{w}_2'$ kannattaa vielä kertoa skalaarilla 7, jotta päästään eroon ikävistä murtoluvuista. Valitaankin kantavektoriksi siis

$$\bar{w}_2 = 7\bar{w}_2' = (-1, 5, -3).$$

Kolmas kantavektorikandidaatti on

$$\begin{aligned} \bar{w}_3' &= \bar{v}_3 - \frac{\bar{v}_3 \cdot \bar{w}_1}{\bar{w}_1 \cdot \bar{w}_1}\bar{w}_1 - \frac{\bar{v}_3 \cdot \bar{w}_2}{\bar{w}_2 \cdot \bar{w}_2}\bar{w}_2 \\ &= \bar{v}_3 - \frac{3}{14}\bar{w}_1 - \frac{-3}{35}\bar{w}_2 = (-3/10,\ 0,\ 1/10) \end{aligned}$$

Vaihdetaan vielä tämänkin vektorin tilalle siistimpi skalaarimonikerta

$$\bar{w}_3 = 10\bar{w}_3' = (-3, 0, 1).$$

Tällöin $(\bar{w}_1, \bar{w}_2, \bar{w}_3)$ on avaruuden $\mathbb{R}^3$ ortogonaalinen kanta.

Tässä esimerkissä ikävät kantavektorit merkittiin kaukonäköisesti pilkulla, jotta haluttuja kantavektoreita voitiin merkitä symbolilla w. Käytännössä ei tietenkään voi etukäteen tietää, onko vektoriin tulossa murtolukuja vai ei, joten pilkkuja voi halutessaan lisätä vektoreiden nimiin jälkikäteen.

Lause 23.32. *Oletetaan, että V on äärellisulotteinen sisätuloavaruus ja W sen aliavaruus. Tällöin*

$$\dim(W) + \dim(W^{\perp}) = \dim(V).$$

Todistus. Oletetaan, että $(\bar{w}_1, \ldots, \bar{w}_k)$ on aliavaruuden W ortogonaalinen kanta ja että $(\bar{v}_1, \ldots, \bar{v}_l)$ on aliavaruuden $W^{\perp}$ ortogonaalinen kanta. Tällaiset kannat ovat olemassa lauseen 23.30 nojalla. Osoitetaan, että $(\bar{w}_1, \ldots, \bar{w}_k, \bar{v}_1, \ldots, \bar{v}_l)$ on avaruuden V kanta, mikä todistaa väitteen.

Havaitaan, että $\langle \bar{w}_i, \bar{v}_j \rangle = 0$ kaikilla $i \in \{1, \ldots, k\}$ ja $j \in \{1, \ldots, l\}$, sillä $\bar{w}_i \in W$ ja $\bar{v}_j \in W^{\perp}$. Lisäksi kummassakin jonossa jonon vekorit ovat kohtisuorassa toisiaan vastaan. Näin jono $(\bar{w}_1, \ldots, \bar{w}_k, \bar{v}_1, \ldots, \bar{v}_l)$ on ortogonaalinen ja siten vapaa lauseen 23.27 nojalla.

Oletetaan, että $\bar{u} \in V$. Lauseen 23.24 mukaan on olemassa yksi sellainen vektori $\bar{w} \in W$ ja yksi sellainen vektori $\bar{w}^{\perp} \in W^{\perp}$, että $\bar{u} = \bar{w} + \bar{w}^{\perp}$. Vektori $\bar{w} \in W$ voidaan kirjoittaa kantavektorien $(\bar{w}_1, \ldots, \bar{w}_k)$ lineaarikombinaationa ja vektori $\bar{w}^{\perp} \in W^{\perp}$ kantavektorien $(\bar{v}_1, \ldots, \bar{v}_l)$ lineaarikombinaationa, joten vektori $\bar{u}$ voidaan kirjoittaa jonon $(\bar{w}_1, \ldots, \bar{w}_k, \bar{v}_1, \ldots, \bar{v}_l)$ vektorien lineaarikombinaationa. Siis $\operatorname{span}(\bar{w}_1, \ldots, \bar{w}_k, \bar{v}_1, \ldots, \bar{v}_l) = V$.

On näytetty, että $(\bar{w}_1, \ldots, \bar{w}_k, \bar{v}_1, \ldots, \bar{v}_l)$ on avaruuden V kanta. Siis

$$\dim(V) = k + l = \dim(W) + \dim(W^{\perp}).$$

□