Lineaarialgebraa lähinnä tasossa hipauksella GNU Octavea

Tommi Sottinen

tommi.sottinen@uwasa.fi

www.uwasa.fi/∼tsottine/mathc1230/la1.pdf

27. lokakuuta 2023

Sisällys

Johdanto

Tämä on luentokirjanen Vaasan yliopiston kurssille Lineaarialgebra I. Kurssi on 2 opintopisteen laajuinen sisältäen noin 20 tuntia (à 45 min) luentoja ja 8 tuntia (à 45 min) harjoituksia. Luentokirjasen jokainen luku vastaa noin 2 tuntia (eli 1 h 30 min) luentoja. Mahdollisesti aivan kaikkea kirjasen tekstiä ei ehditä käymään läpi luennoilla. **Opiskelijoita kannustetaankin lukemaan luentoja vastaava teksti kirjasesta ennen luentoja ja esittää kysymyksiä, jos jokin asia on esitetty epäselvästi.**

Olemme käyttäneet "juoksevaa numerointia": harjoitustehtäviä, esimerkkejä yms. ei ole numeroitu erikseen. Siten esimerkiksi harjoitustehtävään 2.12 johdatteleva esimerkki on 2.11.

Tämä luentokirjanen tarkastelee lähinnä vain tason tapausta; toisin sanoen tarkastelemme kahden muuttujan yhtälöpareja eli (2×2)-matriiseja. Karkeasti otten kurssin ydin on tarkastella (abstraktia) yhtälöparia

$$\begin{cases} A_{11}x_1 & + & A_{12}x_2 & = & b_1 \\ A_{21}x_1 & + & A_{22}x_2 & = & b_2 \end{cases}$$

ja etsiä **kaikki** sen ratkaisut x_1, x_2 tai osoittaa että ratkaisuja ei ole olemassa. Osoittautuu, että ratkaisuja on olemassa tasan yksi pari x_1, x_2, jos determinantti

$$D \;=\; A_{11}A_{22} - A_{12}A_{21} \;\neq\; 0.$$

Tällöin yksikäsitteinen ratkaisu on itse asiassa

$$\begin{cases} x_1 & = & \frac{1}{D}(A_{22}b_1 - A_{12}b_2) \\ x_2 & = & \frac{1}{D}(-A_{21}b_1 + A_{11}b_2) \end{cases}.$$

Jos taas determinantti $D = 0$, niin ratkaisuja x_1, x_2 joko ei ole ollenkaan tai sitten niitä on ääretön määrä (joskaan kaikki tason pisteet x_1, x_2 eivät ole juuri koskaan ratkaisuja).

Luentokirjasen jokaisen luvun lopussa on tähtiosio, joka ei kuulu varsinaisiin oppimistavoitteisiin. Tähtiosiot tarkastelevat yleistä $(m\times n)$-tapausta

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

ja ne antavat vihjeen siitä, mitä opiskellaan jatkokurssilla Lineaarialgebra II.

Vaasassa 27. lokakuuta 2023
T.S.

Luku 1

Yhtälöpari

Yhtälöparin ratkaisu piirtämällä

Kahden **muuttujan** x_1 ja x_2 yleinen **lineaarinen yhtälöpari** on muotoa

$$\left\{\begin{array}{rcrcl} A_{11}x_1 & + & A_{12}x_2 & = & b_1 \\ A_{21}x_1 & + & A_{22}x_2 & = & b_2 \end{array}\right. . \tag{1.1}$$

Yhtälöparin (1.1) ratkaisu on sellainen tason piste (x_1, x_2), joka toteuttaa molemmat yhtälöt samanaikaisesti. Koska molemmat yhtälöt määrittelevät suoran tasoon, on ratkaisu (jos sellainen on, ja jos se on yksikäsitteinen) yhtälöiden määräämien suorien leikkauspiste. Jos yhtälöparin molemmat yhtälöt kuvaavat saman suoran, on ratkaisu kaikki ko. suoralla olevat pisteet. Jos yhtälöparin määräämät suorat ovat yhdensuuntaisia, mutta eivät kuvaa samaa suoraa, niin ratkaisuja ei ole.

1.2 Esimerkki
Ratkaisemme piirtämällä yhtälöparit

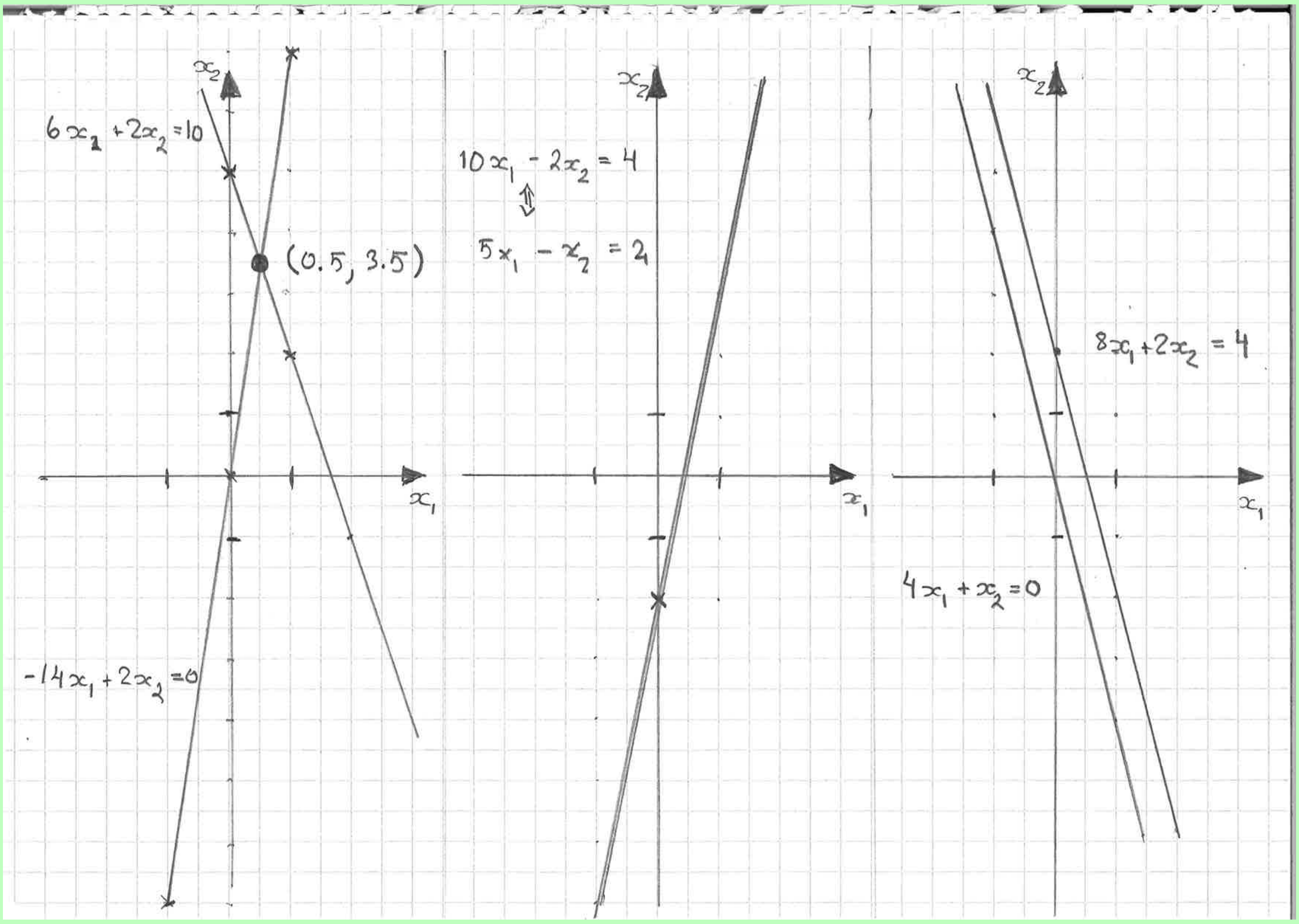
$$\left\{\begin{array}{rcrcr} -14x_1 & + & 2x_2 & = & 0 \\ 6x_1 & + & 2x_2 & = & 10 \end{array}\right. , \tag{1.3}$$

$$\left\{\begin{array}{rcrcr} 10x_1 & - & 2x_2 & = & 4 \\ 5x_1 & - & x_2 & = & 2 \end{array}\right. \tag{1.4}$$

ja

$$\left\{\begin{array}{rcrcr} 8x_1 & + & 2x_2 & = & 4 \\ 4x_1 & + & x_2 & = & 0 \end{array}\right. . \tag{1.5}$$

Alla olevasta piirroksesta näemme, että yhtälöparin (1.3) ratkaisu on $(x_1, x_2) = (0.5, 3, 5)$, yhtälöparin (1.4) ratkaisu on koko suora $(t, 5t-2)$, missä t on vapaa parametri, ja yhtälöparilla (1.5) ei ole ratkaisuja.



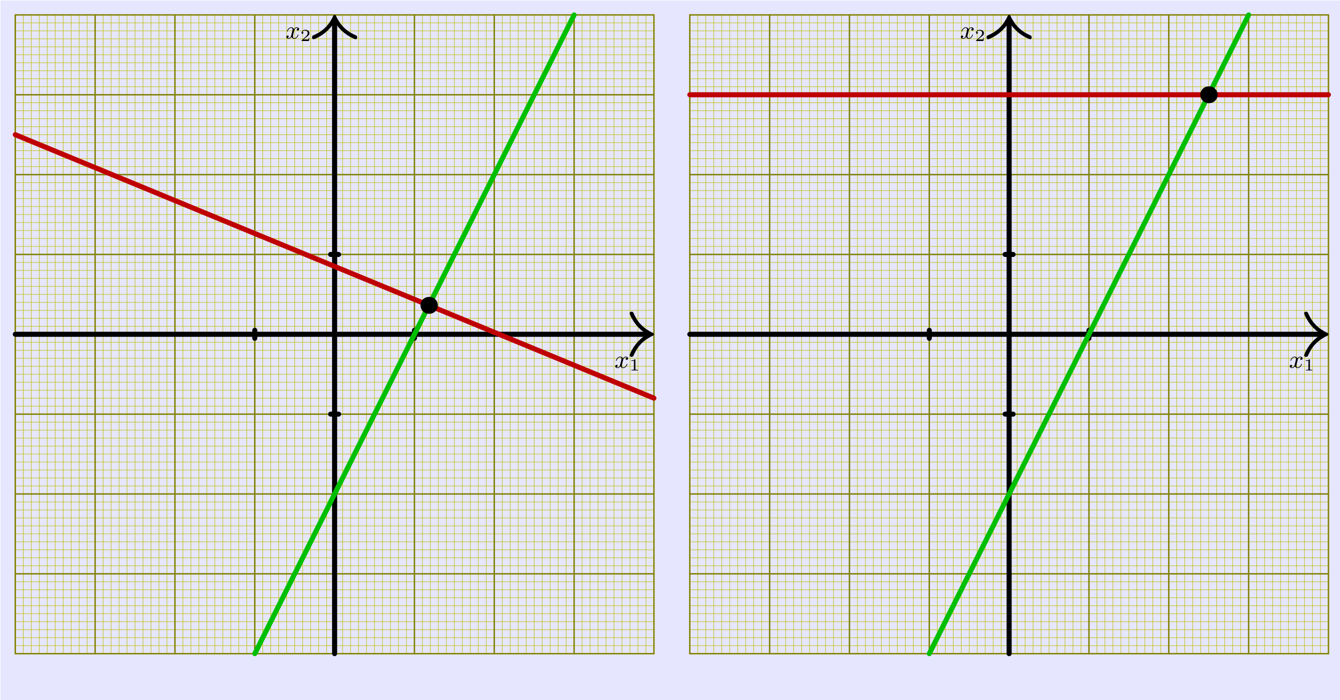
Yhtälöparien (1.3) (vasemmalla), (1.4) (keskellä) ja (1.5) (oikealla) ratkaisut.

1.6 Harjoitustehtävä
Ratkaise seuraavat yhtälöparit piirtämällä:

(i)
$$\begin{cases} -4x_1 & + & 3x_2 & = & 0 \\ 6x_1 & + & 2x_2 & = & 2 \end{cases},$$

(ii)
$$\begin{cases} 4x_1 & - & 2x_2 & = & 2 \\ 6x_1 & + & 2x_2 & = & 3 \end{cases}.$$

1.7 Harjoitustehtävä
Määrää mitä yhtälöpareja seuraavat kuvat kuvaavat.



Yhtälöparin ratkaisu alkeisrivioperaatioilla

1.8 **Esimerkki**
Haluamme ratkaista yhtälöparin

$$\begin{cases} 4x_1 + 2x_2 = 1 \\ 5x_1 - 4x_2 = 6 \end{cases} . \tag{1.9}$$

Eliminoimme muuttujan x_1 jälkimmäisestä yhtälöstä, jolloin saamme muuttujan x_2 ratkaistuksi. Eliminointi onnistuu vähentämällä ensimmäinen yhtälö $5/4$ kertaa jälkimmäisestä yhtälöstä puolittain. Nimittäin tällöin jälkimmäinen yhtälö saa muodon

$$\begin{aligned} 5x_1 - 4x_2 - \frac{5}{4} \cdot (4x_1 + 2x_2) &= 6 - \frac{5}{4} \cdot 1 \\ 5x_1 - 4x_2 - 5x_1 - 2.500x_2 &= 4.750 \\ -6.500x_2 &= 4.750 \\ x_2 &= -0.7308. \end{aligned}$$

Sijoittamalla tämä uusi jälkimmäinen yhtälö paikalleen yhtälöpariin 1.9 saamme yhtälöparin

$$\begin{cases} 4x_1 + 2x_2 = 1 \\ x_2 = -0.7308 \end{cases} \tag{1.10}$$

Nyt vähentämällä vastavuoroisesti jälkimmäinen yhtälö 2 kertaa ensimmäisestä yhtälöstä puolittain yhtälöparissa (1.10) saamme

$$\begin{aligned} 4x_1 + 2x_2 - 2x_2 &= 1 - 2 \cdot (-0.7308) \\ 4x_1 &= 2.4616 \\ x_1 &= 0.6154. \end{aligned}$$

Olemme löytäneet ratkaisun

$$\left\{ \begin{array}{rcr} x_1 \qquad & = & 0.6154 \\ x_2 & = & -0.7308 \end{array} \right. \tag{1.11}$$

1.12

Esimerkissä 1.8 käytimme seuraavia huomiota:

(i) Yhtälön saa kertoa puolittain nollasta poikkeavalla luvulla.
(ii) Yhtälöön saa lisätä puolittain toisen yhtälön.

Näitä huomioita kutsutaan **alkeisrivioperaatioiksi**.

1.13 Esimerkki
Haluamme ratkaista yhtälöparin

$$\left\{ \begin{array}{rcrcr} 4x_1 & + & 2x_2 & = & 1 \\ 8x_1 & + & 4x_2 & = & 3 \end{array} \right. . \tag{1.14}$$

Eliminoimme muuttujan x_1 jälkimmäisestä yhtälöstä vähentämällä ensimmäisen yhtälön 2 kertaa jälkimmäisestä yhtälöstä. Saamme yhtälöparin

$$\left\{ \begin{array}{rcrcr} 4x_1 & + & 2x_2 & = & 1 \\ & & 0 & = & 1 \end{array} \right. .$$

Koska $0 = 1$ on absurdia, näemme että yhtälöparilla (1.14) ei ole ratkaisua.

1.15 Esimerkki
Haluamme ratkaista yhtälöparin

$$\left\{ \begin{array}{rcrcr} 4x_1 & + & 2x_2 & = & 1 \\ 12x_1 & + & 6x_2 & = & 3 \end{array} \right. . \tag{1.16}$$

Eliminoimme muuttujan x_1 jälkimmäisestä yhtälöstä vähentämällä ensimmäisen yhtälön 3 kertaa jälkimmäisestä yhtälöstä. Saamme yhtälöparin

$$\begin{cases} 4x_1 + 2x_2 = 1 \\ 0 = 0 \end{cases}.$$

Koska $0 = 0$ aina, niin toinen yhtälö elimioitui täydellisesti. Tulkinta on, että kaikki pisteet suoralla $4x_1 + 2x_2 = 1$ ovat yhtälöparin ratkaisuja. Toisin sanoen yhtälöparin molemmat yhtälöt määräsivät saman suoran.

1.17 Harjoitustehtävä
Ratkaise harjoitustehtävän 1.6 yhtälöparit alkeisrivioperaatioilla.

1.18 Esimerkki
Haluamme ratkaista yhtälöparin

$$\begin{cases} x_1 + ax_2 = 1 \\ 4x_1 - 2x_2 = 0 \end{cases}, \tag{1.19}$$

missä a on jokin kiinnitetty, mutta mielivaltainen parametri.

Vähentämällä ensimmäinen yhtälö 4 kertaa jälkimmäisestä yhtälöstä eliminoituu x_1 jälkimmäisestä yhtälöstä ja saamme yhtälöparin

$$\begin{cases} x_1 + ax_2 = 1 \\ (-4a-2)x_2 = -4 \end{cases}.$$

Muuttuja x_2 ratkeaa jälkimmäisestä yhtälöstä periaatteessa helposti:

$$\begin{aligned} (-4a-2)x_2 &= -4 \\ x_2 &= \frac{-4}{-4a-2} \\ x_2 &= \frac{2}{2a+1}. \end{aligned}$$

Ainoa ongelma tässä on, että emme saa jakaa nollalla. Siten joudumme olettamaan, että $a \neq -1/2$. Olemme siis saaneet yhtälöparin, jossa x_2 on ratkaistu jälkimmäisessä yhtälössä:

$$\begin{cases} x_1 + ax_2 = 1 \\ x_2 = \frac{2}{2a+1} \end{cases}. \tag{1.20}$$

Nyt pitää eliminoida muuttuja x_2 ensimmäisestä yhtälöstä. Tämä tapahtuu vähentämällä jälkimmäinen yhtälö a kertaa puolittain ensimmäisestä yhtälöstä. Saamme ratkaisun

$$\begin{cases} x_1 = 1 - \frac{2a}{2a+1} \\ x_2 = \frac{2}{2a+1} \end{cases}, \tag{1.21}$$

joka pätee, jos $a \neq -1/2$. Jos $a = -1/2$, niin alkuperäinen yhtälöpari (1.19) on muotoa

$$\begin{cases} x_1 - \frac{1}{2}x_2 = 1 \\ 4x_1 - 2x_2 = 0 \end{cases}$$

Kertomalla ensimmäinen yhtälö 4:llä saamme yhtälöparin

$$\begin{cases} 4x_1 - 2x_2 = 4 \\ 4x_1 - 2x_2 = 0 \end{cases},$$

mistä näemme että tapauksessa $a = -1/2$ yhtälöparilla (1.18) ei ole ratkaisua.

1.22 Harjoitustehtävä
Ratkaise yhtälöpari

$$\begin{cases} x_1 + 6x_2 = b \\ 4x_1 - 2x_2 = 1 \end{cases}$$

käyttämällä alkeisrivioperaatioita.

Tässä yhtälöparissa siis x_1 ja x_2 ovat muuttujia, joiden suhteen se on ratkaistava ja b on kiinnitetty, mutta mielivaltainen parametri.

1.23 Harjoitustehtävä
Millä parametrien a ja b arvoilla seuraavalla yhtälöparilla on (i) yksikäsitteinen ratkaisu, (ii) äärettömästi ratkaisuja ja (iii) ei yhtään ratkaisua

$$\begin{cases} 5x_1 - ax_2 = 0 \\ 10x_1 - 2x_2 = b \end{cases}.$$

Yhtälöryhmän ratkaisu alkeisrivioperaatioilla*

Yleinen lineaarinen n:n muuttujan ja m:n yhtälön yhtälöryhmä on muotoa

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

Jos tällä yhtälöryhmällä on yksikäsitteinen ratkaisu, niin se voidaan löytää alkeisrivioperaatioilla periaatteessa seuraavasti:

(i) Eliminoidaan muuttuja x_1 yhtälöistä $2, \ldots, m$ lisäämällä yhtälö 1 sopivasti kerrottuna niihin. Esimerkiksi yhtälöstä 2 saadaan x_1 eliminoitua vähentämällä siitä puolittain yhtälö 1 kerrottuna vakiolla A_{21}/A_{11}.

(ii) Toistetaan kohta (i) muuttujille $x_2, \ldots, x_{m-1}$, jolloin kaikki muuttujat $x_1, \ldots, x_j$ on eliminoitu yhtälöistä $j+1, \ldots, m$ kaikilla $j = 2, \ldots, m-1$.

(iii) On saatu **porrasmuotoinen** yhtälöryhmä

$$\begin{cases} B_{11}x_1 & + & B_{12}x_2 & + & \cdots & + & B_{1n}x_n & = & c_1 \\ & & B_{22}x_2 & + & \cdots & + & B_{2n}x_n & = & c_2 \\ & & & & & & & \vdots & \\ & & & & & & B_{mn}x_n & = & c_m \end{cases}.$$

Tästä ratkeaa $x_n = c_m/B_{mn}$ ja yleisesti yhtälöryhmä ratkeaa esimerkiksi takaisinsijoittamalla muuttujat $x_n, x_{n-1}, \ldots, x_2$ yhtälöihin $(m-1), \ldots, 1$.

Mikäli yllä kuvattu algoritmi toimii moitteettomasti, löytää se yhtälöryhmän yksikäsitteisen ratkaisun. Muussa tapauksessa on syytä epäillä, että yhtälöryhmällä ei ole ratkaisua tai ratkaisuja on äärettömän paljon. Jos $n \neq m$, niin tyypillisesti yhtälöryhmällä ei ole yksikäsitteistä ratkaisua. Jos muuttujia on liikaa $(n > m)$, niin tyypillisesti ratkasiuja on äärettömästi, koska pelivaraa on liikaa. Vastaavasti jos muuttujia on liian vähän $(n < m)$, niin tyypillisesti ratkaisuja ei ole, koska on liikaa rajoitteita.

Luku 2

Vektorit lähinnä tasossa

Karteesinen muoto ja napamuoto

Samaistamme tason pisteet (x_1, x_2) tason **vektorin** $\mathbf{x} = (x_1, x_2)$ kanssa, joka alkaa origosta ja päättyy pisteeseen (x_1, x_2). Tätä tason pisteen esitysmuotoa kutsutaan **karteesiseksi** ja tällaista vektoria kutsutaan myös nimellä **paikkavektori**.

2.1 Huomautus
Käytämme vektoreille lihavoitua merkintää $\mathbf{x}$. Joissakin insinöörikirjoissa käytetään sen sijaan nuolimerkintää $\overrightarrow{x}$ tai vähemmän rumaa viivamerkintää $\bar{x}$.

Kahden vektorin $\mathbf{x} = (x_1, x_2)$ ja $\mathbf{y} = (y_1, y_2)$ summa on vektori

$$\mathbf{x} + \mathbf{y} \quad = \quad (x_1 + y_1\,,\, x_2 + y_2).$$

Summassa siis lasketaan yhteen komponenteittain.

Vektori $\mathbf{x}$ kerrottuna **skalaarilla** (eli luvulla) t on vektori

$$t\mathbf{x} \quad = \quad (tx_1, tx_2).$$

Skalaarilla kerrottaessa siis venytetään tai typistetään komponenteittain samalla luvulla.

2.2 Esimerkki
Jos $\mathbf{x} = (1, 0)$ ja $\mathbf{y} = (2, 2)$, niin

$$\begin{aligned} \mathbf{x} - 0.5\mathbf{y} &= (1,0) + (-0.5)\cdot(2,2) \\ &= (1,0) + (-1,-1) \\ &= (0,-1) \end{aligned}$$

Tason **koordinaattivektorit** ovat

$$\mathbf{e}_1 = (1,0) \qquad \text{ja} \qquad \mathbf{e}_2 = (0,1).$$

Jokainen vektori $\mathbf{x} = (x_1, x_2)$ voidaan esittää koordinaattivektorien avulla summana

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2. \tag{2.3}$$

2.4 Huomautus
Joissakin toisissa kirjoissa koordinaattivektoreille käytetään toisin merkintöjä. Insinööripuolella on suosittua merkitä $\overrightarrow{i}$, $\bar{i}$ tai $\mathbf{i}$ koordinaattivektorille $\mathbf{e}_1$ ja $\overrightarrow{j}$, $\bar{j}$ tai $\mathbf{j}$ koordinaattivektorille $\mathbf{e}_2$.

2.5 Esimerkki
Olkoon $\mathbf{x} = (4,2)$ ja $\mathbf{y} = (-1,3)$. Laskemme vektorin $\mathbf{z} = 5\mathbf{x} - 2\mathbf{y}$. Käyttämällä koordinaattiesitystä (2.3) voimme laskea

$$\begin{aligned} \mathbf{z} &= 5\mathbf{x} - 2\mathbf{y} \\ &= 5 \cdot (4\mathbf{e}_1 + 2\mathbf{e}_2) - 2 \cdot (-\mathbf{e}_1 + 3\mathbf{e}_2) \\ &= 20\mathbf{e}_1 + 10\mathbf{e}_2 + 2\mathbf{e}_1 - 6\mathbf{e}_2 \\ &= 22\mathbf{e}_1 + 4\mathbf{e}_2, \end{aligned}$$

eli $\mathbf{z} = (22,4)$.

2.6 Harjoitustehtävä
Olkoot $\mathbf{x} = (1,-1)$, $\mathbf{y} = (1,2)$ ja $\mathbf{z} = (0,3)$. Laske

(i) $2\mathbf{x} + \mathbf{y} - \mathbf{z}$,
(ii) $\mathbf{x} - 3\mathbf{y} + 2\mathbf{z}$.

Voit halutessasi käyttää koordinaatiesitystä (2.3).

Vektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ ovat lineaarisesti riippuvia, jos toinen voidaan esittää toisen avulla. Tasossa tämä tarkoittaa sitä, että vektorit ovat samalla suoralla. Toisin sanoen $\mathbf{v}_2 = t\mathbf{v}_1$, missä t on jokin skalaari. Vektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ ovat **lineaarisesti riippumattomia**, jos toista ei voi esittää toisen avulla. Lineaarisesti riippumattomat vektorit **virittävät** koko tason, mikä tarkoittaa sitä että jokainen tason piste $\mathbf{x}$ voidaan esittää niiden avulla:

$$\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2. \tag{2.7}$$

Lineaarisesti riippumattomia vektoreita kutsutaan myös **vapaiksi** ja kaavaa (2.7) kutsutaan vektorin $\mathbf{x}$ esitykseksi **kannassa** $(\mathbf{v}_1, \mathbf{v}_2)$

2.8 Esimerkki
Olkoon $\mathbf{v}_1 = (0,-1)$ ja $\mathbf{v}_2 = (-1,1)$. Haluamme löytää esityksen (2.7) yleiselle tason pisteelle $\mathbf{x} = (x_1, x_2)$.

Yleiselle vektoriparille $(\mathbf{v}_1, \mathbf{v}_2)$ vektoriyhtälö (2.7) tarkoittaa yhtälöparia

$$\begin{cases} a_1 v_{11} + a_2 v_{21} = x_1 \\ a_1 v_{12} + a_2 v_{22} = x_2 \end{cases},$$

missä a_1 ja a_2 ovat muuttujia, joiden suhteen yhtälöpari on ratkaistava. Vektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ on annettu, ja vektori $\mathbf{x}$ on mielivaltainen, mutta kiinnitetty parametri.

Tarkastelemme siis yhtälöparia

$$\begin{cases} - a_2 = x_1 \\ -a_1 + a_2 = x_2 \end{cases}.$$

Vaihtamalla yhtälöiden järjestystä ja kertomalla uusi jälkimmäinen yhtälö -1:llä saamme yhtälöparin

$$\begin{cases} -a_1 + a_2 = x_2 \\ a_2 = -x_1 \end{cases}.$$

Vähentämällä jälkimmäinen yhtälö ensimmäisestä yhtälöstä saamme yhtälöparin

$$\begin{cases} -a_1 = x_2 + x_1 \\ a_2 = -x_1 \end{cases}.$$

Tästä näemmekin jo ratkaisun

$$\begin{cases} a_1 = -(x_1 + x_2) \\ a_2 = -x_1 \end{cases},$$

eli kannassa $(\mathbf{v}_1, \mathbf{v}_2)$ vektorin $\mathbf{x} = (x_1, x_2)$ esitys on

$$\mathbf{x} = -(x_1 + x_2)\mathbf{v}_1 - x_1 \mathbf{v}_2.$$

Vektori $\mathbf{x}$ voidaan karteesisen muodon $\mathbf{x} = (x_1, x_2)$ lisäksi esittää myös **napamuodossa** eli **napakoordinaatistossa** $\mathbf{x} = \langle r, \theta \rangle$, missä r on vektorin $\mathbf{x}$ pituus eli **normi** ja θ on vektorin $\mathbf{x}$ **virittämän** origon kautta kulkevan suoran ja x_1-akselin välinen kulma. Valitsemme kulman mitaksi radiaanit ja tarkastelemme vain positiivisia kulmia. Toisin sanoen $\theta \in [0, 2\pi)$.

Vektorin $\mathbf{x}$ normin r, jolle käytämme myös merkintää $\|\mathbf{x}\|$, saamme tarkastelemalla suorakulmaista kolmiota, jonka kärjet ovat pisteet $(0,0)$, $(x_1, 0)$ ja (x_1, x_2). Tällöin kateettien pituudet ovat x_1 ja x_2, ja hypotenuusan pituus on vektorin pituus. Siten Pythagoraan lauseen nojalla

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}.$$

Jos vektorin pituus on yksi, kutsumme sitä **yksikkövektoriksi**. Karteesisesta muodosta $\mathbf{x} = (x_1, x_2)$ saadaan vektorin $\mathbf{x}$ suuntainen yksikkövektori jakamalla se pituudellaan:

$$\hat{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|}.$$

Napamuodosta $\mathbf{x} = \langle r, \theta \rangle$ vektorin $\mathbf{x}$ suuntaisen yksikkövektorin rakentaminen on helppoa:

$$\hat{\mathbf{x}} = \langle 1, \theta \rangle.$$

2.9 Esimerkki
Vektorin $\mathbf{x} = (2,4)$ normi on

$$\begin{aligned} \|\mathbf{x}\| &= \sqrt{2^2 + 4^2} \\ &= \sqrt{4+16} \\ &= \sqrt{20} \\ &= \sqrt{2^2 \cdot 5} \\ &= 2\sqrt{5} \\ &\approx 4.4721. \end{aligned}$$

Vektorin $\mathbf{x} = (2,4)$ suuntainen yksikkövektori on

$$\begin{aligned} \hat{\mathbf{x}} &= \frac{\mathbf{x}}{\|\mathbf{x}\|} \\ &= \frac{(2,4)}{2\sqrt{5}} \\ &= \left(\frac{2}{2\sqrt{5}}, \frac{4}{2\sqrt{5}} \right) \\ &= \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \\ &\approx (0.44721, 0.89443). \end{aligned}$$

Etsimme nyt napamuodon $\langle r, \theta \rangle$ karteesisesta muodosta (x_1, x_2). Tiedämme jo, että

$$r = \|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}.$$

Olkoon sitten vektori $\mathbf{x} = (x_1, x_2)$ ensimmäisessä neljanneksessä, eli $x_1, x_2 > 0$. Suoraan tangentin määritelmästä seuraa, että

$$\tan\theta = \frac{x_2}{x_1}.$$

Siten

$$\theta = \arctan\frac{x_2}{x_1},$$

Yleinen tapaus, jossa $\mathbf{x}$ on missä tahansa neljänneksessä saadaan geometrisella päättelyllä. Päättelyn tulos on että jos haluamme, että $\theta \in [0, 2\pi)$, voimme laskea sen kaavalla

$$\theta = \begin{cases} \arctan \frac{x_2}{x_1}, & \text{jos } x_1 > 0 \text{ ja } x_2 \geq 0, \\ \arctan \frac{x_2}{x_1} + 2\pi, & \text{jos } x_1 > 0 \text{ ja } x_2 < 0, \\ \arctan \frac{x_2}{x_1} + \pi, & \text{jos } x_1 < 0, \\ \frac{\pi}{2}, & \text{jos } x_1 = 0 \text{ ja } x_2 > 0, \\ \frac{3\pi}{2}, & \text{jos } x_1 = 0 \text{ ja } x_2 < 0, \end{cases}$$

Jos $x_1, x_2 = 0$ molemmat, niin kulma θ ei ole määritelty. Jos taas haluamme että $\theta \in (-\pi, \pi]$, voimme laskea sen kaavalla

$$\theta = \begin{cases} \arctan \frac{x_2}{x_1}, & \text{jos } x_1 > 0, \\ \arctan \frac{x_2}{x_1} + \pi, & \text{jos } x_1 < 0 \text{ ja } x_2 \geq 0, \\ \arctan \frac{x_2}{x_1} - \pi, & \text{jos } x_1 < 0 \text{ ja } x_2 < 0, \\ \frac{\pi}{2}, & \text{jos } x_1 = 0 \text{ ja } x_2 > 0, \\ -\frac{\pi}{2}, & \text{jos } x_1 = 0 \text{ ja } x_2 < 0, \end{cases}$$

Edelleen, jos $x_1, x_2 = 0$ molemmat, ei kulma θ ole määritelty.

2.10 Huomautus
Kulma θ ei ole yksikäsitteinen, vaan kaikki kulmat $\theta + k2\pi$, $k \in \mathbb{Z}$, esittävät samaa kulmaa θ.

Kääntäen karteesinen muoto (x_1, x_2) saadaan napamuodosta $\langle r, \theta \rangle$ tarkastelemalla suorakulmaista kolmiotamme, joka kärjet ovat pisteet $(0,0)$, $(x_1, 0)$ ja (x_1, x_2). Tällöin kolmion hypotenuusan pituus on r, joten sinin ja kosinin määritelmästä seuraa, että

$$\begin{aligned} x_1 &= r \cos \theta, \\ x_2 &= r \sin \theta. \end{aligned}$$

2.11 Esimerkki

(i) Olkoon $\mathbf{x} = (2, -4)$. Etsimme sen napamuodon $\mathbf{x} = \langle r, \theta \rangle$. Vektorin $\mathbf{x}$ normi on

$$\begin{aligned} r &= \|\mathbf{x}\| \\ &= \sqrt{x_1^2 + x_2^2} \\ &= \sqrt{2^2 + (-4)^2} \\ &= \sqrt{4 + 16} \\ &= \sqrt{20} \\ &\approx 4.4721. \end{aligned}$$

Vektorin $\mathbf{x}$ kulma on

$$\begin{aligned} \theta &= \arctan \frac{-4}{2} \\ &\approx -1.1071, \end{aligned}$$

minkä saamme välille $[0, 2\pi)$ vähentämällä se koko kehän kulmasta 2π:

$$\begin{aligned}\theta &\approx 2\pi - 1.1071 \\ &= 5.1761.\end{aligned}$$

Olemme löytäneet napaesityksen

$$\mathbf{x} \approx \langle 4.4721, 5.1761 \rangle.$$

(ii) Olkoon $\mathbf{x} = \langle 3, \pi/4 \rangle$. Etsimme sen karteesisen muodon $\mathbf{x} = (x_1, x_2)$. Ensimmäinen koordinaatti on

$$\begin{aligned}x_1 &= r\cos\theta \\ &= 3\cos\frac{\pi}{4} \\ &\approx 2.1213.\end{aligned}$$

Toinen koordinaatti on

$$\begin{aligned}x_2 &= r\sin\theta \\ &= 3\sin\frac{\pi}{4} \\ &\approx 2.1213.\end{aligned}$$

Olemme löytäneet karteesisen esityksen

$$\mathbf{x} \approx (2.1213, 2.1213).$$

2.12 Harjoitustehtävä

(i) Olkoon $\mathbf{x} = (1, 2.5)$. Etsi $\mathbf{x}$:n napaesitys.

(ii) Olkoon $\mathbf{x} = \langle 4, 1.4 \rangle$. Etsi $\mathbf{x}$:n karteesinen esitys.

Kahden vektorin $\mathbf{x} = (x_1, x_2)$ ja $\mathbf{y} = (y_1, y_2)$ välinen **pistetulo** on

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2. \tag{2.13}$$

2.14 Huomautus

Rakkaalla lapsella on monta nimeä ja symbolia. Pistetuloa kutsutaan myös sisätuloksi, skalaarituloksi ja projektiotuloksi. Sille käytetään myös merkinnän $\mathbf{x} \cdot \mathbf{y}$ lisäksi merkintöjä $(\mathbf{x}, \mathbf{y})$, $(\mathbf{x}|\mathbf{y})$, $\langle \mathbf{x}, \mathbf{y} \rangle$, $\langle \mathbf{x}|\mathbf{y} \rangle$, ja $\mathbf{x}^\top \mathbf{y}$.

2.15 Huomautus
Pistetulo on **symmetrinen** ja **bilineaarinen**. Symmetrisyys tarkoittaa sitä, että

$$\mathbf{x}\cdot\mathbf{y} \;=\; \mathbf{y}\cdot\mathbf{x}.$$

Bilineaarisuus tarkoittaa sitä, että pistetulo on lineaarinen sekä $\mathbf{x}$:n että $\mathbf{y}$:n suhteen. Esimerkiksi lineaarisuus $\mathbf{x}$:n suhteen tarkoittaa sitä, että kaikilla skalaareilla a_1 ja a_2 sekä vektoreilla $\mathbf{x}_1$ ja $\mathbf{x}_2$ sekä $\mathbf{y}$ pätee

$$(a_1\mathbf{x}_1 + a_2\mathbf{x}_2)\cdot\mathbf{y} \;=\; a_1(\mathbf{x}_1\cdot\mathbf{y}) + a_2(\mathbf{x}_2\cdot\mathbf{y}).$$

2.16 Huomautus
Pistetulon yhteys normiin on ilmeinen:

$$\|\mathbf{x}\|^2 \;=\; \mathbf{x}\cdot\mathbf{x}.$$

Kääntäen pätee niin sanotusta suunnikassäännöstä

$$2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 \;=\; \|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2$$

seuraavat polarisaatioidentiteetit:

$$\begin{aligned}
\mathbf{x}\cdot\mathbf{y} &= \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right)\\
&= \frac{1}{2}\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right).
\end{aligned}$$

Pistetulo $\mathbf{x}\cdot\mathbf{y}$ määrää vektorien $\mathbf{x}$ ja $\mathbf{y}$ välisen kulman. Tämä ei ole ihan ilmiselvää pistetulon karteesesta määritelmästä (2.13). Napamuotoinen määritelmä kuitenkin paljastaa totuuden. Olkoot $\mathbf{x} = \langle r_{\mathrm{x}}, \theta_{\mathrm{x}}\rangle$ ja $\mathbf{y} = \langle r_{\mathrm{y}}, \theta_{\mathrm{y}}\rangle$. Tällöin

$$\begin{aligned}
\mathbf{x}\cdot\mathbf{y} &= x_1y_1 + x_2y_2\\
&= r_{\mathrm{x}}\cos\theta_{\mathrm{x}}\; r_{\mathrm{y}}\cos\theta_{\mathrm{y}} + r_{\mathrm{x}}\sin\theta_{\mathrm{x}}\; r_{\mathrm{y}}\sin\theta_{\mathrm{y}}\\
&= r_{\mathrm{x}}r_{\mathrm{y}}\left(\cos\theta_{\mathrm{x}}\cos\theta_{\mathrm{y}} + \sin\theta_{\mathrm{x}}\sin\theta_{\mathrm{y}}\right).
\end{aligned}$$

Nyt muistamme tutun trigonometrisen kaavan

$$\cos(\alpha-\beta) \;=\; \cos\alpha\cos\beta + \sin\alpha\sin\beta.$$

Tämä antaa meille napamuotoisen sisätulon määritelmän

$$\mathbf{x}\cdot\mathbf{y} \;=\; r_{\mathrm{x}}r_{\mathrm{y}}\cos(\theta_{\mathrm{x}}-\theta_{\mathrm{y}}). \tag{2.17}$$

Napamuotoinen sisätulon määritelmä (2.17) antaa meille seuraavan kosinilauseen:

2.18

Pistetulo määrittää vektorien $\mathbf{x}$ ja $\mathbf{y}$ välisen kulman θ **kosinilauseen** kautta:

$$\cos\theta \quad = \quad \hat{\mathbf{x}}\cdot\hat{\mathbf{y}}, \tag{2.19}$$

missä $\hat{\mathbf{x}}$ ja $\hat{\mathbf{y}}$ ovat $\mathbf{x}$:n ja $\mathbf{y}$:n suuntaiset yksikkövektorit: $\hat{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|$ ja $\hat{\mathbf{y}} = \mathbf{y}/\|\mathbf{y}\|$.

Yksi seuraus tästä on se, että vektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat **kohtisuoria** jos ja vain jos

$$\mathbf{x}\cdot\mathbf{y} \quad = \quad 0.$$

Toinen seuraus on se, että vektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat yhdensuuntaisia jos ja vain jos

$$\mathbf{x}\cdot\mathbf{y} \quad = \quad \|\mathbf{x}\|\,\|\mathbf{y}\|.$$

2.20 Harjoitustehtävä

Olkoon $\mathbf{x} = (1,3)$ ja $\mathbf{y} = (2,a)$. Mikä tulee parametrin a olla, jotta vektorit $\mathbf{x}$ ja $\mathbf{y}$ olisivat

(i) yhdensuuntaisia,
(ii) kohtisuoria?

Suorat ja vektorit: parametrimuoto ja normaalivektorimuoto

Jos suora kulkee origon $\mathbf{0} = (0,0)$ ja pisteen $\mathbf{p} = (p_1, p_2)$ kautta, voidaan se esittää muodossa $t\mathbf{p}$, missä t on vapaa parametri. Yleinen suoran **parametrimuoto** on

$$t\mathbf{p} + \mathbf{b}, \tag{2.21}$$

missä t on vapaa parametri, $\mathbf{p}$ on suoran **suuntavektori** ja $\mathbf{b} = (b_1, b_2)$ on **siirtovektori**. Kaava (2.21) tarkoittaa sitä, että kun t käy läpi kaikki luvut, niin kaava (2.21) piirtää tasoon suoran kuvaajan.

2.22 Huomautus

Esityksessä (2.21) ei ole mitään yksikäsitteistä. Esimerkiksi

$$t(1,1) + (0,2), \qquad -t(3,3) + (0,2) \quad \text{ja} \quad t(1,1) + (-4,-2)$$

esittävät samaa suoraa, jonka kulmakerroinmuoto on

$$x_2 \quad = \quad x_1 + 2.$$

Kulmakerroinmuotoa

$$x_2 = kx_1 + \beta$$

vastaa esimerkiksi parametrimuoto

$$t(1, k) + (0, \beta).$$

Normaalimuotoa

$$a_1x_1 + a_2x_2 = b \tag{2.23}$$

vastaavan parametrimuodon saamme vaikkapa käyttämällä apuna kulmakerroinmuotoa:

$$\begin{aligned} a_1x_1 + a_2x_2 &= b \\ a_2x_2 &= -a_1x_1 + b \\ x_2 &= -\frac{a_1}{a_2}x_1 + \frac{b}{a_2}. \end{aligned}$$

Eräs normaalimuotoa (2.23) vastaava parametrimuoto on siis

$$t(1, -a_1/a_2) + (0, b/a_2).$$

Nopeuttamalla parametrin t "juoksuvauhdin" a_2-kertaiseksi, saamme parametrimuodon

$$t(a_2, -a_1) + (0, b/a_2).$$

Jos suora kulkee pisteiden $\mathbf{p} = (p_1, p_2)$ ja $\mathbf{q} = (q_1, q_2)$, niin kaikki pisteet $\mathbf{x}$ pisteiden $\mathbf{p}$ ja $\mathbf{q}$ välisellä suoran pätkällä voidaan kirjoittaa muodossa $\mathbf{x} = t\mathbf{p} + (1-t)\mathbf{q}$, missä t kuuluu välille $[0, 1]$. Tästä laajentamalla näemme että pisteiden $\mathbf{p}$ ja $\mathbf{q}$ kautta kulkevalla suoralla on parametriesitys

$$t\mathbf{p} + (1-t)\mathbf{q}, \tag{2.24}$$

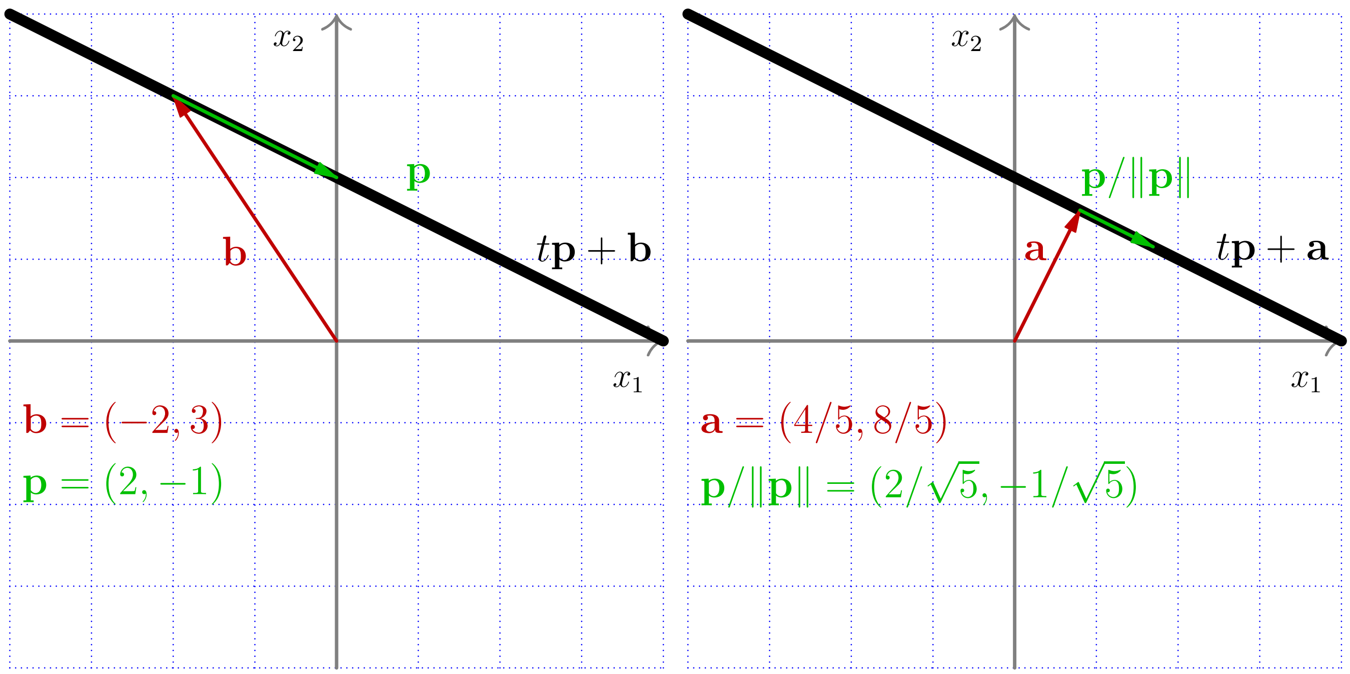
missä t on vapaa parametri.

Kutsuimme esitystä (2.23) suoran **normaalimuodoksi**. Merkitsemme $\mathbf{a} = (a_1, a_2)$. Tällöin huomaamme, että (2.23) voidaan kirjoittaa sisätulon avulla

$$\mathbf{a} \cdot \mathbf{x} = b.$$

Jos $b = 0$ (ja siten suora kulkee origon kautta), niin esitys (2.23) saa muodon $\mathbf{a} \cdot \mathbf{x} = 0$, mikä tarkottaa, että $\mathbf{a}$ ja $\mathbf{x}$ ovat **kohtisuoria**. Toinen tapa sanoa tämä on, että $\mathbf{a}$ on $\mathbf{x}$:n **normaali**. Nimitys normaalimuoto juontuu tästä.

Huomautus: Seuraava tarina on vähän hankala, mutta alla oleva kuva selventänee tarinaa.



Etsimme sitten niin sanotun **normaalivektorimuodon**, kun suora ei kulje origon kautta. Tarkoitus on esittää suora niin, että sen suuntavektori $\mathbf{p}$ on annetun vektorin $\mathbf{a}$ kohtisuora ja suora kulkee pisteen $\mathbf{a}=(a_1,a_2)$ kautta. Tällöin vektori $\mathbf{a}$ määrää suoran yksikäsitteisesti. Erityisesti tällöin $\|\mathbf{a}\|$ on suoran etäisyys origosta. Lähdemme liikkeelle parametrimuodosta (2.21). Jos $\mathbf{p}$ ja $\mathbf{b}$ ovat kohtisuoria, eli $\mathbf{p}\cdot\mathbf{b}=0$, olemme löytäneet normaalivektorin: $\mathbf{a}=\mathbf{b}$. Oletamme sitten, että $\mathbf{p}\cdot\mathbf{b}\neq 0$. Muuttamalla “kulkuvauhtia” t, voimme olettaa että $\mathbf{p}$ on yksikkövektori, eli $\|\mathbf{p}\|=1$. Tämä helpottaa merkintöjä hieman jatkossa. Tarkoituksenamme on siis löytää sellainen suoran (2.21) piste $\mathbf{a}=t_0\mathbf{p}+\mathbf{b}$, että $\mathbf{a}\cdot\mathbf{p}=0$. Saamme siis parametrille t_0 yhtälön

$$(t_0\mathbf{p}+\mathbf{b})\cdot\mathbf{p} \;=\; 0.$$

Tästä t_0 ratkeaa suoraviivaisesti:

$$\begin{aligned}(t_0\mathbf{p}+\mathbf{b})\cdot\mathbf{p} &= 0\\ t_0\mathbf{p}\cdot\mathbf{p}+\mathbf{b}\cdot\mathbf{p} &= 0\\ t_0+\mathbf{b}\cdot\mathbf{p} &= 0\\ t_0 &= -\mathbf{b}\cdot\mathbf{p}.\end{aligned}$$

Yllä käytimme hyväksi ensiksi sisätulon bilinearisuutta, ja toiseksi sitä että $\mathbf{p}$ on yksikkövektori, jolloin $\mathbf{p}\cdot\mathbf{p}=\|\mathbf{p}\|^2=1$. Nyt suoraa (2.21) vastaava normaalivektori $\mathbf{a}$ saadaan sijoittamalla saatu $t_0=-\mathbf{b}\cdot\mathbf{p}$ suoran parametrimuotoon (2.21):

$$\begin{aligned}\mathbf{a} &= t_0\mathbf{p}+\mathbf{b}\\ &= (-\mathbf{b}\cdot\mathbf{p})\mathbf{p}+\mathbf{b}\end{aligned}$$

Näin saatua vektoria $\mathbf{a}$ kutsumme suoran **normaalivektoriksi** ja esitysmuotoa $t\mathbf{p}+\mathbf{a}$ suoran **normaalivektorimuodoksi**.

2.25 **Esimerkki**
Olkoon suora annettu normaalimuodolla

$$x_1 + x_2 \;=\; 1. \tag{2.26}$$

Etsimme sen normaalivektorimuodon. Heti aluksi huomaamme (piirrä kuva), että eräs parametrimuoto $t\mathbf{p} + \mathbf{b}$ suorallemme saadaan asettamalla $\mathbf{p} = (1, -1)$ ja $\mathbf{b} = (0, 1)$. Suuntavektori $\mathbf{p}$ saadaan yksikkövektoriksi jakamalla se pituudellaan $\|\mathbf{p}\| = \sqrt{2}$. Olkoon siis uusi suuntavektori $\mathbf{p} = (1/\sqrt{2}, -1/\sqrt{2})$. Tällöin

$$\begin{aligned}
\mathbf{a} &= (-\mathbf{b}\cdot\mathbf{p})\mathbf{p} + \mathbf{b} \\
&= \big(-(0,1)\cdot(1/\sqrt{2}, -1/\sqrt{2})\big)\mathbf{p} + \mathbf{b} \\
&= \frac{1}{\sqrt{2}}\mathbf{p} + \mathbf{b} \\
&= \frac{1}{\sqrt{2}}(1/\sqrt{2}, -1/\sqrt{2}) + (0,1) \\
&= (1/2, -1/2) + (0,1) \\
&= (1/2, 1/2).
\end{aligned}$$

2.27 **Harjoitustehtävä**
Etsi seuraavien suorien jokin parametriesitys ja normaalivektorimuoto:

(i) $x_1 = 8x_2 - 3$,
(ii) $2x_1 - 3x_2 = 7$.

Vektorit korkeammissa avaruuksissa*

Yleinen n-ulotteinen vektori on muotoa $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Yleisten vektorien $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ ja $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ summa määrätään pisteittäin eli komponenteittain:

$$\mathbf{x} + \mathbf{y} \;=\; (x_1 + y_1,\; x_2 + y_2,\; \ldots,\; x_n + y_n).$$

Skalaarilla t kerrottu vektori on komponenteittain venytys tai typistys

$$t\mathbf{x} \;=\; (tx_1,\; tx_2,\; \ldots\;, tx_n)$$

Vektorit $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ ovat lineaarisesti riippumattomia eli vapaita, jos mikään niistä ei ole minkään muiden lineaarikombinaatio. Eli ei voi olla esimerkiksi niin, että

$$\mathbf{v}_4 \;=\; 2\mathbf{v}_1 - 0.3\mathbf{v}_2 + 10\mathbf{v}_6.$$

Jos vektorit $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ ovat lineaarisesti riippumattomia, niin jokainen niistä määrittelee uuden suunnan, jossa n-ulotteisessa avaruudessa voi kulkea. Yleisessä n-ulotteisessa avaruudessa voi olla korkeintaan n lineaarisesti riippumatonta vektoria.

Vektorin $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ normi saadaan käyttämällä Pythagoraan lausetta $(n-1)$ kertaa peräkkäin:

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Vektorien $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ ja $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ pistetulo on

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Pistetulo on symmetrinen ja bilineaarinen. Kosinikaava pätee. Lisäksi kaava $\|\mathbf{x}\|^2 = \mathbf{x} \cdot \mathbf{x}$ pätee.

Napamuodot ovat viheliäisiä n-ulotteisessa avaruudessa. Niitä on syytä välttää jos vain mahdollista.

Kaikki osion “Suorat ja vektorit: parametrimuoto ja normaalivektorimuoto” tulokset pätevät n-ulotteisessa avaruudessa luonnollisella tavalla yleistäen.

Luku 3

Matriisilaskentaa lähinnä tasossa

Matriisit ja niiden perusoperaatiot

Matriisi on lukuja taulukossa[1]. Tarkastelemme lähinnä ainoastaan (2×2)-matriiseja, eli matriiseja, jotka ovat muotoa

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

Tavallisesti tulkitsemme että vektorit $\mathbf{x} = (x_1, x_2)$ ovat (2×1)-matriiseja eli **pystyvektoreita**

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Matriisin **transpoosi** saadaan, kun sen rivit ja sarakkeet vaihdetaan päittäin. Merkitsemme transpoosioperaatiota yläindeksillä $^\top$.

Jos $\mathbf{x}$ on pystyvektori, niin $\mathbf{x}^\top = [x_1 \ x_2]$ on (1×2)-matriisi eli **vaakavektori**. Tarpeen vaatiessa saatamme tulkita vektorin $\mathbf{x} = (x_1, x_2)$ myös vaakavektoreiksi.

3.1 Esimerkki
Jos

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 30 & 40 \end{bmatrix},$$

niin

$$\mathbf{A}^\top = \begin{bmatrix} 1 & 30 \\ 2 & 40 \end{bmatrix}.$$

[1] ja musiikki on paineaaltoja ilmassa.

Matriisit ovat **samanmuotoisia**, jos niillä on yhtä monta riviä ja saraketta. Kaikki (2×2)-matriisit ovat samanmuotoisia keskenään. Samoin kaikki pystyvektorit ovat keskenään samanmuotoisia ja kaikki vaakavektorit ovat keskenään samanmuotoisia. Sen sijaan (2×2)-matriisit, pystyvektorit ja vaakavektorit ovat toisiinsa nähden erimuotoisia.

Samanmuotoisten **matriisien summa** määritellään alkioittain: jos

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{ja} \quad \mathbf{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

niin

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix}.$$

Vastaavasti, jos $\mathbf{x} = [x_1 \ x_2]$ ja $\mathbf{y} = [y_1 \ y_2]$ ovat vaakavektoreita, niin

$$\mathbf{x} + \mathbf{y} = [x_1 + y_1 \ x_2 + y_2].$$

Jos $\mathbf{x}$ on vaakavektori ja $\mathbf{y}$ on pystyvektori, niin $\mathbf{x} + \mathbf{y}$ ei ole määritelty, koska $\mathbf{x}$ ja $\mathbf{y}$ eivät ole samanmuotoisia. Samoin, jos $\mathbf{A}$ on (2×2)-matriisi ja $\mathbf{x}$ on joko pysty- tai vaakavektori, niin $\mathbf{A} + \mathbf{x}$ ei ole määritelty.[2]

Matriisin **kertominen skalaarilla** määritellään alkioittain samaan tapaan kuin vektoreilla: jos

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

ja λ on skalaari, niin

$$\lambda\mathbf{A} = \begin{bmatrix} \lambda A_{11} & \lambda A_{12} \\ \lambda A_{21} & \lambda A_{22} \end{bmatrix}.$$

3.2 Harjoitustehtävä

Olkoot

$$\mathbf{A} = \begin{bmatrix} 0 & 2 \\ 7 & 1 \end{bmatrix} \quad \text{ja} \quad \mathbf{B} = \begin{bmatrix} 3 & 0 \\ 6 & 0 \end{bmatrix}.$$

Laske

(i) $\mathbf{A} + 2\mathbf{B}$,
(ii) $33\mathbf{A}^\top - 7\mathbf{B}$.

Matriisin riveille ja sarakkeille käytämme **pallonotaatiota**. Matriisin

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

[2] Joskus on tosin luonnollista laajentaa määritelmiä “ilmeisellä tavalla”, mitä se sitten ikinä tarkoittaakaan.

i :s rivi on vaakavektori

$$\mathbf{A}_{i\bullet} \quad = \quad [A_{i1} \;\; A_{i2}]$$

ja sen j :s sarake on pystyvektori

$$\mathbf{A}_{\bullet j} \quad = \quad \begin{bmatrix} A_{1j} \\ A_{2j} \end{bmatrix}.$$

3.3 Huomautus
Pallonotaation voi tulkita niin, että pallo tarkoittaa "vapaana juoksevaa otusta". Esimerkiksi $\mathbf{A}_{\bullet i}$ tarkoittaa sitä, että sarake i on kiinnitetty ja rivi "juoksee vapaana". Toinen tapa tulkita pallonotaatio on lukea pallo sanalla "kaikki" (rivit tai sarakkeet).

3.4 Huomautus
Käyttämällä **lohkomatriisimerkintää** ja transpooseja havaitsemme, että esimerkiksi

$$\mathbf{A} \quad = \quad [\mathbf{A}_{\bullet 1} \;\; \mathbf{A}_{\bullet 2}] \quad = \quad [\mathbf{A}_{1\bullet}^{\top} \;\; \mathbf{A}_{2\bullet}^{\top}]^{\top}.$$

Matriisitulo määritellään hieman kinkkisesti: Jos $\mathbf{A}$ ja $\mathbf{B}$ ovat molemmat (2×2)-matriiseja, niin $\mathbf{AB}$ on (2×2)-matriisi $\mathbf{C}$, jonka alkiot ovat

$$C_{ij} \quad = \quad \mathbf{A}_{i\bullet} \cdot \mathbf{B}_{\bullet j}. \tag{3.5}$$

Toisin sanoen tulomatriisin $\mathbf{C} = \mathbf{AB}$ alkio C_{ij} on matriisin A i :nnen rivin ja matriisin $\mathbf{B}$ j :nnen sarakkeen pistetulo. Vielä toisin sanoen $\mathbf{C} = \mathbf{AB}$ on matriisi jonka alkiot ovat

$$C_{ij} = \sum_{k=1}^{2} A_{ik} B_{kj}.$$

3.6 Huomautus
Jos haluamme pitää tiukasti kiinni tulkinnasta, että vektorit ovat nimenomaan pystyvektoreita, niin silloin kaava (3.5) on formaalisti väärin ja se pitää korvata kaavalla

$$C_{ij} \quad = \quad \mathbf{A}_{i\bullet}^{\top} \cdot \mathbf{B}_{\bullet j}. \tag{3.7}$$

Kaava (3.7) eroaa kaavasta (3.5) ainoastaan siten, että $\mathbf{A}_{i\bullet}$ on vaakavektori ja $\mathbf{A}_{i\bullet}^{\top}$ on pystyvektori.

Yleisemmin, jos matriisilla $\mathbf{A}$ on yhtä monta saraketta kuin matriisilla $\mathbf{B}$ on rivejä, niin tulomatriisin $\mathbf{C} = \mathbf{AB}$ alkio C_{ij} on matriisin $\mathbf{A}$ i :nnen rivin $\mathbf{A}_{i\bullet}$ ja matriisin $\mathbf{B}$ j :nnen sarakkeen $\mathbf{B}_{\bullet j}$ välinen pistetulo. Formaalisti siis samoin kuten kaavassa (3.5).

3.8 Esimerkki
Olkoot

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

Tällöin $\mathbf{C} = \mathbf{AB}$ on (2×2)-matriisi ja se lasketaan seuraavasti:

$$\begin{aligned}
C_{11} &= A_{1\bullet} \cdot B_{\bullet 1} \\
&= (A_{11}, A_{12}) \cdot (B_{11}, B_{21}) \\
&= (0, 2) \cdot (3, 6) \\
&= 0 \cdot 3 + 2 \cdot 6 \\
&= 12, \\
\\
C_{12} &= A_{1\bullet} \cdot B_{\bullet 2} \\
&= (A_{11}, A_{12}) \cdot (B_{12}, B_{22}) \\
&= (0, 2) \cdot (8, 4) \\
&= 0 \cdot 8 + 2 \cdot 4 \\
&= 8, \\
\\
C_{21} &= A_{2\bullet} \cdot B_{\bullet 1} \\
&= (A_{21}, A_{22}) \cdot (B_{11}, B_{21}) \\
&= (7, 1) \cdot (3, 6) \\
&= 7 \cdot 3 + 1 \cdot 6 \\
&= 27, \\
\\
C_{22} &= A_{2\bullet} \cdot B_{\bullet 2} \\
&= (A_{21}, A_{22}) \cdot (B_{12}, B_{22}) \\
&= (7, 1) \cdot (8, 4) \\
&= 7 \cdot 8 + 1 \cdot 4 \\
&= 60.
\end{aligned}$$

Siispä

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 12 & 8 \\ 27 & 60 \end{bmatrix}.$$

3.9

Matriisitulo on siitä ikävä, että se ei ole vaihdannainen: yleisesti ottaen

$$\mathbf{AB} \neq \mathbf{BA}.$$

Muuten matriiseilla voi laskea kuten luvuilla:

(i) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$,
(ii) $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$,
(iii) $\mathbf{C}(\mathbf{A}+\mathbf{B}) = \mathbf{CA}+\mathbf{CB}$.

3.10 Harjoitustehtävä
Olkoot $\mathbf{A}$ ja $\mathbf{B}$ kuten esimerkissä 3.8. Laske $\mathbf{BA}$.

Vektorien välinen matriisitulo voidaan määritellä formaalisti samalla tavalla kuin matriisien. Tällöin osoittautuu, että jos $\mathbf{x} = [x_1\ x_2]^\top$ ja $\mathbf{y} = [y_1\ y_2]^\top$ ovat pystyvektoreja, niin $\mathbf{xy}$ ei ole määritelty eikä $\mathbf{x}^\top\mathbf{y}^\top$ ole määritelty, mutta

$$\mathbf{xy}^\top = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} [y_1 \quad y_2] = \begin{bmatrix} x_1y_1 & x_1y_2 \\ x_2y_1 & x_2y_2 \end{bmatrix}$$

ja

$$\mathbf{x}^\top\mathbf{y} = [x_1 \quad x_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_1 + x_2y_2 = \mathbf{x}\cdot\mathbf{y}.$$

Jos $\mathbf{A}$ on (2×2)-matriisi, niin $\mathbf{xA}$ ei ole määritelty, koska $\mathbf{x}$:llä on yksi sarake ja $\mathbf{A}$:lla on kaksi riviä. Samasta syystä, mutta rivit ja sarakkeet kääntäen, tulo $\mathbf{Ax}^\top$ ei ole määritelty. Sen sijaan

$$\mathbf{x}^\top\mathbf{A} = [x_1 \quad x_2] \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = [A_{11}x_1 + A_{21}x_2 \quad A_{12}x_1 + A_{22}x_2]$$

ja

$$\mathbf{Ax} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_{11}x_1 + A_{12}x_2 \\ A_{21}x_1 + A_{22}x_2 \end{bmatrix}. \tag{3.11}$$

3.12 Huomautus
Kaava (3.11) liittyy yhtälöpareihin. Nimittäin sen nojalla yhtälöpari

$$\begin{cases} A_{11}x_1 & + & A_{12}x_2 & = & b_1 \\ A_{21}x_1 & + & A_{22}x_2 & = & b_2 \end{cases}.$$

voidaan kirjoittaa matriisein ja vektorein kompaktisti muodossa

$$\mathbf{Ax} = \mathbf{b}.$$

Käänteismatriisi ja determinantti

Olkoon $\mathbf{0}$ (nolla) vektori tai matriisi, jonka alkiot ovat nollia.

3.13 Huomautus
Jos haluamme olla tarkkoja, niin meillä on erilaisia nollia:

$$\mathbf{0}_{2\times2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{0}_{2\times1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{ja} \quad \mathbf{0}_{1\times2} = [0\ \ 0].$$

Jatkossa kuitenkin käytämme merkintää $\mathbf{0}$ näille kaikille ja oletamme että asiayhteydestä käy selväksi, mitä nollaa tarkoitamme.

Matriisinolla $\mathbf{0}$ toimii matriisien yhteenlaskun tapaan kuten skalaarinolla 0 toimii skalaarien yhteenlaskussa: $\mathbf{A}+\mathbf{0}=\mathbf{A}=\mathbf{0}+\mathbf{A}$ kaikilla matriiseilla $\mathbf{A}$.

Skalaariykköstä 1 vastaa matriisipuolella **identiteettimatriisi**

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

Nimittäin tällöin $\mathbf{AI}=\mathbf{IA}=\mathbf{A}$ kaikilla (2×2)-matriiseilla $\mathbf{A}$.

Matriisilla jakaminen on kinkkistä. Haluamme määrittää sellaisen **käänteismatriisin** $\mathbf{A}^{-1}$, että $\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$ (jolloin myös automaattisesti $\mathbf{AA}^{-1}=\mathbf{I}$). Tämä vastaa skalaarijakolaskua, missä $a^{-1}a=1$, eli $a^{-1}=1/a$. Skalaaripuolella tiedämme, että nollalla ei saa jakaa. Matriisipuolella on enemmän rajoituksia.

Yritämme seuraavaksi kääntää (2×2)-matriiseja

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

Yksi tapa yrittää kääntää (2×2)-matriisi on tarkastella laajennettua (2×4)-**lohkomatriisia** $[\mathbf{A}\ \mathbf{I}]$ ja yrittää muuttaa tämä **alkeisrivioperaatioilla** laajennetuksi lohkomatriisiksi $[\mathbf{I}\ \mathbf{B}]$. Jos tässä onnistutaan, niin siloin $\mathbf{B}=\mathbf{A}^{-1}$. Menetelmä perustuu seuraavaan ajattelutapaan: Matriisin $\mathbf{A}$ kääntäminen tarkoittaa lineaarisen yhtälöryhmän

$$\mathbf{Ax} = \mathbf{y} \tag{3.14}$$

ratkaisemista, missä $\mathbf{y}$ on mielivaltainen. Tämän yhtälöryhmän yleinen ratkaisu $\mathbf{y}$:n funktiona on

$$\mathbf{A}^{-1}\mathbf{y} = \mathbf{x}. \tag{3.15}$$

Huomaamalla, että $\mathbf{x}=\mathbf{Ix}$ ja $\mathbf{y}=\mathbf{Iy}$, ja kirjoittamalla yhtälöryhmä (3.14) niin että muuttujat $\mathbf{x}$ ja $\mathbf{y}$ vastaavat sarakkeita, saamme laajennetun lohkomatriisin $[\mathbf{A}\ \mathbf{I}]$. Jos tämä laajennettu lohkomatriisi saadaan alkeisrivioperaatioilla muotoon $[\mathbf{I}\ \mathbf{B}]$, niin se vastaa yhtälöryhmää (3.15). Toisin sanoen $\mathbf{B}=\mathbf{A}^{-1}$.

Alkeisrivioperaatiot ovat samat, kuin yhtälöparien tapauksessa:

(i) Rivin saa kertoa (puolittain) nollasta poikkeavalla luvulla.
(ii) Rivin saa lisätä (puolittain) toiseen riviin.

3.16 Esimerkki
Haluamme kääntää matriisin

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

Laajennettu lohkomatriisi kääntämistä varten on

$$[\mathbf{A}\,\mathbf{I}] = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{bmatrix}.$$

Vähentämällä ensimmäinen rivi toisesta rivistä saamme

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

Jakamalla sitten ensimmäinen rivi 2:lla ja toinen rivi 4:llä saamme

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

Tästä päättelemme, että

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

3.17 Esimerkki
Haluamme kääntää matriisin

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

Laajennettu lohkomatriisi kääntämistä varten on

$$[\mathbf{A}\,\mathbf{I}] = \begin{bmatrix} 2 & 3 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{bmatrix}.$$

Vähentämällä ensimmäinen rivi toisesta rivistä saamme

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

Toiselta riviltä eliminoitui liikaa alkioita. Tästä eteenpäin emme voi jatkaa. Tämä antaa ymmärtää, että matriisi $\mathbf{A}$ ei ole kääntyvä.

3.18 Esimerkki
Haluamme kääntää matriisin

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 8 & a \end{bmatrix},$$

missä a on mielivaltainen, mutta kiinnitetty parametri.

Laajennettu lohkomatriisi kääntämistä varten on

$$[\mathbf{A}\ \mathbf{I}] = \begin{bmatrix} 4 & 2 & 1 & 0 \\ 8 & a & 0 & 1 \end{bmatrix}.$$

Vähentämällä ensimmäinen rivi 2 kertaa toisesta rivistä saamme

$$\begin{bmatrix} 4 & 2 & 1 & 0 \\ 8 & a & 0 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 4 & 2 & 1 & 0 \\ 0 & a-4 & -2 & 1 \end{bmatrix}.$$

Vähentämällä toinen rivi $2/(a-4)$ kertaa ensimmäisestä rivistä saamme

$$\begin{bmatrix} 4 & 2 & 1 & 0 \\ 0 & a-4 & -2 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 4 & 0 & 1+\frac{4}{a-4} & -\frac{2}{a-4} \\ 0 & a-4 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & \frac{a}{a-4} & -\frac{2}{a-4} \\ 0 & a-4 & -2 & 1 \end{bmatrix}.$$

Jakamalla toinen rivi luvulla $a-4$ ja sitten ensimmäinen rivi luvulla 4 saamme

$$\begin{bmatrix} 4 & 0 & \frac{a}{a-4} & -\frac{2}{a-4} \\ 0 & a-4 & -2 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 4 & 0 & \frac{a}{a-4} & -\frac{2}{a-4} \\ 0 & 1 & -\frac{2}{a-4} & \frac{1}{a-4} \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & \frac{a}{4a-16} & -\frac{1}{2a-8} \\ 0 & 1 & -\frac{2}{a-4} & \frac{1}{a-4} \end{bmatrix}.$$

Siten, jos $a \neq 4$, niin

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{a}{4a-16} & -\frac{1}{2a-8} \\ -\frac{2}{a-4} & \frac{1}{a-4} \end{bmatrix} = \frac{1}{a-4}\begin{bmatrix} \frac{a}{4} & -\frac{1}{2} \\ -2 & 1 \end{bmatrix}.$$

Jos taas $a = 4$, niin matriisi $\mathbf{A}$ ei ole kääntyvä. Perustelu tälle tulee myöhemmin.

3.19 Esimerkki
Haluamme kääntää matriisin

$$\mathbf{A} = \begin{bmatrix} a & 0 \\ b & 1 \end{bmatrix},$$

missä a ja b ovat mielivaltaisia, mutta kiinnitettyjä parametreja

Laajennettu lohkomatriisi kääntämistä varten on

$$[\mathbf{A}\ \mathbf{I}] = \begin{bmatrix} a & 0 & 1 & 0 \\ b & 1 & 0 & 1 \end{bmatrix}.$$

Jakamalla ensimmäinen rivi a:llä ja vähentämällä näin saatu rivi b kertaa toisesta rivistä saamme

$$\begin{bmatrix} a & 0 & 1 & 0 \\ b & 1 & 0 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & \frac{1}{a} & 0 \\ b & 1 & 0 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & \frac{1}{a} & 0 \\ 0 & 1 & -\frac{b}{a} & 1 \end{bmatrix}.$$

Näemme, että, jos $a \neq 0$, niin

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{a} & 0 \\ -\frac{b}{a} & 1 \end{bmatrix} = \frac{1}{a}\begin{bmatrix} 1 & 0 \\ -b & a \end{bmatrix}.$$

Jos $a = 0$, niin osoittautuu, että $\mathbf{A}$ ei ole kääntyvä.

Yleisen (2×2)-matriisin voi kääntää periaatteessa seuraavasti. Laajennettu lohkomatriisi on

$$[\mathbf{A}\,\mathbf{I}] = \begin{bmatrix} A_{11} & A_{12} & 1 & 0 \\ A_{21} & A_{22} & 0 & 1 \end{bmatrix}.$$

Aluksi vähennämme ensimmäisen rivin toisesta rivistä niin monta kertaa että saamme nollan toisen rivin ensimmäiseksi alkioksi. Seuraavaksi vähennämme toisen rivin ensimmäisestä rivistä niin monta kertaa, että saamme ensimmäisen rivin toisen alkion nollaksi. Lopuksi kerromme rivit sopivilla vakioilla. Merkitsemällä nokkelasti

$$D = A_{11}A_{22} - A_{12}A_{21}.$$

ja jättämällä kiusallisen rasittavat välivaiheet väliin, saamme

$$\begin{bmatrix} A_{11} & A_{12} & 1 & 0 \\ A_{21} & A_{22} & 0 & 1 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{bmatrix} 1 & 0 & \frac{A_{22}}{D} & -\frac{A_{12}}{D} \\ 0 & 1 & -\frac{A_{21}}{D} & \frac{A_{11}}{D} \end{bmatrix}$$

Tästä luemme, että yleinen ratkaisu (2×2)-matriisin $\mathbf{A}$ käänteismatriisille on

$$\mathbf{A}^{-1} = \frac{1}{D}\begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}. \tag{3.20}$$

Lukua

$$D = \det(\mathbf{A}) = A_{11}A_{22} - A_{12}A_{21}$$

kutsutaan matriisin $\mathbf{A}$ **determinantiksi**. Edellinen (sivuutettu) pyörittely antaa oikean lopputuloksen, jos ja vain jos $\det(\mathbf{A}) \neq 0$. Lisäksi, pienellä piirroksella näemme, että $|\det(\mathbf{A})|$ on matriisin $\mathbf{A}$ rivivektorien (tai yhtä hyvin sarakevektorien) virittämän suunnikkaan pinta-ala.

3.21

Seuraavat ovat yhtäpitäviä:

(i) Matriisi $\mathbf{A}$ on kääntyvä.
(ii) $\det(\mathbf{A}) \neq 0$.
(iii) Matriisin $\mathbf{A}$ sarakevektorit $\mathbf{A}_{\bullet 1}$ ja $\mathbf{A}_{\bullet 2}$ ovat lineaarisesti riippumattomia.
(iv) Ei ole olemassa sellaista lukua λ, että $\mathbf{A}_{\bullet 1} = \lambda \mathbf{A}_{\bullet 2}$.
(v) Matriisi $\mathbf{A}^\top$ on kääntyvä.
(vi) $\det(\mathbf{A}^\top) \neq 0$.
(vii) Matriisin $\mathbf{A}$ rivievektorit $\mathbf{A}_{1\bullet}$ ja $\mathbf{A}_{2\bullet}$ ovat lineaarisesti riippumattomia.
(viii) Ei ole olemassa sellaista lukua λ, että $\mathbf{A}_{1\bullet} = \lambda \mathbf{A}_{2\bullet}$.

3.22 Huomautus
Kuten tunnettua rakkaalla lapsella on monta nimeä. Kääntyvälle matriisille on myös nimet **säännöllinen**, **epäsingulaarinen** ja **ei-degeneroitunut**.

3.23 Harjoitustehtävä
Käännä, mikäli mahdollista, matriisit

(i)

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

(ii)

$$\mathbf{B} = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}.$$

3.24 Harjoitustehtävä
Määrää vapaat parametrit a ja b siten, että matriisi $\mathbf{A}$ on kääntyvä

$$\mathbf{A} = \begin{bmatrix} 2 & a \\ 1 & b \end{bmatrix}.$$

Keräämme tähän loppuun vielä matriisien transpoosien ja determinanttien sekä tulojen ja kääntämisen laskusääntöjä:

3.25 Huomautus
Transpoosi toteuttaa seuraavat laskusäännöt:

(i) $(\mathbf{A}^\top)^\top = \mathbf{A}$,
(ii) $(\mathbf{A}+\mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$,
(iii) $(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top$,
(iv) $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$,
(v) $(\lambda\mathbf{A})^\top = \lambda\mathbf{A}^\top$.

Determinantti toteuttaa seuraavat laskusäännöt:

(i) $\det(\mathbf{A}^\top) = \det(\mathbf{A})$,
(ii) $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$,
(iii) $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$,
(iv) $\det(\lambda\mathbf{A}) = \lambda^2\det(\mathbf{A})$.

Tulolle pätee valitettavasti yleisesti $\mathbf{AB} \neq \mathbf{BA}$, mutta onneksi

(i) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$,
(ii) $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$,
(iii) $\mathbf{C}(\mathbf{A}+\mathbf{B}) = \mathbf{CA}+\mathbf{CB}$.

Tulon käänteismatriisille pätee kaava

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

3.26 Harjoitustehtävä
Olkoot a ja b mielivaltaisia parametreja sekä

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

Laske $(\mathbf{AB})^{-1}$, silloin kun se on olemassa.

Matriisit korkeammissa avaruuksissa*

Yleisesti matriisi $\mathbf{A}$ on lukuja taulukossa, $(m \times n)$-matriisissa on m riviä ja n saraketta. Esimerkiksi (7×4)-matriisi on muotoa

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \\ A_{51} & A_{52} & A_{53} & A_{54} \\ A_{61} & A_{62} & A_{63} & A_{64} \\ A_{71} & A_{72} & A_{73} & A_{74} \end{bmatrix}.$$

Kaikki edellä esitetty, paitsi kääntäminen ja determinantti, toimii ilmeisellä tavalla $(m \times n)$-matriiseille. Esimerkiksi matriisitulo määritellään $(n \times p)$-matriisille $\mathbf{A}$ ja $(p \times m)$-matriisille $\mathbf{B}$ asettamalla $\mathbf{C} = \mathbf{AB}$ on se matriisi, jolle

$$C_{ij} = \mathbf{A}_{i\bullet} \cdot B_{\bullet j},$$

tai yhtä hyvin

$$C_{ij} = \sum_{k=1}^{p} A_{ik} B_{kj}.$$

Tämä on formaalisti sama kuin tason määritelmä (3.5).

Matriisin kääntäminen on takkuista: Jotta matriisi olisi käänntyvä, on sen oltava **neilömatriis**, eli muotoa $n \times n$. Neliömatriisin $\mathbf{A}$ käänteismatriisi määritellään toki helposti. Se on se matriisi $\mathbf{A}^{-1}$, joka toteuttaa $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, missä $\mathbf{I}$ on $(n \times n)$-identiteettimatriisi, eli matriisi, jossa on ykköset lävistäjällä ja nollat muualla. Käänteismatriisin voi myös yrittää löytää muuntamalla laajennetun $(n \times 2n)$-lohkomatriisin $[\mathbf{A}\,\mathbf{I}]$ alkeisrivioperaatioilla laajennetuksi lohkomatriisiksi $[\mathbf{I}\,\mathbf{B}]$, jolloin $\mathbf{A}^{-1} = \mathbf{B}$. Tämä onnistuu täsmälleen silloin, kun matriisin $\mathbf{A}$ determinantti ei ole nolla.

Yleisen matriisin determinantin laskeminen, jopa sen esittäminen, $(n \times n)$-tapauksessa on vaikeaa. Esitämme sen tässä vain ja ainoastaan peloittelutarkoituksessa:

$$\det(\mathbf{A}) = \sum_{j_1, j_2, \ldots, j_n} \operatorname{sgn}(j_1, j_2, \ldots, j_n) A_{1j_1} A_{2j_2} \cdots A_{nj_n},$$

missä summa käy yli kaikki lukujen $\{1, 2, \ldots, n\}$ permutaatiot, ja sgn on 1 tai -1 sen mukaan onko permutaatio parillinen vai pariton. Intuitiivisesti determinantti kuvaa (etumerkkiä vaille) matriisin $\mathbf{A}$ sarakevektorien (tai yhtä hyvin rivivektorien) virittämän n-ulotteisen hypersuunnikassärmiön "tilavuutta".

Luku 4

Matriisit ja lineaarikuvaukset

Tähän asti olemme katsoneet matriiseja algebrallisesti: ne ovat olleet vain tiettyjä otuksia, joita lasketaan jollain tietyllä tavalla. Tässä luvusssa esitämme matriiseille tulkinnan, eli katsantokannan, lineaarisina funktioina, eli kuvauksina. Lyhtyesti tämän luvun samoma on seuraava: jos ajattelemme muotoa $\mathbf{Ax}$ muuttujan $\mathbf{x}$ funktiona, on tämä funktio lineaarinen; ja kääntäen jos $\mathbf{f}(\mathbf{x})$ on lineaarinen funktio, niin se on välttämättä muotoa $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ jollekin matriisille $\mathbf{A}$.

Funktioalgebraa

Funktio eli **kuvaus** f joukolta X joukkoon Y on sääntö, joka liittää jokaiseen joukon X alkioon täsmälleen yhden joukon Y alkion. Tälle käytetään usein merkintää $f : X \to Y$. Joukkoa X kutsutaan funktion f **lähtö-** tai **määrittelyjoukoksi** ja joukkoa Y kutsutaan funktion f **maali-** tai **arvojoukoksi**. Määrittelyjoukon X alkioita x kutsutaan usein funktion argumenteiksi. Maalijoukon Y niitä alkioita, joille $y = f(x)$ jollakin x kutsutaan usein funktion **kuviksi**, **kuva-alkioiksi** tai **kuvapisteiksi**. Niitä määrittelyjoukon X pisteitä x, joille $y = f(x)$ kutsutaan kuvapisteen y **alkukuviksi** tai **alkupisteiksi**.

Funktio $f : X \to Y$ on

(i) **injektio** jos jokaisella kuvalla $y = f(x)$ on korkeintaan yksi alkukuva,
(ii) **surjektio** jos jokaisella kuvajoukon Y alkoilla y on jokin alkukuva x joukossa X,
(iii) **bijektio** jos se on sekä injektio että surjektio.

Bijektion $f : X \to Y$ **käänteisfunktio** on funktio $f^{-1} \colon Y \to X$, jolle $x = f^{-1}(y)$ tismalleen silloin kun $y = f(x)$. Jos f ei ole bijektio, niin sille ei ole käänteisfunktiota. Jos taas f on bijektio, niin sen käänteisfunktion käänteisfunktio on se itse: $(f^{-1})^{-1} = f$.

Reaalilukusuoralle käytämme merkintää $\mathbb{R}$. Reaalitaso on $\mathbb{R}^2$.

Vektoriarvoisille funktioille $\mathbf{f} \colon \mathbb{R} \to \mathbb{R}^2$ tai $\mathbf{f} \colon \mathbb{R}^2 \to \mathbb{R}^2$ käytämme tarvittaessa komponenttimerkintää $\mathbf{f} = (f_1, f_2)$ tai pystyvektorimerkintää $\mathbf{f} = [f_1 \;\; f_2]^\top$.

Jatkossa olemme kiinnostuneita lähinnä funktioista suoralta $\mathbb{R}$ tasolle $\mathbb{R}^2$, tasolta $\mathbb{R}^2$ suoralle $\mathbb{R}$ ja tasolta $\mathbb{R}^2$ itselleen.

4.1 Esimerkki
Olkoon funktio $f : \mathbb{R}^2 \to \mathbb{R}$ annettu kaavalla

$$\begin{aligned} f(\mathbf{x}) &= f\big([x_1 \;\; x_2]^\top\big) \\ &= f(x_1, x_2) \\ &= x_1 + 42x_2. \end{aligned}$$

Tällöin f on surjektio, sillä jokaisella kuvajoukon pisteellä $y \in \mathbb{R}$ on alkukuva. Esimerkiksi $\mathbf{x} = [y \;\; 0]^\top$ kelpaa alkukuvaksi $f(\mathbf{x}) = y$. Funktio f ei ole kuitenkaan bijektio, sillä se ei ole injetio. Esimerkiksi pisteen $y = 0$ alkukuvia ovat sekä $\mathbf{x} = [0 \;\; 0]^\top$ ja $\mathbf{x} = [-42 \;\; 1]^\top$, sillä sekä $f(0,0) = 0$ että $f(-42, 1) = 0$. Funktiolla $f : \mathbb{R}^2 \to \mathbb{R}$ ei ole käänteisfunktiota, koska se ei ole bijektio.

4.2 Huomautus
Funktio $f : X \to Y$ on bijektio täsmälleen silloin, kun yhtälöllä $y = f(x)$ on jokaiselle maalijoukon Y alkiolle y täsmälleen yksi ratkaisu x lähtöjoukossa X. Jos f ei ole injektio, niin siitä voi yrittää tehdä bijektion rajoittamalla lähtöjoukkoa X. Vastaavasti jos f ei ole surjektio, siitä voi yrittää tehdä bijektion rajoittamalla maalijoukkoa Y.

4.3 Esimerkki
Olkoon $\mathbf{f} = [f_1 \;\; f_2]^\top : \mathbb{R} \to \mathbb{R}^2$ annettu kaavalla

$$\mathbf{f}(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix} = \begin{bmatrix} x \\ 3x \end{bmatrix}.$$

Tällöin funktio $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$ ei ole surjektio, sillä esimerkiksi pisteellä $\mathbf{y} = [1 \;\; 2]^\top$ ei ole alkukuvaa. Funktio $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$ on kuitenkin injektio, sillä jos $\mathbf{f}(x_1) = \mathbf{y}$ ja $\mathbf{f}(x_2) = \mathbf{y}$, niin $x_1 = x_2 = y_1$

4.4 Esimerkki
Olkoon $\mathbb{R}$ reaalilukusuora, $\mathbb{R}_+$ sen ei-negatiiviset pisteet ja $\mathbb{R}_+^*$ sen aidosti positiiviset pisteet.

(i) Funktio $f : \mathbb{R} \to \mathbb{R}$, joka on annettu kaavalla $f(x) = x^2$ ei ole injektio, sillä esimerkiksi pisteet -2 ja 2 kuvautuvat molemmat pisteeksi 4. Se ei ole myöskään surjektio, sillä esimerkiksi pisteellä -1 ei ole alkukuvaa. Funktio $f : \mathbb{R} \to \mathbb{R}_+$ on surjektio. Funktio $f : \mathbb{R}_+ \to \mathbb{R}_+$ on bijektio ja sen käänteisfunktio $f^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$ on $f^{-1}(x) = \sqrt{x}$.

(ii) Funktio $f : \mathbb{R} \to \mathbb{R}$, joka on annettu kaavalla $f(x) = \mathrm{e}^x$ on injektio, sillä se on aidosti kasvava. Se ei ole kuitenkaan surjektio, sillä esimerkiksi pisteellä 0 ei ole alkukuvaa. Funktio $f : \mathbb{R} \to \mathbb{R}_+^*$ on bijektio ja sen käänteisfunktio $f^{-1} : \mathbb{R}_+^* \to \mathbb{R}$ on $f^{-1}(x) = \ln x$.

4.5 Esimerkki

(i) Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$ annettu kaavalla

$$f(\mathbf{x}) = x_1^2 + x_2^2.$$

Toisin sanoen $f(\mathbf{x})$ on vektorin $\mathbf{x}$ normin neliö. Tällöin f selvästikään ei ole injektio eikä surjektio. Rajoittuma $f : \mathbb{R}^2 \to \mathbb{R}_+$ on surjektio, muttei injektio.

(ii) Olkoon $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$ annettu kaavalla

$$\mathbf{f}(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix} = \begin{bmatrix} -x \\ \mathrm{e}^x \end{bmatrix}.$$

Tällöin $\mathbf{f}$ on injektio, sillä sen komponentit ovat molemmat injektioita, mutta $\mathbf{f}$ ei ole surjektio, sillä esimerkiksi pisteellä $[0\ \ 0]^\top$ ei ole alkukuvaa.

4.6 Esimerkki

Olkoon **polar** funktio, joka kuvaa karteesiset koordinaatit napakoordinaateiksi. Funktion **polar** määrittelyjoukko on taso $\mathbb{R}^2$ ja sen arvojoukko on $\mathbb{R}_+ \times [0, 2\pi)$, missä $\mathbb{R}_+$ tarkoittaa ei-negatiivisia reaalilukuja ja $[0, 2\pi)$ tarkoittaa reaalilukuja θ, joille $0 \le \theta < 2\pi$, ja tulomerkintä $\mathbb{R}_+ \times [0, 2\pi)$ tarkoittaa joukkoa, joka koostuu pareista $(r, \theta) = \langle r, \theta \rangle$, missä $r \in \mathbb{R}_+$ ja $\theta \in [0, 2\pi)$.

Funktion $\mathbf{polar} = [\mathrm{polar}_1\ \mathrm{polar}_2]^\top : \mathbb{R}^2 \to \mathbb{R}_+ \times [0, 2\pi)$ kaava on

$$\mathbf{polar}(\mathbf{x}) = \left\langle \|\mathbf{x}\|\,,\ \arctan\frac{x_2}{x_1} \right\rangle,$$

missä funktiosta arctan pitää valita sopiva haara.

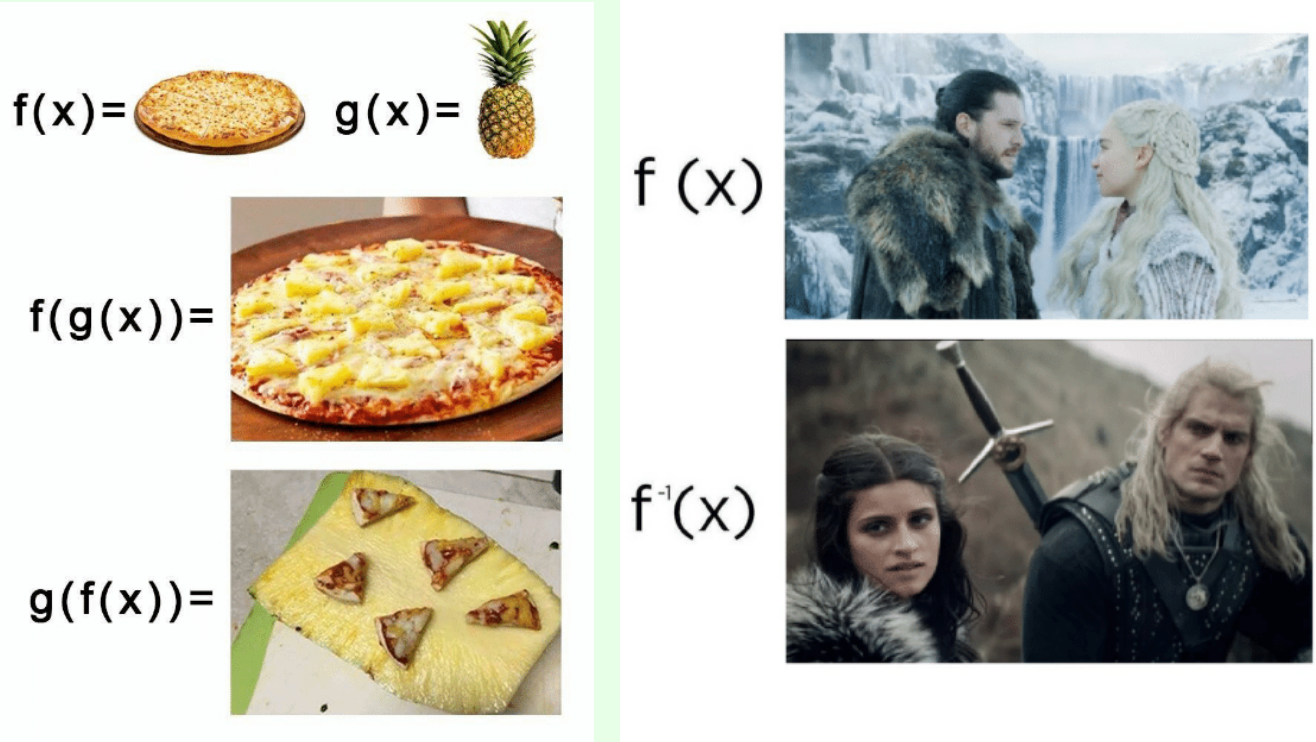
Funktion **polar** käänteisfunktio $\mathbf{polar}^{-1}$ tulee kaavasta

$$\mathbf{polar}^{-1}(r, \theta) = \begin{bmatrix} \mathrm{polar}_1^{-1}(r, \theta) \\ \mathrm{polar}_2^{-1}(r, \theta) \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}.$$

Jos $g : X \to Y$ ja $f : Y \to Z$, niin **yhdistetty funktio** $f \circ g : X \to Z$ on funktio, joka määräytyy kaavasta $(f \circ g)(x) = f(g(x))$. Jos f ja g ovat molemmat bijektioita, niin $f \circ g$ on myös bijektio ja $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

4.7 Huomautus

$f \circ g$ ei ole $g \circ f$, eikä ananaspizza ole pizza-ananas. Sen sijaan $f \circ f^{-1} = f^{-1} \circ f$, kuten Valtaistuinpelistä saadaan Noituri, ja päivastoin.



Yhdistetty kuvaus (vasemmalla) ja käänteiskuvaus (oikealla).

4.8 Harjoitustehtävä

(i) Olkoon $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ annettu kaavalla

$$\mathbf{f}(\mathbf{x}) \quad = \quad \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{bmatrix} \quad = \quad \begin{bmatrix} -x_1 \\ x_2^3 \end{bmatrix}.$$

Onko $\mathbf{f}$ bijektio?

(ii) Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}$, $\mathbf{g}\colon \mathbb{R}^2 \to \mathbb{R}^2$ ja $\mathbf{h}\colon \mathbb{R} \to \mathbb{R}^2$ annettu kaavoilla

$$f(\mathbf{x}) = x_1, \qquad \mathbf{g}(\mathbf{x}) = [2x_1 + x_2 \quad x_2]^\top \quad \text{ja} \quad \mathbf{h}(x) = [x \quad 3x]^\top$$

Onko yhtälöllä $(f \circ \mathbf{g} \circ \mathbf{h})(x) = y$ ratkaisu kaikilla y?

Suorien ja tasojen välisille funktioille f, g, h, jne. (emme nyt käytä lihavoitua merkintää) voidaan määritellä summa ja skalaarilla kertominen pisteittäin luonnollisella tavalla:

$$\begin{aligned} (f+g)(x) &= f(x)+g(x), \\ (\lambda f)(x) &= \lambda f(x). \end{aligned}$$

Jos tulkitsemme yhdistetyn funktion muodostamisen funktiotuloksi, niin saamme funktioille saman kaltaisen algebran kuin matriiseillekin. Nolla-alkoita vastaa nollafunktio $f(x) = 0$ kaikilla x, ykkösalkiota vastaa **identiteettifunktio** $\mathrm{Id}(x) = x$ kaikilla x ja käänteisalkioita vastaa käänteisfunktio f^{-1}.

4.9 Huomautus
Funktiotulolle pätee valitettavasti yleisesti $f \circ g \neq g \circ f$, mutta onneksi

(i) $(f \circ g) \circ h = f \circ (g \circ h)$,
(ii) $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

4.10 Harjoitustehtävä
Perustele Huomautus 4.9.

Matriisialgebra lineaaristen funktioiden algebrana

Funktio $\mathbf{f}$ on **lineaarinen**, jos

$$\mathbf{f}(t\mathbf{x} + s\mathbf{y}) = t\,\mathbf{f}(\mathbf{x}) + s\,\mathbf{f}(\mathbf{y}) \tag{4.11}$$

kaikilla vektoterilla $\mathbf{x}$ ja $\mathbf{y}$ sekä kaikilla skalaareilla t ja s.

4.12 Huomautus
Jos f ja g ovat lineaarisia, niin silloin myös $f \circ g$, $g \circ f$, f^{-1} ja g^{-1} (silloin kun ne ovat olemassa) ovat lineaarisia. Jos taas olisimme määritelleet funktiotulon pisteittäin, eli $(fg)(x) = f(x)g(x)$, niin fg ei olisi yleensä lineaarinen vaikka f ja g olisivatkin. Lisäksi on epäselvää mitä tulo $f(x)g(x)$ tarkoittaa vektorien tapauksessa. Näistä syistä emme määritelleet funktiotuloa pisteittäin.

Olkoon $f : \mathbb{R} \to \mathbb{R}$ linaarinen. Silloin määritelmästä (4.11) seuraa, että

$$f(x) = f(1)x.$$

Erityisesti näemme tästä, että kaikki lineaarikuvaukset $f : \mathbb{R} \to \mathbb{R}$ ovat muotoa

$$f(x) = f_c(x) = c \cdot x$$

jollekin luvulle c.

Olkoon $f : \mathbb{R}^2 \to \mathbb{R}$ lineaarinen. Käyttämällä koordinaattivektoriesitystä $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ määritelmässä (4.11) näemme, että

$$\begin{aligned} f(\mathbf{x}) &= f(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \\ &= x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2). \end{aligned}$$

Merkitsemällä $\mathbf{c} = [f(\mathbf{e}_1)\ \ f(\mathbf{e}_2)]^\top$ näemme että kaikki lineaarikuvaukset $f : \mathbb{R}^2 \to \mathbb{R}$ ovat pistetuloja

$$f(\mathbf{x}) = f_{\mathbf{c}}(\mathbf{x}) = \mathbf{c} \cdot \mathbf{x} = \mathbf{c}^\top \mathbf{x}$$

jollekin vektorille $\mathbf{c} \in \mathbb{R}^2$.

Jos $\mathbf{f} = [f_1\ \ f_2]^\top : \mathbb{R}^2 \to \mathbb{R}^2$ on vektoriarvoinen funktio, niin se on lineaarinen jos ja vain jos sen komponenttikuvaukset ovat molemmat lineaarisia. Tämän näkee koordinaattivektoriesitystä käyttävästä pyörittelystä (se luetaan sekä ylhäältä alas että alhaalta ylös)

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \mathbf{f}(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \\ &= x_1\mathbf{f}(\mathbf{e}_1) + x_2\mathbf{f}(\mathbf{e}_2) \\ &= x_1\begin{bmatrix} f_1(\mathbf{e}_1) \\ f_2(\mathbf{e}_1) \end{bmatrix} + x_2\begin{bmatrix} f_1(\mathbf{e}_2) \\ f_2(\mathbf{e}_2) \end{bmatrix} \\ &= \begin{bmatrix} x_1 f_1(\mathbf{e}_1) + x_2 f_1(\mathbf{e}_2) \\ x_1 f_2(\mathbf{e}_1) + x_2 f_2(\mathbf{e}_2) \end{bmatrix} \\ &= \begin{bmatrix} [f_1(\mathbf{e}_1)\ \ f_1(\mathbf{e}_2)]^\top \cdot \mathbf{x} \\ [f_2(\mathbf{e}_1)\ \ f_2(\mathbf{e}_2)]^\top \cdot \mathbf{x} \end{bmatrix} \\ &= \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{bmatrix}. \end{aligned}$$

Samalla tavalla näemme, että funktio $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^2$ on lineaarinen jos ja vain jos se on muotoa

$$\mathbf{f}(x) = \mathbf{f}_{\mathbf{c}^\top}(x) = \mathbf{c}x = \begin{bmatrix} c_1 \cdot x \\ c_2 \cdot x \end{bmatrix},$$

jollekin vektorille $\mathbf{c} \in \mathbb{R}^2$. Nimittäin lineaarisuuden määritelmästä (4.11) seuraa, että

$$\begin{aligned} \mathbf{f}(x) &= x\mathbf{f}(1) \\ &= x\begin{bmatrix} f_1(1) \\ f_2(1) \end{bmatrix} \\ &= \begin{bmatrix} x f_1(1) \\ x f_2(1) \end{bmatrix} \\ &= \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix}. \end{aligned}$$

Tästä näemme, että $\mathbf{c} = [f_1(1)\ \ f_2(1)]^\top$.

Palaamme lopuksi vielä lineaarisiin funktioihin $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$. Koordinaattivektoriesityksestä (katso edellä)

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} [f_1(\mathbf{e}_1)\ \ f_1(\mathbf{e}_2)]^\top \cdot \mathbf{x} \\ [f_2(\mathbf{e}_1)\ \ f_2(\mathbf{e}_2)]^\top \cdot \mathbf{x} \end{bmatrix} \tag{4.13}$$

päättelemme, että neljä lukua, $f_1(\mathbf{e}_1)$, $f_1(\mathbf{e}_2)$, $f_2(\mathbf{e}_1)$ ja $f_2(\mathbf{e}_2)$ määräävät täysin lineaarikuvauksen $\mathbf{f}$. Huomaamme sitten, että vektori

$$\mathbf{f}(\mathbf{e}_1) = \begin{bmatrix} f_1(\mathbf{e}_1) \\ f_2(\mathbf{e}_1) \end{bmatrix}$$

kertoo, mihin koordinaattivektori $\mathbf{e}_1$ kuvautuu. Samalla tavalla vektori

$$\mathbf{f}(\mathbf{e}_2) = \begin{bmatrix} f_1(\mathbf{e}_2) \\ f_2(\mathbf{e}_2) \end{bmatrix}$$

kertoo, mihin koordinaattivektori $\mathbf{e}_2$ kuvautuu. Jos taas tiedämme, mihin koordinaattivektorit kuvautuvat, tiedämme koko kuvauksen, sillä lineaarisuudesta (4.11) seuraa että

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \mathbf{f}(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \\ &= x_1\mathbf{f}(\mathbf{e}_1) + x_2\mathbf{f}(\mathbf{e}_2). \end{aligned}$$

Nyt jippo on määritellä matriisi

$$\mathbf{A} = \begin{bmatrix} \mathbf{f}(\mathbf{e}_1) & \mathbf{f}(\mathbf{e}_2) \end{bmatrix}.$$

Tällöin, suoraan matriisitulon määritelmästä seuraa, että

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}.$$

Edellä esitetystä seuraa kurssimme yksi keskeisimmistä huomioista:

4.14

Funktio $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ on lineaarinen **jos ja vain jos** se on muotoa

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_\mathbf{A}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

jollekin (2×2)-matriisille $\mathbf{A}$. Matriisin $\mathbf{A}$ sarakkeet kertovat mihin koordinaattivektorit $\mathbf{e}_1 = [1\ 0]^\top$ ja $\mathbf{e}_2 = [0\ 1]^\top$ kuvautuvat.

Funktio $\mathbf{f} = \mathbf{f}_\mathbf{A}$ on bijektio **jos ja vain jos** sitä vastaava matriisi $\mathbf{A}$ on kääntyvä ja käänteiskuvaus tulee kaavasta

$$\mathbf{f}_\mathbf{A}^{-1}(\mathbf{x}) = \mathbf{f}_{\mathbf{A}^{-1}}(\mathbf{x}) = \mathbf{A}^{-1}\mathbf{x}.$$

Jos $\mathbf{f}_\mathbf{A}$ ja $\mathbf{f}_\mathbf{B}$ ovat kaksi lineaarikuvausta tasojen välillä, niin niiden yhdistetyt kuvaukset on lineaarikuvauksia, jotka määräytyvät matriisitulon avulla kaavoilla

$$\begin{aligned} (\mathbf{f}_\mathbf{A} \circ \mathbf{f}_\mathbf{B})(\mathbf{x}) &= \mathbf{A}\mathbf{B}\mathbf{x}, \\ (\mathbf{f}_\mathbf{B} \circ \mathbf{f}_\mathbf{A})(\mathbf{x}) &= \mathbf{B}\mathbf{A}\mathbf{x} \end{aligned}$$

4.15 Harjoitustehtävä
Olkoot lineaariset funktiot $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ ja $\mathbf{g}\colon \mathbb{R}^2 \to \mathbb{R}^2$ annettu kaavoilla

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{x} \quad \text{ja} \quad \mathbf{g}(\mathbf{x}) = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}\mathbf{x}$$

(i) Laske yhdistetyt kuvaukset $\mathbf{f}\circ\mathbf{g}$ ja $\mathbf{g}\circ\mathbf{f}$. Toisin sanoen laske ko. lineaarikuvauksia vastaavat matriisit.
(ii) Laske käänteiskuvaukset $(\mathbf{f}\circ\mathbf{g})^{-1}$ ja $(\mathbf{g}\circ\mathbf{f})^{-1}$. Toisin sanoen laske ko. lineaarikuvauksia vastaavat matriisit.

4.16 Harjoitustehtävä
Olkoot lineaariset funktiot $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ ja $\mathbf{g}\colon \mathbb{R}^2 \to \mathbb{R}^2$ annettu kaavoilla

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & a \end{bmatrix}\mathbf{x} \quad \text{ja} \quad \mathbf{g}(\mathbf{x}) = \begin{bmatrix} 2 & -1 \\ b & 1 \end{bmatrix}\mathbf{x},$$

missä a ja b ovat mielivaltaisia parametreja.

(i) Millä parametrien a ja b arvoilla $\mathbf{f}$ ja $\mathbf{g}$ ovat bijektioita?
(ii) Laske käänteiskuvaukset $(\mathbf{f}\circ\mathbf{g})^{-1}$ ja $(\mathbf{g}\circ\mathbf{f})^{-1}$, silloin kun ne ovat olemassa. Laskeminen tarkoittaa ko. lineaarikuvauksia vastaavien matriisien laskemista.

Esitämme lopuksi muutamia matriiseja eli lineaarikuvauksia, joilla on selkeä tulkinta. Huomautamme, että käytännössä lähes kaikki matriisit saadaan yhdistelemällä alla olevia esimerkkimatriiseja matriisitulon avulla.

4.17 Esimerkki
(i) Matriisi

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

on **venytys**, joka venyttää x_1-akselia 2:n verran ja x_2-akselia 3:n verran.

(ii) Matriisi

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

on **peilaus** x_2-akselin suhteen ja matriisi

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

on **peilaus** x_1-akselin suhteen.

(iii) Matriisi

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

on tason **projektio** x_1-akselille ja matriisi

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

on tason **projektio** x_2-akselille.

(iv) Matriisi

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

on tason θ-radiaanin **kierto** vastapäivään ja matriisi

$$\begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

on tason θ-radiaanin **kierto** myötäpäivään.

Matriisit ja lineaarikuvaukset korkeammissa avaruuksissa*

Olkooon $\mathbb{R}^n$ niin sanottu n-ulotteinen euklidinen avaruus, eli vektoreiden $\mathbf{x} = [x_1\ x_2\ \cdots\ x_n]^\top$ joukko.

Kaikki edellä esitetty laajenee luonnollisella ja ilmeisellä tavalla lineaarikuvauksille $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$. Lineaarikuvausta $\mathbf{f}$ vastaa $(n \times m)$-matriisi $\mathbf{A}$, jonka sarakkeet kertovat mihin koordinaattivektorit $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ kuvautuvat. Yhdistetty kuvaus saadaan matriisitulolla ja käänteiskuvaus käänteismatriisilla. Erityisesti ainoastaan lineaarikuvauksilla avaruudelta itselleen voi olla käänteiskuvaus. Lineaarikuvaus on bijektio täsmälleen silloin, kun sitä vastaavan matriisin determinantti ei ole nolla.

Luku 5

Yhtälöpari matriisein

Ratkaisu matriisimerkinnöin

Palaamme yhtälöparin

$$\left\{\begin{array}{rcrcl} A_{11}x_1 & + & A_{12}x_2 & = & b_1 \\ A_{21}x_1 & + & A_{22}x_2 & = & b_2 \end{array}\right. \tag{5.1}$$

ratkaisemiseen. Toki osaamme ratkaista yhtöparin (5.1) jo koulutiedoilla. Siten olemme kiertäneet ympyrän. Toivottavasti olemme kuitenkin oppineet jotakin kierroksen aikana.

Matriisimerkinnöin voimme kirjoittaa yhtälöparin (5.1) kompaktissa muodossa matriisiyhtälönä

$$\mathbf{A}\mathbf{x} \;=\; \mathbf{b}, \tag{5.2}$$

missä

$$\mathbf{A} \;=\; \begin{bmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{bmatrix}, \quad \mathbf{x} \;=\; \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{ja} \quad \mathbf{b} \;=\; \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Jos matriisi $\mathbf{A}$ on kääntyvä, niin yhtälön (5.2) ratkaisu on

$$\mathbf{x} \;=\; \mathbf{A}^{-1}\mathbf{b} \tag{5.3}$$

Jos kuitenkin tarkoitus on ainoastaan ratkaista $\mathbf{x}$ yhtälöstä (5.2), niin kaavaa (5.3) ei yleensä kannata käyttää, sillä:

(i) Käänteismatriisin $\mathbf{A}^{-1}$ laskeminen on raskasta.
(ii) Yhtälöllä (5.2) saattaa olla ratkaisu, vaikka matriisi $\mathbf{A}$ ei olisikaan kääntyvä.

Käytännössä yhtälö (5.2), eli yhtälöpari (5.1), kannattaa ratkaista muuttamalla laajennettu lohkomatriisi $[\mathbf{A}\ \mathbf{b}]$ alkeisrivioperaatioilla laajennetuksi lohkomatriisiksi $[\mathbf{I}\ \mathbf{b}']$. Tällöin ratkaisu on $\mathbf{x} = \mathbf{b}'$.

5.4 Esimerkki
Ratkaisemme yhtälöparin

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

Tätä yhtälöparia vastaava laajennettu lohkomatriisi on

$$[\mathbf{A}\ \mathbf{b}] = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 1 & 4 \end{bmatrix}.$$

Koska

$$\det(\mathbf{A}) = 1 \cdot 1 - (-1) \cdot 4 = 5 \neq 0,$$

niin yhtälöparillamme on yksikäsitteinen ratkaisu.

Vähentämällä ensimmäinen rivi 4 kertaa toisesta rivistä saamme

$$\begin{bmatrix} 1 & -1 & 2 \\ 4 & 1 & 4 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1-4\cdot(-1) & 4-4\cdot 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & -4 \end{bmatrix}.$$

Lisäämällä toinen rivi $1/5$ kertaa ensimmäiseen riviin saamme

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

Lopuksi kertomalla toinen yhtälö luvulla $1/5$ saamme lopullisen muodon

$$\begin{bmatrix} 1 & 0 & 6/5 \\ 0 & 5 & -4 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 6/5 \\ 0 & 1 & -4/5 \end{bmatrix}.$$

Tästä luemme, että ratkaisu on

$$\begin{cases} x_1 = 6/5 \\ x_2 = -4/5 \end{cases}.$$

5.5 Harjoitustehtävä
Ratkaise yhtälöparit

(i)

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

(ii)

$$\begin{cases} x_1 - x_2 = 10 \\ 4x_1 + x_2 = 20 \end{cases}.$$

5.6 Esimerkki
Ratkaisemme yhtälöparin

$$\begin{cases} \;\; x_1 & - & x_2 & = & b_1 \\ 4x_1 & + & x_2 & = & b_2 \end{cases},$$

missä $\mathbf{b} = [b_1\ b_2]^\top$ on mielivaltainen parametri.

Tätä yhtälöparia vastaava laajennettu lohkomatriisi on

$$[\mathbf{A}\ \mathbf{b}] \;=\; \begin{bmatrix} 1 & -1 & b_1 \\ 4 & 1 & b_2 \end{bmatrix}.$$

Koska

$$\det(\mathbf{A}) \;=\; 1 \cdot 1 - (-1) \cdot 4 \;=\; 5 \;\neq\; 0,$$

niin yhtälöparillamme on yksikäsitteinen ratkaisu.

Vähentämällä ensimmäinen rivi 4 kertaa toisesta rivistä saamme

$$\begin{bmatrix} 1 & -1 & b_1 \\ 4 & 1 & b_2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & -1 & b_1 \\ 0 & 1-4\cdot(-1) & b_2 - 4\cdot b_1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & b_1 \\ 0 & 5 & b_2 - 4b_1 \end{bmatrix}.$$

Lisäämällä toinen rivi $1/5$ kertaa ensimmäiseen riviin saamme

$$\begin{bmatrix} 1 & -1 & b_1 \\ 0 & 5 & b_2 - 4b_1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & b_1/5 + b_2/5 \\ 0 & 5 & -4b_1 + b_2 \end{bmatrix}.$$

Lopuksi kertomalla toinen rivi luvulla $1/5$ saamme lopullisen muodon

$$\begin{bmatrix} 1 & 0 & b_1/5 + b_2/5 \\ 0 & 5 & -4b_1 + b_2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & (1/5)b_1 + (1/5)b_2 \\ 0 & 1 & -(4/5)b_1 + (1/5)b_2 \end{bmatrix}.$$

Tästä luemme, että ratkaisu on

$$\begin{cases} x_1 & = & (1/5)b_1 + (1/5)b_2 \\ x_2 & = & -(4/5)b_1 + (1/5)b_2 \end{cases}.$$

5.7 Huomautus
Esimerkissä 5.6 tulimme itse asiassa kääntäneeksi matriisin

$$\mathbf{A} \;=\; \begin{bmatrix} 1 & -1 \\ 4 & 1 \end{bmatrix}.$$

Nimittäin $\mathbf{b}$:n yleistä muotoa olevasta ratkaisusta

$$\begin{cases} x_1 & = & (1/5)b_1 + (1/5)b_2 \\ x_2 & = & -(4/5)b_1 + (1/5)b_2 \end{cases}.$$

näemme että

$$\mathbf{A}^{-1} = \begin{bmatrix} 1/5 & 1/5 \\ -4/5 & 1/5 \end{bmatrix}.$$

5.8 **Harjoitustehtävä**
Ratkaise yhtälöpari

$$\begin{cases} x_1 & - & 3x_2 & = & b_1 \\ 4x_1 & + & 2x_2 & = & b_2 \end{cases},$$

kun

(i) $[b_1 \ \ b_2]^\top = [-1 \ \ 2]^\top$,
(ii) $[b_1 \ \ b_2]^\top = [-3 \ \ 0]^\top$.

Ratkaisujen lukumäärä ja determinantti

Edellisessä osiossa kaikki yhtälöt (eli yhtälöparit) $\mathbf{Ax} = \mathbf{b}$ olivat sellaisia, että $\det(\mathbf{A}) \neq 0$. Tämä takasi, että ratkaisu oli yksikäsitteinen: $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$, vaikka emme välttämättä laskeneetkaan käänteismatriisia $\mathbf{A}^{-1}$. Tarkastelemme nyt yhtälöpareja, joissa $\det(\mathbf{A}) = 0$.

Aloitamme esimerkeillä.

5.9 **Esimerkki**
Yhtälöllä $\mathbf{0x} = \mathbf{b}$ on ratkaisu täsmällälleen silloin, kun $\mathbf{b} = \mathbf{0}$. Ratkaisu on kaikkea muuta kuin yksikäsitteinen: jokainen $\mathbf{x}$ on ratkaisu.

5.10 **Huomautus**
Koska $\mathbf{0}$-matriisi ei ole kovin mielenkiintoinen, emme käsittele sitä enää, ja oletammekin aina, että matriisimme eivät ole nollamatriiseja.

5.11 **Esimerkki**
Olkoon

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

Tällöin $\mathbf{A}$ on tason **projektio** x_1-suoralle: $\mathbf{Ax} = [x_1 \;\; 0]^\top$. Siispä yhtälöllä $\mathbf{Ax} = \mathbf{b}$ on ratkaisu täsmälleen silloin, kun $b_2 = 0$. Tällöin ratkaisuja on äärettömästi. Nimittäin kaikki suoralla (b_1, α), $\alpha \in \mathbb{R}$, olevat pisteet ovat ratkaisuja.

5.12 **Esimerkki**
Haluamme ratkaista yhtälöparin

$$\begin{cases} 2x_1 & + & x_2 & = & 0 \\ 4x_1 & + & 2x_2 & = & 3 \end{cases}.$$

Tätä yhtälöparia vastaava matriisiyhtälö on

$$\mathbf{Ax} \;=\; \mathbf{b},$$

missä $\mathbf{b} = [0 \;\; 3]^\top$ ja

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

Koska

$$\det(\mathbf{A}) \;=\; 2 \cdot 2 - 1 \cdot 4 \;=\; 0,$$

niin ikävyyksiä voi olla tiedossa!

Yritämme joka tapauksessa ratkaista yhtälöparin. Laajennettu lohkomatriisi on

$$[\mathbf{A}\;\mathbf{b}] \;=\; \begin{bmatrix} 2 & 1 & 0 \\ 4 & 2 & 3 \end{bmatrix}.$$

Eliminoimme x_1-termin toisesta yhtälöstä vähentämällä ensimmäisen yhtälön 2 kertaa toisesta yhtälöstä. Saamme

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

Tämä tarkoittaa, että yhtälöparimme on ekvivalentti seuraavan yhtälöparin kanssa:

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

Tästä näemme, että ratkaisuja ei voi olla olemassa.

5.13 **Esimerkki**
Haluamme ratkaista yhtälöparin

$$\begin{cases} 2x_1 & + & x_2 & = & 10 \\ 4x_1 & + & 2x_2 & = & 20 \end{cases}.$$

Tätä yhtälöparia vastaava matriisiyhtälö on

$$\mathbf{A}\mathbf{x} \;=\; \mathbf{b},$$

missä $\mathbf{b} = [10 \;\; 20]^\top$ ja

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

Koska

$$\det(\mathbf{A}) \;=\; 2\cdot 2 - 1\cdot 4 \;=\; 0,$$

niin ikävyyksiä voi olla tiedossa!

Yritämme joka tapauksessa ratkaista yhtälöparin. Laajennettu lohkomatriisi on

$$[\mathbf{A}\;\mathbf{b}] \;=\; \begin{bmatrix} 2 & 1 & 10 \\ 4 & 2 & 20 \end{bmatrix}.$$

Eliminoimme x_1-termin toisesta yhtälöstä vähentämällä ensimmäisen yhtälön 2 kertaa toisesta yhtälöstä. Saamme

$$\begin{bmatrix} 2 & 1 & 10 \\ 4 & 2 & 20 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 2 & 1 & 10 \\ 0 & 2-2\cdot 1 & 20-2\cdot 10 \end{bmatrix} \;=\; \begin{bmatrix} 2 & 1 & 10 \\ 0 & 0 & 0 \end{bmatrix}.$$

Tämä tarkoittaa, että yhtälöparimme on ekvivalentti seuraavan yhtälöparin kanssa:

$$\begin{cases} 2x_1 & + & x_2 & = & 10 \\ 0x_1 & + & 0x_2 & = & 0 \end{cases}.$$

Tästä näemme, että ratkaisuja on äärettömästi. Nimittäin kaikki tason pisteet, jotka ovat muotoa $(x_1, x_2) = (x_1, 5-(1/2)x_1)$, $x_1 \in \mathbb{R}$, ovat ratkaisuja.

5.14 **Harjoitustehtävä**

Tarkastelemme yhtälöparia

$$\begin{cases} 2x_1 & + & ax_2 & = & 5 \\ 4x_1 & + & 2x_2 & = & 10 \end{cases},$$

missä a on vapaa parametri. Millä a:n arvoilla yhtälöparilla on yksikäsitteinen ratkaisu?

Yritäme nyt selvittää yleisen kuvan. Tarkastelemme yhtälöä $\mathbf{A}\mathbf{x} = \mathbf{b}$, missä $\det(\mathbf{A}) = 0$. Koska $\det(\mathbf{A}) = 0$, niin matriisin $\mathbf{A}$ sarakkeet tai yhtä hyvin rivit ovat lineaarisesti riippuvia. Käytämme rivitulkintaa. Matriisi $\mathbf{A}$ on siis muotoa

$$\mathbf{A} \;=\; [\mathbf{a}\;\;\lambda\mathbf{a}]^\top \;=\; \begin{bmatrix} a_1 & \lambda a_1 \\ a_2 & \lambda a_2 \end{bmatrix}^\top \;=\; \begin{bmatrix} a_1 & a_2 \\ \lambda a_1 & \lambda a_2 \end{bmatrix},$$

missä $\mathbf{a} = [a_1\ a_2]^\top$ on jokin vektori ja λ on jokin skalaari.

Yhtälöä $\mathbf{Ax} = \mathbf{b}$ vastaava laajennettu lohkomatriisi on siis muotoa

$$[\mathbf{A}\ \mathbf{b}] = \begin{bmatrix} a_1 & a_2 & b_1 \\ \lambda a_1 & \lambda a_2 & b_2 \end{bmatrix}.$$

Vähentämällä ensimmäinen rivi λ kertaa toisesta rivistä saamme ekvivalentin muodon

$$\begin{bmatrix} a_1 & a_2 & b_1 \\ \lambda a_1 & \lambda a_2 & b_2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} a_1 & a_2 & b_1 \\ 0 & 0 & b_2 - \lambda b_1 \end{bmatrix}.$$

Tästä näemme välittömästi, että ratkaisu on olemassa täsmälleen silloin, kun $\mathbf{b}$:n koordinaatit toteuttavat yhtälön

$$b_2 = \lambda b_1. \tag{5.15}$$

Tarkastelemme sitten ratkaisujen lukumäärää. Oletamme, että ratkaisuja on olemassa. Toisin sanoen kaava (5.15) pätee. Tämä tarkoittaa sitä, että tarkastelemme yhtälöä

$$\begin{bmatrix} a_1 & a_2 \\ \lambda a_1 & \lambda a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ \lambda b_1 \end{bmatrix}$$

Edellisen nojalla tiedämme, että tämä yhtälö on ekvivalentti seuraavan yhtälön kanssa:

$$\begin{bmatrix} a_1 & a_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ 0 \end{bmatrix}$$

Mutta tämä näennäinen yhtälöpari pelkistyy pelkäksi yhtälöksi

$$a_1 x_1 + a_2 x_2 = b_1. \tag{5.16}$$

Tästä näemme, että ratkaisujoukko on suora. Erityisesti ratkaisuja on äärettömän paljon.

5.17 Harjoitustehtävä

Tarkastelemme yhtälöparia

$$\begin{cases} 2x_1 + a x_2 = 5 \\ 4x_1 - x_2 = b \end{cases},$$

missä a ja b ovat vapaita parametreja. Etsi yhtälöparin **kaikki** ratkaisut.

5.18

Yhtälöparilla $\mathbf{Ax} = \mathbf{b}$ on yksikäsitteinen ratkaisu jokaisella $\mathbf{b}$ jos ja vain jos $\det(\mathbf{A}) \neq 0$. Ratkaisu on tällöin $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$, joskaan käänteismatriisia $\mathbf{A}^{-1}$ ei tarvitse välttämättä ratkaista eksplisiittisesti, vaan ratkaisu voidaan saada muuttamalla laajennettu lohkomatriisi $[\mathbf{A}\ \mathbf{b}]$ alkeisrivioperaatioilla laajennetuksi lohkomatriisiksi $[\mathbf{I}\ \mathbf{b}']$. Tällöin ratkaisu on $\mathbf{x} = \mathbf{b}'$.

Jos $\det(\mathbf{A}) = 0$, niin yhtälöparilla $\mathbf{Ax} = \mathbf{b}$ ei ole ratkaisua, jos $\mathbf{b}$ ei kuulu matriisin $\mathbf{A}$ ensimmäisen (tai yhtä hyvin toisen) sarakkeen virittämälle suoralle. Jos taas $\mathbf{b}$ kuuluu ko. suoralle, on ratkaisuja äärettömän monta.

Yhtälöryhmät ja matriisit*

Kaikki edellä esitetty yleistyy luonnollisella tavalla n:n muuttujan ja n:n yhtälön lineaarisille yhtälöryhmille $\mathbf{Ax} = \mathbf{b}$. Jos muuttujia ja yhtälöitä on eri määrä, voidaan lisätä nolla-yhtälöitä tai nollamuuttujia niin, että niitä on sen jälkeen sama määrä.

Jos $\det(\mathbf{A}) \neq 0$, on ratkaisu yksikäsitteinen $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Jos taas $\det(\mathbf{A}) = 0$, niin ratkaisu on olemmassa jos ja vain jos $\mathbf{b}$ kuuluu matriisin $\mathbf{A}$ sarakkeiden virittämään aliavaruuteen. Tällöin ratkaisuja on myös äärettömän paljon.

Käytännön kannalta keskeinen ongelma on, että matriisin determinantin laskeminen on viheliäistä. Yleensä kannattaakin erikseen "nähdä" onko se nolla vai ei tarkastelemalla matriisin sarakkeiden lineaarista riippuvuutta.

Ratkaisemme esimerkin vuoksi 4 muuttujan ja 4 yhtälön lineaarisen yhtälöryhmän

$$\begin{cases} 2x_1 + x_2 + x_4 = 2 \\ 4x_1 - x_2 + 2x_4 = 1 \\ 4x_1 + 2x_3 = 4 \\ -4x_3 + 2x_4 = 2 \end{cases}.$$

Nyt siis $\mathbf{x} = [x_1\ x_2\ x_3\ x_4]^\top$, $\mathbf{b} = [2\ 1\ 4\ 2]^\top$ ja

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

Voidaan itse asiassa laskea, että $\det(\mathbf{A}) = 24$, mutta tämä on viheliäistä. Sen sijaan yritämme ratkaista yhtälöryhmän laajennetuilla kohkomatriiseilla. Tarkoitus on siis muuttaa alkeisrivioperaatioilla $[\mathbf{A}\ \mathbf{b}]$ muotoon $[\mathbf{I}\ \mathbf{b}']$, missä $\mathbf{I}$ on (4×4)-identiteettimatriisi

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

Aloitamme väännön. Eliminoimme aluksi muuttujan x_1 riveiltä 2 ja 3 (rivillä 4 se ei esiinny):

$$[\mathbf{A}\ \mathbf{b}] = \begin{bmatrix} 2 & 1 & 0 & 1 & 2 \\ 4 & -1 & 0 & 2 & 1 \\ 4 & 0 & 2 & 0 & 4 \\ 0 & 0 & -4 & 2 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 2 & 1 & 0 & 1 & 2 \\ 0 & -3 & 0 & 0 & -3 \\ 4 & 0 & 2 & 0 & 4 \\ 0 & 0 & -4 & 2 & 2 \end{bmatrix}$$

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

Eliminoimme sitten muuttujan x_2 riveiltä 1 ja 3 (rivillä 4 se ei esiinny):

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

Eliminoimme sitten muuttujan x_3 riviltä 4 (riveillä 1 ja 2 se ei esiinny):

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

Eliminoimme sitten muuttujan x_4 riveiltä 1 ja 3 (rivillä 2 se ei esiinny):

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

Lopuksi kerromme jokaisen rivin sitä vastaavan muuttujan kertoimen käänteisluvulla. Näin saamme ratkaistun laajennetun lohkomatriisin

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 4 \\ 0 & -3 & 0 & 0 & -3 \\ 0 & 0 & 2 & 0 & -4 \\ 0 & 0 & 0 & -2 & 6 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix} = [\mathbf{I}\ \mathbf{b}'].$$

Tästä luemme ratkaisun

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ -2 \\ -3 \end{bmatrix}.$$

Yhtälöryhmien ratkaisu Cramerin säännöllä*

Esitämme vielä yhden joidenkin suosiman tavan ratkaista yhtälöryhmiä

$$\mathbf{Ax} = \mathbf{b},$$

kun $\mathbf{A}$ on kääntyvä $(n \times n)$-matriisi. Tämä tapa kantaa nimeä **Cramerin sääntö**. Kaikessa yksinkertaisuudessaan sen mukaan ratkaisu on $\mathbf{x} = [x_1\ x_2\ \cdots\ x_n]^\top$, missä

$$x_i = \frac{\det(\mathbf{A}_{i\to\mathbf{b}})}{\det(\mathbf{A})} \tag{5.19}$$

ja matriisi $\mathbf{A}_{i\to\mathbf{b}}$ saadaan matriisista $\mathbf{A}$ korvaamalla sen i. sarake pystyvektorilla $\mathbf{b}$.

Kaava (5.19) on toki elegantti ja helppo muistaakin, mutta determinanttien laskeminen on työlästä. Jos $\mathbf{A}$ on (3×3)-matriisi, niin determinantin laskeminen on vielä kohtalaisen inhimillistä. Nimittäin

$$\det(\mathbf{A}) = A_{11}\det(\mathbf{A}_{-1}) - A_{12}\det(\mathbf{A}_{-2}) + A_{13}\det(\mathbf{A}_{-3}),$$

missä $\mathbf{A}_{-j}$ on (2×2)-alimatriisi, joka on saatu matriisista $\mathbf{A}$ poistamalla sen ensimmäinen rivi ja j. sarake.

Emme perustele Cramerin sääntöä tässä. Tyydymme laskemaan yhden (3×3)-esimerkin.

5.20 Esimerkki
Ratkaisemme Cramerin säännöllä yhtälöryhmän

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 5 & 4 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 0 \end{bmatrix}.$$

Aloitamme laskemalla $\mathbf{A}$:n determinantin:

$$\begin{aligned} \det(\mathbf{A}) &= 1\det(\mathbf{A}_{-1}) - 0\det(\mathbf{A}_{-2}) + 2\det(\mathbf{A}_{-3}) \\ &= 5\cdot 3 - 4\cdot 1 \ - \ 0 \ + \ 2\cdot(0\cdot 1 - 5\cdot 2) \\ &= 11 - 20 \\ &= -9. \end{aligned}$$

Laskemme sitten determinantit $\mathbf{A}_{i\to\mathbf{b}}$ kaikille $i = 1,2,3$:

$$\begin{aligned} \det(\mathbf{A}_{1\to\mathbf{b}}) &= \det\left(\begin{bmatrix} 6 & 0 & 2 \\ 7 & 5 & 4 \\ 0 & 1 & 3 \end{bmatrix}\right) \\ &= 6\det\left(\begin{bmatrix} 5 & 4 \\ 1 & 3 \end{bmatrix}\right) - 0\det\left(\begin{bmatrix} 7 & 4 \\ 0 & 3 \end{bmatrix}\right) + 2\det\left(\begin{bmatrix} 7 & 5 \\ 0 & 1 \end{bmatrix}\right) \\ &= 6\cdot(5\cdot 3 - 4\cdot 1) \ - \ 0 \ + \ 2\cdot(7\cdot 1 - 5\cdot 0) \\ &= 80. \end{aligned}$$

$$
\begin{aligned}
\det(\mathbf{A}_{2\to\mathbf{b}}) &= \det\left(\begin{bmatrix} 1 & 6 & 2 \\ 0 & 7 & 4 \\ 2 & 0 & 3 \end{bmatrix}\right) \\
&= 1\det\left(\begin{bmatrix} 7 & 4 \\ 0 & 3 \end{bmatrix}\right) - 6\det\left(\begin{bmatrix} 0 & 4 \\ 2 & 3 \end{bmatrix}\right) + 2\det\left(\begin{bmatrix} 0 & 7 \\ 2 & 0 \end{bmatrix}\right) \\
&= 7\cdot 3 - 4\cdot 0 \;-\; 6\cdot(0\cdot 3 - 4\cdot 2) \;+\; 2\cdot(0\cdot 0 - 7\cdot 2) \\
&= 41.
\end{aligned}
$$

$$
\begin{aligned}
\det(\mathbf{A}_{3\to\mathbf{b}}) &= \det\left(\begin{bmatrix} 1 & 0 & 6 \\ 0 & 5 & 7 \\ 2 & 1 & 0 \end{bmatrix}\right) \\
&= 1\det\left(\begin{bmatrix} 5 & 7 \\ 1 & 0 \end{bmatrix}\right) - 0\det\left(\begin{bmatrix} 0 & 7 \\ 2 & 0 \end{bmatrix}\right) + 6\det\left(\begin{bmatrix} 0 & 5 \\ 2 & 1 \end{bmatrix}\right) \\
&= 5\cdot 0 - 7\cdot 1 \;-\; 0 \;+\; 6\cdot(0\cdot 1 - 5\cdot 2) \\
&= -67.
\end{aligned}
$$

Olemme siis saaneet ratkaisun

$$
\begin{aligned}
x_1 &= 80/(-9) &&= -8.8889, \\
x_2 &= 41/(-9) &&= -4.5556, \\
x_3 &= -67/(-9) &&= 7.4444.
\end{aligned}
$$

Lopuksi annamme vauhdikkaan todistuksen Cramerin säännölle. Yksityiskohtien pohtimisen jätämme kiinnostuneen lukijan oman harrastuneisuuden varaan.

$$
\begin{aligned}
x_i &= \det\left(\mathbf{I}_{i\to\mathbf{x}}\right) \\
&= \det\left(\mathbf{A}^{-1}\mathbf{A}_{i\to\mathbf{b}}\right) \\
&= \det(\mathbf{A}^{-1})\det\left(\mathbf{A}_{i\to\mathbf{b}}\right) \\
&= \frac{\det\left(\mathbf{A}_{i\to\mathbf{b}}\right)}{\det(\mathbf{A})}.
\end{aligned}
$$

Luku 6

Kompleksitaso

Kompleksilukujen karteesinen muoto

Eräs tapa ymmärtää kompleksiluvut on tulkita ne tason pisteiksi. Käytämme tässä luvussa edellisistä luvuista poiketen tason pisteille merkintää $z = (x, y)$. Formaalisti kirjoitamme myös

$$z = x + y\mathrm{i}, \tag{6.1}$$

missä $\mathrm{i} = \sqrt{-1}$ on **imaginääriyksikkö**, eli yhtälön

$$\mathrm{i}^2 = -1$$

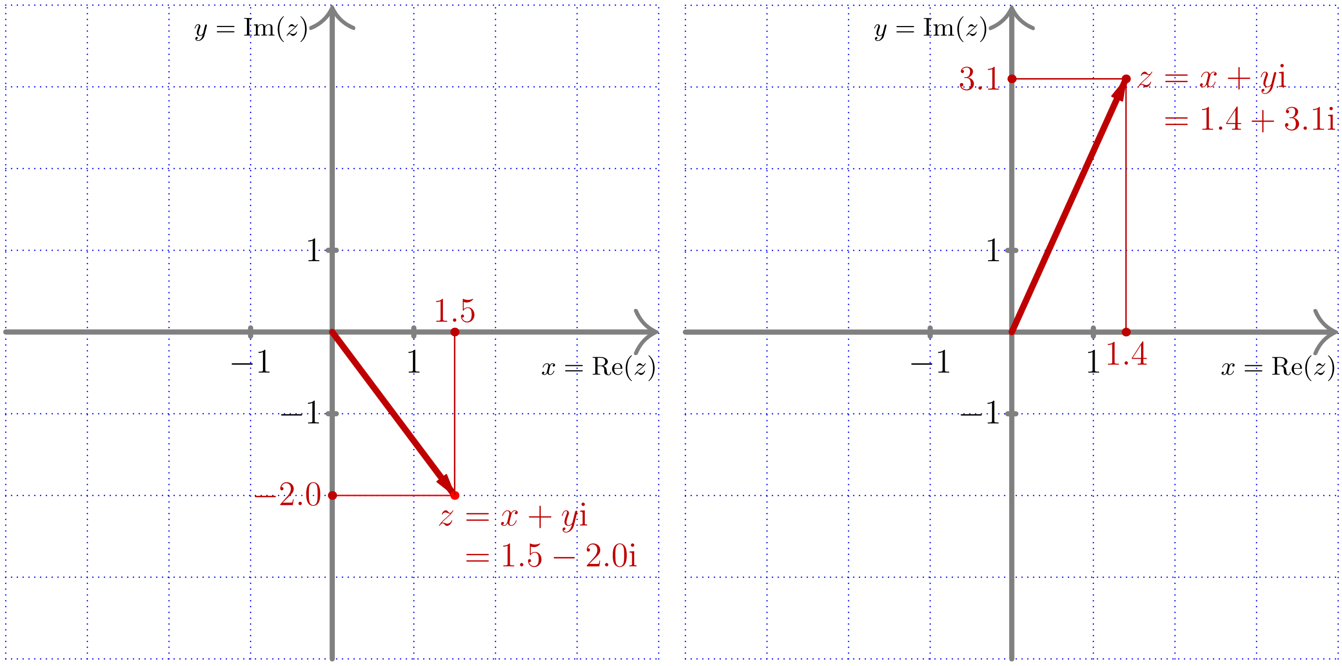
"positiivinen" ratkaisu.

6.2 Huomautus
Formaalisti yhtälöllä $\mathrm{i}^2 = -1$ on ainakin kaksi ratkaisua: $\pm\mathrm{i}$. Nimittäin, jos $\mathrm{i}^2 = -1$, niin myöskin $(-\mathrm{i})^2 = (-1)^2\mathrm{i}^2 = \mathrm{i}^2 = -1$.

Kaava (6.1) on kompleksiluvun **karteesinen muoto**. Myöhemmin opimme ns. polaari- eli napamuodon, joka liittyy yllättäen kompleksiseen eksponenttifunktioon.

Kompleksitaso koostuu pisteistä $z = x + y\mathrm{i}$, missä x ja y ovat reaalilukuja. Lukua x kutsutaan kompleksiluvun z reaaliosaksi ja lukua y kutsutaan kompleksiluvun z imaginääriosaksi. Joskus käytetään myös merkintojä $x = \Re(z) = \mathrm{Re}(z)$ ja $y = \Im(z) = \mathrm{Im}(z)$. Siispä voimme kirjoittaa karteesisen muodon (6.1) myös muodoissa

$$\begin{aligned} z &= \Re(z) + \Im(z)\mathrm{i} \\ &= \mathrm{Re}(z) + \mathrm{Im}(z)\mathrm{i}. \end{aligned}$$



Kompleksiluku $z = x+yi = 1.5-2.0\mathrm{i}$ (vasen kuva) ja kompleksiluku $z = x+yi = 1.4+3.1\mathrm{i}$ (oikea kuva).

Kompleksilukujen $z_1 = x_1 + y_1\mathrm{i}$ ja $z_2 = x_2 + y_2\mathrm{i}$ yhteenlasku toimii täsmälleen samalla tavalla kuin vastaavien tason pisteiden, eli vektorien, yhteenlasku:

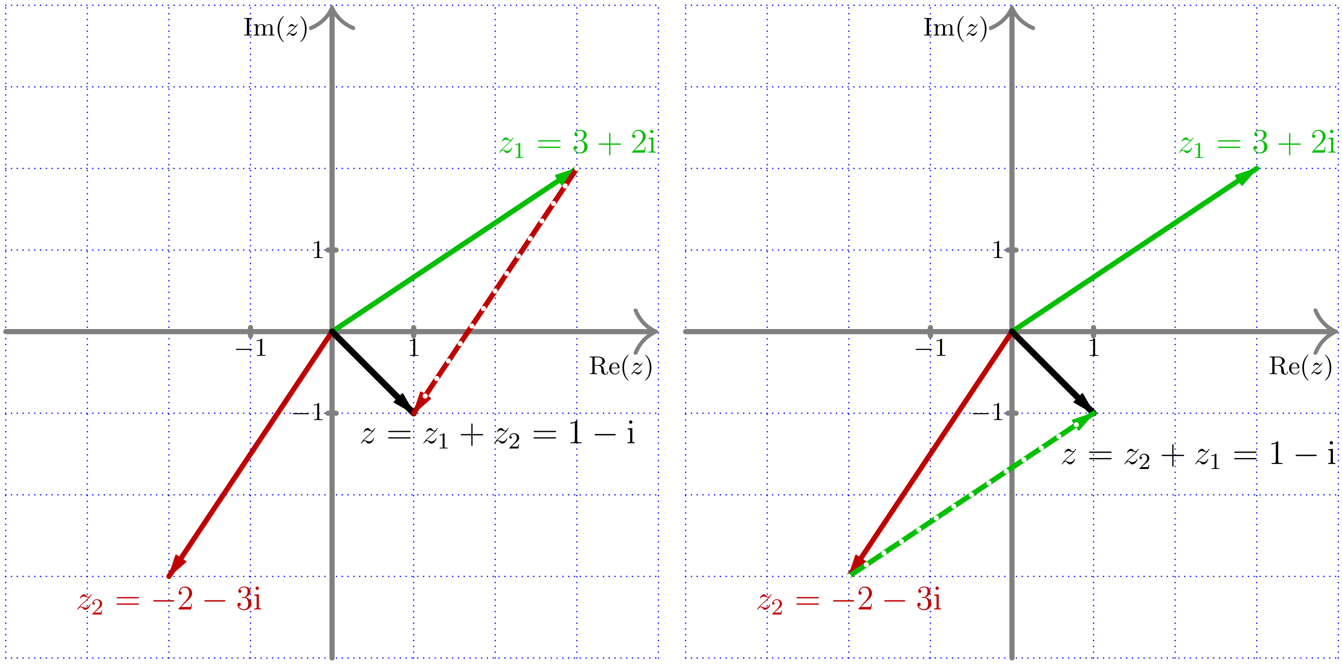
$$\begin{aligned} z_1 + z_2 &= (x_1 + x_2\,,\ y_1 + y_2) \\ &= (x_1 + x_2) + (y_1 + y_2)\mathrm{i}. \end{aligned}$$

6.3 Esimerkki
Olkoon

$$\begin{aligned} z_1 &= 3 + 2\mathrm{i}, \\ z_2 &= -2 - 3\mathrm{i}. \end{aligned}$$

Tällöin

$$\begin{aligned} z_1 + z_2 &= 3 + 2\mathrm{i} + (-2) - 4\mathrm{i} \\ &= 3 - 2 + (2 - 3)\mathrm{i} \\ &= 1 - \mathrm{i}. \end{aligned}$$



Esimerkin 6.3 kompleksilukujen yhteenlaskua karteesisessa muodossa. Sama summa on laskettu (tai pikemminkin piirretty) molemmin päin: $z = z_1 + z_2$ (vasen puoli) ja $z = z_2 + z_1$ (oikea puoli).

Kompleksitason vektorin eli kompleksiluvun $z = (x, y)$ normille käytetään kompleksilaskennassa nimeä **moduuli** ja sitä merkitään (ehkä hieman harhaanjohtavasti) tavallisilla yksinkertaisilla itseisarvomerkeillä:

$$|z| = \sqrt{x^2 + y^2}.$$

6.4 Esimerkki

(i) Olkoon

$$z = -3 + 2\mathrm{i}.$$

Tällöin

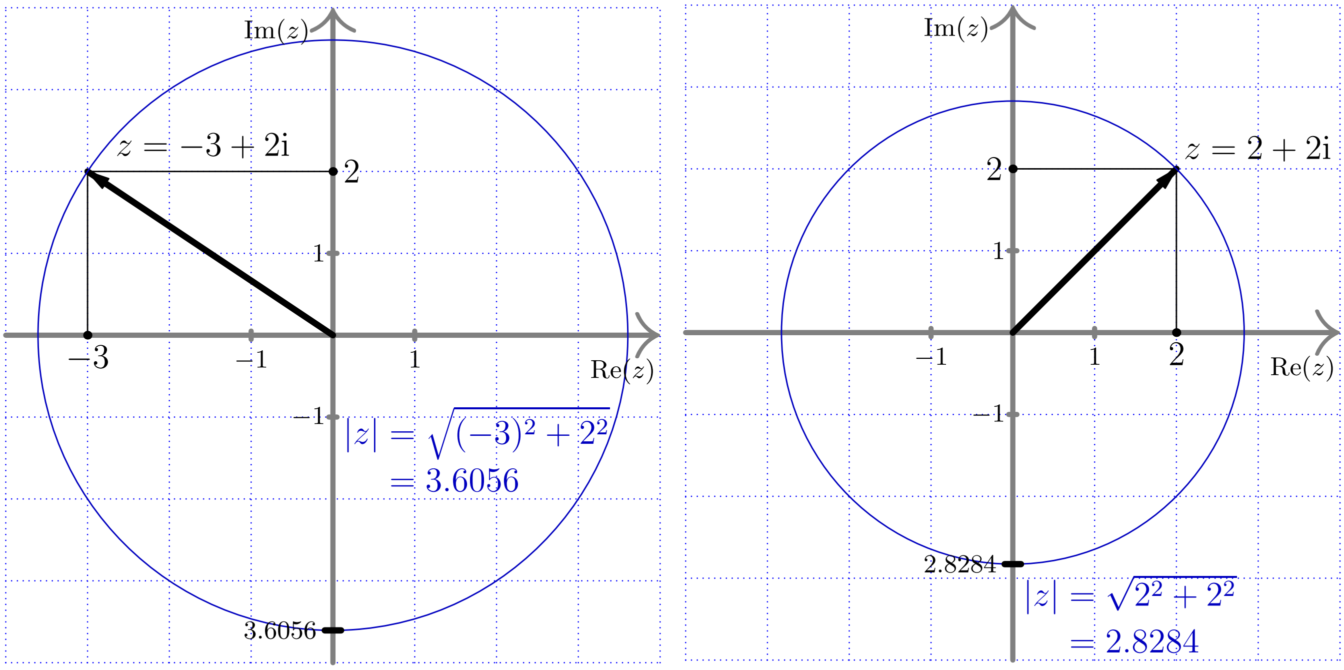
$$\begin{aligned} |z| &= \sqrt{(-3)^2 + 2^2} \\ &= \sqrt{9 + 4} \\ &= \sqrt{13} \\ &= 3.6056. \end{aligned}$$

Olkoon

$$z = 2 + 2\mathrm{i}.$$

Tällöin

$$\begin{aligned} |z| &= \sqrt{2^2 + 2^2} \\ &= \sqrt{8} \\ &= 2.8284. \end{aligned}$$



Esimerkin 6.4 kompleksiluvut ja niiden moduulit. Vasemmalla $z = -3 + 2\mathrm{i}$ ja oikealla $z = 2 + 2\mathrm{i}$.

Reaalitason vektoreita ei voi kertoa keskenään. Pistetulo eli sisätulo ei ole varsinaista kertomista, sillä siinä kaksi vektoria kuvautuu reaaliluvuksi. Kompleksilukuja sen sijaan voidaan kertoa keskenään niin että tulokseksi saadaan kompleksiluku. **Kertolasku** määräytyy formaalisti tavallisen kerto- ja yhteenlaskun osittelulaista sekä kaavasta $\mathrm{i}^2 = -1$:

$$
\begin{aligned}
z_1 z_2 &= (x_1 + y_1\mathrm{i})(x_2 + y_2\mathrm{i}) \\
&= x_1x_2 + x_1y_2\mathrm{i} + y_1x_2\mathrm{i} + y_1y_2\mathrm{i}^2 \\
&= (x_1x_2 - y_1y_2) + (x_1y_2 + y_1x_2)\mathrm{i}.
\end{aligned}
$$

6.5 Esimerkki

(i) Olkoon

$$
\begin{aligned}
z_1 &= -1.5 + 0.5\mathrm{i}, \\
z_2 &= 1 + 2\mathrm{i}.
\end{aligned}
$$

Tällöin

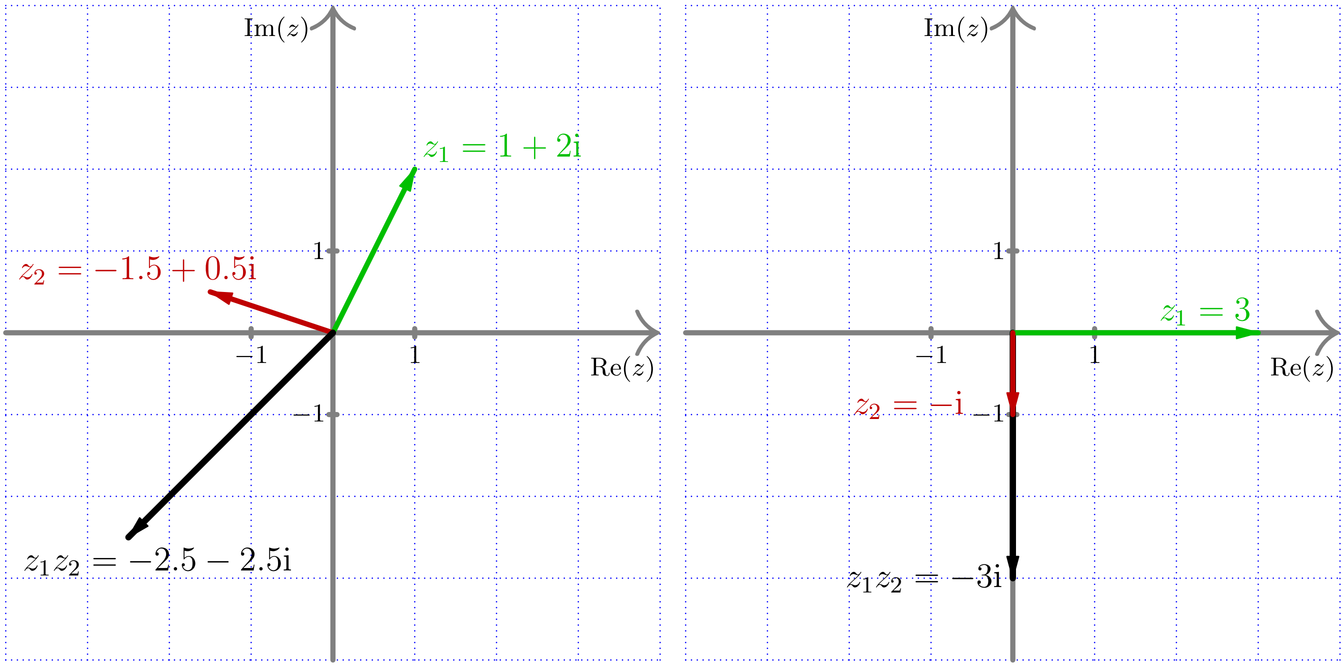
$$
\begin{aligned}
z_1 z_2 &= (-1.5 + 0.5\mathrm{i})(1 + 2\mathrm{i}) \\
&= -1.5 \times 1 - 1.5 \times 2\mathrm{i} + 0.5\mathrm{i} \times 1 + 0.5\mathrm{i} \times 2\mathrm{i} \\
&= -1.5 - 3\mathrm{i} + 0.5\mathrm{i} - 1 \\
&= -2.5 - 2.5\mathrm{i}
\end{aligned}
$$

(ii) Olkoon

$$\begin{aligned} z_1 &= 3, \\ z_2 &= -\mathrm{i}. \end{aligned}$$

Tällöin

$$\begin{aligned} z_1 z_2 &= 3 \times (-\mathrm{i}) \\ &= -3\mathrm{i}. \end{aligned}$$



Esimerkin 6.5 kompleksiluvut ja niiden kertolaskut karteesisessa koordinaatistossa. Vasemmalla (i) ja oikealla (ii). Kuvasta näkee (tai pikemmikin ei näe), että kmpleksilukujen tulo on vaikeammin hahmotettava operaatio kuin kompleksilukujen summa.

Kompleksiluvun $z = x + y\mathrm{i}$ **liittoluku** eli **konjugaatti** on $z^* = x - y\mathrm{i}$. Liitoluvun ottaminen tarkoittaa siis peilausta reaaliakselin (eli x-akselin) suhteen. Erityisesti kompleksiluku z on reaalinen, eli $z = x + 0\mathrm{i}$, jos ja vain jos $z = z^*$.

6.6 Huomautus
Liittoluvulle z^* käytetään joskus myös merkintää $\overline{z}$.

6.7 Esimerkki
Olkoon

$$z_1 = -3 + 2\mathrm{i}.$$

Tällöin

$$z_1^* = -3 - 2\mathrm{i}$$

Olkoon sitten

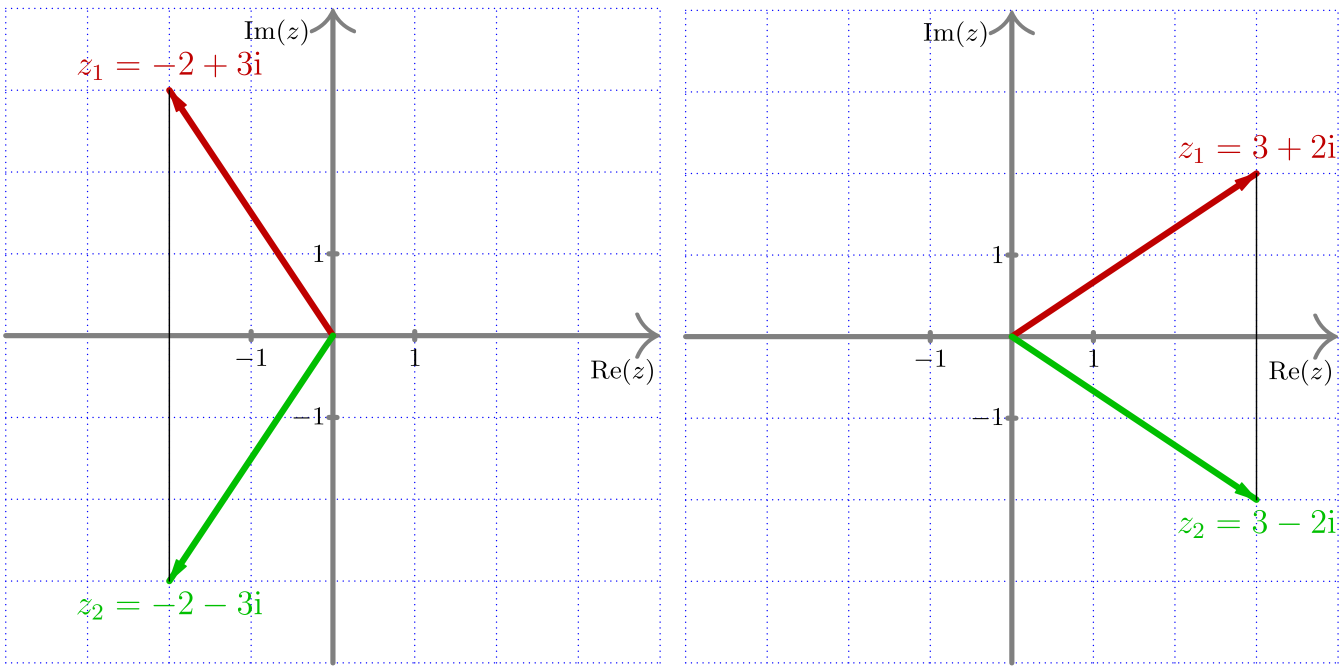
$$\begin{aligned} z_2 &= z_1^* \\ &= -3 - 2\mathrm{i}. \end{aligned}$$

Tällöin

$$\begin{aligned} z_2^* &= -3 - (-2)\mathrm{i} \\ &= -3 + 2\mathrm{i} \\ &= z_1. \end{aligned}$$

Esimerkin 6.7 tavoin on helppo nähdä, että yleisesti mille tahansa kompeksiluvulle z pätee

$$(z^*)^* = z.$$



Kompleksilukujen liittolukuja: $z_2 = z_1^*$, ja sama toisin päin.

Kompleksiluvun moduuli, eli vastaavan kompleksitason vektorin pituus voidaan esittää liittoluvun avulla:

$$|z|^2 = zz^*. \tag{6.8}$$

Kaavan (6.8) näkee suoralla laskulla. Nimittäin

$$\begin{aligned} zz^* &= (x+y\mathrm{i})(x-y\mathrm{i}) \\ &= x^2 - xy\mathrm{i} + xy\mathrm{i} - y^2\mathrm{i}^2 \\ &= x^2 + y^2. \end{aligned}$$

Kompleksilukujen **jakolasku** voidaan nyt määritellä kikkailemalla liittolukujen avulla

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*} = \frac{z_1 z_2^*}{|z_2|^2}. \tag{6.9}$$

6.10 Harjoitustehtävä
Osoita että kikkaileva määritelmä (6.9) on järkevä. Toisin sanoen, jos z_2 ei ole kompleksitason 0-vektori, niin

$$z = \frac{z_1}{z_2}$$

on olemassa ja se on yhtälön

$$z_2 z = z_1$$

ratkaisu muuttujan z suhteen. Käytä perustelussasi ainoastaan edellä esitettyjä karteesisia muotoja ja määritelmiä.

Kompleksilukujen **potenssit** ja **juuret** määritellään samaan tapaan kuin reaalilukujenkin:

$$\begin{aligned} z^n &= z\,z^{n-1}, \\ z^0 &= 1, \end{aligned}$$

ja $z^{1/n}$ on yhtälön

$$z = w^n$$

"positiivinen" ratkaisu kompleksiluvun w suhteen. Käytännössä kompleksilukujen potensseja ja juuria ei kuitenkaan kannata laskea karteesista muotoa $z = x + y\mathrm{i}$ käyttäen (eikä oikeastaan tuloja tai osamääriäkään), vaan käyttämällä napaesitystä ja De Moivren kaavaa, joita käsittelemme seuraavassa osiossa.

6.11 Esimerkki
Olkoon

$$z = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\mathrm{i}$$

Tällöin

$$\begin{aligned} z^2 &= zz \\ &= \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\mathrm{i}\right)\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\mathrm{i}\right) \\ &= \frac{1}{2} + \frac{1}{2}\mathrm{i} + \frac{1}{2}\mathrm{i} + \frac{1}{2}\mathrm{i}^2 \\ &= \frac{1}{2} + \mathrm{i} - \frac{1}{2} \\ &= \mathrm{i}. \end{aligned}$$

6.12 Harjoitustehtävä
Olkoon z kuten esimerkissä 6.11:

$$z = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\mathrm{i}$$

(i) Laske z^n, kun $n = 1, 2, 3, 4, 5, 6, 7, 8$.
(ii) Ratkaise (w:n suhteen) yhtälö $z = w^n$, kun $n = 1, 2, 3, 4, 5, 6, 7, 8$.

6.13 Esimerkki
Olkoon $z_1 = 1 - \mathrm{i}$ ja $z_2 = -2\mathrm{i}$. Laskemme

(i) $z_1 - z_2$,
(ii) $z_1 z_2^2$,
(iii) $z_1 z_2^*$,
(iv) z_1 / z_2.

(i) Tämä on suoraviivaista:

$$\begin{aligned} z_1 - z_2 &= 1 - \mathrm{i} - (-2\mathrm{i}) \\ &= 1 - \mathrm{i} + 2\mathrm{i} \\ &= 1 + \mathrm{i}. \end{aligned}$$

(ii) Laskemme ensin

$$\begin{aligned} z_2^2 &= (-2\mathrm{i})(-2\mathrm{i}) \\ &= 4\mathrm{i}^2 \\ &= -4. \end{aligned}$$

Siten

$$\begin{aligned} z_1 z_2^2 &= (1-\mathrm{i})(-4) \\ &= -4+4\mathrm{i}. \end{aligned}$$

(iii) Koska $z_2^* = 2\mathrm{i}$, niin

$$\begin{aligned} z_1 z_2^* &= (1-\mathrm{i})(2\mathrm{i}) \\ &= 2\mathrm{i}-2\mathrm{i}^2 \\ &= 2+2\mathrm{i}. \end{aligned}$$

(iv) Koska $z_2^* = 2\mathrm{i}$ ja

$$\begin{aligned} |z_2|^2 &= z_2 z_2^* \\ &= (-2\mathrm{i})(2\mathrm{i}) \\ &= 4, \end{aligned}$$

niin

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{z_1 z_2^*}{|z_2|^2} \\ &= \frac{(1-\mathrm{i})(2\mathrm{i})}{4} \\ &= \frac{2\mathrm{i}-2\mathrm{i}^2}{4} \\ &= 0.5+0.5\mathrm{i} \end{aligned}$$

6.14 Harjoitustehtävä
Olkoon $z_1 = 3-7\mathrm{i}$ ja $z_2 = 1+\mathrm{i}$. Laske

(i) $z_1 + 3z_2$,
(ii) $2z_1 z_2^3$,
(iii) $z_1 z_2^*$,
(iv) $0.5z_1/z_2$.

Kompleksilukujen napamuoto

Tason vektori on karteesiessa muodossa (x, y) ja napamuodossa $\langle r, \theta\rangle$, missä r on vektorin normi eli pituus ja θ on vektorin ja vaaka-akselin välinen kulma radiaaneissa. Kompleksitasossa tämä napamuoto voidaan esittää näppärästi kompleksisen eksponenttifunktion avulla.

Olkoon $z = x + y\mathrm{i}$. **Kompleksinen eksponenttifunktio** on

$$\mathrm{e}^z \quad = \quad \exp(z) \quad = \quad \mathrm{e}^x(\cos y + \mathrm{i}\sin y). \tag{6.15}$$

Määritelmä (6.15) ei varmaankaan ole ilmiselvä. Jos tarkastelemme reaalilukua $z = x + 0\mathrm{i}$, niin kaava (6.15) kyllä antaa perinteisen eksponenttifunktion. Siten määritelmä ainakin laajentaa perinteisen reaalisen eksponenttifunktion kompleksitasoon. Mutta miksi juuri tämä laajennus? Ja miten trigonometriset funktiot liittyvät asiaan?

Reaalianalyysin puolelta tiedämme, että

$$\mathrm{e}^x \quad = \quad \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n.$$

Voisimmeko ymmärtää tämän raja-arvon kompleksiselle luvulle z ja määritellä eksponenttifunktion tätä kautta? Itse asiassa voisimme ja saisimme tulokseksi kaavan (6.15). Emme perustele tätä tällä kurssilla.

Reaalianalyysin puolelta tiedämme myös, että

$$\mathrm{e}^x \quad = \quad \sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

Voisimmeko määritellä kompleksisen eksponenttifunktion tämän sarjan avulla. Itse asiassa voisimme ja saisimme tulokseksi kaavan (6.15). Perustelemme tätä hieman ja samalla näemme (joskin hämärästi) miten trigonometriset funktiot astuvat kuvaan. Lähdemme liikkeelle sinin ja kosinin esityksistä Taylorin sarjoina (kehitettyinä pisteen $y_0 = 0$ ympärille):

$$\begin{aligned}
\cos y &= \sum_{k=0}^{\infty}(-1)^k\frac{y^{2k}}{(2k)!},\\
\sin y &= \sum_{k=0}^{\infty}(-1)^k\frac{y^{2k+1}}{(2k+1)!}.
\end{aligned}$$

Siispä

$$\begin{aligned}
\cos y + \mathrm{i}\sin y &= \sum_{k=0}^{\infty}(-1)^k\left[\frac{y^{2k}}{(2k)!} + \mathrm{i}\frac{y^{2k+1}}{(2k+1)!}\right]\\
&= 1 + \mathrm{i}y \; - \; \frac{y^2}{2!} - \mathrm{i}\frac{y^3}{3!} \; + \; \frac{y^4}{4!} + \mathrm{i}\frac{y^5}{5!} \; - \; \cdots\\
&= \frac{\mathrm{i}^0y^0}{0!} + \frac{\mathrm{i}^1y^1}{1!} + \frac{\mathrm{i}^2y^2}{2!} + \frac{\mathrm{i}^3y^3}{3!} + \frac{\mathrm{i}^4y^4}{4!} + \frac{\mathrm{i}^5y^5}{5!} + \cdots\\
&= \sum_{k=0}^{\infty}\frac{(\mathrm{i}y)^k}{k!}.
\end{aligned}$$

Olemme perustelleet niin sanotun **Eulerin kaavan**

$$e^{iy} = \cos y + i \sin y. \tag{6.16}$$

Käyttämällä edellisiä laskuja ja reaalisen eksponenttifunktion sarjakehitelmää saamme

$$\begin{aligned}
e^{z} &= e^{x+iy} \\
&= e^{x} e^{iy} \\
&= e^{x} (\cos y + i \sin y) \\
&= \sum_{\ell=0}^{\infty} \frac{x^{\ell}}{\ell!} \sum_{k=0}^{\infty} \frac{(iy)^{k}}{k!}
\end{aligned}$$

Käyttämällä **Cauchyn kertokaavaa**

$$\sum_{\ell=0}^{\infty} a_{\ell} \sum_{k=0}^{\infty} b_{k} = \sum_{n=0}^{\infty} \sum_{j=0}^{n} a_{j} b_{n-j}$$

ja **binomikaavaa**

$$(x+y)^{n} = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^{k}$$

takaperin saamme

$$\begin{aligned}
e^{z} &= e^{x+iy} \\
&= e^{x} e^{iy} \\
&= e^{x} (\cos y + i \sin y) \\
&= \sum_{\ell=0}^{\infty} \frac{x^{\ell}}{\ell!} \sum_{k=0}^{\infty} \frac{(iy)^{k}}{k!} \\
&= \sum_{n=0}^{\infty} \sum_{j=0}^{n} \frac{x^{j}}{j!} \frac{(iy)^{n-j}}{(n-j)!} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} x^{j} (iy)^{n-j} \\
&= \sum_{n=0}^{\infty} \frac{(x+iy)^{n}}{n!} \\
&= \sum_{n=0}^{\infty} \frac{z^{n}}{n!}
\end{aligned}$$

Olemme siis perustelleet, että

$$\begin{aligned}
e^{z} &= e^{x} (\cos y + i \sin y) \\
&= \sum_{n=0}^{\infty} \frac{z^{n}}{n!}.
\end{aligned}$$

Lopuksi huomautamme vielä, että kuten reaalipuolella myös kompleksipuolella ekpsonenttifunktio voidaan määritellä yksikäsitteisesti fuktiona, jolle $\exp(0) = 1$ ja $\exp'(z) = \exp(z)$. Emme kuitenkaa käsittele tällä kurssilla kompleksista derivointia.

6.17 Huomautus

Sijoittamalla Eulerin kaavaan (6.16) $y = \pi$ saamme matematiikan kauneimman kaavan:

$$\mathrm{e}^{\pi\mathrm{i}} + 1 = 0.$$

6.18

Kompleksinen eksponenttifuktio toimii reaalisen eksponenttifunktion tavoin siinä mielessä, että sille pätee kaavat

(i) $\mathrm{e}^{z_1+z_2} = \mathrm{e}^{z_1}\mathrm{e}^{z_2}$,
(ii) $(\mathrm{e}^{z})^n = \mathrm{e}^{nz}$,
(iii) $\mathrm{e}^{z_1}/\mathrm{e}^{z_2} = \mathrm{e}^{z_1-z_2}$.

Toisin kuin reaalinen eksponenttifunktio, kompleksinen eksponenttifunktio ei ole "positiivinen", ei injektio, eikä kasvava. Itse asiassa jokainen kompleksiluku z voidaan esittää äärettömän monella eri kompleksiluvulla w muodossa $z = \mathrm{e}^w$. Erityisesti kompleksinen eksponenttifuktio on jaksollinen:

$$\mathrm{e}^z = \mathrm{e}^{z+2\pi n\mathrm{i}}$$

kaikilla kokonaisluvuilla n.

Kompleksilukujen napamuoto seuraa nyt Eulerin kaavasta (6.16). Kyseinen kaava nimittäin sanoo, että parametrisoitu käyrä $\mathrm{e}^{\mathrm{i}\theta}$, $\theta \in [0, 2\pi)$, piirtää kompleksitasoon yksikköympyrän (vastapäivään). Siten jokainen kompleksiluku $z = x + y\mathrm{i}$ voidaan esittää **napamuodossa**

$$z = r\mathrm{e}^{\theta\mathrm{i}},$$

missä

$$r = |z| = \sqrt{x^2 + y^2}$$

on kompleksiluvun z **moduuli** ja

$$\theta = \arg(z)$$

on kompleksiluvun z **argumentti**, eli positiivisen reaaliakselin (positiivinen x-akseli) ja kompleksilukua z vastaavan vektorin välinen kulma (vastapäivään kierrettynä). Argumentti θ ei ole yksikäsitteinen. Jos haluamme, että $\theta \in [0, 2\pi)$, voimme laskea sen kaavalla

$$\theta = \begin{cases} \arctan \frac{y}{x}, & \text{jos } x > 0 \text{ ja } y \geq 0, \\ \arctan \frac{y}{x} + 2\pi, & \text{jos } x > 0 \text{ ja } y < 0, \\ \arctan \frac{y}{x} + \pi, & \text{jos } x < 0, \\ \frac{\pi}{2}, & \text{jos } x = 0 \text{ ja } y > 0, \\ \frac{3\pi}{2}, & \text{jos } x = 0 \text{ ja } y < 0, \end{cases}$$

Jos $x, y = 0$ molemmat, niin kulma θ ei ole määritelty. Jos haluamme että $\theta \in (-\pi, \pi]$, voimme laskea sen kaavalla

$$\theta = \begin{cases} \arctan \frac{y}{x}, & \text{jos } x > 0, \\ \arctan \frac{y}{x} + \pi, & \text{jos } x < 0 \text{ ja } y \geq 0, \\ \arctan \frac{y}{x} - \pi, & \text{jos } x < 0 \text{ ja } y < 0, \\ \frac{\pi}{2}, & \text{jos } x = 0 \text{ ja } y > 0, \\ -\frac{\pi}{2}, & \text{jos } x = 0 \text{ ja } y < 0, \end{cases}$$

Edelleen, jos $x, y = 0$ molemmat, ei kulma θ ole määritelty.

Napamuodosta $z = re^{i\theta}$ saadaan karteesinen muoto $z = x + yi$ vaikkapa käyttämällä kaavoja

$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta. \end{aligned}$$

6.19 **Huomautus**

Kertaamme vielä karteesisen muodon ja napamuodon yhteyden kun valitsemme kulmaksi $\theta \in (-\pi, \pi]$. Päästäkseen kompleksiluvun z karteesisesta muodosta, $z = x + yi$, sen napamuotoon, on löydettävä positiivinen luku r ja kulma θ niin, että $z = re^{\theta i}$. Tässä $r = |z| = \sqrt{x^2 + y^2}$ ja kulma θ saadaan ratkaisemalla yhtälö

$$x + yi = z = re^{\theta i} = r(\cos(\theta) + i\sin(\theta) = r\cos(\theta) + ir\sin(\theta).$$

Kulma θ toteuttaa siis seuraavat yhtälöt:

$$x = r\cos\theta \quad \text{ja} \quad y = r\sin\theta$$

Toisin sanoen, on löydettävä θ niin, että

$$\cos\theta = \frac{x}{r} \quad \text{ja} \quad \sin\theta = \frac{y}{r}.$$

Tämä kulma $\theta \in (-\pi, \pi]$ löydetään esimerkiksi seuraavasti seuraavasti:

- Jos $y \geq 0$, niin $\theta = \arccos \frac{x}{r}$.
- Jos $y < 0$, niin $\theta = -\arccos \frac{x}{r}$.

6.20 **Esimerkki**

Etsimme karteesisessa muodossa annetun kompleksiluvun $z = 2 - 5i$ napamuodon. Moduulin saamme suoraan Pytharogaan kaavasta: $r = \sqrt{2^2 + 5^2} = \sqrt{29} = 5.3852$. Argumentin laskemiseksi tarkastelemme normalisoitua, ykkösen pituista, kompleksilukua

$$\begin{aligned} \hat{z} &= \frac{2}{\sqrt{29}} - \frac{5}{\sqrt{29}}i \\ &= 0.37139 - 0.92848i. \end{aligned}$$

Huomaamme, että tämä vektori sojottaa yksikkötason neljänteen neljännekseen. Siten (eräs) vaihekulma, eli argumentti, voidaan laskea esimerkiksi kaavasta

$$-\cos\theta = \frac{2}{\sqrt{29}}.$$

Siten

$$\begin{aligned}\theta &= -\arccos\frac{2}{\sqrt{29}} \\ &= -1.1903.\end{aligned}$$

Olemme siis saaneet karteesisesta muodosta $z = 2-5\mathrm{i}$ (erään) napamuodon

$$z = 5.3852\mathrm{e}^{-1.1904\mathrm{i}}.$$

Etsimme sitten napamuodossa annetun kompleksiluvun $z = 5\mathrm{e}^{\mathrm{i}\pi/4}$ karteesisen muodon. Vaihekulma $\pi/4$ on tunnetusti sama kuin $45°$. Siten kompleksiluvun z reaali- ja imaginääriosat ovat samat, mistä päädymme (reaaliseen) yhtälöön

$$5 = \sqrt{2x^2},$$

jonka ratkaisu on

$$\begin{aligned}x &= \sqrt{\frac{5^2}{2}} \\ &= 3.5355.\end{aligned}$$

Olemme siis saaneet napamuodosta $z = 5\mathrm{e}^{\mathrm{i}\pi/4}$ karteesisen muodon

$$z = 3.5355 + 3.5355\mathrm{i}.$$

6.21 Harjoitustehtävä
Esitä seuraavat karteesisessa muodossa annetut kompleksiluvut napamuodossa:

(i) $3+\mathrm{i}$,
(ii) $-2\mathrm{i}$,

ja esitä seuraavat napamuodossa annetut kompleksiluvut karteesisessa muodossa:

(iii) $3\mathrm{e}^{\mathrm{i}\pi/2}$,
(iv) $4\mathrm{e}^{\mathrm{i}\,317.3}$.

Kompleksiluvun karteesinen muoto $z = x + y\mathrm{i}$ sopii mainiosti summien laskemiseen, sillä reaalija kompleksiosta vain lasketaan erikseen yhteen:

$$\begin{aligned} z_1 + z_2 &= (x_1 + x_2\,,\; y_1 + y_2) \\ &= (x_1 + x_2) + (y_1 + y_2)\mathrm{i}. \end{aligned}$$

Tulojen laskemiseen karteesinen muoto sopii sitten hieman huonommin, koska pitää käyttää osittelulakia ja kaavaa $\mathrm{i}^2 = -1$:

$$\begin{aligned} z_1 z_2 &= (x_1 + y_1\mathrm{i})(x_2 + y_2\mathrm{i}) \\ &= x_1x_2 + x_1y_2\mathrm{i} + y_1x_2\mathrm{i} + y_1y_2\mathrm{i}^2 \\ &= (x_1x_2 - y_1y_2) + (x_1y_2 + y_1x_2)\mathrm{i}. \end{aligned}$$

Komplekisluvun napamuoto $z = r\mathrm{e}^{\mathrm{i}\theta}$ sopii mainioisti kompleksilukujen kertomiseen. Nimittäin

$$\begin{aligned} z_1 z_2 &= r_1\mathrm{e}^{\mathrm{i}\theta_1}\, r_2\mathrm{e}^{\mathrm{i}\theta_2} \\ &= r_1 r_2\, \mathrm{e}^{\mathrm{i}(\theta_1+\theta_2)}. \end{aligned}$$

Tästä kaavasta näemme myös kompleksilukujen kertolaskun geometrisen tulinnan: kompleksilukuja kerrottaessa moduulit r_1 ja r_2 kerrotaan keskenään ja argumentit θ_1 ja θ_2 lasketaan yhteen. Tarkastelaessä kompleksilukuja kompleksitason vektoreina tämä tarkoittaa sitä, että vektorien pituudet kerrotaan keskenään ja vektoreita kierretäään vastapäivään niiden vaihekulmien summan verran.

Lopuksi huomaamme, että kompleksilukujen potenssit voidaan laskea helposti napamuodosta:

$$z^n = r^n\mathrm{e}^{\mathrm{i}n}. \tag{6.22}$$

Kaavaa (6.22) kutsutaan **De Moivren** kaavaksi ja se toimii kaikilla kokonaisluvuilla n. Valitettavasti se ei toimi murtopotensseille (eli rationaaliluvuille) n/m saati reaalisille tai kompleksisille potensseille. Yleisessä tapauksessa, jos haluamme laskea z^w, niin se tulee kirjoittaa muodossa

$$z^w = \mathrm{e}^{w\log z}, \tag{6.23}$$

missä log on kompleksinen logaritmifunktio, joka on kompleksisen eksponenttifunktion käänteiskunktio. Koska kompleksinen eksponenttifunktio on jaksollinen, on kompleksinen logaritmifunktio moniarvoinen. Siten kaavassa (6.23) pitää valita jokin, tilanteesta riippuva, logaritmifunktion haara.

6.24 Harjoitustehtävä
Olkoon $z_1 = 1 - \mathrm{i}$ ja $z_2 = 2\mathrm{i}$. Esitä z_1^4/z_2^*

(i) napamudossa,
(ii) karteesisessa muodossa.

Esitämme lopuksi tämän luvun keskeiset havainnot lyhyesti:

6.25 Kompleksiluvut voidaan esittää kompleksitason vektoreina **karteesisessa muodossa**

$$z \;=\; x + y\mathrm{i},$$

missä x on kompleksiluvun z reaaliosa ja y on kompleksiluvun z imaginääriosa. Kompleksiluvut lasketaan yhteen laskemalla reaali- ja kompleksiosat yhteen erikseen.

Kompleksinen eksponenttifunktio on

$$\mathrm{e}^{z} \;=\; \mathrm{e}^{x}\left(\cos y + \mathrm{i}\sin y\right).$$

Jokainen nollasta poikkeava kompleksiluku voidaan esittää **napamuodossa**

$$z \;=\; r\mathrm{e}^{\mathrm{i}\theta},$$

missä $r = |z|$ on kompleksiluvun z **moduuli**, eli sitä vastaavan kompleksitason vektorin pituus, ja θ on kompleksiluvun z **argumentti**, eli sitä vastaavan kompleksitason vektorin ja positiivisen reaaliakselin välinen kulma.

Kompleksista lineaarialgebraa*

Aiemmissa luvuissa olemme käsitelleet vektoreita ja matriiseja, joiden alkiot ovat reaalilukuja. Mikään ei estä käsittelemästä vektoreita, joiden alkiot ovat kompleksilukuja ja muuttamaan skalaarit (eli aiemmmin reaaliluvut) kompleksiluvuiksi. Periaatteessa kaikki toimii kuten aiemminkin, mutta uutena piirteenä tulee mukaan kompleksikonjugaatit eli liittoluvut. Esimerkiksi kahden kaksiulotteisen kompleksivektorin

$$\mathbf{z}_1 = \begin{bmatrix} z_{11} \\ z_{12} \end{bmatrix} = \begin{bmatrix} x_{11} + y_{11}\mathrm{i} \\ x_{12} + y_{12}\mathrm{i} \end{bmatrix} \quad \text{ja} \quad \mathbf{z}_2 = \begin{bmatrix} z_{21} \\ z_{22} \end{bmatrix} = \begin{bmatrix} x_{21} + y_{21}\mathrm{i} \\ x_{22} + y_{22}\mathrm{i} \end{bmatrix}$$

välinen pistetulo eli sisätulo on

$$\begin{aligned}
\mathbf{z}_1 \cdot \mathbf{z}_2 &= \mathbf{z}_1^{\top}\mathbf{z}_2^{*} \\
&= z_{11}z_{21}^{*} + z_{12}z_{22}^{*} \\
&= (x_{11} + y_{11}\mathrm{i})(x_{21} - y_{21}\mathrm{i}) + (x_{12} + y_{12}\mathrm{i})(x_{22} - y_{22}\mathrm{i}) \\
&= x_{11}x_{21} - x_{11}y_{21}\mathrm{i} + x_{21}y_{11}\mathrm{i} + y_{11}y_{21} + x_{12}x_{22} - x_{12}y_{22}\mathrm{i} + x_{22}y_{12}\mathrm{i} + y_{12}y_{22} \\
&= x_{11}x_{21} + x_{12}x_{22} + y_{11}y_{21} + y_{12}y_{22} + (x_{21}y_{11} - x_{11}y_{21} + x_{22}y_{12} - x_{12}y_{22})\mathrm{i}.
\end{aligned}$$

Vastaavasti reaalipuolen ortogonaalinen matriisi, siis matriisi joka kuvaa ortonormaalin kannan ortonormaalille kannalle, eli matriisi $\mathbf{Q}$, jolle $\mathbf{Q}^{-1} = \mathbf{Q}^{\top}$, pitää korvata **unitaarisella matriisilla** $\mathbf{U}$, jolle $\mathbf{U}^{-1} = \mathbf{U}^{*}$, missä $\mathbf{U}^{*}$ on **konjugaattitranspoosi**:

$$(\mathbf{U}^{*})_{ij} \;=\; U_{ji}^{*}.$$

Luku 7

Tason symmetristen matriisien ominaisarvohajotelma

Diagonaalimatriisit ja ortogonaaliset matriisit

Matriisihajotelmien ideana on esittää annettu matriisi yksinkertaisempien matriisien tulona. Tarkastelemme tässä luvussa tason **symmetrisiä** neliömatriiseja. Toisin sanoen matriiseja

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

joille $A_{12} = A_{21}$ eli $\mathbf{A}^\top = \mathbf{A}$.

Esitämme tässä luvussa, miten symmetrinen matriisi voidaan esittää diagonaalimatriisin ja ortogonaalisen matriisin avulla.

Tason neliömatriisi $\mathbf{\Lambda}$ on **diagonaalimatriisi** eli **lävistäjämatriisi**, jos se on muotoa

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Diagonaalimatriisia vastaavan lineaarisen operaattorin geometrinen tulkinta on yksinkertainen: diagonaalimatriisi vastaa venytystä, jossa x_1-akseli venytetään λ_1-kertaisesti ja x_2-akselia venytetään λ_2-kertaisesti.

Diagonaalimatriisien algebra on myös erittäin yksinkertaista.

Diagonaalimatriisin determinatti on

$$\det(\mathbf{\Lambda}) = \lambda_1 \lambda_2.$$

Siten diagonaalimatriisi on kääntyvä jos ja vain jos molemmat diagonaalialkiot λ_1 ja λ_2 ovat nollasta poikkeavia. Itse käänteismatriisin laskeminen on myös helppoa:

$$\mathbf{\Lambda}^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{bmatrix}.$$

Itse asiassa diagonaalimatriisin potenssien (ja käänteispotenssien) laskeminen on yhtä helppoa:

$$\mathbf{\Lambda}^n = \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix}$$

kaikilla kokonaisluvuilla n.

Diagonaalimatriisien transponointi on poikkeuksellisen helppoa: $\mathbf{\Lambda}^\top = \mathbf{\Lambda}$.

7.1 Esimerkki

(i) Diagonaalimatriisi

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

projisoi tason pisteet $\mathbf{x} = [x_1\ x_2]^\top$ x_1-akselille "hävittämällä" x_2-koordinaatin. Tämä diagonaalimatriisi ei ole kääntyvä.

(ii) Diagonaalimatriisi

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

vaihtaa x_2-akselin suunnan. Tämä diagonaalimatriisi on siis **peilaus** x_2-akselin suhteen. Tämä diagonaalimatriisi on kääntyvä. Kääntäminen tapahtuu vaihtamalla x_2-akselin suunta uudestaan samalla tavalla: $\mathbf{C}^{-1} = \mathbf{C}$.

Toinen mukava luokka matriiseja on ortogonaaliset matriisit. Formaali määritelmä ortogonaaliselle matriisille on seuraava: tason neliömatriisi $\mathbf{Q}$ on **ortogonaalinen**, jos

$$\mathbf{Q}^\top = \mathbf{Q}^{-1}. \tag{7.2}$$

Määritelmä (7.2) on toki elegantti, mutta ehkä hieman läpinäkymätön. Ortogonaalisen matriisin idea on se, että on kannanvaihtomatriisi, joka kuvaa ortonormaalit kannat ortonormaaleiksi kannoiksi. Erityisesti se tarkoittaa sitä, että standardikanta $(\mathbf{e}_1, \mathbf{e}_2)$ kuvautuu ortonormaaliksi kannaksi. Koska $\mathbf{Q}$:n sarakkeet ovat standardikannan kuvat, tarkoittaa tämä sitä, että

$$\|\mathbf{Q}_{\bullet 1}\| = 1, \tag{7.3}$$

$$\|\mathbf{Q}_{\bullet 2}\| = 1, \tag{7.4}$$

$$\mathbf{Q}_{\bullet 1} \cdot \mathbf{Q}_{\bullet 2} = 0. \tag{7.5}$$

7.6 Harjoitustehtävä

Osoita että ehdot (7.3)–(7.5) ja (7.2) tarkoittavat samaa. Voit halutessasi käyttää apuna (jos siitä on apua) seuraavaksi esitettävää ortogonaalisten matriisien hajotelmaa $\mathbf{Q} = \mathbf{CR}$, missä $\mathbf{C}$ on peilaus ja $\mathbf{R}$ on kierto.

Tasossa ortogonaaliset matriisit voidaan esittää poikkeuksellisen konkreettisesti: ne koostuvat peilauksista ja kierroista. Matriisi $\mathbf{C}$ on **peilaus**, jos se on muotoa

$$\mathbf{C} = \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}.$$

Matriisi $\mathbf{R}$ on **kierto** (θ radiaania vastapäivään), jos se on muotoa

$$\mathbf{R} \;=\; \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Pienellä geometrisella pähkäilyllä on suhteellisen helppo nähdä, että matriisi $\mathbf{Q}$ on ortogonaalinen jos ja vain jos se on muotoa $\mathbf{Q} = \mathbf{CR}$, missä $\mathbf{C}$ on peilaus ja $\mathbf{R}$ on kierto.

Peilausmatriisien algebra on helppoa, koska ne ovat diagonaalimatriiseja. Kiertomatriisien algebra on myös helppoa: jos $\mathbf{R}(\theta)$ ja $\mathbf{R}(\phi)$ ovat kiertomatriiseja (θ ja ϕ radiaania vastapäivään), niin

$$\mathbf{R}(\theta)\mathbf{R}(\phi) \;=\; \mathbf{R}(\theta+\phi),$$

mistä seuraa esimerkiksi, että $\mathbf{R}(\theta)^n = \mathbf{R}(n\theta)$ kaikilla kokonaisluvuila n, ja erityisesti käänteismatriisille pätee $\mathbf{R}(\theta)^{-1} = \mathbf{R}(-\theta)$ (kierto myötäpäivään θ radiaania).

7.7 **Harjoitustehtävä**
Olkoon

$$\mathbf{A} \;=\; \begin{bmatrix} 0.7071 & -0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}.$$

Huomaa, että $0.7071 = \sqrt{2}/2$, ja laske

(i) $\mathbf{A}^{-1}$,
(ii) $\mathbf{A}^{8}$.

Ominaisarvohajotelman laskeminen

Entäpä sitten symmetriset matriisit $\mathbf{A}$? Jos $\mathbf{\Lambda}$ on diagonaalimatriisi ja $\mathbf{Q}$ on ortogonaalinen matriisi, niin

$$\mathbf{A} \;=\; \mathbf{Q\Lambda Q}^{-1} \tag{7.8}$$

on symmetrinen. Nimittäin, ensinnäkin

$$\begin{aligned} \mathbf{A}^\top &= (\mathbf{Q\Lambda Q}^{-1})^\top \\ &= (\mathbf{Q\Lambda Q}^\top)^\top, \end{aligned}$$

koska $\mathbf{Q}$ on ortogonaalinen. Toiseksikin, käyttämällä tulon transponointikaavaa ja matriisitulon assosiatiivisuutta, saamme

$$\begin{aligned} (\mathbf{Q\Lambda Q}^\top)^\top &= (\mathbf{Q}(\mathbf{\Lambda Q}^\top))^\top \\ &= (\mathbf{\Lambda Q}^\top)^\top \mathbf{Q}^\top \\ &= (\mathbf{Q}^\top)^\top \mathbf{\Lambda}^\top \mathbf{Q}^\top. \end{aligned}$$

Kolmanneksi, koska $\mathbf{\Lambda}^\top = \mathbf{\Lambda}$ (pätee diagonaalimatriiseille) ja $\mathbf{Q}^\top = \mathbf{Q}^{-1}$ (pätee ortogonaalisille matriiseille) sekä $(\mathbf{Q}^\top)^\top = \mathbf{Q}$ (pätee kaikille matriiseille), niin

$$(\mathbf{Q}^\top)^\top \mathbf{\Lambda}^\top \mathbf{Q}^\top \quad = \quad \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}.$$

Olemme siis osoittaneet, että kaavan (7.8) antamalle matriisille $\mathbf{A}$ pätee $\mathbf{A}^\top = \mathbf{A}$, eli $\mathbf{A}$ on symmetrinen.

Mutta onnistuuko tämä hajotelma toisin päin? Voidaanko annetulle symmetriselle matriisille $\mathbf{A}$ löytää ortogonaalinen matriisi $\mathbf{Q}$ ja diagonaalimatriisi $\mathbf{\Lambda}$ niin, että hajotelma (7.8) pätee? Ja mikä vielä olennaisempaa: miksi ylipäätään haluaisimme löytää hajotelman (7.8)?

7.9 Huomautus
Ennen hajotelman laskemista esitämme yhden ilmiselvän sovelluksen ominaisarvohajotelmalle: käänteismatriisin laskemisen. Jos $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$, niin

$$\begin{aligned} \mathbf{A}^{-1} &= (\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1})^{-1} \\ &= ((\mathbf{Q}\mathbf{\Lambda})\mathbf{Q}^{-1})^{-1} \\ &= (\mathbf{Q}^{-1})^{-1}(\mathbf{Q}\mathbf{\Lambda})^{-1} \\ &= \mathbf{Q}\mathbf{\Lambda}^{-1}\mathbf{Q}^{-1}. \end{aligned}$$

Koska $\mathbf{Q}^{-1} = \mathbf{Q}^\top$ ja $\mathbf{\Lambda}^{-1}$ ovat helppoja laskea, näemme että käänteismatriisin laskeminen on helppoa ominaisarvohajotelmasta.

Ominaisarvohajotelman (7.8) laskeminen kannattaa aloittaa ominaisarvojen, eli matriisin $\mathbf{\Lambda}$ laskemisella. Ominaisarvohajotelman nojalla tarkoitus on löytää sellaiset luvut λ ja vektorit $\mathbf{q}$, että

$$\mathbf{A}\mathbf{q} = \lambda\mathbf{q}. \tag{7.10}$$

Yhtälö (7.10) sanoo, että kuvaus $\mathbf{A}$ on suuntaan $\mathbf{q}$ venytys. Venytyksen voimakkuutta λ kutsutaan matriisin $\mathbf{A}$ **ominaisarvoksi** ja vektoria $\mathbf{q}$ ominaisarvoa λ vastaavaksi **ominaisvektoriksi**. Yhtälön (7.10) ratkaisemiseksi kirjoitamme sen muodossa

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{q} \quad = \quad \mathbf{0}.$$

Tästä muodosta näemme, että ratkaisu on olemassa (kun $\mathbf{q} \neq \mathbf{0}$), jos matriisi $\mathbf{A} - \lambda\mathbf{I}$ ei ole kääntyvä. Täten λ on niin sanotun **karakteristisen polynomin**

$$p(\lambda) \quad = \quad \det(\mathbf{A} - \lambda\mathbf{I})$$

nollakohta. Koska tarkastelemme vain 2×2 symmetrisiä matriiseja (jolloin $A_{12} = A_{21}$), voimme laskea karakteristisen polynomin auki kohtalaisen helposti:

$$\begin{aligned} p(\lambda) &= \det\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} A_{11} - \lambda & A_{12} \\ A_{21} & A_{22} - \lambda \end{bmatrix}\right) \\ &= (A_{11} - \lambda)(A_{22} - \lambda) - A_{12}A_{21} \\ &= (A_{11} - \lambda)(A_{22} - \lambda) - A_{12}^2 \\ &= \lambda^2 - (A_{11} + A_{22})\lambda + A_{11}A_{22} - A_{12}^2. \end{aligned}$$

Näemme, että $p(\lambda)$ on toisen asteen polynomi, joten sen ratkaisut voidaan laskea vanhalla tutulla ratkaisukaavalla:

$$\begin{aligned}\lambda_1 &= \frac{A_{11}+A_{22}+\sqrt{(A_{11}+A_{22})^2-4(A_{11}A_{22}-A_{12}^2)}}{2},\\ \lambda_2 &= \frac{A_{11}+A_{22}-\sqrt{(A_{11}+A_{22})^2-4(A_{11}A_{22}-A_{12}^2)}}{2}.\end{aligned}$$

Tästä näemme, että ratkaisut ovat aina reaalisia, sillä voidaan osoittaa, että aina pätee

$$(A_{11}+A_{22})^2 \geq 4(A_{11}A_{22}-A_{12}^2).$$

Samoin näemme, että ratkaisut λ_1 ja λ_2 ovat samoja jos

$$(A_{11}+A_{22})^2 = 4(A_{11}A_{22}-A_{12}^2),$$

mikä pelkistyy ehdoiksi $A_{12}=0$ ja $A_{11}=A_{22}=\lambda$. Tämä siis tarkoittaa sitä, että $\mathbf{A}$ on jo valmiiksi diagonaalimatriisi, ja erityisesti se on muotoa $\mathbf{A}=\lambda\mathbf{I}$.

Oletamme nyt, että olemme löytäneet kaksi ominaisarvoa $\lambda_1<\lambda_2$ (sillä tilanne $\lambda_1=\lambda_2$ osoittautui tylsäksi).

Etsimme ominaisarvoa λ_1 vastaavan ominaisvektorin $\mathbf{q}_1=[q_{11}\ \ q_{12}]^\top$. Tämä tarkoittaa ominaisarvoyhtälön

$$\mathbf{A}\mathbf{q}_1 = \lambda_1\mathbf{q}_1$$

ratkaisemista. Koska λ_1 on tunnettu, tämä on normaali yhtälöpari, joka voidaan ratkaista normaaliin tapaan esimerkiksi kirjoittamalla se muotoon

$$(\mathbf{A}-\lambda_1\mathbf{I})\mathbf{q}_1 = 0$$

ja ratkaisemalla tätä muotoa vastaava laajennettu lohkomatriisi

$$\left[\begin{array}{ccc} A_{11}-\lambda_1 & A_{12} & 0\\ A_{21} & A_{22}-\lambda_1 & 0\end{array}\right].$$

Ratkaisuja on ääretön määrä (ja nollakin on mitä ilmeisimmin ratkaisu). Valitsemme sellaisen ratkaisun, jolle $\|\mathbf{q}_1\|=1$. Samalla tavalla ominaisarvoa λ_2 vastaava ominaisvektori $\mathbf{q}_2=[q_{21}\ \ q_{22}]^\top$ löytyy ratkaisemalla normaaliin tapaan ominaisarvoyhtälö

$$\mathbf{A}\mathbf{q}_2 = \lambda_2\mathbf{q}_2.$$

Matriisin $\mathbf{A}$ ominaisarvohajotelma (7.8) on nyt itse asiassa löydetty:

$$\mathbf{\Lambda}=\left[\begin{array}{cc}\lambda_1 & 0,\\ 0 & \lambda_2\end{array}\right] \quad \text{ja} \quad \mathbf{Q}=\left[\begin{array}{cc} q_{11} & q_{21}\\ q_{12} & q_{22}\end{array}\right].$$

7.11 Esimerkki
Etsimme symmetrisen matriisin

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -1 & 5 \end{bmatrix}$$

ominaisarvohajotelman.

Nyt

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 3-\lambda & -1 \\ -1 & 5-\lambda \end{bmatrix}.$$

Siten karakakteristinen polynomi on

$$\begin{aligned} p(\lambda) &= (5-\lambda)(3-\lambda)-1 \\ &= \lambda^2 - 8\lambda + 14, \end{aligned}$$

jonka nollakohdat voidaan laskea toisen asteen yhtälön ratkaisukaavasta. Saamme

$$\begin{aligned} \lambda_1 &= 2.5858, \\ \lambda_2 &= 5.4142. \end{aligned}$$

Ominaisarvoa $\lambda_1 = 2.5858$ vastaava ominaisvektori $\mathbf{q}_1$ saadaan yhtälöparista

$$\begin{bmatrix} 3-2.5858 & -1 \\ -1 & 5-2.6858 \end{bmatrix}\mathbf{q}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

eli kaaviomuodosta

$$\begin{bmatrix} 0.4142 & -1.0000 & 0.0000 \\ -1.0000 & 2.4142 & 0.0000 \end{bmatrix}.$$

Tämän kaavion voi ratkaista normaaliin tapaan esimerkiksi eliminoimalla q_{11}- muuttujan toiselta riviltä:

$$\begin{bmatrix} 0.4142 & -1.0000 & 0.0000 \\ -1.0000 & 2.4142 & 0.0000 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0.4142 & -1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$

Toinen rivi eliminoitui täysin (ja näin pitikin käydä, sillä matriisin $\mathbf{A} - \lambda_1\mathbf{I}$ determinantti on nolla. Ensimmäiseltä riviltä luemme, että ratkaisut ovat muotoa

$$q_{12} = 0.4142 \cdot q_{11}.$$

Nyt pitää valita sellainen $\mathbf{q}_1$, että $\|\mathbf{q}_1\|^2 = 1$. Toisin sanoen

$$\begin{aligned} 1 &= q_{11}^2 + q_{12}^2 \\ &= q_{11}^2 + 0.4142^2 q_{11}^2 \\ &= 1.1716 q_{11}^2 \end{aligned}$$

Siten

$$\begin{aligned} q_{11} &= \sqrt{1/1.716} &= 0.9239, \\ q_{12} &= 0.4142 \cdot 0.9239 &= 0.3829 \end{aligned}$$

eli $\mathbf{q}_1 = [0.9239\ 0.3829]^\top$.

Ominaisarvoa $\lambda_2 = 5.4142$ vastaava ominaisvektori $\mathbf{q}_2$ saadaan yhtälöparista

$$\begin{bmatrix} 3-5.4142 & -1 \\ -1 & 5-5.4142 \end{bmatrix} \mathbf{q}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

eli kaaviomuodosta

$$\begin{bmatrix} -2.4142 & -1.0000 & 0.0000 \\ -1.0000 & -0.4142 & 0.0000 \end{bmatrix}.$$

Hävittämällä muuttuja q_{21} toiselta riviltä, eliminoituu toinen rivi täysin (niin kuin pitääkin):

$$\begin{bmatrix} -2.4142 & -1.0000 & 0.0000 \\ -1.0000 & -0.4142 & 0.0000 \end{bmatrix} \rightsquigarrow \begin{bmatrix} -2.4142 & -1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 \end{bmatrix}.$$

Saamme siis ratkaisut, jotka ovat muotoa

$$q_{22} = -2.4142 \cdot q_{21},$$

ja normalisointiehdosta $\|\mathbf{q}_2\|^2 = 1$, saamme ehdon

$$\begin{aligned} 1 &= q_{21}^2 + q_{22}^2 \\ &= q_{21}^2 + 2.4142^2 q_{21}^2 \\ &= 6.8284 \cdot q_{21}^2. \end{aligned}$$

$$\begin{aligned} q_{21} &= \sqrt{1/6.8284} &= 0.3827, \\ q_{22} &= -2.4142 \cdot 0.3827 &= -0.9239, \end{aligned}$$

eli $\mathbf{q}_2 = [0.3827\ \ -0.9239]^\top$.

Olemme siis löytäneet matriisin $\mathbf{A}$ ominaisarvohajotelman $\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$:

$$\begin{aligned} \boldsymbol{\Lambda} &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 2.5858 & 0.0000 \\ 0.0000 & 5.4142 \end{bmatrix} \\ \mathbf{Q} &= \begin{bmatrix} q_{11} & q_{21} \\ q_{12} & q_{22} \end{bmatrix} = \begin{bmatrix} 0.9239 & 0.3827 \\ 0.3829 & -0.9239 \end{bmatrix}. \end{aligned}$$

7.12 Harjoitustehtävä
Esimerkissä 7.11 on pyöristysvirheitä. Laske se tarkoilla arvoilla.

7.13 Harjoitustehtävä
Laske seuraavien matriisien ominaisarvohajotelmat:

(i)

$$\mathbf{A} = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix},$$

(ii)

$$\mathbf{A} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

Keräämme tähän loppuun joitakin keskeisiä havaintojamme ominaisarvohajotelmasta:

7.14

Jokainen symmetrinen matriisi $\mathbf{A}$ voidaan esittää **ominaisarvohajotelmana**

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top,$$

missä $\mathbf{Q}$ on **koordinaatiston vaihto**: se muuttaa tason standardikannan uudeksi ortonormaaliksi kannaksi. Standardikannan saa takaisin käänteismuunnoksella $\mathbf{Q}^{-1} = \mathbf{Q}^\top$. Matriisi $\mathbf{\Lambda}$ on diagonaalimatriisi. Se siis venyttää “$\mathbf{Q}$-koordinaatistoa” eli **ominaisvektoreita** matriisin $\mathbf{\Lambda}$ diagonaalilla olevien **ominaisarvojen** verran.

Ominaisarvohajotelma voidaan laskea ratkaisemalla ensin ominaisarvot λ_i **ominaisarvoyhtälöistä**

$$\mathbf{A}\mathbf{q}_i = \lambda_i \mathbf{q}_i$$

esimerkiksi laskemalla **karakteristisen polynomin**

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$$

nollakohdat. Ominaisvektorit $\mathbf{q}_i$, eli ortogonaalisen matriisin $\mathbf{Q}$ sarakkeet, saadaan ratkaisemalla ominaisarvoyhtälöt, kun niihin on ensin sijoitettu aikaisemmin saadut ominaisarvot. Ratkaisuiksi pitää valita ykkösen pituiset vektorit $\mathbf{q}_i$.

Ominaisarvohajotelma on hyödyllinen esimerkiksi (käänteis)potenssien $\mathbf{A}^k$, missä k on kokonaisluku, laskemisessa.

Ominaisarvohajotelma yleisessä tapauksessa*

Jos $\mathbf{A}$ on symmetrinen $n \times n$-matriisi, niin sillä on ominaisarvohajotelma $\mathbf{A} = \mathbf{Q\Lambda Q}^{-1}$. Tämä ominaisarvohajotelma voidaan laskea teoriassa tismalleen samalla tavalla kuin 2×2-matriisellekin. Käytännössä tämä ei kuitenkaan onnistu, sillä nyt karakteristinen polynomi $p(\lambda)$ on n. astetta, joten sen nollakohtien ratkaiseminen ei onnistu millään ratkaisukaavalla, vaan joudumme turvautumaan numeerisiin menetelmiin.

Jos $\mathbf{A}$ ei ole symmetrinen matriisi, niin sillä ei voi olla hajotelmaa $\mathbf{A} = \mathbf{Q\Lambda Q}^{-1}$, missä $\mathbf{Q}$ on ortogonaalinen ja $\mathbf{\Lambda}$ on diagonaalimatriisi. Periaatteessa voi kuitenkin joskus olla mahdollista löytää hajotelma

$$\mathbf{A} \;=\; \mathbf{P\Lambda P}^{-1} \tag{7.15}$$

missä $\mathbf{\Lambda}$ on edelleen diagonaalimatriisi, mutta $\mathbf{P}$ on vain jokin kääntyvä matriisi. Hajotelmaa (7.15) kutsutaan myös joskus (valitettavasti) ominaisarvohajotelmaksi. Esimerkiksi epäsymmetrisen matriisin

$$\mathbf{A} \;=\; \begin{bmatrix} 1 & 3 \\ -1 & 5 \end{bmatrix}$$

kaavan (7.15) mukainen ominaisarvohajotelma on

$$\mathbf{P} \;=\; \begin{bmatrix} -0.9486 & -0.7071 \\ -0.3162 & -0.7071 \end{bmatrix} \quad \text{ja} \quad \mathbf{\Lambda} \;=\; \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

Tässä siis $\mathbf{P}$ ei ole ortogonaalinen: $\mathbf{P}^{-1} \neq \mathbf{P}^{\top}$.

Joskus voimme rakentaa (ortogonaalistyyppisen) ominaisarvohajotelman piipahtamalla kompleksipuolella. Esimerkiksi antisymmetrinen matriisi

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

voidaan esittää muodossa $\mathbf{U\Lambda U}^{-1}$, missä

$$\mathbf{U} \;=\; \begin{bmatrix} \cos(\pi/4) & \cos(\pi/4) \\ -\cos(\pi/4)\cdot \mathrm{i} & \cos(\pi/4)\cdot \mathrm{i} \end{bmatrix} \quad \text{ja} \quad \mathbf{\Lambda} \;=\; \begin{bmatrix} 2+\mathrm{i} & 0 \\ 0 & 2-\mathrm{i} \end{bmatrix}.$$

Tässä siis $\mathbf{\Lambda}$ on diagonaalimatriisi ja $\mathbf{U}$ on unitaarinen matriisi: $\mathbf{U}^{-1} = \mathbf{U}^{*}$.

Yleisessä tapauksessa, missä $\mathbf{A}$ on vain jokin (kompleksialkioinen) $m \times n$-matriisi, voidaan rakentaa niin sanottu **singulaariarvohajotelma**

$$\mathbf{A} \;=\; \mathbf{U\Sigma V}^{*}, \tag{7.16}$$

missä $\mathbf{U}$ ja $\mathbf{V}$ ovat unitaarisia matriiseja ja $\mathbf{\Sigma}$ on diagonaalimatriisi, jonka diagonaalialkioita kutsutaan matriisin $\mathbf{A}$ **singulaariarvoiksi**. Karkeasti ottaen singulaariarvohajotelma sanoo seuraavaa: ensin pyöritetään koordinaatisto uudeksi koordinaatistoksi käyttämällä muunnosta $\mathbf{V}^{*}$. Tämän jälkeen venytetään akseleita matriisin $\mathbf{\Sigma}$ singulaariarvoilla ja lopuksi vielä pyöräytetään (joskaan ei ehkä takaisin) koordinaatistoa kuvauksella $\mathbf{U}$. Singulaariarvohajotelmaa kutsutaan joskus myös **pääakselihajotelmaksi**. Lopuksi on hyvä huomata, että hajotelmassa (7.16) ainoastaan diagonaalimatriisi $\mathbf{\Sigma}$ on yksikäsitteinen.

Luku 8

Lineaarialgebraa GNU Octavella

GNU Octave on avoimen lähdekoodin versio legendaarisesta Matlab-ohjelmistosta. Sen saa ladattua osoitteesta https://octave.org/. Jos et jostain syystä halua asentaa GNU Octavea, voit käyttää sitä myös verkossa osoitteessa https://octave-online.net/

Näitä luentoja kirjoitettaessa viimeisin stabiili versio GNU Octavesta oli 7.2.0, mutta uudemmat ja vahemmat versiot toimivat tämän kurssin puitteissa varmasti hyvin. Tämän kurssin kannalta myös Matlab toimii täsmälleen samoin kuin GNU Octave.

Esitämme GNU Octaven käyttöä varsin minimaalisesti. Lisää ohjeita GNU Octaven käytöstä löytyy esimerkiksi sen manuaalista, joka löytyy osoitteesta https://docs.octave.org/latest/

Lineaarinen yhtälöryhmä GNU Octavella

GNU Octave on matriisiorientoitunut. Tämä tarkoittaa sitä, että GNU Octavelle enemmän tai vähemmän kaikki otukset tulkitaan matriiseksi. Matriisi esitetään GNU Octavelle hakasulkeissa niin että elementit on erotettu toisistaan joko välilyönnillä tai pilkuilla ja rivit on erotettu toisistaan puolipisteillä. Siten esimerkiksi matriisi

$$\begin{bmatrix} 0.11 & 0.12 \\ 0.21 & 0.22 \end{bmatrix}$$

on GNU Octavessa `[0.11 0.12; 0.21 0.22]` tai `[0.11, 0.12; 0.21, 0.22]`.

Transpoosi $^\top$ on GNU Octavessa heittomerkki `'`. Siten esimerkiksi pystyvektori $[0.1\ \ 0.2]^\top$ on GNU Octavessa joko `[0.1 0.2]'`, `[0.0, 0.2]'` tai `[0.1; 0.2]`.

GNU Octavessa matriisin `A` elementteihin viitataan sulkeilla. Jos esimerkiksi `A = [11 10; 11 23]` niin `A(1,2) = 12`. Vastaavasti `A(2,2)=23`.

Tulomerkki on GNU Octavessa `*`, ja se tarkoittaa matriisituloa. Siten, jos esimerkiksi `A = [11 12; 21 22]` ja `B = [1 0; 0 2]`, niin `A*B = [11 24; 21 44]`.

Matriisin `A` käänteismatriisi on GNU Octavessa joko `inv(A)` tai `A^(-1)`. Siten yhtälöryhmän `A*x = b` voi ratkaista asettamalla `x = A^(-1)*b`. Usein ei kutenkaan kannata ratkaista yhtälöryhmää näin. Kuten tiedämme voi nimittäin olla että yhtälöryhmällä `A*x = b` on ratkaisu, vaikka `A` ei olekaan kääntyvä. GNU Octave osaa ottaa tämän **vain jossain määrin** huomioon ja ratkaista yhtälöryhmän “jakamalla sen vasemmalta” käskyllä `x = A\b`. Tässä siis operaatio \ tarkoittaa “vasemmalta jakamista”. Formaalisti ajatus menee näin: `A*x=b` voidaan jakaa vasemmalla puolittain matriisilla

`A`, jolloin `x = A\A*x=A\b`, koska `A\A` on “identiteetti”. Jätämme lukijan harrastuneisuuden varaan miettiä, mitä “oikealta jakaminen” / tarkoittaa.

Jos yhtälöllä `A*x=b` on useita ratkaisuja komento `A\b` antaa vain yhden ratkaisuista. Jos yhtälöllä `A*x=b` ei ole ratkaisuja, antaa Octaven komento `A\b` sille silti mukaratkaisun. Tämä ei kuitenkaan ole luonnollisestikaan yhtälön oikea ratkaisu, vaan niin sanottu miniminormiratkaisu. Tämä tarkoittaa sitä, että Octave antaa yhtälölle `A*x=b` sellaisen mukaratkaisun `x`, jolle `norm(A*x-b)` on mahdollisimman pieni. Tässä `norm(a)` on verktorin `a` normi, eli `norm(a)` $= \|\mathbf{a}\|$. Vielä kertaalleen toisin sanoen GNU Octaven vasemmalta jakaminen $\mathbf{A}\backslash\mathbf{b}$ antaa yhtälölle $\mathbf{Ax}=\mathbf{b}$ **jonkin** sellaisen mukaratkaisun $\mathbf{x}$, jolle normivirhe $\|\mathbf{Ax}-\mathbf{b}\|$ on mahdollisimman pieni.

Lisää GNU Octavesta ja yhtälöryhmistä löytyy osoitteesta https://docs.octave.org/latest/Simple-Examples.html. Suosittelen lämpimästi lukemaan!

8.1 Esimerkki
Olkoot

$$\mathbf{A}=\begin{bmatrix} 0.51 & 0.02 \\ 1.21 & 2.02 \end{bmatrix} \quad \text{ja} \quad \mathbf{b}=\begin{bmatrix} 0.00 \\ 1.15 \end{bmatrix}.$$

Seuraava konsolikeskustelu (Command Window) ratkaisee yhtälöparin $\mathbf{Ax}=\mathbf{b}$ GNU Octavella:

```
>> A = [0.51 0.02; 1.21 2.02]
A =

   0.510000   0.020000
   1.210000   2.020000

>> b = [0.00 1.15]'
b =

        0
   1.1500

>> x = A\b
x =

  -0.022863
   0.583002
```

8.2 Harjoitustehtävä
Olkoot

$$\mathbf{A}=\begin{bmatrix} 2.50 & 0.00 \\ 7.20 & 2.20 \end{bmatrix} \quad \text{ja} \quad \mathbf{b}=\begin{bmatrix} -2.00 \\ 2.75 \end{bmatrix}.$$

Ratkaise yhtälöpari $\mathbf{Ax}=\mathbf{b}$ GNU Octavella.

8.3 **Harjoitustehtävä**
Etsi sellainen yhtälöpari $\mathbf{Ax} = \mathbf{b}$, jolla on ratkaisu tai ratkaisuja, vaikka $\mathbf{A}$ ei olekaan kääntyvä. Ratkaise $\mathbf{Ax} = \mathbf{b}$ GNU Octavella.

Ominaisarvohajotelma GNU Octavella

GNU Octave löytää matriisin `A` ominaisarvohajotelman funktiolla `eig`, jonka syntaksi on

```
[Q, LAMBDA] = eig(A)
```

missä `Q` on matriisi, jonka sarakkeet ovat ominaisarvovektorit ja `LAMBDA` on lävistäjämatriisi, jonka lävistäjillä on matriisin `Q` ominaisvektoreita vastaavat ominaisarvot. Suosittelen lämpimästi tutustumaan funktion `eig` ohjeeseen kirjoittamalla `help eig` GNU Octaven konsoliin (Command Window).

8.4 **Esimerkki**
Etsimme esimerkin 7.11 ominaisarvohajotelman GNU Octavella. Tämä tapahtuu seuraavan konsolikeskustelun kautta

```
>> A = [3 -1; -1 5]
A =

   3  -1
  -1   5

>> [Q,LAMBDA] = eig(A)
Q =

  -0.9239  -0.3827
  -0.3827   0.9239

LAMBDA =

Diagonal Matrix

   2.5858        0
        0   5.4142
```

Huomamme että saamamme `Q` ei ole sama kuin esimerkin 7.11 $\mathbf{Q}$ (`Q`$=-\mathbf{Q}$). Tämä ei kuitenkaan ole virhe, sillä ominaisarvohajotelma ei ole yksikäsitteinen.

8.5 Harjoitustehtävä
Laske GNU Octavella matriisin

$$\mathbf{A} = \begin{bmatrix} 0.50 & 0.02 \\ 0.02 & 2.00 \end{bmatrix}$$

ominaisarvohajotelma.

8.6 Harjoitustehtävä
Mitä tapahtuu jos yrität laskea GNU Octavella epäsymmetrisen matriisin

$$\mathbf{A} = \begin{bmatrix} 0.50 & 0.02 \\ 3.00 & 2.00 \end{bmatrix}$$

ominaisarvohajotelman? Analysoi tulos.