

Mika Hirvensalo

Insinöörimatematiikka D 2019

Sisältö

1	Lineaarialgebran peruskäsitteitä	3
1.1	Lineaariset yhtälöryhmät	3
1.1.1	Gaussin-Jordanin menetelmä	3
1.1.2	Ratkaisujoukon systemaattinen esittäminen	6
1.2	Vektoriavaruudet	8
1.3	Vektoriavaruuksien rakenteesta	9
2	Matriisit	17
2.1	Kertausta	17
2.2	Lineaarikuvaukset	19
2.3	Käänteismatriisi	24
2.4	Determinantti	26
2.5	Matriisin ominaisarvot ja ominaisvektorit	28
2.6	Matriisin diagonalisointi	31
2.7	Jordanin normaalimuoto	32
2.8	Reaalifunktioiden laajennuksia matriiseille	40
3	Vektoriavaruuksien geometriaa	45
3.1	Etäisyys, normi, sisätulo	45
3.2	Etäisyyden säilyttävät kuvaukset	49
3.3	Korrelaatiokerroin	50
4	Kolmiulotteisen avaruuden geometriaa	53
4.1	Ristitulo	53
4.2	Tasot	54
4.3	Suorat	58
4.4	Etäisyyksien laskemisesta	61
5	Differentiaaliyhtälöt (DY:t)	63
5.1	Yksinkertaiset DY:t	63
5.2	Lineaariset DY:t	64
5.2.1	Vakiokertoimiset lineaariset DY:t	65
5.2.2	Ensimmäisen kertaluvun lineaarinen DY	69
5.2.3	Toisen kertaluvun lineaariset DY:t	71
5.3	Separoituvat DY:t	72
5.4	Eksaktit DY:t	74
5.5	Sekalaisia menetelmiä	76
5.5.1	Sijoitukset $\frac{y}{x}$ ja $ax + by$	76
5.5.2	Bernoullin DY	76
5.5.3	Käänteisfunktioon siirtyminen	77
5.5.4	Muuttujan vaihtaminen	77
5.5.5	Yhtälö $y'' = f(y)$	78
5.6	Differentiaaliyhtälöiden sarjaratkaisut	79
5.7	Lineaarisista vakiokertomisista DY-ryhmistä	83
5.8	Autonomisista DY-pareista	86

5.9	Differentiaaliyhtälöiden numeerisesta ratkaisemisesta	87
-----	---	----

Johdanto

Tämän opintojakson alkupuolella perehdytään lineaarialgebraan, jonka peruskäsitteisiin vektoreihin ja matriiseihin on pintapuolisesti tutustuttu jo kurssilla Insinöörimatematiikka A. Differentiaali- ja integraalilaskentaan sekä niiden sovelluksiin keskittyvissä kursseissa Insinöörimatematiikka B ja C ei erityisesti viitata lineaarialgebraan, mutta osoittautuu, että Fourier-analyysi, erityisesti sen diskreetti muoto voidaan ymmärtää perusteellisesti vain lineaarialgebran käsitteiden kautta.

Kurssin loppuosa keskittyy differentiaaliyhtälöiden käsittelyyn. Differentiaaliyhtälöt ovat keskeisessä roolissa matemaattisten mallien rakentamisessa niin luonnontieteissä kuin muissakin matemaattikkaa soveltavissa tieteissä. Jonkin verran pelkistäen voidaan sanoa, että lähestulkoon kaikki klassisen mekaniikan, sähkö- ja magnetismin, suhteellisuusteorian sekä kvanttimekaniikan mallit voidaan esittää differentiaaliyhtälöiden avulla ja näinollen näistä johdettavat ennusteet perustuvat differentiaaliyhtälöiden ratkaisuihin.

Aiemmissa insinöörimatematiikan kurssikokonaisuuden osissa esitetyt tiedot ovat edellytyksenä tämän opintojakson menestykselle suorittamiselle. Matriiseihin liittyvät peruskäsitteet ja Gaussin-Jordanin menetelmä redusoidun porrasmuodon löytämiseksi tulee kerrata tämän kurssin alkaessa. Differentiaaliyhtälöiden ratkaisemisessa keskeisimmiksi matemaattisen taitamisen lajeiksi nousevat derivointi- ja erityisesti integrointitekniikat.

Luku 1

Lineaarialgebran peruskäsitteitä

Lineaarialgebra on vektoreita ja näiden muodostamia algebrallisia rakenteita, *vektoriavaruuksia* tarkasteleva matematiikan osa-alue. Tärkeän osan lineaarialgebraa muodostavat vektoriavaruuksien funktiot, jotka säilyttävät algebrallisen rakenteen alkukuvan ja kuvan välillä. Näitä kutsutaan lineaarikuvauksiksi.

Lineaarialgebralla on sovelluksia useilla muilla matematiikan aloilla, tietojenkäsittelyssä, luonnontieteissä, tilastotieteessä ja yhteiskuntatieteissä. Lukuisine sovelluksineen lineaarialgebra lieneekin ryhmäteorian ohella matematiikan monikäyttöisin alue. Varhaisimmat viitteet nykyisin tunnistettaviin lineaarialgebran käsitteisiin ovat peräisin 1600-luvun lopulta, mutta nykyään käytössä olevaan lineaarialgebran käsitteistöä on kehitetty vasta 1800-luvun puolivälistä alkaen. Lineaarialgebra on täten differentiaali- ja integraalilaskentaa jonkin verran nuorempi matematiikan ala. Erityisen käytökelpoista molempien alojen kannalta on ollut 1800-luvun lopulla alkanut kehitystyö, jonka myötä on voitu löytää näennäisesti erilaisia matematiikan yhdistäviä rakenteita.

Erilaisten matemaattisten rakenteiden yhdistäminen ei ole edistänyt pelkästään matematiikan kehitystä, vaan on edellytys esimerkiksi kvanttimekaniikan nykyiselle esitystavalle. Kvanttimekaniikan laskennallisesti käyttökelpoisin, ns. Hilbertin avaruuksille perustuva formaalinen esitystapa hyödynää yhtä lailla lineaarialgebran kuin myös differentiaali- ja integraalilaskennan teoriaa.

1.1 Lineaariset yhtälöryhmät

Gaussin-Jordanin menetelmä lineaarisen yhtälöryhmän ratkaisemiseksi on esitetty insinöörimatematiikka A:ssa, mutta kertauksen ja menetelmän lukuisten käyttökohteiden vuoksi selostetaan menetelmä tässä lyhyesti uudelleen.

1.1.1 Gaussin-Jordanin menetelmä

Määritelmä 1 Yhtälöryhmä on lineaarinen, jos siinä on vain ensimmäisen ja nollannen asteen termejä:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (1.1)$$

Huomaa, että tässä määritelmässä muotoa $x_i x_j$ olevia termejä pidetään toisen asteen termeinä vaikka $i \neq j$. Ylläoleva muoto voidaan saavuttaa siirtämällä kaikki vakiotermit oikealle puolelle.

Menetelmä lineaarisen yhtälöryhmän ratkaisemiseksi on hyvin yksinkertainen: Ryhmään sovelletaan seuraavia alkeisoperaatioita: 1) yhtälön kertominen nollasta poikkeavalla vakiolla, 2) yhtälön

lisääminen toiseen vakiolla kerrottuna ja 3) yhtälöjen paikan vaihto, kunnes ryhmä saadaan niin yksinkertaiseen muotoon, että ratkaisu voidaan lukea suoraan.

Koska edellämmainitut operaatiot kohdistuvat muuttujien kertoimiin, itse muuttujia ole tarpeen kirjoittaa näkyviin ollenkaan. Näin ollen ylläoleva yhtälöryhmä voidaan pelkistää matriisiksi

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix},$$

jota sanotaan yhtälöryhmän (1.1) augmentoiduksi eli laajennetuksi matriisiksi. Jos $b_1 = b_2 = \dots = b_m = 0$, sanotaan yhtälöryhmää homogeeniseksi ja tällöin viimeinen sarake voidaan jättää merkitsemättä: Matriisia

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

sanotaan yhtälöryhmän kerroinmatriisiksi. Yhtälöryhmään kohdistettavat operaatiot korvataan tällöin matriisiin kohdistettaviin rivioperaatioihin: 1) rivin kertominen nollasta poikkeavalla vakiolla, 2) rivin lisääminen toiseen vakiolla kerrottuna ja 3) rivien paikan vaihtaminen. Jos matriisi B saadaan A :sta alkeisoperaatioita soveltamalla, merkitään $A \sim B$.

Kertaa kurssilta Insinöörimatematiikka A miten voidaan helposti huomata, että yhtälöryhmän alkeisoperaatiot eivät muuta ratkaisujoukkoa miksiäkään. Siksi matriiseja A ja B vastaavilla yhtälörymillä on sama ratkaisujoukko, jos $A \sim B$. Lisäksi on huomattava, että matriisien rivioperaatiot ovat kääntyviä, mikä tarkoittaa sitä, että jos $A \sim B$, niin myös $B \sim A$.

Strategiana on soveltaa alkuperäisen yhtälöryhmän augmentoituun matriisiin (homogeenisessa tapauksessa kerroinmatriisiin) rivioperaatioita, kunnes päästään niin yksinkertaiseen muotoon, että ratkaisu voidaan lukea suoraan. Tyypillinen päämäärä on *reduoitu porrasmuoto*.

Määritelmä 2 Matriisi A on *porrasmuodossa*, jos sen jokainen rivi alkaa nolilla, joita on enemmän kuin millään ylemmällä rivillä. Ensimmäisen rivin ei tarvitse alkaa nolilla ja jostain rivistä alkaen rivit voivat koostua kokonaan nolista.

Määritelmä 3 Matriisi A on *reduoidussa porrasmuodossa*, jos

- A on porrasmuodossa
- A :n jokaisen rivin ensimmäinen nollasta poikkeava alkio on 1.
- A :n jokaisen rivin ensimmäisen nollasta poikkeavan alkion yläpuolella on vain nollia.

Määritelmä 4 Matriisin A aste $r(A)$ on A :sta saatavan redusoidun porrasmatriisin nollasta eroavien rivien määrä (ns. porraskluku)

Esimerkki 1 Redusoitu porrasmatriisi on siis muotoa

$$\begin{pmatrix} \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \end{pmatrix},$$

missä pisteet merkitsevät nollia ja asteriskit mitä tahansa lukuja. Pelkkä porrasmatriisi taas tarkoittaa sitä, että ykkösten sijasta saa esiintyä mitä hyvänsä nollasta poikkeavia lukuja ja niiden yläpuolella mitä hyvänsä lukuja.

Porrasmuodon saavuttamiseksi voidaan käyttää seuraavaa menetelmää (algoritmia):

1. Jos kaikki rivit ovat nollarivejä, lopeta. Samoin jos rivejä on vain yksi, on A valmiiksi porrasmuotoa ja voidaan lopettaa.
2. Jos A :ssa enemmän kuin yksi rivi, etsi rivi $\mathbf{c} = (0, \dots, c, *, \dots, *)$, jolla on vähiten alkunollia ja vaihda se ylimmäksi. Jos tällaisia rivejä on useampia, valitse niistä mikä hyvänsä. Näin saadaan matriisi

$$A_1 = \begin{pmatrix} 0 & \dots & c & * & \dots & * \\ 0 & \dots & * & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & * & * & \dots & * \end{pmatrix},$$

jossa $c \neq 0$ ja asteriskit ovat mitä hyvänsä lukuja.

3. Jokaista ylimmän rivin alapuolella olevaa riviä $\mathbf{d} = (0, \dots, d, \dots, *)$ kohti toimi seuraavasti (d on siis alkio, joka on ylimmän rivin alkion c alapuolella): lisää ylin rivi tähän riviin kerrottuna vakiolla $-c^{-1}d$. Tällöin d siis korvautuu luvulla $d - c \cdot c^{-1}d = 0$ ja näin saadaan matriisi

$$A_2 = \begin{pmatrix} 0 & \dots & c & * & \dots & * \\ 0 & \dots & 0 & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & * & \dots & * \end{pmatrix} \stackrel{\text{merk.}}{=} \begin{pmatrix} 0 & \dots & c & * & \dots & * \\ 0 & \dots & 0 & & & \\ \vdots & \ddots & \vdots & & B & \\ 0 & \dots & 0 & & & \end{pmatrix},$$

jossa c :n alapuolella on vain nollia.

4. Sovella tätä samaa menetelmää matriisiin B .

Edellä kuvattu algoritmia kutsutaan *Gaussin menetelmäksi* ja se päättyy varmasti, sillä matriisi B , johon menetelmää sovelletaan rekursiivisesti uudelleen, on riviluvultaan pienempi kuin alkuperäinen matriisi. Jossain vaiheessa siis päädytään kokonaan nolliasta koostuvaan tai yksiriviseen matriisiin, jolloin menetelmä päättyy. Menetelmän perusteella on lisäksi selvää, tuloksena on todella porrasmatriisi.

Porrasmatriisista saadaan redusoitu muoto täydentämällä Gaussin menetelmää seuraavasti:

1. Kerro jokainen rivi ensimmäisen nollasta poikkeavan luvun (jos olemassa) käänteisluvulla. Tämän jälkeen jokainen rivi, joka ei kokostu kokonaan nolliasta, alkaa ykkösellä.
2. Jos rivin i aloittavan ykkösen yläpuolella on $d \neq 0$ rivillä j , lisää rivi i riviin $j - d$:llä kerrottuna. Toista tätä uudelleen niin kauan rivien alkuykkösten yläpuolelta löytyy nolasta poikkeavia alkioita.

Esimerkki 2 Ratkaistaan yhtälöryhmä

$$\begin{cases} 2x - y + 3z = 4 \\ 3x + 2y + z = 6 \\ -5x + y - 2z = 3 \end{cases}.$$

Ryhmän augmentoitu matriisi on $\begin{pmatrix} 2 & -1 & 3 & 4 \\ 3 & 2 & 1 & 6 \\ -5 & 1 & -2 & 3 \end{pmatrix}$. Kertomalla ensimmäinen rivi vakiolla $\frac{1}{2}$ saadaan

$\begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 2 \\ 3 & 2 & 1 & 6 \\ -5 & 1 & -2 & 3 \end{pmatrix}$, missä edelleen voidaan lisätä ensimmäinen rivi toiseen -3 :lla kerrottuna ja kolman-

teen 5 :llä kerrottuna. Tuloksena saadaan matriisi $\begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 2 \\ 0 & \frac{7}{2} & -\frac{7}{2} & 0 \\ 0 & -\frac{3}{2} & \frac{11}{2} & 13 \end{pmatrix}$. Kerrotaan näin saadun matriisin

toinen rivi $\frac{2}{7}$:llä. Näin saadaan $\begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 2 \\ 0 & 1 & -1 & 0 \\ 0 & -\frac{3}{2} & \frac{11}{2} & 13 \end{pmatrix}$, minkä jälkeen lisätään toinen rivi ensimmäiseen $\frac{1}{2}$:lla

kerrottuna ja kolmanteen $\frac{3}{2}$:lla kerrottuna. Tuloksena saatava matriisi on $\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 4 & 13 \end{pmatrix}$, josta kolmas

rivi $\frac{1}{4}$:lla kertomalla saadaan $\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & \frac{13}{4} \end{pmatrix}$. Lisäämällä kolmas rivi ensimmäiseen -1 :llä kerrottuna

ensimmäiseen ja 1:llä kerrottuna toiseen saadaan matriisi $\begin{pmatrix} 1 & 0 & 0 & -\frac{5}{4} \\ 0 & 1 & 0 & \frac{13}{4} \\ 0 & 0 & 1 & \frac{13}{4} \end{pmatrix}$, jota vastaava yhtälöryhmä on

$$\begin{cases} x &= -\frac{5}{4} \\ y &= \frac{13}{4} \\ z &= \frac{13}{4} \end{cases},$$

mistä ratkaisu voidaan suoraan lukea: $(x, y, z) = (-\frac{5}{4}, \frac{13}{4}, \frac{13}{4})$. Tässä esimerkissä on siis tasan yksi ratkaisu.

Esimerkki 3 Yhtälöryhmää

$$\begin{cases} x + 3y + 2z = 7 \\ 2x - y - z = 5 \\ -x + 2y + 3z = 4 \end{cases}$$

vastaava augmentoitu matriisi on $\begin{pmatrix} 1 & 3 & 2 & 7 \\ 2 & 1 & -1 & 5 \\ -1 & 2 & 3 & 4 \end{pmatrix}$, mistä alkeisoperaatioilla saadaan matriisi $\begin{pmatrix} 1 & 3 & 2 & 7 \\ 0 & 1 & 1 & \frac{9}{5} \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

Viimeisintä matriisia vastaavan yhtälöryhmän kolmas yhtälö on muotoa $0 \cdot z + 0 \cdot y + 0 \cdot x = 2$, siis $0 = 2$, mikä ei voi toteutua millään muuttujien x , y ja z arvoilla. Koska viimeisimmällä ryhmällä ei ole ratkaisua, ei sellaista voi olla ensimmäiselläkään, sillä ryhmiä vastaavat (augmentoidut) matriisit ovat rivekvivalentteja.

Esimerkki 4 Ratkaistaan yhtälöpari

$$\begin{cases} 2x - y + z = 7 \\ x + 3y + 4z = 0 \end{cases}.$$

Augmentoitu matriisi on

$$\begin{pmatrix} 2 & -1 & 1 & 7 \\ 1 & 3 & 4 & 0 \end{pmatrix},$$

josta Gaussin-Jordanin menetelmällä saadaan matriisi

$$\begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & -1 \end{pmatrix}.$$

Tämä vastaa yhtälöparia

$$\begin{cases} x + z = 3 \\ y + z = -1 \end{cases},$$

jonka ratkaisut ovat $(x, y, z) = (-z + 3, -z - 1, z) = (-z, -z, z) + (3, -1, 0) = z(-1, -1, 1) + (3, -1, 0)$, missä $z \in \mathbb{R}$. Ratkaisuja on siis ääretön määrä.

1.1.2 Ratkaisujoukon systemaattinen esittäminen

Tarkastellaan lähinnä esimerkkien valossa miten redusoidun porrasmatriisin perusteella voidaan yhtälöryhmän kaikki ratkaisut esittää systemaattisessa muodossa. Ensinnäkin esimerkissä 3 todettiin, että mikäli augmentoitu matriisi redusoituu sellaiseen muotoon, jossa viimeisin rivi esittää yhtälöä $0 = c \neq 0$, ei yhtälöryhmällä ole yhtään ratkaisua.

Mikäli ratkaisuja on, esitetään ne systemaattisessa muodossa siten, että redusoidussa porrasmuodossa olevan yhtälöryhmän ”portaan aloittavat” muuttujat esitetään muiden avulla, kun taas muut muuttujat ”jäävät ratkaisuun”.

Esimerkki 5 Olkoon augmentoidusta yhtälöryhmän matriisista saatu redusoitua matriisi

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

Tätä matriisia vastaa yhtälöryhmä

$$\begin{cases} x_1 + 2x_3 = 4 \\ x_2 - x_3 = -3 \\ x_4 = -1 \end{cases},$$

jonka kaikki ratkaisut (x_1, x_2, x_3, x_4) voidaan suoraviivaisesti lukea redusoidusta porrasmuodosta: Ensimmäisen yhtälön perusteella $x_1 = -2x_3 + 4$, toisen yhtälön perusteella $x_2 = x_3 - 3$, ja kolmannen yhtälön perusteella $x_4 = -1$. Täten siis yhtälöryhmän yleinen ratkaisu on muotoa $(x_1, x_2, x_3, x_4) = (-2x_3 + 4, x_3 - 3, x_3, -1) = (-2x_3, x_3, x_3, 0) + (4, -3, 0, -1) = x_3(-2, 1, 1, 0) + (4, -3, 0, -1)$, missä x_3 on mikä hyvänsä reaaliluku.

Esimerkki 6 Tarkastellaan toisena esimerkkinä redusoitua (augmentoitua) porrasmatriisia

$$\begin{pmatrix} 1 & 0 & 0 & -4 & -6 & -2 \\ 0 & 1 & 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 & 7 \end{pmatrix},$$

jota vastaava yhtälöryhmä on

$$\begin{cases} x_1 - 4x_4 - 6x_5 = -2 \\ x_2 + 3x_4 = 3 \\ x_3 + x_4 + 2x_5 = 7 \end{cases}$$

Nyt muuttujat x_1, x_2 ja x_3 ovat kukin ”portaansa aloittavia” muuttujia, joten ne esitetään muiden avulla. Ensimmäisen yhtälön perusteella $x_1 = 4x_4 + 6x_5 - 2$, toisen perusteella $x_2 = -3x_4 + 3$, ja kolmannen perusteella $x_3 = -x_4 - 2x_5 + 7$. Täten siis yleinen ratkaisu $(x_1, x_2, x_3, x_4, x_5)$ voidaan kirjoittaa muodossa

$$\begin{aligned} (x_1, x_2, x_3, x_4, x_5) &= (4x_4 + 6x_5 - 2, -3x_4 + 3, -x_4 - 2x_5 + 7, x_4, x_5) \\ &= (4x_4 + 6x_5, -3x_4, -x_4 - 2x_5, x_4, x_5) + (-2, 3, 7, 0, 0) \\ &= (4x_4, -3x_4, -x_4, x_4, 0) + (6x_5, 0, -2x_5, 0, x_5) + (-2, 3, 7, 0, 0) \\ &= x_4(4, -3, -1, 1, 0) + x_5(6, 0, -2, 0, 1) + (-2, 3, 7, 0, 0), \end{aligned}$$

missä x_4 ja x_5 ovat mitä hyvänsä reaalilukuja.

Esimerkissä 5 yleinen ratkaisu saatiin muodossa $x_3(-2, 1, 1, 0) + (4, -3, 0, -1)$, ja esimerkin 6 tapauksessa yleinen ratkaisu on muotoa $x_4(4, -3, -1, 1, 0) + x_5(6, 0, -2, 0, 1) + (-2, 3, 7, 0, 0)$. Kummassakin tapauksessa ratkaisu voidaan esittää sellaisessa muodossa, jossa äärettömään vektorijoukkoon lisätään yksittäinen vektori: Esimerkin 5 tapauksessa kyseinen yksittäinen vektori on $(4, -3, 0, -1)$ ja esimerkin 6 tapauksessa $(2, 3, 7, 0, 0)$. Lisättyä yksittäistä vektoria kutsutaan *yksittäisratkaisuksi* tai *yksityisratkaisuksi*.

Huomautus 1 Systemaattisen esityksen vapaiden muuttujien määrää voidaan laskea seuraavasti: Oletetaan, että muuttujia on alun perin n kappaletta, ja että yhtälöryhmän kerroinmatriisin redusoidussa porrasmuodossa on $r = r(A)$ porrasta. Jokainen ”portaan aloittava muuttuja” esitetään muiden muuttujien avulla, joten ratkaisuun jää $n - r$ ”vapaata” muuttujaa, joiden arvoa ei kiinnitetä. Näin ollen ratkaisu on muotoa

$$\mathbf{x} = x_1 \mathbf{c}_1 + \dots + x_{n-r} \mathbf{c}_{n-r} + \mathbf{c},$$

missä $\mathbf{c}_1, \dots, \mathbf{c}_{n-r}$ ja $\mathbf{c} \in \mathbb{R}^n$ ja $x_1, \dots, x_{n-r}$ ovat vapaasti valittavia lukuja.

Gaussin-Jordaniin menetelmä on valmiiksi ohjelmoitu useimpiin tavallisiin matematiikkaohjelmiin. Syötetään esimerkin 2 augmentoitu matriisi **Matlab**-ohjelmistoon:

```
>> A=[2, -1, 3, 4; 3, 2, 1, 6; -5, 1, -2, 3]
```

Komento `format rat` asettaa Matlabin laskemaan desimaaliesitysten sijaan murtoluvuilla ja `rref(A)` muuntaa matriisin A redusoituihin porrasmuotoon:

```
>> rref(A)

ans =

     1         0         0    -5/4
     0         1         0    13/4
     0         0         1    13/4

>>
```

1.2 Vektoriavaruudet

Edellisessä pykälässä käytettiin käsitettä matriisi, joka kieltämättä kuuluu lineaarialgebran kulmakiiviin. Toisaalta matriisien ominaisuuksia käytettiin vain minimaalisesti eikä edes mainittu matriisien kertolaskua, joka on perustavaa laatua oleva operaatio lineaarialgebrassa. Siksi käsite matriisi esitellään vielä tulevissa pykälissä uudelleen, tällä kertaa selvittäen myös käsitteen taustat.

Määritelmä 5 Jos A on mikä hyvänsä joukko, merkitään A :n n -kertaista karteesista tuloa itsensä kanssa (karteesinen potenssi) A^n :llä. Se muodostuu n -pituisista järjestetyistä jonoista, joiden kukin alkio on joukon A jäsen:

$$A^n = \{(a_1, a_2, \dots, a_n) \mid a_i \in A\}.$$

Joukon A^n alkioita kutsutaan *vaakavektoreiksi* tai $1 \times n$ -matriiseiksi. Vektoreita on tapana merkitä lihavoidulla kirjasintyyppillä: $\mathbf{a} = (a_1, a_2, \dots, a_n)$ ja alkioita $a_1, a_2, \dots, a_n$ kutsutaan vektorin *koordinaateiksi* tai pelkästään alkioiksi, toisinaan myös komponenteiksi. Fysiikan kirjallisuudessa vektoreista käytetään usein myös yläviivaa $\vec{a}$, ”hattua” $\hat{a}$ tai ylänuolta $\overrightarrow{a}$. Vanhemmassa suomenkielisessä kirjallisuudessa on suosittu myös saksalaisperäistä käytäntöä, jossa vektoreita merkitään fraktuurakirjaimilla $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ jne.

Jos joukossa A ei ole määritelty mitään operaatioita (esim. summa, tulo), jää yllä määritelty karteesinen potenssi A^n pelkästään joukoksi vailla minkäänlaista algebrallista rakennetta. Tämä ei ole lineaarialgebran kannalta erityisen kiintoisa tarkastelukohde, ja tämän kurssin oppimäärän kannalta tärkeimmät tapaukset ovatkin $A = \mathbb{R}$ ja $A = \mathbb{C}$, jolloin puhutaan reaalisesta tai kompleksisesta vektoriavaruudesta. Tällöin A :n alkioita kutsutaan skalaareiksi.

Määritelmä 6 Olkoon $A = \mathbb{R}$ tai $A = \mathbb{C}$. Vektorien $\mathbf{a} = (a_1, \dots, a_n)$ ja $\mathbf{b} = (b_1, \dots, b_n)$ yhteenlasku määritellään

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, \dots, a_n + b_n)$$

ja skalaarikertolasku

$$c\mathbf{a} = (ca_1, \dots, ca_n)$$

Lause 1 Olkoon $A = \mathbb{R}$ tai $A = \mathbb{C}$. Määritellään ns. nollavektori $\mathbf{0} = (0, 0, \dots, 0)$ ja vektorin $\mathbf{a} = (a_1, \dots, a_n)$ vastavektoriksi $-\mathbf{a} = (-a_1, \dots, -a_n)$. Tällöin joukko A^n varustettuna yllä määritellyllä vektorien yhteenlaskulla ja skalaarikertolaskulla toteuttaa seuraavat ehdot:

1. $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$
2. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
3. $\mathbf{a} + \mathbf{0} = \mathbf{a}$
4. $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$
5. $c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}$
6. $(c + d)\mathbf{a} = c\mathbf{a} + d\mathbf{a}$
7. $c(d\mathbf{a}) = (cd)\mathbf{a}$
8. $1\mathbf{a} = \mathbf{a}$

Edellisen lauseen kaikki väittämät on erittäin helppo havaita todeksi käyttämällä reaali- tai kompleksilukujen perusominaisuuksia. Suoraviivaisuudestaan huolimatta lauseen väittämällä on erittäin merkittävä asema lineaarialgebran teoriassa. Vaikka päättelyä tässä kurssissa ei erityisesti perusteta edellisen lauseen väittämiin, on niiden erityisasema syytä tuoda esille:

Määritelmä 7 Vektoriavaruus V on joukko, jossa on määritelty vektoreiden yhteenlasku ja skalaarikertolasku, jotka toteuttavat edellisen lauseen ehdot 1–8. Vektoriavaruuksia kutsutaan myös lineaariavaruuksiksi tai pelkästään avaruuksiksi.

Edellisen määritelmän mukaan $\mathbb{R}^n$ ja $\mathbb{C}^n$ ovat molemmat vektoriavaruuksia, mutta nämä eivät suinkaan ole ainoita esimerkkejä sellaisesta.

Esimerkki 7 Olkoon $C^0[a, b]$ välillä $[a, b]$ määriteltyjen jatkuvien reaalfunktioiden joukko. Määritellään funktioiden yhteenlasku $f + g$ ehdolla $(f + g)(x) = f(x) + g(x)$ ja skalaarikertolasku $c \cdot f$ ehdolla $(cf)(x) = cf(x)$. Totea, että $C^0[a, b]$ toteuttaa lauseessa 1 esitetyt ehdot.

Huomautus 2 Edellisen esimerkin $C^0[a, b]$ ja vektoriavaruuden $\mathbb{R}^n$ välillä on itse asiassa varsin suoraviivainen yhteys. Joka ikinen avaruuden $\mathbb{R}^n$ alkio $\mathbf{f} = (f_1, \dots, f_n)$ voidaan merkinnän $\mathbf{f} = (f(1), \dots, f(n))$ kautta itse asiassa ajatella funktioksi $\mathbf{f} : \{1, \dots, n\} \mapsto \mathbb{R}$, missä $\mathbf{f}(i) = f_i$. Näin ollen vektoriavaruuden $\mathbb{R}^n$ alkio on äärellisellä joukolla $\{1, 2, \dots, n\}$ määritelty funktio joukkoon $\mathbb{R}$, kun taas avaruuden $C^0[a, b]$ alkio on välillä $[a, b]$ määriteltyä funktioita joukkoon $\mathbb{R}$. Voidaan jopa ajatella, että avaruuden $\mathbb{R}^n$ vektorilla $\mathbf{f}$ on n koordinaattia $f_1, \dots, f_n$, kun taas avaruuden $C^0[a, b]$ vektorilla on äärettömän monta koordinaattia $f(t)$ kutakin arvoa $t \in [a, b]$ kohti.

Esimerkki 8 Avaruuksilla $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ ja $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$ on kummallakin ilmeinen visuaalinen tulkinta. Vektori voidaan tulkita geometrisesti kolmi- tai pienempiulotteisen avaruuden pisteinä. Tätä tulkintaa on yleensä tapana laajentaa suuntajanaiksi origosta kyseiseen pisteeseen ja jopa kaikkien samansuuntaisten ja -pituisten suuntajanojen joukoksi. Tulkinnan mukaan $(1, 0, 0)$ on x -akselin suuntainen, $(0, 1, 0)$ y -akselin suuntainen ja $(0, 0, 1)$ z -akselin suuntainen yksikkövektori.

1.3 Vektoriavaruuksien rakenteesta

Ellei toisin mainita, tässä luvussa käsiteltävät vektoriavaruudet ovat joko $V = \mathbb{R}^n$ tai $V = \mathbb{C}^n$. Rakenteen kannalta näiden teoria on samankaltainen, mutta jos käsitellään avaruutta $\mathbb{R}^n$, ovat skalaarit rajoitettu joukkoon $\mathbb{R}$, ja avaruutta $\mathbb{C}^n$ käsiteltäessä otaksutaan skalaarikunnaksi $\mathbb{C}$. Periaatteessa olisi kyllä mahdollista kehittää kompleksisten vektoriavaruuksien teoriaa rajoittamalla skalaarit joukkoon $\mathbb{R}$, mutta tämä ei ole osoittautunut erityisen hedelmälliseksi lähestymistavaksi.

Esimerkki 9 Olkoon $V = \mathbb{R}^3$. Erityisesti fysiikassa on tapana merkitä $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$, jolloin

$$(x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Edellisen esimerkin mukaan kaikki avaruuden $\mathbb{R}^3$ vektorit voidaan esittää kolmen vektorin $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ skalaarimonikertojen ja summien avulla. Tällaista esitystapaa sanotaan *lineaarikombinaatioksi*.

Määritelmä 8 Jos V on vektoriavaruus jonka skalaarikunta on $\mathbb{K}$ (joko $\mathbb{R}$ tai $\mathbb{C}$), ja vektorit $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ovat joitain avaruuden V alkioita, sanotaan, että näiden vektoreiden *linearikombinaatioiden* joukko

$$L(\mathbf{v}_1, \dots, \mathbf{v}_k) = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle = \{c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k \mid c_1, \dots, c_k \in \mathbb{K}\}$$

on niiden vektoreiden joukko, jotka saadaan vektoreista $\mathbf{v}_1, \dots, \mathbf{v}_k$ vektoriavaruuden algebrallisilla operaatioilla: skalaarikertolaskuilla ja vektoreiden yhteenlaskuilla.

Vektoriavaruuksien teoriaa kehitettäessä voidaan huomata, että monissa käyttökohteissa, sekä teoreettisissa, että soveltavissa, usein vektorijoukkoa $A = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ tärkeämpi on joukko $L(\mathbf{v}_1, \dots, \mathbf{v}_k)$, siis kaikki vektorit jotka voidaan esittää joukon A lineaarikombinaatioina.

Esimerkki 10 Määritellään $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, \dots, 0)$, $\dots$, $\mathbf{e}_n = (0, 0, \dots, 1)$. Tällöin

$$(x_1, x_2, \dots, x_n) = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n,$$

joten

$$\mathbb{R}^n = L(\mathbf{e}_1, \dots, \mathbf{e}_n).$$

Määritelmä 9 Jos

$$U = L(\mathbf{v}_1, \dots, \mathbf{v}_k)$$

sanotaan, että vektorit $\mathbf{v}_1, \dots, \mathbf{v}_k$ generoivat joukon U .

Lause 2 Seuraavat tulokset on melko helposti nähtävissä oikeiksi.

1. Kaikilla $\mathbf{v}_i$ on $\mathbf{v}_i \in L(\mathbf{v}_1, \dots, \mathbf{v}_k)$.
2. Jos $\{\mathbf{v}_1, \dots, \mathbf{v}_l\} \subseteq \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, on $L(\mathbf{v}_1, \dots, \mathbf{v}_l) \subseteq L(\mathbf{v}_1, \dots, \mathbf{v}_k)$.
3. Jos $\{\mathbf{u}_1, \dots, \mathbf{u}_l\} \subseteq L(\mathbf{v}_1, \dots, \mathbf{v}_k)$, niin $L(\mathbf{u}_1, \dots, \mathbf{u}_l) \subseteq L(\mathbf{v}_1, \dots, \mathbf{v}_k)$.
4. $\mathbf{u} \in L(\mathbf{v}_1, \dots, \mathbf{v}_k)$, tarkalleen silloin kun $L(\mathbf{v}_1, \dots, \mathbf{v}_k) = L(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u})$.
5. $L(\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \cup \{\mathbf{0}\}) = L(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

Määritelmä 10 Olkoon V on vektoriavaruus ja $U \subseteq V$ sellainen epätyhjä joukko, että aina kun $\mathbf{v}_1, \mathbf{v}_2 \in U$ ja c_1 ja c_2 ovat skalaareja, niin myös $c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 \in U$. Tällöin sanotaan, että U on V :n aliavaruus ja merkitään $U \leq V$

Edellisen määritelmän mukaan aliavaruus on sellainen vektoriavaruuden osajoukko, joka on suljettu vektoriavaruuden operaatioiden (summa ja skalaarimonikerta suhteen) ja siis $L(\mathbf{v}_1, \dots, \mathbf{v}_k)$ on aina V :n aliavaruus. Jokaisella vektoriavaruudella V on ainakin kaksi aliavaruutta: $\{\mathbf{0}\}$ (nolla-avaruus) ja V itse.

Huomautus 3 Koska $\mathbf{0} = 0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2$, on $\mathbf{0} \in V$ aina.

Määritelmän 9 mukainen joukko on aina algebralliselta rakenteeltaan vektoriavaruus:

Lause 3 Jos $\mathbf{v}_1, \dots, \mathbf{v}_k$ ovat jonkin vektoriavaruuden V alkioita, on joukko $L(\mathbf{v}_1, \dots, \mathbf{v}_k)$ vektoriavaruus, joka sisältyy avaruuteen V .

Todistus Suoraviivainen. □

Määritelmä 11 Vektoriavaruus V on äärellisesti generoitu, jos on olemassa sellainen äärellinen joukko $\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V$, että

$$V = L(\mathbf{v}_1, \dots, \mathbf{v}_k)$$

Esimerkki 11 Vektoriavaruus $\mathbb{R}^n$ (samoin kuin $\mathbb{C}^n$) on äärellisesti generoitu:

$$\mathbb{R}^n = L(\mathbf{e}_1, \dots, \mathbf{e}_n).$$

Esimerkki 12 Olkoon $\mathbf{v} = (1, 1) \in \mathbb{R}^2$, jolloin $\langle (1, 1) \rangle = L((1, 1)) = \{t(1, 1) \mid t \in \mathbb{R}\} = \{(t, t) \mid t \in \mathbb{R}\}$

Esimerkki 13 Vektoriavaruutta $\mathbb{R}^2$ on luontevaa sanoa kaksiulotteiseksi ja avaruutta $\mathbb{R}^3$ kolmiulotteiseksi. On helppo huomata että $\mathbb{R}^2 = L(\mathbf{i}, \mathbf{j})$ ja $\mathbb{R}^3 = L(\mathbf{i}, \mathbf{j}, \mathbf{k})$. Esimerkin 12 mukaista avaruutta on luontevaa kutsua yksiulotteiseksi, joten ulottuvuusluku näyttää liittyvän generoitvien vektoreiden määrään. Aivan suoraviivainen ei kytkös ole, sillä esimerkiksi jokainen $\mathbb{R}^2$:n vektori voidaan esittää paitsi muodossa

$$(x, y) = x\mathbf{i} + y\mathbf{j},$$

myös muodossa

$$(x, y) = x\mathbf{i} + y\mathbf{j} + 0 \cdot (\mathbf{i} + \mathbf{j}),$$

joten myös kolmen vektorin joukko $\{\mathbf{i}, \mathbf{j}, \mathbf{i} + \mathbf{j}\}$ generoi avaruuden $\mathbb{R}^2$. Toisaalta myös joukko $\{\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}\}$ generoi avaruuden $\mathbb{R}^2$, sillä

$$(x, y) = \frac{1}{2}(x + y)(\mathbf{i} + \mathbf{j}) + \frac{1}{2}(x - y)(\mathbf{i} - \mathbf{j}).$$

Täten siis avaruudelle $\mathbb{R}^2$ on löytynyt kaksi kahden alkion ja yksi kolmen alkion generoiva joukko. On kuitenkin huomattava että kolmen alkion generoivasta joukosta $\{\mathbf{i}, \mathbf{j}, \mathbf{i} + \mathbf{j}\}$ voidaan yksi alkio jättää pois, sillä esimerkiksi viimeinen alkio $\mathbf{i} + \mathbf{j}$ voidaan esittää kahden ensimmäisen avulla. Yhtä hyvin voitaisiin jättää pois ensimmäinen alkio $\mathbf{i}$, sillä se voidaan esittää kahden jälkimmäisen lineaarikombinaationa. $\mathbf{i} = -1 \cdot \mathbf{j} + 1 \cdot (\mathbf{i} + \mathbf{j})$.

Määritelmä 12 Vektorijoukko $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ on *lineaarisesti riippuva* tai pelkästään *riippuva*, jos jokin sen vektoreista voidaan esittää muiden lineaarikombinaationa: $\mathbf{v}_i \in L(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k)$. Jos vektorijoukko ei ole lineaarisesti riippuva, sanotaan että se on lineaarisesti riippumaton (tai pelkästään riippumaton).

Esimerkki 14 Joukko $\{\mathbf{i}, \mathbf{j}\}$ on selvästi lineaarisesti riippumaton, sillä $(1, 0) = c(0, 1)$ ja $(0, 1) = c(1, 0)$ eivät toteudu millään c :n arvolla. Joukko $\{\mathbf{i}, \mathbf{j}, \mathbf{i} + \mathbf{j}\}$ on lineaarisesti riippuva, sillä $\mathbf{i} + \mathbf{j} = 1 \cdot \mathbf{i} + 1 \cdot \mathbf{j}$.

Huomautus 4 Jokainen vektorijoukko $A = \{\mathbf{0}, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ joka sisältää nollavektorin $\mathbf{v}_1 = \mathbf{0}$, on lineaarisesti riippuva, sillä $\mathbf{0}$ voidaan esittää muiden lineaarikombinaationa:

$$\mathbf{0} = 0 \cdot \mathbf{v}_2 + \dots + 0 \cdot \mathbf{v}_k.$$

Lause 4 Vektorijoukko $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ on lineaarisesti riippumaton tarkalleen silloin kun

$$c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k = \mathbf{0}$$

toteutuu vain ilmeisellä valinnalla jossa kaikki kertoimet ovat nollija: $c_1 = \dots = c_k = 0$.

Todistus Oletetaan ensin, että vektorijoukko $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ on lineaarisesti riippuva. Tällöin jokin vektoreista, $\mathbf{v}_i$ voidaan esittää muiden lineaarikombinaationa:

$$\mathbf{v}_i = c_1 \mathbf{v}_1 + \dots + c_{i-1} \mathbf{v}_{i-1} + c_{i+1} \mathbf{v}_{i+1} + \dots + c_k \mathbf{v}_k,$$

josta nähdään, että

$$\mathbf{0} = c_1 \mathbf{v}_1 + \dots + c_{i-1} \mathbf{v}_{i-1} - 1 \cdot \mathbf{v}_i + c_{i+1} \mathbf{v}_{i+1} + \dots + c_k \mathbf{v}_k.$$

Näin ollen nollavektori voidaan esittää epäilmeisellä tavalla, sillä vektorin $\mathbf{v}_i$ kerroin ylläolevassa esityksessä on -1 .

Oletetaan sitten, että nollavektori voidaan esittää epäilmeisellä tavalla

$$c_1 \mathbf{v}_1 + \dots + c_k \mathbf{v}_k = \mathbf{0},$$

missä ainakin jokin kertoimista, $c_i \neq 0$. Tällöin

$$c_i \mathbf{v}_i = -c_1 \mathbf{v}_1 - \dots - c_{i-1} \mathbf{v}_{i-1} - c_{i+1} \mathbf{v}_{i+1} - \dots - c_k \mathbf{v}_k.$$

ja siis vektori $\mathbf{v}_i$ voidaan esittää muiden lineaarikombinaationa:

$$\mathbf{v}_i = -c_i^{-1} c_1 \mathbf{v}_1 - \dots - c_i^{-1} c_{i-1} \mathbf{v}_{i-1} - c_i^{-1} c_{i+1} \mathbf{v}_{i+1} - \dots - c_i^{-1} c_k \mathbf{v}_k.$$

Esimerkki 15 Vektorijoukko $\{\mathbf{i}, \mathbf{j}, \mathbf{i} + \mathbf{j}\}$ on lineaarisesti riippuva, sillä

$$1 \cdot \mathbf{i} + 1 \cdot \mathbf{j} - 1 \cdot (\mathbf{i} + \mathbf{j}) = \mathbf{0}$$

on nollavektorin ei-ilmeinen esitys.

Esimerkki 16 Vektorijoukko $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ avaruudessa $\mathbb{R}^n$ on lineaarisesti riippumaton, sillä

$$c_1 \mathbf{e}_1 + c_2 \mathbf{e}_2 + \dots + c_n \mathbf{e}_n = (c_1, c_2, \dots, c_n).$$

jotta tämä olisi nollavektori, pitää olla $c_1 = c_2 = \dots = c_n = 0$.

Esimerkki 17 Jokainen vektorijoukko $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, jossa vektorissa $\mathbf{v}_{i+1}$ on enemmän alkunollia kuin vektorissa $\mathbf{v}_i$, mutta joka ei sisällä nollavektoria, on lineaarisesti riippumaton. Jos nimittäin $\mathbf{v}_1 = (0, \dots, v_1, \dots)$ missä $v_1 \neq 0$, on

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k = (0, \dots, c_1 v_1, \dots)$$

ja jotta tämä olisi nollavektori, pitää ensiksikin olla $c_1 v_1 = 0$, josta $c_1 = 0$. Tällöin summan

$$c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k$$

tulee olla nollavektori, ja samoin päättelämällä nähdään, että $c_2 = 0$. Jatkamalla samoin saadaan $c_3 = \dots = c_k = 0$.

Esimerkki 18 Todetaan edellisen esimerkin mukaisesti, että avaruuden $\mathbb{R}^4$ joukko $\{(1, 2, 0, -1), (0, 2, -1, -2), (0, 0, 3, 2), (0, 0, 0, -1)\}$ on lineaarisesti riippumaton:

$$\begin{aligned} & c_1(1, 2, 0, -1) + c_2(0, 2, -1, -2) + c_3(0, 0, 3, 2) + c_4(0, 0, 0, -1) \\ &= (c_1, 2c_1 + 2c_2, -c_2 + 3c_3, -c_1 - 2c_2 + 2c_3 - c_4), \end{aligned}$$

ja jotta tämä olisi nollavektori, pitää olla sen kaikkien koordinaattien olla nolli. Ensinnäkin $c_1 = 0$, ja sijoittamalla tämä saadaan vektori $(0, 2c_2, -c_2 + 3c_3, -2c_2 + 2c_3 - c_4)$, joka voi olla nollavektori vain jos $c_2 = 0$. Sijoittamalla $c_2 = 0$ saadaan $(0, 0, 3c_3, 2c_3 - c_4)$, ja siis pitää olla $c_3 = 0$. Näin saadaan $(0, 0, 0, -c_4)$, josta $c_4 = 0$.

Esimerkki 19 Selvitetään onko avaruuden $\mathbb{R}^2$ vektorijoukko $\{\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}\} = \{(1, 1), (1, -1)\}$ lineaarisesti riippumaton. Yhtälö

$$c_1(1, 1) + c_2(1, -1) = (0, 0)$$

voidaan kirjoittaa yhtälöpariksi

$$\begin{cases} c_1 + c_2 = 0 \\ c_1 - c_2 = 0, \end{cases}$$

jonka ainoa ratkaisu on $(c_1, c_2) = (0, 0)$. Täten vektorit ovat riippumattomat.

Esimerkki 20 Selvitetään, onko avaruuden $\mathbb{R}^3$ vektorijoukko $\{(1, 2, -1), (-2, 3, 1), (4, 1, -3)\}$ lineaarisesti riippumaton. Tätä varten on selvitettävä, onko yhtälöllä

$$c_1(1, 2, -1) + c_2(-2, 3, 1) + c_3(4, 1, -3) = (0, 0, 0)$$

vain ratkaisu $c_1 = c_2 = c_3 = 0$ (jolloin vektorijoukko on lineaarisesti riippumaton) vai onko olemassa sellainen ratkaisu, jossa jokin $c_i \neq 0$ (jolloin vektorijoukko on lineaarisesti riippuva).

Yllä oleva vektoryhtälö voidaan kirjoittaa lukuja koskevaksi yhtälöryhmäksi

$$\begin{cases} c_1 - 2c_2 + 4c_3 = 0 \\ 2c_1 + 3c_2 + c_3 = 0 \\ -c_1 + c_2 - 3c_3 = 0, \end{cases}$$

joka voidaan Gaussin-Jordanin menetelmää käyttämällä saada ekvivalenttiin muotoon

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

Tämän ratkaisut ovat $(c_1, c_2, c_3) = (-2c_3, c_3, c_3) = c_3(-2, 1, 1)$, missä $c_3 \in \mathbb{R}$. Täten valitsemalla esimerkiksi $c_3 = 1$ saadaan nollavektoreille ei-ilmeinen esitys

$$-2 \cdot (1, 2, -1) + 1 \cdot (-2, 3, 1) + 1 \cdot (4, 1, -3) = (0, 0, 0),$$

josta voidaan mikä hyvänsä vasemman puolen vektoreista ratkaista esitettäväksi muiden lineaarikombinaationa.

Esimerkki 21 Yhden vektorin joukko on lineaarisesti riippuva tarkalleen silloin kun kyseinen vektori on $\mathbf{0}$. Kahden vektorin joukko $\{\mathbf{v}_1, \mathbf{v}_2\}$ on lineaarisesti riippuva tarkalleen silloin kun toinen vektoreista on nollavektori tai vektorit ovat toistensa skalaarimonikertoja.

Esimerkki 22 Olkoon C^0 koko reaaliakselilla määriteltyjen funktioiden joukko varustettuna esimerkin 7 mukaisella yhteen- ja skalaarikertolaskulla. Tällöin funktiot 1 , $\sin^2 x$ ja $\cos^2 x$ ovat lineaarisesti riippuvat, sillä

$$1 \cdot \sin^2 x + 1 \cdot \cos^2 x - 1 \cdot 1 = 0$$

Esimerkki 23 Olkoon C kuten edellisessä esimerkissä. Funktiot 1 , $\sin x$ ja $\cos x$ ovat lineaarisesti riippumattomat, sillä jos

$$c_1 \cdot 1 + c_2 \cdot \sin x + c_3 \cdot \cos x \tag{1.2}$$

on nollafunktio, niin myös sen derivaatta

$$c_2 \cos x - c_3 \sin x$$

on nollafunktio. Sijoittamalla tähän $x = 0$ nähdään, että $c_2 = 0$ ja sijoittamalla $x = \frac{\pi}{2}$ nähdään että $c_3 = 0$. Yhtälöstä (1.2) seuraa tällöin $c_1 = 0$.

Lause 5 Olkoot $\mathbf{v}_1, \dots, \mathbf{v}_k$ avaruuden $\mathbb{R}^n$ vektoreita. Muodostetaan $k \times n$ matriisi A , jonka rivejä mainitut vektorit ovat:

$$A = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_k \end{pmatrix}.$$

Jos $A \sim B$ ja

$$B = \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{pmatrix},$$

niin

$$L(\mathbf{v}_1, \dots, \mathbf{v}_k) = L(\mathbf{u}_1, \dots, \mathbf{u}_k)$$

Seuraus 1 Olkoon $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ vektorijoukko avaruudessa $\mathbb{R}^n$ ja A matriisi, jonka riveinä ovat kyseiset vektorit. Olkoon $A \sim B$ ja B redusoitu porrasmuoto. Tällöin vektorijoukko S on lineaarisesti riippuva tarkalleen silloin kun matriisissa B esiintyy nollarivi.

Esimerkki 24 Olkoon $S = \{(1, 2, -1), (-2, 3, 1), (4, 1, -3)\}$ ja A näitä riveinä käyttäen muodostettu matriisi:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ -2 & 3 & 1 \\ 4 & 1 & -3 \end{pmatrix}.$$

Tästä saatava redusoitu porrasmuoto on

$$B = \begin{pmatrix} 1 & 0 & -\frac{5}{7} \\ 0 & 1 & -\frac{1}{7} \\ 0 & 0 & 0 \end{pmatrix}.$$

Näin ollen vektorijoukko S on lineaarisesti riippuva, sillä B :ssä on nollarivi. Lisäksi

$$L((1, 2, -1), (-2, 3, 1), (4, 1, -3)) = L((1, 0, -\frac{5}{7}), (0, 1, -\frac{1}{7})).$$

Määritelmä 13 Joukko B on vektoriavaruuden V kanta, jos

1. $V = L(B)$.
2. B on lineaarisesti riippumaton.

Lause 6 Vektoriavaruuden kaikissa kannoissa on yhtä monta vektoria.

Lause 7 Jos $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ on avaruuden V kanta, niin jokaisella $\mathbf{x} \in V$ on olemassa yksikäsitteinen esitys kantavektoreiden lineaarikombinaationa:

$$\mathbf{x} = x_1 \mathbf{b}_1 + \dots + x_n \mathbf{b}_n$$

Määritelmä 14 Jos $B = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ on äärellinen, sanotaan että V on k -ulotteinen ja merkitään $\dim(V) = k$. Avaruuden $V = \{\mathbf{0}\}$ kannaksi sovitaan tyhjä joukko. Jos vektoriavaruudella V ei ole äärellistä kantaa, sanotaan, että V on ääretönulotteinen.

Esimerkki 25 Vektorit $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$ muodostavat $\mathbb{R}^3$:n kannan. Samoin vektorijoukko $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, missä $\mathbf{e}_i = (0, \dots, 1, \dots, 0)$ (i :s koordinaatti 1) on $\mathbb{R}^n$:n kanta. Tätä kantaa kutsutaan vektoriavaruuden $\mathbb{R}^n$ luonnolliseksi kannaksi.

Määritelmä 15 Jos $\mathbf{x} \in V$ ja $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ on avaruuden V kanta, niin esityksessä

$$\mathbf{x} = x_1 \mathbf{b}_1 + \dots + x_n \mathbf{b}_n$$

lukuja $x_1, \dots, x_n$ kutsutaan vektorin $\mathbf{x}$ koordinaateiksi kannan B suhteen ja niistä muodostuvaa $\mathbb{R}^n$:n vektoria $\mathbf{x}_B = (x_1, \dots, x_n)$ sanotaan $\mathbf{x}$:n koordinaattivektoriksi kannan B suhteen.

Esimerkki 26 Jos $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ ja $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ on luonnollinen kanta, niin

$$\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n,$$

joten

$$\mathbf{x}_B = (x_1, \dots, x_n).$$

Esimerkki 27 Olkoon $B_1 = \{\mathbf{i}, \mathbf{j}\}$ ja $B_2 = \{\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}\}$ sekä $(x, y) \in \mathbb{R}^2$. Esimerkin 13 mukaan $(x, y)_{B_1} = (x, y)$ ja $(x, y)_{B_2} = (\frac{1}{2}(x+y), \frac{1}{2}(x-y))$.

Esimerkki 28 Joukko $\{\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}\}$ on $\mathbb{R}^2$:n kanta, mutta $\{\mathbf{i}, \mathbf{j}, \mathbf{i} + \mathbf{j}\}$ ei ole.

Esimerkki 29 Olkoon $P = \{c_0 + c_1x + c_2x^2 + \dots + c_nx^n \mid c_i \in \mathbb{R}\}$ kaikkien reaalikertoimisten polynomien joukko. P on vektoriavaruus (tarkista), jonka kannaksi voidaan valita polynomit $1, x, x^2, x^3, \dots$. P :llä ei ole äärellistä kantaa.

Esimerkki 30 Olkoon $P_5 = \{c_0 + c_1x + c_2x^2 + c_3x^3 + c_4x^4 \mid c_i \in \mathbb{R}\}$ niiden reaalikertoimisten polynomien joukko, joiden aste on pienempi kuin 5. P_5 on vektoriavaruus (tarkista), jonka kannaksi voidaan valita $\{1, x, x^2, x^3, x^4\}$. Näin ollen $\dim(P_5) = 5$.

Esimerkki 31 Myös joukko $\{1, 1+x, 1+x+x^2, 1+x+x^2+x^3, 1+x+x^2+x^3+x^4\}$ on avaruuden P_5 kanta.

Huomautus 5 Jos $V = L(\mathbf{v}_1, \dots, \mathbf{v}_k)$, mutta $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ ei ole avaruuden V kanta, voidaan generoivasta joukosta poistaa ”turhat” vektorit näin saadaan kanta. Mikäli $\mathbf{v}_k \in L(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})$, on $V = L(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})$ ja prosessi voidaan toistaa kunnes jäljelle jäävä joukko on lineaarisesti riippumaton.

Esimerkki 32 Olkoon $V = L((1, 2, -1), (-2, 3, 1), (4, 1, -3))$ ja etsitään avaruudelle $V \subseteq \mathbb{R}^3$ kanta. Tämä voidaan tehdä esimerkin 24 mukaisesti tai hyödyntämällä esimerkkiä 20.

Lause 8 Jos V on äärellisulotteinen ja U on V :n aliavaruus, niin mikä hyvänsä U :n kanta voidaan täydentää V :n kannaksi.

Todistus Jos $U = L\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, valitaan mikä hyvänsä vektori $\mathbf{v}_{k+1} \in V \setminus U$, jolloin saatava aliavaruus $U_1 = L\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ sisältää aidosti avaruuden U . Jatketaan näin, kunnes ei enää voi valita vektoria aliavaruuden ulkopuolelta, mikä merkitsee sitä, että lopulta saatu avaruus on V . □

Seuraus 2 Jos $U \leq V$, on $\dim(U) \leq \dim(V)$.

Seuraus 3 Olkoon $\dim(V) = n$. Jokainen V :n osajoukko, jossa on enemmän kuin n vektoria, on lineaarisesti riippuva.

Luku 2

Matriisit

2.1 Kertausta

Aiemmissa luvussa ja myös jo insinöörimatematiikka A:ssa on viitattu matriisin käsitteeseen. Aiemmin esitettyä matriisi on kuitenkin käsitetty vain kaaviona, jonka puitteissa voidaan määritellä rivimanipulaatiota, tai alkiokohtaisia tulo- ja summaoperaatioita. Tässä luvussa esitetään syvällisempi, lineaarikuvauxiin perustuva merkitys matriisin käsitteelle. Aluksi kuitenkin kerrataan tarvittavat käsitteet.

Määritelmä 16 Matriisi A on kaavio

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}. \quad (2.1)$$

Matriisissa esiintyviä lukuja A_{ij} sanotaan *matriisin alkioiksi*. Jos $A_{ij} \in \mathbb{C}$, sanotaan, että matriisi on *kompleksinen* tai *yli kompleksilukujen*. Jos taas alkiot ovat reaalisia, voidaan sanoa että A on *reaalinen* tai *yli reaalilukujen*. Matriisin vaakarivejä kutsutaan *riveiksi* ja pystyivejä *sarakkeiksi*. Matriisin *tyyppi* on lukupari $m \times n$, missä m rivien määrä ja n on sarakkeiden määrä. Tyyppiä $m \times n$ olevaa matriisia sanotaan $m \times n$ -matriisiksi.

Määritelmä 17 (Vaaka- ja pystyvektorit)

- $1 \times n$ -matriisia $A = (A_{11} \ A_{12} \ \dots \ A_{1n})$ kutsutaan *vaakavektoriksi* tai *rivivektoriksi*.
- $m \times 1$ -matriisia $A = \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{m1} \end{pmatrix}$ kutsutaan *pystyvektoriksi* tai *sarakevektoriksi*.

Rivi- ja sarakevektoreita kutsutaan yhteisnimityksellä *vektori*. Vektoreiden alkioita kutsutaan myös *koordinaateiksi*.

Mikä hyvänsä $m \times n$ -matriisi voidaan siis esittää rivivektoreiden kokoelmana (m kpl) tai sarakevektoreiden kokoelmana (n kpl). Yksin esiintyvän rivivektorin koordinaatit on tapana erottaa pilkulla. Vektoreita on tapana merkitä myös lihavoiduilla kirjaimilla ja yksinkertaisella indeksoinnilla, esi-

merkiksi $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ja $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$. Sekä rivi- että sarakevektorit voidaan tulkita avaruuden $\mathbb{R}^n$ alkioiksi.

Määritelmä 18 Erityisiä matriiseja ovat mm. seuraavat:

- *Nollamatriisi* O on matriisi, jonka kaikki alkiot ovat nollia
- *Neliömatriisi* on matriisi, jossa rivien määrä on sama kuin sarakkeiden määrä
- *Diagonaalimatriisi* on neliömatriisi D , jolle pätee $i \neq j \Rightarrow D_{ij} = 0$.
- *Identiteettimatriisi* I_n on diagonaalimatriisi, jonka kaikki *diagonaali*alkiot ovat ykkösiä.

Määritelmä 19 Olkoon A $m \times n$ -matriisi. Tällöin määritellään

- *Transpoosi* $(A^T)_{ij} = A_{ji}$.
- Neliömatriisi A on *symmetrinen*, jos $A^T = A$.
- Matriisin A :n *vastamatriisi* $-A$ määritellään asettamalla $(-A)_{ij} = -A_{ij}$.

Määritelmä 20 (Skalaarikertolasku) Jos A on $m \times n$ -matriisi ja c joko kompleks- tai reaali-luku, on cA $m \times n$ -matriisi, jolle pätee $(cA)_{ij} = cA_{ij}$ (kertolasku alkioittain).

Määritelmä 21 (Matriisien yhteenlasku) Jos A on $m \times n$ -matriisi ja B $r \times s$ -matriisi, summa $A + B$ määritellään vain jos $m = r$ ja $n = s$. Tällöin

$$(A + B)_{ij} = A_{ij} + B_{ij}$$

(yhteenlasku alkioittain).

Huomautus 6 On helppo todeta, että $m \times n$ -matriisit muodostavat yhteen- ja skalaarikertolaskun suhteen vektoriavaruuden.

Esimerkki 33 $A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3 -matriisi ja $5A = \begin{pmatrix} 10 & -5 & 0 \\ 0 & 5 & 15 \end{pmatrix}$

Esimerkki 34 $A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3 -matriisi ja $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 3 \end{pmatrix}$ 3×2 -matriisi, joten summaa $A + B$ ei ole määritelty.

Esimerkki 35 $A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3 -matriisi ja $B = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3 -matriisi, joten summa $A + B$ on määritelty:

$$A + B = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 6 \end{pmatrix}$$

Esimerkki 36

$$(-x + 3, -x - 1, x) = (-x, -x, x) + (3, -1, 0) = x(-1, -1, 1) + (3, -1, 0).$$

Matlab-ohjelmistoon (versio 6.5) matriisit syötetään siten, että rivit erotetaan toisistaan puolipisteellä ja alkiot pilkuilla. Esimerkiksi matriisit

$$A = \begin{pmatrix} 1 & 3 & -3 & -2 \\ 0 & 2 & 1 & -2 \\ -4 & 1 & 2 & -3 \end{pmatrix} \quad \text{ja} \quad B = \begin{pmatrix} 2 & 0 & -1 & -1 \\ 1 & 0 & 3 & 3 \\ -1 & 5 & 0 & -2 \end{pmatrix}$$

Syötetään seuraavasti:

```
>> A=[1,3,-3,-2;0,2,1,-2;-4,1,2,-3]
```

```
>> B=[2,0,-1,-1;1,0,3,3;-1,5,0,-2]
```

Matriisien summa lasketaan ilmeisellä tavalla ja transpoosi ilmaistaan heittomerkillä:

```
>> A+B
```

```
ans =
```

```
3      3     -4     -3
1      2      4      1
-5      6      2     -5
```

```
>>
```

```
>> ans'
```

```
ans =
```

```
3      1     -5
3      2      6
-4      4      2
-3      1     -5
```

2.2 Lineaarikuvaukset

Matriisin käsitettä voidaan käyttää, ja käytetäänkin tämän kurssikokonaisuuden puitteissa käytetty jäsentelemään kaavioita: Gaussin-Jordanin menetelmä on käyttökelpoinen.

Käsitteenä matriisi on kuitenkin paljon yleisempi ja käyttökelpoisempi kuin pelkkä kaavio. Matriisi on lineaarikuvauksen esitystapa. Yhdessä lineaarikuvauksen ja matriisin käsitteet muodostavat lineaarialgebran perusta.

Lineaarikuvaus voidaan määritellä ainakin kahdella tavalla, joista kumpikin johtaa samaan käsitteeseen. Kirjallisuudessa tavallisimmin lienee seuraava määritelmä:

Määritelmä 22 Funktio $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ on lineaarinen, jos

$$f(a\mathbf{x} + b\mathbf{y}) = af(\mathbf{x}) + bf(\mathbf{y})$$

aina, kun $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ja a ja b ovat skalaareita.

Yleisesti ottaen sanoja kuvaus ja funktio voidaan käyttää synonyymeinä, mutta erityisesti lineaaristen funktioiden kohdalla on tapana käyttää termiä kuvaus. Lineaarikuvausta on usein myös tapana merkitä isolla kirjaimella ja jättää sulkeet muuttujan ympäriltä merkitsemättä.

Esimerkki 37 $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = \sqrt{x}$ ei ole lineaarinen, sillä esimerkiksi

$$2 = \sqrt{4} = f(4) = f(1 \cdot 2 + 1 \cdot 2) \neq 1 \cdot f(2) + 1 \cdot f(2) = \sqrt{2} + \sqrt{2}$$

Esimerkki 38 $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $T(x, y, z) = (x + y - z, 2x + 3y + z)$ on lineaarinen, sillä

$$\begin{aligned} T(a(x_1, y_1, z_1) + b(x_2, y_2, z_2)) &= T(ax_1 + bx_2, ay_1 + by_2, az_1 + bz_2) \\ &= (ax_1 + bx_2 + ay_1 + by_2 + az_1 + bz_2, 2(ax_1 + bx_2) + 3(ay_1 + by_2) + az_1 + bz_2) \\ &= (a(x_1 + y_1 + z_1) + b(x_2 + y_2 + z_2), a(2x_1 + 3y_1 + z_1) + b(2x_1 + 3y_2 + z_2)) \\ &= a(x_1 + y_1 + z_1, 2x_1 + 3y_1 + z_1) + b(x_2 + y_2 + z_2, 2x_1 + 3y_2 + z_2) \\ &= aT(x_1, y_1, z_1) + bT(x_2, y_2, z_2). \end{aligned}$$

Esimerkki 39 Olkoon lineaarikuvaukselle $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ $f(1, 2) = (2, 2)$ ja $f(2, 3) = (2, 5)$. Määritetään $f(5, 8)$. Koska $(5, 8) = 1 \cdot (1, 2) + 2 \cdot (2, 3)$, on lineaarisuuden perusteella $f(5, 8) = 1 \cdot f(1, 2) + 2 \cdot f(2, 3) = 1 \cdot (2, 2) + 2 \cdot (2, 5) = (6, 12)$.

Esimerkki 40 Olkoon $C^1[a, b]$ välillä $[a, b]$ derivoituvien reaalifunktioiden joukko, joiden derivaatta on jatkuva. Tällöin derivointi

$$\frac{d}{dx} : C^1[a, b] \rightarrow C^0[a, b], \quad \frac{d}{dx} f(x) = f'(x)$$

on lineaarikuvauksella, sillä $\frac{d}{dx}(af_1 + bf_2) = a\frac{d}{dx}f_1 + b\frac{d}{dx}f_2$.

Äärellisulotteisissa avaruuksissa lineaarikuvauksen toinen mahdollinen määritelmä voidaan johdattaa tässä käytetystä määritelmästä suoraviivaisesti. Ensinnäkin voidaan induktiolla todeta, että lineaarikuvaukselle pätee

$$f(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_n\mathbf{x}_n) = a_1f(\mathbf{x}_1) + a_2f(\mathbf{x}_2) + \dots + a_nf(\mathbf{x}_n),$$

ja koska mikä hyvänsä $\mathbf{x} \in \mathbb{R}^n$ voidaan esittää luonnollisen kannan avulla:

$$\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n,$$

saadaan lineaarisuutta käyttämällä

$$f(\mathbf{x}) = x_1f(\mathbf{e}_1) + \dots + x_nf(\mathbf{e}_n)$$

Ylläoleva yhtälö merkitsee sitä, että lineaarikuvauksella määrätty täydellisesti kantavektoreiden kuvien $f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)$ mukaan.

Jos käytetään merkintää $\mathbf{y} = f(\mathbf{x}) \in \mathbb{R}^m$, vektorin $\mathbf{y}$ i :nnessä koordinaatista merkintää y_i , saadaan ylläolevasta yhtälöstä

$$f(\mathbf{x})_i = x_1f(\mathbf{e}_1)_i + x_2f(\mathbf{e}_2)_i + \dots + x_nf(\mathbf{e}_n)_i$$

ja merkitsemällä $a_{ij} = f(\mathbf{e}_j)_i$

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n.$$

Täten lineaarikuvauksessa $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ kuvavektorin $\mathbf{y} = f(\mathbf{x})$ jokainen koordinaatti y_i on ensimmäisen asteen lauseke alkukuvan $\mathbf{x} = (x_1, \dots, x_n)$ koordinaateista:

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n. \quad (2.2)$$

Tämä ominaisuus olisi voitu valita lineaarikuvauksen määritelmäksi luvun alussa esitetyn sijaan.

Edellä esiintyneet, lineaarikuvaukseen $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ liittyvät luvut $a_{ij} = f(\mathbf{e}_j)_i$ voidaan koota matriisiksi

$$M_f = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

jota sanotaan lineaarikuvauksen matriisiksi luonnollisen kannan suhteen.

Esimerkki 41 Esimerkin 38 lineaarikuvaukselle $T(x, y, z) = (x + y - z, 2x + 3y + z)$ $T(\mathbf{e}_1) = (1, 2)$, $T(\mathbf{e}_2) = (1, 3)$ ja $T(\mathbf{e}_3) = (-1, 1)$, joten

$$M_T = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & 1 \end{pmatrix}$$

Esimerkki 42 Olkoon $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ lineaarikuvaus, jolle $f(1, 0, 0) = (2, 1)$, $f(0, 1, 0) = (1, 1)$ ja $f(0, 0, 1) = (0, 2)$. Määritetään $f(1, 2, 3)$. Koska $(1, 2, 3) = 1 \cdot (1, 0, 0) + 2 \cdot (0, 1, 0) + 3 \cdot (0, 0, 1)$, pätee lineaarikuvaukselle f

$$f(1, 2, 3) = 1 \cdot f(1, 0, 0) + 2 \cdot f(0, 1, 0) + 3 \cdot f(0, 0, 1) = 1 \cdot (2, 1) + 2 \cdot (1, 1) + 3 \cdot (0, 2) = (4, 9).$$

Esimerkki 43 Määritetään matriisi A_f edellisen esimerkin lineaarikuvaukselle. Koska

$$\begin{aligned} f(x, y, z) &= f(x \cdot (1, 0, 0) + y \cdot (0, 1, 0) + z \cdot (0, 0, 1)) = x \cdot f(1, 0, 0) + y \cdot f(0, 1, 0) + z \cdot f(0, 0, 1) \\ &= x \cdot (2, 1) + y \cdot (1, 1) + z \cdot (0, 2) = (2x + y, x + y + 2z). \end{aligned}$$

$$\text{Näin ollen } A_f = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}$$

Huomautus 7 Huomaa, että matriisin A_f sarakkeina ovat f :n kuvat vektoreista $(1, 0, 0)$, $(0, 1, 0)$ ja $(0, 0, 1)$.

Seuraavan määritelmän myötä päästään matriisikäsitteen ja lineaarikuvausten välillä vallitsevaan yhteyteen käsiiksi syvällisemmin.

Määritelmä 23 Jos A on $r \times s$ -matriisi ja B $s \times t$ -matriisi, on matriisien A ja B tulo $r \times t$ -matriisi AB , missä

$$(AB)_{ij} = \sum_{k=1}^s A_{ik} B_{kj}.$$

Matriisitulo on siis määritelty vain silloin, kun tulon ensimmäisen tekijän sarakkeiden määrä on yhtä suuri kuin jälkimmäisen tekijän rivien määrä.

Esimerkki 44

$$\begin{aligned} &\begin{pmatrix} 2 & 1 & 1 \\ -1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2 \cdot 1 + 1 \cdot 2 + 1 \cdot 1 & 2 \cdot 0 + 1 \cdot 1 + 1 \cdot 2 & 2 \cdot 0 + 1 \cdot 1 + 1 \cdot 2 \\ -1 \cdot 1 + 2 \cdot 2 + 3 \cdot 1 & -1 \cdot 0 + 2 \cdot 1 + 3 \cdot 2 & -1 \cdot 0 + 2 \cdot 1 + 3 \cdot 2 \end{pmatrix} \\ &= \begin{pmatrix} 5 & 3 & 3 \\ 6 & 8 & 8 \end{pmatrix}. \end{aligned}$$

Esimerkki 45

$$\begin{aligned} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix} &= \begin{pmatrix} 4 & 8 \\ 3 & 5 \end{pmatrix}, \\ \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} 5 & 4 \\ 6 & 4 \end{pmatrix}. \end{aligned}$$

Edellinen esimerkin osoittaa, että matriisitulo ei ole yleisesti *kommutatiivinen* (vaihdannainen), toisin sanoen voi olla $AB \neq BA$.

Esimerkki 46

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

Tämä esimerkki puolestaan osoittaa, että *tulon nollasääntö* ei päde matriiseille, mikä siis merkitsee sitä, että $AB = O$ on mahdollinen, vaikka $A \neq O$ ja $B \neq O$

Lause 9 Matriisitulolle on voimassa seuraavat yhtäsuuruudet:

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $a(AB) = (aA)B = A(aB)$
5. $AO = OA = O$
6. $AI = IA = A$,
7. $(AB)^T = B^T A^T$,

edellyttäen että kunkin yhtälön vasen puoli on määritelty.

Todistus Suoraviivainen, mutta työläs. Jätetään harjoitustehtäväksi. □

Huomautus 8 Suoraan määritelmästä seuraa, että matriisitulo voidaan laskea myös lohkoittain:

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} = \begin{pmatrix} A_1 B_1 + A_2 B_3 & A_1 B_2 + A_2 B_4 \\ A_3 B_1 + A_4 B_3 & A_3 B_2 + A_4 B_4 \end{pmatrix},$$

mikäli lohkojen A_i ja B_i koot on valittu siten, että kaikki tulot ovat määriteltyjä.

Määritelmä 24 Neliömatriiseille määritellään *potenssi* samoin kuin reaali-luvuille:

$$\begin{cases} A^0 = I \\ A^n = A \cdot A \cdot \dots \cdot A \quad (n \text{ kpl}). \end{cases}$$

Neliömatriiseille voidaan siis muodostaa myös *matriisipolynomeja* $P(A) = c_d A^d + \dots + c_1 A + c_0 I$.

Matlab-ohjelmistossa matriisikertolasku ilmaistaan asteriskilla $*$ ja potenssi nuolella $^$

```
>> A*B
```

```
ans =
```

```

-20    12     6
  -9    -1     3
-18    -5     8
```

```
>>
```

Selvitetään seuraavaksi miksi matriisitulo on määritelty juuri sillä tavalla kuin tehtiin. Olkoon $A = (a_{ij})$ $m \times n$ -matriisi ja $\mathbf{x} = (x_1, \dots, x_n)^T$ $n \times 1$ -matriisi (pystyvektori). Tällöin tulo $A\mathbf{x}$ on matriisitulon määritelmän mukaan $m \times 1$ -matriisi (pystyvektori)

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

mikä yhtälön (2.2) perusteella merkitsee sitä, että jos A on lineaarikuvauksen f matriisi, voidaan vektorin $\mathbf{x} = (x_1, \dots, x_n)^T$ kuva $\mathbf{y} = (y_1, \dots, y_m)^T$ esittää muodossa

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Esimerkki 47 Esimerkin 43 lineaarikuvauksen kuvavektori voidaan pystymuodossa esittää seuraavasti:

$$\begin{pmatrix} 2x + y \\ x + y + 2z \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Edellämainitun perusteella seuraava lause on ilmeinen.

Lause 10 Lineaarikuvaus $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ voidaan esittää muodossa $f(\mathbf{x}) = A\mathbf{x}$, missä $\mathbf{x}$ on $n \times 1$ -matriisi (pystyvektori) ja A on $m \times n$ -matriisi. Matriisi $A = A_f$ on lineaarikuvauksen f matriisi luonnollisen kannan suhteen.

Olkoot kuvausten $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ja $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$ matriisiesitykset $f(\mathbf{x}) = A_f\mathbf{x}$ ja $g(\mathbf{y}) = A_g\mathbf{y}$. Tällöin yhdistetty kuvaus saadaan muodossa $(g \circ f)(\mathbf{x}) = g(f(\mathbf{x})) = A_g(A_f\mathbf{x}) = (A_g A_f)\mathbf{x}$. Näin on saatu seuraava tulos:

Lause 11 Jos lineaarikuvausten $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ja $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$ matriisit ovat A_f ja A_g , niin yhdistetyn kuvauksen $g \circ f$ matriisi on tulo $A_g A_f$.

Huomautus 9 Edellisen lauseen tapauksessa matriisi A_f on siis tyyppiä $m \times n$ ja A_g tyyppiä $k \times m$. Tällöin matriisitulo $A_g A_f$ on määritelty ja tyyppiä $k \times n$, kuten kuvauksen $\mathbb{R}^n \rightarrow \mathbb{R}^k$ matriisiin tulee ollakin.

Huomautus 10 Olkoon A $m \times n$ -matriisi, $\mathbf{x}$ n -pituisen pystyvektori ja $\mathbf{y} = A\mathbf{x}$ m -pituisen pystyvektori. Transpoosit laskemalla saadaan $\mathbf{y}^T = \mathbf{x}^T A^T$, mikä antaa vaihtoehdoisen muodon lineaarikuvauksen esittämiseksi matriisien avulla.

Esimerkki 48 Insinöörimatematiikka C:ssä esitetty diskreetti Fourier-muunnos on $\mathbb{C}^N$:ssä tapahtuva operaatio, jossa vektori $(f_0, f_1, \dots, f_{N-1})$ muunnetaan vektoriksi $(F_0, F_1, \dots, F_{N-1})$, missä

$$F_l = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f_k e^{-\frac{2\pi i k l}{N}} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f_k \rho^{kl},$$

missä on merkitty $\rho = e^{-\frac{2\pi i}{N}}$. Tämä voidaan ilmaista myös matriisikertolaskuna

$$\begin{pmatrix} F_0 \\ F_1 \\ F_2 \\ \vdots \\ F_{N-1} \end{pmatrix} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ 1 & \rho^2 & \rho^4 & \dots & \rho^{(N-1)2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \rho^{N-1} & \rho^{2(N-1)} & \dots & \rho^{(N-1)(N-1)} \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{N-1} \end{pmatrix}$$

Esimerkki 49 Joukot $B_1 = \{\mathbf{i}, \mathbf{j}\}$ ja $B_2 = \{\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}\}$ muodostavat avaruuden $\mathbb{R}^2$ kannan. Esimerkin 13 mukaan koordinaattivektorit ovat $(x, y)_{B_1} = (x, y)$ ja $(x, y)_{B_2} = (\frac{1}{2}(x + y), \frac{1}{2}(x - y))$. Tällöin siis

$$(x, y)_{B_2}^T = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} (x, y)_{B_1}^T.$$

Matriisia $M = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ kutsutaan kannanvaihdon matriisiksi.

Esimerkki 50 Lineaarinen yhtälöryhmä

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

voidaan kirjoittaa matriisimuodossa $A\mathbf{x} = \mathbf{b}$, missä

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

$\mathbf{x} = (x_1, \dots, x_n)^T$ ja $\mathbf{b} = (b_1, \dots, b_m)$.

2.3 Käänteismatriisi

Määritelmä 25 Olkoon A $n \times n$ -neliomatriisi. Jos on olemassa sellainen $n \times n$ -neliomatriisi B , että $AB = BA = I_n$ (I_n on n -rivinen identiteettimatriisi, katso määritelmä 18), sanotaan että B on A :n *käänteismatriisi* ja merkitään $B = A^{-1}$. Jos neliomatriisilla A on käänteismatriisi, A :ta kutsutaan *säännölliseksi*, muussa tapauksessa A on *singulaarinen*.

Huomautus 11 Voidaan todistaa, että $n \times n$ -neliomatriisi A on säännöllinen tarkalleen silloin, kun se on *täysiasteinen*, mikä merkitsee sitä, että $r(A) = n$.

Esimerkki 51 Koska $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, on $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

Esimerkki 52 Lasketaan esimerkissä 48 esintyneen matriisin tulo toisen samankaltaisen matriisin kanssa:

$$\frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ 1 & \rho^2 & \rho^4 & \dots & \rho^{(N-1)2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \rho^{N-1} & \rho^{2(N-1)} & \dots & \rho^{(N-1)(N-1)} \end{pmatrix} \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \rho^{-1} & \rho^{-2} & \dots & \rho^{-(N-1)} \\ 1 & \rho^{-2} & \rho^{-4} & \dots & \rho^{-(N-1)2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \rho^{-(N-1)} & \rho^{-2(N-1)} & \dots & \rho^{-(N-1)(N-1)} \end{pmatrix}.$$

Tulon määritelmän mukaan kohdassa rs oleva alkio on

$$\frac{1}{N} \sum_{k=0}^{N-1} \rho^{rk} \rho^{-ks} = \frac{1}{N} \sum_{k=0}^{N-1} e^{\frac{2\pi i}{N} k(r-s)} = \begin{cases} 1, & \text{jos } s \neq r, \\ 0, & \text{jos } s = r. \end{cases}$$

Viimeisin yhtäsuuruus perustellaan kurssilla Insinöörimatematiikka C. Tuloksena on siis identiteettimatriisi ja ylläolevat matriisit ovat siksi toistensa käänteismatriiseja.

Vaikka matriisitulo ei yleensä ole vaihdannainen (kommutatiivinen), siis yleensä $AB \neq BA$, voidaan kuitenkin osoittaa, että jos $AB = I$, niin myös $BA = I$. Täten siis käänteismatriisin määritelmässä riittää, että $AB = I$. Viimeisin yhtälö matriisimuotoon kirjoitettuna on

$$A \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

Mikä matriisitulon määritelmän nojalla voidaan pilkkoa n :ksi yhtälöryhmäksi (matriisitulo voidaan laskea lohkoittain)

$$A \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, A \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, A \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Koska näiden ryhmien kerroinmatriisi (A) on sama, voidaan ryhmät ratkaista samanaikaisesti suoritamalla (useampikertaisesti) augmentoituun matriisiin ($A \ I$) sellaiset alkeisoperaatiot, että A muuntuu identiteettimatriisiksi I (mikäli mahdollista; osoittautuu, että jos A ei ole säännöllinen, ei A :ta voi muuntaa identiteettimatriisiksi alkeisoperaatioin).

Käänteismatriisin etsiminen voidaan toteuttaa alkeisoperaatioin, $(A \ I) \sim \dots \sim (I \ A^{-1})$. Jos alkeisoperaatiot *eivät* muuta A :ta identiteettimatriisiksi, voidaan tämä todeta siten, että A :han tulee jossakin vaiheessa nollarivi. Tämä merkitsee sitä, että A ei ole säännöllinen.

Esimerkki 53 Matriisin $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ käänteismatriisin löytäminen alkeismuunnoksilla tapahtuu Gaussin-Jordanin eliminointimenetelmää soveltaen seuraavasti: Otetaan ensiksi käyttöön (useampikertaisesti) augmentoitu matriisi $(A \ I) = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$, johon sovelletaan alkeismuunnoksia: Ensiksi kerrotaan ensimmäinen rivi vakiolla $\frac{1}{2}$, minkä tuloksena saadaan $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$. Seuraavaksi lisätään ensimmäinen rivi toiseen kerrottuna luvulla -1 , jolloin tuloksena on $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}$. kertomalla tämän matriisin toinen rivi 2:lla saadaan $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & -1 & 2 \end{pmatrix}$, mistä toinen rivi $-\frac{1}{2}$:lla kerrottuna ensimmäiseen lisättyä antaa $\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 2 \end{pmatrix}$. Tällöin siis alkuperäinen matriisi on muunnettu alkeismuunnoksilla identiteettimatriisiksi, ja sen käänteismatriisi $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ voidaan lukea augmentoidun matriisin oikeanpuoleisesta lohkoista.

Huomautus 12 Jos neliömatriisin A käänteismatriisi A^{-1} on helposti saatavilla, voidaan sen avulla ratkaista yhtälöryhmä $A\mathbf{x} = \mathbf{b}$. Tällöin nimittäin yhtälöstä $A\mathbf{x} = \mathbf{b}$ seuraa $A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$, mikä käänteismatriisin määritelmän perusteella voidaan kirjoittaa muotoon $\mathbf{x} = A^{-1}\mathbf{b}$.

Esimerkki 54 Ratkaistaan yhtälöpari

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

käänteismatriisia käyttäen. Matriisitulon määritelmän perusteella kyseinen yhtälöpari voidaan kirjoittaa muodossa

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Kertomalla yhtälö (vasemmalta) matriisilla $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ (joka on siis yllä esiintyvän matriisin käänteismatriisi) saadaan

$$\underbrace{\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}}_{I_2} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

mikä sievenee muotoon

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

On huomattava, että käänteismatriisin etsimiseen perustuvassa yhtälöryhmän ratkaisumenetelmässä on huomattavan paljon rajoitteita verrattuna yleiseen Gaussin-Jordanin menetelmään. On nimittäin oletettava että muuttujia on yhtä monta kuin tuntematonta ja että yhtälöryhmän kerroinmatriisin

käänteismatriisi on ylipäänsä olemassa. Lisäksi on huomattava, että käänteismatriisin etsiminen on jopa työläämpää kuin yhtälöryhmän ratkaiseminen Gaussin-Jordanin menetelmällä (mitään huomattavan paljon helpompaa tapaa kuin yllä esitetty ei tunneta). Tämän vuoksi voidaan perustellusti kysyä mitä mieltä ylipäänsä koko käänteismatriisin käsitteessä on.

Tähän kysymykseen on monenlaisia vastauksia, joista useat liittyvät matriisialgebraan ja ovat tämän kurssin ulottumattomissa, mutta tässä yhteydessä voidaan nähdä yksi relevantti vastaus: Mikäli käänteismatriisi A^{-1} on etsitty, voidaan yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ ratkaisu $\mathbf{x} = A^{-1}\mathbf{b}$ saada helposti esiin millä hyvänsä vektorin $\mathbf{b}$ arvolla. Sen sijaan Gaussin-Jordanin menetelmän käyttö näyttää johtavan siihen, että mikäli $\mathbf{b}$ vaihdetaan, tulee yhtälöryhmä $A\mathbf{x} = \mathbf{b}$ ratkaista uudelleen samaista menetelmää käyttäen.

2.4 Determinantti

Determinantin käsite voidaan määritellä joko algebrallisesta, kombinatorisesta tai geometrisestä näkökulmasta lähtien. Koska lähtökohtia on useita, on ymmärrettävää, että determinantin käsite on erittäin käyttökelpoinen. Tämän kurssin puitteissa ei kuitenkaan voida esitellä determinantin määritelmää tai kytkentöjä lähtökohtiin yksityiskohtaisesti. Siksi tyydytään vain lähinnä esittelemään menetelmä, jolla determinantti lasketaan ja käymään läpi joitakin determinantin ominaisuuksia ilman todistuksia.

Esimerkki 55 Olkoon $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ lineaarikuvaus, jonka matriisi on $A_f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Selvitetään milloin f on injektio.

Injektiivisyys merkitsee sitä, että $\mathbf{x}_1 \neq \mathbf{x}_2 \Rightarrow f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$. Matriisimuotoa käyttäen saadaan ekvivalenssi $f(\mathbf{x}_1) = f(\mathbf{x}_2) \Leftrightarrow A_f \mathbf{x}_1 = A_f \mathbf{x}_2 \Leftrightarrow A_f(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, mikä merkitsee siis sitä, että f on injektio tarkalleen silloin kun ehdosta $A_f \mathbf{x} = \mathbf{0}$ seuraa $\mathbf{x} = \mathbf{0}$. Toisin sanoen, lineaarikuvaus f ei ole injektio, jos on olemassa sellainen $\mathbf{x} \neq \mathbf{0}$, että $A_f \mathbf{x} = \mathbf{0}$. Viimeisin ehto taas merkitsee sitä, että A_f :llä ei voi olla käänteismatriisia, sillä sellaisen olemassaolosta seuraisi $\mathbf{x} = A_f^{-1} \mathbf{0} = \mathbf{0}$.

Koska $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, nähdään, että käänteismatriisi on olemassa jos $ad - bc \neq 0$.

Oletetaan sitten, että $ad - bc = 0$. Jos $c = d = 0$, nähdään, että $A_f(b, -a)^T = \mathbf{0}$ (jos myös $a = b = 0$, mikä tahansa vektori $\mathbf{x} \neq \mathbf{0}$ toteuttaa $A_f \mathbf{x} = \mathbf{0}$), kun taas tapauksessa $c \neq 0$ tai $d \neq 0$ nähdään, että $A_f(d, -c)^T = \mathbf{0}$. Tämä merkitsee sitä, että jos $ad - bc = 0$, ei f voi olla injektio. Kokoamalla kaikki mainitut seikat yhteen nähdään, että f on injektio tarkalleen silloin kun $ad - bc \neq 0$.

Esimerkki 56 Olkoon f kuten edellisessä esimerkissä. Helposti nähdään, että $f(1, 0) = (a, c)$ ja $f(0, 1) = (b, d)$. Lasketaan sellaisen suunnikkaan pinta-ala, jonka sivuina ovat vektorit (a, c) ja (b, d) . Vastaus: $|ad - bc|$. Tällöin voidaan ajatella, että lineaarikuvaus f "venyttää" alkuperäistä suunnikasta, jonka sivuina ovat $(1, 0)$ ja $(0, 1)$ (ja pinta-ala 1) suunnikkaaksi, jonka pinta-ala on $|ad - bc|$.

Määritelmä 26 2×2 -matriisin $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ determinantti on $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

Lause 12 Jos A ja B ovat 2×2 -matriiseja, on $\det(AB) = \det(A) \det(B)$.

Todistus Suoralla laskulla. □

Lineaarikuvausten $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ matriisi A_f on 3×3 -matriisi ja determinantin käsitettä yleistettäessä 3×3 -matriiseille otetaankin lähtökohdaksi 2×2 -tapauksen ominaisuudet. Osoittautuu, että determinantti on mahdollista määritellä kaikille kokonaisluvulle n samoilla ominaisuuksilla:

Kutakin $n \times n$ -matriisia A kohti $\det(A)$ on matriisin A alkioista määräytyvä luku, jolle seuraavat ehdot pätevät:

- $\det(A) \neq 0$ tarkalleen silloin kun A^{-1} on olemassa.
- $|\det(A)|$ esittää sen särmiön ”tilavuutta”, joksi A :n määräämä lineaarikuvaus ”venyttää” kantavektoreiden $\mathbf{e}_1, \dots, \mathbf{e}_n$ määrittämän ”kuution”.
- $\det(AB) = \det(A)\det(B)$.

Tässä kurssissa ei kuitenkaan *määritellä* 3×3 -determinantteja (tai sitä suurempia), ainoastaan selostetaan, miten ne lasketaan.

Esimerkki 57 3-rivinen determinantti voidaan laskea seuraavasti:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

Kyseinen laskutapa on 3×3 -determinantin *kehitemä ensimmäisen rivin mukaan* laskemiselle toimii seuraavasti: Muodostetaan summalauseke, jossa kukin ensimmäisen sarakkeen alkio on kertojana, kertojan merkki kuitenkin varustettuna ”vuorottelusäännöllä”: a :lle merkki $+1$, b :lle merkki -1 , c :lle merkki $+1$, jne. Kerrottavat puolestaan saadaan alkuperäisestä matriisista poistamalla sekä rivi että sarake: alkion a kerroindeterminantti $\begin{vmatrix} e & f \\ h & i \end{vmatrix}$ saadaan poistamalla alkuperäisestä matriisista rivi ja sarake, jolla a on, samoin b :n kerroindeterminantissa $\begin{vmatrix} d & f \\ g & i \end{vmatrix}$ on poistettu lähtökohtaisesta matriisista sen rivi ja sarake, jolla b on, ja sama pätee alkion c kerroindeterminantille.

Samalla tavalla voidaan 4-rivisen determinantin laskeminen esittää *kehitemänä ensimmäisen rivin suhteen*:

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{vmatrix} = a \begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b \begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c \begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d \begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix}$$

Samaa periaatetta käyttäen voidaan palauttaa minkä hyvänsä $n \times n$ -determinantin laskeminen vähempirivisiin determinantteihin. On syytä huomauttaa, että kehitelmää ei tarvitse välttämättä tehdä ensimmäisen rivin suhteen, vaan kehitelmä minkä hyvänsä rivin tai sarakkeen suhteen tuottaa saman tuloksen.

Voidaan todistaa, että neliömatriisin A determinantti ei muutu, mikäli

1. matriisin A rivi lisätään toiseen vakiolla kerrottuna tai
2. matriisi A transponoidaan.

Sen sijaan,

1. Jos matriisin A kahden rivin järjestys vaihdetaan, muuttuu determinantin merkki.
2. Mille tahansa vakiolle c pätee

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ ca_{i1} & ca_{i2} & \dots & ca_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

3. Jos matriisissa on nollarivi, on sen determinantti nolla.
4. Jos matriisissa on kaksi samaa riviä, on sen determinantti nolla.

5. Summahajotelma:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ b_{i1} + c_{i1} & b_{i2} + c_{i2} & \dots & b_{in} + c_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ b_{i1} & b_{i2} & \dots & b_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ c_{i1} & c_{i2} & \dots & c_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Yllä olevat säännöt pätevät myös sarakkeiden suhteen, koska matriisi voidaan transponoida sen determinantin muuttumatta.

Huomautus 13 Aiemmin on käsitelty determinantin laskemista ensimmäisen rivin mukaisen kehittelyn avulla. Yllämainittu perusteella on kuitenkin ilmeistä, että determinantti voitaisiin laskea minkä hyvänsä rivin tai sarakkeen kehittelyn mukaisesti.

Käytännössä determinantin määrittäminen rivikehitelmien mukaan ei kuitenkaan ole laskennallisesti tehokasta, ellei samalla noudateta Gaussin-Jordanin eliminointimenetelmää nollien maksimoinniseksi riveillä.

$$\text{Esimerkki 58} \quad \begin{vmatrix} 1 & 2 & 0 \\ -3 & 0 & 1 \\ 4 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ -3 & 6 & 1 \\ 4 & -10 & 1 \end{vmatrix} = \begin{vmatrix} 6 & 1 \\ -10 & 1 \end{vmatrix} = 6 \cdot 1 - 1 \cdot (-10) = 16.$$

Esimerkki 59

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & b-a & b^2-ba \\ 1 & c-a & c^2-ca \end{vmatrix} = \begin{vmatrix} b-a & b(b-a) \\ c-a & c(c-a) \end{vmatrix} = (b-a)(c-a) \begin{vmatrix} 1 & b \\ 1 & c \end{vmatrix} = (b-a)(c-a)(c-b).$$

Matlabissa matriisin A determinantti, käänteismatriisi, ja aste voidaan laskea komennoilla $\det(A)$, $\text{inv}(A)$ ja $\text{rank}(A)$.

2.5 Matriisin ominaisarvot ja ominaisvektorit

Olkoot A ja B $n \times n$ -neliömatriiseja. Matriisitulo AB on yleensä varsin työläs laskea, sillä alkioita $(AB)_{ij}$ varten pitää laskea summa

$$\sum_{k=1}^n A_{ik} B_{kj},$$

jonka kutakin summattavaa varten pitää vielä laskea kertolasku. Koska tulomatriisissa AB on n^2 alkioita, näyttää edellämainittu merkitsevän, että matriisia AB varten pitää laskea n^3 kertolaskua ja $n^2(n-1)$ yhteenlaskua, siis $O(n^3)$ laskutoimitusta. Matriisitulon laskemista varten on kehitetty parempiakin menetelmiä ja esimerkiksi ensimmäinen epätriviaali menetelmä, ns. Strassenin matriisikertolasku voidaan suorittaa käyttämällä $O(n^{2.8074\dots})$ laskutoimitusta. Parhaat nykyisin tunnetut menetelmät käyttävät $O(n^{2.37\dots})$ laskutoimitusta.

Koska $n \times n$ -matriisissa on n^2 alkioita, ei matriisikertolaskua $A \times B$ yleisesti voida suorittaa vähemmällä määrällä kuin n^2 laskutoimitusta. Poikkeuksena ovat kuitenkin matriisit, joiden rakenne on yksinkertainen

Esimerkki 60 Kahden $n \times n$ -diagonaalimatriisin tulo

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1 & 0 & \dots & 0 \\ 0 & a_2 b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n b_n \end{pmatrix}$$

voidaan laskea n :llä kertolaskulla. Tästä seuraa, että diagonaalimatriisin potenssi on helppo laskea.

Esimerkki 61 Linearisessa diskreettiaikaisessa systeemissä tilaa ajanhetkellä i kuvataan pystyvektorilla $\mathbf{x}_i \in \mathbb{R}^n$ ja systeemin dynamiikkaa esittää matriisi A , mikä merkitsee sitä että $\mathbf{x}_{i+1} = A\mathbf{x}_i$. Tällöin $\mathbf{x}_i = A^i\mathbf{x}_0$, mutta A^i voi olla työläs määrittää.

Edellisen esimerkin vektori $\mathbf{x}_i = A^i\mathbf{x}_0$ olisi helppo määrittää, mikäli A olisi diagonaalimatriisi tai mikäli $\mathbf{x}_0$ voitaisiin esittää sellaisten vektoreiden lineaarikombinaationa, jotka eivät muutu oleellisesti kerrottaessa matriisilla A vasemmalta.

Määritelmä 27 Olkoon A $n \times n$ -matriisi. Luku $\lambda \in \mathbb{C}$ on matriisin A ominaisarvo, jos on olemassa sellainen $\mathbf{x} \neq \mathbf{0}$, että

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Tällöin jokaista ylläolevan yhtälön toteuttavaa vektoria $\mathbf{x}$ sanotaan ominaisarvoon λ liittyväksi ominaisvektoriksi.

Huomautus 14 Nollavektorille pätee joka tapauksessa $A\mathbf{0} = \mathbf{0} = \lambda\mathbf{0}$, olipa λ mikä hyvänsä luku. Edellisessä määritelmässä vaaditaan kuitenkin, että on olemassa jokin nollasta poikkeava vektori $\mathbf{x}$, jolle $A\mathbf{x} = \lambda\mathbf{x}$ toteutuu.

Ominaisvektorit ovat siis sellaisia vektoreita $\mathbf{x}$, joille matriisin A määrittämä lineaarikuvaus $\mathbf{x} \mapsto A\mathbf{x}$ on skalaarikertolasku $\mathbf{x} \mapsto \lambda\mathbf{x}$, mikä on käsitteellisesti ja laskennallisesti yksinkertaisempi asia kuin matriisikertolasku. Yhtälöstä $A\mathbf{x} = \lambda\mathbf{x}$ seuraa helposti induktiolla $A^i\mathbf{x} = \lambda^i\mathbf{x}$.

Huomautus 15 Jos vektori $\mathbf{x}_0$ voidaan esittää matriisin A ominaisvektoreiden lineaarikombinaationa

$$\mathbf{x}_0 = c_1\mathbf{x}_1 + \dots + c_n\mathbf{x}_n,$$

on $A^i\mathbf{x}_0$ esitettävissä muodossa

$$A^i\mathbf{x}_0 = c_1A^i\mathbf{x}_1 + \dots + c_nA^i\mathbf{x}_n = c_1\lambda^i\mathbf{x}_1 + \dots + c_n\lambda^i\mathbf{x}_n$$

Ominaisarvojen määrittämiseksi kirjoitetaan yhtälö $A\mathbf{x} = \lambda\mathbf{x}$ muotoon $A\mathbf{x} = \lambda I\mathbf{x}$ ja tämä edelleen muotoon $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Määritelmän mukaan λ voi olla ominaisarvo vain jos tällä yhtälöllä on nollavektorista poikkeava ratkaisu. Täten välttämättä matriisin $A - \lambda I$ on oltava singulaarinen (ei kääntyvä). Lisäksi on mahdollista näyttää toteen, että aina kun A on singulaarinen, on olemassa sellainen vektori $\mathbf{x} \neq \mathbf{0}$, että $A\mathbf{x} = \mathbf{0}$. Näin saadaan seuraava lause:

Lause 13 Matriisin A ominaisarvot λ ovat tarkalleen kaikki yhtälön

$$\det(A - \lambda I) = 0$$

ratkaisut.

Kun jokin ominaisarvo λ tunnetaan, voidaan siihen liittyvät ominaisvektorit määrittää Gaussin-Jordanin menetelmällä yhtälöstä $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Jokaiseen ominaisarvoon liittyy aina ääretön määrä ominaisvektoreita, kuten seuraava lause osoittaa.

Lause 14 Olkoon A $n \times n$ -matriisi. Matriisin A ominaisarvoon λ liittyvät ominaisvektorit yhdessä nollavektorin kanssa muodostavat $\mathbb{R}^n$:n aliavaruuden.

Todistus On osoitettava, että jos $\mathbf{x}_1$ ja $\mathbf{x}_2$ ovat ominaisarvoon λ liittyviä ominaisvektoreita, niin myös $a_1\mathbf{x}_1 + a_2\mathbf{x}_2$ on sellainen. Matriisikertolaskun ominaisuuksien perusteella tämä on suoraviivaista:

$$A(a_1\mathbf{x}_1 + a_2\mathbf{x}_2) = a_1A\mathbf{x}_1 + a_2A\mathbf{x}_2 = a_1\lambda\mathbf{x}_1 + a_2\lambda\mathbf{x}_2 = \lambda(a_1\mathbf{x}_1 + a_2\mathbf{x}_2).$$

Esimerkki 62 Määritetään matriisin $A = \begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix}$ ominaisarvot ja -vektorit. Ominaisarvoyhtälössä esiintyvä matriisi $A - \lambda I$ saa muodon

$$\begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -13 - \lambda & 30 \\ -9 & 20 - \lambda \end{pmatrix},$$

joten ominaisarvoyhtälö on

$$\det \begin{pmatrix} -13 - \lambda & 30 \\ -9 & 20 - \lambda \end{pmatrix} = 0,$$

mikä voidaan kirjoittaa muotoon

$$(-13 - \lambda)(20 - \lambda) - 30(-9) = 0 \Leftrightarrow \lambda^2 - 7\lambda + 10 = 0 \Leftrightarrow \lambda \in \{2, 5\}.$$

Määritetään ominaisarvoon 2 liittyvät ominaisvektorit. Tämä merkitsee yhtälöryhmän

$$\left(\begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -13 - 2 & 30 \\ -9 & 20 - 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ratkaisujen etsimistä. Gaussin-Jordanin menetelmä muuntaa tämän muotoon

$$\begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

jotka ratkaisut ovat $(x, y) = (2y, y) = y(2, 1)$. Täten siis matriisin A ominaisarvoon 2 liittyvät ominaisvektorit ovat muotoa $y(2, 1)$, missä y on vakio. Jos tahdotaan pidättäytyä reaalisisissa ominaisvektoreissa, voidaan rajoittaa $y \in \mathbb{R}$, mutta yleisesti voidaan myös kompleksiset arvot hyväksyä. Tarkastuksen vuoksi voidaan laskea (valitaan esim $y = 1$)

$$\begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Ominaisarvoon 5 liittyvät ominaisvektorit määritetään samalla tavalla. Tällöin kyse on yhtälöryhmän

$$\begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 5 \begin{pmatrix} x \\ y \end{pmatrix}$$

ratkaisuista, jotka ovat muotoa $(x, y) = (\frac{5}{3}y, y) = \frac{1}{3}y(5, 3)$ ja taas laskemalla voidaan tarkastaa (valitaan esim $y = 3$), että

$$\begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = 5 \begin{pmatrix} 5 \\ 3 \end{pmatrix}.$$

Esimerkki 63 Olkoon matriisi A kuten edellisessä esimerkissä. Selvitetään yleinen lauseke vektorille $A^i(9, 5)^T$. Edellisessä esimerkissä nähtiin, että matriisin $(2, 1)$ ja $(5, 3)$ ovat matriisin A ominaisvektoreita. On helppo huomata, että nämä ovat lineaarisesti riippumattomat, joten ne muodostavat avaruuden $\mathbb{R}^2$ kannan. Näin ollen vektori $(9, 5)$ voidaan esittää vektoreiden $(2, 1)$ ja $(5, 3)$ lineaarikombinaationa:

$$(9, 5) = c_1(2, 1) + c_2(5, 3),$$

joka voidaan kirjoittaa yhtälöparina

$$\begin{cases} 2c_1 + 5c_2 = 9 \\ c_1 + 3c_2 = 5 \end{cases}$$

Tämän ratkaisut saadaan esim. Gaussin-Jordanin menetelmällä: $c_1 = 2$, $c_2 = 1$. Tällöin siis $(9, 5) = 2(2, 1) + (5, 3)$ ja koska molemmat komponenttivektorit ovat matriisin A ominaisvektoreita, saadaan

$$A^i \begin{pmatrix} 9 \\ 5 \end{pmatrix} = A^i \left(2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 5 \\ 3 \end{pmatrix} \right) = 2A^i \begin{pmatrix} 2 \\ 1 \end{pmatrix} + A^i \begin{pmatrix} 5 \\ 3 \end{pmatrix} = 2 \cdot 2^i \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 5^i \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 2^{i+2} + 5^{i+1} \\ 2^{i+1} + 3 \cdot 5^i \end{pmatrix}.$$

Edellä kuvattu menetelmä soveltuu kaikille $n \times n$ -matriiseille, joiden ominaisvektoreiden avulla voidaan muodostaa avaruuden $\mathbb{R}^n$ (tai avaruuden $\mathbb{C}^n$) kanta. Osoittautuu, että jos matriisin kaikki ominaisarvot ovat erisuuret. Toisaalta osoittautuu, että tämä ei ole välttämätön ehto: Matriisilla saattaa

olla vain yksi ominaisarvo, mutta siitä huolimatta sen ominaisvektorit muodostavat $\mathbb{R}^n$:n tai $\mathbb{C}^n$:n kannan.

Esimerkki 64 Identiteettimatriisin I ainoa ominaisarvo on 1, sillä kaikille vektoreille $I\mathbf{x} = 1 \cdot \mathbf{x}$. Toisaalta tämän mukaisesti kaikki vektorit ovat identiteettimatriisin ominaisvektoreita, joten niistä voidaan varmastikin valita vektoriavaruuden kanta.

2.6 Matriisin diagonalisointi

Matriisin potenssien laskeminen palautuu toisinaan yksinkertaisempaan tehtävään, mikäli matriisille voidaan löytää yksinkertaisempi, ns. similaari versio.

Määritelmä 28 Neliömatriisit A ja B ovat *similaarit*, mikäli on olemassa sellainen kääntyvä matriisi P , että $A = P^{-1}BP$.

Jos $A = P^{-1}BP$, on

$$A^i = P^{-1}BP \cdot P^{-1}BP \cdot \dots \cdot P^{-1}BP = P^{-1}B^iP$$

ja jos B on matriisi, jolle B^i on helppo laskea, esimerkiksi diagonaalimatriisi, voidaan A^i laskea ylläolevan kaavan mukaisesti.

Kun pyritään etsimään mahdollisimman yksinkertainen similaari matriisi annetulle matriisille A , nousee esille kysymys siitä, miten matriisi P määritetään. Seuraava havainto on tarpeellinen yksinkertaisen esityksen etsimiselle.

Huomautus 16 Oletetaan, että $n \times n$ -matriisilla A on olemassa n ominaisarvoa $\lambda_1, \dots, \lambda_n$, joihin liittyy ominaisvektorit $\mathbf{x}_1, \dots, \mathbf{x}_n$. Muodostetaan matriisi P käyttämällä ominaisvektoreita $\mathbf{x}_i$ sarakkeina:

$$P = (\mathbf{x}_1 \dots \mathbf{x}_n),$$

jolloin AP voidaan lohkomuodossa laskea seuraavasti:

$$AP = (A\mathbf{x}_1 \dots A\mathbf{x}_n) = (\lambda_1\mathbf{x}_1 \dots \lambda_n\mathbf{x}_n)$$

ja suoraan laskemalla on helppo todeta, että

$$(\lambda_1\mathbf{x}_1 \dots \lambda_n\mathbf{x}_n) = (\mathbf{x}_1 \dots \mathbf{x}_n) \underbrace{\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}}_D = PD,$$

ja kaikki edellä esitetty yhteen kokoamalla saadaan yhtälö $AP = PD$. Jos P on kääntyvä matriisi, saadaan esitys $D = P^{-1}AP$.

Määritelmä 29 Matriisi A on *diagonalisoituva*, jos se on similaari jonkin diagonaalimatriisin D kanssa. Tällöin siis $P^{-1}AP = D$ jollekin kääntyvälle matriisille P . Diagonalisoituva matriisi voidaan esittää muodossa $A = PDP^{-1}$, missä D on diagonaalinen ja P kääntyvä.

Esimerkki 65 Olkoon $A = \begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix}$ kuten esimerkissä 74 todettiin, on matriisilla A ominaisarvoon 2 liittyvä ominaisvektori $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ sekä ominaisarvoon 5 liittyvä ominaisvektori $\begin{pmatrix} 5 \\ 3 \end{pmatrix}$. Näin ollen

$$A \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} = \left(A \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad A \begin{pmatrix} 5 \\ 3 \end{pmatrix} \right) = \left(2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad 5 \begin{pmatrix} 5 \\ 3 \end{pmatrix} \right) = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$$

Matriisin $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ käänteismatriisiin $\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$ määrittäminen onnistuu esimerkiksi aiemmassa pykälässä kuvatulla Gaussin-Jordanin menetelmään perustuvalla prosessilla. Täten matriisi A on diagonalisoituva ja voidaan esittää muodossa.

$$A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1}$$

ja siksi

$$A^i = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}^i \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2^i & 0 \\ 0 & 5^i \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 3 \cdot 2^{i+1} - 5 \cdot 3^i & -5 \cdot 2^{i+1} + 10 \cdot 3^i \\ 3 \cdot 2^i - 3^{i+1} & -5 \cdot 2^i + 2 \cdot 3^{i+1} \end{pmatrix}.$$

2.7 Jordanin normaalmuoto

Edellisessä luvussa kuvattu menetelmä matriisin diagonalisoimiseksi perustuu ominaisarvoihin liittyvien ominaisvektoreiden etsimiseen. On huomattava, että jos $\mathbf{x}_i$ on matriisin A ominaisarvoon λ_i liittyvä, pitää yhtälö $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ määritelmän mukaan paikkansa, ja tällöin matriisille $P = (\mathbf{x}_1 \dots \mathbf{x}_n)$ pätee ilman rajoituksia

$$AP = (A\mathbf{x}_1 \dots A\mathbf{x}_n) = (\lambda_1\mathbf{x}_1 \dots \lambda_n\mathbf{x}_n) = PD, \quad (2.3)$$

missä

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

on diagonaalimatriisi. Sen sijaan diagonaaliesityksen $D = P^{-1}AP$ löytäminen edellyttää, että ominaisvektorit $\mathbf{x}_1, \dots, \mathbf{x}_n$ voidaan valita siten, että $P = (\mathbf{x}_1 \dots \mathbf{x}_n)$ on kääntyvä.

Mikäli $n \times n$ -matriisilla A on n erisuurta lineaarisesti riippumatonta ominaisvektoria, on A diagonalisoituva. Itse asiassa on mahdollista todistaa, että erisuuriin ominaisarvoihin liittyvät ominaisvektorit ovat lineaarisesti riippumattomia, ja tällöin matriisi P on kääntyvä.

On kuitenkin täysin mahdollista, että matriisilla A ei ole n :ää erisuurta ominaisarvoa, ja tällöin on kyseenalaista ovatko ominaisvektorit valittavissa siten, että niistä muodostettu matriisi P olisi kääntyvä.

Esimerkki 66 Matriisin

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

ominaisarvoyhtälö on

$$\begin{vmatrix} -3-\lambda & 3 & 1 \\ -5 & 5-\lambda & 1 \\ -5 & 3 & 3-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0,$$

minkä ratkaisut ovat $\lambda = 1$ (yksinkertainen nollakohta) ja $\lambda = 2$ (kaksinkertainen nollakohta). Nämä voidaan periaatteessa löytää kokeilemalla kaikki mahdolliset rationaalinnollakohdat ja tällä tavoin löytyy polynomille $\lambda^3 - 5\lambda^2 + 8\lambda - 4$ myös muoto $(\lambda - 1)(\lambda - 2)^2$.

Ominaisarvoa 1 vastaavat ominaisvektorit voidaan löytää ratkaisemalla yhtälö

$$(A - 1 \cdot I)\mathbf{x} = \mathbf{0},$$

mikä tässä tapauksessa (merkitsemällä $\mathbf{x} = (x, y, z)^T$) saa muodon

$$\begin{pmatrix} -4 & 3 & 1 \\ -5 & 4 & 1 \\ -5 & 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Kerroinmatriisin redusoitu porrasmuoto on

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

mitä vastaa yhtälöryhmä

$$\begin{cases} x - z = 0 \\ y - z = 0 \end{cases},$$

jonka ratkaisut ovat $(x, y, z) = (z, z, z) = z(1, 1, 1)$. Näin ollen kaikki ominaisarvoon 1 liittyvät ominaisvektorit ovat vektorin $(1, 1, 1)$ skalaarimonikertoja.

Ominaisarvoon 2 liittyvät ominaisvektorit selvitetään samalla tavalla: Yhtälö

$$(A - 2I)\mathbf{x} = \mathbf{0}$$

saa muodon

$$\begin{pmatrix} -5 & 3 & 1 \\ -5 & 3 & 1 \\ -5 & 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Kerroinmatriisin redusoitu porrasmuoto on

$$\begin{pmatrix} 1 & -\frac{3}{5} & -\frac{1}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

mikä vastaa yhtälöä $x - \frac{3}{5}y - \frac{1}{5}z = 0$ ja yleinen ratkaisu voidaan kirjoittaa muotoon

$$(x, y, z) = \left(\frac{3}{5}y + \frac{1}{5}z, y, z\right) = y\left(\frac{3}{5}, 1, 0\right) + z\left(\frac{1}{5}, 0, 1\right) = \frac{1}{5}y(3, 5, 0) + \frac{1}{5}z(1, 0, 5)$$

Ominaisarvoon 2 liittyviksi ominaisvektoreiksi voidaan siis valita esimerkiksi $(3, 5, 0)^T$ ja $(1, 0, 5)^T$, jotka ovat lineaarisesti riippumattomat. Jos siis matriisiksi P valitaan

$$P = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 5 & 0 \\ 1 & 0 & 5 \end{pmatrix},$$

voidaan todeta että

$$P^{-1} = \frac{1}{5} \begin{pmatrix} 25 & -15 & -5 \\ -5 & 4 & 1 \\ -5 & 3 & 2 \end{pmatrix}$$

ja

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Esimerkki 67 Selvitetään matriisin

$$A = \begin{pmatrix} -2 & 1 & 3 \\ -5 & 3 & 4 \\ -4 & 1 & 5 \end{pmatrix}$$

ominaisarvot ja ominaisvektorit. Ominaisarvoyhtälö voidaan kirjoittaa muotoon

$$\begin{vmatrix} -2 - \lambda & 1 & 3 \\ -5 & 3 - \lambda & 4 \\ -4 & 1 & 5 - \lambda \end{vmatrix} = 0 \Leftrightarrow \lambda^3 - 6\lambda^2 + 12\lambda - 8 = 0,$$

jonka ainoa nollakohta on $\lambda = 2$. Tähän liittyvät ominaisvektorit voidaan löytää etsimällä ratkaisut yhtälölle $(A - 2\lambda)\mathbf{x} = \mathbf{0}$, mikä voidaan kirjoittaa muotoon

$$\begin{pmatrix} -4 & 1 & 3 \\ -5 & 1 & 4 \\ -4 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Kerroinmatriisin redusoitu porrasmuoto on

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

mikä vastaa yhtälöparia

$$\begin{cases} x - z = 0 \\ y - z = 0 \end{cases}$$

Jonka ratkaisut ovat muotoa $(x, y, z) = (z, z, z) = z(1, 1, 1)$. Näin ollen yhtälön (2.3) mukaista matriisia P ei voida valita siten että sellaisessa olisi kolme lineaarisesti riippumatonta saraketta.

Edellisen esimerkin tyyppisiä matriiseja sanotaan *diagonalisoitumattomiksi*. Tällaisille matriiseille voidaan kuitenkin löytää myös miltei diagonaalinen esitys, jota kutsutaan *Jordanin normaalimuodoksi*

Määritelmä 30 Muotoa

$$J_\lambda = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

olevaa matriisia sanotaan Jordan-lohkoksi.

Esimerkki 68 Diagonaalimuodossa olevan matriisin potenssit on helppo esittää.

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_k \end{pmatrix}^n = \begin{pmatrix} \lambda_1^n & & & \\ & \lambda_2^n & & \\ & & \ddots & \\ & & & \lambda_k^n \end{pmatrix},$$

mutta myös Jordan-lohkojen potensseista voi suhteellisen helposti löytää säännönmukaisuuksia:

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix},$$

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix},$$

$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \binom{n}{3}\lambda^{n-3} \\ 0 & \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & 0 & \lambda^n \end{pmatrix},$$

ja niin edelleen. Ylläolevat yhtäsuuruudet voidaan todistaa oikeiksi induktiolla.

Lause 15 Olkoon A $n \times n$ -matriisi, jonka erisuuret ominaisarvot ovat $\lambda_1, \dots, \lambda_k$. Tällöin on olemassa sellainen kääntyvä matriisi P , että $P^{-1}AP$ voidaan esittää lohkomuodossa

$$P^{-1}AP = \begin{pmatrix} J_{\lambda_1} & & \\ & J_{\lambda_2} & \\ & & \ddots \\ & & & J_{\lambda_k} \end{pmatrix}.$$

Todistus Ei esitetä, mutta tilannetta valotetaan alla selostuksin ja esimerkein. $\square$

Osoittautuu, että yhtälössä (2.3) esiintyvä matriisi P voidaan aina valita kääntyväksi matriisiksi, mikäli diagonaalisuusvaatimuksesta joustetaan ja vaaditaan ainoastaan matriisin D olevan Jordan-lohko-muodossa. Tällöin matriisin P sarakkeiksi tulee valita alkuperäisen matriisin A ominaisvektorit ja näitä täydentämään ns. *yleistetyt ominaisvektorit*.

Tarkastellaan aluksi mitkä ovat Jordan-lohkon ilmenemisvaatimukset 3×3 -tapauksessa. Jos matriisilla A on ominaisarvoon λ liittyvä ominaisvektori $\mathbf{x}$, mutta ei muita lineaarisesti riippumattomia ominaisvektoreita, merkitään

$$P = (\mathbf{x} \ \mathbf{y} \ \mathbf{z}) = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \text{ ja } J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Tulo $AP = (A\mathbf{x} \ A\mathbf{y} \ A\mathbf{z})$ on helppo esittää, mutta ongelmia ei tuota myöskään

$$PJ = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} x_1\lambda & x_1 + y_1\lambda & y_1 + z_1\lambda \\ x_2\lambda & x_2 + y_2\lambda & y_2 + z_2\lambda \\ x_3\lambda & x_3 + y_3\lambda & y_3 + z_3\lambda \end{pmatrix} = (\lambda\mathbf{x} \ \mathbf{x} + \lambda\mathbf{y} \ \mathbf{y} + \lambda\mathbf{z}).$$

Yhtäsuuruus $AP = PJ$ edellyttää siis, että

$$\begin{cases} A\mathbf{x} = \lambda\mathbf{x} \\ A\mathbf{y} = \mathbf{x} + \lambda\mathbf{y} \\ A\mathbf{z} = \mathbf{y} + \lambda\mathbf{z} \end{cases} \Leftrightarrow \begin{cases} (A - \lambda I)\mathbf{x} = \mathbf{0} \\ (A - \lambda I)\mathbf{y} = \mathbf{x} \\ (A - \lambda I)\mathbf{z} = \mathbf{y} \end{cases}$$

Toisesta yhtälöstä voidaan helppona seurauksena huomata, että $(A - \lambda I)^2\mathbf{y} = (A - \lambda I)(A - \lambda I)\mathbf{y} = (A - \lambda I)\mathbf{x} = \mathbf{0}$ ja samoin kolmannesta että $(A - \lambda I)^3\mathbf{z} = (A - \lambda I)^2\mathbf{y} = \mathbf{0}$.

Määritelmä 31 Olkoon λ matriisin A ominaisarvo ja $k \in \mathbb{N}$. Ominaisarvoon λ liittyvä kertalukua k oleva *yleistetty ominaisvektori* on mikä hyvänsä yhtälön

$$(A - \lambda I)^k \mathbf{x} = \mathbf{0}$$

toteuttava vektori.

Huomautus 17 Kun määritelmässä 31 valitaan $k = 1$, saadaan määritelmän 27 mukaiset tavalliset ominaisvektorit.

Esimerkki 69 Esimerkissä 67 todettiin, että matriisilla

$$A = \begin{pmatrix} -2 & 1 & 3 \\ -5 & 3 & 4 \\ -4 & 1 & 5 \end{pmatrix}$$

on vain yksi ominaisarvo $\lambda = 2$. Siihen kuuluvat ominaisvektorit $\mathbf{x}$ ovat yhtälön

$$(A - 2I)\mathbf{x} = \mathbf{0}$$

ratkaisut. Nämä puolestaan ovat kaikki vektorin $(1, 1, 1)^T$ skalaarimonikertoja, joten yhtälön (2.3) mukaista diagonalisoivaa matriisia ei voida löytää. Tämän sijasta voidaan etsiä ominaisarvoon 2 liittyvät astetta 2 olevat yleistetyt ominaisvektorit $\mathbf{x}$, joilla tässä tapauksessa tarkoitetaan yhtälön

$$(A - 2I)\mathbf{x} = (1, 1, 1)^T$$

ratkaisuja. Kyseinen yhtälö voidaan kirjoittaa muotoon

$$\begin{pmatrix} -4 & 1 & 3 \\ -5 & 1 & 4 \\ -4 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

ja yhtälöryhmän augmentoitu matriisi on

$$\begin{pmatrix} -4 & 1 & 3 & 1 \\ -5 & 1 & 4 & 1 \\ -4 & 1 & 3 & 1 \end{pmatrix},$$

ja edelleen tämän redusoitu porrasmuoto

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

vastaa yhtälöparia

$$\begin{cases} x - z = 0 \\ y - z = 1 \end{cases},$$

jonka ratkaisut voidaan kirjoittaa muotoon

$$(x, y, z) = (z, z + 1, z) = z(1, 1, 1) + (0, 1, 0).$$

Näin ollen astetta 2 olevaksi yleistetyksi ominaisvektoriksi voidaan valita esimerkiksi $(0, 1, 0)$. Toisaalta, valittiinpa asteen 2 yleistetty ominaisvektori miten tahansa, saadaan vasta 2 lineaarisesti riippumatonta (yleistettyä) ominaisvektoria $(1, 1, 1)$ ja esim. $(0, 1, 0)$.

Koska yhtälön (2.3) matriisia P varten tarvitaan 3 lineaarisesti riippumatonta saraketta, etsitään vielä astetta 3 oleva yleistetty ominaisvektori, jolla tässä tapauksessa tarkoitetaan yhtälön

$$(A - 2I)\mathbf{x} = (0, 1, 0)^T$$

ratkaisuja. Tämä yhtälö voidaan kirjoittaa muotoon

$$\begin{pmatrix} -4 & 1 & 3 \\ -5 & 1 & 4 \\ -4 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Kyseisen yhtälöryhmän augmentoitu matriisi on

$$\begin{pmatrix} -4 & 1 & 3 & 0 \\ -5 & 1 & 4 & 1 \\ -4 & 1 & 3 & 0 \end{pmatrix}$$

ja tämän redusoitu porrasmuoto

$$\begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

vastaa yhtälöparia

$$\begin{cases} x - z = -1 \\ y - z = -4 \end{cases},$$

jonka ratkaisut voidaan kirjoittaa muotoon

$$(x, y, z) = (z - 1, z - 4, z) = z(1, 1, 1) + (-1, -4, 0)$$

ja kertalukua 3 olevaksi ominaisvektoriksi voidaan valita esim. $(-1, -4, 0)^T$.

Jos nyt valitaan P :n sarakkeiksi astetta 1, 2 ja 3 olevat yleistetyt ominaisvektorit, saadaan

$$P = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & -4 \\ 1 & 0 & 0 \end{pmatrix}$$

nähdään että

$$P^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ -4 & 1 & 3 \\ -1 & 0 & 1 \end{pmatrix}$$

ja

$$P^{-1}AP = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Tämä esimerkki esittää sen yksinkertaisimman mahdollisen vaihtoehdon, jossa lineaarisesti riippumattomia ominaisvektoreita voidaan valita vain yksi, samoin toisen kertaluvun ominaisvektoreita, ja niinpä myös kolmannen kertaluvun ominaisvektoriksi voidaan valita yksi vektori (tai sen skalaarimonikerta). Valitettavasti tämä esimerkki ei kata kaikenlaisia mahdollisuuksia. On nimittäin mahdollista että alemmaa astetta olevat ominaisvektorit tulee valita sopivalla tavalla, jotta löydetäisiin sopivat ylempää astetta olevat ominaisvektorit. Tämä ongelma voidaan käsitellä tavallisen lineaarialgebran keinoin, valitsemalla määräämättömät kertoimet kertaluvun k ominaisvektoreiden lineaarikombinaatiolle ja ratkaisemalla lineaarisesta yhtälöryhmästä lausekkeet seuraavan kertaluvun ominaisvektoreille.

Esimerkki 70 Olkoon

$$M = \begin{pmatrix} -1 & -1 & -2 \\ 3 & 3 & 2 \\ 3 & 1 & 4 \end{pmatrix}.$$

Samoin kuin edellisissä esimerkeissä nähdään, että tämän matriisin ominaisarvoyhtälö on $8 - 12\lambda + 6\lambda^2 - \lambda^3$, jonka ainoa ratkaisu on $\lambda = 2$.

Ainoaan ominaisarvoon $\lambda = 2$ liittyvät ominaisvektorit saadaan yhtälön $(M - 2I)\mathbf{x} = \mathbf{0}$ ratkaisuna, ja tämä yhtälö puolestaan voidaan saattaa redusoituihin porrasmuotoon

$$\begin{cases} x + \frac{1}{3}y + \frac{2}{3}z = 0 \\ 0 = 0 \\ 0 = 0 \end{cases}. \quad (2.4)$$

Näin ollen matriisin M kaikki ominaisvektorit ovat muotoa

$$(x, y, z) = \left(-\frac{1}{3}y - \frac{2}{3}z, y, z\right) = y\left(-\frac{1}{3}, 1, 0\right) + z\left(-\frac{2}{3}, 0, 1\right). \quad (2.5)$$

Valitsemalla esimerkiksi $(y, z) = (3, 0)$ saadaan vektori $(-1, 3, 0)$ kun taas valinta $(y, z) = (0, 3)$ tuottaa vektorin $(-2, 0, 3)$. Nämä ovat selvästi lineaarisesti riippumattomat (miksi?), mutta normaalimuodon välittävään matriisiin tarvittaisiin vielä kolmas lineaarisesti riippumaton vektori, mutta tällaista ei ole saatavilla (astetta yksi olevien) ominaisvektoreiden joukosta ja täten tulee turvautua yleistettyihin ominaisvektoreihin.

Toisen kertaluvun yleistettyä ominaisvektoria $\mathbf{x}$ voidaan periaatteessa etsiä yhtälön $(M - 2I)\mathbf{x} = (-1, 3, 0)^T$ tai yhtälön $(M - 2I)\mathbf{x} = (-2, 0, 3)^T$ ratkaisusta, mutta ikävä kyllä osoittautuu että kummallakaan yhtälöllä ei ole ratkaisua. Tällöin voidaan tietenkin haluttua yleistettyä ominaisvektoria etsiä yhtälön

$$(M - 2I)\mathbf{x} = a(-1, 3, 0)^T + b(-2, 0, 3)^T \quad (2.6)$$

ratkaisusta ja samalla sisällyttää a ja b muuttujien joukkoon. Eksplisiittisesti kirjoitettuna ylläoleva yhtälö saa muodon

$$\begin{cases} -3x - y - 2z = -a - 2b \\ 3x + y + 2z = 3a \\ 3x + y + 2z = 3b \end{cases},$$

josta Gaussin-Jordanin menetelmällä saadaan yleinen ratkaisu (tarkista!)

$$(x, y, z, a, b) = y\left(-\frac{1}{3}, 1, 0, 0, 0\right) + z\left(-\frac{2}{3}, 0, 1, 0, 0\right) + b(1, 0, 0, 1, 1).$$

Nyt voidaan valita esimerkiksi $(y, z, b) = (0, 0, 1)$, jolloin yleistetyksi toisen kertaluvun ominaisvektori saadaan $(1, 0, 0)$. Tällöin kuitenkin tulee huomata, että $(M - 2I)(1, 0, 0)^T = (-3, 3, 3)^T$, joten ensimmäisen kertaluvun omaisvektori pitää valita $(-3, 3, 3)^T$. Koska esimerkiksi $(-1, 3, 0)^T$ on tämän kanssa lineaarisesti riippumaton (miksi?) voidaan Matriisiin P sarakkeiksi valita $(-1, 3, 0)^T$, $(-3, 3, 3)^T$ ja $(1, 0, 0)^T$, jolloin

$$J = P^{-1}MP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Huomautus 18 Muodollisesti ottaen matriisissa J on kaksi Jordan-lohkoa, vasemman yläkulman 1×1 -lohko 2 ja oikean alakulman 2×2 -lohko $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. Voidaan myös havaita, että jos (yleistettyjen) ominaisvektoreiden järjestystä matriisin P sarakkeina vaihdetaan, voidaan matriisin J muotoa muuttaa yläkolmiomatriisista alakolmiomatriisiksi. Tässä esityksessä priorisoidaan järjestystä, jossa P :n sarakkeet järjestetään vasemmalta oikealle kasvavan kertaluvun mukaan.

Yllä kuvattu menetelmä Jordanin normaalimuodon J ja välittävän matriisin P löytämiseksi voi kuitenkin käydä liian monimutkaiseksi mikäli Jordan-lohkojen koko on suurempi. Tällöin voidaan käyttää vaihtoehtoista menetelmää, jossa yleistetyt ominaisvektorit etsitään korkeammasta kertaluvusta alkaen ja menetelmää sovelletaan jokaiseen lohkoon erikseen.

Seuraavaa esimerkkiä varten otetaan käyttöön määritelmä ja merkintöjä.

Määritelmä 32 Ominaisarvon λ *algebraallinen kertaluku* tarkoittaa λ :n kertalukua itseisarvopolynomin $p(x) = \det(A - xI)$ nollakohtana. Toisin sanoen, λ :n algebraallinen kertaluku ilmaisee kuinka moninkertainen juuri λ on polynomille $p(x)$.

Määritelmä 33 Olkoon λ $n \times n$ -matriisin ominaisarvo. Ominaisarvoon λ liittyvien, kertalukua k olevien ominaisvektoreiden joukosta käytetään merkintää

$$E_\lambda^k = \{\mathbf{x} \mid (A - \lambda I)^k \mathbf{x} = \mathbf{0}\}.$$

Huomautus 19 Määritelmän perusteella on selvää, että jokainen E_λ^k on avaruuden $\mathbb{C}^n$ aliavaruus. Lisäksi on helppo havaita, että $E_\lambda^k \subseteq E_\lambda^{k+1}$ (miksi?), joten ominaisarvoon λ liittyvä aliavaruus laajenee (tai pysyy ennallaan), kun kertalukua kasvatetaan. Tämän havainnon perusteella onkin uskottavaa, että (yleistettyjen) ominaisvektoreiden etsiminen saattaa kannattaa aloittaa korkeimmasta kertaluvusta alkaen.

Merkitään ominaisarvon λ algebraalista kertalukua m :llä. Tähän ominaisarvoon liittyvien riippumattomien (yleistettyjen) ominaisvektoreiden etsiminen voidaan suorittaa seuraavasti:

- Selvitä arvoilla $i \in \{1, 2, 3, \dots\}$ luvut $d_i = \dim(E_\lambda^i)$ (mahdollisesti Gaussin-Jordanin menetelmää käyttämällä) kunnes löytyy avaruus E_λ^k jonka dimensio d_k on ominaisarvon λ algebraallinen kertaluku m .
- Arvosta $i = k$ arvoon $i = 1$ tee seuraavasti:
- Valitse avaruuden E_λ^i kannasta $\dim(E_\lambda^i) - \dim(E_\lambda^{i-1})$ vektoria, jotka eivät kuulu avaruuteen E_λ^{i-1} .
- Jatka seuraavaksi pienempään i :n arvoon ja valitse aina mukaan vektori $(M - \lambda I)\mathbf{x}$, jos $\mathbf{x}$ oli valittu aiemmassa vaiheessa.

Seuraavissa esimerkeissä valaistaan ylläkuvattua menetelmää.

Esimerkki 71 Käydään läpi esimerkin 70 normaalimuodon etsiminen uudelleen lähtien korkeamman kertaluvun ominaisvektoreista. Ominaisarvoyhtälön perusteella ominaisarvon $\lambda = 2$ algebraallinen kertaluku on 3. Yhtälön (2.5) perusteella taas avaruuden E_2^1 dimensio on kaksi, joten tämän jälkeen tulee selvittää avaruuden E_2^2 rakenne. Koska $(M - 2I)^2 = 0$, on selvästikin $E_2^2 = \mathbb{C}^3$ ja ensiksi tulee siis valita $2 - 1 = 1$ sellaista kantavektoria, jotka kuuluvat joukkoon $E_2^2 \setminus E_2^1$.

Yhtälön (2.5) perusteella esimerkiksi $(1, 0, 0)$ kelpaa. Menetelmän mukaisesti seuraavaksi tulee valita $(M - 2I)(1, 0, 0) = (-3, 3, 3) \in E_2^1$, ja koska $\dim(E_2^1) = 2$, pitää valita vielä jokin toinen, riippumaton vektori. Tällaiseksi kelpaa esimerkiksi $(-1, 3, 0)^T$ tai yhtä hyvin $(-2, 0, 3)^T$.

Esimerkki 72 Olkoon

$$A = \begin{pmatrix} -4 & 10 & -17 \\ -1 & 3 & -4 \\ 1 & -2 & 4 \end{pmatrix}.$$

Matriisin A ominaisarvoyhtälö on $1 - 3\lambda + 3\lambda^2 - \lambda^3 = 0 \Leftrightarrow (1 - \lambda)^3 = 0 \Leftrightarrow \lambda = 1$. Koska ominaisarvon 1 algebrallinen kertaluku on 3, tulisi tähän liittää kolme lineaarisesti riippumatonta (yleistettyä) ominaisvektoria.

Selvitetään aluksi avaruuden E_1^1 dimensio. Kyseisen avaruuden muodostavat vektoryhtälön

$$(A - I)\mathbf{x} = \mathbf{0}$$

ratkaisut, jotka voidaan eksplisiittisesti selvittää muodosta

$$\begin{pmatrix} -5 & 10 & -17 \\ -1 & 2 & -4 \\ 1 & -2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

joka puolestaan voidaan Gaussin-Jordanin menetelmää käyttäen kirjoittaa ekvivalenttiin asuun

$$\begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Tämän ratkaisut ovat muotoa

$$(x, y, z) = (2y, y, 0) = y(2, 1, 0), \quad (2.7)$$

missä y voidaan valita vapaasti. Näin ollen E_1^1 on yksiulotteinen avaruus, jonka generoi vektori $(2, 1, 0)$.

Avaruuden E_1^2 ja sen dimension määrittämiseksi pitää selvittää yhtälön

$$(A - I)^2 \mathbf{x} = \mathbf{0}$$

ratkaisujoukko. Eksplisiittisesti ilmaistuna ylläoleva yhtälö on muotoa

$$\begin{pmatrix} -2 & 4 & -6 \\ -1 & 2 & -3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

mikä Gaussin-Jordanin menetelmällä saadaan muotoon

$$\begin{pmatrix} 1 & -2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Tämän ratkaisut puolestaan ovat muotoa

$$(x, y, z) = (2y - 3z, y, z) = (2y, y, 0) + (-3z, 0, z) = y(2, 1, 0) + z(-3, 0, 1), \quad (2.8)$$

missä y ja z voidaan valita vapaasti. Näin ollen E_1^2 on kaksiulotteinen avaruus, joka generoivina vektoreina toimivat esim. $(2, 1, 0)$ ja $(-3, 0, 1)$.

Koska ominaisarvon $\lambda = 1$ algebrallinen kertaluku on 3 ja toistaiseksi löydetyn (yleistetyn) ominaisavaruuden E_1^2 dimensio vasta 2, tulee vielä tarkastella (yleistettyä) ominaisavaruutta E_1^3 , jonka muodostavat kaikki yhtälön

$$(A - I)^3 \mathbf{x} = \mathbf{0}$$

toteuttavat vektorit $\mathbf{x}$. Suora lasku osoittaa, että $(A - I)^3 = \mathbf{0}$, mikä merkitsee sitä, että avaruus E_1^3 on yhtä kuin $\mathbb{C}^3$. Koska näin löydetyn avaruuden E_1^3 dimensio 3 yhtyy ominaisarvon $\lambda = 1$ algebralliseen kertalukuun, voidaan aloittaa (yleistettyjen) ominaisvektoreiden etsiminen siten, että ensimmäiset valitaan joukosta $E_1^3 \setminus E_1^2$.

Koska aiemmin jo todettiin, että $\dim(E_1^3) = 3$ ja $\dim(E_1^2) = 2$, tulee ensiksi valita $3 - 2 = 1$ vektori(a), jotka eivät kuulu avaruuteen E_1^2 . Yhtälön (2.8) perusteella tällaiseksi voidaan valita mikä hyvänsä vektori, joka ei ole muotoa $(2y - 3z, y, z)$, vaikkapa $(1, 0, 0)$.

Menetelmän mukaan tulee valita myös vektori $(A - I)(1, 0, 0)^T = (-5, -1, 1)^T \in E_1^2$. Koska tämä ei kuulu avaruuteen E_1^1 ja avaruudesta E_1^2 piti löytää $\dim(E_1^2) - \dim(E_1^1) = 2 - 1 = 1$ vektoria, voidaan suoraan siirtyä seuraavaan vaiheeseen, jossa avaruudesta E_1^1 pitää valita 1 vektori. Koska joka tapauksessa pitää valita $(A - I)(-5, -1, 1)^T = (-2, -1, 0)^T$, hoituu tämäkin vaihe sillä valinnalla.

Kun järjestetään ominaisvektorit nousevan kertaluvun mukaiseen järjestykseen, saadaan Jordan-muodon välittäväksi matriisiksi

$$P = \begin{pmatrix} 1 & -5 & -2 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

ja voidaan havaita, että

$$P^{-1}AP = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

2.8 Reaalifunktioiden laajennuksia matriiseille

Tässä luvussa tarkastellaan reaalifunktioiden laajentamista neliömatriiseille. Koska kerto- ja yhteenlasku on määritelty neliömatriiseille, voidaan ainakin kaikki polynomifunktiot laajentaa suoraviivaisesti koskemaan mitä hyvänsä neliömatriisia: Jos $p(x) = c_n x^n + c_{n-1} x^{n-1} + \dots + c_1 x + c_0$ ja A on neliömatriisi, määritellään

$$p(A) = c_n A^n + c_{n-1} A^{n-1} + \dots + c_1 A + c_0 I.$$

Muiden kuin polynomifunktioiden laajentaminen neliömatriiseille kannattaakin aloittaa polynomifunktioiden suoraviivaisesta yleistyksestä, potenssisarjoista (käsitellään kurssilla Insinöörimatematiikka C). Mikäli sarja

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \quad (2.9)$$

suppenee jollakin reaalisuoran osalla, voidaan ainakin yrittää määritellä funktio f neliömatriiseille seuraavasti:

$$f(A) = \sum_{k=0}^{\infty} c_k A^k. \quad (2.10)$$

Tällöin kuitenkin joudutaan pohtimaan miten sarja (2.10) matriiseille määritellään. Vaikka tämä voidaan tehdä analogisesti reaalisten potenssisarjojen kanssa, joudutaan silti miettimään mitä käsite raja-arvo matriiseille merkitsee. Tällä kurssilla ei näihin pohdintoihin kuitenkaan pureuduta, vaan asia käsitellään lähinnä esimerkkien kautta.

Esimerkki 73 Diagonaalimatriisille

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

potenssit ja niinkään summan (2.10) alkuosa on täysin suoraviivainen.

$$\sum_{k=0}^M c_k A^k = \sum_{k=0}^M c_k \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix} = \begin{pmatrix} \sum_{k=0}^M c_k \lambda_1^k & 0 & \dots & 0 \\ 0 & \sum_{k=0}^M c_k \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sum_{k=0}^M c_k \lambda_n^k \end{pmatrix}$$

Jos kaikki diagonaali-alkiot $\lambda_1, \dots, \lambda_n$ kuuluvat sarjan (2.9) suppenemisalueeseen, on edellämainitun perusteella täysin luontevaa määritellä

$$f(A) = \begin{pmatrix} f(\lambda_1) & 0 & \dots & 0 \\ 0 & f(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(\lambda_n) \end{pmatrix}.$$

Edellisissä luvuissa on jo pohjustettu mahdollisuuksia siinä tapauksessa, että matriisi A ei ole diagonaalinen. Jos matriisi A kuitenkin on diagonalisoituva, on olemassa määritelmän 29 mukainen matriisi P ja diagonaalimatriisi D , joille pätee $A = PDP^{-1}$. Tällöin selvästikin $A^k = PD^kP^{-1}$, joten funktio f on luontevaa määritellä diagonaalisityksen avulla seuraavasti:

$$f(A) = \sum_{k=0}^{\infty} c_k A^k = \sum_{k=0}^{\infty} c_k P D^k P^{-1} = P \left(\sum_{k=0}^{\infty} c_k D^k \right) P^{-1} = P f(D) P^{-1}.$$

Esimerkki 74 Määritetään

$$\sin \begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix}.$$

Esimerkin

mukaisesti $A = \begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix}$ on diagonalisoituva ja voidaan esittää muodossa

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

Koska sinifunktion määrittelevä sarja suppenee koko reaaliakselilla, voidaan määrittää

$$\sin A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \sin 2 & 0 \\ 0 & \sin 5 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 6 \sin 2 - 5 \sin 5 & -10 \sin 2 + 10 \sin 5 \\ 3 \sin 2 - 3 \sin 5 & -5 \sin 2 + 6 \sin 5 \end{pmatrix}.$$

Esimerkki 75 Kvanttifysiikassa esiintyy ns. Paulin spin-matriisi

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Määritetään $e^{-it\sigma_x}$.

Eksponenttifunktion arvon laskemista varten etsitään ensin mahdollinen diagonaalimuoto ja tätä varten puolestaan ensin ominaisarvot ja -vektorit. Tavanomaisia menetelmiä käyttäen ominaisarvoiksi saadaan ± 1 ja -vektoreiksi $(1, 1)^T$ sekä $(1, -1)^T$. Kun merkitään

$$P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

saadaan yhtälö

$$\sigma_x = P \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} P^{-1},$$

jonka avulla voidaan määrittää

$$e^{-it\sigma_x} = P \begin{pmatrix} e^{-it \cdot 1} & 0 \\ 0 & e^{-it(-1)} \end{pmatrix} P^{-1} = \frac{1}{2} \begin{pmatrix} e^{-it} + e^{it} & e^{-it} - e^{it} \\ e^{-it} - e^{it} & e^{-it} + e^{it} \end{pmatrix} = \begin{pmatrix} \cos t & -i \sin t \\ -i \sin t & \cos t \end{pmatrix}.$$

Viimeisin yhtäsuuruus perustuu Eulerin kaavaan.

Mahdollisuus laajentaa potenssisarjan määrittelemä reaalfunktio sellaisille matriiseille, jotka eivät ole diagonalisoituvia, perustuu viime kädessä Jordanin normaalimuotoon: Jordanin normaali-muodolle voidaan nimittäin hyvin suoraviivaisesti laskea potenssit Esimerkin 68 mukaisesti ja siten laajentaa potenssisarjan määrittely Jordan-muotoisille matriiseille.

Esimerkki 76 Tarkastellaan miten sarjan (2.9) määrittelemä funktio $f(x) = \sum_{k=0}^{\infty} c_k x^k$ voidaan laajen-
taa Esimerkissä 68 esiintyvälle 3×3 -Jordan Jordan-lohkolle

$$J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}. \quad (2.11)$$

Esimerkin 68 mukaisesti

$$J^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{pmatrix}$$

ja siksi sarjaesityksen suoraviivainen laajentaminen antaa määritelmän

$$f(J) = \sum_{k=0}^{\infty} c_k J^k = \begin{pmatrix} \sum_{k=0}^{\infty} c_k \lambda^k & \sum_{k=0}^{\infty} c_k k \lambda^{k-1} & \sum_{k=0}^{\infty} c_k \binom{k}{2} \lambda^{k-2} \\ 0 & \sum_{k=0}^{\infty} c_k \lambda^k & \sum_{k=0}^{\infty} c_k k \lambda^{k-1} \\ 0 & 0 & \sum_{k=0}^{\infty} c_k \lambda^k \end{pmatrix}$$

Ylläolevassa matriisissa diagonaali-alkiot selvästikin esittävät funktion arvoa $f(\lambda)$, joten selvittä-
väksi jää vielä ylimmän rivin toinen ja kolmas alkio. Näihin voidaan pureuta differentiaalilaskennan
keinoin, joskin tarkempi tarkastelu suoritetaan kurssilla Insinöörimatematiikka C, jossa todetaan,
että potenssisarjan määrittelemä funktio voidaan derivoida termeittäin. Jos siis

$$f(x) = \sum_{k=0}^{\infty} c_k x^k,$$

on

$$f'(x) = \sum_{k=0}^{\infty} c_k k x^{k-1}$$

ja

$$f''(x) = \sum_{k=0}^{\infty} c_k k(k-1) x^{k-2}.$$

Näin ollen sarjan (2.9) määrittelemän reaalfunktion laajennus Jordan-lohkolle (2.11) pitäisi tehdä
seuraavasti:

$$f(J) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2} f''(\lambda) \\ 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & f(\lambda) \end{pmatrix}. \quad (2.12)$$

Yleistys Esimerkin 68 mukaiselle 4×4 Jordan-lohkolle on suoraviivainen; potenssisarjan määritte-
lemä funktio tulee 4×4 -jordan lohkolle

$$J = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

määritellä seuraavasti:

$$f(J) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) & \frac{1}{6}f'''(\lambda) \\ 0 & f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) \\ 0 & 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & 0 & f(\lambda) \end{pmatrix}.$$

Aivan samoin kuin jo aiemmin todettiin diagonalisoituvien matriisien kohdalla, voidaan nyt toistaa minkä hyvänsä matriisin A kohdalla: Kaikilla matriiseilla A on olemassa Jordanin normaalimuoto J sekä tämän välittävä kääntyvä matriisi P , joille siis on voimassa yhteys $A = PJP^{-1}$ ja tästä seuraa helposti, että $A^k = PJ^kP^{-1}$. Samoin kuin diagonalisoituville matriiseille, tästä edelleen seuraa, että $f(A) = Pf(J)P^{-1}$.

Täten voidaan kaikki potenssisarjojen avulla määriteltävät reaalifunktiot, kuten $\sin x$, $\cos x$ ja e^x aina laajentaa neliömatriiseille Jordan-lohkomuodon kautta. Tätä laajennusta voidaan käyttää hyödyksi erityisesti lineaaristen differentiaaliyhtälöryhmien ratkaisujen etsimisessä.

Esimerkki 77 Olkoon

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Edellisen perusteella funktio $f(x) = e^{tx}$ arvolla J määritetään seuraavasti: Lasketaan $f'(x) = te^{tx}$ ja käytetään 2.12 mukaista kaavaa, josta saadaan

$$e^{tJ} = \begin{pmatrix} e^{t\lambda} & te^{t\lambda} \\ 0 & e^{t\lambda} \end{pmatrix} = e^{t\lambda} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

Esimerkki 78 Määritetään e^{tJ} 3×3 -Jordan-lohkolle

$$J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Koska funktiolle $f(x) = e^{tx}$ saadaan $f'(x) = te^{tx}$ ja $f''(x) = t^2e^{tx}$, saadaan

$$e^{tJ} = \begin{pmatrix} e^{t\lambda} & te^{t\lambda} & \frac{t^2}{2}e^{t\lambda} \\ 0 & e^{t\lambda} & te^{t\lambda} \\ 0 & 0 & e^{t\lambda} \end{pmatrix} = e^{t\lambda} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Luku 3

Vektoriavaruuksien geometriaa

Aiemmissa pykälissä vektoriavaruuksia $\mathbb{R}^n$ on tarkasteltu lähinnä algebralliselta kannalta: Vektoreille on määritelty yhteenlasku ja skalaarimonikerta, jotka muodostavat monien sovellusten kannalta käyttökelpoisen algebrallisen rakenteen. Tuolloin ei kuitenkaan käsitelty vektoriavaruuksien *geometriaa*, jonka perustavana käsitteenä toimii vektoreiden *etäisyys*. Vektoreiden etäisyys voidaan kuitenkin määritellä useilla eri tavoilla riippuen käyttötarkoituksesta. Tässä luvussa käsitellään pääasiassa pistetulon kautta määriteltävää Euklidista etäisyyttä.

3.1 Etäisyys, normi, sisätulo

Määritelmä 34 Olkoon V vektoriavaruus. *Etäisyys* on funktio $V \times V \rightarrow \mathbb{R}$, joka toteuttaa seuraavat ehdot:

- $d(\mathbf{x}, \mathbf{y}) \geq 0$ ja $d(\mathbf{x}, \mathbf{y}) = 0$ tarkalleen silloin kun $\mathbf{x} = \mathbf{y}$ (epänegativisuus ja epädegeneratiivisuus).
- $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (symmetria).
- $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ (kolmioepäyhtälö).

Etäisyyden määritteleminen on vektoriavaruuden geometrian perusta. Etäisyys ei kuitenkaan määrity yksikäsitteisesti, vaan ylläolevan määritelmän toteuttavia etäisyysfunktioita voidaan määritellä yleensä äärettömän monta. Tämän kurssin kannalta tärkeimmät etäisyyskäsitteet määrittyvät *normin* kautta. Normikaan ei määrity yksikäsitteisesti, mutta normilla on etäisyyttä vahvempi kytkös vektoriavaruuden algebralliseen rakenteeseen.

Määritelmä 35 Olkoon V vektoriavaruus. Vektoriavaruuden *normi* on kuvaus $V \rightarrow \mathbb{R}$, $\mathbf{x} \mapsto \|\mathbf{x}\|$, joka toteuttaa seuraavat ehdot:

- $\|\mathbf{x}\| \geq 0$ ja $\|\mathbf{x}\| = 0$ tarkalleen silloin kun $\mathbf{x} = \mathbf{0}$ (positiivisuus ja epädegeneratiivisuus)
- $\|a\mathbf{x}\| = |a| \|\mathbf{x}\|$ (skalaarin siirto)
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (kolmioepäyhtälö)

Huomautus 20 Jos $\|\mathbf{x}\|$ on mikä hyvänsä ylläolevan määritelmän ehdot toteuttava normi, saadaan sen avulla määriteltyä etäisyysfunktio $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Tällöin sanotaan, että normi indusoi etäisyyden.

Esimerkki 79 Jos $V = \mathbb{R}^n$ tai $V = \mathbb{C}^n$, voidaan vektorille $\mathbf{x} = (x_1, \dots, x_n)$ määritellä

$$\|\mathbf{x}\| = |x_1| + \dots + |x_n|.$$

On helppo todeta, että näin määriteltäessä saadaan todella ylläolevan määritelmän ehdot täyttävä normi. Yleisesti ottaen vektorin normilla on tarkoitus kuvata vektorin suuruutta, mutta kuten mainittiin, normin määritelmä ei ole yksikäsitteinen. Tämän kurssin kannalta tärkeimmät normit ovat *Eukleidisia*, mikä puolestaan tarkoittaa sitä, että ne määrittävät *sisätulon* kautta.

Olkoon V reaalinen vektoriavaruus. Kuvaus $V \times V \rightarrow \mathbb{R}$, on *sisätulo*, jos se toteuttaa seuraavat ehdot:

- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ ja $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ tarkalleen silloin kun $\mathbf{x} = \mathbf{0}$ (epänegatiivisuus ja epädegeneratiivisuus).
- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ (vaihdannaisuus eli kommutatiivisuus).
- $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a\langle \mathbf{x}, \mathbf{z} \rangle + b\langle \mathbf{y}, \mathbf{z} \rangle$ (lineaarisuus).

Sisätulosta käytetään myös merkintöjä $\langle \mathbf{x}, \mathbf{y} \rangle$ ja $\langle \mathbf{x} | \mathbf{y} \rangle$. Sisätulo liittyy vektoriavaruuden algebraliseen rakenteeseen vielä vahvemmin kuin normi. Osoittautuu, että sisätulon avulla pystytään puhumaan vektoreiden etäisyyden lisäksi niiden välisistä kulmista. Sisätulokaan ei ole yksikäsitteisesti määrätty, mutta seuraava määritelmä on erittäin yleinen.

Määritelmä 36 Vektoreiden $\mathbf{x} = (x_1, \dots, x_n)$ ja $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ *pistetulo* määritellään $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \dots + x_n y_n$.

On hyvin helppo todeta, että edellisen määritelmän mukainen pistetulo toteuttaa sisätulon määritelmän ehdot, joten kyseessä on sisätulon erikoistapaus. Huomaa, että pistetulo voidaan kirjoittaa myös matriisitulona $\mathbf{x}\mathbf{y}^T$, jos $\mathbf{x}$ ja $\mathbf{y}$ ymmärretään rivivektoreiksi tai tulona $\mathbf{x}^T \mathbf{y}$, jos ne ymmärretään sarakevektoreiksi.

Esimerkki 80 Välillä $[a, b]$ jatkuvien reaalifunktioiden joukossa $C^0[a, b]$ voidaan määritellä sisätulo ehdolla

$$(f, g) = \int_a^b f g.$$

Kompleksisille vektoriavaruuksille sisätulo on kuitenkin määriteltävä toisin.

Olkoon V kompleksinen vektoriavaruus. Kuvaus $V \times V \rightarrow \mathbb{C}$, on *Hermiten sisätulo*, jos se toteuttaa seuraavat ehdot:

- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ ja $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ tarkalleen silloin kun $\mathbf{x} = \mathbf{0}$ (epänegatiivisuus ja epädegeneratiivisuus).
- $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ (vinosymmetria).
- $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a\langle \mathbf{x}, \mathbf{z} \rangle + b\langle \mathbf{y}, \mathbf{z} \rangle$ (lineaarisuus 1. argumentin suhteen).

Määritelmä 37 Vektoreiden $\mathbf{x} = (x_1, \dots, x_n)$ ja $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{C}^n$ *Hermiten pistetulo* määritellään $\mathbf{x} \cdot \mathbf{y} = x_1 \overline{y_1} + \dots + x_n \overline{y_n}$.

Esimerkki 81 Olkoon $L^2[a, b] \cap C^0[a, b]$ välillä $[a, b]$ määriteltyjen jatkuvien (kompleksi)arvoisten funktioiden joukko, joille $\int_a^b |f|^2$ on olemassa. Tällöin

$$(f, g) = \int_a^b f \overline{g}$$

on Hermiten sisätulo.

Lause 16 *Sisätulo toteuttaa Cauchyn-Schwarzin epäyhtälön*

$$|(\mathbf{x}, \mathbf{y})|^2 \leq (\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y})$$

Lause 17 *Jos $(\mathbf{x}, \mathbf{y})$ on sisätulo, saadaan määrittelemällä*

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$$

aiemman määritelmän ehdot täyttävä normi. Tällöin sanotaan, että sisätulo indusoi normin.

Huomautus 21 Jos normi määritellään käyttämällä pistetuloa, saadaan

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{|x_1|^2 + \dots + |x_n|^2},$$

mikä tapauksissa $n \in \{1, 2, 3\}$ vastaa klassista Eukleidista normia. Ellei toisin mainita, tarkoitetaan tämän jälkeen tässä luvussa normilla juuri tätä nimenomaista normia.

Normilla on siis samankaltainen merkitys vektoreille kuin itseisarvolla reaaliluvuille: normi $\|\mathbf{x}\|$ merkitsee vektorin $\mathbf{x}$ etäisyyttä nollavektorista $\mathbf{0}$ kun taas itseisarvo $|x|$ merkitsee reaaliluvun x etäisyyttä nolasta. Merkintätapojen $\|\mathbf{x}\|$ ja $|\mathbf{x}|$ ero onkin vain sopimuskysymys: Ensimmäistä on sovittu käytettävän vektoreista, jälkimmäistä vain reaali- tai kompleksiluvuista.

Määritelmä 38 Vektoreiden $\mathbf{x}$ ja $\mathbf{y} \in \mathbb{R}^n$ välinen *kulma* tarkoittaa kulmaa $\theta \in [0, \pi]$, joka muodostuu origosta (pisteestä $\mathbf{0}$) pisteeseen $\mathbf{x}$ ja origosta pisteeseen $\mathbf{y}$ kulkevien janojen välille. Huomaa, että kolme pistettä $\mathbf{x}$, $\mathbf{y}$ ja $\mathbf{0}$ kuuluvat aina samalle tasolle, joten vektoreiden välinen kulma on aina *tason* ominaisuus.

Lause 18 *Jos θ on vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välinen kulma, niin*

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos \theta.$$

Todistus Tason kolmioita koskevasta kosinilauseesta saadaan

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos \theta = \|\mathbf{x} - \mathbf{y}\|^2.$$

Yhtälön oikea puoli voidaan edelleen kirjoittaa muotoon

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} - \mathbf{x} \cdot \mathbf{y} - \mathbf{y} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} \\ &= \mathbf{x} \cdot \mathbf{x} - 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x} \cdot \mathbf{y} \end{aligned}$$

Näistä saadaan yhtälö

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos \theta = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x} \cdot \mathbf{y},$$

mistä väite seuraa suoraan. □

Huomautus 22 Edellisen lauseen perusteella kahden vektorin $\mathbf{x}, \mathbf{y}$ (kumpikaan ei nollavektori) välinen kulma θ saadaan esityksestä

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

Cauchy-Schwarzin epäyhtälön mukaan joka tapauksessa on $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$, joten $\cos \theta \in [-1, 1]$ niin kuin pitääkin.

Esimerkki 82 Olkoon $\mathbf{x} = (1, 0, 2, -1)$ ja $\mathbf{y} = (0, 1, 1, -1) \in \mathbb{R}^4$. Tällöin $\|\mathbf{x}\|^2 = 1^2 + 0^2 + 2^2 + (-1)^2 = 6$ ja $\|\mathbf{y}\|^2 = 0^2 + 1^2 + 1^2 + (-1)^2 = 3$, sekä $\mathbf{x} \cdot \mathbf{y} = 1 \cdot 0 + 0 \cdot 1 + 2 \cdot 1 + (-1) \cdot (-1) = 3$. Tällöin siis vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välisen kulman θ kosini on

$$\cos \theta = \frac{3}{\sqrt{6}\sqrt{3}} = \frac{1}{\sqrt{2}},$$

josta $\theta = \frac{\pi}{4}$.

Määritelmä 39 Vektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat *ortogonaalit* eli *kohtisuorat* jos $(\mathbf{x}, \mathbf{y}) = 0$.

Edellisen määritelmän mukaan nollavektori on kohtisuorassa kaikkia muita vektoreita kohtaan: $(\mathbf{0}, \mathbf{x}) = 0$ aina. Jos taas kumpikaan vektoreista $\mathbf{x}, \mathbf{y}$ ei ole nollavektori, voi $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$ tulla nolaksi vain, jos vektoreiden välinen kulma θ toteuttaa $\cos \theta = 0$, mikä rajoituksella $\theta \in [0, \pi]$ tapahtuu vain jos $\theta = \frac{\pi}{2} = 90^\circ$. Geometrisesti vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ (kumpikaan ei nollavektori) ortogonaalisuus tarkoittaa siis sitä, että näiden välinen kulma on suora.

Määritelmä 40 Vektorijoukko $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ on *ortogonaali*, jos $(\mathbf{x}_i, \mathbf{x}_j) = 0$ aina kun $i \neq j$. Ortogonaali joukko $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ on *ortonormaali*, jos lisäksi $(\mathbf{x}_i, \mathbf{x}_i) = 1$.

Huomautus 23 Jokaisesta ortogonaalista joukosta joka ei sisällä nollavektoria voidaan helposti saada ortonormaali kertomalla kukin $\mathbf{x}_i$ norminsa käänteisluvulla $\sqrt{(\mathbf{x}_i, \mathbf{x}_i)}^{-1}$.

Esimerkki 83 $\mathbb{R}^n$:n luonnollinen kanta $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ on ortonormaali.

Lause 19 Olkoon V vektoriavaruus ja $A = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ joukko, joka ei sisällä nollavektoria ja on ortogonaali jonkin sisätulon $(\mathbf{x}, \mathbf{y})$ suhteen. Tällöin A on lineaarisesti riippumaton.

Todistus Oletetaan, että

$$c_1 \mathbf{x}_1 + \dots + c_n \mathbf{x}_n = \mathbf{0}$$

Laskemalla sisätulot puolittain saadaan

$$c_1 (\mathbf{x}_1, \mathbf{x}_i) + \dots + c_i (\mathbf{x}_i, \mathbf{x}_i) + \dots + c_n (\mathbf{x}_n, \mathbf{x}_i) = (\mathbf{0}, \mathbf{x}_i).$$

Ortogonaalisuuden perusteella ylläoleva sievenee muotoon

$$c_i (\mathbf{x}_i, \mathbf{x}_i) = 0,$$

josta $c_i = 0$.

□

Esimerkki 84 Olkoon $L^2[\alpha, \alpha + T] \cap C^0[\alpha, \alpha + T]$ kuten aiemmin ja

$$(f, g) = \int_{\alpha}^{\alpha+T} f \overline{g}$$

Hermiten sisätulo. Funktiot $e_n(x) = e^{\frac{2\pi i n}{T} x}$ ovat lineaarisesti riippumattomat, sillä ne muodostavat ortogonaalisen joukon.

Määritelmä 41 Olkoot $\mathbf{x}$ ja $\mathbf{y} \neq \mathbf{0}$ avaruuden $\mathbb{R}^n$ vektoreita. Tällöin vektorin $\mathbf{x}$ projektio vektorilla $\mathbf{y}$ tarkoittaa sitä origon ja pisteen $\mathbf{y}$ kautta kulkevan suoran $L(\mathbf{y})$ pistettä $\mathbf{x}'$, jolle $\mathbf{x} - \mathbf{x}'$ on kohtisuorassa vektoria $\mathbf{y}$ vastaan.

Selvitetään seuraavaksi, miten vektorin $\mathbf{x}$ projektio $\mathbf{x}'$ vektorille $\mathbf{y}$ lasketaan. Tätä varten voidaan aluksi todeta, että koska $\mathbf{x}'$ on suoran $L(\mathbf{y})$ piste, on määritelmän mukaan $\mathbf{x}' = c\mathbf{y}$ jollekin luvulle c . Koska lisäksi vaaditaan, että $\mathbf{x} - \mathbf{x}'$ on kohtisuorassa $\mathbf{y}$:tä vastaan, on

$$0 = (\mathbf{x} - \mathbf{x}') \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y} - c\mathbf{y} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y} - c\|\mathbf{y}\|^2.$$

Yllä olevasta yhtälöstä voidaan ratkaista $c = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2}$, joten siis vektorin $\mathbf{x}$ projektio $\mathbf{x}'$ vektorilla $\mathbf{y}$ voidaan kirjoittaa muotoon

$$\mathbf{x}' = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{y}.$$

3.2 Etäisyyden säilyttävät kuvaukset

Määritelmä 42 Lineaarikuvaus $f : V \rightarrow V$ säilyttää etäisyyden, jos $d(f(\mathbf{x}), f(\mathbf{y})) = d(\mathbf{x}, \mathbf{y}) \forall \mathbf{x}, \mathbf{y} \in V$. Lineaarikuvaus $f : V \rightarrow V$ säilyttää normin, jos $\|f(\mathbf{x})\| = \|\mathbf{x}\|$. Lineaarikuvaus $f : V \rightarrow V$ säilyttää sisätulon, jos $(f(\mathbf{x}), f(\mathbf{y})) = (\mathbf{x}, \mathbf{y})$.

Oletetaan tässä luvussa, että $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ ja että $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$. Ellei toisin mainita, oletetaan lisäksi, että sisätuloksi on valittu pistetulo tai Hermiten pistetulo.

Lause 20 Jos lineaarikuvaus f säilyttää sisätulon, niin se säilyttää myös normin ja etäisyyden.

Todistus Oletetaan, että f säilyttää sisätulon. Tällöin $\|f(\mathbf{x})\| = \sqrt{(f(\mathbf{x}), f(\mathbf{x}))} = \sqrt{(\mathbf{x}, \mathbf{x})}$ ja

$$d(f(\mathbf{x}), f(\mathbf{y})) = \|f(\mathbf{x}) - f(\mathbf{y})\| = \|f(\mathbf{x} - \mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| = d(\mathbf{x}, \mathbf{y}).$$

Määritelmä 43 $n \times n$ -matriisi A on ortogonaali, jos $A^T = A^{-1}$, siis $A^T A = A A^T = I$.

Lause 21 Jos lineaarikuvausten f matriisi A_f on ortogonaali, niin f säilyttää pistetulon ja etäisyyden.

Todistus Edellämämainitun perusteella riittää osoittaa, f säilyttää pistetulon. Jos $\mathbf{x}$ ymmärretään pystyvektoriksi, niin $\mathbf{y} = f(\mathbf{x}) = A_f \mathbf{x}$ ja näin ollen

$$(f(\mathbf{x}), f(\mathbf{y})) = (A_f \mathbf{x})^T A_f \mathbf{y} = \mathbf{x}^T A_f^T A_f \mathbf{y} = \mathbf{x}^T \mathbf{y} = (\mathbf{x}, \mathbf{y}).$$

Määritelmä 44 $n \times n$ -matriisin A adjugaatti A^* määritellään $A^* = \overline{A}^T$. Sanotaan, että matriisi A on unitaarinen, jos $A^* = A^{-1}$, siis $A^* A = A A^* = I$.

Huomautus 24 Matriisi

$$F = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ 1 & \rho^2 & \rho^4 & \dots & \rho^{(N-1)2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \rho^{N-1} & \rho^{2(N-1)} & \dots & \rho^{(N-1)(N-1)} \end{pmatrix}$$

on unitaarinen, sillä

$$F^* = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \rho^{-1} & \rho^{-2} & \dots & \rho^{-(N-1)} \\ 1 & \rho^{-2} & \rho^{-4} & \dots & \rho^{-(N-1)2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \rho^{-(N-1)} & \rho^{-2(N-1)} & \dots & \rho^{-(N-1)(N-1)} \end{pmatrix} = F^{-1},$$

kuten esimerkissä 48 todettiin.

Lause 22 Jos lineaarikuvauksen $f : \mathbb{C}^n \rightarrow \mathbb{C}^n$ matriisi A_f on unitaarinen, niin f säilyttää Hermiten pistetulon.

Todistus Todetaan ensin, että jos $\mathbf{x}$ ja $\mathbf{y}$ käsitetään pystyvektoreiksi, niin Hermiten pistetulo voidaan kirjoittaa muodossa $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \bar{\mathbf{y}}$. Täten

$$(f(\mathbf{x}), f(\mathbf{y})) = (A_f \mathbf{x}, A_f \mathbf{y}) = (A_f \mathbf{x})^T \overline{A_f \mathbf{y}} = \mathbf{x}^T A_f^T \overline{A_f \mathbf{y}} = \mathbf{x}^T \overline{A_f^* A_f \mathbf{y}} = \mathbf{x}^T \bar{\mathbf{y}} = (\mathbf{x}, \mathbf{y}).$$

Huomautus 25 Olkoot $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})^T$ ja $\mathbf{g} = (g_0, g_1, \dots, g_{N-1})^T \in \mathbb{C}^n$ datavektoreita ja $\mathbf{F} = \mathbf{F}\mathbf{f}$ sekä $\mathbf{G} = \mathbf{F}\mathbf{g}$ niiden Fourier-muunnokset. Yllämainitun perusteella $(\mathbf{F}, \mathbf{G}) = (\mathbf{f}, \mathbf{g})$, mikä voidaan kirjoittaa auki muotoon

$$F_0 \overline{G_0} + F_1 \overline{G_1} + \dots + F_{N-1} \overline{G_{N-1}} = f_0 \overline{g_0} + f_1 \overline{g_1} + \dots + f_{N-1} \overline{g_{N-1}}.$$

Tämä on *Parsevalin kaava* diskreeteille Fourier-muunnoksille.

Huomautus 26 Etäisyyden säilyttävät lineaarikuvaukset ovat avaruuden peilauksia tai kiertoja. Diskreetti Fourier-muunnos voidaan ymmärtää N -ulotteisen kompleksisen avaruuden kiertona tai vaihtoehtoisesti mutta matemaattisesti yhtäpitävästi kannan vaihtona.

3.3 Korrelaatiokerroin

Kun luonnontieteellistä teoriaa muodostetaan havaintoihin perustuen, on yleensä jostakin ilmiöstä kerätty havaintoaineistoa mittaamalla ilmiöön liittyviä eri suureita. Mittaustulosten avulla voidaan pyrkiä päättämään, riippuvatko eri suureet toisistaan jollakin tavalla. Riippuvuussuhteen esiintuominen on yleensä helppoa, mutta sen selvittäminen, johtuuko jokin asia toisesta, on useimmiten varsin hankalaa. Yleensä riippuvuussuhde näyttäytyy siten, että toisen suureen ollessa suuri on keskimäärin myös toinen suuri (tai käänteisen korrelaation tapauksessa pieni), mutta sen selvittäminen, onko molempien suureiden arvoille jokin ulkopuolinen selitys, ei ole mahdollista vain mitattuja suureita vertaamalla.

Esimerkki 85 Jos pyritään selvittämään, millä tavoin jossakin populaatiossa ihmisen pituus (keskimäärin) riippuu painosta, voidaan toimia seuraavasti: valitaan populaatiosta joukko koehenkilöitä, joiden painot $x_1, \dots, x_n$ ja pituudet $y_1, \dots, y_n$ mitataan. Tämän jälkeen voidaan pisteet $(x_1, y_1), \dots, (x_n, y_n)$ piirtää tasolle, ja katsoa onko kuviosta mahdollisesti löydettävissä jonkinlaista johdonmukaisuutta. Mitä paremmin pisteet (x_i, y_i) näyttävät keskittyvän jonkin nousevan suoran ympärille, sitä paremmin voidaan katsoa pituuden riippuvan painosta (lineaarisesti).

Suureiden y_i lineaarista riippuvuutta x_i :stä voidaan arvioida sen mukaan, kuinka ”vahvasti” vektorit $\mathbf{y} = (y_1, \dots, y_n)$ ja $\mathbf{x} = (x_1, \dots, x_n)$ ovat lineaarisesti riippuvat, ja tähän puolestaan arvioksi voidaan valita vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välinen kulma.

Jos nimittäin $\mathbf{y} = c\mathbf{x}$, ovat vektorit $\mathbf{x}$ ja $\mathbf{y}$ lineaarisesti riippuvat ja tällöin

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{\mathbf{x} \cdot c\mathbf{x}}{\|\mathbf{x}\| \|c\mathbf{x}\|} = \frac{c}{|c|} \frac{\mathbf{x} \cdot \mathbf{x}}{\|\mathbf{x}\|^2} = \frac{c}{|c|} = \begin{cases} 1, & \text{jos } c > 0 \\ -1, & \text{jos } c < 0. \end{cases}$$

Ensimmäisessä tapauksessa sanotaan, että koordinaattien y_i ja x_i välillä vallitsee *positiivinen korrelaatio* ja jälkimmäisessä tapauksessa *negatiivinen korrelaatio*. Tapauksessa, jossa $\mathbf{y} = c\mathbf{x}$ ei päde, on aina $|\cos \theta| < 1$ ja mahdollisimman ”riippumattomia” $\mathbf{x}$ ja $\mathbf{y}$ ovat toisistaan kun $\cos \theta = 0$.

Tämänkaltaisen tarkastelu ei kata tapausta $\mathbf{y} = c\mathbf{x} + \mathbf{b}$, missä $\mathbf{b} = (b, b, \dots, b)$ on vakiovektori. Kuitenkin myös tällöin riippuvuus vektoreiden välillä on täydellistä, ja tällöin voidaan menetellä seuraavasti: Sen sijaan, että laskettaisiin vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välisen kulman kosini, laskentaankin tämä kosini vektoreille $\mathbf{x}'$ ja $\mathbf{y}'$, missä $\mathbf{x}'$ saadaan vektorista $\mathbf{x}$ vähentämällä jokaisesta koordinaatista x_i koordinaattien keskiarvo $\mu_{\mathbf{x}}$:

$$\mathbf{x}' = (x_1 - \mu_{\mathbf{x}}, \dots, x_n - \mu_{\mathbf{x}})$$

ja samoin vektorille $\mathbf{y}$:

$$\mathbf{y}' = (y_1 - \mu_{\mathbf{y}}, \dots, y_n - \mu_{\mathbf{y}}),$$

missä $\mu_{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n x_i$ ja $\mu_{\mathbf{y}} = \frac{1}{n} \sum_{i=1}^n y_i$. Tämän korjauksen jälkeen sekä vektorin $\mathbf{x}'$ ja $\mathbf{y}'$ koordinaattien summa on nolla ja suorilla laskuilla voidaan todeta että

$$\mu_{\mathbf{y}} = c\mu_{\mathbf{x}} + b$$

ja

$$y'_i = y_i - c\mu_{\mathbf{x}} - b = cx_i + b - c\mu_{\mathbf{x}} - b = c(x_i - \mu_{\mathbf{x}}) = cx'_i,$$

jolloin siis $\mathbf{y}' = c\mathbf{x}'$.

Määritelmä 45 Vektoreiden $\mathbf{x} = (x_1, \dots, x_n)$ ja $\mathbf{y} = (y_1, \dots, y_n)$ välinen *Pearsonin korrelaatiokerroin* (tunnetaan myös nimellä *tulomomenttikerroin*) on vektoreiden $\mathbf{x}'$ ja $\mathbf{y}'$ välisen kulman kosini, missä $\mathbf{x}'$ ja $\mathbf{y}'$ on saatu vektoreista $\mathbf{x}$ ja $\mathbf{y}$ vähentämällä jokaisesta koordinaatista koordinaattien keskiarvo.

Esimerkki 86 Taulukossa on erälle kurssille osallistuneen 36 opiskelijan demoprosentit ja heidän tentissä saamansa pisteet.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Demo%	40	80	33	81	36	29	38	26	42	26	51	46	48	51	26	39	82	33
Pisteet	25	29	16	30	21	14	10	8	23	10	11	13	14	26	5	6	25	21

i	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Demo%	80	26	43	87	35	56	87	87	71	33	51	60	36	30	33	71	43	32
Pisteet	29	9	9	30	16	27	27	29	26	8	20	22	3	4	15	28	21	15

Lasketaan demoprosentin ja tenttipisteiden välinen korrelaatiokerroin. Tätä varten määritellään x_i = opiskelijan i demoprosentti ja y_i = opiskelijan i tenttipisteet. Tällöin siis $\mathbf{x}$ ja $\mathbf{y}$ ovat 36-ulotteisen avaruuden $\mathbb{R}^{36}$ vektoreita, joiden koordinaatteja korrelaatiokertoimen laskemiseksi aluksi siirretään: vektorin $\mathbf{x}$ koordinaattien keskiarvo on $\mu_{\mathbf{x}} = \frac{1}{36}(40 + 80 + 33 + \dots + 32) = \frac{1768}{36}$ ja vastaavasti $\mu_{\mathbf{y}} = \frac{1}{36}(25 + 29 + 16 + \dots + 15) = \frac{645}{36}$.

Uudet vektorit ovat siis $\mathbf{x}' = (40 - \mu_{\mathbf{x}}, 80 - \mu_{\mathbf{x}}, 33 - \mu_{\mathbf{x}}, \dots, 32 - \mu_{\mathbf{x}})$ ja $\mathbf{y}' = (25 - \mu_{\mathbf{y}}, 29 - \mu_{\mathbf{y}}, 16 - \mu_{\mathbf{y}}, \dots, 15 - \mu_{\mathbf{y}})$, ja näiden sisätulo voidaan laskea suoraviivaisesti:

$$\begin{aligned} \mathbf{x}' \cdot \mathbf{y}' &= (40 - \mu_{\mathbf{x}})(25 - \mu_{\mathbf{y}}) + (80 - \mu_{\mathbf{x}})(29 - \mu_{\mathbf{y}}) + \dots + (32 - \mu_{\mathbf{x}})(15 - \mu_{\mathbf{y}}) \\ &= 4880,333 \dots \end{aligned}$$

Samoin voidaan laskea

$$\|\mathbf{x}'\|^2 = \mathbf{x}' \cdot \mathbf{x}' = 14673,55556 \dots$$

ja

$$\|\mathbf{y}'\|^2 = \mathbf{y}' \cdot \mathbf{y}' = 2566,75 \dots$$

ja lopuksi korrelaatiokerroin saadaan lausekkeesta

$$\frac{\mathbf{x}' \cdot \mathbf{y}'}{\|\mathbf{x}'\| \|\mathbf{y}'\|} = 0,795 \dots$$

Korrelaatiota 0,795 pidetään yleensä hyvin voimakkaana. Yleensä itseisarvoltaan 0,5 ylittävä korrelaatiokerroin tulkitaan voimakkaaksi, mutta on huomattava että korrelaation kutsuminen voimakkaaksi on tulkinnanvaraista: Jos aineisto on saatu mittaamalla fysikaalisia suureita tarkoilla instrumenteilla, voi korrelaatiokertoimen edellyttää ylittävän 0,9 ennen kuin sitä kutsutaan vahvaksi, kun taas yhteiskuntatieteellisessä tutkimuksessa saadussa aineistossa korrelaatiokerroin 0,7 voidaan katsoa hyvin vahvaksi.

Huomautus 27 Havaintotapausten määrän on oltava kyllin suuri, jotta korrelaatiokertoimen tulkinalla olisi merkitystä. Jos esimerkiksi aineistossa on vain kaksi tapausta, joita vastaavat (siiretyt) vektorit $\mathbf{x}' = (x, -x)$ ja $\mathbf{y}' = (y, -y) \in \mathbb{R}^2$, on näillä aina täydellinen korrelaatio joko 1 tai -1 . Yleensä halutaan, että havaintotapausten määrä on vähintään kymmenen, mutta tämäkin on tulkinnanvaraista. Joissakin tilanteissa voidaan kolmeakymmentä havaintotapausta pitää suurena määränä, kun taas esimerkiksi lääkkeen haittavaikutuksia etsittäessä kolmeakymmentä havaintotapausta pidetään aivan riittämättömänä joukkona.

Havaintojoukon koolle on olemassa tilastomatemattisia vaatimuksia, jotka eivät kuitenkaan kuulu tämän kurssin oppimäärään.

Luku 4

Kolmiulotteisen avaruuden geometriaa

Tässä luvussa tutustutaan avaruuden $\mathbb{R}^3$ ”kaksiulotteisiin” ja ”yksiulotteisiin” osajoukkoihin, joista yksinkertaisimmat ovat *tasot* ja *suorat*. Aluksi kuitenkin perehdytään erityiseen kolmiulotteisen avaruuden operaatioon, *ristituloon*.

4.1 Ristitulo

Kuten aiemmin huomattiin, vektoreiden pistetulo (sisätulo) voidaan määritellä miten korkeaulotteiselle avaruudelle $\mathbb{R}^n$ hyvänsä. Sen sijaan tässä kappaleessa käsiteltävä *ristitulo* on pelkästään kolmiulotteisen avaruuden $\mathbb{R}^3$ ominaisuus.

Määritelmä 46 Avaruuden $\mathbb{R}^3$ vektoreiden $\mathbf{x} = (x_1, x_2, x_3)$ ja $\mathbf{y} = (y_1, y_2, y_3)$ *ristitulo* on $\mathbb{R}^3$:n vektori

$$\mathbf{x} \times \mathbf{y} = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1).$$

Ristitulolle voidaan esittää ”lauseke” kolmirivisen determinantin avulla: jos $\mathbf{x} = (x_1, x_2, x_3)$ ja $\mathbf{y} = (y_1, y_2, y_3)$, niin

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \quad (4.1)$$

Tässä yhteydessä ”lauseke” esiintyy lainausmerkeissä, koska determinantti on määritelty vain siinä tapauksessa, että kaikki alkioit ovat lukuja – yllä olevassa ”lausekkeessa” kuitenkin ensimmäisen rivin alkioit ovat $\mathbb{R}^3$:n luonnollisen kannan vektorit $\mathbf{i}, \mathbf{j}, \mathbf{k}$, eikä tällaista determinanttia ole määritelty.

Tästä huolimatta (4.1) voi toimia muistisääntönä:

$$\begin{aligned} \mathbf{x} \times \mathbf{y} &= \mathbf{i} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - \mathbf{j} \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \mathbf{k} \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \\ &= \mathbf{i}(x_2 y_3 - x_3 y_2) - \mathbf{j}(x_1 y_3 - x_3 y_1) + \mathbf{k}(x_1 y_2 - x_2 y_1) \\ &= (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1). \end{aligned}$$

Lause 23 (Skalaarikolmitulo) Olkoot $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$ ja $\mathbf{z} = (z_1, z_2, z_3)$. Tällöin

1. $\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$
2. $\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) = \mathbf{y} \cdot (\mathbf{z} \times \mathbf{x}) = \mathbf{z} \cdot (\mathbf{x} \times \mathbf{y})$

Todistus Ensimmäinen kohta saadaan piste- ja ristitulojen määritelmästä suoraan laskemalla, toinen kohta taas seuraa ensimmäisestä, sillä determinantin merkki muuttuu kun kahden rivin paikkaa vaih-

detaan (mieti kuinka monta kahden rivin vaihtoa tarvitaan, kun determinantti $\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$ muutetaan determinantiksi $\begin{vmatrix} y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \\ x_1 & x_2 & x_3 \end{vmatrix}$). □

Huomautus 28 Skalaarikolmitulo voidaan tulkita sellaisen suuntaissärmiön tilavuutena, jonka yksi kärjistä sijaitsee origossa, ja kolme sivua ovat vektoreiden $\mathbf{x}$, $\mathbf{y}$ ja $\mathbf{z}$ suuntaiset.

Lause 24 Ristitulo toteuttaa seuraavat säännöt:

1. $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$
2. $(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}$
3. $(a\mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (a\mathbf{y}) = a(\mathbf{x} \times \mathbf{y})$
4. $\mathbf{x} \times \mathbf{x} = \mathbf{0}$
5. $\mathbf{x} \cdot (\mathbf{x} \times \mathbf{y}) = 0$ ja $\mathbf{y} \cdot (\mathbf{x} \times \mathbf{y}) = 0$ (ristitulo on kohtisuorassa alkuperäisiin vektoreihin nähden).
6. $\|\mathbf{x} \times \mathbf{y}\|^2 + (\mathbf{x} \cdot \mathbf{y})^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$.
7. $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta$, missä θ on vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välinen kulma.

Todistus Kohdat 1-3 seuraavat ristitulon määritelmästä suoraan laskemalla. Kohta 4 seuraa edellisestä lauseesta, sillä determinantti, jossa on kaksi samaa riviä, on nolla. Samoin kohta 5 saadaan edellisestä lauseesta sijoittamalla $\mathbf{y} = \mathbf{x}$, jolloin todetaan, että $\mathbf{x} \times \mathbf{x}$ on kohtisuorassa kaikkia vektoreita $\mathbf{z}$ vastaan, jolloin siis $\mathbf{x} \times \mathbf{x} = \mathbf{0}$. Väittämä 6 seuraa myös suoralla laskulla:

$$\begin{aligned} \|\mathbf{x} \times \mathbf{y}\|^2 + (\mathbf{x} \cdot \mathbf{y})^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\ &\quad + (x_1y_1 + x_2y_2 + x_3y_3)^2 \\ &= x_2^2y_3^2 - 2x_2x_3y_2y_3 + x_3^2y_2^2 \\ &\quad + x_3^2y_1^2 - 2x_1x_3y_1y_3 + x_1^2y_3^2 \\ &\quad + x_1^2y_2^2 - 2x_1x_2y_1y_2 + x_2^2y_1^2 \\ &\quad + x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2 \\ &\quad + 2x_1x_2y_1y_2 + 2x_1x_3y_1y_3 + 2x_2x_3y_2y_3 \\ &= x_1^2y_1^2 + x_1^2y_2^2 + x_1^2y_3^2 \\ &\quad + x_2^2y_1^2 + x_2^2y_2^2 + x_2^2y_3^2 \\ &\quad + x_3^2y_1^2 + x_3^2y_2^2 + x_3^2y_3^2 \\ &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \end{aligned}$$

Väittämää 7 varten muistetaan, että vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ väliselle kulmalle θ pätee $\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$. Väittämän 6 mukaan voidaan kirjoittaa

$$\begin{aligned} \|\mathbf{x} \times \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \left(1 - \left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}\right)^2\right) \\ &= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 (1 - \cos^2 \theta) = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \sin^2 \theta, \end{aligned}$$

mistä väittämä 7 seuraa suoraan. □

4.2 Tasot

Kahden lineaarisesti riippumattoman vektorin $\mathbf{x}$ ja $\mathbf{y}$ generoima kaksiulotteinen aliavaruus

$$L(\mathbf{x}, \mathbf{y}) = \{c_1\mathbf{x} + c_2\mathbf{y} \mid c_1, c_2 \in \mathbb{R}\} \subseteq \mathbb{R}^3$$

vastaa geometrisesti origon kautta kulkevaa tasoa, johon pisteet $\mathbf{x}$ ja $\mathbf{y}$ kuuluvat.

Määritelmä 47 Olkoon V jokin avaruuden $\mathbb{R}^n$ aliavaruus ja $\mathbf{z} \in \mathbb{R}^n$ jokin vektori. Tällöin joukkoa

$$\mathbf{z} + V = \{\mathbf{z} + \mathbf{w} \mid \mathbf{w} \in V\}$$

kutsutaan vektorin $\mathbf{z}$ määräämäksi aliavaruuden *sivuluokaksi*.

Olkoot $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ lineaarisesti riippumattomia ja merkitään $T = L(\mathbf{x}, \mathbf{y})$ kuten edellä. Olkoon lisäksi $\mathbf{z} \in \mathbb{R}^3$. Tällöin sivuluokka $\mathbf{z} + T$ merkitsee geometrisesti T :n suuntaista tasoa, joka kulkee pisteen $\mathbf{z}$ kautta. Kaikki avaruuden $\mathbb{R}^3$ tasot voidaan esittää kaksiulotteisten aliavaruuksien sivuluokkina $\mathbf{z} + T$. Jatkossa tullaan löytämään myös muita esitystapoja $\mathbb{R}^3$:n tasoille.

Myös korkeampiulotteisissa avaruuksissa $\mathbb{R}^n$ taso määritellään joukkona

$$\{\mathbf{z} + c_1\mathbf{x} + c_2\mathbf{y} \mid c_1, c_2 \in \mathbb{R}\},$$

missä $\mathbf{x}$ ja $\mathbf{y} \in \mathbb{R}^n$ ovat lineaarisesti riippumattomia ja $\mathbf{z} \in \mathbb{R}^n$ on jokin vektori. Tämän perusteella on selvää, avaruuden $\mathbb{R}^n$ taso ei poikkea milteenkään oleellisin osin avaruuden $\mathbb{R}^3$ tasosta: Kun $\mathbf{x}, \mathbf{y}$ ja $\mathbf{z}$ on kiinnitetty, tason piste määräytyy yksikäsitteisesti lukujen c_1 ja c_2 arvoista. Näitä lukuja voidaan pitää vektoreiden $\mathbf{x}, \mathbf{y}$ ja $\mathbf{z}$ määräämän tason pisteen *koordinaatteina*. Tällä tavoin määräytyy yksikäsitteinen vastaavuus pisteen $(c_1, c_2) \in \mathbb{R}^2$ ja tason pisteen $\mathbf{z} + c_1\mathbf{x} + c_2\mathbf{y} \in \mathbb{R}^n$ välille.

Esimerkki 87 Jos $\mathbf{x} \in \mathbb{R}^3$ ei ole nollavektori, on joukko $\{\mathbf{x}\}$ lineaarisesti riippumaton ja geometrisesti yksiulotteinen aliavaruus

$$L(\mathbf{x}) = \{c\mathbf{x} \mid c \in \mathbb{R}\} \subseteq \mathbb{R}^3$$

merkitsee origon kautta kulkevaa suoraa. Avaruuden $L = L(\mathbf{x})$ sivuluokka

$$\mathbf{z} + L = \{\mathbf{z} + c\mathbf{x} \mid c \in \mathbb{R}\} \quad (4.2)$$

puolestaan vastaa L :n suuntaista suoraa, joka kulkee pisteen $\mathbf{z}$ kautta. Avaruuden $\mathbb{R}^3$ kaikki suorat voidaan esittää yksiulotteisten aliavaruuksien sivuluokkina ja jatkossa tutustutaan myös muihin suorien esitystapoihin.

Korkeampiulotteisissa avaruuksissa $\mathbb{R}^n$ suora määritellään joukkona (4.2), missä $\mathbf{x} \in \mathbb{R}^n$ on nollavektorista eroava vektori ja $\mathbf{z} \in \mathbb{R}^n$ on jokin vektori. Kuten tasojen tapauksessa, on selvää, että suora avaruudessa $\mathbb{R}^n$ on oleellisesti samankaltainen objekti kuin suora avaruudessa $\mathbb{R}^3$.

Aiemmin on todettu, että avaruuden $\mathbb{R}^3$ taso vastaa kaksiulotteisen aliavaruuden T sivuluokkaa

$$\mathbf{r} + T = \{\mathbf{r} + \mathbf{w} \mid \mathbf{w} \in T\},$$

ja kaksiulotteisen aliavaruuden T (jota geometrisesti vastaa origon kautta kulkeva taso) puolestaan generoi mikä hyvänsä lineaarisesti riippumaton vektoripari $\{\mathbf{x}, \mathbf{y}\}$:

$$T = \{c_1\mathbf{x} + c_2\mathbf{y} \mid c_1, c_2 \in \mathbb{R}\}. \quad (4.3)$$

Esitysmuodolla (4.3) saadaan origon kautta kulkevat tasot. Kaikki tasot saadaan seuraavasti:

Tason parametriesitys: Olkoot $\mathbf{r}, \mathbf{x}$ ja $\mathbf{y}$ avaruuden $\mathbb{R}^3$ vektoreita. Tällöin esitystä

$$T = \{\mathbf{r} + c_1\mathbf{x} + c_2\mathbf{y} \mid c_1, c_2 \in \mathbb{R}\}, \quad (4.4)$$

kutsutaan tason T *parametriesitykseksi*. Lukuja c_1 ja c_2 kutsutaan *parametreiksi*. Yhtä hyvin paria $(c_1, c_2) \in \mathbb{R}^2$ voidaan kutsua parametrikseksi. Kun parametri (c_1, c_2) käy läpi avaruuden $\mathbb{R}^2$, käy $\mathbf{r} + c_1\mathbf{x} + c_2\mathbf{y}$ läpi kaikki tason T pisteet. Vektoria $\mathbf{r}$ kutsutaan tason *paikkavektori*ksi ja vektoreita $\mathbf{x}$ ja $\mathbf{y}$ sen *suuntavektoreiksi*. Jos $\mathbf{x}$ ja $\mathbf{y}$ ovat lineaarisesti riippuvat, ei (4.4) määrittele tasoa, vaan suoran tai pelkästään yhden pisteen.

Huomautus 29 Parametriesityksen (4.4) avulla voidaan määritellä taso yleisestikin avaruudessa $\mathbb{R}^n$, vaikka n olisi kuinka suuri tahansa. Sen sijaan muut tässä pykälässä esitettävät muodot tasolle

koskevat vain kolmiulotteista avaruutta. Parametriesitys voidaan myös tulkita funktioksi $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f(c_1, c_2) = \mathbf{r} + c_1\mathbf{x} + c_2\mathbf{y}$, jolloin tason $\mathbb{R}^3$:ssa saadaan $\mathbb{R}^2$:n kuvana.

Avaruuden $\mathbb{R}^3$ tasolle on olemassa myös muita esitysmuotoja, joista tärkein on *normaalivektorin* määrittelemä muoto: Tasoa T vastaan kohtisuorassa olevaa vektoria $\mathbf{n}$ kutsutaan tason *normaalivektoriksi*. Jos T on origon kautta kulkeva taso, saadaan normaalivektorin avulla tason esitysmuodoksi

$$T = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} \cdot \mathbf{n} = 0\}, \quad (4.5)$$

sillä sisätulo $\mathbf{x} \cdot \mathbf{n}$ on nolla tarkalleen silloin, kun $\mathbf{x}$ ja $\mathbf{n}$ ovat kohtisuorat.

Yleisempi muoto, jossa T ei välttämättä kulje origon vaan pisteen $\mathbf{x}_0 \in \mathbb{R}^3$ kautta, voidaan kirjoittaa seuraavaan muotoon:

Tason normaalimuoto: Esitystä

$$T = \{\mathbf{x}_0 + \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^3, \mathbf{x} \cdot \mathbf{n} = 0\} = \{\mathbf{x} \in \mathbb{R}^3 \mid (\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{n} = 0\} \quad (4.6)$$

kutsutaan tason *normaalimuodoksi*. $\mathbf{x}_0$ on tason *paikkavektori* ja $\mathbf{n}$ *normaalivektori*.

Jos merkitään $\mathbf{x} = (x, y, z)$, $\mathbf{n} = (a, b, c)$ ja $\mathbf{x}_0 = (x_0, y_0, z_0)$, voidaan yhtälö $(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{n} = 0$ kirjoittaa sisätulon määritelmän mukaan muotoon

$$(x - x_0)a + (y - y_0)b + (z - z_0)c = 0,$$

mikä merkitä $d = ax_0 + by_0 + cz_0 = \mathbf{n} \cdot \mathbf{x}_0$ käyttäen saadaan seuraava esitysmuoto:

Tason koordinaattimuoto: Muotoa

$$T = \{(x, y, z) \mid ax + by + cz = d\}, \quad (4.7)$$

kutsutaan tason *koordinaattimuodoksi* (toisinaan myös normaalimuodoksi).

Avaruuden $\mathbb{R}^3$ tasolle on olemassa vielä yksi luonteva esitystapa:

Kolme pistettä: Olkoot $\mathbf{x}_1$, $\mathbf{x}_2$ ja $\mathbf{x}_3$ avaruuden $\mathbb{R}^3$ pisteitä, jotka eivät ole samalla suoralla. Tällöin ne määrittävän avaruuden $\mathbb{R}^3$ tason. Pisteet $\mathbf{x}_1$, $\mathbf{x}_2$ ja $\mathbf{x}_3$ ovat samalla suoralla tarkalleen silloin kun $\mathbf{x}_2 - \mathbf{x}_1$ ja $\mathbf{x}_3 - \mathbf{x}_1$ ovat lineaarisesti riippuvat, mikä voidaan selvittää esimerkiksi Gaussin-Jordanin menetelmällä. Jos pisteet ovat samalla suoralla, sanotaan että ne ovat *kollineariset*, muutoin *epäkollineariset*.

Tarkastellaan seuraavaksi esimerkkien kautta miten eri esitysmuodoista päästään toiseen.

Esimerkki 88 (parametrimuodosta normaalimuotoon)

Olkoon $\mathbf{r} = (1, 0, 3)$ tason paikkavektori ja $\mathbf{x} = (-1, 2, 2)$ ja $\mathbf{y} = (-3, 2, 1)$ sen suuntavektorit. Tarkastetaan aluksi, että $\mathbf{x}$ ja $\mathbf{y}$ ovat lineaarisesti riippumattomat. Jos

$$c(-1, 2, 2) + d(-3, 2, 1) = (0, 0, 0),$$

saadaan yhtälöryhmä

$$\begin{cases} -c - 3d = 0 \\ 2c + 2d = 0 \\ 2c + d = 0, \end{cases}$$

jonka ainoa ratkaisu on $c = d = 0$, kuten helposti todetaan (esimerkiksi Gaussin-Jordanin menetelmällä). Täten

$$\{(1, 0, 3) + c(-1, 2, 2) + d(-3, 2, 1) \mid c, d \in \mathbb{R}\}$$

todella esittää tasoa kolmiulotteisessa avaruudessa. Normaalimuoto tälle tasolle saadaan etsimällä jokin vektori, joka on kohtisuorassa kumpaakin paikkavektoria vastaan, esimerkiksi $\mathbf{n} = \mathbf{x} \times \mathbf{y}$ kelpaa.

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 2 & 2 \\ -3 & 2 & 1 \end{vmatrix} = \mathbf{i} \begin{vmatrix} 2 & 2 \\ 2 & 1 \end{vmatrix} - \mathbf{j} \begin{vmatrix} -1 & 2 \\ -3 & 1 \end{vmatrix} + \mathbf{k} \begin{vmatrix} -1 & 2 \\ -3 & 2 \end{vmatrix} = -2\mathbf{i} - 5\mathbf{j} + 4\mathbf{k} = (-2, -5, 4).$$

Normaalimuotoa varten tarvitaan myös jokin tason vektori $\mathbf{x}_0$ ja paikkavektori $(1, 0, 3)$ kelpaa tähän erinomaisesti. Tällöin tason normaalimuoto on

$$(-2, -5, 4) \cdot ((x, y, z) - (1, 0, 3)) = 0. \quad (4.8)$$

Esimerkki 89 (normaalimuodosta koordinaattimuotoon)

Esituksen (4.8) muuntaminen koordinaattimuotoon on suoraviivaista, sillä koordinaattimuoto ja normaalimuoto ovat lähes samat. Yhtälössä (4.8) tarvitsee vain laskea sisätulot, jolloin saadaan

$$-2x - 5y + 4z - 10 = 0,$$

siis koordinaattimuoto on

$$-2x - 5y + 4z = 10. \quad (4.9)$$

Huomautus 30 Koordinaattimuodon kertoimista voidaan suoraan lukea tason normaalivektori.

Esimerkki 90 (koordinaattimuodosta kolmeen pisteeseen) Jos tasolle tunnetaan koordinaattimuoto (89), voidaan helposti valita tasolta kolme pistettä. Valitaan esimerkiksi ensin $x = y = 0$, jolloin yhtälön (89) mukaan $z = \frac{5}{2}$. Täten esimerkiksi $\mathbf{x}_1 = (0, 0, \frac{5}{2})$ kuuluu tasolle. Samoin valitsemalla $x = z = 0$ nähdään, että $\mathbf{x}_2 = (0, -2, 0)$ kuuluu tasolle ja lopuksi valitsemalla $y = z = 0$ saadaan tason piste $\mathbf{x}_3 = (-5, 0, 0)$.

Tarkistetaan vielä, etteivät näin saadut pisteet ole samalla suoralla. Tätä varten muodostetaan $\mathbf{x}_3 - \mathbf{x}_1 = (-5, 0, -\frac{5}{2})$ ja $\mathbf{x}_2 - \mathbf{x}_1 = (0, -2, -\frac{5}{2})$, joiden riippumattomuus voidaan helposti todeta.

Huomautus 31 Erityisesti tason parametriesityksestä on helppo löytää kolme epäkollineaarista pistettä. Voidaan nimittäin valita esimerkiksi molemmat parametrit ensin nolliksi, sen jälkeen vuoron perään toinen ykköseksi ja toinen nollassi.

Esimerkki 91 (kolmesta pisteestä parametrimuoto)

Etsitään parametrimuoto tasolle, jonka määrittävät edellisen esimerkin pisteet $\mathbf{x}_1$, $\mathbf{x}_2$ ja $\mathbf{x}_3$. On mahdollista valita mikä hyvänsä näistä pisteistä paikkavektoriksi, esim $\mathbf{r} = \mathbf{x}_1 = (0, 0, \frac{5}{2})$, jolloin suuntavektoreiksi voidaan valita $\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1 = (0, -2, -\frac{5}{2})$ ja $\mathbf{x}_3 - \mathbf{x}_1 = (-5, 0, -\frac{5}{2})$.

Tasojen esitykset eivät yleensä ole yksikäsitteisiä. Parametriesityksessä paikkavektori ja suuntavektorit voidaan valita usealla tavalla. Normaaliesityksessä normaalin pituus voidaan valita vapaasti ja suunta vaihtaa vastakkaiseksi. Tason kolme epäkollineaarista pistettä voidaan myös valita äärettömän monella tavalla, ja koordinaattiesityksen yhtälö voidaan kertoa millä hyvänsä nolasta poikkeavalla luvulla.

Tällöin voidaan kysyä millä perusteilla tasot voidaan havaita samoiksi. Kahden tason havaitseminen yhtäsuuriksi voidaan kyllä perustaa mihin esitykseen hyvänsä, mutta suoraviivaisinta se lienee tehdä koordinaattiesitykseen nojautuen.

Esimerkki 92 Selvitetään, onko esimerkissä (91) saatu parametrimuotoinen taso todella sama kuin esimerkin (88) taso. Tätä varten etsitään molemmille tasoille koordinaattimuoto. Esimerkin (88) koordinaattimuoto on jo selvitetty esimerkissä (4.9): koordinaattiyhtälö

$$-2x - 5y + 4z = 10$$

saadaan yksikäsitteiseen muotoon jakamalla se luvulla 10:

$$-\frac{1}{5}x - \frac{1}{2}y + \frac{2}{5}z = 1. \quad (4.10)$$

Toisen tason koordinaattimuotoa varten määrätään ensin normaalivektori:

$$\mathbf{n} = \mathbf{x} \times \mathbf{y} = (0, -2, -\frac{5}{2}) \times (-5, 0, -\frac{5}{2}) = (5, \frac{25}{2}, -10).$$

Normaalimuodoksi saadaan siis

$$(5, \frac{25}{2}, -10) \cdot ((x, y, z) - (0, 0, \frac{5}{2})) = 0,$$

mistä pistetulo laskemalla saadaan

$$5x + \frac{25}{2}y - 10z = -25.$$

Jakamalla luvulla -25 saadaan

$$-\frac{1}{5} - \frac{1}{2}y + \frac{2}{5}z = 1,$$

sama kuin (4.10).

Esimerkki 93 Tarkastellaan, millä ehdoilla joukko

$$\{(ax + by + c, dx + ey + f, gx + hy + i) \mid x, y \in \mathbb{R}\} \quad (4.11)$$

Määrittelee tason $\mathbb{R}^3$:ssa. Todetaan, että

$$\begin{aligned} & (ax + by + c, dx + ey + f, gx + hy + i) \\ &= (ax, dx, gx) + (by, ey, hy) + (c, f, i) \\ &= x(a, d, g) + y(b, e, h) + (c, f, i), \end{aligned}$$

joten joukko (4.11) määrittelee tason tarkalleen silloin kun vektorit (a, d, g) ja (b, e, h) ovat lineaarisesti riippumattomat. Jos (a, d, g) ja (b, e, h) ovat lineaarisesti riippuvat, määrittelee (4.11) suoran (mikäli $(a, d, g) \neq (0, 0, 0)$).

Määritelmä 48 Tasojen L_1 ja L_2 välisellä kulmalla tarkoitetaan tasojen normaalivektoreiden välistä kulmaa.

Esimerkki 94 Olkoon taso L_1 määritelty yhtälöllä $13x + 5y - 7z = 67$ ja L_2 yhtälöllä $2x - 2y + 3z = -5$. Tällöin tasojen L_1 ja L_2 normaalivektoreiksi voidaan valita $\mathbf{n}_1 = (13, 5, -7)$ ja $\mathbf{n}_2 = (2, -2, 3)$. Tasojen L_1 ja L_2 väliseksi kulma θ voidaan täten määrittää seuraavasti:

$$\cos \theta = \frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{\|\mathbf{n}_1\| \|\mathbf{n}_2\|} = \frac{-5}{\sqrt{243} \sqrt{17}} \approx -0,078,$$

josta $\theta \approx 1,65 = 85,5^\circ$.

4.3 Suorat

Avaruudessa $\mathbb{R}^n$ origon kautta kulkeva, vektorin $\mathbf{x}$ suuntainen suora vastaa yksiulotteista aliavaruutta

$$L = \{c\mathbf{x} \mid c \in \mathbb{R}\}, \quad (4.12)$$

ja muut suoran L suuntaiset suorat avaruudessa $\mathbb{R}^3$ saadaan L :n sivuluokkina

Suoran parametrimuoto: Olkoot $\mathbf{r}$ ja $\mathbf{x} \in \mathbb{R}^3$. Tällöin esitystä

$$L = \{\mathbf{r} + c\mathbf{x} \mid c \in \mathbb{R}\} \quad (4.13)$$

kutsutaan suoran *parametrimuodoksi* ja lukua c *parametriksi*. Jos $\mathbf{x} = \mathbf{0}$, ei esitys (4.13) määrittele suoraa, vaan yhden pisteen.

Huomautus 32 Kuten taso $\mathbb{R}^3$:ssa, voidaan myös suora $\mathbb{R}^3$:ssa käsittää alempiulotteisen avaruuden kuvana ”lähes” lineaarisessa kuvauksessa. Jos nimittäin määritellään $f : \mathbb{R} \rightarrow \mathbb{R}^3$ ehdolla $f(c) = \mathbf{r} + c\mathbf{x}$, saadaan suora (4.13) avaruuden $\mathbb{R}$ kuvana.

Suora on määriteltävissä parametrimuodon avulla miten korkeaulotteisissa avaruuksissa hyvänsä, mutta tässä yhteydessä käsitellään vain kolmiulotteista avaruutta $\mathbb{R}^3$. Tällöin voidaan löytää myös muita esitystapoja.

Avaruuden $\mathbb{R}^3$ tasot voidaan määrätä *yhden* lineaarisen yhtälön $ax + by + cz = d$ perusteella, sillä yhtälö rajoittaa $\mathbb{R}^3$:n vektorin (x, y, z) vapaiden komponenttien määrää yhdellä, ja näin saatava joukko on kaksiulotteinen $\mathbb{R}^3$:n osajoukko. Avaruuden $\mathbb{R}^3$ *suorien* tapauksessa puolestaan tarvitaan oleellisesti *kaksi* lineaarista rajoitetta koordinaateille.

Merkitään $\mathbf{r}_0 = (x_0, y_0, z_0)$ ja $\mathbf{x} = (a, b, c)$, jolloin avaruudessa $\mathbb{R}^3$ kulkevan suoran L parametri-muoto on

$$L = \{(x_0 + ta, y_0 + tb, z_0 + tc) \mid t \in \mathbb{R}\}.$$

Jos kaikki suuntavektorin $\mathbf{x} = (a, b, c)$ koordinaatit ovat nolasta poikkeavia, voidaan parametri t ratkaista yhtälöistä $(x, y, z) = (x_0 + ta, y_0 + tb, z_0 + tc)$, jolloin saadaan suoralle L seuraava esitysmuoto:

Suoran koordinaattimuoto:

$$L = \{(x, y, z) \in \mathbb{R}^3 \mid \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}\}$$

Suoran koordinaattimuotoa sanotaan myös *standardimuodoksi*. Jos esimerkiksi $ab \neq 0$, mutta $c = 0$, on suora xy -tason suuntainen ja sen koordinaattimuodoksi saadaan

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}, \quad z = z_0.$$

Ilmeinen tapa määritellä suora on myös erittäin käyttökelpoinen:

Kaksi pistettä: Avaruuden $\mathbb{R}^3$ kaksi pistettä $\mathbf{x}_1 \neq \mathbf{x}_2$ määrittää yksikäsitteisen suoran, joka kulkee näiden kautta.

Toisinaan voidaan suora esittää myös seuraavalla tavalla:

Kahden tason leikkaus: Jos tasot T_1 ja T_2 eivät ole samansuuntaiset, on niiden leikkaussuora yksikäsitteinen.

Tarkastellaan jälleen esimerkkien valossa miten eri esitysmuodosta päästään toiseen. Koska yleensä ei tarvitse keksiä suoran esitystä kahden tason leikkauksena (itse asiassa tehtävä on ilmeinen jos suorasta on annettu kaksi pistettä: lisätään vain kaksi erillistä pistettä, kumpikin epäkollineaarinen alkuperäisten kanssa, jolloin saadaan kahden eri tason esitystä), jätetään se esitysmuotojen ”syklisiä” pois ja tarkastellaan lopuksi vain miten kahden tason leikkauksesta saadaan suoran parametrimuoto.

Esimerkki 95 (Parametrimuodosta koordinaattimuotoon) Olkoon $\mathbf{r} = (1, 2, 2)$ suoran paikkavektori ja $\mathbf{x} = (2, -3, 1)$ sen suuntavektori. Tällöin suoralle

$$L = \{(1, 2, 2) + t(2, -3, 1) \mid t \in \mathbb{R}\}$$

saadaan koordinaattimuodoksi merkitsemällä

$$(x, y, z) = (1, 2, 2) + t(2, -3, 1) = (1 + 2t, 2 - 3t, 2 + t)$$

ja ratkaisemalla tästä t kunkin koordinaatin lausekkeena:

$$t = \frac{x-1}{2} = \frac{y-2}{-3} = z-2 \quad (4.14)$$

Huomautus 33 Koordinaattimuodosta voidaan helposti lukea suoran paikkavektori (osoittajien vakiot) ja suuntavektori (nimittäjien vakiot).

Esimerkki 96 (Koordinaattimuodosta kaksi pistettä) Yhtälöissä (4.14) voidaan valita t :lle mikä hyvänsä arvo, esim. $t = 0$ tai $t = 1$. Vastaava piste (x, y, z) löytyy ratkaisemalla 1. asteen yhtälö. Jos esimerkiksi $t = 0$ yhtälössä (4.14), on $(x, y, z) = (1, 2, 2)$ ja arvo $t = 1$ antaa $(x, y, z) = (3, -1, 3)$.

Huomautus 34 Parametrimuodosta kahden eri pisteen löytäminen on vielä yksinkertaisempaa: riittää valita parametrin arvoiksi esim 0 ja 1.

Esimerkki 97 (Kahdesta pisteestä parametrimuotoon) Määritetään pisteiden $\mathbf{x}_1 = (3, -1, 3)$ ja $\mathbf{x}_2 = (1, 2, 2)$ kautta kulkevan suoran parametrimuoto. Suuntavektoriksi voidaan valita $\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1 = (-2, 3, -1)$ ja paikkavektoriksi $\mathbf{r} = \mathbf{x}_1 = (3, -1, 3)$. Tällöin suoran parametriesitys on

$$L = \{\mathbf{r} + t\mathbf{x} \mid t \in \mathbb{R}\} = \{(3 - 2t, -1 + 3t, 3 - t) \mid t \in \mathbb{R}\}.$$

Esimerkki 98 (Kahden tason leikkauksesta parametrimuotoon) Olkoot $T_1: 2x + y - z = 2$ ja $T_2: 7x + 3y - 5z = 3$. Koska tasojen T_1 ja T_2 normaalivektorit $(2, 1, -1)$ ja $(7, 3, -5)$ eivät ole samansuuntaiset (miksi?), tasot T_1 ja T_2 leikkaavat. Tasojen leikkaussuoralle saadaan parametrimuoto esimerkiksi käyttämällä Gaussin-Jordanin eliminointimenetelmää yhtälöpariin

$$\begin{cases} 2x + y - z = 2 \\ 7x + 3y - 5z = 3 \end{cases}$$

saadaan pari

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

josta voidaan kirjoittaa ratkaisuksi $(x, y, z) = (2z - 3, -3z + 8, z) = z(2, -3, 1) + (-3, 8, 0)$, missä $z \in \mathbb{R}$. Tämä on suoran parametrimuoto: paikkavektorina toimii $(-3, 8, 0)$, suuntavektorina $(2, -3, 1)$ ja parametrina z .

Kuten tasoja, myös suoria koskevat kysymykset (leikkaus, yhdensuuntaisuus) palautuvat yhtälöitä koskeviin algebrallisiin kysymyksiin. Tarkastellaan taas kysymystä, milloin kaksi suoraa ovat samat. Suoran esityksessä voidaan nimittäin korvata suuntavektori millä tahansa samansuuntaisella vektorilla ja paikkavektoriksi voidaan valita mikä hyvänsä suoralla oleva piste. Tällöin koordinaattimuoto ei tarjoa hyvää tapaa suorien yhtäsuuruuden selvittämiseksi. Sen sijaan käyttökelpoisin tapa lienee tarkastaa, ovatko suorat yhdensuuntaiset (suuntavektorit yhdensuuntaiset) ja onko suorilla yhteinen piste.

Esimerkki 99 Selvitetään, ovatko suorat

$$L_1 = \{(-3, 8, 0) + t(2, -3, 1) \mid t \in \mathbb{R}\}$$

ja

$$L_2 = \{(1, 2, 2) + t(-4, 6, -2) \mid t \in \mathbb{R}\}$$

samat. Ensiksi voidaan todeta, että $(-4, 6, -2) = -2(2, -3, 1)$, joten suorien suuntavektorit ovat yhdensuuntaiset. Valitaan sitten suoralla L_2 piste $(1, 2, 2)$ (saadaan arvolle $t = 0$) ja selvitetään, löytyykö tätä pistettä suoralla L_1 , toisin sanoen, onko sellaista parametrin t arvoa, että $(1, 2, 2) = (-3, 8, 0) + t(2, -3, 1)$. Ensimmäistä koordinaattia tarkastellessa todetaan, että $1 = -3 + 2t$, joten $t = 2$. Todetaan, että tällä t :n valinnalla myös toinen ja kolmas koordinaatti saadaan halutuiksi, joten $(1, 2, 2)$ on myös suoralla L_1 . Täten suorat ovat samat.

Esimerkki 100 Määritetään a ja b siten, että suorat $L_1 = \{(0, -1, 1) + t(1, -7, -3) \mid t \in \mathbb{R}\}$ ja $L_2 = \{(1, -1, 3) + t(-2, a, b) \mid t \in \mathbb{R}\}$ ovat yhdensuuntaiset.

Suorat ovat yhdensuuntaiset silloin, kun niiden suuntavektorit $(1, -7, -3)$ ja $(-2, a, b)$ ovat yhdensuuntaiset. Tämä puolestaan tapahtuu silloin, kun $(1, -7, -3) = k(-2, a, b)$ jollekin luvulle k . Tällöin

on oltava $k = -\frac{1}{2}$, $a = 14$ ja $b = 6$, jolloin siis suoran L_2 suuntavektorina toimii $\mathbf{x} = (-2, 14, 6)$ tai yhtä hyvin $\frac{1}{-2}\mathbf{x} = (1, -7, -3)$.

Selvitetään vielä, ovatko yhdensuuntaiset suorat $L_1 = \{(0, -1, 1) + t(1, -7, -3) \mid t \in \mathbb{R}\}$ ja $L_2 = \{(1, -1, 3) + t(1, -7, -3) \mid t \in \mathbb{R}\}$ erillisiä vai yhtyvätkö nämä suorat. Jos yhdensuuntaiset suorat ovat erilliset, ei niillä voi olla yhtään yhteistä pistettä. Selvitetään siis, onko yhtälöllä

$$(0, -1, 1) + t_1(1, -7, -3) = (1, -1, 3) + t_2(1, -7, -3) \quad (4.15)$$

ratkaisuja $t_1, t_2 \in \mathbb{R}$. Yhtälö (4.15) saadaan muotoon

$$(-1, 0, -2) = (t_2 - t_1)(1, -7, 3),$$

mistä nähdään, että ratkaisuja ei ole. Suorat ovat siis yhdensuuntaiset, mutta erilliset.

Esimerkki 101 Etsitään suorien

$$L_1 : \frac{x-4}{-5} = \frac{y}{3} = \frac{5z-4}{5}$$

ja

$$L_2 : \frac{x}{5} = \frac{y-1}{8} = \frac{z-3}{-12}$$

leikkauspiste (mikäli olemassa).

Muunnetaan aluksi suorat parametrimuotoon: suoran L_1 osalta kirjoitetaan $\frac{x-4}{-5} = t$, $\frac{y}{3} = t$ ja $\frac{5z-4}{5} = t$, mistä $(x, y, z) = (-5t + 4, 3t, t + \frac{4}{5}) = t(-5, 3, 1) + (4, 0, \frac{4}{5})$.

Vastaavasti suoralle L_2 saadaan esitys $(x, y, z) = (5t, 8t + 1, -12t + 3) = t(5, 8, -12) + (0, 1, 3)$, ja leikkauspiste edellyttää, että

$$(-5t_1 + 4, 3t_1, t_1 + \frac{4}{5}) = (5t_2, 8t_2 + 1, -12t_2 + 3)$$

joillakin arvoilla $t_1, t_2 \in \mathbb{R}$. Algebrallisen yhtälöryhmän ratkaisuna saadaan leikkauspiste, jossa $t_1 = \frac{37}{55}$ ja $t_2 = \frac{7}{55}$.

4.4 Etäisyyksien laskemisesta

Kahden tason samoin kuin kahden suorankin etäisyydellä toisistaan tarkoitetaan aina lyhintä mahdollista välimatkaa tasojen tai suorien välillä. Jos tasot tai suorat leikkaavat toisensa, on niiden etäisyys nolla. Samoin pisteen etäisyys tasosta tai suorasta merkitsee lyhintä mahdollista etäisyyttä. Näiden etäisyyksien laskemiseksi voidaan käyttää seuraavia periaatteita:

- Kahden tason välinen etäisyys: Jos tasojen T_1 ja T_2 normaalit $\mathbf{n}_1 = (n_1, n_2, n_3)$ ja $\mathbf{n}_2$ ovat erisuuntaiset, tasot leikkaavat ja niiden etäisyys on nolla. Jos taas normaalit ovat samansuuntaiset (lineaarisesti riippuvat), valitaan tasojen kohtisuoraan oleva suora $S = \{t\mathbf{n}_1 \mid t \in \mathbb{R}\}$, ja ratkaistaan arvo t_1 , jolla $t_1\mathbf{n}_1 \in T_1$ sekä t_2 , jolla $t_2\mathbf{n}_1 \in T_2$. Arvot t_1 löytyy esimerkiksi ratkaisemalla yhtälö $atn_1 + btn_2 + ctn_3 = d$, missä $ax + by + cz = d$ on tason T_1 standardimuoto. Tasojen välinen etäisyys saadaan lopuksi laskemalla pisteiden $t_1\mathbf{n}_1$ ja $t_2\mathbf{n}_1$ välinen etäisyys.
- Tason etäisyys suorasta: Jos tason T normaali $\mathbf{n}$ ei ole kohtisuorassa suoran S suuntavektoria kohtaan, Suora leikkaa tason ja etäisyys on nolla. Muussa tapauksessa suora on tason suuntainen. Tällöin valitaan $\mathbf{r}$ siten, että suora $\mathbf{r} + t\mathbf{n}$ leikkaa suoran S ; vektoriksi $\mathbf{r}$ kelpaa esimerkiksi suoran S paikkavektori. Tämän jälkeen selvitetään, millä arvolla t_1 piste $\mathbf{p}_1 = \mathbf{r} + t_1\mathbf{n}$ kuuluu tasolle T ja millä arvolla t_2 piste $\mathbf{p}_2 = \mathbf{r} + t_2\mathbf{n}$ kuuluu suoralle S . t_1 löytyy käyttämällä tason koordinaattimuotoa ja jos $\mathbf{r}$ valittiin suoran S suuntavektoriksi, on $t_2 = 0$. Lopuksi lasketaan pisteiden $\mathbf{p}_1$ ja $\mathbf{p}_2$ välinen etäisyys.
- Tason etäisyys pisteestä: Olkoon $\mathbf{n}$ tason normaali ja $\mathbf{p}$ kysytty piste. Selvitetään, millä arvolla t_1 piste $\mathbf{p}_1 = \mathbf{p} + t_1\mathbf{n}$ kuuluu tasolle (käyttämällä tason koordinaattimuotoa) ja lasketaan pisteiden $\mathbf{p}_1$ ja $\mathbf{p}$ välinen etäisyys.
- Kahden suoran $S_1 : \mathbf{r}_1 + ts_1$ ja $S_2 : \mathbf{r}_2 + ts_2$ etäisyys: Valitaan suuntavektori $\mathbf{s} \neq \mathbf{0}$, joka on kohtisuorassa kumpaakin suoraa vastaan. Jos $\mathbf{s}_1$ ja $\mathbf{s}_2$ ovat lineaarisesti riippumattomat (eivät toistensa monikerrat), voidaan valita esim. $\mathbf{s} = \mathbf{s}_1 \times \mathbf{s}_2$. Tämän jälkeen selvitetään milloin suorilla

S_1 ja S_2 olevien pisteiden $\mathbf{p}_1 = \mathbf{r}_1 + t\mathbf{s}_1$ ja $\mathbf{p}_2 = \mathbf{r}_2 + s\mathbf{s}_2$ erotus $\mathbf{p}_2 - \mathbf{p}_1$ on vektorin $\mathbf{s}$ suuntainen, siis milloin $\mathbf{p}_2 - \mathbf{p}_1 = c\mathbf{s}$. Tämä on kolmen tuntemattoman ja kolmen yhtälön ryhmä, joka ratkeaa Gaussin-Jordanin menetelmällä. Kysytty etäisyys saadaan pisteiden $\mathbf{p}_1 \in S_1$ ja $\mathbf{p}_2 \in S_2$ välisenä etäisyytenä. Jos alun perin $\mathbf{s}_1$ ja $\mathbf{s}_2$ olivat samansuuntaiset, voidaan valita suoralta S_1 mikä hyvänsä piste ja laskea tämän etäisyys suorasta S_2 (kts. seuraava kohta).

- Suoran etäisyys pisteestä: Minimoidaan differentiaalilaskennan keinoin lauseke $\|\mathbf{r} + t\mathbf{s} - \mathbf{p}\|^2$, missä $S : \mathbf{r} + t\mathbf{s}$ ja $\mathbf{p}$ ovat tarkasteltava suora ja piste. Saatava t :n arvo vastaa suoran pistettä, joka on lähinnä $\mathbf{p}$:tä.

Luku 5

Differentiaaliyhtälöt (DY:t)

Määritelmänsä perusteella funktion derivaatta edustaa funktion kuvaaman ilmiön hetkellistä muutosnopeutta. Tämän vuoksi on ymmärrettävää, että luonnonilmiöitä kuvattaessa on usein käytettävä yhtälöitä, joissa esiintyy sekä funktio että sen derivaatta tai useampikertaisia derivaattoja, siis *differentiaaliyhtälöitä*. Differentiaaliyhtälöitä esiintyy eniten fysikaalisissa, mutta biologisissa, kemiallisissa ja taloustieteellisissä malleissa. Muodollisesti differentiaaliyhtälön käsite määritellään seuraavasti:

Määritelmä 49 Differentiaaliyhtälöllä (DY) tarkoitetaan yhtälöä $F(t, y, y', y'', \dots, y^{(n)}) = 0$, missä F on lauseke, joka sisältää muuttujan t , tuntemattoman funktion $y = y(t)$ sekä tämän yhden- tai useammankertaisia derivaattoja $y', y'' \dots, y^{(n)}$. Lukua n kutsutaan differentiaaliyhtälön *kertaluvuksi*.

Erityisesti differentiaaliyhtälöiden yhteydessä jätetään usein muuttuja merkitsemättä ja siis merkintöjen $y(t), y'(t), \dots, y^{(n)}(t)$ sijaan käytetään merkintöjä $y, y', \dots, y^{(n)}$. Muuttujan merkintä t :n sijaan on perua yleisimmistä fysikaalisista sovelluksista, joissa muuttujana toimii aika.

Määritelmä 50 Differentiaaliyhtälön ratkaisun kuvaajaa kutsutaan *integraalikäyräksi*. Yleensä integraalikäyriä on ääretön määrä, mutta yksikäsitteiseen ratkaisuun voidaan silti yleensä päätyä *reunaehdoilla* $y(0) = c_0, y'(0) = c_1, \dots, y^{(n)}(0) = c_n$.

5.1 Yksinkertaiset DY:t

Yksinkertaiseksi differentiaaliyhtälöksi sanotaan sellaista, jonka ratkaiseminen on pelkkä integrointitehtävä.

Esimerkki 102 Differentiaaliyhtälön $y' = 2t - 1$ ratkaiseminen palautuu pelkäksi antiderivaatan etsimiseksi: Yleinen ratkaisu on muotoa $y = t^2 - t + C$, missä C on jokin vakio.

Yksinkertaisille differentiaaliyhtälöille on tyypillistä se, että etsittävästä funktiosta esiintyy vain derivaatta, eikä itse funktiota. On huomattava, että toisinaan voidaan hankalamman näköinen differentiaaliyhtälö kirjoittaa yksinkertaiseen muotoon (vrt. radioaktiivisen hajoamisen differentiaaliyhtälö kurssilla Insinöörimatematiikka B): Differentiaaliyhtälö

$$N'(t) = -\lambda N(t)$$

voidaan kirjoittaa muotoon

$$\frac{d}{dt} \ln N = -\lambda,$$

joka on yksinkertainen funktion $\ln N$ suhteen.

5.2 Lineaariset DY:t

Määritelmä 51 Muotoa

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_2 y'' + a_1 y' + a_0 y = b(t) \quad (5.1)$$

olevaa differentiaaliyhtälöä kutsutaan *lineaariseksi*. Jos $a_n, a_{n-1}, \dots, a_1$ ja a_0 ovat vakioita, sanotaan yhtälöä *vakiokertomiseksi* lineaariseksi differentiaaliyhtälöksi. Jos $b(t)$ on nollafunktio, sanotaan lineaarista differentiaaliyhtälöä

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_2 y'' + a_1 y' + a_0 y = 0 \quad (5.2)$$

homogeeniseksi.

Lause 25 Olkoon C^n n kertaa derivoituvien funktioiden joukko. Määritellään funktioiden yhteenlasku ja skalaarimonikerta kuten aiemmin. Tällöin homogeenisen differentiaaliyhtälön (5.2) ratkaisut muodostavat avaruuden C^n aliavaruuden.

Todistus Olkoon $V \subseteq C^n$ differentiaaliyhtälön (5.2) ratkaisujen joukko ja $y_1, y_2 \in V$. Derivoinnin lineaarisuuden perusteella

$$\begin{aligned} & a_n(\alpha y_1 + \beta y_2)^{(n)} + a_{n-1}(\alpha y_1 + \beta y_2)^{(n-1)} + \dots + a_1(\alpha y_1 + \beta y_2)' + a_0(\alpha y_1 + \beta y_2) \\ &= \alpha(a_n y_1^{(n)} + a_{n-1} y_1^{(n-1)} + \dots + a_2 y_1'' + a_1 y_1' + a_0 y_1) \\ &+ \beta(a_n y_2^{(n)} + a_{n-1} y_2^{(n-1)} + \dots + a_2 y_2'' + a_1 y_2' + a_0 y_2) = \alpha \cdot 0 + \beta \cdot 0 = 0, \end{aligned}$$

siis myös $\alpha y_1 + \beta y_2 \in V$. Näin ollen V on vektoriavaruuden C^n aliavaruus. □

Seuraava lause esittää luonnehdinnan lineaarisen differentiaaliyhtälön ratkaisujoukosta.

Lause 26 Oletetaan että funktiot a_i ja b ovat jatkuvia. Silloin n :nen kertaluvun homogeenisella differentiaaliyhtälöllä (5.2) on aina n lineaarisesti riippumatonta ratkaisua $y_1, y_2, \dots, y_n$, joiden lineaarikombinaationa saadaan kaikki ratkaisut:

$$y = c_1 y_1 + c_2 y_2 + \dots + c_n y_n.$$

Epähomogeenisen differentiaaliyhtälön (5.1) kaikki ratkaisut ovat muotoa

$$y = c_1 y_1 + c_2 y_2 + \dots + c_n y_n + y_0,$$

missä $y_1, \dots, y_n$ ovat homogeenisen differentiaaliyhtälön (5.2) riippumattomat ratkaisut ja y_0 jokin yksittäinen epähomogeenisen differentiaaliyhtälön (5.1) ratkaisu.

Todistus Sivuutetaan. □

Huomautus 35 Edellinen lause on analoginen sen tuloksen kanssa, että lineaarisen yhtälöryhmän $A\mathbf{x} = \mathbf{b}$ kaikki ratkaisut voidaan kirjoittaa muodossa $c_1 \mathbf{x}_1 + \dots + c_k \mathbf{x}_k + \mathbf{x}_0$, missä c_i ovat mitä hyvänsä

lukuja, $\mathbf{x}_i$ ovat lineaarisesti riippumattomia vastaavan homogeenisen yhtälön $A\mathbf{x} = \mathbf{0}$ ratkaisuja, ja $\mathbf{x}_0$ on alkuperäisen yhtälön yksittäisratkaisu.

Tämän kurssin alkuosassa käytettiin pääasiassa Gaussin-Jordanin menetelmää, jolla voitiin todeta avaruuden $\mathbb{R}^n$ vektoreiden lineaarinen riippumattomuus tai riippuvuus. Menetelmä ei välttämättä sovellu, mikäli tarkasteltava vektoriavaruus on ääretönulotteinen (kuten C^n) tai mikäli ei ole selvää, miten koordinaattivektorit sen alkioille määritetään.

Todetaan siksi lineaarinen riippumattomuus erityistä tyyppiä oleville funktiojoukoille. Yleisesti ottaen vektoreille ei kuitenkaan voida määrittellä äärettömiä summia (sarjoja), joten äärettömiä joukkoja varten tulee asettaa seuraava määritelmä.

Määritelmä 52 Ääretön vektorijoukko on lineaarisesti riippumaton, jos sen jokainen äärellinen osajoukko on sellainen.

Lause 27 Polynomijoukko $\{1, x, x^2, x^3, \dots\}$ on lineaarisesti riippumaton.

Todistus Valitaan jokin äärellinen osajoukko $\{x^{n_1}, \dots, x^{n_k}\}$, ja muodostetaan lineaarikombinaatio $P(x) = c_{n_1}x^{n_1} + \dots + c_{n_k}x^{n_k}$. Mikäli tämä lineaarikombinaatio on nollafunktio, on Insinöörimatematiikka B:n Taylorin sarjoja koskevan tuloksen mukaan $c_{n_1} = \dots = c_{n_k} = 0$. $\square$

Lause 28 Funktiojoukko $\{e^{\lambda x} \mid \lambda \in \mathbb{C}\}$ on lineaarisesti riippumaton.

Todistus Valitaan jokin äärellinen osajoukko

$$\{e^{\lambda_1 x}, \dots, e^{\lambda_k x}\} \quad (5.3)$$

ja osoitetaan induktiolla k :n suhteen, että muotoa (5.3) oleva joukko on lineaarisesti riippumaton.

Jos $k = 1$, on tarkasteltava joukko muotoa $\{e^{\lambda_1 x}\}$ ja lineaarisesti riippumaton, sillä joukon ainoa alkio ei ole nollafunktio. Oletetaan sitten, että jollekin arvolle $k-1$ muotoa (5.3) oleva on lineaarisesti riippumaton, ja osoitetaan, että näin on myös arvolle k .

Oletetaan, että

$$c_1 e^{\lambda_1 x} + \dots + c_{k-1} e^{\lambda_{k-1} x} + c_k e^{\lambda_k x} = 0.$$

Jakamalla $e^{\lambda_k x}$:llä saadaan

$$c_1 e^{(\lambda_1 - \lambda_k)x} + \dots + c_{k-1} e^{(\lambda_{k-1} - \lambda_k)x} + c_k = 0,$$

josta derivoimalla nähdään, että

$$c_1 (\lambda_1 - \lambda_k) e^{(\lambda_1 - \lambda_k)x} + \dots + c_{k-1} (\lambda_{k-1} - \lambda_k) e^{(\lambda_{k-1} - \lambda_k)x} = 0.$$

Induktio-oletuksen perusteella $c_1 (\lambda_1 - \lambda_k) = \dots = c_{k-1} (\lambda_{k-1} - \lambda_k) = 0$, josta $c_1 = \dots = c_{k-1} = 0$ ja edelleen $c_k = 0$. $\square$

5.2.1 Vakiokertoimiset lineaariset DY:t

Vakiokertoimisella differentiaaliyhtälöllä tarkoitetaan muotoa

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y + a_0 y = b \quad (5.4)$$

olevaa differentiaaliyhtälöä, jossa kerroinfunctiot a_i ovat vakioita. Tätä muotoa olevaan differentiaaliyhtälöön voidaan tietenkin soveltaa lausetta 26, jonka mukaan yhtälön (5.4) kaikki ratkaisut saadaan määrittämällä yksi ainoa (yksittäisratkaisu) y_0 yhtälölle 5.4 ja yhdistämällä se homogeenisen yhtälön

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y + a_0 y = 0 \quad (5.5)$$

ratkaisujoukkoon, jonka lauseen 26 mukaan generoi n lineaarisesti riippumatonta ratkaisua $y_1, \dots, y_n$.

Vakiokertoimisen, homogeenisen lineaarisen differentiaaliyhtälön ratkaiseminen on melko helppoa, sillä suoraan laskemalla voidaan todeta, että yhtälöllä

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_0 y = 0 \quad (5.6)$$

on aina muotoa $y = e^{\lambda x}$ oleva ratkaisu. Tätä varten lasketaan derivaatat $y' = \lambda e^{\lambda x}$, $y'' = \lambda^2 e^{\lambda x}$, ..., $y^{(n)} = \lambda^n e^{\lambda x}$ ja sijoitetaan nämä yhtälöön (5.6):

$$a_n \lambda^n e^{\lambda x} + a_{n-1} \lambda^{n-1} e^{\lambda x} + \dots + a_0 e^{\lambda x} = 0,$$

mistä $e^{\lambda x}$:llä jakaminen antaa

$$a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0 = 0. \quad (5.7)$$

Yhtälöllä (5.7), jota kutsutaan differentiaaliyhtälön (5.6) *karakteristiseksi yhtälöksi*, on algebran peruslauseen (kts. Insinöörimatematiikka A) nojalla korkeintaan $k \leq n$ erisuurta ratkaisua $\lambda_1, \dots, \lambda_k$, joista jokaisesta saadaan alkuperäisen yhtälön (5.6) ratkaisu $e^{\lambda_i x}$. Jos juuria λ_i on tarkalleen n , ovat tässä lauseen 28 mukaan kaikki alkuperäisen yhtälön *riippumattomat ratkaisut*, kun taas siinä tapauksessa että $k < n$, on differentiaaliyhtälöllä (5.6) myös muita riippumattomia ratkaisuja.

Voidaan osoittaa, että jos λ_i on karakteristisen yhtälön j -kertainen juuri, niin differentiaaliyhtälöllä (5.6) on ratkaisut

$$e^{\lambda_i x}, x e^{\lambda_i x}, x^2 e^{\lambda_i x}, \dots, x^{j-1} e^{\lambda_i x}.$$

Riippumattomista ratkaisuista $y_1, \dots, y_n$ saadaan kaikki ratkaisut näiden lineaarikombinaatioina

$$C_1 y_1 + \dots + C_n y_n.$$

Tarkastellaan erityisesti sellaista ratkaisua $e^{\lambda x}$, jossa $\lambda = \alpha + i\beta$ on kompleksiluku. Mikäli karakteristinen yhtälö on reaalikeroiminen, on myös $\bar{\lambda}$ yhtälön ratkaisu, ja tällöin kompleksiset ratkaisut $e^{\lambda x}$ ja $e^{\bar{\lambda} x}$ voidaan korvata reaalisilla ratkaisuilla $e^{\alpha x} \cos \beta x$ ja $e^{\alpha x} \sin \beta x$. Tämä johtuu siitä, että tällöin voidaan kirjoittaa $e^{(\alpha \pm i\beta)x} = e^{\alpha x} e^{\pm i\beta x}$, mikä edelleen Eulerin kaavan perusteella on muotoa

$$e^{\alpha x} (\cos \beta x \pm i \sin \beta x).$$

Esimerkki 103 Toisen kertaluvun homogeenisen differentiaaliyhtälön

$$y'' + ay' + by = 0, \quad (5.8)$$

ratkaisut voidaan löytää seuraavasti: Edellämainitun perusteella yhtälöllä on muotoa $y = e^{\lambda x}$ oleva ratkaisu. Tällöin $y' = \lambda e^{\lambda x}$ ja $y'' = \lambda^2 e^{\lambda x}$, ja nämä sijoittamalla saadaan karakteristiseksi yhtälöksi

$$\lambda^2 + a\lambda + b = 0,$$

jonka ratkaisut ovat $\lambda = -\frac{a}{2} \pm \sqrt{(\frac{a}{2})^2 - b}$. Yhtälön käsittely voidaan jakaa kolmeen osaan:

1) $D = (\frac{a}{2})^2 - b > 0$. Tällöin yhtälöllä (5.8) on kaksi ratkaisua $y_1 = e^{\lambda_1 x}$ ja $y_2 = e^{\lambda_2 x}$, missä $\lambda_1 = -\frac{a}{2} + \sqrt{D}$ ja $\lambda_2 = -\frac{a}{2} - \sqrt{D}$.

2) $D = (\frac{a}{2})^2 - b < 0$. Tällöin yhtälöllä on kaksi kompleksista ratkaisua $e^{(\alpha \pm i\beta)x}$, missä $\alpha = -\frac{a}{2}$ ja $\beta = \sqrt{D}$, joita vastaavat reaaliratkaisut $y_1 = e^{\alpha x} \cos \beta x$ ja $y_2 = e^{\alpha x} \sin \beta x$.

3) $D = (\frac{a}{2})^2 - b = 0$ Tällöin yhtälöllä on ratkaisu $y_1 = e^{-\frac{a}{2}x}$. Voidaan todeta, että tällöin myös $y_2 = x e^{-\frac{a}{2}x}$ on myös ratkaisu, joka on edellisestä riippumaton.

Jokaisessa tapauksessa yhtälön kaikki ratkaisut saadaan riippumattomien ratkaisujen lineaarikombinaationa

$$C_1 y_1 + C_2 y_2.$$

Yksikäsitteisen ratkaisun määräävät reunaehdot.

Esimerkki 104 Ratkaistaan differentiaaliyhtälö

$$y'' + 4y = 0. \quad (5.9)$$

Yleinen ratkaisu on muotoa $y = e^{\lambda t}$, josta $y' = \lambda e^{\lambda t}$ ja $y'' = \lambda^2 e^{\lambda t}$, ja sijoitus differentiaaliyhtälöön antaa

$$\lambda^2 e^{\lambda t} + 4e^{\lambda t} = 0,$$

mistä $\lambda^2 = -4$ ja siis $\lambda = \pm 2i$, josta saadaan kompleksiset ratkaisut $e^{\pm i2x}$, joita vastaavat reaaliset ratkaisut ovat $y_1(x) = \cos 2x$ ja $y_2(x) = \sin 2x$.

Yhtälön (5.9) yleinen (reaalinen) ratkaisu on siis

$$y(x) = C_1 \cos 2x + C_2 \sin 2x.$$

Reunaehtojen $y(0) = y_0$, $y'(0) = y_1$ kiinnittämä ratkaisu saadaan laskemalla

$$y'(x) = -2C_1 \sin 2x + 2C_2 \cos 2x,$$

sekä sijoittamalla $x = 0$ kumpaankin yhtälöön: $y_0 = y(0) = C_1$ ja $y_1 = y'(0) = 2C_2$.

Esimerkki 105 Ratkaistaan kurssista Insinöörimatematiikka B tuttu radioaktiivisen hajoamisen differentiaaliyhtälö

$$y' = -\lambda y, \quad (5.10)$$

missä λ on vakio ja alkuehtona toimii $y(0) = y_0$. Käytetään yritettä $y = e^{\alpha t}$, jolloin $y' = \alpha e^{\alpha t}$ ja sijoittamalla yhtälöön saadaan

$$\alpha e^{\alpha t} = -\lambda e^{\alpha t},$$

josta jakamalla saadaan yhtälö $\alpha = -\lambda$. Täten yhtälön yksi ratkaisu on $y_1 = e^{-\lambda t}$, ja kaikki ratkaisut saadaan tämän skalaarimonikertoina

$$y = C e^{-\lambda t}.$$

Vakio C voidaan kiinnittää alkuehdon perusteella: $y_0 = C e^{-\lambda \cdot 0} = C$.

Jos differentiaaliyhtälö ei ole homogeeninen, voidaan huomata että Insinöörimatematiikka C:ssä esitetyt Laplace-muunnokset ovat erittäin käyttökelpoisia nimenomaan vakiokertoimisten lineaaristen differentiaaliyhtälöiden ratkaisemiseksi. Niiden avulla voidaan muuntaa muotoa (5.1) oleva differentiaaliyhtälö *algebralliseksi* yhtälöksi, josta voidaan ratkaista tuntematon funktio $Y(s) = \mathcal{L}[y(t)](s)$. Etsitty funktio $y(t)$ saadaan tämän jälkeen käänteisellä Laplace-muunnoksella. Yleensä oletetaan, että y on oikealta jatkuva nollassa.

Esimerkki 106 Ratkaistaan toisen kertaluvun differentiaaliyhtälö

$$y'' + 4y = t \quad (5.11)$$

alkuehdoilla $y(0) = 1$ ja $y'(0) = 2$. Nyt $\mathcal{L}[y''] = s^2 Y(s) - s - 2$ ja (5.11) saadaan siis Laplace-muunnoksilla muotoon

$$s^2 Y(s) - s - 2 + 4Y(s) = \frac{1}{s^2},$$

mistä voidaan ratkaista

$$Y(s) = \frac{s}{s^2 + 4} + \frac{2}{s^2 + 4} + \frac{1}{s^2(s^2 + 4)} = \frac{s}{s^2 + 4} + \frac{2}{s^2 + 4} + \frac{1}{4} \left(\frac{1}{s^2} - \frac{1}{s^2 + 4} \right)$$

Originaalifunktio saadaan taulukosta: $y(t) = \cos 2t + \sin 2t + \frac{1}{4}(t - \frac{1}{2} \sin 2t) = \cos 2t + \frac{7}{8} \sin 2t + \frac{1}{4}t$.

Edellinen esimerkki sisälsi perustelemattomia oletuksia, kuten sen, että alkuperäisen differentiaaliyhtälön ratkaisu on originaalifunktio ja nollassa jatkuva oikealta. Myöskään ei tarkasteltu sitä, missä alueessa Laplace-muunnos oli määritelty. Sovellusten kannalta näillä heikkouksilla ei kuitenkaan ole yleensä merkitystä, sillä kun differentiaaliyhtälön ratkaisu on saatu, voidaan sen oikeellisuus tarkistaa ja määrittelyjoukko löytää yleensä melko helposti. Ratkaistaessa differentiaaliyhtälöitä Laplace-muunnosten avulla tehdään ratkaisusta yleensä äskeisen esimerkin oletukset.

Esimerkki 107 Ratkaistaan differentiaaliyhtälö

$$y' - 3y = 8e^{2t}$$

alkuehdolla $y(0) = 2$. Laplace-muunnokset laskemalla saadaan tällöin

$$sY(s) - 2 - 3Y(s) = \frac{8}{s-2},$$

mistä voidaan ratkaista $Y(s)$:

$$Y(s) = \frac{2s+4}{(s-3)(s-2)}.$$

Näin saadulle lausekkeelle löytyy osamurtohajotelma:

$$\frac{2s+4}{(s-3)(s-2)} = \frac{10}{s-3} - \frac{8}{s-2},$$

minkä avulla originaalifunktio voidaan määrätä:

$$y(t) = 10e^{3t} - 8e^{2t}.$$

Esimerkki 108 Ratkaistaan

$$y'' + 3y' - 4y = e^{3t}$$

alkuehdoilla $y(0) = 2$, $y'(0) = 1$. Laplace-muunnoksilla saadaan

$$s^2Y(s) - 2s - 1 + 3(sY(s) - 2) - 4Y(s) = \frac{1}{s-3},$$

mikä voidaan kirjoittaa muotoon

$$(s^2 + 3s - 4)Y(s) - 2s - 7 = \frac{1}{s-3}$$

ja edelleen

$$Y(s) = \frac{2s^2 + s - 20}{(s^2 + 3s - 4)(s-3)} = \frac{2s^2 + 2 - 20}{(s-1)(s+4)(s-3)}$$

Tälle löytyy osamurtokehitemä

$$Y(s) = \frac{17}{10} \frac{1}{s-1} + \frac{8}{35} \frac{1}{s+4} + \frac{1}{14} \frac{1}{s-3},$$

josta voidaan määrätä originaalifunktio:

$$y(t) = \frac{17}{10}e^t + \frac{8}{25}e^{-4t} + \frac{1}{14}e^{3t}.$$

Esimerkki 109 Ratkaistaan differentiaaliyhtälö

$$y'' + y = \cos t$$

alkuehdoilla $y(0) = y'(0) = 0$. Laplace-muunnokset laskemalla saadaan

$$s^2Y(s) + Y(s) = \frac{s}{s^2+1},$$

mistä

$$Y(s) = \frac{1}{s^2+1} \frac{s}{s^2+1} = \mathcal{L}[\sin](s)\mathcal{L}[\cos](s) = \mathcal{L}[\sin * \cos](s),$$

jolloin siis $y(t) = (\sin * \cos)(t) = \frac{1}{2}t \sin t$.

Tämän esimerkin originaalifunktio $y(t)$ voidaan selvittää myös suoraan taulukosta.

Esimerkki 110 Ratkaistaan differentiaaliyhtälöpari

$$\begin{cases} y' + 4y + 4z = 0 \\ z' + 2y + 6z = 0 \end{cases}$$

alkuehdoilla $y(0) = 3$, $z(0) = 15$. Laplace-muunnettu pari on muotoa

$$\begin{cases} sY(s) - 3 + 4Y(s) + 4Z(s) = 0 \\ sZ(s) - 15 + 2Y(s) + 6Z(s) = 0, \end{cases}$$

mistä voidaan ratkaista

$$\begin{cases} Y(s) = \frac{3s-42}{s^2+10s+16} = -\frac{8}{s+2} + \frac{11}{s+8} \\ Z(s) = \frac{15s+54}{s^2+10s+16} = \frac{4}{s+2} + \frac{11}{s+8} \end{cases}$$

Vastaavat originaalifunktiot ovat

$$\begin{cases} y(t) = -8e^{-2t} + 11e^{-8t} \\ z(t) = 4e^{-2t} + 11e^{-8t}. \end{cases}$$

Laplace-muunnokset eivät ole laskennallisesti yksinkertaisin tapa ratkaista tällaisia DY-pareja. Laskennallisesti parempaan menetelmään tutustutaan myöhemmin.

Huomautus 36 Seuraava tapa ratkaista vakiokertoiminen lineaarinen DY saattaa olla pelkkiä Laplace-muunnoksia tehokkaampi:

- Etsitään kaikki homogeenisen yhtälön ratkaisut yrittäällä kuten edellä on kuvattu.
- Etsitään yksittäisratkaisu käyttämällä Laplace-muunnoksia ja reunaehtoja, jotka on valittu sopivasti helpottamaan laskutoimituksia. Yleensä $y(0) = y'(0) = y''(0) = \dots = 0$ on käyttökelpoinen valinta.

5.2.2 Ensimmäisen kertaluvun lineaarinen DY

Aiemmin nähtiin esimerkkejä differentiaaliyhtälöistä, jotka ratkeavat Laplace-muunnoksia käyttämällä ja sellaisista, jotka ovat niin yksinkertaisia, etteivät Laplace-muunnoksia tarvitse. Useiden differentiaaliyhtälöiden kohdalla tarvitaan kuitenkin muita menetelmiä.

Esimerkki 111 Laplace-muunnoksella differentiaaliyhtälöstä $y' = -y^2$ saadaan

$$sY(s) - y(0) = -\mathcal{L}[y^2](s),$$

mutta muunnokselle $\mathcal{L}[y^2](s)$ ei ole helppoa esitystapaa muunnoksen $Y(s) = \mathcal{L}[y](s)$ avulla. Kuitenkin alkuperäisellä differentiaaliyhtälöllä on välillä $(0, \infty)$ (ja välillä $(-\infty, 0)$) ratkaisu $y(t) = \frac{1}{t}$ kuten helposti todetaan.

Yleisiä ratkaisumenetelmiä differentiaaliyhtälöille on kuitenkin varsin vähän, mutta 1. kertaluvun lineaariselle DY:lle tunnetaan täydellinen ratkaisumenetelmä.

Ensimmäisen kertaluvun lineaarinen differentiaaliyhtälö voidaan kirjoittaa muotoon

$$y' + a(x)y = b(x), \quad (5.12)$$

ja tämän ratkaisemiseksi tarkastellaan aluksi *homogenisoitua* yhtälöä

$$y' + a(x)y = 0, \quad (5.13)$$

joka voidaan kirjoittaa muotoon

$$\frac{y'}{y} = -a(x). \quad (5.14)$$

Koska $\frac{d}{dx} \ln y = \frac{y'}{y}$, saadaan (5.14) muotoon

$$\frac{d}{dx} \ln y = \frac{d}{dx} \left(- \int a(x) dx \right),$$

josta puolestaan

$$\ln y = - \int a(x) dx + C.$$

Näin ollen homogeenisen yhtälön ratkaisu on

$$y_H = e^C e^{-\int a(x) dx},$$

missä C voidaan valita vapaasti. Jos merkitään $e^C = C_1$, saadaan muoto

$$y_H = C_1 e^{-\int a(x) dx},$$

missä C_1 on vapaasti valittava vakio.

Esimerkki 112 Differentiaaliyhtälö $xy' + y = 0$ voidaan kirjoittaa homogeenisen differentiaaliyhtälön muotoon $y' + \frac{1}{x}y = 0$, jonka yleinen ratkaisu on

$$y(x) = C e^{-\int \frac{1}{x} dx} = C e^{-\ln x} = \frac{C}{x}.$$

Alkuperäisen (epähomogeenisen) differentiaaliyhtälön (5.12) ratkaisu saadaan menettelyllä, jota kutsutaan *vakion varioinniksi*. Vakion varioinnissa käytetään lähtökohtana homogenisoitua differentiaaliyhtälön (5.13) ratkaisua y_H , merkitään $y = C(x)y_H(x)$ ja sijoitetaan näin saatu uusi funktio alkuperäiseen yhtälöön, ja ratkaistaan funktio $C(x)$.

Aluksi todetaan, että $y' = C'y_H + Cy'_H$, joten sijoitus antaa

$$C'y_H + Cy'_H + aCy_H = b.$$

Koska y_H oletettiin homogeenisen differentiaaliyhtälön ratkaisuksi, on $Cy'_H + aCy_H = C(y'_H + ay_H) = 0$ ja siis päädytään yhtälöön $C'y_H = b$, josta C voidaan ratkaista integroimalla:

$$C(x) = \int \frac{b(x)}{y_H(x)} dx + B$$

Tällöin siis alkuperäisen differentiaaliyhtälön (5.12) ratkaisu on

$$y(x) = By_H(x) + y_H(x) \int \frac{b(x)}{y_H(x)} dx$$

missä

$$y_H(x) = e^{-\int a(x) dx}.$$

Esimerkki 113 Ratkaistaan $y' + 2xy = x$. Tarkastellaan aluksi homogenisoitua yhtälöä $y' + 2xy = 0$, joka voidaan kirjoittaa muotoon

$$\frac{d}{dx} \ln y = \frac{y'}{y} = -2x, \quad (5.15)$$

josta $\ln y = -x^2$; tällöin siis $y_H = e^{-x^2}$ on eräs homogenisoitua yhtälön ratkaisu. Alkuperäisen yhtälön ratkaisujen saamiseksi suoritetaan vakion variointi:

$$y(x) = C(x)y_H(x) = C(x)e^{-x^2},$$

mistä nähdään, että $y' = C'e^{-x^2} + Ce^{-x^2}(-2x) = C'y_H - 2xy_HC$. Sijoittamalla tämä alkuperäiseen yhtälöön saadaan

$$C'y_H - 2xy_HC + 2xCy_H = x,$$

mikä voidaan edelleen kirjoittaa muotoon

$$C'(x)y_H(x) = x.$$

Tästä funktiolla $y_H(x) = e^{-x^2}$ jakamalla saadaan differentiaaliyhtälö $C'(x) = xe^{x^2}$. Funktio $C(x)$ saadaan tästä integroimalla

$$C(x) = \int xe^{x^2} dx = \frac{1}{2}e^{x^2} + B.$$

Alkuperäisen differentiaaliyhtälön ratkaisu on siis muotoa

$$y(x) = C(x)y_H(x) = \left(\frac{1}{2}e^{x^2} + B\right)e^{-x^2} = \frac{1}{2} + Be^{-x^2},$$

missä B on vakio.

Kerrataan vielä menetelmä ensimmäisen kertaluvun differentiaaliyhtälön

$$y' + a(x)y = b(x) \quad (5.16)$$

ratkaisemiseksi. Menetelmä jakautuu kahteen vaiheeseen, joista ensimmäinen on yhtälön (5.16) *homogenisointi* ja homogeenisen yhtälön ratkaisu. Homogenisointi merkitsee funktion $b(x)$ korvaamista nollafunktiolla, ja ratkaisu homogeeniselle yhtälölle saadaan jakamalla y :llä ja huomaamalla, että logaritmin derivointisäännön perusteella voidaan kirjoittaa $\frac{y'}{y} = D \ln y$. Tällöin homogeenisen yhtälön ratkaisu palautuu integrointitehtävään, kuten aiemmin on todettu. Kun homogeenisen yhtälön yksi ratkaisu y_H on saatu (tällöin myös Cy_H , missä C on vakio, on homogeenisen yhtälön ratkaisu), siirrytään vaiheeseen, joka tunnetaan nimellä *vakion variointi*. Tämä merkitsee sitä, että homogeenisen yhtälön ratkaisussa Cy_H vakio C korvataan funktiolla $C(x)$. Tämän jälkeen sijoitetaan $y = C(x)y_H$ alkuperäiseen differentiaaliyhtälöön. Kun otetaan huomioon, että $y' = C'(x)y_H + C(x)y'_H$, voidaan todeta että sijoitus tuottaa (sievennysten jälkeen) yhtälön, josta $C(x)$ voidaan ratkaista integroimalla.

Tämänkaltaista täydellistä ratkaisumenetelmää ei kuitenkaan ole esimerkiksi *toisen kertaluvun* lineaariselle differentiaaliyhtälölle

$$y'' + a(x)y' + b(x)y = c(x), \quad (5.17)$$

ei edes homogeenisessa tapauksessa $c(x) = 0$, mutta sen sijaan tunnetaan lukuisia differentiaaliyhtälötyyppejä, jotka voidaan ratkaista.

5.2.3 Toisen kertaluvun lineaariset DY:t

Esimerkki 114 Voidaan todeta, että $y_1(x) = x$ ja $y_2(x) = \frac{1}{x}$ ovat molemmat differentiaaliyhtälön

$$x^2y'' + xy' - y = 0 \quad (5.18)$$

ratkaisuja. Nämä ratkaisut ovat myös (lineaarisesti) riippumattomat, sillä jos kaikilla x :n arvoilla on $C_1x + C_2\frac{1}{x} = 0$, on $C_1x^2 + C_2 = 0$, joten $C_1 = C_2 = 0$. Yleisen lineaarisia differentiaaliyhtälöitä koskevan periaatteen mukaan tällöin yhtälön (5.18) yleinen ratkaisu on siis muotoa $C_1x + C_2\frac{1}{x}$.

Aiemmin on esitetty menetelmiä, joilla toisen kertaluvun differentiaaliyhtälöille saadaan ratkaisuja vakiokertoimisessa tapauksessa. Homogeenisen vakiokertoimisen yhtälön tapauksessa kannattaa käyttää yrittelyä $y = e^{\lambda x}$ ja ratkaista λ . Yksittäisratkaisun löytämiseen epähomogeeniselle yhtälölle taas soveltuvat Laplace-muunnokset.

Yleistä ratkaisumenetelmää toisen kertaluvun differentiaaliyhtälöille ei kuitenkaan ole. Toisaalta, jos toisen kertaluvun lineaariselle homogeeniselle differentiaaliyhtälölle on jollakin tavalla löydetty *jokin* ratkaisu y_1 , voidaan yleiselle yhtälölle saada muita ratkaisuja sijoituksella $y = y_1v$: Tällöin $y' = y_1'v + y_1v'$ ja $y'' = y_1''v + y_1'v' + y_1'v' + y_1v'' = y_1''v + 2y_1'v' + y_1v''$ ja sijoitus yhtälöön

$$y'' + ay' + by = c$$

antaa

$$\begin{aligned} y_1''v + 2y_1'v' + y_1v'' + a(y_1'v + y_1v') + by_1v &= c \\ \Leftrightarrow y_1v'' + 2y_1'v' + ay_1v' + (y_1'' + ay_1' + by_1)v &= c \\ \Leftrightarrow y_1w' + (ay_1 + 2y_1')w &= c, \end{aligned}$$

missä viimeisellä rivillä on merkitty $w = v'$. Viimeksi saatu yhtälö on ensimmäisen kertaluvun lineaarinen differentiaaliyhtälö funktion w suhteen, joten se voidaan ratkaista ja lopuksi v saadaan integroimalla yhtälö $v' = w$.

Esimerkki 115 Olkoon r vakio ja $y = y(t)$ muuttujan t funktio. Etsitään differentiaaliyhtälön

$$y'' - 2ry' + r^2y = t \quad (5.19)$$

ratkaisut ilman Laplace-muunnoksia. Sijoitetaan $y = e^{\lambda t}$ homogenisoituun yhtälöön

$$y'' - 2ry' + r^2y = 0,$$

jolloin saadaan

$$\lambda^2 e^{\lambda t} - 2r\lambda e^{\lambda t} + r^2 e^{\lambda t} = 0,$$

ja eksponenttifunktio jakamalla $\lambda^2 - 2r\lambda + r^2 = 0$, jonka ainoa ratkaisu on $\lambda = r$.

Sijoitus $y = e^{rt}v$ antaa $y' = re^{rt}v + e^{rt}v'$ ja $y'' = r^2e^{rt}v + 2re^{rt}v' + e^{rt}v''$. Sijoittamalla nämä yhtälöön (5.19) saadaan

$$r^2e^{rt}v + 2re^{rt}v' + e^{rt}v'' - 2r^2e^{rt}v - 2re^{rt}v' + r^2e^{rt}v = t,$$

joka sievenee muotoon

$$v'' = te^{-rt}.$$

Tästä saadaan ensin $v' = -\frac{t}{r}e^{-rt} - \frac{1}{r^2}e^{-rt} + C_1$ ja toiseen kertaan integroimalla

$$v = \frac{t}{r^2}e^{-rt} + \frac{2}{r^3}e^{-rt} + C_1t + C_2.$$

Lopuksi

$$y = e^{rt}v = \frac{t}{r^2} + \frac{2}{r^3} + C_1te^{rt} + C_2e^{rt}. \quad (5.20)$$

Voidaan todeta, että te^{rt} ja e^{rt} ovat lineaarisesti riippumattomat homogeenisen yhtälön ratkaisut ja että $\frac{t}{r^2} + \frac{2}{r^3}$ on yhtälön (5.19) yksittäisratkaisu. Lauseke (5.20) sisältää siis *kaikki* differentiaaliyhtälön (5.19) ratkaisut.

5.3 Separoituvat DY:t

Differentiaaliyhtälöä kutsutaan *separoituvaksi*, mikäli se voidaan kirjoittaa muotoon

$$y' = g(x)f(y), \quad (5.21)$$

missä $g(x)$ riippuu ainoastaan muuttujasta x ja $f(y)$ ainoastaan y :stä. Jos merkitään $G(x) = \int g(x) dx$ ja $F(y) = \int \frac{1}{f(y)} dy$, voidaan yhtälö (5.21) kirjoittaa muotoon

$$\frac{d}{dx}F(y(x)) = \frac{d}{dx}G(x), \quad (5.22)$$

mistä saadaan $F(y(x)) = G(x) + C$, missä C on vakio.

Käytännössä separoituva yhtälö $\frac{dy}{dx} = g(x)f(y)$ ratkaistaan kirjoittamalla se muotoon

$$\frac{1}{f(y)} dy = g(x) dx$$

ja integroimalla puolittain:

$$\int \frac{1}{f(y)} dy = \int g(x) dx + C.$$

Esimerkki 116 Tarkastellaan esimerkin 111 differentiaaliyhtälöä $\frac{dy}{dx} = -y^2$ ja kirjoitetaan tämä muotoon

$$-\frac{1}{y^2} dy = dx$$

ja integroidaan:

$$\int -\frac{1}{y^2} dy = \int dx + C,$$

mistä saadaan

$$\frac{1}{y} = x + C,$$

josta edelleen voidaan ratkaista $y = \frac{1}{x+C}$.

Esimerkki 117 Verhulstin populaatiomallissa populaation koko p toteuttaa differentiaaliyhtälön

$$p' = ap - bp^2, \quad (5.23)$$

Missä a ja b ovat positiivisia vakioita. Differentiaaliyhtälöä (5.23) kutsutaan *logistiseksi* yhtälöksi. Logistisen yhtälön

$$\frac{dp}{dt} = ap - bp^2$$

ratkaisu saadaan separoimalla:

$$\int \frac{1}{ap - bp^2} dp = \int dt.$$

Tämä voidaan sieventää muotoon (miten?)

$$\frac{1}{a} \ln \frac{p}{a - bp} = t + C,$$

mistä edelleen saadaan

$$\frac{p}{a - bp} = e^{a(t+C)}, \quad (5.24)$$

josta voidaan ratkaista

$$p = \frac{a}{b + e^{-aC} e^{-at}}. \quad (5.25)$$

Sijoittamalla $t = 0$ yhtälöön (5.24) antaa (merkitään $p_0 = p(0)$) $\frac{p_0}{a - bp_0} = e^{aC}$, ja edelleen sijoittamalla tämä yhtälöön (5.25) saadaan ratkaisuksi

$$p(t) = \frac{ap_0}{bp_0 + (a - bp_0)e^{-at}}.$$

Esimerkki 118 Newtonin liikeyhtälön mukaan kappaleeseen vaikuttavien voimien summa on yhtä kuin kappaleen massa kertaa kiihtyvyys. Putoavaan kappaleeseen vaikuttaa maan vetovoima suuruudella mg ja ilmanvastuksen aiheuttama voima suuruudella $-kv^2$, missä v on kappaleen nopeus ja k on vakio. Tällöin putoavan kappaleen liikettä kuvaa yhtälö

$$mg - kv^2 = ma,$$

missä a on kappaleen kiihtyvyys. Koska $v'(t) = a$, voidaan ylläoleva yhtälö kirjoittaa muotoon

$$\frac{dv}{dt} = g - \frac{k}{m} v^2,$$

josta muuttujat separoimalla ja integroimalla saadaan

$$\int \frac{1}{g - \frac{k}{m} v^2} = \int dt = t + C.$$

Yhtälön vasemman puolen määrittämiseksi käytetään kaavaa

$$\int \frac{dx}{a^2 - x^2} = \frac{1}{2a} \ln \left| \frac{a+x}{a-x} \right|$$

(kaava 28, matematiikan laitoksen kaavakokoelma), jonka mukaan

$$\int \frac{1}{g - \frac{k}{m}v^2} dv = \frac{m}{k} \int \frac{1}{\frac{gm}{k} - v^2} dv = \frac{1}{2} \sqrt{\frac{m}{gk}} \ln \frac{\sqrt{\frac{gm}{k}} + v}{\sqrt{\frac{gm}{k}} - v}.$$

Merkinnällä $\alpha = \sqrt{\frac{gm}{k}}$ saadaan yhtälö muotoon

$$\frac{\alpha}{2g} \ln \frac{\alpha + v}{\alpha - v} = t + C,$$

josta saadaan

$$\frac{\alpha + v}{\alpha - v} = e^{\frac{2g}{\alpha}(t+C)}.$$

Merkittäessä edelleen $\beta(t) = e^{\frac{2g}{\alpha}t}$ ja $C_1 = e^{\frac{2g}{\alpha}C}$ saadaan yhtälön ratkaisuksi

$$v(t) = \alpha \frac{C_1 - \beta(t)^{-1}}{C_1 + \beta(t)^{-1}}.$$

Koska funktio $\beta(t) = e^{\frac{2g}{\alpha}t}$ lähenee ääretöntä t :n kasvaessa, nähdään helposti, että $v(t)$ lähenee arvoa $\alpha = \sqrt{\frac{gm}{k}}$. Kyseistä arvoa kutsutaan nimellä *rajanopeus* (engl. *terminal velocity*).

5.4 Eksaktit DY:t

Jos ensimmäisen kertaluvun differentiaaliyhtälössä ei esiinny derivaatan y' potensseja, voidaan y' ottaa yhteiseksi tekijäksi ja kirjoittaa yhtälö muotoon

$$f(x, y) + g(x, y)y' = 0. \quad (5.26)$$

Tämä yhtälö voidaan toisinaan ratkaista seuraavaa menettelyä käyttäen.

Jos kahden muuttujan funktiossa $F(x, y)$ muuttuja $y = y(x)$ riippuu edelleen muuttujasta x , on kyseessä itse asiassa yhden muuttujan funktio $g(x) = F(x, y(x))$. Tällöin derivaatta $g'(x)$ lasketaan seuraavasti:

$$g'(x) = \frac{\partial F}{\partial x} \frac{dx}{dx} + \frac{\partial F}{\partial y} \frac{dy}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} y' \quad (5.27)$$

(miksi näin on, selvitetään myöhemmin).

Määritelmä 53 Jos on olemassa kahden muuttujan funktio $F(x, y)$, jolle $\frac{\partial F}{\partial x} = f(x, y)$ ja $\frac{\partial F}{\partial y} = g(x, y)$, sanotaan, että differentiaaliyhtälö (5.26) on *eksakti*.

Jos (5.26) on eksakti, on

$$g'(x) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} y' = f(x, y) + g(x, y)y' = 0,$$

jonka ratkaisu on siis $g(x) = F(x, y) = C$ (vakio).

Kokoamalla kaiken yllämainitun yhteen saadaan seuraava tulos: Jos on olemassa sellainen kahden muuttujan funktio $F(x, y)$, että $\frac{\partial F}{\partial x} = f(x, y)$ ja $\frac{\partial F}{\partial y} = g(x, y)$, niin differentiaaliyhtälön (5.26) ratkaisu on

$$F(x, y) = C.$$

Luonnollisesti heräävä kysymys onkin, millä ehdoin voidaan löytää kahden muuttujan funktio $F(x, y)$, jolle ehdot $\frac{\partial F}{\partial x} = f(x, y)$ ja $\frac{\partial F}{\partial y} = g(x, y)$ pätevät.

On mahdollista todistaa, että funktio $F(x, y)$ on olemassa tarkalleen silloin, kun

$$\frac{\partial f(x, y)}{\partial y} = \frac{\partial g(x, y)}{\partial x}, \quad (5.28)$$

kunhan $f(x, y)$ ja $g(x, y)$ ovat riittävän säännöllisiä (toisen kertaluvun osittaisderivaatat jatkuvia).

Esimerkki 119 Differentiaaliyhtälö

$$\frac{x}{\sqrt{x^2 + y^2}} - 1 + \frac{y}{\sqrt{x^2 + y^2}} y' = 0 \quad (5.29)$$

on eksakti, sillä $F(x, y) = \sqrt{x^2 + y^2} - x$ toteuttaa

$$\frac{\partial F}{\partial x} = \frac{2x}{2\sqrt{x^2 + y^2}} - 1 = \frac{x}{\sqrt{x^2 + y^2}} - 1$$

ja

$$\frac{\partial F}{\partial y} = \frac{2y}{2\sqrt{x^2 + y^2}} = \frac{y}{\sqrt{x^2 + y^2}}.$$

Yhtälön (5.29) ratkaisuksi saadaan siis $F(x, y) = C$, mikä voidaan kirjoittaa muotoon

$$\sqrt{x^2 + y^2} - x = C$$

ja edelleen muotoon

$$x = \frac{y^2 - C^2}{2C}.$$

Vaikka yhtälö (5.26) ei olisikaan alun perin eksakti, on toisinaan mahdollista löytää ns. *integroiva tekijä*, jolla kerrottuna yhtälöstä tulee eksakti.

Esimerkki 120 Yhtälö

$$(x + y^2)y' - y = 0$$

ei ole eksakti, koska $\frac{\partial}{\partial x}(x + y^2) = 1 \neq -1 = \frac{\partial}{\partial y}(-y)$. Tällö yhtälöllä on kuitenkin integroiva tekijä $\frac{1}{y^2}$, sillä yhtälö

$$\left(\frac{x}{y^2} + 1\right)y' - \frac{1}{y} = 0 \quad (5.30)$$

on eksakti: $\frac{\partial}{\partial x}\left(\frac{x}{y^2} + 1\right) = \frac{1}{y^2} = \frac{\partial}{\partial y}\left(-\frac{1}{y}\right)$. Etsitään $F(x, y)$, jolle $\frac{\partial}{\partial y}F(x, y) = \frac{x}{y^2} + 1$ ja $\frac{\partial}{\partial x}F(x, y) = -\frac{1}{y}$. Ensimmäisestä yhtälöstä nähdään, että $F(x, y) = -\frac{x}{y} + y + f(x)$ ja toisesta $F(x, y) = -\frac{x}{y} + g(y)$, missä f riippuu vain x :stä ja g vain y :stä. Tällöin $g(y) = y + f(x)$, jolloin on oltava $f(x) = C$ (vakio). Näin ollen

$$F(x, y) = -\frac{x}{y} + y + C$$

ja (5.30) voidaan kirjoittaa muotoon

$$\frac{d}{dx}\left(y - \frac{x}{y} + C\right) = 0,$$

mistä ratkaisuksi saadaan $y - \frac{x}{y} + C = C_1$, toisin sanoen $y - \frac{x}{y} = C_2 = C_1 - C$.

Esimerkki 121 Ensimmäisen kertaluvun lineaarinen differentiaaliyhtälö olisi voitu ratkaista myös seuraavasti: Yritetään löytää yhtälölle

$$y' + a(x)y - b(x) = 0$$

integroiva tekijä $\mu(x)$. Jos tällainen on olemassa, on yhtälö

$$\mu(x)y' + \mu(x)a(x)y - \mu(x)b(x) = 0 \quad (5.31)$$

eksakti, jolloin siis ehdon (5.28) nojalla pitäisi olla

$$\frac{\partial}{\partial y}(\mu(x)a(x)y - \mu(x)b(x)) = \frac{\partial}{\partial x}\mu(x),$$

mikä voidaan kirjoittaa muotoon

$$\mu(x)a(x) = \mu'(x). \quad (5.32)$$

Oletetaan aluksi, että integroiva tekijä $\mu(x)$ löytyy, jolloin siis (5.32) pätee. Tällöin voidaan yhtälö (5.31) kirjoittaa muotoon

$$\mu(x)y' + \mu'(x)y - \mu(x)b(x) = 0,$$

mikä puolestaan voidaan tulon derivointisäännön nojalla kirjoittaa muotoon

$$\frac{d}{dx}(\mu(x)y) = \mu(x)b(x),$$

ja ratkaisuksi saadaan

$$y = \frac{1}{\mu(x)} \left(\int \mu(x)b(x) dx + C \right). \quad (5.33)$$

On vielä selvitettävä miten integroiva tekijä $\mu(x)$ voidaan löytää, mutta tämä selviää helposti yhtälön (5.32) perusteella:

$$\frac{d}{dx} \ln \mu(x) = a(x),$$

minkä ratkaisuna saadaan

$$\mu(x) = \exp\left(\int a(x) dx\right).$$

Suoralla laskulla voidaan todeta, että tällä tavoin saatu ratkaisu (5.33) on sama kuin aiemmin esitetyllä menetelmällä (homogenisointi ja vakion variointi) saatu ratkaisu.

5.5 Sekalaisia menetelmiä

5.5.1 Sijoitukset $\frac{y}{x}$ ja $ax + by$

Muotoa $y' = f(y/x)$ ja $y' = f(ax + by)$ olevat differentiaaliyhtälöt muuntuvat separoituviksi yhtälöiksi sijoituksilla $\frac{y}{x} = z(x)$ ja $ax + by = z(x)$.

Esimerkki 122 Yhtälö $y' = x^2 + 2xy + y^2$ voidaan kirjoittaa muotoon $y' = (x + y)^2$, ja sijoittamalla $z = x + y$ saadaan $z' = 1 + y'$, joten päädytään yhtälöön

$$z' - 1 = z^2,$$

joka saadaan muotoon

$$\frac{1}{z^2 + 1} dz = dx.$$

integroimalla saadaan $\arctan z = x + C$, mistä $z = \tan(x + C)$. Alkuperäinen funktio sijoittamalla saadaan $y = -x + \tan(x + C)$.

5.5.2 Bernoullin DY

Sijoitus $z = y^{1-a}$ ($a \neq 1$) muuntaa Bernoullin differentiaaliyhtälön

$$y' + p(x)y = q(x)y^a \quad (5.34)$$

lineaariseksi differentiaaliyhtälöksi.

Esimerkki 123 Logistinen differentiaaliyhtälö

$$p' = ap - bp^2$$

on separoituva, mutta myös muotoa (5.34). Sijoittamalla $z = p^{-1}$ saadaan $z' = -p^{-2}p'$, ja yhtälö muuntuu muotoon

$$-z'p^2 = ap - bp^2,$$

mistä $-p^2$:lla jakamalla saadaan

$$z' + az = b. \quad (5.35)$$

Näin saadun lineaarisen differentiaaliyhtälön ratkaisut saadaan aluksi homogenisoimalla: $z_H = Ce^{-at}$ on homogeenisen yhtälön ratkaisu, minkä jälkeen vakion variointi tuottaa yhtälön (5.35) ratkaisuksi

$$z = \frac{b}{a} + Ce^{-at}.$$

Alkuperäisen yhtälön ratkaisuksi saadaan tällöin

$$p = \frac{1}{\frac{b}{a} + Ce^{-at}} = \frac{a}{b + aCe^{-at}},$$

mikä saadaan muotoon

$$p = \frac{a}{b + a(\frac{1}{p_0} - \frac{b}{a})e^{-at}} = \frac{ap_0}{bp_0 + (a - bp_0)e^{-at}}$$

merkitsemällä $p_0 = p(0)$.

5.5.3 Käänteisfunktioon siirtyminen

Differentiaaliyhtälö saattaa toisinaan olla helpommin ratkaistavissa funktion $y = y(x)$ sijaan tämän käänteisfunktion $x = x(y)$ differentiaaliyhtälönä. Tällöin on muistettava, että $x'(y) = \frac{1}{y'(x)}$ ($\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$).

Esimerkki 124 Esimerkin 120 differentiaaliyhtälö voidaan kirjoittaa muodossa

$$x' - \frac{1}{y}x = y,$$

mikä on lineaarinen differentiaaliyhtälö funktion $x(y)$ suhteen.

5.5.4 Muuttujan vaihtaminen

Eulerin differentiaaliyhtälö on muotoa

$$x^2y'' + pxy' + qy = f(x), \quad (5.36)$$

missä p ja q ovat vakioita. Jos merkitään $t = \ln x$ ja $y_1(t) = y(e^t)$ (jolloin $y(x) = y_1(\ln x)$), saadaan ketjusäännöllä

$$\frac{dy}{dx} = \frac{dy_1}{dt} \frac{dt}{dx} = \frac{1}{x} \frac{dy_1}{dt}$$

ja

$$\frac{d^2y}{dx^2} = -\frac{1}{x^2} \frac{dy_1}{dt} + \frac{d^2y_1}{dt^2} \frac{dt}{dx} = -\frac{1}{x^2} y_1'(t) + \frac{1}{x^2} y_1''(t).$$

Sijoittamalla nämä alkuperäiseen yhtälöön saadaan

$$y_1''(t) + (p-1)y_1'(t) + qy_1(t) = f(e^t),$$

jonka ratkaisut voidaan ainakin periaatteessa määrittää Laplace-muunnosta käyttämällä.

5.5.5 Yhtälö $y'' = f(y)$

Toisen kertaluvun differentiaaliyhtälö

$$y'' = f(y) \quad (5.37)$$

voidaan palauttaa ensimmäisen kertaluvun differentiaaliyhtälöksi kertomalla (5.37) tekijällä $2y'$. Tällöin saadaan $2y'y'' = 2f(y)y'$, mikä voidaan kirjoittaa muotoon

$$\frac{d}{dx}(y')^2 = 2 \frac{d}{dx} \int f(y) dy,$$

mistä

$$(y')^2 = 2 \int f(y) dy + C. \quad (5.38)$$

Ottamalla tästä neliöjuuri puolittain saadaan separoituva differentiaaliyhtälö.

Esimerkki 125 Newtonin liikeyhtälön mukaan kappaleeseen vaikuttava voima on $F = m \frac{d^2s}{dt^2}$, missä s on kappaleen paikkakoordinaatti. Toisaalta Newtonin gravitaatiolain mukaan m ja M -massaisten kappaleiden välillä vallitsee vetovoima, joka on verrannollinen kappaleiden massaun ja kääntäen verrannollinen niiden etäisyyden neliöön s :

$$F = G \frac{mM}{s^2}.$$

Tällöin siis liike painovoimakentässä esitetään yhtälöllä

$$ms''(t) = -G \frac{mM}{s^2},$$

(liikkeen ajatellaan suuntautuvan pois päin massasta M) joka kirjoittamalla $GM = k$ saadaan muotoon

$$s'' = -\frac{k}{s^2}.$$

Kertomalla näin saatu yhtälö $2s'$:llä saadaan

$$2s's'' - 2\frac{k}{s^2}s',$$

mikä on yhtäpitävä yhtälön

$$\frac{d}{dt}(s')^2 = 2k \frac{d}{dt} \frac{1}{s} \quad (5.39)$$

kanssa. Yhtälöstä (5.39) saadaan

$$(s')^2 = \frac{2k}{s} + C. \quad (5.40)$$

Alkuehdot $s(0) = s_0 > 0$ ja $s'(0) = v_0 > 0$ antavat $C = v_0^2 - \frac{2k}{s_0}$ ja siis

$$s' = \sqrt{\frac{2k}{s} + v_0^2 - \frac{2k}{s_0}}, \quad (5.41)$$

mistä s voidaan ratkaista (johtaa tosin työlääseen integrointitehtävään). Vaikka yhtälöä (5.41) ei ratkaistakaan, sisältävät siihen johtaneet välivaiheet jo paljon informaatiota.

Yhtälöstä (5.40) saadaan

$$\frac{1}{2}mv^2 - G \frac{Mm}{s} = \frac{Cm}{2} = \frac{1}{2}mv_0^2 - G \frac{Mm}{s_0},$$

joka on muodoltaan energian säilymislaki: Liike-energian $\frac{1}{2}mv^2$ ja potentiaalienergian $-G\frac{Mm}{s}$ summa pysyy vakiona. Yhtälöstä (5.41) voidaan myös määrittää *pakonopeus*, ts. se nopeus, jolla maan pinnalta lähetetty kappale ei enää palaa. Ehtona voidaan käyttää sitä, että nopeudella s' ei ole nollakohtaa, mikä toteutuu, jos

$$v_0^2 - \frac{2k}{s_0} > 0.$$

Sijoittamalla tähän vakiot $s_0 = 6,378 \cdot 10^6$ m (maan säde) $G = 6,674 \cdot 10^{-11} \frac{\text{m}^3}{\text{kg s}^2}$ (gravitaatiovakio) ja $M = 5,9737 \cdot 10^{24}$ kg (maan massa) saadaan ehdoksi $v_0 > 11,19 \frac{\text{km}}{\text{s}}$.

Yleisemminkin voidaan todeta, että itse differentiaaliyhtälö tarjoaa toisinaan tietoa integraalikäyrästä, vaikka sitä ei olisikaan ratkaistu.

Esimerkki 126 Yhtälöstä $y' = y(1 - y)$ nähdään, että $y' > 0$, kun $0 < y < 1$, joten tällä välillä y on kasvava. Kertaalleen derivoimalla saadaan

$$y'' = y'(1 - y) - yy' = y'(1 - 2y) = y(1 - y)(1 - 2y),$$

mistä voidaan päätellä, että integraalikäyrä on alueessa $0 < y < \frac{1}{2}$ alaspäin kupera ($y'' > 0$) ja alueessa $\frac{1}{2} < y < 1$ ylöspäin kupera ($y'' < 0$).

5.6 Differentiaaliyhtälöiden sarjaratkaisut

Aiemmin tässä luvussa on todettu, että melko harvoille differentiaaliyhtälötyypeille on olemassa yleistä ratkaisumenetelmää. Sarjaoppi tarjoaa mahdollisuuden ratkaisujen löytämiseen joissakin selkeissä tapauksissa, joissa aiemmat menetelmät eivät toimi.

Luonnontieteissä, erityisesti fysiikassa erityisen tärkeitä ovat toisen kertaluvun differentiaaliyhtälöt

$$y'' + a(x)y' + b(x)y = c(x),$$

joille ei kuitenkaan ole olemassa yleistä ratkaisumenetelmää. Aiemmin kuitenkin huomattiin, että jos vastavalle homogeeniselle yhtälölle

$$y'' + a(x)y' + b(x)y = 0$$

löytyy jokin ratkaisu, niin alkuperäisen yhtälön kaikki ratkaisut voidaan löytää. Tällöin homogeenisen yhtälön ratkaisua voidaan etsiä merkitsemällä tuntematonta funktiota y potenssisarjalla

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots,$$

kehittämällä $a(x)$ ja $b(x)$ sarjoiksi, ja selvittämällä näin minkälainen on jono $a_0, a_1, a_2, \dots$

Esimerkki 127 Valon diffraktioon liittyvä Airyn differentiaaliyhtälö

$$y'' + xy = 0$$

on muodoltaan yksinkertainen, mutta mikään Insinöörimatematiikka aiemmin esitetty ratkaisumenetelmä ei siihen pysty.

Valitaan alkuehdoiksi $y(0) = 0$ ja $y'(0) = 1$ ja merkitään

$$y(x) = \sum_{n=0}^{\infty} a_n x^n,$$

jolloin alkuehdot merkitsevät, että $a_0 = 0$, $a_1 = 1$ ja

$$y'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

ja

$$y''(x) = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}.$$

Nyt siis

$$xy = \sum_{n=0}^{\infty} a_n x^{n+1} = \sum_{n=3}^{\infty} a_{n-3} x^{n-2}$$

ja differentiaaliyhtälön vasen puoli saa muodon

$$y'' + xy = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} + \sum_{n=3}^{\infty} a_{n-3} x^{n-2} = 2a_2 + \sum_{n=3}^{\infty} (n(n-1)a_n + a_{n-3})x^{n-2}.$$

Koska tämän pitää olla nolla kaikilla x :n arvoilla, on oltava $a_2 = 0$ ja

$$a_n = -\frac{1}{n(n-1)}a_{n-3}$$

kaikilla $n \geq 3$. Tällöin siis $a_3 = -\frac{1}{3 \cdot 2}a_0 = 0$, $a_4 = -\frac{1}{4 \cdot 3}a_1 = -\frac{1}{4 \cdot 3}$, $a_5 = -\frac{1}{5 \cdot 4}a_2 = 0$, $a_6 = -\frac{1}{6 \cdot 5}a_3 = 0$, $a_7 = -\frac{1}{7 \cdot 6}a_4 = \frac{1}{7 \cdot 6 \cdot 4 \cdot 3}$, jne ja

$$y(x) = x - \frac{1}{4 \cdot 3}x^4 + \frac{1}{7 \cdot 6 \cdot 4 \cdot 3}x^7 - \dots$$

Näin saatu sarja suppenee kaikilla x :n arvoilla (miksi?), joten saatu sarja edustaa Airyn yhtälön ratkaisua koko reaaliakselilla.

Esimerkki 128 (Vetyatomin elektroni) Kvanttimekaniikassa hiukkasen liiketilaa kuvaan Schrödingere-
rin yhtälö, joka voidaan abstraktilla tasolla kirjoittaa muotoon

$$\frac{d}{dt}\psi(t) = -iH\psi(t),$$

missä ψ on hiukkasen käyttäytymistä kuvaava aaltofunktio ja H ns. Hamiltonin operaattori, joka esittää hiukkasen kokonaisenergiaa. Monissa esimerkeissä H sisältää liike-energian, jonka esittäminen puolestaan tuo yhtälöön toisen kertaluvun derivoiminnin paikkakoordinaatin suhteen.

Kvanttimekaniikka on luonteeltaan probabilistinen teoria, mikä merkitsee sitä, että aaltofunktio ψ edustaa hiukkasen tarkan paikan sijasta paikan todennäköisyysjakaumaa. Koska hiukkasen paikka ei ole eksaktisti määrätty minään ajanhetkenä, on täysin mielekäästä tarkastella myös ns. ajasta riippumattomaa Schrödingerin yhtälöä, joka saadaan aikariippuvaisesta Schrödingerin yhtälöstä muuttujat erottamalla. Ajasta riippumattoman Schrödingerin abstrakti muoto on

$$H\psi = E\psi,$$

missä H on Hamiltonin operaattori ja E hiukkasen kokonaisenergia. Tyypillisesti H :n esityksessä on mukana liike-energia $E = \frac{1}{2}mv^2 = \frac{1}{2m}(mv)^2 = \frac{1}{2m}p^2$, ja liikemäärän p kvanttimekaanisessa esitysmuodossa esiintyy derivaatta paikkakoordinaatin suhteen.

Ajasta riippumaton vetyatomin elektronin liiketilaa kuvaava Schrödingerin yhtälö on mahdollista esittää napakoordinaatiston yleistävässä pallokoordinaatistossa, ja jakaa yhtälöiksi, joissa esiintyy etäisyys ytimestä r , sekä kulmat θ ja ϕ . Pallokoordinaatistoa käsitellään tarkemmin seuraavassa luvussa, mutta tässä yhteydessä käsitellään yhtälöä, joka esittää elektronin etäisyyttä ytimestä. Olkoon tätä kuvaava ns. radiaalinen aaltofunktio R .

Yksikköjärjestelmä huomioituna radiaalinen Schrödingerin yhtälö saa muodon

$$-\frac{\hbar}{2\mu} \frac{d^2 R}{dr^2} + \left(-\frac{Ze^2}{4\pi\epsilon_0 r} + \frac{l(l+1)\hbar^2}{2\mu r^2} \right) R = ER, \quad (5.42)$$

missä $\hbar = \frac{h}{2\pi}$, h on Planckin vakio, $\mu = \frac{m_e m_y}{m_e + m_y}$ ns. redusoitu massa, jossa m_e ja m_y ovat elektronin ja ytimen massat, Z ytimen varausten määrä (vetyatomille $Z = 1$), ϵ_0 on tyhjiön permittiivisyys ja $l \geq 0$ on kokonaisluku, ns. sivukvanttiluku.

Fysikaalisen tulkinnan vuoksi differentiaaliyhtälölle (5.42) etsitään ratkaisuja, jotka toteuttavat ehdon

$$\int_0^\infty \frac{1}{r^2} |R|^2 dr = 1. \quad (5.43)$$

Erityisesti pitää siis olla $R(r) \xrightarrow{r \rightarrow 0} 0$ ja $\frac{|R|^2}{r^2} \xrightarrow{r \rightarrow \infty} 0$. Lisäksi tavanomaisesti tulkittuna potentiaalienergia on negatiivinen, ja nolla vasta ”äärettömän kaukana” ytimeistä. Tällöin myös elektronin energia E ajatellaan negatiiviseksi.

Yhtälön (5.42) mutkikkaita vakiokertoimia voidaan yksinkertaistaa ottamalla käyttöön uusi muuttuja $\rho = \sqrt{\frac{-8\mu E}{\hbar^2}} r$, jolloin

$$\frac{dR}{dr} = \frac{dR}{d\rho} \frac{d\rho}{dr} = \frac{dR}{d\rho} \sqrt{\frac{-8\mu E}{\hbar^2}} \quad \text{ja} \quad \frac{d^2 R}{dr^2} = \frac{d^2 R}{d\rho^2} \cdot \frac{-8\mu E}{\hbar^2},$$

ja sijoittamalla tämä yhtälöön (5.42) saadaan

$$\frac{d^2 R}{d\rho^2} + \left(\frac{\beta}{\rho} - \frac{l(l+1)}{\rho^2} \right) R = \frac{1}{4} R, \quad (5.44)$$

missä on merkitty $\beta = \sqrt{-\frac{\mu}{2E}} \frac{Ze^2}{4\pi\epsilon_0\hbar}$. Tälle yhtälölle on periaatteessa mahdollista etsiä sarjaratkaisu suoraan, mutta osoittautuu, että saatavasta sarjasta on vaikeaa määrittää ehdot (5.43) toteuttavaa ratkaisua.

Tämän ongelman kiertämiseksi tarkastellaan ensin yhtälöä, joka näyttää määräävän dynamiikan suurin piirtein: Jos ρ on suuri, on yhtälö (5.44) likimain muotoa

$$\frac{d^2 R}{d\rho^2} = \frac{1}{4} R. \quad (5.45)$$

Tämän ratkaisut saadaan yritteestä $R = e^{\lambda\rho}$, mikä johtaa algebralliseen yhtälöön

$$\lambda^2 = \frac{1}{4} \Rightarrow \lambda = \pm \frac{1}{2}.$$

Vaatumuksen (5.43) seuraus sulkee vaihtoehdon $\lambda = \frac{1}{2}$ pois, jolloin $R = e^{-\frac{1}{2}\rho}$ on yhtälön (5.45) ainoa vaadittua muotoa oleva ratkaisu. Toimitaan tämän jälkeen 1. kertaluvun lineaarisen yhtälön ratkaisumenetelmästä tutulla tavalla, siis esitetään yrite

$$R = e^{-\frac{1}{2}\rho} G(\rho),$$

joka yhtälöön (5.44) sijoittamalla tuottaa

$$G''(\rho) - G'(\rho) + \left(\frac{\beta}{\rho} - \frac{l(l+1)}{\rho^2} \right) G(\rho) = 0, \quad (5.46)$$

jolle etsitään ratkaisu sarjan

$$G(\rho) = \sum_{k=0}^{\infty} a_k \rho^k$$

avulla. Ehtojen (5.43) perusteella $a_0 = 0$ ja suoraan laskemalla

$$G' = \sum_{k=1}^{\infty} a_k k \rho^{k-1}$$

ja

$$G'' = \sum_{k=1}^{\infty} a_k k(k-1) \rho^{k-2} = \sum_{k=1}^{\infty} a_{k+1} (k+1) k \rho^{k-1}.$$

Huomataan vielä että

$$\frac{G}{\rho^2} = a_1 \rho^{-1} + \sum_{k=1}^{\infty} a_{k+1} \rho^{k+1},$$

ja sijoittamalla nämä yhtälöön (5.46) saadaan

$$\sum_{k=1}^{\infty} (a_{k+1}(k+1)k - a_k k + \beta a_k k - l(l+1)a_{k+1})\rho^{k-1} - a_1 \rho^{-1} l(l+1) = 0.$$

tällöin $a_1 = 0$ (ellei $l = 0$) ja

$$a_{k+1}(k(k+1) - l(l+1)) = a_k(k - \beta).$$

Jos $\beta \notin \mathbb{Z}$, huomataan sijoittamalla $k = l$ että $a_l = 0$. Edelleen sijoittamalla $k = l-1$ saadaan $a_{l-1} = 0$ ja näin jatkaen $a_{l-2} = \dots = a_1 = 0$, mutta kun $k > l$, on

$$\frac{a_{k+1}}{a_k} = \frac{k - \beta}{k(k+1) - l(l+1)} \approx \frac{1}{k}.$$

Käyttämällä vertailukohteen eksponenttifunktion sarjaan $e^\rho = \sum_{k=0}^{\infty} \frac{1}{k!} \rho^k$ havaitaan, että $G(\rho) \xrightarrow{\rho \rightarrow \infty} \infty$, jolloin ehdot (5.43) täyttäviä ratkaisuja ei löydy.

Ehdot (5.43) täyttävä ratkaisu löytyy siis vain, jos $\beta = n \in \mathbb{Z}$, jolloin vastaavat energiat ovat kvantittuneita:

$$E = E_n = -\frac{\mu Z^2 e^4}{2(4\pi\epsilon_0)^2 \hbar n^2}$$

Korkeammasta energiatilasta E_{n_1} matalampaan E_{n_2} siirtyessä elektroni vapauttaa fotonin, jonka energia on

$$E_{n_1} - E_{n_2} = -\frac{\mu Z^2 e^4}{2(4\pi\epsilon_0)^2 \hbar} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right). \quad (5.47)$$

Ylläoleva yhtälö tunnetaan nimellä Rydbergin kaava. Vapautuvien fotonien energia $E_{n_1} - E_{n_2}$ nähdään kokeellisesti vetyatomien emittoimassa sähkömagneettisessa säteilyssä spektriviivoina. Historiallisesti mielenkiintoisena seikkana on syytä mainita, että tämän kaavan perusteella tunnistettiin ”kaksi kertaa raskaampi vety” vuonna 1932, ennen neutronin löytymistä. Tunnistus perustui siihen, että deuteriumilla μ on likimain kaksinkertainen vetyyn nähden, muutoin kaikki on kuten kaavassa (5.47).

Kokonaislukua n kutsutaan *pääkvanttiluvuksi*. Sarjakehitelmän avulla on mahdollista määrittää funktion $R(r)$ muoto eri pää- ja sivukvanttiluvun arvoilla. Jos radiaalisen yhtälön 5.42 lisäksi ratkaistaan Schrödingerin yhtälön pallokoordinaatiston kulmista riippuvat osat, on mahdollista määrittää elektronin orbitaalien muoto eri energiatasoilla.

Esimerkki 129 Laplacen lämmönjohtumisyhtälön mukaan ohuen levyn lämpötilalle $T(x, y)$ pätee lämpötilan tasaannuttua

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \quad (5.48)$$

Tämänkaltaisten osittaisdifferentiaaliyhtälöiden ratkaisun löytämiseksi toimitaan yleensä siten, että aluksi etsitään yhtälölle (5.48) muotoa

$$T(x, y) = X(x)Y(y) \quad (5.49)$$

olevia ratkaisuja, missä X riippuu pelkästään x :stä ja Y pelkästään y :stä. Sijoittamalla (5.49) yhtälöön (5.48) saadaan

$$Y \frac{d^2 X}{dx^2} + X \frac{d^2 Y}{dy^2} = 0 \leftrightarrow \frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} = 0$$

Koska ensimmäinen summattava riippuu vain x :stä ja toinen vain y :stä, on oltava

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -\frac{1}{Y} \frac{d^2 Y}{dy^2} = C,$$

missä C on vakio. Merkitään $C = -k^2$, ($k > 0$) jolloin saadaan kaksi differentiaaliyhtälöä

$$X'' = -k^2 X \quad \text{ja} \quad Y'' = k^2 Y,$$

joiden ratkaisut ovat

$$X(x) = C_1 \sin kx + C_2 \cos kx \quad \text{ja} \quad Y(y) = D_1 e^{ky} + D_2 e^{-ky},$$

missä C_1, C_2, D_1 ja D_2 ovat vakioita. Tällöin siis $T(x, y)$ koostuu ratkaisujen

$$e^{ky} \sin kx, \quad e^{-ky} \sin kx, \quad e^{ky} \cos kx, \quad \text{ja} \quad e^{-ky} \cos kx$$

lineaarikombinaatioista, joissa arvo k tulee valita sopivasti. Valinta tapahtuu reunaehtojen perusteella: Olkoon $T(x, 0) = 100$, kun $x \in [0, 10]$ ja $T(0, y) = T(10, y) = 0$, kun $y > 0$, sekä $T(x, y) \rightarrow 0$ kun $x \in [0, 10]$ ja $y \rightarrow \infty$. Viimeisimmän ehdon perusteella 1. ja 3. muoto esiintyy ratkaisussa vain kertoimella 0. Lisäksi ehdon $T(0, y) = 0$ perusteella 4. ratkaisu esiintyy myös kertoimella 0.

Täten siis vain muotoa $e^{-ky} \sin kx$ oleva ratkaisu esiintyy nollasta poikkeavalla kertoimella, ja ehdon $T(10, y) = 0$ nojalla tulee olla $\sin(10 \cdot k) = 0$, mistä $10k = n\pi$, $n \in \mathbb{N}$, siis $k = \frac{n\pi}{10}$. Yhtälön lineaarisuuden perusteella taas kaikki näiden ratkaisujen lineaarikombinaatiot ovat ratkaisuja, jolloin ratkaisuksi saadaan

$$T(x, y) = \sum_{n=1}^{\infty} b_n e^{-\frac{n\pi}{10}y} \sin \frac{n\pi}{10}x,$$

ja määrättäväksi jäävät kertoimet $b_1, b_2, b_3, \dots$. Nämä taas saadaan reunaehdosta $T(x, 0) = 100$, kun $x \in [0, 10]$, jolloin siis

$$100 = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi}{10}x.$$

Yllä oleva yhtälö puolestaan sanoo, että kertoimet $b_1, b_2, b_3, \dots$ ovat vakiofunktion 100 fourier-kertoimet, kun funktion ajatellaan olevan 20-jaksoinen. Kertoimet b_n saadaan siis Fourier-sarjojen teorian avulla, lopputuloksena

$$b_n = \begin{cases} \frac{400}{n\pi}, & \text{jos } n \text{ on pariton,} \\ 0, & \text{jos } n \text{ on parillinen.} \end{cases}$$

Tällöin siis

$$T(x, y) = \frac{400}{\pi} \left(e^{-\frac{\pi}{10}y} \sin \frac{\pi}{10}x + \frac{1}{3} e^{-\frac{3\pi}{10}y} \sin \frac{3\pi}{10}x + \frac{1}{5} e^{-\frac{5\pi}{10}y} \sin \frac{5\pi}{10}x + \dots \right).$$

Ratkaisun löytäminen Laplaceen lämmönjohtumisyhtälölle on siis varsin työlästä, mutta vielä työläämpää on näyttää toteen, että kaikki ratkaisut todella saadaan tällä tavalla.

5.7 Lineaarisista vakiokertomisista DY-ryhmistä

Lineaaristen, vakiokertoimisten differentiaaliyhtälöryhmien ratkaisemiseksi jopa eksaktisti on olemassa periaatteessa toimiva menetelmä, joka kuvaillaan tässä luvussa. Menetelmän käytännön toimivuus riippuu kuitenkin mahdollisuudesta löytää diagonaalinen tai Jordan-esitys DY-ryhmän matriisille, kuin myös mahdollisuudesta löytää eksplisiittisesti antiderivaatta yhtälöryhmässä tai sen käsittelyssä esiintyvillä funktioilla.

Menetelmä ei ole mitenkään yllättävä. Päinvastoin, se on täysin suoraviivainen yleistys menetelmästä, jolla ratkaistaan 1. kertaluvun lineaarinen (vakiokertoiminen) differentiaaliyhtälö. Yllättävänä voisi pikemminkin pitää sitä kuinka paljon mekaanista laskentaa voisi periaatteessa vähentää analysoimalla menetelmää tarkemmin, mutta se ei ole tässä yhteydessä ensisijainen tavoite. Tällä kurssilla pyritään esittämään lähinnä periaate lineaaristen DY-ryhmien ratkaisemiseksi, vaikkakin ratkaisujen löytäminen laskennallisesti onkin työlästä.

Lineaarinen, vakiokertoiminen DY-ryhmä on muotoa

$$\begin{cases} x'_1 = a_{11}x_1 + \dots + a_{1n}x_n + f_1(t) \\ x'_2 = a_{21}x_1 + \dots + a_{2n}x_n + f_2(t) \\ \vdots \\ x'_n = a_{n1}x_1 + \dots + a_{nn}x_n + f_n(t) \end{cases}. \quad (5.50)$$

Tämä voidaan kirjoittaa abstraktimpaan muotoon

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{f},$$

missä $\mathbf{x} = (x_1, \dots, x_n)^T$ ja $\mathbf{f} = (f_1, \dots, f_n)^T$ ovat pystyvektoreita, joiden jokainen koordinaatti on t :n funktio ja A on $n \times n$ vakiomatriisi.

Ratkaisujen löytämiseksi turvaudutaan analogiaan: Jos $x' = ax + f$ on ensimmäisen kertaluvun lineaarinen differentiaaliyhtälö, jossa lisäksi a on vakio, voidaan aluksi ratkaista homogeeninen yhtälö $x' = ax$ ja tämän jälkeen voidaan löytää ratkaisu alkuperäiselle yhtälölle vakion varioinnilla. Homogeenisen yhtälön $x' = ax$ ratkaisu taas on tunnettu: Yhtälö voidaan kirjoittaa muotoon

$$\frac{x'}{x} = a \Leftrightarrow \frac{d}{dt} \ln x = a \Leftrightarrow \ln x = \int a \, dt \Leftrightarrow x = e^{\int a \, dt}.$$

Kertoimen a ollessa vakio on $\int a \, dt = at + C_0$ ja ratkaisu on siis muotoa

$$x = e^{at+C_0} = e^{C_0} e^{at} = C e^{at}.$$

Tarkastellaan seuraavaksi olisiko mahdollista löytää homogeeniselle differentiaaliyhtälöryhmälle ratkaisu analogisesti. Tätä varten turvaudutaan sarjaan perustuvaan määritelmään

$$e^{tA} = \sum_{k=0}^{\infty} \frac{1}{k!} (tA)^k = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k,$$

josta muodollisesti derivoimalla saadaan

$$\frac{d}{dt} e^{tA} = \sum_{k=1}^{\infty} \frac{k}{k!} t^{k-1} A^k = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} A^{k-1} A = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k A = e^{tA} A.$$

Samoin voidaan johtaa yhtälö $\frac{d}{dt} e^{tA} = A e^{tA}$, mutta tekniset yksityiskohdat sivuutetaan tässä yhteydessä. Näin ollen, jos $\mathbf{x} = e^{tA} \mathbf{c}$, missä $\mathbf{c}$ on vakiovektori, on

$$\mathbf{x}' = A e^{tA} \mathbf{c} = A \mathbf{x}.$$

Itse asiassa voidaan osoittaa, että homogeenisella yhtälöryhmällä ei ole muunlaisia ratkaisuja, mutta tämä sivuutetaan tällä kurssilla. Yhteenvetona esitetään seuraava tulos:

Lause 29 Homogeenisen differentiaaliyhtälöryhmän $\mathbf{x}' = A\mathbf{x}$ ratkaisu on $\mathbf{x} = e^{tA} \mathbf{c}$, missä $\mathbf{c}$ on vakiovektori.

Huomautus 37 Matriisin e^{tA} määrittäminen käytännössä ei kuitenkaan ole helpointa suoraan sarjaesitykseen nojautuen, vaan luvussa 2.8 esitetyllä tavalla.

Esimerkki 130 Etsitään uudelleen esimerkin 110 DY-parin

$$\begin{cases} y' + 4y + 4z = 0 \\ z' + 2y + 6z = 0 \end{cases}$$

ratkaisut alkuehdoilla $(y(0), z(0)) = (3, 15)$. DY-pari voidaan kirjoittaa muotoon

$$\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} -4 & -4 \\ -2 & -6 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}.$$

Käyttämällä kurssin alkuosan menetelmiä voidaan matriisin

$$A = \begin{pmatrix} -4 & -4 \\ -2 & -6 \end{pmatrix}$$

ominaisarvoiksi löytää -8 ja -2 sekä näitä vastaaviksi ominaisvektoreiksi $(1, 1)$ ja $(-2, 1)$. Näin ollen matriisi A on diagonalisoituva ja diagonalisaation välittäväksi matriisiksi voidaan valita

$$P = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}.$$

Tällöin

$$A = P \begin{pmatrix} -8 & 0 \\ 0 & -2 \end{pmatrix} P^{-1}$$

ja

$$e^{tA} = P \begin{pmatrix} e^{-8t} & 0 \\ 0 & e^{-2t} \end{pmatrix} P^{-1}$$

ja yleinen ratkaisu on muotoa

$$\begin{pmatrix} y \\ z \end{pmatrix} = e^{tA} \mathbf{c} = P \begin{pmatrix} e^{-8t} & 0 \\ 0 & e^{-2t} \end{pmatrix} P^{-1} \mathbf{c}.$$

Huomioimalla alkuehto $\mathbf{c} = (3, 15)^T$ voidaan ylläoleva yhtälö kirjoittaa muotoon

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{-8t} & 0 \\ 0 & e^{-2t} \end{pmatrix} \begin{pmatrix} 11 \\ 4 \end{pmatrix} = 11 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-8t} + 4 \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{-2t}.$$

Esimerkki 131 Etsitään DY-ryhmän

$$\begin{cases} x' = 2x - 2y - z \\ y' = -2x + 2y + z \\ z' = -x + y + 5z \end{cases}$$

ratkaisut (Harjoitustehtävä).

Muotoa

$$\mathbf{x}' = A\mathbf{x} + \mathbf{f} \quad (5.51)$$

olevan lineaarisen epähomogeenisen differentiaaliyhtälöryhmän ratkaisemiseksi voidaan soveltaa aiemmin esiintynyttä tekniikkaa: Homogeenisen yhtälöryhmän $\mathbf{x}' = A\mathbf{x}$ ratkaisu on muotoa $\mathbf{x} = e^{tA}\mathbf{c}$, ja epähomogeenisen yhtälöryhmän ratkaisua voidaan etsiä *vakion varioinnilla*, ts. yritelmällä, jossa vakiovektori $\mathbf{c}$ korvataan funktiolla $\mathbf{c}(t)$, siis $\mathbf{x} = e^{tA}\mathbf{c}(t)$.

Tällöin

$$\mathbf{x}' = e^{tA}A\mathbf{c}(t) + e^{tA}\mathbf{c}'(t) = A\mathbf{x} + e^{tA}\mathbf{c}'(t)^1$$

ja kun tämä sijoitetaan yhtälöön (5.51) saadaan

$$A\mathbf{x} + e^{tA}\mathbf{c}'(t) = A\mathbf{x} + \mathbf{f},$$

josta

$$\mathbf{c}'(t) = e^{-tA}\mathbf{f}, \quad (5.52)$$

mistä $\mathbf{c}(t)$ saadaan periaatteessa integroimalla: $\mathbf{c}(t) = \int e^{-tA}\mathbf{f} dt + \mathbf{c}$. Mikäli tämä osataan laskea, saadaan

$$\mathbf{x}(t) = e^{tA}\mathbf{c}(t) = e^{tA} \int e^{-tA}\mathbf{f} dt + e^{tA}\mathbf{c}.$$

Esimerkki 132 Ratkaistaan DY-pari

$$\begin{cases} x' = 3x - 2y + t \\ y' = 2x - 2y + 3e^t. \end{cases}$$

Edellä kuvatulla menetelmällä.

DY-pari voidaan esittää muodossa

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 3 & -2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t \\ 3e^t \end{pmatrix},$$

¹ Matriisien summan ja tulon derivoinnille on voimassa samankaltaiset säännöt kuin funktioillekin, mutta näitä ei tässä yhteydessä selvitetä yksityiskohtaisesti.

josta homogenisoimalla saadaan muoto

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \underbrace{\begin{pmatrix} 3 & -2 \\ 2 & -2 \end{pmatrix}}_A \begin{pmatrix} x \\ y \end{pmatrix}.$$

Homogeenisen DY-parin ratkaisut saadaan aiemmin esitetyllä tavalla: Etsimällä diagonaalisitys matriisille A saadaan

$$e^{tA} = P \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{-t} \end{pmatrix} P^{-1},$$

missä

$$P = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Nyt siis $\mathbf{x}(t) = e^{tA}\mathbf{c}(t)$, missä $\mathbf{c}(t)$ saadaan yhtälöstä (5.52), joka tässä tapauksessa saa muodon

$$\begin{aligned} \mathbf{c}'(t) &= \left(P \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{-t} \end{pmatrix} P^{-1} \right)^{-1} \begin{pmatrix} t \\ 3e^t \end{pmatrix} = P \begin{pmatrix} e^{-2t} & 0 \\ 0 & e^t \end{pmatrix} P^{-1} \begin{pmatrix} t \\ 3e^t \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} -6e^{-t} + 6e^{2t} + 4te^{-2t} - te^t \\ -3e^{-t} + 12e^{2t} + 2te^{-2t} - 2e^t \end{pmatrix} \end{aligned}$$

Tästä voidaan $\mathbf{c}(t)$ määrittää integroimalla:

$$\mathbf{c}(t) = \frac{1}{3} \begin{pmatrix} 6e^{-6} + 3e^{2t} + e^{-2t}(-1 - 2t) + e^t(1 - t) \\ 3e^{-t} + 6e^{2t} + e^t(2 - 2t) + e^{-2t}(-\frac{1}{2} - t) \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Alkuperäisen ryhmän ratkaisu saadaan tästä sijoittamalla

$$\mathbf{x}(t) = e^{tA}\mathbf{c}(t).$$

5.8 Autonomisista DY-pareista

Sanotaan, että differentiaaliyhtälöpari

$$\begin{cases} x'(t) = f_1(x, y) \\ y'(t) = f_2(x, y) \end{cases}$$

on *autonominen*, jos f_1 :n ja f_2 :n lausekkeessa ei esiinny muuttujaa t . Autonomisesta differentiaaliyhtälöparista voidaan eliminoida muuttuja t soveltamalla derivointia parametrin suhteen:

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{y'(t)}{x'(t)}$$

(katso Insinöörimatematiikka B). Tällöin saadaan differentiaaliyhtälö

$$\frac{dy}{dx} = \frac{f_2(x, y)}{f_1(x, y)}. \quad (5.53)$$

On kuitenkin huomattava, että vaikka (5.53) voitaisiin ratkaista, saadaan ainoastaan sen käyrän yhtälö, jota pitkin $(x(t), y(t))$ kulkee, kun t kasvaa. Itse liikeilmiö jää muilla keinoilla selvitetäväksi (millä nopeudella piste (x, y) kulkee, hidastuuko kulku, onko olemassa rajapistettä, jne.). Käyrää $(x(t), y(t))$ sanotaan *ratakäyräksi*.

Esimerkki 133 Differentiaaliyhtälöpari

$$\begin{cases} x'(t) = x - y \\ y'(t) = x + y \end{cases}$$

on autonominen ja sen ratakäyrien differentiaaliyhtälö on $y' = \frac{x+y}{x-y} = \frac{1+\frac{y}{x}}{1-\frac{y}{x}}$. Tähän voidaan soveltaa sijoitusta $z = \frac{y}{x}$ (harjoitustehtävä).

Esimerkki 134 Lotka-Volterra differentiaaliyhtälöpari kuvaa kahden lajin (peto ja saalis) yhteiseloa. Yksin eläessään x (saaliseläinten kanta) kasvaisi, kun taas y (petoeläinten kanta) vähenisi. Yhdessä eläessä x kärsii y :stä ja y hyötyy x :stä. Tämänkaltaista dynamiikkaa yksinkertaistetussa muodossa kuvaa differentiaaliyhtälöpari

$$\begin{cases} x'(t) = \alpha x - \beta xy \\ y'(t) = -\gamma y + \delta xy, \end{cases} \quad (5.54)$$

missä vakiot α , β , γ ja δ ovat positiivisia ja termi xy kuvaa lajien kohtaamislaajuutta.

Differentiaaliyhtälöpari (5.54) on autonominen ja ratakäyrien yhtälö on muotoa

$$y' = \frac{y}{x} \frac{\delta x - \gamma}{\alpha - \beta y},$$

minkä ratkaisu

$$\alpha \ln y - \beta y = \delta x - \gamma \ln x + C \quad (5.55)$$

voidaan saada separoimalla yhtälö. Käyrät (5.55) ovat pisteen $(\frac{\gamma}{\delta}, \frac{\alpha}{\beta})$ ympäri kiertäviä suljettuja käyriä, ja parin (5.54) ratkaisut jaksollisia: On olemassa sellainen luku T , että $(x(T), y(T)) = (x(0), y(0))$.

Jos parin (5.54) ensimmäinen yhtälö kirjoitetaan muotoon

$$\frac{x'(t)}{x(t)} = \alpha - \beta y,$$

saadaan integroimalla

$$0 = \ln \frac{x(T)}{x(0)} = \int_0^T \frac{x'}{x} dt = \int_0^T (\alpha - \beta y) dt = \alpha T - \beta \int_0^T y(t) dt,$$

mistä nähdään, että

$$\frac{1}{T} \int_0^T y(t) dt = \frac{\alpha}{\beta}. \quad (5.56)$$

Yhtälön (5.56) vasemman puolen lauseketta kutsutaan funktion $y(t)$ *keskiarvoksi välillä* $[0, T]$ ja merkitään $\bar{y}$.

Samalla tavalla funktion $x(t)$ keskiarvoksi välillä $[0, T]$ saadaan $\bar{x} = \frac{\gamma}{\delta}$.

Jos ulkopuolinen toimija hävittää *kumpaakin eläinkantaa* tehokkuudella e , tulee yhtälöparin (5.54) oikealle puolelle lisätä termit $-ex$ ja $-ey$, siis α ja γ korvautuvat luvuilla $\alpha - e$ ja $\gamma + e$. Uudet keskiarvot tulevat tällöin olemaan

$$\bar{x} = \frac{\gamma + e}{\delta}, \quad \text{ja} \quad \bar{y} = \frac{\alpha - e}{\beta}.$$

Täten siis kohtuullinen ($e < \alpha$) ulkopuolinen toiminta itse asiassa lisää saaliseläinten määrän keskiarvoa.

5.9 Differentiaaliyhtälöiden numeerisesta ratkaisemisesta

Differentiaaliyhtälölle

$$y' = f(t, y)$$

alkuarvolla $y(0) = y_0$ ei tunneta yleistä ratkaisukaavaa, vaikka useita erikoistapauksia voidaankin ratkaista. Lisäksi ratkaisun *olemassaolo* ja yksikäsitteisyys voidaan taata melko vaatimattomin oletuksin. Jos yhden differentiaaliyhtälön sijasta on tarkasteltava differentiaaliyhtälöryhmiä

$$\begin{cases} y'_1 = f_1(t, y_1, y_2, \dots, y_n) \\ y'_2 = f_2(t, y_1, y_2, \dots, y_n) \\ \dots \quad \dots \\ y'_n = f_n(t, y_1, y_2, \dots, y_n) \end{cases}, \quad (5.57)$$

on eksplisiittisiä ratkaisumenetelmiä entistä vähemmän. Tällöin voidaan turvautua numeerisiin menetelmiin, joissa määritetään funktioiden likimääräiset arvot ajanhetkillä $0, h, 2h, 3h, 4h, \dots$. Differentiaaliyhtälöryhmän muodosta ja alkuarvoista riippuu kuinka hyvin saadut likiarvot kuvaavat todellisia ratkaisuja. Näitä kysymyksiä tarkastellaan kaaosteorian piirissä.

Huomautus 38 Kaikki differentiaaliyhtälöryhmät voidaan saattaa muotoon (5.57) muotoon ottamalla käyttöön uusia funktioita. Esimerkiksi korkeamman kertaluvun derivaatat $y^{(n)}$ voidaan palauttaa matalampiin uuden funktion $z = y'$ avulla, jolloin $z^{(n-1)} = y^{(n)}$, mutta yhtälöiden määrää pitää kasvattaa: uudeksi yhtälöksi tulee ottaa funktion z määrittelevä $y' = z$.

Yksinkertaisin numeerinen menetelmä, ns. *Eulerin menetelmä* perustuu siihen, että derivaatta edustaa funktion lineaarista approksimaatiota. Tällöin funktion y arvoja $y_0, y_1, y_2, y_3, \dots$ aikavälein h lasketaan seuraavasti:

$$y_{i+1} = y_i + hy'(t_i) = y_i + h \cdot f(t_i, y_i).$$

Menetelmä yleistyy suoraviivaisesti differentiaaliyhtälöryhmille, mutta sitä voidaan parantaa oleellisesti käyttämällä lineaarisen approksimaation sijaan Taylorin polynomeja. Näin saatavia menetelmiä kutsutaan Runge–Kutta -menetelmiksi. Nämä perustuvat korkeamman kertaluvun approksimaatioihin, joissa derivaattojen arvoja approksimoidaan laskemalla funktioiden differenssejä.

Näiden johtaminen on kuitenkin teknisesti mutkikasta ja sivuutetaan tässä yhteydessä. Menetelmien esittäminen sen sijaan niiden esittäminen on melko suoraviivaista: Hyväksi havaitussa *Runge-Kutta 4:ssä* funktion y arvoja $y_0, y_1, y_2, y_3, \dots$ lasketaan seuraavasti:

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

missä

$$\begin{cases} k_1 = f(t_i, y_i)h \\ k_2 = f(t_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1)h \\ k_3 = f(t_i + \frac{1}{2}h, y_i + \frac{1}{2}k_2)h \\ k_4 = f(t_i + h, y_i + k_3)h \end{cases}.$$

Runge-Kutta 4 toimii myös differentiaaliyhtälöryhmillä. Se on ylläolevan esityksen perusteella helppo ohjelmoida, mutta valmiita ohjelmia Runge-Kutta 4:n käyttämiseksi on runsaasti saatavilla.

Esimerkki 135 Tarkastellaan differentiaaliyhtälöparia

$$\begin{cases} \frac{dx}{dt} = -2x^2 \\ \frac{dy}{dt} = -\frac{1}{2}xy \end{cases} \quad (5.58)$$

alkuehdoilla $x(0) = 2, y(0) = 1$.

Matlabiin ohjelmoitu `ode23` on muunnelma Runge-Kutta -menetelmästä. Tätä käyttäen DY-parin likimääräisratkaisu välillä $t \in [0, 10]$ voidaan löytää seuraavasti: `[t,xy]=ode23(@esimpari,[0 10],[2 1])` laskee vektorin `[t,xy]`, jossa `t` on lista ajanhetkistä välillä $[0, 10]$ ja `xy` lista funktioparin (x, y) likiarvoista. `esimpari` on pystyvektori, joka alkioina ovat DY-parin (5.58) oikean puolen funktiot. Vektori `[2 1]` määrittelee alkuarvot $x(0) = 2$ ja $y(0) = 1$.

Seuraavat rivit voidaan kirjoittaa Matlabiin *M*-tiedostoksi:

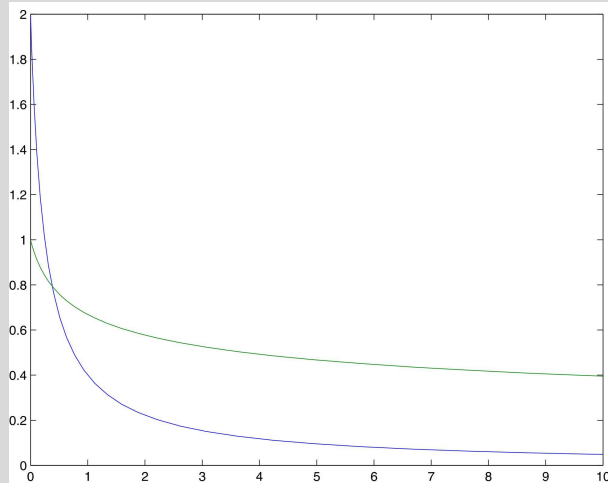
```
function esimerkki
[t,xy]=ode23(@esimpari,[0 10],[2 1])
plot(t,xy)
return

%Yllä olevat rivit käyttävät ode23-ohjelmaa DY-parin
%likimääräiseen ratkaisemiseen ja kuvaajan piirtämiseen

function arvo=esimpari(t,xy)
x=xy(1);
y=xy(2);
```

```
arvo(1,1)=-2*x*x;
arvo(2,1)=-0.5*x*y;
return
```

%Yllä oleva määrittelee funktion esimpari. Funktion arvo
 %on 2-pituinen pystyvektori, joka Matlabissa voidaan
 %esittää 2*1-matriisina: arvo(1,1) on 1. alkio ja
 %arvo(2,1) pystyvektorin 2. alkio.



Kuviossa funktion x kuvaaja alkaa arvosta 2 ja y :n arvosta 1. Tarkempi analyysi osoittaa, että molempien funktioiden kuvaajat lähestyvät nollaa t :n kasvaessa.

Huomautus 39 Edellisen esimerkin differentiaaliyhtälöpari voidaan ratkaista myös eksplisiittisesti. Ensinnäkin, autonomiselle parille saadaan

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{-\frac{1}{2}xy}{-2x^2} = \frac{1}{4} \frac{y}{x},$$

mikä on separoituva yhtälö:

$$4 \int \frac{1}{y} dy = \int \frac{1}{x} dx + C,$$

mikä voidaan kirjoittaa muotoon

$$4 \ln y = \ln x + C,$$

ja edelleen $y^4 = e^C x$. Vakio e^C saadaan alkuehdoista sijoittamalla $t = 0$: $y(0) = 1$ ja $x(0) = 2$. Tällöin $1^4 = e^C 2$, josta $e^C = \frac{1}{2}$. Tällöin siis $y = \sqrt[4]{\frac{x}{2}}$, mutta miten pari (x, y) käyttäytyy t :n funktiona, jää selvittämättä tällä tavalla.

Toisaalta taas parin ensimmäinen yhtälö on separoituva:

$$-\frac{1}{2} \frac{1}{x^2} dx = dt,$$

mistä integroimalla saadaan

$$\frac{1}{2} \frac{1}{x} = t + C_2,$$

ja jälleen alkuehto kertoo, että $\frac{1}{2} \cdot \frac{1}{2} = 0 + C_2$. Näin ollen $x = \frac{2}{4t+1}$ ja $y = \sqrt[4]{\frac{1}{4t+1}}$.