

Mika Hirvensalo

Insinöörimatematiikka D 2013

Osa 1: Lineaarialgebra

Sisältö

1	Lineaarialgebran peruskäsitteitä	5
1.1	Lineaariset yhtälöryhmät	5
1.1.1	Gaussin-Jordanin menetelmä	5
1.1.2	Ratkaisujoukon systemaattinen esittäminen	8
1.2	Vektoriavaruudet	10
1.3	Vektoriavaruuksien rakenteesta	11
1.4	Matriisit	17
1.4.1	Kertausta	17
1.4.2	Lineaarikuvaukset	19
1.4.3	Käänteismatriisi	24
1.4.4	Determinantti	26
1.5	Matriisin ominaisarvot ja ominaisvektorit	28
2	Vektoriavaruuksien geometriaa	33
2.1	Pistetulo, normi	33
2.2	Sovellus: Korrelaatiokerroin	35
3	Kolmiulotteisen avaruuden geometriaa	39
3.1	Ristitulo	39
3.2	Tasot	41
3.3	Suorat	44
3.4	Etäisyyksien laskemisesta	47
4	Differentiaaliyhtälöt (DY:t)	49
4.1	Yksinkertaiset DY:t	49
4.2	Lineaariset DY:t	50
4.2.1	Vakiokertoimiset lineaariset DY:t	51
4.2.2	Vakiokertoimiset homogeeniset DY:t	53
4.2.3	Ensimmäisen kertaluvun lineaarinen DY	54
4.2.4	Toisen kertaluvun lineaariset DY:t	56
4.3	Separoituvat DY:t	57
4.4	Eksaktit DY:t	59
4.5	Sekalaisia menetelmiä	61
4.5.1	Sijoitukset $\frac{y}{x}$ ja $ax + by$	61
4.5.2	Bernoullin DY	62
4.5.3	Käänteisfunktioon siirtyminen	62
4.5.4	Muuttujan vaihtaminen	62
4.5.5	Yhtälö $y'' = f(y)$	63
4.6	Autonomisista DY-pareista	64
4.7	DY:iden numeerisesta ratkaisemisesta	65

5	Usean muuttujan funktioiden differentiaalilaskentaa	69
5.1	Jatkuvuus	70
5.2	Osittaisderivaatat	74
5.3	Differentioituvuus	75
5.4	Suunnattu derivaatta	80
5.5	Ketjusaanto	81
5.6	Usean muuttujan väliarvolause	84
5.7	Ääriarvot, Lagrangen menetelmä	85
5.8	Funktion ääriarvojen etsiminen yleisesti	88
6	Usean muuttujan funktioiden integraalilaskentaa	91
6.1	Iteroidut integraalit	94
6.2	Sijoitus integraaliin	96
7	Vektorianalyysiä	101
7.1	Vektoriavaruuden funktion derivointi	101
7.2	Vektorifunktion käyräintegraali	101
7.3	Skalaarifunktion käyräintegraali	105
7.4	Pinta-ala	105
7.5	Skalaarifunktion pintaintegraali	108
7.6	Vektorikentän pintaintegraali	109
7.7	Stokesin lause	110

Johdanto

Tässä opintojaksossa perehdytään lineaarialgebraan, jonka peruskäsitteisiin vektoreihin ja matriiseihin on hyvin pintapuolisesti tutustuttu kursilla 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.

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. Insinöörimatematiikka B:ssä esitetty derivaatan määritelmä funktion lineaarisena approksimaationa yleistyy usean muuttujan funktioille, ja integraalin käsite voidaan luontevasti määritellä reaalisuoran osien karteesisilla tuloilla. Näihin yleistyksiin viitataan lyhyesti kurssin viimeisessä osassa.

Luku 1

Lineaarialgebran peruskäsitteitä

Lineaarialgebra käsittelee vektoreita ja näiden muodostamia algebrallisia rakenteita, *vektoriavaruuksia*. Tärkeä tarkastelykohde on myös sellaiset 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 nykyaikaisen lineaarialgebran käsitteisiin ovat peräisin 1600-luvun lopulta, mutta varsinaisen lineaarialgebran käsitteistöä on kehitetty 1800-luvun puolivälistä alkaen. Lineaarialgebra on täten differentiaali- ja integraalilaskentaa jonkin verran nuorempi matematiikan ala. Erityisen käyttökelpoista 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.

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 nolasta 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 edellämäinitut 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 kertomonen nollasta poikkeavalla vakiolla, 2) rivin lisääminen toiseen vakiolla kerrottuna ja 3) rivien paikan vaihtaminen. Jos matriisi B saadaan A :sta alkeisoperaatioita soveltamalla, merkitään $A \sim B$.

On helppo huomata, että yhtälöryhmän alkeisoperaatiot eivät muuta ratkaisujoukkoa miksiäkään. Siksi matriiseja A ja B vastaavilla yhtälöryhmillä 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

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

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 $\frac{1}{2}$ saa-

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

ja kolmanteen 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 $\begin{pmatrix} 1 & 3 & 2 & 7 \\ 0 & 1 & 1 & \frac{9}{5} \\ 0 & 0 & 2 \end{pmatrix}$.

Viimeisintä matriisia vastaavan yhtälöryhmän kolmas yhtälö on muotoa $0 \cdot z + 0 \cdot y + 0 \cdot z = 2$, 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 vektori-joukkoon 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 52 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 kirjasintyypillä: $\mathbf{a} = (a_1, a_2, \dots, a_n)$ ja alkioita $a_1, a_2, \dots, a_n$ kutsutaan vektorin *koordinaateiksi* tai pelkästään alkioiksi, toisinaan myös komponenteiksi. Fysikaalisten tieteiden kirjallisuudessa vektoreista käytetään usein myös yläviivaa $\vec{a}$, ”hattua” $\hat{a}$ tai ylänuolta $\overrightarrow{a}$. Vanhemmassa suomenkielisessä matematiikan 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 reaal- tai kompleksilukujen perusominaisuuksia. Suoraviivaisuudesta 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 $C[a, b]$ 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 olleen vektoriavaruuden $\mathbb{R}^n$ alkio on äärellisellä joukolla $\{1, 2, \dots, n\}$ määritelty funktio joukkoon $\mathbb{R}$, kun taas avaruuden $C[a, b]$ 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 $C[a, b]$ 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 pisteenä. Tätä tulkintaa on yleensä tapana laajentaa suuntajanaksi 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 *lineaarikombinaatioiden* 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 $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, \dots, 0)$, ..., $\mathbf{e}_n = (0, 0, \dots, 1)$. 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 $\mathbf{0} \in V_1$ 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 generoitavien 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 vektori joukko $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 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. Vektori joukko $\{\mathbf{i}, \mathbf{j}, \mathbf{i} + \mathbf{j}\}$ 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. Vektori joukko $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ 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 vektori joukko $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, 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 nollia. 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 $\mathbb{R}^2$ vektori joukko $\{\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}\} = \{(1, 1), (1, -1)\}$ 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 esimerkiksi 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.

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.

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:

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

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 14. 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 saada 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 $V = \{(1, 2, -1), (-2, 3, 1), (4, 1, -3)\}$ 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.

1.4 Matriisit

1.4.1 Kertausta

Kerrataan ensin insinöörimatematiikka A:ssa esitetyt matriiseja koskevat määritelmät ja selvitetään sen jälkeen varsinainen motivaatio matriisikäsitteelle.

Määritelmä 15. Matriisi A on kaavio

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}. \quad (1.3)$$

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 16. (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 17. 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 18. Olkoon A $m \times n$ -matriisi. Tällöin määritellään

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

Määritelmä 19 (Skalaarikertolasku). Jos A on $m \times n$ -matriisi ja c joko kompleksi- tai reaaliluku, on cA $m \times n$ -matriisi, jolle pätee $(cA)_{ij} = cA_{ij}$ (kertolasku alkioittain).

Määritelmä 20 (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
```

1.4.2 Lineaarikuvaukset

Matriisin käsite liittyy läheisesti lineaarikuvauksiin. Periaatteessa voidaan sanoa, että matriisi on lineaarikuvauksen esitystapa. Yhdessä lineaarikuvauksen ja matriisin käsitteet muodostavat lineaarialgebran perustan, jolle koko koneisto rakentuu.

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

Määritelmä 21. 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

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

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 lineaarikuvaus, 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 johtaa 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 lineaarikuvaus määräytyy 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 $\mathbf{y}$:nnestä 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:

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n. \quad (1.4)$$

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 käsiksi syvällisemmin.

Määritelmä 22. 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 23. 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 (1.4) 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 vaihtoehtoisen muodon lineaarikuvauksen esittämiseksi matriisien avulla.

Esimerkki 48. Insinöörimatematiikka C:ssä esitetty diskreetti Fourier-muunnos on $\mathbb{C}^N$:ssä tapahtuva operaatio, jossa vektori $(f_0, f_1, \dots, f_N)$ 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. 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 $\mathbf{b} = (b_1, \dots, b_m)$.

1.4.3 Käänteismatriisi

Määritelmä 24. 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 17), 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 50. 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 51. Lasketaan matriisitulo

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

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 on perusteltu 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 52. 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 53. 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.

1.4.4 Determinantti

Determinantin käsite on erittäin käyttökelpoinen lineaarialgebrassa, ja siksi on ymmärrettävää, että lähtökohtia determinantin määrittelemiseksi on useita. Tämän kurssin puitteissa ei kuitenkaan voida esitellä determinantin määritelmää tai kytkentöjä muihin käsitteisiin yksityiskohtaisesti. Siksi tyydytään lähinnä esittelemään menetelmä, jolla determinantti lasketaan ja käymään läpi joitakin determinantin ominaisuuksia ilman todistuksia.

Esimerkki 54. 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.

Injektiivisuus 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 55. 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 suunnikkaa, jonka sivuina ovat $(1, 0)$ ja $(0, 1)$ (ja pinta-ala 1) suunnikkaaksi, jonka pinta-ala on $|ad - bc|$.

Määritelmä 25. 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* kokonaisluvuille 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 56. 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 determinanteihin. 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. Yllämainitun 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 maksimimiseksi riveillä.

$$\text{Esimerkki 57. } \begin{vmatrix} 1 & 2 & 0 \\ -3 & 0 & 1 \\ 4 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ -3 & 6 & 1 \\ 4 & -10 & 1 \end{vmatrix} = \begin{vmatrix} 6 & 1 \\ -10 & 1 \end{vmatrix} = 6 \cdot 1 - 1 \cdot (-10) = 16.$$

Esimerkki 58.

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

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

1.5 Matriisin ominaisarvot ja ominaisvektorit

Olkoot A ja B $n \times n$ -neliomatriiseja. 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 59. 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.

Esimerkki 60. Lineaarisessa 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 26. 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 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 61. 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

$$\begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{pmatrix} -13-2 & 30 \\ -9 & 20-2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ratkaisujen etsimistä. Gaussin-Jordanin menetelmä muuntaa tämän muotoon

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

jonka 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 62. 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 63. 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.

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

Määritelmä 27. Neliömatriisit A ja B ovat *similaarit*, mikäli on olemassa sellainen kääntyvä matriisi P , että $A = P^{-1}BP$.

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

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

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

Esimerkki 64. Olkoon $A = \begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix}$ kuten esimerkissä 61 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 $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ käänteismatriisin $\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$ määrittäminen onnistuu esimerkiksi aiemmassa pykälässä kuvatulla Gaussin-Jordanin menetelmään perustuvalla prosessilla. Täten

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

ja siksi

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

Luku 2

Vektoriavaruuksien geometriaa

Aiemmissa pykälissä vektoriavaruuksia $\mathbb{R}^n$ lähinnä algebralliselta kannalta: Vektoreilla 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.

2.1 Pistetulo, normi

Määritelmä 28. Vektoreiden $\mathbf{x} = (x_1, \dots, x_n)$ ja $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ *pistetulo* määritellään $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \dots + x_n y_n$. Pistetulosta käytetään myös nimitystä *sisätulo* ja merkintöjä $\langle \mathbf{x}, \mathbf{y} \rangle$, $\langle \mathbf{x}, \mathbf{y} \rangle$ ja $\langle \mathbf{x} | \mathbf{y} \rangle$.

Lause 15. *Pistetulo toteuttaa ominaisuudet*

1. $\mathbf{x} \cdot \mathbf{x} \geq 0$ ja $\mathbf{x} \cdot \mathbf{x} = 0$ tarkalleen silloin kun $\mathbf{x} = \mathbf{0}$ (epänegativisuus ja epädegeneratiivisuus).
2. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ (vaihdannaisuus eli kommutatiivisuus).
3. $(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = a\mathbf{x} \cdot \mathbf{z} + b\mathbf{y} \cdot \mathbf{z}$ (lineaarisuus).

Todistus. Harjoitustehtävä.

Lause 16. *Pistetulo toteuttaa Cauchyn-Schwarzin epäyhtälön*

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

Todistus. Harjoitustehtävä.

Määritelmä 29. Vektorin $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ *normi* määritellään

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

Vektorin normista käytetään myös merkintää $|\mathbf{x}|$.

Lause 17. Normi toteuttaa seuraavat ominaisuudet:

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

Todistus. Todistetaan kolmas väittämä ja jätetään muut harjoitustehtäväksi.

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

mistä väite seuraa suoraan. Toisen ja kolmannen rivin välinen epäyhtälö on saatu soveltamalla Cauchyn-Schwarzin epäyhtälöä.

Pythagoraan lauseen perusteella $\mathbb{R}^3$:ssa $\|\mathbf{x}\|$ merkitsee pisteen $\mathbf{x}$ etäisyyttä origosta, siis vektorin $\mathbf{x}$ visuaalisen esityksen pituutta. Normia $\mathbf{x}$ kutsutaankin toisinaan myös vektorin $\mathbf{x}$ *pituudeksi*. Tällöin pisteiden $\mathbf{x}$ ja $\mathbf{y}$ välinen etäisyys on $\|\mathbf{x} - \mathbf{y}\|$. Korkeampiulotteisille avaruuksille $\mathbb{R}^n$ pisteen $\mathbf{x}$ etäisyys origosta *määritellään* vektorin normiksi $\|\mathbf{x}\|$ ja pisteiden $\mathbf{x}$ ja $\mathbf{y}$ välinen etäisyys *määritellään* lausekkeen $\|\mathbf{x} - \mathbf{y}\|$ avulla.

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 $|x|$ ero onkin vain sopimuskysymys: Ensimmäistä on sovittu käytettävän vektoreista, jälkimmäistä vain reaali- tai kompleksiluvuista.

Määritelmä 30. 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 17. Edellisen lauseen perusteella kahden vektorin $\mathbf{x}$, $\mathbf{y}$ (kumpikaan ei nollavektori) välinen kulma θ saadaan esityksestä

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

Cauchyn-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 65. 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 31. Vektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat *ortogonaalit* eli *kohtisuorat* jos $\mathbf{x} \cdot \mathbf{y} = 0$.

Edellisen määritelmän mukaan nollavektori on kohtisuorassa kaikkia muita vektoreita kohtaan: $\mathbf{0} \cdot \mathbf{x} = 0$ 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 32. 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

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

2.2 Sovellus: 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 66. 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 $\mathbf{y} = c\mathbf{x}$, ovan 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 μ_x :

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

ja samoin vektorille $\mathbf{y}$:

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

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

$$\mu_y = c\mu_x + b$$

ja

$$y'_i = y_i - c\mu_x - b = cx_i + b - c\mu_x - b = c(x_i - \mu_x) = cx'_i,$$

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

Määritelmä 33. 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 67. Taulukossa on eräälle 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 $\mu_x = \frac{1}{36}(40 + 80 + 33 + \dots + 32) = \frac{1768}{36}$ ja vastaavasti $\mu_y = \frac{1}{36}(25 + 29 + 16 + \dots + 15) = \frac{645}{36}$.

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

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

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 18. 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 3

Kolmiulotteisen avaruuden geometriaa

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

3.1 Ristitulo

Kuten aiemmin huomattiin, vektoreiden pistetulo (eli 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 34. 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

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

Kun palautetaan mieleen *determinantin* käsite (Insinöörimatematiikka I), saadaan ristitulon määritelmälle muistisääntö. Kaksirivinen determinanttihan määriteltiin suoraan

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

ja todettiin, että kolmirivinen determinantti voidaan laskea kehittämää

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

käyttämällä. Yllä oleva kehitemä on muodostettu yleistä determinantin kehittelysääntöä käyttämällä ja sääntö toimii seuraavasti: Muodostetaan summalauseke, jossa kukin ensimmäisen rivin alkio on kertojana, kertojan merkki kuitenkin varustettuna ”vuorottelusäännöllä”: a :lle merkki $+1$, b :lle merkki -1 , c :lle merkki $+1$. 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.

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

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \quad (3.1)$$

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 $\mathbf{i}$, $\mathbf{j}$ ja $\mathbf{k}$, eikä tällaista determinanttia ole määritelty.

Tästä huolimatta (3.1) voi toimia muistisääntönä:

$$\begin{aligned}\mathbf{x} \times \mathbf{y} &= \mathbf{i} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - \mathbf{j} \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \mathbf{k} \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \\ &= \mathbf{i}(x_2y_3 - x_3y_2) - \mathbf{j}(x_1y_3 - x_3y_1) + \mathbf{k}(x_1y_2 - x_2y_1) \\ &= (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).\end{aligned}$$

Lause 19 (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

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

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 $\begin{vmatrix} y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \\ x_1 & x_2 & x_3 \end{vmatrix}$).

Lause 20. 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.

3.2 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 35. 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 miltyään 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 68. 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

$$\mathbf{z} + L = \{\mathbf{z} + c\mathbf{x} \mid c \in \mathbb{R}\} \quad (3.2)$$

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 (3.2), 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 vektori pari $\{\mathbf{x}, \mathbf{y}\}$:

$$T = \{c_1\mathbf{x} + c_2\mathbf{y} \mid c_1, c_2 \in \mathbb{R}\}. \quad (3.3)$$

Esitysmuodolla (3.3) 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

$$T = \{\mathbf{r} + c_1\mathbf{x} + c_2\mathbf{y} \mid c_1, c_2 \in \mathbb{R}\}, \quad (3.4)$$

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 parametrikksi. 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 *paikkavektoriksi* ja vektoreita $\mathbf{x}$ ja $\mathbf{y}$ sen *suuntavektoreiksi*. Jos $\mathbf{x}$ ja $\mathbf{y}$ ovat lineaarisesti riippuvat, ei (3.4) määrittele tasoa, vaan suoran tai pelkästään yhden pisteen.

Huomautus 19. Parametriesityksen (3.4) 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

$$T = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} \cdot \mathbf{n} = 0\}, \quad (3.5)$$

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

$$T = \{\mathbf{x}_0 + \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^3, \mathbf{x} \cdot \mathbf{n} = 0\} = \{\mathbf{x} \in \mathbb{R}^3 \mid (\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{n} = 0\} \quad (3.6)$$

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

Tason koordinaattimuoto: Muotoa

$$T = \{(x, y, z) \mid ax + by + cz = d\}, \quad (3.7)$$

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 *kollineaariset*, muutoin *epäkollineaariset*.

Tarkastellaan seuraavaksi esimerkkien kautta miten eri esitysmuodoista päästään toiseen.

Esimerkki 69 (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

$$(-2, -5, 4) \cdot ((x, y, z) - (1, 0, 3)) = 0. \quad (3.8)$$

Esimerkki 70 (normaalimuodosta koordinaattimuotoon). Esituksen (3.8) muuntaminen koordinaattimuotoon on suoraviivaista, sillä koordinaattimuoto ja normaalimuoto ovat lähes samat. Yhtälössä (3.8) tarvitsee vain laskea sisätulot, jolloin saadaan

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

siis koordinaattimuoto on

$$-2x - 5y + 4z = 10. \quad (3.9)$$

Huomautus 20. Koordinaattimuodon kertoimista voidaan suoraan lukea tason normaalivektori.

Esimerkki 71 (koordinaattimuodosta kolmeen pisteeseen). Jos tasolle tunnetaan koordinaattimuoto (70), voidaan helposti valita tasolta kolme pistettä. Valitaan esimerkiksi ensin $x = y = 0$, jolloin yhtälön (70) 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 21. 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 nolllaksi.

Esimerkki 72 (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 paikkavektori, 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 nolllasta poikkeavalla luvulla.

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

Esimerkki 73. Selvitetään, onko esimerkissä (72) saatu parametrimuotoinen taso todella sama kuin esimerkin (69) taso. Tätä varten etsitään molemmille tasoille koordinaattimuoto. Esimerkin (69) koordinaattimuoto on jo selvitetty esimerkissä (3.9): koordinaattiyhtälö

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

saadaan yksikäsitteiseen muotoon jakamalla se luvulla 10:

$$-\frac{1}{5}x - \frac{1}{2}y + \frac{2}{5}z = 1. \quad (3.10)$$

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 pistetulot laskemalla saadaan

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

Jakamalla luvulla -25 saadaan

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

sama kuin (3.10).

Esimerkki 74. Tarkastellaan, millä ehdoilla joukko

$$\{(ax + by + c, dx + ey + f, gx + hy + i) \mid x, y \in \mathbb{R}\} \quad (3.11)$$

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 (3.11) 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 (3.11) suoran (mikäli $(a, d, g) \neq (0, 0, 0)$).

Määritelmä 36. Tasojen L_1 ja L_2 välisellä kulmalla tarkoitetaan tasojen normaalivektoreiden välistä kulmaa.

Esimerkki 75. 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 $\theta \approx 1,65 = 85,5^\circ$.

3.3 Suorat

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

$$L = \{c\mathbf{x} \mid c \in \mathbb{R}\}, \quad (3.12)$$

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

$$L = \{\mathbf{r} + c\mathbf{x} \mid c \in \mathbb{R}\} \quad (3.13)$$

kutsutaan suoran *parametrimuodoksi* ja lukua c *parametriksi*. Jos $\mathbf{x} = \mathbf{0}$, ei esitys (3.13) määrittele suoraa, vaan yhden pisteen.

Huomautus 22. 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 (3.13) avaruuden $\mathbb{R}^3$ 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 parametrimuoto 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 nollasta 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 76 (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:

$$t = \frac{x-1}{2} = \frac{y-2}{-3} = z-2 \quad (3.14)$$

Huomautus 23. Koordinaattimuodosta voidaan helposti lukea suoran paikkavektori (osoittajien vakiot) ja suuntavektori (nimitäjien vakiot).

Esimerkki 77 (Koordinaattimuodosta kaksi pistettä). Yhtälöissä (3.14) 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 (3.14), on $(x, y, z) = (1, 2, 2)$ ja arvo $t = 1$ antaa $(x, y, z) = (3, -1, 3)$.

Huomautus 24. Parametrimuodosta kahden eri pisteen löytäminen on vielä yksinkertaisempaa: riittää valita parametrin arvoiksi esim 0 ja 1.

Esimerkki 78 (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 79 (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 80. 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 suoralta 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 81. 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

$$(0, -1, 1) + t_1(1, -7, -3) = (1, -1, 3) + t_2(1, -7, -3) \quad (3.15)$$

ratkaisuja $t_1, t_2 \in \mathbb{R}$. Yhtälö (3.15) 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 82. 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 $t_2 = \frac{7}{55}$.

3.4 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 tasoja 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 $S_1 : \mathbf{r}_1 + t\mathbf{s}_1$ ja $S_2 : \mathbf{r}_2 + t\mathbf{s}_2$ 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 $\mathbf{p}_2 = \mathbf{r}_2 + t\mathbf{s}_2$ 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 4

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 myös kemiallisissa, biologisissa ja taloustieteellisissä malleissa. Muodollisesti differentiaaliyhtälön käsite määritellään seuraavasti:

Määritelmä 37. 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 38. Differentiaaliyhtälön ratkaisun kuvaajaa kutsutaan *integraalikäyräksi*. Yleensä integraalikäyriä on ääretön määrä, mutta yksikäsitteiseen ratkaisuun voidaan silti yleensä päätyä *reunaehdoilla* $y(0) = c_0, y'(0) = c_1, \dots, y^{(n)}(0) = c_n$.

4.1 Yksinkertaiset DY:t

Yksinkertaiseksi differentiaaliyhtälöksi sanotaan sellaista, jonka ratkaiseminen on pelkkä integrointitehtävä.

Esimerkki 83. 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 IB): DY

$$N'(t) = -\lambda N(t)$$

voidaan kirjoittaa muotoon

$$\frac{d}{dt} \ln N = -\lambda,$$

joka on yksinkertainen funktion $\ln N$ suhteen.

4.2 Lineaariset DY:t

Määritelmä 39. Muotoa

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_2 y'' + a_1 y' + a_0 y = b(t) \quad (4.1)$$

olevaa differentiaaliyhtälöä kutsutaan *lineaariseksi*. Jos $a_n, a_{n-1}, \dots, a_1$ ja a_0 ovat vakioita, sanotaan yhtälöä *vakiokertoisiksi* lineaariseksi differentiaaliyhtälöksi. Jos $b(t)$ on nollafunktio, sanotaan lineaarista differentiaaliyhtälöä *homogeeniseksi*.

Huomautus 25. Suoralla laskulla huomataan, että homogeenisilla (ei välttämättä vakiokertoisilla) differentiaaliyhtälöillä on seuraava ominaisuus: jos y_1 ja y_2 ovat ratkaisuja, niin myös $c_1 y_1 + c_2 y_2$ on ratkaisu. Vertaa tätä ominaisuutta lineaarisen yhtälöryhmän ratkaisuihin (Insinöörimatematiikka IA).

Huomautus 26. On myös mahdollista todistaa, että (mikäli kerroinfunktiot a_i ja b ovat jatkuvia), on n :nen kertaluvun homogeenisella differentiaaliyhtälöllä aina n lineaarisesti riippumatonta ratkaisua $y_1, y_2, \dots, y_n$, joiden kombinaationa saadaan yleinen ratkaisu:

$$y = c_1 y_1 + c_2 y_2 + \dots + c_n y_n.$$

Edellä käytettiin kurssista Insinöörimatematiikka IA tuttua lineaarisen riippumattomuuden käsitettä, joka reaalfunktioille sovellettuna on seuraavanlainen:

Määritelmä 40. Funktiot $f_1, \dots, f_k$ ovat *lineaarisesti riippumattomat*, jos ehdosta

$$(\forall x)(c_1 f_1(x) + \dots + c_k f_k(x) = 0)$$

seuraa $c_1 = \dots = c_k = 0$.

Esimerkki 84. funktiot $f_0(x) = 1$ ja $f_1(x) = x$ ja $f_2(x) = x^2$ ovat lineaarisesti riippumattomat: Jos nimittäin

$$c_0 \cdot 1 + c_1 x + c_2 x^2 = 0$$

kaikilla x :n arvoilla, on välttämättä $c_0 = c_1 = c_2 = 0$. Yleisemmin voidaan todeta, että polynomi-joukko $1, x, x^2, x^3, \dots, x^k$ on lineaarisesti riippumaton.

Esimerkki 85. Jos luvut $\lambda_1, \dots, \lambda_k$ ovat erisuuret, ovat funktiot $e^{\lambda_1 x}, \dots, e^{\lambda_k x}$ lineaarisesti riippumattomat. Tämä voidaan näyttää toteen induktiolla luvun k suhteen.

Huomautus 27. On myös mahdollista näyttää toteen, että yleisen n :nnen kertaluvun lineaarisen DY:n kaikki ratkaisut voidaan esittää muodossa

$$y = c_1 y_1 + c_2 y_2 + \dots + c_n y_n + y_0,$$

missä $y_1, y_2, \dots, y_n$ ovat homogenisoidun yhtälön riippumattomat ratkaisut ja y_0 jokin yhtälön yksittäisratkaisu.

Vertaa yllämainittuja käsitteitä Insinöörimatematiikka IA:ssa esitettyihin lineaarisen yhtälöryhmän ratkaisuihin.

4.2.1 Vakiokertoimiset lineaariset DY:t

Laplace-muunnokset ovat erittäin käyttökelpoisia nimenomaan vakiokertoimisten lineaaristen differentiaaliyhtälöiden ratkaisemiseksi, sillä niiden avulla voidaan muuntaa muotoa (4.1) 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.

Esimerkki 86. Ratkaistaan kurssista Insinöörimatematiikka I tuttu differentiaaliyhtälö

$$y' = -\lambda y, \quad (4.2)$$

missä λ on vakio ja alkuehtona toimii $y(0) = y_0$. Jos y on originaalifunktio, merkitään sen Laplace-muunnosta $Y(s)$:llä. Oletetaan lisäksi, että y on oikealta jatkuva nollassa. Taulukon mukaan tällöin funktion $y'(t)$ Laplace-muunnos on $sY(s) - y_0$, joten differentiaaliyhtälö (53) muuttuu Laplace-muunnoksilla muotoon

$$sY(s) - y_0 = -\lambda Y(s). \quad (4.3)$$

Näin saatu yhtälö on Laplace-muunnoksen $Y(s)$ suhteen algebrallinen yhtälö, josta $Y(s)$ voidaan ratkaista:

$$Y(s) = y_0 \frac{1}{s + \lambda},$$

ja originaalifunktio $y(t)$ saadaan taulukosta: $y(t) = y_0 e^{-\lambda t}$.

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 87. Ratkaistaan toisen kertaluvun differentiaaliyhtälö

$$y'' + 4y = 0 \quad (4.4)$$

alkuehdoilla $y(0) = y_0$ ja $y'(0) = 0$. Nyt $\mathcal{L}[y''] = s^2 Y(s) - y(0)s - y'(0) = s^2 Y(s) - y_0 s$ ja (4.4) saadaan siis Laplace-muunnoksilla muotoon

$$s^2 Y(s) - y_0 s + 4Y(s) = 0,$$

mistä voidaan ratkaista

$$Y(s) = \frac{y_0 s}{s^2 + 4}.$$

Originaalifunktio saadaan taulukosta: $y(t) = y_0 \cos 2t$.

Toisinaan Laplace-muunnokset soveltuvat myös muuhunkin kuin differentiaaliyhtälöiden ratkaisemiseen.

Esimerkki 88. Ratkaistaan integraaliyhtälö

$$y(t) = t + \int_0^t \sin(t-u)y(u) du.$$

Konvoluution määritelmän perusteella yhtälö voidaan kirjoittaa muotoon

$$y(t) = t + \sin t * y(t),$$

mistä Laplace-muunnokset ottamalla saadaan

$$Y(s) = \mathcal{L}[t](s) + \mathcal{L}[\sin t](s)Y(s) = \frac{1}{s^2} + \frac{1}{s^2 + 1}Y(s).$$

Tästä voidaan ratkaista $Y(s)$:

$$Y(s) = \frac{1}{s^2} + \frac{1}{s^4},$$

ja esimerkin 132 perusteella $y(t) = t + \frac{1}{6}t^3$.

Esimerkki 89. 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 90. 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 91. 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 92. 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

$$\begin{cases} y(t) = -8e^{-2t} + 11e^{-8t} \\ x(t) = 4e^{-2t} + 11e^{-8t}. \end{cases}$$

4.2.2 Vakiokertoimiset homogeeniset DY:t

Vakiokertoimiset homogeeniset differentiaaliyhtälöt ovat erikoistapauksia vakiokertoimisista lineaarisista differentiaaliyhtälöistä. Ne voidaan periaatteessa ratkaista ilman Laplace-munnoksia. Todetaan aluksi, että yhtälöllä

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = 0 \quad (4.5)$$

on aina muotoa $y = e^{\lambda x}$ oleva ratkaisu. Tätä varten lasketaan derivaatat $y' = \lambda e^{\lambda x}, y'' = \lambda^2 e^{\lambda x}, \dots, y^{(n)} = \lambda^n e^{\lambda x}$ ja sijoitetaan nämä yhtälöön (4.5):

$$\lambda^n e^{\lambda x} + a_{n-1} \lambda^{n-1} e^{\lambda x} + \dots + a_0 e^{\lambda x} = 0,$$

mistä $e^{\lambda x}$:llä jakaminen antaa

$$\lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0 = 0. \quad (4.6)$$

Yhtälöllä (4.6), jota kutsutaan differentiaaliyhtälön (4.5) *karakteristiseksi yhtälöksi*, on algebran peruslauseeseen (kts. Insinöörimatematiikka IA) nojalla korkeintaan $k \leq n$ erisuurta ratkaisua $\lambda_1, \dots, \lambda_k$, joista jokaisesta saadaan alkuperäisen yhtälön (4.5) ratkaisu $e^{\lambda_i x}$. Jos juuria λ_i on tarkalleen n , ovat tässä kaikki alkuperäisen yhtälön *riippumattomat ratkaisut*, kun taas siinä tapauksessa että $k < n$, on differentiaaliyhtälöllä (4.5) myös muita riippumattomia ratkaisuja.

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. Tällöin voidaan kirjoittaa $e^{(\alpha+i\beta)x} = e^{\alpha x} e^{i\beta x}$, mikä edelleen Eulerin kaavan perusteella on muotoa

$$e^{\alpha x} (\cos \beta x + i \sin \beta x)$$

Tällöin siis kukin kompleksinen ratkaisu $e^{\lambda x}$ voidaan hajottaa kahdeksi reaalisiksi ratkaisuksi $e^{\alpha x} \cos \beta x$ ja $e^{\alpha x} \sin \beta x$.

Esitetään tämän kappaleen lopuksi miten toisen kertaluvun lineaarista yhtälöä

$$y'' + ay' + by = 0, \quad (4.7)$$

käsitellään. Yhtälöllä on edellämainitun perusteella 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) $D = \left(\frac{a}{2}\right)^2 - b > 0$. Tällöin yhtälöllä (4.7) 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) $D = \left(\frac{a}{2}\right)^2 - b < 0$. 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) $D = \left(\frac{a}{2}\right)^2 - b = 0$ 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 93. Ratkaistaan uudelleen esimerkin (4.4) differentiaaliyhtälö

$$y'' + 4y = 0. \quad (4.8)$$

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 (4.8) yleinen (reaalinen) ratkaisu on siis

$$y(x) = C_1 \cos 2x + C_2 \sin 2x,$$

mistä saadaan

$$y'(x) = -2C_1 \sin 2x + 2C_2 \cos 2x,$$

jolloin vakiot C_1 ja C_2 voidaan määrittää reunaehdoista $y(0) = y_0$, $y'(0) = 0$, siis $y_0 = C_1 \cos 2 \cdot 0 + C_2 \sin 2 \cdot 0 = C_1$ ja $0 = 2C_2$. Täten alkuehtojen määrittämä ratkaisu on

$$y(x) = y_0 \cos 2x.$$

4.2.3 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 94. 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 on siis muotoa

$$y' + a(x)y = b(x), \quad (4.9)$$

ja tämän ratkaisemiseksi tarkastellaan aluksi *homogenisoitua* yhtälöä

$$y' + a(x)y = 0, \quad (4.10)$$

joka voidaan kirjoittaa muotoon

$$\frac{y'}{y} = -a(x). \quad (4.11)$$

Koska $\frac{d}{dx} \ln y = \frac{y'}{y}$, saadaan (4.11) muotoon

$$\frac{d}{dx} \ln y = \frac{d}{dx} \left(- \int a(x) dx \right),$$

josta puolestaan

$$\ln y = - \int a(x) dx + C.$$

Valitsemalla $C = 0$ saadaan homogenisoidun yhtälön (4.10) ratkaisuksi $y_H = e^{-\int a(x) dx}$, ja tämä onkin vakiokerrointa lukuunottamatta yhtälön (4.10) *ainoa* ratkaisu: jos nimittäin y on jokin toinen ratkaisu, saadaan

$$\frac{d}{dx} \frac{y}{y_H} = \frac{y'y_H - yy'_H}{y_H^2} = \frac{-ayy_H + yay_H}{y_H^2} = 0,$$

jolloin siis $\frac{y}{y_H} = C$ ja $y = Cy_H$.

Esimerkki 95. Differentiaaliyhtälö $xy' + y = 0$ voidaan kirjoittaa homogeenisen differentiaaliyhtälön muotoon $y' + \frac{1}{x}y = 0$, jonka yleinen ratkaisu on

$$y(x) = Ce^{-\int \frac{1}{x} dx} = Ce^{-\ln x} = \frac{C}{x}.$$

Alkuperäisen (epähomogeenisen) differentiaaliyhtälön (4.9) ratkaisu saadaan menettelyllä, jota kutsutaan *vakion varioinniksi*. Vakion varioinnissa käytetään lähtökohtana homogenisoidun differentiaaliyhtälön (4.10) 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 josta C voidaan ratkaista integroimalla:

$$C(x) = \int \frac{b(x)}{y_H(x)} dx + B$$

Tällöin siis alkuperäisen differentiaaliyhtälön (4.9) 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 96. Ratkaistaan $y' + 2xy = x$. Tarkastellaan aluksi homogenisoitua yhtälöä $y' + 2xy = 0$, joka voidaan kirjoittaa muotoon

$$\frac{d}{dx} \ln y = \frac{y'}{y} = -2x, \quad (4.12)$$

josta $\ln y = -x^2$; tällöin siis $y_H = e^{-x^2}$ on eräs homogenisoidun yhtälön ratkaisu. Alkuperäisen yhtälön ratkaisujen saamiseksi suoritetaan vakion variointi:

$$y(x) = C(x)y_H(x) = C(x)e^{-x^2},$$

mistä nähdään, että $y' = C'e^{-x^2} + Ce^{-x^2}(-2x) = C'y_H - 2xy_H C$. Sijoittamalla tämä alkuperäiseen yhtälöön saadaan

$$C'y_H - 2xy_H C + 2xCy_H = x,$$

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

$$y' + a(x)y = b(x) \quad (4.13)$$

ratkaisemiseksi. Menetelmä jakautuu kahteen vaiheeseen, joista ensimmäinen on yhtälön (4.13) *homogenointi* 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 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

$$y'' + a(x)y' + b(x)y = c(x), \quad (4.14)$$

ei edes homogeenisessa tapauksessa $c(x) = 0$, mutta sen sijaan tunnetaan lukuisia differentiaaliyhtälötyyppejä, jotka voidaan ratkaista.

4.2.4 Toisen kertaluvun lineaariset DY:t

Esimerkki 97. Voidaan todeta, että $y_1(x) = x$ ja $y_2(x) = \frac{1}{x}$ ovat molemmat differentiaaliyhtälön

$$x^2 y'' + xy' - y = 0 \quad (4.15)$$

ratkaisuja. Nämä ratkaisut ovat myös (lineaarisesti) riippumattomat, sillä jos kaikilla x :n arvoilla on $C_1 x + C_2 \frac{1}{x} = 0$, on $C_1 x^2 + C_2 = 0$, joten $C_1 = C_2 = 0$. Yleisen lineaarisia differentiaaliyhtälöitä koskevan periaatteen mukaan tällöin yhtälön (4.15) yleinen ratkaisu on siis muotoa $C_1 x + C_2 \frac{1}{x}$.

Aiemmin on esitetty menetelmiä, joilla toisen kertaluvun differentiaaliyhtälöille saadaan ratkaisuja vakiokertoimisessa tapauksessa. Homogeenisen vakiokertoimisen yhtälön tapauksessa kannattaa käyttää yritelmää $y = e^{\lambda x}$ ja ratkaista λ . Yksittäisratkaisun löytämiseen epähomogeeniselle yhtälölle taas soveltuvat Laplace-muunnokset.

Huomautus 28. On huomattava, että yksittäisratkaisua Laplace-muunnosten avulla etsittäessä kannattaa yleensä kiinnittää alkuehdot $y(0) = 0$ ja $y'(0) = 0$.

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 $y = y_1 v$: Tällöin $y' = y_1' v + y_1 v'$ ja $y'' = y_1'' v + y_1' v' + y_1' v' + y_1 v'' = y_1'' v + 2y_1' v' + y_1 v''$ ja sijoitus yhtälöön

$$y'' + ay' + by = c$$

antaa

$$\begin{aligned} & y_1'' v + 2y_1' v' + y_1 v'' + a(y_1' v + y_1 v') + by_1 v = c \\ \iff & y_1 v'' + 2y_1' v' + ay_1 v' + (y_1'' + ay_1' + by_1) v = c \\ \iff & y_1 v' + (ay_1 + 2y_1') w = c, \end{aligned}$$

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 98. Olkoon r vakio ja $y = y(t)$ muuttujan t funktio. Etsitään differentiaaliyhtälön

$$y'' - 2ry' + r^2 y = t \quad (4.16)$$

ratkaisut ilman Laplace-muunnoksia. Sijoitetaan $y = e^{\lambda t}$ homogenisoituun yhtälöön

$$y'' - 2ry' + r^2 y = 0,$$

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 $y'' = r^2 e^{rt} v + 2re^{rt} v' + e^{rt} v''$. Sijoittamalla nämä yhtälöön (4.16) saadaan

$$r^2 e^{rt} v + 2re^{rt} v' + e^{rt} v'' - 2r^2 e^{rt} v - 2re^{rt} v' + r^2 e^{rt} v = t,$$

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

$$v = \frac{t}{r^2}e^{-rt} + \frac{2}{r^3}e^{-rt} + C_1 t + C_2.$$

Lopuksi

$$y = e^{rt} v = \frac{t}{r^2} + \frac{2}{r^3} + C_1 t e^{rt} + C_2 e^{rt}. \quad (4.17)$$

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 (4.16) yksittäisratkaisu. Lauseke (4.17) sisältää siis *kaikki* differentiaaliyhtälön (4.16) ratkaisut.

4.3 Separoituvat DY:t

Differentiaaliyhtälöä kutsutaan *separoituvaksi*, mikäli se voidaan kirjoittaa muotoon

$$y' = g(x)f(y), \quad (4.18)$$

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 (4.18) kirjoittaa muotoon

$$\frac{d}{dx}F(y(x)) = \frac{d}{dx}G(x), \quad (4.19)$$

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 99. Tarkastellaan esimerkin 94 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 100. Verhulstin populaatiomallissa populaation koko p toteuttaa differentiaaliyhtälön

$$p' = ap - bp^2, \quad (4.20)$$

Missä a ja b ovat positiivisia vakioita. Differentiaaliyhtälöä (4.20) 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

$$\frac{p}{a - bp} = e^{a(t+C)}, \quad (4.21)$$

josta voidaan ratkaista

$$p = \frac{a}{b + e^{-aC}e^{-at}}. \quad (4.22)$$

Sijoittamalla $t = 0$ yhtälöön (4.21) antaa (merkitään $p_0 = p(0)$) $\frac{p_0}{a - bp_0} = e^{aC}$, ja edelleen sijoittamalla tämä yhtälöön (4.22) saadaan ratkaisuksi

$$p(t) = \frac{ap_0}{bp_0 + (a - bp_0)e^{-at}}.$$

Esimerkki 101. 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*).

4.4 Eksaktit DY:t

Jos ensimmäisen kertaluvun differentiaaliyhtälössä ei esiinny derivaatan y' potensseja, voidaan y' ottaa yhteiseksi tekijäksi ja kirjoittaa yhtälö muotoon

$$f(x, y) + g(x, y)y' = 0. \quad (4.23)$$

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:

$$g'(x) = \frac{\partial F}{\partial x} \frac{dx}{dx} + \frac{\partial F}{\partial y} \frac{dy}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} y' \quad (4.24)$$

(miksi näin on, selvitetään osiossa Insinöörimatematiikka IIB).

Määritelmä 41. 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 (4.23) on *eksakti*.

Jos (4.23) 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 (4.23) 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. **Osoittautuu, että funktio $F(x, y)$ on olemassa tarkalleen silloin**, kun

$$\frac{\partial f(x, y)}{\partial y} = \frac{\partial g(x, y)}{\partial x}, \quad (4.25)$$

kunhan $f(x, y)$ ja $g(x, y)$ ovat riittävän säännöllisiä (toisen kertaluvun osittaisderivaatat jatkuvia).

Esimerkki 102. Differentiaaliyhtälö

$$\frac{x}{\sqrt{x^2 + y^2}} - 1 + \frac{y}{\sqrt{x^2 + y^2}} y' = 0 \quad (4.26)$$

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 (4.26) 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 (4.23) ei olisikaan alun perin eksakti, on toisinaan mahdollista löytää ns. *integroiva tekijä*, jolla kerrottuna yhtälöstä tulee eksakti.

Esimerkki 103. Yhtälö

$$(x + y^2)y' - y = 0$$

ei ole eksakti, koska $\frac{\partial}{\partial x}(x + y^2) = 1 \neq -1 = \frac{\partial}{\partial y}(-y)$. Tällä yhtälöllä on kuitenkin integroiva tekijä $\frac{1}{y^2}$, sillä yhtälö

$$\left(\frac{x}{y^2} + 1\right)y' - \frac{1}{y} = 0 \quad (4.27)$$

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 (4.27) 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 104. 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

$$\mu(x)y' + \mu(x)a(x)y - \mu(x)b(x) = 0 \quad (4.28)$$

eksakti, jolloin siis ehdon (4.25) 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

$$\mu(x)a(x) = \mu'(x). \quad (4.29)$$

Oletetaan aluksi, että integroiva tekijä $\mu(x)$ löytyy, jolloin siis (4.29) pätee. Tällöin voidaan yhtälö (4.28) 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

$$y = \frac{1}{\mu(x)} \left(\int \mu(x)b(x) dx + C \right). \quad (4.30)$$

On vielä selvitettävä miten integroiva tekijä $\mu(x)$ voidaan löytää, mutta tämä selviää helposti yhtälön (4.29) 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 (4.30) on sama kuin aiemmin esitetyllä menetelmällä (homogenisointi ja vakion variointi) saatu ratkaisu.

4.5 Sekalaisia menetelmiä

4.5.1 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 105. 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 $y = -x + \tan(x + C)$.

4.5.2 Bernoullin DY

Sijoitus $z = y^{1-a}$ ($a \neq 1$) muuntaa Bernoullin differentiaaliyhtälön

$$y' + p(x)y = q(x)y^a \quad (4.31)$$

lineaariseksi differentiaaliyhtälöksi.

Esimerkki 106. Logistinen differentiaaliyhtälö

$$p' = ap - bp^2$$

on separoituva, mutta myös muotoa (4.31). 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

$$z' + az = b. \quad (4.32)$$

Näin saadun lineaarisen differentiaaliyhtälön ratkaisut saadaan aluksi homogenisoimalla: $z_H = Ce^{-at}$ on homogeenisen yhtälön ratkaisu, minkä jälkeen vakion variointi tuottaa yhtälön (4.32) 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 $p_0 = p(0)$.

4.5.3 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 107. Esimerkin 103 differentiaaliyhtälö voidaan kirjoittaa muodossa

$$x' - \frac{1}{y}x = y,$$

mikä on lineaarinen differentiaaliyhtälö funktion $x(y)$ suhteen.

4.5.4 Muuttujan vaihtaminen

Eulerin differentiaaliyhtälö on muotoa

$$x^2y'' + pxy' + qy = f(x), \quad (4.33)$$

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 ketjusaännöllä

$$\frac{dy}{dx} = \frac{dy_1}{dt} \frac{dt}{dx} = \frac{1}{x} \frac{dy_1}{dt}$$

ja

$$\frac{d^2y}{dx^2} = -\frac{1}{x^2} \frac{dy_1}{dt} + \frac{d^2y_1}{dt} \frac{dt}{dx} = -\frac{1}{x^2} y_1'(t) + \frac{1}{x^2} y_1''(t).$$

Sijoittamalla nämä alkuperäiseen yhtälöön saadaan

$$y_1''(t) + (p-1)y_1'(t) + qy_1(t) = f(e'),$$

jonka ratkaisut voidaan ainakin periaatteessa määrittää Laplace-muunnosta käyttämällä.

4.5.5 Yhtälö $y'' = f(y)$

Toisen kertaluvun differentiaaliyhtälö

$$y'' = f(y) \quad (4.34)$$

voidaan palauttaa ensimmäisen kertaluvun differentiaaliyhtälöksi kertomalla (4.34) 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

$$(y')^2 = 2 \int f(y) dy + C. \quad (4.35)$$

Ottamalla tästä neliöjuuri puolittain saadaan separoituva differentiaaliyhtälö.

Esimerkki 108. 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

$$\frac{d}{dt}(s')^2 = 2k \frac{d}{dt} \frac{1}{s} \quad (4.36)$$

kanssa. Yhtälöstä (4.36) saadaan

$$(s')^2 = \frac{2k}{s} + C. \quad (4.37)$$

Alkuehdot $s(0) = s_0 > 0$ ja $s'(0) = v_0 > 0$ antavat $C = v_0^2 - \frac{2k}{s_0}$ ja siis

$$s' = \sqrt{\frac{2k}{s} + v_0^2 - \frac{2k}{s_0}}, \quad (4.38)$$

mistä s voidaan ratkaista (johtaa tosin työlääseen integrointitehtävään). Vaikka yhtälöä (108) ei ratkaistakaan, sisältävät siihen johtaneet välivaiheet jo paljon informaatiota.

Yhtälöstä (4.37) 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 (4.38) 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) $G = 6,673 \cdot 10^{-11} \frac{\text{m}^3}{\text{kg s}^2}$ (gravitaatiovakio) ja $M = 5,9736 \cdot 10^{24}$ kg (maan massa) saadaan ehdoksi $v_0 > 11,18 \frac{\text{km}}{\text{s}}$.

Yleisemminkin voidaan todeta, että itse differentiaaliyhtälö tarjoaa toisinaan tietoa integraalikäyrästä, vaikka sitä ei olisikaan ratkaistu.

Esimerkki 109. 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 ($y'' < 0$).

4.6 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 IB). Tällöin saadaan differentiaaliyhtälö

$$\frac{dy}{dx} = \frac{f_2(x, y)}{f_1(x, y)}. \quad (4.39)$$

On kuitenkin huomattava, että vaikka (4.39) voitaisiin ratkaista, saadaan ainoastaan sen käyrän yhtälö, jota pitkin $(x(t), y(t))$ kulkee, kun t kasvaa. Itse liikeilmiö jää muilla keinoin selvitettäväksi (millä nopeudella piste (x, y) kulkee, hidastuuko kulku, onko olemassa rajapistettä, jne.). Käyrää $(x(t), y(t))$ sanotaan *ratakäyräksi*.

Esimerkki 110. 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 111. 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 hyöttyy x :stä. Tämänkaltaista dynamiikkaa yksinkertaistetussa muodossa kuvaa differentiaaliyhtälöpari

$$\begin{cases} x'(t) = \alpha x - \beta xy \\ y'(t) = -\gamma x + \delta xy, \end{cases} \quad (4.40)$$

missä vakiot α , β , γ ja δ ovat positiivisia ja termi xy kuvaa lajien kohtaamislaajuutta.

Differentiaaliyhtälöpari (4.40) on autonominen ja ratakäyrien yhtälö on muotoa

$$y' = \frac{y}{x} \frac{\delta x - \gamma}{\alpha - \beta y},$$

minkä ratkaisu

$$\alpha \ln y - \beta y = \delta x - \gamma \ln x + C \quad (4.41)$$

voidaan saada separoimalla yhtälö. Käyrät (4.41) ovat pisteen $(\frac{\gamma}{\delta}, \frac{\alpha}{\beta})$ ympäri kiertäviä suljettuja käyriä, ja parin (4.40) ratkaisut jaksollisia: On olemassa sellainen luku T , että $(x(T), y(T)) = (x(0), y(0))$.

Jos parin (4.40) 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

$$\frac{1}{T} \int_0^T y(t) dt = \frac{\alpha}{\beta}. \quad (4.42)$$

Yhtälön (4.42) 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 *kumppaakin eläinkantaa* tehokkuudella e , tulee yhtälöparin (4.40) 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.

4.7 DY:iden numeerisesta ratkaisemisesta

Differentiaaliyhtälölle

$$y' = f(t, y)$$

alkuarvolla $y(0) = y_0$ ei tunneta yleistä ratkaisukaavaa, vaikka useita erikoistapauksia voidaan ratkaista. Lisäksi ratkaisun *olemassaolo* ja yksikäsitteisyys voidaan taata melko vaatimattomin oletuksin. Jos yhden differentiaaliyhtälön sijasta on tarkasteltava differentiaaliyhtälöryhmiä

$$\begin{cases} y'_1 = f_1(t, y_1, y_2, \dots, y_n) \\ y'_2 = f_2(t, y_1, y_2, \dots, y_n) \\ \dots \dots \\ y'_n = f_n(t, y_1, y_2, \dots, y_n) \end{cases}, \quad (4.43)$$

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 29. Kaikki differentiaaliyhtälöryhmät voidaan saattaa muotoon (4.43) 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, ja niiden johtaminen on melko mutkikasta (sivuutetaan), mutta esittäminen suoraviivaista: Hyväksi havaittussa *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 112. Tarkastellaan differentiaaliyhtälöparia

$$\begin{cases} \frac{dx}{dt} = -2x^2 \\ \frac{dy}{dt} = -\frac{1}{2}xy \end{cases} \quad (4.44)$$

alkuehdoilla $x(0) = 2, y(0) = 1$.

Matlabiin ohjelmoitu `ode23` on muunnella 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 (4.44) oikean puolen funktiot. Vektori `[2 1]` määrittelee alkuarvot $x(0) = 2$ ja $y(0) = 1$.

Seuraavat rivit voidaan kirjoittaa Matlabiin *M*-tiedostoksi:

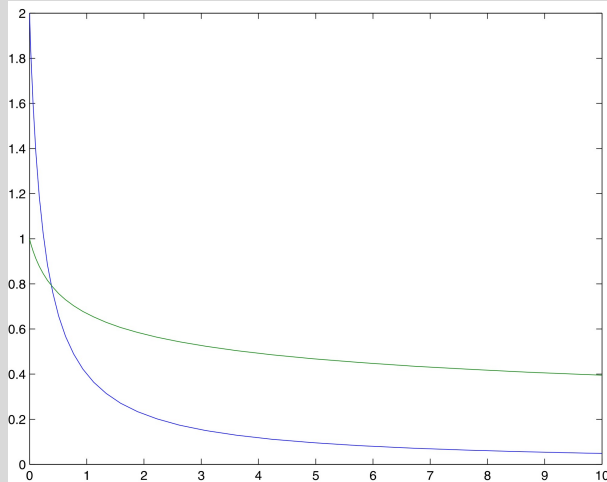
```
function esimerkki
[t, xy]=ode23(@esimpari, [0 10], [2 1])
plot(t, xy)
return

%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 30. 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 $y = \sqrt[4]{\frac{1}{4t+1}}$.

Luku 5

Usean muuttujan funktioiden differentiaalilaskentaa

Tässä luvussa tarkastellaan yhden muuttujan funktioiden differentiaali- ja integraalilaskennan yleistyksiä funktioille, joilla on useampia muuttujia kuin myös funktiolle, joiden arvoa ei kuvata yhdellä vaan useammalla reaalityylillä.

Jos funktiolla ajatellaan olevan n reaalista muuttujaa ja funktion arvon ilmaisee m reaalilukua, voidaan ajatella, että funktio on määritelty jossakin $\mathbb{R}^n$:n osajoukossa ja että funktio saa arvokseen $\mathbb{R}^m$:n vektorin.

Luonteva tapa käsitellä usean muuttujan funktioita on siis tarkastella funktioita $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, jolloin muodollisesti funktiolla f on vain yksi muuttuja $\mathbf{x}$, johon kuitenkin sisältyy n koordinaattia: $\mathbf{x} = (x_1, \dots, x_n)$.

On myös huomattava, että funktion $\mathbb{R}^n \rightarrow \mathbb{R}^m$ voidaan jakaa m :ksi funktioksi $\mathbb{R}^n \rightarrow \mathbb{R}$, jolloin siis funktion arvo ($\mathbb{R}^m$:n vektori) pisteessä $\mathbf{x}$ ($\mathbb{R}^n$:n vektori) voidaan esittää muodossa $(f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$, missä kukin f_i on funktio $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$. Näin esitettynä funktioita f_i sanotaan alkuperäisen funktion f *komponenttifunktioiksi* tai joskus myös *koordinaattifunktioiksi*. Tällöin merkitään myös $f = (f_1, \dots, f_m)$.

Esimerkki 113. Tason esitys napakoordinaateista xy -koordinaatteihin määrittelee kahden muuttujan funktion, jolla on myös kaksi reaalista arvoa. Napakoordinaattiesityksessä

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

r merkitsee pisteen (x, y) etäisyyttä origosta ja θ origosta pisteeseen (x, y) kulkevan janan ja x -akselin välistä kulmaa. Muunnos napakoordinaateista xy -koordinaatteihin voidaan siis kirjoittaa funktiona $\mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$f(r, \theta) = (r \cos \theta, r \sin \theta).$$

Komponenttifunktiot $f_1(r, \theta) = r \cos \theta$ ja $f_2(r, \theta) = r \sin \theta$, ja näillä merkinnöillä

$$f(r, \theta) = (f_1(r, \theta), f_2(r, \theta)).$$

Esimerkki 114. Pallokoordinaateissa (r, θ, ϕ) $r \in [0, \infty)$ merkitsee pisteen etäisyyttä origosta, $\theta \in [0, 2\pi)$ kulmaa xz -tason ja $\phi \in [0, \pi)$ kulmaa z -akselin välillä. Yhteys xyz -koordinaatiston välillä voidaan esittää seuraavasti:

$$\begin{cases} x = r \cos \theta \sin \phi \\ y = r \sin \theta \sin \phi \\ z = r \cos \phi. \end{cases}$$

Jos merkitään $f_1(r, \theta, \phi) = r \cos \theta \sin \phi$, $f_2(r, \theta, \phi) = r \sin \theta \sin \phi$ ja $f_3(r, \theta, \phi) = r \cos \phi$, on muunnos pallokoordinaateista suorakulmaisiin koordinaatteihin esitettävissä funktiona $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $f(r, \theta, \phi) = (f_1(r, \theta, \phi), f_2(r, \theta, \phi), f_3(r, \theta, \phi))$.

Esimerkki 115. Jos A on $m \times n$ -matriisi, määrittelee matriisikertolasku $\mathbf{x} \rightarrow A\mathbf{x}$ funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Tälle funktiolle pätee $f(a\mathbf{x} + b\mathbf{y}) = A(a\mathbf{x} + b\mathbf{y}) = aA\mathbf{x} + bA\mathbf{y} = af(\mathbf{x}) + bf(\mathbf{y})$. Tämän ehdon toteuttavia funktiota sanotaan *lineaariseksi*. Lineaariset funktiot ovat kaikkein yksinkertaisimpia usean muuttujan funktioita, ja niiden rooli usean muuttujan differentiaalilaskennassa on samankaltainen kuin käyrän approksimointi suoralla yhden muuttujan differentiaalilaskennassa.

Funktioilla, joiden määrittelyjoukko on jokin kolmiulotteisen avaruuden $\mathbb{R}^3$ osajoukko, saattaa olla fysikaalinen tulkinta, olipa maalijoukko $\mathbb{R}$:n, $\mathbb{R}^2$:n tai $\mathbb{R}^3$:n osajoukko.

Esimerkki 116. Jatkuva funktio (jatkuvuus määritellään myöhemmin) $\mathbb{R}^2 \rightarrow \mathbb{R}$ esittää pintaa $z = f(x, y)$. Esimerkiksi $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 - y^2$ määrittelee hyperbolisen paraboloidin $z = f(x, y) = x^2 - y^2$.

Esimerkki 117. Jatkuva funktio $f = (f_1, f_2, f_3), \mathbb{R} \rightarrow \mathbb{R}^3$ esittää käyrää kolmiulotteisessa avaruudessa. Jos jokainen komponenttifunktio f_i on lisäksi derivoituva, sanotaan, että käyrä on *sileä*.

Esimerkki 118. Funktio $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ voi esittää fysikaalista *skalaarisuuretta*. Esimerkiksi kolmiulotteisen kappaleen lämpötilaa T pisteessä $\mathbf{x} = (x, y, z)$ voidaan esittää funktiolla $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, $T = f(\mathbf{x})$.

Esimerkki 119. Niin sanottuihin *vektorisuureisiin* liittyy suuruuden (itseisarvon) lisäksi myös suunta. Esimerkiksi magneettikentän voimakkuutta ja suuntaa kussakin avaruuden $\mathbb{R}^3$ pisteessä esittää kolmiulotteinen vektori – siis avaruuden $\mathbb{R}^3$ alkio. Täten magneettikenttää voidaan esittää funktiolla $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, missä $\mathbf{y} = f(\mathbf{x})$ on magneettikenttään suuntaa ja voimakkuutta kuvaava vektori $\mathbf{y} \in \mathbb{R}^3$ pisteessä $\mathbf{x} \in \mathbb{R}^3$.

5.1 Jatkuvuus

Usean muuttujan funktioiden yhteydessä tarvitaan Insinöörimatematiikka IIA:ssa esitettyjä käsitteitä. Vektoreiden $\mathbf{x} = (x_1, \dots, x_n)$ ja $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ *pistetulo* (*sisätulo*) määriteltiin $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + \dots + x_ny_n$, ja tämän perusteella vektorin $\mathbf{x} \in \mathbb{R}^n$ *pituus* (*normi*) $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$.

Seuraava lause esittää kvantitatiivisesti intuitiivisen idean, jonka mukaan vektorin koordinaatit ovat pienet, mikäli vektorilla on pieni normi, ja päinvastoin, vektorin normi on pieni, jos sen koordinaatit ovat pienet.

Lause 21. Vektorille $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ pätee

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

Todistus. Jätetään harjoitustehtäväksi.

Normikäsitteiden perusteella määritellään avaruudessa $\mathbb{R}^n$ vektoreiden etäisyys.

Määritelmä 42. Vektoreiden $\mathbf{x}$ ja $\mathbf{y} \in \mathbb{R}^n$ *etäisyys* määritellään

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

Aiemmasta lauseesta seuraa tällöin suoraan, että jos kaksi vektoria ovat lähellä toisiaan, ovat niiden kaikki vastinkoordinaatit lähellä toisiaan ja päinvastoin.

Edellisen määritelmän mukainen etäisyys vastaa $\mathbb{R}^3$:ssa ja $\mathbb{R}^2$:ssa tavanomaista eukliidistä etäisyyttä ja normin kolmioepäyhtälöstä seuraa, että

$$d(\mathbf{x}, \mathbf{z}) = \|\mathbf{x} - \mathbf{z}\| = \|\mathbf{x} - \mathbf{y} + \mathbf{y} - \mathbf{z}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\| = d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}).$$

Myös käsitteillä *raja-arvo* ja *jatkuvuus* on välitön vastineensa usean muuttujan funktioiden teoriassa.

Määritelmä 43 (Raja-arvo). Olkoon f funktio $\mathbb{R}^n \rightarrow \mathbb{R}^m$. Sanotaan, että *funktion f raja-arvo pisteessä $\mathbf{a} \in \mathbb{R}^n$ on $\mathbf{y} \in \mathbb{R}^m$* ja merkitään

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \mathbf{y},$$

mikäli $f(\mathbf{x})$ saadaan miten tahansa lähelle pistettä $\mathbf{y}$, kunhan vain $\mathbf{x} \neq \mathbf{a}$ valitaan riittävän läheltä pistettä $\mathbf{a}$.

Täsmällisemmin ilmaistuna: Funktio $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ raja-arvo pisteessä $\mathbf{a}$ on $\mathbf{y}$, jos jokaista positiivilukua ε on olemassa luku $\delta_\varepsilon > 0$ siten että

$$d(f(\mathbf{x}), \mathbf{y}) < \varepsilon,$$

mikäli

$$0 < d(\mathbf{x}, \mathbf{a}) < \delta_\varepsilon.$$

Myös jatkuvuuden määritelmä voidaan esittää samankaltaisena kuin reaali-lukujen tapauksessa:

Määritelmä 44. Funktio $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ on *jatkuva* pisteessä $\mathbf{a} \in \mathbb{R}^n$, jos f on ensinnäkin määritelty pisteessä $\mathbf{a}$ ja jos jokaista positiivilukua ε kohti on olemassa positiiviluku δ_ε siten, että

$$d(f(\mathbf{x}), f(\mathbf{a})) < \varepsilon,$$

kunhan $\mathbf{x}$ on valittu niin läheltä $\mathbf{a}$:ta, että

$$d(\mathbf{x}, \mathbf{a}) < \delta_\varepsilon.$$

Huomautus 31. Merkintää $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ käyttämällä jatkuvuuden määritelmä voidaan kirjoittaa seuraavasti: Jokaista positiivilukua ε kohti on olemassa $\delta_\varepsilon > 0$ siten, että

$$\|f(\mathbf{x}) - f(\mathbf{a})\| < \varepsilon \quad \text{aina kun} \quad \|\mathbf{x} - \mathbf{a}\| < \delta_\varepsilon.$$

Usean muuttujan funktion jatkuvuus voidaan päätellä komponenttifunktioiden jatkuvuudesta, kuten seuraava lause asian ilmaisee.

Lause 22. Jos funktiot $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ovat kaikki jatkuvia pisteessä $\mathbf{a}$, niin myös funktio $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ on jatkuva pisteessä $\mathbf{a}$. Kääntäen, jos funktio $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ on jatkuva pisteessä $\mathbf{a}$, ovat kaikki komponenttifunktiot f_i jatkuvia pisteessä $\mathbf{a}$.

Todistus. Jos kukin komponenttifunktio $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ on jatkuva, on positiivilukua $\frac{\varepsilon}{n}$ kohti olemassa luku $\delta_i > 0$ jolle pätee $\|\mathbf{x} - \mathbf{a}\| < \delta_i \implies |f_i(\mathbf{x}) - f_i(\mathbf{a})| < \frac{\varepsilon}{n}$. Kun valitaan $\delta = \min\{\delta_1, \dots, \delta_n\}$, saadaan

$$\begin{aligned} \|f(\mathbf{x}) - f(\mathbf{a})\| &\leq |f_1(\mathbf{x}) - f_1(\mathbf{a})| + \dots + |f_n(\mathbf{x}) - f_n(\mathbf{a})| \\ &< \frac{\varepsilon}{n} + \dots + \frac{\varepsilon}{n} = \varepsilon, \end{aligned}$$

mikä kertoo sen, että $\|f(\mathbf{x}) - f(\mathbf{a})\|$ saadaan kyllä pienemmäksi kuin ε , kunhan vain $\|\mathbf{x} - \mathbf{a}\|$ on pienempi kuin $\delta = \min\{\delta_1, \dots, \delta_n\}$.

Oletetaan sitten, että $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ on jatkuva ja näytetään, että kaikki komponenttifunktiot f_i ovat jatkuvia. Tämä seuraa suoraan epäyhtälöstä

$$|f_i(\mathbf{x}) - f_i(\mathbf{a})| \leq \|f(\mathbf{x}) - f(\mathbf{a})\|. \square$$

Edellisen lauseen mukaan funktion $\mathbb{R}^n \rightarrow \mathbb{R}^m$ jatkuvuus on sama asia kuin komponenttifunktioiden $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ jatkuvuus. Tällöin siis on vielä tarkasteltava, miten funktioiden $f: \mathbb{R}^n \rightarrow \mathbb{R}$ jatkuvuutta voitaisiin käsitellä.

Lause 23. Projektiofunktiot $\mathbb{R}^n \rightarrow \mathbb{R}$, $p_i(x_1, \dots, x_n) = x_i$ ovat jatkuvia. Lisäksi

1. Jos f ja $g: \mathbb{R}^n \rightarrow \mathbb{R}$ ovat jatkuvia pisteessä $\mathbf{a} \in \mathbb{R}^n$, niin myös funktio $\alpha f + \beta g: \mathbb{R}^n \rightarrow \mathbb{R}$ on.
2. Funktio fg on jatkuva pisteessä $\mathbf{a}$ jos f ja g ovat.
3. Funktio $\frac{f}{g}$ on jatkuva pisteessä $\mathbf{a}$, jos $g(\mathbf{a}) \neq 0$.

Todistus. Projektiofunktioiden jatkuvuus seuraa suoraan epäyhtälöstä $|x_i - a_i| \leq \|\mathbf{x} - \mathbf{a}\|$. Lauseen muut kohdat todistetaan oikeaksi samalla tavalla kuin vastaavat tulokset reaalfunktioille. $\square$

Lause 24 (Yhdistetty funktio). Jos $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ on jatkuva pisteessä $\mathbf{a} \in \mathbb{R}^n$ ja $g: \mathbb{R}^m \rightarrow \mathbb{R}^k$ jatkuva pisteessä $\mathbf{b} = f(\mathbf{a})$, niin yhdistetty funktio $g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^k$ on jatkuva pisteessä $\mathbf{a}$.

Todistus. Todistus on täsmälleen samankaltainen kuin vastaavan, reaalfunktioita koskevan tuloksen todistus: On näytettävä toteen, että $\|g(f(\mathbf{x})) - g(f(\mathbf{a}))\|$ saadaan mielivaltaisen pieneksi, kunhan $\|\mathbf{x} - \mathbf{a}\|$ valitaan kyllin pieneksi (harjoitustehtävä). $\square$

Esimerkki 120. Näytetään, että esimerkin 113 funktio

$$f(r, \theta) = (r \cos \theta, r \sin \theta)$$

on jatkuva. Lauseen 22 perusteella riittää osoittaa, että komponenttifunktiot $f_1(r, \theta) = r \cos \theta$ ja $f_2(r, \theta) = r \sin \theta$ ovat jatkuvia.

Ensinnäkin voidaan todeta, että funktiot $(r, \theta) \rightarrow r$ ja $(r, \theta) \rightarrow \theta$ ovat projektiofunktioina jatkuvia (lause 23). Funktio $(r, \theta) \rightarrow \cos \theta$ taasen on jatkuva jatkuvien funktioiden $(r, \theta) \rightarrow \theta$ ja $\theta \rightarrow \cos \theta$ yhdistettynä funktiona. Edelleen, funktio $f_1(r, \theta) = r \cos \theta$ on jatkuva jatkuvien funktioiden $(r, \theta) \rightarrow r$ ja $(r, \theta) \rightarrow \cos \theta$ tulona jatkuva. Samoin voidaan todeta, että komponenttifunktio f_2 on jatkuva.

Usean muuttujan funktioiden raja-arvo on kuitenkin ongelmallisempi käsite kuin vastaava yhden muuttujan tapauksessa. Voidaan nimittäin ajatella, että jotakin pistettä $\mathbf{a}$ lähestytään eri suunnista.

Esimerkki 121. Tarkastellaan funktiota $f: \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$,

$$f(x, y) = \frac{xy}{x^2 + y^2}. \quad (5.1)$$

Funktio ei ole määritelty origossa, mutta jos origoa lähestytään suoraa $y = kx$ pitkin, on tällä suoralla

$$f(x, kx) = \frac{x \cdot kx}{x^2 + (kx)^2} = \frac{k}{1 + k^2}.$$

Tästä seuraa, että funktion (5.1) arvoa origossa ei voi määritellä siten, että funktiosta tulisi origossa jatkuva, sillä kutakin k :n arvoa kohti saadaan oma raja-arvonsa.

Täten siis jo $\mathbb{R}^2$:ssa voi olla äärettömän monta eri raja-arvoa lähestyttäessä tarkasteltavaa pistettä eri suunnista. Koska reaaliakselilla $\mathbb{R}$ suuntia oli vain kaksi, oli mahdollisia raja-arvojakin vain kaksi: oikeanpuoleinen tai vasemmanpuoleinen.

Eri lähestymissuuntien lisäksi on myös mahdollisesti olemassa eri lähestymistapoja (esim. pistettä lähestyvää spiraalaa pitkin), joilla saadaan erilaisia raja-arvoja funktiolle f . Jos kuitenkin määritelmän 43 ehto pätee, tuottavat kaikki mahdolliset lähestymistavat saman raja-arvon.

Jatkuvuuden käsite yleistettiin aiemmin ja derivoituvuuden käsite usean muuttujan funktioille esitetään tuota pikaa. Myös suljetun välin käsite on melko suoraviivaista yleistää $\mathbb{R}^n$:ään: suljettu väli $[a, b]$ voidaan yleistää karteesiseksi tuloksi.

Määritelmä 45. Avaruuden $\mathbb{R}^n$ suljettu väli on karteesinen tulo

$$[a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]. \quad (5.2)$$

Avaruudessa $\mathbb{R}^2$ välin geometrinen vastine on suorakaide ja avaruudessa $\mathbb{R}^3$ suorakulmio. Kun $n > 3$, visuaalinen esitysmahdollisuus puuttuu, mutta tästä huolimatta joukkoa (5.2) kutsutaan myös avaruuden $\mathbb{R}^n$ suorakulmioksi.

Näin yleistetty välin käsite ei kuitenkaan ole tarkoituksenmukaisin funktioiden ääriarvoja etsittäessä. Osoittautuu, että tällöin luontevin suljetun välin yleistys on *kompakti* joukko. Intuitiivisesti kompakti joukko merkitsee avaruudessa $\mathbb{R}^n$ rajoitettua joukkoa, joka sisältää reunansa. Täsmällistä määritelmää varten on otettava käyttöön seuraavat käsitteet.

Määritelmä 46. R -säteinen, $\mathbf{x}$ -keskinen avoin pallo avaruudessa $\mathbb{R}^n$ on joukko

$$B(\mathbf{x}, R) = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| < R\}.$$

Pallo $B(\mathbf{x}, R)$ siis koostuu niistä avaruuden $\mathbb{R}^n$ pisteistä, jotka ovat alle R :n etäisyydellä pisteestä $\mathbf{x}$.

Määritelmä 47. Joukko $A \subseteq \mathbb{R}^n$ on *kompakti* jos se on sekä suljettu että rajoitettu (nämä käsitteet määritellään seuraavaksi).

Määritelmä 48. Joukko $A \subseteq \mathbb{R}^n$ on *rajoitettu* jos on olemassa sellainen M , että $\|\mathbf{x}\| \leq M$ aina, kun $\mathbf{x} \in A$. Toisin sanoen, A on rajoitettu joukko, jos se sisältyy johonkin origokeskiseen, M -säteiseen palloon.

Esimerkki 122. x -akseli $\{(x, 0, 0) \mid x \in \mathbb{R}\}$ avaruudessa $\mathbb{R}^3$ ei ole rajoitettu, sillä $\|(x, 0, 0)\| = |x|$ joten siis x -akselilla on pisteitä, joilla on miten suuri normi tahansa.

Esimerkki 123. Ellipsoidi $\left(\frac{x}{2}\right)^2 + y^2 + 9z^2 = 1$ on rajoitettu, sillä jokainen ellipsoidin piste toteuttaa $1 = \left(\frac{x}{2}\right)^2 + y^2 + 9z^2 \geq \frac{1}{4}x^2 + \frac{1}{4}y^2 + \frac{1}{4}z^2 = \frac{1}{4}\|(x, y, z)\|^2$, joten $\|(x, y, z)\| \leq 2$.

Esimerkki 124. Yksivaippainen hyperboloidi $x^2 + y^2 - z^2 = 1$ ei ole rajoitettu, sillä siihen kuluu pisteitä $(a, 1, a)$, joiden normi $\|(a, 1, a)\| = \sqrt{2a^2 + 1}$ voi olla miten suuri tahansa.

Määritelmä 49. Joukko $A \subseteq \mathbb{R}^n$ on *suljettu* jos sen komplementti $\mathbb{R}^n \setminus A$ on avoin.

Määritelmä 50. Joukko $A \subseteq \mathbb{R}^n$ on *avoin*, jos sen jokaisen pisteen $\mathbf{x}$ ympärille voidaan asettaa jokin ε -säteinen pallo $B(\mathbf{x}, \varepsilon)$, joka sisältyy joukkoon A .

Esimerkki 125. Pallo $B(\mathbf{x}, r)$ on avoin joukko, sillä jos $\mathbf{y} \in B(\mathbf{x}, r)$, on määritelmän mukaan $d = \|\mathbf{y} - \mathbf{x}\| < r$, joten pisteen $\mathbf{y}$ ympärille piirtää $\frac{r-d}{2}$ -säteinen, joukkoon $B(\mathbf{x}, r)$ sisältyvä avoin pallo.

Esimerkki 126. Joukko

$$\bar{B}(\mathbf{x}, R) = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{x}\| \leq R\}$$

on suljettu joukko, sillä sen komplementti $\mathbb{R}^n \setminus \bar{B}(\mathbf{x}, R)$ on avoin. Tämä nähdään siten, että jos $\mathbf{y} \in \mathbb{R}^n \setminus \bar{B}(\mathbf{x}, R)$, on $d\|\mathbf{y} - \mathbf{x}\| > R$ ja tällöin pisteen $\mathbf{y}$ ympärille voidaan piirtää $\frac{R-d}{2}$ -säteinen, joukkoon $\mathbb{R}^n \setminus \bar{B}(\mathbf{x}, R)$ kuuluva pallo.

Esimerkki 127. $\mathbb{R}$:ssä suljetut välit $[a, b]$ ovat suljettuja joukkoja ja avoimet välit (a, b) avoimia joukkoja. Puoliavoimet välit $[a, b)$ tai $(a, b]$ eivät ole suljettuja eivätkä avoimia joukkoja.

Avaruuden $\mathbb{R}^n$ joukko voidaan usein nähdä suljetuksi käyttämällä seuraavaa lausetta, jonka todistus sivuutetaan.

Lause 25. Jos $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on jatkuva ja A on suljettu $\mathbb{R}$:n joukko, niin myös $f^{-1}(A)$ on suljettu.

Lisäksi pätee seuraava tulos:

Lause 26. Suljettujen joukkojen leikkaus on suljettu. Äärellisen monen suljetun joukon unioni on suljettu. Suljettujen joukkojen karteesinen tulo on suljettu.

Todistus. Harjoitustehtävä.

Esimerkki 128. Avaruuden $\mathbb{R}^3$ yksikköpallon pinta $x^2 + y^2 + z^2 = 1$ on suljettu joukko, sillä on suljetun joukon $[1, 1] = \{1\}$ alkukuva jatkuvassa kuvauksessa $f(x, y, z) = \|(x, y, z)\|$ (miksi tämä kuvaus on jatkuva?).

Myös yksikköpallo $\bar{B}(\mathbf{0}, 1)$, jonka määrittää epäyhtälö $x^2 + y^2 + z^2 \leq 1$ on suljettu joukko, sillä se on suljetun joukon $[0, 1]$ alkukuva $f^{-1}([0, 1])$.

Lause 27. Kompaktissa joukossa $A \subseteq \mathbb{R}^n$ määritellyllä jatkuvalla funktiolla $f : A \rightarrow \mathbb{R}$ on sekä maksimi että minimi.

Todistus. Sivuuutetaan.

5.2 Osittaisderivaatat

Osittaisderivaatan käsitettä on käsitelty jo aiemmissa Insinöörimatematiikan osioissa. Kerrataan ja yleistetään tämä käsite ja esitetään jatkossa tarvittava tulos osittaisderivoinnin järjestyksen vaihtamisesta.

Määritelmä 51. Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ osittaisderivaatta i :n muuttujan x_i suhteen pisteessä $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ on

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h},$$

mikäli raja-arvo on olemassa. Osittaisderivaatasta i :n muuttujan suhteen käytetään myös merkintöjä $f_{x_i}(\mathbf{a})$ ja $D_{x_i}f(\mathbf{a})$. Jos muuttujien $x_1, x_2, \dots, x_n$ on kiinteä, niin osittaisderivaatasta $D_{x_i}f(\mathbf{a})$ käytetään myös merkintää $D_i f(\mathbf{a})$.

Määritelmä 52. Toisen kertaluvun osittaisderivaatoista käytetään merkintöjä $\frac{\partial^2 f}{\partial x_j \partial x_i}$, $f_{x_i x_j}$, $D_{x_i x_j} f$, tai $D_{ij} f$.

Lause 28. Jos osittaisderivaatat $\frac{\partial^2 f}{\partial x_i \partial x_j}$ ja $\frac{\partial^2 f}{\partial x_j \partial x_i}$ ovat olemassa jossakin pisteen $\mathbf{a}$ ympäristössä ja jatkuvia pisteessä $\mathbf{a}$, niin

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{a}).$$

Todistus. Pitkä, sivuutetaan.

5.3 Differentioituvuus

Palautetaan mieleen kurssilta Insinöörimatematiikka IB käsite derivoituvuus: Sanotaan, että reaalifunktio f on *derivoituva* pisteessä x , jos on olemassa sellainen luku k ja funktio $\varepsilon(h)$, että $\lim_{h \rightarrow 0} \varepsilon(h) = \varepsilon(0) = 0$ ja

$$f(x+h) - f(x) = kh + \varepsilon(h)h. \quad (5.3)$$

Yhtälö (5.3) ilmaisee sen, että funktion arvon muutosta $f(x+h) - f(x)$ (pisteessä x) voidaan arvioida x :n muutoksen h ensimmäisen asteen polynomifunktiona $h \mapsto k \cdot h$:

$$f(x+h) - f(x) \approx k \cdot h,$$

kun h on pieni. Yhtälössä (5.3) esiintyvää lukua k kutsutaan funktion f *derivaataksi* pisteessä x ja merkitään $k = f'(x)$. On siis huomattava, että yhtälössä (5.3) esiintyvä luku k (samoin kuin funktio $\varepsilon(h)$) voi riippua pisteen x arvosta.

Esimerkki 129. Olkoon $f(x) = x^2 - x + 2$. Tällöin

$$\begin{aligned} f(x+h) - f(x) &= (x+h)^2 - (x+h) + 2 - (x^2 - x + 2) \\ &= x^2 + 2xh + h^2 - x - h + 2 - x^2 + x - 2 \\ &= 2xh - h + h^2 = (2x-1)h + h^2 \\ &= (2x-1)h + \varepsilon(h)h, \end{aligned}$$

kun merkitään $\varepsilon(h) = h$. Täten siis $f'(x) = 2x - 1$ ja yllä olevan yhtälön mukaisesti lauseketta $f(x+h) - f(x)$ (funktion f muutosta) voidaan approksimoida lausekkeella $(2x-1)h$ h :n ollessa pieni.

On myös huomattava, että derivaatan määritelmässä esiintyvä funktio $f(h) = k \cdot h$ on *lineaarinen*, toisin sanoen $f(ah_1 + bh_2) = k(ah_1 + bh_2) = akh_1 + bkh_2 = af(h_1) + bf(h_2)$.

Lineaarikuvausten $f: \mathbb{R} \rightarrow \mathbb{R}$ suoraviivainen yleistys on käsitelty Insinöörimatematiikka IA:ssa: lineaarikuvaus $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ voidaan esittää $m \times n$ -matriisin A avulla, jolloin

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

kun $\mathbf{x}$ ajatellaan sarakevektoriksi ($n \times 1$ -matriisiksi).

Derivaatan käsite yleistetään käyttämällä lineaarisen approksimaation vaatimusta. Tällöin myös termi ”derivoituvuus” korvataan termillä ”differentioituvuus”.

Määritelmä 53. Funktio $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ on *differentioituva* pisteessä $\mathbf{x}$, jos on olemassa sellainen lineaarikuvaus $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, että

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = T\mathbf{h} + \varepsilon(\mathbf{h})\|\mathbf{h}\|,$$

missä $\varepsilon(\mathbf{h})$ on ehdon $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \varepsilon(\mathbf{h}) = \varepsilon(\mathbf{0}) = \mathbf{0}$ toteuttava funktio $\mathbb{R}^n \rightarrow \mathbb{R}^m$.

Määritelmä 54. Jos $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ on differentioituva, kutsutaan määritelmän (53) lineaarikuvausta $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funktion f (Fréchet'n) *derivaataksi* pisteessä $\mathbf{x}$ ja siitä käytetään merkinä $Df(\mathbf{x})$. Derivaatan $Df(\mathbf{x})$ matriisia kutsutaan *Jacobin matriisiksi* (pisteessä $\mathbf{x}$) ja merkitään $J_f(\mathbf{x})$.

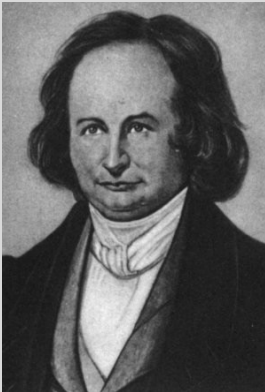
Taustatietoa



Maurice Fréchet (1878–1973) oli ranskalainen matemaatikko. Hän kehitti topologiaa, ns. metristen avaruuksien käsitteen ja esitti yleisen määritelmän derivaatalle metrisissä avaruuksissa. Hän esitti mm. tällä kurssilla käytettävän kompaktiuden käsitteen.

(kuva: Wikimedia Commons)

Taustatietoa



Carl Gustav Jacob Jacobi (1804–1851) oli saksalainen matemaatikko, joka kehitti ns. elliptisten funktioiden teoriaa, differentiaaliyhtälöiden teoriaa, sekä Newtonin mekaniikkaa yleisempään muotoon. Jacobia pidettiin myös erinomaisena opettajana.

(kuva: Wikimedia Commons)

Differentioituvuus pisteessä $\mathbf{x}$ tarkoittaa siis sitä, että funktion f muutosta $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$ voidaan approksimoida lineaarikuvauksella niin hyvin, että approksimaation virhetermi lähestyy nollaa hyvin voimakkaasti, kun $\mathbf{h} \rightarrow \mathbf{0}$.

Välittömästi voidaan nähdä, että $f(\mathbf{x} + \mathbf{0}) - f(\mathbf{x}) = \mathbf{0}$, joten $T\mathbf{0} = \mathbf{0}$ – ominaisuus, joka on oltava voimassa kaikille lineaarikuvauksille. Jos funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ komponenttifunktiot voidaan esittää Taylorin polynomien avulla, voidaan määritelmässä esiintyvä lineaarikuvaus saada esille suoraviivaisesti erottamalla lausekkeen $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$ 1. asteen termit korkeamman asteen termeistä.

Esimerkki 130. Tarkastellaan esimerkin 113 funktion $f(r, \theta) = (r \cos \theta, r \sin \theta)$ differentioituvuutta pisteessä $\mathbf{0} = (0, 0)$.

Tällöin on siis arvioitava muutosta

$$f(\mathbf{0} + \mathbf{h}) - f(\mathbf{0}) = f(\mathbf{h}) = (h_1 \cosh_2, h_1 \sin h_2)$$

kun vektori $\mathbf{h} = (h_1, h_2)$ on lähellä nollaa. Kirjoitetaan

$$\begin{aligned} (h_1 \cosh_2, h_1 \sin h_2) &= (h_1(1 - O(h_2)), h_1 \cdot O(h_2)) = (h_1 - h_1 O(h_2), h_1 O(h_2)) \\ &= (h_1, 0) + (h_1 O(h_2), h_1 O(h_2)). \end{aligned}$$

Nyt siis lineaarikuvaus $T(h_1, h_2) = (h_1, 0)$ on funktion f Fréchet'n derivaatta pisteessä $(0, 0)$. Muodollista todistusta varten tulee vielä arvioida jäännöstermiä: $(h_1 O(h_2), h_1 O(h_2))$. Merkitään

$$\varepsilon(\mathbf{h}) = \|\mathbf{h}\|^{-1}(h_1 O(h_2), h_1 O(h_2)),$$

jolle saadaan arvio

$$\begin{aligned} d(\varepsilon(\mathbf{h}), \mathbf{0})^2 &= \|\varepsilon(\mathbf{h})\|^2 = \|\mathbf{h}\|^{-2}((h_1 O(h_2))^2 + (h_1 O(h_2))^2) \\ &\leq \|\mathbf{h}\|^{-2} C h_1^2 h_2^2 \leq C \|\mathbf{h}\|^{-2} \|\mathbf{h}\|^4 = C \|\mathbf{h}\|^2 = C d(\mathbf{h}, \mathbf{0})^2, \end{aligned}$$

joten $\varepsilon(\mathbf{h}) \rightarrow \mathbf{0}$, kun $\mathbf{h} \rightarrow \mathbf{0}$. Yllä esiintyvä vakio C saadaan ordomerkinnän vakioista.

Huomautus 32. Lineaarikuvauksen $T(h_1, h_2) = (h_1, 0)$ matriisi on $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, sillä $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ 0 \end{pmatrix}$. Lineaarikuvauksen $\mathbb{R}^n \rightarrow \mathbb{R}$ matriisi puolestaan on muotoa $1 \times n$, siis vaakavektori, jolloin siis lineaarikuvaus $\mathbb{R}^n \rightarrow \mathbb{R}$ on muotoa

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto (a_1 \ a_2 \ \dots \ a_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + a_2 x_2 + \dots + a_n x_n = \mathbf{a} \cdot \mathbf{x},$$

missä $\mathbf{a} \cdot \mathbf{x}$ tarkoittaa vektoreiden $\mathbf{a} = (a_1, a_2, \dots, a_n)$ ja $\mathbf{x} = (x_1, x_2, \dots, x_n)$ pistetuloa. Täten lineaarikuvaus $\mathbb{R}^n \rightarrow \mathbb{R}$ voidaan siis yhtä hyvin kirjoittaa muotoon $\mathbf{x} \mapsto \mathbf{a} \cdot \mathbf{x}$.

Huomautus 33. Funktioille $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ lineaarikuvaus saa muodon

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto (a \ b) \begin{pmatrix} x \\ y \end{pmatrix} = ax + by.$$

Tällöin siis differentioituvuus pisteessä $\mathbf{x}_0$ funktiolle $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ voidaan kirjoittaa muotoon

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = ah + bk + \varepsilon(\mathbf{h})\|\mathbf{h}\|, \quad (5.4)$$

missä on merkitty $\mathbf{h} = (h, k)$. Jos edelleen merkitään $\mathbf{x}_0 = (x_0, y_0)$, voidaan (5.4) kirjoittaa muotoon

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = ah + bk + \varepsilon(h, k)\sqrt{h^2 + k^2}, \quad (5.5)$$

missä $\varepsilon(h, k) \rightarrow (0, 0)$ kun $(h, k) \rightarrow (0, 0)$. Yhtälön (5.5) mukaan muutosta pisteestä $z_0 = f(x_0, y_0)$ pisteeseen $z = f(x_0 + h, y_0 + k)$ voidaan arvioida lineaarisen lausekkeen avulla:

$$z - z_0 \approx ah + bk.$$

Kun merkitään $(x, y) = (x_0 + h, y_0 + k)$, voidaan ylläoleva approksimaatio kirjoittaa muotoon

$$z - z_0 \approx a(x - x_0) + b(y - y_0).$$

Tasoa

$$a(x - x_0) + b(y - y_0) - (z - z_0) = 0 \quad (5.6)$$

sanotaan pinnan $z = f(x, y)$ tangenttitasoksi pisteessä (x_0, y_0) .

Samoin kuin yhden muuttujan funktioiden kohdalla, differentioituvuus takaa jatkuvuuden:

Lause 29. Pisteessä $\mathbf{x}$ differentioituva funktio on myös jatkuva pisteessä $\mathbf{x}$.

Todistus. Helppo, jätetään harjoitustehtäväksi.

Lause 30. Jos $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f = (f_1, \dots, f_m)$ on differentioituva pisteessä $\mathbf{x}$, ovat kaikki osittaisderivaatat $\frac{\partial f_i}{\partial x_j}$ pisteessä $\mathbf{x}$ olemassa, ja lineaarikuvauksen T matriisi (Jacobian matriisi) on

$$(J_f(\mathbf{x}))_{ij} = T_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}).$$

Todistus. Väittämä seuraa tarkastelemalla muotoa $\mathbf{h} = (0, \dots, h_j, \dots, 0)$ olevia muutosvektoreita. Yksityiskohdat jätetään harjoitustehtäväksi. $\square$

Huomautus 34. Edellisen lauseen mukaan huomautuksessa 33 on

$$(a \ b) = \left(\frac{\partial f}{\partial x}(x_0) \ \frac{\partial f}{\partial y}(y_0) \right),$$

jolloin siis tangenttitason yhtälö (5.6) saa muodon

$$\frac{\partial f(x_0)}{\partial x}(x - x_0) + \frac{\partial f(y_0)}{\partial y}(y - y_0) - (z - z_0) = 0.$$

Pinnan $z = f(x, y)$ tangenttitason normaalivektori pisteessä (x_0, y_0) on siis

$$\left(\frac{\partial f(x_0)}{\partial x}, \frac{\partial f(y_0)}{\partial y}, -1 \right).$$

Yleisemmin voidaan osoittaa, että pinnan $F(x, y, z) = 0$ normaaliksi voidaan valita vektori

$$\left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right)$$

(tätä asiaa käsitellään myöhemmin).

Esimerkki 131. Pisteessä (r, θ) esimerkin (130) funktion $f(r, \theta) = (f_1(r, \theta), f_2(r, \theta)) = (r \cos \theta, r \sin \theta)$ Jacobin matriisi on

$$\begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \theta} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}.$$

Jos valitaan $(r, \theta) = (0, 0)$ saadaan Jacobin matriisiksi

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

mitä vastaava lineaarikuvaus on

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ 0 \end{pmatrix},$$

joka löydettiin jo esimerkin 130 yhteydessä.

Edellisessä lauseessa mainittujen osittaisderivaattojen olemassaolosta pisteessä $\mathbf{x}$ ei kuitenkaan seuraa välttämättä funktion f differentioituvuus. Differentioituvuus voidaan kuitenkin taata mikäli osittaisderivaatat ovat jatkuvia.

Lause 31. Oletetaan, että funktiolla $f = (f_1, \dots, f_m)$ on jossakin pisteen $\mathbf{x}$ ympäristössä olemassa osittaisderivaatat $\frac{\partial f_i}{\partial x_j}$, jotka ovat jatkuvia pisteessä $\mathbf{x}$. Tällöin f on differentioituva pisteessä $\mathbf{x}$.

Määritelmä 55. Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$, Jacobin matriisia ($1 \times n$ -matriisi eli vaakavektori) kutsutaan *gradientiksi* ja merkitään

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right).$$

Merkintä ∇f luetaan ” f :n gradientti” tai ”nabla f ”.

Huomautus 35. Edellisen määritelmän merkintää käyttäen funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f = (f_1, f_2, \dots, f_m)$ Jacobin matriisi voidaan kirjoittaa muotoon

$$\begin{pmatrix} \nabla f_1 \\ \nabla f_2 \\ \vdots \\ \nabla f_m \end{pmatrix}$$

Huomautus 36. Huomautuksen 34 mukaan gradientilla on geometrinen tulkinta: pinnan $F(x, y, z) = 0$ normaaliksi voidaan valita ∇F . Myöhemmin nähdään, että muissa yhteyksissä gradientilla on myös toisenlaisia geometrisia tulkintoja.

Esimerkki 132. Olkoon $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = \sin\left(\frac{xy}{z}\right)$, jolloin gradientti pisteessä (x, y, z) on

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = \left(\cos\left(\frac{xy}{z}\right) \cdot \frac{y}{z}, \cos\left(\frac{xy}{z}\right) \cdot \frac{x}{z}, -\cos\left(\frac{xy}{z}\right) \cdot \frac{xy}{z^2} \right).$$

Määritelmä 56 (Kokonaisdifferentiaali). Funktiolle $f: \mathbb{R}^n \rightarrow \mathbb{R}$ differentioituvuusehto voidaan edellämainitun perusteella kirjoittaa muotoon

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{h} + \|\mathbf{h}\| \varepsilon(\mathbf{h}),$$

missä $\varepsilon(\mathbf{h}) \rightarrow \varepsilon(\mathbf{0}) = 0$ kun $\mathbf{h} \mapsto \mathbf{0}$.

Differentioituvuuden määritelmässä esiintyvää lineaarista osaa

$$\nabla f(\mathbf{x}) \cdot \mathbf{h} = \frac{\partial f(\mathbf{x})}{\partial x_1} h_1 + \dots + \frac{\partial f(\mathbf{x})}{\partial x_n} h_n$$

kutsutaan funktion f *kokonaisdifferentiaaliksi* pisteessä $\mathbf{x}$. Jos merkitään $h_i = dx_i$ ja $df = \nabla f(\mathbf{x}) \cdot \mathbf{h}$, voidaan kokonaisdifferentiaalin lauseke kirjoittaa muotoon

$$df = \frac{\partial f}{\partial x_1} dx_1 + \dots + \frac{\partial f}{\partial x_n} dx_n.$$

Koska differentioituvuuden ehdossa jälkimmäinen epälineaarinen osa $\|\mathbf{h}\| \varepsilon(\mathbf{h})$ lähestyy nollaa epsilonin lähestyessä nollaa, on kokonaisdifferentiaali tarkentuva approksimaatio muutokselle

$$f(x_1 + dx_1, \dots, x_n + dx_n) - f(x_1, \dots, x_n);$$

sitä tarkempi mitä lähempänä nollavektoria vektori $(dx_1, \dots, dx_n)$ on.

Esimerkki 133. Arvioidaan, kuinka paljon tulo xy muuttuu, jos mitatuissa arvoissa $x = 10$ ja $y = 50$ on virhe (dx, dy) , jolle pätee $|dx| \leq 0,1$ ja $|dy| \leq 0,2$. Tällä siis tarkoitetaan, että todelliset arvot ovat välillä $x \in [10 - \frac{1}{10}, 10 + \frac{1}{10}]$ ja $y \in [50 - \frac{2}{10}, 50 + \frac{2}{10}]$.

Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(xy) = xy$ kokonaisdifferentiaali on

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = y dx + x dy.$$

Jos tällöin esimerkiksi $(x, y) = (10, 50)$ ja $|dx| \leq 0,1$ sekä $|dy| \leq 0,2$ on

$$|df| = |50 \cdot dx + 10 \cdot dy| \leq 50|dx| + 10|dy| \leq 50 \cdot 0,1 + 10 \cdot 0,2 = 7.$$

On kuitenkin huomattava, että näin saatu yläraja (7) virheelle koskee vain *kokonaisdifferentiaalin* df itseisarvoa, ei *todellisen muutoksen* itseisarvoa. Toisaalta taas kokonaisdifferentiaali on hyvä likimääräistys todelliselle muutokselle, kun muutos on pieni suhteessa alkuperäiseen arvoon (10, 50). Tämän esimerkin lukuarvoilla kokonaisdifferentiaali antaa melko hyvän arvion funktion todellisen arvon muutoksesta.

Tarkempi arvio todellisesta muutoksesta saadaan väliarvolauseen yhteydessä.

5.4 Suunnattu derivaatta

Määritelmä 57. Olkoon $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ja $\mathbf{u} \in \mathbb{R}^n$ sellainen, että $\|\mathbf{u}\| = 1$. Jos raja-arvo

$$\lim_{h \rightarrow 0} \frac{1}{h} (f(\mathbf{x} + h\mathbf{u}) - f(\mathbf{x}))$$

on olemassa, sitä sanotaan f :n *suunnatuksi derivaataksi* tai *Gâteauxin derivaataksi* pisteessä $\mathbf{x}$ suuntaan $\mathbf{u}$. Suunnatusta derivaatasta käytetään merkintöjä $\partial_{\mathbf{u}} f(\mathbf{x})$, $\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x})$ ja $D_{\mathbf{u}} f(\mathbf{x})$.

Taustatietoa

René Gâteaux (1889–1914) oli ranskalainen matemaatikko, joka tunnetaan suunnatun derivaatan käsitteestä. Hän kuoli taistelussa 1. maailmansodan alkupäivinä

Suunnattu derivaatta yleistää osittaisderivaatan käsitteen: Jos nimittäin valitaan $\mathbf{u} = \mathbf{e}_i = (0, \dots, 1, \dots, 0)$ (luonnollisen kannan i :s vektori), huomataan että suunnattu derivaatta vektorin $\mathbf{e}_i$ suhteen on sama kuin osittaisderivaatta muuttujan x_i suhteen.

Jos tarkasteltava funktio on differentioituva (kuten yleensä on laita tässä kurssissa), on suunnatulla derivaatalla derivaatalla yhteys gradienttiin:

Lause 32. Jos $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on differentioituva pisteessä $\mathbf{x}$ ja $\|\mathbf{u}\| = 1$, niin suunnattu derivaatta $D_{\mathbf{u}} f(\mathbf{x})$ on olemassa ja

$$D_{\mathbf{u}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u}$$

(vektorien $\nabla f(\mathbf{x})$ ja $\mathbf{u}$ pistetulo).

Todistus. Differentioituvuuden määritelmän perusteella

$$f(\mathbf{x} + h\mathbf{u}) - f(\mathbf{x}) = Df(\mathbf{x})h\mathbf{u} + \|h\mathbf{u}\|\varepsilon(h\mathbf{u}),$$

missä $\varepsilon(h\mathbf{u}) \rightarrow 0$, kun $h\mathbf{u} \rightarrow \mathbf{0}$.

Lauseen 30 perusteella lineaarikuvauksen $Df(\mathbf{x})$ matriisi on $\nabla f(\mathbf{x})$ ja huomautuksen 32 mukaan

$$Df(\mathbf{x})h\mathbf{u} = h\nabla f(\mathbf{x}) \cdot \mathbf{u}.$$

Koska lisäksi $\|h\mathbf{u}\| = |h|$, saadaan

$$\frac{1}{h} (f(\mathbf{x} + h\mathbf{u}) - f(\mathbf{x})) = \nabla f(\mathbf{x}) \cdot \mathbf{u} + \frac{|h|}{h} \varepsilon(h\mathbf{u})$$

js väite seuraa suoraan tästä, koska $\varepsilon(h\mathbf{u}) \rightarrow 0$, kun $h \rightarrow 0$. $\square$

Funktiolle $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ suunnatun derivaatan geometrinen tulkinta on samankaltainen kuin reaali-funktioille: $D_{\mathbf{u}} f(\mathbf{x})$ kuvaa funktion f kasvunopeutta pisteessä $\mathbf{x}$ vektorin $\mathbf{u}$ suuntaan, ja tulkinta on

sama kun $\mathbb{R}^2$:n sijasta on $\mathbb{R}^n$, mutta tällöin visuaalinen esitys ei onnistu kuten $\mathbb{R}^2$:n tapauksessa. Niemenomaan tämän tulkinnan vuoksi on tärkeää, että suunnatun derivaatan määritelmässä käytetään yksikkövektoria. Käytettäessä muuta kuin yksikkövektoria olisi funktion kasvunopeudessa ”väärä mittakaava”.

Gradientin osoittamalla suunnalla on erikoisasema suunnattujen derivaattojen joukossa, kuten seuraava lause osoittaa.

Lause 33. Olkoon $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differentioituva pisteessä $\mathbf{x}$ ja $\|\mathbf{u}\| = 1$. Tällöin $|D_{\mathbf{u}}f(\mathbf{x})| \leq \|\nabla f(\mathbf{x})\|$ ja $|D_{\mathbf{u}}f(\mathbf{x})| = \|\nabla f(\mathbf{x})\|$ tarkalleen silloin kun $\nabla f(\mathbf{x}) = \lambda \mathbf{u}$ jollekin luvulle λ . Tällöin lisäksi $D_{\mathbf{u}}f(\mathbf{x}) = \|\nabla f(\mathbf{x})\|$ jos $\lambda > 0$ ja $D_{\mathbf{u}}f(\mathbf{x}) = -\|\nabla f(\mathbf{x})\|$, jos $\lambda < 0$.

Todistus. Lauseen 32 perusteella

$$|D_{\mathbf{u}}f(\mathbf{x})| = |\nabla f(\mathbf{x}) \cdot \mathbf{u}| \leq \|\nabla f(\mathbf{x})\| \cdot \|\mathbf{u}\| = \|\nabla f(\mathbf{x})\|.$$

Yllä on käytetty Cauchyn-Schwarzin epäyhtälöä $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$. Lauseen jälkimmäinen osa seuraa siitä, että Cauchyn-Schwarzin epäyhtälössä esiintyy yhtäsuuruus tarkalleen silloin kun vektorit ovat lineaarisesti riippuva, sekä analysoimalla tapaukset $\lambda > 0$ ja $\lambda < 0$ erikseen. $\square$

Edellinen lause siis osoittaa, että differentioituva funktio kasvaa jyrkimmin gradientin osoittamassa suunnassa (ja vastaavasti vähenee jyrkimmin gradientin suuntaa vastaan).

Seuraus 4. Jos $f(x, y, z) = 0$ esittää pintaa avaruudessa $\mathbb{R}^3$, on pinnan pisteeseen (x, y, z) asetettu gradientti $\nabla f(x, y, z)$ kohtisuorassa pintaa $f(x, y, z) = 0$ vastaan. Sama tulos pätee myös $\mathbb{R}^2$:n käyrille: Jos $f(x, y) = 0$ esittää avaruuden $\mathbb{R}^2$ käyrää, on käyrän pisteeseen (x, y) asetettu gradientti $\nabla f(x, y)$ kohtisuorassa käyrää vastaan.

Todistus. Todistuksen tekniset yksityiskohdat ovat työläitä, mutta ideaa ei ole vaikea löytää. Mieti miten tämä tulos seuraa siitä, että funktion muutos on nopeinta juuri gradientin suuntaan.

5.5 Ketjusääntö

Yhden muuttujan funktioille pätee $D(g(f(x))) = Dg(f(x))Df(x)$, ja seuraava lause yleistää tämän usean muuttujan funktioille.

Lause 34 (Ketjusääntö). Olkoot $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ja $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$ differentioituvia kuvauksia. Tällöin yhdistetty kuvaus $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ on differentioituva ja

$$D(g \circ f)(\mathbf{x}) = Dg(f(\mathbf{x})) \circ Df(\mathbf{x}).$$

Todistus. Helppo mutta pitkä, jätetään harjoitustehtäväksi.

Huomautus 37. Kurssilla Insinöörimatematiikka IA todettiin, että jos matriisit T ja S esittävät lineaarikuvauksia $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ja $S : \mathbb{R}^m \rightarrow \mathbb{R}^k$ ovat lineaarikuvauksia, saadaan yhdistetyn kuvauksen $S \circ T$ matriisi tulona ST . Jos siis merkitään $h = (h_1, \dots, h_k) = g \circ f$, on kuvauksen h Jacobin matriisi

$$J_{g \circ f}(\mathbf{x}) = \begin{pmatrix} \frac{\partial h_1(\mathbf{x})}{\partial x_1} & \frac{\partial h_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial h_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial h_2(\mathbf{x})}{\partial x_1} & \frac{\partial h_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial h_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_k(\mathbf{x})}{\partial x_1} & \frac{\partial h_k(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial h_k(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

sama kuin matriisitulo (merkitään $\mathbf{y} = f(\mathbf{x})$)

$$\begin{pmatrix} \frac{\partial g_1(\mathbf{y})}{\partial y_1} & \frac{\partial g_1(\mathbf{y})}{\partial y_2} & \cdots & \frac{\partial g_1(\mathbf{y})}{\partial y_m} \\ \frac{\partial g_2(\mathbf{y})}{\partial y_1} & \frac{\partial g_2(\mathbf{y})}{\partial y_2} & \cdots & \frac{\partial g_2(\mathbf{y})}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_k(\mathbf{y})}{\partial y_1} & \frac{\partial g_k(\mathbf{y})}{\partial y_2} & \cdots & \frac{\partial g_k(\mathbf{y})}{\partial y_m} \end{pmatrix} \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \frac{\partial f_m(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{pmatrix}.$$

Huomautus 38. Edellisestä esimerkistä selviää, että yhdistetyn funktion $g \circ f = (h_1, \dots, h_k)$ r :nnen komponenttifunktion osittaisderivaatta $\frac{\partial h_r}{\partial x_s}$ voidaan laskea matriisitulon määritelmän perusteella seuraavasti:

$$\begin{aligned} \frac{\partial h_r}{\partial x_s}(\mathbf{x}) &= \frac{\partial g_r}{\partial y_1}(f(\mathbf{x})) \frac{\partial f_1}{\partial x_s}(\mathbf{x}) + \frac{\partial g_r}{\partial y_2}(f(\mathbf{x})) \frac{\partial f_2}{\partial x_s}(\mathbf{x}) + \cdots + \frac{\partial g_r}{\partial y_m}(f(\mathbf{x})) \frac{\partial f_m}{\partial x_s}(\mathbf{x}) \\ &= \nabla g_r(f(\mathbf{x})) \cdot \left(\frac{\partial f_1}{\partial x_s}(\mathbf{x}), \dots, \frac{\partial f_m}{\partial x_s}(\mathbf{x}) \right). \end{aligned} \quad (5.7)$$

Jos erityisesti $k = 1$, siis g on funktio $\mathbb{R}^m \rightarrow \mathbb{R}$, on $h = g \circ f$ funktio $\mathbb{R}^n \rightarrow \mathbb{R}$ ja tällöin voidaan kirjoittaa

$$\frac{\partial (g \circ f)}{\partial x_s} = \nabla g(f(\mathbf{x})) \cdot \left(\frac{\partial f_1}{\partial x_s}(\mathbf{x}), \dots, \frac{\partial f_m}{\partial x_s}(\mathbf{x}) \right), \quad (5.8)$$

siis funktion $g \circ f$ osittaisderivaatta $\frac{\partial (g \circ f)}{\partial x_s}$ on gradientin $\nabla g(f(\mathbf{x}))$ ja osittaisderivaatoista muodostetun vektorin

$$\left(\frac{\partial f_1}{\partial x_s}(\mathbf{x}), \dots, \frac{\partial f_m}{\partial x_s}(\mathbf{x}) \right)$$

sisätulo.

Jos lisäksi $n = 1$, on $f = (f_1, \dots, f_m) : \mathbb{R} \rightarrow \mathbb{R}^m$ ja $g : \mathbb{R}^m \rightarrow \mathbb{R}$, ja yhdistetty funktio $h = g \circ f$ reaalifunktio $\mathbb{R} \rightarrow \mathbb{R}$ ja tällöin 5.8 saa muodon

$$\begin{aligned} D(g \circ f)(x) &= \nabla g(f(x)) \cdot (f'_1(x), f'_2(x), \dots, f'_m(x)) \\ &= \frac{\partial g}{\partial x_1} f'_1(x) + \cdots + \frac{\partial g}{\partial x_m} f'_m(x) \end{aligned}$$

Esimerkki 134. Olkoon $h(t) = \sin \frac{e^t \ln^2 t}{t^2}$. Derivaattaa $h'(t)$ laskettaessa voidaan tietysti soveltaa tavallista ketjusääntöä tai vaihtoehtoisesti kirjoittaa $x = e^t$, $y = \ln^2 t$ ja $z = t^2$ ja laskea yhdistetyn funktion $h = f \circ g$ derivaatta, missä $g : \mathbb{R} \rightarrow \mathbb{R}^3$ on määritelty $g(t) = (e^t, \ln^2 t, t^2)$ ja $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ $f(x, y, z) = \sin(\frac{xy}{z})$ kuten esimerkissä 132. Esimerkissä 132 laskettiin, että

$$\nabla f = \left(\cos\left(\frac{xy}{z}\right) \cdot \frac{y}{z}, \cos\left(\frac{xy}{z}\right) \cdot \frac{x}{z}, -\cos\left(\frac{xy}{z}\right) \cdot \frac{xy}{z^2} \right),$$

jolloin

$$\begin{aligned} f'(t) &= \left(\cos\left(\frac{xy}{z}\right) \cdot \frac{y}{z}, \cos\left(\frac{xy}{z}\right) \cdot \frac{x}{z}, -\cos\left(\frac{xy}{z}\right) \cdot \frac{xy}{z^2} \right) \cdot (e^t, 2 \ln t \frac{1}{t}, 2t) \\ &= \cos\left(\frac{xy}{z}\right) \left(\frac{y}{z} \cdot e^t + \frac{x}{z} \cdot 2 \ln t \frac{1}{t} - \frac{xy}{z^2} \cdot 2t \right) \\ &= \cos\left(\frac{e^t \ln^2 t}{t^2}\right) \left(\frac{e^t \ln^2 t}{t^2} + \frac{2e^t \ln t}{t^3} - \frac{2e^t \ln^2 t}{t^3} \right) \\ &= \cos\left(\frac{e^t \ln^2 t}{t^2}\right) \frac{e^t \ln t}{t^2} \left(\ln t + \frac{2}{t} - \frac{2 \ln t}{t} \right). \end{aligned}$$

Esimerkki 135. Selvitetään, miten osittaisderivaatat $\frac{\partial f}{\partial x}$ ja $\frac{\partial f}{\partial y}$ voidaan määrittää osittaisderivaattojen $\frac{\partial f}{\partial r}$ ja $\frac{\partial f}{\partial \theta}$ avulla, kun $x = r \cos \theta$, $y = r \sin \theta$ on tavallinen napakoordinaattimuunnos. Lasketaan ensin (5.7):n mukaan

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial r} = \frac{\partial f}{\partial x} \cdot \cos \theta + \frac{\partial f}{\partial y} \cdot \sin \theta. \quad (5.9)$$

ja

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial \theta} = \frac{\partial f}{\partial x} \cdot (-r \sin \theta) + \frac{\partial f}{\partial y} \cdot r \cos \theta. \quad (5.10)$$

Näin saadusta yhtälöparista

$$\begin{cases} \cos \theta \cdot \frac{\partial f}{\partial x} + \sin \theta \cdot \frac{\partial f}{\partial y} = \frac{\partial f}{\partial r} \\ -r \sin \theta \cdot \frac{\partial f}{\partial x} + r \cos \theta \cdot \frac{\partial f}{\partial y} = \frac{\partial f}{\partial \theta} \end{cases}$$

voidaan ratkaista Gaussin menetelmää käyttäen $\frac{\partial f}{\partial x}$ ja $\frac{\partial f}{\partial x} \frac{\partial f}{\partial r}$:n ja $\frac{\partial f}{\partial \theta}$:n avulla:

$$\frac{\partial f}{\partial x} = \cos \theta \cdot \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \cdot \frac{\partial f}{\partial \theta} \quad (5.11)$$

$$\frac{\partial f}{\partial y} = \sin \theta \cdot \frac{\partial f}{\partial r} + \frac{\cos \theta}{r} \cdot \frac{\partial f}{\partial \theta}. \quad (5.12)$$

Esimerkki 136. Jos osittaisderivaatta $\frac{\partial f}{\partial x} : [a, b] \times [c, d] \rightarrow \mathbb{R}$ on jatkuva, niin

$$F(x) = \int_a^b f(x, t) dt$$

voidaan derivoida ”termeittäin”, ts. ”integraalimerkin alta”, kuten aiemmassa luvussa mainittiin:

$$F'(x) = \int_a^b \frac{\partial}{\partial x} f(x, t) dt.$$

Toisaalta taas analyysin peruslauseen (Insinöörimatematiikka IB) nojalla tiedetään, että

$$\frac{d}{dx} \int_a^x f(t) dt = f(x),$$

kun f on jatkuva funktio. Määritetään sitten x :n suhteen laskettu derivaatta

$$\frac{d}{dx} \int_{g(x)}^{h(x)} f(x, t) dt,$$

kun otaksutaan, että $f(x, t)$ on kummankin muuttujan suhteen jatkuva ja x :n suhteen derivoituva funktio. Merkitään ensin

$$I(x, g, h) = \int_g^h f(x, t) dt$$

jolloin $I : \mathbb{R}^3 \rightarrow \mathbb{R}$ riippuu muuttujista x , g ja h , mutta koska g ja h riippuvat x :stä, riippuu I viime kädessä pelkästään x :stä. Derivaatan $I'(x)$ laskemiseen voidaan soveltaa ketjusääntöä:

$$\begin{aligned} \frac{d}{dx} I(x, g, h) &= \frac{\partial I}{\partial x} \frac{\partial x}{\partial x} + \frac{\partial I}{\partial g} \frac{\partial g}{\partial x} + \frac{\partial I}{\partial h} \frac{\partial h}{\partial x} \\ &= \int_g^h \frac{\partial}{\partial x} f(x, t) dt - f(x, g)g'(x) + f(x, h)h'(x). \end{aligned}$$

Näin saatua yhtälöä

$$\frac{d}{dx} \int_{g(x)}^{h(x)} f(x, t) dt = \int_{g(x)}^{h(x)} \frac{\partial}{\partial x} f(x, t) dt + f(x, h(x))h'(x) - f(x, g(x))g'(x)$$

kutsutaan *Leibnizin säännöksi*.

Esimerkki 137. Olkoon

$$g(x) = \int_x^{x^2} \frac{\sin xt}{t} dt.$$

Leibnizin säännön mukaan

$$\begin{aligned}
 g'(x) &= \int_x^{x^2} \frac{\partial}{\partial x} \frac{\sin xt}{t} dt + \frac{\sin x \cdot x^2}{x^2} \cdot 2x - \frac{\sin x \cdot x}{x} \cdot 1 \\
 &= \int_x^{x^2} \cos xt dt + \frac{2 \sin x^3}{x} - \frac{\sin x^2}{x} \\
 &= \left[\frac{1}{x} \sin xt + \frac{2 \sin x^3}{x} - \frac{\sin x^2}{x} \right]_x^{x^2} \\
 &= \frac{\sin x^3 - \sin x^2}{x} + \frac{2 \sin x^3}{x} - \frac{\sin x^2}{x} = \frac{3 \sin x^3 - 2 \sin x^2}{x}.
 \end{aligned}$$

5.6 Usean muuttujan väliarvolause

Differentiaalilaskennan väliarvolauseen (Insinöörimatematiikka IB) mukaan derivoituvalle funktiolle f pätee

$$f(b) - f(a) = f'(\xi)(b - a),$$

missä $\xi \in (a, b)$. Väliarvolause voidaan yleistää usean muuttujan funktioille seuraavasti:

Lause 35 (Usean muuttujan väliarvolause). *Olkoon $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differentioituva funktio. Tällöin*

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\xi) \cdot (\mathbf{b} - \mathbf{a}),$$

missä ξ on jokin pisteitä $\mathbf{a}$ ja $\mathbf{b}$ yhdistävän janan $\{(1-t)\mathbf{a} + t\mathbf{b} \mid t \in [0, 1]\}$ piste.

Todistus. Merkitsemällä $g(t) = f((1-t)\mathbf{a} + t\mathbf{b})$ saadaan yhden muuttujan funktio $g : \mathbb{R} \rightarrow \mathbb{R}$, jolle $g(0) = f(\mathbf{a})$ ja $g(1) = f(\mathbf{b})$. Tavanomaisen väliarvolauseen mukaan välillä $(0, 1)$ on luku ξ , jolle pätee

$$f(\mathbf{b}) - f(\mathbf{a}) = g(1) - g(0) = g'(\xi)(1 - 0).$$

Tällöin väite saadaan yleistetystä ketjusäännöstä: yleistetyn ketjusäännön perusteella nimittäin

$$g'(\xi) = \nabla f(\xi) \cdot (\mathbf{b} - \mathbf{a}),$$

missä $\xi = (1 - \xi)\mathbf{a} + \xi\mathbf{b}$. $\square$

Väliarvolausesta voidaan soveltaa myös virheen arviointiin kuten kokonaisdifferentiaaliakin. Väliarvolauseen etuna on kuitenkin se, että tällöin approksimaativirheet voidaan arvioida kattavammin.

Esimerkki 138. Käsitellään esimerkin 133 funktiota $f(x, y) = xy$ ja sen approksimaatiota väliarvolauseen avulla. Merkitään suureiden x ja y mitattua arvoa $\mathbf{a}$:lla ($\mathbf{a} = (a_1, a_2)$) ja oikeaa arvoa $\mathbf{b}$:llä ($\mathbf{b} = (b_1, b_2)$), jolloin siis oikean ja mitatun arvon erotus voidaan väliarvolauseen mukaan esittää muodossa

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\xi) \cdot (\mathbf{b} - \mathbf{a}), \quad (5.13)$$

missä ξ on jokin pisteitä $\mathbf{a}$ ja $\mathbf{b}$ yhdistävän janan piste. Nyt $\nabla f(x, y) = (y, x)$ ja jos merkitään $\xi = (\xi_1, \xi_2)$, voidaan (5.13) kirjoittaa muotoon

$$f(\mathbf{b}) - f(\mathbf{a}) = (\xi_2, \xi_1) \cdot (b_1 - a_1, b_2 - a_2) = \xi_2(b_1 - a_1) + \xi_1(b_2 - a_2).$$

Oletuksen mukaan $|b_1 - a_1| \leq 0,1$ ja $|b_2 - a_2| \leq 0,2$ ja koska $\xi = (1-t)\mathbf{a} + t\mathbf{b}$, jollekin luvulle $t \in (0, 1)$. Tällöin myös $\xi_1 = (1-t)a_1 + tb_1$ ja $\xi_2 = (1-t)a_2 + tb_2$, mikä merkitsee, että $\xi_1 \in (a_1, b_1)$ ja $\xi_2 \in (a_2, b_2)$, jolloin siis

$$\begin{aligned}
 |f(\mathbf{b}) - f(\mathbf{a})| &\leq |\xi_2| |b_1 - a_1| + |\xi_1| |b_2 - a_2| \\
 &\leq \max\{|a_2|, |b_2|\} |b_1 - a_1| + \max\{|a_1|, |b_1|\} |b_2 - a_2|.
 \end{aligned} \quad (5.14)$$

Esimerkin 133 numeeriset arvot ovat $a_1 = 10$, $a_2 = 50$, sekä $|b_1 - a_1| \leq 0,1$ ja $|b_2 - a_2| \leq 0,2$. Tällöin $|b_1| \leq 10,1$ ja $|b_2| \leq 50,2$. Sijoittamalla nämä ylärajaan (5.14) saadaan

$$|f(\mathbf{b}) - f(\mathbf{a})| \leq 50,2 \cdot 0,1 + 10,1 \cdot 0,2 = 7,04.$$

Näin saatu arvio, joka ei lukuarvoltaan paljoakaan poikkea esimerkin (133) arviosta, poikkeaa kuitenkin esimerkiksi (133) periaatteellisesti, sillä tässä saatu yläraja on *ehdoton* yläraja mittausvääristymistä johtuvalle virheelle.

5.7 Ääriarvot, Lagrangen menetelmä

Aiemmin (Insinöörimatematiikka IB) on todettu, että suljetulla välillä I määritellyllä jatkuvalla reaalifunktiolla on olemassa sekä minimi että maksimi. Jos f on lisäksi derivoituva, löytyvät ääriarvot (maksimit tai minimi) kohdista, joissa derivaatan arvo on nolla *tai* välin päätepisteistä. Tässä luvussa nähdään, että samankaltainen tulos pätee myös differentioituville funktioille $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Tässä yhteydessä käsitellään myös hieman yleisempää ongelmaa, ns. *sidottujen ääriarvojen* määrittämistä. Sidottujen ääriarvojen vastine yhden muuttujan funktioiden teoriassa olisi etsiä ääriarvoja x -akselin rajatuilta osilta, mutta jo $\mathbb{R}^2$:ssa rajatut osat voivat olla huomattavasti monimuotoisemmat kuin reaaliksiakselilla, joten sidotuilla ääriarvoilla ei oikeastaan ole mielekäästä vastinetta yhden muuttujan funktioiden teoriassa.

Määritelmä 58. Jos f on funktion $\mathbb{R}^n \rightarrow \mathbb{R}$ ja on olemassa sellainen pisteen $\mathbf{a}$ ympäristö B , että $f(\mathbf{x}) \leq f(\mathbf{a})$, aina, kun $\mathbf{x} \in B$, sanotaan, että funktiolla f on *lokaali maksimi* pisteessä $\mathbf{a}$. Jos B voidaan valita funktion f koko määrittelyjoukoksi, sanotaan, että maksimi on *absoluuttinen* tai *globaali*. Samoin määritellään lokaali ja absoluuttinen minimi.

Lause 36. Oletetaan, että $\|\mathbf{u}\| = 1$ ja että funktiolla $f: \mathbb{R}^n \rightarrow \mathbb{R}$ on olemassa suunnattu derivaatta $D_{\mathbf{u}}f(\mathbf{x})$. Tällöin ääriarvopisteissä $\mathbf{x}$ pätee

$$D_{\mathbf{u}}f(\mathbf{x}) = 0.$$

Todistus. Merkitään $g(t) = f(\mathbf{x} + t\mathbf{u})$, jolloin siis funktiolla g on ääriarvo pisteessä $t = 0$. Tällöin siis $g'(0) = 0$ ja $D_{\mathbf{u}}f(\mathbf{x}) = g'(0) = 0$. $\square$

Edellisen lauseen mukaan funktion f ääriarvopisteissä *kaikki* suunnatut derivaatat ovat nollia. Koska osittaisderivaatat ovat suunnattujen derivaattojen erikoistapauksia saadaan seuraava lause.

Lause 37. Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ääriarvopisteissä $\mathbf{a}$ on

$$\nabla f(\mathbf{a}) = \mathbf{0}.$$

Edellisen lauseen mukaan siis funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ääriarvopisteet voivat esiintyä ainoastaan pisteissä, joissa gradientti on nollavektori – olettaen, että gradientti on olemassa. Käytännössä vastaan voi kuitenkin tulla tilanteita, joissa joukko, josta ääriarvoa etsitään ei ole mikään hyvänsä, vaan hyvinkin selvästi rajoitettu. Näiden ns. sidottujen ääriarvojen käsittelyä varten tarvitaan määritelmä *säännöllisestä pisteestä*.

Määritelmä 59. Olkoon $m < n$ ja $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funktio $g = (g_1, \dots, g_m)$, jonka jokaisella komponenttifunktiolla g_i on jatkuvat osittaisderivaatat. Sanotaan, että piste $\mathbf{x}$ on funktion g *säännöllinen piste*, mikäli Jacobin matriisin

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial g_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(\mathbf{x}) & \frac{\partial g_m}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial g_m}{\partial x_n}(\mathbf{x}) \end{pmatrix} \quad (5.15)$$

aste on m , toisin sanoen mikäli matriisin (5.15) rivit muodostavat lineaarisesti riippumattoman joukon. Vielä toisella tavalla ilmaistuna, piste $\mathbf{x}$ on funktion g säännöllinen piste, mikäli matriisi (5.15) on täysiasteinen.

Tapauksessa $m = 1$, jolloin g on funktio $\mathbb{R}^n \rightarrow \mathbb{R}$, merkitsee pisteen $\mathbf{x}$ säännöllisyys siis sitä, että $\nabla g(\mathbf{x}) \neq \mathbf{0}$.

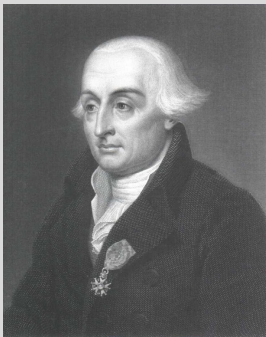
Ääriarvojen etsimisessä voidaan usein käyttää *Lagrangen menetelmää*, joka perustuu seuraavaan lauseeseen.

Lause 38. Olkoon $m < n$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ funktioita, joiden osittaisderivaatat ovat jatkuvia. Jos piste $\mathbf{a}$ on funktion f ääriarvo joukossa $g^{-1}(\mathbf{0}) = \{\mathbf{x} \mid g(\mathbf{x}) = \mathbf{0}\}$ ja $\mathbf{a}$ on lisäksi funktion g säännöllinen piste, niin tällöin on olemassa luvut $\lambda_1, \dots, \lambda_m$, joille pätee

$$\nabla(f(\mathbf{a}) + \lambda_1 g_1(\mathbf{a}) + \dots + \lambda_m g_m(\mathbf{a})) = \mathbf{0}.$$

Todistus. Vaikea, sivuutetaan suosiolla.

Taustatietoa



Joseph-Louis Lagrange (1736–1813) oli italialaissyntyinen matemaatikko ja astronomi. Hän oli yksi 1700-luvun merkittävimmistä tiedemiehistä ja kehitti ns. variaatiolaskentaa, menetelmän ääriarvojen löytämiseen, vakion variointi -menetelmän differentiaaliyhtälöille sekä lukuteoriaa. Hän kehitti myös Newtonin mekaniikkaa ja löysi stabiilit ratkaisut taivaanmekaniikassa tunnetulle ns. kolmen kappaleen ongelmalle. Hänen mukaansa on nimetty sijainnit, joista yhteen mm. Euroopan avaruusjärjestön Herschel-havaintoasema on lähetetty.

(kuva: Wikimedia Commons)

Huomautus 39. Edellisen lauseen lukuja $\lambda_1, \dots, \lambda_m$ kutsutaan *Lagrangen kertoimiksi*, ja joukosta $g^{-1}(\mathbf{0})$ löytyviä ääriarvoja *sidotuiksi* ääriarvoiksi. Tällöin kaikkia mahdollisia ääriarvoja rajoittavaa ehtoa $g(\mathbf{x}) = \mathbf{0}$ kutsutaan *side-ehdoksi*. Ääriarvot löytyvät tällöin niiden pisteiden $\mathbf{a} = (a_1, \dots, a_n)$ joukosta, jotka ovat joko g :n epäsäännöllisiä pisteitä tai jotka joillakin lukujen $\lambda_1, \dots, \lambda_m$ valinnoilla toteuttavat yhtälöryhmän

$$\begin{cases} g(\mathbf{a}) = \mathbf{0} \\ \nabla(f(\mathbf{a}) + \lambda_1 g_1(\mathbf{a}) + \dots + \lambda_m g_m(\mathbf{a})) = \mathbf{0}, \end{cases}$$

mikä voidaan yksityiskohtaisemmin kirjoittaa muotoon

$$\left\{ \begin{array}{l} g_1(\mathbf{a}) = 0 \\ g_2(\mathbf{a}) = 0 \\ \vdots \\ g_m(\mathbf{a}) = 0 \\ \frac{\partial}{\partial x_1} (f(\mathbf{a}) + \lambda_1 g_1(\mathbf{a}) + \dots + \lambda_m g_m(\mathbf{a})) = 0 \\ \frac{\partial}{\partial x_2} (f(\mathbf{a}) + \lambda_1 g_1(\mathbf{a}) + \dots + \lambda_m g_m(\mathbf{a})) = 0 \\ \vdots \\ \frac{\partial}{\partial x_n} (f(\mathbf{a}) + \lambda_1 g_1(\mathbf{a}) + \dots + \lambda_m g_m(\mathbf{a})) = 0 \end{array} \right.$$

Yhtälöryhmässä tuntemattomia ovat siis vektorin $\mathbf{a}$ koordinaatit $a_1, \dots, a_n$ sekä luvut $\lambda_1, \dots, \lambda_m$. Tuntemattomien määrä $n + m$ on siis sama kuin yhtälöiden määrä. Optimointitehtävän ratkaisemiseksi ei yleensä kuitenkaan ole tarpeen tietää lukuja $\lambda_1, \dots, \lambda_m$, joten ne kannattaa pyrkiä eliminoimaan.

Esimerkki 139. Funktiolla $f(x, y, z) = x^2 + x + y^2 + z^2$ ei ole suurinta arvoa, vaan helposti nähdään, että f saa miten suuria arvoja tahansa.

Tarkastellaan sitten f :n arvoja rajoitetussa joukossa. Valitaan side-ehdoiksi yksikköympyrä ja taso $x + y + z = 1$, jolloin siis tarkastellaan f :n ääriarvoja yksikköympyrän ja mainitun tason leikkauksessa. Side-ehdon määrääväksi funktioksi voidaan tällöin valita $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $g = (g_1, g_2)$, missä

$$\begin{aligned} g_1(x, y, z) &= x^2 + y^2 + z^2 - 1 \quad \text{ja} \\ g_2(x, y, z) &= x + y + z - 1. \end{aligned}$$

Tällä tavoin määritelty joukko on kompakti (miksi?), joten f saa suurimman ja pienimmän arvonsa tässä joukossa.

Ääriarvot löytyvät siis joko g :n epäsäännöllisistä pisteiden tai niiden pisteiden joukosta, jotka toteuttavat yhtälöryhmän

$$\left\{ \begin{array}{l} g_1(x, y, z) = 0 \\ g_2(x, y, z) = 0 \\ \frac{\partial}{\partial x} (f(x, y, z) + \lambda_1 g_1(x, y, z) + \lambda_2 g_2(x, y, z)) = 0 \\ \frac{\partial}{\partial y} (f(x, y, z) + \lambda_1 g_1(x, y, z) + \lambda_2 g_2(x, y, z)) = 0 \\ \frac{\partial}{\partial z} (f(x, y, z) + \lambda_1 g_1(x, y, z) + \lambda_2 g_2(x, y, z)) = 0, \end{array} \right.$$

mikä voidaan kirjoittaa muotoon

$$\left\{ \begin{array}{l} x^2 + y^2 + z^2 = 1 \\ x + y + z = 1 \\ 2x + 1 + 2\lambda_1 x + \lambda_2 = 0 \\ 2y + 2\lambda_1 y + \lambda_2 = 0 \\ 2z + 2\lambda_1 z + \lambda_2 = 0 \end{array} \right.$$

Kahdesta viimeisimmästä yhtälöstä seuraa $2(1 + \lambda_1)(y - z) = 0$, joten $y = z$ tai $\lambda_1 = -1$. Jälkimmäisessä tapauksessa neljännestä yhtälöstä seuraa $\lambda_2 = 0$ ja kolmannelta taas $\lambda_2 = -1$, mikä on ristiriita. Tällöin siis on oltava $y = z$, jolloin kaksi ensimmäistä yhtälöä saavat muodon

$$\left\{ \begin{array}{l} x^2 + 2y^2 = 1 \\ x + 2y = 1, \end{array} \right.$$

minkä ainoat ratkaisut ovat $(x, y) = (1, 0)$ ja $(x, y) = (-\frac{1}{3}, \frac{2}{3})$. Ainoat pisteet, joissa f :llä voi olla ehtojen $g_1(x, y, z) = 0$ ja $g_2(x, y, z) = 0$ sitoma ääriarvo, ovat siis $(1, 0, 0)$ ja $(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$ sekä g :n epäsäännölliset pisteet. Suoraan laskemalla saadaan $f(1, 0, 0) = 2$ ja $f(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}) = \frac{2}{3}$. Koska tarkasteltava joukko (yksikköpallon pinnan ja tason leikkaus) on kompakti, ovat ääriarvot joka tapauksessa olemassa, joten ne ovat näin löydettyissä pisteissä tai funktion g epäsäännöllisissä pisteissä.

Selvitetään sitten mahdolliset funktion g epäsäännölliset pisteet. Funktion g Jacobin matriisi on

$$\begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} & \frac{\partial g_1}{\partial z} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} & \frac{\partial g_2}{\partial z} \end{pmatrix} = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix},$$

jonka aste on pienempi kuin kaksi tarkalleen silloin kun $x = y = z$. Tällaisia pisteitä ei kuitenkaan ole joukossa, jossa $g_1(x, y, z) = 0$ ja $g_2(x, y, z) = 0$, joten funktion f g :n sitomiksi ääriarvopisteiksi jäävät löydettyt: maksimi 2 pisteessä $(1, 0, 0)$ ja minimi $\frac{2}{3}$ pisteessä $(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

5.8 Funktion ääriarvojen etsiminen yleisesti

Jos $A \subseteq \mathbb{R}^n$ on epätyhjä kompakti joukko ja $f: \mathbb{R}^n \rightarrow \mathbb{R}$ jatkuva funktio, on lauseen 27 mukaan f :llä on joukossa A sekä maksimi että minimi. Jos optimointitehtävän joukkoa ei voida määrittää luonteavasti side-ehdoilla, voidaan maksimi ja minimi määrätä 1) etsimällä ne pisteet, joissa f :n gradientti on nolla tai 2) etsimällä ääriarvot joukon A reunoilta.

Avaruuden $\mathbb{R}^n$ riittävän säännöllisen joukon reunat ovat ”alempiulotteisia” kuin joukko A ja tällöin ääriarvojen etsiminen reunoilta palautuu toisinaan optimointitehtävään, jossa optimoitava funktio on määritelty avaruudessa, jonka dimensio on pienempi kuin n .

Esimerkki 140. Määritetään funktion $f(x, y) = x^2 + y^2 - 6x$ suurin ja pienin arvo kompaktissa (miksi?) joukossa

$$A = \{(x, y) \mid -2 \leq x \leq 2, x^2 \leq y \leq 4\}$$

Jaetaan joukko A kolmeen osaan:

$$A_1 = \{(x, y) \mid -2 < x < 2, x^2 < y < 4\} \quad (\text{sisäosa}),$$

$$A_2 = \{(x, y) \mid -2 \leq x \leq 2, y = x^2\} \quad (\text{reuna}),$$

$$A_3 = \{(x, y) \mid -2 \leq x \leq 2, y = 4\} \quad (\text{reuna})$$

ja selvitetään aluksi missä f :n gradientti on nolla.

$$\nabla f(x, y) = (2x - 6, 2y) = (0, 0)$$

ainoastaa, jos $(x, y) = (3, 0)$, mutta tämä piste ei kuulu joukkoon A . Tällöin siis ääriarvot löytyvät joukon reunoilta. Todetaan, että

$$\{f(x, y) \mid (x, y) \in A_2\} = \{x^2 + x^4 - 6x \mid -2 \leq x \leq 2\}$$

ja että funktion $h_1(x) = x^2 + x^4 - 6x$ derivaatan ainoa nollakohta välillä $[-2, 2]$ on 1. Koska $h_1(1) = -4$, $h_1(-2) = 32$, $h_1(2) = 18$, on funktion f pienin arvo reunalla A_2 -4 ja suurin 32. Tarkastellaan sitten reunaa A_3 .

$$\{f(x, y) \mid (x, y) \in A_3\} = \{x^2 - 6x + 16 \mid -2 \leq x \leq 2\},$$

eikä funktiolla $h_2(x) = x^2 - 6x + 16$ ole derivaatan nollakohtia välillä $[-2, 2]$. Päätepisteissä saatavat arvot ovat $h_2(-2) = 32$ ja $h_2(2) = 18$. Täten funktion f pienin arvo reunalla A_3 on 18 ja suurin 32.

Kokoamalla näin saadut arvot yhteen voidaan todeta, että funktion f suurin arvo koko A :ssa on 32 ja pienin -4 .

Toisinaan voidaan myös päätellä, että funktio saa suurimman ja/tai pienimmän arvonsa joukossa A , vaikka A ei olisikaan kompakti. Esim. pienintä arvoa määrittäessä pyritään etsimään jokin kompakti osajoukko $A_1 \subseteq A$ ja löytämään A_1 :ssä funktion arvo, jota pienempää se ei voi saada A_1 :n ulkopuolelta.

Esimerkki 141. Suorakulmaisen särmiön muotoisen kannettoman astian tilavuudeksi halutaan 32. Selvitetään, miten seinien ja pohjan mittasuhteet tulee valita, jotta materiaalin tarve (kokonaispinta-ala) olisi mahdollisimman pieni.

Tätä varten merkitään pituutta, leveyttä ja korkeutta x :llä, y :llä ja z :lla, jolloin astian pinta-ala on $2xz + 2yz + xy$. Astian tilavuus on $xyz = 32$, joten $z = \frac{32}{xy}$, jolloin minimoitavaksi jää funktio $f(x, y) = 2x \cdot \frac{32}{xy} + 2y \cdot \frac{32}{xy} + xy = \frac{64}{y} + \frac{64}{x} + xy$.

Tässä optimointitehtävässä vaatimuksena on $x > 0$ ja $y > 0$, mutta näiden ehtojen määräämä neljännessä ei ole suljettu eikä rajoitettu joukko, eikä näin ollen siis myöskään kompakti. Ääriarvopisteissä on voimassa

$$\nabla f(x, y) = \left(-\frac{64}{x^2} + y, -\frac{64}{y^2} + x\right) = (0, 0),$$

Gradientin nollakohdat löytyvät siis yhtälöparin $y = \frac{64}{x^2}$, $x = \frac{64}{y^2}$ ratkaisujoukosta. Sijoittamalla x :n lauseke ensimmäiseen yhtälöön saadaan $y = \frac{y^4}{64}$, mistä saadaan $y^3 = 64$ ja siis $y = 4$. Tällöin $x = \frac{64}{4^2} = 4$ ja $z = \frac{32}{4 \cdot 4} = 2$.

Näin löydettyssä pisteessä $(x, y) = (4, 4)$ on itse asiassa funktion f absoluuttinen minimi $f(4, 4) = 48$ (neljännessä osassa $x > 0$, $y > 0$). Tätä voidaan perustella nojautuen geometriseen intuitioon tai seuraavasti:

Joukkoa $x > 0$, $y > 0$ voidaan rajoittaa siten että lähellä x - ja y -akseleita funktio saa varmasti suurempia arvoja kuin 48. Voidaan esimerkiksi todeta, että jos $0 < x < 1$, on

$$f(x, y) = \frac{64}{y} + \frac{64}{x} + xy \geq \frac{64}{x} \geq 64.$$

vastaavasti jos $0 < y < 1$, on myös $f(x, y) \geq 64$. Jos taas $x \geq 1$ ja $y \geq 64$, on

$$f(x, y) = \frac{64}{y} + \frac{64}{x} + xy \geq xy \geq y \geq 64$$

ja samoin tapauksessa $y \geq 1$ ja $x \geq 64$ $f(x, y) \geq 64$. Tällöin siis kompaktin (miksi) joukon

$$\{(x, y) \mid 1 \leq x, y \leq 64\}$$

joukon reunalla tai sen ulkopuolella funktio saa arvoja, jotka ovat suurempia kuin 64. Koska kompaktissa joukossa jatkuva funktio joka tapauksessa saa minimiarvonsa ja $f(4, 4) = 48$ on edellisen mukaan ainoa mahdollisuus, täytyy tämän olla funktion f absoluuttinen minimi koko neljännessä osassa $x > 0$, $y > 0$.

Esimerkki 142. Olkoon $a > 1$ ja tarkastellaan pintaa $z = e^{-x^2 - ay^2}$. Koska $x^2 + ay^2 \geq 0$, on pinnan korkein piste kohdassa $(x, y) = (0, 0)$. Selvitetään, missä pinnan pisteissä voidaan valita mahdollisimman jyrkkä laskusuunta. Tätä varten merkitään $f(x, y) = e^{-x^2 - ay^2}$ ja todetaan, että

$$\nabla f(x, y) = (-2xf(x, y), -2ayf(x, y)).$$

Pisteessä (x, y) jyrkin nousu tapahtuu gradientin suuntaan (vrt. seuraus 4).

Toisaalta suunnattu derivaatta gradientin suuntaisen yksikkövektorin suuntaan on (kun $\nabla f(x, y) \neq 0$)

$$\nabla f(x, y) \cdot \|\nabla f(x, y)\|^{-1} \nabla f(x, y) = \|\nabla f(x, y)\|.$$

Määritetään siis missä funktio

$$g(x, y) = \|\nabla f(x, y)\|^2 = (4x^2 + 4a^2y^2)(e^{-x^2 - ay^2})^2$$

saa suurimman arvonsa. Funktion g gradientin nollakohtien määrittämiseksi tulee selvittää yhtälöparin

$$\begin{cases} x - 2x^3 - 2a^2xy^2 = 0 \\ a^2y - 2ax^2y - 2a^3y^3 = 0 \end{cases}$$

ratkaisut, jotka ovat $(0, 0)$, $(0, \pm \frac{1}{\sqrt{2a}})$, $(\pm \frac{1}{\sqrt{2}}, 0)$. Laskemalla todetaan, että $g(0, 0) = 0$, $g(0, \pm \frac{1}{\sqrt{2a}}) = \frac{2a}{e}$ ja $g(\pm \frac{1}{\sqrt{2}}, 0) = \frac{2}{e}$. On mahdollista valita kompakti joukko, jonka ulkopuolella g saa vain näitä pienempiä arvoja (miksi?). Täten $\frac{2a}{e}$ on g :n suurin arvo ja siis pisteissä $(0, \pm \frac{1}{\sqrt{2a}})$ voidaan valita mahdollisimman jyrkkä laskusuunta.

Luku 6

Usean muuttujan funktioiden integraalilaskentaa

Riemann-integraalin määritelmä (Insinöörimatematiikka IB) voidaan yleistää suoraviivaisesti funktioille $\mathbb{R}^n \rightarrow \mathbb{R}$. Jos nimittäin

$$I = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$$

on avaruuden $\mathbb{R}^n$ väli (kts. määritelmä 45), voidaan aluksi määritellä välin I *geometrisen mittaa* (välin pituuden yleistys $\mathbb{R}^n$:ssä)

$$\mu(I) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \dots \cdot (b_n - a_n).$$

Tällä tavoin määritelty mitta esittää pinta-alaa avaruudessa $\mathbb{R}^2$ ja tilavuutta avaruudessa $\mathbb{R}^3$.

Kuten reaalfunktioiden tapauksessa, voidaan kukin väli $[a_i, b_i]$ jakaa osiin (kts Insinöörimatematiikka IB). Jos jokaiselle $i \in \{1, \dots, n\}$ D_i on välin $[a_i, b_i]$ jako, sanotaan, että $D = (D_1, D_2, \dots, D_n)$ on välin $I = [a_1, b_1] \times \dots \times [a_n, b_n]$ jako. Jos kussakin jaossa D_i on k_i välin $[a_i, b_i]$ osaväliä, määrää jako $D = (D_1, D_2, \dots, D_n)$ $k = k_1 \cdot k_2 \cdot \dots \cdot k_n$ kappaletta välin I osavälejä $I_1, \dots, I_k$, joiden unioni on I . Merkitään

$$m_i = \inf_{\mathbf{x} \in I_i} f(\mathbf{x}) \quad \text{ja} \quad M_i = \sup_{\mathbf{x} \in I_i} f(\mathbf{x})$$

Kutakin välin I jakoa D kohti määritellään *alasumma*

$$S_D = \sum_{i=1}^k m_i \cdot \mu(I_i)$$

sekä *yläsumma*

$$\bar{S}_D = \sum_{i=1}^k M_i \cdot \mu(I_i).$$

Samoin kuin reaalfunktioiden kohdalla, voidaan osoittaa, että kaikki alasummat ovat pienempiä kuin mikä hyvänsä yläsumma ja että jakoa tihennettäessä alasumma ei pienene eikä yläsumma kasva.

Määritelmä 60. Funktio $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on *Riemann-integroituva* välillä I , jos

$$\sup_D S_D = \inf_D \bar{S}_D,$$

missä D käy läpi kaikkien välin I jakojen. Tällöin sanotaan yläsummien infimumia (mikä siis on yhtäsuuri kuin alasummien supremum) funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ *Riemann-integraaliksi* yli välin I ja merkitään

$$\int_I f = \int_I f(x_1, \dots, x_n) dx_1 \dots dx_n = \int_I f(\mathbf{x}) d\mathbf{x} = \sup_D S_D = \inf_D \bar{S}_D.$$

Usein joudutaan kuitenkin tarkastelemaan integraalia yli sellaisen joukon $A \subseteq \mathbb{R}^n$, joka ei ole väli. Integrointi yli rajoitetun joukon A joka ei ole väli, käsitellään teoreettisesti määrittelemällä uusi

funktio $f_A : \mathbb{R}^n \rightarrow \mathbb{R}$ seuraavasti:

$$f_A(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{jos } \mathbf{x} \in A, \\ 0, & \text{muulloin,} \end{cases}$$

minkä jälkeen funktion f integraali yli joukon A määritellään valitsemalla jokin joukon A sisältävä väli I ja määrittelemällä

$$\int_A f = \int_I f_A.$$

Funktioiden $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integraaleille pätee samankaltainen tulos kuin reaalfunktioidenkin:

Lause 39. *Funktio $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on integroitava yli välin I tarkalleen silloin kun jokaista positiivilukua ε on olemassa Väin I jako D , jolle pätee*

$$\bar{S}_D - \underline{S}_D < \varepsilon.$$

Todistus. Helppo mutta pitkä, sivuutetaan.

Tunnetusti jatkuvat reaalfunktiot ovat integroituvia, kuin myös funktiot, joilla on vain äärellinen määrä epäjatkuvuuskohtia (Insinöörimatematiikka IA). Vastaava tulos on voimassa myös usean muuttujan funktioiden reaaliarvoisille funktioille, mutta tällöin joukko, jossa funktion epäjatkuvuuskohtien joukko voi olla monimutkaisempi. Tätä varten määritellään *nollamittaisuuden* käsite. Joukon A nollamittaisuus tarkoittaa sitä, että A voidaan peittää $\mathbb{R}^n$:n väleillä, joiden yhteenlaskettu mitta saadaan miten pieneksi hyvänsä. Tämä ilmaistaan täsmällisesti seuraavassa määritelmässä.

Määritelmä 61. Joukko $A \subseteq \mathbb{R}^n$ on *nollamittainen* (eli *mitätön* eli *Lebesgue-nollajoukko*), jos jokaista lukua $\varepsilon > 0$ on olemassa sellainen jono $I_1, I_2, I_3, \dots$ avaruuden $\mathbb{R}^n$ välejä, että

$$A \subseteq \bigcup_{n=1}^{\infty} I_n$$

ja

$$\sum_{n=1}^{\infty} \mu(I_n) < \varepsilon.$$

Esimerkki 143. Joukko $A \subseteq \mathbb{R}^n$ on numeroituva, jos on olemassa bijektio $f : \mathbb{N} \rightarrow A$, mikä puolestaan merkitsee sitä, että joukon A alkiot voidaan asettaa (päättymättömään) jonoon: $A = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots\}$. Tällöin kutakin pistettä $\mathbf{a} = (a_1, \dots, a_n)$ kohti voidaan asettaa väli $I = [a_1, a_1] \times \dots \times [a_n, a_n]$, joka sisältää vain pisteen $\mathbf{a}$ ja jonka mitta on nolla. Näin ollen numeroituva ja myös äärellinen joukko on nollamittainen.

Esimerkki 144. Avaruuden $\mathbb{R}^2$ suora $y = x$ on nollamittainen, sillä sen 1. neljännekseen kuuluva osa voidaan peittää väleillä $I_1 = [0, \varepsilon] \times [0, \varepsilon]$, $I_2 = [\varepsilon, \varepsilon + \frac{\varepsilon}{2}] \times [\varepsilon, \varepsilon + \frac{\varepsilon}{2}]$, $I_3 = [\varepsilon + \frac{\varepsilon}{2}, \varepsilon + \frac{\varepsilon}{2} + \frac{\varepsilon}{3}] \times [\varepsilon + \frac{\varepsilon}{2}, \varepsilon + \frac{\varepsilon}{2} + \frac{\varepsilon}{3}]$, Tällöin välit I_n todella peittävät koko puolisuoran, sillä $\varepsilon + \frac{\varepsilon}{2} + \frac{\varepsilon}{3} + \dots + \frac{\varepsilon}{n} = \varepsilon(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n})$ tulee miten suureksi tahansa kun n valitaan sopivasti (harmoninen sarja hajaantuu) olipa $\varepsilon > 0$ valittu miten hyvänsä. Toisaalta taas $\mu(I_1) = \varepsilon^2$, $\mu(I_2) = \frac{\varepsilon^2}{2^2}$, $\mu(I_3) = \frac{\varepsilon^2}{3^2}$, ..., josta välien yhteenlaskettu mitta saadaan laskettua:

$$\sum_{n=1}^{\infty} \frac{\varepsilon^2}{n^2} = \varepsilon^2 \frac{\pi^2}{6}.$$

Näin saatu lauseke saadaan miten pieneksi hyvänsä valitsemalla $\varepsilon > 0$ riittävän pieneksi.

Edellisen esimerkin nollamittainen joukko on hyvin tyypillinen: Yksiulotteisessa avaruudessa pisteet ovat nollamittaisia, kaksiulotteisessa pisteiden lisäksi suorat, kolmiulotteisessa avaruudessa myös tasot ovat nollamittaisia. Yleisesti ottaen avaruudessa $\mathbb{R}^n$ kaikki ”alempiulotteiset” joukot ovat

nollamittaisia. Näiden ”alempiulotteisten” joukkojen ei kuitenkaan tarvitse olla lineaarisia osajoukkoja, vaan myös ”alempiulotteiset pinnat” ovat nollamittaisia. Seuraava lause ilmaisee tämän hieman täsmällisemmin.

Lause 40. Olkoon $m < n$, $U \subseteq \mathbb{R}^m$ ja $f : U \rightarrow \mathbb{R}^n$ differentioituva kuvaus, jonka kaikki osittaisderivaatat ovat jatkuvia. Tällöin $f(U)$ on nollamittainen.

Todistus. Sivuutetaan.

Esimerkki 145. Avaruuden $\mathbb{R}^n$ mikä hyvänsä suora $\{\mathbf{r} + t\mathbf{s} \mid t \in \mathbb{R}\}$ voidaan esittää reaaliakselin (yksiulotteisen avaruuden) $\mathbb{R}$ kuvana differentioituvassa kuvauksessa $f(x) = \mathbf{r} + t\mathbf{s}$ (osittaisderivaatat ovat vakioita ja täten jatkuvia), jolloin suorat ovat nollamittaisia, kun $n \geq 2$.

Esimerkki 146. suorakaiteen $[0, 2\pi) \times [0, \pi)$ kuva kuvauksessa

$$f(\theta, \phi) = (\cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi)$$

on yksikköpallon pinta. Koska kuvauksen f osittaisderivaatat ovat jatkuvia, on yksikköpallon pinta on nollamittainen avaruuden $\mathbb{R}^3$ joukko.

Lause 41 (Lebesgue). Olkoon I avaruuden $\mathbb{R}^n$ kompakti väli ja funktio $f : I \rightarrow \mathbb{R}$ rajoitettu funktio. Funktio f on Riemann-integroituva yli välin I tarkalleen silloin, kun joukko

$$\{\mathbf{x} \in I \mid f \text{ on epäjatkuva } \mathbf{x}: \text{ssä}\}$$

on nollamittainen.

Todistus. Sivuutetaan.

Taustatietoa



Henri Lebesgue (1875–1941) oli ranskalainen matemaatikko, joka väitöskirjatyössään 1902 kehitti hyvin yleisen integraalikäsitteen. Insinöörimatematiikan kursseissa käytetty Riemann-integraali on erikoistapaus Lebesguen integraalista.

(kuva: Wikimedia Commons)

Yksi tärkeä Riemann-integraalin sovelluksista on mitan käsitteen yleistäminen avaruuden $\mathbb{R}^n$ välejä yleisemmille joukoille. Tämä merkitsee samalla *tilavuuden* käsitteen yleistämistä.

Määritelmä 62. Jos vakiofunktio $f(\mathbf{x}) = 1$ on integroituva yli avaruuden $\mathbb{R}^n$ joukon A , sanotaan, että A on *Jordan-mitallinen* ja integraalin arvoa kutsutaan joukon A *Jordan-mitaksi*. Ja merkitään

$$\mu(A) = \int_A d\mathbf{x}.$$

Tapauksessa $n = 1$ puhutaan *pituudesta*, $n = 2$ *pinta-alasta* ja tapauksessa $n = 3$ *tilavuudesta*.

Usean muuttujan funktioiden integraaleille pätee samankaltainen tulos kuin reaalifunktiollekin.

Lause 42. Jos $A \subseteq \mathbb{R}^n$ on rajoitettu ja $f, g : A \rightarrow \mathbb{R}$ ovat integroituvia, on

- $\int_A (\alpha f + \beta g) = \alpha \int_A f + \beta \int_A g$,
- Jos $f(\mathbf{x}) \leq g(\mathbf{x})$ kaikilla $\mathbf{x} \in A$, niin $\int_A f \leq \int_A g$,
- Funktio $|f|$ on integroituva ja $\left| \int_A f \right| \leq \int_A |f|$.

Todistus. Samoin kuin reaalifunktioiden tapauksessa, väitetyt ominaisuudet pätevät äärellisille summille. Ominaisuuksien periytyminen äärellisiltä summilta integraaleille on näytettävissä toteen periaatteessa melko yksinkertaisesti, mutta yksityiskohdat vaativat työtä, joka sivuutetaan. $\square$

Myös joukko, jonka yli integroidaan, voidaan pilkkoa osiin samoin kuin reaalifunktioiden tapauksessa.

Lause 43. Oletetaan, että jossakin avaruuden $\mathbb{R}^n$ osajoukossa määritelty reaaliarvoinen funktio f on integroituva yli joukkojen $A_1, \dots, A_k$, joille $A_i \cap A_j$ on nollamittainen aina, kun $i \neq j$. Tällöin f on integroituva yli joukon $A = A_1 \cup \dots \cup A_k$ ja

$$\int_A f = \int_{A_1} f + \dots + \int_{A_k} f.$$

Todistus. Sivuutetaan.

6.1 Iteroidut integraalit

Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ integraali lasketaan usein palauttamalla integrointi n -ulotteisesta avaruudesta matalampiulotteiseen tapaukseen. Tällaista palauttamisprosessia kutsutaan *iteroinniksi*.

Palauttamista varten tarkastellaan vektorin $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ hajotelmaa $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$, missä $\mathbf{x}_1 = (x_1, \dots, x_{n_1}) \in \mathbb{R}^{n_1}$ (ensimmäiset n_1 koordinaattia) ja $\mathbf{x}_2 = (x_{n_1+1}, \dots, x_n) \in \mathbb{R}^{n_2}$ (lopun n_2 koordinaattia) ja $n = n_1 + n_2$. Tällöin myös avaruuden $\mathbb{R}^n$ suljettu väli I voidaan kirjoittaa muotoon $I_1 \times I_2$, missä I_1 on avaruuden $\mathbb{R}^{n_1}$ ja I_2 avaruuden $\mathbb{R}^{n_2}$ suljettu väli.

Lause 44 (Fubini). Jos $I = I_1 \times I_2$ kuten yllä ja $f : I \rightarrow \mathbb{R}$ integroituva funktio. Jos lisäksi jokaista $\mathbf{x} \in I_1$ kohti funktio $\mathbf{y} \mapsto f(\mathbf{x}, \mathbf{y})$ on integroituva välin I_2 yli, on funktio

$$\mathbf{x} \mapsto \int_{I_2} f(\mathbf{x}, \mathbf{y}) d\mathbf{y}$$

integroituva yli välin I_1 ja

$$\int_I f = \int_{I_1} \left(\int_{I_2} f(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) d\mathbf{x}$$

Todistus. Perustuu suoraan Riemann-integraalin määritelmään, jätetään harjoitustehtäväksi. Huomautetaan lisäksi, että edellisessä lauseessa voidaan $\mathbf{x}$:n ja $\mathbf{y}$:n roolit tietenkin vaihtaa keskenään. $\square$

Erikoistapauksessa, jossa f voidaan kirjoittaa muotoon $f(\mathbf{x}, \mathbf{y}) = f_1(\mathbf{x})f_2(\mathbf{y})$, missä f_1 ei riipu $\mathbf{y}$:stä eikä f_2 $\mathbf{x}$:stä, saadaan tulos, jota myös kutsutaan Fubinin lauseeksi.

Lause 45. Olkoot I , I_1 , I_2 ja f kuten lauseessa 44. Jos $f(\mathbf{x}, \mathbf{y}) = f_1(\mathbf{x})f_2(\mathbf{y})$, missä f_1 on integroituva yli välin I_1 ja f_2 yli I_2 :n, on

$$\int_I f = \int_{I_1} f_1(\mathbf{x}) d\mathbf{x} \int_{I_2} f_2(\mathbf{y}) d\mathbf{y}.$$

Soveltamalla $n - 1$ kertaa lausetta 44 voidaan integraali n -ulotteisessa avaruudessa palauttaa yhden muuttujan funktion integroimiseksi, mikäli tarkasteltava funktio on riittävän säännöllinen. Lebesguen lauseen (lause 41) mukaan esimerkiksi jatkuville funktioille voidaan näin tehdä.

Esimerkki 147. Jos $I = [0, 1] \times [0, 2] \times [0, 3]$ ja $f(x, y, z) = x^2 + y^2 + z^2$, niin

$$\begin{aligned} \int_I f &= \int_{z=0}^3 \int_{y=0}^2 \int_{x=0}^1 (x^2 + y^2 + z^2) dx dy dz \\ &= \int_{z=0}^3 \int_{y=0}^2 \left/ \int_0^1 \left(\frac{1}{3}x^3 + xy^2 + xz^2 \right) dy dz \right. \\ &= \int_{z=0}^3 \int_{y=0}^2 \left(\frac{1}{3} + y^2 + z^2 \right) dy dz \\ &= \int_0^3 \left/ \int_0^2 \left(\frac{1}{3}y + \frac{1}{3}y^3 + yz^2 \right) dz \right. \\ &= \int_0^3 \left(\frac{10}{3} + 2z^2 \right) dz \\ &= \left/ \int_0^3 \left(\frac{10}{3}z + \frac{2}{3}z^3 \right) dz \right. = 10 + 2 \cdot 9 = 28. \end{aligned}$$

Myös väliä yleisemmän joukon iteroitu integraali voidaan laskea samalla periaatteella. Tällöin on tosin huomattava joukon asettamat rajoitukset integraalien ylä- ja alarajoille. Edellisessä esimerkissä integrointimuuttujien järjestyksen vaihtaminen ei oleellisesti muuta tehtävää helpommaksi toisin kuin monissa muissa tapauksissa.

Esimerkki 148. Määritetään sen kappaleen A tilavuus, jota rajoittavat ehdot $0 \leq x \leq 1$, $1 \leq y \leq 2$ sekä $z \geq 0$ ja $z = 4 - x - y$. Tilavuudeksi saadaan

$$\begin{aligned} &\int_{y=1}^2 \int_{x=0}^1 (4 - x - y) dx dy \\ &= \int_{y=1}^2 \left/ \int_0^1 \left(4x - \frac{1}{2}x^2 - xy \right) dy \right. \\ &= \int_{y=1}^2 \left(\frac{7}{2} - y \right) dy = \left/ \int_1^2 \left(\frac{7}{2}y - \frac{1}{2}y^2 \right) dy \right. = 2. \end{aligned}$$

Esimerkki 149. Lasketaan funktion $f(x, y) = xy$ integraali yli kolmion, jonka kärkinä ovat pisteet $(0, 0)$, $(1, 0)$ ja $(1, 1)$. Kolmio voidaan kattaa, kun x käy välin $[0, 1]$ (ulompi integraali) ja y kutakin x :n arvoa kohti välin $[0, x]$ (sisempi integraali). Tällöin saadaan

$$\int_0^1 \int_0^x xy \, dx \, dy = \int_0^1 \left[\frac{1}{2} x^2 y \right]_0^x dy = \int_0^1 \frac{1}{2} x^3 dy = \left[\frac{1}{8} x^4 \right]_0^1 = \frac{1}{8}.$$

Esimerkki 150. Lasketaan paraboloidin $z = 4 - x^2 - y^2$ ja tason $z = 4 - 2x$ rajoittaman joukon A tilavuus.

Kuvion piirtämisen perusteella selviää, että kysytyn kappaleen sisällä x -koordinaatti vaihtelee välillä $[0, 2]$ (uloin integraali). Kutakin x :n arvoa kohti y -koordinaatti vaihtelee myös tietyn välin, jonka selvittämiseksi eliminoidaan muuttuja z : yhtälöistä nimittäin seuraa, että $4 - x^2 - y^2 = 4 - 2x$, siis $y^2 = 2x - x^2$. Tämä on xy -tasossa ympyrän yhtälö, keskipisteenä $(1, 0)$ ja säteenä 1 ; tämän ympyrän kautta kulkeva z -akselin suuntainen sylinteri sisältää paraboloidin ja tason leikkauskäyrän. Tällöin siis kutakin x :n arvoa kohti y :n annetaan kulkea väli $[-\sqrt{2x - x^2}, \sqrt{2x - x^2}]$ (sisempi integraali). Lopuksi z :n annetaan kulkea tason ja paraboloidin väli $[4 - 2x, 4 - x^2 - y^2]$. Kysytty tilavuus saadaan siis integraalista

$$\begin{aligned} \mu(A) &= \int_{x=0}^2 \int_{y=-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} \int_{z=4-2x}^{4-x^2-y^2} dz \, dy \, dx \\ &= \int_{x=0}^2 \int_{y=-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} (4 - x^2 - y^2 - (4 - 2x)) \, dy \, dx \\ &= \int_{x=0}^2 \int_{y=-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} (2xy - x^2y - \frac{1}{3}y^3) \, dy \, dx \\ &= \frac{4}{3} \int_0^2 (2x - x^2)^{\frac{3}{2}} \, dx. \end{aligned}$$

Näin saatu integraali muuntuu sijoituksella $t = x - 1$ muotoon

$$\frac{4}{3} \int_{-1}^1 (1 - t^2)^{\frac{3}{2}} \, dt,$$

mikä saadaan edelleen sijoituksella $t = \sin \theta$ muotoon

$$\frac{4}{3} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^4 \theta \, d\theta = \frac{8}{3} \int_0^{\frac{\pi}{2}} \cos^4 \theta \, d\theta = \frac{\pi}{2}.$$

Viimeisimmän integraalin arvo saadaan osittaisintegroimalla useita kertoja tai suoraan kaavakokoelmasta.

6.2 Sijoitus integraaliin

Seuraava lause vastaa yhden muuttujan funktion integraaliin sijoittamista.

Lause 46. Olkoot U ja V ovat rajoitettuja, avoimia avaruuden $\mathbb{R}^n$ Jordan-mitallisia joukkoja ja $g : U \rightarrow V$ jatkuva bijektio. Jos lisäksi g :n osittaisderivaatat ovat jatkuvia ja kaikissa pisteissä $\mathbf{x} \in U$ funktion g Jacobin matriisin

$$Dg(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial g_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1}(\mathbf{x}) & \frac{\partial g_n}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial g_n}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

determinantti $\det(J_g(\mathbf{x}))$ (ns. Jacobin determinantti) on nollasta poikkeava, on

$$\int_V f(\mathbf{y}) d\mathbf{y} = \int_U f(g(\mathbf{x})) |\det(J_g(\mathbf{x}))| d\mathbf{x}.$$

Todistus. Vaativa, sivuutetaan.

Esimerkki 151. Olkoon $g(r, \theta) = (r \cos \theta, r \sin \theta)$ napakoordinaateista xy -koordinaatteihin välittävä muunnos ja $U = (0, a) \times (0, 2\pi)$ sekä $V = \{(x, y) \mid x^2 + y^2 < a^2\} \setminus \{(x, y) \mid x \geq 0, y = 0\}$, jolloin $g : U \rightarrow V$ on bijektio; kumpikin joukoista U ja V kuvaa siis origokeskistä a -säteistä ympyrää, josta on poistettu ei-negatiivisen x -akselin osuus. Funktion g Jacobin matriisi on

$$Dg(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

ja tämän determinanti on $\det(Dg(r, \theta)) = r \cos^2 \theta + r \sin^2 \theta = r$. Tämä on nollasta poikkeava koko tarkastelualueella, joten tällöin avaruuden $\mathbb{R}^2$ jossakin osajoukossa määritellyn jatkuvan reaalifunktion integraali voidaan laskea sijoituksella

$$\begin{aligned} \int_V f(x, y) dx dy &= \int_U f(r \cos \theta, r \sin \theta) r dr d\theta \\ &= \int_0^{2\pi} \int_0^a f(r \cos \theta, r \sin \theta) r dr d\theta \end{aligned}$$

On siis huomattava, että sijoitettaessa napakoordinaattiesitys $x = r \cos \theta$, $y = r \sin \theta$ paitsi integrointialue muuttuu V :stä U :ksi, $dx dy$:n sijaan kirjoitetaan $r dr d\theta$. Tämä on analoginen yhden muuttujan integraaleihin tehtävän sijoituksen $x = g(t)$ kanssa, jolloin dx :n sijaan kirjoitetaan $g'(t) dt$. Usean muuttujan funktioiden tapauksessa derivaattaa $g'(t)$ vastaa sijoitusfunktion g Jacobin matriisin $J_g(\mathbf{x})$ determinantin itseisarvo $|\det(J_g(\mathbf{x}))|$.

Esimerkki 152. Pallokoordinaattisijoitus. Funktion

$$(r, \theta, \phi) \mapsto (r \cos \theta \sin \phi, r \sin \theta \sin \phi, r \cos \phi) = (x, y, z)$$

Jacobin matriisi on

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos \theta \sin \phi & -r \sin \theta \sin \phi & r \cos \theta \cos \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi & r \sin \theta \cos \phi \\ \cos \phi & 0 & -r \sin \phi \end{pmatrix},$$

jonka determinantiksi saadaan viimeisen rivin mukaan kehittämällä

$$\begin{aligned} & \cos \phi \begin{vmatrix} -r \sin \theta \sin \phi & r \cos \theta \cos \phi \\ r \cos \theta \sin \phi & r \sin \theta \cos \phi \end{vmatrix} - r \sin \phi \begin{vmatrix} \cos \theta \sin \phi & -r \sin \theta \sin \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi \end{vmatrix} \\ &= r^2 \cos^2 \theta \sin \phi \begin{vmatrix} -\sin \theta & r \cos \theta \\ \cos \theta & r \sin \theta \end{vmatrix} - r^2 \sin^3 \phi \begin{vmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{vmatrix} \\ &= r^2 \cos^2 \phi \sin \phi (-\sin^2 \theta - \cos^2 \theta) - r^2 \sin^3 \phi (\cos^2 \theta + \sin^2 \theta) \\ &= -r^2 \sin \phi (\cos^2 \phi + \sin^2 \phi) = -r^2 \sin \phi. \end{aligned}$$

Kun $0 \leq \phi \leq \pi$, saadaan Jacobin matriisin determinantin itseisarvoksi $r^2 \sin \phi$. Täten funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ integrointi yli R -säteisen pallon $\bar{B}(\mathbf{0}, R)$ voidaan suorittaa seuraavasti:

$$\begin{aligned} & \int_{\bar{B}(\mathbf{0}, R)} f(x, y, z) dx dy dz \\ &= \int_{\phi=0}^{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^R f(r \cos \theta \sin \phi, r \sin \theta \sin \phi, r \cos \phi) r^2 \sin \phi dr d\theta d\phi. \end{aligned}$$

Esimerkki 153. Lasketaan integraalin $I = \int_0^\infty e^{-x^2} dx$ arvo. Integrandille e^{-x^2} ei löydy antiderivaattaa suljetussa muodossa, mutta Fubinin lauseen perusteella (itse asiassa sen yleistyksen perusteella)

$$I^2 = \int_0^\infty e^{-x^2} dx \int_0^\infty e^{-y^2} dy = \int_0^\infty \int_0^\infty e^{-x^2-y^2} dx dy. \quad (6.1)$$

Kun tähän tehdään napakoordinaattisijoitus $x = r \cos \theta$, $y = r \sin \theta$, muuttuvat integroimisrajat: $r : 0 \rightarrow \infty$ ja $\theta : 0 \rightarrow \frac{\pi}{2}$, sekä $dx dy$ korvataan $r dr d\theta$:lla ja siis integraali (6.1) muuttuu muotoon (huomaa että napakoordinaatistossa $x^2 + y^2 = r^2$)

$$\int_{\theta=0}^{\frac{\pi}{2}} \int_{r=0}^{\infty} e^{-r^2} r dr d\theta.$$

Tällöin nimittäin integrandiksi on saatu $e^{-r^2} r$, minkä antiderivaatta r :n suhteen on helppo löytää, koska $\frac{d}{dr} e^{-r^2} = -2r e^{-r^2}$. Sisempi integraali saadaan tämän perusteella lasketuksi:

$$\int_0^\infty e^{-r^2} r dr = - \int_0^\infty \frac{1}{2} e^{-r^2} = \frac{1}{2}.$$

Tällöin

$$\int_{\theta=0}^{\frac{\pi}{2}} \int_{r=0}^{\infty} e^{-r^2} r dr d\theta = \int_0^{\frac{\pi}{2}} \frac{1}{2} d\theta = \frac{\pi}{4}.$$

Koska $I^2 = \frac{\pi}{4}$ ja integrandi e^{-x^2} on koko reaaliakselilla positiivinen, on myös integraalin arvon oltava positiivinen. Täten $I = \frac{\sqrt{\pi}}{2}$.

Esimerkki 154. Selvitä harjoitustehtävänä, miten integraalin

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

arvosta voidaan selvittää $\Gamma(\frac{1}{2})$.

Esimerkki 155. Määritetään sen kappaleen tilavuus, jota rajoittavat yksikköpallo $x^2 + y^2 + z^2 = 4$ ja sylinteri $x^2 + y^2 - 2y = 0$. Vaihdettaessa napakoordinaatistoon saadaan pallon yhtälöksi $r^2 + z^2 = 4$ ja sylinterin yhtälöksi $r = 2 \sin \theta$. Kokonaistilavuus on neljä kertaa sen osan tilavuus, jota lisäksi rajoittavat ehdot $x \geq 0$, $y \geq 0$, $z \geq 0$. Tällöin tilavuus on siis

$$\begin{aligned} & 4 \int_{\theta=0}^{\frac{\pi}{2}} \int_{r=0}^{2 \sin \theta} \int_{z=0}^{\sqrt{4-r^2}} dz r dr d\theta \\ &= 4 \int_{\theta=0}^{\frac{\pi}{2}} \int_{r=0}^{2 \sin \theta} \sqrt{4-r^2} r dr d\theta \end{aligned}$$

Näin saatuun integraaliin sijoitus $u = 4 - r^2$ antaa

$$2 \int_{\theta=0}^{\frac{\pi}{2}} \int_{u=4 \cos^2 \theta}^4 \sqrt{u} du d\theta = \frac{4}{3} \int_0^{\frac{\pi}{2}} (8 - 8 \cos^3 \theta) d\theta,$$

johon sijoittamalla $v = \sin \theta$ saadaan

$$\frac{16}{3} \pi - \frac{32}{3} \int_0^1 (1 - v^2) dv = \frac{16}{3} \pi - \frac{64}{9}.$$

Esimerkki 156. Funktioiden f ja g Laplace-muunnokset ovat

$$F(s) = \int_0^\infty f(t) e^{-st} dt \quad \text{ja} \quad G(s) = \int_0^\infty g(t) e^{-st} dt.$$

Näiden tulo voidaan kirjoittaa muotoon

$$\begin{aligned} F(s)G(s) &= \int_0^\infty f(x) e^{-sx} dx \int_0^\infty g(y) e^{-sy} dy \\ &= \int_0^\infty \int_0^\infty f(x) g(y) e^{-s(x+y)} dx dy. \end{aligned}$$

Tason $\mathbb{R}^2$ ensimmäinen neljännes voidaan käydä läpi kuten ylläolevassa integraalissa: Kutakin x :n arvoa $x \in [0, \infty)$ kohti y käy läpi puolisuoran $[0, \infty)$. Tason 1. neljännes voidaan käydä läpi myös siten, että kutakin suoran $y = x$ pistettä (u, u) kohti kuljetaan tätä suoraa vastaan kohtisuorassa oleva jana $y = -x + 2u$ pisteestä $(2u, 0)$ pisteeseen $(0, 2u)$. Tätä vastaa sijoitus $x = u + v$ ja $y = u - v$ (jolloin $u = \frac{1}{2}(x + y)$ ja $v = \frac{1}{2}(x - y)$) ja Jacobin matriisin determinantiksi saadaan

$$\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 2.$$

Tällöin integraali saadaan muotoon

$$\int_{u=0}^{\infty} \int_{v=-u}^u f(u+v)g(u-v)e^{-s \cdot 2u} 2 dv du,$$

joka sisempään integraaliin sijoittamalla $v = w - u$ muuttuu muotoon

$$\int_{u=0}^{\infty} \int_{w=0}^{2u} f(w)g(2u-w)e^{-s \cdot 2u} 2 dw du.$$

Kun vielä ulompaan integraaliin tehdään sijoitus $u = \frac{1}{2}t$, saadaan integraali muotoon

$$\int_{t=0}^{\infty} \int_{w=0}^t f(w)g(t-w) dw e^{-st} dt.$$

Näin saatu integraali voidaan tulkita funktion

$$(f * g)(t) = \int_0^t f(w)g(t-w) dw$$

Laplace-muunnokseksi.

Luku 7

Vektorianalyysiä

Vektorianalyysillä tarkoitetaan differentiaali- ja integraalilaskennan soveltamista vektoriarvoisiin funktioihin ja oikeastaan vektorianalyysistä on nähty jo esimerkkejä aiemminkin mm. gradienteista ja suunnatuista derivaatoista puhuttaessa. Vektorianalyysissä käytetään sekä differentiaali- ja integraalilaskennan että lineaarialgebran käsitteistöä. Perinteinen vektorianalyysi keskittyy kuitenkin lähinnä kolmiulotteisen avaruuden $\mathbb{R}^3$ ominaisuuksiin.

7.1 Vektoriarvoisen funktion derivointi

Ajan suhteen muuttuva vektori voidaan mallintaa kuvauksella $f : \mathbb{R} \rightarrow \mathbb{R}^3$, jolloin siis $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t)) \in \mathbb{R}^3$ on tarkasteltava vektori ajanhetkellä t . Aiemmin on todettu, että mikäli komponenttifunktiot f_1, f_2, f_3 , ovat jatkuvia, on funktiolla f selkeä geometrinen merkitys: kun t käy läpi jonkin reaalilukuvälin, piirtää $\mathbf{f}(t)$ avaruuteen $\mathbb{R}^3$ käyrän. Jos lisäksi komponenttifunktiot ovat derivoituvia, sanotaan, että funktion f määräämä käyrä on *sileä*. Tässä luvussa tarkastellaan yleensä tilanteita, joissa lisäksi komponenttifunktioiden derivaatat ovat jatkuvia, jolloin sanotaan, että funktion f määräämä käyrä on *jatkuvasti derivoituva*.

Hieman yleisemmin voidaan tarkastella vektoriarvoista funktiota $\mathbb{R} \rightarrow \mathbb{R}^n$, jonka Jacobin matriisi on helppo määrittää: Kyseessä on $n \times 1$ -matriisi (pystyvektori) $(f'_1(t), f'_2(t), \dots, f'_n(t))^T$. Tällöin funktion f Fréchet'n derivaatta pisteessä t on lineaarikuvaus $\mathbb{R} \rightarrow \mathbb{R}^n$, joka kuvaa luvun h vektoriksi

$$(f'_1(t)h, f'_2(t)h, \dots, f'_n(t)h),$$

josta käytetään merkintää $h\mathbf{f}'(t)$. Kyseinen lineaarikuvaus on siis avaruuden $\mathbb{R}^n$ origon kautta kulkeva suora, jonka suuntavektori on $\mathbf{f}'(t)$. Koska

$$\mathbf{f}(t+h) - \mathbf{f}(t) = h\mathbf{f}'(t) + |h|\varepsilon(h),$$

voidaan todeta että suora $\mathbf{f}(t) + h\mathbf{f}'(t)$ on funktion f paras mahdollinen lineaarinen approksimaatio pisteessä t . Suoraa

$$\mathbf{r} + h\mathbf{s}, \quad h \in \mathbb{R}$$

missä paikkavektori $\mathbf{r} = \mathbf{f}(t)$ ja suuntavektori $\mathbf{s} = \mathbf{f}'(t)$ kutsutaankin funktion f määräämän käyrän *tangentiksi* pisteessä t . Vektoriarvoisen funktion $f : \mathbb{R} \rightarrow \mathbb{R}^n$ derivointi komponentteittain tuottaa siis f :n piirtämän käyrän tangentin suuntavektorin.

7.2 Vektorifunktion käyräintegraali

Klassisessa mekaniikassa työn käsite ymmärretään tulona Fs , missä s merkitsee matkaa, jonka kappale on siirtynyt voiman F vaikutuksesta voiman F suuntaan. Jos kuitenkin voima F ja sen suunta vaihtelee kappaleen siirron ajan, pitää mekaanisen työn käsite määritellä laajemmin.

Yksinkertaisessa tapauksessa, jossa voima ei vaihtele ajan mukana, pitäisi mekaanisen työn lausekkeen olla muotoa $\|\mathbf{F}_s\| \cdot \|s\|$, missä $\mathbf{F}_s$ merkitsee voiman $\mathbf{F}$ komponenttia liikesuunnassa s .

Komponentin $\mathbf{F}_s$ suuruus projektiona on selvitetty aiemmin (Insinöörimatematiikka IIA)

$$\mathbf{F}_s = \frac{\mathbf{F} \cdot \mathbf{s}}{\|\mathbf{s}\|^2} \mathbf{s},$$

missä $\mathbf{F} \cdot \mathbf{s}$ tarkoittaa vektoreiden pistetuloa. Tällöin

$$\|\mathbf{F}_s\| = \frac{|\mathbf{F} \cdot \mathbf{s}|}{\|\mathbf{s}\|^2} \|\mathbf{s}\|,$$

mistä nähdään, että

$$\|\mathbf{F}_s\| \|\mathbf{s}\| = |\mathbf{F} \cdot \mathbf{s}|. \quad (7.1)$$

jos halutaan sisällyttää työn määritelmään mukaan myös se mahdollisuus, että vektoreiden $\mathbf{F}$ ja $\mathbf{s}$ kulma on suurempi kuin 90° , voidaan itseisarvomerkki poistaa lausekkeesta (7.1) ja todeta, että mikäli voima $\mathbf{F}$ siirtää kappaletta vektorin $\mathbf{s}$ verran, on tehty työ $\mathbf{F} \cdot \mathbf{s}$.

Yleistettäessä mekaanisen työn käsitettä tapaukseen, jossa kappaleen kulkusuunta ja voiman suuruus vaihtelee käytetään tässä yhteydessä Newtonin ja Leibnizin ”infinitesimaalisia” suureita. Ajateltaessa, että voima $\mathbf{F}$ siirtää kappaletta ”äärettömän” pienen vektorin $d\mathbf{s}$ verran, on tehty työ $dW = \mathbf{F} \cdot d\mathbf{s}$, ja kokonaistyö saadaan integroimalla ”äärettömän” pienet työt dW yhteen:

$$W = \int_{\gamma} dW = \int_{\gamma} \mathbf{F} \cdot d\mathbf{s}. \quad (7.2)$$

Integrointi tapahtuu yli käyrän γ , jonka kappale kulkee voiman $\mathbf{F}$ vaikuttaessa siihen. Integraalia (7.2) kutsutaan *vektorikentän $\mathbf{F} = (F_1, F_2, F_3)$ käyräintegraaliksi yli käyrän γ* .

Integraali (7.2) tulkitaan siten, että kappale, johon voima vaikuttaa, kulkee yli käyrän γ . Jos merkitään $\mathbf{s} = (x, y, z)$, saadaan käyrän tangentiksi edellisen pykälän mukaan $\frac{d\mathbf{s}}{dt} = (x'(t), y'(t), z'(t))$, ja muodollisesti voidaan merkitä $d\mathbf{s} = (x'(t) dt, y'(t) dt, z'(t) dt)$. Jos vielä merkitään $x'(t) dt = dx$, $y'(t) dt = dy$ ja $z'(t) dt = dz$, saadaan integraali muotoon

$$W = \int_{\gamma} dW = \int_{\gamma} \mathbf{F} \cdot d\mathbf{s} = \int_{\gamma} F_1 \cdot dx + F_2 \cdot dy + F_3 \cdot dz \quad (7.3)$$

Käyräintegraali (7.2) voidaan laskea esittämällä käyrä $\gamma: [a, b] \rightarrow \mathbb{R}^3$ koordinaattifunktioiden avulla. Olkoon $\gamma(a)$ käyrän alkupiste, $\gamma(b)$ sen loppupiste ja $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, käyrän γ koordinaattifunktiot. Näillä merkinnöillä $\mathbf{s} = \gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$ on kappaleen paikka ajanhetkellä t ja

$$d\mathbf{s} = (\gamma_1'(t) dt, \gamma_2'(t) dt, \gamma_3'(t) dt),$$

jolloin

$$\mathbf{F} \cdot d\mathbf{s} = \mathbf{F}(\gamma(t)) \cdot d\mathbf{s} = (F_1(\gamma(t))\gamma_1'(t) + F_2(\gamma(t))\gamma_2'(t) + F_3(\gamma(t))\gamma_3'(t)) dt$$

ja integraali (7.3) saadaan muotoon

$$W = \int_{\gamma} \mathbf{F} \cdot d\mathbf{s} = \int_a^b (F_1(\gamma(t))\gamma_1'(t) + F_2(\gamma(t))\gamma_2'(t) + F_3(\gamma(t))\gamma_3'(t)) dt \quad (7.4)$$

Kolmiulotteista tapausta yleisempi määritelmä on suoraviivaista esittää.

Määritelmä 63 (Vektorifunktion käyräintegraali). Olkoon $\gamma: [a, b] \rightarrow \mathbb{R}^n$ sileä käyrä (komponenttifunktioiden derivaatat jatkuvia) ja $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\mathbf{f} = (f_1, \dots, f_n)$. Tällöin funktion $\mathbf{f}$ käyräintegraali pitkin käyrää γ on

$$\int_{\gamma} \mathbf{f} \cdot d\mathbf{x} = \int_{\gamma} f_1 dx_1 + \dots + f_n dx_n = \int_a^b \mathbf{f}(\gamma(t)) \cdot \gamma'(t) dt.$$

Käyräintegraalin määrittelevässä lausekkeessa $\mathbf{f}(\gamma(t)) \cdot \gamma'(t)$ on muodoltaan samanlainen kuin määrättyyn integraaliin sijoitus, mikä antaa muistisäännön käyräintegraalin laskemiseksi.

Esimerkki 157. Olkoon $\gamma: [-1, 1] \rightarrow \mathbb{R}^2$ käyrä $\gamma(t) = (x(t), y(t)) = (t^2, t^3)$ ja $f(x, y) = (x^2 + y, x + y^3)$. Lasketaan funktion f käyräintegraali käyrää γ pitkin.

Koska käyrällä γ on $x = x(t) = t^2$, on käyrällä $dx = 2t dt$ ja samoin $y = y(t) = t^3$ sekä $dy = 3t^2 dt$. Tällöin

$$\begin{aligned} \int_{\gamma} \mathbf{f} \cdot d\mathbf{x} &= \int_{\gamma} (x^2 + y) dx + (x + y^3) dy \\ &= \int_{-1}^1 ((t^2)^2 + t^3) \cdot 2t dt + (t^2 + (t^3)^3) \cdot 3t^2 dt \\ &= \int_{-1}^1 (3t^{11} + 2t^5 + 5t^4) dt = \int_{-1}^1 \left(\frac{1}{4}t^{12} + \frac{1}{3}t^6 + t^5 \right) dt = 2 \end{aligned}$$

Esimerkki 158. Olkoon $\gamma(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$, jolloin $\gamma(0) = (1, 0) = \gamma(2\pi)$, siis γ on sulkeutuva käyrä (origokeskinen yksikköympyrä). Tällöin

$$\begin{aligned} &\int_{\gamma} \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy \\ &= \int_0^{2\pi} \frac{-\sin t}{\cos^2 t + \sin^2 t} (-\sin t) dt + \frac{\cos t}{\cos^2 t + \sin^2 t} (\cos t) dt \\ &= \int_0^{2\pi} dt = 2\pi. \end{aligned}$$

Käyräintegraalin määritelmästä seuraa melko suoraan seuraava lause:

Lause 47. Käyräintegraaleille pätee

1. Jos käyrän γ_1 päätepiste on sama kuin käyrän γ_2 alkupiste, ja yhdistettyä käyrää $\gamma_1 \gamma_2$ merkitään γ :lla, on

$$\int_{\gamma} \mathbf{f} \cdot d\mathbf{x} = \int_{\gamma_1} \mathbf{f} \cdot d\mathbf{x} + \int_{\gamma_2} \mathbf{f} \cdot d\mathbf{x}.$$

2. Merkitään $-\gamma$:lla käyrää, joka saadaan kulkemalla γ lopusta alkuun. Tällöin

$$\int_{-\gamma} \mathbf{f} \cdot d\mathbf{x} = - \int_{\gamma} \mathbf{f} \cdot d\mathbf{x}$$

- 3.

$$\int_{\gamma} (\alpha \mathbf{f} + \beta \mathbf{g}) \cdot d\mathbf{x} = \alpha \int_{\gamma} \mathbf{f} \cdot d\mathbf{x} + \beta \int_{\gamma} \mathbf{g} \cdot d\mathbf{x}$$

Jatkossa otaksutaan, että U on avaruuden $\mathbb{R}^n$ käyräyhtenäinen joukko, ts. sellainen joukko jonka mitkä tahansa kaksi pistettä voidaan yhdistää joukkoon U sisältyvällä käyrällä.

Määritelmä 64. Olkoon $U \subseteq \mathbb{R}^n$. Vektorianalyysissä funktiota $U \rightarrow \mathbb{R}^n$, $f = (f_1, \dots, f_n)$ kutsutaan *vektorikentäksi*. Jos on olemassa funktio $V : U \rightarrow \mathbb{R}$ jolle pätee $f = \nabla V$ joukossa U , sanotaan, että V on vektorikentän f *potentiaali*. Tällöin sanotaan myös, että *differentiaalimuoto* $f_1 dx_1 + \dots + f_n dx_n$ on *eksakti* ja että sen *integraalifunktio* (joukossa U) on V .

Lause 48. Oletetaan, että $f_1 dx_1 + \dots + f_n dx_n$ on eksakti alueessa U ja että $f = \nabla V$. Jos $\gamma : [a, b] \rightarrow U$ on alueeseen U sisältyvä sileä käyrä jonka derivaatta on jatkuva ja jolle $\gamma(a) = \mathbf{a}$ ja $\gamma(b) = \mathbf{b}$, on

$$\int_{\gamma} f_1 dx_1 + \dots + f_n dx_n = V(\mathbf{b}) - V(\mathbf{a}). \quad (7.5)$$

Todistus. Oletus $f = \nabla V$ merkitsee sitä, että $f_i(\mathbf{x}) = \frac{\partial V}{\partial x_i}$. Tällöin

$$\begin{aligned} \int_{\gamma} f_1 dx_1 + \dots + f_n dx_n &= \int_a^b \mathbf{f}(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_a^b \nabla V(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b (V \circ \gamma)'(t) dt \\ &= \int_a^b (V \circ \gamma)(t) dt = V(\gamma(b)) - V(\gamma(a)) = V(\mathbf{b}) - V(\mathbf{a}). \end{aligned}$$

□

Huomautus 40. Edellinen lause sanoo, että mikäli vektorikentällä on alueessa U potentiaali (jolloin sanotaan myös, että vektorikenttä on *konservatiivinen*), riippuu käyräintegraalin arvo vain käyrän päätepisteistä, eikä käyrästä itsestään. Esimerkiksi Newtonin mekaniikassa maan gravitaation määrittämä voimakenttä on konservatiivinen.

Huomautus 41. Huomaa yhtäläisyys kaavan (7.5) ja Newton-Leibnizin säännön

$$\int_a^b f(t) dt = F(b) - F(a),$$

missä $F'(t) = f(t)$ välillä. Molemmat kaavat ovat itse asiassa erikoistapauksia ns. abstraktista Stokesin lauseesta, jonka erikoistapauksia on vektorianalyysissä lukuisia muitakin.

Jos vektorikentän $f = (f_1, \dots, f_n)$ potentiaalin V toisen kertaluvun osittaisderivaatat ovat jatkuvia, on $\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 V}{\partial x_j \partial x_i} = \frac{\partial^2 V}{\partial x_i \partial x_j} = \frac{\partial f_j}{\partial x_i}$. Tietyin edellytyksin ehto $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ kaikille i, j on myös *riittävä* potentiaalin olemassaololle.

Sanotaan, että alue U on *tähtimäinen*, jos alueessa U on piste $\mathbf{p}$, josta voidaan piirtää alueeseen U sisältyvä jana mihin tahansa U :n pisteeseen. On mahdollista todistaa seuraava lause:

Lause 49. Jos U on $\mathbb{R}^n$:n tähtimäinen alue ja $f = (f_1, \dots, f_n)$ vektorikenttä, jonka osittaisderivaatat ovat jatkuvia alueessa U . Jos $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ kaikille i, j , niin f :llä on potentiaali U :ssa.

Todistus. Siivutetaan.

Esimerkki 159. Selvitä, miksi esimerkin 158 funktiolla ei ole potentiaalia tarkastelualueessa.

Esimerkki 160. Selvitetään, onko vektorikentällä $f(x, y) = (f_1, f_2) = (2x + 3y, 3x + 4y)$ potentiaalia $\mathbb{R}^2$:n jossakin alueessa. Todetaan ensiksi, että $\frac{\partial f_1}{\partial y} = 3 = \frac{\partial f_2}{\partial x}$ ja että $\mathbb{R}^2$ on tähtimäinen alue. Näin ollen vektorikentällä on $\mathbb{R}^2$:ssa potentiaali $V(x, y)$, ja tämän määrittämiseksi ratkaistaan osittaisdifferentiaaliyhtälöt $\frac{\partial V}{\partial x} = f_1$ ja $\frac{\partial V}{\partial y} = f_2$. Osittaisdifferentiaaliyhtälöstä

$$\frac{\partial V}{\partial x} = 2x + 3y$$

seuraa, että $V(x, y) = x^2 + 3xy + h(y)$, missä $h(y)$ on vain y :stä riippuva funktio. Derivoimalla näin saatu lauseke y :n suhteen todetaan, että

$$3x + 4y = \frac{\partial V}{\partial y} = 3x + h'(y),$$

joten $h'(y) = 4y$, mistä $h(y) = 2y^2 + C$. Täten potentiaali on

$$V(x, y) = x^2 + 3xy + 2y^2 + C,$$

missä C on vakio.

Esimerkki 161. Laskettaessa edellisen tehtävän vektorikentän käyräintegraalia $\int_{\gamma} \mathbf{f} \cdot d\mathbf{x}$ yli käyrän, joka alkaa pisteestä $(0, 0)$ ja päättyy pisteeseen $(2, 1)$ ei käyrää tarvitse määritellä erikseen: potentiaalin olemassaolosta seuraa, että käyräintegraalin arvo riippuu vain päätepisteistä ja

$$\int_{\gamma} \mathbf{f} \cdot d\mathbf{x} = V(2, 1) - V(0, 0) = 12 + C - C = 12.$$

7.3 Skalaarifunktion käyräintegraali

Edellä on puhuttu vain vektoriarvoisen funktion käyräintegraaleista, kun myös skalaariarvoisen funktion käyräintegraali on mahdollista määritellä luontevasti. Tarkastellaan aluksi käyräintegraalin määrittelevää lauseketta

$$\int_{\gamma} \mathbf{f} \cdot d\mathbf{x} = \int_a^b \mathbf{f}(\gamma(t)) \cdot \gamma'(t) dt$$

ja merkitään

$$\mathbf{T} = \|\gamma'(t)\|^{-1} \gamma'(t),$$

jolloin $\mathbf{T}$ on käyrän tangentin suuntainen yksikkövektori. Jos lisäksi merkitään $\|\gamma'(t)\| dt = ds$, voidaan käyräintegraali kirjoittaa muotoon

$$\int_{\gamma} \mathbf{f} \cdot d\mathbf{x} = \int_a^b \mathbf{f} \cdot \mathbf{T} ds = \int_a^b \mathbf{f} \cdot \mathbf{T} \|\gamma'(t)\| dt.$$

Koska $\mathbf{T}$ on yksikkövektori, merkitsee yllä oleva lauseke sitä, että vektorifunktion käyräintegraalissa itse asiassa integroidaan vektorifunktion $\mathbf{f}$ käyrän tangentin suuntaisen komponentin skalaarikerrointa ja lisäksi mukaan on otettu painokerroin $\|\gamma'(t)\|$.

Määritelmä 65. Skalaariarvoisen funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ käyräintegraali käyrän γ yli määritellään lausekkeella

$$\int_{\gamma} f ds = \int_a^b f(\gamma(t)) \|\gamma'(t)\| dt,$$

missä $\gamma(a)$ on käyrän alkupiste ja $\gamma(b)$ sen päätepiste.

Tässä yhteydessä ei käsitellä skalaariarvoisten funktioiden käyräintegraaleja sen tarkemmin, vaan todetaan, että skalaarifunktion $f = 1$ käyräintegraali tuottaa lausekkeen

$$\int_a^b \|\gamma'(t)\| dt,$$

joka on täsmälleen sama kuin aiemmin saatu lauseke käyrän pituudelle (Insinöörimatematiikka IB).

7.4 Pinta-ala

Pinta avaruudessa $\mathbb{R}^3$ on tason $\mathbb{R}^2$ alueen A kuva jatkuvassa kuvauksessa $r : A \rightarrow \mathbb{R}^3$ (tarkemmin sanoen kyseessä on pinnan parametriesitys). Pinta on *sileä*, mikäli f on differentioituva. Tässä yhteydessä tarkastellaan lähinnä sellaisia pintoja $r : A \rightarrow \mathbb{R}^3$, missä r :n osittaisderivaatat ovat jatkuvia.

Esimerkki 162. Suorakaide $[0, 2\pi] \times [0, \pi]$ kuvautuu R -säteiseksi, origokeskisen pallon pinnaksi kuvauksessa $\mathbf{r}(\theta, \phi) = (R \cos \theta \sin \phi, R \sin \theta \sin \phi, R \cos \phi)$.

Esimerkki 163. Yhtälön $z = f(u, v)$ määrittämälle pinnalle saadaan parametriesitys $\mathbf{r}(u, v) = (u, v, f(u, v))$.

Esimerkki 164. Ellipsoidin $(\frac{x}{a})^2 + (\frac{y}{b})^2 + (\frac{z}{c})^2 = 1$ puolikkaalle $z \geq 0$ saadaan $z = c \sqrt{1 - (\frac{x}{a})^2 - (\frac{y}{b})^2}$ ja tästä edelleen parametriesitys

$$\mathbf{r}(x, y) = (x, y, c \sqrt{1 - (\frac{x}{a})^2 - (\frac{y}{b})^2}).$$

Esimerkki 165. Edellisen esimerkin ellipsoidille saadaan myös parametriesitys

$$\mathbf{r}(\theta, \phi) = (a \cos \theta \sin \phi, b \sin \theta \sin \phi, c \cos \phi),$$

missä $(\theta, \phi) \in [0, 2\pi] \times [0, \pi]$.

Selvitetään aluksi, miten pinnan $\{\mathbf{r}(u, v) \mid (u, v) \in A\}$ ala määrätään. Tätä varten huomataan että koska $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta$ (insinöörimatematiikka IIA), missä θ on vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välinen kulma, merkitsee $\|\mathbf{x} \times \mathbf{y}\|$ vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ määrittämän suunnikkaan pinta-alaa. Funktion r Jacobin matriisi on

$$\begin{pmatrix} \frac{\partial r_1}{\partial u} & \frac{\partial r_1}{\partial v} \\ \frac{\partial r_2}{\partial u} & \frac{\partial r_2}{\partial v} \\ \frac{\partial r_3}{\partial u} & \frac{\partial r_3}{\partial v} \end{pmatrix},$$

joten funktion r Fréchet'n derivaatta pisteessä on kuvaus

$$(h_1, h_2) \mapsto h_1 \mathbf{s}_1 + h_2 \mathbf{s}_2,$$

missä

$$\mathbf{s}_1 = \left(\frac{\partial r_1}{\partial u}, \frac{\partial r_2}{\partial u}, \frac{\partial r_3}{\partial u} \right) \quad \text{ja} \quad \mathbf{s}_2 = \left(\frac{\partial r_1}{\partial v}, \frac{\partial r_2}{\partial v}, \frac{\partial r_3}{\partial v} \right)$$

ovat pinnan tangenttitason suuntavektorit pisteessä $\mathbf{r}(u_0, v_0)$.

Huomautus 42. Tason $z = f(x, y)$ parametriesityksessä $\mathbf{r}(u, v) = (u, v, f(u, v))$ saadaan tangenttitason suuntavektoreiksi

$$\mathbf{s}_1 = (1, 0, \frac{\partial f}{\partial u}(u_0, v_0)) \quad \text{ja} \quad \mathbf{s}_2 = (0, 1, \frac{\partial f}{\partial v}(u_0, v_0))$$

ja paikkavektoriksi (u_0, v_0, z_0) , missä $z_0 = f(u_0, v_0)$. Selvitä miten tästä saadaan huomautuksessa 33 mainittu muoto.

Mikäli muutosvektori (du, dv) on ”äärettömän” pieni, siirryttäessä pisteestä (u, v) pisteeseen $(u + du, v + dv)$ kuvaa funktion r muutosta $\mathbf{r}(u + du, v + dv) - \mathbf{r}(u, v)$ lauseke $du \mathbf{s}_1 + dv \mathbf{s}_2$ ($\mathbf{s}_1$ ja $\mathbf{s}_2$ ovat kuten edellä) ja suorakaide $[u, u + du] \times [v, v + dv]$ kuvautuu siis suorakaiteeksi, jonka sivuina ovat vektorit $\mathbf{s}_1 du$ ja $\mathbf{s}_2 dv$. Tämän suorakaiteen pinta-ala on $\|\mathbf{s}_1 du \times \mathbf{s}_2 dv\|$, mikä voidaan kirjoittaa myös muotoon

$$\left\| \left(\frac{\partial r_1}{\partial u}, \frac{\partial r_2}{\partial u}, \frac{\partial r_3}{\partial u} \right) \times \left(\frac{\partial r_1}{\partial v}, \frac{\partial r_2}{\partial v}, \frac{\partial r_3}{\partial v} \right) \right\| du dv = \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| du dv.$$

näin ollen pinnan $\{\mathbf{r}(u, v) \mid (u, v) \in A\}$ alaa esittää integraali

$$A = \iint_{(u,v) \in A} \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| du dv.$$

Tuota pikaa nähdään, että lauseke $\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\|$ toimii kertoimena, jolla kerrottuna saadaan uv -tason $du \cdot dv$ -alaisen suorakulmion alasta tämän yläpuolella tai (alapuolella) olevan pinnan S tangenttitasolla olevan suunnikkaan ala. Tämä vuoksi on erityisesti fysiikan oppikirjoissa tapana merkitä $\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| du dv = dA$, jolloin pinta-alan määrittelevä integraali voidaan kirjoittaa muotoon

$$A = \iint_{(u,v) \in A} dA.$$

Huomautus 43. Jos pinta on annettu muodossa $z = f(x, y)$, voidaan parametriesitykselle $\mathbf{r}(u, v) = (u, v, f(u, v))$ laskea suoraan (miten?)

$$\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| = \sqrt{\left(\frac{\partial f}{\partial u} \right)^2 + \left(\frac{\partial f}{\partial v} \right)^2 + 1}.$$

Esimerkki 166. Määritetään sen pinnan ala, jonka sylinteri $x^2 + y^2 - y = 0$ leikkaa yksikköpalloa $x^2 + y^2 + z^2 = 1$ ylemmästä puolikkaasta ($z > 0$). Sylinterin yhtälö voidaan kirjoittaa muotoon $x^2 + y^2 - 2\frac{1}{2}y + \frac{1}{4} = \frac{1}{4} \iff x^2 + (y - \frac{1}{2})^2 = (\frac{1}{2})^2$, jolloin siis sylinterin ja xy -tason leikkaus on $\frac{1}{2}$ -säteinen ympyrä A , jonka keskipiste on $(0, \frac{1}{2})$. Tarkasteltavaa pintaa esittää siis funktio $f(x, y) = \sqrt{1 - x^2 - y^2}$, missä x ja y kulkevat yli ympyrän A .

Näillä valinnoilla $\frac{\partial f}{\partial u} = -\frac{u}{\sqrt{1-u^2-v^2}}$ ja $\frac{\partial f}{\partial v} = \frac{v}{\sqrt{1-u^2-v^2}}$, jolloin

$$\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| = \sqrt{\left(\frac{\partial f}{\partial u} \right)^2 + \left(\frac{\partial f}{\partial v} \right)^2 + 1} = \frac{1}{1 - u^2 - v^2}.$$

kysytty pinta-ala on siis

$$A = \iint_{(u,v) \in A} \frac{1}{\sqrt{1-u^2-v^2}} du dv. \quad (7.6)$$

Symmetrian nojalla voidaan integroida yli ensimmäisen neljänneksen ja siis

$$A = 2 \int_{v=0}^1 \int_{u=0}^{\sqrt{u-u^2}} \frac{1}{\sqrt{1-u^2-v^2}} du dv. \quad (7.7)$$

Integraali (7.7) puolestaan voidaan määrittää napakoordinaattimuunnoksella: Kun $u = r \cos \theta$ ja $v = r \sin \theta$, on sen puoliympyrän, jonka yli integroidaan, reunan yhtälö $r = \sin \theta$ (miksi?). Tällöin integraali (7.7) saadaan muotoon

$$\begin{aligned} A &= 2 \int_0^{\frac{\pi}{2}} \int_0^{\sin \theta} \frac{r}{\sqrt{1-r^2}} dr d\theta = 2 \int_0^{\frac{\pi}{2}} \left/ \begin{matrix} \sin \theta \\ 0 \end{matrix} \right. (-\sqrt{1-r^2}) d\theta \\ &= -2 \int_0^{\frac{\pi}{2}} (\sqrt{1-\sin^2 \theta} - 1) d\theta = 2 \int_0^{\frac{\pi}{2}} (1 - \cos \theta) d\theta \\ &= \left/ \begin{matrix} \frac{\pi}{2} \\ 0 \end{matrix} \right. 2(\theta - \sin \theta) = 2\left(\frac{\pi}{2} - 1\right) = \pi - 2. \end{aligned}$$

Huomautus 44. Lausekkeelle $\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\|$ on mahdollista löytää myös esitysasua, joka on käyttökelpoinen, vaikka parametriesitystä $\mathbf{r}(u, v)$ ei olisikaan mahdollista ratkaista u :n ja v :n lausekkeena. Saattaa nimittäin olla, että kolmen muuttujan funktio $G: \mathbb{R}^3 \rightarrow \mathbb{R}$ ja yhtälö $G(u, v, w) = 0$ määrittelevät funktion $w = f(u, v)$, vaikka f :n lausekkeen löytäminen olisikin toivotonta.

Tällöin itse asiassa saadaan funktio $\mathbb{R}^2 \rightarrow \mathbb{R}$ yhdistämällä kuvaukset

$$(u, v) \mapsto (u, v, f(u, v)) \mapsto G(u, v, f(u, v)).$$

Ehdon $G(u, v, f(u, v)) = 0$ toteutuessa on kyseinen funktio vakio ja siksi myös sen Fréchet'n derivaatta on nollafunktio. Fréchet'n derivaatta voidaan laskea ketjusäännön avulla: Ulkofunktion derivaatta $(\nabla G = (\frac{\partial G}{\partial u}, \frac{\partial G}{\partial v}, \frac{\partial G}{\partial w}))$ arvolla sisäfunktion $(u, v) \mapsto (u, v, f(u, v))$ derivaatta. Matriisiesitys yhdistetyn funktion derivaatalle on siis 1×2 -matriisi

$$\left(\frac{\partial G}{\partial u}, \frac{\partial G}{\partial v}, \frac{\partial G}{\partial w}\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix} = \left(\frac{\partial G}{\partial u} + \frac{\partial G}{\partial w} \frac{\partial f}{\partial u}, \frac{\partial G}{\partial v} + \frac{\partial G}{\partial w} \frac{\partial f}{\partial v}\right),$$

ja koska näin saatu lineaarikuvaus on nollakuvaus, voidaan todeta, että

$$\frac{\partial f}{\partial u} = -\frac{\frac{\partial G}{\partial u}}{\frac{\partial G}{\partial w}} \quad \text{ja} \quad \frac{\partial f}{\partial v} = -\frac{\frac{\partial G}{\partial v}}{\frac{\partial G}{\partial w}}.$$

Tällöin

$$\begin{aligned} \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2 + 1 &= \left(\frac{\frac{\partial G}{\partial u}}{\frac{\partial G}{\partial w}}\right)^2 + \left(\frac{\frac{\partial G}{\partial v}}{\frac{\partial G}{\partial w}}\right)^2 + \left(\frac{\frac{\partial G}{\partial w}}{\frac{\partial G}{\partial w}}\right)^2 \\ &= \frac{1}{\left(\frac{\partial G}{\partial w}\right)^2} \left(\left(\frac{\partial G}{\partial u}\right)^2 + \left(\frac{\partial G}{\partial v}\right)^2 + \left(\frac{\partial G}{\partial w}\right)^2 \right) = \frac{1}{\left(\frac{\partial G}{\partial w}\right)^2} \|\nabla G\|^2 \end{aligned}$$

ja

$$\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| = \sqrt{\left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2 + 1} = \frac{1}{\left|\frac{\partial G}{\partial w}\right|} \|\nabla G\|.$$

Sovellettaessa näin saatua esitystä esimerkkiin 166 voidaan valita $G(u, v, w) = u^2 + v^2 + w^2 - 1$ ja tästä saadaan $\frac{\partial G}{\partial w} = 2w$ sekä $\|\nabla G\| = \|(2u, 2v, 2w)\| = 2\sqrt{u^2 + v^2 + w^2} = 2$, jolloin siis

$$\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| = \frac{1}{w} = \frac{1}{\sqrt{1 - u^2 - v^2}}.$$

7.5 Skalaarifunktion pintaintegraali

Määritelmä 66. Olkoon U avaruuden $\mathbb{R}^3$ alue ja $f : A \rightarrow \mathbb{R}$ reaaliarvoinen funktio ja $S = \{\mathbf{r}(u, v) \mid (u, v) \in A\}$ pinta avaruudessa $\mathbb{R}^3$. Tällöin integraalia

$$\iint_{(u,v) \in A} f(\mathbf{r}(u, v)) \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| du dv$$

sanotaan funktion f integraaliksi yli pinnan S .

Merkitsemällä $\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| du dv = dA$ voidaan pintaintegraali kirjoittaa muotoon

$$\iint_{(u,v) \in A} f(\mathbf{r}(u, v)) dA.$$

Esimerkki 167. Etäisyydellä r pyörimisakselista olevan pistemäisen massan hitausmomentti on mr^2 . Yleisesti kappaleen hitausmomentti selvitetään jakamalla kappale pieniin pistemäisiin massoihin ja laskemalla yhteen osien hitausmomentit. Jakoa tihentämällä päädytään integrointitehtävään. Selvitetään R -säteisen ohuen pallonkuoren hitausmomentti (pyörimisakselina z -akseli), kun pallonkuoren pintatiheys on ρ (kg/m^3).

Pallonkuoren S parametriesitykseksi voidaan valita

$$\mathbf{r}(u, v) = (R \cos u \sin v, R \sin u \sin v, R \cos v),$$

missä $u \in [0, 2\pi]$ ja $v \in [0, \pi]$. Tällöin

$$\frac{\partial \mathbf{r}}{\partial u} = (-R \sin u \sin v, R \cos u \sin v, 0)$$

ja $\frac{\partial \mathbf{r}}{\partial v} = (R \cos u \cos v, R \sin u \cos v, -R \sin v)$. Suoralla laskulla saadaan

$$\left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| = |R^2 \sin v| = R^2 \sin v,$$

kun $v \in [0, \pi]$. Pallon pinnan pisteen (x, y, z) etäisyyden neliö pyörimisakselista (z -akseli) on

$$r^2 = \|(x, y, z) - (0, 0, z)\|^2 = x^2 + y^2,$$

joten kysytyn hitausmomentin lausekkeeksi saadaan

$$J = \int_S \rho(x^2 + y^2) dA.$$

Edellä on jo selvitetty lauseke dA :lle ja parametriesitys pinnalle, joten integraali voidaan laskea suoraan:

$$\begin{aligned} J &= \rho \int_{u=0}^{2\pi} \int_{v=0}^{\pi} (R^2 \cos^2 u \sin^2 v + R^2 \sin^2 u \sin^2 v) R^2 \sin v dv du \\ &= \rho R^4 \int_{u=0}^{2\pi} \int_{v=0}^{\pi} \sin^3 v dv du = \rho R^4 \cdot 2\pi \int_{v=0}^{\pi} \sin^3 v dv = \rho R^4 \cdot 2\pi \cdot \frac{4}{3}. \end{aligned}$$

Jälkimmäisin integraali voidaan laskea osittaisintegroinnilla. Täten

$$J = \frac{8\pi}{3} \rho R^4 = \frac{2}{3} R^2 \cdot 4\pi R^2 \rho = \frac{2}{3} R^2 m,$$

missä $m = 4\pi R^2 \rho$ on pallonkuoren massa (ala kertaa pintatiheys).

7.6 Vektorikentän pintaintegraali

Määritelmä 67. Jos $\mathbf{f}$ avaruuden $\mathbb{R}^3$ alueessa U määritelty vektorikenttä (siis funktio $\mathbf{f} : U \rightarrow \mathbb{R}^3$) ja $S = \{\mathbf{r}(u, v) \mid (u, v) \in A\}$ pinta $\mathbb{R}^3$:ssa, on

$$\iint_{(u,v) \in A} \mathbf{f}(\mathbf{r}(u, v)) \cdot \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} du dv$$

vektorikentän $\mathbf{f}$ pintaintegraali yli pinnan S .

Merkitään lisäksi

$$\mathbf{n}(\mathbf{r}(u, v)) = \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\|^{-1} \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v},$$

jolloin $\mathbf{n}(\mathbf{r}(u, v))$ on siis pisteessä $\mathbf{r}(u, v)$ pintaa vastaan kohtisuorassa oleva yksikkövektori. Näillä merkinnöillä pintaintegraali voidaan kirjoittaa muotoon

$$\iint_{(u,v) \in A} \mathbf{f}(\mathbf{r}(u, v)) \cdot \mathbf{n}(\mathbf{r}(u, v)) dA, \quad (7.8)$$

jolloin siis vektorikentän pintaintegraalissa itse asiassa on kyseessä vektoriarvoisen funktion pintaa vastaan kohtisuoran komponentin (skalaari) integrointi yli pinnan.

Tutkittaessa virtausilmiöitä annetun pinnan läpi (esim. nesteen virtaus, magneettivuo, jne.) katsotaan normaalivektorin suuntainen virtaus positiiviseksi ja sille vastakkainen suunta negatiiviseksi. Tämä tietysti edellyttää, että pinta on kaksipuolinen (esim. ns. Möbiuksen nauha on yksipuolinen pinta), jolloin voidaan sopia kumpaa puolta pidetään ”sisäpuolena” ja kumpaa ”ulkopuolena”. Integraalia (7.8) kutsutaankin nimellä *vektorikentän $\mathbf{f}$ vuo pinnan S läpi*.

Esimerkki 168. Olkoon $T(x, y, z)$ lämpötila pisteessä $(x, y, z) \in W$, missä W on $\mathbb{R}^3$:n alue. Termodynamiikan mukaan lämpö siirtyy korkeammasta lämpötilasta matalampaan, ja lämmön johtumista kuvaa ns. Fourierin laki, jonka mukaan lämmön virtausnopeutta edustaa vektorikenttä $\mathbf{F} = -k\nabla T$, missä vakio $k > 0$ kuvaa materiaalin lämmönjohtavuutta.

Jos $T(x, y, z) = x^2 + y^2 + z^2$ ja pinta S yksikköpallon $x^2 + y^2 + z^2 = 1$ pinta, saadaan $\mathbf{F} = -2k(x, y, z)$. Pinnan ulospäin suuntautunut yksikkönormaalivektori saadaan seurauksen 4 mukaan jakamalla funktion $F(x, y, z) = x^2 + y^2 + z^2 - 1$ gradientti normillaan, jolloin

$$\mathbf{n} = \sqrt{(4x^2 + 4y^2 + 4z^2)}^{-1} (2x, 2y, 2z) = (x, y, z)$$

pinnalla S . Tällöin $\mathbf{F} \cdot \mathbf{n} = -2k(x, y, z) \cdot (x, y, z) = -2k(x^2 + y^2 + z^2) = -2k$ pinnalla S . Näin ollen

$$\int_S \mathbf{F} \cdot \mathbf{n} dA = \int_S (-2k) dA = -2k \cdot 4\pi = -8\pi k.$$

7.7 Stokesin lause

Newton-Leibnizin säännön (kts. Insinöörimatematiikka IB) mukaan

$$\int_a^b f'(x) dx = f(b) - f(a), \quad (7.9)$$

jos f on välillä $[a, b]$ derivoituva funktio. Tämän yleistämistä ajatellen huomionarvoista on, että sääntö kytkee toisiinsa integraalin yli *välin* $[a, b]$ (vasen puoli) ja funktion arvojen yhdistelmän välin *reunalla* (= päätepisteissä) (oikea puoli). Yleisemmässä terminologiassa oikean puolen summalausketta $f(b) - f(a)$ voidaan pitää eräänlaisena ”integraalina” yli *välin* $[a, b]$ *reunan*. Newton-Leibnizin säännön yleistys, joka tunnetaan (abstraktina) *Stokesin lauseena*, sanookin intuitiivisesti ilmaistuna, että differentiaali muodon df integraali yli A :n on sama kuin alkuperäisen muodon f integraali yli A :n reunan ∂A :

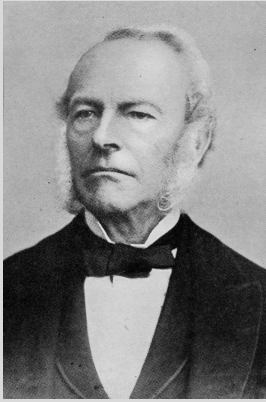
Lause 50 (Stokesin lause).

$$\int_A df = \int_{\partial A} f, \quad (7.10)$$

kunhan riittävät säännöllisyys ehdot täyttyvät.

Tällä kurssilla ei esitetä abstraktin Stokesin lauseen todistusta, eikä täsmennetä käsitettä *differentiaalimuoto*. Sen sijaan esitetään todistukset erikoistapauksia abstraktista Stokesin lauseesta kolmiulotteisessa avaruudessa. Muistutettakoon myös, että (7.9) on erikoistapaus Stokesin lauseesta (7.10).

Taustatietoa



Sir George Gabriel Stokes (1819–1903) oli brittiläinen matemaatikko ja fyysikko, joka saavutti huomattavia tuloksia erityisesti virtausdynamiikan ja optiikan tutkimuksessa. Virtausdynamiikkaan kuuluva Navier-Stokesin väittämä on yksi seitsemästä matemaattisesta ongelmasta, joiden ratkaisemisesta Clay Mathematics Institute on luvannut miljoonan dollarin palkkion.

(kuva: Wikimedia Commons)

Tavallisimpia $\mathbb{R}^3$:ssa esiintyviä Stokesin lauseen erikoistapauksia varten otetaan käyttöön symbolinen *nabla*-vektori

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right),$$

ja tämän operoinnit:

- Jos $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ on skalaarikenttä, on sen *gradientti* vektorikenttä

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

- Jos $\mathbf{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\mathbf{f} = (f_1, f_2, f_3)$ on vektorikenttä, on sen *divergenssi* symbolisesti pistetulona saatava skalaarikenttä

$$\nabla \cdot \mathbf{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}.$$

- Jos $\mathbf{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\mathbf{f} = (f_1, f_2, f_3)$ on vektorikenttä, on sen *roottori* symbolisesti ristitulona saatava vektorikenttä

$$\nabla \times \mathbf{f} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{vmatrix} = \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right)$$

Abstraktin Stokesin lauseen erikoistapauksia ovat mm. seuraavat tulokset:

Lause 51 (Klassinen Stokesin lause). Olkoon A pinta $\mathbb{R}^3$:ssa ja ∂A sen (suljettu) reunakäyrä, sekä $\mathbf{f}$ vektorikenttä, jonka komponenttifunktioiden osittaisderivaatat ovat jatkuvia. Tällöin

$$\iint_A (\nabla \times \mathbf{f}) \cdot \mathbf{n} dA = \int_{\partial A} \mathbf{f} \cdot d\mathbf{x}.$$

Toisin sanoen, roottorin $\nabla \times \mathbf{f}$ pintaintegraali yli pinnan A on sama kuin kentän $\mathbf{f}$ käyräintegraali yli suljetun käyrän ∂A .

Rajoitettaessa tarkastelu xy -tasolle, jolloin vektorikentän z -koordinaatti on siis nolla, saadaan Stokesin lauseesta ns. *Greenin lause*

$$\iint_A \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right) dx dy = \int_{\partial A} f_1 dx + f_2 dy,$$

Lause 52 (Gaussin lause, divergenssilause). Olkoon V alue $\mathbb{R}^3$:ssa ja ∂V sen (suljettu) reunapinta, sekä $\mathbf{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ vektorikenttä, jonka komponenttifunktioiden osittaisderivaatat ovat jatkuvia. Tällöin

$$\iiint_V \nabla \cdot \mathbf{f} dx dy dz = \iint_{\partial V} \mathbf{f} \cdot \mathbf{n} dA.$$

Toisin sanoen, divergenssin $\nabla \cdot \mathbf{f}$ (skalaarikenttä) integraali yli tilavuuden V on sama kuin vektorikentän $\mathbf{f}$ pintaintegraali yli pinnan ∂V .

Aiemmin on saatu gradientille intuitiivinen tulkinta skalaarikentän voimakkaimman muutoksen suuntana. Edellä mainitut lauseet antavat myös intuitiivisen tulkinnan divergenssille ja roottorille.

Tarkastellaan divergenssin tulkintaa, ja tätä varten oletetaan, että $\nabla \cdot \mathbf{f}$ on pisteessä $\mathbf{x}$ jatkuva skalaarikenttä. Olkoon V pieni avaruuden $\mathbb{R}^3$ kappale, johon $\mathbf{x}$ sisältyy ja jonka tilavuus on $|V|$. Mitä pienempi V on, sitä lähempänä divergenssi $\nabla \cdot \mathbf{f}$ on (jatkuvuuden perusteella) sen arvoa pisteessä $\mathbf{x}$, ja siksi integraali

$$\iiint_V \nabla \cdot \mathbf{f} dV \quad (7.11)$$

on sitä lähempänä arvoa $|V| \nabla \cdot \mathbf{f}(\mathbf{x})$, mitä pienempi V on. Toisaalta taas divergenssilauseen mukaan (7.11) on yhtä kuin vektorikentän $\mathbf{f}$ vuo yli V :tä rajoittavan pinnan. Kun annetaan tilavuuden V kutistua kohti nollaa, nähdään siis, että divergenssi $\nabla \cdot \mathbf{f}(\mathbf{x})$ edustaa vektorikentän $\mathbf{f}$ lähteisyyttä pisteessä $\mathbf{x}$, ts. sitä, kuinka suuri kentän $\mathbf{f}$ vuo pisteestä $\mathbf{x}$ lähtee.

Samoin tarkastelemalla pientä pintaa A ja sen reunakäyrää ∂A nähdään, että roottori $\nabla \times \mathbf{f}$ edustaa vektorikentän $\mathbf{f}$ pyörteisyyttä. Sanotaan, että vektorikenttä $\mathbf{f}$ on lähteetön, jos $\nabla \cdot \mathbf{f} = 0$ ja että $\mathbf{f}$ on pyörteetön, jos $\nabla \times \mathbf{f} = \mathbf{0}$.

Esimerkki 169. Lasketaan käyräintegraali

$$\int_{\gamma} xy dx - y^2 dy$$

yli suljetun käyrän γ , jonka osat ovat $\{(0, y) \mid 0 \leq y \leq 1\}$, $\{(x, 1) \mid 0 \leq x \leq 2\}$ ja $\{(2\sqrt{y}, y) \mid 0 \leq y \leq 1\}$. Greenin lausetta käyttämällä voidaan käyräintegraali muuttaa tavalliseksi integraaliksi yli $\mathbb{R}^2$:n alueen A , jota rajoittaa käyrä γ .

$$\begin{aligned} \int_{\partial A} xy dx - y^2 dy &= \int_A \left(\frac{\partial}{\partial x}(-y^2) - \frac{\partial}{\partial y}(xy) \right) dx dy \\ &= - \int_{y=0}^1 \int_{x=0}^{2\sqrt{y}} x dx dy = - \int_0^1 \frac{1}{2} (2\sqrt{y})^2 dy = - \int_0^1 2y dy = -1. \end{aligned}$$