Matriisi ja determinantti

Noora Mehtälä

Matematiikan pro gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Talvi 2012

Sisältö

Johdantoa ja taustaa 1

Luku 1. Matriisien ja determinanttien historiaa 5

Luku 2. Matriisi 9
1. Lineaarinen yhtälöryhmä 9
2. Matriisi ja lineaarikuvaus 10
3. Käänteismatriisi ja transponoitu matriisi 15

Luku 3. Determinantti 18
1. Determinantti-käsite ja determinantin perusominaisuudet 18
2. Determinantin ominaisuuksia 26
3. Alideterminantti ja determinantin kehityssääntö 28
4. Käänteismatriisi ja determinantti, Cramerin sääntö 29
5. Integraalilaskennan muuttujanvaihtolause 32

Luku 4. Matriisin aste 34
1. Matriisin astelause 34
2. Matriisin dimensiolause 37

Luku 5. Ominaisarvoteoriaa 39
1. Ominaisarvot ja ominaisvektorit 39
2. Yhtäläiset matriisit ja diagonalisoituvuus 40
3. Symmetriset matriisit ja ortogonaalinen diagonalisoituvuus 42
4. Neliömuodot 44
5. Differentiaalilaskennan ääriarvotarkasteluja 47

Luku 6. Ehdotus lukion pitkän matematiikan *Numeerisia ja algebrallisia menetelmiä* -kurssin sisältöön 49
1. Lineaarista algebraa ja geometriaa lukiossa 50

Viitteet 61

Johdantoa ja taustaa

Pro gradu -tutkielmani koostuu kuudesta luvusta, joissa käsitellään lineaarista algebraa ja geometriaa pääpainona lineaarikuvaukset, matriisit, determinantit ja matriisien ominaisarvot. Tässä johdannossa esitellään lyhyesti, mitä tutkielman luvut sisältävät sekä otetaan muutamia esimerkkejä, joissa voidaan soveltaa lineaarisia yhtälöryhmiä ja matriiseja. Näitä sovellusesimerkkejä löytyy hyvin paljon, koska lineaarialgebraa voidaan hyödyntää monilla eri tieteen aloilla. Myös lineaarialgebraa käsittelevää kirjallisuutta löytyy suorastaan valtavasti, koska se on paljon tutkittu matematiikan osa-alue, jota vielä nykyäänkin tutkitaan ja opetetaan korkeakouluissa. Tutkielmani tärkeimmät lähteet ovat [1], [6] ja [7]. Lähteisiin [6] ja [7] liittyen olen hyödyntänyt syksyllä 2007 ja keväällä 2008 Veikko Purmosen pitämien lineaarisen algebran ja geometrian kurssien aikana tekemiäni muistiinpanoja ja luennoilla esitettyjä esimerkkejä. Tutkielmani esimerkit pohjautuvat pääasiassa lähteen [1] esimerkkeihin, joita olen tarvittaessa hieman muokannut. Luvun 5 viimeisen kappaleen esimerkit olen muokannut Veikko Purmosen syksyllä 2009 pitämän kurssin *Differentiaalilaskenta 1* harjoitustehtävistä ja luvun 6 esimerkki 6.13 on lähteen [8] harjoitustehtävä 234. Työni lauseista olen todistanut tai johtanut vain osan tarkasti. Lauseisiin, joista puuttuu tarkka todistus, olen laittanut näkyville viitteen, josta todistus löytyy. Tämän kappaleen sovellusesimerkit pohjautuvat lähteisiin [3] ja [10].

Ensimmäisessä luvussa kerrotaan, kuinka matriisit ja determinantit ovat lähteneet kehittymään lineaaristen yhtälöryhmien pohjalta. Niiden historia ulottuu jopa 300-100 -luvuille eKr, mutta niiden varsinainen kehittyminen alkoi vasta 1600-luvulla. Varhaisessa matriisiteoriassa determinantteja tutkittiin enemmän kuin matriiseja, koska determinantille oli silloin enemmän käyttöä, eikä matriisien ja determinanttien välistä yhteyttä ymmärretty. Determinantti on myös käsitteenä puoli vuosisataa vanhempi kuin matriisi. Tärkeimpiä henkilöitä lineaarialgebran kehittymisen kannalta ovat mm. Leibniz, Gauss, Cauchy ja Cayley. Tässä luvussa keskitytään tarkastelemaan niiden käsitteiden ja tulosten kehittymistä, joita tarvitaan tutkielman muissa luvuissa. Luku pohjautuu lähteisiin [1], [6], [10], [11] ja [12].

Toisessa luvussa käsitellään lyhyesti lineaarisia yhtälöryhmiä, lineaarisista yhtälöryhmistä muodostuvia matriiseja ja selvitetään matriisien yhteyttä lineaarikuvauksiin. Matriisi on lineaarisen yhtälöryhmän reaalisista kertoimista muodostuva lukukaavio, joka koostuu rivi- ja sarakevektoreista. Lineaarikuvausta $L : \mathbb{R}^n \to \mathbb{R}^m$ vastaa kiinnitettyjen kantojen suhteen yksikäsitteinen $m \times n$-matriisi, mikä pätee kääntäenkin eli matriisi määrää tietyn lineaarikuvauksen. Tämä lineaarikuvauksen ja matriisin välinen yhteys on tärkeä, koska sen avulla voidaan johtaa monia tuloksia. Esimerkiksi kohdassa 2.13 matriisien tulo johdetaan käyttämällä lineaarikuvauksien ja matriisien välistä yhteyttä. Laskennallisestikin tämä yhteys on hyödyllinen, koska lineaarikuvausten laskut voidaan ratkaista helpommin niitä vastaavien matriisien avulla. Kohdassa 2.21 johdetaan käänteismatriisi, jonka yhteyttä lineaarikuvauksen käänteiskuvaukseen tarkastellaan lauseessa 2.26, joka sanoo, että matriisille löytyy käänteismatriisi, jos ja vain jos sitä vastaava lineaarikuvaus on bijektio. Luku pohjautuu lähteisiin [1] ja [6].

Selvitetäänpä seuraavaksi, mihin näitä lineaarisia yhtälöryhmiä ja matriiseja todellisuudessa tarvitaan. Lineaarisia yhtälöryhmiä käytetään hyvin laajalti esimerkiksi fysiikassa, kemiassa ja taloustieteessä. Myös kaupunkisuunnittelijat ja liikenneinsinöörit hyödyntävät niitä liikennevirtauksia ja niissä ilmeneviä toistuvuuksia mallintaessaan. Usein eri ongelmista saatavat tiedot ovat lineaarisia, jolloin tutkittavan asian matemaattinen mallintaminen onnistuu hyvin lineaaristen yhtälöryhmien avulla. Tälläiset lineaariset yhtälöryhmät koostuvat tavallisesti sadoista jopa tuhansista muuttujista ja yhtälöistä, jolloin niiden tehokas käsittely onnistuu vain tietokoneiden avulla. Matriisit ja vektorit ovatkin tarpeellisia juuri matemaattisessa mallintamisessa, koska matriiseilla voidaan havainnollistaa selkeästi lineaarisia tilanteita. Tällaisesta mallintamisesta löytyy väestön muuttoliikennettä käsittelevä esimerkki lähteen [3] sivuilta 87-89. Myös ravitsemustieteessä hyödynnetään vektoreita ja matriiseja tehokkaita ja turvallisia ruokavalioita muodostaessa. Tästä löytyy esimerkki lähteen [3] sivuilta 86-87.

Tietokoneiden grafiikka pohjautuu myös matriiseihin. Siinä matriiseja käytetään heijastamaan 3-ulotteisia kuvia 2-ulotteiselle näytölle ja luomaan todelliselta näyttävää liikettä. Lineaarialgebran sanotaankin olevan graafisen tietokoneohjelmiston ”sielu”, koska kaikki tietokoneella tehtävät kuvien muokkaukset saadaan aikaan lineaarialgebran menetelmien kautta. Esimerkiksi autoteollisuudessa insinöörit rakentavat tietokoneilla matriiseista koostuvia matemaattisia rautakehikkoautoja, joita sitten testaillaan tietokoneilla. Näiden matemaattisten autojen avulla saadaan paljon hyödyllistä tietoa tulevasta autosta ja ne mahdollistavat sen, että valmistettavaan autoon voidaan tehdä tarvittavia muutoksia ennen kuin sitä on edes alettu käytännössä rakentaa. Tästä aiheesta löytyy lisää tietoa lähteen [3] sivuilta 95-96.

Kemiassa matriisien avulla voidaan esimerkiksi tasapainottaa reaktioyhtälöitä. Tasapainotetaan malliksi reaktioyhtälö

$$a\mathrm{Cu_2S} + b\mathrm{O_2} \to c\mathrm{Cu} + d\mathrm{SO_2}.$$

Kertoimet a, b, c ja d löydetään käsittelemällä jokainen alkuaine erikseen omana homogeenisena yhtälönä

$$\begin{array}{cl} \mathrm{Cu}: & 2a + 0b - 1c - 0d = 0 \\ \mathrm{S}: & 1a + 0b - 0c - 1d = 0 \\ \mathrm{O}: & 0a + 2b - 0c - 2d = 0. \end{array}$$

Homogeenista yhtälöryhmää vastaa laajennettu kerroinmatriisi

$$A = \left[\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & -2 & 0 \end{array}\right],$$

josta saadaan yhtälöryhmälle ratkaisu $t(1,1,2,1)$, $t \in \mathbb{R}$, Gaussin-Jordanin menetelmällä. Reaktioyhtälön kertoimiksi halutaan mahdollisimman pienet positiiviset kokonaisluvut, joten valitaan $t = 1$, jolloin tasapainotetuksi reaktioyhtälöksi saadaan

$$\mathrm{Cu_2S} + \mathrm{O_2} \to 2\mathrm{Cu} + \mathrm{SO_2}.$$

Luvussa 3 tarkastellaan determinatti-käsitettä Cauchyn geometrisen tilavuustulkinnan avulla: 2×2-matriisin sarakkeet virittävät suunnikkaan, jonka pinta-ala on yhtä suuri kuin saman matriisin determinantin itseisarvo, 3×3-matriisin sarakkeet virittävät vastaavasti suuuntaissärmiön, jonka tilavuus vastaa matriisin determinantin itseisarvoa ja $n \times n$-matriisin sarakkeet virittävät n-ulotteisen suuntaissärmiön, jonka tilavuus on taas matriisin determinantin itseisarvo. Kohdan 3.1 tarkastelu 3×3-matriisille on itseni tekemä ja samoin kohdan 3.3 tarkastelut 2×2 ja 3×3 -matriiseille. Determinantin yleiseen määritelmään päästään determinantin karakterististen ominaisuuksien kautta. Tätä määritelmää ei tavallisesti käytetä determinanttien laskemiseen, koska se onnistuu helpommin determinantin kehityssäännön 3.11 ja lauseessa 3.9 lueteltujen determinantin ominaisuuksien avulla. Näitä ominaisuuksia olen perustellut lauseen todistuksessa käyttäen apuna determinantin tilavuustulkintaa. Yksi lineaarialgebran tunnetuimpia tuloksia on Cramerin sääntö, joka johdetaan kohdassa 3.17. Cramerin sääntöä perustellaan geometrisesti 2×2-matriisille hieman myöhemmin. Lopuksi otetaan integraalilaskennan muuttujanvaihtolause 3.22 esimerkkinä matriisien ja determinanttien soveltuvuudesta muihin matematiikan aloihin. Luku pohjautuu lähteisiin [5], [6] ja [9].

Luvussa 4 määritellään matriisin aste, jolla tarkoitetaan matriisin sarakeavaruuden dimensiota. Luvun tärkeimpiä lauseita ovat matriisin astelause 4.11, joka sanoo, että matriisin riviavaruuden ja sarakeavaruuden dimensiot ovat samat, vaikka ne olisivatkin eri avaruuksien aliavaruuksia, ja dimensiolause 4.14, jonka mukaan matriisin $A = A_{m\times n}$ ytimen dimensiolle $\dim N(A)$ ja matriisin asteelle $\operatorname{rank} A$ pätee $\dim N(A) + \operatorname{rank} A = n$. Dimensiolause on tehokas työkalu lineaaristen yhtälöryhmien ja niistä muodostuvien matriisien tietojen käsittelemisessä. Luvun kolmas tärkeä lause liittää lineaarikuvauksen, matriisin ja determinantin toisiinsa. Tämä lause 4.19 sanoo, että matriisia $A = A_{n\times n}$ vastaava lineaarikuvaus $L : \mathbb{R}^n \to \mathbb{R}^n$ on bijektio, jos ja vain jos determinantti $\det A$ ei ole nolla. Tulosta voidaan selittää geometrisesti determinantin tilavuustulkintaa käyttäen. Luku pohjautuu lähteisiin [1] ja [6].

Luvussa 5 määritellään lineaarikuvaukselle ja matriisille käsitteet ominaisvektori ja ominaisarvo. Matriisin ominaisarvot saadaan ratkaistua lauseen 5.4 avulla, joka sanoo, että luku λ on matriisin $A = A_{n\times n}$ ominaisarvo, jos ja vain jos se toteuttaa matriisin A karakteristisen yhtälön $\det(A - \lambda I) = 0$, missä $I = I_n$ on yksikkömatriisi. Luvussa määritellään mm. myös käsitteet symmetrinen matriisi ja matriisin diagonalisoituvuus: matriisi $A = A_{n\times n}$ on symmetrinen, jos sille pätee $A^T = A$, ja matriisi $A_{n\times n}$ on diagonalisoituva, jos on olemassa kääntyvä matriisi $P = P_{n\times n}$, jolle $D = P^{-1}AP$ on diagonaalimatriisi. Nämä määritelmät liitetään toisiinsa lauseessa 5.19, jonka mukaan matriisi A on symmetrinen, jos ja vain jos löytyy ortogonaalinen matriisi Q, jolle $D = Q^TAQ$ on diagonaalinen. Kohdassa 5.22 määritellään avaruuden $\mathbb{R}^n$ neliömuoto, joka on funktio $q : \mathbb{R}^n \to \mathbb{R} : x \mapsto x^TAx$ täsmälleen yhdelle symmetriselle matriisille A. Neliömuodon definiittisyyttä tutkitaan määritelmässä 5.26 ja lauseessa 5.28. Matriisin ominaisarvoja voidaan soveltaa differentiaalilaskennassa ääriarvotarkasteluihin. Ominaisarvojen avulla saadaan helposti kriittistä tietoa differentiaaliyhtälöistä. Muutenkin ominaisvektorit ja ominaisarvot ovat hyödyllisiä kaikkialla teoreettisessa ja sovelletussa matematiikassa. Käytännössä niitä hyödynnetään teknisessä suunnittelussa sekä luonnollisesti fysiikassa ja kemiassa. Neliömuodot esiintyvät mm. lineaarialgebran tekniikan sovellutuksissa, kuten kriittisessä suunnittelussa ja optimoinnissa. Luku pohjautuu lähteisiin [1], [3], [4] ja [7].

Luvussa 6 annan ehdotuksen, kuinka lineaarista algebraa ja geometriaa voitaisiin käsitellä lukiossa pitkän matematiikan soveltavalla kurssilla *Numeerisia ja algebrallisia menetelmiä*. Kurssilla voitaisiin käsitellä 2×2 ja 3×3 -matriiseja, niiden determinatteja, determinantin geometrista tulkintaa sekä Cramerin sääntöä. Tarkempi johdanto löytyy luvun alusta. Luku pohjautuu lähteisiin [1], [2], [6] ja [13].

LUKU 1

Matriisien ja determinanttien historiaa

Matematiikan historian alkuvaiheista huomataan, että matematiikan kehittyminen on lähtenyt liikkeelle ihmisten tarpeesta ja halusta ratkaista käytännön ongelmia tarkasti. Myös lineaariset yhtälöt ja yhtälöryhmät ilmaantuivat juuri käytännön ongelmien kautta. Yhtälöryhmien tutkimisen tuloksena saatiin sääntöjä niiden ratkeavuudelle ja ratkaisumenetelmiä, joiden yhteydessä matriisit ja determinantit alkoivat muodostua ja kehittyä. Matriisien ja determinanttien kehittyminen alkoi kunnolla 1600-luvulla, vaikka niiden historia ulottuukin jopa 300-100 -luvuille eKr. Varhainen matriisiteoria korosti determinantteja vahvemmin kuin matriiseja.

Jo babylonialaiset törmäsivät lineaarisiin yhtälöryhmiin juuri käytännön ongelmien kautta. Heillä oli tapana kirjoittaa asioita ylös hyvin säilyviin savitauluihin, joista noin vuodelta 300 eKr säilynyt taulu sisältää seuraavan ongelman:

> Kahden pellon kokonaispinta-ala on 1800 pinta-alayksikköä. Ensimmäinen pelto tuottaa viljaa 2 kannullista/pinta-alayksikkö ja toinen pelto tuottaa 1/2 kannullista/pinta-alayksikkö. Jos kokonaistuotto on 1100 kannullista, niin minkä kokoisia pellot ovat.

Kiinassa vuosina 200-100 eKr matriisit tulivat tarkemmin esille taas lineaarisia yhtälöryhmiä ratkaistaessa. Tältä ajalta säilynyt teksti sisältää ensimmäisen esimerkin matriisimenetelmästä. Tämän esimerkin ongelma on vastaavanlainen kuin babylonialaisten:

> Kun otetaan kolme lyhdettä ensimmäistä jyvälajia, kaksi lyhdettä toista jyvälajia ja yksi lyhde kolmatta, saadaan yhteensä 39 mitallista. Kaksi lyhdettä ensimmäistä lajia, kolme toista lajia ja yksi kolmatta lajia ovat yhteensä 34 mitallista. Yksi lyhde ensimmäistä, kaksi toista ja kolme kolmatta ovat yhteensä 26 mitallista. Kuinka monta jyvämitallista yksi lyhde jokaista lajia sisältää?

Ratkaistakseen ongelman kirjoittaja asetteli kolmen lineaarisen yhtälön yhtälöryhmän kertoimet seuraavasti taulukkoon:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \\ 26 & 34 & 39 \end{array}$$

Tämä kirjoitustapa vastaa muuten nykyistä tapaa kirjoittaa lineaariset yhtälöt matriiseina paitsi, että nykyinen tapa on kirjoittaa yhtälöt riveinä, tässä ne ovat sarakkeina. Vielä merkittävämpää on se, että kirjoittaja neuvoo lukijaa kertomaan keskimmäisen sarakkeen numerolla 3 ja vähentämään siitä oikeanpuoleisimman sarakkeen niin monesti kuin se on mahdollista. Vastaavanlaisia operaatioita pyydetään tekemään, kunnes taulu on seuraavassa muodossa

$$\begin{array}{ccc} 0 & 0 & 3 \\ 0 & 5 & 2 \\ 36 & 1 & 1 \\ 99 & 24 & 39, \end{array}$$

josta saadaan ratkaisu jyväongelmalle. Menetelmä vastaa Gaussin eliminointimenetelmää, joka tosin tuli lännessä järjestelmällisesti tutkituksi ja tunnetuksi vasta 1800-luvulla.

Monet matriisiteorian tulokset ovat ilmaantuneet aivan kuin sivutuotteena, kun on tutkittu muita asioita. Determinanttien kehittymisen alkuvaiheessa esille tulleiden tulosten ja nykyisten modernien ideoiden välillä on joskus vaikea nähdä selkeää yhteyttyä, koska aina ei ole käytetty tarkkoja määritelmiä. Myös yleisten merkintätapojen puuttuminen on hankaloittanut niiden tutkimista. Aluksi determinantteja käytettiin ilman niiden yhteyttä matriiseihin: tavallisesti determinantin ajateltiin olevan lineaarisen yhtälöryhmän ominaisuus, joka määrää, onko yhtälöryhmällä ratkaisua. Näissä yhtälöryhmissä tuli olla yhtä paljon yhtälöitä ja muuttujia. Tämä on vanhin determinantin käyttötarkoituksista matematiikan historiassa ja tässä mielessä determinantteja käytettiinkin jo 200-luvulla eKr Kiinassa. Toinen determinantin varhaisista käyttötarkoituksista on tuntemattomien poistaminen yhtälöryhmästä, jossa on $n + 1$ yhtälöä ja n tuntematonta, ja kolmantena ovat lineaarikuvaukset.

Determinantin idea esiintyi Japanissa ja Euroopassa melkein täsmälleen samaan aikaan, vaikka Seki Kōwa Japanissa luultavamminkin julkaisi ensimmäisenä siihen liittyvää materiaalia. Vuonna 1683 Seki käsitteli matriiseja taulukkoina, kuten jo kiinalaiset huomattavasti aiemmin. Hän ei käyttänyt mitään determinanttiin viittaavaa sanaa, mutta kuitenkin hän esitteli ne ja antoi yleiset menetelmät niiden määrittämiseen esimerkkien avulla. Käyttäen "determinanttejansa" Seki löysi 2×2, 3×3, 4×4 ja 5×5 -matriisien determinantit ja sovelsi niitä yhtälöiden ratkaisuissa.

Euroopassa samana vuonna 1683 G. Leibniz kirjoitti G. de l´Hôpitalille ja selitti kirjeessään, että yhtälöryhmällä

$$\begin{cases} 10 + 11x + 12y = 0 \\ 20 + 21x + 22y = 0 \\ 30 + 31x + 32y = 0 \end{cases}$$

on ratkaisu, koska pätee $10{\cdot}21{\cdot}32+11{\cdot}22{\cdot}30+12{\cdot}20{\cdot}31 = 10{\cdot}22{\cdot}31+11{\cdot}20{\cdot}32+12{\cdot}21{\cdot}30$. Leibniz ei käyttänyt numeerisia kertoimia vaan selitti asiansa sanallisesti. Hänen 21 tarkoittaa samaa kuin nykyään alkio a_{21}. Determinatti on siis liittynyt ensin kahden yhtälön yhtälöryhmän ja sitten kolmen yhtälön yhtälöryhmän ratkeavuusehtoon. Leibniz oli vakuuttunut siitä, että hyvä matemaattinen merkintätapa on tärkeä tekijä tutkimusten etenemisen kannalta, joten hän kokeili todella monia eri tapoja kirjoittaa yhtälöryhmien kertoimet ylös. Lisäksi hän tutki neliömuotojen kertoimia, mikä luonnollisesti johti kohti matriisiteoriaa.

Useat matemaatikot todistivat erilaisia tuloksia 1700-luvulla matriiseille ja determinanteille, vaikka eivät käyttäneet samoja termejä, joita nykyään käytetään. Tämän ajan yksi kuuluisimmista tuloksista on Cramerin sääntö, joka oli aikanaan hyvin käyttökelpoinen ja tärkeä tulos. G. Cramer julkaisi säännön vuonna 1750, mutta väitetään, että säännön tiesi C. Maclaurin jo vuonna 1729. Tärkeä saavutus tältä ajalta on myös

J.-L. Lagrangen käsitys asiasta, joka nykyään tunnetaan determinantin tilavuustulkintana. Determinantti lähti myös käsitteenä kehittymään 1700-luvulla. P.-S. Laplace käytti determinantille käsitettä "jonkin tulos", joka vastaa juuri sitä, mitä nykyään tarkoitetaan determinantilla. Yllättävää on se, että samaa termiä käytti Leibniz jo aiemmin tutkimuksissaan, joista Laplacen täytyi vielä olla täysin tietämätön.

Yksi matematiikan historian tärkeimmistä henkilöistä on C. Gauss, joka esitteli neliömuotoja käsittelevässä julkaisussaan vuonna 1801 ensimmäisenä termin "determinantti". Hän käytti termiä, koska determinantti määrittelee neliömuodon ominaisuuksia. Determinantti on siis käsitteenä vanhempi kuin matriisi, vaikkakaan Gaussin käyttämä "determinantti" ei vastaa täysin vastaavaa termiä, jota nykyään käytetään. Gauss tutki neliömuotoja ja havainnollisti toisen asteen neliömuotoja $ax^2 + 2bxy + cy^2$ symmetrisen matriisin

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

avulla. Lisäksi hän pääsi hyvin lähelle matriisien tulon määritelmää, jonka J. Binet esitti virallisesti 1800-luvun alussa.

A.-L. Cauchy oli myös hyvin tärkeä henkilö matriisien ja determinanttien kehittymisen kannalta, koska hänen tutkimuksensa olivat aikaisempiin verrattuna huomattavasti täydellisempiä. Hän käytti ensimmäisenä determinanttia sen modernissa merkityksessä. Hän toi determinantteihin mukaan geometrisen näkökulman liittämällä käsitteen tason $\mathbb{R}^2$ vektoreiden $u = (u_1, u_2)$ ja $v = (v_1, v_2)$ virittämän suunnikkaan pinta-alaan

$$\pm \operatorname{ala}(u, v) = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}$$

ja vastaavasti avaruudessa $\mathbb{R}^3$ kolmen vektorin u, v ja w virittämän suuntaissärmiön tilavuuteen $\pm \operatorname{vol}(u, v, w) = \det[\bar{u}\ \bar{v}\ \bar{w}]$. Cauchy todisti useita tuloksia ensimmäistä kertaa, joista yksi tärkeimmistä oli determinantin tuloa käsittelevän tuloksen $\det(AB) = (\det A)(\det B)$ todistus. Ominaisarvoteoriaa hän kehitti neliömuotoja tutkiessaan määrittelemällä matriisin A karakteristisen yhtälön $\det(A - \lambda I) = 0$, löytämällä matriisin ominaisarvot ja antamalla tuloksia matriisin diagonalisoinnille. Hän myös esitteli yhtäläisten matriisien idean ja näytti, että jos kaksi matriisia ovat yhtäläisiä, niin niillä on silloin sama karakteristinen yhtälö. Neliömuotojen yhteydessä hän käytti kerroinmatriisille termiä "taulu" sekä löysi ja todisti tuloksen, joka sanoo, että jokainen reaalinen symmetrinen matriisi on diagonalisoituva ja sen ominaisarvot ovat reaalisia. Lisäksi hän todisti mm. determinanttien ominaisuuksia ja antoi tarkemman todistuksen matriisien tulolle kuin Binet.

J. Sturm antoi ominaisarvoille yleistyksen tavallisten differentiaaliyhtälöiden yhtälöryhmien ratkaisemisen yhteydessä. Ominaisarvo käsitteenä tosin oli ilmestynyt jo 80 vuotta aiemmin J.-B. d'Alembertilla lineaaristen differentiaaliyhtälöiden yhtälöryhmien yhteydessä. On tärkeää huomata, ettei Cauchy eikä Sturm ymmärtänyt ominaisarvoja yleisesti, vaan vain niissä erityisissä tapauksissa, joita he tutkivat.

C. Jacobi, L. Kronecker ja K. Weierstrass tutkivat matriiseja lineaarikuvausten yhteydessä. Jacobi julkaisi vuonna 1841 kolme tutkielmaa, joissa hän määritteli ensimmäistä kertaa determinantin algoritmisesti ja käsitteli tuloksia niin numeroiden kuin funktioiden avulla. Nämä tutkielmat toivat determinantin idean laajalti tunnetuksi. Jacobi myös tunsi ja käytti matriiseja ja determinantteja, jotka nykyään tunnetaan Jacobin matriisina ja Jacobin determinanttina.

Determinantin merkintätapa kehittyi sellaiseksi kuin se nykyään on, kun A. Cayley vuonna 1841 piirsi matriisin alkioiden molemmille puolille pystysuorat viivat. Eisenstein vuonna 1844 otti käyttöön merkinnöissään yksittäiset kirjaimet matriisien symboleiksi ja näytti, että niillä voidaan laskea yhteen ja kertoa vastaavasti kuin tavallisilla numeroilla. Vuonna 1850 J. Sylvester käytti ensimmäisenä termiä "matriisi" ja hän halusi sen tarkoittavan "determinanttien äitiä". Hän ratkaisi determinantit vastaavalla tavalla, miten nykyään tehdään, kun lasketaan determinantti alideterminanttien avulla. Hän ymmärsi matriisin apuvälineenä, jonka avulla saadaan determinantti ratkaistua.

Vuonna 1858 Cayley antoi ensimmäisen abstraktin määritelmän matriisille ja näytti, että aiemmin tutkitut neliömuodot ja lineaariset yhtälöt ovat hänen käsitteensä erikoistapauksia. Hän määritteli algebrallisesti matriisien summan, tulon ja vakiolla kertomisen sekä konstruoi käänteismatriisit.

1800-luvun lopussa ilmestyi kaksi tärkeää määritelmää, kun F. Frobenius määritteli matriisin asteen ja Sylvester määritteli neliömatriisin ytimen. 1900-luvun alussa julkaistiin Weierstrassin ja Kroneckerin luentojen pohjalta teoksia, jotka sisälsivät determinanttien modernia teoriaa. Matriisiteoria tuli täysin hyväksytyksi teoriaksi hieman myöhemmin ja niin matriisit saavuttivat keskeisen roolin lineaarialgebrassa.

LUKU 2

Matriisi

1. Lineaarinen yhtälöryhmä

MÄÄRITELMÄ 2.1. Reaalimuuttujien $x_1, \ldots, x_n$ yhtälö on *lineaarinen*, jos se on muotoa

$$a_1x_1 + \cdots + a_nx_n = b, \tag{2.1}$$

missä kertoimet $a_1, \ldots, a_n$ ja termi b ovat reaalilukuja. Yhtälön (2.1) *ratkaisu* on lukujono $(z_1, \ldots, z_n)$, jos yhtälö toteutuu näillä arvoilla $x_1 = z_1, \ldots, x_n = z_n$. Yhtälö (2.1) on *homogeeninen*, jos on voimassa $b = 0$. Muuttujien $x_1, \ldots, x_n$ *lineaarinen yhtälöryhmä*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \quad\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.2}$$

saadaan yhdistämällä $m \geq 2$ yhtälön (2.1) kaltaista yhtälöä. Lukujono $(z_1, \ldots, z_n)$ on yhtälöryhmän (2.2) ratkaisu, jos se toteuttaa ryhmän kaikki m reaalista yhtälöä. Jos taas on voimassa $b_j = 0$ kaikille $j = 1, \ldots, m$, niin yhtälöryhmä (2.2) on *homogeeninen*.

2.2. Yhtälön (2.1) vasen puoli määrittelee kuvauksen

$$F_1 : \mathbb{R}^n \to \mathbb{R} : x = (x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n$$

ja (2.1) on yhtäpitävä yhtälön

$$F_1(x) = b$$

kanssa, kun $b \in \mathbb{R}$. Kuvaus F_1 on muuttujan x komponenttien $x_1, \ldots, x_n$ *ensimmäisen asteen polynomi* ilman vakiotermiä eli *homogeeninen polynomi*, ja tällöin sillä on seuraavat *lineaarikuvauksen* määrittelevät ominaisuudet:

(i) $F(x + y) = F(x) + F(y)$ kaikille $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$,
(ii) $F(\lambda x) = \lambda F(x)$ kaikille $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Näin on kaikille yhtälöryhmän (2.2) yhtälöille, joten (2.2) määrittelee kuvauksen

$$F_2 : \mathbb{R}^n \to \mathbb{R}^m : x \mapsto (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n),$$

ja tällöin (2.2) on yhtäpitävä yhtälön

$$F_2(x) = b$$

kanssa, kun $b := (b_1, \ldots, b_m) \in \mathbb{R}^m$. Myös kuvauksella F_2 on ominaisuudet (i) ja (ii), joten F_2 on *lineaarinen kuvaus*. Huomataan, että yhtälön $F_2(x) = 0$ ratkaisuista muodostuu aliavaruus

$$N = \{x \in \mathbb{R}^n : F_2(x) = 0\}.$$

2.3. Lineaarikuvauksen $L : \mathbb{R}^n \to \mathbb{R}^m$ homogeenisen yhtälön ratkaisujoukko

$$N(L) = \{x \in \mathbb{R}^n : Lx = 0\}$$

on sen *ydin* ja

$$R(L) = L(\mathbb{R}^n) = \{y \in \mathbb{R}^m : y = Lx \text{ jollekin } x \in \mathbb{R}^n\}$$

on sen *arvojoukko.*

Ydin $N(L)$ on lähtöavaruuden $\mathbb{R}^n$ aliavaruus. Vastaavasti arvojoukko $R(L)$ on maaliavaruuden $\mathbb{R}^m$ aliavaruus, koska itse asiassa avaruuden $\mathbb{R}^n$ luonnollisille kantavektoreille $\bar{e}_1, \dots, \bar{e}_n$ pätee

$$R(L) = \langle L\bar{e}_1, \dots, L\bar{e}_n \rangle \subset \mathbb{R}^m.$$

Jos $x = \sum_{i=1}^n x_i \bar{e}_i \in \mathbb{R}^n$, niin lineaarisuuden nojalla pätee

$$Lx = L(\sum_{i=1}^n x_i \bar{e}_i) = \sum_{i=1}^n x_i (L\bar{e}_i) \in \langle L\bar{e}_1, \dots, L\bar{e}_n \rangle ,$$

ja jos $y = \sum_{i=1}^n \lambda_i L\bar{e}_i$, niin $y = L(\sum_{i=1}^n \lambda_i \bar{e}_i) \in R(L)$.

Lineaarikuvauksen injektiivisyydellä ja surjektiivisyydellä on voimassa seuraava tulos.

Lause 2.4. *Lineaarikuvaus $L : \mathbb{R}^n \to \mathbb{R}^m$ on*

(i) *injektio jos ja vain jos $N(L) = \{0\}$, ja*
(ii) *surjektio jos ja vain jos $L(\mathbb{R}^n) = \mathbb{R}^m$.*

2. Matriisi ja lineaarikuvaus

2.5. Lineaarinen yhtälöryhmä

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.3}$$

voidaan kirjoittaa seuraavaan muotoon

$$\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} =: Ax. \tag{2.4}$$

Tällöin saadaan *matriisi*

$$A = [a_{ki}] = [a_{ki}]_{k,i} = [a_{ki}]_{\substack{k=1,\dots,m \\ i=1,\dots,n}} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix},$$

joka on siis yhtälöryhmän (2.3) kertoimista $a_{ki} = a_{k,i}$ muodostuva lukukaavio. Matriisi A on $m \times n$-matriisi, jonka alkiot a_{ki} ovat reaalilukuja, ja sitä merkitään $A =: A_{m\times n}$. Merkinnästä nähdään selkeästi, että matriisissa A on m *riviviektoria* eli *riviä* ja n *sarakevektoria* eli *saraketta.*

Matriisimuoto (2.4) on usein sopivampi ja selkeämpi käyttää kuin avoin muoto (2.3), ja siinä on kaikki olennainen tieto näkyvillä.

2.6. Matriisien yksi tärkeä ominaisuus on niiden yhteys lineaarikuvauksiin. Kohdan 2.2 mukaan yhtälöryhmä (2.3) on yhtäpitävä myös yhtälön

$$L_A x = b$$

kanssa, missä $L_A : \mathbb{R}^n \to \mathbb{R}^m$ on lineaarikuvaus, jolle on

$$L_A x = (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n).$$

Tällaista lineaarikuvausta L_A sanotaan *matriisin A määräämäksi*, koska se liittyy yksikäsitteisesti matriisiin A. Tämä pätee myös kääntäen eli lineaarikuvausta $L : \mathbb{R}^n \to \mathbb{R}^m$ vastaa yksikäsitteisesti matriisi $A = A_{m\times n}$, jolle on $L_A = L$. Tämä on tosi, sillä jos kannat $E = \{\bar{e}_1, \ldots, \bar{e}_n\}$ ja $F = \{\bar{f}_1, \ldots, \bar{f}_m\}$ ovat vastaavasti euklidisten avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^m$ euklidisia kantoja, niin kaikille $x = x_1\bar{e}_1 + \cdots + x_n\bar{e}_n \in \mathbb{R}^n$ pätee

$$\begin{aligned} Lx = L\left(\sum_{i=1}^n x_i\bar{e}_i\right) &= \sum_{i=1}^n x_i(L\bar{e}_i) \\ &= \sum_{i=1}^n x_i\left(\sum_{k=1}^m a_{ki}\bar{f}_k\right) \\ &= \sum_{k=1}^m \left(\sum_{i=1}^n a_{ki}x_i\right)\bar{f}_k, \end{aligned}$$

kun $a_{ki} = (L\bar{e}_i|\bar{f}_k) = (\bar{f}_k|L\bar{e}_i)$. Tällöin on voimassa yhteys $L = L_A$ matriisille $A = [a_{ki}]$. Tällaista matriisia A sanotaan *lineaarikuvausta L vastaavaksi matriisiksi kannoissa E ja F* ja sitä merkitään $\operatorname{mat} L = \operatorname{mat} L(E,F) := A$.

Esimerkki 2.7. (a) Matriisia

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

vastaa lineaarikuvaus $L : \mathbb{R}^2 \to \mathbb{R}^2$, jolle on

$$Lx = Ax = \begin{bmatrix} 1 & -2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 2x_2 \\ 2x_1 - 2x_2 \end{bmatrix}$$

eli kuvaus $L : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (x_1 - 2x_2, 2x_1 - 2x_2)$. Tämä pätee myös kääntäen eli on $\operatorname{mat} L = A$.

(b) Lineaarikuvausta

$$L : \mathbb{R}^4 \to \mathbb{R}^2 : x \mapsto (x_1 + x_2 + 2x_3 + 2x_4, -x_1 - 2x_3)$$

vastaa matriisi

$$A = \operatorname{mat} L = \begin{bmatrix} 1 & 1 & 2 & 2 \\ -1 & 0 & -2 & 0 \end{bmatrix},$$

mikä pätee taas kääntäenkin.

Määritelmä 2.8. Matriisi

$$A = A_{m\times n} = [a_{ki}] = \begin{bmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{k1} & \dots & a_{ki} & \dots & a_{kn} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mn} \end{bmatrix}$$

koostuu *sarakkeista*

$$\begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix} \leftrightarrow \bar{a}_i = (a_{1i}, \dots, a_{mi}) \in \mathbb{R}^m, \quad i = 1, \dots, n,$$

jotka ovat vektoreina *sarakevektoreita*, ja *riveistä*

$$[a_{k1} \ \dots \ a_{kn}] \leftrightarrow \vec{a}_k = (a_{k1}, \dots, a_{kn}) \in \mathbb{R}^n, \quad k = 1, \dots, m,$$

eli *rivivektoreista*. Näiden avulla voidaan kirjoittaa

$$A = [\bar{a}_1 \ \dots \ \bar{a}_n] = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix}.$$

Lause 2.9. *Olkoon matriisi*

$$A = A_{m\times n} = [a_{ki}] = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

ja olkoon $L = L_A : \mathbb{R}^n \to \mathbb{R}^m$ *sitä vastaava lineaarikuvaus. Tällöin avaruuden* $\mathbb{R}^n$ *luonnollisille kantavektoreille* $\bar{e}_1, \dots, \bar{e}_n$ *pätee*

$$L\bar{e}_i = \bar{a}_i, \quad i = 1, \dots, n,$$

ja siten on myös voimassa

$$L(\mathbb{R}^n) = \langle \bar{a}_1, \dots, \bar{a}_n \rangle \,,$$

kun $\bar{a}_i$ *on matriisin* A *sarakevektori. Samoin lineaarikuvaukselle* $L : \mathbb{R}^n \to \mathbb{R}^m$ *ja sitä vastaavalle matriisille pätee*

$$\operatorname{mat} L = A = [a_{ki}].$$

Todistus. Lause seuraa suoraan kohdasta 2.3 ja määritelmästä 2.8. □

Esimerkki 2.10. Esimerkin 2.7 (b)-kohdan matriisin

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

sarakevektorit ovat $\bar{a}_1 = (1, -1) \in \mathbb{R}^2$, $\bar{a}_2 = (1, 0) \in \mathbb{R}^2$, $\bar{a}_3 = (2, -2) \in \mathbb{R}^2$ ja $\bar{a}_4 = (2, 0) \in \mathbb{R}^2$, ja rivivektorit ovat $\vec{a}_1 = (1, 1, 2, 2) \in \mathbb{R}^4$ ja $\vec{a}_2 = (-1, 0, -2, 0) \in \mathbb{R}^4$. Lineaarikuvauksen $L : \mathbb{R}^4 \to \mathbb{R}^2$ arvojoukko on

$$L(\mathbb{R}^4) = \langle (1, -1), (1, 0), (2, -2), (2, 0) \rangle = \mathbb{R}^2.$$

Koska matriisilla on selvä yhteys sitä vastaavaan lineaarikuvaukseen, on järkevää asettaa seuraava määritelmä. (Vrt. huom. 2.12.)

MÄÄRITELMÄ 2.11. Matriisien $A = A_{m\times n} = [a_{ki}]$ ja $B = B_{m\times n} = [b_{ki}]$ *summa* ja *kertominen reaaliluvulla* λ määritellään asettamalla

$$A + B := [a_{ki} + b_{ki}]_{k,i} \quad \text{ja} \quad \lambda A := [\lambda a_{ki}]_{k,i}.$$

HUOMAUTUS 2.12.

(a) Edellisen määritelmän 2.11 nojalla *kaikkien* $m \times n$*-matriisien joukko* $M(m, n)$ on lineaariavaruus (ks. määritelmä [6, s. 2]). Lineaarikuvausten joukko

$$L(\mathbb{R}^n; \mathbb{R}^m) := \{L : \mathbb{R}^n \to \mathbb{R}^m : L \text{ lineaarinen}\}$$

on myös lineaarinen avaruus, jolloin matriisivastaavuuden nojalla lineaariavaruuksien $L(\mathbb{R}^n; \mathbb{R}^m)$ ja $M(m, n)$ alkiot vastaavat bijektiivisesti toisiaan.

(b) Käyttämällä määritelmää 2.11 ja matriisin yhteyttä sitä vastaavaan lineaarikuvaukseen saadaan yhteys, joka motivoi määritelmän 2.11:

$$L_{A+B} = L_A + L_B \quad \text{ja} \quad L_{\lambda A} = \lambda L_A.$$

Tämä pätee kääntäenkin eli lineaarikuvauksille $L : \mathbb{R}^n \to \mathbb{R}^m$ ja $K : \mathbb{R}^n \to \mathbb{R}^m$ pätee

$$\operatorname{mat}(L + K) = \operatorname{mat} L + \operatorname{mat} K \quad \text{ja} \quad \operatorname{mat}(\lambda L) = \lambda \operatorname{mat} L,$$

kun $\lambda \in \mathbb{R}$.

Seuraava lineaarikuvauksien avulla tehty tarkastelu johtaa *matriisitulon* määritelmään 2.14.

2.13. Olkoot $L : \mathbb{R}^n \to \mathbb{R}^m$ ja $K : \mathbb{R}^m \to \mathbb{R}^l$ lineaarikuvauksia ja $KL : \mathbb{R}^n \to \mathbb{R}^l$ niiden yhdistetty kuvaus. Olkoot euklidisten avaruuksien $\mathbb{R}^n$, $\mathbb{R}^m$ ja $\mathbb{R}^l$ euklidiset kannat vastaavasti $\{\bar{e}_1, \ldots, \bar{e}_n\}$, $\{\bar{f}_1, \ldots, \bar{f}_m\}$ ja $\{\bar{g}_1, \ldots, \bar{g}_l\}$. Merkitään nyt näiden kantojen avulla

$$A = \operatorname{mat} K = [a_{jk}], \;\; B = \operatorname{mat} L = [b_{ki}] \;\; \text{ja} \;\; C = \operatorname{mat}(KL) = [c_{ji}].$$

Tällöin voidaan kohdan 2.6 nojalla kirjoittaa

$$c_{ji} = (KL\bar{e}_i|\bar{g}_j),$$

ja edelleen

$$\begin{aligned} KL\bar{e}_i &= K\left(\sum_{k=1}^{m} b_{ki}\bar{f}_k\right) = \sum_{k=1}^{m} b_{ki}(K\bar{f}_k) \\ &= \sum_{k=1}^{m} b_{ki}\left(\sum_{p=1}^{l} a_{pk}\bar{g}_p\right) \\ &= \sum_{p=1}^{l}\left(\sum_{k=1}^{m} a_{pk}b_{ki}\right)\bar{g}_p, \end{aligned}$$

joten on

$$\begin{aligned} c_{ji} &= (KLe_i|\bar{g}_j) \\ &= \left((\sum_{p=1}^{l} \left(\sum_{k=1}^{m} a_{pk} b_{ki} \right) \bar{g}_p)|\bar{g}_j \right) \\ &= \sum_{k=1}^{m} a_{jk} b_{ki}. \end{aligned}$$

Tämän takia on järkevää asettaa seuraava matriisitulon määritelmä.

MÄÄRITELMÄ 2.14. Kahden matriisin

$$A = A_{l\times m} = [a_{jk}] \quad \text{ja} \quad B = B_{m\times n} = [b_{ki}]$$

tulo AB on $l \times n$-matriisi $AB = [c_{ji}]$, missä

$$c_{ji} = \sum_{k=1}^{m} a_{jk} b_{ki}, \quad j = 1, \dots, l, \quad i = 1 \dots, n.$$

HUOMAUTUS 2.15. Jotta matriisien tulo voidaan muodostaa, tulee matriisin A sarakkeiden ja matriisin B rivien lukumäärän olla sama. Edellisessä määritelmässä 2.14 kyseinen lukumäärä on m, jolloin matriisin A rivivektorit $\vec{a}_j$ ja matriisin B sarakevektorit $\bar{b}_i$ ovat avaruuden $\mathbb{R}^m$ vektoreita. Nyt vektoreiden sisätulon avulla saadaan matriisien A ja B tulo

$$AB = \begin{bmatrix} (\vec{a}_1|\bar{b}_1) & \dots & (\vec{a}_1|\bar{b}_n) \\ \vdots & & \vdots \\ (\vec{a}_l|\bar{b}_1) & \dots & (\vec{a}_l|\bar{b}_n) \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_l \end{bmatrix} [\bar{b}_1 \ \dots \ \bar{b}_n],$$

ja erityisesti on $AB = [A\bar{b}_1 \ \dots \ A\bar{b}_n]$.

ESIMERKKI 2.16. Matriisien

$$A = A_{2\times 3} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \quad \text{ja} \quad B = B_{3\times 1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

tulo on

$$AB = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 14 \\ 10 \end{bmatrix},$$

mutta tuloa BA ei ole määritelty.

HUOMAUTUS 2.17. Kohdan 2.13 oletuksin ja tarkastelun mukaan saadaan

$$\operatorname{mat}(KL) = (\operatorname{mat} K)(\operatorname{mat} L).$$

Tämä pätee myös kääntäen eli matriiseille

$$A = A_{l\times m} = [a_{jk}], \quad B = B_{m\times n} = [b_{ki}] \quad \text{ja} \quad AB = [c_{ji}]$$

on

$$L_{AB} = L_A L_B,$$

koska kaikille $x \in \mathbb{R}^n$ pätee

$$\begin{aligned}
L_A L_B x &= L_A \left(\sum_{k=1}^{m} \left(\sum_{i=1}^{n} b_{ki} x_i \right) \bar{f}_k \right) \\
&= \sum_{j=1}^{l} \left(\sum_{k=1}^{m} a_{jk} \left(\sum_{i=1}^{n} b_{ki} x_i \right) \right) \bar{g}_j \\
&= \sum_{j=1}^{l} \left(\sum_{i=1}^{n} \left(\sum_{k=1}^{m} a_{jk} b_{ki} \right) x_i \right) \bar{g}_j \\
&= \sum_{j=1}^{l} \left(\sum_{i=1}^{n} c_{ji} x_i \right) \bar{g}_j \\
&= L_{AB} x.
\end{aligned}$$

Seuraus 2.18. *Matriisille $A = A_{m \times n}$ pätee $AI_n = A = I_m A$.*

Todistus. Ks. [6, s. 18]. □

Huomautus 2.19. Merkinnällä $I = I_n$ tarkoitetaan *yksikkömatriisia*

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

jota vastaava lineaarikuvaus on *identtinen kuvaus*

$$L_I = Id : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto x.$$

2.20. Matriisitulon määritelmää 2.14 toistaen saadaan matriiseille $A = A_{l \times m}$, $B = B_{m \times n}$ ja $C = C_{n \times p}$ tulot $(AB)C$ ja $A(BC)$. Huomatuksen 2.17 kautta voidaan todeta, että matriisitulo ei riipu ryhmittelystä, vaan pätee $(AB)C = A(BC)$ ja siten tulo voidaan kirjoittaa myös muodossa ABC.

3. Käänteismatriisi ja transponoitu matriisi

Käänteismatriisin määritelmään 2.23 johdutaan seuraavan lineaarikuvauksen bijektiivisyyden tarkastelun kautta.

2.21. Olkoot $L : \mathbb{R}^n \to \mathbb{R}^m$ lineaarikuvaus ja $A = \operatorname{mat} L$ sitä vastaava matriisi. Jotta kuvaus L olisi surjektio, tulisi yhtälöllä

$$Ax = y$$

olla ratkaisu kaikille $y \in \mathbb{R}^m$. Tämä on mahdollista vain jos $m \leq n$.

Jotta kuvaus L olisi injektio, tulisi yhtälöllä

$$Ax = 0$$

olla vain triviaaliratkaisu. Tämä on mahdollista vain jos $m \geq n$. Nyt siis kuvaus L voi olla bijektio vain jos $m = n$ eli kuvauksena $L : \mathbb{R}^n \to \mathbb{R}^n$.

Jos nyt kuvaus $L : \mathbb{R}^n \to \mathbb{R}^n$ on bijektio, niin sillä on käänteiskuvaus $L^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, joka on myös lineaarikuvaus. Tällöin saadaan

$$L^{-1} L = Id \quad \text{ja} \quad LL^{-1} = Id.$$

Huomaa, että vain bijektiivisellä kuvauksella on käänteiskuvaus.

Lineaarikuvauksia vastaaville neliömatriiseille

$$A = A_{n\times n} = \operatorname{mat} L \quad \text{ja} \quad B = B_{n\times n} = \operatorname{mat}(L^{-1})$$

saadaan

$$BA = I_n = I \quad \text{ja} \quad AB = I_n = I.$$

Esimerkki 2.22. Esimerkin 2.7 (a)-kohdan lineaarikuvaus

$$L : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (x_1 - 2x_2, 2x_1 - 2x_2)$$

on bijektio, joten sille löytyy käänteiskuvaus

$$L^{-1} : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (-x_1 + x_2, -x_1 + \frac{1}{2}x_2),$$

ja pätee

$$LL^{-1} : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (x_1, x_2) \quad \text{ja} \quad L^{-1}L : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (x_1, x_2).$$

Lineaarikuvauksia L ja L^{-1} vastaavat matriisit ovat

$$\operatorname{mat} L = A = \begin{bmatrix} 1 & -2 \\ 2 & -2 \end{bmatrix} \quad \text{ja} \quad \operatorname{mat}(L^{-1}) = B = \begin{bmatrix} -1 & 1 \\ -1 & \frac{1}{2} \end{bmatrix},$$

joiden tuloille pätee

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

ja vastaavasti on

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

Määritelmä 2.23. Neliömatriisi $A = A_{n\times n}$ on *kääntyvä*, jos löytyy sellainen matriisi $B = B_{n\times n}$, että

$$AB = I \quad \text{ja} \quad BA = I.$$

Tällöin merkitään $B =: A^{-1}$ ja sanotaan, että A^{-1} on matriisin A *käänteismatriisi.*

Huomautus 2.24. Käänteismatriisi on yksikäsitteinen ja siten sen määritelmä on hyvin asetettu.

Todistus. Ks. [6, s. 23]. □

Esimerkki 2.25. Esimerkin 2.22 matriisi A on kääntyvä ja sen käänteismatriisi on $B = A^{-1}$.

Lineaarikuvauksen käänteiskuvauksella ja matriisin käänteismatriisilla on taas yhteys toisiinsa. Tätä yhteyttä käsitellään seuraavassa lauseessa.

Lause 2.26. *Olkoot $L : \mathbb{R}^n \to \mathbb{R}^n$ lineaarikuvaus ja $A = \operatorname{mat} L$ sitä vastaava matriisi. Tällöin on voimassa seuraavat väitteet:*

(i) *Jos lineaarikuvaus $L : \mathbb{R}^n \to \mathbb{R}^n$ on bijektio, niin $\operatorname{mat} L$ on kääntyvä ja*

$$(\operatorname{mat} L)^{-1} = \operatorname{mat}(L^{-1}).$$

(ii) *Jos matriisi $A = A_{n\times n}$ on kääntyvä, niin lineaarikuvaus L_A on bijektio ja*

$$(L_A)^{-1} = L_{A^{-1}}.$$

Todistus. (Ks. [6, s. 23].)

(i) Tämä on käsitelty jo kohdassa 2.21, sillä pätee

$$(\operatorname{mat} L)^{-1} = A^{-1} = B = \operatorname{mat}(L^{-1}).$$

(ii) Kääntyvälle matriisille A on voimassa seuraavat yhtälöt $A^{-1}A = I$ ja $AA^{-1} = I$. Huomautuksen 2.17 nojalla saadaan

$$Id = L_{A^{-1}A} = L_{A^{-1}}L_A \quad \text{ja} \quad Id = L_{AA^{-1}} = L_A L_{A^{-1}},$$

eli L_A on bijektio ja $(L_A)^{-1} = L_{A^{-1}}$.

□

Seuraava *transponoidun matriisin* määritelmä on tarpeen, kun tarkastellaan matriisin ominaisuuksia sekä rivivektoreiden että sarakevektoreiden suhteen. Tällaisia tapauksia on mm. käänteismatriisin ja determinantin (käsitellään myöhemmin) yhteydessä. Määritelmän jälkeen tulevaa lausetta tarvitaan myös myöhemmin.

Määritelmä 2.27. Matriisin $A = A_{m\times n}$ *transponoitu matriisi* eli *peilattu matriisi* on $n \times m$-matriisi $A^T = A^T_{n\times m}$. Sen alkioille pätee

$$[A^T]_{ik} = [A]_{ki} := a_{ki} \quad \text{kaikille} \quad i = 1, \dots, n, \quad k = 1, \dots, m.$$

Jos on matriisi

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix},$$

niin sen transponoitu matriisi on siis

$$A^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}.$$

Huomataan, että matriisin A rivi k on transponoidun matriisin A^T sarake k ja matriisin A sarake i on transponoidun matriisin A^T rivi i.

Lause 2.28. *Olkoot matriisit $A = A_{m\times n}$ ja $B = B_{n\times l}$. Tällöin on voimassa seuraavat väitteet:*

(i) *$(A^T)^T = A$ ja*

(ii) *$(AB)^T = B^T A^T$.*

Todistus. Ensimmäinen väite seuraa suoraan määritelmästä 2.27 ja toinen väite saadaan kirjoittamalla

$$\begin{aligned}[(AB)^T]_{ij} &= [AB]_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b_{ki}a_{jk} \\ &= \sum_{k=1}^{n} [B^T]_{ik}[A^T]_{kj} = [B^T A^T]_{ij}.\end{aligned}$$

□

LUKU 3

Determinantti

1. Determinantti-käsite ja determinantin perusominaisuudet

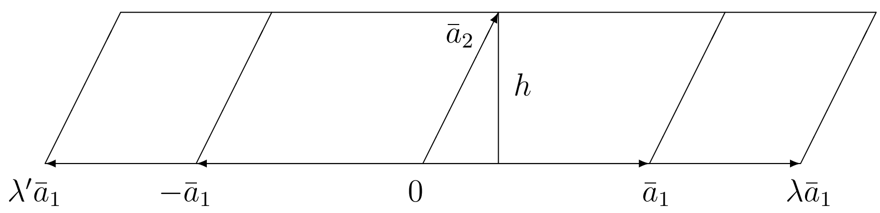
Determinantti-käsitetteeseen voidaan tutustua esimerkiksi Cauchyn geometrisen tilavuustulkinnan avulla. Determinantin yleiseen määritelmään päästään tarkastelemalla determinantin karakteristisia ominaisuuksia, jotka ovat samat kuin merkillä varustetun sunntaissärmiön tilavuuden ominaisuudet. Aloitetaan geometrialla.

3.1. **Kaksiulotteinen suuntaissärmiö eli suunnikas.** Olkoon $D : A \mapsto D(A)$ reaaliarvoinen funktio, joka liittää matriisiin $A = A_{2\times 2} = [\bar{a}_1\ \bar{a}_2]$ luvun

$$D(\bar{a}_1, \bar{a}_2) := D(A) = \pm \operatorname{ala}(\bar{a}_1, \bar{a}_2),$$

missä etumerkki tulee kiertokulmasta $\sphericalangle(\bar{a}_1, \bar{a}_2)$ (väliltä $[0, \pi]$) ja merkinnällä $\operatorname{ala}(\bar{a}_1, \bar{a}_2)$ tarkoitetaan vektoreiden $\bar{a}_1$ ja $\bar{a}_2$ virittämän suunnikkaan pinta-alaa. Tällöin pätee

(i) $D(\lambda\bar{a}_1, \bar{a}_2) = \lambda D(\bar{a}_1, \bar{a}_2)$ kaikille $\lambda \in \mathbb{R}$ ja
(ii) $D(\bar{u} + \bar{v}, \bar{a}_2) = D(\bar{u}, \bar{a}_2) + D(\bar{v}, \bar{a}_2)$, kun $\bar{u} \in \mathbb{R}^2$ ja $\bar{v} \in \mathbb{R}^2$.



Kuva (1)

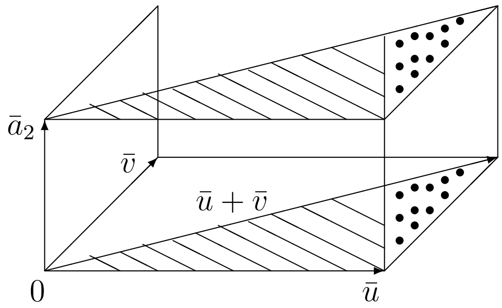
Perustelu:

(i) Kuvasta (1) nähdään, että kun λ on positiivinen reaaliluku, pinta-ala $\operatorname{ala}(\lambda\bar{a}_1, \bar{a}_2)$ saadaan kertomalla kannan $\lambda\bar{a}_1$ pituus eli normi $\|\lambda\bar{a}_1\|$ korkeudella h, jolloin pätee $\operatorname{ala}(\lambda\bar{a}_1, \bar{a}_2) = \|\lambda\bar{a}_1\|\, h = \lambda\, \|\bar{a}_1\|\, h$. Vastaavasti saadaan $\lambda \operatorname{ala}(\bar{a}_1, \bar{a}_2) = \lambda\, \|\bar{a}_1\|\, h$, jolloin on voimassa $\operatorname{ala}(\lambda\bar{a}_1, \bar{a}_2) = \lambda \operatorname{ala}(\bar{a}_1, \bar{a}_2)$. Koska tässä on $\sphericalangle(\bar{a}_1, \bar{a}_2) = \sphericalangle(\lambda\bar{a}_1, \bar{a}_2)$, pätee $D(\lambda\bar{a}_1, \bar{a}_2) = \lambda D(\bar{a}_1, \bar{a}_2)$.

Jos λ' on negatiivinen reaaliluku, niin tällöinkin pätee $\operatorname{ala}(\lambda'\bar{a}_1, \bar{a}_2) = \|\lambda'\bar{a}_1\|\, h$. Vastaavasti on voimassa

$$|\lambda' \operatorname{ala}(\bar{a}_1, \bar{a}_2)| = |\lambda'\, \|\bar{a}_1\|\, h| = |\lambda'|\, \|\bar{a}_1\|\, h = \|\lambda'\bar{a}_1\|\, h,$$

jolloin pätee $\operatorname{ala}(\lambda'\bar{a}_1, \bar{a}_2) = |\lambda' \operatorname{ala}(\bar{a}_1, \bar{a}_2)| = |\lambda'| \operatorname{ala}(\bar{a}_1, \bar{a}_2)$. Tällöin (i)-kohta on tosi myös negatiiviselle reaaliluvulle λ', sillä kulman suunta eli merkki vaihtuu: $\sphericalangle(\bar{a}_1, \bar{a}_2)$ on eri suuntainen kuin $\sphericalangle(-\bar{a}_1, \bar{a}_2) = \sphericalangle(\lambda\bar{a}_1, \bar{a}_2)$.

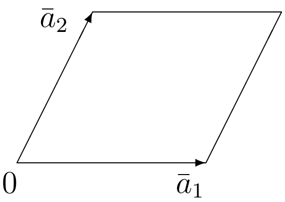


Kuva (2)

(ii) Kuvasta (2) nähdään, että on $\mathrm{ala}(\bar{u}+\bar{v},\bar{a}_2) = \mathrm{ala}(\bar{u},\bar{a}_2) + \mathrm{ala}(\bar{v},\bar{a}_2)$, koska viivoitetut pinta-alat vastaavat toisiaan ja samoin pisteelliset pinta-alat vastaavat toisiaan. Tällöin (ii)-kohta on tosi.

Vastaavasti kohdat (i) ja (ii) pätevät myös toisen sarakkeen suhteen, joten saadaan:

(1) Funktio D on matriisin A molempien sarakkeiden suhteen erikseen lineaarinen.



Kuva (3)

Kuvasta (3) nähdään, että kulmat $\sphericalangle(\bar{a}_1,\bar{a}_2)$ ja $\sphericalangle(\bar{a}_2,\bar{a}_1)$ ovat vastakkaismerkkiset. Lisäksi tietenkin on

$$\mathrm{ala}(\bar{a}_2,\bar{a}_1) = \mathrm{ala}(\bar{a}_1,\bar{a}_2),$$

josta seuraa suoraan, että

(2) $D(\bar{a}_2,\bar{a}_1) = -D(\bar{a}_1,\bar{a}_2)$.

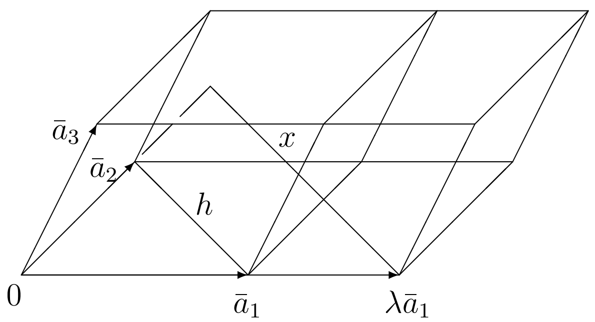
3.2. **Suuntaissärmiö.** Tarkastellaan lyhyesti matriisia $A = A_{3\times 3} = [\bar{a}_1\ \bar{a}_2\ \bar{a}_3]$, jonka avulla kaikki oleellinen saadaan näkyville. Tarkastellaan reaaliarvoista funktiota $D : A \mapsto D(A)$ siten, että pätee

$$D(\bar{a}_1,\bar{a}_2,\bar{a}_3) := D(A) = \pm\,\mathrm{vol}(\bar{a}_1,\bar{a}_2,\bar{a}_3),$$

missä etumerkki riippuu vektoreiden järjestyksestä ja merkinnällä $\mathrm{vol}(\bar{a}_1,\bar{a}_2,\bar{a}_3)$ tarkoitetaan vektoreiden $\bar{a}_1$, $\bar{a}_2$ ja $\bar{a}_3$ virittämän suuntaissärmiön tilavuutta. Osoitetaan, että funktion D määritelmässä esiintyvän merkin $\pm$ voi valita, siten että pätee:

(i) $D(\lambda\bar{a}_1,\bar{a}_2,\bar{a}_3) = \lambda D(\bar{a}_1,\bar{a}_2,\bar{a}_3)$ kaikille $\lambda\in\mathbb{R}$ ja
(ii) $D(\bar{u}+\bar{v},\bar{a}_2,\bar{a}_3) = D(\bar{u},\bar{a}_2,\bar{a}_3) + D(\bar{v},\bar{a}_2,\bar{a}_3)$, kun $\bar{u}\in\mathbb{R}^3$ ja $\bar{v}\in\mathbb{R}^3$.

Perustellaan ensin, miksi (i) ja (ii) pätevät ainakin molempien puolien itseisarvoille.



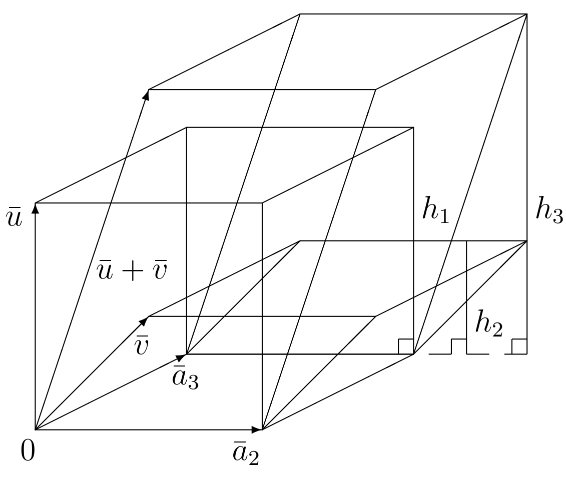
Kuva (4)

Kuvassa (4), kun λ on positiivinen reaaliluku, vektoreiden $\lambda\bar{a}_1$, $\bar{a}_2$ ja $\bar{a}_3$ sekä vektoreiden $\bar{a}_1$, $\bar{a}_2$ ja $\bar{a}_3$ muodostamien suuntaissärmiöiden tilavuudet riippuvat samasta pohjan pinta-alasta, jonka virittävät vektorit $\bar{a}_2$ ja $\bar{a}_3$. Korkeus h poikkeaa vakiolla λ kertomisen verran, koska muodostuu yhdenmuotoiset kolmiot, joiden sivujen pituuksien suhteet vastaavat toisiaan. Tällöin on voimassa

$$\frac{h}{\|\bar{a}_1\|} = \frac{x}{\lambda\,\|\bar{a}_1\|},$$

josta saadaan $x = \lambda h$. Nyt jos tilavuus $\mathrm{vol}(\bar{a}_1, \bar{a}_2, \bar{a}_3)$ kerrotaan vakiolla λ, saadaan $\lambda\,\mathrm{vol}(\bar{a}_1, \bar{a}_2, \bar{a}_3) = \mathrm{vol}(\lambda\bar{a}_1, \bar{a}_2, \bar{a}_3)$. Merkki $\pm$ on siis viisasta sopia valittavaksi siten, että se ei muutu, kun vektorit säilyttävät suuntansa, jolloin pätee
$\lambda D(\bar{a}_1, \bar{a}_2, \bar{a}_3) = D(\lambda\bar{a}_1, \bar{a}_2, \bar{a}_3)$.

Jotta tapaus menisi negatiiviselle λ samalla tavalla kuin 2×2-matriisin tapauksessa, täytyy valita, että yhden (sarake)vektorin suunnan kääntäminen päinvastaiseksi vaihtaa merkin myös 3×3-matriisin tapauksessa. Tämä on geometrisesti perusteltavissa: vektorikolmikon $(\bar{a}_1, \bar{a}_2, \bar{a}_3)$ "kätisyys" vaihtuu tässä toimenpiteessä.



Kuva (5)

Kuvassa (5) vektoreiden $\bar{u} + \bar{v}$, $\bar{a}_2$, $\bar{a}_3$, ja $\bar{u}$, $\bar{a}_2$, $\bar{a}_3$ sekä $\bar{v}$, $\bar{a}_2$, $\bar{a}_3$ virittämillä suuntaissärmiöillä on sama pohjan pinta-ala $\mathrm{ala}(\bar{a}_2, \bar{a}_3)$ ja vektorikolmikoilla $(\bar{u}, \bar{a}_2, \bar{a}_3)$ ja

$(\bar{v}, \bar{a}_2, \bar{a}_3)$ on sama kiertosuunta. Suuntaissärmiön $(\bar{u}, \bar{a}_2, \bar{a}_3)$ tilavuus saadaan kertomalla ala$(\bar{a}_2, \bar{a}_3)$ korkeudella h_1, joka on vektorin $\bar{u}$ ortogonaalinen komponentti. Vastaavasti suuntaissärmiön $(\bar{v}, \bar{a}_2, \bar{a}_3)$ tilavuus riippuu pohjan pinta-alasta ala$(\bar{a}_2, \bar{a}_3)$ ja korkeudesta h_2, joka on vektorin $\bar{v}$ ortogonaalinen komponentti, ja suuntaissärmiön $(\bar{u}+\bar{v}, \bar{a}_2, \bar{a}_3)$ tilavuus tulee pohjan pinta-alasta ala$(\bar{a}_2, \bar{a}_3)$ ja korkeudesta h_3, joka on vektorin $\bar{u}+\bar{v}$ ortogonaalinen komponentti. Huomataan, että korkeuksille h_1, h_2 ja h_3 pätee $h_3 = h_1 + h_2$. Tällöin saadaan

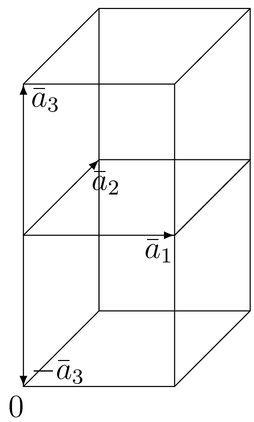
$$\operatorname{vol}(\bar{u}+\bar{v}, \bar{a}_2, \bar{a}_3) = \operatorname{vol}(\bar{u}, \bar{a}_2, \bar{a}_3) + \operatorname{vol}(\bar{v}, \bar{a}_2, \bar{a}_3)$$

ja siten on

$$D(\bar{u}+\bar{v}, \bar{a}_2, \bar{a}_3) = D(\bar{u}, \bar{a}_2, \bar{a}_3) + D(\bar{v}, \bar{a}_2, \bar{a}_3).$$

Vastaava pätee myös toisen ja kolmannen sarakkeen suhteen, joten saadaan:

(1) Funktio D on matriisin A kaikkien sarakkeiden suhteen erikseen lineaarinen.



Kuva (6)

Kuvasta (6) nähdään, että on voimassa

$$-\operatorname{vol}(\bar{a}_3, \bar{a}_2, \bar{a}_1) = \operatorname{vol}(-\bar{a}_3, \bar{a}_1, \bar{a}_2) = \operatorname{vol}(\bar{a}_1, \bar{a}_2, \bar{a}_3),$$

jolloin pätee

$$\operatorname{vol}(\bar{a}_3, \bar{a}_2, \bar{a}_1) = -\operatorname{vol}(\bar{a}_1, \bar{a}_2, \bar{a}_3).$$

Vastaavasti voidaan todeta, että pätee

$$\operatorname{vol}(\bar{a}_1, \bar{a}_2, \bar{a}_3) = \operatorname{vol}(\bar{a}_3, \bar{a}_1, \bar{a}_2) = \operatorname{vol}(\bar{a}_2, \bar{a}_3, \bar{a}_1).$$

Näistä seuraa

(2) $D(\bar{a}_3, \bar{a}_2, \bar{a}_1) = -D(\bar{a}_1, \bar{a}_2, \bar{a}_3)$ ja
$D(\bar{a}_1, \bar{a}_2, \bar{a}_3) = D(\bar{a}_3, \bar{a}_1, \bar{a}_2) = D(\bar{a}_2, \bar{a}_3, \bar{a}_1)$.

3.3. **Karakteristiset ominaisuudet.** Tarkastellaan determinantin karakteristisia ominaisuuksia 2×2, 3×3 ja sitten yleisesti $n \times n$ -matriiseille.

Oletetaan, että D on neliömatriiseille $A = A_{2\times 2} = [\bar{a}_1\ \bar{a}_2] \in M(2,2)$ määritelty reaaliarvoinen funktio

$$D : A \mapsto D(A) =: D(\bar{a}_1, \bar{a}_2),$$

joka toteuttaa ehdot (D1) ja (D2):

(D1) D on *multilineaarinen* eli se on erikseen molempien sarakkeiden suhteen lineaarinen;

(D2) D on *alternoiva* eli *vuorotteleva*, jolloin lausekkeen $D(\bar{a}_1, \bar{a}_2)$ merkki vaihtuu, jos sarakkeet vaihdetaan keskenään.

Huomataan, että ehdon (D2) nojalla pätee $D(A) = 0$, jos matriisin A molemmat sarakkeet ovat samat. Kirjoitetaan sarakkeet $\bar{a}_i$ jatkoa varten seuraavaan muotoon

$$\bar{a}_i = (a_{1,i}, a_{2,i}) = \sum_{k_i=1}^{2} a_{k_i,i}\bar{e}_{k_i}.$$

Nyt ehdon (D1) nojalla saadaan

$$\begin{aligned} D(\bar{a}_1, \bar{a}_2) &= D(a_{1,1}\bar{e}_1 + a_{2,1}\bar{e}_2, a_{1,2}\bar{e}_1 + a_{2,2}\bar{e}_2) \\ &= D\left(\sum_{k_1=1}^{2} a_{k_1,1}\bar{e}_{k_1}, \sum_{k_2=1}^{2} a_{k_2,2}, \bar{e}_{k_2}\right) \\ &= \sum_{k_1=1}^{2}\sum_{k_2=1}^{2} a_{k_1,1}a_{k_2,2}D(\bar{e}_{k_1}, \bar{e}_{k_2}), \end{aligned} \tag{3.1}$$

missä $D(\bar{e}_{k_1}, \bar{e}_{k_2}) = 0$, jos toinen k_i esiintyy kaksi kertaa. Nyt riittää, että summataan yli niiden indeksien k_1, k_2, jotka saadaan luvuista $1, 2$ peräkkäisillä vaihdoilla ts. permutoimalla, eli summataan yli indeksien $(k_1, k_2) \in P_2$, kun $P_2 = \mathrm{Perm}(1, 2)$ merkitsee lukujen 1 ja 2 permutaatioiden joukkoa. Siten lauseke (3.1) saadaan ominaisuuden (D2) nojalla muotoon

$$D(\bar{a}_1, \bar{a}_2) = \sum_{(k_1,k_2)\in P_2} a_{k_1,1}a_{k_2,2}\,\mathrm{sign}(k_1, k_2)D(\bar{e}_1, \bar{e}_2), \tag{3.2}$$

missä *merkkifunktio* $\mathrm{sign}(k_1, k_2)$ on 1 vast. -1, kun (k_1, k_2) saadaan perusjonosta $(1, 2)$ parillisella vast. parittomalla määrällä vaihtoja. Lausekkeen (3.2) termi $D(\bar{e}_1, \bar{e}_2)$ on vakio. Determinantin tilavuustulkinnassa tämä vakio luonnollisesti kiinnitetään *normeeraavalla ehdolla* (D3) seuraavasti:

(D3) $D(\bar{e}_1, \bar{e}_2) = 1$.

Tällöin matriisille $A \in M(2, 2)$ määritelty funktio $D : A \mapsto D(A)$ varustettuna ominaisuuksilla (D1)-(D3) on muotoa

$$D(\bar{a}_1, \bar{a}_2) = \sum_{(k_1,k_2)\in P_2} \mathrm{sign}(k_1, k_2)a_{k_1,1}a_{k_2,2} = (a_{11}a_{22} - a_{12}a_{21}).$$

Käsitellään lyhyesti karakteristisia ominaisuuksia 3×3-matriisille. Oletetaan, että D on neliömatriiseille $A = A_{3\times 3} = [\bar{a}_1\ \bar{a}_2\ \bar{a}_3] \in M(3, 3)$ määritelty reaaliarvoinen funktio

$$D : A \mapsto D(A) =: D(\bar{a}_1, \bar{a}_2, \bar{a}_3),$$

joka toteuttaa ehdot (D1) ja (D2):

(D1) D on *multilineaarinen* eli se on erikseen jokaisen sarakkeiden suhteen lineaarinen;

(D2) D on *alternoiva* eli *vuorotteleva*, jolloin lausekkeen $D(\bar{a}_1, \bar{a}_2, \bar{a}_3)$ merkki vaihtuu, jos kaksi saraketta vaihdetaan keskenään.

Huomataan, että ehdon (D2) nojalla pätee $D(A) = 0$, jos matriisissa A on kaksi tai kolme samaa saraketta. Kirjoitetaan sarakkeet $\bar{a}_i$ jatkoa varten seuraavaan muotoon

$$\bar{a}_i = (a_{1,i}, a_{2,i}, a_{3,i}) = \sum_{k_i=1}^{3} a_{k_i,i} \bar{e}_{k_i}.$$

Nyt ehdon (D1) nojalla saadaan

$$\begin{aligned} D(\bar{a}_1, \bar{a}_2, \bar{a}_3) &= D(a_{1,1}\bar{e}_1 + a_{2,1}\bar{e}_2 + a_{3,1}\bar{e}_3, \dots, a_{1,3}\bar{e}_1 + a_{2,3}\bar{e}_2 + a_{3,3}\bar{e}_3) \\ &= D\left(\sum_{k_1=1}^{3} a_{k_1,1}\bar{e}_{k_1}, \sum_{k_2=1}^{3} a_{k_2,2}\bar{e}_{k_2}, \sum_{k_3=1}^{3} a_{k_3,3}\bar{e}_{k_3} \right) \\ &= \sum_{k_1=1}^{3} \sum_{k_2=1}^{3} \sum_{k_3=1}^{3} a_{k_1,1} a_{k_2,2} a_{k_3,3} D(\bar{e}_{k_1}, \bar{e}_{k_2}, \bar{e}_{k_3}). \end{aligned}$$

Kiinnitetään ehto (D3) seuraavasti:

(D3) $D(\bar{e}_1, \bar{e}_2, \bar{e}_3) = 1$,

jolloin ehtojen (D1)-(D3) nojalla saadaan

$$D(\bar{a}_1, \bar{a}_2, \bar{a}_3) = \sum_{(k_1,k_2,k_3) \in P_3} \operatorname{sign}(k_1, k_2, k_3) a_{k_1,1} a_{k_2,2} a_{k_3,3}.$$

Otetaan yleinen tapaus $n \times n$-matriiseille. Oletetaan, että D on neliömatriiseille

$$A = A_{n \times n} = [\bar{a}_1 \ \dots \ \bar{a}_n] \in M(n, n)$$

määritelty reaaliarvoinen funktio

$$D : A \mapsto D(A) =: D(\bar{a}_1, \dots, \bar{a}_n),$$

joka toteuttaa ehdot (D1) ja (D2):

(D1) D on *multilineaarinen* eli se on erikseen jokaisen sarakkeen suhteen lineaarinen;

(D2) D on *alternoiva* eli *vuorotteleva*, jolloin lausekkeen $D(\bar{a}_1, \dots, \bar{a}_n)$ merkki vaihtuu, jos kaksi saraketta vaihdetaan keskenään.

Huomataan, että ehdon (D2) nojalla pätee $D(A) = 0$, jos matriisissa A on kaksi tai useampi samaa saraketta. Kirjoitetaan sarakkeet $\bar{a}_i$ seuraavaan muotoon

$$\bar{a}_i = (a_{1,i}, a_{2,i}, \dots, a_{n,i}) = \sum_{k_i=1}^{n} a_{k_i,i} \bar{e}_{k_i}.$$

Nyt ehdon (D1) nojalla saadaan

$$\begin{aligned} D(\bar{a}_1, \dots, \bar{a}_n) &= D(a_{1,1}\bar{e}_1 + \dots + a_{n,1}\bar{e}_n, \dots, a_{1,n}\bar{e}_1 + \dots + a_{n,n}\bar{e}_n) \\ &= D\left(\sum_{k_1=1}^{n} a_{k_1,1}\bar{e}_{k_1}, \dots, \sum_{k_n=1}^{n} a_{k_n,n}\bar{e}_{k_n} \right) \\ &= \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} a_{k_1,1} \cdots a_{k_n,n} D(\bar{e}_{k_1}, \dots, \bar{e}_{k_n}), \end{aligned} \tag{3.3}$$

missä $D(\bar{e}_{k_1},\ldots,\bar{e}_{k_n}) = 0$, jos jokin k_i esiintyy useammin kuin kerran. Nyt riittää, että summataan yli niiden indeksien $k_1,\ldots,k_n$, jotka saadaan luvuista $1,\ldots,n$ permutoimalla, eli yli indeksien $(k_1,\ldots,k_n) \in P_n$, kun $P_n = \mathrm{Perm}(1,\ldots,n)$ merkitsee lukujen $1,\ldots,n$ permutaatioiden joukkoa. Siten lauseke (3.3) saadaan ominaisuuden (D2) nojalla muotoon

$$D(\bar{a}_1,\ldots,\bar{a}_n) = \sum_{(k_1,\ldots,k_n)\in P_n} a_{k_1,1}\cdots a_{k_n,n}\,\mathrm{sign}(k_1,\ldots,k_n) D(\bar{e}_1,\ldots,\bar{e}_n), \tag{3.4}$$

missä merkkifunktio $\mathrm{sign}(k_1,\ldots,k_n)$ on 1 vast. -1, kun $(k_1,\ldots,k_n)$ saadaan perusjonosta $(1,\ldots,n)$ parillisella vast. parittomalla määrällä vaihtoja. Lausekkeen (3.4) termi $D(\bar{e}_1,\ldots,\bar{e}_n)$ on vakio. Kiinnitetään tämä vakio ehdolla (D3) seuraavasti:

(D3) $D(\bar{e}_1,\ldots,\bar{e}_n) = 1$.

Tällöin matriisille $A \in M(n,n)$ määritelty funktio $D : A \mapsto D(A)$ varustettuna ominaisuuksilla (D1)-(D3) on muotoa

$$D(\bar{a}_1,\ldots,\bar{a}_n) = \sum_{(k_1,\ldots,k_n)\in P_n} \mathrm{sign}(k_1,\ldots,k_n) a_{k_1,1}\cdots a_{k_n,n}.$$

Tämä lauseke otetaan sitten determinantin yleiseksi määritelmäksi.

Määritelmä 3.4. Matriisin

$$A = A_{n\times n} = [a_{ki}] \in M(n,n)$$

determinantti on luku

$$\det A := \sum_{(k_1,\ldots,k_n)\in P_n} \mathrm{sign}(k_1,\ldots,k_n) a_{k_1,1}\cdots a_{k_n,n},$$

jota myös usein merkitään

$$\det A = |A| = \begin{vmatrix} a_{11} & \ldots & a_{1i} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{k1} & \ldots & a_{ki} & \ldots & a_{kn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & a_{ni} & \ldots & a_{nn} \end{vmatrix}.$$

Determinantin rivit/sarakkeet ovat matriisin A rivejä/sarakkeita.

Esimerkki 3.5. Matriisin

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

determinantti on

$$\det A = 1\cdot 5\cdot 9 - 1\cdot 8\cdot 6 - 4\cdot 2\cdot 9 + 4\cdot 8\cdot 3 + 7\cdot 2\cdot 6 - 7\cdot 5\cdot 3 = 0.$$

3.6. **Perusominaisuudet.** Determinantti määrittelee funktion

$$\det : M(n,n) \to \mathbb{R} : A = [\bar{a}_1 \ \ldots \ \bar{a}_n] \mapsto \det A = \det[\bar{a}_1 \ \ldots \ \bar{a}_n],$$

jolla on voimassa aiemmin käsitellyt karakteristiset ominaisuudet:

(D1) Determinantti on lineaarinen jokaisen sarakkeen suhteen erikseen. Esimerkiksi ensimmäisen sarakkeen suhteen pätee

$$\det[\bar{u}+\bar{v}\ \bar{a}_2\ \ldots\ \bar{a}_n] = \det[\bar{u}\ \bar{a}_2\ \ldots\ \bar{a}_n] + \det[\bar{v}\ \bar{a}_2\ \ldots\ \bar{a}_n]$$

ja

$$\det[\lambda\bar{a}_1\ \bar{a}_2\ \ldots\ \bar{a}_n] = \lambda\det[\bar{a}_1\ \bar{a}_2\ \ldots\ \bar{a}_n],\ \text{kun } \lambda\in\mathbb{R},$$

ja vastaavasti muille sarakkeille.

(D2) Jos matriisissa $A = [\bar{a}_1\ \ldots\ \bar{a}_n]$ vaihdetaan keskenään kaksi saraketta $\bar{a}_i$ ja $\bar{a}_j$, $i\neq j$, niin saadulle matriisille $\tilde{A}$ on $\det\tilde{A} = -\det A$.

(D3) Yksikkömatriisille I_n on $\det I_n = 1$.

Esimerkkejä 3.7. (a) Kun käytetään ominaisuutta (D1) saadaan

$$\begin{vmatrix}1&2\\3&4\end{vmatrix} = \begin{vmatrix}1+0&2\\2+1&4\end{vmatrix} = \begin{vmatrix}1&2\\2&4\end{vmatrix} + \begin{vmatrix}0&2\\1&4\end{vmatrix} = 0+(-2) = -2.$$

(b) Käyttämällä ominaisuuksia (D1) ja (D2) saadaan

$$\begin{vmatrix}1&2\\3&4\end{vmatrix} = \begin{vmatrix}1+0&2\\2+1&4\end{vmatrix} = \begin{vmatrix}1&2\\2&4\end{vmatrix} + \begin{vmatrix}0&2\\1&4\end{vmatrix} = \begin{vmatrix}1&2\\2&4\end{vmatrix} - \begin{vmatrix}2&0\\4&1\end{vmatrix} = 0-2 = -2.$$

3.8. Determinantti voidaan johtaa myös käyttämällä rivejä sarakkeiden tilalla, koska tiedetään, että matriisin $A = A_{n\times n}$ rivivektorit $\vec{a}_1,\ldots,\vec{a}_n$ ovat samat kuin sen transponoidun matriisin A^T sarakevektorit. Nyt determinantin määritelmän 3.4 nojalla pätee yhtälö

$$\det A^T = \det A,$$

koska jos on matriisi

$$A = [a_{kj}] = [a_{kj}]_{k,j} = \begin{bmatrix}\vec{a}_1\\ \vdots\\ \vec{a}_n\end{bmatrix},$$

niin sen transponoitu matriisi on $A^T = [\alpha_{jk}] = [\alpha_{jk}]_{j,k} = [\bar{\alpha}_1\ \ldots\ \bar{\alpha}_n]$, missä $\alpha_{jk} = a_{kj}$ ja

$$\bar{\alpha}_k = (\alpha_{1k},\ldots,\alpha_{nk}) = (a_{k1},\ldots,a_{kn}) = \vec{a}_k.$$

Tällöin saadaan

$$\begin{aligned}\det A^T &= \sum_{(j_1,\ldots,j_n)\in P_n} \operatorname{sign}(j_1,\ldots,j_n)\alpha_{j_1,1}\cdots\alpha_{j_n,n}\\ &= \sum_{(j_1,\ldots,j_n)\in P_n} \operatorname{sign}(j_1,\ldots,j_n)a_{1,j_1}\cdots a_{n,j_n}.\end{aligned}$$

Koska on $j_{k_1} = 1,\ldots,j_{k_n} = n$ jollekin $(k_1,\ldots,k_n)\in P_n$, niin on voimassa

$$\operatorname{sign}(k_1,\ldots,k_n) = \operatorname{sign}(j_1,\ldots,j_n).$$

Siispä saadaan

$$\det A^T = \sum_{(k_1,\ldots,k_n)\in P_n} \operatorname{sign}(k_1,\ldots,k_n)a_{k_1,1}\cdots a_{k_n,n} = \det A.$$

2. Determinantin ominaisuuksia

LAUSE 3.9. *Olkoon matriisi*

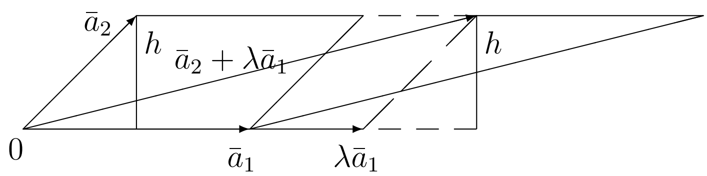
$$A = A_{n\times n} = [a_{ki}] = \begin{bmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{k1} & \dots & a_{ki} & \dots & a_{kn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nn} \end{bmatrix}.$$

(a) *Jos matriisin A sarake/rivi kerrotaan vakiolla λ, niin saadun matriisin determinantti on $\lambda \det A$.*
(b) *Jos matriisissa A on nollasarake/nollarivi, niin $\det A = 0$.*
(c) *Jos matriisissa A on kaksi samaa saraketta/riviä, niin $\det A = 0$.*
(d) *Determinantti pysyy samana, jos sarake i/rivi k kerrottuna vakiolla λ lisätään toiseen sarakkeeseen/riviin j.*
(e) *Jos sarakkeessa i/rivillä k vain alkio $a_{ki} \neq 0$, niin rivin k/sarakkeen i muiden alkioiden vaihtaminen nollaksi ei muuta determinanttia.*

TODISTUKSESTA. Motivoidaan väitteitä determinantin tilavuustulkinnan avulla.

(a) Matriisin A yhden sarakevektorin pituus muuttuu vakiolla λ kertomisen verran, kun matriisin A sarake kerrotaan vakiolla λ. Tällöin suuntaissärmiön tilavuus $\mathrm{vol}\,(\bar{a}_1, \dots, \bar{a}_n)$ muuttuu tilavuuteen $\pm\lambda \,\mathrm{vol}\,(\bar{a}_1, \dots, \bar{a}_n)$, joka siis riippuu vakion λ etumerkistä. Tällöin vakiolla λ kerrotun matriisin determinantti on $\lambda \det A$, koska determinantin tilavuustulkinnan mukaan pätee $\det A = \pm \,\mathrm{vol}\,(\bar{a}_1, \dots, \bar{a}_n)$. Tätä käsiteltiin jo kohdassa 3.1 2×2-matriisille.
(b) Tässä tilanteessa yksi sarakevektori arvoavaruudessa on surkastunut ja sen pituus on nolla. Tällöin sarakevektoreiden virättämän suuntaissärmiön tilavuus on myös nolla ja siten myös on $\det A = 0$.
(c) Jos matriisissa A on kaksi samaa saraketta, niin sen virittämän suuntaissärmiön tilavuus on nolla, joten on $\det A = 0$.

    Vastaavasti käy myös tilanteessa, jossa kaksi saraketta ovat muutoin samat, mutta toinen on kerrottu vakiolla.
(d) Esimerkiksi avaruudessa $\mathbb{R}^2$, kun ensimmäinen sarake kerrotaan vakiolla λ ja lisätään toiseen sarakkeeseen, saadaan $\mathrm{ala}[\bar{a}_1\ \bar{a}_2] = \mathrm{ala}[\bar{a}_1\ \bar{a}_2 + \lambda\bar{a}_1]$, koska suunnikkailla on sama kanta $\bar{a}_1$ ja sama korkeus h.



Kuva (7)

(e) Jos rivillä k vain alkio $a_{ki} \neq 0$, niin tällöin sarakkeen i muodostama vektori ei ole muiden sarakkeiden virittämässä tasossa. Tällöin sarakevektoreiden virittämän suuntassärmiön tilavuus riippuu vain sarakevektorin $\bar{a}_i$ alkiosta a_{ki}, jolloin sarakkeen i muiden alkioiden nollaaminen ei vaikuta suuntaissärmiön tilavuuteen, eikä siis myöskään determinanttiin.

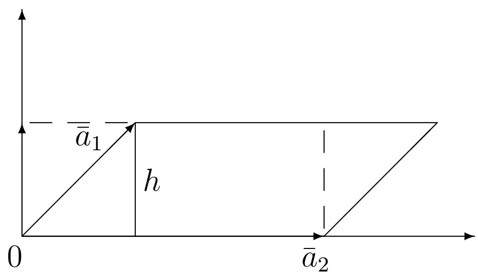
Esimerkiksi 2×2-matriisille

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{bmatrix}$$

saadaan

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{vmatrix} = \begin{vmatrix} 0 & a_{12} \\ a_{21} & 0 \end{vmatrix},$$

koska muodostuvien suunnikkaiden pinta-alat ovat yhtä suuret. Näin on, sillä suunnikkailla on sama kanta $\bar{a}_2$ ja korkeus h, joka riippuu vain alkiosta a_{21}.



Kuva (8)

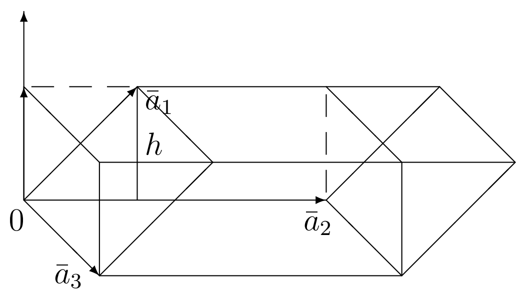
Vastaavasti 3×3-matriisille

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

saadaan

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{23} & a_{33} \end{vmatrix},$$

koska muodostuvien suuntaissärmiöiden tilavuudet ovat yhtä suuret. Vektorit $\bar{a}_2$ ja $\bar{a}_3$ virittävät suuntaissärmiöiden pohjat ja korkeus h riippuu vain vektorin $\bar{a}_1$ alkiosta a_{21}. Tällöin suuntaissärmiöillä on sama pohjan pinta-ala ja sama korkeus, jolloin tilavuudet ovat yhtä suuret.



Kuva (9)

3. Alideterminantti ja determinantin kehityssääntö

3.10. Determinanttia voidaan kehittää laskennallisesti käyttökelpoisempaan muotoon seuraavasti. Olkoon matriisi

$$A = \begin{bmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{k1} & \dots & a_{ki} & \dots & a_{kn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_k \\ \vdots \\ \vec{a}_n \end{bmatrix} = [\bar{a}_1 \ \dots \ \bar{a}_i \ \dots \ \bar{a}_n].$$

Poistamalla matriisista A rivi k ja sarake i, saadaan $(n-1)\times(n-1)$-matriisi $A_{\hat{k},\hat{\imath}}$, jota sanotaan *alkioon* a_{ki} *liittyväksi alimatriisiksi* ja sen determinanttia *alideterminantiksi.* Huomataan, että transponoidulle matriisille $A^T = [\alpha_{ik}]$, $\alpha_{ik} = a_{ki}$, pätee yhtälö

$$\left(A^T\right)_{\hat{\imath},\hat{k}} = \left(A_{\hat{k},\hat{\imath}}\right)^T. \tag{3.5}$$

Tämä nähdään helposti todeksi kirjoittamalla matriisit auki alkiomuodossa. Lauseen 3.9 (e)-kohdan ja ominaisuuden (D1) nojalla pätee seuraava päättely ensimmäisen sarakkeen suhteen:

$$\begin{aligned}\det A &= \sum_{(k_1,\dots,k_n)\in P_n} \operatorname{sign}(k_1,\dots,k_n) a_{k_1,1}\cdots a_{k_n,n} \\ &= \sum_{k_1=1}^{n} a_{k_1,1} \sum_{\substack{k_2,\dots,k_n \\ (k_1,\dots,k_n)\in P_n}} \operatorname{sign}(k_1,\dots,k_n) a_{k_2,2}\cdots a_{k_n,n} \\ &= \sum_{k_1=1}^{n} a_{k_1,1}(-1)^{k_1-1}\det A_{\hat{k}_1,\hat{1}} \\ &= \sum_{k=1}^{n} a_{k1}(-1)^{k+1}\det A_{\hat{k},\hat{1}}.\end{aligned}$$

Vastaavasti voidaan johtaa tulos sarakkeen i suhteen, jolloin saadaan

$$\det A = \sum_{k=1}^{n} a_{ki}(-1)^{k+i}\det A_{\hat{k},\hat{\imath}}. \tag{3.6}$$

Nyt yhtälön (3.5) ja kohdan 3.8 nojalla pätee

$$\begin{aligned}\det A = \det A^T &= \sum_{i=1}^{n} \alpha_{ik}(-1)^{i+k}\det\left(A^T\right)_{\hat{\imath},\hat{k}} \\ &= \sum_{i=1}^{n} a_{ki}(-1)^{k+i}\det\left(A_{\hat{k},\hat{\imath}}\right)^T \\ &= \sum_{i=1}^{n} a_{ki}(-1)^{k+i}\det A_{\hat{k},\hat{\imath}},\end{aligned} \tag{3.7}$$

jota voidaan selventää merkitsemällä

$$(-1)^{k+i}\det A_{\hat{k},\hat{\imath}} =: \operatorname{cof} a_{ki},$$

jota sanotaan *alkion* a_{ki} *kofaktoriksi.* Tämän avulla saadaan yhtälöille (3.6) ja (3.7) seuraava tulos, jonka mukaan determinantti voidaan *kehittää sarakkeen vast. rivin suhteen* (ks. determinantin kehityssäännön johtaminen [6, s. 49-50]).

LAUSE 3.11. *Matriisin* $A = A_{n\times n} = [a_{ki}]$ *determinantille pätee kehityssääntö*

$$\det A = \sum_{k=1}^{n} (-1)^{k+i} a_{ki} \det A_{\hat{k},\hat{\imath}} = \sum_{k=1}^{n} a_{ki} \operatorname{cof} a_{ki}$$

jokaisen sarakkeen $i = 1, \dots, n$ *suhteen, ja vastaavasti*

$$\det A = \sum_{i=1}^{n} (-1)^{k+i} a_{ki} \det A_{\hat{k},\hat{\imath}} = \sum_{i=1}^{n} a_{ki} \operatorname{cof} a_{ki}$$

jokaisen rivin $k = 1, \dots, n$ *suhteen.*

HUOMAUTUS 3.12. Determinantin laskeminen perustuu tavallisesti lauseeseen 3.11, jota sovelletaan toistuvasti sopivan sarakkeen tai rivin suhteen. Myös lauseen 3.9 tuloksia kannattaa hyödyntää laskujen eri vaiheissa.

ESIMERKKI 3.13. Esimerkin 3.5 matriisille

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

saadaan alimatriisit

$$A_{\hat{1},\hat{1}} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}, \quad A_{\hat{2},\hat{1}} = \begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix} \quad \text{ja} \quad A_{\hat{3},\hat{1}} = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix},$$

joiden avulla voidaan laskea matriisin A determinantti

$$\det A = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 4 \cdot \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 7 \cdot \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3 + 24 - 21 = 0.$$

Seuraava lause antaa tulon determinantille tärkeän tuloksen, joka helpottaa determinanttien määrittämistä.

LAUSE 3.14. *Matriiseille* $A = A_{n\times n}$ *ja* $B = B_{n\times n}$ *pätee*

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

TODISTUS. Ks. [6, s. 50-51]. □

4. Käänteismatriisi ja determinantti, Cramerin sääntö

Seuraavaa lausetta tarvitaan, kun johdetaan Cramerin säännön nimellä tunnettu tulos kohdassa 3.17. Cramerin säännön avulla voidaan ratkaista yhtälöryhmiä, joissa on saman verran tuntemattomia ja ratkaistavia yhtälöitä.

LAUSE 3.15. *Kääntyvän matriisin* $A = A_{n\times n}$ *käänteismatriisille* A^{-1} *pätee*

$$A^{-1} = \frac{1}{\det A} (\operatorname{cof} A)^T,$$

missä $\operatorname{cof} A := [\operatorname{cof} a_{ki}]_{k,i}$ *on matriisin* A *kofaktorimatriisi.*

TODISTUS. Ks. [6, s. 52-53]. □

ESIMERKKI 3.16. Matriisin

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

kofaktorimatriisi on

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

ja determinantti on $\det A = 4$, jolloin sen käänteismatriisiksi saadaan

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

3.17. **Cramerin säännön johtaminen.** Olkoon $A = A_{n\times n} = [a_{ki}] = [\bar{a}_1 \ \dots \ \bar{a}_n]$ kääntyvä matriisi ja olkoon $b = (b_1, \dots, b_n) \in \mathbb{R}^n$ annettu. Tällöin muuttujan $x \in \mathbb{R}^n$ yhtälöryhmällä

$$Ax = b \tag{3.8}$$

on yksikäsitteinen ratkaisu $x = A^{-1}b$, joka saa muodon

$$x = \frac{1}{\det A} (\operatorname{cof} A)^T b$$

lauseen 3.15 nojalla. Nyt on

$$(\operatorname{cof} A)^T b = \begin{bmatrix} \operatorname{cof} a_{11} & \dots & \operatorname{cof} a_{n1} \\ \vdots & & \vdots \\ \operatorname{cof} a_{1n} & \dots & \operatorname{cof} a_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Tällöin lauseen 3.11 nojalla pätee esimerkiksi

$$\begin{aligned}(\det A) x_1 &= b_1 \operatorname{cof} a_{11} + b_2 \operatorname{cof} a_{21} + \dots + b_n \operatorname{cof} a_{n1} \\ &= \det[b \ \bar{a}_2 \ \dots \ \bar{a}_n].\end{aligned}$$

Matriisista A saadaan matriisi $A_i(b)$, kun vaihdetaan sarakkeen $\bar{a}_i$ paikalle b. Käyttämällä tätä matriisia $A_i(b)$ saadaan yhtälölle (3.8) *Cramerin sääntö* seuraavassa lauseessa. (Ks. tuloksen johtaminen [6, s. 53-54].)

LAUSE 3.18. *Olkoon $A = A_{n\times n}$ kääntyvä matriisi ja olkoon $b \in \mathbb{R}^n$ annettu. Tällöin yhtälön*

$$Ax = b$$

yksikäsitteiselle ratkaisulle $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ pätee

$$x_i = \frac{\det A_i(b)}{\det A}, \quad i = 1, \dots, n.$$

HUOMAUTUS 3.19. Selkeästä kaavamaisuudestaan huolimatta lauseessa 3.18 saatu tulos ei ole laskennallisesti kovinkaan käyttökelpoinen.

Esimerkki 3.20. Matriisin

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

yhtälön $Ax = b$ ratkaisulle $x = (x_1, x_2, x_3)$ pätee

$$x_1 = \frac{1}{4} \begin{vmatrix} 1 & 0 & 1 \\ 2 & 2 & 0 \\ 3 & 1 & 2 \end{vmatrix} = 0, \quad x_2 = \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 0 \\ 1 & 3 & 2 \end{vmatrix} = 1 \quad \text{ja} \quad x_3 = \frac{1}{4} \begin{vmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 1 & 1 & 3 \end{vmatrix} = 1,$$

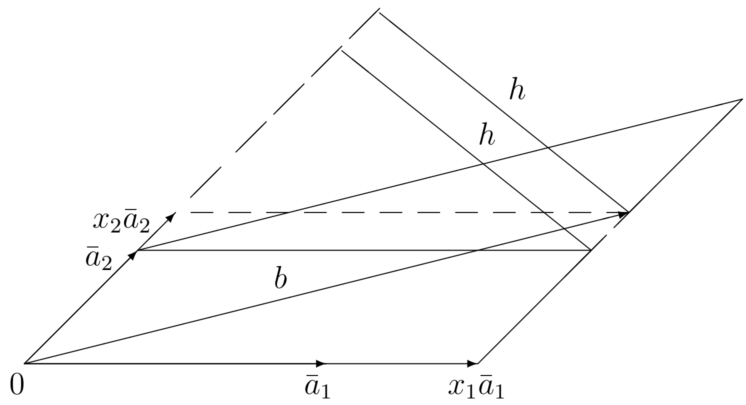
kun matriisin A determinantti on $\det A = 4$.

3.21. Cramerin sääntöä voidaan selittää myös geometrisesti determinantin tilavuustulkinnan avulla. Tarkastellaan nyt kääntyvää matriisia $A = A_{2\times 2} = [\bar{a}_1 \ \bar{a}_2]$, kun $b \in \mathbb{R}^2$ on annettu. Tällöin yhtälöryhmä

$$Ax = b$$

voidaan ratkaista geometrisesti seuraavasti:

Kirjoitetaan aluksi $b = Ax = x_1\bar{a}_1 + x_2\bar{a}_2$.



Kuva (10)

Kuvasta (10) nähdään, että vektoreiden $x_1\bar{a}_1$ ja $\bar{a}_2$ sekä vektoreiden b ja $\bar{a}_2$ virittämien suunnikkaiden pinta-alat ovat yhtä suuret, koska niillä on sama kanta $\bar{a}_2$ ja sama korkeus h. Näin huomattu yhtälö

$$\text{ala}(x_1\bar{a}_1, \bar{a}_2) = \text{ala}(b, \bar{a}_2)$$

voidaan kirjoittaa determinanttien avulla muotoon

$$\det[x_1\bar{a}_1 \ \bar{a}_2] = \det[b \ \bar{a}_2],$$

koska kiertokulmat $\sphericalangle(x_1\bar{a}_1, \bar{a}_2)$ ja $\sphericalangle(b, \bar{a}_2)$ ovat samanmerkkiset. Toisin sanoen pätee

$$x_1 \det[\bar{a}_1 \ \bar{a}_2] = \det[b \ \bar{a}_2].$$

Vastaavasti voidaan todeta, että on voimassa yhtälö

$$\text{ala}(\bar{a}_1, x_2\bar{a}_2) = \text{ala}(\bar{a}_1, b),$$

joka voidaan kirjoittaa muotoon

$$x_2 \det[\bar{a}_1 \ \bar{a}_2] = \det[\bar{a}_1 \ b],$$

koska kiertokulmat $\sphericalangle(\bar{a}_1, x_2\bar{a}_2)$ ja $\sphericalangle(\bar{a}_1, b)$ ovat samanmerkkiset. Tällöin saadaan

$$x_1 = \frac{\det[b\ \bar{a}_2]}{\det[\bar{a}_1\ \bar{a}_2]} = \frac{\det A_1(b)}{\det A} \quad \text{ja} \quad x_2 = \frac{\det[\bar{a}_1\ b]}{\det[\bar{a}_1\ \bar{a}_2]} = \frac{\det A_2(b)}{\det A}.$$

5. Integraalilaskennan muuttujanvaihtolause

Matriiseja ja determinantteja voidaan soveltaa mm. integraalilaskentaan. Esimerkiksi integraalin muuttujanvaihtolauseessa derivaattaa vastaava matriisi on Jacobin matriisi ja sen determinantti on Jacobin determinantti J_g. Tämä on esimerkki tilanteesta, jossa matriisi ja determinantti muuttuvat tarkastelupisteen mukana.

LAUSE 3.22. (*Muuttujanvaihtolause.*) *Oletetaan, että*

1) *$g : W \to \mathbb{R}^n$ on jatkuvasti differentioituva kuvaus avoimessa joukossa $W \subset \mathbb{R}^n$,*
2) *$U \subset W$ on rajoitettu Jordan-joukko (ts. rajoitettu joukko, jonka reuna on nollamittainen), jolle sulkeuma $\overline{U} \subset W$, ja*
3) *jollekin suljetulle nollamittaiselle joukolle $N \subset \overline{U}$ avoimeen joukkoon $U' := (\operatorname{int} U) \setminus N$ rajoitettuna kuvaus $g : U' \to g(U')$ on diffeomorfismi (ts. kuvaus on molempiin suuntiin jatkuvasti differentioituva bijektio).*

Tällöin $g(U)$ on Jordan-joukko, ja rajoitetulle integroituvalle funktiolle $f : g(U) \to \mathbb{R}$ myös funktio

$$(f \circ g)|J_g| : U \to \mathbb{R}$$

on integroituva, ja on voimassa muuttujanvaihtoyhtälö

$$\int_{g(U)} f = \int_U (f \circ g)|J_g|.$$

TODISTUS. Ks. [9, s. 424-428.] □

ESIMERKKI 3.23. Olkoon joukko P suuntaissärmiö, joka on esitetty kuvassa (11), ja olkoon funktio

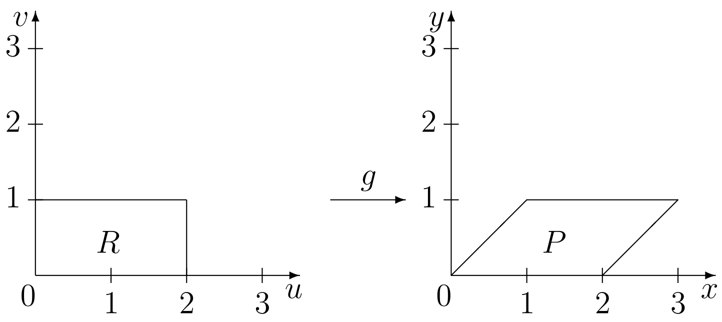
$$f : P \to \mathbb{R} : (x, y) \mapsto x + y.$$

Nyt joukolle R, joka myös on esitetty kuvassa (11), pätee $P = g(R)$ kuvauksessa $g : (u, v) \mapsto (u + v, v)$, jonka Jacobin determinantti on

$$J_g = \begin{vmatrix} \partial_1 g_1 & \partial_2 g_1 \\ \partial_1 g_2 & \partial_2 g_2 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1.$$

Lisäksi kuvaus g on lineaarisena bijektiona diffeomorfismi, jolloin muuttujanvaihtolauseen 3.22 nojalla saadaan

$$\int_P f = \int_R (f \circ g)|J_g| = \int_0^2 du \int_0^1 (u + 2v) dv = 4.$$



Kuva (11)

Esimerkki 3.24. Olkoon joukko $A = B^2(0,1)$ ja olkoon funktio

$$f : A \to \mathbb{R} : (x,y) \mapsto x^2 + y^2.$$

Nyt joukolle $U = [0,1[\times[0,2\pi[$ pätee $A = g(U)$ napakoordinaattikuvauksessa

$$g : (r,\theta) \mapsto (r\cos\theta, r\sin\theta),$$

jonka Jacobin determinatti on $J_g = r$. Lisäksi kuvaus $g : \operatorname{int} U \to g(\operatorname{int} U)$ on diffeomorfismi, jolloin muuttujanvaihtolauseen 3.22 nojalla saadaan

$$\int_A f = \int_U (f \circ g)|J_g| = \int_0^{2\pi} \int_0^1 r^2 r \; dr d\theta = \frac{\pi}{2}.$$

LUKU 4

Matriisin aste

1. Matriisin astelause

Seuraavat kaksi määritelmää sisältävät matriisin sarake- ja riviavaruuden sekä matriisin asteen määritelmät. Matriisin aste voidaan laskea esimerkiksi alideterminanttien avulla.

MÄÄRITELMÄ 4.1. Matriisin

$$A = A_{m\times n} = [\bar{a}_1 \ \dots \ \bar{a}_n] = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix}$$

sarakevektoreiden $\bar{a}_1, \dots, \bar{a}_n$ virittämää avaruutta

$$\langle \bar{a}_1, \dots, \bar{a}_n \rangle \subset \mathbb{R}^m$$

sanotaan sen *sarakeavaruudeksi*, ja vastaavasti

$$\langle \vec{a}_1, \dots, \vec{a}_m \rangle \subset \mathbb{R}^n$$

on sen *riviavaruus.*

MÄÄRITELMÄ 4.2. Matriisin

$$A = A_{m\times n} = [\bar{a}_1 \ \dots \ \bar{a}_n]$$

aste $\operatorname{rank} A$ määritellään asettamalla

$$\operatorname{rank} A := \dim \langle \bar{a}_1, \dots, \bar{a}_n \rangle .$$

Matriisin aste on siis matriisin sarakeavaruuden dimensio, jolla tarkoitetaan sarakeavaruuden kannan alkioiden lukumäärää.

HUOMAUTUS 4.3. Olkoon matriisia $A = A_{m\times n}$ vastaava lineaarikuvaus

$$L = L_A : \mathbb{R}^n \to \mathbb{R}^m.$$

Tällöin on voimassa

$$R(L) = L(\mathbb{R}^n) = \langle \bar{a}_1, \dots, \bar{a}_n \rangle ,$$

koska sarakevektorit $\bar{a}_1, \dots, \bar{a}_n$ virittävät arvoavaruuden, ja siten pätee

$$\operatorname{rank} A = \dim R(L).$$

ESIMERKKI 4.4. Matriisin

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

$$\text{sarakeavaruus on } \langle (1,2), (2,-1), (-1,3) \rangle = \mathbb{R}^2$$

ja riviavaruus on $\langle(1,2,-1),(2,-1,3)\rangle \subset \mathbb{R}^3$. Kuvaukselle $L = L_A : \mathbb{R}^3 \to \mathbb{R}^2$,

$$Lx = (x_1 + 2x_2 - x_3, 2x_1 - x_2 + 3x_3),$$

on $\dim R(L) = 2 = \operatorname{rank} A$.

Matriisin aste voidaan määrittää *alideterminanttien* avulla seuraavan lauseen mukaisesti, kun alideterminantilla tarkoitetaan $r \times r$-alimatriisin determinanttia, ja *alimatriisilla* matriisia, joka saadaan alkuperäisestä matriisista poistamalla siitä yksi tai useampi rivi tai sarake.

LAUSE 4.5. *Olkoon matriisi $A = A_{m\times n}$. Tällöin matriisin aste on* $\operatorname{rank} A = r > 0$ *jos ja vain jos matriisilla A on ainakin yksi $r \times r$-alimatriisi, jonka determinantti ei ole nolla, ja jokaisen $(r+1)\times(r+1)$-alimatriisin determinantti on nolla.*

TODISTUS. Ks. [1, s. 268-269]. □

HUOMAUTUS 4.6. Jos matriisin A aste on $\operatorname{rank} A = r > 0$, niin ainakin yhden sen $r \times r$-alimatriisin aste on täsmälleen r.

ESIMERKKI 4.7. (a) Esimerkin 4.4 matriisin

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

2×2-alideterminantit ovat

$$\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} = -5, \quad \begin{vmatrix} 1 & -1 \\ 2 & 3 \end{vmatrix} = 5 \quad \text{ja} \quad \begin{vmatrix} 2 & -1 \\ -1 & 3 \end{vmatrix} = 5.$$

Näistä mikään ei ole nolla, joten matriisin A aste on $\operatorname{rank} A = 2$.

(b) Matriisilla

$$A = \begin{bmatrix} 1 & 2 & -1 & 4 \\ 3 & 1 & 5 & 2 \\ 1 & -3 & 7 & -6 \end{bmatrix}$$

on 2×2-alimatriisi, jonka determinantti

$$\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5$$

ei ole nolla. Lisäksi matriisin A kaikki neljä 3×3-alideterminanttia ovat nollia, joten matriisin aste on $\operatorname{rank} A = 2$.

4.8. Kohdan 2.5 yhtälöryhmä (2.3) voidaan kirjoittaa myös *laajennettuna kerroinmatriisina*

$$[A|b] = \left[\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right],$$

jossa on kaikki tarvittava tieto selkeästi näkyvillä. Tämän ja lauseen 4.5 avulla voidaan seuraavan lauseen mukaisesti selvittää, onko yhtälöryhmällä olemassa ratkaisua.

LAUSE 4.9. *Yhtälöryhmällä $Ax = b$ on ainakin yksi ratkaisu jos ja vain jos b kuuluu matriisin A sarakeavaruuteen. Tämä tapahtuu jos ja vain jos matriisilla A ja laajennetulla kerroinmatriisilla $[A|b]$ on sama aste.*

TODISTUS. Ks. [1, s. 261]. □

ESIMERKKI 4.10. Ratkaistaan, onko yhtälöryhmällä

$$\begin{cases} 2x_1 + 4x_2 + 6x_3 = 18 \\ 4x_1 + 5x_2 + 6x_3 = 24 \\ 2x_1 + 7x_2 + 12x_3 = 40 \end{cases} \tag{4.1}$$

ratkaisua. Matriisin

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 2 & 7 & 12 \end{bmatrix}$$

determinantti on $\det A = 0$ ja sen alimatriisin

$$\begin{bmatrix} 2 & 4 \\ 4 & 5 \end{bmatrix}$$

determinantti on $-6 \neq 0$, jolloin lauseen 4.5 nojalla matriisin A aste on $\operatorname{rank} A = 2$. Vastaavasti laajennetun kerroinmatriisin

$$\left[\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 40 \end{array}\right]$$

alimatriisin

$$\begin{bmatrix} 4 & 6 & 18 \\ 5 & 6 & 24 \\ 7 & 12 & 40 \end{bmatrix}$$

determinantti on $-60 \neq 0$, joten lauseen 4.5 nojalla laajennetun kerroinmatriisin aste on 3. Nyt lauseen 4.9 nojalla yhtälöryhmällä (4.1) ei ole ratkaisua.

Seuraava tulos tunnetaan astelauseena. Sen mukaan matriisin sarakeavaruuden ja riviavaruuden dimensiot ovat samat, vaikka ne olisivat eri avaruuksien aliavaruuksia.

LAUSE 4.11. (*Astelause.*) *Matriisin*

$$A = A_{m\times n} = [\bar{a}_1 \ \dots \ \bar{a}_n] = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix}$$

sarakeavaruudella ja riviavaruudella on sama dimensio $\operatorname{rank} A$*,*

$$\operatorname{rank} A = \dim \langle \bar{a}_1, \dots, \bar{a}_n \rangle = \dim \langle \vec{a}_1, \dots, \vec{a}_m \rangle .$$

TODISTUS. Ks. [6, s. 57]. □

Matriisin sarakeavaruus on sama kuin sen transponoidun matriisin riviavaruus, jolloin astelauseesta 4.11 seuraa välittömästi seuraava tulos.

SEURAUS 4.12. *Matriisille* $A = A_{m\times n}$ *on* $\operatorname{rank} A^T = \operatorname{rank} A$*.*

2. Matriisin dimensiolause

Matriiseille saadaan keskeinen tulos dimensiolauseessa, joka antaa yhteyden matriisin asteen ja ytimen dimension välille. Tämän lauseen todistamiseen tarvitaan astelausetta ja seuraavaa lausetta.

LAUSE 4.13. *Olkoon $H \subset \mathbb{R}^n$ aliavaruus ja olkoon*

$$H^{\perp} := \{z \in \mathbb{R}^n : (z|y) = 0 \text{ kaikille } y \in H\}$$

sen ortogonaalinen komplementti. Tällöin pätee

$$\dim H + \dim H^{\perp} = n = \dim \mathbb{R}^n.$$

TODISTUS. Ks. [6, s. 39-40]. □

LAUSE 4.14. (*Dimensiolause.*) *Matriisin $A = A_{m\times n}$ ytimen dimensiolle* $\dim N(A)$ *ja asteelle* $\operatorname{rank} A$ *on voimassa dimensioyhtälö*

$$\dim N(A) + \operatorname{rank} A = n.$$

TODISTUS. (Ks. [6, s. 58].) Olkoon matriisi

$$A = [\bar{a}_1 \ \dots \ \bar{a}_n] = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix}.$$

Koska pätee

$$Ax = \begin{bmatrix} (\vec{a}_1|x) \\ \vdots \\ (\vec{a}_m|x) \end{bmatrix}, \quad \text{kun } x \in \mathbb{R}^n,$$

niin on $x \in N(A)$ jos ja vain jos on voimassa $x \perp \langle \vec{a}_1, \dots, \vec{a}_m \rangle$. Tällöin matriisin A ydin on

$$N(A) = \langle \vec{a}_1, \dots, \vec{a}_m \rangle^{\perp}.$$

Astelauseen 4.11 mukaan on voimassa

$$\dim \langle \vec{a}_1, \dots, \vec{a}_m \rangle = \dim \langle \bar{a}_1, \dots, \bar{a}_n \rangle,$$

jolloin tästä ja lauseesta 4.13 seuraa, että pätee

$$\begin{aligned} n &= \dim \langle \vec{a}_1, \dots, \vec{a}_m \rangle + \dim \langle \vec{a}_1, \dots, \vec{a}_m \rangle^{\perp} \\ &= \dim \langle \vec{a}_1, \dots, \vec{a}_m \rangle + \dim N(A) \\ &= \dim \langle \bar{a}_1, \dots, \bar{a}_n \rangle + \dim N(A) \\ &= \operatorname{rank} A + \dim N(A). \end{aligned}$$

□

ESIMERKKI 4.15. Esimerkin 4.4 matriisille

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

on $\dim N(A) = \dim \mathbb{R}^3 - \operatorname{rank} A = 3 - 2 = 1$.

Huomautus 4.16. (a) Matriisin $A_{m\times n}$ ytimen ortogonaalinen komplementti on $N(A)^\perp = \langle \vec{a}_1, \dots, \vec{a}_m \rangle$ ja matriisia A vastaavan lineaarikuvauksen $L : \mathbb{R}^n \to \mathbb{R}^m$ arvojoukko on $R(L) = \langle \bar{a}_1, \dots, \bar{a}_n \rangle$. Matriisin A rivivektorit virittävät siis ytimen ortogonaalisen komplementin $N(A)^\perp$.

(b) Ytimen ortogonaalisen komplementin $N(A)^\perp$ dimensio on $\dim N(A)^\perp = \operatorname{rank} A$, koska pätee

$$\dim N(A)^\perp = \dim \langle \vec{a}_1, \dots, \vec{a}_m \rangle = \dim \langle \bar{a}_1, \dots, \bar{a}_n \rangle = \operatorname{rank} A.$$

(c) Seurauksen 4.12 nojalla matriisin $A_{m\times n}$ transponoidun matriisin $A^T_{n\times m}$ ytimen dimensio on $\dim N(A^T) = m - \operatorname{rank} A^T = m - \operatorname{rank} A$. Jos $A = A_{n\times n}$ on neliömatriisi, niin pätee $\dim N(A^T) = \dim N(A)$.

Lause 4.17. *Matriisille $A = A_{n\times n}$ on voimassa seuraavat neljä väitettä, jos niistä yksikin on tosi.*

(i) *A on kääntyvä.*
(ii) $\det A \neq 0$.
(iii) $\dim N(A) = 0$.
(iv) $\operatorname{rank} A = n$.

Todistus. Ks. [6, s. 52], [1, s. 261], lause 2.4 ja dimensiolause 4.14. Näitä ja matriisin määräämää lineaarikuvausta käyttämällä nähdään, että väitteet ovat ekvivalentteja keskenään. □

4.18. Matriisia $A = A_{m\times n}$ vastaava lineaarikuvaus $L : \mathbb{R}^n \to \mathbb{R}^m$ on injektio lauseen 2.4 nojalla täsmälleen silloin, kun on ydin $N(A) = \{0\}$. Nyt dimensiolauseen 4.14 nojalla pätee

$$\operatorname{rank} A = n.$$

Tämä johtaa tilanteessa $m = n$ seuraavaan lauseeseen, joka sitoo lineaarikuvauksen, matriisin ja determinantin toisiinsa.

Lause 4.19. *Matriisia $A = A_{n\times n}$ vastaava lineaarikuvaus $L = L_A : \mathbb{R}^n \to \mathbb{R}^n$ on injektio ja surjektio eli bijektio jos ja vain jos pätee* $\det A \neq 0$.

Todistus. Bijektiivinen lineaarikuvaus L on kääntyvä ja sitä vastaava matriisi A on myös kääntyvä. Nyt lauseen 4.17 nojalla väite on tosi molempiin suuntiin. □

Esimerkki 4.20. (a) Kuvaus

$$L : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (x_1 - x_2, 2x_1 - 2x_2)$$

ei ole bijektio ja matriisin

$$\operatorname{mat} L = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}$$

determinantti on $\det(\operatorname{mat} L) = 0$.

(b) Kuvaus

$$L : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto (x_1 - 2x_2, 2x_1 - 2x_2)$$

on bijektio ja matriisin

$$\operatorname{mat} L = \begin{bmatrix} 1 & -2 \\ 2 & -2 \end{bmatrix}$$

determinantti on $\det(\operatorname{mat} L) = 2 \neq 0$.

LUKU 5

Ominaisarvoteoriaa

1. Ominaisarvot ja ominaisvektorit

5.1. Olkoon $L : V \to V$ lineaarikuvaus. Jos on olemassa luku λ ja vektori $v \in V$, $v \neq 0$, joille pätee yhtälö

$$Lv = \lambda v,$$

niin sanotaan, että λ on kuvauksen L *ominaisarvo* ja v siihen liittyvä *ominaisvektori*. Kuvaus L venyttää tai kutistaa ominaisvektoria v suoralla $\langle v \rangle$ jonkin vakiokertoimen verran. Tämä vakio on kuvauksen L ominaisarvo. Lineaarikuvaus siis kertoo ominaisvektorin ominaisarvollaan. Lisäksi, jos lineaarikuvauksen L ominaisvektorit muodostavat avaruuden V kannan, niin tässä kannassa $\operatorname{mat} L$ on diagonaalimatriisi, jonka alkioina ovat kuvauksen L ominaisarvot.

Kuvauksella L ei tarvitse olla yhtään ominaisarvoa, jos on $\dim V = \infty$. Jos taas on $\dim V < \infty$, niin ominaisarvoja löytyy ainakin, jos kerroinlukualue $\mathbb{R}$ laajennetaan kompleksilukualueeksi $\mathbb{C}$. Lisäksi, jos on $\dim V = n < \infty$, niin lineaarikuvausta $L : V \to V$ vastaa kiinnitettyjen kantojen suhteen yksikäsitteinen neliömatriisi $A = A_{n\times n}$, jolle voidaan määritellä ominaisarvo ja ominaisvektori kuten sitä vastaavalle lineaarikuvaukselle. Tästä syystä voidaan rajoittua tarkastelemaan matriiseja $A_{n\times n}$ ja tarvittaessa käyttää matriisien vastaavuutta lineaarikuvauksiin.

MÄÄRITELMÄ 5.2. Olkoon matriisi $A = A_{n\times n}$. Luku $\lambda \in \mathbb{C}$ on matriisin A *ominaisarvo* ja $v \in \mathbb{C}^n$, $v \neq 0$, siihen liittyvä *ominaisvektori*, jos pätee yhtälö

$$Av = \lambda v.$$

ESIMERKKI 5.3. Olkoon $A = I$ yksikkömatriisi. Tällöin jokaiselle $v \in \mathbb{C}^n$ pätee $Av = Iv = v$. Matriisin A ominaisarvo on $\lambda = 1$ ja jokainen vektori $v \neq 0$ on yksikkömatriisin I ominaisvektori.

Seuraava lause antaa säännön, jolla matriisin ominaisarvot ja ominaisvektorit saadaan ratkaistua.

LAUSE 5.4. *Luku λ on matriisin $A = A_{n\times n}$ ominaisarvo jos ja vain jos se toteuttaa matriisin A karakteristisen yhtälön*

$$\det(A - \lambda I) = 0, \quad I = I_n.$$

TODISTUS. (Ks. [1, s. 387].) Olkoon λ matriisin $A = A_{n\times n}$ ominaisarvo. Tällöin löytyy ominaisarvoon liittyvä ominaisvektori $v \neq 0$, ja pätee yhtälö

$$Av = \lambda v = \lambda I v,$$

joka voidaan kirjoittaa muotoon

$$(A - \lambda I)v = 0. \tag{5.1}$$

Jos matriisi A on $n \times n$-matriisi, niin yhtälö (5.1) on homogeeninen yhtälöryhmä, jossa on n yhtälöä ja n tuntematonta. Tällöin yhtälöryhmällä on ei-triviaali ratkaisu, joten pätee $\det(A - \lambda I) = 0$.

Jos taas on $\det(A - \lambda I) = 0$, niin yhtälöllä (5.1) on ei-triviaali ratkaisu ja matriisin A ominaisarvo on λ. □

ESIMERKKI 5.5. Matriisin

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

karakteristisesta yhtälöstä

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & 2 & 0 \\ 2 & 2-\lambda & 2 \\ 0 & 2 & 1-\lambda \end{vmatrix} \\ &= (3-\lambda)[(2-\lambda)(1-\lambda) - 4] - 4(1-\lambda) \\ &= -(\lambda - 5)(\lambda - 2)(\lambda + 1) = 0, \end{aligned}$$

nähdään, että sen ominaisarvot ovat $\lambda_1 = 5$, $\lambda_2 = 2$ ja $\lambda_3 = -1$.

Kun ominaisarvo on $\lambda_1 = 5$, niin pätee $(A - 5I)v_1 = 0$ eli

$$\begin{bmatrix} -2 & 2 & 0 \\ 2 & -3 & 2 \\ 0 & 2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

josta saadaan ratkaistua ominaisvektori $v_1 = (2, 2, 1)$.

Vastaavasti, kun ominaisarvo on $\lambda_2 = 2$, niin on voimassa $(A - 2I)v_2 = 0$ ominaisvektorilla $v_2 = (2, -1, -2)$.

Ominaisarvolle $\lambda_3 = -1$ saadaan vastaavasti ominaisvektori $v_3 = (1, -2, 2)$.

Seuraavaa lausetta tarvitaan myöhemmin.

LAUSE 5.6. *Matriisin $A = A_{n\times n}$ erisuuria ominaisarvoja vastaavat ominaisvektorit ovat lineaarisesti riippumattomia.*

TODISTUS. Ks. [1, s. 389]. □

2. Yhtäläiset matriisit ja diagonalisoituvuus

Seuraava yhtäläisten matriisien määritelmä antaa hyödyllisen yhteyden kahden matriisin välille. Tämän jälkeen tuleva lause sitten yhdistää yhtäläisen matriisin ja lineaarikuvauksen.

MÄÄRITELMÄ 5.7. Matriisit $A = A_{n\times n}$ ja $B = B_{n\times n}$ ovat *yhtäläisiä*, jos pätee yhtälö

$$B = P^{-1}AP$$

jollakin kääntyvällä matriisilla $P = P_{n\times n}$.

LAUSE 5.8. *Olkoon $L : V \to V$ lineaarikuvaus ja olkoot äärellisulotteisen avaruuden V kannat $B = \{v_1, \dots, v_n\}$ ja $C = \{w_1, \dots, w_n\}$. Jos* $\operatorname{mat} L(B, B)$ *on lineaarikuvausta L vastaava matriisi kannassa B ja jos* $\operatorname{mat} L(C, C)$ *on lineaarikuvausta L*

vastaava matriisi kannassa C, niin matriisit $\operatorname{mat} L(B,B)$ *ja* $\operatorname{mat} L(C,C)$ *ovat yhtäläisiä.*

TODISTUS. (Ks. [1, s. 411-412] ja [7, s. 19-20].) Matriisi $\operatorname{mat} L(C,C)$ voidaan kirjoittaa muotoon

$$\operatorname{mat} L(C,C) = K(B,C) \operatorname{mat} L(B,B) K(C,B),$$

kun $K(C,B)$ on kannanvaihtomatriisi kannasta C kantaan B. Koska matriisi $P = K(C,B)$ on kääntyvä ja pätee $P^{-1} = K(B,C)$, niin on voimassa

$$\operatorname{mat} L(C,C) = P^{-1} \operatorname{mat} L(B,B) P.$$

Tällöin matriisit $\operatorname{mat} L(B,B)$ ja $\operatorname{mat} L(C,C)$ ovat yhtäläisiä. □

HUOMAUTUS 5.9. Lineaarikuvausta eri kannoissa vastaavat matriisit ovat aina yhtäläisiä.

Seuraavan lauseen mukaan yhtäläisillä matriiseilla on samat ominaisarvot.

LAUSE 5.10. *Jos matriisit $A = A_{n\times n}$ ja $B = B_{n\times n}$ ovat yhtäläisiä, niin niillä on sama karakteristinen yhtälö ja samat ominaisarvot.*

TODISTUS. (Ks. [1, s. 406].) Yhtäläisille matriiseille $A = A_{n\times n}$ ja $B = B_{n\times n}$ pätee $B = P^{-1}AP$, kun matriisi $P = P_{n\times n}$ on kääntyvä. Tällöin on voimassa

$$\begin{aligned}\det(B - \lambda I) &= \det(P^{-1}AP - \lambda P^{-1}P) = \det(P^{-1}(A - \lambda I)P) \\ &= (\det P^{-1}) \det(A - \lambda I)(\det P) \\ &= \det(A - \lambda I),\end{aligned}$$

jolloin matriiseilla A ja B on sama karakteristinen yhtälö ja samat ominaisarvot. □

MÄÄRITELMÄ 5.11. Matriisi $A = A_{n\times n}$ on diagonalisoituva, jos matriisi A on yhtäläinen diagonaalimatriisin D kanssa.

HUOMAUTUS 5.12. Jos D on diagonaalimatriisi, niin sen ominaisarvot ovat sen diagonaalilla. Lisäksi, jos matriisi A on yhtäläinen matriisin D kanssa, niin niillä on samat ominaisarvot lauseen 5.10 nojalla. Nämä yhdistämällä huomataan, että jos matriisi A on diagonalisoituva, niin se on yhtäläinen diagonaalimatriisin kanssa, jonka diagonaalilla on matriisin A ominaisarvot.

LAUSE 5.13. *Matriisi $A = A_{n\times n}$ on diagonalisoituva jos ja vain jos sillä on n lineaarisesti riippumatonta ominaisvektoria. Jos matriisille A on voimassa $A = P^{-1}DP$, missä matriisi D diagonaalinen, niin matriisin D diagonaalilla on matriisin A ominaisarvot ja kääntyvän matriisin P sarakkeet muodostuvat matriisin A ominaisvektoreista.*

TODISTUS. Ks. [3, s. 290]. □

HUOMAUTUS 5.14. Edellinen lause sanoo, että matriisi $A = A_{n\times n}$ on diagonalisoituva jos ja vain jos sillä on riittävästi ominaisvektoreita muodostamaan avaruuden $\mathbb{R}^n$ kannan. Tätä kantaa sanotaan *ominaisvektorikannaksi.*

Esimerkki 5.15. Kun esimerkin 5.5 matriisin

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

ominaisvektorit $v_1 = (2, 2, 1)$, $v_2 = (2, -1, -2)$ ja $v_3 = (1, -2, 2)$ laitetaan matriisin P sarakevektoreiksi, saadaan kääntyvä matriisi

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

Nyt matriisille P ja sen käänteismatriisille P^{-1} pätee

$$P^{-1}AP = \frac{1}{9} \begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} =: D,$$

missä matriisi D on diagonaalinen ja sen diagonaalilla on matriisin A ominaisarvot. Matriisi A on diagonalisoituva ja sen ominaisvektorit v_1, v_2 ja v_3 antavat avaruudelle $\mathbb{R}^3$ ominaisvektorikannan $\{v_1, v_2, v_3\} = \{(2, 2, 1), (2, -1, -2), (1, -2, 2)\}$.

3. Symmetriset matriisit ja ortogonaalinen diagonalisoituvuus

Lauseen 5.6 mukaan matriisin erisuuria ominaisarvoja vastaavat ominaisvektorit ovat lineaarisesti riippumattomia. Seuraava lause sanoo, että symmetriselle matriisille tämä tulos on vahvempi: symmetrisen matriisin ominaisarvoja vastaavat ominaisvektorit ovat ortogonaalisia.

Lause 5.16. *Olkoon matriisi $A = A_{n\times n}$ symmetrinen, eli pätee $A^T = A$. Tällöin matriisin A ominaisarvot ovat reaalisia, eri ominaisarvoja vastaavat ominaisvektorit ovat ortogonaalisia ja on olemassa ortonormaali ominaisvektorikanta.*

Todistus. Kahden ensimmäisen väittämän todistukset löytyvät esimerkiksi lähteen [1] sivuilta 413-414. Ortogonaalisista ominaisvektoreista saadaan normeeraamalla ortonormaaleja. Niiden kantaominaisuuden osoittaminen on vaativaa, joten viimeisen väittämän todistaminen ohitetaan. □

Esimerkki 5.17. Esimerkin 5.15 matriisi

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

on symmetrinen, koska pätee $A = A^T$. Sen ominaisarvot $\lambda_1 = 5$, $\lambda_2 = 2$ ja $\lambda_3 = -1$ ovat reaalisia ja ominaisvektorit $v_1 = (2, 2, 1)$, $v_2 = (2, -1, -2)$ ja $v_3 = (1, -2, 2)$ ovat ortogonaalisia. Lisäksi ortonormaalit vektorit $u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{3}(2, 2, 1)$, $u_2 = \frac{v_2}{\|v_2\|} = \frac{1}{3}(2, -1, -2)$ ja $u_3 = \frac{v_3}{\|v_3\|} = \frac{1}{3}(1, -2, 2)$ muodostavat ortonormaalin ominaisvektorikannan $\{u_1, u_2, u_3\}$.

Määritelmä 5.18. Matriisi $A = A_{n\times n}$ on ortogonaalisesti diagonalisoituva, jos on olemassa ortogonaalinen matriisi Q, jolle pätee

$$D = Q^T AQ,$$

missä D on diagonaalimatriisi, jonka diagonaalilla ovat matriisin A ominaisarvot.

LAUSE 5.19. *Matriisi $A = A_{n\times n}$ on symmetrinen jos ja vain jos se on ortogonaalisesti diagonalisoituva.*

TODISTUS. (Ks. [7, s. 29].) Jos matriisi $A = A_{n\times n}$ on symmetrinen, niin sen ominaisarvoihin $\lambda_1, \dots, \lambda_n$ liittyy lauseen 5.16 nojalla ortonormaali ominaisvektorikanta $\{u_1, \dots, u_n\}$ ja pätee

$$Au_i = \lambda_i u_i, \quad i = 1, \dots, n. \tag{5.2}$$

Tällöin matriisi $Q = [u_1 \ \dots \ u_n]$ on ortogonaalinen, koska sen sarakevektorit $u_1, \dots, u_n$ ovat ortonormaaleja. Yhtälöt (5.2) yhdistämällä saadaan

$$AQ = QD \tag{5.3}$$

diagonaalimatriisille D, jonka diagonaalilla on matriisin A ominaisarvot. Yhtälö (5.3) on yhtäpitävä yhtälön

$$Q^T AQ = D$$

kanssa, koska ortogonaaliselle matriisille pätee $Q^{-1} = Q^T$.

Jos taas matriisi A on ortogonaalisesti diagonalisoituva, niin löytyy ortogonaalinen matriisi Q, jolle pätee

$$D = Q^T AQ, \tag{5.4}$$

missä matriisi D on diagonaalinen. Yhtälö (5.4) voidaan kirjoittaa muotoon $A = QDQ^T$, jolloin lauseen 2.28 nojalla pätee

$$A^T = (QDQ^T)^T = QD^TQ^T = QDQ^T = A,$$

koska diagonaalimatriisille D on $D^T = D$. $\square$

HUOMAUTUS 5.20. Matriisin Q sarakkeet muodostuvat matriisin A ortonormaalista ominaisvektorikannasta.

ESIMERKKI 5.21. Esimerkin 5.15 symmetrisen matriisin

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

ortonormaalista ominaisvektorikannasta

$$\{u_1, u_2, u_3\} = \left\{\frac{1}{3}(2,2,1), \frac{1}{3}(2,-1,-2), \frac{1}{3}(1,-2,2)\right\}$$

muodostuu matriisi

$$Q = \frac{1}{3}\begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{bmatrix},$$

jolle pätee

$$Q^T AQ = \frac{1}{3}\begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{bmatrix}\begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} =: D.$$

4. Neliömuodot

Symmetrisellä matriisilla $A = A_{n\times n}$ ja sen diagonalisoinnilla $D = Q^T AQ$ on samat ominaisarvot. Tätä tietoa voidaan käyttää hyväksi seuraavaksi määriteltävien *neliömuotojen* kautta differentiaalilaskennassa, kun tutkitaan esimerkiksi reaaliarvoisten funktioiden ääriarvoja.

5.22. Avaruuden $\mathbb{R}^n$ *neliömuoto* on funktio

$$q : \mathbb{R}^n \to \mathbb{R} : x = (x_1, \dots, x_n) \mapsto \sum_{k,i=1}^{n} q_{ki} x_k x_i$$

kertoimilla $q_{ki} \in \mathbb{R}$. Sen karakteristinen ominaisuus on *homogeenisuus astetta kaksi* eli on voimassa

$$q(tx) = t^2 q(x) \quad \text{kaikille } t \in \mathbb{R}. \tag{5.5}$$

Tällöin neliömuoto q ei vaihda etumerkkiä suoralla $\langle x\rangle$, kun $x \neq 0$. Neliömuoto q voidaan aina esittää muodossa

$$q(x) = \sum_{k,i=1}^{n} a_{ki} x_k x_i = x^T Ax = (x|Ax) \tag{5.6}$$

symmetriselle matriisille $A = A_{n\times n} = [a_{ki}]$, kun on $a_{ki} = \frac{1}{2}(q_{ki} + q_{ik})$.

Huomautus 5.23. Esitys (5.6) on voimassa vain yhdelle symmetriselle matriisille A. Tätä matriisia A sanotaan neliömuotoa q vastaavaksi tai siihen liittyväksi matriisiksi.

Esimerkki 5.24. Toisen asteen käyrä

$$ax_1^2 + bx_1x_2 + cx_2^2 = d \tag{5.7}$$

voidaan tunnistaa kirjoittamalla se muotoon $x^T Ax$, missä symmetrinen matriisi on

$$A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}.$$

Olkoon $d \neq 0$. Tällöin yhtälön (5.7) ratkaisujen joukossa on

(i) hyperbeli, jos $\det A < 0$,
(ii) ellipsi, ympyrä, piste tai tyhjä joukko, jos $\det A > 0$,
(iii) kaksi yhdensuuntaista suoraa, jos $\det A = 0$.

Toisaalta, jos pätee $d = 0$, niin ratkaisujen joukko antaa suoran, jos $\det A = 0$, ja kaksi toisiaan leikkaavaa suoraa, jos $\det A \neq 0$.

Todistus. Ks. [1, s. 427]. □

Tarkastellaan toisen asteen käyrän yhtälöä

$$x_1^2 - 4x_1x_2 + 3x_2^2 = 6. \tag{5.8}$$

Kirjoitetaan nyt yhtälö (5.8) muotoon $x^T Ax$, missä

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

on symmetrinen matriisi, jonka determinantti on $\det A = -1 < 0$. Tällöin (5.8) on hyperbelin yhtälö.

5.25. Ehdon (5.5) mukaan neliömuoto q ei vaihda merkkiä suoralla $\langle x \rangle$, $x \neq 0$. Neliömuodon q käyttäytyminen määräytyy myös olennaisesti sitä vastaavan matriin A ominaisvektorikannan mukaan. Jos $B = \{u_1, \ldots, u_n\}$ on matriisin A ominaisarvoihin $\lambda_1, \ldots, \lambda_n$ liittyvä ortonormaali ominaisvektorikanta, niin lauseen 5.19 nojalla on

$$A = QDQ^T,$$

missä $Q = [u_1 \ \ldots \ u_n]$ ja D diagonaalimatriisi, jonka diagonaalilla on matriisin A ominaisarvot. Tällöin avaruuden $\mathbb{R}^n$ luonnollisessa kannassa E annetulla vektorilla $x = (x_1, \ldots, x_n) = [x]_E$ on kannassa B esitys

$$y = (y_1, \ldots, y_n)_B := [x]_B = K(E, B)x = Q^T x,$$

missä $K(E, B)$ on kannanvaihtomatriisi kannasta E kantaan B. Tällöin neliömuodolle q saadaan kannan B suhteen esitys

$$\begin{aligned} q(x) &= x^T A x = x^T Q D Q^T x \\ &= (Q^T x)^T D (Q^T x) \\ &= y^T D y \\ &= \sum_{i=1}^n \lambda_i y_i^2 =: \tilde{q}(y), \end{aligned}$$

missä neliömuodon $\tilde{q} = \tilde{q}(y)$ lauseke sisältää vain puhtaita neliötermejä. Erityisesti pätee $q(u_i) = \lambda_i$, $i = 1, \ldots, n$. Tämä johtaa seuraavan määritelmän 5.26 luokitukseen ja lauseeseen 5.28.

Määritelmä 5.26. Sanotaan, että neliömuoto $q : \mathbb{R}^n \to \mathbb{R}$ on

(i) *positiivisesti/negatiivisesti definiitti*, jos

$$q(x) > 0/q(x) < 0 \text{ kaikille } x \neq 0;$$

(ii) *positiivisesti/negatiivisesti semidefiniitti*, jos

$$q(x) \geq 0/q(x) \leq 0 \text{ kaikille } x \in \mathbb{R} \text{ ja } q(x) = 0 \text{ jollekin } x \neq 0;$$

(iii) *indefiniitti*, jos

$$q(x) > 0 \text{ ja } q(x') < 0 \text{ joillekin } x \in \mathbb{R}^n \text{ ja } x' \in \mathbb{R}^n.$$

Esimerkki 5.27. (a) Neliömuoto $q : \mathbb{R}^2 \to \mathbb{R}$,

$$q(x) = 4x_1^2 - 4x_1x_2 + 4x_2^2 = (x_1 - 2x_2)^2 + 3x_1^2 > 0,$$

on selvästi positiivisesti definiitti.

(b) Neliömuoto $q : \mathbb{R}^2 \to \mathbb{R}$,

$$q(x) = x_1^2 - 2x_1x_2 + x_2^2 = (x_1 - x_2)^2 \geq 0,$$

on positiivisesti semidefiniitti, koska esimerkiksi pisteessä $(1, 1) \neq 0$ on $q(1, 1) = 0$.

(c) Vastaavasti neliömuoto $q : \mathbb{R}^3 \to \mathbb{R}$,

$$q(x) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3,$$

on indefiniitti, koska esimerkiksi pisteissä $(0, 1, -1)$ ja $(0, 0, 1)$ on $q(0, 1, -1) = -1 < 0$ ja $q(0, 0, 1) = 1 > 0$.

LAUSE 5.28. *Oletetaan, että neliömuodon $q : \mathbb{R}^n \to \mathbb{R}$ esityksessä $q(x) = x^T Ax$ symmetrisen matriisin A ominaisarvot ovat $\lambda_1, \dots, \lambda_n$. Tällöin neliömuoto q on*

(i) *positiivisesti/negatiivisesti definiitti jos ja vain jos*

$$\lambda_i > 0/\lambda_i < 0 \text{ jokaiselle } i = 1, \dots, n;$$

(ii) *positiivisesti/negatiivisesti semidefiniitti jos ja vain jos*

$$\lambda_i \geq 0/\lambda_i \leq 0 \text{ jokaiselle } i = 1, \dots, n \text{ ja } \lambda_i = 0 \text{ jollekin } i;$$

(iii) *indefiniitti jos ja vain jos $\lambda_i > 0$ ja $\lambda_j < 0$ joillekin i ja j.*

TODISTUS. Kohdan 5.25 mukaan pätee yhtälö $q(x) = x^T Ax = y^T Dy = \sum_{i=1}^n \lambda_i y_i^2$, mistä väitteen vaihtoehdot seuraavat (ks. [3, s. 417]). □

ESIMERKKI 5.29. Esimerkin 5.27 (c)-kohdan neliömuoto

$$q(x) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3 = x^T \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix} x = x^T Ax$$

on indefiniitti, koska matriisin A ominaisarvot ovat $\lambda_1 = 5 > 0$, $\lambda_2 = 2 > 0$ ja $\lambda_3 = -1 < 0$.

HUOMAUTUS 5.30. Neliömuodon $q : \mathbb{R}^n \to \mathbb{R}$, $q(x) = x^T Ax$, vastinmatriisin $A = [a_{ki}]$ *(pää)alimatriisit*

$$A_j = [a_{ki}]_{k,i=1}^j, \quad j = 1, \dots, n,$$

ovat myös symmetrisiä. Osoittautuu, että neliömuodon mahdollinen definiittisyys voidaan selvittää myös seuraavalla determinanttiehdolla: neliömuoto q on positiivisesti/negatiivisesti definiitti jos ja vain jos pätee

$$\det A_j > 0/(-1)^j \det A_j > 0 \quad \text{kaikilla} \quad j = 1, \dots, n.$$

ESIMERKKI 5.31. Neliömuoto

$$q(x) = 3x_1^2 + 3x_2^2 + 2x_3^2 + 2x_1x_2 + 4x_2x_3 = x^T Ax$$

on positiivisesti definiitti, koska symmetrisen matriisin

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

pääalimatriisien determinantit ovat

$$\det A_1 = 3 > 0, \quad \det A_2 = \begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} = 8 > 0 \quad \text{ja} \quad \det A_3 = \begin{vmatrix} 3 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & 2 & 2 \end{vmatrix} = 4 > 0.$$

Neliömuoto

$$-q(x) = -3x_1^2 - 3x_2^2 - 2x_3^2 - 2x_1x_2 - 4x_2x_3$$

on negatiivisesti definiitti.

5. Differentiaalilaskennan ääriarvotarkasteluja

5.32. Ominaisarvoja voidaan soveltaa differentiaalilaskennassa ääriarvotarkasteluihin, joissa kriittisiä pisteitä voidaan luokitella esimerkiksi ominaisarvojen avulla.

Avoimessa joukossa $G \subset \mathbb{R}^n$ kaksi kertaa jatkuvasti differentioituvan funktion $f : G \to \mathbb{R}$ toisen kertaluvun differentiaali pisteessä $a \in G$

$$u \mapsto d_a^2 f(u) = \sum_{i,j=1}^{n} \partial_i \partial_j f(a) u_i u_j$$

on neliömuoto, jota vastaa funktion f matriisi

$$\begin{bmatrix} \partial_1\partial_1 f(a) & \dots & \partial_1\partial_n f(a) \\ \vdots & & \vdots \\ \partial_n\partial_1 f(a) & \dots & \partial_n\partial_n f(a) \end{bmatrix},$$

ns. *Hessen matriisi.* Sanotaan, että $a \in G$ on funktion f kriittinen piste, jos pätee

$$\nabla f(a) := (\partial_1 f(a), \dots, \partial_n f(a)) = 0.$$

Jos kriittinen piste ei ole ääriarvopiste, niin se on *satulapiste.*

LAUSE 5.33. *Olkoon $f : G \to \mathbb{R}$ kaksi kertaa jatkuvasti differentioituva funktio avoimessa joukossa $G \subset \mathbb{R}^n$ ja olkoon $a \in G$ sen kriittinen piste. Jos nyt neliömuotoa $u \mapsto d_a^2 f(u)$ vastaavan symmetrisen matriisin $A = [\partial_i \partial_j f(a)]$ ominaisarvoille pätee*

(i) $\lambda_i > 0$ *jokaiselle* $i = 1, \dots, n$, *niin a on lokaali aito minimipiste;*
(ii) $\lambda_i < 0$ *jokaiselle* $i = 1, \dots, n$, *niin a on lokaali aito maksimipiste;*
(iii) $\lambda_i > 0$ *ja* $\lambda_j < 0$ *joillekin i ja j, niin a on satulapiste.*

TODISTUS. Ks. lause 5.28 ja [4, s. 73-74]. □

HUOMAUTUS 5.34. Lauseen soveltamiseksi riittää selvittää matriisin $A = [\partial_i \partial_j f(a)]$ ominaisarvot ehdosta $\det(A - \lambda I) = 0$, mikä on pelkkää polynomin juurien määritystä.

ESIMERKKI 5.35. Funktion

$$f : \mathbb{R}^3 \to \mathbb{R} : x \mapsto x_1^3 + x_3^3 + x_2^2 - 2x_1x_2 - x_1 - 3x_3$$

yhtälö

$$\nabla f(x) = (3x_1^2 - 2x_2 - 1,\ -2x_1 + 2x_2,\ 3x_3^2 - 3) = 0$$

toteutuu, kun $x = (1,1,1)$, $(1,1,-1)$, $(-\frac{1}{3}, -\frac{1}{3}, 1)$ tai $(-\frac{1}{3}, -\frac{1}{3}, -1)$, jolloin funktion kriittiset pisteet ovat $a_1 = (1,1,1)$, $a_2 = (1,1,-1)$, $a_3 = (-\frac{1}{3}, -\frac{1}{3}, 1)$ ja $a_4 = (-\frac{1}{3}, -\frac{1}{3}, -1)$.

Funktion f neliömuoto on

$$d_x^2 f(u) = 6x_1u_1^2 + 2u_2^2 + 6x_3u_3^2 - 4u_1u_2.$$

Pisteessä $a_1 = (1,1,1)$ neliömuotoa

$$d_{a_1}^2 f(u) = 6u_1^2 + 2u_2^2 + 6u_3^2 - 4u_1u_2$$

vastaavan matriisin kaikki kolme ominaisarvoa ovat aidosti positiivisia, jolloin a_1 on lokaali aito minimipiste.

Pisteessä $a_2 = (1, 1, -1)$ neliömuotoon

$$d^2_{a_2} f(u) = 6u_1^2 + 2u_2^2 - 6u_3^2 - 4u_1u_2$$

liittyvän matriisin kaksi ominaisarvoa ovat aidosti positiivisia ja yksi aidosti negatiivinen, jolloin a_2 on satulapiste.

Samoin pisteessä $a_3 = (-\frac{1}{3}, -\frac{1}{3}, 1)$ neliömuotoon

$$d^2_{a_3} f(u) = -2u_1^2 + 2u_2^2 + 6u_3^2 - 4u_1u_2$$

liittyvän matriisin ominaisarvoista kaksi ovat aidosti positiivisia ja yksi aidosti negatiivinen, jolloin a_3 on satulapiste.

Vastaavasti $a_4 = (-\frac{1}{3}, -\frac{1}{3}, -1)$ on satulapiste, koska neliömuotoa

$$d^2_{a_4} f(u) = -2u_1^2 + 2u_2^2 - 6u_3^2 - 4u_1u_2$$

vastaavan matriisin ominaisarvoista yksi on aidosti positiivinen ja kaksi aidosti negatiivista.

LAUSE 5.36. *Oletetaan, että edellisen lauseen 5.33 oletukset ovat voimassa, ja merkitään*

$$\Delta_k(x) := \det[\partial_i \partial_j f(x)]_{i,j=1}^k, \quad k = 1, \dots, n,$$

joka on funktion f Hessen matriisin pääalideterminantti. Tällöin kriittinen piste a on

(i) *lokaali aito minimipiste, jos*

$$\Delta_k(a) > 0 \quad \textit{jokaiselle } k = 1, \dots, n;$$

(ii) *lokaali aito maksimipiste, jos*

$$(-1)^k \Delta_k(a) > 0 \quad \textit{jokaiselle } k = 1, \dots, n.$$

TODISTUS. Seuraa olennaisesti huomautuksesta 5.30. □

ESIMERKKI 5.37. Esimerkin 5.35 funktion

$$f : \mathbb{R}^3 \to \mathbb{R} : x \mapsto x_1^3 + x_3^3 + x_2^2 - 2x_1x_2 - x_1 - 3x_3$$

Hessen matriisi kriittisessä pisteessä $a_1 = (1, 1, 1)$ on

$$[\partial_i \partial_j f(a_1)] = \begin{bmatrix} 6 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

Sen pääalideterminantit ovat

$$\Delta_1(a_1) = 6 > 0, \quad \Delta_2(a_1) = \begin{vmatrix} 6 & -2 \\ -2 & 2 \end{vmatrix} = 8 > 0 \quad \text{ja}$$

$$\Delta_3(a_1) = \begin{vmatrix} 6 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 6 \end{vmatrix} = 48 > 0,$$

joten a_1 lokaali aito minimipiste.

LUKU 6

Ehdotus lukion pitkän matematiikan *Numeerisia ja algebrallisia menetelmiä* -kurssin sisältöön

Lukion opetussuunnitelman perusteet (2003) sisältää kolme pitkän matematiikan syventävää kurssia. Yksi näistä on *Numeerisia ja algebrallisia menetelmiä* (MAA12), jonka tavoitteisiin kuuluu muun muassa, että opiskelija oppii ratkaisemaan yhtälöryhmiä numeerisesti sekä harjaantuu käyttämään nykyaikaisia matemaattisia välineitä. Lukion pitkän matematiikan opetussuunnitelman perusteissa ei vaadita minkään kurssin sisältöön matriiseja ja determinantteja. Mielestäni ne kuitenkin sopisivat juuri tähän kurssiin (MAA12), jossa voitaisiin syventää kurssin *Analyyttinen geometria* (MAA4) tietoja lineaarisista yhtälöistä ja yhtälöryhmistä. Kurssiin sopii lineaarista algebraa ja geometriaa käsittelevä osio, joka koostuu lineaaristen yhtälöryhmien kertauksesta, matriiseista ja determinanteista sekä Cramerin säännöstä. Kurssilla rajoitutaan käsittelemään 2×2 ja 3×3 -matriiseja. Kurssin asiat eivät ole laskennallisesti vaikeita, mutta mukana olevat geometriset tulkinnat voivat vaatia tarkentavaa selitystä opetustilanteessa, eikä muutenkaan käsiteltäviä asioita tulisi jättää opiskelijan lukemisen varaan. Kurssi on syventävä, joten asiat voivatkin olla hieman vaikeampia kuin pakollisilla kursseilla.

Cramerin sääntö otetaan 2×2 ja 3×3 -matriiseille yksinkertaisessa muodossa. Cramerin sääntö vaatii, että matriisin A tulee olla kääntyvä, jotta sääntö olisi voimassa. Tämä ehto voidaan korvata sillä tiedolla, että matriisin determinantti ei saa olla nolla, koska tiedetään, että matriisi A on kääntyvä, jos ja vain jos pätee $\det A \neq 0$. Cramerin sääntöä yksinkertaistetaan myös siten, että lineaarisia yhtälöryhmiä ei kirjoiteta matriisimuotoon $Ax = b$, koska tällöin olisi hyvä käsitellä selvyyden vuoksi matriiseilla kertominenkin. Tästä syystä rajoitutaan käsittelemään lineaaristen yhtälöryhmien ratkaisemista hieman muokatulla Cramerin säännöllä, jota perustellaan geometrisesti 2×2-matriisille. Lukion pitkän matematiikan opetussuunnitelman perusteissa ei vaadita todistamista pakollisiin kursseihin. Kuitenkin kirjasarjojen teoriaosioissa tuloksia pyritään jossain määrin johtamaan ja todistamaan. Monet lukiossa käytetyt tulokset ja niiden todistukset joudutaan yksinkertaistamaan, koska yleensä tarkkoihin tuloksiin ja todistuksiin tarvitaan sellaisia tietoja, joita opiskelijalla ei vielä ole.

Kurssiin voidaan sisällyttää matemaattiset tietokoneohjelmat, kuten *Mathematica* ja *Maxima*, joiden avulla esimerkiksi determinantteja voidaan määrittää. Myös suorien piirtäminen onnistuu hyvin, jolloin yhtälöryhmien ratkaisuja voidaan tarkastaa kuvaajien avulla. Matemaattisten ohjelmien ottaminen opetukseen ja opiskeluun mukaan motivoi useita opiskelijoita. Tietokoneohjelmat ovat hyviä nykyaikaisia matemaattisia apuvälineitä graafisen laskimen ohella. Näillä opiskelija pystyy tarkastamaan monien tehtävien ratkaisuja sekä soveltamaan ja kehittämään tietojaan ja taitojaan.

1. Lineaarista algebraa ja geometriaa lukiossa

6.1. **Lyhyt kertaus lineaarisista yhtälöryhmistä.** Kahden lineaarisen yhtälön yhtälöryhmä on muotoa

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2, \end{cases} \tag{6.1}$$

missä kertoimet a_{ij} ja vakiotermit b_i ovat reaalilukuja kaikilla $i = 1, 2$ ja $j = 1, 2$, ja x ja y ovat tuntemattomia muuttujia. Yhtälö on lineaarinen, jos se on jokaisen muuttujan suhteen ensimmäistä astetta.

Vastaavasti kolmen lineaarisen yhtälön yhtälöryhmä on muotoa

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3, \end{cases} \tag{6.2}$$

missä kertoimet a_{ij} ja vakiotermit b_i ovat reaalilukuja kaikilla $i = 1, 2, 3$ ja $j = 1, 2, 3$, ja x, y ja z ovat tuntemattomia muuttujia.

Yhtälöryhmän (6.1) piste (x, y) on sen ratkaisu, jos se toteuttaa molemmat yhtälöryhmän yhtälöt. Vastaavasti yhtälöryhmän (6.2) piste (x, y, z) on sen ratkaisu, jos se toteuttaa yhtälöryhmän kaikki kolme yhtälöä. Yhtälöryhmien lineaariset yhtälöt ovat suorien yhtälöitä. Tällöin yhtälöryhmällä on yksi ratkaisu, jos yhtälöryhmän suorat leikkaavat toisensa yhdessä pisteessä. Jos suorat ovat yhdensuuntaisia ja vakiotermit b_i ovat kaikki eri lukuja, niin suorat eivät leikkaa toisiaan eikä yhtälöryhmällä ole ratkaisua. Jos taas suorat ovat yhdensuuntaisia ja vakiotermit b_i ovat kaikki samoja lukuja, niin suorat yhtyvät samaksi suoraksi, jolloin kaikki suorien pisteet toteuttavat yhtälöryhmän ja yhtälöryhmällä on äärettömän monta ratkaisua.

6.2. **Lineaarisesta yhtälöryhmästä muodostuva matriisi.** Tarkastellaan kohdan 6.1 yhtälöryhmiä. Yhtälöryhmän (6.1) kertoimet a_{ij} voidaan kirjoittaa taulukkoon seuraavasti

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Tätä sanotaan 2×2-*matriisiksi* ja merkitään $A = A_{2\times 2}$. Tämä matriisi koostuu kahdesta *sarakkeesta* ja kahdesta *rivistä*.

Vastaavasti yhtälöryhmän (6.2) kertoimet a_{ij} voidaan kirjoittaa seuraavasti taulukkoon

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

jota sanotaan 3×3-*matriisiksi* ja merkitään $A = A_{3\times 3}$. Huomaa, että tämä matriisi koostuu kolmesta sarakkeesta ja kolmesta rivistä.

Matriisimuoto on monessa tilanteessa selkeämpi käyttää kuin lineaarinen yhtälöryhmä.

Esimerkki 6.3. (a) Yhtälöryhmän

$$\begin{cases} x + 2y = 1 \\ 3x + 4y = 2, \end{cases}$$

kertoimista saadaan matriisi

$$A_{2\times 2} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

(b) Yhtälöryhmän

$$\begin{cases} 3x + 4y + 5z = 1 \\ 6x + 7y + 8z = 2 \\ 9x + y + 2z = 3 \end{cases}$$

kertoimista saadaan matriisi

$$A_{3\times 3} = \begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \\ 9 & 1 & 2 \end{bmatrix}.$$

6.4. **Matriisin determinantti.** Oletetaan, että yhtälöryhmän

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases} \tag{6.3}$$

kertoimille pätee $a_{ij} \neq 0$. Kun kerrotaan yhtälöryhmän (6.3) ensimmäinen yhtälö reaaliluvulla a_{22} ja toinen yhtälö reaaliluvulla a_{12}, saadaan yhtälöryhmä

$$\begin{cases} a_{11}a_{22}x + a_{12}a_{22}y = a_{22}b_1 \\ a_{12}a_{21}x + a_{12}a_{22}y = a_{12}b_2. \end{cases}$$

Vähennetään nämä yhtälöt toisistaan, jolloin saadaan

$$(a_{11}a_{22} - a_{12}a_{21})x = a_{22}b_1 - a_{12}b_2,$$

josta saadaan ratkaistua

$$x = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} \quad \text{ja} \quad y = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}},$$

kun $(a_{11}a_{22} - a_{12}a_{21}) \neq 0$.

Yhtälöparista (6.3) saatua lauseketta $a_{11}a_{22} - a_{12}a_{21}$ sanotaan yhtälöparista saatavan matriisin

$$A = A_{2\times 2} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

determinantiksi ja sitä merkitään

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}a_{22} - a_{12}a_{21}).$$

Huomataan, että yhtälöparilla (6.3) on yksi ratkaisu, jos pätee $(a_{11}a_{22} - a_{12}a_{21}) \neq 0$, eli jos matriisin A determinantille on voimassa $\det A \neq 0$.

Määritellään seuraavaksi 3×3-matriisin determinantti. Olkoon matriisi

$$A = A_{3\times 3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Tällöin sen determinantti on

$$\begin{aligned}\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}\left(a_{22}a_{33} - a_{23}a_{32}\right) - a_{21}\left(a_{12}a_{33} - a_{13}a_{32}\right) + a_{31}\left(a_{12}a_{23} - a_{13}a_{22}\right).\end{aligned}$$

Determinantti on siis luku.

Huomautus 6.5. Jos matriisin A determinantti on $\det A \neq 0$, niin tällöin sillä yhtälöryhmällä, josta matriisi on saatu, on yksi ratkaisu. Jos pätee $\det A = 0$, niin yhtälöryhmällä ei ole ratkaisua tai niitä on äärettömän monta. Tätä tietoa hyödynnettiin Kiinassa jo 200-luvulla eKr yhtälöryhmiä ratkaistaessa.

Esimerkki 6.6. Esimerkin 6.3 matriiseille voidaan laskea determinantit seuraavasti:

(a) Matriisin

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

determinantti on

$$\det A = |A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2,$$

jolloin yhtälöryhmällä

$$\begin{cases} x + 2y = 1 \\ 3x + 4y = 2 \end{cases}$$

on ratkaisu.

(b) Matriisin

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \\ 9 & 1 & 2 \end{bmatrix}$$

determinantti on

$$\begin{aligned}\det A = |A| = \begin{vmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \\ 9 & 1 & 2 \end{vmatrix} &= 3 \begin{vmatrix} 7 & 8 \\ 1 & 2 \end{vmatrix} - 6 \begin{vmatrix} 4 & 5 \\ 1 & 2 \end{vmatrix} + 9 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\ &= 3\left(7 \cdot 2 - 8 \cdot 1\right) - 6\left(4 \cdot 2 - 5 \cdot 1\right) + 9\left(4 \cdot 8 - 5 \cdot 7\right) \\ &= -27,\end{aligned}$$

jolloin yhtälöryhmällä

$$\begin{cases} 3x + 4y + 5z = 1 \\ 6x + 7y + 8z = 2 \\ 9x + y + 2z = 3 \end{cases}$$

on ratkaisu.

Huomautus 6.7. Determinantteja voidaan määrittää tietokoneella matemaattisilla ohjelmilla kuten *Mathematicalla* ja *Maximalla*. Näitä ohjelmia kannattaa hyödyntää esimerkiksi determinanttien arvojen tarkastamisessa.

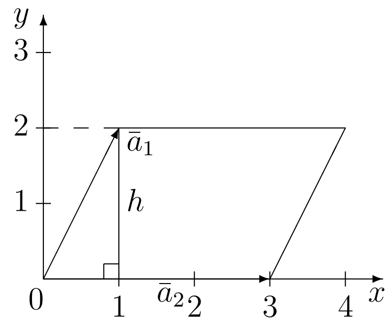
6.8. **Determinantin geometrista tulkintaa.** Tarkastellaan, mitä 2×2-matriisin determinantti tarkoittaa geometrisesti. Olkoon matriisi

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

jonka sarakkeet ovat *sarakevektoreita* $\bar{a}_1 = (a_{11}, a_{21})$ ja $\bar{a}_2 = (a_{12}, a_{22})$, jotka lähtevät aina origosta eli pisteestä $(0,0)$. Kaksi sarakevektoria virittävät yhdessä suoran tai suunnikkaan. Esimerkiksi matriisin

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

sarakevektorit ovat $\bar{a}_1 = (1,2)$ ja $\bar{a}_2 = (3,0)$. Nämä molemmat lähtevät origosta ja sarakevektori $\bar{a}_1$ päättyy pisteeseen $(1,2)$ ja sarakevektori $\bar{a}_2$ päättyy pisteeseen $(3,0)$. Piirretään nämä vektorit kuvaan ja huomataan, että ne virittävät suunnikkaan, jonka kärkipisteet ovat $(0,0)$, $(1,2)$, $(3,0)$ ja neljäs kärkipiste saadaan laskemalla yhteen vektorit $\bar{a}_1$ ja $\bar{a}_2$ eli $\bar{a}_1 + \bar{a}_2 = (1,2) + (3,0) = (1+3, 2+0) = (4,2)$.



Kuva (12)

Vektoreiden $\bar{a}_1$ ja $\bar{a}_2$ virittämän suunnikkaan pinta-ala saadaan kertomalla kannan pituus eli vektorin $\bar{a}_2$ pituus $|\bar{a}_2| = \sqrt{3^2 + 0^2} = 3$ korkeudella $h = 2$, jolloin on voimassa

$$\text{ala}(\bar{a}_1, \bar{a}_2) = |\bar{a}_2| \cdot h = 3 \cdot 2 = 6.$$

Merkinnällä $\text{ala}(\bar{a}_1, \bar{a}_2)$ tarkoitetaan vektoreiden $\bar{a}_1$ ja $\bar{a}_2$ virittämän suunnikkaan pinta-alaa.

Matriisin A determinantti on

$$\det A = \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} = 1 \cdot 0 - 3 \cdot 2 = -6.$$

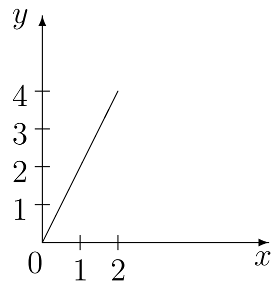
Matriisin $A = A_{2\times 2}$ sarakevektoreiden virittämälle pinta-alalle ja determinantille päteekin tulos

$$\text{ala}(\bar{a}_1, \bar{a}_2) = |\det A|.$$

Jos matriisin A determinantti on $\det A = 0$, niin tällöin sarakevektorit virittävät suoran, jonka pinta-ala on nolla. Näin käy esimerkiksi matriisin

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

sarakevektoreille $\bar{a}_1 = (1,2)$ ja $\bar{a}_2 = (2,4)$. Kuvasta (13) nähdään, että ne virittävät suoran.



Kuva (13)

Matriisin A determinantti voidaan myös laskea ja saadaan

$$\det A = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 2 = 0.$$

6.9. Aiemmin todettiin, että yhtälöryhmästä

$$(6.4) \qquad \begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

saadaan matriisi

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

Jos matriisin A ensimmäiseen sarakkeeseen sijoitetaan yhtälöryhmän (6.4) vakiotermit b_1 ja b_2, saadaan matriisi

$$A_1(b) = \begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}.$$

Sijoittamalla vakiotermit matriisin A toiseen sarakkeeseen saadaan matriisi

$$A_2(b) = \begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix}.$$

Aiemmin todettiin myös, että yhtälöryhmästä

$$(6.5) \qquad \begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases}$$

saadaan matriisi

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Jos yhtälöryhmän (6.5) vakiotermit b_1, b_2 ja b_3 sijoitetaan matriisin A ensimmäiseen sarakkeeseen, saadaan matriisi

$$A_1(b) = \begin{bmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{bmatrix}.$$

Vastaavasti vakiotermit voidaan sijoittaa myös toiseen ja kolmanteen sarakkeeseen, joilloin saadaan matriisit $A_2(b)$ ja $A_3(b)$. Näitä vakiotermien avulla saatuja matriiseja tarvitaan seuraavassa lauseessa, joka antaa Cramerin säännön nimellä tunnetun tuloksen. Tämän tuloksen avulla voidaan ratkaista lineaarisia yhtälöryhmiä. Sääntöä perustellaan geometrisesti kohdassa 6.14.

LAUSE 6.10. *Jos yhtälöryhmän*

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

matriisille $A = A_{2\times 2}$ *on* $\det A \neq 0$, *niin sen ratkaisulle* (x, y) *pätee*

$$x = \frac{\det A_1(b)}{\det A} \quad \text{ja} \quad y = \frac{\det A_2(b)}{\det A}.$$

Vastaavasti, jos yhtälöryhmän

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases}$$

matriisille $A = A_{3\times 3}$ *on* $\det A \neq 0$, *niin sen ratkaisu* (x, y, z) *on*

$$x = \frac{\det A_1(b)}{\det A}, \quad y = \frac{\det A_2(b)}{\det A} \quad \text{ja} \quad z = \frac{\det A_3(b)}{\det A}.$$

HUOMAUTUS 6.11. *Mathematicaa*, *Maximaa* tai graafista laskinta kannattaa hyödyntää myös suorien piirtämisessä, jolloin yhtälöryhmien ratkaisut saadaan tarkastettua.

ESIMERKKI 6.12. Ratkaistaan esimerkin 6.6 lineaariset yhtälöryhmät.

(a) Lineaariselle yhtälöryhmälle

$$\begin{cases} x + 2y = 1 \\ 3x + 4y = 2 \end{cases} \tag{6.6}$$

saadaan ratkaisu Cramerin säännön nojalla, koska esimerkissä 6.6 saatiin, että tämän lineaarisen yhtälöryhmän matriisin

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

determinantti on $\det A = -2 \neq 0$. Jos yhtälöryhmän (6.6) vakiotermit 1 ja 2 sijoitetaan matriisin A ensimmäiseen sarakkeeseen, saadaan matriisi

$$A_1(b) = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix},$$

jonka determinantti on

$$\det A_1(b) = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 2 = 0.$$

Jos vakiotermit 1 ja 2 sijoitetaan matriisin A toiseen sarakkeeseen, saadaan matriisi

$$A_2(b) = \begin{bmatrix} 1 & 1 \\ 3 & 2 \end{bmatrix},$$

jonka determinantti on

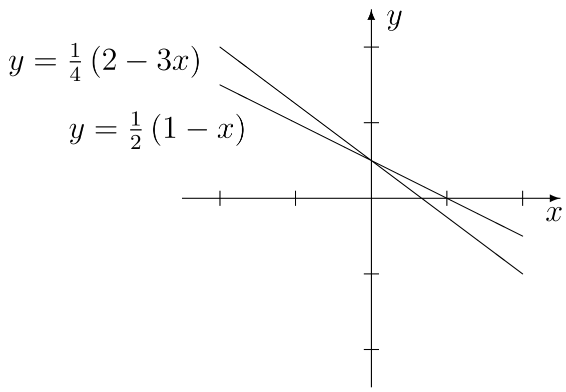
$$\det A_2(b) = \begin{vmatrix} 1 & 1 \\ 3 & 2 \end{vmatrix} = 1 \cdot 2 - 1 \cdot 3 = -1.$$

Tällöin Cramerin säännön nojalla saadaan yhtälöryhmän (6.6) ratkaisuksi

$$x = \frac{\det A_1(b)}{\det A} = \frac{0}{-2} = 0 \quad \text{ja} \quad y = \frac{\det A_2(b)}{\det A} = \frac{-1}{-2} = \frac{1}{2}$$

eli piste $(0, \frac{1}{2})$.

Yhtälöryhmän (6.6) suorat $x + 2y = 1$ ja $3x + 4y = 2$ voidaan kirjoittaa muotoon $y = \frac{1}{2}(1 - x)$ ja $y = \frac{1}{4}(2 - 3x)$. Piirretään nämä suorat kuvaan.



Kuva (14)

Kuvasta (14) nähdään, että suorat leikkaavat pisteessä $\left(0, \frac{1}{2}\right)$ ja tästäkin syystä tämä piste $\left(0, \frac{1}{2}\right)$ on yhtälöryhmän ratkaisu.

(b) Myös lineaariselle yhtälöryhmälle

$$\begin{cases} 3x + 4y + 5z = 1 \\ 6x + 7y + 8z = 2 \\ 9x + y + 2z = 3 \end{cases} \tag{6.7}$$

saadaan ratkaisu Cramerin säännön nojalla, koska esimerkissä 6.6 saatiin, että tämän lineaarisen yhtälöryhmän matriisin

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \\ 9 & 1 & 2 \end{bmatrix}$$

determinantti on $\det A = -27 \neq 0$. Jos yhtälöryhmän (6.7) vakiotermit 1, 2 ja 3 sijoitetaan matriisin A ensimmäiseen sarakkeeseen, saadaan matriisi

$$A_1(b) = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 7 & 8 \\ 3 & 1 & 2 \end{bmatrix},$$

jonka determinantti on

$$\begin{aligned}\det A_1(b) = |A_1(b)| = \begin{vmatrix} 1 & 4 & 5 \\ 2 & 7 & 8 \\ 3 & 1 & 2 \end{vmatrix} &= 1\begin{vmatrix} 7 & 8 \\ 1 & 2 \end{vmatrix} - 2\begin{vmatrix} 4 & 5 \\ 1 & 2 \end{vmatrix} + 3\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\ &= 1(7\cdot 2 - 8\cdot 1) - 2(4\cdot 2 - 5\cdot 1) + 3(4\cdot 8 - 5\cdot 7) \\ &= -9.\end{aligned}$$

Jos vakiotermit 1, 2 ja 3 sijoitetaan matriisin A toiseen sarakkeeseen, saadaan matriisi

$$A_2(b) = \begin{bmatrix} 3 & 1 & 5 \\ 6 & 2 & 8 \\ 9 & 3 & 2 \end{bmatrix},$$

jonka determinantti on

$$\begin{aligned}\det A_2(b) = |A_2(b)| = \begin{vmatrix} 3 & 1 & 5 \\ 6 & 2 & 8 \\ 9 & 3 & 2 \end{vmatrix} &= 3\begin{vmatrix} 2 & 8 \\ 3 & 2 \end{vmatrix} - 6\begin{vmatrix} 1 & 5 \\ 3 & 2 \end{vmatrix} + 9\begin{vmatrix} 1 & 5 \\ 2 & 8 \end{vmatrix} \\ &= 3(2\cdot 2 - 8\cdot 3) - 6(1\cdot 2 - 5\cdot 3) + 9(1\cdot 8 - 5\cdot 2) \\ &= 0.\end{aligned}$$

Sijoittamalla vakiotermit 1, 2 ja 3 matriisin A kolmanteen sarakkeeseen, saadaan matriisi

$$A_3(b) = \begin{bmatrix} 3 & 4 & 1 \\ 6 & 7 & 2 \\ 9 & 1 & 3 \end{bmatrix},$$

jonka determinantti on

$$\begin{aligned}\det A_3(b) = |A_3(b)| = \begin{vmatrix} 3 & 4 & 1 \\ 6 & 7 & 2 \\ 9 & 1 & 3 \end{vmatrix} &= 3\begin{vmatrix} 7 & 2 \\ 1 & 3 \end{vmatrix} - 6\begin{vmatrix} 4 & 1 \\ 1 & 3 \end{vmatrix} + 9\begin{vmatrix} 4 & 1 \\ 7 & 2 \end{vmatrix} \\ &= 3(7\cdot 3 - 2\cdot 1) - 6(4\cdot 3 - 1\cdot 1) + 9(4\cdot 2 - 1\cdot 7) \\ &= 0.\end{aligned}$$

Tällöin Cramerin säännön nojalla saadaan yhtälöryhmän (6.7) ratkaisuksi

$$x = \frac{\det A_1(b)}{\det A} = \frac{-9}{-27} = \frac{1}{3}, \quad y = \frac{\det A_2(b)}{\det A} = \frac{0}{-27} = 0 \quad \text{ja}$$

$$z = \frac{\det A_3(b)}{\det A} = \frac{0}{-27} = 0$$

eli piste $(\frac{1}{3}, 0, 0)$.

Esimerkki 6.13. Yrittäjä sijoittaa 100000 euroa pankkiin kolmelle eri tilille. Sekkitilille talletettu erä tuottaa 5 prosentin vuotuisen koron, säästötilille talletettu erä tuottaa 6 prosentin vuotuisen koron ja vuoden määräaikaistilille sijoitettu erä tuottaa 7 prosentin vuotuisen koron. Sijoitus säästötilille on kolminkertainen verrattuna sijoitukseen sekkitilille. Kuinka suuret ovat eri talletuserät, kun vuoden korkotuotto on yhteensä 6500 euroa?

Ratkaisu. Merkitään vakiolla x sekkitilille talletetun erän rahamäärää ja vastaavasti vakioilla y ja z säästötilille ja vuoden määräaikaistilille talletettujen erien rahamääriä. Kirjoitetaan näiden avulla ongelma lineaariseksi yhtälöryhmäksi

$$\begin{cases} x+y+z=100000 \\ -3x+y=0 \\ 0,05x+0,06y+0,07z=6500, \end{cases} \tag{6.8}$$

jonka alkiot voidaan kirjoittaa matriisiksi

$$A=\begin{bmatrix} 1 & 1 & 1 \\ -3 & 1 & 0 \\ 0,05 & 0,06 & 0,07 \end{bmatrix}.$$

Matriisin A determinantti on

$$\begin{aligned} \det A=|A| &= \begin{vmatrix} 1 & 1 & 1 \\ -3 & 1 & 0 \\ 0,05 & 0,06 & 0,07 \end{vmatrix} \\ &= 1\begin{vmatrix} 1 & 0 \\ 0,06 & 0,07 \end{vmatrix} -(-3)\begin{vmatrix} 1 & 1 \\ 0,06 & 0,07 \end{vmatrix} +0,05\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} \\ &= 0,07+0,03-0,05 \\ &= 0,05\neq 0. \end{aligned}$$

Sijoitetaan yhtälöryhmän (6.8) vakiotermit $b_1 = 100000$, $b_2 = 0$ ja $b_3 = 6500$ matriisin A ensimmäiseen sarakkeeseen, sitten toiseen ja lopuksi kolmanteen sarakkeeseen, jolloin saadaan matriisit

$$A_1(b)=\begin{bmatrix} 100000 & 1 & 1 \\ 0 & 1 & 0 \\ 6500 & 0,06 & 0,07 \end{bmatrix}, \quad A_2(b)=\begin{bmatrix} 1 & 100000 & 1 \\ -3 & 0 & 0 \\ 0,05 & 6500 & 0,07 \end{bmatrix} \quad \text{ja}$$

$$A_3(b)=\begin{bmatrix} 1 & 1 & 100000 \\ -3 & 1 & 0 \\ 0,05 & 0,06 & 6500 \end{bmatrix}.$$

Lasketaan näiden matriisien determinantit:

$$\begin{aligned} \det A_1(b) &= \begin{vmatrix} 100000 & 1 & 1 \\ 0 & 1 & 0 \\ 6500 & 0,06 & 0,07 \end{vmatrix} \\ &= 100000\begin{vmatrix} 1 & 0 \\ 0,06 & 0,07 \end{vmatrix} +6500\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} \\ &= 500, \end{aligned}$$

$$\begin{aligned} \det A_2(b) &= \begin{vmatrix} 1 & 100000 & 1 \\ -3 & 0 & 0 \\ 0,05 & 6500 & 0,07 \end{vmatrix} \\ &= 1\begin{vmatrix} 0 & 0 \\ 6500 & 0,07 \end{vmatrix} -(-3)\begin{vmatrix} 100000 & 1 \\ 6500 & 0,07 \end{vmatrix} +0,05\begin{vmatrix} 100000 & 1 \\ 0 & 0 \end{vmatrix} \\ &= 1500 \end{aligned}$$

ja

$$\begin{aligned}\det A_3\,(b) &= \begin{vmatrix} 1 & 1 & 100000 \\ -3 & 1 & 0 \\ 0,05 & 0,06 & 6500 \end{vmatrix} \\ &= 1\begin{vmatrix} 1 & 0 \\ 0,06 & 6500 \end{vmatrix} - (-3)\begin{vmatrix} 1 & 100000 \\ 0,06 & 6500 \end{vmatrix} + 0,05\begin{vmatrix} 1 & 100000 \\ 1 & 0 \end{vmatrix} \\ &= 3000.\end{aligned}$$

Tällöin Cramerin säännön nojalla saadaan yhtälöryhmän (6.8) ratkaisuksi

$$x = \frac{\det A_1\,(b)}{\det A} = \frac{500}{0,05} = 10000, \quad y = \frac{\det A_2\,(b)}{\det A} = \frac{1500}{0,05} = 30000 \quad \text{ja}$$

$$z = \frac{\det A_3\,(b)}{\det A} = \frac{3000}{0,05} = 60000$$

eli sekkitilillä on 10000 euroa, säästötilillä on 30000 euroa ja vuoden määräaikaistilillä on 60000 euroa.

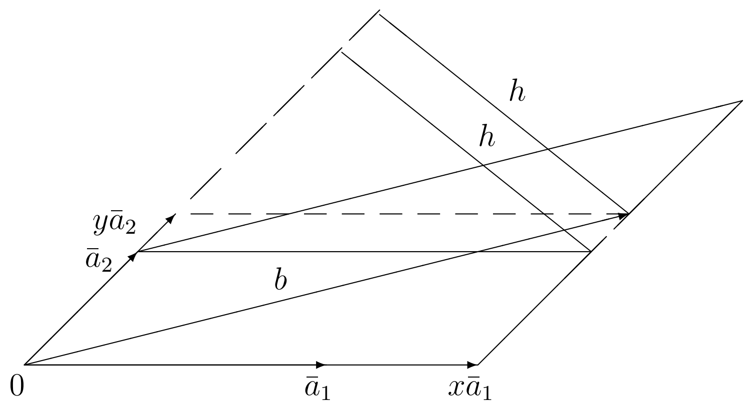
6.14. **Cramerin säännön geometrista tulkintaa.** Tarkastellaan, miten yhtälöryhmä

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases} \tag{6.9}$$

voidaan ratkaista geometrisesti. Yhtälöryhmästä (6.9) saadaan kahdesta sarakevektorista $\bar{a}_1 = (a_{11}, a_{21})$ ja $\bar{a}_2 = (a_{12}, a_{22})$ koostuva matriisi A, joka voidaan kirjoittaa sarakevektoreiden avulla muotoon $A = [\bar{a}_1\ \bar{a}_2]$. Huomataan, että yhtälöryhmässä (6.9) sarakevektorin $\bar{a}_1$ kertoimet on kerrottu muuttujalla x ja vastaavasti sarakevektorin $\bar{a}_2$ kertoimet on kerrottu muuttujalla y. Huomataan lisäksi, että vakiotermit b_1 ja b_2 voidaan kirjoittaa sarakevektorina $b = (b_1, b_2)$. Nyt yhtälöryhmä (6.9) voidaan kirjoittaa muotoon

$$x\bar{a}_1 + y\bar{a}_2 = b,$$

jolloin vektori b on kahden vektorin summa. Piirretään sitten nämä vektorit $\bar{a}_1$, $\bar{a}_2$, $x\bar{a}_1$, $y\bar{a}_2$ ja b kuvaan.



Kuva (15)

Tarkastelemalla vektoreiden virittämien suunnikkaiden pinta-aloja huomataan, että pätee

$$\text{ala}(x\bar{a}_1, \bar{a}_2) = \text{ala}(b, \bar{a}_2). \tag{6.10}$$

Tämä on tosi, koska vektoreiden $x\bar{a}_1$ ja $\bar{a}_2$ virittämän suunnikkaan kannan pituus on vektorin $\bar{a}_2$ pituus $|\bar{a}_2|$ ja korkeus on h, jolloin saadaan $\text{ala}(x\bar{a}_1, \bar{a}_2) = |\bar{a}_2| \cdot h$. Samoin vektoreiden b ja $\bar{a}_2$ virittämän suunnikkaan pinta-alaksi saadaan $\text{ala}(b, \bar{a}_2) = |\bar{a}_2| \cdot h$.

Muistetaan sitten, että pätee

$$\text{ala}(\bar{a}_1, \bar{a}_2) = |\det A| = |\det[\bar{a}_1\ \bar{a}_2]|,$$

joka voidaan kirjoittaa myös muodossa

$$\pm\,\text{ala}(\bar{a}_1, \bar{a}_2) = \det A = \det[\bar{a}_1\ \bar{a}_2].$$

Näin voidaan tehdä, koska pinta-ala on aina positiivista ja determinantin $\det[\bar{a}_1\ \bar{a}_2]$ etumerkki riippuu vektoreiden järjestyksestä eli kiertokulmasta. Valitaan $\sphericalangle(\bar{a}_1, \bar{a}_2)$ positiiviseksi kiertokulmaksi, kun kierretään vastapäivään. Ts. vastapäivään kierrettäessä vektoreiden järjestys on $\bar{a}_1$, $\bar{a}_2$. Tällöin $\det[\bar{a}_1\ \bar{a}_2]$ on positiivinen ja $\det[\bar{a}_2\ \bar{a}_1]$ on negatiivinen.

Huomataan, että kiertokulmat $\sphericalangle(x\bar{a}_1, \bar{a}_2)$ ja $\sphericalangle(b, \bar{a}_2)$ ovat samanmerkkisiä positiiviselle ja negatiiviselle x. Yhtälö (6.10) voidaan kirjoittaa muotoon

$$\det[x\bar{a}_1\ \bar{a}_2] = \det[b\ \bar{a}_2],$$

jolle pätee

$$x \cdot \det[\bar{a}_1\ \bar{a}_2] = \det[b\ \bar{a}_2],$$

jolloin saadaan

$$x = \frac{\det[b\ \bar{a}_2]}{\det[\bar{a}_1\ \bar{a}_2]} = \frac{\det A_1(b)}{\det A}.$$

Vastaavasti voidaan näyttää, että on

$$y = \frac{\det A_2(b)}{\det A}.$$

Näin saatiin yhtälöryhmälle (6.9) ratkaisu (x, y), joka on juuri Cramerin sääntö 2×2-matriisille.

Viitteet

[1] S. I. Grossman, *Elementary Linear Algebra*, Saunders College Publishing, 4th edition, 1991.
[2] P. Jäppinen, A. Kupiainen ja M. Räsänen, *Lukion Calculus 2*, Otava, 1.-2. painos, 2004.
[3] D. C. Lay, *Linear Algebra and Its Applications*, Addison-Wesley Publishing Company, 1st edition, 1994.
[4] V. T. Purmonen, *Differentiaalilaskentaa 1*, Jyväskylän yliopistopaino, 2009.
[5] V. T. Purmonen, *Integraalilaskentaa 1*, Jyväskylän yliopistopaino, 3. painos, 2009.
[6] V. T. Purmonen, *Lineaarinen algebra ja geometria 1*, Jyväskylän yliopistopaino, 2007.
[7] V. T. Purmonen, *Lineaarinen algebra ja geometria 2*, Jyväskylän yliopistopaino, 2008.
[8] H. Silfverberg, L. Pippola, M.-L. Viilo ja W. Söderström, *Matematiikan taito 5, Analyyttinen geometria*, WSOY, 4. painos, 2000.
[9] R. E. Williamson, R. H. Crowell, and H. F. Trotter, *Calculus of vector functions*, Prentice-Hall, Inc., 2nd edition, 1968.
[10] http://en.wikipedia.org/wiki/Matrix_(mathematics), luettu 31.1.2012.
[11] http://www.gap-system.org/~history/HistTopics/Matrices_and_determinants.html, luettu 31.1.2012.
[12] http://www.jstor.org/stable/2299112, luettu 31.1.2012.
[13] http://www.oph.fi/download/47345_lukion_opetussuunnitelman_perusteet_2003.pdf, luettu 31.1.2012.