Mika Hirvensalo

Insinöörimatematiikka: Lineaarialgebra 2023

Sisällys

Huomioita sisällöstä: Insinöörimatematiikan opintokokonaisuuden tarkoitus on esittää perustiedot valikoiduista matematiikan työkaluista, joita sovelletaan teknillisillä aloilla. Matematiikan Perustiedot on osin lukiomatematiikan kurssien kanssa päällekkäinen ja näin on myös Differentiaali- ja integraalilaskennan laita.

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

Lineaarialgebralla on sovelluksia useilla muilla matematiikan aloilla, tietojenkäsittelyssä, luonnontieteissä, tilastotieteessä ja yhteiskuntatieteissä. Lukuisine sovelluksineen lineaarialgebra lieneekin ryhmäteorian ohella matematiikan monikäyttöisin alue.

Varhaisimmat viitteet nykyisin tunnistettaviin lineaarialgebran käsitteisiin vaikuttavat olevan peräisin 1600-luvulta, mutta nykyään käytössä olevaan lineaarialgebran käsitteistöä on varsinaisesti kehitetty vasta 1800-luvun puolivälistä alkaen. Lineaarialgebra on täten differentiaali- ja integraalilaskentaa jonkin verran nuorempi matematiikan ala. Erityisen käyttökelpoista molempien alojen kannalta on ollut 1800-luvun lopulla alkanut kehitystyö, jonka myötä on voitu löytää näennäisesti erilaisia matematiikan yhdistäviä rakenteita. Erityisesti usean muuttujien funktiot muodostavat matematiikan osa-alueen, joka rakentuu sekä differentiaali- ja integraalilaskennan että lineaarialgebran pohjalle.

Erilaisten matemaattisten rakenteiden yhdistäminen ei ole edistänyt pelkästään matematiikan kehitystä, vaan on edellytys esimerkiksi kvanttimekaniikan nykyiselle esitystavalle. Sovellusten kannalta kvanttimekaniikan käyttökelpoisin, ns. Hilbertin avaruuksille perustuva formaalinen esitystapa hyödyntää sekä lineaarialgebraa että differentiaali- ja integraalilaskentaa.

Monet tilastotieteelliset menetelmät, esimerkkinä vaikkapa pearsonin korrelaatiokerroin ja regressioanalyysi perustuvat suoraan lineaarialgebran käsitteille. Tekoälyn pohjana toimivista koneoppimisen menetelmistä erityisesti ns. tukivektorikone perustuu niin ikään lineaarialgebraan. Käytännössä voi olla haastavaa löytää mitään vähänkään korkeampaan teknologiaan liittyvää tuotetta, jonka kehityksen yhdessäkään vaiheessa ei ole tarvittu ollenkaan lineaarialgebran käsitteitä.

Luku 1
Vektoriavaruus

1.1 Vektorit

Lineaarialgebran keskeinen käsite on *vektori*, jonka 2 ja 3-ulotteiset (reaaliset) versiot ja erityisesti niiden visuaaliset esitykset suuntajanoina lienevät jokaiselle tuttuja jo lukiokursseista. Yleensä visuaalisessa esityksessä samaistetaan samanpituiset- ja suuntaiset suuntajanat, jolloin matematiikan terminologiassa puhutaan suuntajanojen *ekvivalenssiluokasta*.

Suuntajanoille **a** ja **b** määritellään tyypillisesti yhteenlasku siirtämällä **b** alkamaan suuntajanan **a** alkupisteestä ja muodostamalla uusi suuntajana **c** **a**:n alkupisteestä **b**:n loppupisteeseen.

Tällöin tulee samalla määriteltyä suuntajanan kertominen luonnollisella luvulla: $1\mathbf{a} = \mathbf{a}$, $2\mathbf{a} = \mathbf{a}+\mathbf{a}$, $3\mathbf{a} = \mathbf{a}+2\mathbf{a}$, jne. Visuaalisen esityksen perustella $n\mathbf{a}$ on samansuuntainen kuin n, mutta pituudeltaan n-kertainen. Tätä ajatusta laajentamalla saadaan aikaan skalaarikertolasku, jossa $c\mathbf{a}$ tarkoittaa **a**:n suuntaista janaa, jonka pituus on c kertaa **a**:n pituus kun $c > 0$. Jos taas $c < 0$, tarkoittaa $c\mathbf{a}$ suuntajanaa, jonka pituus on $|c|$ kertaa **a**:n pituus, mutta suunta **a**:han nähden vastakkainen. Otetaan käsitteistöön mukaan vielä nollamittainen **0**, jolla ei ole suuntaa.

Näin saadaan algebrallinen järjestelmä, joka toteuttaa seuraavat ehdot:

- $(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$
- $(\exists\mathbf{0})(\forall\mathbf{a})(\mathbf{a}+\mathbf{0} = \mathbf{a})$
- $(\forall\mathbf{a})(\exists-\mathbf{a})(\mathbf{a}+(-\mathbf{a}) = \mathbf{0})$
- $\mathbf{a}+\mathbf{b} = \mathbf{b}+\mathbf{a}$.

Ylläolevat ehdot toteuttavaa algebrallista järjestelmää sanotaan *kommutatiiviseksi ryhmäksi* eli *Abelin ryhmäksi*.

Myös seuraavat, skalaarikertolaskua koskevat ehdot toteutuvat:

- $1\cdot\mathbf{a} = \mathbf{a}$
- $c(\mathbf{a}+\mathbf{b}) = c\mathbf{a}+c\mathbf{b}$
- $(c+d)\mathbf{a} = c\mathbf{a}+d\mathbf{a}$
- $c(d\mathbf{a}) = (cd)\mathbf{a}$

Algebrallista systeemiä, joka toteuttaa molemmat ylläolevat ehtokokoelmat, sanotaan *vektoriavaruudeksi*. Näitä ehtokokoelmia yhdessä sanotaan *vektoriavaruuden aksioomiksi*.

Täsmällisempää käsittelyä varten on syytä siirtyä suuntajanoista analyyttisen geometrian mukaiseen kuvaukseen, joka voidaan toteuttaa karteesisessa koordinaatistossa. Koska ylläkuvattu vektorin käsite perustui suuntajanojen luokkaan yksittäisen suuntajanan sijaan, voidaan jokaiselle vektorille valita *edustaja*, jonka alkupiste on karteesisen koordinaatiston origo. Tällöin vektori voidaan esittää käyttämällä ainoastaan päätepistettä.

Kaksiulotteisessa tapauksessa tämä tarkoittaa joukkoa $\mathbb{R}^2 = \{(x,y) \mid x,y \in \mathbb{R}\}$ ja vektoreita tässä tapauksessa edustavat lukuparit (x,y). Niiden yhteenlasku toteutuu suoraviivaisesti: $(x_1,y_1)+$

$(x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ ja skalaarikertolasku saa muodon $c(x, y) = (cx, cy)$. Nollavektorina toimii $(0, 0)$ ja vektorin (x, y) vastavektorina $(-x, -y)$.

On hyvin suoraviivaista todeta, että jos joukossa $\mathbb{R}^2$ määritellään tällä tavalla yhteenlasku ja skalaarikertolasku, saadaan algebrallinen järjestelmä, joka toteuttaa ylläolevat vektoriavaruuden aksioomat. Kolmiulotteisessa tapauksessa joukon $\mathbb{R}^2$ sijasta valitaan joukko $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$, mutta vektoreiden yhteenlasku ja skalaarikertolasku määritellään samoin kuin kaksiulotteisessa tapauksessa.

Vaikka visualisointi hankaloituu tai tulee käytännössä mahdottomaksi suurempien ulottuvuuksien kohdalla, ei idealisointi kuitenkaan pysähdy kolmeen ulottuvuuteen. Seuraava abstrahointi voidaankin tehdä pelkästään matemaattisiin käsitteisiin tukeutuen. Tässä yleistyksessä reaalilukujen kunta korvataan millä tahansa joukolla A.

Määritelmä 1. Joukon A karteesinen potenssi A^n tarkoittaa n-pituisia järjestettyjä jonoja, joiden kukin alkio on joukon A jäsen:

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

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

Seuraava lause on helppo osoittaa oikeaksi:

Lause 1. *Olkoon $\mathbb{K}$ jokin kunta. Määritellään joukossa $\mathbb{K}^n$ yhteenlasku ehdolla*

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$$

ja skalaarikertolasku ehdolla $c(a_1, a_2, \dots, a_n) = (ca_1, ca_2, \dots, ca_n)$. Nollavektoriksi valitaan $(0, 0, \dots, 0)$ ja vektorin $(a_1, a_2, \dots, a_n)$ vastavektoriksi $-(a_1, a_2, \dots, a_n) = (-a_1, -a_2, \dots, -a_n)$. Näin saatu algebrallinen järjestelmä toteuttaa vektoriavaruuden aksioomat.

Hyvin monissa tekniikan alan matematiikkaan liittyvissä sovelluksissa kunta $\mathbb{K}$ valitaan joko reaalilukujen kunnaksi $\mathbb{R}$ tai kompleksilukujen kunnaksi $\mathbb{C}$. Toisaalta taas tiedonsiirrossa käytettävissä virheitä korjaavissa koodeissa voi $\mathbb{K}$ olla jokin äärellinen kunta.

Joka tapauksessa ylläolevassa lauseessa määritelty vektoriavaruus muodostaa ns. prototyypin äärellisulotteisille vektoriavaruuksille, mutta jo yli sata vuotta sitten on osoittautunut, että on olemassa muunlaisiakin vektoriavaruuksia, jotka ovat käyttökelpoisia mallinnettaessa esimerkiksi fysikaalisia systemejä tai signaalinkäsittelyä.

Esimerkki 1. Olkoon $C^0[a, b]$ välillä $[a, b]$ määriteltyjen jatkuvien reaalifunktioden joukko. Määritellään funktioden yhteenlasku $f + g$ ehdolla $(f + g)(x) = f(x) + g(x)$ ja skalaarikertolasku $c \cdot f$ ehdolla $(cf)(x) = cf(x)$. On suoraviivaista todeta, että $C^0[a, b]$ toteuttaa vektoriavaruuden aksioomat.

Huomautus 1. Edellisen esimerkin $C^0[a, b]$ ja vektoriavaruuden $\mathbb{R}^n$ välillä on itse asiassa varsin suoraviivainen yhteys. Joka ikinen avaruuden $\mathbb{R}^n$ alkio on järjestetty n-jono $\mathbf{f} = (f_1, \dots, f_n)$, joka voidaan merkinnän $\mathbf{f} = (f(1), \dots, f(n))$ kautta itse asiassa ajatella funktioksi $\mathbf{f} : \{1, \dots, n\} \mapsto \mathbb{R}$, missä $\mathbf{f}(i) = f_i$. Näin ollen vektoriavaruuden $\mathbb{R}^n$ alkio on äärellisessä joukossa $\{1, 2, \dots, n\}$ määritelty funktio joukkoon $\mathbb{R}$, kun taas avaruuden $C^0[a, b]$ alkiot ovat välillä $[a, b]$ määriteltyjä funktioita joukkoon $\mathbb{R}$. Siinä missä avaruuden $\mathbb{R}^n$ vektorilla $\mathbf{f}$ on n koordinaattia f_1, …, f_n, voidaan avaruuden $C^0[a, b]$ vektorilla ajatella olevan äärettömän monta koordinaattia $f(t)$ kutakin arvoa $t \in [a, b]$ kohti.

Huomautus 2. Edellisen huomautuksen erikoistapauksena ovat mm. polynomifunktiot $f(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n$. Kaikkien polynomifunktioiden joukko muodostaa siis vektoriavaruuden, mutta polynomien yhteydessä kytkös lineaarialgebran käsitteisiin on yleensä erilainen, nimittäin polynomin kertoimet voidaan ajatella koordinaatteina. Jos rajoittaudutaan polynomeihin, joiden aste on korkeintaan jokin kiinteä luku N, saadaan $N + 1$-ulotteinen vektoriavaruus: Polynomi $f(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_N x^N$ voidaan ymmärtää $N + 1$-ulotteisen avaruuden vektorina $(c_0, c_1, \dots, c_N)$.

Huomautus 3. Vektoreita on tapana merkitä lihavoidulla kirjasintyypillä: $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ ja alkioita a_1, a_2, …, a_n kutsutaan vektorin *koordinaateiksi* tai pelkästään alkioiksi, toisinaan myös komponenteiksi. Fysiikan kirjallisuudessa vektoreista käytetään usein myös yläviivaa $\bar{a}$, "hattua" $\hat{a}$ (varsinkin ns. yksikkövektoreista puhuttaessa) tai ylänuolta $\overrightarrow{a}$. Vanhemmassa suomenkielisessä kirjallisuudessa on suosittu myös saksalaisperäistä käytäntöä, jossa vektoreita merkitään fraktuurakirjaimilla $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ jne. Liitutaululle (ja piirtopöydälle) kirjoitettaessa on tapana merkitä vektoreita alaviivalla: $\underline{a}$.

1.2 Vektoriavaruuden algebrallinen rakenne

Esimerkki 2. Avaruudessa $\mathbb{R}^2$ merkitään yleensä $\mathbf{i} = (1,0)$, $\mathbf{j} = (0,1)$ ja avaruudessa $\mathbb{R}^3$ puolestaan merkitään $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$ ja $\mathbf{k} = (0,0,1)$. Visuaalisessa esityksessä vektorit $\mathbf{i}$, $\mathbf{j}$ ja $\mathbf{k}$ vastaavat x-, y- ja z-akselin suuntaisia yksikkövektoreita.

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

Edellisen esimerkin mukaisesti kaikki avaruuden $\mathbb{R}^2$ alkiot voidaan esittää valittujen vektoreiden $\mathbf{i} = (1,0)$ ja $\mathbf{j} = (0,1)$ yhdistelminä. Tällaista yhdistelmää kutsutaan *lineaarikombinaatioksi*.

Määritelmä 2. Vektoreiden $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$ *lineaarikombinaatio* on muotoa

$$c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n$$

oleva lauseke, missä $c_1, \ldots c_n$ ovat skalaarikunnan $\mathbb{K}$ alkioita.

Kaikkien vektorijoukosta $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ muodostettujen lineaarikombinaatioiden joukosta käytetään merkintää

$$L(\mathbf{x}_1, \ldots, \mathbf{x}_n) \quad \text{tai} \quad \langle \mathbf{x}_1, \ldots, \mathbf{x}_n \rangle$$

Lause 2. *Olkoot $\mathbf{x}_1, \ldots, \mathbf{x}_n$ vektoriavaruuden V alkioita ja $U = L(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Jos $\mathbf{x}$, $\mathbf{y} \in U$, niin myös $c\mathbf{x} \in U$, ja $\mathbf{x} + \mathbf{y} \in U$. Toisin sanoen, U on suljettu skalaarikertolaskun ja vektoriyhteenlaskun suhteen.*

Huomautus 4. Edellisen lauseen tilanteessa on selvästi $U \subseteq V$. Lisäksi voidaan todeta, että $0 \cdot \mathbf{x} = \mathbf{0} \in U$, $-1 \cdot \mathbf{x} \in U$, jne., siis myös U toteuttaa vektoriavaruuden ehdot. Näin ollen U on vektoriavaruus toisen vektoriavaruuden V sisällä, niin kutsuttu *aliavaruus*.

Määritelmä 3. Jos U ja V ovat sellaisia vektoriavaruuksia että $U \subseteq V$, ja avaruuden U skalaarikertolasku ja vektoriyhteenlasku on sama kuin V:n vastaavat operaatiot, sanotaan että U on V:n aliavaruus. Tällöin merkitään myös $U \leq V$.

Aliavaruussuhde $U \leq V$ merkitsee erityisesti sitä, että U on suljettu skalaarikertolaskun ja vektoreiden yhteenlaskun suhteen. Jokaisella vektoriavaruudella V on olemassa ns. triviaalit aliavaruudet V itse ja pelkän nollavektorin muodostama $\{\mathbf{0}\}$.

Esimerkki 3. Valitaan vektoriavaruudesta $\mathbb{R}^3$ vektorit $\mathbf{i} = (1,0,0)$ ja $\mathbf{j} = (0,1,0)$. Tällöin

$$L(\mathbf{i},\mathbf{j}) = \{x(1,0,0) + y(0,1,0) \mid x,y \in \mathbb{R}\} = \{(x,y,0) \mid x,y \in \mathbb{R}\}$$

Tämä on kolmiulotteisen avaruuden $\mathbb{R}^3$ kaksiulotteinen aliavaruus, jonka visuaalinen esitys on xy-taso.

Esimerkki 4. Olkoon $\mathscr{P}$ kaikkien reaalikertoimisten polynomien joukko:

$$\mathscr{P} = \{c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n \mid n \in \mathbb{N}, c_i \in \mathbb{R}\}.$$

Kun polynomien yhteenlasku ja kertominen vakiolla määritellään tavalliseen tapaan, muodostaa $\mathscr{P}$ vektoriavaruuden. Jos tarkastellaan esimerkiksi korkeintaan kolmatta astetta olevien polynomien joukkoa

$$\mathscr{P}_3 = \{c_0 + c_1 x + c_2 x^2 + c_3 x^3 \mid c_i \in \mathbb{R}\}$$

huomataan että tämä on avaruuden $\mathscr{P}$ aliavaruus.

Määritelmä 4. Joukko $B = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ *generoi* vektoriavaruuden V, jos $V = L(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.

Esimerkki 5. Joukko $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ generoi vektoriavaruuden $\mathbb{R}^3$, koska

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

Esimerkki 6. Joukko $\{\mathbf{i}, \mathbf{k}\}$ generoi vektoriavaruuden

$$x\mathbf{i} + z\mathbf{k} = \{x(1,0,0) + z(0,0,1) \mid x, z \in \mathbb{R}\} = \{(x, 0, z) \mid x, z \in \mathbb{R}\}.$$

Visualisaatiossa tämä vastaa xz-tasoa kolmiulotteisessa avaruudessa.

Esimerkki 7. Valitaan $\mathbf{a} = (1, 2) \in \mathbb{R}^2$, jolloin

$$L(\mathbf{a}) = \{t(1, 2) \mid t \in \mathbb{R}\}.$$

Joukon $L(\mathbf{a})$ alkiot ovat siis muotoa $(t, 2t)$. Jos merkitään $x = t$ ja $y = 2t$, saadaan $y = 2x$. Geometrinen tulkinta joukosta $L(\mathbf{a})$ on siis suora, joka kulkee origon ja pisteen $(1, 2)$ kautta.

Esimerkki 8. Polynomijoukko $\{1, x, x^2, x^3\}$ generoi aiemmassa esimerkissä esiintyneen vektoriavaruuden $\mathscr{P}_3$, koska jokainen korkeintaan kolmatta astetta oleva polynomi voidaan esittää muodossa

$$c_0 \cdot 1 + c_1 \cdot x + c_2 \cdot x^2 + c_3 \cdot x^3 \in L(1, x, x^2, x^3)$$

Esimerkki 9. Avaruudella $\mathbb{R}^n$ (samoin kuin avaruudella $\mathbb{C}^n$) on generoiva joukko

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, \ldots, 1),$$

koska jokainen $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ voidaan esittää muodossa

$$\begin{aligned}
&(x_1, x_2, \ldots, x_n) \\
&= (x_1, 0, \ldots, 0) + (0, x_2, \ldots, 0) + \ldots + (0, 0, \ldots, x_n) \\
&= x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + \ldots + x_n(0, 0, \ldots, 1) \\
&= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n.
\end{aligned}$$

Esimerkki 10. Joukko $\{\mathbf{i}, \mathbf{i} + \mathbf{j}\}$ generoi avaruuden $\mathbb{R}^2$, sillä

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

Esimerkki 11. Myös joukko $\{\mathbf{i}, \mathbf{j}, \mathbf{i} + \mathbf{j}\}$ generoi avaruuden $\mathbb{R}^2$, sillä

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

Esimerkki 12. Vektoriavaruutta $\mathbb{R}^2$ on luontevaa sanoa kaksiulotteiseksi ja avaruutta $\mathbb{R}^3$ kolmiulotteiseksi. On helppo huomata että $\mathbb{R}^2 = L(\mathbf{i}, \mathbf{j})$ ja $\mathbb{R}^3 = L(\mathbf{i}, \mathbf{j}, \mathbf{k})$. Esimerkin 7 mukaista avaruutta on luontevaa kutsua yksiulotteiseksi, joten ulottuvuusluku näyttää liittyvän generoivien vektoreiden määrään. Aivan suoraviivainen ei kytkös ole, sillä esimerkiksi jokainen $\mathbb{R}^2$:n vektori voidaan esittää paitsi muodossa

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

myös Esimerkin 11 muodossa

$$(x,y) = x\mathbf{i} + y\mathbf{j} + 0 \cdot (\mathbf{i} + \mathbf{j}) \tag{1.1}$$

ja myös esitys

$$(x,y) = 0 \cdot \mathbf{i} + (y-x)\mathbf{j} + x(\mathbf{i} + \mathbf{j}) \tag{1.2}$$

pitää paikkansa (totea laskemalla).

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

$$(x,y) = \frac{1}{2}(x+y)(\mathbf{i}+\mathbf{j}) + \frac{1}{2}(x-y)(\mathbf{i}-\mathbf{j}) \tag{1.3}$$

(totea laskemalla).

Täten siis avaruudella $\mathbb{R}^2$ on ainakin kaksi kahden alkion ja yksi kolmen alkion generoiva joukko. On kuitenkin huomattava että kolmen alkion generoivasta joukosta $\{\mathbf{i},\mathbf{j},\mathbf{i}+\mathbf{j}\}$ esimerkiksi viimeinen alkio $\mathbf{i}+\mathbf{j}$ jättää pois yhtälön (1.1) perusteella. Toisin sanoen, aina kun on tarve lisätä vektori $\mathbf{i}+\mathbf{j}$ johonkin lineaarikombinaatioon, voidaan yhtä hyvin lisätä $\mathbf{i}$ ja $\mathbf{j}$ siihen erikseen. Toisaalta yhtä hyvin yhtälön (1.2) perusteella voitaisiin jättää pois ensimmäinen alkio $\mathbf{i}$.

Generoivassa joukossa voi siis olla turhia" vektoreita siinä mielessä, että jokin osajoukko generoi saman vektoriavaruuden. Esimerkiksi Yhtälön (1.1) perusteella $L(\mathbf{i},\mathbf{j},\mathbf{i}+\mathbf{j}) = L(\mathbf{i},\mathbf{j})$. Toisaalta ei ole yksiselitteistä tapaa määrittää mikä vektoreista on ”turha”, sillä yhtälön (1.2) perusteella $L(\mathbf{i},\mathbf{j},\mathbf{i}+\mathbf{j}) = L(\mathbf{j},\mathbf{i}+\mathbf{j})$.

Generoinnin kannalta ”turhia” vektoreita pyritään lähestymään seuraavan määritelmän kautta.

Määritelmä 5. Vektorijoukko $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_k\}$ on *lineaarisesti riippuva* tai pelkästään *riippuva*, jos jokin sen vektoreista voidaan esittää muiden lineaarikombinaationa: $\mathbf{v}_i \in L(\mathbf{v}_1,\ldots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\ldots,\mathbf{v}_k)$. Jos vektorijoukko ei ole lineaarisesti riippuva, sanotaan että se on lineaarisesti riippumaton (tai pelkästään riippumaton).

Esimerkki 13. Joukko $\{\mathbf{i},\mathbf{j}\}$ on selvästi lineaarisesti riippumaton, sillä $(1,0) = c(0,1)$ tai $(0,1) = c(1,0)$ eivät toteudu millään c:n arvolla. Joukko $\{\mathbf{i},\mathbf{j},\mathbf{i}+\mathbf{j}\}$ on lineaarisesti riippuva, sillä $\mathbf{i}+\mathbf{j} = 1\cdot\mathbf{i} + 1\cdot\mathbf{j}$.

Huomautus 5. Jokainen vektorijoukko $A = \{\mathbf{0},\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ joka sisältää nollavektorin $\mathbf{v}_1 = \mathbf{0}$, on lineaarisesti riippuva, sillä $\mathbf{0}$ voidaan esittää muiden lineaarikombinaationa:

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

Lineaariselle riippumattomuudelle voidaan esittää seuraava symmetrinen kriteeri.

Lause 3. *Vektorijoukko $\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ on lineaarisesti riippumaton tarkalleen silloin kun*

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

toteutuu vain ilmeisellä (=triviaalilla) valinnalla jossa kaikki kertoimet ovat nollia: $c_1 = \ldots = c_k = 0$.

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

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

josta nähdään, että

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

Näin ollen nollavektori voidaan esittää epäilmeisellä tavalla, sillä vektorin $\mathbf{v}_i$ kerroin ylläolevassa esityksessä on $-1 \neq 0$.

Oletetaan sitten, että nollavektori voidaan esittää epäilmeisellä tavalla

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

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

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

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

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

Esimerkki 14. Vektorijoukko $\{\mathbf{i},\mathbf{j},\mathbf{i}+\mathbf{j}\}$ on lineaarisesti riippuva, sillä

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

on nollavektorin ei-ilmeinen esitys.

Esimerkki 15. Vektorijoukko $\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$ avaruudessa $\mathbb{R}^n$ on lineaarisesti riippumaton, sillä

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

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

Esimerkki 16. Selvitetään onko avaruuden $\mathbb{R}^2$ vektorijoukko $\{\mathbf{i}+\mathbf{j},\mathbf{i}-\mathbf{j}\} = \{(1,1),(1,-1)\}$ lineaarisesti riippumaton. Yhtälö

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

voidaan kirjoittaa yhtälöpariksi

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

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

Esimerkki 17. Selvitetään, onko avaruuden $\mathbb{R}^3$ vektorijoukko $\{(1,2,-1),(-2,3,1),(4,1,-3)\}$ lineaarisesti riippumaton. Tätä varten on selvitettävä, onko yhtälöllä

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

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

Yllä oleva vektoriyhtälö voidaan kirjoittaa lukuja koskevaksi yhtälöryhmäksi

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

Seuraavassa luvussa nähdään miten ylläoleva yhtälöryhmä voidaan ns. *Gaussin-Jordanin prosessilla* saattaa ekvivalenttiin muotoon

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

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

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

jolloin siis vektorijoukko on lineaarisesti riippuva. Yllä olevan yhtälön perusteella voidaan mikä hyvänsä vasemman puolen vektoreista esittää muiden lineaarikombinaationa.

Huomautus 6. Yhden vektorin joukko on lineaarisesti riippuva tarkalleen silloin kun kyseinen vektori on $\mathbf{0}$. Kahden vektorin joukko $\{\mathbf{v}_1, \mathbf{v}_2\}$ on lineaarisesti riippuva tarkalleen silloin kun toinen vektoreista on nollavektori tai vektorit ovat toistensa skalaarimonikertoja.

Esimerkki 18. Jokainen vektorijoukko $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, jossa vektorissa $\mathbf{v}_{i+1}$ on enemmän alkunollia kuin vektorissa $\mathbf{v}_i$, mutta joka ei sisällä nollavektoria, on lineaarisesti riippumaton. Jos nimittäin $\mathbf{v}_1 = (0, \ldots, v_1, \ldots)$ missä $v_1 \neq 0$, on

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = (0, \ldots, c_1v_1, \ldots)$$

ja jotta tämä olisi nollavektori, pitää ensiksikin olla $c_1v_1 = 0$, josta $c_1 = 0$. Tällöin siis summan

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

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

Esimerkki 19. Todetaan edellisen esimerkin mukaisesti, että avaruuden $\mathbb{R}^4$ joukko $\{(1,2,0,-1), (0,2,-1,-2), (0,0,3,2), (0,0,0,-1)\}$ on lineaarisesti riippumaton:

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

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

Esimerkki 20. Olkoon C^0 koko reaaliakselilla määriteltyjen funktioiden joukko varustettuna esimerkin 1 mukaisella yhteen- ja skalaarikertolaskulla. Tällöin funktiot 1, $\sin^2 x$ ja $\cos^2 x$ ovat lineaarisesti riippuvat, sillä

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

Esimerkki 21. Olkoon C^0 kuten edellisessä esimerkissä. Funktiot 1, $\sin x$ ja $\cos x$ ovat lineaarisesti riippumattomat, sillä jos

$$c_1 \cdot 1 + c_2 \cdot \sin x + c_3 \cdot \cos x = 0 \tag{1.4}$$

kaikilla x:n arvoilla, saadaan derivoimalla että

$$c_2 \cos x - c_3 \sin x = 0$$

kaikilla x:n arvoilla. Sijoittamalla tähän $x = 0$ nähdään, että $c_2 = 0$ ja sijoittamalla $x = \frac{\pi}{2}$ nähdään että $c_3 = 0$. Yhtälöstä (1.4) seuraa tällöin $c_1 = 0$.

Vektoriavaruuden *kanta* on yksi lineaarialgebran peruskäsitteistä. Sanallisesti luonnehdittuna kanta on vektoriavaruuden minimaalinen generoiva joukko.

Määritelmä 6. Joukko B on vektoriavaruuden V *kanta*, jos

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

Lause 4. *Vektoriavaruuden kaikissa kannoissa on yhtä monta vektoria.*

Todistus. Sivuutetaan.

Lause 5. *Jos $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ on avaruuden V kanta, niin jokaisella $\mathbf{x} \in V$ on olemassa yksikäsitteinen esitys kantavektoreiden lineaarikombinaationa:*

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

Todistus. Esitys on olemassa sen perusteella, että B generoi vektoriavaruuden V ja yksikäsitteisyys seuraa siitä, että kanta on lineaarisesti riippumaton: Jos nimittäin olisi kaksi erilaista esitystä

$$\mathbf{x} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n = x_1'\mathbf{b}_1 + \ldots + x_n'\mathbf{b}_n,$$

missä ainakin jokin $x_i \neq x_i'$, niin tällöin

$$(x_1 - x_1')\mathbf{b}_1 + \ldots + (x_n - x_n')\mathbf{b}_n = \mathbf{0}$$

olisi nollavektorin ei-ilmeinen esitys koska $x_i - x_i' \neq 0$. Tämä on vastoin kannan lineaarista riippumattomuutta.

Määritelmä 7. Jos $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ on äärellinen kanta, sanotaan että $V = L(B)$ on k-ulotteinen ja merkitään $\dim(V) = k$. Avaruuden $V = \{\mathbf{0}\}$ kannaksi sovitaan tyhjä joukko. Jos vektoriavaruudella V ei ole äärellistä kantaa, sanotaan, että V on ääretönulotteinen.

Esimerkki 22. Vektorit $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$, $\mathbf{k} = (0,0,1)$ muodostavat $\mathbb{R}^3$:n kannan. Samoin vektorijoukko $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, missä $\mathbf{e}_i = (0, \ldots, 1, \ldots, 0)$ (i:s koordinaatti 1) on $\mathbb{R}^n$:n kanta. Tätä kantaa kutsutaan vektoriavaruuden $\mathbb{R}^n$ luonnolliseksi kannaksi. Näin ollen $\mathbb{R}^n$ on n-ulotteinen vektoriavaruus.

Vektoriavaruudella voi olla monia kantoja, kuten esimerksiksi seuraavassa esimerkissä nähdään.

Esimerkki 23. Joukot $B_1 = \{\mathbf{i}, \mathbf{j}\}$ ja $B_2 = \{\mathbf{i}+\mathbf{j}, \mathbf{i}-\mathbf{j}\}$ ovat molemmat avaruuden $\mathbb{R}^2$ kantoja. Joukon B_1 generointi ja lineaarinen riippuvuus on selvä, ja joukolle B_2 nämä on todettu aiemmin esimerkissä 12).

Esimerkki 24. Joukko $\{\mathbf{i}, \mathbf{j}, \mathbf{i}+\mathbf{j}\}$ ei ole avaruuden $\mathbb{R}^2$ kanta. Se tosin generoi avaruuden $\mathbb{R}^2$, mutta ylläolevien esimerkkien perusteella se ei ole lineaarisesti riippumaton.

Määritelmä 8. Jos $\mathbf{x} \in V$ ja $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ on avaruuden V kanta, niin esityksessä

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

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

Esimerkki 25. Jos $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ja $B = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ on luonnollinen kanta, niin

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

joten

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

Tässä tapauksessa siis $\mathbf{x}_B = \mathbf{x}$ ja tämä on syy miksi kantaa B kutsutaan avaruuden $\mathbb{R}^n$ luonnolliseksi kannaksi.

Esimerkki 26. Olkoon $B_1 = \{\mathbf{i}, \mathbf{j}\}$ avaruuden $\mathbb{R}_2$ luonnollinen kanta ja $B_2 = \{\mathbf{i}+\mathbf{j}, \mathbf{i}-\mathbf{j}\}$. Tällöin $(x,y)_{B_1} = (x,y)$ ja yhtälön (1.3) perusteella $(x,y)_{B_2} = (\frac{1}{2}(x+y), \frac{1}{2}(x-y))$.

Esimerkki 27. Olkoon $\mathscr{P}_5 = \{c_0 + c_1x + c_2x^2 + c_3x^3 + c_4x^4 \mid c_i \in \mathbb{R}\}$ niiden reaalikertoimisten polynomien joukko, joiden aste on pienempi kuin 5. P_5 on vektoriavaruus (tarkista), jonka kannaksi voidaan valita $\{1, x, x^2, x^3, x^4\}$. Näin ollen $\dim(P_5) = 5$.

Esimerkki 28. Myös joukko $\{1,\ 1+x,\ 1+x+x^2,\ 1+x+x^2+x^3,\ 1+x+x^2+x^3+x^4\}$ on avaruuden P_5 kanta.

Huomautus 7. Jos $V = L(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, mutta $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ ei ole avaruuden V kanta, voidaan generoivasta joukosta poistaa "turhat" vektorit näin saadaan kanta. Mikäli $\mathbf{v}_k \in L(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$, on $V = L(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$ ja prosessi voidaan toistaa kunnes jäljelle jäävä joukko on lineaarisesti riippumaton.

Lause 6. *Jos V on äärellisulotteinen ja U on V:n aliavaruus, niin mikä hyvänsä U:n kanta voidaan täydentää V:n kannaksi.*

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

Seuraus 1. Jos $U \leq V$, on $\dim(U) \leq \dim(V)$.

Seuraus 2. Olkoon $\dim(V) = n$. Jokainen V:n osajoukko, jossa on enemmän kuin n vektoria, on lineaarisesti riippuva.

Vektoriavaruuden V jokainen aito aliavaruus $U \lneq V$ hajottaa vektoriavaruuden kahteen osaan, jotka muodostavat koko avaruuden ns. *suorana summana*.

Määritelmä 9 (Suora summa). Olkoot U_1 ja U_2 vektoriavaruuden V aliavaruuksia. Jos jokainen $\mathbf{v} \in V$ voidaan esittää yksikäsitteisesti muodossa $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$, sanotaan että V on aliavaruuksien U_1 ja U_2 suora summa ja merkitään $V = U_1 \oplus U_2$.

Esimerkki 29. Vektoriavaruus $\mathbb{R}^2$ on yksiulotteisten aliavaruuksiensa $U_1 = \langle (1,0) \rangle$ ja $U_2 = \langle (0,1) \rangle$, suora summa, koska jokainen $(x,y) \in \mathbb{R}^2$ voidaan yksikäsitteisesti esittää muodossa

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

jossa ensimmäinen summattava kuuluu aliavaruuteen U_1 ja toinen aliavaruuteen U_2.

Esimerkki 30. Vektoriavaruus $\mathbb{R}^3$ on yksi- ja kaksiulotteisten aliavaruuksiensa $U_1 = \langle (1,0,0) \rangle$ ja $U_2 = \langle (0,1,0), (0,0,1) \rangle$, koska jokainen $(x,y) \in \mathbb{R}^2$ voidaan yksikäsitteisesti esittää muodossa

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

jossa $x(1,0,0) \in U_1$ ja $y(0,1,0) + z(0,0,1) \in U_2$

Lause 7. *Jos V on äärellisulotteinen ja $U_1 \lneq V$, on olemassa sellainen aliavaruus $U_2 \leq V$, että $V = U_1 \oplus U_2$. Lisäksi* $\dim(V) = \dim(U_1) + \dim(U_2)$.

Todistus. Lauseen 6 mukaisesti aliavaruuden U_1 kanta $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ voidaan täydentää koko vektoriavaruuden V kannaksi $\{\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$. Kun määritellään $U_2 = \langle \mathbf{b}_{k+1}, \ldots, \mathbf{b}_k \rangle$, voidaan helposti todeta että $V = U_1 \oplus U_2$ ja että $\dim(V) = n = k + (n-k) = \dim(U_1) + \dim(U_2)$.

Luku 2
Gaussin-Jordanin menetelmä

Äärellinen lineaarinen yhtälöryhmä voidaan aina ratkaista Gaussin-Jordanin menetelmällä. Siksi tämä menetelmä on avain moniin lineaarialgebran ongelmiin ja siihen on syytä perehtyä perusteellisesti.

Määritelmä 10. Yhtälöryhmä on lineaarinen, jos siinä on vain ensimmäisen ja nollannen asteen termejä (vakiotermejä):

$$\begin{cases} A_{11}x_1 & + A_{12}x_2 & + \ldots + & A_{1n}x_n & = b_1 \\ A_{21}x_1 & + A_{22}x_2 & + \ldots + & A_{2n}x_n & = b_2 \\ & \vdots & \ddots & \vdots & \\ A_{m1}x_1 & + A_{m2}x_2 & + \ldots + & A_{mn}x_n & = b_m \end{cases} . \tag{2.1}$$

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

Lineaarinen yhtälöryhmä on *homogeeninen*, jos oikean puolen vakiot ovat kaikki nollia: $b_1 = b_2 = \ldots = b_m = 0$.

Lineaaristen yhtälöryhmien ratkaisuja voidaan löytää lukiokurssista tutulla *muuttujien eliminointimenetelmällä*, jossa yhtälö kerrotaan jollakin luvulla ja lisätään toiseen muuttujien määrän vähentämiseksi. Vaihtoehtoisesti voidaan jokin muuttuja esittää muiden muuttujien avulla ja sijoittaa ryhmään. Kumpikin tapa toimii, mikäli alkuperäisellä yhtälöryhmällä on yksikäsitteinen ratkaisu, mutta ratkaisujoukon ollessa ääretön kumpikaan ei välttämättä johda selkeästi päättyvään prosessiin. Tässä luvussa esiteltävä Gaussin-Jordanin eliminointimenetelmä ei tuo mitään uutta käytettäviin operaatioihin nähden, vaan sen merkittävin piirre on selkeästi hallitava, arvioitava ja päättyvä prosessi.

2.1 Alkeisoperaatiot

Menetelmä lineaarisen yhtälöryhmän ratkaisemiseksi on hyvin yksinkertainen: Ryhmään sovelletaan seuraavia alkeisoperaatiota, kunnes ryhmä saadaan niin yksinkertaiseen muotoon, että ratkaisu voidaan lukea suoraan.

Määritelmä 11. Lineaaristen yhtälöryhmien *alkeisoperaatiot* ovat

1. Kahden yhtälön järjestyksen vaihtaminen,
2. Jonkin yhtälön kertominen nollasta eroavalla vakiolla, ja
3. Jokin yhtälön lisääminen toiseen toiseen vakiolla kerrottuna.

Esimerkki 31. Lisätään yhtälöparin

$$\begin{cases} 2x_1 & -x_2 & +x_3 & = 7 \\ x_1 & +3x_2 & +4x_3 & = 0 \end{cases}$$

toinen yhtälö ensimmäiseen -2:lla kerrottuna, jolloin saadaan

$$\begin{cases} & -7x_2 & -7x_3 & = 7 \\ x_1 & +3x_2 & +4x_3 & = 0 \end{cases}.$$

Kerrotaan tämän jälkeen ensimmäinen yhtälö luvulla $-\frac{1}{7}$ ja vaihdetaan yhtälöiden järjestystä. Näin saadaan

$$\begin{cases} x_1 & +3x_2 & +4x_3 & = & 0 \\ & x_2 & +x_3 & = & -1 \end{cases}.$$

Lisätään nyt toinen yhtälö ensimmäiseen luvulla -3 kerrottuna. Tuloksena saadaan yhtälöpari

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

jonka ratkaisut (x_1,x_2,x_3) ovat ilmeiset: ensimmäisen yhtälön perusteella $x_1 = -x_3 + 3$ ja toisen yhtälön nojalla $x_2 = -x_3 - 1$. Täten siis viimeisimmän yhtälöparin ratkaisut ovat muotoa $(x_1,x_2,x_3) = (-x_3+3,-x_3-1,x_3) = x_3(-1,-1,1)+(3,-1,0)$, missä x_3 on mikä tahansa reaaliluku. Ratkaisut ovat siis muotoa $t(-1,-1,1)+(3,-1,0)$, missä $t \in \mathbb{R}$ ja täten ratkaisuja on ääretön määrä.

Nyt kuitenkin puhuttiin *viimeisimmän* yhtälöparin ratkaisuista, mutta helposti voidaan nähdä, että nämä ovat myös alkuperäisen yhtälöparin ratkaisut. Ensinnäkin, alkeisoperaatiot ovat selvästi totuusarvon säilyttäviä siinä mielessä, että jos yhtälöryhmä S' saadaan yhtälöryhmästä S jotakin näistä operaatioista hyödyntämällä, on ryhmällä S' välttämättä *ainakin* ne ratkaisut, jotka ryhmällä S on. Toisin sanoen, jos ryhmä S toteutuu joillakin muuttujien arvoilla, niin myös ryhmä S' toteutuu samoilla arvoilla.

Toisaalta taas alkeisoperaatiot ovat *kääntyviä*, sillä ryhmästä S' voidaan saada jälleen S alkeisoperaatioilla. Jos nimittäin suoritettu alkeisoperaatio oli yhtälöiden järjestyksen vaihto, voidaan samainen vaihto suorittaa uudelleen alkuperäisen yhtälöryhmän palauttamiseksi. Jos taas alkeisoperaatio oli yhtälön kertominen nollasta eroavalla vakiolla, saadaan alkuperäinen ryhmä kertomalla vakion käänteisluvulla. Yhtälön lisääminen toiseen vakiolla kerrottuna voidaan puolestaan kääntää lisäämällä sama yhtälö vakion vastaluvulla kerrottuna.

Kuten aiemmin todettiin, ryhmällä S' on *ainakin* ne ratkaisut, jotka ryhmällä S on, mutta koska ryhmästä S' voidaan alkeisoperaatiota käyttäen saada jälleen ryhmä S, saadaan seuraava lause:

Lause 8. *Jos yhtälöryhmään sovelletaan mitä hyvänsä alkeisoperaatiota, on tuloksena yhtälöryhmä, jolla on tarkalleen samat ratkaisut kuin alkuperäisellä ryhmällä. Tällöin sanotaan, että ryhmät ovat ekvivalentit.*

Lauseen 8 mukaan yhtälöryhmän ratkaisemiseksi voidaan ryhmään soveltaa alkeisoperaatiota, kunnes saadaan niin yksinkertainen ryhmä, että sen ratkaisut nähdään suoraan. Juuri näin tehtiin esimerkissä 31.

Menettelytapaa tarkastelemalla huomataan helposti, että muuttujien ja yhtäsuuruusmerkkien kirjoittaminen joka vaiheeseen ei anna mitään lisäinformaatiota. Koska edellämainitut operaatiot kohdistuvat muuttujien kertoimiin, itse muuttujia ole tarpeen kirjoittaa näkyviin ollenkaan. Näin ollen ylläoleva yhtälöryhmä voidaan pelkistää matriisiksi. Esimerkissä 31 voidaan yhtälöparin sijaan yhtä hyvin siis käyttää ns. *kerroinmatriisia*.

Määritelmä 12. Lineaarisen yhtälöryhmän

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

kerroinmatriisi on $m \times n$-kaavio

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

Matriisin vaakasuoria rivejä kutsutaan *riveiksi* ja pystysuoria rivejä *sarakkeiksi*.

Saman yhtälöryhmän *augmentoitu matriisi* on $m \times (n+1)$-kaavio

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

jossa matriisin viimeiseksi sarakkeeksi on kirjoitettu yhtälöryhmän oikealla puolella esiintyvät vakiot.

Välittömästi huomataan, että esitettäessä yhtälöryhmiä tällä tavalla matriisimuodossa ilman muuttujia alkeisoperaatiot voidaan tulkita *matriisien alkeisoperaatioiksi* jotka siis määritellään seuraavasti:

Määritelmä 13. Matriisien rivien alkeisoperaatioita ovat

1. Kahden rivin järjestyksen vaihtaminen,
2. Rivin kertominen nollasta eroavalla vakiolla, ja
3. Rivin lisääminen toiseen vakiolla kerrottuna.

Esimerkki 32. Esimerkin 31 kerroinmatriisi on $\begin{pmatrix} 2 & -1 & 1 \\ 1 & 3 & 4 \end{pmatrix}$. Lisäämällä tähän oikealla puolella esiintyvät vakiot omaksi sarakkeekseen saadaan ns. *augmentoitu* eli laajennettu kerroinmatriisi $\begin{pmatrix} 2 & -1 & 1 & 7 \\ 1 & 3 & 4 & 0 \end{pmatrix}$.

Huomaa, että esimerkin 31 toisen yhtälöparin (augmentoitu) matriisi on

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

mikä saadaan ensimmäisen yhtälöparin augmentoidusta matriisista lisäämällä toinen rivi ensimmäiseen -2:lla kerrottuna. Vastaavasti esimerkin toinen yhtälöpari saatiin ensimmäisestä lisäämällä toinen yhtälö ensimmäiseen -2:lla kerrottuna. Tämä on varsin ymmärrettävää, koska *augmentoidut matriisit esittävät yhtälöryhmiä täsmällisesti*, ainoastaan muuttujat ja yhtäsuuruusmerkit on jätetty kirjoittamatta. Täten siis yhtälöryhmiä ratkaistaessa voidaan yhtä hyvin tarkastella augmentoituja matriiseja ja suorittaa niille alkeisoperaatioita.

Määritelmä 14. Matriisit A ja B ovat *(rivi)ekvivalentteja* (merkitään $A \sim B$), jos ne saadaan toisistaan alkeisoperaatiota suorittamalla.

Ylläolevasta tarkastelusta seuraa välittömästi seuraava tulos:

Lause 9. *Jos lineaaristen yhtälöryhmien S_1 ja S_2 augmentoidut matriisit A ja B ovat riviekvivalentteja, on ryhmillä S_1 ja S_2 samat ratkaisut.*

Yhtälöryhmien yleinen ratkaisustrategia onkin etsiä ryhmän alkuperäisen (augmentoidun) matriisin A kanssa riviekvivalentti matriisi B, jossa on mahdollisimman paljon nollia. Tällöin ratkaisut on helppo selvittää. Mitä sitten tarkoittaa ”mahdollisimman paljon nollia”? Tätä ryhdytään selvittämään seuraavaksi.

Esimerkki 33. Aloittamalla edellisen yhtälöryhmän augmentoidusta matriisista saadaan vaiheittain alkeisoperaatioilla

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

Viimeisin muoto vastaa yhtälöparia

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

Strategiana on soveltaa alkuperäisen yhtälöryhmän augmentoituun matriisiin (homogeenisessa tapauksessa kerroinmatriisiin) rivioperaatioita, kunnes päästään niin yksinkertaiseen muotoon, että ratkaisu voidaan lukea suoraan. Aiemmin mainittu tavoite ”mahdollisimman paljon nollia” voidaan nyt täsmentää asettamalla päämääräksi *redusoitu porrasmuoto*.

2.2 Porrasmuodot

Määritelmä 15. Matriisi A on *porrasmuodossa*, jos sen jokainen rivi alkaa nollilla, joita on enemmän kuin millään ylemmällä rivillä. Ensimmäisen rivin ei tarvitse alkaa nollalla ja jostain rivistä alkaen rivit voivat koostua kokonaan nollista.

Määritelmä 16. Matriisi A on *redusoidussa porrasmuodossa*, jos

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

Esimerkki 34. Redusoitu porrasmatriisi on siis muotoa

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

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

Gaussin-Jordanin menetelmä

Porrasmuodon saavuttamiseksi voidaan käyttää seuraavaa menetelmää (algoritmia):

1. Jos kaikki rivit ovat nollarivejä, lopeta. Samoin jos rivejä on vain yksi, on A valmiiksi porrasmuotoa ja voidaan lopettaa.
2. Jos A:ssa enemmän kuin yksi rivi, etsi rivi $\mathbf{c} = (0,\dots,c,*,\dots,*)$, jossa on vähiten alkunollia ja vaihda se ylimmäksi. Jos tällaisia rivejä on useampia, valitse niistä mikä hyvänsä. Näin saadaan matriisi
$$B = \begin{pmatrix} 0 & \dots & c & * & \dots & * \\ 0 & \dots & * & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & * & * & \dots & * \end{pmatrix},$$
jossa $c \neq 0$ ja asteriskit ovat mitä hyvänsä lukuja.
3. Jokaista ylimmän rivin alapuolella olevaa riviä $\mathbf{d} = (0,\dots,d,\dots,*)$ kohti toimi seuraavasti (d on siis alkio, joka on ylimmän rivin alkion c alapuolella): lisää ylin rivi tähän riviin kerrottuna vakiolla $-c^{-1}d$. Tällöin d siis korvautuu luvulla $d - c \cdot c^{-1}d = 0$ ja näin saadaan matriisi
$$C = \begin{pmatrix} 0 & \dots & c & * & \dots & * \\ 0 & \dots & 0 & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & * & \dots & * \end{pmatrix} \overset{\text{merk.}}{=} \begin{pmatrix} 0 & \dots & c & * \dots * \\ 0 & \dots & 0 & \\ \vdots & \ddots & \vdots & D \\ 0 & \dots & 0 & \end{pmatrix},$$
jossa c:n alapuolella on vain nollia.
4. Sovella tätä samaa menetelmää matriisiin D.

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

Porrasmatriisista saadaan redusoitu muoto täydentämällä Gaussin menetelmää *Gaussin-Jordanin menetelmäksi* seuraavasti:

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

Gaussin-Jordanin menetelmällä voidaan mikä hyvänsä matriisi, jonka alkiot kuuluvat johonkin kuntaan (yleensä $\mathbb{R}$ tai $\mathbb{C}$) saattaa redusoituun porrasmuotoon. Tämän on yksi lineaarialgebran tärkeimmistä algoritmeista ja on valmiiksi ohjelmoitu useimpii tavallisiin matematiikkaohjelmiin.

Esimerkki 35. Ratkaistaan yhtälöryhmä
$$\begin{cases} 2x & -y & +3z = 4 \\ 3x & +2y & +z = 6 \\ -5x & +y & -2z = 3 \end{cases}.$$

Ryhmän augmentoitu matriisi on $\begin{pmatrix} 2 & -1 & 3 & 4 \\ 3 & 2 & 1 & 6 \\ -5 & 1 & -2 & 3 \end{pmatrix}$. Kertomalla ensimmäinen rivi vakiolla $\frac{1}{2}$ saadaan $\begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 2 \\ 3 & 2 & 1 & 6 \\ -5 & 1 & -2 & 3 \end{pmatrix}$, missä edelleen voidaan lisätä ensimmäinen rivi toiseen -3:lla kerrottuna ja kolmanteen 5:llä kerrottuna. Tuloksena saadaan matriisi $\begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 2 \\ 0 & \frac{7}{2} & -\frac{7}{2} & 0 \\ 0 & -\frac{3}{2} & \frac{11}{2} & 13 \end{pmatrix}$. Kerrotaan näin saadun matriisin toinen rivi $\frac{2}{7}$:llä. Näin saadaan $\begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 2 \\ 0 & 1 & -1 & 0 \\ 0 & -\frac{3}{2} & \frac{11}{2} & 13 \end{pmatrix}$, minkä jälkeen lisätään toinen rivi ensimmäiseen $\frac{1}{2}$:lla kerrottuna ja kolmanteen $\frac{3}{2}$:lla kerrottuna. Tuloksena saatava matriisi on $\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 4 & 13 \end{pmatrix}$, josta kolmas rivi $\frac{1}{4}$:lla kertomalla saadaan $\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & \frac{13}{4} \end{pmatrix}$. Lisäämällä kolmas rivi ensimmäiseen -1:llä kerrottuna ensimmäiseen ja 1:llä kerrottuna toiseen saadaan matriisi $\begin{pmatrix} 1 & 0 & 0 & -\frac{5}{4} \\ 0 & 1 & 0 & \frac{13}{4} \\ 0 & 0 & 1 & \frac{13}{4} \end{pmatrix}$, jota vastaava yhtälöryhmä on

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

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

Esimerkki 36. Yhtälöryhmää

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

vastaava agmentoitu matriisi on $\begin{pmatrix} 1 & 3 & 2 & 7 \\ 2 & 1 & -1 & 5 \\ -1 & 2 & 3 & 4 \end{pmatrix}$, mistä alkeisoperaatioilla saadaan matriisi $\begin{pmatrix} 1 & 3 & 2 & 7 \\ 0 & 1 & 1 & \frac{9}{5} \\ 0 & 0 & 0 & 2 \end{pmatrix}$. Viimeisintä matriisia vastaavan yhtälöryhmän kolmas yhtälö on muotoa $0 \cdot z + 0 \cdot y + 0 \cdot z = 2$, siis $0 = 2$, mikä ei voi toteutua millään muuttujien x, y ja z arvoilla. Koska viimeisimmällä ryhmällä ei ole ratkaisua, ei sellaista voi olla ensimmäiselläkään, sillä ryhmiä vastaavat (augmentoidut) matriisit ovat riviekvivalentteja.

Taustatietoa

(Johann) Carl Friedrich Gauss (Gauß) (latin. *Carolus Fridericus Gauss* 1777–1855) oli saksalainen matemaatikko, jota Arkhimedeen ja Newtonin ohella pidetään yhtenä maailmanhistorian merkittävimpämä matematiikan kehittäjänä. Gaussia pidetään mm. ensimmäisen *algebran peruslauseen* aukottoman todistuksen esittäjänä. Gauss saavutti huomattavia tuloksia suurella osalla nykymatematiikan aloista.

(kuva: Wikimedia Commons)

Syötetään esimerkin 35 augmentoitu matriisi **Matlab**-ohjelmistoon:

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

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

```
>> rref(A)

ans =

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

>>
```

Lause 10. *Redusoidut porrasmatriisit ovat riviekvivalentteja keskenään vain sillon kun ne ovat samat.*

Todistus. Sivuutetaan. Todistus on monivaiheinen, mutta ei erityisen syvällinen.

Edellisestä lauseesta seuraa, että jokainen matriisi A on riviekvivalentti *tarkalleen yhden* redusoidun porrasmatriisin kanssa.

Määritelmä 17. Matriisiin A *aste (rank)* $r(A)$ on sen (redusoidun) porrasmatriisin porrasluku (portaiden määrä), joka saadaan A:sta rivien alkeisoperaatioilla.

Määritelmän perusteella on selvää, että matriisin aste on korkeintaan sen rivien määrä, mutta ei myöskään ole vaikeaa nähdä, että matriisin astetta rajoittaa lisäksi sarakkeiden määrä, toisin sanoen $m \times n$-matriisille pätee $r(A) \leq \min\{m,n\}$.

Erityistä huomiota ansaitsee kuitenkin matriisin rivien tulkitseminen vektoreiksi ja nollarivin ilmaantuminen Gaussin-Jordanin prosessissa.

Lause 11. *Muodostetaan matriisi A valitsemalla sen riveiksi vektorit* $\mathbf{x}_1, \ldots, \mathbf{x}_k$. *Jos Gaussin-Jordanin prosessi tuottaa matriisin A redusoituun porrasmuotoon B nollarivin, on vektorijoukko* $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ *lineaarisesti riippuva. Jos* $\{\mathbf{y}_1, \ldots, \mathbf{y}_l\}$ *ovat nollarivistä poikkeavat rivit redusoidussa porrasmuodossa, ovat ne ensinnäkin lineaarisesti riippumattomat ja toiseksi* $L(\mathbf{x}_1, \ldots, \mathbf{x}_k) = L(\mathbf{y}_1, \ldots, \mathbf{y}_l)$.

Todistus (Idea). Joka tapauksessa Gaussin-Jordanin prosessi muodostaa matriisien riveistä lineaarikombinaatioita, koska siinä esiintyy joko järjestyksen vaihto, vakiolla $\neq 0$ kertominen, tai rivin lisääminen toiseen skalaarilla kerrottuna ($\mathbf{x_i} \to \mathbf{x}_i + c\mathbf{x}_i$). Näin ollen $L(\mathbf{y}_1, \ldots, \mathbf{y}_l) \subseteq L(\mathbf{x}_1, \ldots, \mathbf{x}_k)$. Koska prosessi on kääntyvä (voidaan peruuttaa lopusta alkuun), on myös $L(\mathbf{x}_1, \ldots, \mathbf{x}_k) \subseteq L(\mathbf{y}_1, \ldots, \mathbf{y}_l)$, siksi $L(\mathbf{x}_1, \ldots, \mathbf{x}_k) = L(\mathbf{y}_1, \ldots, \mathbf{y}_k)$.

Prosessin aikana ilmaantuva nollarivi on niinikään lineaarikombinaatio alkuperäisistä riveistä ja tällöin pitäisi vain varmistua siitä, että jokin alkuperäisistä riveistä esiintyy nollarivin esityksessä nollasta poikkeavalla kertoimella. Tästä voidaan kuitenkin helposti varmistua: Gaussin-Jordanin prosessin 1. vaiheessa siirretään ylimmäksi rivi, jossa on vähiten alkunollia, tämän jälkeen vakiolla kertomalla saatetaan ylimmän rivin 1. nollasta poikkeavaksi alkioksi 1, ja sen jälkeen lisätään 1. rivi muihin sopivalla vakiolla kerrottuna, jolloin saadaan 1. rivin alkion 1 alapuolelle pelkkiä nollia. Tämän jälkeen prosessi jatkuu rekursiivisesti, eikä 1. riviä enää lisätä mihinkään. Siksi alussa ylimmäksi siirretty rivi ainakin esiintyy nollasta poikkeavalla kertoimella.

Redusoidussa porrasmuodossa (tai pelkässä porrasmuodossa) olevan matriisin nollasta poikkeavat rivit ovat aina lineaarisesti riippumattomat, koska seuraavassa on enemmän alkunollia kuin edellisessä (kts. Esimerkit 18 ja 19).

2.3 Ratkaisujoukon systemaattinen esittäminen

Tarkastellaan lähinnä esimerkkien valossa miten redusoidusta porrasmatriisista saadaan *kaikki* yhtälöryhmän ratkaisut. Ensinnäkin voidaan todeta, että ratkaisujen puuttuminen havaitaan, mikäli (augmentoitu) matriisi redusoituu esimerkissä 36 olevaan muotoon, siis muotoon, jossa esiintyy yhtälö $0 = c \neq 0$.

Ratkaisujen systemaattinen esittäminen perustuu yksinkertaisesti siihen, että redusoidussa porrasmuodossa olevan yhtälöryhmän ”portaan aloittavat” muuttujat esitetään muiden avulla, kun taas muut muuttujat ”jäävät ratkaisuun”.

Esimerkki 37. Olkoon augmentoidusta yhtälöryhmän matriisista saatu redusoitua matriisi

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

Tätä matriisia vastaa yhtälöryhmä

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

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

Esimerkki 38. Tarkastellaan toisena esimerkkinä redusoitua (augmentoitua) porrasmatriisia

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

jota vastaava yhtälöryhmä on

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

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

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

missä x_4 ja x_5 ovat mitä hyvänsä reaalilukuja.

Esimerkissä 37 yleinen ratkaisu saatiin muodossa $x_3(-2,1,1,0)+(4,-3,0,-1)$, ja esimerkin 38 tapauksessa yleinen ratkaisu on muotoa $x_4(4,-3,-1,1,0) + x_5(6,0,-2,0,1) + (-2,3,7,0,0)$. Kummassakin tapauksessa ratkaisu voidaan esittää sellaisessa muodossa, jossa äärettömään vektorijoukkoon lisätään yksittäinen vektori: Esimerkin 37 tapauksessa kyseinen yksittäinen vektori on $(4,-3,0,-1)$ ja esimerkin $(2,3,7,0,0)$. Lisättyä yksittäistä vektoria kutsutaan *yksityisratkaisuksi*.

Kuten aiemmin huomattiin, esimerkin 38 ratkaisut ovat muotoa

$$t(4,-3,-1,1,0) + s(6,0,-2,0,1) + (-2,3,7,0,0),$$

missä $t, s \in \mathbb{R}$. Voidaan lisäksi huomata, että esimerkin 38 ääretön ratkaisujoukko $\{t(4,-3,-1,1,0) + s(6,0,-2,0,1) \mid s,t \in \mathbb{R}\}$ on ns. *homogenisoidun* yhtälöryhmän

$$\begin{cases} x_1 \qquad -4x_4 \ -6x_5 = 0 \\ \quad x_2 \qquad 3x_4 \qquad\quad = 0 \\ \qquad x_3 \quad x_4 \ +2x_5 = 0 \end{cases}$$

ratkaisujoukko. Tämänkaltainen jäsentely pätee yleisestikin. Seuraava lause onkin edellisten esimerkkien yleistys.

Lause 12. *Jos A on lineaarisen yhtälöryhmän $m \times n$-matriisi (ei siis augmentoitu), jonka aste on r, voidaan ryhmän ratkaisut esittää muodossa $\mathbf{x} = x_{r+1}\mathbf{c}_{r+1} + \ldots + x_n\mathbf{c}_n + \mathbf{c}$, missä $\mathbf{c}_i, \mathbf{c} \in \mathbb{R}^n$ ovat vakiovektoreita ja $x_i \in \mathbb{R}$. Kaikilla kertoimien x_i valinnoilla $x_{r+1}\mathbf{c}_{r+1} + \ldots + x_n\mathbf{c}_n$ on alkuperäisestä yhtälöryhmästä saatavan homogeenisen yhtälöryhmän ratkaisu.*

Todistus (Idea). Muistetaan, että aste r merkitsee A:sta saatavan redusoidun porrasmatriisin P porraslukua (määritelmä 17), ja että jokainen lineaarinen yhtälöryhmä on ekvivalentti sellaisen yhtälöryhmän kanssa jonka matriisi on redusoitua porrasmuotoa. Kuten esimerkeissä 37 ja 38 nähtiin, voidaan kunkin portaan aloittava muuttuja $x_1, \ldots, x_r$ esittää muiden muuttujien $x_{r+1}, \ldots, x_n$ avulla (muuttujat voidaan numeroida uudelleen siten, että nimenomaan $x_1, \ldots, x_r$ aloittavat portaan). Samoin kuin esimerkeissä 37 ja 38, voidaan ratkaisuvektori $\mathbf{x}$ (jossa siis esiintyy $n-r$ muuttujaa $x_{r+1}, \ldots, x_n$) hajoittaa ensin muotoon $\mathbf{x} = \mathbf{x}_{r+1} + \ldots + \mathbf{x}_n + \mathbf{c}$, missä vektorissa $\mathbf{x}_i$ ainoa esiintyvä muuttuja on x_i ja $\mathbf{c}$ on vakiovektori. Tämän jälkeen muuttuja x_i voidaan ottaa vektorin $\mathbf{x}_i$ kertoimeksi, joten $\mathbf{x}_i = x_i\mathbf{c}_i$, missä $\mathbf{c}_i \in \mathbb{R}^n$ on vakiovektori. Näin saadaan haluttu esitys $\mathbf{x} = x_{r+1}\mathbf{c}_{r+1} + \ldots + x_n\mathbf{c}_n + \mathbf{c}$. Jälkimmäinen väite seuraa suoraan siitä, että ratkaisussa esiintyvä vakiovektori $\mathbf{c}$ on nolla tarkalleen silloin kun yhtälöryhmän oikean puolen luvut ovat nollia.

Edellinen lause voidaan ilmaista myös seuraavassa muodossa:

Lause 13. *Lineaarisen yhtälöryhmän kaikki ratkaisut ovat muotoa $\mathbf{x} = \mathbf{x}_1 + \mathbf{c}$, missä $\mathbf{c}$ on jokin yksittäinen ratkaisu ja $\mathbf{x}_1$ on mikä hyvänsä alkuperäisestä ryhmästä saatavan homogeenisen ryhmän ratkaisu.*

Voidaan huomata, että esimerkin 37 yhtälöryhmässä homogenisoidun ryhmän ratkaisut saadaan yhden vektorin skalaarimonikertoina $\{t(-2,1,1,0) \mid t \in \mathbb{R}\}$, kun taas yhtälöryhmää 38 vastaavan homogeenisen ryhmän ratkaisut ovat muotoa $\{t(4,-3,-1,1,0) + s(6,0,-2,0,1) \mid s,t \in \mathbb{R}\}$.

Ensiksi mainitussa yhtälöryhmässä 37 on neljä tuntematonta ja kolme yhtälöä. Redusoidusta porrasmatriisista nähdään, että sen aste on 3. Vastaavasti yhtälöryhmässä 38 on viisi tuntematonta ja kolme yhtälöä, redusoidun porrasmatriisin asteen ollessa 3.

Etsittäessä ratkaisuja näille esimerkeille olemme huomanneet, että ratkaisujen joukko homogenisoidulle yhtälöryhmälle voidaan esimerkin 37 tapauksessa ilmoittaa $4-3=1$ vektorin avulla, ja esimerkin 38 tapauksessa $5-3=2$ vektorin avulla. Nämä lukumäärät eivät ole sattumaa, vaan vastaava yleinen sääntö pätee:

Huomautus 8. Lauseen 12 mukaan homogeenisella yhtälöryhmällä (m yhtälöä ja n tuntematonta) on $n - r(A)$ *(toisistaan) riippumatonta ratkaisua* $\mathbf{c}_{r+1}, \ldots, \mathbf{c}_n$, joiden avulla kaikki muut ratkaisut voidaan esittää.

Esimerkin 37 kerroinmatriisi (ei augmentoitu versio) on 3×4-matriisi, jonka aste on 3; täten kyseistä ryhmää vastaavalla homogeenisella ryhmällä $4-1=1$ riippumatonta ratkaisua. Vastaavasti esimerkin 38 ei-augmentoitu matriisi on 3×5-matriisi, jonka porrasluku on 3; täten ryhmää vastaavalla homogeenisella ryhmällä on $5-3=2$ riippumatonta ratkaisua.

Luvun oleellisia asioita:

- Gaussin-Jordanin menetelmä
- Yhtälöryhmän ratkaisun esittäminen systemaattisesti.

Luku 3
Matriisit

Matriisin käsitettä voidaan käyttää, ja on jo käytettykin jäsentelemään kaavioita: Gaussin-Jordanin menetelmä on hyvin käyttökelpoinen lineaarisen yhtälöryhmän ratkaisemiseksi.

Käsitteenä matriisi on kuitenkin paljon yleisempi kuin pelkkä suorakulmainen lukukaavio, jossa on m riviä n saraketta. Matriisi on lineaarikuvauksen esitystapa ja lineaarikuvaus puolestaan edustaa yksinkertaista (lineaarista) funktiota vektoriavaruudelta toiselle. Lineaarikuvaksen ja matriisin käsitteet ovat saman asian eri esitystapoja ja yhdessä vektoreiden kanssa muodostavat lineaarialgebran rungon.

3.1 Lineaarikuvaus

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

Määritelmä 18. Funktio $f : \mathbb{R}^n \to \mathbb{R}^m$ on lineaarinen, jos

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

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

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

Esimerkki 39. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \sqrt{x}$ ei ole lineaarinen, sillä esimerkiksi

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

Esimerkki 40. $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x,y,z) = (x + y - z, 2x + 3y + z)$ on lineaarinen, sillä

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

Esimerkki 41. Olkoon lineaarikuvaukselle $f : \mathbb{R}^2 \to \mathbb{R}^2$ $f(1,2) = (2,2)$ ja $f(2,3) = (2,5)$ ja selvitetään mitä on $f(5,8)$. Koska $(5,8) = 1 \cdot (1,2) + 2 \cdot (2,3)$, on lineaarisuuden perusteella $f(5,8) = 1 \cdot f(1,2) + 2 \cdot f(2,3) = 1 \cdot (2,2) + 2 \cdot (2,5) = (6,12)$.

Esimerkki 42. Olkoon $C^1[a,b]$ välillä $[a,b]$ derivoituvien reaalifunktioiden joukko, joiden derivaatta on jatkuva. Tällöin derivointi

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

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

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

$$f(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \ldots + a_n\mathbf{x}_n) = a_1 f(\mathbf{x}_1) + a_2 f(\mathbf{x}_2) \ldots + a_n f(\mathbf{x}_n),$$

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

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

saadaan lineaarisuutta käyttämällä

$$f(\mathbf{x}) = x_1 f(\mathbf{e}_1) + \ldots + x_n f(\mathbf{e}_n)$$

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

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

$$f(\mathbf{x})_i = x_1 f(\mathbf{e}_1)_i + x_2 f(\mathbf{e}_2)_i + \ldots + x_n f(\mathbf{e}_n)_i$$

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

$$y_i = A_{i1}x_1 + A_{i2}x_2 + \ldots + A_{in}x_n.$$

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

$$y_i = A_{i1}x_1 + A_{i2}x_2 + \ldots + A_{in}x_n. \tag{3.1}$$

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

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

$$A_f = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \ldots & A_{mn} \end{pmatrix},$$

jota sanotaan lineaarikuvauksen matriisiksi luonnollisen kannan suhteen.

Esimerkki 43. Esimerkin 40 lineaarikuvaukselle $T(x,y,z) = (x+y-z, 2x+3y+z)$ $T(\mathbf{e}_1) = (1,2)$, $T(\mathbf{e}_2) = (1,3)$ ja $T(\mathbf{e}_3) = (-1,1)$, joten

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

Esimerkki 44. Olkoon $f : \mathbb{R}^3 \to \mathbb{R}^2$ lineaarikuvaus, jolle $f(1,0,0) = (2,1)$, $f(0,1,0) = (1,1)$ ja $f(0,0,1) = (0,2)$. Määritetään $f(1,2,3)$. Koska $(1,2,3) = 1\cdot(1,0,0) + 2\cdot(0,1,0) + 3\cdot(0,0,1)$, pätee lineaarikuvaukselle f

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

Esimerkki 45. Määritetään matriisi A_f edellisen esimerkin lineaarikuvaukselle. Koska

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

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

Huomautus 9. Huomaa, että matriisin A_f sarakkeina ovat f:n kuvat vektoreista $(1,0,0)$, $(0,1,0)$ ja $(0,0,1)$.

Jo aiemmissa luvuissa on viitattu matriisin käsitteeseen, mutta se on kuitenkin lähinnä käsitetty vain kaaviona, jonka puitteissa voidaan määritellä rivimanipulaatiota, tai alkiokohtaisia tulo- ja summaoperaatioita. Tässä luvussa esitetään syvällisempi, lineaarikuvauksiin perustuva merkitys matriisin käsitteelle. Aluksi kuitenkin kerrataan tarvittavat käsitteet.

Määritelmä 19. Matriisi A on kaavio

$$
A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}. \tag{3.2}
$$

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

Määritelmä 20. (Vaaka- ja pystyvektorit)

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

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

Mikä hyvänsä $m\times n$-matriisi voidaan siis esittää rivivektoreiden kokoelmana (m kpl) tai sarakevektoreiden kokoelmana (n kpl). Yksin esiintyvän rivivektorin koordinaatit on tapana erottaa pilkulla. Vektoreita on tapana merkitä myös lihavoiduilla kirjaimilla ja yksinkertaisella indeksoinnilla, esimerkiksi $\mathbf{x} = (x_1,x_2,\dots,x_n)$ ja $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$. Sekä rivi- että sarakevektorit voidaan tulkita avaruuden $\mathbb{R}^n$ alkioiksi.

Määritelmä 21. Erityisiä matriiseja ovat mm. seuraavat:

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

Määritelmä 22. Olkoon A $m \times n$-matriisi. Tällöin määritellään

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

Määritelmä 23 (Skalaarikertolasku). Jos A on $m \times n$-matriisi ja c joko kompleksi- tai reaaliluku, on cA $m \times n$-matriisi, jolle pätee $(cA)_{ij} = cA_{ij}$ (kertolasku alkioittain).

Määritelmä 24 (Matriisien yhteenlasku). Jos A on $m \times n$-matriisi ja B $r \times s$-matriisi, summa $A+B$ määritellään vain jos $m = r$ ja $n = s$. Tällöin

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

(yhteenlasku alkioittain).

Huomautus 10. On helppo todeta, että $m \times n$-matriisit muodostavat yhteen- ja skalaarikertolaskun suhteen vektoriavaruuden.

3.2 Matriisikertolasku

Matriisikertolasku määritellään aluksi vain siinä tapauksessa, että oikeanpuoleinen matriisi on sarakevektori (siis $n \times 1$-matriisi) ja sen jälkeen määritelmää laajennetaan johdonmukaisella tavalla.

Tarve määritellä matriisikertolaskun käsite juontaa juurensa siitä, että lineaarikuvauksen $\mathbb{R}^n \to \mathbb{R}^m$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \to \begin{pmatrix} A_{11}x_1 + A_{12}x_2 + \ldots + A_{1n}x_n \\ A_{21}x_1 + A_{22}x_2 + \ldots + A_{2n}x_n \\ \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \ldots + A_{mn}x_n \end{pmatrix}$$

kuvavektorissa tahdotaan erottaa toisistaan kuvauksen määrittelevät luvut A_{ij} (matriisialkiot) ja alkukuvavektorin koordinaatit x_i ja siis määritellä sellainen $m \times n$ ja $n \times 1$-matriisin tulo, että

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

Ylläoleva yhtäsuuruus voidaan esittää seuraavan määritelmän muodossa.

Määritelmä 25. Olkoon A $m \times n$-matriisi ja $\mathbf{x}$ $n \times 1$-matriisi Tällöin matriisisitulo $A\mathbf{x}$ on $m \times 1$-matriisi jonka koordinaatit ovat

$$(A\mathbf{x})_i = \sum_{l=1}^{n} A_{il}\mathbf{x}_l$$

kaikille $i \in \{1, 2, \ldots, m\}$.

Esimerkki 46. Esimerkin 45 lineaarikuvauksen kuvavektori voidaan sarakevektorina esittää seuraavasti:

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

Edellämainitun perusteella seuraava lause on ilmeinen.

Lause 14. *Lineaarikuvaus $f : \mathbb{R}^n \to \mathbb{R}^m$ voidaan esittää muodossa $f(\mathbf{x}) = A\mathbf{x}$, missä $\mathbf{x}$ on $n \times 1$-matriisi (pystyvektori) ja A on $m \times n$-matriisi. Matriisi $A = A_f$ on lineaarikuvauksen f matriisi luonnollisen kannan suhteen.*

Matriisitulo voidaan välittömästi laajentaa myös sellaiseen tapaukseen, jossa oikeanpuoleinen matriisi ei ole sarakevektori. Tällöin tulee kuitenkin huolehtia siitä, että vasemmanpuoleisen matriisin sarakkeiden määrä on yhtäsuuri kuin oikeanpuoleisen matriisin rivien määrä.

Määritelmä 26. Olkoon A $m \times n$-matriisi ja B $n \times k$-matriisi. Tällöin B voidaan esittää muodossa $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$, missä jokainen $\mathbf{b}_j$ on $n \times 1$ sarakevektori. Tällöin matriisitulo AB määritellään

$$AB = (A\mathbf{b}_1, \ldots, A\mathbf{b}_k)$$

Mikäli ylläolevassa määritelmässä merkitään

$$\mathbf{b}_j = \begin{pmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{pmatrix},$$

voidaan matriisitulon määritelmä kirjoittaa symmetrisempään, mutta yhtäpitävään muotoon:

Määritelmä 27. Jos A on $m \times n$-matriisi ja B $n \times k$-matriisi, on matriisien A ja B *tulo* $m \times k$-matriisi AB, missä

$$(AB)_{ij} = \sum_{l=1}^{n} A_{il} B_{lj}.$$

Huomautus 11. Matriisitulo on siis määritelty vain silloin, kun tulon ensimmäisen tekijän sarakkeiden määrä on yhtä suuri kuin jälkimmäisen tekijän rivien määrä.

Esimerkki 47.

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

Esimerkki 48.

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

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

Esimerkki 49.

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

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

Esimerkki 50. $A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3-matriisi ja $5A = \begin{pmatrix} 10 & -5 & 0 \\ 0 & 5 & 15 \end{pmatrix}$

Esimerkki 51. $A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3-matriisi ja $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 3 \end{pmatrix}$ 3×2-matriisi, joten summaa $A+B$ ei ole määritelty.

Esimerkki 52. $A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3-matriisi ja $B = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \end{pmatrix}$ on 2×3-matriisi, joten summa $A+B$ on määritelty:

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

Esimerkki 53.

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

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

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

Syötetään seuraavasti:

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

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

Matriisien summa lasketaan ilmeisellä tavalla ja transpoosi ilmaistaan heittomerkillä:

```
>> A+B

ans =
     3     3    -4    -3
     1     2     4     1
    -5     6     2    -5

>>
>> ans'

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

Lause 15. *Matriisitulolle on voimassa seuraavat yhtäsuuruudet:*

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

edellyttäen että kunkin yhtälön vasen puoli on määritelty. edellyttäen että vasemmat puolet ovat määriteltyjä.

O on nollamatriisi ja I identiteettimatriisi, jonka kaikki diagonaalialkiot ovat ykkösiä ja muut nollia. A^T merkitsee matriisin A transpoosia: $(A^T)_{ij} = A_{ji}$, siis A^T saadaan matriisista A vaihtamalla rivit sarakkeiksi.

Todistus. Suoraviivainen, mutta työläs. Jätetään harjoitustehtäväksi.

Huomautus 12. Suoraan määritelmästä seuraa, että matriisitulo voidaan laskea myös lohkoittain:

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} = \begin{pmatrix} A_1B_1+A_2B_3 & A_1B_2+A_2B_4 \\ A_3B_1+A_4B_3 & A_3B_2+A_4B_4 \end{pmatrix},$$

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

Määritelmä 28. Neliömatriiseile määritellään *potenssi* samoin kuin reaaliluvuille:

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

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

Matlab-ohjelmistossa matriisikertolasku ilmaistaan asteriskilla $\star$ ja potenssi nuolella ˆ

```
>> A*B

ans =

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

>>
```

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

Lause 16. *Jos lineaarikuvausten $f : \mathbb{R}^n \to \mathbb{R}^m$ ja $g : \mathbb{R}^m \to \mathbb{R}^k$ matriisit ovat A_f ja A_g, niin yhdistetyn kuvauksen $g \circ f$ matriisi on tulo A_gA_f.*

Huomautus 13. Edellisen lauseen tapauksessa matriisi A_f on siis tyyppiä $m \times n$ ja A_g tyyppiä $k \times m$. Tällöin matriisitulo A_gA_f on määritelty ja tyyppiä $k \times n$, kuten kuvauksen $\mathbb{R}^n \to \mathbb{R}^k$ matriisin tulee ollakin.

Huomautus 14. Olkoon A $m\times n$-matriisi, $\mathbf{x}$ n-pituinen pystyvektori ja $\mathbf{y}=A\mathbf{x}$ m-pituinen pystyvektori. Transpoosit laskemalla saadaan $\mathbf{y}^T=\mathbf{x}^TA^T$, mikä antaa vaihtoehtoisen muodon lineaarikuvauksen esittämiseksi matriisien avulla.

Esimerkki 54. Insinöörimatematiikan myöhemmässä osassa esitettävä diskreetti Fourier-muunnos on $\mathbb{C}^N$:ssä tapahtuva operaatio, jossa vektori $(f_0,f_1,\dots,f_{N-1})$ muunnetaan vektoriksi $(F_0,F_1,\dots,F_{N-1})$, missä

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

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

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

Esimerkki 55. Joukot $B_1=\{\mathbf{i},\mathbf{j}\}$ ja $B_2=\{\mathbf{i}+\mathbf{j},\mathbf{i}-\mathbf{j}\}$ muodostavat avaruuden $\mathbb{R}^2$ kannan. Voidaan todeta että koordinaattivektorit ovat $(x,y)_{B_1}=(x,y)$ ja $(x,y)_{B_2}=(\frac{1}{2}(x+y),\frac{1}{2}(x-y))$. Tällöin siis

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

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

Esimerkki 56. Lineaarinen yhtälöryhmä

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

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

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

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

3.3 Käänteismatriisi

Määritelmä 29. Olkoon A $n\times n$-neliömatriisi. Jos on olemassa sellainen $n\times n$-neliömatriisi B, että $AB=BA=I_n$ (I_n on n-rivinen identiteettimatriisi), sanotaan että B on A:n *käänteismatriisi* ja merkitään $B=A^{-1}$. Jos neliömatriisilla A on käänteismatriisi, A:ta kutsutaan *säännölliseksi*, muussa tapauksessa A on *singulaarinen*.

Huomautus 15. Voidaan todistaa, että $n\times n$-neliömatriisi A on säännöllinen tarkalleen silloin, kun se on *täysiasteinen*, mikä merkitsee sitä, että $r(A)=n$.

Esimerkki 57. Koska $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, on $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

Esimerkki 58. Lasketaan esimerkissä 54 esintyneen matriisin tulo toisen samankaltaisen matriisin kanssa:

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

Tulon määritelmän mukaan kohdassa rs oleva alkio on

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

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

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

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

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

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

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

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

Esimerkki 59. Matriisin $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ käänteismatriisin löytäminen alkeismuunnoksilla tapahtuu Gaussin-Jordanin eliminointimenetelmää soveltaen seuraavasti: Otetaan ensiksi käyttöön (useampikertaisesti) augmentoitu matriisi $(A\ I) = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$, johon sovelletaan alkeismuunnoksia: Ensiksi kerrotaan ensimmäinen rivi vakiolla $\frac{1}{2}$, minkä tuloksena saadaan $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$. Seuraavaksi lisätään ensimmäinen rivi toiseen kerrottuna luvulla -1, jolloin tuloksena on $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}$. kertomalla tämän matriisin toinen rivi 2:lla saadaan $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & -1 & 2 \end{pmatrix}$, mistä toinen rivi $-\frac{1}{2}$:lla kerrottuna ensimmäiseen lisättynä antaa $\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 2 \end{pmatrix}$. Tällöin siis alkuperäinen matriisi on muunnettu alkeismuunnoksilla identiteettimatriisiksi, ja sen käänteismatriisi $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ voidaan lukea augmentoidun matriisin oikeanpuoleisesta lohkosta.

Huomautus 16. Jos neliömatriisin A käänteismatriisi A^{-1} on helposti saatavilla, voidaan sen avulla ratkaista yhtälöryhmä $A\mathbf{x} = \mathbf{b}$. Tällöin nimittäin yhtälöstä $A\mathbf{x} = \mathbf{b}$ seuraa $A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$, mikä käänteismatriisin määritelmän perusteella voidaan kirjoittaa muotoon $\mathbf{x} = A^{-1}\mathbf{b}$.

Esimerkki 60. Ratkaistaan yhtälöpari

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

käänteismatriisia käyttäen. Matriisitulon määritelmän perusteella kyseinen yhtälöpari voidaan kirjoittaa muodossa

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

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

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

mikä sievenee muotoon

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

On huomattava, että käänteismatriisin etsimiseen perustuvassa yhtälöryhmän ratkaisumenetelmässä on huomattavan paljon rajoitteita verrattuna yleiseen Gaussin-Jordanin menetelmään. On nimittäin oletettava että muuttujia on yhtä monta kuin tuntematonta ja että yhtälöryhmän kerroinmatriisin käänteismatriisi on ylipäänsä olemassa. Lisäksi on huomattava, että käänteismatriisin etsiminen on jopa työläämpää kuin yhtälöryhmän ratkaiseminen Gaussin-Jordanin menetelmällä (mitään huomattavan paljon helpompaa tapaa kuin yllä esitetty ei tunneta). Tämän vuoksi voidaan perustellusti kysyä mitä mieltä ylipäänsä koko käänteismatriisin käsitteessä on.

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

3.4 Determinantti

Determinantin käsite voidaan määritellä joko algebrallisesta, kombinatorisesta tai geometrisesta näkökulmasta lähtien. Koska lähtökohtia on useita, on ymmärrettävää, että determinantin käsite on erittäin käyttökelpoinen. Tämän kurssin puitteissa ei kuitenkaan voida esitellä determinantin kytkentöjä kaikkiin osa-alueihin yksityiskohtaisesti.

Tällä kurssilla valitaan determinantin käsitteen perustaksi algebrallinen näkökulma ja asetetaan multilineaarisiin kuvauksiin perustuva määritelmä.

Määritelmä 30. Kuvaus $f : V \times \ldots \times V \to U$ on *multilineaarinen*, jos se on lineaarinen jokaisen muuttujan suhteen, siis

$$f(\ldots, a\mathbf{x}, \ldots) = af(\ldots, \mathbf{x}, \ldots)$$

ja

$$f(\ldots, \mathbf{x} + \mathbf{y}, \ldots) = f(\ldots, \mathbf{x}, \ldots) + f(\ldots, \mathbf{y}, \ldots).$$

Huomautus 17. Jos kuvaus on multilineaarinen, on sen arvo välttämättä $\mathbf{0}$ mikäli yksikin muuttujista on $\mathbf{0}$. Tämä nähdään seuraavasti:

$$f(\ldots,\mathbf{0},\ldots) = f(\ldots,\mathbf{0}+\mathbf{0},\ldots) = f(\ldots,\mathbf{0},\ldots) + f(\ldots,\mathbf{0},\ldots),$$

joten vähentämällä $f(\ldots,\mathbf{0},\ldots)$ nähdään että $f(\ldots,\mathbf{0},\ldots) = \mathbf{0}$. Tämä pätee tietysti myös yhden vektorimuuttujan lineaarikuvauksille: $f(\mathbf{0}) = \mathbf{0}$ aina kun f on lineaarikuvaus.

Määritelmä 31. Multilineaarinen kuvaus on *alternoiva*, jos sen merkki vaihtuu aina kun kahden vektorin paikkaa vaihdetaan:

$$f(\ldots,\mathbf{x},\ldots,\mathbf{y},\ldots) = -f(\ldots,\mathbf{y},\ldots,\mathbf{x},\ldots)$$

Seuraus 3. Jos alternoivassa kuvauksessa esiintyy kaksi samaa vektoria, sen arvo $\mathbf{0}$.

Tämä seuraa yksinkertaisesti sijoittamalla $\mathbf{y} = \mathbf{x}$ yllä olevaan yhtälöön, jolloin

$$f(\ldots,\mathbf{x},\ldots,\mathbf{x},\ldots) = -f(\ldots,\mathbf{x},\ldots,\mathbf{x},\ldots),$$

joten $f(\ldots,\mathbf{x},\ldots,\mathbf{x},\ldots) = \mathbf{0}$.

Lause 17. *Jos multilineaarisessa kuvauksessa lisätään jokin vektori toiseen skalaarilla kerrottuna, ei kuvauksen arvo muutu. Tämä nähdään seuraavasti:*

$$f(\ldots,\mathbf{x},\ldots,\mathbf{y}+a\mathbf{x},\ldots) = f(\ldots,\mathbf{x},\ldots,\mathbf{y},\ldots) + a\underbrace{f(\ldots,\mathbf{x},\ldots,\mathbf{x},\ldots)}_{\mathbf{0}} = f(\ldots,\mathbf{x},\ldots,\mathbf{y},\ldots)$$

Lause 18. *Multilineaarinen alternoiva kuvaus määräytyy yksikäsitteisesti arvosta $f(\mathbf{b}_1,\ldots,\mathbf{b}_n)$, missä $\{\mathbf{b}_1,\ldots,\mathbf{b}_n\}$ on kanta.*

Todistus. Alternoivuuden perusteella $f(\ldots,\mathbf{b}_i,\ldots,\mathbf{b}_i,\ldots) = \mathbf{0}$, jos kantavektori $\mathbf{b}_i$ esiintyy kahteen kertaan. Tämän vuoksi riittää tarkastella ainoastaan sellaisia arvoja $f(\mathbf{b}_{i_1},\ldots,\mathbf{b}_{i_n})$, jossa kaikki kantavektorit esiintyvät vain kerran.

Koska vektoreiden paikan vaihto vaihtaa funktion f merkin, voidaan todeta, että

$$f(\mathbf{b}_{i_1},\ldots,\mathbf{b}_{i_n}) = \pm f(\mathbf{b}_1,\ldots,\mathbf{b}_n),$$

missä etumerkki riippuu siitä kuinka monta kahden vektorin paikanvaihtoa on tarvittu aikaansaamaan järjestys $(\mathbf{b}_{i_1},\ldots,\mathbf{b}_{i_n})$.

Esittämällä jokainen vektori $\mathbf{v}_i$ kantavektoreiden avulla

$$\mathbf{v}_i = \sum_{j=1}^{m} v_{ij}\mathbf{b}_j$$

saadaan

$$\begin{aligned}
f(\mathbf{v}_1,\ldots,\mathbf{v}_n) &= f\Big(\sum_{j_1=1}^{m} v_{1j_1}\mathbf{b}_{j_1},\ldots,\sum_{j_n=1}^{m} v_{nj_n}\mathbf{b}_{j_n}\Big)\\
&= \sum_{j_1=1}^{m}\cdots\sum_{j_n=1}^{m} v_{1j_1}\cdots v_{nj_n} f(\mathbf{b}_{j_1},\ldots,\mathbf{b}_{j_n})\\
&= \underbrace{\sum_{j_1=1}^{m}\cdots\sum_{j_n=1}^{m}}_{(j_1,\ldots,j_n)\text{ kaikki erisuuret}} v_{1j_1}\cdots v_{nj_n} f(\mathbf{b}_{j_1},\ldots,\mathbf{b}_{j_n})\\
&= \underbrace{\sum_{j_1=1}^{m}\cdots\sum_{j_n=1}^{m}}_{(j_1,\ldots,j_n)\text{ kaikki erisuuret}} v_{1j_1}\cdots v_{nj_n}(\pm 1) f(\mathbf{b}_1,\ldots,\mathbf{b_n}).
\end{aligned}$$

Toiseksi viimeisin yhtäsuuruus perustuu siihen, että saman kantavektorin esiintyessä kaksi kertaa on funktion f arvo $\mathbf{0}$. Viimeisin yhtäsuuruus perustuu siihen, että järjestyksessä $(\mathbf{b}_{i_1},\ldots,\mathbf{b}_{i_n})$ esiintyvät kantavektorit voidaan toistuvasti kahden vektorin paikkaa vaihtamalla saattaa järjestykseen

$(\mathbf{b}_1,\ldots,\mathbf{b}_n)$ ja jokaisessa kahden vektorin välisessä paikanvaihdossa kuvauksen etumerkki muuttuu. □

Esimerkki 61. Kolmen kantavektorin tapauksessa voidaan helposti laskea

$$f(\mathbf{b}_3,\mathbf{b}_1,\mathbf{b}_2) = -f(\mathbf{b}_1,\mathbf{b}_3,\mathbf{b}_2) = f(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3),$$

$$f(\mathbf{b}_3,\mathbf{b}_2,\mathbf{b}_1) = -f(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3),$$

ja

$$f(\mathbf{b}_2,\mathbf{b}_1,\mathbf{b}_3) = -f(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3).$$

Etumerkin tarkempaa tarkastelua varten otetaan käyttöön seuraava käsite, joka myös helpottaa yllä esiintyvien summien merkitsemistä.

Määritelmä 32. Permutaatio tarkoittaa jonon $(1,2,\ldots,n)$ uudelleenjärjestystä $(i_1,i_2,\ldots,i_n)$, jossa kaikki luvut 1, 2, ..., n esiintyvät jossain järjestyksessä. Jonon $(1,2,\ldots,n)$ kaikkien permutaatioiden joukosta käytetään merkintää S_n.
Permutaation merkki $\operatorname{sgn}(i_1,i_2,\ldots,i_n)$ on $+1$ ja permutaatiota sanotaan *parilliseksi*, jos se on saatu jonosta $(1,2,\ldots,n)$ parillisella määrällä kahden alkion vaihtoja. Permutaation merkki on -1 ja sitä sanotaan *parittomaksi*, jos tarvittavien vaihtojen määrä on pariton. Kahden alkion vaihtoa sanotaan *transpositioksi*.

Esimerkki 62. Joukon $\{1,2,3\}$ kaikki permutaatiot voidaan muodostaa valitsemalla 1. alkioksi ensin 1, sitten 2 ja lopulta 3. Mikäli 1 on 1. alkio, voidaan toiseksi alkioksi valita joko 2 tai 3, jolloin saadaan 2 erilaista permutaatiota $(1,2,3)$ ja $(1,3,2)$. Samoin jatkamalla voidaan todeta, että kolmen alkion joukosta $\{1,2,3\}$ voidaan laatia $6 = 3\cdot 2\cdot 1 = 3!$ permutaatiota

$$S_3 = \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}.$$

Yleisemmin voidaan todeta, että joukosta $\{1,2,\ldots,n\}$ voidaan laatia $n\cdot(n-1)\cdot\ldots\cdot 2\cdot 1 = n!$ permutaatiota, siis

$$|S_n| = n!.$$

Esimerkki 63. Kahdella transpositiolla saadaan $(1,2,3)\to(2,1,3)\to(3,1,2)$. Näin ollen $\operatorname{sgn}(3,1,2) = +1$.

Esimerkki 64. Yhdellä transpositiolla saadaan $(1,2,3)\to(3,2,1)$, joten $\operatorname{sgn}(3,2,1) = -1$.

Esimerkki 65. Kahdella transpositiolla saadaan $(1,2,3)\to(3,2,1)\to(3,1,2)$, joten $\operatorname{sgn}(3,1,2) = +1$.

Huomautus 18. Koska jokainen permutaatio voidaan muodostaa monella tavalla käyttämällä transpositioita peräkkäin, on perusteltua kysyä onko permutaation merkki hyvin määritelty käsite, siis voitaisiinko sama permutaatio saada sekä parillisella että parittomalla määrällä transpositioita. Pienellä pohdinnalla voidaan päätyä siihen tulokseen, että näin ei voi käydä, vaan tarvittavien transpositioiden määrän pariteetti on aina sama riippumatta siitä millä tavalla permutaatio muodostetaan.

Lause 19. *Olkoon $(i_1,\ldots,i_n)$ jokin jonon $(1,2,\ldots,n)$ permutaatio. Sanotaan, että pari (i_{k_1},i_{k_2}) esiintyy* käännetyssä järjestyksessä, *jos $i_{k_1} > i_{k_2}$. Olkoon m käännetyssä järjestyksessä esiintyvien parien määrä. Tällöin* $\operatorname{sgn}(i_1,\ldots,i_n) = (-1)^m$.

Esimerkki 66. Permutaatiossa $(3,1,2)$ parit $(3,1)$ ja $(3,2)$ esiintyvät käännetyssä järjestyksessä. Näin ollen $\operatorname{sgn}(3,1,2) = (-1)^2 = +1$.

Esimerkki 67. Permutaatiossa $(3,2,1)$ parit $(3,2)$, $(3,1)$ ja $(2,1)$ esiintyvät käännetyssä järjestyksessä. Näin ollen $\operatorname{sgn}(3,2,1) = (-1)^3 = -1$.

Esimerkki 68. Permutaatiossa $(3,1,2)$ parit $(3,1)$ ja $(3,2)$ esiintyvät käännetyssä järjestyksessä. Näin ollen $\operatorname{sgn}(3,2,1) = (-1)^2 = +1$.

Käyttämällä permutaation sgn-merkintää voidaan aiemman lauseen mukainen esitys multilineaarisen alternoivan kuvauksen arvolle kirjoittaa muodossa

$$f(\mathbf{v}_1,\ldots,\mathbf{v}_n) = \sum_{(j_1,\ldots,j_n)\in S_n} \operatorname{sgn}(j_1,\ldots,j_n) v_{1j_1}\ldots\cdot v_{nj_n} f(\mathbf{b}_1,\ldots,\mathbf{b_n}). \tag{3.3}$$

Tässä merkinnässä siis v_{ij_1} edustaa vektorin $\mathbf{v}_i$ koordinaatteja kannassa $B = \{\mathbf{b}_1,\ldots,\mathbf{b}_n\}$ ja jono $(j_1,\ldots,j_n)$ valitaan siten, että kahta samaa arvoa ei esiinny.

Olkoon $\mathbb{K}$ skalaarikunta (yleensä joko $\mathbb{R}$ tai $\mathbb{C}$) ja $D : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$ multilineaarinen, alternoiva kuvaus, jolle pätee

$$D(\mathbf{e}_1^T,\ldots,\mathbf{e}_n^T) = 1. \tag{3.4}$$

Tässä $\mathbf{e}_1^T$, ..., $\mathbf{e}_n^T$ ovat luonnollisen kannan vektoreita (sarakemuodossa kirjoitettuna).

Määritelmä 33. Esitetään neliömatriisi $A = (\mathbf{a}_1 \ldots \mathbf{a}_n)$ sarakkeiden $\mathbf{a}_i$ kokoelmana. Tällöin determinantti $\det(A)$ määritellään

$$\det(A) = D(\mathbf{a}_1,\ldots,\mathbf{a}_n).$$

Determinantista käytetään myös merkintää $|A|$. Tässä merkinnässä jätetään pois matriisin A molemmat kaarisulkeet. Lyhyyden vuoksi toisinaan myös samaistetaan merkinnät $\det(A)$ ja $D(\mathbf{x}_1,\ldots,\mathbf{x}_n)$, kun $A = (\mathbf{x}_1,\ldots,\mathbf{x}_n)$ ja merkitään $\det(A) = \det(\mathbf{x}_1,\ldots,\mathbf{x}_n)$.

Seuraus 4. Ylläolevan määritelmän välittömiä seurauksia ovat

- Identiteettimatriisin determinantti $\det(I) = 1$. Tämä seuraa vaatimuksesta (3.4) ja siitä, että identiteettimatriisin sarakkeet koostuvat luonnollisen kannan vektoreista: $I = (\mathbf{e}_1^T,\ldots,\mathbf{e}_n^T)$.
- Jos matriisissa on nollasarake, on sen determinantti 0. Tämä seuraa Huomautuksesta 17.
- Jos matriisissa on kaksi samaa saraketta, on sen determinantti 0 Tämä johtuu seurauksesta 3.
- Jos matriisin sarake lisätään toiseen vakiolla kerrottuna, determinantti ei muutu. Tämä seuraa Lauseesta 17
- Jos matriisin kahden sarakkeen paikkaa vaihdetaan, muuttuu determinantin merkki. Tämä seuraa siitä, että determinantti on määritelty sarakkeiden alternoivana kuvauksena.
- Determinantin eteen voidaan ottaa vakiokerroin miltä hyvänsä sarakkeelta, siis $\det(\mathbf{x_1},\ldots,a\mathbf{x}_i,\ldots,\mathbf{x}_n) = a\det(\mathbf{x_1},\ldots,\mathbf{x}_i,\ldots,\mathbf{x}_n)$. Tämä seuraa (multi)lineaarisuudesta.
- $\det(\mathbf{x_1},\ldots,\mathbf{x}+\mathbf{y},\ldots,\mathbf{x}_n) = \det(\mathbf{x_1},\ldots,\mathbf{x},\ldots,\mathbf{x}_n) + \det(\mathbf{x_1},\ldots,\mathbf{y},\ldots,\mathbf{x}_n)$. Tämä seuraa myös lineaarisuudesta.

Determinantin ominaisuuksista seuraa myös hyvin tärkeä geometrinen tulkinta (selitetään luennolla)

Lause 20. *Kun $A = (\mathbf{a}_1,\ldots,\mathbf{a}_n)$ (sarakkeet), determinantin itseisarvo $|\det(A)|$ merkitsee vektoreiden $\{\mathbf{a}_1,\ldots,\mathbf{a}_n\}$ generoiman särmiön tilavuutta*

Seuraavalla lauseella on myös teorian kannalta hyvin kauaskantoisia seurauksia.

Lause 21. *Jos A ja B ovat $n\times n$-matriiseja, on*

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

Todistus. Sekä $f(X) = \det(AX)$ ja $g(X) = \det(A)\det(X)$ ovat matriisin X sarakkeiden alternoivia multilineaarikuvauksia, joille $f(I) = g(I) = \det(A)$. Väite seuraa nyt suoraan aiemmin todistetusta alternoivien multilineaarikuvasten yksikäsitteisyydestä: $f(X) = g(X)$, myös kun $X = B$.

Kaavan (3.3) seurauksena saadaan matriisin

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \ldots & A_{nn} \end{pmatrix}$$

determinantille esitys

$$\det(A) = \sum_{(i_1,i_2,\dots,i_n)\in S_n} \operatorname{sgn}(i_1,i_2,\dots,i_n)A_{1i_1}A_{2i_2}\dots A_{ni_n}, \tag{3.5}$$

jolla on hyvin tärkeitä seurauksia. Ennen näihin tutustumista voidaan tarkastella pienten neliömatriisien determinantteja.

Esimerkki 69. Tapauksessa $n = 2$ on permutaatioiden joukko $S_2 = \{(1,2),(2,1)\}$, joille $\operatorname{sgn}(1,2) = +1$ ja $\operatorname{sgn}(2,1) = -1$. Laskettaessa kaavan (3.5) mukaan matriisin

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

determinanttia saadaan summaan vain kaksi termiä:

$$\det(A) = \operatorname{sgn}(1,2)A_{11}A_{22} + \operatorname{sgn}(2,1)A_{12}A_{21} = A_{11}A_{22} - A_{12}A_{21}.$$

Esimerkki 70. Tapauksessa $n = 3$ on kuusi permutaatiota: $S_3 = \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$, joiden merkit ovat $+1$, -1, -1, $+1$, $+1$ ja -1. Täten

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

Edelliset esimerkit havainnollistavat kaavaa (3.5), josta voidaan kuitenkin havaita, että se ei tuota erityisen käyttökelpoista laskennallista menetelmää determinantin määrittämiseksi, sillä summassa (3.5) on $n!$ yhteenlaskettavaa. Kertomafunktio $n!$ kasvaa nopeammin kuin mikään eksponenttifunktio, ja jo pienillä arvoilla $3! = 6$, $4! = 24$ $5! = 120, 6! = 720$ voidaan todeta summattavien määrän kasvavan hyvin suureksi.

Kaava (3.5) ei siis suoraan tuota tehokasta laskentamenetelmää determinantille, mutta tuota pikaa nähdään, että välillisesti sellainen löyty kaavan (3.5) avulla. Sitä ennen voidaan kuitenkin todeta, että jokaisessa yhteenlaskettavassa

$$A_{1i_1}A_{2i_2}\cdot\ldots\cdot A_{ni_n} \tag{3.6}$$

on valittu yksi alkio jokaiselta riviltä ja jokaiselta sarakkeelta. Merkinnässä (3.6) on alkiot valittu kasvavan rivijärjestyksen (1. alaindeksi) mukaan, mutta sarake vaihtelee järjestyksessä $(i_1,i_2,\dots,i_n)$. Tulo (3.6) on kuitenkin mahdollista järjestää myös niin, että *sarakkeiden* järjestys on kasvava ja rivien järjestys vaihtelee:

$$A_{1i_1}A_{2i_2}\cdot\ldots\cdot A_{ni_n} = A_{j_i1}A_{j_22}\cdot\ldots\cdot A_{j_nn}, \tag{3.7}$$

missä $(j_1,j_2,\dots,j_n)$ on jokin rivien permutaatio.

Esimerkki 71. Permutaatio $(1,3,2)$ tuottaa determinantin (70) lausekkeeseen yhteenlaskettavan $-A_{11}A_{23}A_{32} = -A_{11}A_{32}A_{23}$. Jälkimmäisessä lausekkeessa sarakenumero kasvaa $1 \to 2 \to 3$ ja rivinumerot määräytyvät permutaation $(1,3,2)$ mukaan.

Permutaatio $(3,1,2)$ puolestaan tuottaa determinantin (70) lausekkeeseen termin $A_{13}A_{21}A_{32} = A_{21}A_{32}A_{13}$, jossa rivit määräytyvät permutaation $(2,3,1)$ mukaan. Voidaan huomata, että $\operatorname{sgn}(3,1,2) = +1$, koska $(3,1,2)$ saadaan perusjärjestyksestä $(1,2,3)$ kahdella transpositiolla, ja samoin on permutaation $(2,3,1)$ laita, siis $\operatorname{sgn}(2,3,1) = +1$.

On verrattain helppoa todistaa seuraava lause:

Lause 22. *Olkoot merkinnät kuten yhtälössä (3.7). Tällöin*

$$\operatorname{sgn}(i_1,i_2,\dots,i_n) = \operatorname{sgn}(j_1,j_2,\dots,j_n).$$

Determinantin summakaava (3.5) tuottaa tällöin

$$\begin{aligned}
\det(A) &= \sum_{(i_1,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(i_1,i_2,\ldots,i_n)A_{1i_1}A_{2i_2}\cdot\ldots\cdot A_{ni_n} \\
&= \sum_{(j_1,j_2,\ldots,j_n)\in S_n} \operatorname{sgn}(j_1,j_2,\ldots,j_n)A_{j_1 1}A_{j_2 2}\cdot\ldots\cdot A_{j_n n} \\
&= \sum_{(j_1,j_2,\ldots,j_n)\in S_n} \operatorname{sgn}(j_1,j_2,\ldots,j_n)A^T_{1j_1}A^T_{2j_2}\cdot\ldots\cdot A^T_{nj_n} \\
&= \det(A^T),
\end{aligned}$$

joka voidaan ilmaista kompaktisti yhdessä lauseessa:

Lause 23.

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

Huomautus 19. Edellisen lauseen perustella determinantti voidaan käsittää multilineaarisena kuvauksena matriisien *rivien* esittäessä vektoreita. Seurauksen 4 kaikki sarakkeita koskevat väittämät soveltuvat myös siis matriisin riveihin.

Käytännöllinen menetelmä determinantin arvon laskemiseksi voidaan löytää kun käytetään yhdistelmänä Seurauksen 4 huomioita sekä summaesityksen (3.5) seurauksia, joihin seuraavaksi tutustutaan lyhyesti.

Määritelmä 34. Matriisin A *alimatriisi* $A[ij]$ saadaan matriisista A pyyhkimällä pois rivi i ja sarake j.

Esimerkki 72. Olkoon $A = \begin{pmatrix} -1 & 2 & 0 \\ 3 & -5 & 1 \\ 2 & 2 & 3 \end{pmatrix}$. Tällöin $A[11] = \begin{pmatrix} -5 & 1 \\ 2 & 3 \end{pmatrix}$, $A[12] = \begin{pmatrix} 3 & 1 \\ 2 & 3 \end{pmatrix}$, $A[13] = \begin{pmatrix} 3 & -5 \\ 2 & 2 \end{pmatrix}$ ja $A[32] = \begin{pmatrix} -1 & 0 \\ 3 & 1 \end{pmatrix}$.

Määritelmä 35. Matriisin A alkion A_{ij} *komplementti* $C_{ij} = (-1)^{i+j}\det A[ij]$.

Esimerkki 73. Olkoon A Esimerkin 72 matriisi. Tällöin $C_{11} = (-1)^{1+1}\begin{vmatrix} -5 & 1 \\ 2 & 3 \end{vmatrix} = -17$, $C_{12} = (-1)^{1+2}\begin{vmatrix} 3 & 1 \\ 2 & 3 \end{vmatrix} = -7$, $C_{13} = (-1)^{1+3}\begin{vmatrix} 3 & -5 \\ 2 & 2 \end{vmatrix} = 16$ ja $C_{32} = (-1)^{3+2}\begin{vmatrix} -1 & 0 \\ 3 & 1 \end{vmatrix} = 4$.

Syy komplementin käsitteen esille tuomiseen selviää, kun huomataan, että determinantin summakaavassa (3.5) voidaan jokaisesta tulon tekijästä ottaa yhteiseksi tekijäksi jonkin rivin tai sarakkeen alkioita. Jos esimerkiksi ryhmitellään joukon S_n permutaatiot 1. alkion mukaisesti, saadaan determinantin lauseke jaoteltua seuraavasti:

$$\begin{aligned}
\det(A) &= \sum_{(i_1,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(i_1,i_2,\ldots,i_n)A_{1i_1}A_{2i_2}\ldots A_{ni_n} \\
&= \sum_{(1,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(1,i_2,\ldots,i_n)A_{11}A_{2i_2}\ldots A_{ni_n} \\
&= \sum_{(2,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(2,i_2,\ldots,i_n)A_{12}A_{2i_2}\ldots A_{ni_n} \\
&\quad +\cdots+ \\
&= \sum_{(n,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(n,i_2,\ldots,i_n)A_{1n}A_{2i_2}\ldots A_{ni_n}
\end{aligned}$$

Ottamalla summattavista yhteiset tekijät $A_{11}, A_{12}, \ldots, A_{1n}$ tästä saadaan

$$\begin{aligned}
\det(A) &= A_{11} \sum_{(1,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(1,i_2,\ldots,i_n)A_{2i_2}\ldots A_{ni_n} \\
&= A_{12} \sum_{(2,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(2,i_2,\ldots,i_n)A_{2i_2}\ldots A_{ni_n} \\
&\quad +\cdots+ \\
&= A_{1n} \sum_{(n,i_2,\ldots,i_n)\in S_n} \operatorname{sgn}(n,i_2,\ldots,i_n)A_{2i_2}\ldots A_{ni_n}
\end{aligned}$$

Esimerkki 74. Käyttämällä Seurauksen 4 huomioita 3-rivinen determinantti voidaan laskea seuraavasti:

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

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

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

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

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

Huomautus 20. Aiemmissa esimerkeissä on käsitelty determinantin laskemista ensimmäisen rivin mukaisen kehitelmän avulla. Seurauksen (4) perusteella on kuitenkin ilmeistä, että determinantti voitaisiin laskea minkä hyvänsä rivin tai sarakkeen kehitelmän mukaisesti.

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

Esimerkki 75. $\begin{vmatrix} 1 & 2 & 0 \\ -3 & 0 & 1 \\ 4 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ -3 & 6 & 1 \\ 4 & -10 & 1 \end{vmatrix} = \begin{vmatrix} 6 & 1 \\ -10 & 1 \end{vmatrix} = 6 \cdot 1 - 1 \cdot (-10) = 16.$

Esimerkki 76.

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

Seuraava lause on erittäin tärkeä ja perustellaan luennolla:

Lause 24. *Neliömatriisilla A on olemassa käänteismatriisi A^{-1} tarkalleen silloin kun* $\det(A) \neq 0$.

Lause 25. *Jos neliömatriisilla A on käänteismatriisi A^{-1}, on* $\det(A^{-1}) = \det(A)^{-1}$.

Todistus.

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

Väite seuraa tästä välittömästi.

Huomautus 21. Muistutuksena edellisestä luvusta: Matriisilla on käänteismatriisi tarkalleen silloin kun matriisi on täysiasteinen, mikä puolestaan toteutuu tarkalleen silloin kun matriisin rivit ovat lineaarisesti riippumattomat.

Yhteenvetona voidaan esittää seuraava lause.

Lause 26. *Seuraavat ehdot ovat neliömatriisille A ekvivalentteja*

1. *Käänteismatriisi A^{-1} on olemassa.*
2. *Matriisin A rivit ovat lineaarisesti riippumattomat.*
3. *Matriisin A sarakkeet ovat lineaarisesti riippumattomat.*
4. *Matriisin A rivit muodostavat n-ulotteisen vektoriavaruuden kannan.*
5. *Matriisin A sarakkeet muodostavat n-ulotteisen vektoriavaruuden kannan.*
6. $\det(A) \neq 0$.
7. $r(A) = n$.

Matlabissa matriisin A determinantti, käänteismatriisi, ja aste voidaan laskea komennoilla `det(A)`, `inv(A)` ja `rank(A)`.

Luku 4
Ominaisarvot ja -vektorit

4.1 Matriisin ominaisarvot ja ominaisvektorit

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

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

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

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

Esimerkki 77. Kahden $n \times n$-diagonaalimatriisin tulo

$$\begin{pmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \ldots & 0 \\ 0 & b_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & 0 & \ldots & 0 \\ 0 & a_2b_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_nb_n \end{pmatrix}$$

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

Esimerkki 78. Lineaarisessa diskreettiaikaisessa systeemissä tilaa ajanhetkellä i kuvataan pystyvektorilla $\mathbf{x}_i \in \mathbb{R}^n$ ja systeemin dynamiikkaa esittää matriisi A, mikä merkitsee sitä että $\mathbf{x}_{i+1} = A\mathbf{x}_i$. Tällöin $\mathbf{x}_i = A^i \mathbf{x}_0$, mutta A^i voi olla työläs määrittää.

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

Määritelmä 36. Olkoon A $n \times n$-matriisi. Luku $\lambda \in \mathbb{C}$ on matriisin A ominaisarvo, jos on olemassa sellainen $\mathbf{x} \neq 0$, että

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

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

Huomautus 22. Nollavektorille pätee joka tapauksessa $A\mathbf{0} = \mathbf{0} = \lambda \mathbf{0}$, olipa λ mikä hyvänsä luku. Edellisessä määritelmässä vaaditaan kuitenkin, että on olemassa jokin nollasta poikkeava vektori $\mathbf{x}$, jolle $A\mathbf{x} = \lambda \mathbf{x}$ toteutuu.

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

Huomautus 23. Jos vektori $\mathbf{x}_0$ voidaan esittää matriisin A ominaisvektoreiden lineaarikombinaationa

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

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

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

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

Lause 27. *Matriisin A ominaisarvot λ ovat tarkalleen kaikki yhtälön*

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

ratkaisut.

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

Lause 28. *Olkoon A $n \times n$-matriisi. Matriisin A ominaisarvoon λ liittyvät ominaisvektorit yhdessä nollavektorin kanssa muodostavat $\mathbb{R}^n$:n aliavaruuden.*

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

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

Esimerkki 79. Määritetään matriisin $A = \begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix}$ ominaisarvot ja -vektorit. Ominaisarvoyhtälössä esiintyvä matriisi $A - \lambda I$ saa muodon

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

joten ominaisarvoyhtälö on

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

mikä voidaan kirjoittaa muotoon

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

Määritetään ominaisarvoon 2 liittyvät ominaisvektorit. Tämä merkitsee yhtälöryhmän

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

ratkaisujen etsimistä. Gaussin-Jordanin menetelmä muuntaa tämän muotoon

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

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

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

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

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

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

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

Esimerkki 80. Olkoon matriisi A kuten edellisessä esimerkissä. Selvitetään yleinen lauseke vektorille $A^i(9,5)^T$. Edellisessä esimerkissä nähtiin, että matriisin $(2,1)$ ja $(5,3)$ ovat matriisin A ominaisvektoreita. On helppo huomata, että nämä ovat lineaarisesti riippumattomat, joten ne muodostavat avaruuden $\mathbb{R}^2$ kannan. Näin ollen vektori $(9,5)$ voidaan esittää vektoreiden $(2,1)$ ja $(5,3)$ lineaarikombinaationa:

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

joka voidaan kirjoittaa yhtälöparina

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

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

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

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

Esimerkki 81. Identiteettimatriisin I ainoa ominaisarvo on 1, sillä kaikille vektoreille $I\mathbf{x} = 1 \cdot \mathbf{x}$. Toisaalta tämän mukaisesti kaikki vektorit ovat identiteettimatriisin ominaisvektoreita, joten niistä voidaan varmastikin valita vektoriavaruuden kanta.

4.2 Matriisin diagonalisointi

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

Määritelmä 37. Neliömatriisit A ja B ovat *similaarit*, mikäli on olemassa sellainen kääntyvä matriisi P, että $A = P^{-1}BP$.

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

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

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

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

Huomautus 24. Oletetaan, että $n \times n$-matriisilla A on olemassa n ominaisarvoa $\lambda_1, \ldots, \lambda_n$, joihin liittyy ominaisvektorit $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Muodostetaan matriisi P käyttämällä ominaisvektoreita $\mathbf{x}_i$ sarakkeina:

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

jolloin AP voidaan lohkomuodossa laskea seuraavasti:

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

ja suoraan laskemalla on helppo todeta, että

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

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

Määritelmä 38. Matriisi A on *diagonalisoituva*, jos se on similaari jonkin diagonaalimatriisin D kanssa. Tällöin siis $P^{-1}AP = D$ jollekin kääntyvälle matriisille P. Diagonalisoituva matriisi voidaan esittää muodossa $A = PDP^{-1}$, missä D on diagonaalinen ja P kääntyvä.

Esimerkki 82. Olkoon $A = \begin{pmatrix} -13 & 30 \\ -9 & 20 \end{pmatrix}$ kuten esimerkissä 91 todettiin, on matriisilla A ominaisarvoon 2 liittyvä ominaisvektori $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ sekä ominaisarvoon 5 liittyvä ominaisvektori $\begin{pmatrix} 5 \\ 3 \end{pmatrix}$. Näin ollen

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

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

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

ja siksi

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

4.3 Jordanin normaalimuoto

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

$$AP = (A\mathbf{x}_1 \dots A\mathbf{x}_n) = (\lambda_1 \mathbf{x}_1 \dots \lambda_n \mathbf{x}_n) = PD, \tag{4.1}$$

missä

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

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

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

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

Esimerkki 83. Matriisin

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

ominaisarvoyhtälö on

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

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

Ominaisarvoa 1 vastaavat ominaisvektorit voidaan löytää ratkaisemalla yhtälö

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

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

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

Kerroinmatriisin redusoitu porrasmuoto on

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

mitä vastaa yhtälöryhmä

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

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

Ominaisarvoon 2 liittyvät ominaisvektorit selvitetään samalla tavalla: Yhtälö

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

saa muodon

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

Kerroinmatriisin redusoitu porrasmuoto on

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

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

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

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

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

voidaan todeta että

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

ja

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

Esimerkki 84. Selvitetään matriisin

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

ominaisarvot ja ominaisvektorit. Ominaisarvoyhtälö voidaan kirjoittaa muotoon

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

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

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

Kerroinmatriisin redusoitu porrasmuoto on

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

mikä vastaa yhtälöparia

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

Jonka ratkaisut ovat muotoa $(x,y,z) = (z,z,z) = z(1,1,1)$. Näin ollen yhtälön (4.1) mukaista matriisia P ei voida valita siten että sellaisessa olisi kolme lineaarisesti riippumatonta saraketta.

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

Määritelmä 39. Muotoa

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

olevaa matriisia sanotaan Jordan-lohkoksi.

Esimerkki 85. Diagonaalimuodossa olevan matriisin potenssit on helppo esittää.

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

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

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

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

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

ja niin edelleen. Ylläolevat yhtäsuuruudet voidaan todistaa oikeiksi induktiolla.

Lause 29. *Olkoon A $n \times n$-matriisi, jonka erisuuret ominaisarvot ovat $\lambda_1, \dots, \lambda_k$. Tällöin on olemassa sellainen kääntyvä matriisi P, että $P^{-1}AP$ voidaan esittää lohkomuodossa*

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

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

Osoittautuu, että yhtälössä (4.1) esiintyvä matriisi P voidaan aina valita kääntyväksi matriisiksi, mikäli diagonaalisuusvaatimuksesta joustetaan ja vaaditaan ainoastaan matriisin D olevan Jordan-lohko -muodossa. Tällöin matriisin P sarakkeiksi tulee valita alkuperäisen matriisin A ominaisvektorit ja näitä täydentämään ns. *yleistetyt ominaisvektorit*.

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

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

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

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

Yhtäsuuruus $AP = PJ$ edellyttää siis, että

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

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

Määritelmä 40. Olkoon λ matriisin A ominaisarvo ja $k \in \mathbb{N}$. Ominaisarvoon λ liittyvä kertalukua k oleva *yleistetty ominaisvektori* on mikä hyvänsä yhtälön

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

toteuttava vektori.

Huomautus 25. Kun määritelmässä 40 valitaan $k = 1$, saadaan määritelmän 36 mukaiset tavalliset ominaisvektorit.

Esimerkki 86. Esimerkissä 84 todettiin, että matriisilla

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

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

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

ratkaisut. Nämä puolestaan ovat kaikki vektorin $(1,1,1)^T$ skalaarimonikertoja, joten yhtälön (4.1) mukaista diagonalisoivaa matriisia ei voida löytää. Tämän sijasta voidaan etsiä ominaisarvoon 2 liittyvät astetta 2 olevat yleistetyt ominaisvektorit $\mathbf{x}$, joilla tässä tapauksessa tarkoitetaan yhtälön

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

ratkaisuja. Kyseinen yhtälö voidaan kirjoittaa muotoon

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

ja yhtälöryhmän augmentoitu matriisi on

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

ja edelleen tämän redusoitu porrasmuoto

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

vastaa yhtälöparia

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

jonka ratkaisut voidaan kirjoittaa muotoon

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

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

Koska yhtälön (4.1) matriisia P varten tarvitaan 3 lineaarisesti riippumatonta saraketta, etsitään vielä astetta 3 oleva yleistetty ominaisvektori, jolla tässä tapauksessa tarkoitetaan yhtälön

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

ratkaisuja. Tämä yhtälö voidaan kirjoittaa muotoon

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

Kyseisen yhtälöryhmän augmentoitu matriisi on

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

ja tämän redusoitu porrasmuoto

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

vastaa yhtälöparia

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

jonka ratkaisut voidaan kirjoittaa muotoon

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

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

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

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

nähdään että

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

ja

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

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

Esimerkki 87. Olkoon

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

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

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

$$\begin{cases} x + \frac{1}{3}y + \frac{2}{3}z = 0 \\ \qquad\qquad 0 = 0 \\ \qquad\qquad 0 = 0 \end{cases}. \tag{4.2}$$

Näin ollen matriisin M kaikki ominaisvektorit ovat muotoa

$$(x,y,z) = (-\frac{1}{3}y - \frac{2}{3}z, y, z) = y(-\frac{1}{3}, 1, 0) + z(-\frac{2}{3}, 0, 1). \tag{4.3}$$

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

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

$$(M - 2I)\mathbf{x} = a(-1,3,0)^T + b(-2,0,3)^T \tag{4.4}$$

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

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

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

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

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

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

Huomautus 26. Muodollisesti ottaen matriisissa J on kaksi Jordan-lohkoa, vasemman yläkulman 1×1-lohko 2 ja oikean alakulman 2×2-lohko $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. Voidaan myös havaita, että jos (yleistettyjen) ominaisvektoreiden järjestystä matriisin P sarakkeina vaihdetaan, voidaan matriisin J muotoa muuttaa yläkolmiomatriisista alakolmiomatriisiksi. Tässä esityksessä priorisoidaan järjestystä, jossa P:n sarakkeet järjestetään vasemmalta oikealle kasvavan kertaluvun mukaan.

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

Seuraavaa esimerkkiä varten otetaan käyttöön määritelmä ja merkintöjä.

Määritelmä 41. Ominaisarvon λ *algebrallinen kertaluku* tarkoittaa λ:n kertalukua itseisarvopolynomin $p(x) = \det(A - xI)$ nollakohtana. Toisin sanoen, λ:n algebrallinen kertaluku ilmaisee kuinka moninkertainen juuri λ on polynomille $p(x)$.

Määritelmä 42. Olkoon λ $n \times n$-matriisin ominaisarvo. Ominaisarvoon λ liittyvien, kertalukua k olevien ominaisvektoreiden joukosta käytetään merkintää

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

Huomautus 27. Määritelmän perusteella on selvää, että jokainen E_λ^k on avaruuden $\mathbb{C}^n$ aliavaruus. Lisäksi on helppo havaita, että $E_\lambda^k \subseteq E_\lambda^{k+1}$ (miksi?), joten ominaisarvoon λ liittyvä aliavaruus laajenee (tai pysyy ennallaan), kun kertalukua kasvatetaan. Tämän havainnon perusteella onkin uskottavaa, että (yleistettyjen) ominaisvektoreiden etsiminen saattaa kannattaa aloittaa korkeimmasta kertaluvusta alkaen.

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

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

Seuraavissa esimerkeissä valaistaan ylläkuvattua menetelmää.

Esimerkki 88. Käydään läpi esimerkin 87 normaalimuodon etsiminen uudelleen lähtien korkeamman kertaluvun ominaisvektoreista. Ominaisarvoyhtälön perusteella ominaisarvon $\lambda = 2$ algebrallinen kertaluku on 3. Yhtälön (4.3) perusteella taas avaruuden E_2^1 dimensio on kaksi, joten tämän jälkeen tulee selvitää avaruuden E_2^2 rakenne. Koska $(M - 2I)^2 = 0$, on selvästikin $E_2^2 = \mathbb{C}^3$ ja ensiksi tulee siis valita $2 - 1 = 1$ sellaista kantavektoria, jotka kuuluvat joukkoon $E_2^2 \setminus E_2^1$.

Yhtälön (4.3) perusteella esimerkiksi $(1, 0, 0)$ kelpaa. Menetelmän mukaisesti seuraavaksi tulee valita $(M - 2I)(1, 0, 0) = (-3, 3, 3) \in E_2^1$, ja koska $\dim(E_2^1) = 2$, pitää valita vielä jokin toinen, riippumaton vektori. Tällaiseksi kelpaa esimerkiksi $(-1, 3, 0)^T$ tai yhtä hyvin $(-2, 0, 3)^T$.

Esimerkki 89. Olkoon

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

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

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

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

ratkaisut, jotka voidaan eksplisiittisesti selvittää muodosta

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

joka puolestaan voidaan Gaussin-Jordanin menetelmää käyttäen kirjoittaa ekvivalenttiin asuun

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

Tämän ratkaisut ovat muotoa

$$(x,y,z) = (2y,y,0) = y(2,1,0), \tag{4.5}$$

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

Avaruuden E_1^2 ja sen dimension määrittämiseksi pitää selvittää yhtälön

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

ratkaisujoukko. Eksiplisiittisesti ilmaistuna yläoleva yhtälö on muotoa

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

mikä Gaussin-Jordanin menetelmällä saadaan muotoon

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

Tämän ratkaisut puolestaan ovat muotoa

$$(x,y,z) = (2y-3z,y,z) = (2y,y,0) + (-3z,0,z) = y(2,1,0) + z(-3,0,1), \tag{4.6}$$

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

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

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

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

Koska aiemmin jo todettiin, että $\dim(E_1^3) = 3$ ja $\dim(E_1^2) = 2$, tulee ensiksi valita $3-2=1$ vektori(a), jotka eivät kuulu avaruuteen E_1^2. Yhtälön (4.6) perusteella tällaiseksi voidaan valita mikä hyvänsä vektori, joka ei ole muotoa $(2y-3z,y,z)$, vaikkapa $(1,0,0)$.

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

Kun järjestetään ominaisvektorit nousevan kertaluvun mukaiseen järjestykseen, saadaan Jordanmuodon välittäväksi matriisiksi

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

ja voidaan havaita, että

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

4.4 Reaalifunktioiden laajennuksia matriiseille

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

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

Muiden kuin polynomifunktioiden laajentaminen neliömatriiseile kannattaakin aloittaa polynomifunktioiden suoraviivaisesta yleistyksestä, potenssisarjoista (käsitellään jonkin verran kurssilla Insinöörimatematiikka: Differentiaali- ja integraalilaskenta). Mikäli sarja

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \tag{4.7}$$

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

$$f(A) = \sum_{k=0}^{\infty} c_k A^k. \tag{4.8}$$

Tällöin kuitenkin joudutaan pohtimaan miten sarja (4.8) matriiseille määritellään. Vaikka tämä voidaan tehdä analogisesti reaalisten potenssisarjojen kanssa, joudutaan silti miettimään mitä käsite raja-arvo matriiseille merkitsee. Tällä kurssilla ei näihin pohdintoihin kuitenkaan pureuduta, vaan asia käsitellään lähinnä esimerkkien kautta.

Esimerkki 90. Diagonaalimatriisille

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

potenssit ja niinikään summan (4.8) alkuosa on täysin suoraviivainen.

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

Jos kaikki diagonaalialkiot $\lambda_1, \ldots, \lambda_n$ kuuluvat sarjan (4.7) suppenemisalueeseen, on edellämainitun perusteella täysin luontevaa määritellä

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

Edellisissä luvuissa on jo pohjustettu mahdollisuuksia siinä tapauksessa, että matriisi A ei ole diagonaalinen. Jos matriisi A kuitenkin on diagonalisoituva, on olemassa määritelmän 38 mukainen matriisi P ja diagonaalimatriisi D, joille pätee $A = PDP^{-1}$. Tällöin selvästikin $A^k = PD^kP^{-1}$, joten funktio f on luontevaa määritellä diagonaaliesityksen avulla seuraavasti:

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

Esimerkki 91. Määritetään

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

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

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

Koska sinifunktion määrittelevä sarja suppenee koko reaaliakselilla, voidaan määrittää

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

Esimerkki 92. Kvanttifysiikassa esiintyy ns. Paulin spin-matriisi

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

Määritetään $e^{-it\sigma_x}$.

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

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

saadaan yhtälö

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

jonka avulla voidaan määrittää

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

Viimeisin yhtäsuuruus perustuu Eulerin kaavaan.

Mahdollisuus laajentaa potenssisarjan määrittelemä reaalifunktio sellaisille matriiseille, jotka eivät ole diagonalisoituvia, perustuu viime kädessä Jordanin normaalimuotoon: Jordanin normaalimuodolle voidaan nimittäin hyvin suoraviivasesti laskea potenssit Esimerkin 85 mukaisesti ja siten laajentaa potenssisarjan määrittely Jordan-muotoisille matriiseille.

Esimerkki 93. Tarkastellaan miten sarjan (4.7) määrittelemä funktio $f(x) = \sum_{k=0}^{\infty} c_k x^k$ voidaan laajentaa Esimerkissä 85 esiintyvälle 3×3 -Jordan Jordan-lohkolle

$$J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}. \tag{4.9}$$

Esimerkin 85 mukaisesti

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

ja siksi sarjaesityksen suoraviivainen laajentaminen antaa määritelmän

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

Ylläolevassa matriisissa diagonaalialkiot selvästikin esittävät funktion arvoa $f(\lambda)$, joten selvitettäväksi jää vielä ylimmän rivin toinen ja kolmas alkio. Näihin voidaan pureuta differentiaalilaskennan keinoin, joskin tarkempi tarkastelu sivuutetaan. Jos siis

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

on

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

ja

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

Näin ollen sarjan (4.7) määrittelemän reaalifunktion laajennus Jordan-lohkolle (4.9) pitäisi tehdä seuraavasti:

$$f(J) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) \\ 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & f(\lambda) \end{pmatrix}. \tag{4.10}$$

Yleistys Esimerkin 85 mukaiselle 4×4 Jordan-lohkolle on suoraviivainen; potenssisarjan määrittelemä funktio tulee 4×4-jordan lohkolle

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

määritellä seuraavasti:

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

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

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

Esimerkki 94. Olkoon

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

Edellisen perusteella funktio $f(x) = e^{tx}$ arvolla J määritetään seuraavasti: Lasketaan $f'(x) = te^{tx}$ ja käytetään 4.10 mukaista kaavaa, josta saadaan

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

Esimerkki 95. Määritetään e^{tJ} 3×3-Jordan-lohkolle

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

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

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

Luku 5
Vektoriavaruuksien geometriaa

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

5.1 Etäisyys, normi, sisätulo

Määritelmä 43. Olkoon V vektoriavaruus. *Etäisyys* on funktio $V \times V \to \mathbb{R}$, joka toteuttaa seuraavat ehdot:

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

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

Määritelmä 44. Olkoon V vektoriavaruus. Vektoriavaruuden *normi* on kuvaus $V \to \mathbb{R}$, $\mathbf{x} \mapsto ||\mathbf{x}||$, joka toteuttaa seuraavat ehdot:

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

Huomautus 28. Jos $||\mathbf{x}||$ on mikä hyvänsä ylläolevan määritelmän ehdot toteuttava normi, saadaan sen avulla määriteltyä etäisyysfunktio $d(\mathbf{x},\mathbf{y}) = ||\mathbf{x}-\mathbf{y}||$. Tällöin sanotaan, että normi indusoi etäisyyden.

Esimerkki 96. Jos $V = \mathbb{R}^n$ tai $V = \mathbb{C}^n$, voidaan vektorille $\mathbf{x} = (x_1, \ldots, x_n)$ määritellä

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

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

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

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

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

Määritelmä 45. Vektoreiden $\mathbf{x} = (x_1,\ldots,x_n)$ ja $\mathbf{y} = (y_1,\ldots,y_n) \in \mathbb{R}^n$ *pistetulo* määritellään $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + \ldots + x_ny_n$.

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

Esimerkki 97. Välilä $[a,b]$ jatkuvien reaalifunktioiden joukossa $C^0[a,b]$ voidaan määritellä sisätulo ehdolla

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

Kompleksisille vektoriavaruuksille sisätulo on kuitenkin määriteltävä toisin.

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

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

Määritelmä 46. Vektoreiden $\mathbf{x} = (x_1,\ldots,x_n)$ ja $\mathbf{y} = (y_1,\ldots,y_n) \in \mathbb{C}^n$ *Hermiten pistetulo* määritellään $\mathbf{x} \cdot \mathbf{y} = x_1\overline{y_1} + \ldots + x_n\overline{y_n}$.

Esimerkki 98. Olkoon $L^2[a,b] \cap C^0[a,b]$ välillä $[a,b]$ määriteltyjen jatkuvien (kompleksiarvoisten) funktioiden joukko, joille $\int_a^b |f|^2$ on olemassa. Tällöin

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

on Hermiten sisätulo.

Lause 30. *Sisätulo toteuttaa* Cauchyn-Schwarzin epäyhtälön

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

Lause 31. *Jos* $(\mathbf{x},\mathbf{y})$ *on sisätulo, saadaan määrittelemällä*

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

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

Huomautus 29. Jos normi määritellään käyttämällä pistetuloa, saadaan

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

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

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

Määritelmä 47. Vektoreiden $\mathbf{x}$ ja $\mathbf{y} \in \mathbb{R}^n$ välinen *kulma* tarkoittaa kulmaa $\theta \in [0,\pi]$, joka muodostuu origosta (pisteestä $\mathbf{0}$) pisteeseen $\mathbf{x}$ ja origosta pisteeseen $\mathbf{y}$ kulkevien janojen välille. Huomaa, että kolme pistettä $\mathbf{x}$, $\mathbf{y}$ ja $\mathbf{0}$ kuuluvat aina samalle tasolle, joten vektoreiden välinen kulma on aina *tason* ominaisuus.

Lause 32. *Jos* θ *on vektoreiden* $\mathbf{x}$ *ja* $\mathbf{y}$ *välinen kulma, niin*

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

Todistus. Tason kolmioita koskevasta kosinilauseesta saadaan

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

Yhtälön oikea puoli voidaan edelleen kirjoittaa muotoon

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

Näistä saadaan yhtälö

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

mistä väite seuraa suoraan.

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

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

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

Esimerkki 99. Olkoon $\mathbf{x} = (1,0,2,-1)$ ja $\mathbf{y} = (0,1,1,-1) \in \mathbb{R}^4$. Tällöin $||\mathbf{x}||^2 = 1^2+0^2+2^2+(-1)^2 = 6$ ja $||\mathbf{y}||^2 = 0^2+1^2+1^2+(-1)^2 = 3$, sekä $\mathbf{x}\cdot\mathbf{y} = 1\cdot 0+0\cdot 1+2\cdot 1+(-1)\cdot(-1) = 3$. Tällöin siis vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välisen kulman θ kosini on

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

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

Määritelmä 48. Vektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat *ortogonaalit* eli *kohtisuorat* jos $(\mathbf{x},\mathbf{y}) = 0$.

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

Määritelmä 49. Vektorijoukko $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ on *ortogonaali*, jos $(\mathbf{x}_i,\mathbf{x}_j) = 0$ aina kun $i \ne j$. Ortogonaali joukko $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ on *ortonormaali*, jos lisäksi $(\mathbf{x}_i,\mathbf{x}_i) = 1$.

Huomautus 31. Jokaisesta ortogonaalista joukosta joka ei sisällä nollavektoria voidaan helposti saada ortonormaali kertomalla kukin $\mathbf{x}_i$ norminsa käänteisluvulla $\sqrt{(\mathbf{x}_i,\mathbf{x}_i)}^{-1}$.

Esimerkki 100. $\mathbb{R}^n$:n luonnollinen kanta $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ on ortonormaali.

Lause 33. *Olkoon V vektoriavaruus ja $A = \{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ joukko, joka ei sisällä nollavektoria ja on ortogonaali jonkin sisätulon $(\mathbf{x},\mathbf{y})$ suhteen. Tällöin A on lineaarisesti riippumaton.*

Todistus. Oletetaan, että

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

Laskemalla sisätulot puolittain saadaan

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

Ortogonaalisuuden perusteella ylläoleva sievenee muotoon

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

josta $c_i = 0$.

Esimerkki 101. Olkoon $L^2[\alpha,\alpha+T]\cap C^0[\alpha,\alpha+T]$ kuten aiemmin ja

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

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

Määritelmä 50. Olkoot $\mathbf{x}$ ja $\mathbf{y}\neq\mathbf{0}$ avaruuden $\mathbb{R}^n$ vektoreita. Tällöin *vektorin* $\mathbf{x}$ *projektio vektorilla* $\mathbf{y}$ tarkoittaa sitä origon ja pisteen $\mathbf{y}$ kautta kulkevan suoran $L(\mathbf{y})$ pistettä $\mathbf{x}'$, jolle $\mathbf{x}-\mathbf{x}'$ on kohtisuorassa vektoria $\mathbf{y}$ vastaan.

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

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

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

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

5.2 Etäisyyden säilyttävät kuvaukset

Määritelmä 51. Lineaarikuvaus $f:V\to V$ säilyttää etäisyyden, jos $d(f(\mathbf{x}),f(\mathbf{y}))=d(\mathbf{x},\mathbf{y})\ \forall\mathbf{x},\mathbf{y}\in V$. Lineaarikuvaus $f:V\to V$ säilyttää normin, jos $||f(\mathbf{x})||=||\mathbf{x}||$. Lineaarikuvaus $f:V\to V$ säilyttää sisätulon, jos $(f(\mathbf{x}),f(\mathbf{y}))=(\mathbf{x},\mathbf{y})$.

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

Lause 34. *Jos lineaarikuvaus f säilyttää sisätulon, niin se säilyttää myös normin ja etäisyyden.*

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

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

Määritelmä 52. $n\times n$-matriisi A on *ortogonaali*, jos $A^T=A^{-1}$, siis $A^TA=AA^T=I$.

Lause 35. *Jos lineaarikuvauksen f matriisi A_f on ortogonaali, niin f säilyttää pistetulon ja etäisyyden.*

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

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

Määritelmä 53. $n\times n$-matriisin A adjugaatti A^* määritellään $A^*=\overline{A}^T$. Sanotaan, että matriisi A on *unitaarinen*, jos $A^*=A^{-1}$, siis $A^*A=AA^*=I$.

Huomautus 32. Matriisi

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

on unitaarinen, sillä

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

kuten esimerkissä 54 todettiin.

Lause 36. *Jos lineaarikuvauksen $f : \mathbb{C}^n \to \mathbb{C}^n$ matriisi A_f on unitaarinen, niin f säilyttää Hermiten pistetulon.*

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

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

Huomautus 33. Olkoot $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})^T$ ja $\mathbf{g} = (g_0, g_1, \dots, g_{N-1})^T \in \mathbb{C}^n$ datavektoreita ja $\mathbf{F} = F\mathbf{f}$ sekä $\mathbf{G} = F\mathbf{g}$ niiden Fourier-muunnokset. Yllämainitun perusteella $(\mathbf{F}, \mathbf{G}) = (\mathbf{f}, \mathbf{g})$, mikä voidaan kirjoittaa auki muotoon

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

Tämä on *Parsevalin kaava* diskreeteille Fourier-muunnoksille.

Huomautus 34. Etäisyyden säilyttävät lineaarikuvaukset ovat avaruuden peilauksia tai kiertoja. Diskreetti Fourier-muunnos voidaan ymmärtää N-ulotteisen kompleksisen avaruuden kiertona tai vaihtoehtoisesti mutta matemaattisesti yhtäpitävästi kannan vaihtona.

5.3 Korrelaatiokerroin

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

Esimerkki 102. Jos pyritään selvittämään, millä tavoin jossakin populaatiossa ihmisen pituus (keskimäärin) riippuu painosta, voidaan toimia seuraavasti: valitaan populaatiosta joukko koehenkilöitä, joiden painot x_1, …, x_n ja pituudet y_1, …, y_n mitataan. Tämän jälkeen voidaan pisteet (x_1, y_1), …, (x_n, y_n) piirtää tasolle, ja katsoa onko kuviosta mahdollisesti löydettävissä jonkinlaista johdonmukaisuutta. Mitä paremmin pisteet (x_i, y_i) näyttävät keskittyvän jonkin nousevan suoran ympärille, sitä paremmin voidaan katsoa pituuden riippuvan painosta (lineaarisesti).

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

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

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

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

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

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

ja samoin vektorille $\mathbf{y}$:

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

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

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

ja

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

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

Määritelmä 54. Vektoreiden $\mathbf{x} = (x_1, \ldots, x_n)$ ja $\mathbf{y} = (y_1, \ldots, y_n)$ välinen *Pearsonin korrelaatiokerroin* (tunnetaan myös nimellä *tulomomenttikerroin*) on vektoreiden $\mathbf{x}'$ ja $\mathbf{y}'$ välisen kulman kosini, missä $\mathbf{x}'$ ja $\mathbf{y}'$ on saatu vektoreista $\mathbf{x}$ ja $\mathbf{y}$ vähentämällä jokaisesta koordinaatista koordinaattien keskiarvo.

Esimerkki 103. Taulukossa on eräälle kurssille osallistuneen 36 opiskelijan demoprosentit ja heidän tentissä saamansa pisteet.

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

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

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

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

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

Samoin voidaan laskea

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

ja

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

ja lopuksi korrelaatiokerroin saadaan lausekkeesta

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

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

Huomautus 35. Havaintotapausten määrän on oltava kyllin suuri, jotta korrelaatiokertoimen tulkinnalla olisi merkitystä. Jos esimerkiksi aineistossa on vain kaksi tapausta, joita vastaavat (siiretyt) vektorit $\mathbf{x}' = (x, -x)$ ja $\mathbf{y}' = (y, -y) \in \mathbb{R}^2$, on näillä aina täydellinen korrelaatio joko 1 tai -1. Yleensä halutaan, että havaintotapausten määrä on vähintään kymmenen, mutta tämäkin on tulkinnanvaraista. Joissakin tilanteissa voidaan kolmeakymmentä havaintotapausta pitää suurena määränä, kun taas esimerkiksi lääkkeen haittavaikutuksia etsittäessä kolmeakymmentä havaintotapausta pidetään aivan riittämättömänä joukkona.

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

Luku 6
Kolmiulotteisen avaruuden geometriaa

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

6.1 Ristitulo

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

Määritelmä 55. Avaruuden $\mathbb{R}^3$ vektoreiden $\mathbf{x} = (x_1, x_2, x_3)$ ja $\mathbf{y} = (y_1, y_2, y_3)$ *ristitulo* on $\mathbb{R}^3$:n vektori

$$\mathbf{x} \times \mathbf{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

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

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \tag{6.1}$$

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

Tästä huolimatta (6.1) voi toimia muistisääntönä:

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

Lause 37 (Skalaarikolmitulo). *Olkoot* $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$ *ja* $\mathbf{z} = (z_1, z_2, z_3)$. *Tällöin*

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

2. $\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) = \mathbf{y} \cdot (\mathbf{z} \times \mathbf{x}) = \mathbf{z} \cdot (\mathbf{x} \times \mathbf{y})$

Todistus. Ensimmäinen kohta saadaan piste- ja ristitulojen määritelmästä suoraan laskemalla, toinen kohta taas seuraa ensimmäisestä, sillä determinantin merkki muuttuu kun kahden rivin paikkaa vaihdetaan (mieti kuinka monta kahden rivin vaihtoa tarvitaan, kun determinantti $\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$ muutetaan determinantiksi $\begin{vmatrix} y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \\ x_1 & x_2 & x_3 \end{vmatrix}$).

Huomautus 36. Skalaarikolmitulo voidaan tulkita sellaisen suuntaissärmiön tilavuutena, jonka yksi kärjistä sijaitsee origossa, ja kolme sivua ovat vektoreiden $\mathbf{x}$, $\mathbf{y}$ ja $\mathbf{z}$ suuntaiset.

Lause 38. *Ristitulo toteuttaa seuraavat säännöt:*

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

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

$$\begin{aligned}
||\mathbf{x} \times \mathbf{y}||^2 + (\mathbf{x} \cdot \mathbf{y})^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\
&+ (x_1y_1 + x_2y_2 + x_3y_3)^2 \\
&= x_2^2y_3^2 - 2x_2x_3y_2y_2 + x_3^2y_2^2 \\
&+ x_3^2y_1^2 - 2x_1x_3y_1y_3 + x_1^2y_3^2 \\
&+ x_1^2y_2^2 - 2x_1x_2y_1y_2 + x_2^2y_1^2 \\
&+ x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2 \\
&+ 2x_1x_2y_1y_2 + 2x_1x_3y_1y_3 + 2x_2x_3y_2y_3 \\
&= x_1^2y_1^2 + x_1^2y_2^2 + x_1^2y_3^2 \\
&+ x_2^2y_1^2 + x_2^2y_2^2 + x_2^2y_2^2 \\
&+ x_3^2y_1^2 + x_3^2y_2^2 + x_3^2y_3^2 \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) = ||\mathbf{x}||^2\, ||\mathbf{y}||^2
\end{aligned}$$

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

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

mistä väittämä 7 seuraa suoraan.

6.2 Tasot

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

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

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

Määritelmä 56. Olkoon V jokin avaruuden $\mathbb{R}^n$ aliavaruus ja $\mathbf{z} \in \mathbb{R}^n$ jokin vektori. Tällöin joukkoa
$$\mathbf{z}+V = \{\mathbf{z}+\mathbf{w} \mid \mathbf{w} \in V\}$$
kutsutaan vektorin $\mathbf{z}$ määräämäksi aliavaruuden *sivuluokaksi*.

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

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

Esimerkki 104. Jos $\mathbf{x} \in \mathbb{R}^3$ ei ole nollavektori, on joukko $\{\mathbf{x}\}$ lineaarisesti riippumaton ja geometrisesti yksiulotteinen aliavaruus
$$L(\mathbf{x}) = \{c\mathbf{x} \mid c \in \mathbb{R}\} \subseteq \mathbb{R}^3$$
merkitsee origon kautta kulkevaa suoraa. Avaruuden $L = L(\mathbf{x})$ sivuluokka
$$\mathbf{z}+L = \{\mathbf{z}+c\mathbf{x} \mid c \in \mathbb{R}\} \tag{6.2}$$
puolestaan vastaa L:n suuntaista suoraa, joka kulkee pisteen $\mathbf{z}$ kautta. Avaruuden $\mathbb{R}^3$ kaikki suorat voidaan esittää yksiulotteisten aliavaruuksien sivuluokkina ja jatkossa tutustutaan myös muihin suorien esitystapoihin.

Korkeampiulotteisissa avaruuksissa $\mathbb{R}^n$ suora määritellään joukkona (6.2), missä $\mathbf{x} \in \mathbb{R}^n$ on nollavektorista eroava vektori ja $\mathbf{z} \in \mathbb{R}^n$ on jokin vektori. Kuten tasojen tapauksessa, on selvää, että suora avaruudessa $\mathbb{R}^n$ on oleellisesti samankaltainen objekti kuin suora avaruudessa $\mathbb{R}^3$.

Aiemmin on todettu, että avaruuden $\mathbb{R}^3$ *taso* vastaa kaksiulotteisen aliavaruuden T sivuluokkaa
$$\mathbf{r}+T = \{\mathbf{r}+\mathbf{w} \mid \mathbf{w} \in T\},$$
ja kaksiulotteisen aliavaruuden T (jota geometrisesti vastaa origon kautta kulkeva taso) puolestaan generoi mikä hyvänsä lineaarisesti riippumaton vektoripari $\{\mathbf{x},\mathbf{y}\}$:
$$T = \{c_1\mathbf{x}+c_2\mathbf{y} \mid c_1,c_2 \in \mathbb{R}\}. \tag{6.3}$$
Esitysmuodolla (6.3) saadaan origon kautta kulkevat tasot. Kaikki tasot saadaan seuraavasti:

Tason parametriesitys: Olkoot $\mathbf{r}$, $\mathbf{x}$ ja $\mathbf{y}$ avaruuden $\mathbb{R}^3$ vektoreita. Tällöin esitystä
$$T = \{\mathbf{r}+c_1\mathbf{x}+c_2\mathbf{y} \mid c_1,c_2 \in \mathbb{R}\}, \tag{6.4}$$
kutsutaan tason T *parametriesitykseksi*. Lukuja c_1 ja c_2 kutsutaan *parametreiksi*. Yhtä hyvin paria $(c_1,c_2) \in \mathbb{R}^2$ voidaan kutsua parametriksi. Kun parametri (c_1,c_2) käy läpi avaruuden $\mathbb{R}^2$, käy $\mathbf{r}+c_1\mathbf{x}+c_2\mathbf{y}$ läpi kaikki tason T pisteet. Vektoria $\mathbf{r}$ kutsutaan tason *paikkavektoriksi* ja vektoreita $\mathbf{x}$ ja $\mathbf{y}$ sen *suuntavektoreiksi*. Jos $\mathbf{x}$ ja $\mathbf{y}$ ovat lineaarisesti riippuvat, ei (6.4) määrittele tasoa, vaan suoran tai pelkästään yhden pisteen.

Huomautus 37. Parametriesityksen (6.4) avulla voidaan määritellä taso yleisestikin avaruudessa $\mathbb{R}^n$, vaikka n olisi kuinka suuri tahansa. Sen sijaan muut tässä pykälässä esitettävät muodot tasolle koskevat vain kolmiulotteista avaruutta. Parametriesitys voidaan myös tulkita funktioksi $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(c_1, c_2) = \mathbf{r} + c_1\mathbf{x} + c_2\mathbf{y}$, jolloin tason $\mathbb{R}^3$:ssa saadaan $\mathbb{R}^2$:n kuvana.

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

$$T = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{x} \cdot \mathbf{n} = 0\}, \tag{6.5}$$

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

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

Tason normaalimuoto: Esitystä

$$T = \{\mathbf{x}_0 + \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^3, \mathbf{x} \cdot \mathbf{n} = \mathbf{0}\} = \{\mathbf{x} \in \mathbb{R}^3 \mid (\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{n} = 0\} \tag{6.6}$$

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

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

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

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

Tason koordinaattimuoto: Muotoa

$$T = \{(x, y, x) \mid ax + by + cz = d\}, \tag{6.7}$$

kutsutaan tason *koordinaattimuodoksi* (toisinaan myös normaalimuodoksi).

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

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

Tarkastellaan seuraavaksi esimerkkien kautta miten eri esitysmuodoista päästään toiseen.

Esimerkki 105 (parametrimuodosta normaalimuotoon). Olkoon $\mathbf{r} = (1, 0, 3)$ tason paikkavektori ja $\mathbf{x} = (-1, 2, 2)$ ja $\mathbf{y} = (-3, 2, 1)$ sen suuntavektorit. Tarkastetaan aluksi, että $\mathbf{x}$ ja $\mathbf{y}$ ovat lineaarisesti riippumattomat. Jos

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

saadaan yhtälöryhmä

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

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

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

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

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

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

$$(-2,-5,4)\cdot((x,y,z)-(1,0,3))=0. \tag{6.8}$$

Esimerkki 106 (normaalimuodosta koordinaattimuotoon). Esituksen (6.8) muuntaminen koordinaattimuotoon on suoraviivaista, sillä koordinaattimuoto ja normaalimuoto ovat lähes samat. Yhtälössä (6.8) tarvitsee vain laskea sisätulot, jolloin saadaan

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

siis koordinattimuoto on

$$-2x-5y+4z=10. \tag{6.9}$$

Huomautus 38. Koordinaattimuodon kertoimista voidaan suoraan lukea tason normaalivektori.

Esimerkki 107 (koordinaattimuodosta kolmeen pisteeseen). Jos tasolle tunnetaan koordinaattimuoto (106), voidaan helposti valita tasolta kolme pistettä. Valitaan esimerkiksi ensin $x=y=0$, jolloin yhtälön (106) mukaan $z=\frac{5}{2}$. Täten esimerkiksi $\mathbf{x}_1=(0,0,\frac{5}{2})$ kuuluu tasolle. Samoin valitsemalla $x=z=0$ nähdään, että $\mathbf{x}_2=(0,-2,0)$ kuuluu tasolle ja lopuksi valitsemalla $y=z=0$ saadaan tason piste $\mathbf{x}_3=(-5,0,0)$.

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

Huomautus 39. Erityisesti tason paremetriesityksestä on helppo löytää kolme epäkollineaarista pistettä. Voidaan nimittäin valita esimerkiksi molemmat parametrit ensin nolliksi, sen jälkeen vuoron perään toinen ykköseksi ja toinen nollaksi.

Esimerkki 108 (kolmesta pisteestä parametrimuoto). Etsitään parametrimuoto tasolle, jonka määrittävät edellisen esimerkin pisteet $\mathbf{x}_1$, $\mathbf{x}_2$ ja $\mathbf{x}_3$. On mahdollista valita mikä hyvänsä näistä pisteistä paikkavektoriksi, esim $\mathbf{r}=\mathbf{x}_1=(0,0,\frac{5}{2})$, jolloin suuntavektoreiksi voidaan valita $\mathbf{x}=\mathbf{x}_2-\mathbf{x}_1=(0,-2,-\frac{5}{2})$ ja $\mathbf{x}_3-\mathbf{x}_1=(-5,0,-\frac{5}{2})$.

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

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

Esimerkki 109. Selvitetään, onko esimerkissä (108) saatu parametrimuotoinen taso todella sama kuin esimerkin (105) taso. Tätä varten etsitään molemmille tasoille koordinaattimuoto. Esimerkin (105) koordinaattimuoto on jo selvitetty esimerkissä (6.9): koordinaattiyhtälö

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

saadaan yksikäsitteiseen muotoon jakamalla se luvulla 10:

$$-\frac{1}{5}x-\frac{1}{2}y+\frac{2}{5}z=1. \tag{6.10}$$

Toisen tason koordinaattimuotoa varten määrätään ensin normaalivektori:

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

Normaalimuodoksi saadaan siis

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

mistä pistetulot laskemalla saadaan

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

Jakamalla luvulla -25 saadaan

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

sama kuin (6.10).

Esimerkki 110. Tarkastellaan, millä ehdoilla joukko

$$\{(ax+by+c, dx+ey+f, gx+hy+i) \mid x,y \in \mathbb{R}\} \tag{6.11}$$

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

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

joten joukko (6.11) määrittelee tason tarkalleen silloin kun vektorit (a,d,g) ja (b,e,h) ovat lineaarisesti riippumattomat. Jos (a,d,g) ja (b,e,h) ovat lineaarisesti riippuvat, määrittelee (6.11) suoran (mikäli $(a,d,g) \neq (0,0,0)$).

Määritelmä 57. Tasojen L_1 ja L_2 välisellä kulmalla tarkoitetaan tasojen normaalivektoreiden välistä kulmaa.

Esimerkki 111. Olkoon taso L_1 määritelty yhtälöllä $13x+5y-7z=67$ ja L_2 yhtälöllä $2x-2y+3z=-5$. Tällöin tasojen L_1 ja L_2 normaalivektoreiksi voidaan valita $\mathbf{n}_1 = (13,5,-7)$ ja $\mathbf{n}_2 = (2,-2,3)$. Tasojen L_1 ja L_2 väliseksi kulma θ voidaan täten määrittää seuraavasti:

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

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

6.3 Suorat

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

$$L = \{c\mathbf{x} \mid c \in \mathbb{R}\}, \tag{6.12}$$

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

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

$$L = \{\mathbf{r} + c\mathbf{x} \mid c \in \mathbb{R}\} \tag{6.13}$$

kutsutaan suoran *parametrimuodoksi* ja lukua *c parametriksi*. Jos $\mathbf{x} = \mathbf{0}$, ei esitys (6.13) määrittele suoraa, vaan yhden pisteen.

Huomautus 40. Kuten taso $\mathbb{R}^3$:ssa, voidaan myös suora $\mathbb{R}^3$:ssa käsittää alempiulotteisen avaruuden kuvana ”lähes” lineaarisessa kuvauksessa. Jos nimittäin määritellään $f : \mathbb{R} \to \mathbb{R}^3$ ehdolla $f(c) = \mathbf{r} + c\mathbf{x}$, saadaan suora (6.13) avaruuden $\mathbb{R}$ kuvana.

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

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

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

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

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

Suoran koordinaattimuoto:

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

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

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

Ilmeinen tapa määritellä suora on myös erittäin käyttökelpoinen:

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

Toisinaan voidaan suora esittää myös seuraavalla tavalla:

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

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

Esimerkki 112 (Parametrimuodosta koordinaattimuotoon). Olkoon $\mathbf{r} = (1,2,2)$ suoran paikkavektori ja $\mathbf{x} = (2,-3,1)$ sen suuntavektori. Tällöin suoralle

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

saadaan koordinaattimuodoksi merkitsemällä

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

ja ratkaisemalla tästä t kunkin koordinaatin lausekkeena:

$$t = \frac{x-1}{2} = \frac{y-2}{-3} = z-2 \tag{6.14}$$

Huomautus 41. Koordinaattimuodosta voidaan helposti lukea suoran paikkavektori (osoittajien vakiot) ja suuntavektori (nimittäjien vakiot).

Esimerkki 113 (Koordinaattimuodosta kaksi pistettä). Yhtälöissä (6.14) voidaan valita t:lle mikä hyvänsä arvo, esim. $t=0$ tai $t=1$. Vastaava piste (x,y,z) löytyy ratkaisemalla 1. asteen yhtälö. Jos esimerkiksi $t=0$ yhtälössä (6.14), on $(x,y,z)=(1,2,2)$ ja arvo $t=1$ antaa $(x,y,z)=(3,-1,3)$.

Huomautus 42. Parametrimuodosta kahden eri pisteen löytäminen on vielä yksinkertaisempaa: riittää valita parametrin arvoiksi esim 0 ja 1.

Esimerkki 114 (Kahdesta pisteestä parametrimuotoon). Määritetään pisteiden $\mathbf{x}_1=(3,-1,3)$ ja $\mathbf{x}_2=(1,2,2)$ kautta kulkevan suoran parametrimuoto. Suuntavektoriksi voidaan valita $\mathbf{x}=\mathbf{x}_2-\mathbf{x}_1=(-2,3,-1)$ ja paikkavektoriksi $\mathbf{r}=\mathbf{x}_1=(3,-1,3)$. Tällöin suoran parametriesitys on

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

Esimerkki 115 (Kahden tason leikkauksesta parametrimuotoon). Olkoot $T_1\colon 2x+y-z=2$ ja $T_2\colon 7x+3y-5z=3$. Koska tasojen T_1 ja T_2 normaalivektorit $(2,1,-1)$ ja $(7,3,-5)$ eivät ole samansuuntaiset (miksi?), tasot T_1 ja T_2 leikkaavat. Tasojen leikkaussuoralle saadaan parametrimuoto esimerkiksi käyttämällä Gaussin-Jordanin eliminointimenetelmää yhtälöpariin

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

saadaan pari

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

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

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

Esimerkki 116. Selvitetään, ovatko suorat

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

ja

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

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

Esimerkki 117. Määritetään a ja b siten, että suorat $L_1=\{(0,-1,1)+t(1,-7,-3)\mid t\in\mathbb{R}\}$ ja $L_2=\{(1,-1,3)+t(-2,a,b)\mid t\in\mathbb{R}\}$ ovat yhdensuuntaiset.

Suorat ovat yhdensuuntaiset silloin, kun niiden suuntavektorit $(1,-7,-3)$ ja $(-2,a,b)$ ovat yhdensuuntaiset. Tämä puolestaan tapahtuu silloin, kun $(1,-7,-3)=k(-2,a,b)$ jollekin luvulle k. Tällöin on oltava $k=-\frac{1}{2}$, $a=14$ ja $b=6$, jolloin siis suoran L_2 suuntavektorina toimii $\mathbf{x}=(-2,14,6)$ tai yhtä hyvin $\frac{1}{-2}\mathbf{x}=(1,-7,-3)$.

Selvitetään vielä, ovatko yhdensuuntaiset suorat $L_1=\{(0,-1,1)+t(1,-7,-3)\mid t\in\mathbb{R}\}$ ja $L_2=\{(1,-1,3)+t(1,-7,-3)\mid t\in\mathbb{R}\}$ *erillisiä* vai yhtyvätkö nämä suorat. Jos yhdensuuntaiset suorat ovat erilliset, ei niillä voi olla yhtään yhteistä pistettä. Selvitetään siis, onko yhtälöllä

$$(0,-1,1)+t_1(1,-7,-3)=(1,-1,3)+t_2(1,-7,-3) \tag{6.15}$$

ratkaisuja t_1, $t_2 \in \mathbb{R}$. Yhtälö (6.15) saadaan muotoon

$$(-1,0,-2) = (t_2 - t_1)(1,-7,3),$$

mistä nähdään, että ratkaisuja ei ole. Suorat ovat siis yhdensuuntaiset, mutta erilliset.

Esimerkki 118. Etsitään suorien

$$L_1 : \frac{x-4}{-5} = \frac{y}{3} = \frac{5z-4}{5}$$

ja

$$L_2 : \frac{x}{5} = \frac{y-1}{8} = \frac{z-3}{-12}$$

leikkauspiste (mikäli olemassa).

Muunnetaan aluksi suorat parametrimuotoon: suoran L_1 osalta kirjotetaan $\frac{x-4}{-5} = t$, $\frac{y}{3} = t$ ja $\frac{5z-4}{5} = t$, mistä $(x,y,z) = (-5t+4, 3t, t+\frac{4}{5}) = t(-5,3,1) + (4,0,\frac{4}{5})$.

Vastaavasti suoralle L_2 saadaan esitys $(x,y,z) = (5t, 8t+1, -12t+3) = t(5,8,-12) + (0,1,3)$, ja leikkauspiste edellyttää, että

$$(-5t_1+4, 3t_1, t_1 + \frac{4}{5}) = (5t_2, 8t_2+1, -12t_2+3)$$

joillakin arvoilla $t_1, t_2 \in \mathbb{R}$. Algebrallisen yhtälöryhmän ratkaisuna saadaan leikkauspiste, jossa $t_1 = \frac{37}{55}$ ja $t_2 = \frac{7}{55}$.

6.4 Etäisyyksien laskemisesta

Kahden tason samoin kuin kahden suorankin etäisyydellä toisistaan tarkoitetaan aina lyhintä mahdollista välimatkaa tasojen tai suorien välillä. Jos tasot tai suorat leikkaavat toisensa, on niiden etäisyys nolla. Samoin pisteen etäisyys tasosta tai suorasta merkitsee lyhintä mahdollista etäisyyttä. Näiden etäisyyksien laskemiseksi voidaan käyttää seuraavia periaatteita:

- Kahden tason välinen etäisyys: Jos tasojen T_1 ja T_2 normaalit $\mathbf{n}_1 = (n_1, n_2, n_3)$ ja $\mathbf{n}_2$ ovat erisuuntaiset, tasot leikkaavat ja niiden etäisyys on nolla. Jos taas normaalit ovat samansuuntaiset (lineaarisesti riippuvat), valitaan tasoja kohtisuoraan oleva suora $S = \{t\mathbf{n}_1 \mid t \in \mathbb{R}\}$, ja ratkaistaan arvo t_1, jolla $t_1\mathbf{n}_1 \in T_1$ sekä t_2, jolla $t_2\mathbf{n}_1 \in T_2$. Arvot t_1 löytyy esimerkiksi ratkaisemalla yhtälö $atn_1 + btn_2 + ctn_3 = d$, missä $ax + by + cz = d$ on tason T_1 standardimuoto. Tasojen välinen etäisyys saadaan lopuksi laskemalla pisteiden $t_1\mathbf{n}_1$ ja $t_2\mathbf{n}_1$ välinen etäisyys.
- Tason etäisyys suorasta: Jos tason T normaali $\mathbf{n}$ ei ole kohtisuorassa suoran S suuntavektoria kohtaan, Suora leikkaa tason ja etäisyys on nolla. Muussa tapauksessa suora on tason suuntainen. Tällöin valitaan $\mathbf{r}$ siten, että suora $\mathbf{r} + t\mathbf{n}$ leikkaa suoran S; vektoriksi $\mathbf{r}$ kelpaa esimerkiksi suoran S paikkavektori. Tämän jälkeen selvitetään, millä arvolla t_1 piste $\mathbf{p}_1 = \mathbf{r} + t_1\mathbf{n}$ kuuluu tasolle T ja millä arvolla t_2 piste $\mathbf{p}_2 = \mathbf{r} + t_2\mathbf{n}$ kuuluu suoralle S. t_1 löytyy käyttämällä tason koordinaattimuotoa ja jos $\mathbf{r}$ valittiin suoran S suuntavektoriksi, on $t_2 = 0$. Lopuksi lasketaan pisteiden $\mathbf{p}_1$ ja $\mathbf{p}_2$ välinen etäisyys.
- Tason etäisyys pisteestä: Olkoon $\mathbf{n}$ tason normaali ja $\mathbf{p}$ kysytty piste. Selvitetään, millä arvolla t_1 piste $\mathbf{p}_1 = \mathbf{p} + t_1\mathbf{n}$ kuuluu tasolle (käyttämällä tason koordinaattimuotoa) ja lasketaan pisteiden $\mathbf{p}_1$ ja $\mathbf{p}$ välinen etäisyys.
- Kahden suoran $S_1 : \mathbf{r}_1 + t\mathbf{s}_1$ ja $S_2 : \mathbf{r}_2 + t\mathbf{s}_2$ etäisyys: Valitaan suuntavektori $\mathbf{s} \neq \mathbf{0}$, joka on kohtisuorassa kumpaakin suoraa vastaan. Jos $\mathbf{s}_1$ ja $\mathbf{s}_2$ ovat lineaarisesti riippumattomat (eivät toistensa monikerrat), voidaan valita esim. $\mathbf{s} = \mathbf{s}_1 \times \mathbf{s}_2$. Tämän jälkeen selvitetään milloin suorilla S_1 ja S_2 olevien pisteiden $\mathbf{p}_1 = \mathbf{r}_1 + t\mathbf{s}_1$ ja $\mathbf{p}_2 = \mathbf{r}_2 + s\mathbf{s}_2$ erotus $\mathbf{p}_2 - \mathbf{p}_1$ on vektorin $\mathbf{s}$ suuntainen, siis milloin $\mathbf{p}_2 - \mathbf{p}_1 = c\mathbf{s}$. Tämä on kolmen tuntemattoman ja kolmen yhtälön ryhmä, joka ratkeaa Gaussin-Jordanin menetelmällä. Kysytty etäisyys saadaan pisteiden $\mathbf{p}_1 \in S_1$ ja $\mathbf{p}_2 \in S_2$ välisenä etäisyytenä. Jos alun perin $\mathbf{s}_1$ ja $\mathbf{s}_2$ olivat samansuuntaiset, voidaan valita suoralta S_1 mikä hyvänsä piste ja laskea tämän etäisyys suorasta S_2 (kts. seuraava kohta).
- Suoran etäisyys pisteestä: Minimoidaan differentiaalilaskennan keinoin lauseke $||\mathbf{r} + t\mathbf{s} - \mathbf{p}||^2$, missä $S : \mathbf{r} + t\mathbf{s}$ ja $\mathbf{p}$ ovat tarkasteltava suora ja piste. Saatava t:n arvo vastaa suoran pistettä, joka on lähinnä $\mathbf{p}$:tä.