Lineaarialgebra ja geometria

Lineaarialgebraa ja geometriaa käsittelevä vektorien ja matriisien kurssi yleistää koulusta tutut tason ja suoran sekä vektorin käsitteet useampaan kuin kolmeen ulotteisuuteen ja teknisesti myös vinokulmaiseen koordinaatistoon. Algebrallisesti tämä vastaa useamman kuin kolmen lineaarisen yhtälön käsittelyä.

Korkekouluissa opetettava differentiaalilaskenta käsittelee mielivaltaisen moniulotteisia pinnan ja käyrän yleistyksiä. Lukiossa differentiaalilaskennan keskeinen idea on korvata käyrä likimäärin tangentillaan. Moniulotteisessa differentiaalilaskennassa moniulotteinen pinta korvataan moniulotteisella tangenttitasollaan, ja viime kädessä korvataan mahdollisimman yleinen funktio lineaarisella funktiolla. Lineaarifunktion kuvaaja on yksi tason ja suoran moniulotteisista vastineista.

Lineaarialgebrassa selviää, mitä täsmälleen tarkoitetaan lineaarisilla funktioilla ja opetellaan laskemaan niillä. Osoittautuu, että tämä on helppoa.

Aloitamme sukelluksen ääretönulotteiseen avaruuteen kodikkaasta kolmiulotteisesta maailmastamme, osaammehan jo ennestään käsitellä tavallisia tasoja ja suoria.

Kiitos

Emme tietenkään ole kirjoittaneet tätä monistetta ilman työtovereidemme ja oppilaidemme apua. Kiitämme heitä kaikkia.

Kolmiulotteisessa kodikkaassa Mattilanniemessä 12.1.1999

Tekijät

Esipuhe nollapainokseen

Tämä malli monisteesta on painettu hieman keskeneräisenä. Mm. kiitos on kovin vajavainen. Ainakin Antti-Juhani Kaijanaho ja Pasi Mikkonen ansaitsisivat tulla nimeltä mainituiksi. Kiire johtuu siitä, että haluamme monisteen käytettäväksi kevätlukukauden 1999 opetuksessa.

Otamme edelleen mielellämme vastaan ilmoituksia puutteista.

12.1.1999

Tekijät

Esipuhe korjailtuun painokseen

Ilmeisesti jokainen malli monisteesta on oleva hieman keskeneräinen. Tähän laitokseen olemme koettanet korjata Pekka Kekäläisen keräämät painovirheet, mutta tekstiä ei ole sisällöltään muuteltu. Otamme kyllä edelleen mielellämme vastaan parannusehdotuksia.

25.9.2004

Tekijät

Typeset by $\mathcal{AMS}$-TEX

0. Koulutietojen kertaus ja täydennys

0.1. Avaruuden, tason ja lukusuoran vektorit.

Lukiossa ei määritellä täsmällisesti, mitä tarkoitetaan sanalla **vektori**. Tälläkin kurssilla määritelmä annetaan vasta myöhemmin ja on aika yllättävä. Tämän johdanto-osan ajaksi asetumme sille kannalle, että tiedämme mitä kolmi-, kaksi- ja yksiulotteiset avaruudet ja niiden vektorit ovat. Vektorit ovat meille siis hetken aikaa nuolia, suuntajanoja tai avaruuden siirtoja; tämähän on mukavaa vektoreiden laskutoimitusten tulkinnan kannalta. Toisaalta vektorit voi tunnetusti ymmärtää myös avaruuden (tai tason tai suoran) pisteiksi , paikkavektoreiksi, ja nämä edelleen lukukolmikoiksi. Tällöin avaruus on tosin ajateltava varustetuksi koordinaatistolla. Samalla pisteellä voi olla eri koordinaatit eri koordinaatistoissa.[1]

Lykkäämme epäilykset tuonnemmaksi ja merkitsemme avaruuden vektoreita tuttuun tapaan *koordinaatein* eli *komponentein*

$$\mathbf{v} = \overrightarrow{v} = (v_1, v_2, v_3) = v_1 \overrightarrow{i} + v_2 \overrightarrow{j} + v_3 \overrightarrow{k}.$$

tai — kuten matematiikassa yleensä on tapana[2] — ilman vektoriviivaa tai lihavointia ja milloin milläkin aakkosten kirjaimella. *Standardikantavektoreita* $\overrightarrow{i}$, $\overrightarrow{j}$ ja $\overrightarrow{k}$ merkitään usein yhtenäisesti $\overrightarrow{e}_i$ tai viivatta e_i (i=1,2,3). Vektoria merkitään siis usein esimerkiksi näin:

$$x = (x_1, x_2, x_3) = x_1 \overrightarrow{i} + x_2 \overrightarrow{j} + x_3 \overrightarrow{k}. = \sum_{i=1}^{3} x_i e_i.$$

Käytämme vastaavaa merkintätapaa myös alempiulotteisessa vektorilaskennassa:

$$x = (x_1, x_2) = x_1 \overrightarrow{i} + x_2 \overrightarrow{j} = x_1 e_1 + x_2 e_2.$$

Tässä on pieni tarkkana olon ja myöhemmän korjailun paikka: Riippuu tilanteesta, onko $(x_1, x_2) = x_1 \overrightarrow{i} + x_2 \overrightarrow{j}$ sama kuin (x_1, x_2) vai $(x_1, x_2, 0)$. Niin tai näin, voimme asettaa alustavan määritelmän vektorille:

Määritelmä 0.1.1. Reaaliluvuista $x_1, x_2, \ldots, x_n$ muodostettua järjestettyä joukkoa $x = (x_1, x_2, \ldots, x_n)$ sanotaan *n–komponenttiseksi vektoriksi* — ainakin jos n on 1, 2 tai 3.

1-komponenttiset vektorit ovat oleellisesti sama asia kuin lukusuoran pisteet, mikäli sovimme, että lukusuora on varustettu origolla ja asteikolla, jolloin sen pisteet toisaalta vastaavat reaalilukuja.

On hyvä myös ymmärtää, että koordinaattien käyttöön perustuvan määritelmämme mukaan kaksi vektoria ovat samat, jos niillä on samat komponentit: $x = y$, jos $x_i = y_i$ kaikilla i.

[1] Koordinaatiston ja vektorin itsensä suhde ei ole koulutietojen pohjalta ihan selvä. Onko vektori lukukolmikko vai ei? Koordinaatiston olemus ja koordinaatiston vaihdon tekniikka onkin oleellinen osa lineaarialgebraa.

[2] Matemaatikot tutkivat niin monenlaisia olioita lukujen ja vektoreiden lisäksi, että on toivotonta varata joka lajille omaa merkintäsysteemiä. On vain opittava elämään sen tosiasian kanssa, että sama kirjain x voi toisinaan tarkoittaa yhtälön tuntematonta, toisinaan vektoria ja toisinaan jonkin ihan muun joukon alkiota eli pistettä. Matemaattista tekstiä lukiessa on tästä huolimatta aina pidettävä mielessä, mikä kirjain juuri tässä edustaa minkäkin tyyppistä otusta!

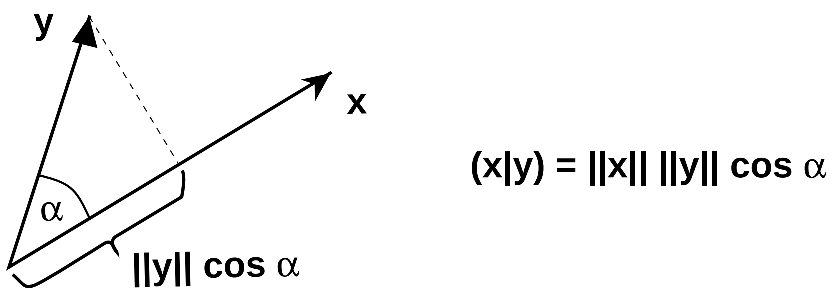
0.2. Laskutoimituksia vektoreilla.

Seuraavissa määritelmissä n on 1, 2 tai 3 riippuen siitä, minkäulotteista "luonnollista avaruutta" ajattelemme.

MÄÄRITELMÄ 0.2.1. Olkoot $x = (x_1, x_2, \ldots, x_n)$ ja $y = (y_1, y_2, \ldots, y_n)$ vektoreita ja $\lambda \in \mathbb{R}$. Määritellään *vektoreiden summa*, *kertominen reaaliluvulla* ja *sisätulo*:

$$
\begin{aligned}
&x + y = (x_1 + y_1,\, x_2 + y_2, \ldots, x_n + y_n),\\
&\lambda x = (\lambda x_1,\, \lambda x_2, \ldots, \lambda x_n),\\
&(x|y) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i.
\end{aligned}
$$

HUOMAUTUKSIA. Nämä laskutoimitukset ovat juuri samat, joihin lukiossa on totuttu. Summa muodostetaan suunnikkaan halkaisijan avulla, positiiviluvulla kertominen muuttaa vektorin "pituutta" ja negatiiviluvulla kertominen kääntää sen suunnan päinvastaiseksi. Sisätulo on merkintää vaille sama asia kuin *pistetulo* $x \cdot y$, joka koulutietojen mukaan on vektoreiden pituuksien tulo kerrottuna niiden välisen kulman kosinilla.



Antamillamme määritelmillä on se hyvä puoli, että ei tarvitse ensin selitellä, mitä kulmalla tai pituudella tarkoitetaan. Sisätulon avulla voi jälkeenpäin määritellä vektorin *pituuden*.[3]

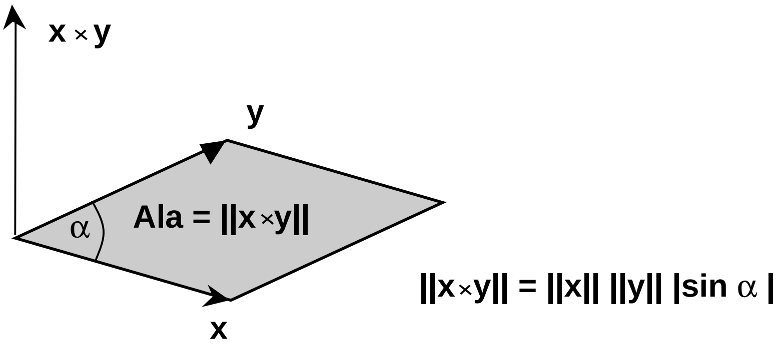
$$\|u\| = \sqrt{(u|u)}.$$

Kolmiulotteisessa avaruudessa on yleisesti käytössä vielä kolmas kertolaskun kaltainen laskutoimitus, nimittäin kahden vektorin *ristitulo*:

$$x \times y = (x_2 y_3 - x_3 y_2)\,\overrightarrow{i} + (x_3 y_1 - x_1 y_3)\,\overrightarrow{j} + (x_1 y_2 - x_2 y_1)\,\overrightarrow{k}\,.$$

Tavallisessa geometrisessa mielessä kahden vektorin ristitulo on niitä molempia vastaan *suorassa kulmassa* oleva eli *ortogonaalinen* vektori ja sen pituus on alkuperäisten vektoreiden pituuksien tulo kerrottuna niiden välisen kulman sinillä. Miten tämä asia perusteltiin koulussa?

[3]Samoin kulman. Asiaan palataan myöhemmin.



Koska ristitulon käyttö on nimenomaan kolmiulotteisen avaruuden erikoistekniikkaa, sen merkitys on vähäinen tällä mielivaltaiseen ulotteisuuteen pyrkivällä kurssilla.

Ristitulon symboli $\times$ on englanninkielisessä kirjallisuudessa usein tavallisen lukujen kertolaskun merkkinä. Tämä aiheuttaa sekaannusta aika harvoin. Myöskään kolmannesta samanlaisen ruksin $\times$ käyttömerkityksestä ei juuri saa syntymään kunnon väärinkäsitystä. Ristiähän käytetään myös merkitsemään kahden joukon A ja B *karteesista* eli *joukko-opillista tuloa*

$$A \times B = \{(a, b) \mid a \in A \text{ ja } b \in B\}.$$

Tason tavalliset koordinaatit antavat esimerkin karteesisesta tulosta:

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x_1, x_2) \mid x_1 \in \mathbb{R} \text{ ja } x_2 \in \mathbb{R}\}.$$

Kolmen joukon A,B ja C karteesinen tulo muodostuu vastaavasti alkiokolmikoista jne. Lukukolmikoiden joukkoa, tavallista avaruutta on siten mukava merkitä $\mathbb{R}^3$. Seikkaperäinen selostus karteesisesta tulosta on JOHDATUS MATEMATIIKKAAN–monisteessa.

Vektoreiden laskusääntöjä on harjoiteltu muualla. Emme kertaakaan niitä nyt, vaan kiinnitämme vastaisen varalle huomiota asian joukko-opilliseen puoleen. Kaikki laskutoimitukset ja myös vektorin pituuden määrääminen ovat nimittäin joukko-opillisessa mielessä funktioita eli kuvauksia. Pitää vain ymmärtää missä joukoissa ne on määritelty ja missä joukoissa ne saavat arvoja. Jos merkitsemme reaalilukujen joukkoa tavalliseen tapaan $\mathbb{R}$:llä, lukuparien joukkoa $\mathbb{R}^2$:lla ja kolmikoiden joukkoa $\mathbb{R}^3$:lla, niin huomaamme, että esimerkiksi avaruuden vektorin pituuden laskeminen on kuvaus

$$\mathbb{R}^3 \to \mathbb{R}.$$

Avaruuden vektorin kertominen reaaliluvulla on kuvaus

$$\mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3,$$

liittäähän se luvun ja vektorin muodostamaan pariin (λ, x) vektorin λx. Vektoreiden summa ja ristitulo ovat samassa mielessä kuvauksia

$$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3,$$

ja sisätulo on kuvaus

$$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}.$$

Tason vektoreiden laskutoimitukset (joihin ristitulo ei lukeudu) tulkitaan kuvauksiksi samalla tavalla, mutta $\mathbb{R}^3$ korvataan tietysti joukolla $\mathbb{R}^2$. On tärkeätä opetella ajattelemaan mm. laskutoimituksia funktioina, sillä erilaisten funktioiden — nimenomaan muiden kuin koulusta tuttujen reaalifunktioiden — käsittelytaito on matematiikassa aivan keskeistä.

Miettikäämme hetken ajan, mitä vektorifunktiot oikeastaan voivat olla geometriselta kannalta ja missä niitä esiintyy.

0.3. Vektorimuuttujan vektoriarvoinen funktio.

Lukiossa ja viimeistään JOHDATUKSESSA MATEMATIIKKAAN on esiintynyt yleinen *funktion* eli *kuvauksen* käsite. Kertaamme pääkohdat.

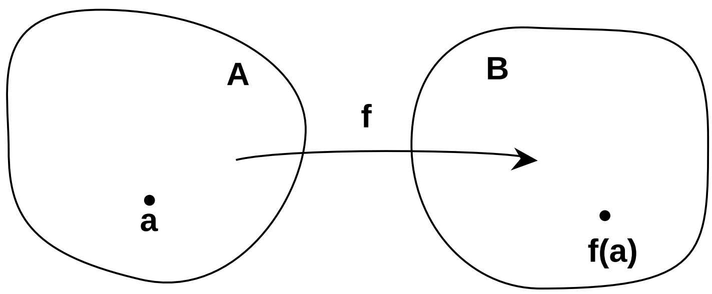
Funktio on aina määritelty kahden joukon välille:

$$f : A \to B$$

Funktio **liittää kuhunkin** *lähtö-* **eli** *määrittelyjoukon* A *alkioon* **eli** *pisteeseen* $a \in A$ **tasan yhden alkion** *maalijoukosta* B;

$$f : a \mapsto f(a).$$

Periaatteessa joukot A ja B voivat olla mitä tahansa ja funktioksi kelpaa mikä tahansa sääntö, kunhan se kertoo jokaisesta A:n alkiosta a, mikä sen *kuva* eli *arvo* $f(a)$ on. Huomaa, että funktion arvo on joukon B alkio, yleensä siis ei luku.

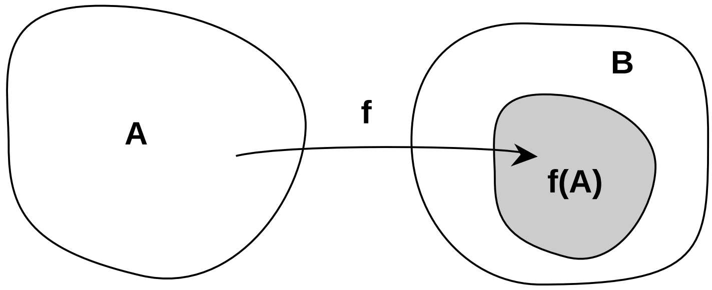


Esimerkki funktiosta on vaikkapa se kuvaus f, joka liittää kuhunkin Suomen kansalaiseen hänen nimensä. Toinen funktio, olkoon se vaikka g, liittää jokaiseen kansalaiseen hänen henkilötunnuksensa. Sekä f että g ovat määriteltyjä kansalaisten joukossa A ja saavat arvoja kaikkien mahdollisten kirjaimista, miinus- ja plusmerkeistä, numeroista ja sananväleistä muodostettujen enintään 100-merkkisten sanojen joukossa B.

Joukon A kuva eli kuvajoukko on kaikkien A :n alkioiden kuvien joukko, siis

$$f(A) = \{f(x) \mid x \in A\}.$$

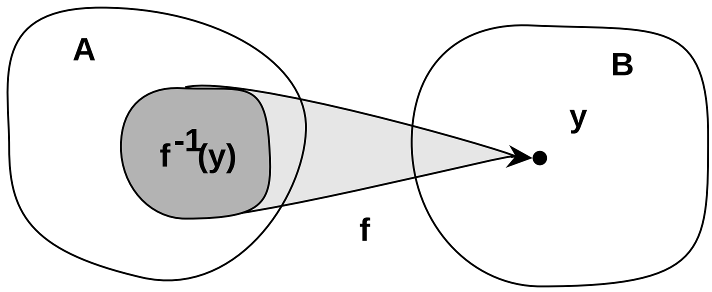
Selvästi $f(A) \subset B$, mutta yleensä $f(A) \neq B$. Jos sattuu olemaan $f(A) = B$, niin sanotaan, että f on *surjektio* joukolta A joukolle B.



Alkion $y \in B$ *alkukuva* on niiden alkioiden $x \in A$ joukko, joilla $f(x) = y$.

$$f^{-1}(\{y\}) = \{x \in A \mid f(x) = y\}$$

Toisin sanoen pisteen $y \in B$ alkukuva on *yhtälön $f(x) = y$ ratkaisujen* joukko. Huomaa, että eri pisteiden alkukuvat ovat *erillisiä joukkoja*, ts. niillä ei ole yhteisiä alkioita. Tämä johtuu siitä, että funktiolle esittämämme vaatimuksen mukaan f kuvaa jokaisen luvun vain yhdeksi arvoksi.



Pisteen alkukuvan merkintätapa aiheuttaa joskus hämmennystä. Se saattaa seota *käänteiskuvaukseen*, jota merkitään melkein samalla tavalla. Ero on siinä, että pisteen alkukuva $f^{-1}(\{x\})$ on joukko. Toisin kuin käänteiskuvaus alkukuva on myös aina olemassa, olipa f mikä funktio tahansa. Pisteen alkukuva voi tietysti olla tyhjä joukko.

Joukon kuvan ja pisteen alkukuvan välillä on sellainen yhteys, että jos y kuuluu kuvajoukkoon $f(A)$, niin sen alkukuva $f^{-1}(\{y\})$ on epätyhjä, mutta muuten tyhjä joukko. Erityisesti siis f on surjektio täsmälleen sillä ehdolla, että minkään pisteen $y \in B$ alkukuva $f^{-1}(\{y\})$ ei ole tyhjä, vaan yhtälöllä $f(x) = y$ on ratkaisu, olipa y mikä tahansa B:n alkio. Huomaamme, että "surjektio" ja "kuvajoukko" ovat sanoja, joita kannattaa käyttää tutkittaessa yhtälöiden ratkeavuutta, siis sitä, milloin ratkaisuja on olemassa **ainakin yksi**. Entä milloin yhtälöllä $f(x) = y$ on "yksikäsitteinen" eli **enintään yksi** ratkaisu? Näin käy ilmeisesti tasan silloin, kun alkukuvaan $f^{-1}(\{y\})$ kuuluu enintään yksi piste. Erityisesti näin käy, jos f on sellainen kuvaus, joka kuvaa eri pisteet aina eri pisteiksi, ts. jos pätee

$$f(a) = f(b) \implies a = b$$

Tällaista funktiota f sanotaan *injektioksi.* Jos f on sekä sur– että injektio, niin f:llä on olemassa käänteiskuvaus. Tällaista funktiota sanotaan *bijektioksi* $A \to B$.

Lähes kaikki koulumatematiikassa esiintyneet funktiot liittävät reaalilukuihin lukuja. Monilla näistä funktioista on "nimi" tai "lauseke". Tällaisia ovat esimerkiksi sini, kosini, eksponenttifunktio tai eräästä polynomista ja eksponenttifunktiosta yhdistetty kuvaus

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto e^{2x^2+3x+2}.$$

Vektoreiden laskutoimitusten ajatteleminen funktioina auttaa vapautumaan vanhentuneesta[4] ja varmasti ainakin lineaarialgebran kannalta haitallisesta tavasta ajatella, että sana "funktio" tarkoittaisi yleensä vain näitä "lausekkeita".

Tällä kurssilla kohtaamme pääasiassa vektorimuuttujan vektoriarvoisia funktioita — itse asiassa lähes yksinomaan ns. *lineaarikuvauksia.* Lineaarisuuden käsitteen kaikinpuolinen selittäminen on lineaarialgebran ja geometrian kurssin varsinainen sisältö. Sanomme lineaarisuudesta jotakin alustavaa seuraavassa kohdassa, mutta kannattaa ensin miettiä hiukan sitä, millaisia olioita kuvaukset

$$\mathbb{R}^n \to \mathbb{R}^m$$

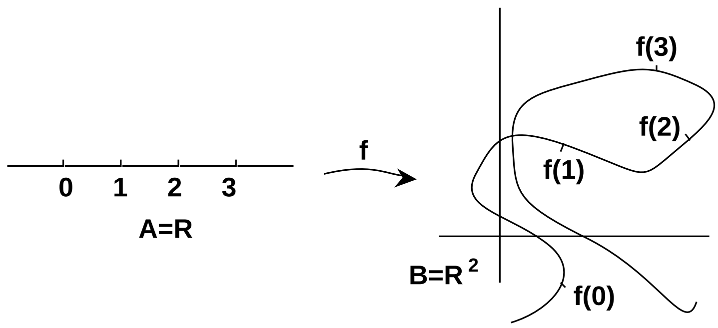
yleensäkin ovat ($n,\ m \in \{1, 2, 3\}$).

Tapaus $m = n = 1$ on koulussa käsitelty; kuvausta $f : \mathbb{R} \to \mathbb{R}$ voi havainnollistaa esimerkiksi kuvaajallaan, tai sitten ajatella, että $f(t)$ on vaikkapa "ajasta t riippuva luku".

Seuraavaksi helpoimmassa tapauksessa $n = 1,\ m = 2$, jolloin f on reaalimuuttujan vektoriarvoinen funktio $f : \mathbb{R}^1 \to \mathbb{R}^2$. Tällaisen voi mukavasti ajatella ajasta riippuvaksi tason pisteeksi, eli pisteen liikkeeksi tasossa. Koordinaatein ilmaistuna

$$f(t) = (f_1(t), f_2(t)),$$

missä f_1 ja f_2 ovat tavallisia reaaliarvoisia funktioita, "parametrisoidun käyrän" eli "liikkeen" f *komponenttifunktiot.*



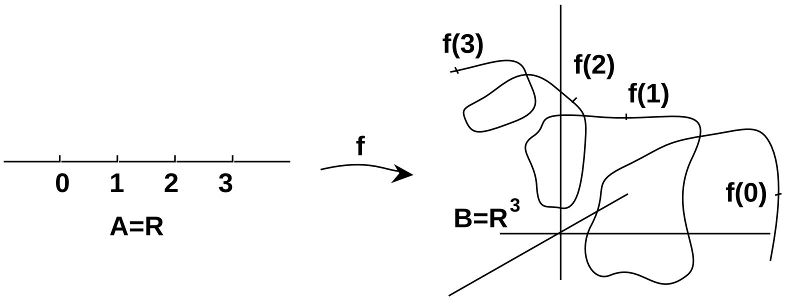
Lähtöjoukkona olevan lukusuoran $\mathbb{R}$ *kuvajoukkoa*

$$f(\mathbb{R}^1) = \{f(t) \mid t \in \mathbb{R}^1\}$$

voi nyt hyvällä syyllä sanoa *tasokäyräksi.* Lineaarialgebrassa tutkitaan erityisesti tilannetta, jossa $f(\mathbb{R}^1)$ on suora.

Tapaus $n = 1,\ m = 3$ johtaa kolmikomponenttisten yhden muuttujan funktioiden teoriaan. Geometriselta kannalta näitä voi ajatella luonnollisessa avaruudessa sijaitsevina käyrinä.

[4]Aivan niin. Funktiolla on ennen vanhaan tarkoitetettu juuri näitä "yhden reaalimuuttujan alkeisfunktioita"



Tapauksessa $n = 2$, $m = 1$ kuvaus $f : \mathbb{R}^2 \to \mathbb{R}$ liittää jokaiseen tason pisteeseen luvun, ja tapauksessa $n = 3$, $m = 1$ $f : \mathbb{R}^3 \to \mathbb{R}$ jokaiseen avaruuden pisteeseen luvun. Fysikaalisessa todellisuudessa luku $f((x_1, x_2))$ — lyhemmin $f(x_1, x_2)$ tai vain $f(x)$ — voi olla vaikkapa tämänhetkinen ilmanpaine tai lämpötila paikkakunnalla, jonka maantieteelliset koordinaatit ovat x ja y. Lämpötila paikan funktiona on hyvä myös esimerkkinä **avaruuden** vektorin reaaliarvoisesta funktiosta eli *kolmen reaalimuuttujan* funktiosta, jollaista fyysikot usein sanovat *skalaarikentäksi.*

Tarkastellessamme edellä yhden (n=1) reaalimuuttujan — ”ajan” — vektoriarvoisia funktioita huomasimme, että lähtöavaruuden kuvajoukko $f(\mathbb{R}^1)$ oli geometrisesti mielenkiintoinen, itse asiassa ”käyrä” maalijoukossa $\mathbb{R}^2$ tai vastaavasti $\mathbb{R}^3$.

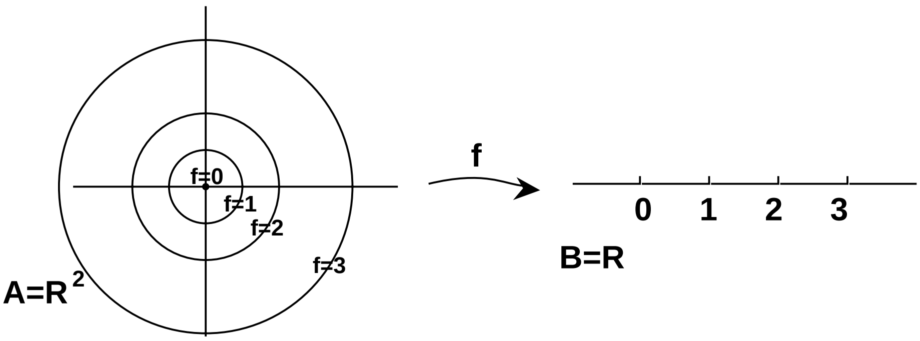
Tarkastellessamme nyt toiseen suuntaan menevää kuvausta $f : \mathbb{R}^2 \to \mathbb{R}$ ovat pisteiden alkukuvat eli funktion f *tasa-arvojoukot*

$$f^{-1}(\{y\}) = \{x \in \mathbb{R}^2 \bigm| f(x) = y\}$$

geometrisesti mielenkiintoisia. Esimerkiksi kuvauksessa

$$f : \mathbb{R}^2 \to \mathbb{R} : (x_1, x_2) \mapsto x_1^2 + x_2^2$$

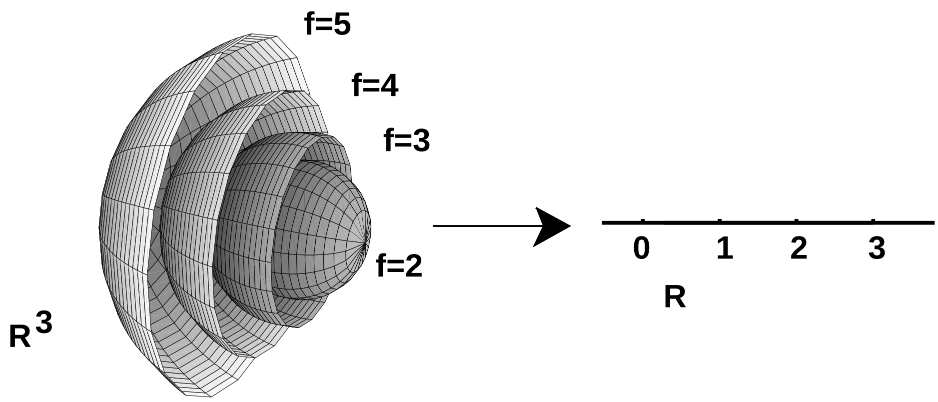
pisteen $y = 1 \in \mathbb{R}$ alkukuva, siis niiden pisteiden joukko, joissa $f(x) = f(x_1, x_2) = 1$ on origokeskinen 1-säteinen ympyrä. Klassinen analyyttinen geometria käsittelee tähän tapaan määriteltyjä käyriä, *tasa-arvokäyriä.*



Kolmiulotteista tilannetta sivutaan koulukurssissa vain vähän. Kuvauksessa $f : \mathbb{R}^3 \to \mathbb{R}$ pisteiden eli lukujen alkukuvia on järkevää sanoa funktion f *tasa-arvopinnoiksi.* Esimerkiksi kuvauksessa

$$f : \mathbb{R}^3 \to \mathbb{R} : (x_1, x_2, x_3) \mapsto x_1^2 + x_2^2 + x_3^2$$

pisteen $1 \in \mathbb{R}$ alkukuva — siis niiden pisteiden joukko, joissa $f(x_1, x_2, x_3) = 1$ — on origokeskinen 1-säteinen pallo, oikeastaan pallonkuori. Tässä esimerkissä muutkin funktion positiivisa arvoja $y > 0$ vastaavat tasa-arvopinnat $f^{-1}(\{y\})$ ovat origokeskisiä, siis sisäkkäisiä palloja. Nollas tasa-arvopinta kutistuu pelkäksi origoksi ja negatiivisten lukujen alkukuvat tyhjäksi joukoksi.



Palloesimerkki on siinä mielessä edustava, että yleensäkin vektorimuuttujan säännöllisen reaaliarvoisen funktion tasa-arvojoukot ovat yhdensuuntaisia pintoja ikäänkuin sipulin kuoret. Kun tunnemme tasa-arvopinnat ja funktion arvot kullakin niistä, tunnemme koko funktion. Mielikuva sipulinkuorista on lämpöjakauman tavoin hyvä keino kuvitella mielessään kolmen reaalimuuttujan funktiota eli skalaarikenttää.

Millaisia sitten ovat vektorimuuttujan vektoriarvoiset funktiot? Ainakin on selvää, että kuten edellä tapauksessa $n = 1$ funktiolla

$$f : \mathbb{R}^n \to \mathbb{R}^m$$

on joka tapauksessa m komponenttifunktiota, mutta nämä ovat yleisen n tapaukssa *vektorimuuttujan* reaaliarvoisia funktioita, siis juuri sellaisia, joita äsken kuvailin.

Palaamme tähän tapaukseen tuotapikaa, mutta varmistelemme ensin äsken opittuja asioita kertaamalla niitä ”lineaarisessa” tapauksessa, lukiossahan on jo tutkiskeltu tasojen ja suorien yhtälöitä.

0.4. Tason ja suoran yhtälöitä.

Lukiossa esitetään monenlaisia suoran ja tason yhtälöitä. Näitä ovat ainakin seuraavat:

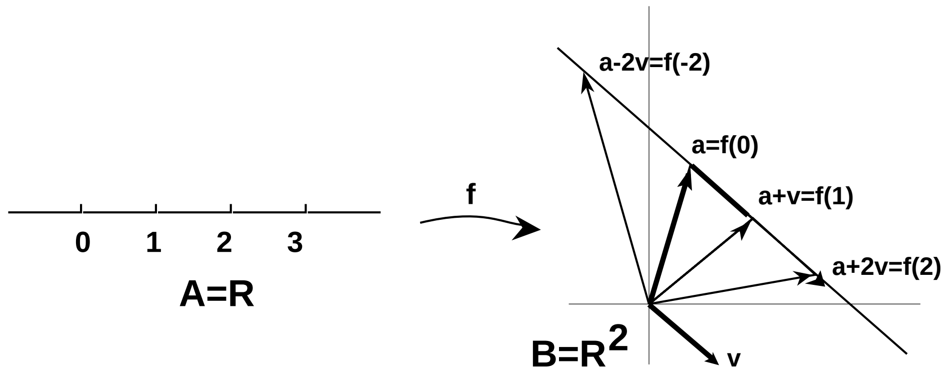
1) *Suoran parametrimuotoinen yhtälö (Suora ”käyränä”).*

Olkoot a ja v m–komponenttisia vektoreita, joista v ei nollavektori[5]. Funktion

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}^m \\ t &\mapsto a + tv \end{aligned}$$

kuvajoukko $\{a + tv \mid t \in \mathbb{R}\}$ on suora. Parametrimuotoinen yhtälö on helppo muodostaa, kun suorasta tunnetaan piste ja suunta tai kaksi pistettä. Jälkimmäinen tapaus palautuu edelliseen ottamalla suunnaksi pisteiden erotusvektori.

[5]Mieti, miksi nollaa ei haluta.



On helppoa piirtää suora, kun sen parametrimuotoinen yhtälö on annettu. Sen sijaan parametrimuotoisesta yhtälöstä ei heti laskuitta näe, onko jokin annettu piste $x \in \mathbb{R}^m$ suorallamme. Tietysti piste a kuuluu suoralle. Erityisesti, jos $a = \overrightarrow{0}$, niin suora kulkee origon kautta; ehkä muutenkin. Huomaa, että maaliavaruuden $\mathbb{R}^m$ ulotteisuus m ei juuri vaikuta näihin päättelyihin.

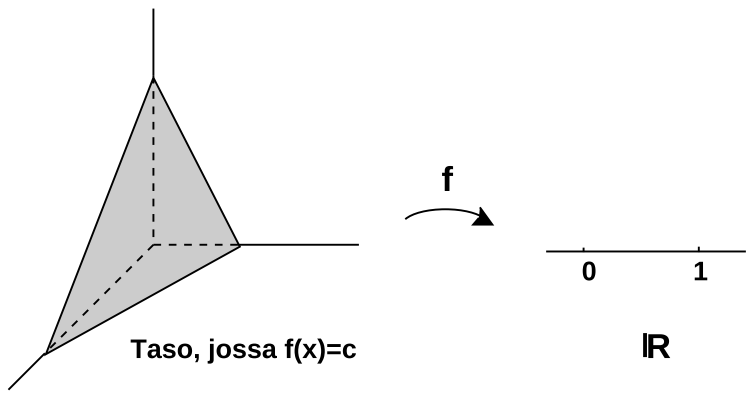
2) *Tason lineaarinen yhtälö avaruudessa. (Taso "tasa-arvopintana")*

Olkoon $a = (a_1, a_2, a_3) \in \mathbb{R}^3$ nollasta eroava vektori ja c luku, ja olkoon f funktio

$$\begin{aligned} \mathbb{R}^3 &\to \mathbb{R} \\ (x_1, x_2, x_3) &\mapsto a_1x_1 + a_2x_2 + a_3x_3. \end{aligned}$$

Pisteen $c \in \mathbb{R}$ alkukuva $f^{-1}(\{c\})$ on taso[6]

$$\{(x_1, x_2, x_3) \mid a_1x_1 + a_2x_2 + a_3x_3 = c\}.$$



Tason alkukuvatyyppisellä yhtälöllä pystyy helposti testaamaan, onko jokin annettu piste x tasossa vai ei. Tarvitsee vain laskea $f(x)$ ja todeta saatiinko c vai jokin muu luku — toteutuuko yhtälö. Yhtälön avulla voi myös helposti selvittää, missä kohdissa tasomme leikkaa jonkin käyrän tai paramterimuotoisen suoran, esimerkiksi koordinaattiakselit. Lasketaan malliksi leikkaus x–akselin kanssa. x–akseli on parametrimuodossa esimerkiksi kuvajoukko $S = \varphi(\mathbb{R})$, missä

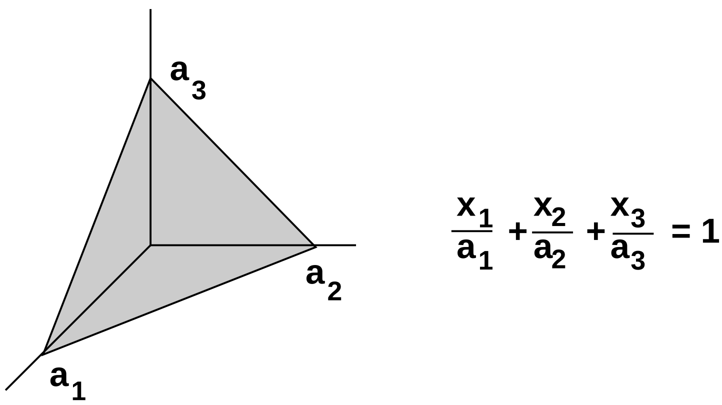
[6]Tässä tarvitaan tietoa, että jokin $a_i \neq 0$. Miksi?

$\varphi(t) = 0 + t(1,0,0) = (t,0,0)$. Piste $\varphi(t) = (t,0,0)$ kuuluu tasoon $f^{-1}(\{c\})$ tasan sillä ehdolla, että $f(\varphi(t)) = c$. Sijoitamme ja sievennämme:

$$\begin{aligned} f(\varphi(t)) &= c \\ f(t,0,0) &= c \\ a_1 t &= c \\ t &= \frac{c}{a_1}, \text{ mikäli } a_1 \neq 0. \text{ Muuten ratkaisua ei ole.} \end{aligned}$$

Leikkauskohta on $\varphi(t) = (\frac{c}{a_1}, 0, 0)$.

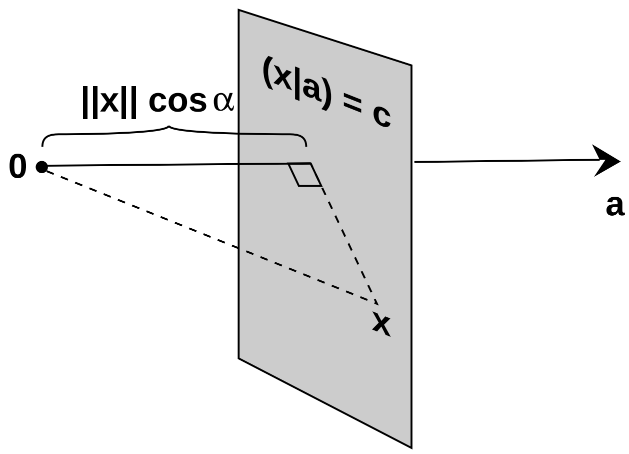
Laskemalla toiset vastaavalla tavalla huomaamme, että taso $a_1x_1 + a_2x_2 + a_3x_3 = c$ leikkaa akselit kohdissa $\frac{c}{a_i}$, mikäli luvut a_i ovat nollasta eroavia ($i = 1, 2, 3$). Muuten se on ilmeisesti jonkin akselin suuntainen. Tätä voi käyttää toisinkin päin: tasolla, joka leikkaa akselit nollasta eroavissa kohdissa a_1, a_2 ja a_3 on ilmeisesti yhtälö $\frac{x_1}{a_1} + \frac{x_2}{a_2} + \frac{x_2}{a_2} = 1$, koska olemme juuri laskeneet, että tämä taso leikkaa akselit oikeissa paikoissa. Taso on helppo piirtää, kun tunnetaan sen leikkaukset akselien kanssa.



Päätämme tason alkukuvatyyppisen yhtälön esittelyn kahteen havaintoon.

Ensinnäkin, jos pidetään lukuja a_i kiinteinä, siis funktiota f ei muutella, niin eri lukujen c alkukuvat ovat erillisiä tasoja, ovathan ne eri pisteiden alkukuvia. Ne ovat siis yhdensuuntaisten tasojen parvi. ”Sipulinkuoremme” ovat tässä tapauksessa pikemminkin ”kirjan lehtiä”!

Toisekseen on niin, että edellä tehdyillä laskuilla on luonnollinen geometrinen tulkinta. Funktion $f : (x_1, x_2, x_3) \mapsto a_1x_1 + a_2x_2 + a_3x_3$ arvo $f(x) = \sum_{i=1}^{3} a_i x_1$ on yksinkertaisesti vektoreiden $a = (a_1, a_2, a_3)$ ja x sisätulo! Tasomme muodostuu kaikista niistä vektoreista, joilla sisätulo kiinteän vektorin a kanssa saa saman arvon, siis yhtälön $(a|x) = c$ ratkaisuista. Siksi tasomme on kohtisuorassa vektoria a vastaan. Olemme löytäneet kaikki vektorin a normaalitasot. Nyt ymmärrämme geometrisestikin, millä ehdolla tasomme on jonkin koordinaattiakselin suuntainen; tasan silloin kun $a = (a_1, a_2, a_3)$ on kohtisuorassa akselia vastaan.

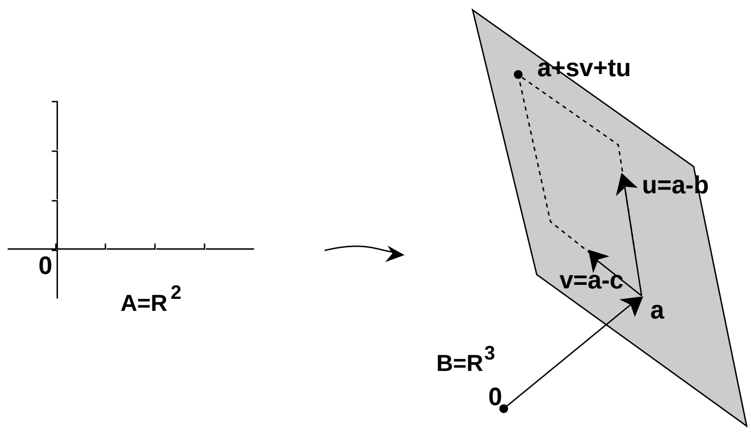


3) *Tason parametrimuotoinen yhtälö (Taso "parametrisoituna pintana").* Yleistämällä kohtaa 1) saa tasollekin parametrimuotoisen yhtälön. Tämä on helppo kirjoittaa tietokoneen siirrä teksti- toiminnolla. Pitää vain korjata lähtöjoukoksi $\mathbb{R}^2$. Kas noin:

Olkoot a ja v sekä u m–komponenttisia vektoreita, joista v ja u erisuuntaisia.[7] Funktion

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R}^m \\ (t,s) &\mapsto a + tv + su \end{aligned}$$

kuvajoukko $\{a + tv + su \mid (t,s) \in \mathbb{R}^2\}$ on taso. Parametrimuotoinen yhtälö on helppo muodostaa, kun tasosta tunnetaan piste ja kaksi sen suuntaista vektoria tai kolme eri pistettä. Jälkimmäinen tapaus palautuu edelliseen ottamalla suunniksi kaksi pisteiden erotusvektoria.



On myös helppoa piirtää taso, kun sen parametrimuotoinen yhtälö on annettu. Sen sijaan parametrimuotoisesta yhtälöstä ei heti laskuitta näe, onko jokin annettu piste $x \in \mathbb{R}^m$ tasollamme. Tietysti piste a kuuluu tasolle. Erityisesti, jos $a = \overrightarrow{0}$, niin taso kulkee origon kautta, ehkä muutenkin.

[7] Mieti, miksi yhdensuuntaiset eivät kelpaa.

Ongelma: Miten muutat tason yhtälön muodosta toiseen?

4) *Suoran yhtälöryhmä (suora tasojen leikkauksena).* Edellisen innoittamana korjataan kohtaa 1) muuttamalla maaliavaruus kaksiulotteiseksi.

Olkoot $a = (a_1, a_2, a_3)$ ja $b = (b_1, b_2, b_3)$ kaksi erisuuntaista vektoria ja c_1 ja c_2 kaksi lukua, toisin sanoen $c = (c_1, c_2) \in \mathbb{R}^2$. Määritellään funktio

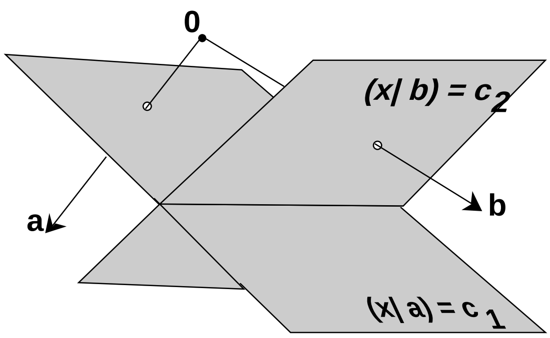
$$\begin{aligned} \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x_1, x_2, x_3) &\mapsto (a_1x_1 + a_2x_2 + a_3x_3, b_1x_1 + b_2x_2 + b_3x_3) \end{aligned}$$

Pisteen $c \in \mathbb{R}^2$ alkukuva $f^{-1}(\{c\})$ on kahden erisuuntaisen tason leikkaus, siis suora. Tämän voi perustella näin:

$$f^{-1}(\{c\}) = \{(x_1, x_2, x_3) \mid a_1x_1 + a_2x_2 + a_3x_3 = c_1 \text{ ja } b_1x_1 + b_2x_2 + b_3x_3 = c_2\} =$$
$$\{(x_1, x_2, x_3) \mid a_1x_1 + a_2x_2 + a_3x_3 = c_1\} \cap \{(x_1, x_2, x_3) \mid b_1x_1 + b_2x_2 + b_3x_3 = c_2\}.$$

Suoralle $f^{-1}(\{c\})$ on saatu tutunnäköinen yhtälöpari.

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 &= c_1 \\ b_1x_1 + b_2x_2 + b_3x_3 &= c_2 \end{cases}$$



Suora (x |a) = c$_1$ ja (x | b) = c$_2$

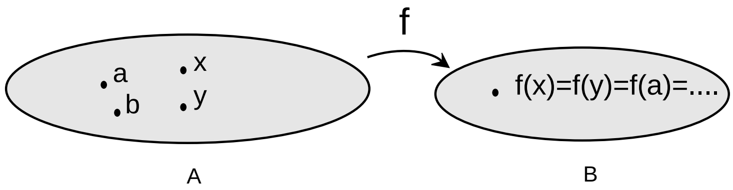
Emme nyt laske enää enempää. Sen sijaan on aika miettiä tuloksia. Ainakin on selvää, että tason yhtälön hyvä hallinta antaa sivutuotteena tuntuman suoran yhtälöpariin. Toinen asia, johon kannattaa kiinnittää huomiota, on se, että avaruudessa $\mathbb{R}^3$ määritellyn $\mathbb{R}^2$–arvoisen funktion tasa-arvojoukot, siis pisteiden alkukuvat näyttävät olevan pintojen leikkauksia, siis siinä mielessä ”käyriä”.

0.5. Lineaarikuvaukset.

Täydennämme nyt koulukurssia. Edellisen luvun pitkät laskut ja mutkikkaat merkinnät sekä eri tapausten erittelyt ovat hankalia. Parempi olisi katsella asioita yhtenäisellä tavalla ennen siirtymistä mielivaltaisen moneen ulotteisuuteen. Kaikki tutkimamme neljä tason ja suoran yhtälöä olivat pohjimmiltaan kuva- tai alkukuvajoukkoja, kunhan tarkasteltiin sopivaa funktiota f. Lineaarialgebran kannalta olennaista on, että f oli tapauksissa (2) ja (4) lineaarikuvaus ja tapauksissa (1) ja (3) lineaarikuvauksen ja vakiofunktion summa.

MÄÄRITELMÄ. Kuvaus $f : A \to B$ on *vakiofunktio*, mikäli se saa saman arvon kaikissa pisteissä, toisin sanoen jos on olemassa piste $b \in B$ siten, että jokaisella $a \in A$ pätee

$$f(a) = b.$$



Vakiokuvauksen kuvajoukko on yksipisteinen: $f(A) = \{b\}$.

MÄÄRITELMÄ. Kuvaus $f : \mathbb{R}^n \to \mathbb{R}^m$ on *lineaarinen*, mikäli se toteuttaa seuraavat aksioomat:

(L-1) $f(x + y) = f(x) + f(y)$ kaikille x ja $y \in \mathbb{R}^n$.
(L-2) $f(\lambda x) = \lambda f(x)$ kaikille $\lambda \in \mathbb{R}$ ja kaikille $x \in \mathbb{R}^n$.

Osoittautuu, että jokainen lineaarikuvaus kuvaa origon origoksi. Lisäksi se kuvaa suorat suoriksi tai pisteiksi. Taso kuvautuu tasoksi, suoraksi tai pisteeksi. Myös pisteen, suoran tai tason lineaariset alkukuvat luonnollisessa avaruudessa $\mathbb{R}^3$ ovat joitakin näistä tai koko avaruus. Pohdimme seuraavassa kohdassa, miksi näin on. Ensin tarkastellaan esimerkkejä.

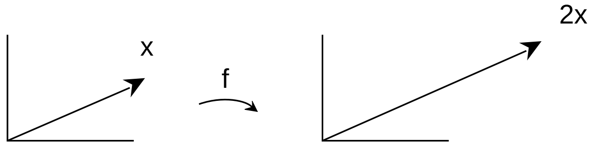
ESIMERKKEJÄ LINEAARIKUVAUKSISTA. Perusesimerkkejä lineaarikuvauksista:

(1) *Luvulla $\mu \in \mathbb{R}$ kertominen*, toisin sanoen *μ–venytys* eli *μ–skaalaus*

$$f : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto \mu x$$

on tietysti lineaarikuvaus, onhan

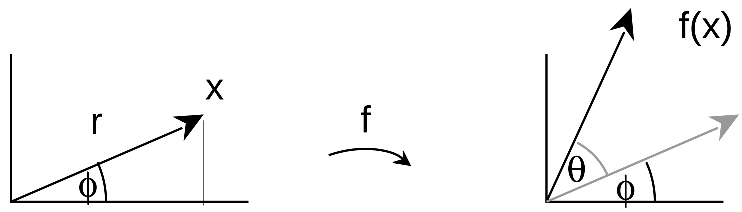
$$\begin{aligned} &f(x + y) = \mu(x + y) = \mu x + \mu y = f(x) + f(y) \\ &\text{ja } f(\lambda x) = \mu(\lambda x) = \mu\lambda x = \lambda\mu x = \lambda f(x). \end{aligned}$$



(2) Tason *kierto* (origon suhteen) kulman θ verran on myös lineaarikuvaus.[8] Laskennallinen käsittely on hankalampaa kuin venytyksen tapauksessa, mutta ainakin kierron voi ilmaista napakoordinaateissa ja sieventää tuloksen.

$$\begin{aligned} f : \mathbb{R}^2 \to \mathbb{R}^2 : (x_1, x_2) &= (r\cos\phi, r\sin\phi) \\ &\mapsto (r\cos(\phi + \theta), r\sin(\phi + \theta)) \\ &= (r(\cos\phi\cos\theta - \sin\phi\sin\theta), r(\sin\phi\cos\theta + \cos\phi\sin\theta)) \\ &= (x_1\cos\theta - x_2\sin\theta, x_1\sin\theta + x_2\cos\theta). \end{aligned}$$

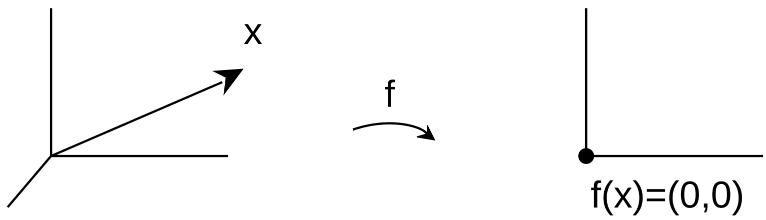
[8] $F(x + y) = F(x) + F(y)$ ja $F(\lambda x) = \lambda F(x)$ pätevät tietty aina, mikäli F on kierto origon ympäri!



(3) *Nollakuvaus*

$$f : \mathbb{R}^n \to \mathbb{R}^m : x \mapsto 0$$

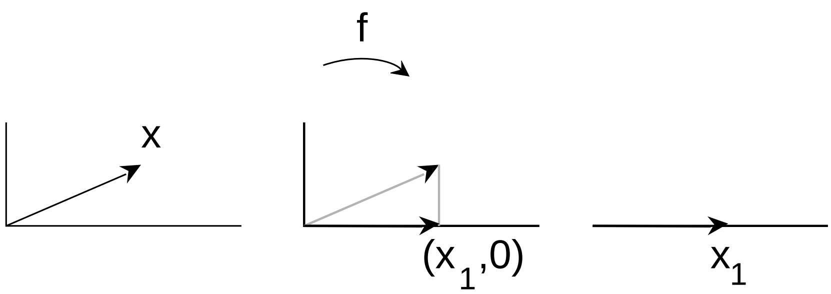
on lineaarikuvaus, olivatpa ulotteisuudet n ja m mitä tahansa.



(4) *Ensimmäinen koordinaattifunktio* eli projektio x_1-akselille

$$f : \mathbb{R}^n \to \mathbb{R} : x = (x_1, \ldots, x_n) \mapsto x_1$$

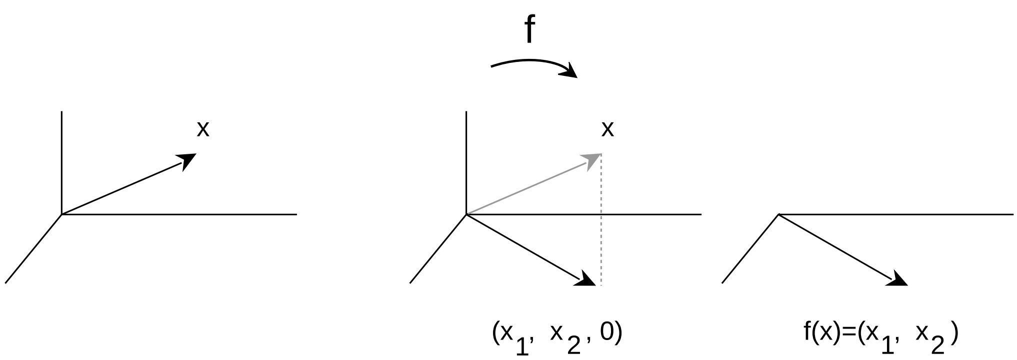
on tietysti lineaarinen. Sama koskee muita koordinaattifunktioita $f : \mathbb{R}^n \to \mathbb{R} : x = (x_1, \ldots, x_n) \mapsto x_i$, missä $(i = 1, 2 \ldots, n)$.



(5) Vastaavasti myös *projektio* x_1, x_2*–tasoon*

$$f : \mathbb{R}^3 \to \mathbb{R}^2 : x = (x_1, x_2, x_3) \mapsto (x_1, x_2)$$

eli x_3–koordinaatin poisto on sekin lineaarinen — samoin projektiot muille koordinaattitasoille.



Esimerkkejä lineaarikuvauksista ja vakioista. Myös tasojen ja suorien yhtälöitä tutkiessamme esiintyneet funktiot f, joista oikeastaan saimme aiheen tutkia lineaarikuvauksia, ovat lineaarisia tapauksissa (2) ja (4). Tämän toteamiseksi tarkastetaan huolellisesti toteutuvatko määritelmän kaksi ehtoa näille funktioille f. Aloitetaan suoran alkukuvamuotoisessa yhtälössä esiintyneellä funktiolla

$$\begin{aligned} \mathbb{R}^3 &\to \mathbb{R} \\ (x_1, x_2, x_3) &\mapsto a_1x_1 + a_2x_2 + a_3x_3. \end{aligned}$$

Tarkastamme lineaarisuuden:

$$\begin{aligned} f(x+y) &= f(x_1+y_1, x_2+y_2, x_3+y_3) \\ &= a_1(x_1+y_1) + a_2(x_2+y_2) + a_3(x_3+y_3) \\ &= a_1x_1 + a_1y_1 + a_2x_2 + a_2y_2 + a_3x_3 + a_3y_3 \\ &= (a_1x_1 + a_2x_2 + a_3x_3) + (a_1y_1 + a_2y_2 + a_3y_3) \\ &= f(x) + f(y) \\ f(\lambda x) &= f(\lambda x_1, \lambda x_2, \lambda x_3) \\ &= a_1\lambda x_1 + a_2\lambda x_2 + a_3\lambda x_3 \\ &= \lambda(a_1x_1 + a_2x_2 + a_3x_3) \\ &= \lambda f(x). \end{aligned}$$

Tässä oli liikaa laskemista. Saman tuloksen saa suoraan käyttämällä havaintoa, että kyseessä on sisätulo, ja että tunnemme sisätulon laskusäännöt.

$$\begin{aligned} f(x+y) &= (a|x+y) = (a|x) + (a|y) = f(x) + f(y) \\ f(\lambda x) &= (a|\lambda x) = \lambda(a|x) = \lambda f(x). \end{aligned}$$

Myös kohdan (4) kaksikomponenttisen funktion voi kirjoittaa sisätulon avulla.

$$\begin{aligned} f : \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x_1, x_2, x_3) &\mapsto (a_1x_1 + a_2x_2 + a_3x_3, b_1x_1 + b_2x_2 + b_3x_3) = ((a|x), (b|x)) \end{aligned}$$

Lineaarisuus on nyt helppo todeta.

$$\begin{aligned} f(x+y) &= ((a|x+y), (b|x+y)) \\ &= ((a|x) + (a|y), (b|x) + (b|y)) \\ &= ((a|x), (b|x)) + ((a|y), (b|y)) \\ &= f(x) + f(y) \\ f(\lambda x) &= ((a|\lambda x), (b|\lambda x)) \\ &= (\lambda(a|x), \lambda(b|x)) \\ &= \lambda((a|x), (b|x)) \\ &= \lambda f(x). \end{aligned}$$

Kohdissa (1) ja (3) käsitellyt funktiot eivät ole lineaarisia, paitsi tapauksessa $a = 0$, jolloin saadaan origon kautta kulkevaa suoraa ja tasoa parametrisoivat funktiot

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}^m \\ t &\mapsto tv \end{aligned}$$

ja

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R}^m \\ (t, s) &\mapsto tv + su. \end{aligned}$$

Kummankin lineaarisuus on ilmiselvää.

Vakiofunktio

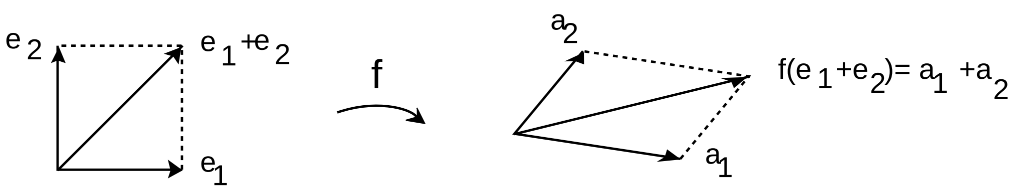
$$\begin{aligned} f_c : \mathbb{R}^n &\to \mathbb{R}^m \\ x &\mapsto c \end{aligned}$$

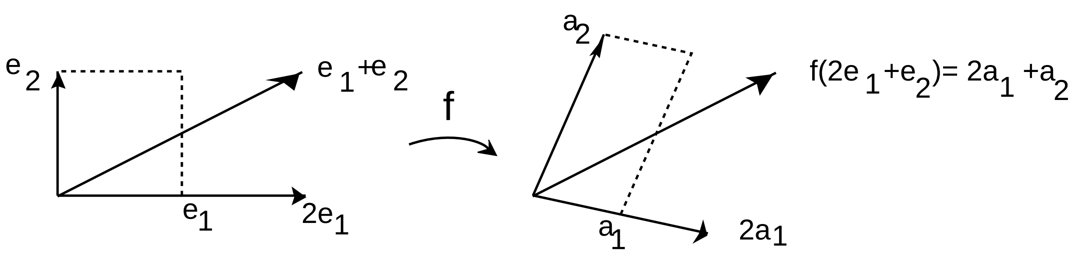
ei sen sijaan ole lineaarinen muulloin kuin tapauksessa $c = 0$.

0.6. Matriisit.

Lineaarikuvausten valtava menestys matemaatikon yleistyökaluna perustuu hyvien geometristen ominaisuuksien lisäksi siihen, että[9] lineaarikuvauksen arvon voi laskea helposti tavallisella nelilaskimella. Tämän ymmärtää parhaiten geometriselta kannalta. Siksi tarkastellaan kaksiulotteista tapausta, joka on helppo piirtää.

LINEAARIKUVAUKSEN GEOMETRIA. Olkoon $F : \mathbb{R}^2 \to \mathbb{R}^2$ lineaarikuvaus.[10] Olkoot standardikantavektoreiden kuvat $F(e_1) = \overrightarrow{a}_1$ ja $F(e_2) = \overrightarrow{a}_2$. Osoittautuu, että **muuta tietoa kuin kantavektoreiden kuvat ei tarvitakaan kaikkien pisteiden $x \in \mathbb{R}^2$ kuvien laskemiseksi tai piirtämiseksi.** Aloitamme piirtämällä muutamien pisteiden kuvat.





[9] Äärellisulotteisessa avaruudessa.

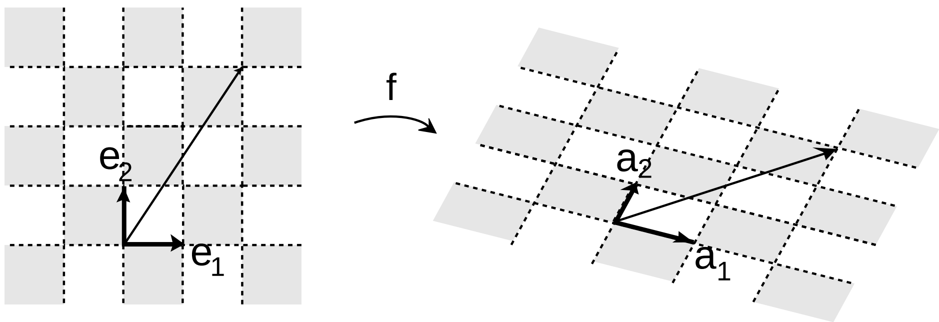
[10] Lineaarikuvauksia merkitään useimmiten isoilla kirjaimilla, esim. F, L, A tai T. Lineaarikuvauksen arvoa $F(x)$ merkitään usein sulkeitta Fx. Tässä johdannossa käytetään vielä sulkeita.

Näemme, että kantavektorin e_1 suuntaisen vektorin λe_1 kuva on yksinkertaisesti $F(\lambda e_1) = \lambda \overrightarrow{a}_1$ ja kantavektorin e_2 suuntaisen vektorin μe_2 kuva on $F(\mu e_2) = \mu \overrightarrow{a}_2$.

Myös eräiden muiden vektoreiden kuva on helppo päätellä, esimerkiksi $F(e_1 + e_2) = F(e_1) + F(e_2) = \overrightarrow{a}_1 + \overrightarrow{a}_2$. Yhdistämällä nämä kaksi ideaa saamme selville, että minkä tahansa vektorin $x = (x_1, x_2) \in \mathbb{R}^2$ kuva on:

$$\begin{aligned} F(x_1, x_2) &= F(x_1 e_1 + x_2 e_2) \\ &= F(x_1 e_1) + F(x_2 e_2) \\ &= x_1 F(e_1) + x_2 F(e_2) \\ &= x_1 \overrightarrow{a}_1 + x_2 \overrightarrow{a}_2. \end{aligned}$$

Tämä on todella helppo piirtää: lue x:n koordinaatit x_1 ja x_2 ja piirrä kuvapisteeksi piste, jolla on nämä **samat** koordinaatit siinä *vinokulmaisessa koordinaatistossa*, jonka *kantavektorit* ovat $\overrightarrow{a}_1$ ja $\overrightarrow{a}_2$, siis standardikantavektoreiden kuvat.



Näiden tarkastelujen jälkeen on kuvan perusteella uskottavaa, että ainakin 2-ulotteisessa tapauksessamme lineaarikuvaus $F : \mathbb{R}^2 \to \mathbb{R}^2$ kuvaa origon origoksi ja suorat viivat suoriksi viivoiksi tai mahdollisesti pisteiksi.[11] Olemme myös huomanneet, että kun kantavektoreiden kuvat on piirretty, niin minkä tahansa pisteen $x = (x_1, x_2)$ kuvapisteen $F(x) = x_1 \overrightarrow{a}_1 + x_2 \overrightarrow{a}_2$ voi piirtää mitään laskematta.

Lineaarikuvauksen algebra. Lineaarikuvauksessa $F : \mathbb{R}^2 \to \mathbb{R}^2$ vektorin x kuvapisteen $F(x)$ voi toisaalta myös **laskea** tavallisissa koordinaateissa hyvin siististi, kun kantavektoreiden kuvapisteet tunnetaan. Laskua varten tarvitsemme tietysti annettujen vektoreiden koordinaatit. Olkoot esimerkiksi

$$\begin{aligned} & F(e_1) = \overrightarrow{a}_1 = (1, 2) \\ \text{ja} \quad & F(e_2) = \overrightarrow{a}_2 = (3, 4), \end{aligned}$$

jolloin

$$\begin{aligned} F(x) = F(x_1, x_2) &= x_1 \overrightarrow{a}_1 + x_2 \overrightarrow{a}_2 \\ &= x_1(1, 2) + x_2(3, 4) \\ &= (x_1 + 3x_2,\ 2x_1 + 4x_2) \end{aligned}$$

[11] Ainakin nollakuvaus kuvaa kaiken pisteeksi.

Vastaavan tuloksen saa tietysti joka tapauksessa, olivatpa kantavektoreiden kuvat $\overrightarrow{a}_1$ ja $\overrightarrow{a}_2$ mitä tahansa vektoreita: Merkiten

$$\begin{aligned} &\overrightarrow{a}_1 = (a_{11}, a_{21}) \\ \text{ja} \quad &\overrightarrow{a}_2 = (a_{12}, a_{22}), \end{aligned}$$

saa

$$\begin{aligned} F(x) = F(x_1, x_2) &= x_1 \overrightarrow{a}_1 + x_2 \overrightarrow{a}_2 \\ &= x_1(a_{11}, a_{21}) + x_2(a_{12}, a_{22}) \\ &= (x_1 a_{11} + x_2 a_{12},\ x_1 a_{21} + x_2 a_{22}). \end{aligned}$$

Kun katselee tulosta, voisi ensi silmäyksellä ehkä kuvitella, että tuloksen komponentit olisivat vektoreiden x ja $\overrightarrow{a}_i$ sisätuloja, mutta **näin ei asia sentään ole**. Tarkkaan katsoen huomaa, että a_{12} ja a_{21} ovat vaihtuneet toisikseen. Oikean tuloksen muistamiseksi kannattaakin kirjata kantavektoreiden kuvien $F(e_1) = \overrightarrow{a}_1 = (a_{11}, a_{21})$ ja $F(e_2) = \overrightarrow{a}_2 = (a_{12}, a_{22})$ koordinaatit kaavioksi, joita sanotaan lineaarikuvauksen F *matriisiksi.*

$$\mathrm{Mat}(F) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Kuvan $F(x)$ koordinaatit ovat matriisin ***rivien*** **ja kuvattavan vektorin x sisätulot. Matriisi on järjestelty siten, että matriisin sarakkeet ovat kantavektoreiden kuvat.**

$$\begin{aligned} &F(e_1) = \overrightarrow{a}_1 = (a_{11}, a_{21}) \\ \text{ja} \quad &F(e_2) = \overrightarrow{a}_2 = (a_{12}, a_{22}) \end{aligned}$$

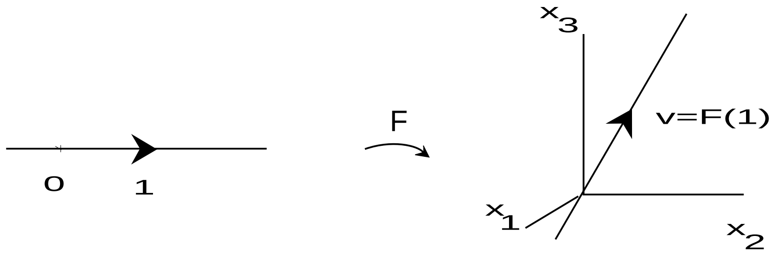
Matriisi sisältää kaiken tiedon lineaarikuvauksesta $F : \mathbb{R}^2 \to \mathbb{R}^2$. Kaikenulotteisilla lineaarikuvauksilla $\mathbb{R}^n \to \mathbb{R}^m$ on matriisi samaan tapaan kuin tapauksessa $n = m = 2$. *Sarakkeiden* — kantavektoreiden kuvien — määrä on ilmeisesti n ja rivien — kantavektorin kuvan komponenttien — luku m. Muodostetaan malliksi matriisit niille funktioille, jotka saimme edellä tason ja suoran yhtälöistä.

Esimerkki 1.

$$F : \mathbb{R} \to \mathbb{R}^3 : \quad t \mapsto F(t) = tv = (tv_1, tv_2, tv_3)$$

Määrittelyjoukkona olevan avaruuden $\mathbb{R} = \mathbb{R}^1$ ainoa standardikantavektori $e_1 \in \mathbb{R}^1$ on luku 1. Sen kuva on $F(1) = (v_1, v_2, v_3)$, ja tällaisen yksiulotteisessa avaruudessa määritellyn lineaarikuvauksen F matriisissa on siis vain tämä yksi ainoa sarake. $\mathrm{Mat}(F)$ on *sarakematriisi*:

$$Mat(F) = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

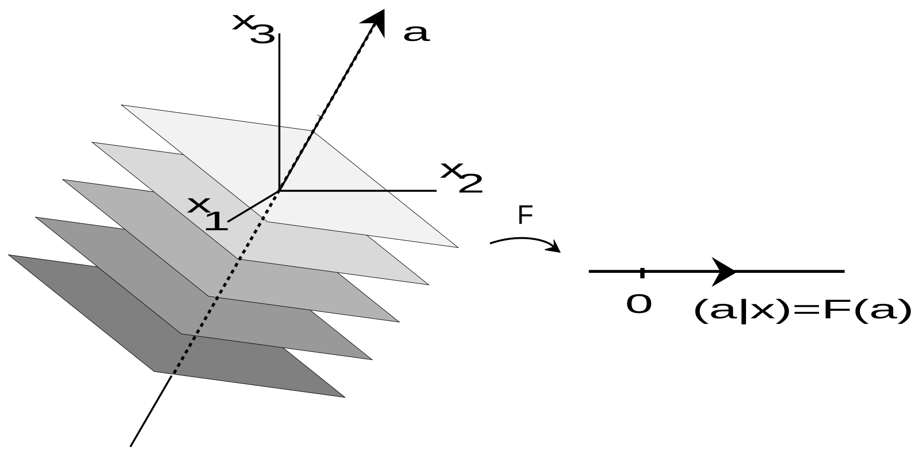


Esimerkki 2.

$$F:\mathbb{R}^3\to\mathbb{R}:\quad (x_1,x_2,x_3)\mapsto F(x_1,x_2,x_3)=a_1x_1+a_2x_2+a_3x_3=(a|x)$$

Määrittelyjoukkona olevan avaruuden $\mathbb{R}^3$ standardikantavektoreiden e_1, e_2, $e_3\in\mathbb{R}^3$ kuvat ovat $F(e_1)=(a|e_1)=a_1$, $F(e_2)=(a|e_2)=a_2$, ja $F(e_3)=(a|e_3)=a_3$. Nämä ovat reaalilukuja, siis avaruuden $\mathbb{R}^1$ alkioita. Reaaliarvoisen kuvauksen F matriisissa on vain yksi ainoa rivi. Mat(F) on *rivimatriisi*:

$$\mathrm{Mat}(F)=\begin{bmatrix} a_1 & a_2 & a_3\end{bmatrix}$$

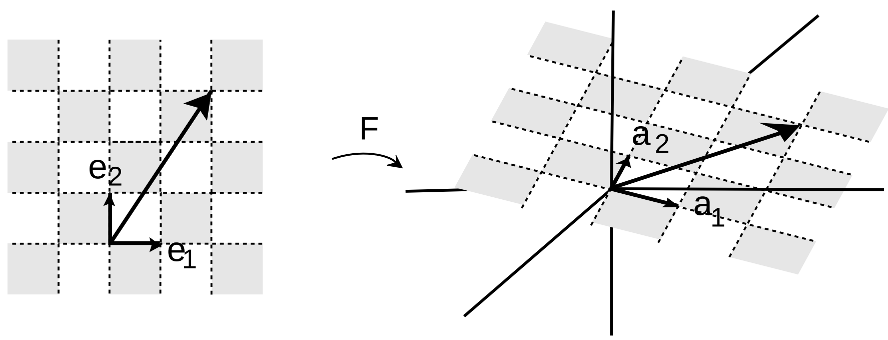


Esimerkki 3.

$$F:\mathbb{R}^2\to\mathbb{R}^3:\quad (t,s)\mapsto F(t,s)=tv+su$$

Määrittelyjoukkona olevan avaruuden $\mathbb{R}^2$ standardikantavektoreiden $e_1=(1,0)$ ja $e_2=(0,1)\in\mathbb{R}^2$ kuvat ovat $F(e_1)=f(1,0)=v=(v_1,v_2,v_3)$ ja $F(e_2)=F(0,1)=u=(u_1,u_2,u_3)\in\mathbb{R}^3$. Kuvauksen $F:\mathbb{R}^2\to\mathbb{R}^3$ matriisissa on 2 saraketta ja 3 riviä.

$$\mathrm{Mat}(F)=\begin{bmatrix} v_1 & u_1\\ v_2 & u_2\\ v_3 & u_3\end{bmatrix}.$$



Esimerkki 4.

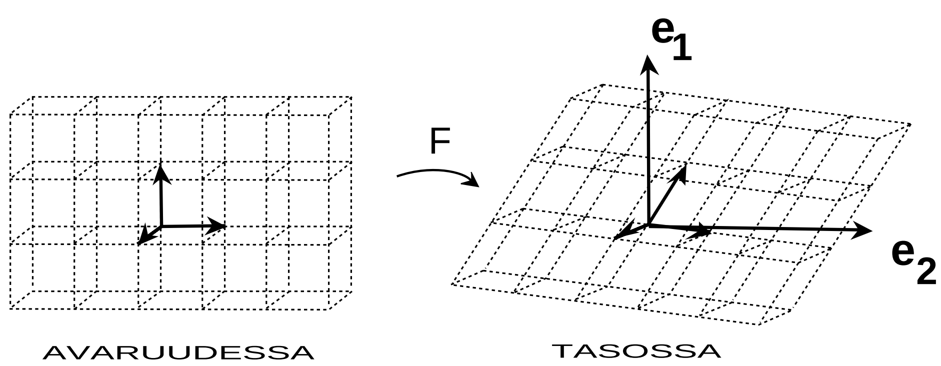
$$F : \mathbb{R}^3 \to \mathbb{R}^2 : \quad (x_1, x_2, x_3) \mapsto (a_1x_1+a_2x_2+a_3x_3, b_1x_1+b_2x_2+b_3x_3) = ((a|x), (b|x))$$

Määrittelyjoukkona olevan avaruuden $\mathbb{R}^3$ standardikantavektoreiden $e_1 = (1, 0, 0), e_2 = (0, 1, 0)$ ja $e_3 = (0, 0, 1) \in \mathbb{R}^2$ kuvat ovat

$$\begin{aligned} F(e_1) &= ((a|e_1), (b|e_1)) = (a_1, b_1) \\ F(e_2) &= ((a|e_2), (b|e_2)) = (a_2, b_2) \\ \text{ja } F(e_3) &= ((a|e_3), (b|e_3)) = (a_3, b_3) \in \mathbb{R}^2. \end{aligned}$$

Kuvauksen $F : \mathbb{R}^3 \to \mathbb{R}^2$ matriisissa on 3 saraketta ja 2 riviä.

$$\mathrm{Mat}(F) = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$



0.7. Tason kierto.

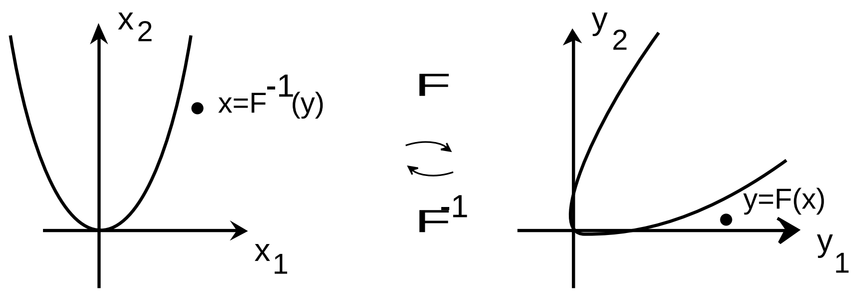
Lineaarikuvausta ja sen matriisia voi käyttää tasojen ja suorien yhtälöiden ymmärtämiseen ja yleistämiseen, kuten edellisessä kohdassa annoimme aavistaa. Samalla on saatu väline, jolla päästään käsiksi koordinaatiston kiertoon. Koordinaatiston kiertäminen on toisaalta analyyttisen geometrian kannalta suunnilleen sama asia kuin "kuvioiden kiertäminen vastakkaiseen suuntaan". Havainnollistamme tätä johtamalla 45 astetta eli $\frac{\pi}{4}$ radiaania kierretyn paraabelin yhtälön.[12]

Esimerkki. Asetamme tavoitteeksi keksiä yhtälön, joka kuvaa paraabelia

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid f(x) = x_2 - x_1^2 = 0\}$$

kierrettynä $\frac{\pi}{4}$:n verran myötäpäivään eli matemaattisesti negatiiviseen suuntaan

[12]Koulussahan kartioleikkauksia, siis paraabelia, ellipsiä ja hyperbeliä, käsitellään lähes aina ns. pääakselikoordinaatistossaan, jonka akselit ovat kuvion symmetria-akselit. Hyperbelille on lukiossa yhtälöt $xy = a$ ja $x^2 - y^2 = b$ asennosta ja koosta riippuen. Kierto muuntaa nämä toisikseen!



Periaatteessahan kaikki on ihan selvää. Kierto on lineaarikuvaus $F : \mathbb{R}^2 \to \mathbb{R}^2$. Kierto on myös bijektio ja sen käänteiskuvaus F^{-1} on kierto $\frac{\pi}{4}$:n verran vastapäivään, eli matemaattisesti positiiviseen suuntaan. Käänteiskuvausta tarvitaan siksi, että piste $F(x) = y = (y_1, y_2) \in \mathbb{R}^2$ kuuluu **kierrettyyn** paraabeliin tasan silloin, kun piste

$$F^{-1}(y)$$

on **alkuperäisellä** paraabelilla eli toteuttaa alkuperäisen paraabelin yhtälön:

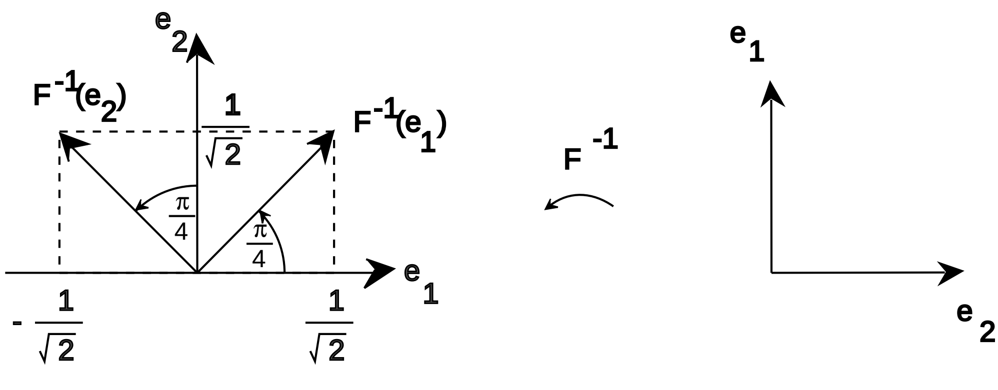
$$f(F^{-1}(y)) = 0.$$

Tämä on haluttu yhtälö tiiviissä muodossa.[13] Lineaarikuvauksen $F^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ lauseke pitää vielä keksiä ja sijoittaa tiiviiseen yhtälöön. Sittenhän loppu on pelkkää sieventelyä.

Kierron F^{-1} lauseke onkin tiedossamme. Muistamme nimittäin sivulta 18, että koska kierto F^{-1} on lineaarikuvaus $\mathbb{R}^2 \to \mathbb{R}^2$, sen täydelliseen hallintaan riittää tietää kantavektoreiden kuvat, tarkemmin sanoen

$$F^{-1}(y) = F^{-1}(y_1, y_2) = y_1 \overrightarrow{a}_1 + y_2 \overrightarrow{a}_2$$

missä $\overrightarrow{a}_1 = (a_{11}, a_{21})$ ja $\overrightarrow{a}_2 = (a_{12}, a_{22})$ ovat standardikantavektoreiden kuvat kuvauksessa F^{-1}, siis tässä tehtävässä kantavektorit kierrettyinä $\frac{\pi}{4}$:n verran. Nämä tiedetään:



[13]Korvaamalla f jonkin muun käyrän määrittelevällä funktiolla saisi kierrettyä jotakin muuta käyrää.

$$\overrightarrow{a}_1 = (a_{11}, a_{21}) = F^{-1}(e_1) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$
$$\overrightarrow{a}_2 = (a_{12}, a_{22}) = F^{-1}(e_2) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right),.$$

Siispä

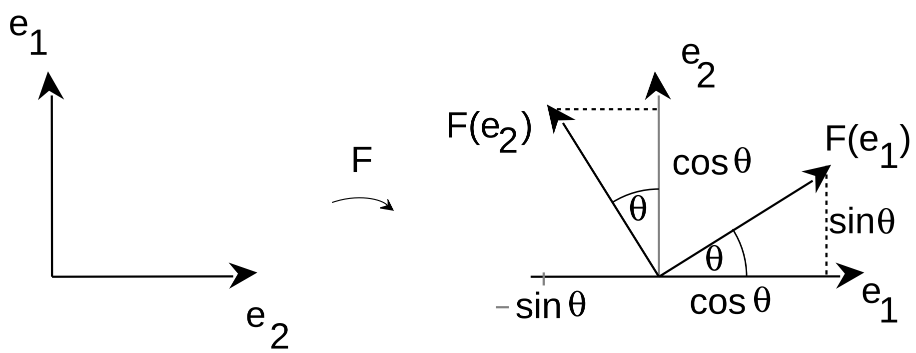
$$\begin{aligned} F^{-1}(y) = F^{-1}(y_1, y_2) &= \\ &= y_1 \overrightarrow{a}_1 + y_2 \overrightarrow{a}_2 \\ &= y_1 \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) + y_2 \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \\ &= \left(\frac{y_1 - y_2}{\sqrt{2}}, \frac{y_1 + y_2}{\sqrt{2}}\right). \end{aligned}$$

Sijoittamalla tämä f:n lausekkeeseen saadaan kierretyn paraabelin yhtälöksi

$$\begin{aligned} 0 &= f(F^{-1}(y)) \\ &= \left(\frac{y_1 + y_2}{\sqrt{2}}\right) - \left(\frac{y_1 - y_2}{\sqrt{2}})^2\right) \\ &= \frac{1}{\sqrt{2}} y_1 + \frac{1}{\sqrt{2}} y_2 + y_1 y_2 - \frac{1}{2} y_1^2 - \frac{1}{2} y_2^2. \end{aligned}$$

Kootaan tulokset:

- Kierretynäkin paraabeli on toisen asteen kahden muuttujan polynomin korkeuskäyrä, *toisen asteen käyrä.*
- Kierto on lineaarikuvaus.



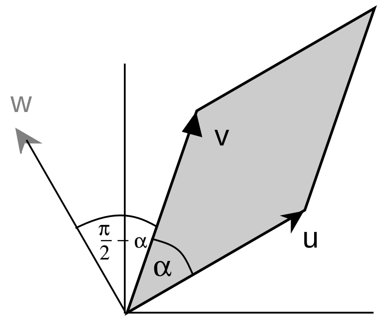
- Lineaarikuvauksia koskevan luvun alussa käytimme napakoordinaatteja laskeaksemme lausekkeen mielivaltaista kulmaa θ vastaavalle kierrolle. Kiertokuvausta $(x_1, x_2) \mapsto (x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta)$ vastaava matriisi on
$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

0.8. Suunnikkaan ja suuntaissärmiön tilavuus.

Tason vektoreiden u ja $v \in \mathbb{R}^2$ virittämän suunnikkaan *ala* eli *kaksiulotteinen tilavuus* on niiden pituuksien tulo kerrottuna niiden välisen kulman sinillä, siis $\|u\|\|v\|\sin(\alpha)$. Tämä on — mahdollisesti merkkiä vaille — sama kuin *kaksirivinen determinantti*

$$\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} = u_1v_2 - v_1u_2.$$

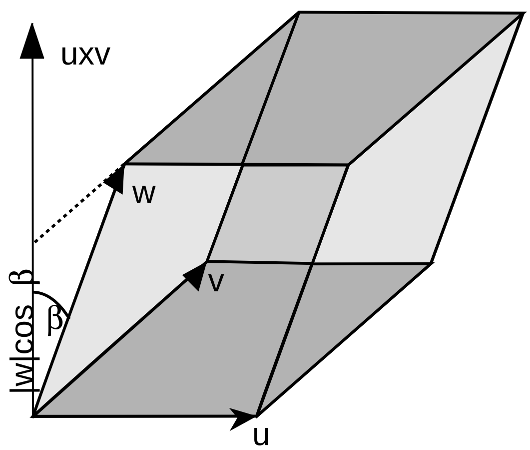
Perustelu. $u_1v_2 - v_1u_2$ on sisätulo vektorista v ja vektorista w, joka on saatu kiertämällä vektoria u kulman $\frac{\pi}{2}$ verran. Laskemalla kulmat huomaa, että sisätulolle pätee $(v|w) = \|v\|\|w\|\cos(\frac{\pi}{2} - \alpha) = \|v\|\|u\|\sin\alpha$.



□

Suunnikkaan alalle saatiin yllättävän helppo lauseke, ja asia vain paranee, kun pannaan merkille, että on olemassa myös kolmiulotteinen malli samasta ideasta. Avaruuden vektoreiden u, v ja w virittämän suuntaissärmiön tilavuus on — mahdollisesti merkkiä vaille — sama kuin *kolmirivinen determinantti*

$$\begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} = u_1 \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} - u_2 \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix} + u_3 \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix}$$



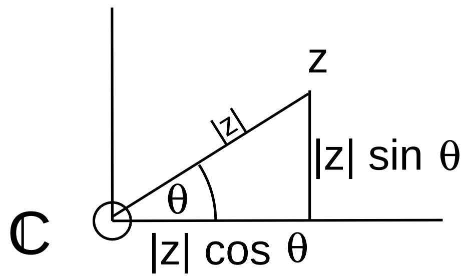
Perustelu. Ainakin kuvan tilanteessa, jossa vektoreiden väliin jää terävät kulmat, pohjana olevan suunnikkaan ala on $\|u \times v\|$ ja särmiön korkeus on $\|w\|\cos\beta$, joten sen tilavuus on $\|u\times v\|\|w\|\cos\beta$, siis sama kuin "skalaarikolmitulo" $((u\times v)|w)$.

Laskemalla koordinaateissa voi todeta, että tämä on juuri kolmirivinen determinanttimme.

Esitämme determinanttia käsittelevässä luvussa 2 todistuksen, joka pätee kaikissa tapauksissa ja joka myös toimii kaikenulotteisissa avaruuksissa. $\square$

0.9. Kompleksiluvut tulkittuna tasoksi $\mathbb{R}^2$ ja suoraksi $\mathbb{C}^1$.

Kompleksilukujen joukko $\mathbb{C}$ voidaan tunnetusti samaistaa tasoon $\mathbb{R}^2$ pitämällä kompleksilukua $z = x + iy$ pisteenä (x, y). Kompleksiluvun z *itseisarvo* $|z|$ on vastaavan pisteen etäisyys origosta. Käyttämällä napakoordinaatteja saa kompleksiluvulle esityksen $z = (x, y) = (r \cos \theta, r \sin \theta)$.



Kompleksilukujen $z = (x, y)$ ja $w = (a, b)$ *tulo* lasketaan näin käyttäen ehtoa $i^2 = -1$:

$$zw = (x + iy)(a + ib) = xa + xib + iya + i^2yb = (xa - yb) + i(xb + ya).$$

Kompleksilukujen tulo on sisä- ja ristitulosta eroava tapa kertoa keskenään kaksi vektoria. Se toimii vain kaksiulotteisessa tasossa.

Kompleksilukujen algebralliset laskutoimitukset noudattavat täysin samoja laskulakeja kuin reaalilukujen laskutoimituksetkin.[14] Siksi voisimme muodostaa vektoreita yhtä hyvin lähtemällä kompleksiluvuista. Kompleksilukujen pareja $(z, w) \in \mathbb{C}^2$ voisi sanoa *kompleksisiksi kaksikomponenttisiksi vektoreiksi* jne. Palaamme tähän ideaan vasta kirjan lopussa.[15]

0.10. Trigonometrian kaavoja.

Ei voi olla vahingoksi pitää mielessä, että $\sin^2 x + \cos^2 x = 1$, kuten jo Pythagoraan kerrotaan tienneen. Viisas muistaa lisäksi sinin ja kosinin summakaavat, joita on edellä jo käytettykin laskettaessa kierrolle matriisia

$$\sin(x + y) = \sin x \cos y + \cos x \sin y$$

$$\cos(x + y) = \cos x \cos y - \sin x \sin y,$$

[14] Yhteisiä laskulakeja sanotaan *kunta-aksioomiksi.*

[15] Viimeisessä luvussa laskeskellaan kompleksiluvuilla. Siltä varalta, että lukijakin näin tekee, huomautan siitä, että kompleksiluvun neliö voi olla negatiivinen ja siksi myös kompleksisen vektorin "sisätulo" itsensä kanssa voisi olla negatiivinen, jolloin "vektorin pituus" tulisi imaginaariseksi. Ongelmalta on tapana välttyä määrittelemällä *kompleksisten vektoreiden sisätulo* lausekkeeksi $(x|y) = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n} = \sum_{i=1}^{n} x_i\overline{y_i}$, missä ylleviivauksella on merkitty kompleksikonjugointia eli liittolukua. Kompleksisessa lineaarialgebrassa on luonnollista sanoa "kompleksitasoa" $\mathbb{C}$ "kompleksiseksi suoraksi".

eikä opettele erotuksen kaavoja, koska summassa tietysti saa olla negatiivinen y. Sen sijaan sinin parittomuus

$$\sin(-x) = -\sin x$$

ja kosinin parillisuus

$$\cos(-x) = \cos x$$

on parasta pitää mielessä.

1. Lineaarinen yhtälöryhmä ja matriisi

Lineaarialgebra ja tasojen ja suorien geometria yleistyksineen ovat saman asian kaksi hyvin erinäköistä puolta. Geometria antaa asialle havainnollisen ja myös syvällisen sisällön ja runsaasti sovelluksia. Myös tässä luvussa esiteltävä algebrallinen puoli sisältää syviä oivalluksia.[16] Lisäksi algebrallinen puoli sisältää tehokkaita laskennallisia menettelytapoja. Tässä luvussa esitellään keskeinen ongelma — lineaarisen yhtälöryhmän ratkaiseminen, kun siinä on mielivaltaisen monta yhtälöä ja tuntematonta.

Johdantoluvussa emme perustelleet kaikkia väitteitä aukottomasti. Koetamme nyt olla huolellisempia, koska tarkoituksena on esittää yhtälöryhmän täydellinen ratkaisu kaikissa tapauksissa. Siksi noudatamme tästä alkaen matemaattisen tekstin vakiintunutta muotoa. Lausumme tarkastetut tosiasiat numeroituina "lauseina" merkkiin $\square$ päättyvine todistuksineen ja otamme käyttöön uusia käsitteitä numeroiduin "määritelmin". Välissä on lukemisen helpottamiseksi paljon muutakin tekstiä, mutta teorian kannalta sen voisi periaatteessa jättää pois. Varoen sekoamasta johdantoluvun todistamattomin väitteisiin esitämme periaatteessa kaiken alusta alkaen uudelleen.

1.1. Alustava yritelmä — kaksi yhtälöä ja kaksi tuntematonta.

Tarkastelemme aluksi helponnäköistä erikoistapausta, yhtälöparia

$$\text{(1)} \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2, \end{cases}$$

missä kertoimet $a_{11}, \ldots, a_{22}$, b_1 ja b_2 ovat reaalilukuja sekä x_1 ja x_2 ovat tuntemattomia. Lukion kurssin perusteella muistamme, että yhtälöparin ratkaisun luonne riippuu luvusta $D = a_{11}a_{22} - a_{12}a_{21}$.

1.1.1. Lause.

a) *Jos $D \neq 0$, niin yhtälöparilla (1) on täsmälleen yksi ratkaisu, nimittäin*

$$\begin{cases} x_1 = \dfrac{a_{22}b_1 - a_{12}b_2}{D} \\ x_2 = \dfrac{a_{11}b_2 - a_{21}b_1}{D}. \end{cases}$$

b) *Jos taas $D = 0$, niin yhtälöparilla (1) joko ei ole ratkaisua, tai sillä on äärettömän monta ratkaisua.*

Todistushahmotelma. Tapauksessa $D \neq 0$ voi sijoittaa annetut lausekkeet yhtälöihin ja todeta sieventämällä, että ne toteutuvat. Toisaalta voi pienellä vaivannäöllä todistaa sen, että muita ratkaisuja ei ole. Tämä tapahtuu vaikkapa ratkaisemalla ensimmäisestä yhtälöstä x_1 toisen tuntemattoman lausekkeena

$$x_1 = \frac{b_1 - a_{12}x_2}{a_{11}},$$

[16] Matriisilaskenta on yllättävän uutta, vasta viime vuosisadan loppupuolella kehittynyttä.

sijoittamalla se toiseen yhtälöön ja ratkaisemalla x_2. Laskussa joutuu lopuksi jakamaan luvulla D, joka esiintyy ratkaisun nimittäjänä, mutta se ei haittaa, koska oletettiin, että $D \neq 0$. Itse asiassa jaoimme myös luvulla a_{11}. Jos sattuu olemaan $a_{11} = 0$, niin tämä oli laitonta ja joudumme ratkaisemaan yhtälöryhmän esimerkiksi lausumalla ensin toisen tuntemattoman x_2 vastaavasti ensimmäisen avulla, jolloin jaetaan kertoimella a_{12}. Toinen vaihtoehto olisi ratkaista x_1 tai sitten x_2 ensin alemmasta yhtälöstä, jolloin jaettaisiin a_{21}:lla tai a_{22}:lla. Jokin kertoimista on varmasti nollasta eroava, koska $D \neq 0$, niin että laskun voi aina suorittaa tavalla tai toisella. Kussakin tapauksessa johtopäätökseksi saadaan, että kaavan ilmoittama ratkaisu on ainoa.

Tapauksessa $D = 0$ tulee mieleen ainakin aluksi menetellä periaatteessa samalla tavalla kuin tapauksessa $D \neq 0$, siis jakaa ylempi yhtälö puolittain kertoimella a_{11}, jos se ei satu olemaan 0. Näin saadaan yhtälöpari

$$\begin{cases} x_1 = \dfrac{b_1 - a_{12}x_2}{a_{11}} \\ \underbrace{\dfrac{-a_{21}a_{12} + a_{22}a_{11}}{a_{11}}}_{0} x_2 = b_2 - \frac{a_{21}b_1}{a_{11}}. \end{cases}$$

Jos alemman yhtälön oikea puoli eroaa nollasta, ei yhtään ratkaisua ole. Muuten ratkaisuun voidaan valita x_2 vapaasti ja sitten x_1 siten, että ylempi yhtälö toteutuu. Ratkaisujen määrä on todella nolla tai ääretön. Tapauksessa $a_{11} = 0$ jaetaan jollain muulla kertoimella ja tehdään vastaavat johtopäätökset. Jos kaikki kertoimet ovat nollia, yhtälöryhmä on helppo. $\square$

Huomaamme todistaneemme lauseen kömpelösti. Tapauksia tuli paljon. Menettelyn yleistäminen kolmelle tai useammalle yhtälölle näyttää hankalalta. Tulos puolestaan tuntuu järkevältä, onhan sillä geometrinen tulkinta. Kahden muuttujan lineaarisen yhtälön ratkaisujen (x_1, x_2) joukkohan on suora tasossa $\mathbb{R}^2$, elleivät yhtälössä molempien tuntemattomien kertoimet ole nollia. Yhtäløparin ratkaisujen joukko on näin ollen kahden suoran leikkaus. Pienellä laskulla voi tarkastaa, että yhtälöiden ratkaisujoukot ovat tapauksesta riippuen

- toisiaan leikkaavat suorat, jos $D \neq 0$;
- yhdensuuntaiset (ei ratkaisua) tai yhtyvät (∞ monta ratkaisua) suorat, jos $D = 0$.

Tarkoituksena on nyt kuitenkin harrastaa nimenomaan laskentoa. On ilmeisesti tarpeen kehittää järjestelmällinen lasku- ja kirjanpitomenetelmä laajempien yhtälöryhmien ratkaisemiseksi.

1.2. Yleinen lineaarinen yhtälöryhmä ja sen ratkaiseminen GAUSSIN **ja** JORDANIN[17] **eliminointimenetelmällä.**

Tarkastelemme lineaarista yhtälöryhmää

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

missä m ja n ovat luonnollisia lukuja, kertoimet a_{ij} ja b_i reaalilukuja sekä $x_1, \ldots, x_n$ tuntemattomia. Yhtälöryhmässä (1) voi olettaa ja oletetaan, että kullakin sarakkeella on ainakin yksi nollasta eroava kerroin, $a_{ij} \neq 0$, muutenhan tuntematon x_j ei esiinny lainkaan vaan voi saada mitä tahansa arvoja. Yhtälöryhmän (1) toteuttavaa lukujonoa $(x_1, x_2, \ldots, x_n)$ sanotaan sen *ratkaisuksi* ja kaikkien ratkaisujen joukkoa sen *ratkaisujoukoksi*, toisinaan myös sen *yleiseksi ratkaisuksi*. Yhtälön *ratkaiseminen* on ratkaisujoukon määräämistä.

Gaussin ja Jordanin eliminointimenetelmä (1):n ratkaisemiseksi perustuu yhtälöryhmän sieventämiseen yksinkertaisin laskutoimituksin, kunnes se on muuttunut hyvin helpoksi ratkaista[18]. Sieventämiseen käytetään kolmea *rivioperaatiota*, jotka ovat seuraavat

- Operaatio P_{ij}: Vaihdetaan yhtälöt i ja j keskenään.
- Operaatio $M_i(c)$: Kerrotaan yhtälö i nollasta eroavalla luvulla c.
- Operaatio $A_{ij}(c)$: Lisätään jollakin luvulla c kerrottu yhtälö i yhtälöön j, missä $i \neq j$.

1.2.1. LAUSE. *Edellä mainitut operaatiot eivät muuta (1):n ratkaisuja, ts. uusi yhtälöryhmä on yhtäpitävä eli ekvivalentti (1):n kanssa.*

TODISTUS. Jätetään harjoitustehtäväksi. □

Rivioperaatioissa tarvitsee käsitellä vain yhtälöryhmän kertoimia. Siksi yhtälöryhmä kannattaa kirjoittaa laskelmien ajaksi ilman muuttujia muotoon

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right],$$

jota sanotaan *laajennetuksi kerroinmatriisiksi.*

Ratkaisumenettely eli *algoritmi* on nyt seuraava:[19] Aloitetaan laajennetun kerroinmatriisin **sarakkeen 1** käsittelyllä:

- Jos laajennetun kerroinmatriisin vasemmassa yläkulmassa eli paikassa (1,1) on $a_{11} = 0$, vaihdetaan ensimmäiseksi jokin sellainen yhtälö, jossa ensimmäinen kerroin ei ole 0. (Operaatio P_{1j}.)

[17] CARL FRIEDRICH GAUSS (1777–1855) "Matemaatikoiden kuningas"; CAMILLE JORDAN (1838–1922), ranskalainen algebrikko.

[18] Ks. tilannearvio 1.2.2.

[19] Algoritmin kuvailua on varmasti ikävä lukea sellaisenaan. Kannattaa mieluummin soveltaa ohjetta vaikkapa umpimähkään valittuun neljän yhtälön ja tuntemattoman ryhmään. Harjoitustehtävät valaisevat eri mahdollisuuksia, miten laskussa voi käydä.

- Kerrotaan saadun yhtälöryhmän ylin yhtälö paikassa (1,1) olevan kertoimen käänteisluvulla, jolloin saadaan paikkaan (1,1) luku 1. (Operaatio M_1.)
- Operaatioilla A_{1j} järjestetään luku 0 kerroinmatriisin ensimmäisen sarakkeen muihin paikkoihin, siis paikkoihin (2,1), (3,1),...,(m,1).

Yhtälöryhmän laajennetun kerroinmatriisin ensimmäisessä sarakkeessa on nyt ylimpänä ykkönen ja muuten pelkkiä nollia. Muuntamista jatketaan pyrkien saattamaan toinenkin sarake lähes samaan muotoon.

Sarakkeen 2 käsittely riippuu tilanteesta. Jos paikoissa (2,2),...,(m,2) on 0, niin siirrytään suoraan sarakkeen 3 käsittelyyn. Tällöin toisen sarakkeen ylimmässä paikassa (1,2) oleva kerroin jää siksi luvuksi mikä se sattuu olemaan ensimmäisen sarakkeen käsittelyn jälkeen.

Jos taas toisessa sarakkeessa esiintyy rivillä 2 tai alempana edes yksi nollasta eroava kerroin, niin menetellään seuraavasti:

- Jos paikassa (2,2) on 0, toinen yhtälö vaihdetaan johonkin alempaan, jossa toinen kerroin on nollasta eroava. (Operaatio P_{2j}; tässä on oltava $j \geq 3$.)
- Operaatiolla M_2 saadaan paikkaan (2,2) luku 1.
- Operaatioilla A_{ij} saadaan kaikkiin[20] paikkoihin (1,2), (3,2),...,(m,2) luku 0.

Selvästi mikään toisen sarakkeen käsittelytavoista ei enää muuta yhtälöryhmän ensimmäistä saraketta. Laajennetun kerroinmatriisin toisessa sarakkeessa on nyt kohdan (2,2) alapuolella pelkkiä nollia. Kohdassa (2,2) on joko 1 tai 0. Jos siinä on 1, niin toisen sarakkeen kaikki muut kertoimet ovat nollia — myös ylin kerroin.

Sarakkeen 3 käsittely jätetään taas tekemättä, mikäli kohdassa (3,3) ja sen alapuolella on pelkkiä nollia. Muuten menetellään kuten sarakkeen 2 kohdalla, paitsi että tietenkään ei toiseksi, vaan kolmanneksi yhtälöksi vaihdetaan tarvittaessa jokin alempi, jossa kolmas kerroin on nollasta eroava. Kun kolmannen yhtälön kolmas kerroin on muutettu ykköseksi operaatiolla M_3, järjestetään kaikkiin muihin yhtälöihin nollat operaatiolla A_{3j} — myös ylimpään kahteen.

Selvästi mikään kolmannen sarakkeen käsittelytavoista ei enää muuta kahta ensimmäistä saraketta. Laajennetun kerroinmatriisin kolmannessa sarakkeessa on käsittelyn jälkeen kohdan (3,3) alapuolella pelkkiä nollia. Kohdassa (3,3) on joko 1 tai 0. Jos siinä on 1, niin kolmannen sarakkeen kaikki muut kertoimet ovat nollia.

Edellä selostettua menettelyä sovelletaan järjestyksessä kaikkiin pystyviivan vasemmalla puolella oleviin sarakkeisiin. Käsittelyn jälkeen laajennetussa kerroinmatriisissa pystyviivan vasemmalla puolella on seuraavaa:

1. Puhtaat nollarivit (siis sellaiset, joissa pystyviivan vasemmalla puolella on vain nollia) ovat alimmaisina. Voimme huolehtia siitä, että näistä alimmaisina ovat ne rivit, joissa myös vakio b_n on 0.
2. Riveillä, jotka eivät ole nollarivejä, on voimassa:
   - Rivin ensimmäinen nollasta eroava luku on 1.
   - Alemman rivin ensimmäinen 1 on ylemmän rivin ensimmäisen 1:n oikealla puolella, ei kuitenkaan aina viereisellä sarakkeella.
   - Rivin ensimmäisen 1:n ylä– ja alapuolella on vain nollia.

[20] Huomaa, että tässä toisessa tapauksessa myös paikkaan (1,2), siis ensimmäiselle riville järjestetään nolla. Jos tätä ei tehdä, on kyseessä *Gaussin eliminointimenetelmä*, joka myös johtaa yhtälöryhmän ratkaisemiseen. Selostamassamme Gaussin ja Jordanin menetelmässä sievennetään mahdollisimman paljon. Riippuu yhtälöistä, kumpi menetelmä on sujuvampi käyttää.

Vastaava yhtälöryhmä on pysynyt koko ajan ekvivalenttina alkuperäisen kanssa ja on helppo ratkaista aloittamalla alimmasta yhtälöstä ja etenemällä ylöspäin.

1.2.2. Tilannearvio. Käsittelyn jälkeen esiintyy joku seuraavista tilanteista:

a) Yhtälöt ovat muokkautuneet muotoon

$$\left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & d_1 \\ 0 & 1 & \dots & 0 & d_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & d_n \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array}\right], \quad \text{eli} \quad \begin{cases} x_1 = d_1 \\ x_2 = d_2 \\ \vdots \\ x_n = d_n \end{cases},$$

jolloin yhtälöryhmällä on selvästikin täsmälleen yksi ratkaisu.

b) Viimeisin yhtälö, joka ei ole muotoa $0 = 0$, on muotoa

$$x_j + c_{j\,(j+1)} x_{j+1} + \cdots + c_{jn} x_n = \text{ vakio.}$$

Tällöin yhtälöryhmällä on äärettömän monta ratkaisua. Esimerkiksi yhtälöryhmä

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 4 & 2 \\ 0 & 1 & 0 & 8 & 7 \\ 0 & 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad \text{eli} \quad \begin{cases} x_1 + 4x_4 = 2 \\ x_2 + 8x_4 = 7 \\ x_3 + 3x_4 = 3 \end{cases}$$

on tätä tyyppiä ja vaikkapa x_4 voidaan valita vapaasti.

Erikoistapauksena tapauksesta b) on tilanne, jossa viimeisin yhtälö, joka ei ole muotoa $0 = 0$, on muotoa $x_p =$ vakio, mutta yhtälöt eivät ole muokkautuneet muotoon a). Tällaista tilannetta edustavat esimerkiksi

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad \text{eli} \quad \begin{cases} x_1 + x_2 = 2 \\ x_3 = 8, \end{cases}$$

jossa vaikkapa x_2 voidaan valita vapaasti, ja

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right], \quad \text{eli} \quad \begin{cases} x_1 = 2 \\ x_2 = 8 \\ x_3 = 0, \end{cases}$$

jossa x_4 voidaan valita vapaasti.

c) Viimeisin yhtälö, joka ei ole muotoa $0 = 0$, on muotoa $0 = c$, missä c on nollasta eroava luku. Tällöin yhtälöryhmällä ei ole ratkaisua. Esimerkki tästä on

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 4 & 2 \\ 0 & 1 & 0 & 8 & 7 \\ 0 & 0 & 0 & 0 & 4567 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad \text{eli} \quad \begin{cases} x_1 + 4x_4 = 2 \\ x_2 + 8x_4 = 7 \\ 0 = 4567\,. \end{cases}$$

Jos yhtälöitä on vähemmän kuin tuntemattomia, siis jos $m < n$, niin vain tapaukset b) tai c) voivat esiintyä. Ratkaisuja on tässä tilanteessa äärettömän monta tai ei yhtään. Jos taas yhtälöitä on vähintään yhtä monta kuin tuntemattomia, niin voi esiintyä mikä tahansa eo. tapauksista a), b) tai c).

1.2.3. LAUSE. *Gaussin–Jordanin menetelmällä on mahdollista ratkaista yleinen lineaarinen yhtälöryhmä (1) kaikissa tapauksissa.*

TODISTUS. Asia on jo selvä. □

1.2.4. HUOMAUTUS. Menetelmä on tuottanut ratkaisun ongelmaamme. Kaivamaan jäi kysymys siitä, miten kahden yhtälön tapauksessa esiintynyt luku, determinantti D, jolla voitiin testata ratkaisujen lukumäärää, yleistyy useamman yhtälön tapaukseen. Gauss–Jordanilla ei ratkaisujen pelkkä määrä näy selviävän ilman, että joutuu ratkaisemaan koko yhtälöryhmän. Tähän asiaan palataan vielä perusteellisesti luvuissa 2 ja 4.

1.3. Homogeeninen yhtälöryhmä.

1.3.1. MÄÄRITELMÄ. Lineaarinen yhtälöryhmä (1) on *homogeeninen*, jos $b_1 = b_2 = \cdots = 0$ ts. yhtälöryhmä on muotoa

$$(2) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

Homogeenisen yhtälöryhmän ratkaiseminen on hieman helpompaa kuin yleisen yhtälöryhmän, sillä homogeenisella yhtälöryhmällä (2) on **aina** yhtenä ratkaisuna ns. *triviaaliratkaisu*[21] $x_1 = x_2 = \cdots = x_n = 0$, eikä siis tapaus c) edellä voi esiintyä homogeeniyhtälölle. Vaihtoehdot a) ja b) merkitsevät, että jos epätriviaaleja ratkaisuja on, niitä on äärettömän monta. Erityisesti, jos $n > m$, niin (2):lla on aina äärettömän monta ratkaisua.

1.3.2. YHTEENVETO. *Seuraavat matriisin A ominaisuudet ovat yhtäpitäviä keskenään*

- *Homogeeniyhtälöllä $Ax = 0$ on ainoastaan triviaaliratkaisu $x = 0$.*
- *Millään $b \in \mathbb{R}^m$ ei yhtälöllä $Ax = b$ ole enempää kuin yksi ratkaisu.*
- *Kuvaus $x \mapsto Ax$ on injektio $\mathbb{R}^n \to \mathbb{R}^m$.*
- *Gaussin ja Jordanin menettely homogeeniyhtälölle johtaa tapaukseen a), toisin sanoen matriisi A muokkautuu rivioperaatioilla muotoon*

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

[21]"Triviaali" merkitsee matematiikassa alkeellista tai itsestäänselvää.

1.4. Vektorit ja matriisit.

On helppoa ymmärtää tilanne, jossa yhtälöryhmällä ei ole yhtään ratkaisua tai vain yksi ratkaisu, joka on yksi lukujono $(x_1, x_2, \ldots, x_n)$. Sen sijaan äärettömän monen ratkaisun muodostaman joukon kuvailu saattaa olla hankalampaa. Mitä yhtälöryhmän ratkaiseminen oikestaan on? Voisihan näsäviisaasti sanoa, että alkuperäinen yhtälöryhmä itsessään jo kuvailee, mitkä lukujonot $(x_1, x_2, \ldots, x_n)$ kuuluvat rakaisujoukkoon — tietysti ne, joilla yhtälöt toteutuvat. Johdantoluvun lukeneina ymmärrämme kuitenkin, että Gaussin–Jordanin menettely ratkaisee yhtälöryhmän siinä mielessä, että se antaa meille ratkaisujoukon parametrisoinnin, jonka avulla sen voi tyypillisesti vaikka piirtää. Lineaarisen yhtälön ratkaisujoukon geometrian selvittäminen on yksi tämän kurssin tavoitteista, jonka saavutamme vähitellen. Ensimmäinen askel tähän suuntaan on kunkin ratkaisun $(x_1, x_2, \ldots, x_n)$ tulkinta n–ulotteisen avaruuden vektoriksi.

Määritelmä 1.4.1. Reaaliluvuista $x_1, x_2, \ldots, x_n$ muodostettua järjestettyä joukkoa eli jonoa $x = (x_1, x_2, \ldots, x_n)$ sanotaan *n–komponenttiseksi vektoriksi*, usein epätäsmällisesti vain *vektoriksi*. Kaikkien $n-$komponenttisten vektoreiden joukkoa sanotaan *avaruudeksi* $\mathbb{R}^n$.

Huomaa, että vektorit x ja y ovat samoja, ts. $x = y$, jos ja vain jos $x_i = y_i$ kaikilla i. Yksi vektoreiden välinen yhtälö vastaa siis tavallista yhtälöryhmää.

Määritelmä 1.4.2. Olkoot $x = (x_1, x_2, \ldots, x_n)$ ja $y = (y_1, y_2, \ldots, y_n)$ vektoreita ja $\lambda \in \mathbb{R}$. Määritellään *vektoreiden summa*, *kertominen reaaliluvulla* ja *sisätulo*

$$\begin{aligned}
&x + y = (x_1 + y_1,\, x_2 + y_2, \ldots, x_n + y_n),\\
&\lambda x = (\lambda x_1,\, \lambda x_2, \ldots, \lambda x_n),\\
&(x|y) = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i.
\end{aligned}$$

Määritelmä 1.4.3. Reaaliluvuista a_{ij} muodostettua kaaviota

$$A_{m\times n} = \begin{bmatrix}
a_{11} & a_{12} & a_{13} & \ldots & a_{1j} & \ldots & a_{1n}\\
a_{21} & a_{22} & a_{23} & \ldots & a_{2j} & \ldots & a_{2n}\\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots\\
a_{i1} & a_{i2} & a_{i3} & \ldots & a_{ij} & \ldots & a_{in}\\
\vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots\\
a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mj} & \ldots & a_{mn}
\end{bmatrix}$$

sanotaan *$m\times n$–matriisiksi* ja merkitään myös lyhyemmin $A = [a_{ij}]$. Lukuja a_{ij} sanotaan matriisin A *alkioiksi*[22] ja merkitään joskus myös A_{ij}, etenkin jos matriisilla sattuu olemaan pitkä nimi, esimerkiksi $[\text{pitkäniminen matriisi}]_{ij}$.

Matriisit $A = A_{n\times m}$ ja $B = B_{s\times t}$ ovat samat, ts. $A = B$, jos ja vain jos $n = s$, $m = t$ ja $a_{ij} = b_{ij}$ kaikilla i ja j.

Kuten vektoreille myös matriiseille on tapana määritellä yhteenlasku ja luvulla kertominen alkioittain.

[22]Engl. *matrix element*; joskus myös *entry*.

Määritelmä 1.4.4. Olkoon $A_{m\times n} = [a_{ij}]$, $B_{m\times n} = [b_{ij}]$ ja $\lambda \in \mathbb{R}$. Määritellään matriisien *summa* ja *kertominen reaaliluvulla*:

$$\begin{aligned} A + B &= [a_{ij} + b_{ij}]\,, \\ \lambda A &= [\lambda a_{ij}]\,. \end{aligned}$$

Määritelmässä 1.4.1. tarkasteltiin vektoreita lukujonoina. Matriiseita käytettäessä on yleensä edullista kirjoittaa vektori *sarakevektorina, pystyvektorina* eli *sarakematriisina* muotoon

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

Harvemmin vektori samaistetaan *riviv ektoriin, vaakavektoriin* eli *rivimatriisiin.*

$$x = [x_1, x_2, \ldots, x_n]\,.$$

Vektori on siis kahdella eri tavalla matriisin erikoistapaus, $n\times 1$ – tai $1\times n$ –matriisi. Matriisi A esitetään usein rivi– tai sarakevektoreidensa eli *riviensä* ja *sarakkeidensa* avulla, esim.[23]

$$A = \begin{bmatrix} \overrightarrow{a}_{1*} \\ \overrightarrow{a}_{2*} \\ \vdots \\ \overrightarrow{a}_{m*} \end{bmatrix}, \text{ missä } \begin{cases} \overrightarrow{a}_{1*} = [a_{11}, a_{12}, \ldots, a_{1n}] \\ \overrightarrow{a}_{2*} = [a_{21}, a_{22}, \ldots, a_{2n}] \\ \vdots \\ \overrightarrow{a}_{m*} = [a_{m1}, a_{m2}, \ldots, a_{mn}]\,. \end{cases}$$

tai

$$B = \left[\overrightarrow{b}_{*1}, \overrightarrow{b}_{*2}, \ldots, \overrightarrow{b}_{*n}\right], \quad \text{missä} \quad \overrightarrow{b}_{*1} = \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{bmatrix}, \quad \overrightarrow{b}_{*2} = \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{m2} \end{bmatrix}, \cdots \overrightarrow{b}_{*n} = \begin{bmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{mn} \end{bmatrix}.$$

Kahden matriisin tulo — matriisitulo — määritellään niiden rivi- ja sarakevektoreiden sisätulojen avulla. Matriisitulokin on matriisi. Johdannon lukeneina tiedämme, että matriisit on suunniteltu esittämään lineaarikuvauksia ja niiden tulo kannattaa määritellä siten, että se vastaa kuvaustan yhdistämistä. Emme johdattele aiheeseen enää uudelleen, vaan asetamme muodollisen määritelmän. Selvittelemme myöhemmin, mihin se kelpaa:

[23]Tässä olen taas kerran varmuuden vuoksi merkinnyt sarakkeiden symbolien päälle nuolet osoittamaan, että kyseessä ovat vektorit. Yleensä näin ei menetellä, vaan lukijan on itse pidettävä kirjaa siitä, mikä merkki edustaa minkäkintyyppistä objektia.

Määritelmä 1.4.5. Olkoon

$$A_{m\times n} = [a_{ij}] = \begin{bmatrix} \overrightarrow{a}_{1*} \\ \overrightarrow{a}_{2*} \\ \vdots \\ \overrightarrow{a}_{m*} \end{bmatrix},$$

missä $\overrightarrow{a}_{i*}$ on A:n i:s rivivektori ja olkoon

$$B_{n\times p} = [b_{ij}] = \left[\overrightarrow{b}_{*1}, \dots, \overrightarrow{b}_{*j}, \dots, \overrightarrow{b}_{*p}\right],$$

missä $\overrightarrow{b}_{*j}$ on B:n j:s sarakevektori. Määritellään matriisien A ja B *matriisitulo* eli *tulo*:

$$AB = AB_{m\times p} = \begin{bmatrix} \left(\overrightarrow{a}_{1*}|\overrightarrow{b}_{*1}\right) & \dots & \left(\overrightarrow{a}_{1*}|\overrightarrow{b}_{*j}\right) & \dots & \left(\overrightarrow{a}_{1*}|\overrightarrow{b}_{*p}\right) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \left(\overrightarrow{a}_{i*}|\overrightarrow{b}_{*1}\right) & \dots & \left(\overrightarrow{a}_{i*}|\overrightarrow{b}_{*j}\right) & \dots & \left(\overrightarrow{a}_{i*}|\overrightarrow{b}_{*p}\right) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \left(\overrightarrow{a}_{m*}|\overrightarrow{b}_{*1}\right) & \dots & \left(\overrightarrow{a}_{m*}|\overrightarrow{b}_{*j}\right) & \dots & \left(\overrightarrow{a}_{m*}|\overrightarrow{b}_{*p}\right) \end{bmatrix}.$$

Matriisitulo AB on siis matriisi, jonka i:nnen rivin j:s alkio on vasemman kerrottavan matriisin A i:nnen rivin ja oikeanpuoleisen kerrottavan matriisin B j:nnen sarakkeen sisätulo

$$[AB]_{ij} = (\overrightarrow{a}_{i*}|\overrightarrow{b}_{*j}) = \sum_{k=1}^{n} a_{ik}b_{kj}. \quad \text{(Summausindeksi } k \text{ keskimmäisenä.)}$$

Tulomatriisi muodostetaan siis kertomalla ensimmäisen tekijän rivit toisen tekijän sarakkeilla ja kirjoittamalla saadut sisätulot rivien risteysten kohdalle. Tämä on helppo ja — kuten pian näemme — hyödyllinen asia oppia.

Matriisitulo on siitä kätevä, että se noudattaa monia samoja laskulakeja kuin tavallinen lukujen tulo — ei kuitenkaan kaikkia. Merkittävä ero on vaihdannaisuuden puute.

Huomautus 1.4.6. Matriiseille $A_{m\times n}$ ja $B_{n\times p}$ on yleensä määritelty vain AB, mutta ei BA. Jos kuitenkin $p = m$, niin on olemassa sekä AB että BA. Nämä ovat kuitenkin yleensä eri matriisit, eikä edes ole olemassa mitään kaavaa toisen laskemiseksi toisesta. Matriisitulo **ei** siis ole *vaihdannainen* eli *kommutatiivinen* laskutoimitus.

Tämän näyttää vielä erityisen selvästi seuraava esimerkki. Rivimatriisin $A = A_{1\times m}$ ja sarakematriisin $B = B_{n\times 1}$ matriisitulo AB on määritelty ainoastaan, kun $n = m$, mutta toisessa järjestyksessä rivin ja sarakkeen matriisitulo BA on aina olemassa.

Määritelmä antaa tuloiksi

$$AB = [a_1, a_2, \dots, a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \left[\sum_{k=1}^{n} a_k b_k \right], \quad \text{kun } n = m,$$

$$BA = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} [a_1, a_2, \dots, a_m] = \begin{bmatrix} a_1 b_1 & a_2 b_1 & \dots & a_m b_1 \\ a_1 b_2 & a_2 b_2 & \dots & a_m b_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1 b_n & a_2 b_n & \dots & a_m b_n \end{bmatrix}.$$

AB on 1×1 –matriisi, jonka ainoa alkio on annettujen vektoreiden sisätulo. BA on $n \times m$ –matriisi.[24]

Monessa muussa suhteessa matriisialgebra kuitenkin muistuttaa reaaliluvuilla laskemista. Seuraavat lauseet sisältävät yhteisiä laskussääntöjä.

Lause 1.4.7. *Matriiseille A, B ja C ovat voimassa tavalliset osittelulait, edellyttäen että summat ja tulot ovat määriteltyjä:*

$$\begin{aligned} &a) \qquad A(B+C) = AB + AC, \\ &b) \qquad (A+B)C = AC + BC. \end{aligned}$$

Todistus. Harjoitustehtävä $\square$

Lause 1.4.8. *Matriiseille $A_{m\times n}$, $B_{n\times p}$ ja $C_{p\times q}$ on voimassa liitännäisyys- eli assosiatiivisuuslaki:*

$$A(BC) = (AB)C.$$

Tulo on $m \times q$–matriisi. Voidaan merkitä sulkeitta ABC.

Todistus.

$$\begin{aligned} [A(BC)]_{ij} &= \sum_{k=1}^{n} a_{ik} (BC)_{kj} \\ &= \sum_{k=1}^{n} a_{ik} \left(\sum_{\ell=1}^{p} b_{k\ell} c_{\ell j} \right) \\ &= \sum_{k=1}^{n} \left(\sum_{\ell=1}^{p} a_{ik} b_{k\ell} c_{\ell j} \right) \\ &= \sum_{\ell=1}^{p} \left(\sum_{k=1}^{n} a_{ik} b_{k\ell} c_{\ell j} \right) \\ &= \sum_{\ell=1}^{p} \left(\sum_{k=1}^{n} a_{ik} b_{k\ell} \right) c_{\ell j} \\ &= \sum_{\ell=1}^{p} (AB)_{i\ell}\, c_{\ell j} \\ &= [(AB)C]_{ij}. \quad \square \end{aligned}$$

[24] Sitä sanotaan vaaka- ja pystyvektorin *tensorituloksi*.

Taustaa. Johdannon mukaan $m \times n$–matriisista voi muodostaa lineaarikuvauksen

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$

valitsemalla kantavektoreiden kuviksi sarakkeet.

$$T_A(e_i) = \overrightarrow{a}_{*i} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} \in \mathbb{R}^m.$$

On totta, että lineaarikuvausten $T_A : \mathbb{R}^n \to \mathbb{R}^m$ ja $T_B : \mathbb{R}^p \to \mathbb{R}^n$ yhdistetty kuvaus on lineaarikuvaus, jonka matriisi on matriisitulo AB, toisin sanoen

$$T_{AB} = T_A \circ T_B.$$

Jos tämän tietää[25], niin ymmärtää, että matriisitulon väitetty liitännäisyys seuraa laskuitta siitä, että kuvausten yhdistäminen on liitännäistä — siis $(f \circ g) \circ h = f \circ (g \circ h)$. Palaamme asiaan luvussa 4.

1.5. Neliömatriisin käänteismatriisi.
Jos merkitään

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \; \text{ja } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

niin yleinen lineaarinen yhtälöryhmä

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

voidaan kirjoittaa matriisitulon avulla tiiviiseen muotoon

$$(2) \qquad Ax = b.$$

Kun nyt olemme ajatelleet lauseketta Ax kahden matriisin tulona, herää ajatus, että yhtälön (2) voisi ehkä ratkaista ”jakamalla molemmat puolet matriisilla A”. Osoittautuu, että tämä idea ei ole täysin järjetön, vaan kuten luvuilla laskettaessa, matriisienkin kertolaskussa on olemassa ykkönen ja monilla matriiseilla A — joskaan ei kaikilla — on todella käänteismatriisi, jolla kertomalla A:sta tulee matriisiykkönen ja yhtälö (2) ratkeaa muodossa $x = A^{-1}Ax = A^{-1}b$.

[25] Vaikkapa luettuaan lauseen 4.3.4.

Määritelmä 1.5.1.
$A_{n\times n}$ on *neliömatriisi*.
Neliömatriisi A on *lävistäjä-* eli *diagonaalimatriisi*, jos $a_{ij} = 0$ kaikille $i \neq j$, eli kun vain matriisin *päälävistäjällä* eli *diagonaalilla* mahdollisesti on nollasta eroavia alkioita.
Neliömatriisi A on *yläkolmiomatriisi*, jos $a_{ij} = 0$ kaikille $i < j$, eli kun päälävistäjän alapuolella on vain nollia.
Neliömatriisi A on *alakolmiomatriisi*, jos $a_{ij} = 0$ kaikille $i < j$, eli kun päälävistäjän yläpuolella on vain nollia.

$$\underset{\text{yläkolmiomatriisi}}{\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}} \quad \underset{\text{alakolmiomatriisi}}{\begin{bmatrix} 1 & 0 & 0 \\ 8 & 1 & 0 \\ 7 & 6 & 2 \end{bmatrix}} \quad \underset{\text{diagonaalimatriisi}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}}$$

Diagonaalimatriisia

$$I = I_{n\times n} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = [\delta_{ij}]\,, \text{ missä } \begin{cases} \delta_{ij} = 1, \text{ kun } i = j, \\ \delta_{ij} = 0, \text{ kun } i \neq j, \end{cases}$$

sanotaan *identtiseksi matriisiksi, ykkösmatriisiksi* tai *yksikkömatriisiksi*. Ykkösmatriisilla on matriisialgebrassa ykkösen rooli.

Lause 1.5.2. *Jokaiselle matriisille $A_{m\times n}$ ja erityisesti siis vektorille $x \in \mathbb{R}^n$ on*

$$AI_{n\times n} = I_{m\times m}A = A,$$
$$I_{n\times n}x = x.$$

Todistus. Helppo. $\square$

Määritelmä 1.5.3. Neliömatriisi $A_{n\times n}$ on *kääntyvä* eli *säännöllinen*, jos on olemassa neliömatriisi $B_{n\times n}$ siten, että

$$AB = BA = I.$$

Tällöin merkitään $B = A^{-1}$ ja $A = B^{-1}$ ja sanotaan, että A^{-1} on A:n *käänteismatriisi*, jolloin A ja B ovat toistensa käänteismatriisit.

Lause 1.5.4. *Olkoot matriisit $A_{n\times n}$ ja $B_{n\times n}$ kääntyviä. Silloin tulo AB on kääntyvä ja*

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

Todistus.

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I \quad \text{ja}$$
$$(B^{-1}A^{-1})(AB) = B(AA^{-1})B^{-1} = BIB^{-1} = BB^{-1} = I.$$

$\square$

LAUSE 1.5.5. *Matriisilla $A_{n\times n}$ on vain yksi käänteismatriisi ja*

$$\left(A^{-1}\right)^{-1} = A.$$

TODISTUS. Olkoot B ja C matriisin A käänteismatriiseja, jolloin siis erityisesti $AB = CA = I$. Nyt pätee: $C = CI = C(AB) = (CA)B = IB = B$. Siis $C = B$ ja A on B:n käänteinen. □

TILANNEKATSAUS 1.5.6. Jos neliömatriisin A käänteismatriisi tunnetaan, niin lineaarisen yhtälöryhmän

$$Ax = b$$

voi todella ratkaista yksinkertaisesti kertomalla sen puolittain A^{-1}:lla:

$$\begin{aligned} A^{-1}Ax &= A^{-1}b \\ Ix &= A^{-1}b \\ x &= A^{-1}b. \end{aligned}$$

Ongelmaksi jää käänteismatriisin olemassaolon selvittäminen ja myönteisessä tapauksessa sen laskeminen. Osoittautuu, että tämä voidaan tehdä Gaussin ja Jordanin menetelmällä suunnilleen samalla vaivalla kuin yhtälön ratkaiseminen.[26] Laskemalla käänteismatriisi saadaan se hyöty, että yhtälöryhmä (1) on tullut ratkaistuksi saman tien kaikille vakioille b.

ALGORITMI 1.5.7 MATRIISIN KÄÄNTÄMISEKSI. *Näin se käy:*

- *Kirjoitetaan laajennettu kerroinmatriisi $(A|I)$. (Viivan oikealla puolella on nyt vektorin b tilalla identtinen matriisi.)*
- *Viedään läpi Gaussin ja Jordanin rivimenettely.*
- *Jos A muuttui identtiseksi matriisiksi, on viivan oikealla puolella A^{-1}; muussa tapauksessa A ei ole kääntyvä.*

TODISTUS ALGORITMIN 1.5.7 TOIMIVUUDELLE. Matriisin kääntämiseksi on ratkaistava yhtälöstä $AX = I$ neliömatriisi X. Auki kirjoitettuna tämä yhtälö on

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Matriisitulon määritelmän mukaan tämä merkitsee, että $A\,x_{*1} = e_1$, $A\,x_{*2} = e_2$ ja yleensäkin $A\,x_{*n} = e_n$, missä x_{*j} on etsityn matriisin j:s sarake ja e_i on standardikantavektori. Ratkaisemme saadut n yhtälöryhmää $A\,x_{*j} = e_j$ Gaussin–Jordanin menetelmällä. Koska kaikissa on sama kerroinmatriisi A, toimitus voidaan suorittaa kaikille kerralla. Juuri näin algoritmi käskee tekemään. Koska A on neliömatriisi, on yhtälöryhmän ratkaisu tilannearvion 1.2.2 a) mukaisesti olemassa ja yksikäsitteinen

[26]Tehokkaiden matriisinkääntömenetelmien keksiminen on hyvin tarpeellista. Suuria lineaarisia yhtälöryhmiä — tuhansia yhtälöitä — esiintyy matematiikan sovelluksissa useinkin.

tasan silloin, kun A saadaan rivioperaatioilla lopulta muunnetua ykkösmatriisiksi. Tällöin siis on löydetty matriisi X, jolle $AX = I$. Käänteismatriisin määritelmä vaatii, että myös $XA = I$. Tämä on algoritmin käytön jälkeen hyvä tarkastaa aina erikseen, mutta syynä ei ole se, että XA voisi olla muuta kuin I, vaan yksinomaan tarve varmistaa, että ei ole tehty laskuvirheitä. Seuraavan lauseen mukaan nimittäin $AX = I$ riittää takaamaan, että neliömatriisi A on kääntyvä. □

LAUSE 1.5.8. *Neliömatriisille $A = A_{n\times n}$ seuraavat ehdot ovat yhtäpitäviä keskenään:*

(1) *Jokaisella yhtälöryhmällä $Ax = b$ ($b \in \mathbb{R}^n$) on täsmälleen yksi ratkaisu*
(2) *Homogeenisella yhtälöryhmällä $Ax = 0$ on vain triviaaliratkaisu.*
(3) *Gaussin ja Jordanin menettely muuntaa matriisin $A = A_{n\times n}$ yksikkömatriisiksi $I = I_{n\times n}$.*
(4) *On olemassa $X = X_{n\times n}$ siten, että $AX = I$, ts. X on A:n oikeanpuoleinen käänteismatriisi.*
(5) *On olemassa $X = X_{n\times n}$ siten, että $XA = I$. X on A:n vasemmanpuoleinen käänteismatriisi.*
(6) *A on kääntyvä, ts. on olemassa $X = X_{n\times n}$ siten, että $XA = AX = I$.*

TODISTUKSESTA. Väitteiden (1), (2) ja (3) yhtäpitävyys saadaan Gaussin ja Jordanin menettelyä koskevasta lauseesta 1.3.2 tapauksena $m = n$. Ehtojen (3) ja (4) yhtäpitävyys tarkastettiin juuri äsken.

Toispuoleisen kääntyvyyden riittävyys kääntyvyydelle, siis ehdot (4) $\Longrightarrow$ (5) ja (5) $\Longrightarrow$ (6) selviävät kuitenkin parhaiten vasta[27] luvuissa 3 ja 4.

Käytännön laskujen kannalta on kuitenkin jo nyt hyvä tietää, että haluttaessa todistaa, että neliömatriisi Xon neliömatriisin A käänteinen, riittää laskea AX tai XA, eikä molempia tarvita.

TAUSTAA 1.5.9. Edellisen lauseen ehto (1) merkitsee sitä, että kuvaus $x \mapsto Ax$ on bijektio $\mathbb{R}^n$ to $\mathbb{R}^n$ ja käänteiskuvaus $b \mapsto x$ saadaan ratkaisemalla yhtälö $Ax = b$. On arvattavissa, että tämä käänteiskuvaus $b \mapsto x$ on lineaarikuvaus, jolloin sen matriisi muodostuu kantavektoreiden alkukuvista ja on siis täsmälleen edellä löytämämme oikeanpuoleinen käänteismatriisi X. Yhtälön $XA = I$ laskennallinen tarkastaminen vaatii kuitenkin tätä — sinänsä luvussa 4 helposti pääteltävää — tietoa lineaarikuvauksen käänteiskuvauksen lineaarisuudesta. Luvussa 3 ongelma ratkaistaan toisella tavalla kiertämällä ongelmat käyttämällä determinantteja ja transponointia.

1.6. Transponoitu matriisi.

Toisinaan joutuu matriisissa vaihtamaan rivit ja sarakkeet keskenään, *transponoimaan matriisin.* Itse asiassa olemme törmänneet tähän ilmiöön rivi- ja sarakevektoreiden yhteydessä. Vastaisen varalle esitetään seuraava määritelmä.

MÄÄRITELMÄ 1.6.1. Matriisin $A_{m\times n} = [a_{ij}]$ *transpoosi* on

$$A^T = [b_{ji}]\,, \text{ missä } b_{ji} = a_{ij} \text{ kaikilla } i \text{ ja } j.$$

[27]Ks. lause 3.1.6.

Esimerkiksi
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

Osoittautuu, että alkuperäisellä ja transponoidulla matriisilla on hämmästyttävän paljon yhteisiä ominaisuuksia. Ainakin pätee:

LAUSE 1.6.2.

a) $\left(A^T\right)^T = A$

b) $(A+B)^T = A^T + B^T \qquad (A_{m\times n}\ ja\ B_{m\times n})$

c) $(\lambda A)^T = \lambda A^T \qquad (\lambda \in \mathbb{R})$

d) $(AB)^T = B^T A^T \qquad (A_{m\times n}\ ja\ B_{n\times p})$ **(Huomaa järjestys!)**

e) *Jos A on kääntyvä, niin myös A^T on kääntyvä ja $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$.*

TODISTUS. Väitteet a), b) ja c) on helppo tarkastaa. Myös d) on tylsä suora lasku

$$(AB)^T_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk}B_{ki} = \sum_{k=1}^{n} B_{ki}A_{jk} = \sum_{k=1}^{n} B^T_{ik}A^T_{kj} = (B^TA^T)_{ij}.$$

Kohta e) palautuu kohtaan d): $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$ ja $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$,
□

MÄÄRITELMÄ 1.6.3. Neliömatriisi $A = A_{n\times n}$ on *symmetrinen*, jos $A^T = A$ eli $a_{ij} = a_{ji}$ kaikilla i, j.

Esimerkiksi $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ on symmetrinen, mutta kumpikaan matriiseista

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ ja } \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix}$$

ei ole. Palaamme tutkimaan symmetristen matriisien erikoisominaisuuksia luvun 5 lopulla ja luvussa 6.

2. Determinantti

2.1. Kaksiriviset determinantit.

MÄÄRITELMÄ 2.1.1. Olkoon

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

Määritellään *kaksirivinen determinantti*

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

LAUSE 2.1.2. *Edellisen määritelmän 2×2–matriisille A on voimassa:*

$$A \text{ on kääntyvä} \iff \det A \neq 0,$$

ja kun A on kääntyvä, niin

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

PERUSTELU. Väite esiintyy toisinaan koulukurssissa. Perustelemme sen kuitenkin:

Jos nimittäjä $\det A$ ei ole 0, niin matriisi $\frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$ on olemassa ja suora lasku osoittaa, että tosiaan

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \cdots = I_{2\times 2}$$

$$\text{ja (siis) myös} \quad \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \cdots = I_{2\times 2}.$$

Ehto on siis riittävä. Nyt on vielä todistettava, että jokaisen kääntyvän kaksirivisen matriisin determinantti todella on nollasta eroava. Teemme sen epäsuorasti. Vastaoletus on, että kuitenkin $\det A = 0$ eli $a_{11}a_{22} = a_{12}a_{21}$, ja haluamme tästä todistaa, että A ei ole kääntyvä. Riittää näyttää, että A:ta vastaava lineaarikuvaus ei ole bijektio. Se ei olekaan edes injektio, vaan voimme todella löytää ainakin 2 eri vektoria $x \neq y \in \mathbb{R}^2$, jotka se kuvaa samaksi: $Ax = Ay$. Tapauksessa $A = 0$ voidaan tietysti valita vaikkapa $x = e_1$ ja $y = e_2$. Myös tapauksessa $A \neq 0$ tarkastellaan kantavektoreiden kuvia, siis A :n sarakkeita $Ae_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$ ja $Ae_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$. Vastaoletuksemme $a_{11}a_{22} = a_{12}a_{21}$ ilmaisee, että nämä ovat yhdensuuntaiset, siis toinen saadaan toisesta kertomalla jollakin luvulla[28]; esimerkiksi $Ae_1 = \lambda Ae_2$. Nyt $Ae_1 = \lambda Ae_2 = A(\lambda e_2)$, joten vektoreiksi x ja y kelpaavat e_1 ja λe_2. $\square$

[28] Kun $A \neq 0$, jokin elementti eroaa nollasta, esimerkiksi $a_{12} \neq 0$, jolloin ehdosta seuraa heti, että $a_{21} = \frac{a_{11}}{a_{12}} a_{22}$ ja siis $Ae_1 = \frac{a_{11}}{a_{12}} Ae_2$.

Todistuksen voi helposti ymmärtää myös geometriselta kannalta. Determinanttihan on (mahdollista merkkiä vaille) sama kuin kantavektoreiden kuvien virittämän suunnikkaan pinta-ala. Se on epäilemättä nolla juuri silloin, kun kantavektoreiden kuvat ovat yhdensuuntaiset. Tällöin niiden virittämä taso eli kuvajoukko $A(\mathbb{R}^2)$ degeneroituu suoraksi, tapauksessa $A = 0$ jopa pisteeksi. Tällaisessa tilanteessa A ei edusta surjektiota, ei siis ainakaan bijektiota $\mathbb{R}^2 \to \mathbb{R}^2$.

Tässä luvussa näytetään miten nämä tarkastelut yleistyvät $n \times n$–matriiseille. Aloitetaan varovasti.

2.2 Kolmiriviset determinantit.

Määritelmä 2.2.1. Olkoon

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

Määritellään A:n *kolmirivinen determinantti* (**Huomaa merkit!**)

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

Muistamme johdannosta, että geometriselta kannalta kolmirivinen determinantti on itseisarvoltaan sarakkeiden virittämän suuntaissärmiön tilavuus. Siksi myös 3×3–matriisi on kääntyvä tasan silloin, kun sen determinantti on nollasta eroava. Käänteismatriisillekin on olemassa samantapainen lauseke kuin 2×2–tapauksessa. Emme todista näitä väiteitä vielä, vaan kiinnitämme huomiota niihin kolmirivisen determinantin ominaisuuksiin, jotka ovat olennaisia n–rivisen determinantin määritelmässä.

Koska determinantti on sarakevektoreiden virittämän särmiön tilavuus, sitä on luonnollista tarkastella **sarakkeiden funktiona**. Kyseessä on kolmen vektorimuuttujan reaaliarvoinen funktio; determinantti liittää kolmeen 3–komponenttiseen[29] vektoriin u, v ja w luvun

$$\det(u, v, w) = \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} \in \mathbb{R}.$$

Lause 2.2.2. *Matriisin $A_{3\times 3}$ determinantille on voimassa:*

(ml-1)

$$\det(\lambda u, v, w) = \det(u, \lambda v, w) = \det(u, v, \lambda w) = \lambda \det(u, v, w).$$

Esimerkiksi

$$\det(\lambda u, v, w) = \begin{vmatrix} \lambda u_1 & v_1 & w_1 \\ \lambda u_2 & v_2 & w_2 \\ \lambda u_3 & v_3 & w_3 \end{vmatrix} = \lambda \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} = \lambda \det(u, v, w).$$

[29] Muista, että toistaiseksi n on 3.

Sama sanoin: Jos A:n jokin sarake kerrotaan luvulla λ, niin $\det A$ *tulee kerrotuksi luvulla λ.*

(ML-2)

$$\begin{aligned}\det(u+u',v,w) &= \det(u,v,w)+\det(u',v,w),\\ \det(u,v+v',w) &= \det(u,v,w)+\det(u,v',w) \qquad \textit{ja}\\ \det(u,v,w+w') &= \det(u,v,w)+\det(u,v,w'),\end{aligned}$$

siis esimerkiksi

$$\begin{vmatrix} u_1+u_1' & v_1 & w_1\\ u_2+u_2' & v_2 & w_2\\ u_3+u_3' & v_3 & w_3\end{vmatrix} = \begin{vmatrix} u_1 & v_1 & w_1\\ u_2 & v_2 & w_2\\ u_3 & v_3 & w_3\end{vmatrix} + \begin{vmatrix} u_1' & v_1 & w_1\\ u_2' & v_2 & w_2\\ u_3' & v_3 & w_3\end{vmatrix}.$$

Sama sanoin: Olkoot A ja B **samoja** *lukuunottamatta mahdollisesti yhtä saraketta — olkoon se i:s sarake. Esimerkiksi tapauksessa $i=1$ siis $A=(u,v,w)$ ja $B=(u',v,w)$. Olkoon C matriisi, jolla on muuten nämä samat sarakkeet, mutta i:nnellä sarakkeella A :n ja B :n i:nsien sarakkeiden summa. Tällöin* $\det C=\det A+\det B$.

Yhdessä ehdot (ML-1) ja (ML-2)merkitsevät helposti muistettavaa asiaa: determinantti on **kunkin sarakkeen suhteen lineaarikuvaus.**

(ALT) *Jos A:n kaksi saraketta vaihdetaan keskenään, niin determinantti tulee kerrotuksi luvulla -1.* Toisin sanoen determinantin **merkki vaihtuu** päinvastaiseksi, mutta tietysti determinantin itseisarvo säilyy.

MÄÄRITELMÄ 2.2.3. Sanomme kolmen vektorimuuttujan funktiota, jolla on edellä listatut determinantin lineaarisuusominaisuudet (ML-1) ja (ML-2) *kolmilineaariseksi* kuvaukseksi. Jos myös merkin vaihtumista koskeva ehto (ALT) toteutuu, kuvaus on lisäksi *vuorotteleva* eli *alternoiva*. Olemme huomanneet, että kolmirivinen determinantti on sarakkeidensa alternoiva 3–lineaarifunktio.

Seuraava lause listaa alternoivan 3–lineaarifunktion — erityisesti kolmirivisen determinantin — ominaisuuksia.

LAUSE 2.2.4. *Jokaisella alternoivalla kolmilineaarikuvauksella $\Delta : \mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}$, erityisesti kolmirivisellä determinantilla* $\det : \mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}$, *on seuraavat ominaisuudet*

(1) *Jos yksi muuttujista (siis esimerkiksi determinantin sarakkeista) u,v,w on sama kuin jokin toinen, mahdollisesti kerrottuna jollakin luvulla, niin $\Delta(u,v,w)=0$. Esimerkiksi*

$$\det(u,\lambda u,w) = \begin{vmatrix} u_1 & \lambda u_1 & w_1\\ u_2 & \lambda u_2 & w_2\\ u_3 & \lambda u_3 & w_3\end{vmatrix} = 0,$$

ja

$$\det(u,0,w) = \begin{vmatrix} u_1 & 0 & w_1\\ u_2 & 0 & w_2\\ u_3 & 0 & w_3\end{vmatrix} = \begin{vmatrix} u_1 & 0u_1 & w_1\\ u_2 & 0u_2 & w_2\\ u_3 & 0u_3 & w_3\end{vmatrix} = 0.$$

Erityisesti, jos determinantissa on kaksi samaa saraketta tai nollasarake, sen arvo on 0.

(2) *Jos muuttujaan (sarakkeeseen) lisätään toinen kerrottuna vakiolla, niin kolmilineaarikuvauksen arvo ei muutu, siis esimerkiksi*

$$\det(u, v+\lambda u, w) = \begin{vmatrix} u_1 & v_1+\lambda u_1 & w_1 \\ u_2 & v_2+\lambda u_2 & w_2 \\ u_3 & v_3+\lambda u_3 & w_3 \end{vmatrix} = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}.$$

Perustelu. Harjoitustehtävä □

Seuraava lause sanoo, että voimme saada edellisistä uusia laskusääntöjä determinantille korvaamalla joka paikassa sanan ”sarake” sanalla ”rivi”.

Lause 2.2.5. *Kolmirivisen determinantin* $\det : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ *arvo ei muutu, jos sen rivit ja sarakkeet vaihdetaan toisikseen, siis vastaavan matriisin transponoinnissa, vaan* $\det A = \det A^T$*. Toisin sanoen*

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

Perustelu. Asian voi tarkastaa laskemalla molemmat determinantit.

Determinantin lineaarisuus ja alternoivuus sekä sarakkeiden että rivien suhteen antavat nopean tavan laskea determinantin Gauss–Jordanin menetelmän kaltaisella algoritmilla, jossa determinantin sisään muotoillaan kolmiomatriisi muuttamatta determinantin arvoa. Tästä on hyötyä, sillä kolmiomatriisin determinantti on tietysti helppo laskea.

Lause 2.2.6. *Olkoon* $A_{3\times 3} =$ *kolmiomatriisi. Tällöin*

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}.$$

Erityisesti $\det I = 1$.

Esimerkki 2.2.7. a) $\begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} =$
$-\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = 15$

b) $\begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -10 & 70 & 0 \end{vmatrix} = 10 \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} = \cdots = 10 \cdot 15 = 150$

Determinantin laskemisessa on oleellista hyötyä myös seuraavasta lauseesta

LAUSE 2.2.8. *Kolmirivisille neliömatriiseille A ja B on voimassa:*

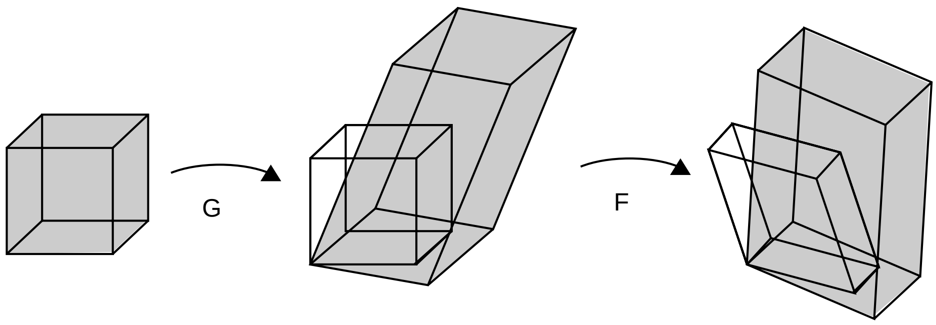
$$\det(AB) = \det A \cdot \det B.$$

PERUSTELU. Molemmat puolet voi laskea auki. Lausekkeista tulee pitkät, mutta samat. $\square$

Geometrian kannalta kyseessä on merkkiä lukuunottamatta ilmeinen asia, sillä $\det A$ ilmaisee tilavuuden suurennussuhteen siinä lineaarikuvauksessa $F = F_A : \mathbb{R}^3 \to \mathbb{R}^3$, jonka matriisi on A, ja vastaavasti $\det B$ ilmaisee tilavuuden suurennussuhteen siinä lineaarikuvauksessa $G = G_B : \mathbb{R}^3 \to \mathbb{R}^3$, jonka matriisi on B. Näistä yhdistetty lineaarikuvaus $F \circ G$ suurentaa tietysti tilavuuksia kertoimella, joka on edellisten tulo, siis $\det A \cdot \det B$, onhan esimerkiksi yksikkökuution $K \subset \mathbb{R}^3$ kuvajoukon tilavuus, jota merkitsemme hetken aikaa lyhenteellä Vol.

$$Vol(F \circ G)(K) = Vol(F(G(K)) = \det A \cdot Vol(G(K)) = \det A \cdot (\det B \cdot Vol(K)).$$

Kuvauksen $F \circ G$ suurennussuhde on siis $\det A \cdot \det B$.



Toisaalta lineaarikuvauksen $F \circ G$ matriisi on AB, ja se suurentaa siis tilavuuksia kertoimella $\det(AB)$. Ainakin itseisarvot $|\det A \cdot \det B|$ ja $|\det(AB)|$ ovat samat, koska kumpikin kuvaa samaa asiaa.

SEURAUS 2.2.9. *Matriisi $A_{3\times 3}$ on kääntyvä, jos ja vain jos $\det A \neq 0$. Tällöin*

$$\det A^{-1} = \frac{1}{\det A}.$$

TODISTUKSESTA. Lauseen 1.5.9 mukaan[30] 3×3–matriisi A on kääntyvä tasan silloin, kun se muokkautuu Gaussin ja Jordanin ratkaisumenettelyllä identtiseksi matriisiksi $I_{3\times 3}$. Kääntymätön taas muokkautuu muotoon

$$\begin{bmatrix} \dots & \dots & \dots \\ 0 & 0 & 0 \end{bmatrix},$$

jossa on pelkkiä nollia alarivinä. Gaussin–Jordanin käsittelyn jälkeen saadulla matriisilla on siis determinantti 1 tai 0 sen mukaan oliko alkuperäinen A kääntyvä vai ei. Toisaalta mikään Gaussin–Jordanin rivioperaatioista ei voi muuttaa nollasta eroavaa determinanttia nollaksi eikä päinvastoin.

Determinantin lauseke saadaan lopuksi edellisen lauseen avulla, siis yhtälöstä

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

$\square$

[30] Lauseen 1.5.9 todistus on kesken. Tarkkaan ottaen saamme tässä vain tuloksen, että A:lla on oikeanpuoleinen käänteismatriisi tasan silloin, kun $\det A \neq 0$.

2.3. Yleinen n–rivinen determinantti.

Nyt määriteltävän determinantin halutaan tuottavan moniuotteisen vastineen suuntaissärmiön tilavuuden käsitteelle. Useimmista lineaarialgebran oppikirjoista löytyy seuraava määritelmä:

MÄÄRITELMÄ 2.3.1. Matriisin $A_{n\times n}$ ij:s *alimatriisi* A_{ij} saadaan poistamalla A:sta rivi i ja sarake j.

MÄÄRITELMÄ 2.3.2. Määritellään matriisin $A_{n\times n}$ *determinantti* kaksivaiheisesti[31]:

I $\det A_{1\times 1} = a_{11}$.
II $\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13} - \cdots + (-1)^{1+n}a_{1n}\det A_{1n}$
$= \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det A_{1j}$.

Lukijan on syytä testata, että määritelmä toimii tapauksessa $n = 2$ tai 3. Kirjoissa määritelmä kirjoitetaan usein näennäisesti lyhemmin ottamalla ensin käyttöön kofaktorin käsite.

MÄÄRITELMÄ 2.3.3. Alimatriisin A_{ij} determinantti $\det A_{ij}$ on nimeltään ij:s *alideterminantti*. Matriisin $A_{n\times n}$ alkion a_{ij} *kofaktori* on

$$\operatorname{cof} a_{ij} = (-1)^{i+j}\det A_{ij},$$

jolloin

$$\det A = a_{11}\operatorname{cof} a_{11} + a_{12}\operatorname{cof} a_{12} + \cdots + a_{1n}\operatorname{cof} a_{1n} = \sum_{j=1}^{n} a_{1j}\operatorname{cof} a_{1j}.$$

Samalla kun alamme tarkastaa yleisen determinantin ominaisuuksia, pohdimme mitä ominaisuuksia tilavuutta kuvaavalta determinantilta kannattaa vaatia. On myös aihetta miettiä, miten yllä annettuun määritelmään on päädytty ja voisiko determinantin määritellä jotenkin toisin kuin yllä ja silti saada halutut ominaisuudet.

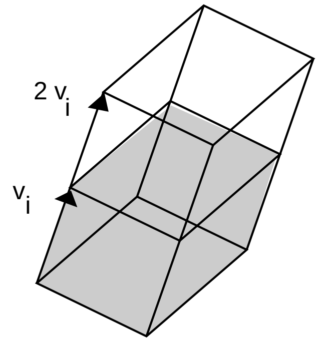
VAATIMUKSIA. Koska halutaan, että n–rivinen determinantti $\det A_{n\times n}$ on $n\times n$-matriisin A sarakevektoreiden virittämän särmiön ”tilavuus”, sitä on luonnollista tarkastella **sarakkeiden funktiona**. Determinantti liittää n–komponenttisiin vektoreihin

$$v_1 = a_{*1} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad v_2 = a_{*2} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \ldots \text{ ja } v_n = a_{*n} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}$$

$$\text{luvun } \det(v_1, v_2, \ldots, v_n) = \det \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}.$$

On luonnollista vaatia särmiön tilavuuden kaksinkertaistuvan, kun yksi särmä venytetään kaksinkertaiseksi.

[31] Induktiostahan tässä on kysymys. Ks. esim. JOHDATUS MATEMATIIKKAAN.

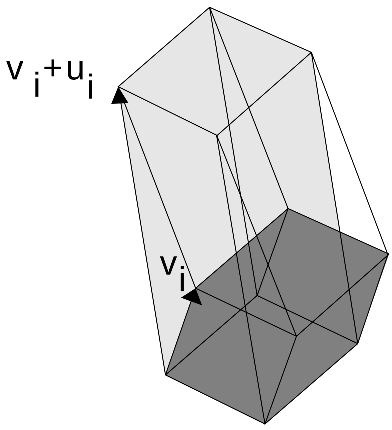


Tämä antaa aiheen vaatia, että jos A:n jokin sarake kerrotaan luvulla λ, niin $\det A$ tulee kerrotuksi samalla luvulla λ, ts. että (ML-1) pätee. Esimerkiksi

$$\det(\lambda v_1, v_2, \ldots, v_n) = \lambda \det(v_1, v_2, \ldots, v_n).$$

Myös yhteenlaskuominaisuus (ML-2) kaikkien särmien suhteen on luonnollinen. Ensimmäisen särmän suhteenhan se sanoo, että

$$\det(v_1 + u_1, v_2, \ldots, v_n) = \det(v_1, v_2, \ldots, v_n) + \det(u_1, v_2, \ldots, v_n)$$



On siis aihetta vaatia, että yleinenkin determinantti on kunkin sarakkeen suhteen lineaarikuvaus. Ilmaisemme tämän sanomalla, että determinantin kuuluu olla *moni–* eli *multilineaarikuvaus.*[32]

Kolmirivisellä determinantilla on lisäksi alternoivuusominaisuus (ALT), jonka mukaan kahden sarakkeen vaihto vaihtaa determinantin merkin. Tämän geometrinen sisältö ei ole aivan heti ilmeinen, mutta on kyllä luonnollista vaatia, että suuntaissärmiön tilavuus on 0, jos kaksi sen särmistä yhtyy. Korvatkaamme siis ehto (ALT) tällä vaatimuksella. Nyt käy niin onnellisesti, että ehto (ALT) seuraa ehdoista (ML-1), (ML-2) ja tästä lievemmästä oletuksesta. Tämän toteaminen on sopiva harjoitustehtävä. Huomaamme siis, että alternoivuusvaatimuskin on geometrisesti luonnollinen.

Esitettyämme vaatimukset voimme tietysti kysyä, onko olemassa funktiota $\Delta : (\mathbb{R}^n)^n \to \mathbb{R}$, joka toteuttaisi ne. Kuinka monta erilaista ehdokasta determinantiksi

[32] Joudumme maksamaan lineaarisuusvaatimuksesta (ML-1) sen hinnan, että täytyy hyväksyä "negatiiviset tilavuudet".

on jäljellä, kun vaadimme, että determinantin kuuluu olla sarakkeiden alternoiva n–lineaarikuvaus? Onko yhtään? Onko siis ollenkaan järkevää puhua ”tilavuudesta” avaruudessa $\mathbb{R}^n$, kun $n \geq 4$? Voimmeko toisaalta ehkä vaatia ”tilavuudelta” eli determinanttifunktiolta lisää ominaisuuksia?

Osoittautuu, että voimme vielä vaatia, että ykkösmatriisin determinantti eli yksikkökuution tilavuus $\det(e_1, e_2, \ldots, e_n)$ on tasan yksi, mutta että tätä ”mittyksikön valintaa” vaille on olemassa ainoastaan yksi alternoiva n–lineaarikuvaus

$$\Delta : \mathbb{R}^n \times \mathbb{R}^n \times \ldots \mathbb{R}^n \to \mathbb{R}.$$

Kuten arvata saattaa, se on määritelmissä 2.3.1 ja 2.3.2 mainittu determinantti. On todistettava, että se kelpaa ja että muita ei ole. Seuraava lause 2.3.4. sanoo, että se kelpaa. Yksikäsitteisyys — periaate 2 — todistetaan lauseen 2.4.6. jälkeen.

Lause 2.3.4 (Determinantin perusominaisuus). *Determinantti on sarakkeiden alternoiva n–lineaarifunktio.*

Perustelu. Induktiotodistus:

Vaihe I: ”Lause pätee, kun n on 1”. Totta kai näin on. Tiedämme itse asiassa ennestään myös, että lause pätee n:n arvoilla 2 ja 3.

Vaihe II: Oletetaan, että lause pätee $n-1$-riviselle determinantille ja todistetaan se n–riviselle. Merkintöjen yksinkertaistamiseksi tarkastetaan lineaarisuus ensimmäisen muuttujan suhteen. Ehto (ML-1) tarkastetaan näin:

$$\det(\lambda v_1, v_2, \ldots, v_n) = \begin{vmatrix} \lambda a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

$$= \lambda a_{11} \begin{vmatrix} a_{22} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} \lambda a_{21} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & a_{n3} & \ldots & a_{nn} \end{vmatrix} + \cdots \pm a_{1n} \begin{vmatrix} \lambda a_{21} & a_{22} & \ldots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & a_{n2} & \ldots & a_{n(n-1)} \end{vmatrix}$$

$$= \lambda a_{11} \begin{vmatrix} a_{22} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix} - \lambda a_{12} \begin{vmatrix} a_{21} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \ldots & a_{nn} \end{vmatrix} + \cdots \pm \lambda a_{1n} \begin{vmatrix} a_{21} & a_{22} & \ldots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{n(n-1)} \end{vmatrix} =$$

$$= \lambda \det(v_1, \ldots, v_n).$$

Huomaa, missä käytettiin induktioehtoa.

Ehto (ML-2) tarkastetaan samaan tapaan: Olkoon $u_1 = (b_{11}, \dots, b_{n1})$.

$$\det(v_1 + u_1, v_2, \dots, v_n) = \begin{vmatrix} a_{11} + b_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= (a_{11} + b_{11}) \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$- a_{12} \begin{vmatrix} a_{21} + b_{21} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix} + \dots \pm a_{1n} \begin{vmatrix} a_{21} + b_{21} & a_{22} & \dots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} & \dots & a_{n(n-1)} \end{vmatrix}$$

$$= a_{11} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix} + \dots \pm a_{1n} \begin{vmatrix} a_{21} & a_{22} & \dots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n(n-1)} \end{vmatrix} +$$

$$+ b_{11} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} b_{21} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix} + \dots \pm a_{1n} \begin{vmatrix} b_{21} & a_{22} & \dots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ b_{n1} & a_{n2} & \dots & a_{n(n-1)} \end{vmatrix}$$

$$= \det(v_1, \dots, v_n) + \det(u_1, \dots, v_n).$$

Induktioehtoa (ML-2) käytettiin alideterminantteihin. Tarkastetaan vielä vaihdantaehto (ALT) kahdelle ensimmäiselle muuttujalle. Määritelmän 2.3.1 mukaan

$$\det(v_1, v_2, \dots, v_n) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix} + \dots \pm a_{1n} \begin{vmatrix} a_{21} & a_{22} & \dots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n(n-1)} \end{vmatrix}$$

Vaihtamalla kaksi ensimmäistä termiä ja muissa termeissä kaksi ensimmäistä saraketta tämän saa alideterminattien alternoivuutta käyttäen muotoon

$$- a_{12} \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix} + a_{11} \begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} - \dots \mp a_{1n} \begin{vmatrix} a_{22} & a_{21} & \dots & a_{2(n-1)} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n1} & \dots & a_{n(n-1)} \end{vmatrix}$$

$$= -\det(v_2, v_1, \dots, v_n)$$

□

2.4. Yleisen determinantin ominaisuuksia.

Yleisellä determinantilla on periaatteessa samat ominaisuudet kuin kolmiriviselläkin. Determinantin tärkeimmät ominaisuudethan ovat lineaarisuus kunkin sarakkeen suhteen ja merkinvaihto sarakkeiden vaihdossa. Muut ominaisuudet seuraavat näistä.

LAUSE 2.4.1. *Matriisin $A_{n\times n}$ determinantti voidaan laskea "kehittämällä" determinantti mielivaltaisen rivin i tai sarakkeen j mukaan ts.*

$$\begin{aligned}\det A &= a_{i1}\operatorname{cof} a_{i1} + a_{i2}\operatorname{cof} a_{i2} + \cdots + a_{in}\operatorname{cof} a_{in} = \sum_{j=1}^{n} a_{ij}\operatorname{cof} a_{ij}\\ &= a_{1j}\operatorname{cof} a_{1j} + a_{2j}\operatorname{cof} a_{2j} + \cdots + a_{nj}\operatorname{cof} a_{nj} = \sum_{i=1}^{n} a_{ij}\operatorname{cof} a_{ij}.\end{aligned}$$

LAUSE 2.4.2. *Olkoon $A_{n\times n}$ ylä- tai alakolmiomatriisi tai erikoistapauksena edellisistä diagonaalimatriisi. Tällöin determinantti saadaan kertomalla keskenään A:n diagonaali- eli lävistäjäalkiot*

$$\det A = a_{11}a_{22}\dots a_{nn}.$$

LAUSE 2.4.3. *Matriisin $A_{n\times n}$ determinantille on voimassa:*

a) *Jos A:ssa on nollarivi (nollasarake), niin* $\det A = 0$.
b) *Jos A:n rivi (sarake) kerrotaan vakiolla c, niin* $\det A$ *tulee kerrotuksi samalla vakiolla.*
c) *Olkoot A ja B samoja lukuunottamatta riviä i (saraketta j) ts.*

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{i1} & \dots & a_{in}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ b_{i1} & \dots & b_{in}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{bmatrix}.$$

Olkoon lisäksi

$$C = \begin{bmatrix} a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{i1}+b_{i1} & \dots & a_{in}+b_{in}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{bmatrix}.$$

Tällöin

$$\det C = \det A + \det B.$$

d) *Jos A:n kaksi riviä (saraketta) vaihdetaan keskenään, niin determinantin merkki vaihtuu (determinantin itseisarvo säilyy).*
e) *Jos A:n kaksi riviä (saraketta) ovat samoja, niin* $\det A = 0$.
f) *Jos A:n rivi (sarake) on toinen rivi (sarake) kerrottuna vakiolla, niin* $\det A = 0$.
g) *Jos A:n riviin (sarakkeeseen) lisätään toinen rivi (sarake) kerrottuna vakiolla, niin determinantti ei muutu.*

Lause 2.4.4. *Matriisille $A_{n\times n}$ on voimassa:*

$$a_{i1}\operatorname{cof} a_{j1} + a_{i2}\operatorname{cof} a_{j2} + \cdots + a_{in}\operatorname{cof} a_{jn} = 0, \ \text{jos } i \neq j.$$

Lauseen 2.4.4 merkitys on siinä, että sen avulla saadaan klassinen lauseke käänteismatriisille.

Lause 2.4.5. *Matriisille $A_{n\times n}$ on voimassa:*

$$\det A = \det A^T.$$

Lause 2.4.6. *Matriiseille $A_{n\times n}$ ja $B_{n\times n}$ on voimassa:*

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

Todistuksista. Lause 2.4.2 on ilmeinen tosiasia. Lause 2.4.3 sanoo sarakkeiden osalta, että determinantti on alternoiva multilineaarikuvaus ja luettelee muutamia välittömiä seurauksia. Rivien osalta väitteet saadaan saman tien, **mikäli transponointia koskeva lause 2.4.5. on tosi.** Transponoinnin luvallisuus on nimittäin sama asia kuin lupa ”kehittää” determinantti ensimmäisen sarakkeen mukaan. Toisaalta sarakkeita vaihtamalla saa minkä tahansa sarakkeen ensimmäiseksi (pidä kirjaa merkeistä!), joten matriisin voi ”kehittää” minkä tahansa sarakkeen mukaan. Vielä transponointi, ja huomaamme, että ”kehittäminen” onnistuu minkä tahansa rivinkin mukaan. Lause 2.4.4 on harjoitustehtäväksi sopiva, kunhan tunnetaan 2.4.3 ja 2.4.1. Varsinaiseksi ongelmaksi jää siis todistaa transponointia koskeva lause 2.4.5 ja tulon determinanttia koskeva lause 2.4.6.

Osoittautuu, että molemmat ovat seurauksia niistä kahdesta syvällisestä periaatteesta, jotka otimme koko determinanttiteorian lähtökohdiksi:

Periaate 1. *n-rivinen determinantti on sarakkeidensa alternoiva n–lineaarifunktio ja* $\det I_{n\times n} = 1$.

Periaate 2. *Jokainen alternoiva n–lineaarifunktio $\Delta : (\mathbb{R}^n)^n \to \mathbb{R}$ on determinantti kerrottuna jollakin vakiolla, siis muotoa*

$$\Delta(v_1, \dots, v_n) = \lambda \det(v_1, \dots, v_n).$$

Yhdessä periaatteet sanovat, että determinantin muodostaminen on vakiokerrointa vaille ainoa alternoiva n–komponenttisten vektoreiden n–lineaarikuvaus.

Periaatteiden todistukset. Periaate 1 on jo todistettu, kyseessähän on determinantin perusominaisuus 2.4.4.

Periaate 2 voidaan tarkastaa suoralla laskulla: Olkoon $\Delta : (\mathbb{R}^n)^n \to \mathbb{R}$ alternoiva n–lineaarifunktio. Olkoon $\ell = \Delta(e_1, \dots, e_n)$, missä $e_1, \dots, e_n$ ovat standardikantavektorit. Lasketaan $\Delta(v_1, \dots, v_n)$ mielivaltaisille vektoreille

$$v_1 = \sum_{i=1}^{n} a_{i1} e_i = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \dots \text{ ja } \ v_n = \sum_{j=1}^{n} a_{jn} e_j = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}.$$

Huomaamme aluksi, että alternoivuuden takia $\Delta(e_i, \ldots, e_j) = 0$ aina, kun kaksi kantavektoreista $e_i \ldots, e_j$ ovat samoja. Muuten $\Delta(e_i, \ldots, e_j)$ on ℓ tai $-\ell$ sen mukaan, saadaanko vektoreiden jono $e_i, \ldots, e_j$ jonosta $e_1, \ldots, e_n$ parillisella vai parittomalla määrällä vaihtoja.[33] Merkitään $\epsilon_{i\ldots j} =$

$$\begin{cases} 0, & \text{jos ainakin kaksi indekseistä } i, \ldots, j \in \{1, \ldots, n\} \text{ ovat samoja} \\ 1, & \text{jos jono } (i, \ldots, j) \text{ saadaan jonosta } (1, \ldots, n) \text{ parillisella määrällä vaihtoja} \\ -1, & \text{jos jono } (i, \ldots, j) \text{ saadaan jonosta } (1, \ldots, n) \text{ parittomalla määrällä vaihtoja.} \end{cases}$$

$$\begin{aligned}
\Delta(v_1, \ldots, v_n) &= \Delta(\sum_{i=1}^n a_{i1} e_i, \ldots, \sum_{j=1}^n a_{jn} e_j) \\
&\overset{ML-1}{=} \sum_{i=1}^n \Delta(a_{i1} e_i, \ldots, \sum_{j=1}^n a_{jn} e_j) \\
&\overset{ML-2}{=} \sum_{i=1}^n a_{i1} \Delta(e_i, \ldots, \sum_{j=1}^n a_{jn} e_j) = \ldots \\
&\overset{ML-1}{=} \sum_{i=1}^n a_{i1} \cdots \sum_{j=1}^n \Delta(e_i, \ldots, a_{jn} e_j) \\
&\overset{ML-2}{=} \sum_{i=1}^n a_{i1} \cdots \sum_{j=1}^n a_{jn} \Delta(e_i, \ldots, e_j) \\
&= \sum_{i=1}^n a_{i1} \cdots \sum_{j=1}^n a_{jn}\, \ell\, \epsilon_{i\ldots j} \\
&= \ell \left(\sum_{i=1}^n a_{i1} \cdots \sum_{j=1}^n a_{jn}\, \epsilon_{i\ldots j} \right) \\
&= \ell \left(\sum_{i=1}^n \cdots \sum_{j=1}^n a_{i1} \ldots a_{jn}\, \epsilon_{i\ldots j} \right)
\end{aligned}$$

Täsmälleen sama lasku sovellettuna alternoivaan n–lineaarikuvaukseen det antaa

$$\begin{aligned}
\det(v_1, \ldots, v_n) &= \det(\sum_{i=1}^n a_{i1} e_i, \ldots, \sum_{j=1}^n a_{jn} e_j) \\
&= \cdots = \sum_{i=1}^n \cdots \sum_{j=1}^n a_{i1} \ldots a_{jn}\, \epsilon_{i\ldots j},
\end{aligned}$$

sillä $\det(e_1, \ldots, e_n) = \det I = 1$. Koska lausekkeet ovat kerrointa vaille samat, on lause todistettu. Periaatteet on todettu oikeiksi. □

[33] Vain toinen on kulloinkin mahdollista. Tämän todistus vie sivuun aiheestamme.

Periaatteen 2 yhteydessä mainitulle kertoimelle λ on todistuksen sivutuotteena saatu arvoksi

$$\lambda = \ell = \Delta(e_1, \dots, e_n).$$

Todistus tuotti myös uuden lausekkeen determinantille.

$$\det(v_1, \dots, v_n) = \sum_{i=1}^{n} \cdots \sum_{j=1}^{n} a_{i1} \dots a_{jn} \, \epsilon_{i \dots j}.$$

Koska kaikki n summausindeksiä saavat arvot $1, \dots, n$, summassa on periaatteessa n^n termiä. Koska kuitenkin $\epsilon_{i \dots j} = 0$ aina, kun kaksi indekseistä on samoja, jäljellä ovat vain ne termit, joiden indekseinä ovat kaikki luvut $1, \dots, n$, kukin tasan yhden kerran. Toisin sanoen summassa esiintyy yksi termi jokaista järjestystä kohti, johon jono $1, \dots, n$ voidaan permutoida. Näitä on tunnetusti vain(?) $n!$ kappaletta. Edelleenkin determinantin laskeminen tällä kaavalla — tai suoraan määritelmällä 2.3.2, joka johtaa olennaisesti samaan laskuun — on erittäin työlästä. Suositeltavampaa on käyttää esimerkiksi Gauss–Jordan tyyppistä algoritmia. Laskutoimitusten mutkikkuuden arviointi — kompleksisuusteoria — on oma mielenkiintoinen matematiikan alansa.

Käytämme lopuksi periaatetta 2 esittääksemme todistukset tulon determinanttia koskevalle lauseelle 2.3.6 ja transponointia koskevalle lauseelle 2.3.5.

Todistus lauseelle 2.3.6. Olkoon $B = B_{n \times n}$ jokin neliömatriisi, jonka rivejä merkitsemme $b_{1*}, \dots, b_{n*}$, ja olkoon Δ se funktio, joka liittää matriisin A sarakkeisiin $v_1, \dots, v_n$ luvun $\Delta(v_1, \dots, v_n) = \det(BA)$. Lausekkeesta

$$\det(BA) = \begin{vmatrix} (b_{1*}|v_1) & \dots & (b_{1*}|v_i) & \dots & (b_{1*}|v_m) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ (b_{j*}|v_1) & \dots & (b_{j*}|v_i) & \dots & (b_{j*}|v_m) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ (b_{n*}|v_1) & \dots & (b_{n*}|v_i) & \dots & (b_{n*}|v_m) \end{vmatrix}$$

voi päätellä, että tämäkin Δ on muuttujien $v_1, \dots, v_n$ alternoiva n–lineaarikuvaus, joten periaatteiden 1 ja 2 mukaan

$$\det BA = \Delta(A) = \Delta(e_1, \dots, e_n) \det A = \det B \det A.$$

□

Todistusidea lauseelle 2.3.5. Olkoon Δ se funktio, joka liittää matriisin A sarakkeisiin $v_1, \dots, v_n$ luvun $\Delta(v_1, \dots, v_n) = \det A^T$. (Huomaa transpoosi.) Jäljittelemällä perusominaisuuslauseen 2.3.4 todistusta voi induktiolla — joskin hiukan mutkikkaammalla — todistaa, että Δ on alternoiva n–lineaarikuvaus, joten periaatteiden 1 ja 2 mukaan

$$\Delta = \Delta(e_1, \dots, e_n) \cdot \det = 1 \cdot \det = \det .$$

□

2.5. Determinantin yhteys käänteismatriisiin.

LAUSE 2.5.1. *Neliömatriisi $A_{n\times n}$ on kääntyvä jos ja vain jos* $\det A \neq 0$. *Tällöin*

$$\det A^{-1} = \frac{1}{\det A}.$$

TODISTUS. Todistus on periaatteessa sama kuin kolmirivisessä tapauksessa 2.2.9. Kertaamme sen nyt ja paikkaamme samalla todistukseen jääneen aukon.

Lauseen 1.5.9 yhteydessä on selvinnyt, että Gaussin ja Jordanin menettely johtaa siihen, että neliömatriisilla A:lla on oikeanpuoleinen käänteismatriisi — siis $X = X_{n\times n}$, jolla $AX = I$ — tasan silloin, kun se muokkautuu Gaussin ja Jordanin ratkaisumenettelyllä identtiseksi matriisiksi $I_{n\times n}$, jolla on determinantti 1. Kääntymätön muokkautuu muotoon

$$\begin{bmatrix} \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix},$$

jossa on pelkkiä nollia alarivinä ja siis determinantti 0. Koska mikään Gaussin–Jordanin rivioperaatioista ei voi muuttaa nollasta eroavaa determinanttia nollaksi eikä päinvastoin, on determinatti siis nollasta eroava tasan silloin, kun A:lla on oikeanpuoleinen käänteismatriisi.

Erityisesti on selvää, että kääntyvällä matriisilla deterinantti on nollasta eroava, mutta emme ole vielä selvittäneet, takaako tämä determinanttiehto (molemminpuolisen) kääntyvyyden. Asia voidaan nyt ratkaista vetoamalla transpoosin determinanttiin, joka on[34] $\det A^T = \det A$, siis oletuksen mukaan nollasta eroava. Transpoosilla on siis edellisen tarkastelun mukaan olemassa oikeanpuoleinen käänteismatriisi $Y = Y_{n\times n}$, jolle $A^T Y = I$. Nytpä Y^T on A:n vasemmanpuoleinen käänteismatriisi, sillä

$$Y^T A = (A^T Y)^T = I^T = I.$$

□

MÄÄRITELMÄ 2.5.2. Matriisin $A_{n\times n} = [a_{ij}]$ *kofaktorimatriisi*[35] on

$$\operatorname{cof} A = [\operatorname{cof} a_{ij}] = \begin{bmatrix} \operatorname{cof} a_{11} & \dots & \operatorname{cof} a_{1n} \\ \vdots & \ddots & \vdots \\ \operatorname{cof} a_{n1} & \dots & \operatorname{cof} a_{nn} \end{bmatrix}.$$

LAUSE 2.5.3. *Matriisille $A_{n\times n}$ on voimassa:*

$$A\,[\operatorname{cof} A]^T = \det A \cdot I.$$

TODISTUS. Seuraa välittömästi lauseista 2.4.1 ja 2.4.4. □

SEURAUS 2.5.4. *Olkoon matriisi $A_{n\times n}$ kääntyvä. Tällöin*

$$A^{-1} = \frac{1}{\det A}\,[\operatorname{cof} A]^T.$$

[34] Tässä ratkaiseva kohta.

[35] Toisinaan myös *liittomatriisi*.

2.6. Cramerin[36] sääntö.

Tarkastellaan lineaarista yhtälöryhmää $Ax = b$, missä A on $n \times n$–matriisi. Merkitään A_j:llä matriisia, joka on saatu A:sta korvaamalla sarake j vektorilla b, esimerkiksi

$$A_1 = \begin{bmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Nyt on voimassa ns. *Cramerin sääntö*:

LAUSE 2.6.1. *Olkoon* $\det A \neq 0$. *Tällöin yhtälöryhmällä* $Ax = b$ *on täsmälleen yksi ratkaisu*

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \dots, \quad x_n = \frac{\det A_n}{\det A}.$$

TODISTUS. $x = A^{-1}(b)$. Sijoita A^{-1}:n lauseke lauseesta 2.5.4 ja laske. $\square$

[36]GABRIEL CRAMER (1704-1752), sveitsiläinen matemaatikko ja fyysikko.

3. Vektoriavaruus ja sen dimensio eli ulotteisuus

Tässä luvussa alamme tutkia tasojen ja suorien korkeampiulotteisia vastineita. Samalla selviää, mitä oikein tarkoitetaan sillä, että jokin avaruus on vaikkapa 8–ulotteinen.

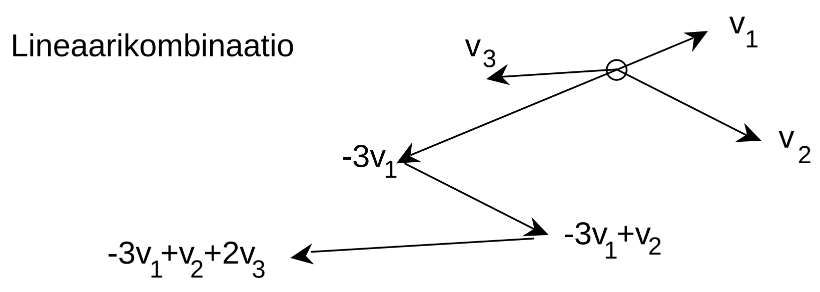
3.1. Lineaarikombinaatiot ja lineaarinen riippumattomuus.

Kahdesta vektorista on koulukurssin perusteella helppo sanoa, milloin ne ovat samansuuntaisia, milloin eivät. Korkeintaan voi epäilystä aiheuttaa nollavektori, jonka sovitaan olevan kaikkien vektoreiden suuntainen. Myös avaruudessa kolme vektoria voivat tietysti olla samansuuntaisia tai sitten kokonaan erisuuntaisia, kuten vaikkapa kantavektorit e_1, e_2 ja e_3. Mutta on kolmaskin mahdollisuus: kolme vektoria voivat olla saman tason suuntaisia olematta samansuuntaisia. Näin on asian laita esimerkiksi vektoreille e_1, e_2 ja e_1+e_2. Korkeampiulotteisissa avaruuksissa vektoreiden "erisuuntaisuuden aste" eli niiden virittämän avaruuden dimensio vaihtelee vielä enemmän. Aloitamme dimensiota koskevat tarkastelumme tilanteesta, jossa tutkittavat vektorit ovat täysin erisuuntaisia eli **lineaarisesti riippumattomia**.

Määritelmä 3.1.1. Avaruuden $\mathbb{R}^n$ vektoreiden $v_1, v_2, \dots, v_k$ *lineaarikombinaatioiksi* sanotaan vektoreita

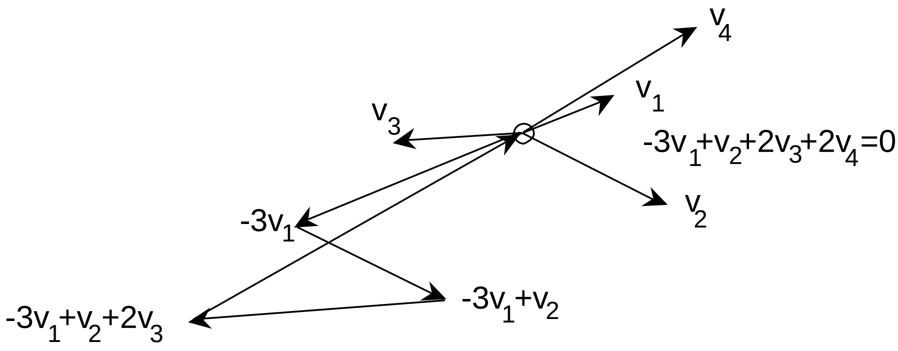
$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k,$$

missä luvut $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$ saavat olla mitä tahansa.



Esimerkiksi avaruudessa $\mathbb{R}^3$ kahden kantavektorin e_1 ja e_2 kaikkien lineaarikombinaatioiden joukko $\{\lambda_1 e_1 + \lambda_2 e_2 \mid \lambda_1 \in \mathbb{R}, \lambda_2 \in \mathbb{R}\}$ on sama kuin taso $\{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 = 0\}$.

MÄÄRITELMÄ 3.1.2. Avaruuden $\mathbb{R}^n$ vektorit $v_1, v_2, \ldots, v_k$ ovat (toisistaan) *lineaarisesti riippuvat*[37], jos niiden suuntaisista vektoreista voidaan muodostaa sulkeutuva silmukka,



Lineaarisesti riippuvat vektorit

eli jos on olemassa reaaliluvut $\lambda_1, \lambda_2, \ldots, \lambda_k$ siten, että

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0,$$

ja ainakin yksi **eroaa nollasta.** Jos tällaista silmukkaa ei voi muodosta, niin vektorit $v_1, v_2, \ldots, v_k$ ovat (toisistaan) *lineaarisesti riippumattomat.*[38] Vektorit $v_1, v_2, \ldots, v_k$ ovat siis lineaarisesti riippumattomat täsmälleen silloin, kun ehdosta

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0.$$

seuraa, että $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$.

Määritelmään saa tuntuman todistamalla sen perusteella seuraavat väitteet:

LAUSE 3.1.3.

a) *Jos jokin vektoreista $v_1, v_2, \ldots, v_k$ on nollavektori, ne ovat LD.*
b) *Yksinäinen vektori v_1 on LD ainoastaan, kun $v_1 = 0$.*
c) *Vektorit $v_1, v_2, \ldots, v_k$ ovat lineaarisesti riippumattomat täsmälleen silloin, kun yksi niistä*[39] *voidaan lausua muiden lineaarikombinaationa, siis esimerkiksi*
$$v_1 = \sum_{j \neq 1} \lambda_j v_j.$$
d) *Jos vektorit $v_1, v_2, \ldots, v_k$ ovat LD, niin $v_1, v_2, \ldots, v_k, u$ ovat LD, olipa u mikä tahansa vektori.*

Edellisen lauseen kohta d) sanoo, että lineaarisesti riippumattomat vektorit pysyvät riippumattomina, vaikka niistä jättäisi pois yhden tai vaikka useammankin (induktio!).

[37]Yleisesti tälle on käytössä lyhenne LD englannin sanoista linearly dependent.
[38]Tämä on juuri etsimäämme täydellistä erisuuntaisuutta. Lineaariselle riippumattomuudelle on käytössä lyhenne LI englannin sanoista linearly independent.
[39]Mutta ei välttämättä jokainen!

LAUSE 3.1.4. *Olkoon*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}; \qquad a_{ij} \in \mathbb{R}.$$

Tällöin A:n sarakevektorit ovat LI , *jos ja vain jos yhtälöryhmällä* $Ax = 0$ *on pelkkä triviaaliratkaisu.*

TODISTUS. Kumpikin väite sanoo saman:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

vain, kun jokainen x_i on 0, ts.

$$\sum_{i=1}^{n} x_i a_{*i} = 0 \implies x_1 = \dots = x_n = 0,$$

missä symbolilla a_{*i} on luonnollisesti merkitty matriisin i:nnettä saraketta. □

LAUSE 3.1.5. *Olkoon* $D : (\mathbb{R}^n)^k \to \mathbb{R}$ *alternoiva* $k-$*lineaarikuvaus ja vektorit* $v_1, \dots, v_k \in \mathbb{R}^n$ *lineaarisesti riippuvat, siis LD. Tällöin* $D(v_1, \dots, v_k) = 0$.

Erityisesti determinantti

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

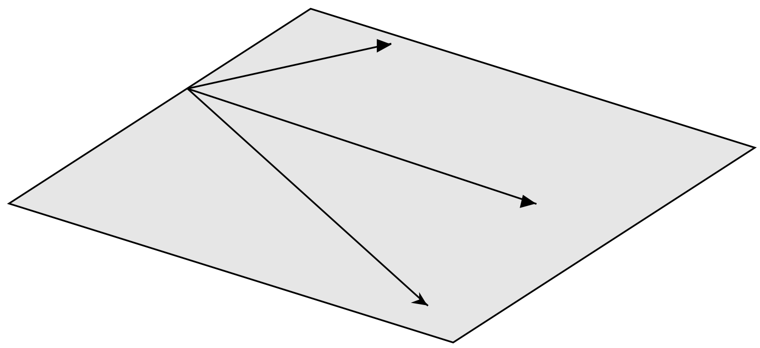
on 0, jos sen sarakkeet riippuvat toisistaan lineaarisesti.

TODISTUS. Jos vektorit $v_1, \dots, v_k$ riippuvat toisistaan lineaarisesti, niin ainakin yksi voidaan lausua muiden lineaarikombinaationa. Koska sarakkeiden vaihto vain vaihtaa multilineaarikuvauksen arvon merkin, voimme olettaa, että ensimmäinen sarake on muiden lineaarikombinaatio, toisin sanoen sopivin kertoimin λ_i pätee $v_1 = \sum_{i=2}^{k} \lambda_i v_i$ ja siis

$$\begin{aligned} D(v_1, v_2, \dots, v_k) &= D(\sum_{i=2}^{k} \lambda_i v_i \, , v_2, \dots, v_k) \\ &= \sum_{i=2}^{k} \lambda_i \underbrace{D(v_i, v_2, \dots, v_k)}_{0 \text{ (kaksi samaa)}} = 0. \end{aligned}$$

□

Tosiasiassa determinantin häviäminen myös takaa, että sen sarakkeet ovat lineaarisesti riippuvat. Tämä on osa seuraavaa lausetta, johon on koottu neliömatriisin kääntyvyysehtoja.



Kolme lineaarisesti riippuvaa vektoria ovat samassa tasossa

Lause 3.1.6 (Jatkoa lauseille 1.5.8 ja 2.5.1). *Neliömatriisille $A_{n\times n}$ seuraavat ovat yhtäpitäviä:*

(a) $\det A \neq 0$.
(b) *A on kääntyvä, ts. on olemassa $X = X_{n\times n}$ siten, että $XA = AX = I$.*
(c) *Yhtälöryhmällä $Ax = b$ on (kaikilla $b \in \mathbb{R}^n$!) täsmälleen yksi ratkaisu, nimittäin $x = A^{-1}b$.*
(d) *Yhtälöryhmällä $Ax = b$ on (kaikilla $b \in \mathbb{R}^n$!) täsmälleen yksi ratkaisu.*
(e) *Kuvaus $T_A : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto Ax$ on bijektio.*
(f) *Yhtälöryhmällä $Ax = b$ on (kaikilla $b \in \mathbb{R}^n$!) enintään yksi ratkaisu.*
(g) *Kuvaus $T_A : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto Ax$ on injektio.*
(h) *Homogeenisella yhtälöryhmällä $Ax = 0$ on vain triviaaliratkaisu.*
(i) *A:n sarakevektorit ovat* LI.
(j) *On olemassa $X = X_{n\times n}$ siten, että $AX = I$.*
(k) *On olemassa $X = X_{n\times n}$ siten, että $XA = I$.*
(l) *A:n rivivektorit ovat* LI.
(m) *A^T on kääntyvä.*

Todistus. Väite on yhdistelmä lauseista 1.5.9, 2.2.9, 2.5.9, 3.1.4 ja 3.1.5. □

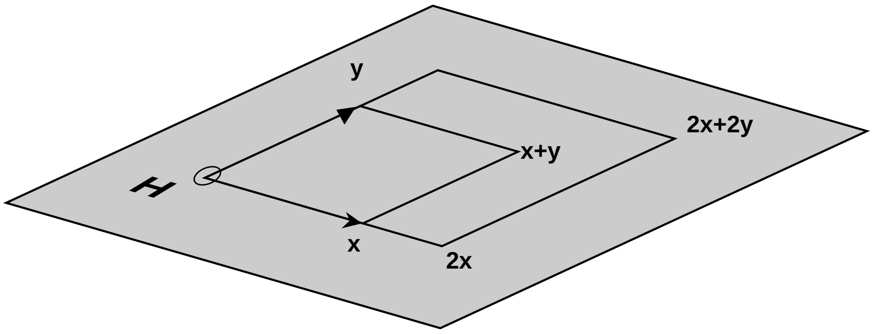
Lineaarinen riippumattomuus ja riippuvuus voidaan määritellä myös äärettömän monelle vektorille. Palaamme tähän kohdassa 3.7.

3.2. Aliavaruus.

Yleistämme seuraavassa suoran ja tason käsitteet aliavaruuden käsitteeksi. Tarkoituksena on, että kolmiulotteisen avaruuden $\mathbb{R}^3$ aliavaruuksia olisivat origon kautta kulkevat suorat ja tasot, joita usein ajattelemme alempiulotteisten avaruuksien $\mathbb{R}^1$ ja $\mathbb{R}^2$ kopioina. Vastaavasti esimerkiksi $\mathbb{R}^4$:llä on myös kolmiulotteisia aitoja aliavaruuksia, origon kautta kulkevia avaruuden $\mathbb{R}^3$ kopioita. Sen sijaan muita kuin origon kautta kulkevia suoria ja tasoja ei tulla kelpuuttamaan linaarisiksi aliavaruuksiksi.[40]

Määritelmä 3.2.1. Avaruuden $\mathbb{R}^n$ epätyhjä osajoukko H on $\mathbb{R}^n$:n *vektorialivavaruus* eli *lineaarinen aliavaruus*, lyhyesti *aliavaruus*, jos H sisältää kaikki alkioidensa lineaarikombinaatiot.

[40]Tilanne on samanlainen kuin lineaarifunktioidenkin kohdalla. Muista, että esimerkiksi funktio $f : \mathbb{R} \to \mathbb{R} : x \mapsto f(x) = 2x + 3$ ei lineaarialgebran terminologiassa ole lineaarinen, vaikka sen kuvaaja on suora viiva ja koulussa sitä on tapana sanoa lineaariseksi.



On geometrisesti ilmeistä, että kolmiulotteisen avaruuden $\mathbb{R}^3$ aliavaruuksia ovat ainakin origon kautta kulkevat suorat ja tasot.

Tarkastaessamme onko jokin joukko $H \subset \mathbb{R}^n$ aliavaruus vai ei joudumme periaatteessa tutkimaan kaikkia H:n alkioiden lineaarikombinaatioita. Käytännössä ei kuitenkaan tarvitse menetellä näin hankalalla tavalla, vaan riittää tutkia kaikki kahden vektorin summat ja kaikki vektoreiden reaaliset monikerrat, sillä pätee seuraava aliavaruuden karakterisointilause:

Lause 3.2.2. *Avaruuden $\mathbb{R}^n$ epätyhjä osajoukko H on $\mathbb{R}^n$:n aliavaruus jos ja vain jos seuraavat ehdot pätevät*

(A-1) $$u, v \in H \implies u + v \in H$$

(A-2) $$\lambda \in \mathbb{R} \text{ ja } u \in H \implies \lambda u \in H$$

Todistus. Tietysti jokaisella aliavaruudella $H \subset \mathbb{R}^n$ on ominaisuudet (A-1) ja (A-2), ovathan $u + v$ ja λu joukon H alkioiden lineaarikombinaatioita, kun $\lambda \in \mathbb{R}$ ja $u, v \in H$.

Olkoon H avaruuden $\mathbb{R}^n$ epätyhjä osajoukko, jolla on ominaisuudet (A-1) ja (A-2). On osoitettava, että H sisältää kaikki alkioidensa lineaarikombinaatiot. Olkoon $u = \sum_{i=1}^{k} \lambda_i v_i = \lambda_1 v_1 + \cdots + \lambda_k v_k$ sellainen, ts. jokainen $\lambda_i \in \mathbb{R}$ ja $v_i \in H$. Nyt oletuksen (A-2) mukaan jokainen $\lambda_i v_i$ kuuluu joukkoon H ja siis oletuksen (A-1) mukaan $\lambda_1 v_1 + \lambda_2 v_2 \in H$, samoin siis $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = (\lambda_1 v_1 + \lambda_2 v_2) + \lambda_3 v_3 \in H$ jne. kunnes $\lambda_1 v_1 + \cdots + \lambda_k v_k \in H$. (Induktio!) $\square$

Todistamalla seuraavan lauseen voi testata onko ymmärtänyt aliavaruuden määritelmän oikein.

Lause 3.2.3.

a) *$\mathbb{R}^n$:n jokainen aliavaruus sisältää $\mathbb{R}^n$:n nollavektorin,*
b) *$\mathbb{R}^n$:n triviaali aliavaruus $\{0\}$ ja koko avaruus $\mathbb{R}^n$ ovat $\mathbb{R}^n$:n aliavaruuksia.*

Todistus. Harjoitustehtävä. $\square$

Lause 3.2.4. *Olkoot H ja K avaruuden $\mathbb{R}^n$ aliavaruuksia. Silloin leikkaus $H \cap K$ on $\mathbb{R}^n$:n aliavaruus, mutta yhdiste $H \cup K$ ei yleensä ole $\mathbb{R}^n$:n aliavaruus.*

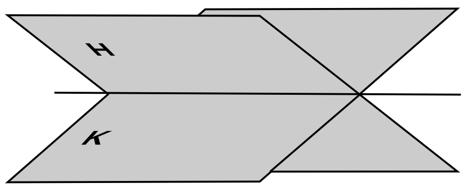
Todistus. Olkoot H ja K avaruuden $\mathbb{R}^n$ aliavaruuksia. Testaamme lauseen 3.2.2. kriteereillä onko $H \cap K$ avaruuden $\mathbb{R}^n$ aliavaruus. Käytämme tietoa, että H

ja K ovat aliavaruuksia:

$$\begin{aligned} u, v \in H \cap K &\implies u, v \in H \text{ ja } u, v \in K \\ &\implies u+v \in H \text{ ja } u+v \in K \\ &\implies u+v \in H \cap K. \end{aligned} \tag{A-1}$$

$$\text{ja} \qquad \begin{aligned} \lambda \in \mathbb{R} \text{ ja } u \in H \cap K &\implies \lambda \in \mathbb{R} \text{ ja } u \in H \text{ ja } u \in K \\ &\implies \lambda u \in H \text{ ja } \lambda u \in K \\ &\implies \lambda u \in H \cap K \end{aligned} \tag{A-2}$$

Leikkaus on siis aliavaruus. On vielä keksittävä esimerkki kahdesta aliavaruudesta, joiden yhdiste ei ole aliavaruus. Tällaiseksi kelpaavat avaruuden $\mathbb{R}^2$ koordinaattiakselit, joiden yhdiste ei tietenkään sisällä kaikkia alkioidensa summia. Esimerkiksi kantavektorit $e_1$ja e_2 kuuluvat tähän yhdisteeseen, mutta summa $e_1 + e_2$ ei kuulu. $\square$

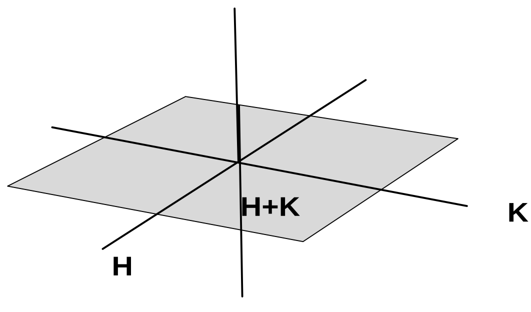


Perusesimerkki kahden aliavaruuden leikkauksesta avaruudessa $\mathbb{R}^3$ on tietysti kahden origon kautta kulkevan tason leikkaus, joka on origon kautta kulkeva suora, elleivät tasot yhdy.

Kahdesta aliavaruudesta voi luonnollisella tavalla rakentaa aliavaruuden, joka sisältää molemmat annetut aliavaruudet, vaikka yhdiste ei tähän tarkoitukseen yleensä kelpaakaan.

Määritelmä 3.2.5. Olkoot H ja K avaruuden $\mathbb{R}^n$ aliavaruuksia. Määritellään *H:n ja K:n summa*:

$$H + K = \{u + v \mid u \in H \text{ ja } v \in K\}.$$



Perusesimerkki aliavauuksien summasta on avaruuden $\mathbb{R}^3$ x–akselin ja y–akselin summa, joka on xy–taso — aliavaruus sekin. Aliavaruuksien summan määritelmässä on ideana luoda mahdollisimman pieni aliavaruus, joka sisältää annetut aliavaruudet H ja K. Seuraavat lauseet sanovat, että onnistuimme.

LAUSE 3.2.6. *Olkoot H ja K avaruuden $\mathbb{R}^n$ aliavaruuksia. Silloin $H + K$ on $\mathbb{R}^n$:n aliavaruus.*

TODISTUS. Samantapainen kuin lauseen 3.2.4. todistus. $\square$

LAUSE 3.2.7. *Olkoot H ja K avaruuden $\mathbb{R}^n$ aliavaruuksia. Silloin*

a) *$H \cup K \subset H + K$,*

b) *$H + K$ on suppein $\mathbb{R}^n$:n aliavaruus, joka sisältää $H \cup K$:n.*

TODISTUS. a) $H \subset H + K = \{u + v \mid u \in H \text{ ja } v \in K\}$, sillä jokainen vektori $h \in H$ on muotoa $h = h + 0$ ja tiedämme, että $0 \in K$. Vastaavasti tarkastetaan, että $K \subset H + K$. Siis $H \cup K \subset H + K$.

b) Meidän on nyt todistettava, että $H + K$ on suppein $\mathbb{R}^n$:n aliavaruus, joka sisältää yhdisteen $H \cup K$ eli molemmat alkuperäiset aliavaruudet. Otetaan tutkittavaksi aliavaruus L, jolle pätee $H \subset L$ ja $K \subset L$ ja väitetään, että $H + K \subset L$. Tämä väite tarkoittaa, että jokainen joukon $H + K$ alkio kuuluu myös joukkoon L. Testaamme onko asian laita näin. Otetaan tutkittavaksi joukon $H + K$ alkio, siis summavektori $h + k$, missä $h \in H$ ja $k \in K$. Koska L on aliavaruus, niin ehdon (A-1) nojalla:

$$\left.\begin{array}{l} h \in H \subset L \implies h \in L \\ k \in K \subset L \implies k \in L \end{array}\right\} \implies h + k \in L. \quad \square$$

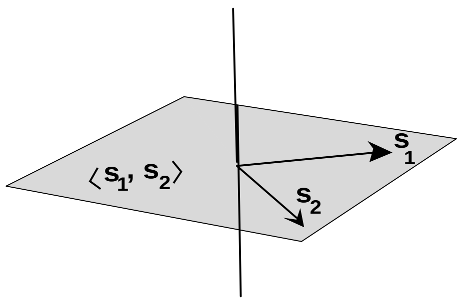
Kolmen tai useammankin aliavaruuden summa muodostetaan vastaavalla tavalla. Erityisesti voi aliavaruuden muodostaa aloittamalla joistakin vektoreista, jatkamalla ne origon kautta kulkeviksi suoriksi ja muodostamalla suorien summan. Seuraavassa pykälässä näin tehdään sallien jopa, että aloitetaan äärettömän monella vektorilla.

3.3. Lineaarinen verho.

MÄÄRITELMÄ 3.3.1. Olkoon S avaruuden $\mathbb{R}^n$ epätyhjä osajoukko. Merkitään

$$\begin{aligned} < S > &= \{u \in \mathbb{R}^n \mid u \text{ on äärellisen monen } S\text{:n alkion lineaarikombinaatio}\} \\ &= \{\lambda_1 s_1 + \cdots + \lambda_k s_k \mid k \in \mathbb{N}, \lambda_1, \ldots, \lambda_k \in \mathbb{R} \text{ ja } s_1, \ldots, s_k \in S\} \end{aligned}$$

Joukko $< S >$ on nimeltään *joukon S lineaarinen verho.* Jos S on äärellinen joukko $\{s_1, s_2, \ldots, s_n\}$, niin merkitään $< S >=< s_1, s_2, \ldots, s_n >$.



Lineaarisesta verhosta $< S >$ käytetään myös nimeä *joukon S virittämä aliavaruus* tai *kaikkien vektoreiden $v \in S$ virittämä aliavaruus*. Tämä on oikeutettua, sillä pätee:

LAUSE 3.3.2. *Avaruuden $\mathbb{R}^n$ epätyhjän osajoukon S lineaarinen verho $< S >$ on $\mathbb{R}^n$:n aliavaruus, vieläpä suppein joukon S sisältävä aliavaruus. Erityisesti, jos H ja K ovat $\mathbb{R}^n$:n aliavaruuksia, niin $< H \cup K >= H + K$.*

TODISTUS. Joukon $< S >$ voi todistaa aliavaruudeksi jäljittelemällä lauseen 3.2.6. todistusta. Tietysti $S \subset< S >$. Jokainen joukkoa S laajempi aliavaruus $L \supset S$ sisältää kaikki alkioidensa lineaarikombinaatiot, erityisesti S:n alkoiden linaarikombinaatiot, joiden joukko on $< S >$. $\square$

LAUSE 3.3.3. *Olkoon $S \subset \mathbb{R}^n$ epätyhjä joukko vektoreita ja $x \in \mathbb{R}^n$. Tällöin*

$$< S \cup \{x\} >=< S > \iff x \in< S > .$$

TODISTUS. Väite seuraa suoraan määritelmästä 3.3.1 $\square$

On ilmeistä, että kahdesta vektorista ei voi lineaarikombinaatioina saada kolmea lineaarisesti riippumatonta vektoria. Seuraava tärkeä lause sanoo, että vastaava asiantila vallitsee useammankin vektorin kohdalla. Todistus on suora lasku ja siksi pitkän näköinen.

LAUSE 3.3.4. *Olkoon $S = \{v_1, \ldots, v_k\}$ k–alkioinen vektorijoukko. Tällöin lineaarisen verhon $< v_1, \ldots, v_k >$ jokainen $(k+1)$–alkioinen osajoukko on* LD.

TODISTUS. Tehtävänä on todistaa, että — riippumatta matriisista (a_{ij}) — vektorit

$$\begin{aligned}
u_1 &= a_{11}v_1 + a_{12}v_2 + \cdots + a_{1k}v_k \\
&\vdots \\
u_k &= a_{k1}v_1 + a_{k2}v_2 + \cdots + a_{kk}v_k \\
\text{ja} \quad u_{(k+1)} &= a_{(k+1)1}v_1 + a_{(k+1)2}v_2 + \cdots + a_{(k+1)k}v_k
\end{aligned}$$

ovat LD, eli että vektoriyhtälöllä

$$(*) \qquad \lambda_1 u_1 + \cdots + \lambda_{(k+1)}u_{(k+1)} = 0$$

on epätriviaali ratkaisu $\lambda = (\lambda_1, \ldots, \lambda_{(k+1)})$. Auki kirjoitettuna yhtälö $(*)$ on

$$\begin{array}{lllll}
 & a_{11}\lambda_1 v_1 + & a_{12}\lambda_1 v_2 + \cdots + & a_{1k}\lambda_1 v_k & \\
 & \vdots & \vdots & \vdots & \\
+ & a_{k1}\lambda_k v_1 + & a_{k2}\lambda_k v_2 + \cdots + & a_{kk}\lambda_k v_k, & \\
+ & a_{(k+1)1}\lambda_{(k+1)}v_1 + & a_{(k+1)2}\lambda_{(k+1)}v_2 + \cdots + & a_{(k+1)k}\lambda_{(k+1)}v_k & = \quad 0
\end{array}$$

eli sarakkeittain summaten

$$\begin{aligned}
&\left(a_{11}\lambda_1 + \cdots + a_{k1}\lambda_k + a_{(k+1)1}\lambda_{(k+1)}\right) v_1 + \\
&\left(a_{12}\lambda_1 + \cdots + a_{k2}\lambda_k + a_{(k+1)2}\lambda_{(k+1)}\right) v_2 + \\
&\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\
&\left(a_{1k}\lambda_1 + \cdots + a_{kk}\lambda_k + a_{(k+1)k}\lambda_{(k+1)}\right) v_k = 0.
\end{aligned}$$

Tämä yhtälö toteutuu varmasti ainakin silloin, kun jokaisen v_j:n kerroin on 0, eli kun homogeeninen lineaarinen yhtälöryhmä

$$\begin{cases} a_{11}\lambda_1 + \cdots + a_{k1}\lambda_k + a_{(k+1)1}\lambda_{(k+1)} = 0 \\ a_{12}\lambda_1 + \cdots + a_{k2}\lambda_k + a_{(k+1)2}\lambda_{(k+1)} = 0 \\ \vdots \qquad\qquad\qquad \vdots \\ a_{1k}\lambda_1 + \cdots + a_{kk}\lambda_k + a_{(k+1)k}\lambda_{(k+1)} = 0. \end{cases}$$

toteutuu. Mutta tässäpä on enemmän tuntemattomia kuin yhtälöitä, joten epätriviaali ratkaisu on olemassa.

□

SEURAUS 3.3.5. *Olkoon* $\{v_1, v_2, \ldots, v_n\} \subset \mathbb{R}^m$ *ja* $n > m$. *Tällöin vektorit* $v_1, v_2, \ldots, v_n$ *ovat* LD.

TODISTUS. Tunnemme tuloksen ennestään, mutta se seuraa myös edellisestä, sillä $\mathbb{R}^m$ on kantavektoreidensa joukon lineaarinen verho.

□

3.4. Kanta, koordinaatit ja dimensio.

Valpas lukija on epäilemättä arvannut, että tarkoituksenamme on rakentaa koordinaatistoja aliavaruuksiin. Jotakin tämäntapaista on jo saatu aikaankin: Vektoreiden $v_1, \ldots, v_k$ virittämän aliavaruuden alkiot ovat muotoa $x = x_1v_1 + \cdots + x_kv_k$ ja tuntuu luonnolliselta pitää kertoimia x_i "vinokulmaisina koordinaatteina". Vinokulmaisetkaan koordinaattiakselit eivät tietenkään saa olla missään mielessä samansuuntaisia. Siksi asetetaan seuraava määritelmä.

MÄÄRITELMÄ 3.4.1. Aliavaruuden $H \subset \mathbb{R}^m$ osajoukko $S = \{v_1, v_2, \ldots, v_n\} \subset H$ on H:n *kanta*, jos

(K-1)
$v_1, v_2, \ldots, v_n$ virittävät avaruuden H, ts. $< S >= H$ ja

(K-2)
$v_1, v_2, \ldots, v_n$ ovat lineaarisesti riippumattomia (erityisesti eri vektoreita)

Perusesimerkki kannasta on avaruuden $\mathbb{R}^n$ standardikanta $\{e_1, \ldots, e_n\}$. Jokainen lineaarisesti riippumaton vektorijoukko on virittämänsä avaruuden $< v_1, v_2, \ldots, v_n >$ eli lineaarisen verhonsa kanta.

Kanta tuottaa virittämäänsä aliavaruuteen koordinaatit[41] seuraavassa mielessä.

LAUSE 3.4.2. *Jos* $v_1, \ldots, v_k$ *ovat lineaarisesti riippumattomia, niin jokainen niiden virittämän aliavaruuden* $< v_1, \ldots, v_k >$ *vektori voidaan lausua* **vain yhdellä tavalla** *lineaarikombinaationa*

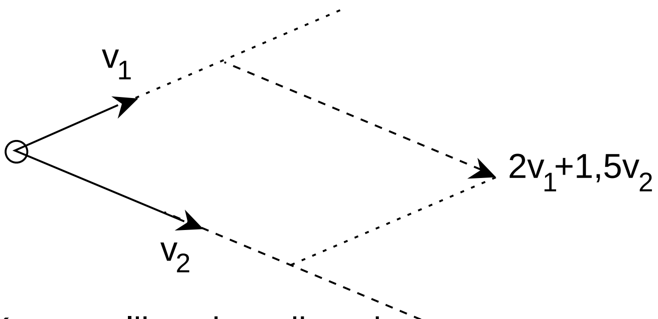
$$v = \sum_{i=1}^{k} \lambda_i v_i.$$

[41] Yleensä "vinokulmaiset"

TODISTUS. Lause sanoo, että lineaarisesti riippumattomille vektoreille $v_1, v_2, \ldots, v_n$ pätee

$$\sum_{i=1}^{k} \lambda_i v_i = \sum_{i=1}^{k} \mu_i v_i \iff \lambda_i = \mu_i \text{ kaikilla } i = 1, \ldots k.$$

Tutkitaan, onko näin. Jos $\sum_{i=1}^{k} \lambda_i v_i = \sum_{i=1}^{k} \mu_i v_i$, niin $\sum_{i=1}^{k} (\lambda_i - \mu_i) v_i = 0$, joten lineaarisen riippumattomuuden nojalla $\lambda_i - \mu_i = 0$ kaikilla i. $\square$



Kantaan liittyy koordinaatisto

Vektori v lausuttuna aliavaruuden H kannassa $F = (v_1, \ldots, v_k)$ on summa $v = \sum_{i=1}^{k} \lambda_i v_i$. Summa määräytyy kertoimien jonosta $(\lambda_1, \ldots \lambda_k)$. Vektorin v voi siis ilmaista luettelemalla kertoimet ja mainitsemalla kannan, siis vaikka kantamerkillä varustettuna sarakevektorina

$$v = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{bmatrix}_F.$$

Kertoimet λ_j ovat luonnollisesti nimeltään vektorin v *koordinaatit* aliavaruuden $H =< v_1, v_2, \ldots, v_n >$ kannassa[42] $F = (v_1, \ldots, v_k)$.

Kanta on siis tarpeen, kun halutaan laskea koordinaatein aliavaruudessa. Lisäksi kanta kertoo olennaisen tiedon aliavaruuden rakenteesta, sen ulotteisuuden eli dimension. Dimensio on kantavektoreiden lukumäärä. Tämä olisi mieletön määritelmä, ellei pätisi seuraava lohdullinen lause.

LAUSE 3.4.3. *Aliavaruuden $H \subset \mathbb{R}^n$ jokainen kanta sisältää yhtä monta vektoria. Erityisesti koko avaruuden $\mathbb{R}^n$ jokainen kanta sisältää n vektoria kuten standardikanta.*

TODISTUS. Seuraa välittömästi lauseesta 3.3.4. $\square$

MÄÄRITELMÄ 3.4.4. Jos aliavaruudella $H \subset \mathbb{R}^n$ on m–alkioinen kanta, niin H:n *dimensio* eli *ulotteisuus* on m, merkitään $\dim H = m$. Jos $H = \{0\}$, merkitään $\dim H = 0$.

Dimensio kertoo jotakin aliavaruuden koosta:

[42]Kantavektorit on tästä alkaen lueteltu kaarisulkeissa. Tällä korostetaan, että luettelon järjestys on otettava huomioon. Muutenhan koordinaatit vaihtuvat keskenään.

LAUSE 3.4.5. *Olkoot* $H \subset K \subset \mathbb{R}^n$ *aliavaruuksia.*
a) $\dim H \leq \dim K \leq n$.
b) $\dim H = \dim K \iff H = K$.

TODISTUS. Olkoon $S = (v_1, \ldots, v_k)$ pienemmän aliavaruuden H kanta. Erikoistapauksessa $H = K$ se on samalla K:n kanta, joten tällöin $\dim H = \dim K$ eikä siitä sen enempää, mutta muuten $H =< S >$ on K:n aito aliavaruus, jolloin on olemassa ainakin yksi vektori $u_1 \in K \setminus < S >$. Olkoon $S_1 = S \cup \{u_1\}$ joukko, joka saadaan lisäämällä H:n kantaan siitä lineaarisesti riippumaton vektori $u \in K$. Jos näin muodostettu S_1 on K:n kanta, on tehtävä suoritettu ja $\dim K = \dim H + 1$, mutta muussa tapauksessa toistetaan menettelyä, kunnes joukosta K ei enää löydy uutta lineaarisesti riippumatonta vektoria, vaan jokainen K:n vektori on lineaarikombinaatio jo saaduista, jotka siis muodostavat etsityn kannan. Tämä tapahtuu viimeistään n:n askeleen jälkeen, sillä totesimme edellä, että edes koko avaruudessa $\mathbb{R}^n$ ei ole $n+1$:tä lineaarisesti riippumatonta vektoria. Huomaamme siis, että $\dim H < \dim K \leq n$, ellei H ole koko K. □

Edellisen lauseen todistuksessa jouduttiin laajentamaan lineaarisesti riippumaton joukko kannaksi. Tämän joutuu useinkin tekemään. Siksi kirjaamme jo todistetun lauseen:

LAUSE 3.4.6. *Olkoon* $H \subset \mathbb{R}^n$ *aliavaruus ja* $(v_1, v_2, \ldots, v_k)$ *jono* H*:n lineaarisesti riippumattomia vektoreita. Tällöin on olemassa* H*:n kanta* T *siten, että jokainen* v_j *on kantavektori. Toisin sanoen jokainen aliavaruuden* H *lineaarisesti riippumaton osajoukko voidaan laajentaa sen kannaksi.*

Lauseella on luonnollinen pari, joka koskee lineaarista verhoa:

LAUSE 3.4.7. *Olkoon* $S \subset \mathbb{R}^n$ *ja* $H =< S >$*. Tällöin on olemassa* H*:n kanta, jonka kantavektorit kuuluvat joukkoon* S*. Toisin sanoen jokainen aliavaruuden* H *virittäjäjoukko voidaan supistaa sen kannaksi.*

Kantaa on hiukan helpompi tutkia ja käsitellä, jos jo tietää avaruuden dimension. Seuraava lause sanoo tästä jotakin.

SEURAUS 3.4.8. *Mitkä tahansa* n *lineaarisesti riippumatonta vektoria* $v_1, \ldots, v_n \in \mathbb{R}^n$ *virittävät* $n-$*vektoreiden avaruuden* $\mathbb{R}^n$*, joten* n *lineaarisesti riippumatonta vektoria* $v_1, \ldots, v_n \in \mathbb{R}^n$ *muodostavat koko avaruuden* $\mathbb{R}^n$ *kannan.*

TODISTUS. Lauseen 3.4.6 mukaan $v_1, v_2, \ldots, v_n$ voidaan laajentaa avaruuden $\mathbb{R}^n$ kannaksi, mutta lauseen 3.4.3 nojalla laajennettuun kantaan ei tule yhtään uutta vektoria. □

Kootaan vielä yhteen perusasiat n–ulotteisen avaruuden $\mathbb{R}^n$ kannasta:

SEURAUS 3.4.9. *Olkoon* S n*–ulotteisen avaruuden* $\mathbb{R}^n$ n*–alkioinen osajoukko. Tällöin seuraavat ehdot ovat yhtäpitäviä keskenään:*

(1) S *on* $\mathbb{R}^n$*:n kanta*
(2) S*:n vektorit ovat lineaarisesti riippumattomat*
(3) $< S >= \mathbb{R}^n$.

Todistus. Ehto (1) merkitsee samaa kuin (2) ja (3) yhdessä, joten riittää todistaa, että (2) $\iff$ (3). Lause 3.4.3. sanookin jo, että n lineaarisesti riippumatonta vektoria virittävät koko avaruuden $\mathbb{R}^n$. Toisaalta n vektoria, jotka virittävät koko avaruuden ovat lineaarisesti riippumattomat, sillä lauseen 3.4.7 mukaan virittäjäjoukko voidaan supistaa avaruuden kannaksi, mutta lauseen 3.4.3 mukaan kantaan on jätettävä n vektoria, siis kaikki. $\square$

3.5. Vektoriavaruus.

Tähän asti olemme tarkastelleet yksinomaan **tavallisia n–komponenttisia vektoreita** $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Tuntuu toisaalta selvältä, että jokainen $\mathbb{R}^n$:n aliavaruus $H \subset \mathbb{R}^n$ on itsessään ihan "hyvä vektoriavaruus", jonkinlainen kopio jostain alempiulotteisesta avaruudesta $\mathbb{R}^k$. Esimerkiksi tavallisen avaruuden $\mathbb{R}^3$ jokainen origon kautta kulkeva taso[43] on tietyssä mielessä kopio kaksiulotteisesta avaruudesta $\mathbb{R}^2$. Näiden aika ilmeisten kohteiden lisäksi on mahdollista soveltaa kaikkea tähän asti oppimaamme valtavan paljon suurempaan joukkoon matemaattisia objekteja, joita sitten myös voi sanoa "vektoreiksi", mahdollisesti "ääretönulotteisissa" avaruuksissa. Tähän pyrkii seuraava määritelmä ja koko tämä luku, joka nostaa tarkastelumme uudelle abstraktiotasolle määrittelemällä vektorin aksiomaattisesti muodollisten ominaisuuksiensa avulla. Tämä on alussa lupaamamme ihmeellinen vektorin määritelmä. Vektori on siis **mikä tahansa** olio, kunhan sillä lasketaan "kuten vektorilla lasketaan"!

Määritelmä 3.5.1. *Vektoriavaruus* eli *reaalinen lineaariavaruus* on kolmikko $(V, +, \cdot)$, missä V on epätyhjä joukko ja $+$ ja $\cdot$ ovat laskutoimitukset nimeltään *summa* ja *tulo*

$$\begin{aligned} + &: V \times V \to V \\ \cdot &: \mathbb{R} \times V \to V, \end{aligned}$$

jotka toteuttavat seuraavat **vektoriavaruuden aksioomat**

1. 1.1. $u + (v + w) = (u + v) + w$ kaikille $u, v, w \in V$.
   1.2. $u + v = v + u$ kaikille $u, v \in V$.
   1.3. On olemassa ns. *nollavektori*, merkitään $0 \in V$, siten, että $0 + u = u$ kaikille $u \in V$.
   1.4. Kaikille $u \in V$ on olemassa ns. *vastavektori*, merkitään $-u \in V$, siten, että $u + (-u) = 0$.
2. 2.1. $(\lambda\mu)u = \lambda(\mu u)$ kaikille $\lambda, \mu \in \mathbb{R}$ ja $u \in V$.
   2.2. $1 \cdot u = u$ kaikille $u \in V$.
3. 3.1. $\lambda(u + v) = \lambda u + \lambda v$ kaikille $\lambda \in \mathbb{R}$ ja $u, v \in V$.
   3.2. $(\lambda + \mu)u = \lambda u + \mu u$ kaikille $\lambda, \mu \in \mathbb{R}$ ja $u \in V$.

V:n alkioita sanotaan seuraavassa *vektoreiksi.*

Vektoriavaruuden määritelmän idea on se, että aksioomina on lueteltu riittävästi yhteenlaskun ja tulon ominaisuuksia, jotta muut "vektoriavaruuksissa tarvitut" lauseet seuraavat niistä. On tietysti periaatteessa makuasia, mitkä ominaisuudet on haluttu mukaan. Loogiselta kannalta on niin, että valinta on tehty laadittaessa

[43]Toki muutkin tasot, mutta kopioita hieman eri mielessä.

yo. lista ja voimme nyt vain todeta, mitkä $\mathbb{R}^n$:ssä tutuista väitteistä seuraavat näistä aksioomista eli **ovat voimassa jokaisessa vektoriavaruudessa** V. Tällaisia ovat mm. seuraavat lauseet:

LAUSE 3.5.2.

a) *On olemassa vain yksi nollavektori.*
b) *Kullakin vektorilla on vain yksi vastavektori.*
c) *Olkoot $u \in V$ ja $v \in V$. Tällöin on olemassa täsmälleen yksi vektori $w \in V$ siten, että $u + w = v$. Tämä vektori on $w = (v + (-u))$ ja sitä sanotaan v:n ja u:n erotukseksi ja merkitään $w = v - u$.*

TODISTUS. a) Olkoot Φ ja Ψ nollavektoreita vektoriavaruudessa V. Tällöin yhtälöt

$$(*) \qquad x = \Phi + x$$

ja

$$(**) \qquad y + \Psi = y$$

pätevät kaikille vektoreille x ja $y \in V$. Valitsemalla x :ksi Ψ ja y:ksi Φ saadaan molemmat yhdistämällä tulos:

$$\Phi = \Phi + \Psi = \Phi.$$

b) Samantapainen todistus

c) Ainakin lauseessa mainittu vektori $w\,(= v + (-u))$ tekee sen mitä väitetään, sillä

$$u + w = u + (v + (-u)) \overset{1.2.}{=} u + ((-u) + v) \overset{1.1.}{=} (u + (-u)) + v \overset{1.4.}{=} 0 + v \overset{1.3.}{=} v.$$

Erotuksen yksikäsitteisyys todistetaan suunnilleen samalla tavalla kuin nollan yksikäsitteisyys, siis esimerkiksi näin: Jos $u + w = v$ ja $u + w' = v$, niin $u + w$ ja $u + w'$ ovat sama vektori (nimittäin v), eli

$$(*) \qquad u + w = u + w'.$$

Aksiooman 1.4. mukaan on olemassa vastavektori $(-u)$. Lisätään se yhtälöön $(*)$ puolittain ja saadaan

$$\begin{aligned} & (-u) + (u + w) = (-u) + (u + w'), \\ \text{josta aksiooman 1.1. nojalla} \quad & (-u + u) + w = (-u + u) + w', \\ \text{josta aksiooman 1.4. nojalla} \quad & 0 + w = 0 + w', \\ \text{josta aksiooman 1.3. nojalla} \quad & w = w'. \end{aligned}$$

□

Lause 3.5.3. *Kaikille $u, v \in V$ ja $\lambda, \mu \in \mathbb{R}$ on voimassa:*

a) $-(-u) = u$.
b) $-(u + v) = -u - v$.
c) $-u = (-1)u$.
d) $-(\lambda u) = (-\lambda)u = \lambda(-u)$.
e) $(-\lambda)(-u) = \lambda u$.
f) $\lambda(u - v) = \lambda u - \lambda v$.
g) $(\lambda - \mu)u = \lambda u - \mu u$.
h) $0 \cdot u = \lambda \cdot 0 = 0$.
i) $\lambda u = 0 \iff \lambda = 0$ *tai* $u = 0$.
j) $\lambda u = \lambda v$ *ja* $\lambda \neq 0 \implies u = v$.
k) $\lambda u = \mu u$ *ja* $u \neq 0 \implies \lambda = \mu$.

Todistuksesta. Aksioomista asti todistaminen tapahtuu samaan tyyliin kuin edellisen lauseen kohdalla. Sitä mukaa kuin lauseita todistetaan, niitä voi käyttää oletuksina todistettaessa lisää. Esimerkiksi erotuksen olemassolo ja yksikäsitteisyys ovat siis nyt käytettävissä aksioomien lisäksi. Työn säästämisen kannalta on oleellista, missä järjestyksessä vektoriavaruuksien ominaisuudet todistaa. Yllä esitettyä listaa ei ole optimoitu tältä kannalta. Varsinainen urakka jää harjoitustehtäväksi. $\square$

Esittelemme pian lisää vektoriavaruuksissa yleisesti päteviä lauseita. Kaikkia ei kannata luetella missään kirjassa, vaan ideana on, että lähes kaikki $\mathbb{R}^n$:stä tuttu asia toimii myös mielivaltaisessa vektoriavaruudessa ja on parasta oppia tarkastamaan asia tarpeen tullen. Poikkeuksen muodostavat väitteet, jotka liittyvät äärellisulotteisuuteen, siis kantaan. On ehkä syytä miettiä hetken aikaa myös sitä, miten voisi tarvittaessa todistaa, että jokin lause tai kaava **ei** seuraa aksioomista eikä päde jokaisessa vektoriavaruudessa.[44] Tämä tapahtuu esittämällä esimerkki avaruudesta, jossa lause on epätosi. Yksinkertainen esimerkki: Lause, jonka mukaan avaruudella V on 2–alkioinen kanta, on tosi avaruudessa $V = \mathbb{R}^2$, mutta epätosi avaruudessa $V = \mathbb{R}^3$. Kaksialkioisen kannan olemassaoloa ei siis voi todistaa vektoriavaruuden aksioomista lähtemällä, koska se ei seuraa niistä, vaan on olemassa avaruus — nimittäin esimerkiksi $\mathbb{R}^3$ — jossa aksioomat pätevät, mutta kaksialkioista kantaa ei ole.

Aksiomaattisen lähestymistavan etu on, että aksioomista alkaen todistetut lauseet pätevät kaikissa vektoriavaruuksissa. Uudessa avaruudessa jokin $\mathbb{R}^n$:n tuttu ominaisuus saattaa saada aivan uuden sisällön ja kertoa meille jotakin sellaista tutkittavasta avaruudesta, jota ei ilman $\mathbb{R}^n$:n geometriaan vertaamista olisi tullut ajatelleeksi. Kaikissa vektoriavaruuksissa tosien yleisten lauseiden lisäksi on siksi hyvä tuntea runsaasti konkreettisia esimerkkejä erilaisista vektoriavaruuksista.

3.6. Esimerkkejä vektoriavaruuksista ja laskemisesta niissä.

Alun perin tutkimamme avaruudet $\mathbb{R}^n$ ovat kaikki vektoriavaruuksia, samoin niiden aliavaruudet. Itse asiassa jokainen vektoriavaruuden aliavaruus on itsekin vektoriavaruus. Aliavaruudessa käytetään tällöin samoja laskutoimituksia kuin alkuperäisessä avaruudessa. Kaikki aliavaruudet ovat siis samalla esimerkkejä vektoriavaruuksista. Antamamme yleisen vektoriavaruuden aksiomaattinen määritelmä

[44] Ei ole itsestäänselvää. että nämä ovat sama asia.

ei kuitenkaan olisi kovinkaan hyödyllinen, jos ei olisi olemassa aidosti erilaisia vektoriavaruuksia kuin jo tuntemamme $\mathbb{R}^n$ aliavaruuksineen. Asia onkin niin, että matematiikassa ja sen sovellusalueilla käytetään runsasta valikoimaa erityyppisiä vektoriavaruuksia. Seuraavat esimerkit ovat vain pieni näyte.

Esimerkki 3.6.1. Funktioavaruuksia.

a) $\mathcal{F}(\mathbb{R},\mathbb{R}) = \{f \mid f \text{ on funktio } \mathbb{R} \to \mathbb{R}\}$.
b) $\mathcal{P}(\mathbb{R},\mathbb{R}) = \{f \mid f \text{ on polynomi } \mathbb{R} \to \mathbb{R}\}$.
c) $\mathcal{T}_1(\mathbb{R},\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : x \mapsto a\sin x + b\cos x \mid a, b \in \mathbb{R}\}$.
d) $\mathcal{P}_n(\mathbb{R},\mathbb{R}) = \{f \mid f \text{ on asteen } \leq n \text{ polynomi } \mathbb{R} \to \mathbb{R}\}$, missä $n \in \mathbb{N}$.
e) $\mathcal{F}([0,1],\mathbb{R}) = \{f \mid f \text{ on funktio } [0,1] \to \mathbb{R}\}$.
f) $\mathcal{F}(A,\mathbb{R}) = \mathbb{R}^A = \{f \mid f \text{ on funktio } A \to \mathbb{R}\}$, missä A on jokin joukko

Näissä kaikissa on tietysti käytössä tavalliset pisteittäiset laskutoimitukset:

$$(f+g)(x) = f(x) + g(x)$$
$$(\lambda f)(x) = \lambda \cdot f(x).$$

Esimerkki 3.6.2. Jonoavaruuksia.

a) $\mathcal{F}(\mathbb{N},\mathbb{R}) = \mathbb{R}^{\mathbb{N}} = \{f \mid f \text{ on funktio } \mathbb{N} \to \mathbb{R}\} = \{\text{lukujonot}\}$.
b) $\mathbb{R}^{(\mathbb{N})} = \{f \in \mathbb{R}^{\mathbb{N}} \mid f(n) \neq 0 \text{ vain äärellisen monella } n \in \mathbb{N}\} = \{\text{äärelliset lukujonot}\}$.
c) $\{f \in \mathbb{R}^{\mathbb{N}} \mid f(k) = 0 \text{ aina, kun } k > n\}$, missä $n \in \mathbb{N}$. (Olennaisesti $\mathbb{R}^n$)

Esimerkki 3.6.3. Matriisiavaruuksia.

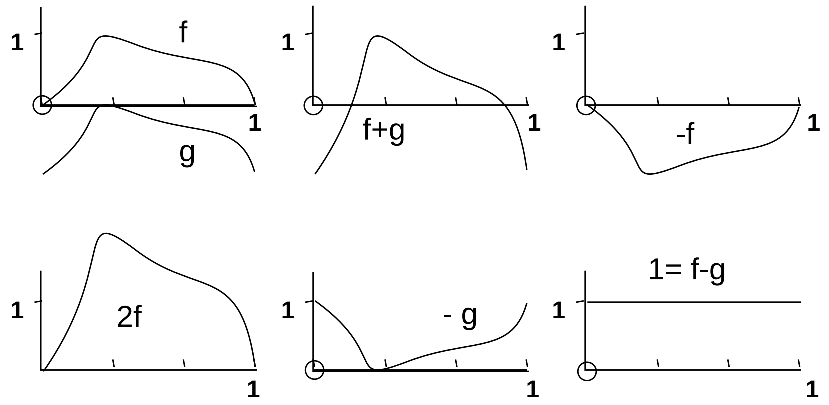
$\mathcal{M}at_{m\times n} = \{f \mid f \text{ on } m \times n\text{–matriisi}\}$. (Olennaisesti $\mathbb{R}^{mn}$)

Kopioimalla lukuja 3.1. – 3.4. soveltuvin osin saamme myös yleisessä vektoriavaruudessa määritelmät mm. aliavaruudelle $H \subset V$ ja äärellisen monen vektorin lineaariselle riippumattomuudelle ja niiden virittämälle aliavaruudelle. Näistä saadaan kannan käsite. Sanotaan, että vektoriavaruus on *n–ulotteinen*, jos sillä on n:stä vektorista muodostuva kanta. Muunkokoisia kantoja ei tällöin avaruudessa V ole. Jos mitään äärellistä kantaa ei löydy, avaruus on *ääretönulotteinen*. Kun kanta on olemassa ja valittu, se tuottaa vektoriavaruuteen *koordinaatit*. Äärellisulotteisen vektoriavaruuden alkioita $v \in V$ voi siis kuvailla n:llä luvulla — koordinaateillaan — mutta riippuu kannasta, millä vektorilla on mitkäkin koordinaatit. Siksi kanta on syytä tarvittaessa merkitä näkyviin. Jos V:n kanta on esimerkiksi $G = (g_1, g_2 \ldots, g_k)$, niin vektorin $x = \sum x_i g_i \in V$ voi tarvittaessa kirjoittaa sarakevektorina

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}_G .$$

Nämä ovat matematiikassa hyvinkin usein esiintyviä käsitteitä, joilla kannattaa opetella laskemaan sujuvasti eri avaruuksissa. Harjoittelemme hieman.

Esimerkki 3.6.4. lineaarisesti riippuvia vektoreita funktioavaruudessa $\mathcal{F}(\mathbb{R},\mathbb{R})$.



Kuvan funktiot f, g ja vakiofunktio 1 ovat lineaarisesti riippuvat koska $f + g = 1$, mutta funktiot f ja g ovat lineaarisesti riippumattomat, koska ehdosta

$$\lambda f + \mu g = 0 \quad (\text{=nollafunktio})$$

seuraa muun muassa, että

$$\lambda \underbrace{f(0)}_{0} + \mu \underbrace{g(0)}_{-1} = 0 \quad (\text{, joten } \mu = 0.)$$

$$\text{ja } \lambda \underbrace{f(\frac{1}{3})}_{1} + \mu \underbrace{g(\frac{1}{3})}_{0} = 0 \quad (\text{, joten myös } \lambda = 0.))$$

Samaan tyyliin voi todistaa, että myös esimerkiksi funktiot $1, x, x^2, \ldots x^{n-1}$ ja x^n ovat lineaarisesti riippumattomia ja muodostavat siis kannan virittämälleen avaruudelle $\mathcal{P}_n(\mathbb{R},\mathbb{R})$. Tuntuisi tietysti järkevältä sanoa, että kaikkien monomien jono $1, x, x^2, \ldots$ on "kanta" kaikenasteisten polynomien avaruudessa $\mathcal{P}(\mathbb{R},\mathbb{R})$. Tämä vaatii lineaarisen riippumattomuuden käsitteen määrittelemistä myös äärettömän monelle vektorille. Teemme sen seuraavassa pykälässä.

Myös funktiot $\sin x$ ja $\cos x$ ovat lineaarisesti riippumattomat. Totta ja tärkeää[45], mutta hieman vaikeampaa todistaa on, että myös funktiot $\sin x$, $\cos x$, $\sin 2x$, $\cos 2x$, $\sin 3x$, $\cos 3x, \ldots, \cos nx$ ovat lineaarisesti riippumattomia, olipa luku n kuinka suuri tahansa. Niiden lineaarikombinaatioita sanotaan *trigonometrisiksi polynomeiksi.*

Esimerkki 3.6.5. Aliavaruuksia vektoriavaruuksissa. Seuraavat ovat toistensa aliavaruuksia:

a) $\mathcal{P}_0(\mathbb{R},\mathbb{R}) \subset \mathcal{P}_1(\mathbb{R},\mathbb{R}) \subset \mathcal{P}_2(\mathbb{R},\mathbb{R}) \subset \cdots \subset \mathcal{P}(\mathbb{R},\mathbb{R}) \subset \mathcal{F}(\mathbb{R},\mathbb{R})$
b) $\mathcal{T}_1(\mathbb{R},\mathbb{R}) \subset \mathcal{F}(\mathbb{R},\mathbb{R})$
c) $\mathbb{R}^{(\mathbb{N})} \subset \mathcal{F}(\mathbb{N},\mathbb{R})$.

[45]Tähän perustuvat Fourier'n sarjat

Huomaa, että esimerkiksi $\mathcal{F}([0,1],\mathbb{R})$ ei tietenkään ole avaruuden $\mathcal{F}(\mathbb{R},\mathbb{R})$ aliavaruus, ei edes sen osajoukko.

Näiden esimerkkien jälkeen täydennämme kantateoriaamme koskemaan myös äärетönulotteisia vektoriavaruuksia.

3.7. Hamelin kanta ääretönulotteisessa vektoriavaruudessa V. Totesimme jo edellisessä kohdassa, että lineaarinen riippuvuus, aliavaruus ja lineaarinen verho voidaan määritellä yleisessä vektoriavaruudessa samalla tavalla kuin aluksi käsitellyssä esimerkkiavaruudessa $\mathbb{R}^n$. Täydellisyyden vuoksi[46] kerrotaan nyt, miten myös äärettömän monen vektorin lineaarinen riippumattomuus ja riippuvuus voidaan määritellä.

MÄÄRITELMÄ 3.7.1. Vektoriavaruuden V osajoukko $S \subset V$ on *vapaa* eli *lineaarisesti riippumaton* — LI — jos mitkä tahansa sen eri alkiot $v_1, v_2, \ldots, v_k$ ovat *lineaarisesti riippumattomat*. Muussa tapauksessa $S \subset V$ on *sidottu* eli *lineaarisesti riippuva*, LD, jolloin siis joistakin joukon S vektoreista voidaan muodostaa nollavektori epätriviaalina lineaarikombinaationa.

Esimerkiksi monomien jono $1, x, x^2, \cdots \subset \mathcal{P}(\mathbb{R},\mathbb{R})$ on tämän määritelmän mukaan lineaarisesti riippumaton eli vapaa.

Äärettömän monen vektorin lineaariseen riippuvuuteen saa tuntuman todistamalla sen perusteella tutunnäköiset väitteet:

LAUSE 3.7.2. *Olkoon V vektoriavaruus ja $S \subset V$.*

a) *Jos jokin vektoreista $v \in S$ on nollavektori, S on LD.*
b) *Joukko S on LD täsmälleen silloin, kun jokin sen alkio*[47] *voidaan lausua muiden lineaarikombinaationa.*
c) *Jos S on LD, niin $S \cup R$ on LD, olipa $R \subset V$ mikä tahansa vektorijoukko avaruudessa V.*

On syytä varmistaa, että emme ole luoneet äärelliselle vektorijoukolle kahta erilaista lineaarisen riippumattomuuden käsitettä. Asian kanssa saa olla tarkkana:

LAUSE 3.7.3. *Olkoon S äärellinen, epätyhjä joukko* **eri** *vektoreita ts. $S = \{v_1, v_2, \ldots, v_n\}$; $v_j \neq v_k$, kun $j \neq k$. Tällöin joukko S on LD tai LI (määritelmä 3.1.3), jos ja vain jos vektorit $v_1, v_2, \ldots, v_n$ ovat LD tai LI (määritelmä 3.1.1)*

Voidaan viimein määritellä äärettömän monesta vektorista muodostuva kanta ja muodostaa ainakin yksi esimerkki sellaisesta. Määritelmän voi kopioida aikaisemmasta, kunhan sallii äärettömän joukon ja muistaa, mitä lineaarinen riippumattomuus tässä yhteydessä merkitsee.

MÄÄRITELMÄ 3.7.4. Vektoriavaruuden V osajoukko $S \subset V$ on V:n *kanta* — tarkemmin sanoen HAMELIN[48] *kanta* — jos

(K-1) S virittää avaruuden V, ts. $< S >= V$

(K-2) S on LI.

[46] Sama selväkielisenä: Luvun loppu ei kuulu kurssiin.
[47] Mutta ei välttämättä jokainen!
[48] GEORG HAMEL (1877–1954) saksalainen matemaatikko ja fyysikko.

Jokainen lineaarisesti riippumaton vektorijoukko S on siis virittämänsä avaruuden eli lineaarisen verhonsa kanta. Perusesimerkki äärettömästä kannasta on avaruuden $\mathcal{P}(\mathbb{R},\mathbb{R})$ standardikanta $(1, x, x^2, \dots)$. Lukija voi testata taitoaan äärettömän kannan käsittelyssä tutkimalla, muodostavatko standardivektorit

$$\begin{aligned} e_1 &= (1, 0, \dots) \\ e_2 &= (0, 1, 0, \dots) \\ e_1 &= (0, 0, 1, 0, \dots) \\ &\vdots \end{aligned}$$

kannan mihinkään jonoavaruuteen.

Ääretönkin kanta S tuottaa virittämäänsä aliavaruuteen $< S >$ koordinaatiston, nimittäin yleistetyt eli *Hamelin koordinaatit* seuraavassa mielessä.

LAUSE 3.7.5. *Jos S on LI, niin jokainen aliavaruuden $< S >$ vektori voidaan lausua tasan yhdellä tavalla äärellisen monen S:n alkion lineaarikombinaationa: $v = \sum_{i=1}^{k} \lambda_i v_i$, missä $\lambda_i \in \mathbb{R}$ ja $v_i \in S$. Yksikäsitteisyys merkitsee, että jos μ_i, $\lambda_i \in \mathbb{R}$ ja $v_i \in S$, niin*

$$\sum_{i=1}^{k} \lambda_i v_i = \sum_{i=1}^{k} \mu_i v_i \iff \lambda_i = \mu_i \text{ kaikilla } i = 1, \dots k.$$

TODISTUS. Jätetään kiinnostuneen lukijan pohdittavaksi.

Seuraavakin lause kopioituu äärellisestä teoriasta, mutta sen todistus ei enää.

LAUSE 3.7.6. (HAMELIN KANTALAUSE). *Jokaisella vektoriavaruudella V on olemassa kanta.*

TODISTUKSESTA. Lause on yhtäpitävä logiikan ja joukko-opin alaan kuuluvan *Zornin lemman* ja myös *valinta-aksiooman* kanssa, joka ei kuulu tähän kirjaan. Lukija voi vakuuttua asian hankaluudesta yrittämällä keksiä kannan vaikkapa funktioavaruuteen $\mathcal{F}(\mathbb{R},\mathbb{R})$ tai jonoavaruuteen $\mathbb{R}^{\mathbb{N}}$. Hamelin kanta[49] voi olla ylinumeroituvasti ääretön.

Hamelin lauseen ihmeellisestä todistuksesta saadaan seuraavat vahvemmat tulokset, jotka itsessäänkin ovat kiinnostavia:

LAUSE 3.7.7. *Olkoon S vektoriavaruuden V lineaarisesti riippumaton osajoukko. Tällöin on olemassa V:n kanta, joka sisältää joukon S. Tämä merkitsee, että mikä tahansa V:n LI osajoukko voidaan laajentaa koko avaruuden V kannaksi.*

LAUSE 3.7.8. *Olkoon S vektoriavaruuden V virittäjäjoukko, ts. $V =< S >$. Tällöin on olemassa V:n kanta, joka sisältyy joukkoon S, ts. V:n mikä tahansa virittävä osajoukko voidaan supistaa avaruuden kannaksi.*

[49] Ääretönulotteisissa avaruuksissa käytetään Hamelin kannan lisäksi muitakin kantakäsitteitä. Yleensä niissä on ideana esittää vektori äärettömän monen kantavektorin ”lineaarikombinaationa”, siis esimerkiksi sarjana.

SEURAUS 3.7.9. *Olkoon S vektoriavaruuden V osajoukko. Tällöin seuraavat ehdot ovat keskenään yhtäpitäviä:*

(1) *S on V:n kanta*
(2) *S on LI, eikä mikään sitä laajempi joukko V:n vektoreita ole LI*
(3) *$< S >= V$, eikä mikään S:n aito osajoukko viritä avaruutta V.*

PERUSTELUISTA. Lause sanoo, että V:n kanta on toisaalta sen maksimaalinen vapaa joukko, toisaalta sen minimaalinen virittäjäjoukko ja että nämä siis ovat sama asia. Äärellisessä tapauksessa tämä on jo todettu kohdissa 3.4.6 ja 3.4.7. Ääretönkin tapaus on todistettavissa ilman esitietoja vaikeahkona harjoitustehtävänä. Yritä!

4. Lineaarikuvaukset

Tähän asti tutkimamme matriisien ja lineaarikuvausten välinen yhteys perustuu standardikannan käyttöön avaruudessa $\mathbb{R}^n$. Kohdassa 0.6. on myös vilahtanut ajatus, että kaikki olisi ehkä paljon helpompaa, jos sallisimme lineaarikuvauksista puhuessamme muitakin kantoja, myös vinokulmaisia. Nyt sellaisia osataan käsitellä, ja on mahdollista esittää teoria kaikkien vektoriavaruuksien välisille lineaarikuvauksille. äärellisulotteisissa avaruuksissa alamme käyttää matriiseja mielivaltaisten kantojen suhteen.

4.1. Kuvaukseen liittyvien peruskäsitteiden kertaus ja täydennys. Kertaamme funktio-opin perusteet, koska niitä tarvitaan nyt.

- *Funktio* eli *kuvaus*
$$f : A \to B$$
liittää lähtö- eli määrittelyjoukon A alkioon eli pisteeseen eli *kohtaan* $x \in A$ tasan yhden alkion maalijoukosta B. Sitä merkitään $f(a) \in B$ ja sanotaan f :n *arvoksi* kohdassa x tai *pisteen* x *kuvaksi.*
- *Lähtöjoukon* A *kuva* eli funktion f *arvojoukko* on joukko $f(A) = \{f(x) \mid x \in A\} = \{y \in B \mid \exists x \in A : y = f(x)\}$.
- Osajoukon $E \subset A$ *kuvajoukko* on $f(E) = \{f(x) \mid x \in E\}$. Käyrän tai pinnan esitys parametrimuodossa on sama asia kuin sen esitys kuvajoukkona.
- *Pisteen* $y \in B$ *alkukuva* on joukko $f^{-1}(\{y\}) = \{x \in A \mid f(x) = y\}$. **Alkukuva on siis yhtälön** $f(x) = y$ **ratkaisujoukko**. Huomaa, että $A = \bigcup_{y \in B} f^{-1}(\{y\})$ ja että eri pisteiden alkukuvat ovat erillisiä eli pistevieraita joukkoja. Pisteen alkukuva voi olla tyhjä joukko. Tällöin vastaavalla yhtälöllä ei ole ratkaisua. Pisteiden alkukuvia sanotaan myös funktion f *tasa-arvojoukoiksi.* Tämä on järkevää erityisesti silloin, kun $B = \mathbb{R}$. Analyyttisestä geometriasta tutut yhtälöt esittävät suoria ja muita käyriä sopivien funktioiden tasa-arvojoukkoina.
- *Osajoukon* $G \subset B$ *alkukuva* on joukko $f^{-1}(G) = \{x \in A \mid f(x) \in G\} = \bigcup_{y \in G} f^{-1}(\{y\})$.
- Funktion $f : A \to B$ *rajoittuma* osajoukkoon $E \subset A$ on funktio
$$f_{|E} : E \to B : x \mapsto f(x).$$
- Funktioista $f : A \to B$ ja $g : B \to C$ *yhdistetty kuvaus* on funktio
$$g \circ f : A \to C : x \mapsto (g \circ f)(x) = g(f(x)).$$
- Joukon A *identtinen kuvaus* on funktio
$$Id = Id_A : A \to A : x \mapsto x.$$
- Kun $A \subset B$, niin joukon A *upotus* eli *inkluusio* joukkoon B on funktio
$$i_{AB} : A \to B : x \mapsto x.$$
Inkluusio i_{AB} on siis identtisen kuvauksen Id_B rajoittuma joukkoon A.

- a) f on *injektio*, jos se kuvaa eri alkiot aina eri alkioiksi: $x \neq y \implies f(x) \neq f(y)$.
  b) f on *surjektio*, jos $f(A) = B$.
  c) f on *bijektio*, jos se on sekä injektio että surjektio. Tällöin ja vain tällöin on olemassa f:n *käänteiskuvaus* $f^{-1} : B \to A$, joka on määritelmän mukaan sellainen funktio, että $f^{-1} \circ f = Id_A$ ja $f \circ f^{-1} = Id_B$, toisin sanoen

$$f(f^{-1}(y)) = y \qquad \forall\, y \in B \text{ ja}$$
$$f^{-1}(f(x)) = x \qquad \forall\, x \in A.$$

- Jos f ja g ovat funktioita $A \to \mathbb{R}$, niin funktio $(f,g) : A \to \mathbb{R}^2 : a \mapsto (f(x), g(x))$ on yksinkertainen esimerkki vektoriarvoisesta kuvauksesta. f ja g ovat sen komponentit standardikannassa. Pisteen alkukuva tällaisen kuvauksen (f,g) suhteen saadaan leikkauksena

$$(f,g)^{-1}(\{(a,b)\}) = f^{-1}(\{a\}) \cap g^{-1}(\{b\}).$$

  Esimerkki tästä on tuttu tapa esittää avaruuden $\mathbb{R}^3$ suora kahden tason leikkauksena eli yhtälöparin avulla.

4.2. Lineaarikuvauksen perusominaisuuksia. Lineaarikuvauksen määritelmä kuuluu niihin asioihin, jotka voi esittää missä tahansa vektoriavaruudessa, sanalla sanoen varsinaiseen lineaarialgebraan:

MÄÄRITELMÄ 4.2.1. Olkoot V ja W vektoriavaruuksia.
Kuvaus $L : V \to W$ on *lineaarikuvaus*, jos se toteuttaa ehdot:
(L-1) $L(u+v) = Lu + Lv$ kaikille $u, v \in V$,
(L-2) $L(\lambda u) = \lambda Lu$ kaikille $\lambda \in \mathbb{R}$ ja $u \in V$.

Yleisen tavan mukaisesti on sulut jätetty lineaarikuvauksen muuttujan ympäriltä merkitsemättä, kun niitä ei selvyyssyistä tarvita.

ESIMERKKI 4.2.2. Kuvaus, joka funktioon $f : \mathbb{R} \to \mathbb{R}$ liittää sen arvon pisteessä 3 — erikoistapaus *evaluaatiokuvauksesta* — on lineaarikuvaus funktioavaruudelta luvuille:

$$L_3 : \mathcal{F}(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$$
$$f \mapsto L_3(f) = f(3).$$

LAUSE 4.2.3. *Lineaarikuvaus kuvaa aina origon eli nollavektorin origoksi. Lisäksi jokaiselle lineaarikuvaukselle $L : V \to W$ on voimassa:*

$$L\left(\sum_{k=1}^{n} \lambda_k u_k\right) = \sum_{k=1}^{n} \lambda_k L u_k.$$

Olemme nollaluvussa käsitelleet lineaarikuvausten ja matriisien välistä yhteyttä käyttäen standardikantaa avaruudessa $\mathbb{R}^n$. Emme todistaneet kaikkia väitteitämme. Nyt tutkimme lineaarikuvauksia ensin yleisesti ilman koordinaatteja. Seuraavassa pykälässä tarkastelemme lineaarikuvauksia äärellisulotteisessa avaruudessa matriisein käyttäen apuna kantaa.

Tarkoituksena on taas miettiä, milloin ja miten yhtälö

$$Lx = b$$

voidaan ratkaista. Pitää siis tutkia lineaarikuvauksen kuvajoukkoa (surjektiivisuus) ja pisteiden alkukuvia (injektiivisyys) ja erityisesti selvittää, milloin lineaarikuvaus on bijektio ja millainen lineaarisen bijektion käänteiskuvaus on. Äärellisulotteisissa avaruuksissa lineaarikuvausten yhdistäminen ja kääntäminen on suhteellisen helppoa, koska siihen voi käyttää matriisien laskutoimituksia. Olennaista tässä on yhdistetyn kuvauksen ja käänteiskuvauksen lineaarisuus. Mahdollisesti ääretönulotteisessa tapauksessa matriisit eivät ole käytettävissä, mutta lineaarisuus sentään on:

LAUSE 4.2.4. *Olkoot $L : V \to W$ ja $L' : W \to U$ lineaarikuvauksia. Tällöin yhdistetty kuvaus $L' \circ L : V \to U$ on lineaarikuvaus.*

TODISTUS. Helppo juttu! □

LAUSE 4.2.5. *Bijektiivisen lineaarikuvauksen $L : V \to W$ käänteiskuvaus $L^{-1} : W \to V$ on lineaarikuvaus.*

TODISTUS. Avaruuden $\mathbb{R}^2$ tapauksessa asia onkin jo tuttu: käänteiskuvausta vastaa käänteismatriisi. Lineaarisen bijektion käänteiskuvauksen lineaarisuus on kuitenkin yleinen tosiasia. Sen voi päätellä esimerkiksi näin: Olkoot w ja $w' \in W$ ja $v = L^{-1}(w)$ ja $v' = L^{-1}(w')$. Oletimme, että L on lineaarinen. Ehto

$$L(v + v') = Lv + L(v')$$

on toisin merkiten

$$L(L^{-1}w + L^{-1}(w')) = w + w'.$$

Sovelletaan molempiin puoliin kuvausta L^{-1} ja saadaan

$$L^{-1}w + L^{-1}(w') = L^{-1}(w + w').$$

Kaava $\lambda L^{-1}w = L^{-1}(\lambda w)$ todistetaan vastaavasti. □

Lineaarikuvaukseen liittyy luonnollisella tavalla kaksi aliavaruutta, ydin ja arvojoukko eli kuva-avaruus.

MÄÄRITELMÄ 4.2.6. a) Lineaarikuvauksen $L : V \to W$ *ydin*[50] $N(L)$ on *homogeenisen yhtälön* $Lx = 0$ ratkaisujen joukko, siis

$$N(L) = L^{-1}(\{0\}) = \{u \in V \mid Lu = 0\}.$$

b) Lineaarikuvauksen $L : V \to W$ arvojoukkoa sanotaan myös sen *kuva-avaruudeksi.*[51] $L(V)$.

[50]Kuvauksen L ydintä $N(L)$ merkitään toisinaan germaanisperäisellä lyhenteellä $\mathrm{Ker}(L)$.

[51]Kuvauksen L arvojoukkoa $L(V)$ merkitään toisinaan germaanisperäisellä lyhenteellä $\mathrm{R}(L)$ tai $\mathrm{W}(L)$.

LAUSE 4.2.7. *Olkoon $L : V \to W$ lineaarikuvaus.*

a) *Aliavaruuden $H \subset V$ kuvajoukko on aliavaruus $L(H) \subset W$.*
b) *Erityisesti koko arvojoukko $L(V)$ on aliavaruus.*
c) *Aliavaruuden $K \subset W$ alkukuvajoukko on aliavaruus $L^{-1}(K) \subset V$.*
d) *Erityisesti ydin $N(L) = L^{-1}(0)$ on aliavaruus.*

TODISTUS. a) Olkoot w ja $w' \in L(H)$ sekä $\lambda \in \mathbb{R}$. Kuvajoukkoon kuuluminen merkitsee, että on olemassa vektorit v ja $v' \in H$, joille $w = Lv$ ja $w' = L(v')$. Lineaarisuuden ehdon (L-1) ja aliavaruuden määritelmän ehdon (A-1) nojalla on

$$w + w' = Lv + L(v') = L(v + v') \in L(H).$$

Vastaavalla tavalla todistetaan, että $\lambda w \in L(H)$, kun $\lambda \in \mathbb{R}$ ja $w \in L(H)$.
c) Olkoot u ja $v \in L^{-1}(K)$ sekä $\lambda \in \mathbb{R}$. Alkukuvajoukkoon kuuluminen merkitsee, että Lu ja $Lv \in K$. Lineaarisuuden ja aliavaruuden määritelmistä saadaan taas:

$$L(u + v) = Lu + Lv \in K.$$

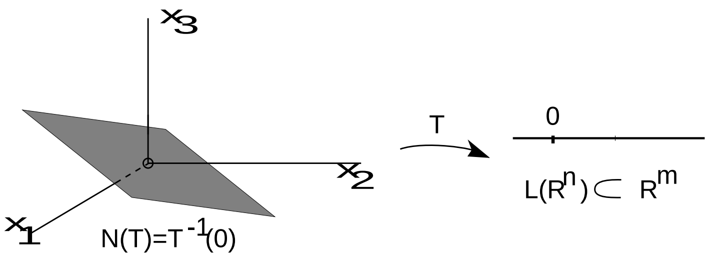
Ehto (A-2) tarkastetaan taas vastaavasti. $\square$

Edellisestä lauseesta seuraa erityisesti, että lineaarikuvauksen ydin ja kuvajoukko sisältävät asianomaisten avaruuksien nollavektorit eli origot, minkä jo tiesimmekin, onhan $L0 = 0$.

ESIMERKKI 4.2.8. a) Origon kautta kulkeva taso tavallisessa yhtälömuodossaan, vaikkapa

$$T = \{(x, y, z) \in \mathbb{R}^3 \mid 4x + 5y - 456z = 0\}$$

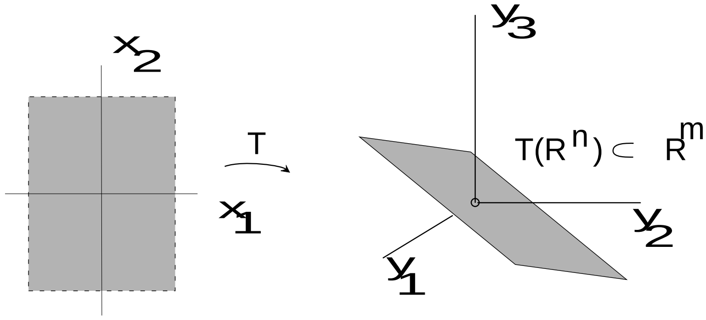
on perusesimerkki lineaarikuvauksen ytimestä.



b) Origon kautta kulkevan tason parametriesitys, vaikkapa

$$T' = \{(s + 2t, 3s + 4t, 9s + 99t) \mid (s, t) \in \mathbb{R}^2\}$$

on perusesimerkki lineaarikuvauksen kuvajoukosta. Mitkä ovat ao. lineaarikuvaukset? Kumpikin taso on perusesimerkki aliavaruudesta. Pieni perustehtävä: Määrää kummallekin kanta.



Lineaarikuvaus $L : V \to W$ on tietysti surjektio, jos sen arvojoukko on koko maaliavaruus eli $\dim L(V) = \dim W$. Kiinnostavampaa on, että injektiivisyyden voi puolestaan tarkastaa ytimen dimension avulla:

LAUSE 4.2.9. *Lineaarikuvaus $L : V \to W$ on injektio, jos ja vain jos sen ydin on nollaulotteinen eli triviaali, ts. $N(L) = \{0\}$.*

TODISTUS. On tietysti selvää, että origo kuuluu aina ytimeen ja että ydin ei voi sisältää muita pisteitä, mikäli L on injektio. Todistettavaksi jää lauseen mielenkiintoisempi puoli, siis että injektiivisyydelle riittää, että origolla on vain yksi alkukuvapiste.

Olkoon siis $N(L) = \{0\}$ ja olkoot u ja $v \in V$ siten, että $Lu = Lv$. On osoitettava, että $u = v$. Väite merkitsee samaa kuin $u - v = 0$, eli $(u - v) \in \{0\}$. Oletuksen mukaan $\{0\} = N(L)$, joten riittää todistaa, että $L(u - v) = 0$. Tämä puolestaan onkin totta, sillä oletuksen mukaan $L(u) - L(v) = 0$ ja L on lineaarinen. $\square$

Injektio ei voi kuvata kahta vektoria samaksi. Itse asiassa lineaarinen injektio ei voi edes kuvata kahta erisuuntaista vektoria samansuuntaisiksi eikä yleensäkään riippumattomia vektoreita riippuviksi. Tämän takaa seuraava lause:

LAUSE 4.2.10. *Olkoon $L : V \to W$ lineaarinen injektio, $S \subset V$ ja S lineaarisesti riippumaton joukko. Tällöin myös kuvajoukko $L(S)$ on lineaarisesti riippumaton.*

TODISTUS. Kopioimme todistusidean edellisestä lauseesta. On näytettävä, että ehdoista $\sum_{j=1}^{k} \lambda_j w_j = 0$, $w_j \in L(S)$ ja $\lambda_j \in \mathbb{R}$ seuraa, että jokainen λ_j on 0. Kuvajoukon määritelmän mukaan jokainen $w_j \in L(S)$ on muotoa $L(v_j)$, missä $v_j \in S$. Siis $\sum_{j=1}^{k} \lambda_j L(v_j) = 0$, eli $L(\sum_{j=1}^{k} \lambda_j v_j) = 0$. Lineaarikombinaatio $\sum_{j=1}^{k} \lambda_j v_j$ kuuluu siis L:n ytimeen, joka on injektiivisyyden takia pelkkä $\{0\}$. Tässä on siis lineaarisesti riippumattomista vektoreista tehty lineaarikombinaationa nollavektori. Kertoimet ovat näin ollen nollia. $\square$

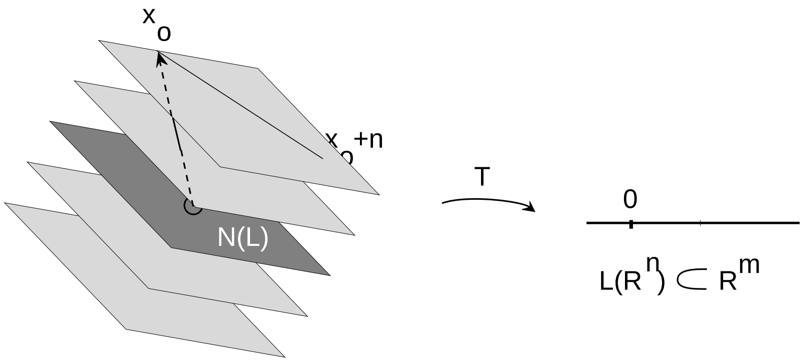
Samaan tapaan todistetaan myös seuraava lineaaristen yhtälöryhmien ja myös differenssi- ja differentiaaliyhtälöiden ratkaisemisessa hyödyllinen lause, joka kertoo, että ytimen tunteminen auttaa yleisen lineaariyhtälön ratkaisemisessa.

LAUSE 4.2.11. *Olkoon $L : V \to W$ lineaarikuvaus ja $b \in W$ sekä x_0 yhtälön*

$$(*) \qquad Lx = b$$

ratkaisu, ts. $Lx_0 = b$. *Tällöin yhtälön yleinen ratkaisu, siis kaikkien ratkaisujen joukko eli alkukuva* $L^{-1}(\{b\})$ *on*

$$x_0 + N(L) = \{x_0 + n \mid n \in N(L)\}.$$



Voimme todeta, että lause 4.2.9. tietysti seuraa lauseesta 4.2.11. Itse asiassa huomaamme enemmänkin: lineaarikuvaus on injektio, jos edes yhden pisteen alkukuva on yksiö.

4.3. Lineaarikuvaukset ja kannat.

Äärellisulotteisten vektoriavaruuksien välisiä lineaarikuvauksia voi käsitellä tuttuun tapaan matriisien avulla kunhan varustaa avaruudet kannalla. Kuvauksen matriisi riippuu tällöin myös kantojen valinnasta. Kaiken takana on edelleen tieto, että lineaarikuvaus määräytyy täysin kantavektoreiden kuvista, jotka puolestaan voivat olla mitä tahansa vektoreita.

Lause 4.3.1. *Olkoot* V *ja* W *vektoriavaruuksia,* $(e_1, e_2, \ldots, e_n)$ *avaruuden* V *kanta sekä* $v_1, v_2, \ldots, v_n$ *avaruuden* W *vektoreita. Silloin on olemassa täsmälleen yksi lineaarikuvaus* $L : V \to W$ *siten, että*

$$Le_k = v_k \quad \textit{kaikilla } k = 1, 2, \ldots, n,$$

nimittäin

$$L\left(\sum_{k=1}^{n} \lambda_k e_k\right) = \sum_{k=1}^{n} \lambda_k v_k.$$

Edellisessä lauseessa ei ole paljon todistamista eikä mielenkiintoa. Asia käy tutun näköiseksi, kun vektorit $v_1, v_2, \ldots, v_n$ lausutaan jossakin maalipuolen kannassa.

Lause 4.3.2. *Olkoot* V *ja* W *vektoriavaruuksia,* $\dim V = n$, $\dim W = m$, $E = (e_1, e_2, \ldots, e_n)$ *avaruuden* V*:n kanta ja* $E' = (e'_1, e'_2, \ldots, e'_m)$ *avaruuden* W *kanta. Tällöin lineaarikuvaukset* $L : V \to W$ *ja matriisit* $A_{m\times n} = [a_{ij}]$ *vastaavat kääntäen yksikäsitteisesti toisiaan siten, että*

$$\text{(M-1)} \qquad Le_j = \sum_{i=1}^{m} a_{ij} e'_i; \quad j = 1, 2, \ldots, n.$$

Vektorin $u = \sum_{k=1}^{n} \lambda_k e_k$ *kuva on*

$$Lu = L\left(\sum_{k=1}^{n} \lambda_k e_k\right) = \sum_{i=1}^{m} \mu_i e'_i, \ \ \textit{missä}\ \mu_i = \sum_{j=1}^{n} a_{ij}\lambda_j.$$

Tämä merkitsee, että erityisesti E–kantavektoreiden kuvien koordinaatit E'–koordinaatistossa ovat lineaarikuvausta L vastaavan matriisin $Mat(L) = (a_{ij})$ eli täydellisemmin merkiten $Mat(L; E, E')$ sarakkeet a_{*j}. Kannassa E lausutusta vektorista

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_E \in V$$

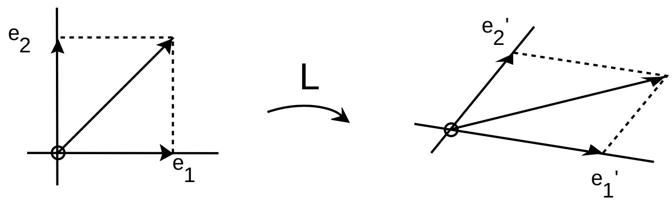
saadaan kannassa E' lausuttu vektori

$$Lx = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}_{E'} \in W$$

tavalliseen tapaan matriisilla $Mat(L; E, E')$ kertomalla. Pystymme näin lausumaan vektoriavaruuksien välisen lineaarikuvauksen matriisin avulla, kunhan lähtö- ja maaliavaruudessa on äärellinen kanta.

Jos L on lineaarikuvaus joltakin vektoriavaruudelta V itselleen, ja jos V:ssä on käytössä vain yksi kanta E, niin kuvauksen $L : V \to V$ matriisille käytetään toisinaan lyhennettyä merkintää $Mat(L; E)$.

Esimerkki 4.3.3. Kiinnostava on mm. tilanne, jossa tarkastellaan yhdessä avaruudessa toimivan kuvauksen $L : V \to V$ matriisia kahden eri kannan suhteen. Asian hyödyllisyyden arvaa esimerkiksi tarkastelemalla mielivaltaista lineaarista bijektiota eli *lineaari-isomorfismia* $L : V \to V$. Valitaanpa V:lle ensimmäiseksi kannaksi $E = \{e_1, e_2, \ldots, e_n\}$ mikä tahansa kanta ja toiseksi kannaksi E' alkuperäisten kantavektoreiden kuvat $E' = \{e'_1, e'_2, \ldots, e'_n\} = \{Le_1, Le_2, \ldots, Le_n\}$, jotka todella ovat riippumattomia, siis kanta. Näissä kannoissa lineaari-isomorfismin L matriisiksi $Mat(L, E, E')$ tulee tietysti ykkösmatriisi. Mahdolliset vaikeudet voi näin toisinaan siirtää matriisista kannan valintaan.



Ajatusta voi jatkaa: Jos kerran ykkösmatriisi ei saman avaruuden V kahdessa eri kannassa E ja E' kuvaakaan identtistä kuvausta, vaan kuvausta, joka vie toisen kannan kantavektorit toisen kannan kantavektoreiksi, niin kiinnostaa tietää, mikä on identtisen kuvauksen $Id : E \to E$ matriisi $Mat(Id, E, E')$ kahden eri kannan suhteen. Vastaus on helppo keksiä muistamalla, mitä on määritelty. Minkä tahansa lineaarikuvauksen matriisin sarakkeet ovat lähtöavaruuden kantavektoreiden kuvat — tässä siis ne itse — lausuttuna maaliavaruuden kannassa. Etsitty matriisi on

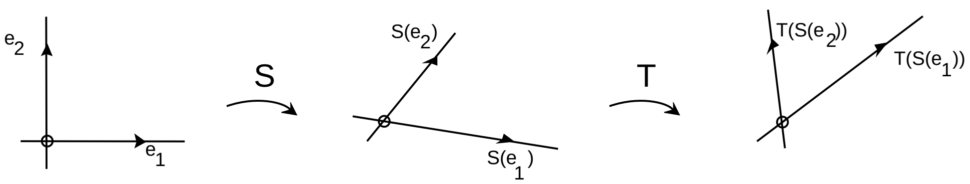
$$Mat(Id, E, E') = ((e_1)_{E'}, (e_2)_{E'}, \ldots, (e_n)_{E'}) .$$

Matriisien käytön tärkein puoli on, että matriisitulo esittää kuvausten yhdistämistä. Näin on myös mielivaltaisia kantoja käytettäessä. Matriisitulo on yhdistetyn kuvauksen matriisi, kunhan kannat huomioidaan.

LAUSE 4.3.4. *Olkoot V, V' ja V'' äärellisulotteisia vektoriavaruuksia, niissä kannat E, E' ja E'' sekä $S : V \to V'$ ja $T : V' \to V''$ lineaarikuvauksia. Tällöin*

$$Mat(T \circ S; E, E'') = Mat(T; E', E'')Mat(S; E, E').$$

TODISTUS. Tarkkaan ottaen emme ole todistaneet tätä lausetta vielä myöskään standardikannan erikoistapauksessa, joten emme voi vedota analogiaan. Koska lauseen väite on pelkkä laskukaava, asian voi selvittää suoralla laskulla. Vähemmällä pääsee ja enemmän ymmärtää päättelemällä seuraavasti: Koska tiedämme, että lineaarikuvauksista yhdistetty kuvaus on lineaarinen, on etsitty matriisi olemassa ja muodostuu kantavektoreiden kuvista $(T \circ S)e_j$. Riittää siis osoittaa, että V:n kantavektorin $e_j \in E$ kuva $(T \circ S)(e_j)$ on V'':n koordinaateissa sama kuin tulomatriisin $Mat(T; E', E'')Mat(S; E, E')$ sarake numero j. Tiedämme, että $S(e_j)$ on V':n koordinaateissa sama kuin $Mat(S; E, E')$:n sarake numero j. Sen kuva $T(S(e_j))$ lausuttuna kannassa E'' saadaan kertomalla tämä sarake matriisilla $Mat(T; E', E'')$. Väitämme, että tämä on sama kuin matriisin $Mat(T; E', E'')Mat(S; E, E')$ sarake numero j. Näin onkin, sillä matriisien AB tulon j:s sarake on matriisitulon määritelmän mukaan tulo matriisista A ja B:n sarakkeesta numero j. $\square$



Käänteiskuvauksen ja käänteismatriisin yhteys seuraa yhdistämistä koskevasta lauseesta ja tiedosta, että lineaarikuvauksen käänteinen on lineaarikuvaus. Pätee siis:

SEURAUS 4.3.5. *Olkoot V ja W äärellisulotteisia vektoriavaruuksia ja E ja F niiden kantoja sekä $L : V \to W$ lineaarikuvaus. Kuvaus L on bijektio jos ja vain jos $Mat(L; E, F)$ on kääntyvä. Tällöin*

$$Mat\left(L^{-1}; F, E\right) = [Mat(L; E, F)]^{-1} .$$

4.4. Dimensiolause ja Gaussin ja Jordanin menetelmä.

Matriisien avulla hallitaan myös monia lineaarikuvauksen geometriaan liittyviä asioita. Esimerkiksi seuraava on luonnollisesti totta:

LAUSE 4.4.1. *Olkoon $L : V \to W$ lineaarikuvaus ja $\{e_1, e_2, \ldots, e_n\}$ V:n kanta. Tällöin arvojoukko $R(L) =< Le_1, Le_2, \ldots, Le_n >$.*

Lineaarikuvauksen arvojoukko on siis helppo määrätä siinä mielessä, että sille on jo matriisin sarakkeina annettu virittäjäjoukko. Tämä ei kuitenkaan kerro kovin paljoa kuvan luonteesta. Ennen kaikkea kuvajoukon dimensio jää hämärän peittoon, voivathan matriisin sarakkeet olla mitä tahansa, erityisesti siis lineaarisesti riippuvia, vaikkapa kaikki samassa tasossa. Kysymys kuuluu, miten kuva-avaruudelle löytyisi kanta. Luonnollista on jättää sarakkeista turhat pois siten, että jäljelle jääneistä tulee kanta. Lauseen 3.4.7 mukaan tämä on mahdollista ja voisi tapahtua esimerkiksi sellaisella työläällä tavalla, että kustakin sarakkeesta vuorollaan tarkastetaan, onko se muiden lineaarikombinaatio. Kun sellainen löytyy, se poistetaan ja seuraavasta testataan, onko se jäljelle jääneiden kombinaatio. Tätä jatketaan, kunnes mikään sarake ei ole muiden lineaarikombinaatio.

Dimensiolause kertoo meille tavan määrätä kuva–avaruuden dimension tutkimalla ydintä, siis esimerkiksi Gaussin ja Jordanin menettelyllä, jolla loppujen lopuksi saadaan kanta sekä ytimelle että kuva-avaruudelle.

LAUSE 4.4.2. DIMENSIOLAUSE. *Olkoot V ja W vektoriavaruuksia, V äärellisulotteinen ja $L : V \to W$ lineaarikuvaus. Tällöin*

$$\dim V = \dim N(L) + \dim L(V).$$

DIMENSIOLAUSEEN KOORDINAATTIVAPAA TODISTUS.

Merkitään tutkittavia dimensioita lyhyesti[52]

$$n = \dim N(L) \quad \text{ja} \quad r = \dim L(V).$$

Tehtävänä on osoittaa, että V on $n + r-$ulotteinen. Tämä tapahtuu tietystikin konstruoimalla siihen $n + r-$alkioinen kanta. Käytettävissä on $N(L)$:n kanta — olkoon se $(e_1, e_2, \ldots, e_n)$ — ja $L(V)$:n kanta — olkoon se $(f_1, f_2, \ldots, f_r)$. Yhteensä näissä on oikea määrä vektoreita, mutta ne eivät kelpaa kannaksi jo siitäkään syystä, että f_j:llä merkityt eivät ole avaruudessa V. Ensin mainitut kuitenkin ovat oikeassa avaruudessa, onhan $N(L) \subset V$. Valitaan ainakin ne kantavektoriehdokkaiksi. Kuva-avaruuden kantavektoreista saadaan lisää kantavektoriehdokkaita "siirtämällä ne avaruuteen V jollakin luonnollisella tavalla." Luonnollista lienee valita V:n loput kantavektoriehdokkaat $e_{n+1}, \ldots, e_{n+r}$ siten, että

$$L(e_{n+j}) = f_j \quad \forall j = 1, \ldots r.$$

Kantaehdokas $(e_1, \ldots, e_{n+r})$ on valittu. Todistetaan kantavektoriehdokkaiden olevan lineaarisesti riippumattomia ja virittävän koko avaruuden V. Riippumattomuuden tarkastamiseksi oletetaan, että

$$(*) \qquad \sum_{i=1}^{n+r} \lambda_i e_i = 0$$

[52] n viittaa nolla-avaruuteen $N(L)$, kun taas r viittaa kuva-avaruuden dimensioon, kuvauksen ja matriisin *rankiin.*

ja todistetaan, että jokainen λ_i on 0. Oletukset saa parhaiten käyttöön siirtymällä kuvauksen L avulla takaisin avaruuteen W:

$$L(\sum_{i=1}^{n+r} \lambda_i e_i) = L(0) = 0.$$

Vasen puoli on lineaarisuuden nojalla

$$\sum_{i=1}^{n+r} \lambda_i L(e_i).$$

Ytimeen kuulumisen takia $L(e_i) = 0$, kun $i = 1, \dots, n$. Siksi vasemmalla puolella vain summan r jälkimmäistä termiä eroavat nollasta. Muut $L(e_i)$:t ovat W:n kantavektorit f_j, ja vasen puoli saa siis muodon

$$\sum_{i=1}^{r} \lambda_{n+i} f_i = 0,$$

josta tietysti seuraa $\lambda_{n+i} = 0$ kaikille $i = 1, \dots r$. Kun tämä tiedetään, jää oletuksestamme $(*)$ vain

$$\sum_{i=1}^{n} \lambda_i e_i = 0,$$

josta seuraa, että myös ensimmäiset λ_i:t ovat nollia. Riippumattomuus on todettu.

Virittämisen tarkastamiseksi valitaan mielivaltainen vektori $v \in V$ ja koetetaan esittää se lineaarikombinaationa kantavektoriehdokkaista, siis muodossa

$$\sum_{i=1}^{n+r} \mu_i e_i = v.$$

Kertoimien arvaamista helpottaa taas kuvaaminen avaruuteen W. Tutkittavan vektorin kuva $L(v)$ voidaan lausua kuvan kannassa:

$$L(v) = \sum_{i=1}^{m} \mu_{n+i} L e_{n+i}.$$

Voisi arvata, että ehkä v on

$$\sum_{i=1}^{r} \mu_{n+i} e_{n+i},$$

mutta tästä puuttuvat ilmeisesti ensimmäiset koordinaatit. Ne keksitään, kun huomataan, että

$$(**) \qquad v - \sum_{i=1}^{r} \mu_{n+i} e_{n+i} \in N(L),$$

onhan

$$L(v - \sum_{i=1}^{r} \mu_{n+i} e_{n+i}) = L(v) - \sum_{i=1}^{r} \mu_{n+i} L e_{n+i} = L(v) - L(v) = 0.$$

Erotus kuuluu siis ytimeen ja voidaan näin ollen lausua ytimen kannan avulla:

$$v - \sum_{i=1}^{r} \mu_{n+i} e_{n+i} = \sum_{i=1}^{n} \mu_i e_i,$$

joten

$$v = \sum_{i=1}^{n} \mu_i e_i + \sum_{i=1}^{r} \mu_{n+i} e_{n+i} = \sum_{i=1}^{n+r} \mu_i e_i.$$

□

Dimensiolauseen matriisitodistus. Tässä oletetaan, että myös kuva-avaruus W on äärellisulotteinen, jolloin käytettävissä on matriisi. Mikäli Gaussin ja Jordanin menettelyssä yhtälöryhmän ratkaisemiseksi sallitaan myös sarakkeita vaihdeltavan — mikä merkitsee vain koordinaattiakselien uudelleen numerointia kuvapuolella — niin voidaan aina päästä lopulta *puolisuunnikastyyppiseen* matriisiin

$$\begin{bmatrix} 1 & \dots & 0 & * & \dots & * \\ 0 & \dots & 0 & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & * & \dots & * \\ 0 & \dots & 0 & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & * & \dots & * \end{bmatrix}$$

On ilmeistä, että puolisuunnikastyyppiselle matriisille — oikeastaan vastaavalle lineaarikuvaukselle — dimensiolause pätee. Riittää siis todistaa, etteivät matriisin ytimen ja kuva-avaruuden dimensio muutu Gaussin ja Jordanin sievennysprosessissa; sarakkeiden vaihtohan ei missään tapauksessa vaikuta dimensioihin. Ytimen osalta tämä on tietysti selvää, koska ydin on yhtälön $Ax = 0$ ratkaisujen joukko, joka ei muutu Gaussin–Jordanin rivioperaatioissa.[53]

Matriisiin A kohdistetut rivioperaatiot $P_{ij}, M_i(\lambda)$ ja $A_{ij}(1)$ muuttelevat kyllä kuvauksen $x \mapsto Ax$ kuvajoukkoa, mutta eivät sen dimensiota. Tämä johtuu siitä, että jos jokin A:n sarake on muiden sarakkeiden lineaarikombinaatio joillakin kertoimilla λ_i, ja A:han kohdistetaan jokin em. operaatioista, niin näin syntyneessä matriisissakin vastaava sarake on lineaarikombinaatio muista samoin kertoimin. □

Sivutuotteena saadaan, että itse asiassa Gauss–Jordan antaa tavan löytää kanta avaruudelle $R(A)$. Muokatun matriisin (älä vaihda sarakkeita) kuva-avaruuden kanta näkyy helposti. Alkuperäisen matriisin kuva-avaruuden kantavektoreiksi käyvät samoista kohdista valitut sarakkeet.

Myös seuraava yllättävä lause on tullut todistetuksi:

[53]Sitä vartenhan ne alunperin keksittiin.

LAUSE 4.4.3. *Matriisin A riviviektoreiden ja saman matriisin sarakevektoreiden virittämällä aliavaruudella on sama dimensio, matriisin ranki*[54] *rank(*A*).*

TODISTUS. Lause pätee selvästikin puolisuunnikastyyppiselle, loppuun muokatulle matriisille. Sarakkeiden virittämän kuva-avaruuden dimensio ei muutu riviooperaatioissa, mutta myöskään rivien virittämän avaruuden dimensio ei tietenkään muutu rivejä vaihdettaessa, vakiolla kerrottaessa, toisiinsa lisättäessä tai sarakkeita vaihdettaessa. □

Dimensiolauseesta saadaan eräissä erikoistapauksissa hyödyllisiä siistinnäköisiä seurauksia

SEURAUS 4.4.4. *Dimensiolauseen olettamusten vallitessa on voimassa*

$$L \text{ on injektio} \iff \dim V = \dim L(V).$$

Jos myös dim(V*)=dim(*W*), niin L on bijektio.*

SEURAUS 4.4.5. *Olkoot vektoriavaruudet V ja W äärellisulotteisia ja* $\dim V = \dim W$*, sekä $L : V \to W$ lineaarikuvaus. Tällöin*

$$L \text{ on bijektio} \iff L \text{ on injektio} \iff L \text{ on surjektio.}$$

Kannattaa muuten pohtia, mitä dimensiolause sanoo, kun toinen avaruuksista on yksiulotteinen.

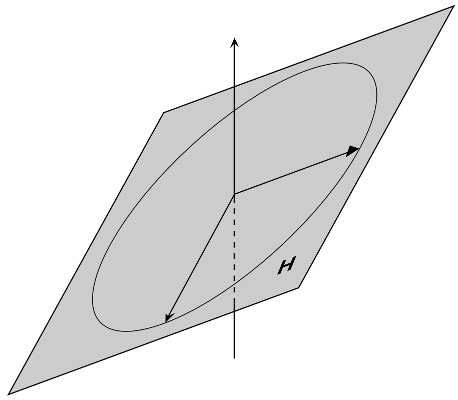
4.5. Kannan vaihtamisen yleinen teoria.

Äärellisulotteisessa vektoriavaruudessa vektorit, lineaarikuvaukset ja myös aliavaruudet ja muutkin osajoukot ilmaistaan usein koordinaattien avulla, ja nämä riippuvat käytetystä kannasta. Kantaa saattaa olla aihetta vaihtaa eri syistä. Jokin lasku saattaa olla paljon helpompi sopivissa $\mathbb{R}^n$:n vinokulmaisissa koordinaateissa tai sitten tilanteen geometria houkuttelee sopivaan kannan valintaan. Esimerkiksi avaruuden $\mathbb{R}^3$ tasossa $x + y - z = 0$ olevan origokeskisen ympyrän yhtälöpari (mikä se lienee?) saadaan standardimuotoon

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

valitsemalla avaruuteen R^3 sopiva kanta: kaksi kohtisuoraa, säteen mittaista vektoria tasoon ja kolmanneksi mikä tahansa niistä riippumaton. Yhtälöpari standardikannassa pitäisi mielellään voida laskea tästä.

[54]Ranki–sanan hyvän suomennoksen keksijälle lupaan pienen palkinnon. Rankki on jotakin muuta.



Merkintöjä 4.5.1. Tällaisiin kannanvaihtoihin tarvittava teoria on kehitelty edellä. Katsokaamme, mihin se kelpaa. Olkoot sitä varten seuraavassa koko ajan $E = (e_1, e_2, \ldots, e_k)$ ja $F = (f_1, f_2, \ldots, f_k)$ kaksi kantaa vektoriavaruudessa V. Itse kantojen välisestä yhteydestä pitää tietysti tietää jotakin, jotta eri objektien kannanvaihtoja voisi tehdä. Siksi oletamme, että toisen kannan — olkoon se F — kantavektorit on annettu toisesssa kannassa — siis kannassa E:

$$f_j = \sum_{i=1}^{k} c_{ij} e_i = \begin{bmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{kj} \end{bmatrix}_E .$$

Olemme jo esimerkin 4.3.3 yhteydessä kinnittäneet huomiota siihen, että edellä sanottu merkitsee samaa kuin

$$C = (c_{ij}) = Mat(Id; F, E),$$

missä on huomattava suunta, ovathan kantaa F kuvaavan matriisin $C = (c_{ij})$ sarakeet nimenomaan kantavektorit f_i. Koska identtinen kuvaus on itsensä käänteiskuvaus, on selvää, että

$$C^{-1} = Mat(Id; E, F),$$

joka siis ilmaisee e_i– kantavektorit kannassa F. Olisimme siis matriisin kääntämiseen tarvittavaa vaivaa lukuunottamatta voineet aloittaa tarkastelut olettamalla, että tunnemme kannan E vektoreiden koordinaatit kannassa F. Huomaamme pian, että eri objektien kannanvaihtoon tarvitaan toisinaan matriisia C, toisinaan matriisia C^{-1} ja usein molempia.

Lause 4.5.2. Kannanvaihto vektorille. *Olkoon*

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}_E \in V$$

jokin vektori lausuttuna kannassa E. Tällöin sama vektori lausuttuna kannassa F on

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix}_F = C^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}.$$

Vastaavasti $E-$ koordinaatit saadaan $F-$koordinaateista kertomalla matriisilla C.

Lauseen tuloksen muistaa helposti suuntaa vaille. Itse asiassa suunnan arvaa helposti väärin. Varmempaa onkin toistaa tarvittaessa seuraava lyhyt todistus:

TODISTUS. Merkitään $x_E = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$ ja $x_F = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix}$.

$$\begin{aligned} & x = Id(x), \\ \text{joten} \quad & x_F = Mat(Id; E, F)x_E = C^{-1}x_E. \end{aligned}$$

□

Seuraavaksi vaihdamme kantaa lineaarikuvaukselle $L : V \to W$. Sitä varten tarvitaan kaksi kantaa lähtöavaruuteen — olkoot ne E ja G — ja kaksi kantaa maaliavaruuteen W — olkoot ne F ja H.

LAUSE 4.5.3. KANNANVAIHTO LINEAARIKUVAUKSELLE. *Olkoon $L : V \to W$ lineaarikuvaus ja $A = Mat(L; E, F)$ sitä vastaava matriisi kannoissa E ja F. Saman kuvauksen L matriisi kannoissa G ja H on*

$$Mat(L; G, H) = \widehat{C}^{-1}AC,$$

missä $C = Mat(Id; G, E)$ ja $\widehat{C} = Mat(Id; H, F)$.

Myös tämän lauseen tuloksen muistaa helposti suuntaa vaille ja suunnan saa toistamalla tarvittaessa seuraavan lyhyen todistuksen:

TODISTUS.

$$\begin{tikzcd} V_E \arrow[r, "L", "A"'] & W_F \arrow[d, "(\widehat{C}\uparrow)"', "\widehat{C}^{-1}=\mathrm{Mat}(Id_W;F,H)"] \\ V_G \arrow[u, "\mathrm{Mat}(Id_V;G,E)=C"] \arrow[r, "L", "\widehat{C}^{-1}AC"'] & W_H \end{tikzcd}$$

$L = Id_W \circ L \circ Id_V$, joten matriisituloa koskevan lauseen 4.3.4 mukaan

$$\mathrm{Mat}(L; G, H) = Mat(Id, F, H) \circ \mathrm{Mat}(L; E, F) \circ \mathrm{Mat}(Id; G, E) = \widehat{C}^{-1}AC.$$

□

Lausetta käytettäessä kannattaa muistaa, että matriisin C sarakkeet ovat V:n uudet kantavektorit g_i lausuttuina alkuperäisessä kannassa E ja matriisin $\widehat{C}$ sarakkeet ovat W:n uudet kantavektorit h_i lausuttuina alkuperäisessä kannassa F.

Erikoistapaus, jossa $V = W$, $E = F$ ja $G = H$ vastaa kannanvaihtoa yhdessä avaruudessa, jota lineaarikuvauksemme kuvaa itselleen. Nyt $\widehat{C} = C$. Tämä tilanne on erityisen yleinen. Siksi sitä kuvaamaan on käytössä omaa sanastoa.

MÄÄRITELMÄ 4.5.4. Matriisit A ja B ovat *similaareja*, jos on olemassa kääntyvä matriisi C siten, että $B = C^{-1}AC$.

LAUSE 4.5.5.

(1) *A ja B ovat similaareja* $\Longrightarrow$ $\det A = \det B$.
(2) *Jokainen neliömatriisi A on similaari itsensä kanssa.*
(3) *A ja B ovat similaareja* $\Longrightarrow$ *on olemassa D siten, että $A = D^{-1}BD$.*
(4) *A ja B ovat similaareja ja lisäksi B ja C ovat similaareja* $\Longrightarrow$ *A ja C ovat similaareja.*

TODISTUS. Harjoitustehtävä. $\square$

SEURAUS 4.5.6. *Matriisit A ja B ovat similaareja jos ja vain jos ne vastaavat samaa lineaarikuvausta $L : V \to V$, joskin tietysti eri kannassa.*

5. Ominaisarvoprobleema

5.1. Diagonalisoituvuus.

Tässä luvussa V on äärellisulotteinen vektoriavaruus ja $n = \dim V$.

LAUSE 5.1.1. *Olkoon* $E = (e_1, e_2, \dots, e_n)$ *avaruuden* V *kanta ja* $L : V \to V$ *lineaarikuvaus. Seuraavat ehdot ovat yhtäpitäviä*

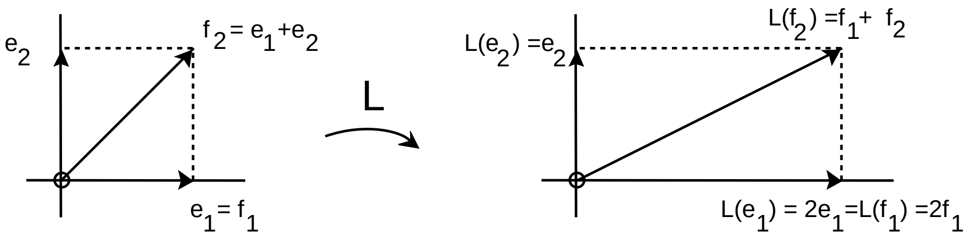
(1) *Kuvauksen* L *matriisi kannassa* E *on diagonaalimatriisi, siis*

$$Mat(L;E) = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

(2) *On olemassa luvut* $\lambda_1, \dots \lambda_n \in \mathbb{R}$, *joille pätee*

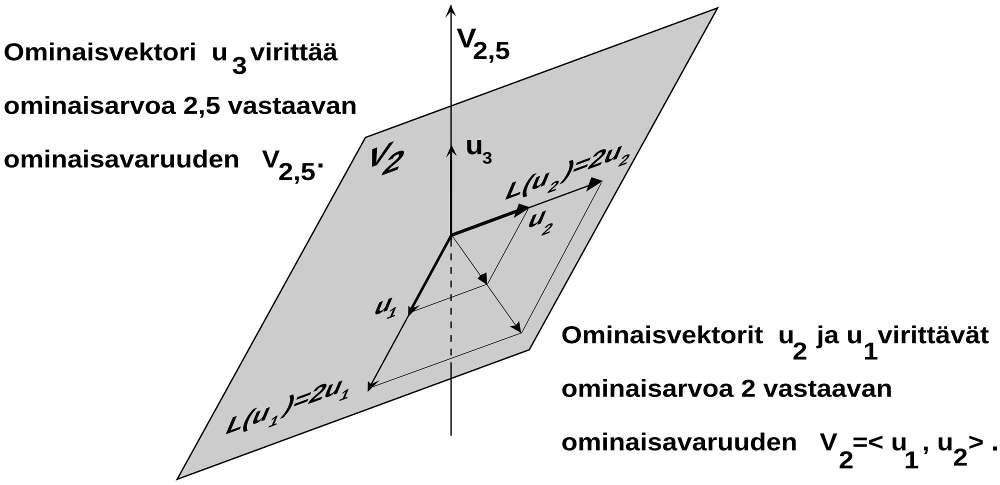
$$Le_j = \lambda_j e_j \quad \textit{kaikille } j = 1, 2, \dots, n.$$

Lineaarikuvaus $L : V \to V$ on *diagonalisoituva*, jos sen matriisi jossakin kannassa on diagonaalimatriisi. Kantaa, jossa L:n matriisi on diagonaalinen, sanotaan kuvaksen L *ominaiskannaksi.* Diagonalisoituvan kuvauksen matriisi ei yleensä ole diagonaalinen muissa kannoissa, vaan diagonaalimatriisi menettää diagonaalisuutensa kannanvaihdossa, kuten heti huomaa:



Kuvassa samaa lineaarikuvausta esittää $E-$kannassa diagonaalimatriisi $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ ja $F-$kannassa matriisi $\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$, onhan $Lf_1 = 2f_1$ ja $Lf_2 = f_1 + f_2$. Sanomme, että *matriisi* A on *diagonalisoituva*, jos sen esittämä lineaarikuvaus on diagonalisoituva, eli jos A esittää diagonalisoituvaa lineaarikuvausta. Matriisi $A = A_{n\times n}$ on siis diagonalisoituva, jos se on similaarinen jonkin diagonaalimatriisin kanssa, ts. on olemassa kääntyvä matriisi $C = C_{n\times n}$ siten, että $C^{-1}AC$ on jokin diagonaalimatriisi D. On siis hyvin helppoa tuottaa diagonalisoituvia matriiseja. Tarvitsee vain kertoa diagonaalimatriisi toiselta puolelta jollakin neliömatriisilla ja toiselta puolelta sen käänteisellä. Mutta miten voi suoraan nähdä matriisista, onko se diagonalisoituva ja mistä diagonaalimatriisista se siinä tapauksessa on saatu? Koska diagonaalimuotoinen matriisi on erittäin helppo käsiteltävä ja geometrisesti selkeä, on kiinnostavaa selvittää, mitkä matriisit ovat diagonalisoituvia ja miten ominaiskanta voidaan laskea.

Alamme ratkaista diagonalisoimisongelmaa etsimällä edes yhden vektorin, jota lineaarikuvauksemme vain venyttää — ominaisvektorin. Ominaisarvojen ja -vektorien etsimistehtävää sanotaan *ominaisarvo-ongelmaksi.*[55]



MÄÄRITELMÄ 5.1.2. Olkoon $L : V \to V$ lineaarikuvaus. Jos on olemassa $\lambda \in \mathbb{R}$ ja $u \in V \setminus \{0\}$ siten, että

$$Lu = \lambda u,$$

niin λ on L:n *ominaisarvo* ja u siihen liittyvä *ominaisvektori* eli $\lambda-$*ominaisvektori*. Ominaisarvoon λ liittyvä *ominaisavaruus* on kaikkien $\lambda-$ominaisvektoreiden joukko

$$V_\lambda(L) = \{u \in V \mid Lu = \lambda u\}.$$

Nimi vihjaisee, että ominaisavaruus on aliavaruus:

LAUSE 5.1.3. *$V_\lambda(L)$ on V:n aliavaruus ja* $\dim V_\lambda(L) \geq 1$.

Tämä asia onkin helppo tarkastaa. Diagonalisoimisongelman kannalta kiinnostavaa on yrittää muodostaa koko avaruuden kanta L :n eri ominaisarvoihin liittyvistä ominaisvektoreista, sillä juuri sellaisessa ominaiskannassa kuvauksemme diagonalisoituisi, diagonaalimatriisin lävistäjäelementteinä ominaisarvot. Siksi on kiintoisaa huomata, että pätee

LAUSE 5.1.4. *Lineaarikuvauksen $L : V \to V$* **eri** *ominaisarvoihin liittyvät ominaisvektorit ovat lineaarisesti riippumattomia.*

TODISTUS. Tarkastamme asian ensin kahden vektorin tapauksessa. Olkoot u ja v ominaisarvoihin $\alpha \neq \beta$ liittyvät ominaisvektorit ja λ ja μ lukuja, joille

$$(*) \qquad \lambda u + \mu v = 0.$$

Todistetaan, että kertoimet λ ja μ ovat nollia. Yhtälöstä $(*)$ saadaan kuvauksen L lineaarisuuden nojalla

$$\lambda L u + \mu L v = 0,$$

[55]Ominaisvektoreilla on paljon muutakin käyttöä kuin diagonalisointi.

josta ominaisvektorin määritelmän perusteella

$$(**)\qquad \lambda\alpha u + \mu\beta v = 0.$$

Suoraan luvulla α kertomalla yhtälöstä $(*)$ seuraa toisaalta

$$(***)\qquad \lambda\alpha u + \mu\alpha v = 0,$$

ja erotuksena saadaan $\mu(\beta - \alpha)v = 0$. Oletusten mukaan $\beta - \alpha \neq 0$, eikä vektori v ole nolla, sillä nollavektoria ei kelpuuteta ominaisvektoriksi. Näin ollen $\mu = 0$. Yhtälöstä $(*)$ seuraa tämän jälkeen, että myös λ on nolla.

Seuraavaksi tarkastamme lineaarisen riippumattomuuden kolmen eri ominaisarvon tapauksessa:

Olkoot v_1, v_2 ja v_3 eri ominaisarvoihin α_1, α_2 ja α_3 liittyvät ominaisvektorit ja λ_1, λ_2, λ_3 lukuja, joille

$$(*)\qquad \sum_{j=1}^{3} \lambda_j v_j = 0.$$

Todistetaan, että kertoimet λ_j ovat nollia. Yhtälöstä $(*)$ saadaan kuvauksen L lineaarisuuden nojalla

$$\sum_{j=1}^{3} \lambda_j L v_j = 0,$$

josta ominaisvektorin määritelmän perusteella

$$(**)\qquad \sum_{j=1}^{3} \lambda_j \alpha_j v_j = 0.$$

Suoraan luvulla α_1 kertomalla yhtälöstä $(*)$ seuraa toisaalta

$$(***)\qquad \sum_{j=1}^{3} \lambda_j \alpha_1 v_j = 0$$

ja erotuksesta häviää taas ensimmäinen termi, jolloin siihen jää kaksi termiä:

$$\sum_{j=2}^{3} \lambda_j (\alpha_1 - \alpha_j) v_j = 0.$$

v_2 ja v_3 ovat kaksi eri ominaisarvoihin liittyvää ominaisvektoria, siis todistuksen alkuosan nojalla lineaarisesti riippumattomat. Molemmat kertoimet $\lambda_j(\alpha_1 - \alpha_j)$ ovat siis nollia. Tästä seuraa, että nimenomaan kumpikin λ_j on nolla, sillä oli oletettu, että $\alpha_1 \neq \alpha_j$.

Huomaamme, että olemme keksineet induktiotodistuksen lauseen väitteelle. Esittämämme todistuksen ensimmäinen osa antaa induktion alun ja toinen osa yleistyy luonnollisella tavalla induktioaskeleeksi, jolla lause todistetaan k:lle vektorille, jos se tunnetaan $k-1$:lle. $\square$

Edellisen lauseen todistus oli hiukan mutkikas, tarvittiinhan induktiota. Vaivan palkka on kuitenkin hyvä, sillä seuraava lause seuraa suoraan äskeisestä:

Lause 5.1.5. *a) Olkoon* $L : V \to V$ *lineaarikuvaus ja* $\dim V = n$. *Tällöin* L*:llä on korkeintaan* n *eri ominaisarvoa.*

b) Jos **eri** *ominaisarvoja on maksimaaliset* n *kappaletta, niin niitä vastaavat ominaisvektorit muodostavat* V*:n kannan ja* L*:n matriisi tämän kannan suhteen on diagonaalimatriisi, jossa diagonaalialkioina ovat* L*:n ominaisarvot. Tällainen lineaarikuvaus on siis diagonalisoituva.*

Jos ominaisarvoja on vähemmän, lineaarikuvaus voi kuitenkin olla diagonalisoituva tai sitten ei. Esimerkki diagonalisoitumattomasta lineaarikuvauksesta on tason $\mathbb{R}^2$ kierto vaikkapa kulman $\frac{\pi}{4}$ verran.

Vaikka eri ominaisarvoja olisi vähemmn kuin n kappaletta, voi diagonalisoituvuuden kuitenkin päätellä, jos ominaisavaruudet ovat tarpeeksi suuria. Määritelmän mukaanhan matriisi $A_{n\times n}$ on diagonalisoituva, jos ja vain jos sen ominaisvektoreista voidaan muodostaa $\mathbb{R}^n$:n kanta, ts. jos A:lla on n kappaletta lineaarisesti riippumattomia ominaisvektoreita:

Lause 5.1.6. *Matriisi* $A_{n\times n}$ *on diagonalisoituva, jos ja vain jos sen eri ominaisarvoja vastaavien ominaisavaruuksien dimensioiden summa* $\sum_\lambda \dim V_\lambda$ *on* n*. Tällöin diagonalisoivan matriisin* C *sarakkeina ovat* A*:n ominaisvektorit ja diagonaalimatriisin* $D = C^{-1}AC$ *lävistäjäalkioina* A*:n ominaisarvot vastaavassa järjestyksessä.*

Todistus. Valitaan kullekin ominaisavaruudelle V_λ kanta $(v_{\lambda,1}, \dots, v_{\lambda,k_\lambda})$, missä $k_\lambda = \dim V_\lambda$. Riittää todistaa, että näin on yhteensä saatu avaruuden $\mathbb{R}^n$ kanta. Vektoreita $v_{\lambda,j}$ on oletuksen mukaan oikea määrä, siis n kappaletta, joten riittää todistaa, että ne ovat lineaarisesti riippumattomia. Olkoon siis

$$\sum_\lambda \sum_{j=1}^{k_\lambda} \mu_{\lambda,j} v_{\lambda,j} = 0.$$

Olemme ryhmitelleet summan ominaisarvojen mukaan ja huomaamme, että jokainen osa

$$\sum_{j=1}^{k_\lambda} \mu_{\lambda,j} v_{\lambda,j}$$

on ominaisavaruudessa V_λ, siis ominaisarvoon λ liittyvä ominaisvektori. Koska eri ominaisarvoihin liittyvät ominaisvektorit ovat lineaarisesti riippumattomia, on kukin tällainen osa siis $\sum_{j=1}^{k_\lambda} \mu_{\lambda,j} v_{\lambda,j} = 0$. Koska tässä nolla on lineaarikombinaationa ominaisavaruuden V_λ kantavektoreista, kertoimet $\mu_{\lambda,j}$ ovat nollia. □

5.2. Matriisin ominaisvektoreiden ja -arvojen laskemisesta.

Matriisin A ominaisvektoreiden ja ominaisarvojen löytäminen on sama asia kuin yhtälön

$$(*) \qquad Ax = \lambda x$$

ratkaiseminen, missä tuntemattomia ovat vektori x ja luku λ. Alamme selvittää, miten tämän yhtälön voi ratkaista.

Kirjoittamalla yhtälön täydelliseen muotoon

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = \lambda x_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = \lambda x_2 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = \lambda x_n \end{cases}$$

huomaa, että se on epälineaarinen — sisältäähän se tuntemattomien tuloja λx_j. Yhtälö muuttuu kuitenkin tuntemattomien lukujen x_j lineaariseksi yhtälöryhmäksi heti, kun tuntematon luku λ on jollakin tavalla onnistuttu löytämään. Tarkemmin ajatellen yhtälöryhmällä tietysti on aina triviaaliratkaisu $x = 0$, olipa λ mikä tahansa, mutta triviaaliratkaisu on kielletty ominaisvektorin määritelmässä. Havainnosta on kuitenkin se ilo, että tulemme nyt ajatelleeksi, että yhtälö $(*)$ on itse asiassa muuttujien x osalta homogeeninen lineaariyhtälö, voihan sen kirjoittaa myös muotoon

$$Ax - \lambda x = 0,$$

eli

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \cdots + a_{2n}x_n = 0 \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \lambda)x_n = 0, \end{cases}$$

eli lyhyesti

$$(A - \lambda I)x = 0, \tag{$*$}$$

missä I on identtinen eli ykkösmatriisi. Ja nyt suuri OIVALLUS! **Luku λ on matriisin A ominaisarvo täsmälleen sillä ehdolla, että yhtälöllä $(*)$ on olemassa epätriviaali ratkaisu x.** Tälle taas olemme kehitelleet ehtoja tutkiskellessamme lineaarikuvauksia. Vastaava matriisi $A - \lambda I$ ei saa olla kääntyvä, vaan on oltava

$$\det(A - \lambda I) = 0. \tag{!}$$

Lineaarikuvauksen ominaisarvojen määrääminen on siis palautunut **yhden muuttujan λ** yhtälön ratkaisemiseksi. Kokeilemalla jollakin matriisilla huomaa, että $\det(A - \lambda I)$ on muuttujan λ polynomi. Asia on tietysti yleisestikin juuri näin. Ongelmana on siis enää (?) tämän n:nnen asteen polynomin nollakohtien löytäminen. Tulos on siinä määrin merkittävä, että on aihetta nimetä muutamia aiheeseen liittyviä käsitteitä, jotta niistä olisi helpompi puhua jatkaessamme ongelman purkamista.

MÄÄRITELMÄ 5.2.1. Olkoon $L : V \to V$ lineaarikuvaus ja $A = Mat(L, E)$ sitä vastaava matriisi jossakin V:n kannassa. Lineaarikuvauksen L ja matriisin A *karakteristinen polynomi* on

$$f_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix}$$

ja *karakteristinen yhtälö* on

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

Luvun alun tarkastelut ovat johtaneet oivallukseen, jonka toistamme lauseena:

LAUSE 5.2.2. *L:n ominaisarvot ovat karakteristisen polynomin* $\det(A - \lambda I)$ *nollakohdat.*

Lukijaa saattaa ärsyttää, että puhumme lineaarikuvauksen karakteristisesta polynomista, vaikka polynomi on määritelty kannan valinnasta riippuvan matriisin avulla. Puhetapamme on kuitenkin oikeutettu, sillä samaa lineaarikuvausta eri kannoissa esittävät matriisit ovat similaareja ja similaareilla matriiseilla on lauseen sama determinantti, joten

$$\det(C^{-1}AC - \lambda I) = \det(C^{-1}(A - \lambda I)C) = \det(A - \lambda I).$$

Näin olemme todistaneet, että pätee:

LAUSE 5.2.3. *Lineaarikuvauksen* $L : V \to V$ *karakteristinen polynomi ei riipu avaruuden* V *kannan valinnasta, vaan ainoastaan lineaarikuvauksesta* L*.*

Olemme nyt löytäneet menetelmän, jolla ominaisarvoprobleema voidaan periaatteessa ratkaista.

MALLI 5.2.4. PERUSMENTTELY MATRIISIN A OMINAISVEKTOREIDEN LÖYTÄMISEKSI.

- Kirjoitetaan näkyviin matriisi $A - \lambda I$. (λ on tuntematon.)
- Lasketaan $\det(A - \lambda I)$. (Tähän on käytettävissä luvun 2 opit.)
- Ratkaistaan yhtälö $\det(A - \lambda I) = 0$. (Kun A on $n \times n$–matriisi, $\det(A - \lambda I)$ on tuntemattoman λ polynomi, jonka aste on n. Sen nollakohtien löytäminen on siten periaatteessa vaikea paikka, mutta onneksi se ei kuulu lineaarialgebran piiriin, joten asiaa ei tarvitse pohtia juuri tässä. Onnea vain yhtälönratkaisemiseen!)[56]
- Listataan näkyville — mielellään esimerkiksi suuruusjärjestyksessä — karakteristisen polynomin reaaliset nollakohdat, jotka ovat etsityt ominaisarvot $\lambda_1, \ldots, \lambda_k$. Nollakohtien kertalukuja sanotaan ominaisarvojen *algebrallisiksi kertaluvuiksi.*
- Kutakin ominaisarvoa λ_j kohti etsitään vastaavat ominaisvektorit ratkaisemalla homogeeninen yhtälöryhmä $(A - \lambda_j I)x = 0$. Ratkaisujen joukko, matriisin $A - \lambda_j I$ ydin, on ominaisarvoon λ_j liittyvien ominaisvektoreiden joukko eli ominaisavaruus. Ominaisavaruuksien dimensioita sanotaan

[56]Kannattaa kuitenkin kerrata, mitä on joskus osannut toisen ja kolmannen asteen yhtälöistä ja funktion nollakohtien numeerisesta määräämisestä. On hyvä muistaa, että n:nnen asteen polynomilla on aina kertaluvut huomioiden n nollakohtaa, jotka tosin saattavat olla kompleksilukuja. Me olemme etsimässä reaalisia nollakohtia, mutta huomaamme, että koko teoriaa kannattaisi ehkä laajentaa kompleksilukujen suuntaan, jolloin tämä kohta olisi helpompi. Kompleksisilla ominaisarvoilla on sitä paitsi reaalisillekin matriiseille huomattava geometrinen merkitys; ne liittyvät avaruuden kiertämiseen ja ovat siten fysiikassa yhtä tärkeitä kuin reaalisetkin. Perusesimerkki on tason kierto kulmalla θ, jonka kompleksiset ominaisarvot ovat $exp(\pm i\theta)$. Kompleksista lineaarialgebraa käsitellään tämän kirjan viimeisillä sivuilla. Hyvä selostus avaruuden kierroista on ainakin kirjassa LAY: LINEAR ALGEBRA.

vastaavien ominaisarvojen *geometrisiksi kertaluvuiksi.* (Olemme opetelleet määräämään ytimelle kannan. Näin kannattaa yleensä menetellä, sillä kannan ilmoittaminen on kätevä tapa kuvata aliavaruutta.)

- Kirjoitetaan näkyviin luettelo kunkin ominaisavaruuden kantavektoreista.

5.3. Lineaarikuvauksen diagonalisointi.

Olemme oikeastaan jo selittäneet, miten lineaarikuvaus diagonalisoidaan, jos se on mahdollista. Kertaamme vielä pääkohdat. Lineaarikuvaus $n-$ulotteiselta vektoriavaruudelta itselleen $L : V \to V$ on diagonalisoituva, jos sen ominaisvektoreita riittää kannaksi asti. Diagonalisointi tarkoittaa, että L lausutaan ominaisvektoreista muodostuvassa kannassa, siis diagonaalimatriisina. Myös kanta on yleensä syytä ilmoittaa, mutta toisinaan kiinnostaa pelkkä diagonaalimatriisi, jonka diagonaalilla ovat L:n ominaisarvot toistettuina vastaavien ominaisavaruuksien dimension mukaan.

Yleensä diagonalisoitava kuvaus L on annettu matriisina A jossakin (standardi-) kannassa. Olemme edellä kuvailleet, miten sille löydetään kaikki ominaisavaruudet ja kullekin kanta. Kaiken kaikkiaan on löydetty mahdollisimman monta lineaarisesti riippumatonta ominaisvektoria. Jos näitä vektoreita on yhteensä $n = \dim V$ kappaletta — enemmän niitä ei voi n–ulotteisessa avaruudessa ollakaan — niin ne muodostavat etsityn ominaiskannan, jossa L diagonalisoituu. Muuten L ei diagonalisoidu.

Myönteisessä tapauksessa tehtäväksi jää kannanvaihto matriisille $A = Mat(L, E)$. Kannanvaihtoon tarvittava matriisi C on se, jonka sarakkeina ovat listaamamme ominaiskantavektorit. Jos laskuvirheitä ei ole, kannanvaihto antaa diagonaalimatriisin $D = C^{-1}AC$, jonka diagonaalilla ovat ominaisarvot.

Jää pohdittavaksi, voisiko matriisista A jotenkin etukäteen nähdä, onko se diagonalisoituva. Diagonalisointihan on aika työlästä, eikä sitä soisi tekevänsä turhan päiten. Tunnistaminen ei ole helppoa, mutta todistamme myöhemmin, että ainakin kaikki symmetriset matriisit ovat diagonalisoituvia. Muitakin diagonalisoituvia matriiseja on kuitenkin paljon.

5.4. Kartioleikkausten tunnistamisesta.

Tässä luvussa sovellamme ominaiskantateoriaa seuraavaan ongelmaan, kartioleikkauksen *pääakselikoordinaatiston* etsimiseen. Voisimme kysyä, mitä tasokäyrää esittää yhtälö

$$3x^2 + 4xy + 7y^2 - 88y + 9x = 20 \ ?$$

Entäpä $x^2 + 5xy - 55y^2 + y = 8$ tai yleisemmin mikä tahansa kahden muuttujan toisen asteen yhtälö

$$Ax^2 + Bxy + Cy^2 + Dx + Ey = F \ ?$$

Käsittelemme vain tapauksen, jossa lineaariset termit Dx ja Ey puuttuvat ja vakio F on 1. Asia on nimittäin niin, että yleisen yhtälön voi lähes aina palauttaa tähän muotoon ”siirtämällä origon keskelle käyrää”.[57]

[57] On olemassa poikkeus — paraabelilla ei ole keskipistettä — mutta ongelma ei ole vakava.

Ongelma 5.4.1. Mitä käyrää esittää toisen asteen yhtälö

$$ax^2 + 2bxy + cy^2 = 1 \ ?$$

Esitämme ratkaisun tähän kysymykseen vaiheittain:

Idea 5.4.2. Tehtävän muunto matriisimuotoon.

Ongelmamme kytkeytyy lineaarialgebraan, kun huomaamme, että jos $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, niin matriisitulona saadaan

$$[x\, y]A \begin{bmatrix} x \\ y \end{bmatrix} = [ax^2 + 2bxy + cy^2].$$

Tämä on syy sille, että $xy-$termin kerroin kirjoitettiin jo tehtävää laadittaessa muotoon $2b$. Idea on nyt seuraava: **siirrytään A:n ominaiskantaan**. Sellainen on todella olemassa. Ominaiskannassa A diagonalisoituu ja ongelma on ehkä ratkennut. Toiveita herättää, että jos A on diagonaalinen eli $b = 0$, niin tutkittava yhtälö on tyyppiä

$$ax^2 + cy^2 = 1,$$

joka tunnetusti[58] esittää ellipsiä tai hyperbeliä sen mukaan, ovatko a ja c saman- vai erimerkkisiä.

Vaihe 5.4.3. Ominaisarvot ja ominaiskannan olemassaolo.

Tehtävänä on diagonalisoida symmetrinen $2 \times 2-$ matriisi $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Karakteristinen polynomi on

$$f_A(\lambda) = \begin{vmatrix} a - \lambda & b \\ b & c - \lambda \end{vmatrix} = \lambda^2 - (a + c)\lambda + ac - b^2,$$

jonka nollakohdat ovat toisen asteen yhtälön ratkaisukaavan mukaan

$$\lambda_{1,2} = \frac{a + c}{2} \pm \sqrt{b^2 + \left(\frac{a - c}{2}\right)^2}.$$

Diskriminantti $b^2 + \left(\frac{a-c}{2}\right)^2$ on kahden neliön summa, siis varmastikaan ei negatiivinen. Nollaksi diskriminantti voi kuitenkin mennä, nimittäin tapauksessa, jossa $b = 0$ ja lisäksi $c = a$, eli kun $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$. Tässä tapauksessa A diagonalisoituu, onhan se jo valmiiksi diagonaalinen. Itse asiassa käyräkin on helppo tunnistaa:

$$ax^2 + ay^2 = 1$$

on epäilemättä $\frac{1}{\sqrt{a}}-$säteisen ympyrän ($a > 0$), pisteen ($a = 0$) tai tyhjän joukon ($a < 0$) yhtälö.

[58] Paras opetella!

Muuten A:lla on kaksi eri ominaisarvoa $\lambda_1 \neq \lambda_2$, jotka juuri laskimme. Muistamme matriisin diagonalisoinnista sen, että näistä ominaisarvoista saatavassa ominaiskannassa matriisia A vastaa diagonaalimatriisi

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

ja että tämä myös merkitsee sitä, että

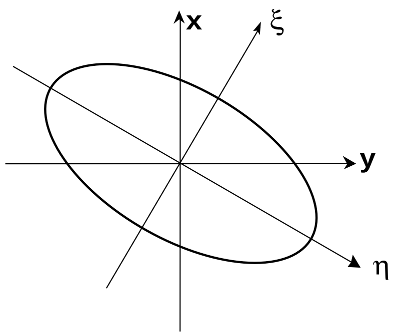
$$D = C^{-1}AC,$$

missä

$$C = [v_1 v_2]$$

on matriisi, jonka sarakkeet ovat A:n eri ominaisarvoihin liittyvät ominaisvektorit.

Vaihe 5.4.4. Huoleton arvaus.



Arvataan, että ominaiskantaa vastaavassa koordinaatistossa — koordinaatit olkoot ξ ja η — tutkittavan käyrän yhtälö on

$$[\xi\ \eta] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} = [\lambda_1 \xi^2 + \lambda_2 \eta^2] = 0,$$

jonka tunnistamme kuten kohdassa 5.4.2. Arvaus osuu oikeaan, mutta asia ei ole ihan niin vaaraton kuin ensi silmäykseltä näyttää. Emme ole vielä todistaneet, että yhtälö muuntuu toivomallamme tavalla. Laskussa onkin vielä pieni mutka. Ominaiskanta voi sitä paitsi yleensä olla vinokulmainen, eikä ole vielä ihan selvää, että koulusta tutut ellipsin ja hyperbelin yhtälöt toimivat. Todistamme seuraavassa luvussa, että jokaisen symmetrisen matriisin ominaiskantavektorit ovat tavallisessa mielessä kohtisuorassa toisiaan vastaan, mutta emme tarvitse tässä esimerkissä vielä mitään teoriaa: Asiahan on helppo tarkastaa tässä laskemalla ominaisvektorit ja toteamalla niiden sisätulo sitten nollaksi:

Työvaihe 5.4.5. Ominaiskannan löytäminen.
Kun λ on A:n ominaisrvo, vastaava ominaisvektori saadaan ratkaisemalla yhtälö $Ax = \lambda x$ eli yhtälöpari

$$\begin{cases} (a-\lambda)x + by = 0 \\ bx + (c-\lambda)y = 0 \end{cases},$$

jonka rivit ovat LD, koska ratkaisuna on muutakin kuin 0. Riittää siis ratkaista vain jompi kumpi, esimerkiksi ensimmäinen yhtälö, jolloin saadaan ominaisvektori

$$v_\lambda = \begin{bmatrix} b \\ \lambda - a \end{bmatrix}, \tag{$*$}$$

tai sitten toinen, jolloin sama saadaan muodossa

$$v_\lambda = \begin{bmatrix} \lambda - c \\ b \end{bmatrix}. \tag{$**$}$$

Ominaisarvoihin λ_1 ja λ_2 liittyvät ominaisvektorit saadaan kummasta tahansa sijoittamalla $\lambda = \lambda_1$ ja $\lambda = \lambda_2$.

Tarkastus 5.4.6. Ominaiskannan suorakulmaisuus.
Jos kirjoitamme toisen ominaisvektorin muotoon ($*$) ja toisen muotoon ($**$), voimme helposti laskea sisätulon

$$(v_1|v_2) = b(\lambda_1 - a) + b(\lambda_2 - c) = b(\lambda_1 + \lambda_2 - (a + c)) = 0.$$

Ominaisvektorit ovat siis suorakulmaiset toisiinsa nähden. Ne voi lisäksi valita **ykkösen mittaisiksi**, ja niin teemmekin, jolloin on löydetty *ortonormaali* ominaiskanta.

Tarkastus 5.4.7.
On lopuksi vielä tarkastettava, että yhtälö

$$[\xi\ \eta]D \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \lambda_1 \xi^2 + \lambda_2 \eta^2 = 0, \tag{1}$$

todella esittää ominaiskannassa samaa käyrää kuin

$$[x\ y]A \begin{bmatrix} x \\ y \end{bmatrix} = 0 \tag{2}$$

alkuperäisessä. Tämän tarkastaminen on koordinaatistonvaihto: Uudet koordinaatit saadaan vanhoista kaavalla

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = C^{-1} \begin{bmatrix} x \\ y \end{bmatrix},$$

joten

$$[\xi\ \eta] = \begin{bmatrix} \xi \\ \eta \end{bmatrix}^T = \left(C^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \right)^T = [x\ y] \left(C^{-1} \right)^T.$$

Tavoittelemamme yhtälö (1) sanoo siis, että

$$[x\ y] \left(C^{-1} \right)^T D C^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = 0. \tag{3}$$

Mutta, mutta — alkuperäinen yhtälö (2) on

$$[x\ y]\underbrace{CDC^{-1}}_{A}\begin{bmatrix} x \\ y \end{bmatrix} = 0. \tag{2}$$

Onko siis sittenkin arvattu väärin? Ei ole! Matriisilla C on yhtälöt (2) ja (3) samaksi tekevä ominaisuus $\left(C^{-1}\right)^T = C$ eli

$$C^{-1} = C^T.$$

Tämä johtuu yksinkertaisesti siitä, että C:n sarakkeet ovat ortonormaaleja, siis $C = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$, missä $\alpha^2 + \beta^2 = 1$. Lukija todetkoon laskemalla, että tällä ehdolla todella

$$C^T C = I_{2\times 2}.$$

□

Arviointi 5.4.8. Olemme kehittäneet menetelmän origokeskisen ellipsin ja hyperbelin kiertämiseksi symmetria- eli pääakselikoordinaatistoonsa. Tämä oli kaksiulotteinen ongelma. Samoja ideoita voi soveltaa kolmi- tai useampiulotteisena versiona esimerkiksi toisen asteen pintojen tutkimiseen avaruudessa $\mathbb{R}^3$ ja paljoon muuhunkin. Kootkaamme oleelliset havainnot:

- Päädyimme diagonalisoimaan symmetristä matriisia $A = A^T$.
- Käytimme sisätuloa ensimmäisen kerran johdantoluvun jälkeen.
- Kannanvaihtomatriisin C sarakkeet, siis symmetrisen matriisin A ominaisvektorit ovat sisätulon mielessä kohtisuorassa toisiaan vastaan ja ykkösen pituisia.
- Kannanvaihtomatriisin C sarakkeiden ortonormaalius merkitsee, että $C^{-1} = C^T$.

Näillä ominaisuuksilla on tärkeä geometrinen merkitys. Palaamme symmetrisen matriisin diagonalisointiin seuraavissa luvuissa.

6. Sisätuloavaruudet

Edellisen luvun lopussa käytimme tavallisen tason $\mathbb{R}^2$ vektoreiden sisätuloa. Standardisisätulo on määritelty koordinaattien avulla, eikä sitä siis ole olemassa yleisessä vektoriavaruudessa. Standardisisätulon tärkeimmät ominaisuudet voi kuitenkin listata aksioomiksi samaan tapaan kuin tason vektoreiden ominaisuudet koottiin vektoriavaruuden määritelmäksi. Näin saadaan abstraktin sisätuloavaruuden käsite.

6.1. Sisätulo.

Määritelmä 6.1.1. *(Reaalinen) sisätuloavaruus* on vektoriavaruus V varustettuna yhteenlaskun ja luvulla kertomisen lisäksi kolmannella laskutoimituksella, *sisätulolla*

$$(\cdot|\cdot) : V \times V \to \mathbb{R},$$

jos sillä on seuraavat ominaisuudet: Kaikilla $u, v, w \in V$ ja $\lambda \in \mathbb{R}$

(ML-2) $(u + v|w) = (u|w) + (v|w)$
(ML-1) $(\lambda u|v) = \lambda(u|v)$
(SYM) $(u|v) = (v|u)$
(POS) $(u|u) \geq 0$
(DEF) $u = 0 \iff (u|u) = 0.$

Sisätulo on siis bilineaarikuvaus, joka lisäksi on *symmetrinen* ja *positiividefiniitti.* Huomaa, että symmetria (SYM) on determinanttien yhteydessä käsittelmällämme alternoivuudelle (ALT) tavallaan vastakkainen ominaisuus.

Esimerkkejä 6.1.2. a) Standardisisätulo avaruudessa $\mathbb{R}^n$

$$(x|y) = x \cdot y = \sum_{j=1}^{n} x_j y_j$$

on tunnetusti esimerkki sisätulosta. Ominaisuudet on mainittu jo johdantoluvussa. Standardisisätulolla on läheinen yhteys matriiseihin. Matriisien kertolaskuhan määritellään kertomalla vasemman matriisin rivit oikeanpuoleisen sarakkeilla juuri standardisisätulon mielessä. Vektoreiden x ja y standardisisätulon voi myös — sulkeita vaille — ilmaista matriisitulona, kun vektorit on annettu standardikannassa, siis jos

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ ja } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \text{ niin}$$

$$[(x|y)] = \left[\sum_{j=1}^{n} x_j y_j \right] = [x_1, x_2, \ldots, x_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x^T y = y^T x.$$

b) *Kantaan E liittyvä standardisisätulo* äärellisulotteisessa avaruudessa V määritellään samalla tavalla: Olkoon $E = (v_1, \ldots, v_n)$ jokin kanta avaruudessa V. Jos

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_E \quad \text{ja} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_E, \text{ niin}$$

$$(x|y)_E = \sum_{j=1}^{n} x_j y_j.$$

c) Olkoon $V = \mathcal{C}([0,1], \mathbb{R}) = \{f : [0,1] \to \mathbb{R} |\ f \text{ on jatkuva.}\}$ Sisätuloksi asetamme:

$$(f|g)_{\mathcal{C}} = \int_0^1 f(t)g(t)\, dt.$$

d) Olkoon $V = \mathbb{R}^{(\mathbb{N})} = \{x = (x_1, x_2, x_3, \ldots) |\ \exists n_0 \text{ siten, että } x_n = 0\, \forall\, n \geq n_0.\}$ Sisätuloksi asetamme äärellisen summan:

$$(x|y) = \sum_{i=0}^{\infty} x_i y_i$$

Lukija todistakoon tarkastamalla kaikkien ehtojen pätevyyden, että nämä todella ovat sisätuloja. Keinotekoisesti voi heti kehittää muitakin esimerkkejä, mutta edellä luettelemamme ovat perusesimerkit sisätuloista.[59]

Varmistettuamme, että muitakin sisätuloja kuin tavallinen pistetulo on olemassa saamme aiheen tarkastaa, mitkä tavallisen pistetulon ominaisuudet seuraavat aksioomista eli ovat voimassa jokaiselle sisätulolle. Pääsemme alkuun heti:

LAUSE 6.1.3. *Sisätuloavaruudessa V on voimassa kaikille vektoreille ja luvuille:*

1° $(0|v) = (v|0) = 0$
2° $(u|v + w) = (u|v) + (u|w)$
3° $(u|\lambda v) = \lambda(u|v)$
4° $\left(\sum_{i=1}^{k} \lambda_i u_i \middle| \sum_{j=1}^{n} \mu_j v_j\right) = \sum_{i=1}^{k} \sum_{j=1}^{n} \lambda_i \mu_j (u_i|v_j)$.

TODISTUS. Helppoja □

Standardisisätulon tärkein ominaisuus on, että sen avulla voi laskea vektoreiden pituuksia, pisteiden etäisyyksiä ja vektoreiden välisiä kulmia. Mielivaltaisessa sisätuloavaruudessa V nämä käsitteet määritellään sisätulon avulla:

6.2. Normi, etäisyys ja kulma.

Seuraavassa V on sisätuloavaruus.

[59] Muitakin tärkeitä sisätuloja on kyllä olemassa, mutta niitä ei käsitellä cl-opinnoissa.

MÄÄRITELMÄ 6.2.1. Vektorin $v \in V$ *normi* eli *pituus* on luku

$$\|v\| = \sqrt{(v|v)}.$$

Nollasta eroavien vektoreiden u ja v välinen *kulma* on luku

$$\alpha(u,v) = \arccos \frac{(v|u)}{\|v\|\|u\|}.$$

Määritelmä on järkevä siinä mielessä, että vektorin pituus voidaan aina laskea — onhan neliöjuuren alla ei-negatiivinen luku. Pituus on positiivinen, paitsi nollavektorilla nolla. Kulman määritelmäkin on mielekäs, sillä seuraavan lauseen — CAUCHYN, SCHWARZIN ja BUNJAKOVSKIN[60] epäyhtälön — mukaan luku $\frac{(v|u)}{\|v\|\|u\|}$ on tosiaankin -1:n ja 1:n välissä, joten arcus kosini siitä on olemassa. Standardisisätulolla varustetussa tasossa $\mathbb{R}^2$ määritelmä antaa tavallisen pituus- ja kulmakäsitteen.

LAUSE 6.2.2. (CAUCHYN, SCHWARZIN JA BUNJAKOVSKIN EPÄYHTÄLÖ). *Kaikille vektoreille u, $v \in V$ sisätulo on itseisarvoltaan pienempi tai sama kuin normien tulo:*

$$|(u|v)| \leq \|u\|\|v\|.$$

TODISTUS. Lause pätee tietysti ainakin silloin, kun toinen vektoreista on nolla. Voimme siis olettaa, että molemmat eroavat nollasta. Olkoon

$$\lambda = -\frac{(u|v)}{\|v\|^2}.$$

Tätä lukua λ ei helpolla muista, mutta sen keksii uudelleen etsimällä u:n kärjen kautta v:n suuntaan kulkevalta suoralta sen pisteen $u + \lambda v$, joka on lähimpänä origoa normin mielessä. Totesimme juuri, että jokaisen vektorin sisätulo itsensä kanssa on ei-negatiivinen. Erityisesti siis

$$\begin{aligned} 0 \leq (u + \lambda v | u + \lambda v) &= (u|u) + 2\lambda(u|v) + \lambda^2(v|v) \\ &= \|u\|^2 + \frac{(u|v)^2}{\|v\|^4}\|v\|^2 - 2\frac{(u|v)}{\|v\|^2}(u|v) \\ &= \|u\|^2 - \frac{(u|v)^2}{\|v\|^2}. \end{aligned}$$

□

CSB:n epäyhtälö on geometrian peruslähtökohta, koska se merkitsee, että kulmia on olemassa. On syytä vielä tarkastaa erikseen ääritilanteet: Milloin $(u|v)$ on 0, milloin suurin mahdollinen, siis $\|u\|\|v\|$, milloin pienin mahdollinen, siis -$\|u\|\|v\|$. Tee se![61]

Vektorin normia sanottiin edellä myös sen pituudeksi. Voi myös ajatella, että vektorin normi on sen kärjen etäisyys origosta. Tavallisesta tason pistetulosta saamme idean mitata sisätuloavaruuden pisteiden **etäisyyttä** niiden välisen vektorin pituudella:

[60] Augustin Louis Cauchy (1789–1857), ranskalainen modernin analyysin isä, Karl Herman Schwarz (1843–1921) saksalainen funktioteoreetikko, Viktor Jakovlevits Bunjakovski (1804–1889), ukrainalainen mekaniikan ja matematiikan professori, yhtälön (1859) varsinainen keksijä.

[61] Osviittaa tuloksesta antaa $V = \mathbb{R}^2$ tavallisella pistetulolla varustettuna.

6.2.3. Määritelmä. Vektoreiden u ja $v \in V$ välinen *etäisyys* on

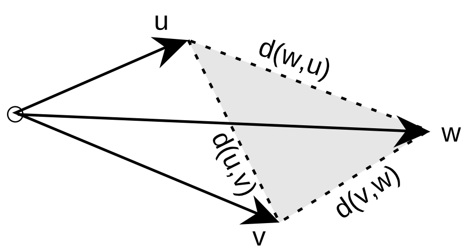
$$d(u,v) = \|u - v\|.$$

Kun osaamme mitata pisteiden välisiä etäisyyksiä, voimme yrittää harrastaa geometriaa sisätuloavaruudessa V. Voimme määritellä esimerkiksi pallonpinnan niiden pisteiden joukoksi, jotka ovat samalla etäisyydellä annetusta pisteestä ja kahden pisteen välisen janan keskinormaalitason niiden pisteiden joukoksi, jotka ovat yhtä kaukana kummastakin.[62]

Näin saadaankin geometriaa, joka monessa suhteessa muistuttaa kaksiulotteisesta avaruudesta tuntemaamme. Kuvailemme seuraavassa ensisijaisesti juuri tällaisia tuttuja asioita, mutta moniulotteisessa sisätuloavaruudessa saattaa kyllä esiintyä sellaisiakin ilmiöitä, joihin emme ole tottuneet tasossa.[63] Seuraava lause sanoo, että määrittelemällämme etäisyyskäsitteellä on etäisyyden[64] oikeat ominaisuudet:

Lause 6.2.4. (Metriikan ominaisuudet). *Kaikille $u, v \in V$ pätee*

(KOL) $d(u,w) \leq d(u,v) + d(v,w)$
(SYM) $d(u,v) = d(v,u)$
(REF) $d(u,u) = 0$
(POS) $d(u,v) \geq 0$
(DEF) $d(u,v) = 0 \implies u = v$



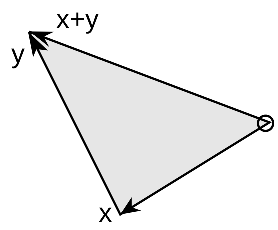
Todistus. Neljä jälkimmäistä ehtoa, symmetria (Etäisyys täältä sinne on yhtä suuri kuin etäisyys sieltä tänne.), refleksiivisyys (Etäisyys täältä tänne on 0.), positiivisuus (Etäisyydet eivät ole negatiivisia.) ja definiittisyys (Etäisyys täältä muualle ei ole nolla.) ovat funktiolle $d(u,v) = \|u - v\|$ ilmeisiä tosiasioita, mutta samaa ei suinkaan voi sanoa ensimmäisestä ominaisuudesta, kolmioepäyhtälöstä, joka vaatii, että kolmion yksi sivu on lyhempi kuin kahden muun summa. Itse asiassa kolmioepäyhtälö on aika hankala todistettava jo avaruudessa $\mathbb{R}^3$, ja yksinkertaisin tapa lienee jo tuossa klassisessa tapauksessa sama, jota käytämme nyt — vetoaminen CSB-epäyhtälöön.

Aloitamme erikoistapauksella, joka itsessäänkin on kiinnostava, nimittäin kolmiolla, jonka kärjet ovat 0, jokin $x \in V$ ja jokin $x + y$.

[62] Miten määrittelisit ellipsoidin sisätuloavaruudessa? Entä paraboloidin?

[63] Tervetuloa lineaarianalyysin tai kvanttimekaniikan kurssille.

[64] Jokaista funktiota $d : V \times V \to \mathbb{R}$, jolla on nämä ominaisuudet, sanotaan *metriikaksi* eli etäisyysfunktioksi. On olemassa monenlaisia muitakin metriikoita eli etäisyyskäsitteitä kuin sisätulosta normin kautta saatavat *euklidiset* metriikat, esimerkiksi etäisyys paikasta toiseen pitkin maapallon pintaa.



Väite

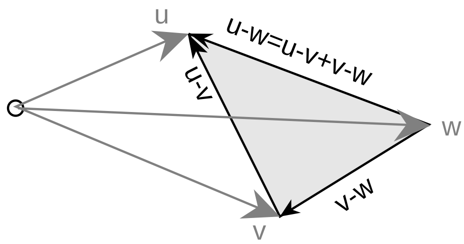
$$\|x+y\| \leq \|x\| + \|y\|$$

merkitsee samaa kuin

$$\|x+y\|^2 \leq (\|x\| + \|y\|)^2$$
$$\text{eli} \quad \|x\|^2 + \|y\|^2 + 2(x|y) \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$
$$\text{eli} \quad (x|y) \leq \|x\|\|y\|,$$

joka on CSB-epäyhtälön nojalla tosi.

Yleinen etäisyyksien kolmioepäyhtälö saadaan nyt heti: Olkoot u, v ja $w \in V$.



Juuri todistamamme epäyhtälön nojalla

$$\|u-w\| \leq \|u-v\| + \|v-w\|,$$

sillä $(u-w) = (u-v) + (v-w)$. $\square$

LAUSE 6.2.5. (NORMIN OMINAISUUDET). *Sisätuloavaruuden normilla on ominaisuudet*[65]*:*

(POS) $\|u\| \geq 0$ *kaikilla* $u \in V$,
(DEF) $\|u\| = 0 \iff u = 0$,
(P-HOM) $\|\lambda u\| = |\lambda|\|u\|$ *kaikille* $\lambda \in \mathbb{R}$ *ja* $u \in V$,
(KOL) $\|u+v\| \leq \|u\| + \|v\|$ *kaikille* $u, v \in V$,

TODISTUS. Harjoitustehtävä. $\square$

[65]Nämä ominaisuudet voi ottaa abstraktin normin määritteleviksi aksioomiksi. Abstraktilla normilla varustettua vektoriavaruutta sanotaan *normiavaruudeksi*. On olemassa paljon muitakin normiavaruuksia kuin sisätuloavaruudet. Esimerkin muusta normiavaruudesta tarjoaa $\mathbb{R}^2$ varustettuna normilla $\|(x,y)\|_1 = |x| + |y|$. Normiavaruuksia käsitellään lineaarianalyysin kurssilla.

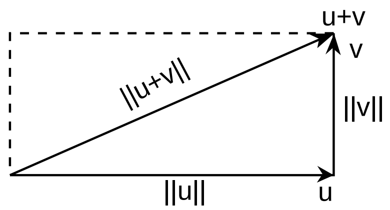
6.3. Ortogonaalisuus.

MÄÄRITELMÄ 6.3.1. Sisätuloavaruuden V vektorit u ja v ovat *ortogonaalisia* eli kohtisuorassa toisiaan vastaan, jos $(u|v) = 0$. Ortogonaalisuutta merkitään $u \perp v$.

Etäisyyksiä mittaavan geometrian peruslähtökohta on suorakulmaista kolmiota koskeva Pythagoraan lause. Se pätee jokaisessa sisätuloavaruudessa!

LAUSE 6.3.2. (PYTHAGORAS). *Sisätuloavaruuden vektoreille $u, v \in V$ on voimassa*

$$u \perp v \iff \|u+v\|^2 = \|u\|^2 + \|v\|^2.$$



TODISTUS.

$$\|u+v\|^2 = (u+v|u+v) = \|u\|^2 + \|v\|^2 + 2(u|v),$$

joten

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 \iff (u|v) = 0.$$

□

MÄÄRITELMÄ 6.3.3. Sisätuloavaruuden V osajoukko S on *ortogonaalinen joukko*, jos sen jokainen vektoripari on ortogonaalinen ts.

$$(u|v) = 0 \text{ kaikille } u, v \in S, \text{ joille } u \neq v.$$

Joukko S on *ortonormaali*, jos se on ortogonaalinen ja lisäksi

$$\|u\| = 1 \text{ kaikille } u \in S.$$

LAUSE 6.3.4. *Olkoon S ortogonaalinen joukko sisätuloavaruudessa V ja $0 \notin S$. Tällöin S on* LI. *Erityisesti ortonormaali joukko on aina* LI.

TODISTUS (HUOMATTAVA TEKNIIKKA!). Olkoon S ortogonaalinen ja $v_1, \ldots, v_k \in S$ sekä $\lambda_1, \ldots \lambda_k \in \mathbb{R}$ siten, että

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0.$$

Kerrotaan yhtälö puolittain vektorilla v_1 sisätulon mielessä ja saadaan

$$(\lambda_1 v_1 + \cdots + \lambda_k v_k \, |v_1) = (0|v_1) = 0, \text{ eli}$$
$$\lambda_1 \underbrace{(v_1|v_1)}_{\neq 0} + \lambda_2 \underbrace{(v_2|v_1)}_{0} + \cdots + \lambda_k \underbrace{(v_k|v_1)}_{0} = 0,$$

josta näkyy, että $\lambda_1 = 0$. Vastaavasti todetaan muidenkin kertoimien olevan nollia. □

6.4. Ortonormaali kanta.

Määritelmä 6.4.1. Sisätuloavaruuden V osajoukko S on V:n *ortonormaali kanta*, jos S on sekä ortonormaali joukko että V:n kanta.

Esimerkki 6.4.2. Avaruuden $\mathbb{R}^n$ standardikanta on ortonormaali standardisisätulon suhteen. Itse asiassa minkä tahansa kannan kantavektorit ovat ortonormaalit kannan itsensä määrittelemässä sisätulossa. (Vrt esimerkki 6.1.2. b)) Tätä asiaa voi ajatella niin, että ei ole olemassa vinokulmaisia koordinaatteja vaan ainoastaan kantaan sopimattomia sisätuloja[66].

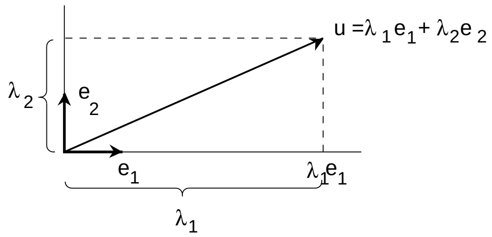
Ortonormaalissa kannassa vektorin koordinaatit on helppo laskea "projisoimalla vektori kantavektorille" tavallisesta tasogeometriasta tuttuun tapaan.

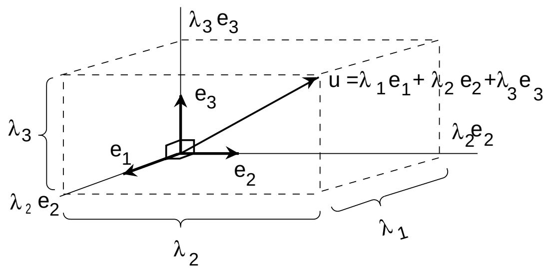
Lause 6.4.3. *Olkoon V äärellisulotteinen sisätuloavaruus,* $(e_1, e_2, \dots, e_n)$ *sen* **ortonormaali** *kanta ja $u \in V$. Tällöin esityksen*

$$u = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n$$

kertoimet — siis vektorin u koordinaatit — saadaan kaavoista

$$(*) \qquad \lambda_1 = (u|e_1), \quad \lambda_2 = (u|e_2), \dots, \lambda_n = (u|e_n).$$





Jos kannasta oletetaan ainoastaan ortogonaalisuus, mutta kantavektoreiden pituuksista ei oleteta mitään, niin pätee kuitenkin

$$(**) \qquad \lambda_1 = \frac{(u|e_1)}{\|e_1\|^2}, \quad \lambda_2 = \frac{(u|e_2)}{\|e_2\|^2}, \dots, \lambda_n = \frac{(u|e_n)}{\|e_n\|^2}.$$

Todistus. Riittää todistaa väite $(**)$. Tämä tehdään yksinkertaisesti laskemalla sisätulot. Esimerkiksi ensimmäinen saadaan näin

$$(u|e_1) = (\lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n | e_1) = \lambda_1 \underbrace{(e_1|e_1)}_{\|e_1\|^2} + \lambda_2 \underbrace{(e_2|e_1)}_{0} + \cdots + \lambda_n \underbrace{(e_n|e_1)}_{0} = \lambda_1,$$

ja muut vastaavalla tavalla. □

Ortonormaalissa kannassa annettujen vektoreiden sisätulo ja normi lasketaan samanlaisella kaavalla kuin standardisisätulo. Tämä on seuraavien lauseiden sisältö:

[66]Tunnetusti ei ole olemassa huonoa säätäkään, vaan ainoastaan epätarkoituksenmukaista vaatetusta.

LAUSE 6.4.4. *Olkoon V äärellisulotteinen sisätuloavaruus, $(e_1, e_2, \ldots, e_n)$ sen ortonormaali kanta ja $u = \sum_{i=1}^{n} \lambda_i e_i$ ja $v = \sum_{i=1}^{n} \mu_i e_i$. Tällöin on voimassa* PARSEVALIN[67] *yhtälö:*

$$
\begin{aligned}
&(u|v) = \sum_{k=1}^{n} (u|e_k)(v|e_k), \\
\textit{eli} \quad &(u|v) = \sum_{k=1}^{n} \lambda_k \mu_k.
\end{aligned}
$$

Äärellisulotteisessa avaruudessa sisätulo on siis ortonormaaliin kantaan liittyvä "pistetulo". (Vrt. 6.1.2.b ja 6.5.4.)

TODISTUS.

$$
\begin{aligned}
(u|v) &= (\sum_{i=1}^{n} \lambda_i e_i \mid \sum_{k=1}^{n} \mu_k e_k) \\
&= \sum_{i=1}^{n} \lambda_i (e_i \mid \sum_{k=1}^{n} \mu_k e_k) \\
&= \sum_{i=1}^{n} \lambda_i \left(\sum_{k=1}^{n} \mu_k \underbrace{(e_i|e_k)}_{0, \text{ kun } i \neq k} \right) \\
&= \sum_{k=1}^{n} \lambda_k \mu_1 \underbrace{(e_k|e_k)}_{1}. \qquad \square
\end{aligned}
$$

SEURAUS 6.4.5. *Edellisen lauseen oletuksin on*

$$
\|u\|^2 = \sum_{k=1}^{n} (u|e_k)^2 = \sum_{k=1}^{n} \lambda_k^2.
$$

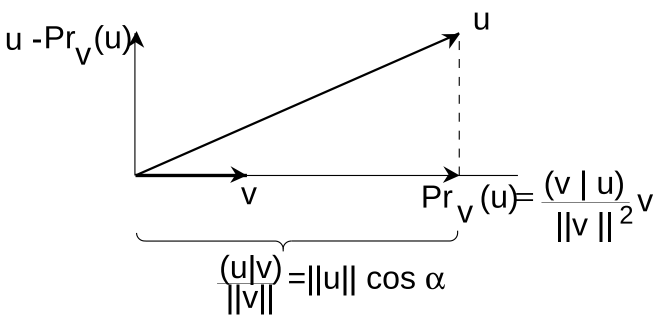
6.5. Gramin–Schmidtin ortogonalisointimenetelmä.

MÄÄRITELMÄ 6.5.1. Olkoon V sisätuloavaruus, $v \in V$, $u \in V$ ja $v \neq 0$. Vektori

$$
Pr_v(u) = \frac{(u|v)}{\|v\|^2} v
$$

on vektorin u *ortogonaalinen projektio vektorille* v.

[67] MARC–ANTOINE PARSEVAL DES CHÊNES (1755–1836), ranskalainen matemaatikko ja Napoleonin vastainen runoilija.



Erityisesti, jos $\|v\| = 1$, niin $Pr_v(u) = (u|v)v$

HUOMAUTUS 6.5.2. *Kaikilla u, v, jhoilla $v \neq 0$, on*

$$u - Pr_v(u) \perp v,$$

toisin sanoen

$$(u, v) = (Pr_v(u), v).$$

TODISTUS.

$$\begin{aligned}(u - Pr_v(u)|v) &= (u|v) - (Pr_v(u)|v) \\ &= (u|v) - (\frac{(u|v)}{\|v\|^2}v|v) \\ &= (u|v) - \frac{(u|v)}{\|v\|^2}(v|v) \\ &= (u|v) - (u|v) \\ &= 0.\end{aligned}$$

□

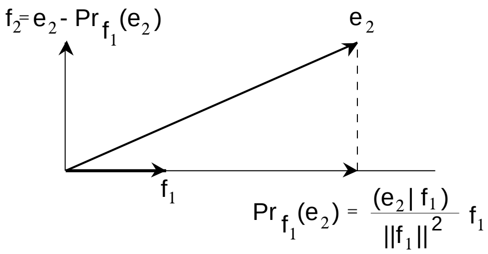
LAUSE 6.5.3. *Olkoot sisätuloavaruuden V vektorit $u_1, u_2, \ldots, u_n$ lineaarisesti riippumattomia. Tällöin on olemassa ortogonaaliset vektorit $v_1, v_2, \ldots, v_n \in V$ siten, että*

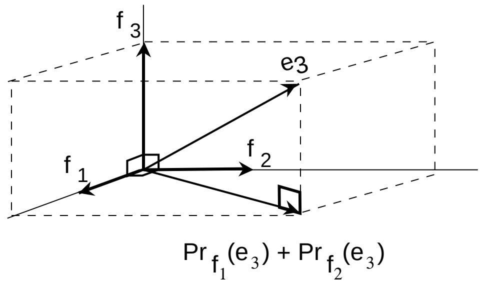
$$< v_1, v_2, \ldots, v_k >=< u_1, u_2, \ldots, u_k > \text{ kaikilla } k = 1, 2, \ldots, n.$$

Vektorit v_k saadaan esimerkiksi GRAMIN–SCHMIDTIN[68] *ortogonalisointikaavoista:*

$$\begin{cases} v_1 = u_1 \\ v_{k+1} = u_{k+1} - \sum_{i=1}^{k} Pr_{v_i}(u_{k+1}); \quad k = 1, 2, \ldots, n-1. \end{cases}$$

[68] JØRGEN GRAM (1850-1916), tanskalainen vakuutusmatemaatikko — Ortogonalisointimenetelmän keksijä lienee kuitenkin Cauchy. Saksalainen /smc Erhard Schmidt (1856–1959) yleisti menetelmän äärettönulotteiseen avaruuteen.





Todistus. Kuten itse ortogonalisointi etenee todistuskin induktioperiaatteella. Todistamme, että kaikilla $k \in \{1, \dots, n-1\}$ pätee $v_k \perp v_\ell$ jokaisella $\ell < k$, että $v_k \neq 0$ ja

$$< v_1, v_2, \dots, v_k > \subset < u_1, u_2, \dots, u_k > .$$

Tapaus $k = 1$ on kunnossa. Induktio-oletus sanokoon, että väitteet pätevät luvulle k. Osoitamme, että ne pätevät luvulle $k+1$.

Koska $Pr_{v_i}(u_{k+1}) \in < v_i >$ ja $v_i \in < u_1, \dots, u_k >$, niin $Pr_{v_i}(u_{k+1}) \in < u_1, \dots, u_k >$ ja siis

$$v_{k+1} = u_{k+1} - \sum_{i=1}^{k} Pr_{v_i}(u_{k+1}) \in < u_1, \dots, u_{k+1} > .$$

Koska alkuperäiset vektorit u_j ovat lineaarisesti riippumattomia, ei u_{k+1} ole muiden lineaarikombinaatio, vaan $u_{k+1} \notin < u_1, \dots, u_k >$, erityisesti $u_{k+1} \notin < v_1, \dots, v_k >$, ja siis $v_{k+1} \neq 0$.

Olkoon $\ell \leq k$. Todistetaan, että $v_{k+1} \perp v_\ell$.

$$\begin{aligned}(v_{k+1}|v_\ell) &= \Big(v_{k+1} - \sum_{i=1}^{k} Pr_{v_i}(u_{k+1}) \;\Big|\; v_\ell\Big) \\ &= \big(v_{k+1} \;\big|\; v_\ell\big) - \sum_{i=1}^{k} \big(Pr_{v_i}(u_{k+1}) \;\big|\; v_\ell\big).\end{aligned}$$

induktio-oletuksen mukaan $v_i \perp v_\ell$, ja siis $\big(Pr_{v_i}(u_{k+1}) \;\big|\; v_\ell\big) = 0$, kun $i \neq \ell$, joten summaan jää vain indeksiä ℓ vastaava termi ja $(v_{k+1}|v_\ell) = \big(v_{k+1} \;\big|\; v_\ell\big) - \big(Pr_{v_\ell}(u_{k+1}) \;\big|\; v_\ell\big)$. Tämä on edellisen huomautuksen 6.5.2 nojalla nolla. $\square$

Seuraus 6.5.4. *Jokaisella äärellisulotteisella sisätuloavaruudella on ortonormaali kanta.*

Todistus. Ortogonalisoimalla jokin kanta saadaan ortogonaalinen kanta. Jakamalla kukin kantavektori normillaan saadaan ortonormaali kanta. $\square$

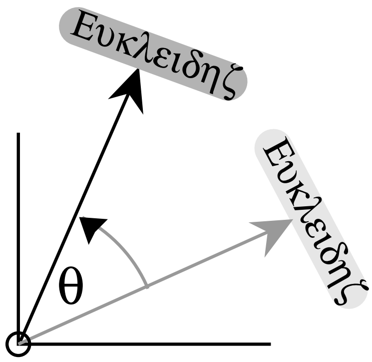
6.6. Isometriset lineaarikuvaukset eli ortogonaalikuvaukset.

Määritelmä 6.6.1. Lineaarikuvaus $L : V \to V$ on *isometrinen*, jos se säilyttää etäisyydet, eli jos kaikilla $u \in V$ on

$$\|Lu\| = \|u\|.$$

isometrinen lineaarikuvaus L on tietysti injektio. Jos V on äärellisulotteinen, niin $L : V \to V$ siis on myös bijektio ja tietysti tällöin myös L^{-1} on isometrinen lineaarikuvaus. Tällöin L on *isometrinen lineaari-isomorfismi.*[69]

Isometrinen lineaarikuvaus on yhteneväisyyskuvaus geometrian mielessä, säilyttäähän se kaikkien kuvioiden koon ja siis myös muodon. Muodon säilymisestä tärkein esimerkki on kulman säilyminen: Vektoreiden Lx ja Ly välinen kulma on sama kuin vektoreiden x ja y välinen, koska kolmion sivujen pituudet määräävät koko kolmion, erityisesti sen kulmat. Itse asiassa isometrinen lineaarikuvaus säilyttää myös vektoreiden väliset sisätulot. Erityisesti suorat kulmat säilyvät ja isomorfismissa ortonormaali kanta kuvautuu ortonormaaliksi kannaksi. Toisaalta on helppo arvata, että jokainen lineaarikuvaus, joka näin säilyttää ortonormaaliuden, on isometrinen. Siksi lineaarista isometrista isomorfismia yleensä sanotaan *ortogonaalikuvaukseksi.* Kokoamme äärellisulotteisen sisätuloavaruuden ortogonaalikuvauksen ominaisuuksia seuraavaksi luetteloksi:



Tason kierto on lineaarinen yhteneväisyyskuvaus, siis isometria

LAUSE 6.6.2. *Olkoon L lineaarikuvaus äärellisulotteiselta sisätuloavaruudelta V itselleen. Seuraavat ehdot ovat yhtäpitäviä keskenään ja sen kanssa, että L on isometrinen lineaarikuvaus eli ortogonaalikuvaus:*

(1) $\|Lu\| = \|u\| \quad \forall u \in V$
(2) $d(Lu, Lv) = d(u, v). \quad \forall u, v \in V$
(3) $(Lu|Lv) = (u|v). \quad \forall u, v \in V$
(4) *Jollakin/jokaisella ortonormaalilla kannalla $E = (e_1, \ldots, e_n)$ myös $E' = (Le_1, \ldots, Le_n)$ on ortonormaali kanta.*
(5) *Jollakin/jokaisella ortonormaalilla kannalla matriisin $C = Mat(L; E)$ sarakkeet ovat keskenään ortonormaaleita, toisin sanoen $C^T C = I_{n\times n}$.*
(6) *Jollakin/jokaisella ortonormaalilla kannalla matriisin $C = Mat(L; E)$ rivit ovat keskenään ortonormaaleita, toisin sanoen $CC^T = I_{n\times n}$.*

TODISTUS.

[69]Yleisemminkin sisätuloavaruuksien välistä lineaarista bijektiota $L : V \to W$, joka säilyttää pituudet, sanotaan sisätuloavaruuksien väliseksi *isomorfismiksi.* Kun tällainen on olemassa, avaruudet V ja W ovat *sisätuloavaruuksina isomorfiset* ja ne voi samaistaa.

" 1) $\Longleftrightarrow$ 2)" Ilmeinen tosiasia, onhan $d(Lu, Lv) = \|L(u-v)\|$.

" 1) $\Longleftrightarrow$ 3)" Normi on määritelty sisätulon avulla: $\|u\| = \sqrt{(u|u)}$, joten sisätulon säilyminen takaa tietysti normin säilymisen. Toinen suunta perustuu siihen, että myös sisätulon voi lausua normin avulla:

$$\underbrace{(u+v|u+v)}_{\|u+v\|^2} = \underbrace{(u|u)}_{\|u\|^2} + 2(u|v) + \underbrace{(v|v)}_{\|v\|^2},$$

joten

$$(u|v) = \frac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2}.$$

Jos siis $\forall u \in V: \quad \|Lu\| = \|u\|$, niin myös $\forall u, v \in V$:

$$(Lu|Lv) = \frac{\|L(u+v)\|^2 - \|Lu\|^2 - \|Lv\|^2}{2} = \frac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2} = (u|v).$$

" 3) $\Longrightarrow$ 4)" Olkoon $E = (e_1, \ldots, e_n)$ ortonormaali kanta, ts. $(e_i|e_j) = 1$, kun $i = j$ ja 0 muuten. Jos L säilyttää sisätulot, niin $(Le_i|Le_j) = (e_i|e_j) = 1$, kun $i = j$ ja 0 muuten, eli $E' = (Le_1, \ldots, Le_n)$ on ortonormaali kanta.

" 4) $\Longleftrightarrow$ 5)" Olkoon $E = (e_1, \ldots, e_n)$ ortonormaali kanta ja $C = Mat(L; E)$. Matriisin C sarakkeet ovat vektorit $Le_1, \ldots, Le_n$. Näiden keskinäiset sisätulot ovat toisaalta juuri matriisin $C^T C$ alkiot:

$$(C^T C)_{ij} = (Le_i|Le_j).$$

Toisaalta vektorit $(Le_1, \ldots, Le_n)$ ovat ortonormaalit täsmälleen silloin, kun

$$(Le_i|Le_j) = \begin{cases} 1, \text{ kun } i = j \\ 0, \text{ kun } i \neq j \end{cases} = (I_{n \times n})_{ij}.$$

" 5) $\Longleftrightarrow$ 6)" C on neliömatriisi. Jos $C^T C = I_{n\times n}$, tai $CC^T = I_{n\times n}$, niin C on välttämättä kääntyvä (Lause 3.1.6) ja C^T ja C ovat toistensa käänteismatriisit.

" 4) $\Longrightarrow$ 1)" On vielä näytettävä, että kannan ortonormaalisuuden säilyttävä lineaarikuvaus on isometrinen. Tämä seuraa suoraan lauseesta (6.4.4), jonka mukaan mikä tahansa sisätulo on ortonormaaliin kantaansa liittyvä sisätulo, ja silloin normi lasketaan kuten standardiavaruudessa R^n. Jos siis on olemassa ortonormaali kanta $E = (e_1, \ldots, e_n)$, jolla myös $E' = (Le_1, \ldots, Le_n)$ on ortonormaali kanta, niin vektorin $x = \sum_{j=1}^n \lambda_j e_j$ kuvan normin neliö on

$$\|Lx\|^2 = \|L(\sum_{j=1}^n \lambda_j e_j)\|^2 = \|\sum_{j=1}^n \lambda_j Le_j\|^2 = \sum_{j=1}^n \lambda_j^2 = \|x\|^2.$$

$\square$

Määritelmä 6.6.3. Neliömatriisi C on *ortogonaalinen*, jos C on kääntyvä ja

$$C^{-1} = C^T.$$

Lause 6.6.2 sanoo, että lineaarikuvaus äärellisulotteiselta sisätuloavaruudelta itselleen on isometrinen isomorfismi tasan silloin, kun sen matriisi $C_{n\times n} = Mat(L; E)$ ortonormaalissa kannassa E on ortogonaalinen.

Avaruuden $\mathbb{R}^n$ ortogonaalikuvausten joukkoa merkitään useimmiten symbolilla $\mathcal{O}(n)$.[70]

6.7. Symmetrisen lineaarikuvauksen diagonalisointi.

Lineaarikuvausta, jonka matriisi ortonormaalissa kannassa on symmetrinen, sanotaan symmetriseksi lineaarikuvaukseksi. Ominaisarvoista puhuessamme huomasimme, että symmetrisen $2\times 2-$matriisin voi diagonalisoida — vieläpä ortogonaalisella kannanvaihdolla — ja että tämän avulla pystyttiin löytämään kartioleikkauksen pääakselikoordinaatisto. Toisen asteen pintojen ja niiden korkeampiulotteisten vastineiden tutkiminen samalla tavalla on mahdollista, mikäli onnistumme diagonalisoimaan minkä tahansa symmetrisen matriisin ortogonaalisesti.

Määritelmä 6.7.1. Sisätuloavaruudessa V määritelty lineaarikuvaus $L : V \to V$ on *symmetrinen*[71], jos

$$(Lu|v) = (u|Lv) \quad \text{kaikille } u \text{ ja } v \in V.$$

Lause 6.7.2. *Olkoon V sisätuloavaruus, $E = (e_1, e_2, \ldots, e_n)$ sen* **ortonormaali** *kanta ja $L : V \to V$ lineaarikuvaus. Tällöin L on symmetrinen, jos ja vain jos vastaava matriisi $A = Mat(L; E)$ on symmetrinen eli $A = A^T$.*

Todistus. Lineaarikuvauksen L matriisin alkiot ovat kantavektoreiden kuvien Le_j koordinaatit, sisätuloavaruudessa siis luvut $a_{ij} = (e_i|Le_j)$. Jos L on symmetrinen, on siis $a_{ij} = (e_i|Le_j) = (Le_i|e_j) = (e_j|Le_i) = a_{ji}$, matriisi A on symmetrinen.

[70] Koska ortogonaalikuvausten yhdistetty kuvaus ja käänteiskuvauskin ovat ortogonaalikuvauksia, on kuvausten yhdistäminen laskutoimitus joukossa $\mathcal{O}(n)$, ja ne tekevät siitä ns. ortogonaaliryhmän. Ryhmiä käsitellään algebran kurssilla.

[71] Kompleksisten sisätulojen teoriassa esiintyy samassa roolissa lähisukuinen käsite. Kuvaus on *hermiittinen*, jos $(Lu|v) = (u|Lv)$:n liittoluku kaikille u ja $v \in V$..

Jos taas oletetaan matriisin symmetria, niin voidaan laskea kannassa E

$$\begin{aligned}
(Lu|v) &= (L\sum_{i=1}^{n} u_i e_i \mid \sum_{j=1}^{n} v_j e_j) \\
&= \sum_{i=1}^{n} u_i \sum_{j=1}^{n} v_j (Le_i|e_j) \\
&= \sum_{i=1}^{n} u_i \sum_{j=1}^{n} v_j a_{ij} \\
&= \sum_{i=1}^{n} u_i \sum_{j=1}^{n} v_j a_{ji} \\
&= \sum_{j=1}^{n} v_j \sum_{i=1}^{n} u_i (Le_j|e_i) \\
&= (Lv|u).
\end{aligned}$$

□

Juuri todistetun lauseen merkitys on seuraava: Matriisin symmetria ortonormaalissa kannassa on vastaavan lineaarikuvauksen ominaisuus, joka liittyy sisätuloon. Lauseesta 6.7.2. seuraa, että sisätuloavaruudessa V lineaarikuvauksen $L : V \to V$ matriisi on joko symmetrinen jokaisessa ortonormaalikannassa tai ei yhdessäkään sen mukaan, onko itse L symmetrinen lineaarikuvaus. Tämä merkitsee, että matriisin symmetria säilyy vaihdettaessa ortonormaalista kannasta toiseen, ts.

Lause 6.7.3. *$C^T AC$ on symmetrinen, jos A on symmetrinen ja C on ortogonaalinen.*

Koska kaikki diagonaalimatriisit ovat symmetrisiä, on nyt selvää, että ortogonaalinen kannanvaihto ei voi tehdä diagonaalimatriisista muuta kuin symmetrisen. Tämä merkitsee, että ainoastaan symmetrisen matriisin voi mahdollisesti diagonalisoida ortogonaalimatriisilla C. Ortonormaalissa kannassa diagonaalinen matriisi on toisaalta geometrisesti erittäin selkeä objekti — samanlainen kuin diagonaalimatriisi standardikannssa. Siksi on kiinnostavaa, että **kaikki symmetriset matriisit ovat tätä tyyppiä**, siis ortogonaalisella matriisilla diagonalisoituvia.

Lause 6.7.4. (Spektraalilause) *Olkoon V äärellisulotteinen sisätuloavaruus ja $L : V \to V$ symmetrinen lineaarikuvaus. Tällöin L:n ominaisvektoreista voidaan muodostaa V:n ortonormaali kanta E.*

Todistus. Todistus edellyttää kompleksilukujen käyttöä ja esitetään siksi vasta luvussa 7. □

Seuraus 6.7.5. *Symmetrinen matriisi on diagonalisoituva ja ominaiskanta voidaan valita ortonormaaliksi.*

SEURAUS 6.7.6. *Symmetrisen lineaarikuvauksen eri ominaisarvoihin liittyvät ominaisvektorit ovat kohtisuorassa toisiaan vastaan.*

TODISTUS. Tämä seuraa tietysti suoraan spektraalilauseesta, mutta esitämme myös suoran todistuksen. Itse asiassa tätä tietoa nimittäin tarvitaan spektraalilauseen todistuksessa.

Jos $Lu = \lambda u$ ja $Lv = \mu v$, ja toinen kerroin — vaikkapa μ — eroaa nollasta, niin

$$(Lu|Lv) = \begin{cases} (\lambda u|\mu v) = \lambda\mu(u|v) \\ (u|L(Lv)) = (u|L(\mu v)) = (u|\mu\mu v) = \mu^2(u|v). \end{cases}$$

Näin ollen $\frac{\lambda}{\mu}(u|v) = (u|v)$, joten $(u|v) = 0$, koska $\frac{\lambda}{\mu} \neq 1$. $\square$

6.8 Neliömuoto.

Olemme selvittäneet luvussa 5, että kaksiulotteisessa avaruudessa symmetriset matriisit kelpasivat kartioleikkausten tunnistamiseen. Yleistämme tilanteen $n-$ulotteiseksi: Olkoon V sisätuloavaruus, $n = \dim V$ ja $E = (e_1, e_2, \ldots, e_n)$ avaruuden V ortonormaali kanta.

MÄÄRITELMÄ 6.8.1. Symmetriseen[72] *lineaarikuvaukseen* $L : V \to V$ *liittyvä neliömuoto* eli *kvadraattinen muoto* on kuvaus $Q : V \to \mathbb{R}$;

$$Q(x) = (Lx|x).$$

Neliömuoto Q on helppo laskea kannassa:

LAUSE 6.8.2. *Olkoot L ja Q kuten edellä ja $A = (a_{ij}) = Mat(L;E)$. Tällöin*

$$(*) \qquad Q(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \quad \textit{kaikille } x = \sum_{i=1}^{n} x_ie_i \in V.$$

TODISTUS. Harjoitustehtävä $\square$

MÄÄRITELMÄ 6.8.3. Jos on annettu symmetrinen matriisi $A = A_{n\times n} = (a_{ij})$, niin edellisen lauseen lauseke $(*)$ määrittelee sitä kannassa E vastaavaan lineaarikuvaukseen liittyvän neliömuodon, jota sanotaan *matriisiin $A_{n\times n}$ liittyväksi neliömuodoksi* $Q : \mathbb{R}^n \to \mathbb{R}$;

$$Q(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \quad \text{kaikille} \quad x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

Matriisisulkeita vaille tämä on sama kuin matriisitulo

$$= x^TAx, \qquad \text{missä } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n.$$

[72]Määritelmä voidaan asetaa, vaikka L ei olisi symmetrinen. Tästä ei kuitenkaan olisi juuri etua.

Diagonaalimatriisiin $\begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$ liittyvää neliömuotoa

$$Q(x) = \sum_{i=1}^{n} \lambda_i x_i^2$$

sanotaan *diagonaaliseksi neliömuodoksi.*

LAUSE 6.8.4. *Olkoon matriisi $A_{n\times n}$ symmetrinen ja matriisi $C_{n\times n}$ ortogonaalinen siten, että $C^{-1}AC = D =$ diagonaalimatriisi*

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

ts. D:n lävistäjällä on A:n ominaisarvot. Olkoon $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ ja olkoon

$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = C^{-1}x$. Tällöin

$$x^T Ax = y^T Dy = \sum_{k=i}^{n} \lambda_k y_k^2.$$

TODISTUS.

$$x^T Ax = x^T CC^{-1}ACC^{-1}x = (C^{-1}x)^T D(C^{-1}x) = yDy^T.$$

□

Edellinen lause merkitsee, että neliömuoto diagonalisoituu toivotulla tavalla:

MÄÄRITELMÄ 6.8.5. Neliömuoto

$$Q(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$$

on

- *positiivisesti definiitti*, jos $Q(x) > 0$ kaikilla $x \in \mathbb{R}^n \setminus \{0\}$,
- *negatiivisesti definiitti*, jos $Q(x) < 0$ kaikilla $x \in \mathbb{R}^n \setminus \{0\}$,
- *positiivisesti semidefiniitti*, jos $Q(x) \geq 0$ kaikilla $x \in \mathbb{R}^n$ ja $Q(x) = 0$ jollakin $x \neq 0$,
- *negatiivisesti semidefiniitti*, jos $Q(x) \leq 0$ kaikilla $x \in \mathbb{R}^n$ ja $Q(x) = 0$ jollakin $x \neq 0$,
- *indefiniitti*, jos on olemassa $x \in \mathbb{R}^n$ siten, että $Q(x) > 0$ ja $y \in \mathbb{R}^n$ siten, että $Q(y) < 0$.

Lause 6.8.6. *Olkoon $Q(x)$ neliömuoto ja A vastaava symmetrinen matriisi. Tällöin*

- *$Q(x)$ on positiivisesti definiitti jos ja vain jos A:n kaikki ominaisarvot ovat positiivisia.*
- *$Q(x)$ on negatiivisesti definiitti jos ja vain jos A:n kaikki ominaisarvot ovat negatiivisia.*
- *$Q(x)$ on positiivisesti semidefiniitti jos ja vain jos A:n kaikki ominaisarvot ovat ei–negatiivisia ja joku ominaisarvo on nolla.*
- *$Q(x)$ on negatiivisesti semidefiniitti jos ja vain jos A:n kaikki ominaisarvot ovat ei–positiivisia ja joku ominaisarvo on nolla.*
- *$Q(x)$ on indefiniitti jos ja vain jos A:lla on sekä positiivisia että negatiivisia ominaisarvoja.*

Todistus. Edellisen lauseen mukaan neliömuoto diagonalisoituu säilyttäen ominaisarvonsa. Toisaalta diagonaaliselle neliömuodolle lauseen väite on ilmeinen. □

Neliömuodon tyypin määrääminen lauseen 6.8.5 keinoin edellyttää ominaisarvojen määräämistä. On kuitenkin olemassa helpompikin keino:

Lause 6.8.7. *Olkoon $Q(x)$ neliömuoto, A vastaava symmetrinen matriisi ja merkitään*

$$A_k = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}; \quad k = 1, 2, \dots, n$$

toisin sanoen

$$A_1 = a_{11}, \quad A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, A_n = \det A.$$

Tällöin

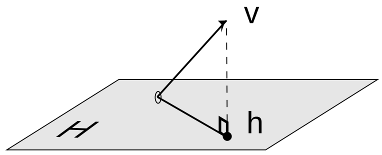
- *$Q(x)$ on positiivisesti definiitti jos ja vain jos $A_k > 0$ kaikilla k,*
- *$Q(x)$ on negatiivisesti definiitti jos ja vain jos $(-1)^k A_k > 0$ kaikilla k (ts. $A_1 < 0$, $A_2 > 0$,$A_3 < 0$ jne.).*

Perustelu. Sivuutetaan

6.9. Ääriarvotehtäviä sisätuloavaruudessa. Pienin neliösumma.
Äärellisulotteisen sisätuloavaruuden V geometrialle ominainen piirre on ortonormaalin kannan olemassaolo. Kannan avulla on mahdollista ratkaista mm. seuraavat kaksi toisilleen lähisukuista ääriarvotehtävää:

Tehtävä 6.9.1.
Annettuna on aliavaruus $H \subset V$ ja piste $v \in V$. Tehtävänä on löytää aliavaruudesta H piste, joka on kaikkein lähimpänä pistettä v, ts. ratkaista joukossa H ääriarvotehtävä

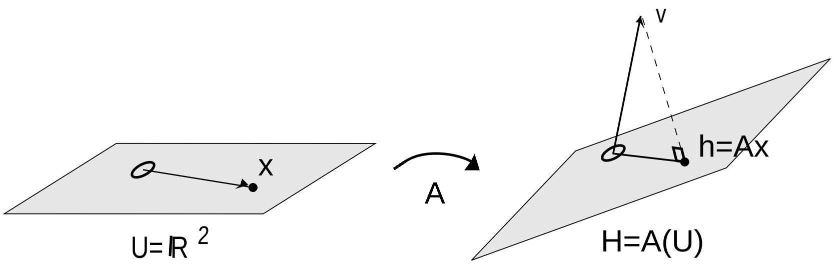
$$\|v - b\| = min!$$



TEHTÄVÄ 6.9.2.
Annettuna on piste $v \in V$ ja lineaarikuvaus $A : U \to V$. Tehtävänä on löytää avaruudesta U pisteet $\hat{x} \in U$, joiden kuva on kaikkein lähimpänä pistettä v, ts. ratkaista joukossa U ääriarvotehtävä

$$\|Ax - v\| = min!$$

Tehtävänä on siis ratkaista yhtälö $Ax = v$ likimain, kun tarkkaa ratkaisua ei ole olemassa. Tällaisen tilanteen eteen voi joutua isomman tehtävän osana, kun v on esimerkiksi pyöristys- tai mittausvirheiden vuoksi ajautunut hieman syrjään joukosta $A(U)$, jossa sen oikeastaan pitäisi olla.



Näitä ongelmia sanotaan joskus pienimmän neliösumman (PNS) ongelmiksi, koska yleensä sisätuloavaruutena on tavallinen $\mathbb{R}^n$, ja minimoitavana on oikeastaan etäisyyden neliö

$$\|v - b\|^2 = \sum (v_i - b_i)^2.$$

Todistamme, että molemmilla ongelmilla on ratkaisu ja esitämme keinoja sen löytämiseksi. Koska äärellisulotteinen sisätuloavaruus on isomorfinen avaruuden $\mathbb{R}^n$ kanssa, oletamme seuraavassa, että kaikki avaruudet ovat tätä tyyppiä. Matriisit samaistamme lineaarikuvauksiin standardikannan avulla.

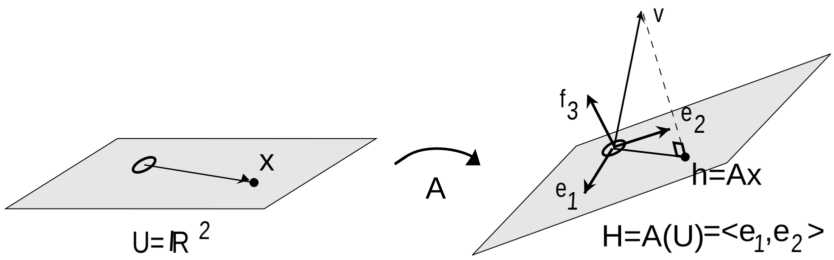
LAUSE 6.9.3. *Tehtävän 6.9.1 tilanteessa olkoon $E = (e_1, \ldots, e_m)$ jokin aliavaruuden H* **ortonormaali** *kanta. Tällöin vektori*

$$(*) \qquad \hat{b} = \sum_{i=1}^{m} (v|e_i)e_i.$$

on tehtävän 6.9.1 ainoa ratkaisu — siis etsitty lähin piste — ja sillä on seuraavat suorakulmaiselle projektiolle tyypilliset ominaisuudet:

a) $(v - \hat{b}) \perp H$, *ts.* $(v - \hat{b}) \perp h \quad \forall h \in H$,
b) $\hat{b}$ *on ainoa aliavaruuden H piste, jolla on ominaisuus a)*

TODISTUS. Täydennetään E koko avaruuden ortonormaaliksi kannaksi $\hat{E} = (e_1, \ldots, e_m, f_{m+1}, \ldots, f_{m+k})$ lisäämäällä siihen maksimaalinen määrä LI vektoreita ja ortonormeeraamalla saatu kanta.



Tässä kannassa lausuttuna on

$$v = \underbrace{\sum_{i=1}^{m}(v|e_i)e_i}_{\hat{b}} + \underbrace{\sum_{i=1}^{k}(v|f_{m+i})f_{m+i}}_{v-\hat{b}},$$

joten $(v-\hat{b}) \perp H$ ja $\|v-\hat{b}\|^2 = \sum_{i=1}^{k}(v|f_{m+i})^2$. Jos nyt $h = \sum_{i=1}^{m} h_i e_i$ on jokin toinen H:n vektori, niin

$$v - h = \underbrace{\sum_{i=1}^{m}(v-h|e_i)e_i}_{\hat{b}-h} + \underbrace{\sum_{i=1}^{k}(v|f_{m+i})e_i}_{v-\hat{b}},$$

joten

$$\|v-h\|^2 = \sum_{i=1}^{m}(v-h|e_i)^2 + \sum_{i=1}^{k}(v|f_{m+i})^2 > \sum_{i=1}^{k}(v|f_{m+i})^2 = \|v-\hat{b}\|^2$$

ja

$$(v-h|\hat{b}-h) = (v|\hat{b})-(v|h)-(h|\hat{b})+(h|h) = (\hat{b}|\hat{b})-(\hat{b}|h)-(h|\hat{b})+(h|h) = \|\hat{b}-h\|^2 \neq 0$$

Siis $\hat{b}$ on lähin piste ja ainoa normaalin kantapiste. □

Edellisestä lauseesta seuraa, että tehtävällä 6.9.1 on tasan yksi ratkaisupiste, ja että siis kaavan $(*)$ vektori $\hat{b}$ — v:n *ortoprojektio* eli *komponentti* aliavaruudella H — ei riipu aliavaruuden H ortonormaalin kannan E valinnasta.

Pisteen $\hat{b}$ löytäminen näillä ohjeilla on kuitenkin laskuteknisesti hankalaa, ellei avaruudessa H ole valmiiksi annettuna ortonormaalia kantaa, sillä kannan ortonormeeraus on pitkällinen toimenpide. Tehtävän 6.9.2 ratkaisun yhteydessä keksimme elegantimman ratkaisun myös ensimmäiselle tehtävälle.

Tarkastelu 6.9.4. Toisen ongelman voi tietysti periaatteessa ratkaista ensimmäisen avulla: Etsitään kuva-avaruudesta $A(U)$ pistettä v lähin kohta $\hat{b}$ ja ratkaistaan sitten esimerkiksi Gaussin ja Jordanin menetelmällä yhtälö $Ax = \hat{b}$. Totesimme kuitenkin jo, että ortonormeeraukseen perustuva lasku on raskas. Onneksi se voidaan välttää seuraavalla tavalla: Tehtävänä on löytää $\hat{x}$, jolle

$(*)$ $$A\hat{x} = \hat{b}, \text{ ja}$$

$(**)$ $$(v-\hat{b}) \perp A(U)$$

Koska $A(U)$ on A:n sarakevektoreiden virittämä avaruus, ehto $(**)$ merkitsee sitä, että $v - \hat{b}$ on jokaista A:n saraketta eli A^T:n riviä vastaan kohtisuorassa, toisin sanoen

$$A^T(v - \hat{b}) = 0.$$

Muistaen ehdon $(*)$ saamme tästä

$$(\bullet) \qquad A^T(v - A\hat{x}) = 0, \text{ eli } A^T A\hat{x} = A^T v.$$

Tämän ratkaisemiseksi tarvitaan vain yksi Gaussin–Jordanin menettely!

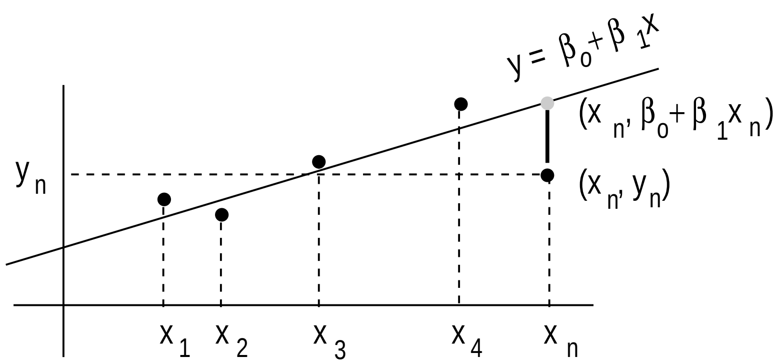
Lause 6.9.5. *Yhtälön* $(\bullet)$ *ratkaisut ovat samat kuin tehtävän 6.9.2 ratkaisut.*

Todistus. Olemme juuri päätelleet, että jokainen tehtävän 6.9.2 ratkaisu toteuttaa myös yhtälön $(\bullet)$. Toisaalta yhtälöstä $(\bullet)$ seuraa yhtälö $(**)$, kun muistetaan merkintä $(*)$. Voimme siis päätellä myös takaperin ja huomata, että alkuperäinen tehtävä on ratkaistu. □

Huomaamme lopuksi, että tehtävän 6.9.2 ratkaisu on yksikäsitteinen täsmälleen silloin, kun A on injektio, mutta injektiivisyyden testaamiseen ei ole käytettävissä determinanttia ellei A satu olemaan neliömatriisi. Ei kuitenkaan ole vaikeaa todistaa, että A on injektio tasan silloin, kun neliömatriisi $A^T A$ on kääntyvä, sillä $A^T(Ax) = 0$ tasan silloin, kun Ax on kohtisuorassa A:n sarakkeita vastaan, mutta Ax on sarakkeiden lineaarikombinaatio.

Tehtävä 6.9.1 on tietysti[73]erikoistapaus tehtävästä 6.9.2. Pitää vain valita A:ksi lineaarikuvaus $x \mapsto x$, joka upottaa aliavaruuden $U = H$ avaruuteen V. Jos $F = (v_1, \ldots, v_n)$ on H:n kanta, niin A:n matriisissa $Mat(A; F, E)$ on sarakkeet $v_1, \ldots, v_n$. Ratkaisu löytyy siis soveltamalla tehtävän 6.9.2 ratkaisun ideaa tähän matriisiin ja yhtälön $A^T\hat{b} = A^T v$ ratkaisu on etsitty $\hat{b}$ lausuttuna kannassa F.

Esimerkki 6.9.6. Näytämme miten tilastotieteestä tuttu pienimmän neliösumman suoran eli regressiosuoran löytäminen palautuu tehtävään 6.9.2.



On annettu lukuparit $(x_1, y_1), \ldots, (x_n, y_n)$. Käytännössä ne on yleensä saatu mittaamalla jotakin suuretta y parametriarvoilla x_j. Haetaan sellaista suoraa

[73]Muuta tekstin järjestystä

$y(x) = \beta_0 + \beta_1(x)$, joka kulkee y-suunnassa mahdollisimman läheltä annettuja datapisteitä, tarkemmin sanoen suoraa, jolla *neliösumma*

$$\sum_{i=1}^{n}(\beta_0 + \beta_1 x_i - y_i)^2$$

on mahdollisimman pieni. Jos pisteet $(x_1, y_1), \ldots, (x_n, y_n)$ sattuvat todella olemaan samalla suoralla, niin neliösumma saadaan nollaksi valitsemalla luvut β_0 ja β_1 siten, että

$$\beta_1 x_i = y_i - \beta_0 \quad \forall i,$$

toisin sanoen $A\beta = y$, missä on merkitty

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Pienimmän neliösumman suoran löytäminen on selvästikin sama asia kuin yhtälöryhmän

$$A\beta = y$$

likimääräinen ratkaiseminen tehtävän 6.9.2 mielessä.

7. Kompleksista lineaarialgebraa

Tutkiessamme symmetrisen $n \times n-$matriisin ominaisarvo-ongelmaa jouduimme ratkaisemaan n:nnen asteen algebrallista yhtälöä eli etsimään polynomin nollakohtia. Nollakohtia on *algebran peruslauseen* mukaan olemassa kertaluvut huomioiden tasan n kappaletta — mutta ne ovat yleensä kompleksilukuja. Tämä antaa aiheen miettiä, voisiko kompleksiluvun tulkita matriisin ominaisarvoksi. Tähän onkin useita mahdollisuuksia. Selvin tapa lienee hyväksyä kompleksiluvut "luvuiksi" koko lineaarialgebrassa alusta asti. Tässä liitteessä kuvaillaan lyhyesti kompleksista lineaarialgebraa ja todistetaan sen avulla symmetrisen matriisin spektraalilause eli diagonalisoituminen tavallisessa reaalisessa mielessä.

7.1. Kompleksiset vektoriavaruudet.

Kompleksiluvut muodostavat ns. *kunnan*, toisin sanoen niillä on samat laskulait kuin reaaliluvuilla tai rationaaliluvuilla: Yhteen- ja kertolasku ovat vaihdannaisia ja liitännäisiä, osittelulait vallitsevat, nolla ja ykkönen ovat olemassa ja jokaisella luvulla on vasta- ja käänteisluku, paitsi että 0 ei käänny. Nämä ominaisuudet riittävät lineaarialgebran perusteiden läpiviemiseen. Algebralliselta kannalta olennainen ero tulee sopivalla hetkellä, kun joudutaan epälineaarisiin yhtälöihin.

Voimme jäljentää tämän kirjan luvut 1-4 sellaisinaan korvaamalla reaaliluvut kompleksiluvuilla ja tällöin kaikki määritelmät ja lauseet todistuksineen säilyvät mielekkäinä ja tosina. Koska todistukset siis jo on kirjoitettu, voimme tyytyä kirjaamaan joitakin havaintoja:

Kompleksinen vektoriavaruus eli $\mathbb{C}-$vektoriavaruus on joukko V, jossa on yhteenlaskun lisäksi toisena laskutoimituksena kompleksiluvulla kertominen

$$\cdot : \mathbb{C} \times V \to V$$

ja jossa kaikki vektoriavaruuden aksioomat pätevät. Erityisesti $\mathbb{C}^n$ on $\mathbb{C}-$vektoriavaruus. Esimerkkinä laskutoimituksista lasketaan vaikkapa luvun $3 - 2i \in \mathbb{C}$ ja vektorin $\begin{pmatrix} 2+i \\ i \end{pmatrix} \in \mathbb{C}^2$ tulo:

$$(3-2i)\begin{pmatrix} 2+i \\ i \end{pmatrix} = \begin{pmatrix} (3-2i)(2+i) \\ (3-2i)i \end{pmatrix} = \begin{pmatrix} 8-i \\ 2+3i \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix} + i\begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Tulos on sievennetty "reaali- ja imaginaariosaksi". Näin voi ilmeisesti tehdä avaruudessa $\mathbb{C}^n$, mutta on syytä panna merkille, että kompleksisen vektorin jako koordinaattien "reaali- ja imaginaariosiin" riippuu käytetystä kannasta eikä siis onnistu abstraktissa avaruudessa. Lukija vakuuttukoon tästä esimerkiksi korvaamalla $\mathbb{C}^2$:n standardikantavekorit $(1,0)$ ja $(0,1)$ kantavektoreilla $(i,0)$ ja $(0,i)$.

$\mathbb{C}-$vektoriavaruudessa voi määritellä aliavaruuden ja lineaarikombinaation käsitteet kuten reaalisessakin, mutta nyt sallitaan tietysti kompleksiluvut kertoimina ja määritellään siis esimerkiksi lineaarinen verho asettamalla

$$< S > = < S >_{\mathbb{C}} = \{\lambda_1 x_1 + \cdots + \lambda_k x_k \bigm| k \in \mathbb{N}, \lambda_j \in \mathbb{C}, x_j \in S\}.$$

Lineaarinen riippumattomuus ja riippuvuus määritellään tuttuun tapaan ja saadaan dimensiokäsite — luonnollisesti kompleksisessa mielessä. On hyvä kiinnittää huomiota siihen, että kompleksilukujen kunta $\mathbb{C}$ on yksiulotteinen kompleksinen, mutta kaksiulotteinen reaalinen vektoriavaruus ja että jokainen $\mathbb{C}^n$ — itse asiassa jokainen $n-$ulotteinen kompleksinen vektoriavaruus — on samalla $2n-$ulotteinen reaalinen vektoriavaruus.

7.2. Kompleksiset lineaarikuvaukset ja matriisit.

Kompleksiselta lineaarikuvaukselta vaaditaan tietysti homogeenisuus myös kompleksikertoimien suhteen, siis

$$L(\lambda v) = \lambda L v \quad \forall \lambda \in \mathbb{C},\, v \in V.$$

Kompleksisilla matriiseilla lasketaan kuten reaalisilla. Myös matriisialkiot ovat kompleksilukuja ja matriisin voi siten jakaa reaali- ja imaginaariosaan, kuten edellä jaoimme vektorin, siis esimerkiksi

$$\begin{bmatrix} 1 & 2+2i \\ 3+3i & 4+4i \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 2i \\ 3i & 4i \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + i \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix}.$$

Näin saatava lineaarikuvauksen näennäinen jako reaali- ja imaginaariosaan riippuu sekin kannasta.

Lineaarisia yhtälöryhmiä, ytimiä ja kuva-avaruuksia koskevat tarkastelut kelpaavat sellaisinaan kompleksiseenkin tilanteeseen.

7.3. Kompleksiset ominaisarvot.

Ominaisarvo määritellään kompleksisessa vektoriavaruudessa kuten reaalisessa. Algebran peruslauseesta seuraa, että jokaisella kompleksisella matriisilla on ainakin yksi kompleksinen ominaisarvo, onhan sen karakteristisella polynomilla nollakohta kompleksilukujen joukossa.

Esimerkki

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \lambda = i, \quad x = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

osoittaa, että reaalisella matriisilla voi olla kompleksinen ominaisarvo ja kompleksinen ominaisvektori. Mutta reaalisen matriisin kompleksiseen ominaisarvoon ei voi liittyä reaalinen ominaisvektori. Jos nimittäin A ja x ovat reaalisia, niin Ax on tietysti reaalinen, mutta selvästikään $(a+ib)x$ ei voi olla reaalinen ellei b tai x ole 0.

7.4. Kompleksiset sisätuloavaruudet.

Kompleksiluvun *itseisarvo* eli *moduli* määritellään tunnetusti kaavalla

$$\|z\| = \|a+ib\| = \sqrt{a^2+b^2} = \sqrt{(z\overline{z})},$$

missä $\overline{z}$ on kompleksiluvun $z = a+ib$ *liittoluku* eli *kompleksikonjugaatti* $a - ib$. Kompleksiluvun neliö ei yleensä ole reaalinen, saati positiivinen. Kompleksisten vektoreiden $x = (x_1, \ldots, x_n) = (a_1 + ib_1, \ldots, a_n + ib_n)$ ja $y = (y_1, \ldots, y_n) = (c_1 + i\,d_1, \ldots, c_n + i\,d_n)$ standardisisätulo määritellään kaavalla

$$(x|y) = \sum_{j=1}^{n} x_j \overline{y}_j,$$

auki kirjoitettuna siis

$$\begin{aligned}(x|y) &= \big((a_1 + ib_1, \dots, a_n + ib_n) \bigm| (c_1 + id_1, \dots, c_n + id_n)\big) \\ &= (a_1 + ib_1)(c_1 - id_1) + \dots + (a_n + ib_n)(c_n - id_n) \\ &= a_1c_1 + b_1d_1 + \dots + a_nc_n + b_nd_n + i(b_1c_1 - a_1d_1 + \dots + b_nc_n - a_nd_n).\end{aligned}$$

Liittoluvun käytön tarkoituksena on huolehtia siitä, että vektorin sisätulo itsensä kanssa $(x|x)$ on myös kompleksisessa tapauksessa positiivinen, paitsi nollavektorin tapauksessa 0. Näin kompleksisen vektorin normi $\|x\| = \sqrt{(x|x)}$ on reaaliluku, vieläpä ei negatiivinen.

Nämä kompleksisen standardisisätulon ominaisuudet otetaan huomioon asetettaessa sisätuloavaruuden aksioomia kompleksisen vektoriavaruuden pohjalta. Sisätulon aksioomat ovat muuten samat kuin reaalisessa tapauksessa, mutta sisätulon sallitaan olevan kompleksiluku ja symmetria-aksiooma SYM korvataan *Hermiten aksioomalla*

(HER) $:(x|y) = \overline{(y|x)} \quad \forall x, y \in V.$

Todisteltaessa erilaisia sisätulon ominaisuuksia on seurattava, minne tästä aiheutuu eroja reaaliseen tapaukseen verrattuna. Huomattava ero on ainakin siinä, että kompleksinen sisätulo **ei ole bilineaarinen**, vaan *seskilineaarinen*, toisin sanoen lineaarinen ensimmäisen tekijän suhteen ja *antilineaarinen* jälkimmäisen suhteen:

$$\begin{aligned}(\lambda x + \mu y|z) &= \lambda(x|z) + \mu(y|z) \quad \forall x, y, z \in V,\ \lambda, \mu \in \mathbb{C} \\ (z|\lambda x + \mu y) &= \overline{\lambda}(z|x) + \overline{\mu}(z|y) \quad \forall x, y, z \in V,\ \lambda, \mu \in \mathbb{C}.\end{aligned}$$

Tästä ei aiheudu kovin vakavia seurauksia teorialle. Vektorin pituuden voi taas määritellä asettamalla

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

ja Cauchyn, Schwarzin epäyhtälö ja kolmioepäyhtäkö pätevät ja todistuvat samaan tapaan kuin reaalisessa sisätuloavaruudessakin. Ortogonaalisuus merkitsee taas sitä, että $(x|y) = 0$, mutta muuten kahden vektorin välisen kulman määrittely jätetään tekemättä, koska tulisi otettavaksi arcus kosini kompleksiluvusta. Ortogonaalisuuden tultua määritellyksi saamme kuitenkin ortonormaalin kannan olemassaolon todistetuksi Gramin ja Schmidtin konstruktiolla ja ortogonaalisen projektion minimaalietäisyysominaisuuden johdettua sen avulla kuten reaalisessa tapauksessa.

Kahden ortonormaalissa kannassa lausutun vektorin sisätulo on koordinaattien *kompleksinen pistetulo*, siis (Todista tämä parsevalin kaavan kompleksiversio! Huomaa liittoluvut.)

$$(*) \qquad \left(\sum_{j=1}^{n} \lambda_j e_j \mid \sum_{j=1}^{n} \mu_j e_j\right) = \sum_{j=1}^{n} \lambda_j \overline{\mu}_j.$$

7.5. Unitaariset ja hermiittiset operaattorit.

Kompleksisen sisätuloavaruuden isometrisia isomorfismeja sanotaan unitaarisiksi operaattoreiksi. *Unitaarinen* operaattori on siis lineaarikuvaus

$$\begin{aligned}L : V \to V, &\quad \text{jolle} \\ L(x) = \|x\| &\quad \forall x \in V.\end{aligned}$$

Samalla tavalla kuin reaalisessa tapauksessa voi todistaa, että tämä on yhtäpitävää etäisyyksien ja myös sisätulojen säilymisen kanssa. Unitaariselle operaattorille L pätee siis

$$(Lx|Ly) = (x|y) \quad \forall x, y \in V,$$

Unitaarinen operaattori on siis sama asia kuin kompleksinen *sisätuloavaruusisomorfismi*. Erityisesti se kuvaa ortonormaalin kannan ortonormaaliksi kannaksi.

Sisätulon seskilineaarisuudesta aiheutuu, että unitaarista operaattoria ortonormaalissa kannassa E esittävä *unitaarinen matriisi* toteuttaa yhtälön

$$C^{-1} = C^{\dagger},$$

missä $C^{\dagger}$ on matriisin C *hermiittinen konjugaatti*, ts. saatu C:stä transponoimalla ja vaihtamalla sen lisäksi kaikki alkiot liittoluvuikseen. Erityisesti reaalinen matriisi on unitaarinen tasan ollessaan ortogonaalinen.

Seuraavat ehdot ovat yhtäpitäviä keskenään ja sen kanssa, että L on unitaarikuvaus:

(1) $\forall u \in V: \quad \|Lu\| = \|u\|$
(2) $\forall u, v \in V : d(Lu, Lv) = d(u, v)$.
(3) $\forall u, v \in V: \quad (Lu|Lv) = (u|v)$.
(4) Jollakin/jokaisella ortonormaalilla kannalla $E = (e_1, \ldots, e_n)$ myös $E' = (Le_1, \ldots, Le_n)$ on ortonormaali kanta
(5) Jollakin/jokaisella ortonormaalilla kannalla matriisin $C = Mat(L; E)$ sarakkeet ovat keskenään ortonormaaleita, toisin sanoen $C^{\dagger}C = I_{n\times n}$.
(6) Jollakin/jokaisella ortonormaalilla kannalla matriisin $C = Mat(L; E)$ rivit ovat keskenään ortonormaaleita, toisin sanoen $CC^{\dagger} = I_{n\times n}$.

Perustelu. Näiden väitteiden todistus on olennaisesti samanlainen kuin reaalisessa tapauksessa, paitsi matriisin liittolukuja koskevassa kohdassa, joka nyt on muutettava muotoon:

" 4) $\Longleftrightarrow$ 5) " Olkoon $E = (e_1, \ldots, e_n)$ ortonormaali kanta ja $C = Mat(L; E)$. Matriisin C sarakkeet ovat vektorit $Le_1, \ldots, Le_n$. Näiden keskinäiset sisätulot ovat kaavan $(*)$ mukaisesti toisaalta juuri matriisin $C^{\dagger}C$ alkiot:

$$(C^{\dagger}C)_{ij} = (Le_i|Le_j)$$

Toisaalta vektorit $(Le_1, \ldots, Le_n)$ ovat ortonormaalit täsmälleen silloin, kun

$$(Le_i|Le_j) = \begin{cases} 1, \text{ kun } i = j \\ 0, \text{ kun } i \neq j \end{cases} = (I_{n\times n})_{ij}.$$

Unitaariset matriisit ovat kompleksisessa teoriassa samassa roolissa kuin ortogonaalimatriisit reaalisessa teoriassa: niillä vaihdetaan ortonormaalia kantaa. Ei ole vaikeata arvata, että myös symmetrisen matriisin määritelmää kannattaa muuttaa lisäämällä liittoluvun otto: Sanomme, että kompleksinen neliömatriisi A on *hermiittisesti symmetrinen* eli *hermiittinen*[74], mikäli

$$A^{\dagger} = A,$$

[74]Ranskalaisen matematiikon Charles Hermiten (1822-1901) mukaan.

eli $a_{ij} = \overline{a_{ji}} \quad \forall i, j$, myös, kun $i = j$. Reaalinen matriisi on siis hermiittinen tasan ollessaan symmetrinen.

Hermiittisyys on itse asiassa lineaarikuvauksen sisätuloon liittävä ominaisuus. Voimme määritellä, että lineaarikuvaus $L : V \to V$ on *hermiittinen*, jos

$$(Lx|y) = (x|Ly) \quad \forall x, y \in V.$$

Tämä on yhtäpitävää sen kanssa, että L:n matriisi ortonormaalikannassa on hermiittinen.

7.6. Hermiittisen operaattorin unitaarinen diagonalisointi. Todistamme tässä luvussa, että hermiittisen matriisin A voi diagonalisoida unitaarisella kannanvaihtomatriisilla C ja että tästä seuraa vastaava lause symmetrisille ja ortogonaalimatriiseille. Diagonalisointi perustuu luonnollisesti ominaisarvojen ja -vektoreiden tutkimiseen. Aluksi voi kuitenkin todeta, että tietysti $C^{-1}AC$ on hermiittinen aina, kun A on hermiittinen ja C unitaarinen.

LAUSE. *Hermiittisen operaattorin ominaisarvot ovat reaalisia.*

TODISTUS. Olkoon $Lx = \lambda x$, $x \neq 0$, L hermiittinen.

$$\lambda\|x\|^2 = \lambda(x|x) = (\lambda x|x) = (Lx|x) = (x|Lx) = (x|\lambda x) = \overline{\lambda}(x|x) = \overline{\lambda}\|x\|^2.$$

□

LAUSE. *Hermiittisen operaattorin eri ominaisarvoihin liittyvät ominaisvektorit ovat ortogonaalisia.*

TODISTUS. Oleellisesti kuten reaalinen versio. □

LAUSE. *Jos* $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n$ *on* **reaalisen** *matriisin A ominaisvektori, niin myös sen "liittovektori"* $\overline{x} = \begin{bmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \vdots \\ \overline{x}_n \end{bmatrix}$ *on samaan ominaisarvoon $\lambda \in \mathbb{R}$ liittyvä ominaisvektori, samoin siis x:n reaaliosa* $\mathrm{Re}(x) = \frac{1}{2}(x + \overline{x})$

TODISTUS. Olkoon $Ax = \lambda x$, $x \neq 0$.

$$A\overline{x} = \overline{Ax} = \overline{\lambda x} = \overline{\lambda}\overline{x} = \lambda\overline{x}.$$

□

LAUSE. *Symmetrisellä matriisilla on reaalinen ominaisarvo ja siihen liittyvä reaalinen ominaisvektori, toisin sanoen symmetrisellä matriisilla on ominaisarvo reaalisen teorian mielessä.*

TODISTUS. Algebran peruslauseesta seuraa, että jokaisella matriisilla A on ainakin yksi kompleksinen ominaisarvo. Koska symmetrinen matriisi on hermiittinen, ominaisarvo λ on reaalinen. Koska symmetrinen matriisi on reaalinen, vastaavan ominaisvektorin reaaliosakin on saman ominaisarvoon liittyvä ominaisvektori, siis etsitty reaalinen ominaisvektori. □

Lause. *Hermiittinen operaattori diagonalisoituu unitaarisella matriisilla.*

Todistus. Olkoon $\lambda_1 \in \mathbb{R}$ hermiittisen operaattorin L ominaisarvo ja $V_1 \subset V$ vastaava ominaisavaruus. Hermiittisyydestä seuraa helposti, että ominaisavaruuden ortogonaalinen komplementti $V_1^\perp = \{x \in V \mid x \perp y \forall y \in V_1\}$ on *invariantti aliavaruus*, toisin sanoen

$$Lx \in V_1^\perp \quad \forall x \in V_1^\perp.$$

Valitaan ortonormaali kanta kummallekin aliavaruudelle V_1 ja $V_1^\perp$. Saadussa koko avaruuden ortogonaalikannassa $Mat(L)$ hajoaa neljään lohkoon, joista ainakin kaksi on varmasti nollia:

$$\mathrm{Mat}(L) = \begin{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_1 \end{bmatrix} & 0 \\ 0 & A_1 \end{bmatrix},$$

ja koska $\mathrm{Mat}\, L$ on hermiittinen, on myös lohko A_1 hermiittinen matriisi. Päättely voidaan siis toistaa ja lohkoa myös A_1:

$$A_1 = \begin{bmatrix} \begin{bmatrix} \lambda_2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_2 \end{bmatrix} & 0 \\ 0 & A_2 \end{bmatrix},$$

missä A_2 on hermiittinen. Yhteensä siis

$$\mathrm{Mat}(L) = \begin{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_1 \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} \lambda_2 & \dots & 0 & \\ \vdots & \ddots & \vdots & 0 \\ 0 & \dots & \lambda_2 & \\ & 0 & & A_2 \end{bmatrix} \end{bmatrix},$$

Toistamme päättelyä kunnes dimensiot loppuvat ja L on diagonalisoitu ortonormaalissa kannassa. $\square$

Seuraus. *Yhdistämällä edelliset lauseet saadaan haluttu symmetristä matriisia koskeva väite.* $\square$