2 Johdanto

Tällä kurssilla tutustumme vektoreihin ja matriiseihin jotka ovat tärkeä osa matematiikan ja tieteellisen laskennan peruskieltä. Käyttämällä vektoreita ja matriiseja voidaan monimutkaisia geometrian käsitteitä, yhtälöryhmiä ja lineaarisia kuvauksia kirjoittaa lyhyesti sekä manipuloida helposti. Oikealla "kielellä" kirjoitettuna monet muuten monimutkaiset käsitteet muuttuvat yksinkertaisiksi ja niiden taustalla olevat matemaattiset ideat tulevat selvästi näkyviin.

Yksinkertaisimmillaan vektori on n-alkion lista, esimerkiksi

$$\boldsymbol{a} = (1, 3, 1) \quad \text{ja} \quad \boldsymbol{b} = (1, 4, 2, 5, 6).$$

Vektori eroaa listasta siten, että kahden saman kokoisen vektorin välille on määritelty yhteenlasku ja vektoria voidaan kertoa skalaarilla s. Vektoreilla voidaan esittää hyvin erilaisia asioita, vaikkapa polynomin kertoimia tai koordinaatiston pisteitä. Käytetään seuraavassa vektorin $\boldsymbol{a}$ alkiosta i merkintätapaa a_i.

Määritelmä 2.1. *Olkoot s mielivaltainen reaaliluku ja $(x_1, \ldots, x_n), (y_1, \ldots, y_n)$ kaksi n-alkion vektoria. Tällöin*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

ja

$$s(x_1, \ldots, x_n) = (sx_1, \ldots, x_n).$$

Esimerkki 2.1. *Mielivaltainen ensimmäisen asteen polynomi*

$$p(x) = a_0 + a_1 x$$

voidaan esittää vektorina

$$(a_0, a_1)$$

Vektorien yhteenlasku vastaa tällöin kahden polynomin yhteenlaskua ja skalaarilla kertominen polynomin kertomista skalaarilla.

Matriisi on kaksiulotteinen $n \times m$-alkion taulukko, esimerkiksi

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

Käytetään seuraavassa matriisin A alkiosta ij merkintätapaa a_{ij}. Matriisien avulla voidaan esittää mm. yhtälöryhmiä tai funktioita jotka liittävät kuhunkin vektoriin toisen vektorin ja ovat lineaarisia. Kuten vektoreja, matriiseja voidan myös kertoa sekä laskea yhteen.

Esimerkki 2.2. *Olkoot f funktio, joka vastaa tason kiertoa θ-astetta. Funktio f saa argumentikseen koordinaatin (x_1, x_2) ja palauttaa pisteen (y_1, y_2) johon (x_1, x_2) kuvautuu, kun tasoa kierretään θ-astetta. Piste (y_1, y_2) voidaan määrittää helposti geometrisen argumentin perusteella.*

Esimerkki 2.3. *Tarkastellaan yhtälöryhmää*

$$\begin{cases} x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 1 \\ x_1 + x_2 + 2x_3 &= 3. \end{cases}$$

Kaikki tähän yhtälöryhmään liittyvä informaatio on muuttujien x_1, x_2, x_3 kertoimissa ja yhtälön oikeassa puolessa. Muuttujien kertoimet voidaan esittää taulukkona

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

ja yhtälön oikea puoli vektorina

$$\boldsymbol{b} = (0, 1, 3).$$

Taulukko A ja vektori $\boldsymbol{b}$ voidaan helposti kirjoittaa tietokoneen muistiin ja yhtälöryhmä voidaan ratkaista käyttämällä sopivaa menetelmää. Matlabissa tämä onnistuu seuraavien komentojen avulla

```
A = [ 1 1 1 ; 1 2 1 ; 1 1 2];
b = [0 ;1 ;3];
x = A\b
```

vektori $\boldsymbol{x}$ sisältää muuttujien x_1, x_2, x_3 arvot.

2.1 Esimerkki käytännön ongelmasta: lämmönjohtuminen seinärakenteessa

Tieteellisessä laskennassa matriiseja ja vektoreita tarvitaan eritysesti kun käytännön ongelmasta muodostettu matemaattinen malli halutaan ratkaista tietokoneella. Jos käytettävä malli on lineaarinen, muodostetaan tietokoneen muistiin suuri yhtälöryhmä jolle etsitään ratkaisu. Saadusta tuloksesta voidaan piirtää kuva tai laskea käsiteltävän ongelman kannalta tärkeitä tunnuslukuja.

Ongelmanratkaisu tieteellisessä laskennassa voidaan jakaa karkeasti seuraaviin vaiheisiin :

1. **Mallin muodostaminen käytännön ongelmasta**

2. **Mallin ratkaiseminen tietokoneen avulla**

3. **Saatujen tulosten käsittely**

Tämän kurssin materiaali on tärkeää kaikkien ylläolevien vaiheiden kannalta. Ensimmäinen ja viimeinen askel vaativat kuitenkin tutkittavan ongelman syvällistä ymmärtämistä, joten emme juurikaan käsittele niitä tällä kurssilla. Jotta voimme paremmin tieteelliseen laskentaan liittyviä haasteita tarkastellaan esimerkin lämpötilan muuttumista rakennuksen ulkoseinän sisällä.

Olkoot huonetilan lämpötila T_{in} ja ulkotilan lämpötila T_{out}. Olemme kiinnostuneita lämpötilasta ulkoseinän sisällä sekä seinän läpi virtaavan lämpöenergian määrästä. Koska seinärakenne on suuri ja tasomainen voimme olettaa että lämpötila rakenteen sisällä säikyy vakiona kun liikutaan seinän tasossa. Kyseessä on siis yksiulotteinen lämmönjohtuvuusongelma. Kiinnitetään origo seinän sisäpinnalle ja etsitään seuraavassa lämpötilaa $u(x)$ seinän sisällä, jossa x on etäisyys seinän sisäpinnasta ulkopintaan päin.

Ulkoseinän sisällä tapahtuu lämpöenergian siirtymistä asunnon sisätilasta ulkotilaan. Tekemiemme yksinkertaistusten vuoksi lämpöenergiaa ei siirry seinän tasossa. Lämpöenergian siirtymistä kuvataan lämpövuon avulla. Yksinkertaistetusti, lämpövuo kertoo mikä on lämpöenergian siirtymisnopeus pinta-ala yksikköä kohde.

Seinän sisällä lämpövuo q noudattaa Fourierin lämmönjohtumislakia,

$$q(x) = -ku'(x)$$

jossa k on materiaalille ominainen lämmönjohtumiskerroin ja u on lämpötila. Kun tarkastelemme tasapainotilaa lämpövuo on (tuntematon) vakio koko seinärakenteessa,

$$q(x) = q_0$$

Olkoot x_i seinärakenteen materiaalirajapinnan i etäisyys sisäseinästä ja k_i lämmönjohtuvuuskerroin materiaalissa i. Kussakin materiaalissa $1 \leq i \leq N$ pätee siis

$$k_i u_i'(x) = q_0 \quad x \in (x_i, x_{i+1}). \tag{1}$$

Rajapinnoilla $i = 1, N$ lämpötila on tunnettu, eli

$$u(x_i) = T_i, i = 1, N$$

Yhtälöstä (1) seuraa, että u on kussakin seinärakenteen materiaalissa ensimmäisen asteen polynomi. Esitetään $u(x)$ sen solmuarvojen $u(x_i), i = 1, \ldots, N$ avulla. Tällöin välillä (x_i, x_{i+1}) pätee

$$u(x) = u(x_i)\frac{x_{i+1} - x}{x_{i+1} - x_i} + u(x_{i+1})\frac{x_i - x}{x_i - x_{i+1}}.$$

Joten lämpövuo välillä (x_i, x_{i+1}) on

$$q(x) = ku(x_i)\frac{1}{x_{i+1} - x_i} + ku(x_{i+1})\frac{1}{x_i - x_{i+1}}.$$

Lämpövuon on jatkuva, joten kullakin rajapinnalla $i = 2, \ldots, (N-1)$ pätee

$$\lim_{x \to x_i^-} q(x) = \lim_{x \to x_i^+} q(x).$$

Eli,

$$\frac{k_{i-1}}{x_i - x_{i-1}}u(x_{i-1}) + \frac{k_{i-1}}{x_{i-1} - x_i}u(x_i) = \frac{k_i}{x_{i+1} - x_i}u(x_i) + \frac{k_i}{x_i - x_{i+1}}u(x_{i+1})$$

Käytetään merkintää $h_i = x_{i+1} - x_i$, joten saadaan

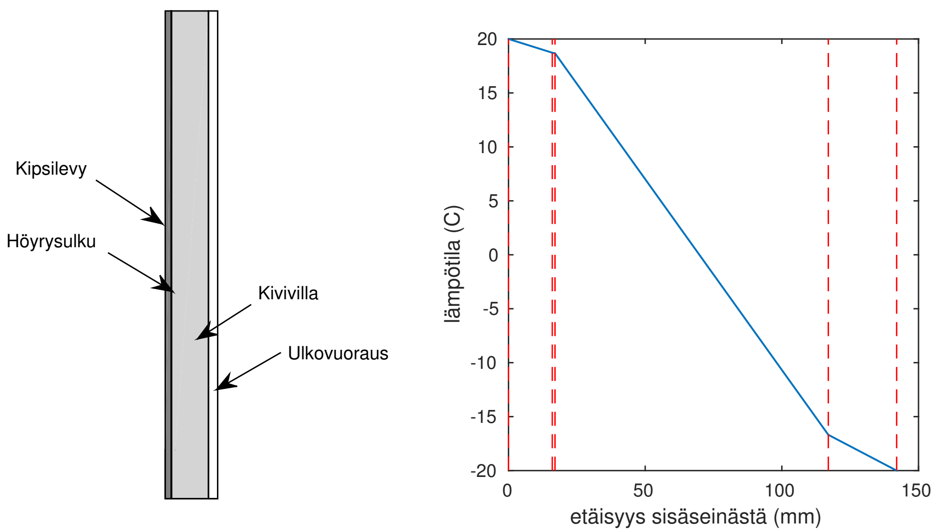
$$\frac{k_{i-1}}{h_{i-1}}u(x_{i-1}) - \left(\frac{k_{i-1}}{h_{i-1}} + \frac{k_i}{h_i}\right)u(x_i) + \frac{k_i}{h_i}u(x_{i+1}) = 0.$$

Huonetilan ja ulkotilan lämpötila huomioidaan rajapintoihin $i = 2$ ja $i = (N-1)$ liittyvissä yhtälöissä,

$$-\left(\frac{k_1}{h_1} + \frac{k_2}{h_2}\right)u(x_2) + \frac{k_2}{h_2}u(x_3) = -\frac{k_1}{h_1}T_{in}$$

$$\frac{k_{N-2}}{h_{N-2}}u(x_{N-3}) - \left(\frac{k_{N-1}}{h_{N-1}} + \frac{k_{N-2}}{h_{N-2}}\right)u(x_{N-1}) = -\frac{k_{N-1}}{h_{N-1}}T_{out}$$

Kyseessä on lineaarinen yhtälöryhmä joka koostuu $N-2$ yhtälöstä.



Kuva 1: Oikealla esimerkki yksinkertaisesta seinärakenteesta. Vasemmalla lämpötila seinärakenteen sisällä, kun $T_{in} = 20$ ja $T_{out} = -20$.

Tarkastellaan esimerkin vuoksi seinärakennetta, jonka rakenne on seuraava :

1. 16 mm Kipsilevyä, $k_1 = 0,2$
2. 1 mm Höyrysulkumuovi, $k_2 = 0.33$
3. 100 mm Kivivillaa, $k_3 = 0,045$
4. 25 mm Puinen ulkoverhous $k_5 = 0.12$.

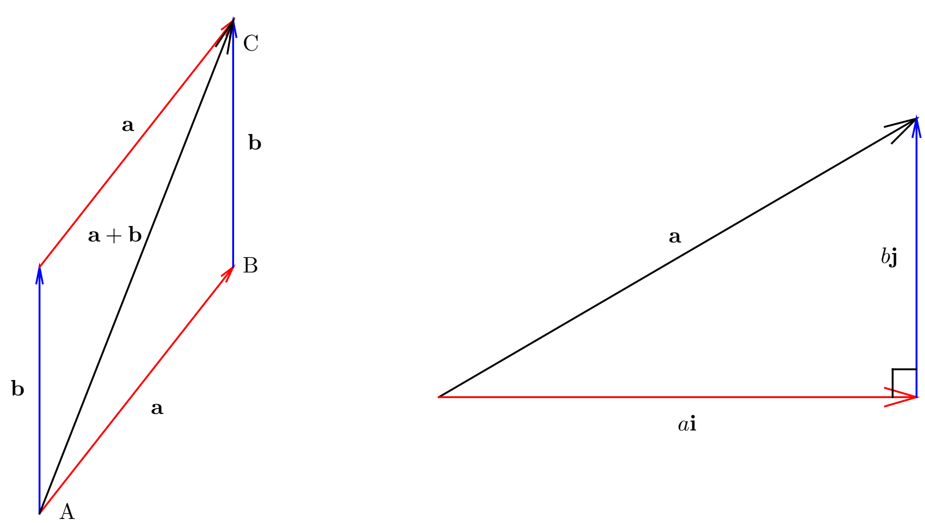
Sisälämpötila on $T_{in} = 20$ astetta ja ulkolämpötila $T_{out} = -20$ astetta. Lämpötila seinän sisällä voidaan ratkaista ylläolevan menettelyn avulla ja oheisen Matlab-ohjelman avulla. Näin laskettu lämpötila $u(x)$ on esitetty kuvassa 1.

```
% Maaritellaan seinarakenne
k = [0.2 0.33 0.045 0.12];
h = [16 1 100 25];

% Maaritellaan ulko ja sisalampotila
Tin = 20;
Tout = -20;

% lasketaan materiaalirajapintojen koordinaatit
x = [0 cumsum(h)];

N = length(x);

% Muodosta yhtaloryhma
A = zeros(N-2,N-2);
b = zeros(N-2,1);

for i=1:(N-2)

    A(i,i) = -( k(i)/h(i) + k(i+1)/h(i+1));

    switch(i)
        case 1
            A(i,i+1) = k(i+1)/h(i+1);
            b(i) = -k(i)/h(i)*Tin;
        case N-2
            A(i,i-1) =  k(i)/h(i);
            b(i) = -k(i+1)/h(i+1)*Tout;

        otherwise
            A(i,i-1) = k(i)/h(i);
            A(i,i+1) = k(i+1)/h(i+1);
    end
end

% Ratkaistaan lampotila u
u = [Tin ; A\b ; Tout];

% piirretaan u
P = plot(x,u);
set(P,'linewidth',1);
3

% Lasketaan lampovuo sisalta ulos
q_01 = u(1)*k(1)/h(1) - u(2)*k(1)/h(1);
```

3 Vektori

Vektorilla voidaan tarkoittaa montaa eri asiaa. Lukiossa vektoreihin tutustutaan sekä fysiikan että matematiikan kursseilla. Fysiikassa vektoreilla kuvataan suureita, esimerkiksi voimaa, nopeutta tai liikemäärää, joilla on sekä suuruus että suunta. Lukiossa vektoria ajatellaa geometrian kautta, eli taustalla on ajatus paperille piirretystä nuolesta. Tällä kurssilla vektori käsitetään aikaisempaa abstraktimmassa mielessä.

Fysiikassa käytettävien vektorien laskutoimitukset on määritelty siten että ne vastaavat niillä esitettävien suureiden kokeellisesti määrättyä käyttäytymistä. Esimerkiksi kahden vektorin $\boldsymbol{a}$ ja $\boldsymbol{b}$ yhteenlasku määritellään puolisuunnikassäännön kautta : Olkoot A, B, C tason pisteitä ja vektori $\boldsymbol{a} = AB$ ja $\boldsymbol{b} = BC$. Tällöin summavektori $\boldsymbol{a}+\boldsymbol{b} = AC$, katso Kuva 2. Yhteenlaskun määritelmä on puhtaan geometrinen, eli se voidaan suorittaa paperilla piirtämällä kaksi vektoria peräkkäin. Summavektori saadaan yhdistämällä ensimmäisen vektorin alkupiste toisen vektorin loppupisteeseen.



Kuva 2: Geometristen vektorien yhteenlasku ja vektorin $\boldsymbol{a}$ jako kantavektorien $\boldsymbol{i}$ ja $\boldsymbol{j}$ summaksi.

Lukion matematiikan kursseilla vektoriin törmätään analyyttisessä geometriassa, jossa vektori esitetään yksikkövektorien avulla. Olemme tottuneet kirjoittamaan kolmiulotteisen avaruuden vektorit muodossa

$$\boldsymbol{a} = a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k},$$

jossa $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ ovat x, y, z-akselien suuntaiset yksikkövektorit ja a, b, c niiden kertoimet, katso Kuva 2. Kun vektorit esitetään tässä muodossa, voidaan moni geometrian tehtävä ratkaista algebrallisesti, ainoastaan vektoriin liittyvien kertoimien avulla. Siirtyminen puhtaasti geometrisesta vektorista yksikkövektorien avulla esitettyyn vektorin on ollut aikanaan erittäin merkittävä matemaattinen edistysaskel.

Esimerkki 3.1. *Kolmiulotteisen avaruuden taso voidaan määritellä kolmen vektorin* $\boldsymbol{a}$,$\boldsymbol{b}$ *ja* $\boldsymbol{c}$ *avulla muodossa*

$$\{\ \boldsymbol{c} + t\boldsymbol{a} + s\boldsymbol{b} \mid t \in \mathbb{R}, s \in \mathbb{R}\ \}.$$

Tällöin esimerkiksi kysymys: "kuuluuko piste $\boldsymbol{x}$ *tasoon ?" - voidaan muotoilla uudelleen : onko olemassa* t *ja* s *siten että*

$$\boldsymbol{c} + t\boldsymbol{a} + s\boldsymbol{b} = \boldsymbol{x}\ ?$$

Alkuperäinen kysymys on siis ekvivalentti ylläolevan yhtälöryhmän ratkaisun olemssaolon kanssa.

Matematiikassa vektori ymmärretään aikasempaa yleisemmässä mielessä. Termiä vektori käytetään, kun annetun joukon alkioden välille on määritelty yhteenlasku sekä skalaarilla kertominen. Joukon alkiot voivat esittää mitä vain, kyseessä voi olla vaikka joukko funktioita. Kyseessä on matematiikassa laajasti käytetty perusrakenne, jolla voidaan taata, että tietyt laskutoimitukset toimiva odotetulla tavalla.

3.1 Avaruuden $\mathbb{R}^2$:n vektorit

Aloitetaan määrittelemällä vektoriavaruus $\mathbb{R}^2$. Löyhästi määriteltynä vektoriavaruus on matemaattinen rakenne, joka koostuu

(i). Vektorien joukosta V.

(ii). Vektorien yhteenlaskusta.

(iii). Vektorin ja skalaarin tulosta.

Tälläisen rakenteen ajatuksena on taata, että skalaareilla opitut laskutoimitukset toimivat odotetulla tavalla kun käsitellään esimerkiksi n-alkion listoja ja laskutoimitukset määritellään kuten Luvussa 2. Kyseessä on hyvin formaali rakenne, jota harvoin tarvitaan käytännön laskennassa. Vektoriavaruuksia tarkastellaan tarkemmin matematiikan jatkokursseilla.

Vektoriavaruuksista puhuttaessa laskutoimitukset ajatellaan kuvauksina joukon V alkioilta toisilleen. Esimerkiksi yhteenlasku voidaan tulkita kuvauksena f, joka liittää kuhunkin joukon V alkioon yhden V:n alkion. Formaalisti plusmerkki tulee "lukea"muodossa

$$\boldsymbol{a} + \boldsymbol{b} = f(\boldsymbol{a}, \boldsymbol{b}).$$

Vektoriavaruus on suljettu yhteenlaskun ja skaarilla kertomisen suhteen, eli nämä operaatiot tuottavat jälleen avaruuden V alkion. Jokaiselle skalaarille s ja alkiolle $\boldsymbol{a}, \boldsymbol{b} \in V$ pätee

$$\boldsymbol{a} + \boldsymbol{b} \in V \quad \text{ja} \quad s\boldsymbol{a} \in V.$$

Yhteenlaskun ja skalaarilla kertomisen tulee olla määritelty järkevällä tavalla, esimerkiksi yhteenlaskun tulee olla assosiatiivinen ja kommutatiivinen, eli jokaiselle $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in V$ pätee

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a} \quad \text{and} \quad (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}).$$

Ajatuksen on taata, että vektoriavaruudessa suoritetut laskutoimitukset käyttäytyvät samalla tavalla kuin skalaareille. Kahdella joukon V alkion tulolla ei tyypillisesti ole järkevää tulkintaa tai se ei enää kuulu joukkoon V. Tästä syystä vektoriavaruuden alkioiden välille ei ole määritelty tuloa.

Vektoriavaruus $\mathbb{R}^2$ rakentuu järjestettyjen parien joukosta

$$V = \{ \ (x_1, x_2) \mid x_1 \in \mathbb{R}, x_2 \in \mathbb{R} \ \},$$

jossa yhteenlasku ja skalaarilla kertominen on määritelty seuraavalla tavalla :

Määritelmä 3.1. *Olkoon s mielivaltainen reaaliluku ja $(a_1, a_2), (a_2, b_2) \in \mathbb{R}^2$. Tällöin*

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$$

ja

$$s(a_1, b_1) = (sa_1, sb_1).$$

Kyseessä on yleinen matemaattinen rakenne, jota käyttämällä voidaan helposti esittää useita erilaisia asioita, esimerkiksi ensimmäisen asteen polynomeja.

Esimerkki 3.2. *Mielivaltainen ensimmäisen asteen polynomi*

$$p(x) = a_0 + a_1 x$$

voidaan esittää vektorina

$$(a_0, a_1)$$

Vektorien yhteenlasku vastaa tällöin kahden polynomin yhteenlaskua ja skalaarilla kertominen polynomin kertomista skalaarilla. Tälläistä esitystapaa käytetään polynomien esittämiseen esimerkiksi MATLABissa. Huomaa, että kahden ensimmäisen asteen polynomin kertolaskun on toisen asteen polynomi, joten tulon puuttuminen avaruuden $\mathbb{R}^2$ määritelmästä on tässä esimerkissä täysin luonnollista. Itse asiassa ensimmäisen asteen polynomit muodostavat itsessään vektoriavaruuden jolle pätee $V = \{a_0 + a_1 x \mid a_0, a_1 \in \mathbb{R} \ \}$. Polynomien avaruudessa yhteenlasku ja skalaarilla kertominen on määritelty luonnollisella tavalla.

Käytetään seuraavassa avaruuden $\mathbb{R}^2$ vektoreista merkintätapaa $\boldsymbol{a} \in \mathbb{R}^2$. Vektorin komponenteilla a_1 ja a_2 tarkoitetaan tarkoitetaan järjestetyn parin $\boldsymbol{a}$ ensimmäistä ja toista alkioa, eli,

$$\boldsymbol{a} = (a_1, a_2).$$

Merkitään nolla-vektoria $(0, 0)$ lyhyesti 0:lla ja käytetään merkintätapaa

$$-\boldsymbol{a} = (-a_1, -a_2).$$

Kahden vektorin $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$ välinen vähennyslasku tulkitaan laskutoimituksena

$$\boldsymbol{a} - \boldsymbol{b} = \boldsymbol{a} + (-\boldsymbol{b}) = (a_1 - b_1, a_2 - b_2).$$

Avaruuden $\mathbb{R}^2$ alkiot voidaan samaistaa geometrisiin vektoreihin. Tätä varten tarvitaan kaksi geometrista vektoria $\boldsymbol{a}$ ja $\boldsymbol{b}$ joiden tulee poiketa nollavektorista ja jotka eivät saa olla yhdensuuntaisia, eli

$$\boldsymbol{a}, \boldsymbol{b} \neq 0 \quad \text{ja} \quad \boldsymbol{a} \neq \lambda \boldsymbol{b} \quad \forall \lambda \in \mathbb{R}.$$

Käyttämällä näitä kahta vetoria, tason geometriset vektorit voidaan esittää muodossa

$$\{ \, x\boldsymbol{a} + y\boldsymbol{b} \mid x, y \in \mathbb{R} \, \}.$$

Käyttämällä ylläolevaa esitystä, kukin tason vektori voidaan samaistaa tiettyyn avaruuden $\mathbb{R}^2$:n alkioon. Näin vektorien väliset laskutoimitukset voidaan suorittaa täysin algebrallisesti. Esimerkiksi kahden pisteitä $(x_1, y_1), (x_2, y_2)$ vastaavan tason vektorin summa voidaan laskea helposti :

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2),$$

joka vastaa laskutoimitusta

$$(x_1\boldsymbol{a} + y_1\boldsymbol{b}) + (x_2\boldsymbol{a} + y_2\boldsymbol{b}) = (x_1 + x_2)\boldsymbol{a} + (y_1 + y_2)\boldsymbol{b}.$$

Kun kiinnitetään origo O ja valitaan tason vektorit $\boldsymbol{a}$ ja $\boldsymbol{b}$ siten että niiden pituus on yksi ja ne ovat toisiinsa nähden suorassa kulmassa voidaan avaruuden $\mathbb{R}^2$ alkiot samaistaa tason pisteiden kanssa. Tällöin $\mathbb{R}^2$:n vektori (x_1, y_1) on vastaavan pisteen koordinaatti.

3.2 Tason geometriaa

Edellisessä luvussa määrittelimme vektoriavaruuden $\mathbb{R}^2$. Avaruuden määritelmän taustalla oli ajatus geometriselle vektorille määritellyistä laskutoimituksista. Suurin osa vektoriavaruuksiin määritellyistä toimituksista pohjaa samalla tavalla geometrisiin käsitteisiin. Tässä luvussa käydään läpi Euklidiseen tasoon määriteltyjä operaatioita, jotka yleistämme myöhemmin muihin vektoriavaruuksiin. Seuraavassa samaistamme avaruuden $\mathbb{R}^2$ Euklidiseen tasoon, eli tulkitsemme vektorin (x_1, x_2) muodossa

$$x_1\boldsymbol{i} + x_2\boldsymbol{j},$$

jossa $\boldsymbol{i}, \boldsymbol{j}$ ovat x- ja y-akselien suuntaisia yksikkövektoreita.

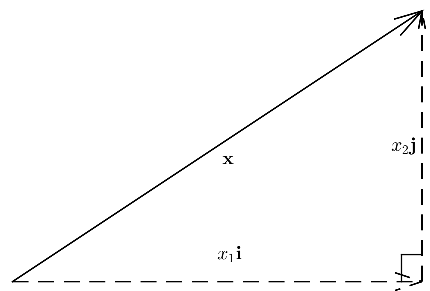
3.3 Vektorin pituus

Määritellään vektorin $\boldsymbol{x} \in \mathbb{R}^2$ pituus tutulla tavalla,

$$|\boldsymbol{x}| = \left(x_1^2 + x_2^2\right)^{1/2}.$$

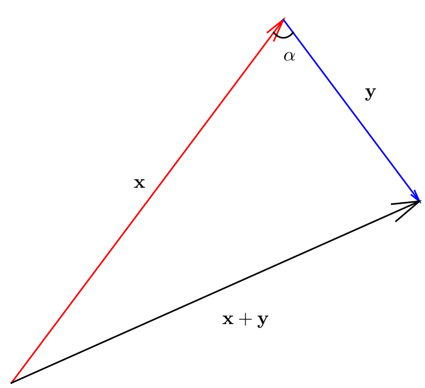
Tämän määritelmän taustalla on vektorien geometrinen tulkinta sekä Pythagoran lause, katso Kuva 3. Määritelmästä seuraa suoraan kolmioepäyhtälö: olkoot $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$. Tällöin

$$|\boldsymbol{x} + \boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{y}|.$$



Kuva 3: Vektorin $\boldsymbol{x}$ jako komponentteihin. Vektorin pituus voidaan laskea Pythagoran lauseen avulla.

Kolmioepäyhtälö voidaan perustella geometrisen argumentin avulla. Vektorit $\boldsymbol{x}$,$\boldsymbol{y}$ ja $\boldsymbol{x}+\boldsymbol{y}$ muodostavat kolmion, katso Kuva 4. Oletetaan nyt, että kolmioepäyhtälö ei päde. Tällöin kolmion kahden sivun pituuksien summa on pienempi, kuin kolmannen sivun pituus. Koska tälläistä kolmioa ei voi olla olemassa, kolmioepäyhtälön tulee olla totta. Todistamme kolmioepäyhtälön myöhemmin algebrallisesti.



Kuva 4: Vektorien $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}$ muodostama kolmio.

Suoraan määritelmästä nähdään miten vektorin pituus käyttäytyy kun sitä kerrotaan mielivaltaisella skalaarilla λ. Saadaan,

$$|\lambda \boldsymbol{x}| = \left((\lambda \boldsymbol{x}_1)^2 + (\lambda \boldsymbol{x}_y)^2\right)^{1/2} = |\lambda||\boldsymbol{x}|.$$

Kyseessä on vektorin pituudelle luonnollinen ominaisuus : kun mitattava vektori kasvaa kaksinkertaiseksi, kasvaa myös sen pituus kaksinkertaiseksi. Suoraan määritelmästä seuraa, että vektorin pituudelle pätee

$$|\boldsymbol{x}| = 0 \Rightarrow \boldsymbol{x} = 0,$$

eli ainoastaan nollvektorin pituus on nolla. Tämän lisäksi $|\boldsymbol{x}| > 0, \boldsymbol{x} \neq 0$.

Yllä olemme määritelleet erään tavan mitata vektorin pituutta, jolla on neljä tärkeää ominaisuutta:

(i). $|\lambda \boldsymbol{x}| = |\lambda||\boldsymbol{x}|$ (Skaalaus)

(ii). $|\boldsymbol{x}| > 0, \boldsymbol{x} \neq 0$ (Positiivisuus)

(iii). $|\boldsymbol{x}| = 0 \Rightarrow \boldsymbol{x} = 0$ (Nollan pituus on nolla)

(iv). $|\boldsymbol{x} + \boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{y}|$. (Kolmioepäyhtälö)

Toinen lähtökohta vektorin pituusmitan määrittelyyn on ajatella sitä kuvauksena, joka liittää jokaiseen avaruuden $\mathbb{R}^2$ pisteeseen skalaarin. Tälläistä kuvausta kutsutaan normiksi, jos se täyttää ehdot (i)-(iv).

Esimerkki 3.3. *Tarkastellaan kuvausta $f : \mathbb{R}^2 \to \mathbb{R}$,*

$$f(\boldsymbol{x}) = max\{|x_1|, |x_2|\}.$$

Todistataan että f toteuttaa ominaisuudet (i)-(iv) eli on normi. Aloitetaan skaalausominaisuudesta (i). Suoraan laskemalla,

$$f(\lambda \boldsymbol{x}) = max\{|\lambda x_1|, |\lambda x_2|\} = \begin{cases} |\lambda x_1| & , |\lambda x_1| \geq |\lambda x_2| \\ |\lambda x_2| & , |\lambda x_1| < |\lambda x_2| \end{cases}$$

Itseisarvon laskusääntöjen nojalla saadaan

$$f(\lambda \boldsymbol{x}) = \begin{cases} |\lambda||x_1| & , |x_1| \geq |x_2| \\ |\lambda||x_2| & , |x_1| < |x_2| \end{cases} = |\lambda| max\{|x_1|, |x_2|\} = |\lambda| f(\boldsymbol{x}).$$

Ominaisuus (ii) seuraa helposti määritelmästä. Oletetaan, että $\boldsymbol{x} \neq 0$. Käsitellään erikseen tapauksia: $x_1 \neq 0$ tai $x_2 \neq 0$. Suoraan määritelmästä nähdään, että

$$f(\boldsymbol{x}) \geq |x_1| > 0, \text{ kun } x_1 \neq 0 \quad \text{ja} \quad f(\boldsymbol{x}) \geq |x_2| > 0, \text{ kun } x_2 \neq 0.$$

Eli, $f(\boldsymbol{x}) > 0$, kun $\boldsymbol{x} \neq 0$. Ominaisuus (iii) seuraa suoraan määritelmästä. Todistetaan vielä kolmioepäyhtälö (iv). Tutkitaan kaksi vaihtoehtoista tapausta, $f(\boldsymbol{x} + \boldsymbol{y}) = |x_1 + y_1|$ ja $f(\boldsymbol{x} + \boldsymbol{y}) = |x_2 + y_2|$. Iteseisarvon laskusääntöjen nojalla

$$|x_1 + y_1| \leq |x_1| + |y_1| \quad \text{ja} \quad |x_2 + y_2| \leq |x_2| + |y_2|$$

Koska $|x_i| \leq \max\{|x_1|, |x_2|\}, i = 1, 2$ ja $|y_i| \leq \max\{|y_1|, |y_2|\}, i = 1, 2$ kolmioepäyhtälö pätee.

Sopiva normi vektorin pituuden mittaamiseen voidaan valita tutkittavan tilanteen mukaan. Esimerkiksi geometrisia vektoreita kannattaa mitata eri tavalla kun vektoreita, jotka on samaistettu ensimmäisen asteen polynomeihin. Normin laskusäännöt mahdollistavat tiettyjen perustulosten todistamisen ilman tietoa siitä, mitä konkreettista normia käsitellään.

Esimerkki 3.4. *Tarkastellaan ensimmäisen asteen polynomeja. Yksi mahdollinen mitta polynomin suuruudelle on sen neliöintegraali välillä* $(0, 1)$*,*

$$\int_0^1 p^2(t)\, dt.$$

Esitetään polynomi muodossa $p = \alpha_0 + \alpha_1 x$*. Sijoittamalla tämä ylläolevaan lausekkeeseen saadaan*

$$\int_0^1 \alpha_0^2 + 2\alpha_0\alpha_1 t + \alpha_1^2 t^2\, dt = \alpha_0^2 + \alpha_0\alpha_1 + \frac{1}{3}\alpha_1^2.$$

Valitaan ylläolevan mukaan $\mathbb{R}^2$*:n pituusmitaksi*

$$f(\boldsymbol{x}) = x_1^2 + x_1x_2 + \frac{1}{3}x_2^2.$$

Myöhemmin todistamme, että tämä kuvaus toteuttaa ominaisuudet (i)-(iv) eli on normi.

3.4 Vektorin polaarimuoto

Kukin avaruuden $\mathbb{R}^2$ vektori $\boldsymbol{x}$ voidaan esittää polaarimuodossa sen pituuden r ja suunnan avulla. Vektorin suunnaksi valitaan yleensä sen x-akselin kanssa muodostama kulma $\theta \in [\pi, -\pi)$. Polaariesityksestä tuttuun komponenttiesitykseen voidaan siirtyä käyttämällä trigonometristen funktioiden määritelmiä. Näin saadaan,

$$\boldsymbol{x} = r\cos\theta\boldsymbol{i} + r\sin\theta\boldsymbol{j}.$$

Tämä muunnos voidaan tehdä myös toiseen suuntaan. Käytännön haasteena on valita kulman merkki oikein.

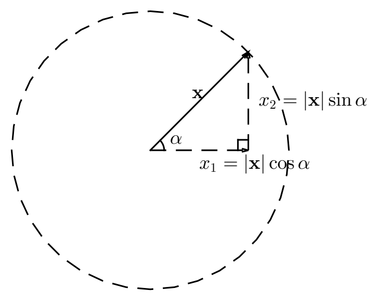
Polaariesitystä käyttämällä voidaan helposti todistaa joitan tason vektoreihin liittyviä perustuloksia. Insinöörityössä tälläistä esitysmuotoa käytetään, kun mallinnetaan sylinterinmuotoisia kappaleita, esimerkiksi pyöreitä johtimia. Tällöin esimerkiksi a-säteinen ympyrä voidaan esittää helposti joukkona

$$\{\, \boldsymbol{x} \in \mathbb{R}^2 \mid r = a \,\}.$$

3.5 Pistetulo

Määritellään kahden vektorin $\boldsymbol{x}$ ja $\boldsymbol{y}$ välinen pistetulo tutussa muodossa,

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1y_1 + x_2y_2.$$



Kuva 5: Vektorin esitys polaarikoordinaatistossa. Vektorin $\boldsymbol{x}$ pituus on $|\boldsymbol{x}|$ ja se muodostaa kulman α x-akselin kanssa.

Geometrisesti pistetulo jakaa vektorin $\boldsymbol{x}$ vektorin $\boldsymbol{y}$ suuntaiseen sekä sitä vastaan kohtisuoraan komponenttiin, katso Kuva 6. Vektorin $\boldsymbol{x}$ $\boldsymbol{y}$:n suuntaista komponenttia kutsutaan $\boldsymbol{x}$:n kohtisuoraksi projektioksi $\boldsymbol{y}$:n suuntaa, ja se voidaan laskea muodossa

$$\frac{\boldsymbol{x} \cdot \boldsymbol{y}}{|\boldsymbol{y}|^2}\boldsymbol{y}.$$

Vektorin $\boldsymbol{x}$ $\boldsymbol{y}$:n suuntaa vastaan kohtisuora komponentti saadaan vektorien yhteenlaskua käyttämällä muodossa

$$\boldsymbol{x} - \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{|\boldsymbol{y}|^2}\boldsymbol{y}.$$

Pistetuloon littyvä geometrinen tulkinta voidaan todistaa avaruudessa $\mathbb{R}^2$ käyttämällä vektorien polaariesitystä.

Lemma 3.1. *Olkoot $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ ja θ niiden välinen kulma. Tällöin pätee*

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

Todistus. Kirjoitetaan vektorit $\boldsymbol{x}$ ja $\boldsymbol{y}$ polaarikoordinaatistossa. Tällöin,

$$\boldsymbol{x} = (|\boldsymbol{x}| \cos \alpha, |\boldsymbol{x}| \sin \alpha) \quad \text{ja} \quad \boldsymbol{y} = (|\boldsymbol{y}| \cos \beta, |\boldsymbol{y}| \sin \beta).$$

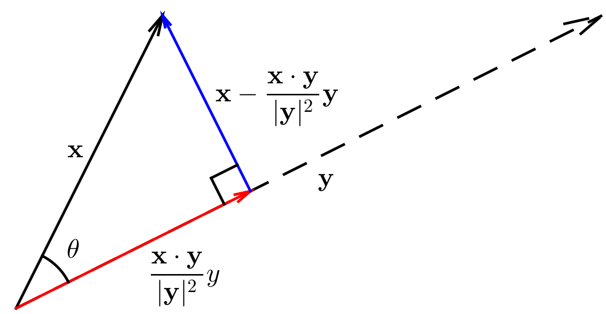
Pistetulolle pätee

$$\boldsymbol{x} \cdot \boldsymbol{y} = |\boldsymbol{x}||\boldsymbol{y}| \left(\cos \alpha \cos \beta + \sin \alpha \sin \beta\right).$$

Trigonometriasta tiedetään, että $\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$, joten

$$\boldsymbol{x} \cdot \boldsymbol{y} = |\boldsymbol{x}||\boldsymbol{y}| \cos(\alpha - \beta),$$

jossa $\alpha - \beta$ on vektorien $\boldsymbol{x}$ ja $\boldsymbol{y}$ välinen kulma. Kosinifunktion ominaisuuksista seuraa, ettei kulman $\alpha - \beta$ merkillä ole ylläolevassa lausekkeessa väliä, joten kulma voidaan valita aina positiiviseksi. $\square$



Kuva 6: Pistetulon geometrinen tulkinta.

Pistetulon projektio-ominaisuus voidaan myös kirjoittaa toisella tavalla. Huomataan, ensin että Euklidisen normin ja pistetulon välillä on yhteys $|\boldsymbol{x}|^2 = \boldsymbol{x} \cdot \boldsymbol{x}$. Lisäksi pistetulon määritelmän mukaan pätee, että

(i). $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$ (Symmetria)

(ii). $(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \cdot \boldsymbol{z} = \alpha\boldsymbol{x} \cdot \boldsymbol{z} + \beta\boldsymbol{y} \cdot \boldsymbol{z}$ (Bilineaarisuus)

Tarkastellaan nyt minimointitehtävää,

$$\min_{\alpha\in\mathbb{R}} \left| \boldsymbol{x} - \alpha\frac{\boldsymbol{y}}{|\boldsymbol{y}|} \right|^2 \tag{2}$$

Minimi voidaan laskea derivoimalla ylläolevaa lauseketta α:n suhteen. Minimi saavutetaan, kun $\alpha = \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{|\boldsymbol{y}|}$. Minimointitehtävä (2) on vaihtoehtoinen määritelmä vektorin $\boldsymbol{x}$ kohtisuoralle projektiolle vektorin $\boldsymbol{y}$ suuntaan.

Kuten pituusmitan tapauksessa, voidaan pistetulo yleistää skalaarituloksi. Tällöin tarkastellaan kuvausta $\langle\cdot,\cdot\rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ joka liittää kahteen avaruuden $\mathbb{R}^2$ vektoriin reaaliluvun. Skalaaritulota vaaditaan ominaisuudet (i)-(ii). Jotta skalaaritulo voi määritellä normin lausekkeen

$$\|\boldsymbol{x}\|^2 = \langle\boldsymbol{x},\boldsymbol{x}\rangle \tag{3}$$

kautta, vaaditaan siltä lisäominaisuudet

(iii). $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0, \boldsymbol{x} \neq 0$ (Positiivisuus)

(iv). $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \Rightarrow \boldsymbol{x} = 0$ (Vain nolla projektio itselleen on nolla)

Jotta yhtälö (3) määrittelee normin, tulee sille todistaa ominaisuuksien (iii)-(iv) lisäksi todistaa skaalausominaisuus sekä kolmioepäyhtälö. Nämä todistetaan tämän luvun loppupuolella.

Esimerkki 3.5. *Samaistetaan $\mathbb{R}^2$:n vektorit geometristen vektorien kanssa siten että $\boldsymbol{x} \in \mathbb{R}^2$ vastaa geometrista vektoria*

$$x_1 \boldsymbol{a} + x_2 \boldsymbol{b},$$

jossa $\boldsymbol{a}, \boldsymbol{b} \neq 0$ ja $\boldsymbol{a} \neq \lambda \boldsymbol{b} \forall \lambda \in \mathbb{R}$. Tässä tapauksessa luonnollinen valinta $\mathbb{R}^2$:n sisätuloksi on vastaavien geometristen vektorien pistetulo. Olkoot $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ ja valitaan

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = (x_1 \boldsymbol{a} + x_2 \boldsymbol{b}) \cdot (y_1 \boldsymbol{a} + y_2 \boldsymbol{b}).$$

Laskuharjoituksissa tarkistamme, että ylläoleva kuvaus toteuttaa ehdot (i)-(iv) ja on näinollen skalaaritulo. Myöhemmin kurssilla todistamme, että kaikki avaruuden $\mathbb{R}^2$ skalaaritulot voidaan muodostaa valitsemalla kaikki mahdolliset geometriset vektori $\boldsymbol{a}, \boldsymbol{b}$.

Ehdot (i)-(iv) täyttävälle kuvaukselle voidaan todistaa samankaltainen projektio-ominaisuus kuin pistetulolle. Tarkastellaan minimointitehtävää

$$\min_{\alpha \in \mathbb{R}} \left\| \boldsymbol{x} - \alpha \frac{\boldsymbol{y}}{|\boldsymbol{y}|} \right\|^2.$$

Derivoimalla ylläolevaa lauseketta, nähdään että minimi saavutetaan kun $\alpha = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|}$. Kyseessä on eräs vektorin $\boldsymbol{x}$ projektio vektorin $\boldsymbol{y}$ suuntaan. Voidaankin sanoa, että yleistetty sisätulo määrittelee, mitä tarkoitetaan kun kaksi vektoria ovat toisiaan vastaan kohtisuorassa.

Skalaaritulon projektio-ominasuudesta seuraa Cauchy-Schwartzin epäyhtälö. Tämä epäyhtälön tulkinta on, että järkevästi määritellyn projektion tulee aina olla alkuperäistä vektoria lyhyempi.

Lemma 3.2. *Olkoot $\langle \cdot, \cdot \rangle$ sisätulo ja $\| \cdot \|$ siihen liittyvä normi. Tällöin jokaisella $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ pätee*

$$| \langle \boldsymbol{x}, \boldsymbol{y} \rangle | \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\|.$$

Todistus. Tarkastellaan sisätuloa $\left\langle \frac{x}{\|\boldsymbol{x}\|}, \frac{y}{\|\boldsymbol{y}\|} \right\rangle$. Normin määritelmästä seuraa, että

$$0 \leq \min_{\alpha \in \mathbb{R}^2} \left\| \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} - \alpha \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \right\|^2 = \min_{\alpha \in \mathbb{R}^2} \left(1 + \alpha^2 - 2\alpha \left\langle \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}, \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \right\rangle \right)$$

Koska minimointitehtävän ratkaisu on $\alpha = \left\langle \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}, \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \right\rangle$, saadaan ylläolevasta

$$0 \leq 1 - \left\langle \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}, \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \right\rangle^2.$$

Eli,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 \leq \|\boldsymbol{x}\|^2 \|\boldsymbol{y}\|^2.$$

□

Ominaisuudet (i)-(iv) yhdessä skaaritulon ja vektorin normin välisen yhteyden kanssa ovat erittäin tärkeitä identiteettejä joiden avulla voidaan todistaa helposti useita vektorilaskennan perustuloksia. Esimerkiksi kolmioepäyhtälön todistus seuraa suoraan laskusäännöistä sekä Cauchy-Schwartzin epäyhtälöstä.

Lemma 3.3. *Olkoot $\langle \cdot, \cdot \rangle$ sisätulo ja $\| \cdot \|$ siihen liittyvä normi. Tällöin jokaisella $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$ pätee*

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|.$$

Todistus. Käyttämällä sisätulon ja normin välistä yhteyttä sekä laskusääntöjä (i)-(ii) saadaan

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + 2 \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2,$$

Arvoimalla sisätuloa käyttämällä Cauchy-Schwartzin epäyhtälöä saamme

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 \leq \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| = (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2.$$

□

3.6 Suora

Lukiossa tason suorat esitettiin muodossa

$$ay + bx + c = 0. \tag{4}$$

Kun tätä esitysmuotoa käytetään geometristen tehtävien ratkaisuun, tulee erikoistapauksia kuten $y = 0$ tai $x = 0$ käsitellä erikseen. Erikoistapaukset voidaan yleensä sivuuttaa, kun avaruuden $\mathbb{R}^2$ suorat esitetään muodossa

$$\{ \, t\boldsymbol{a} + \boldsymbol{b} \mid t \in \mathbb{R} \, \},$$

jossa $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$.

Esimerkki 3.6. *Tarkastellaan suoraa (4). Etsitään suoralta kaksi pistettä. Oletetaan, että $a \neq 0$ ja asetetaan $x = 0$ sekä $x = 1$. Tällöin saadaan*

$$y = \frac{c}{a} \quad \text{ja} \quad y = \frac{c - b}{a}.$$

vektorin $\boldsymbol{a}$ *voidaan kirjoittaa muodossa*

$$\boldsymbol{a} = \left(1, \frac{c-b}{a}\right) - \left(0, \frac{c}{a}\right) = \left(1, -\frac{b}{a}\right)$$

ja vektori $\boldsymbol{b}$ *muodossa*

$$\boldsymbol{b} = \left(0, \frac{c}{a}\right).$$

Kukin suoran piste voidaan siis esittää muodossa

$$\left\{ t\left(1, -\frac{b}{a}\right) + \left(0, \frac{c}{a}\right) \;\middle|\; t \in \mathbb{R} \right\}.$$

Kun $\boldsymbol{b} = 0$, origon kautta kulkeva suora

$$V = \{ \, t\boldsymbol{a} \mid t \in \mathbb{R} \, \},$$

yhdessä normaalien $\mathbb{R}^2$:n laskusääntöjen kanssa muodostaa itsessään vektoriavaruuden. Tälläistä avaruutta kutsutaan aliavaruudeksi. Koska muodostettu voidaan samaistaa avaruuden $\mathbb{R}$ kanssa, sanotaan että sen dimensio on yksi. Aliavaruuksiin törmätään matematiikassa usein.

3.7 Avaruuden $\mathbb{R}^n$ vektorit

Koska avaruuden $\mathbb{R}^2$ vektorit eivät suoraan liity geometrisiin käsitteisiin, tämä vektorikäsite voidaan helposti yleistää korkeampiin ulottuvuuksiin. Avaruus $\mathbb{R}^n$ koostuu n alkion järjestetyistä listoista,

$$V = \{ \, (x_1, \ldots, x_n) \, | x_1, \ldots, x_n \in \mathbb{R} \, \}.$$

Laskutoimitukset määritellään avaruudessa $\mathbb{R}^n$ seuraavalla tavalla

Määritelmä 3.2. *Olkoot* s *mielivaltainen reaaliluku ja* $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$. *Tällöin*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$
$$s(x_1, \ldots, x_n) = (sx_1, \ldots, x_n).$$

Tapauksessa $n = 2$ ylläoleva määritelmä vastaa yhteenlaskun ja skaarilla kertomisen määritelmää avaruudessa $\mathbb{R}^2$. Samalla tavalla kun avaruus $\mathbb{R}^2$, voidaan avaruus $\mathbb{R}^n$ samaistaa esimerkiksi $n-1$-asteisten polynomien kanssa

Esimerkki 3.7. *Kuten aikaisemmassa esimerkissä (3.2) voidaan* $n-1$ *- asteen polynomit*

$$p(x) = \sum_{i=1}^{n} a_i x^{i-1}$$

samaistaa avaruuden $\mathbb{R}^n$ *vektorien kassa. Tällöin polynomit esitetään muodossa*

$$\boldsymbol{a} = (a_1, \ldots, a_n).$$

Kuten aikaisemmin, määritemän 3.2 mukaiset laskutoimitukset vastaavat polynomien yhteenlaskua tai skalaarilla kertomista.

3.8 Vektorien lineaariyhdistely, lineaarinen riippumattomuus ja kanta

Lineaarialgebrassa tutkitaan usein korkeaulotteisen avaruuden osajoukkoja, aliavaruuksia. Tarkastellaan vektoriavaruutta, joka koostuu vektorien joukosta V sekä yhteenlaskun ja skalaarilla kertomisen määritelmistä. Olkoot $W \subset V$. Jos W on suljettu V:ssä määritellyn yhteenlaskun ja skalaarilla kertomisen suhteen, kyseessä on V:hen liittyvän vektoriavaruuden aliavaruus.

Olkoot $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^n$ ja $\alpha_1, \ldots, \alpha_N$ skalaareita. Tällöin vektorien $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ lineaarikombinaatio tai lineaariyhdistely on vektori

$$\sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i \in \mathbb{R}^n.$$

Lineaariyhdistelyä apuna käyttämällä voidaan esittää avaruuden $\mathbb{R}^n$ aliavaruuksia. Määritellään $\mathbb{R}^n$ - osajoukko,

$$\left\{ \sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i \mid \alpha_1, \ldots, \alpha_N \in \mathbb{R} \right\}.$$

Tämä osajoukko on selvästi suljettu $\mathbb{R}^n$ yhteenlaskun ja skalaarilla kertomisen suhteen joten se on $\mathbb{R}^n$:n aliavaruus. Vektorien $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ muodostamasta joukosta käytetään lineaarialgebrassa merkintää

$$\operatorname{span}\{ \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \} := \left\{ \sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i \mid \alpha_1, \ldots, \alpha_N \in \mathbb{R} \right\}.$$

Esimerkki 3.8. *Tarkastellaan 2 -asteen polynomiavaruutta. Polynomit, joiden vakiotermi on nolla voidaan esittää vektorien*

$$(1,0,0) \text{ ja } (0,1,0)$$

lineaarikombinaationa.

Esimerkki 3.9. *Tarkastellaan 2 -asteen polynomiavaruutta. Kaikki polynomit, jotka voidaan kirjoittaa muodossa $\beta(x-2)(x-\alpha)$ voidaan esittää vektorien*

$$(1,-2,0) \text{ ja } (0,1,-2)$$

lineaarikombinaationa.

Vektorien $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ lineaariyhdistely saattaa tuottaa saman vektorin eri kertoimilla $\alpha_1, \ldots, \alpha_2$. Tällöin vektoreita kutsutaan lineaarisesti riippuviksi. Aliavaruudet esitetään lähes aina käyttämällä lineaarisesti riippumattomia vektoreita.

Määritelmä 3.3. *Vektorit* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^n$ *ovat lineaarisesti riippumattomia, jos pätee*

$$\sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i = 0 \Rightarrow a_i = 0 \quad \forall i = 1, \ldots, N \tag{5}$$

Esimerkki 3.10. *Olkoot* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$. *Tutkitaan, millä ehdolla* $\boldsymbol{a}$ *ja* $\boldsymbol{b}$ *ovat lineaarisesti riippumattomia. Ehto (5) voidaan kirjoittaa muodossa*

$$\alpha_1 \boldsymbol{a} + \alpha_2 \boldsymbol{b} = 0 \Rightarrow \alpha_1 = 0 \quad \textit{ja} \quad \alpha_2 = 0.$$

Tämä yhtälö määrittää kahden origon kautta kulkevan suoran,

$$\{\, \alpha_1 \boldsymbol{a} \mid \alpha_1 \in \mathbb{R} \,\} \quad \textit{ja} \quad \{\, \alpha_2 \boldsymbol{b} \mid \alpha_2 \in \mathbb{R} \,\}$$

leikkauspisteen. Tehtävä voidaan ratkaista geometrisesti :

(i). Suorat leikkaavat vain origossa

(ii). Suorat ovat yhdensuuntaiset ja leikkaavat kaikissa pisteissä.

Näinollen, vektorit ovat lineaarisesti riippumattomia, jos ehdot $\boldsymbol{a}, \boldsymbol{b} \neq 0$ *ja* $\boldsymbol{a} \neq \lambda \boldsymbol{b} \quad \forall \lambda \in \mathbb{R}$ *pätevät.*

Palaamme myöhemmin lineaarisen riippumattomuuden käsitteeseen ja liitämme sen yhtälöryhmän ratkaisun olemassaoloon. Seuraavassa esitämme $\mathbb{R}^n$ vektorit usein tietyn kanna avulla. Kannalla tarkoitetaan n lineaarisesti riippumatonta vektoria, joiden lineaarikombinaationa voimme kirjoittaa kunkin $\mathbb{R}^n$ alkion.

Yksinkertaisin $\mathbb{R}^n$ kanta on ns. luonnollinen kanta, $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N$, jolle pätee

$$(\boldsymbol{e}_i)_j = \begin{cases} 1 & , i = j \\ 0 & \text{muulloin} \end{cases}.$$

Kyseessä on selvästi lineaarisesti riippumaton joukko vektoreita.

3.9 Sisätulo ja normi

Määrittelimme aikaisemmin avaruudessa $\mathbb{R}^2$ vektorin pituuden geometrisen argumentin kautta. Ajatellaan tässä vektorin pituutta kuvauksena, joka liittää kuhunkin avaruuden $\mathbb{R}^n$ alkioon reaaliluvun. Jotta kyseessä on järkevä pituusmitta, vaaditaan tältä kuvaukselta samat ominausuudet (i)-(iv) kuin Luvussa 3.3. Kuten aikaisemmin tälläistä kuvausta kutsutaan normiksi.

Tällä kurssilla tarvitsemme käytämme pääasiassa Euklidista normia, joka on geometrisesti määritellyn vektorin pituusmitan yleistys. Euklidista normia käytetään, kun ei ole tietoa siitä mihin käsiteltävät vektorit liittyvät.

Määritelmä 3.4. *Vektorin* $\boldsymbol{x} \in \mathbb{R}^n$ *Euklidinen normi* $\|\cdot\|_2 : \mathbb{R}^n \to \mathbb{R}$ *on kuvaus*

$$\|\boldsymbol{x}\|_2 = \left(\sum x_i^2\right)^{1/2}.$$

Tapauksessa $n = 2$ Euklidinen normi vastaa avaruudessa $\mathbb{R}^2$ määriteltyä vektorin pituusmittaa. Kun määritellään ensin määritellään pistetulon yleistys Eulkidinen sisätulo,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_2 := \sum_i x_i y_i,$$

Voidaan Normilta vaadittavat ominaisuudet (i)-(iv) voidaan todistaa samalla menettelyllä, kun luvussa 3.5. Koska käytämme seuraavassa lähes ainoastaa Euklidista sisätuloa, käytetään siitä tästä lähtien merkintää $\langle \cdot, \cdot \rangle$. Kuten avaruudessa $\mathbb{R}^2$, sisätulon tehtävänä on luoda avaruuteen $\mathbb{R}^n$ kohtisuoruuden käsite. Kun $n = 2$, vastaa ylläoleva määritelmä pistetuloa. Kuten aikaisemmin, Euklidisen normin ja sisätulon välillä on yhteys

$$\|\boldsymbol{x}\|_2^2 = \langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

4 Matriisi

Kuten vektorilla, matriisilla on monta erilaista käyttötarkoitusta ja tulkintaa. Matriisi voi esimerkiksi toimia lineaarisen yhtälöryhmän ”kirjoitusmuotona” tai esittää avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^m$ välisiä lineaarikuvauksia. Aloitetaan matriiseihin tutustuminen tarkastelemalla yhtälöryhmää

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m & = b_2 \\ \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m & = b_n \end{cases} \tag{6}$$

joka koostuu n-yhtälöstä ja m-tuntemattomasta x_i. Kaikki yhtälöryhmän (6) sisältämä informaatio voidaan lukea tuntemattomien $x_1, x_2, \ldots x_m$ kertoimista sekä oikean käden arvoista. Yhtälöryhmä (6) voidaan siis esittää esimerkiksi kaksiulotteisen $n \times m$ - taulukon ja n - alkion listan

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

avulla.

Matriisien avulla voidaan esittää lineaarisia kuvauksia avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^m$ välillä. Tälläiset kuvaukset liittävät kuhunkin $\mathbb{R}^n$:n vektoriin $\mathbb{R}^m$:n vektorin sekä toteuttavat Määritelmän 4.1 mukaiset ominaisuudet.

Määritelmä 4.1. *Olkoot kuvaus $f : \mathbb{R}^n \to \mathbb{R}^m$ lineaarinen. Tällöin jokaisella $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ ja reaaliluvulla α, β pätee*

$$f(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y}).$$

Esimerkki 4.1. *Tarkastellaan kuvausta $f : \mathbb{R}^2 \to \mathbb{R}^2$, joka kiertää koordinaatistoa θ-astetta. Tälläinen kuvaus on havainnollistettu Kuvassa 7. Kierrossa piste (x_1, x_2) kuvataan pisteeksi (y_1, y_2). Koordinaatiston kierrossa näitä kahta pistettä esittävien vektorien välinen kulma sekä pituus säilyy, eli*

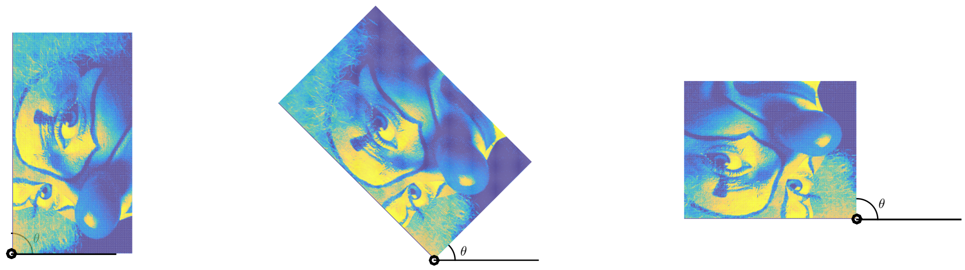
$$\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle \quad \text{ja} \quad \|f(\boldsymbol{x})\| = \|\boldsymbol{x}\|, \|f(\boldsymbol{y})\| = \|\boldsymbol{y}\| \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2. \tag{7}$$

Todistetaan, että koordinaatiston kierto on lineaarinen kuvaus.

Lause 4.1. *Olkoot $f : \mathbb{R}^n \to \mathbb{R}^m$ kuvaus siten, että*

$$\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle \quad \text{ja} \quad \|f(\boldsymbol{x})\| = \|\boldsymbol{x}\|, \|f(\boldsymbol{y})\| = \|\boldsymbol{y}\| \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n. \tag{8}$$

Tällöin f on lineaarinen.



Kuva 7: Koordinaatiston kierto : 0, 45 ja 90 astetta.

Todistus. Ehto kuvauksen f lineaarisuudesta on yhtäpitävää ehtojen

$$\|f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})-f(\boldsymbol{y})\|^2 = 0 \quad \forall \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n \tag{9}$$

ja

$$\|f(\alpha\boldsymbol{x})-\alpha f(\boldsymbol{x})\|^2 = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n, \alpha \in \mathbb{R}. \tag{10}$$

kanssa. Koska tiedämme tässä vaiheessa kuvauksesta f ainoastaan ominaisuudet (8) soveltuvat ylläolevat ehdot lineaarisuuden määritelmää paremmin tähän tehtävään. Määritelmän perusteella saadaan

$$\|f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})-f(\boldsymbol{y})\|^2 = \|f(\boldsymbol{x}+\boldsymbol{y})\|^2 - 2\left\langle f(\boldsymbol{x}+\boldsymbol{y}), f(\boldsymbol{x})+f(\boldsymbol{y})\right\rangle + \|f(\boldsymbol{x})+f(\boldsymbol{y})\|^2.$$

Käyttämällä ominaisuuksia (8) sekä normin laskusääntöjä nähdään, että

$$\begin{aligned}
\|f(\boldsymbol{x}+\boldsymbol{y})\|^2 &= \|\boldsymbol{x}+\boldsymbol{y}\|^2\\
\langle f(\boldsymbol{x}+\boldsymbol{y}), f(\boldsymbol{x})+f(\boldsymbol{y})\rangle &= \|\boldsymbol{x}+\boldsymbol{y}\|^2\\
\|f(\boldsymbol{x})+f(\boldsymbol{y})\|^2 &= \|\boldsymbol{x}+\boldsymbol{y}\|^2.
\end{aligned}$$

Joka todistaa ehdon 9 . Ehto 10 voidaan todistaa saman kaltaisella laskulla. □

Kaikki avaruuden $\mathbb{R}^n$ sekä $\mathbb{R}^m$ vektorit voidaan esittää kantavektorien lineaarikombinaationa. Olkoot $\{\boldsymbol{e}_i\}$ avaruuden $\mathbb{R}^n$ luonnollinen kanta ja $\{\boldsymbol{v}_i\}$ avaruuden $\mathbb{R}^m$ luonnollinen kanta. Tällöin kukin avaruuden $\mathbb{R}^n$ vektori voidaan kirjoittaa vektorien $\{\boldsymbol{e}_i\}$ lineaarikombinaationa,

$$\boldsymbol{x} = \sum_i^n x_i \boldsymbol{e}_i.$$

Huomataan, että pätee $\langle \boldsymbol{x}, \boldsymbol{e}_i\rangle = x_i$, eli vektorin $\boldsymbol{x}$ i:s kompontti saadaan laskemalla vektorin $\boldsymbol{x}$ sisätulo vektorin $\boldsymbol{e}_i$ kanssa. Toinen tapa kirjoittaa $\mathbb{R}^n$ - vektorit on

$$\boldsymbol{x} = \sum_i^n \langle \boldsymbol{x}, \boldsymbol{e}_i\rangle \boldsymbol{e}_i. \tag{11}$$

Tämä esitystapa pätee aina, kun käytetään luonnollista kantaa. Kun vektoriin $\boldsymbol{x}$ operoidaan lineaarikuvauksella f saadaan

$$f(\boldsymbol{x}) = f(\sum_{i=1}^{n} x_i \boldsymbol{e}_i).$$

Käyttämällä funktion f - lineaarisuutta $n-1$ - kertaa saadaan

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} x_i f(\boldsymbol{e}_i).$$

Kun tämä vektori esitetään $\mathbb{R}^m$:n luonnollisessa kannassa käyttämällä esitysmuotoa (11), saadaan

$$f(\boldsymbol{x}) = \sum_{j=1}^{m} \sum_{i=1}^{n} x_i \langle f(\boldsymbol{e}_i), \boldsymbol{v_j} \rangle \boldsymbol{v_j}.$$

Lineaarikuvaus f tunnetaan täysin, kun tunnetaan alkiot $\langle f(\boldsymbol{e}_i), \boldsymbol{v_j} \rangle$. Luonnollinen tapa esittää lineaarikuvausta on jälleen $m \times n$- taulukko

$$\begin{bmatrix} \langle f(\boldsymbol{e}_1), \boldsymbol{v_1} \rangle & \langle f(\boldsymbol{e}_2), \boldsymbol{v_1} \rangle & \cdots & \langle f(\boldsymbol{e}_n), \boldsymbol{v_1} \rangle \\ \vdots & \vdots & & \vdots \\ \langle f(\boldsymbol{e}_1), \boldsymbol{v_m} \rangle & \langle f(\boldsymbol{e}_2), \boldsymbol{v_m} \rangle & \cdots & \langle f(\boldsymbol{e}_n), \boldsymbol{v_m} \rangle \end{bmatrix}.$$

Kun määrittelemme lineaarikuvauksia esittävien taulukoiden välille yhteenlaskun, skalaarilla kertomisen ja kertolaskun, voimme muodostaa lineaarikuvausten summia sekä yhdistettyjä kuvauksia helposti. Tarkastellaan seuraavassa kaksiulotteisista $n \times m$-taulukoista koostuvan vektoriavaruutta $\mathbb{R}^{n \times m}$. Tämän avaruuden alkioita kutsutaan matriiseiksi. Tavallisten vektoriavaruuden operaatioiden lisäksi määrittelemme matriisien välisen tulon. Matriisien väliset laskuoperaatiot ovat luonnollisia, kun samaistamme $\mathbb{R}^{n \times m}$ matriisit avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^m$ välisten lineaarikuvausten kanssa.

Esimerkki 4.2. *Tarkastellaan koordinaatiston kiertoa θ-astetta. Olemme aikaisemmin todistaneet, että kyseessä on lineaarinen kuvaus $f : \mathbb{R}^2 \to \mathbb{R}^2$. Kiertoon liittyvä matriisi $A \in \mathbb{R}^{2 \times 2}$ voidaan määrittää laskemalla arvot*

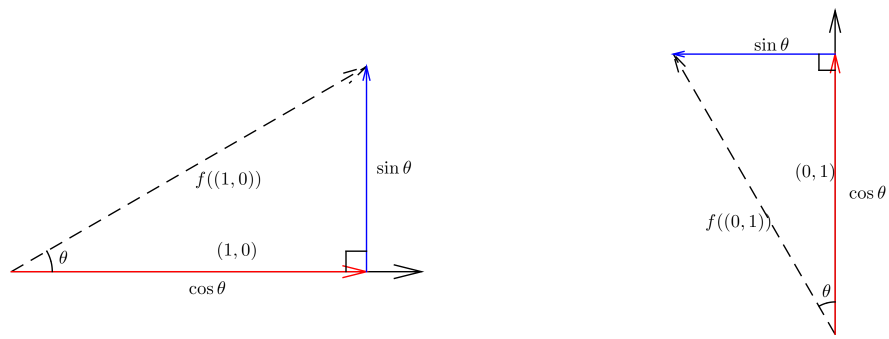
$$f((1,0)) \quad \textit{ja} \quad f((0,1)).$$

Nämä arvot voidaan laskea geometrisen argumentin avulla, Kuva 8. Saadaan

$$f((1,0)) = (\cos\theta, \sin\theta) \quad \textit{and} \quad f((0,1)) = (-\sin\theta, \cos\theta).$$

Kierto voidaan kirjoittaa matriisimuodossa:

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$



Kuva 8: Vektorien $(1,0)$ ja $(0,1)$ kierto $\pi/6$-astetta.

Määritelmä 4.2. *Olkoot* $A \in \mathbb{R}^{n\times m}$ *ja* $B \in \mathbb{R}^{m\times p}$. *Tällöin matriisien tulo* $AB \in \mathbb{R}^{n\times p}$ *on määritelty siten että*

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

Ylläolevaa määritelmää ei yleensä käytetä matriisien tulon laskentaa. Tämän sijaan käytetään muistisääntöjä. Määritelmä 4.2 on hyödyllinen, kun todistetaan matriisituloon liittyviä ominaisuuksia.

Määritelmä 4.3. *Olkoot* $A, B \in \mathbb{R}^{n\times m}$ *ja* s *mielivaltainen reaaliluku. Tällöin* $A + B \in \mathbb{R}^{n\times m}$

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

ja

$$(sA)_{ij} = sA_{ij}.$$

Matriisitulon määritelmä on valittu siten, että sen avulla voidaan muodostaa helposti yhdistettyjä kuvauksia.

Lemma 4.1. *Olkoot* f, g *lineaarikuvauksia joita esittävät matriisit* A *ja* B. *Tällöin yhdistetty kuvaus* $f \circ g$ *on lineaarikuvaus, jota esittää matriisi*

$$AB$$

Yhdessä Määritelmän 4.3 kanssa matriisit muodostavat vektoriavaruuden. Matriisien yhteenlasku ja skalaarilla kertominen ”toimivat” kuten reaaliluvuilla. Sama ei päde kahden matriisin kertolaskulle. Olkoot $A, B \in \mathbb{R}^{n\times m}$ ja $C \in \mathbb{R}^{m\times p}$ sekä $\alpha, \beta \in \mathbb{R}$. Tällöin pätee

$$(\alpha A + \beta B)C = \alpha(AC) + \beta(BC).$$

Matriisitulo ei kommutoi. Olkoot $A, B \in \mathbb{R}^{n\times n}$. Tällöin

$$AB = BA$$

ei päde jokaisella A, B.

Avaruuden $\mathbb{R}^n$ vektorit voidaan esittää joko avaruuden $\mathbb{R}^{n\times 1}$ tai $\mathbb{R}^{1\times n}$ matriisena. Tällöin puhutaan joko sarake- tai rivivektoreista. Tässä luentomonisteessa vektorit esitetään tästä lähtien sarakevektoreina. Monessa tilanteessa matriisien väliset operaatiot voidaan ymmärtää helpommin, jos matriisit $A \in \mathbb{R}^{n\times m}$ kirjoitetaan sarakevektorien avulla muodossa

$$A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{bmatrix} \tag{12}$$

jossa kukin $\boldsymbol{a}_i \in \mathbb{R}^{n\times 1}$. Matriisitulon määritelmästä seuraa suoraan, tulkinta matriisin ja vektorin väliselle tulolle,

$$A\boldsymbol{x} = \sum_{i}^{n} \boldsymbol{a}_i x_i. \tag{13}$$

Kyseessään on matriisin A sarakevektorien lineaariyhdistely.

Esimerkki 4.3. *Matriisin ja vektorin tulo (13) on helppo muistaa matriisin sarakevektorien lineaariyhdistelynä. Kun lasketaan matriisi-matriisi tuloja, voidaan käyttää seuraavaa menettelyä. Olkoot $A \in \mathbb{R}^{n\times m}$ ja $B \in \mathbb{R}^{m\times p}$.*

$$A = \begin{bmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_m \end{bmatrix} \quad ja \quad B = \begin{bmatrix} \boldsymbol{b}_1 & \dots & \boldsymbol{b}_p \end{bmatrix}$$

Matriisi-Matriisi tulolle (harjoitustehtävä) pätee

$$AB = \begin{bmatrix} A\boldsymbol{b}_1 & \dots & A\boldsymbol{b}_p \end{bmatrix}.$$

Ylläoleva tarjoaa yhden muistisäännön matriisitulon laskentaan.

Esimerkki 4.4. *Olkoot*

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

Tällöin

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

Lasketaan ylläolevat matriisi-vektorin tulot. Esimerkiksi,

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

Saadaan,

$$AB = \begin{bmatrix}2 & 0 & 1\\2 & 2 & 0\\0 & 0 & 1\end{bmatrix} \quad ja \quad BA = \begin{bmatrix}1 & 0 & 0\\1 & 2 & 4\\1 & 0 & 2\end{bmatrix}$$

Koska $AB \neq BA$, matriisitulo ei kommutoi.

Kuhunkin matriisiin A liittyy sen transpoosi A^T.

Määritelmä 4.4. *Olkoot $A \in \mathbb{R}^{n\times m}$. Tällöin A:n transpoosi on matriisi $A^T \in \mathbb{R}^{m\times n}$, siten että*

$$A^T_{ij} = A_{ji} \quad i = 1, \dots, m \text{ ja } j = 1, \dots, n.$$

Kahden vektorin $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ välinen pistetulo voidaan kirjoittaa transpoosin sekä matriisitulon avulla muodossa

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \boldsymbol{y}.$$

Vektorin Euklidiselle normille pätee siis, $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$. Matriisin transpoosi liittyy sisätuloon. Olkoot $A \in \mathbb{R}^{n\times m}, \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m$. Tällöin pätee

$$\langle A\boldsymbol{x}, \boldsymbol{y} \rangle = \left\langle \boldsymbol{x}, A^T\boldsymbol{y} \right\rangle.$$

Matriisin ja vektorin tulo voidaan tulkita sarjana pistetuloja. Kirjoitetaan matriisi muodossa

$$A = \begin{bmatrix}\boldsymbol{a}_1^T\\ \boldsymbol{a}_2^T\\ \vdots\\ \boldsymbol{a}_m^T\end{bmatrix}.$$

jossa kukin $\boldsymbol{a}_i \in \mathbb{R}^{m\times 1}$. Tällöin matriisitulo vastaa operaatioa

$$A\boldsymbol{x} = \begin{bmatrix}\boldsymbol{a}_1^T\boldsymbol{x}\\ \boldsymbol{a}_2^T\boldsymbol{x}\\ \vdots\\ \boldsymbol{a}_m^T\boldsymbol{x}\end{bmatrix}$$

Kyseessä on siis kaikkien matriisin A rivivektorien ja vektori $\boldsymbol{x}$ välisistä pistetuloista.

Osalla matriiseista on erityinen tulkinta. Avaruuksien $\mathbb{R}^n$ ja $\mathbb{R}^n$ väliseen identiteettikuvaukseen liittyvää matriisia merkitään symbolilla I. Kyseessä on matriisi jolle pätee $I \in \mathbb{R}^{n\times n}, I_{ij} = \delta_{ij}$. Merkitään nollalla matriisia jonka jokainen alkio on nolla. Kyseessä on yhteenlaskun nolla-alkio. Identiteettialkiolle pätee,

$$AI = IA = A. \forall A \in \mathbb{R}^{n\times n}.$$

Matriisia $A \in \mathbb{R}^{n\times n}$ kutsutaan symmetriseksi, jos $A = A^T$.

5 Lineaarisen yhtälöryhmän ratkaiseminen

Olkoot $A \in \mathbb{R}^{n\times n}$ ja $\boldsymbol{b} \in \mathbb{R}^n$. Tässä luvussa tarkastellaan yhtälöryhmän : Etsi $\boldsymbol{x} \in \mathbb{R}^n$ siten että

$$A\boldsymbol{x} = \boldsymbol{b} \tag{14}$$

ratkaisemista. Tutustumme ensin Gaussin eliminaatiomenetelmään, jossa alkuperäisestä yhtälöryhmästä muodostetaan uusi, helposti ratkaistava yhtälöryhmä. Tämän jälkeen keskustelemme ratkaisun olemassaolosta, LU-hajotelmasta sekä singulaarisen yhtälöryhmän ratkaisemisesta.

5.1 Gaussin eliminaation

Gaussin eliminaatio on algoritmi lineaarisen yhtälöryhmän ratkaisemiseen. Algoritmi muuntaa alkuperäisen yhtälöryhmän (14) helposti ratkaistavaan muotoon : Etsi $\boldsymbol{x} \in \mathbb{R}^n$ siten, että

$$U\boldsymbol{x} = \tilde{\boldsymbol{b}},$$

jossa kerroinmatriisi $U \in \mathbb{R}^{n\times n}$ on yläkolmiomatriisi ja $\tilde{\boldsymbol{b}} \in \mathbb{R}^n$.

Määritelmä 5.1. *Matriisi $U \in \mathbb{R}^{n\times n}$ on yläkolmiomatriisi jos*

$$U_{ij} = 0, \text{ kun } i > j.$$

Yhtälöryhmät, joiden kerroinmatriisi on yläkolmiomatriisiin voidaan ratkaista käyttämällä takaisinsijoitus-menettelyä. Kirjoitetaan matriisi U blokkimuodossa,

$$\begin{bmatrix} \tilde{U} & \boldsymbol{u}_{12} \\ 0 & u_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ b_2 \end{bmatrix}$$

jossa x_2, b_2 ja u_{22} ovat skalaareja, $\tilde{U} \in \mathbb{R}^{(n-1)\times(n-1)}$ ja $\boldsymbol{u}_{12}, \boldsymbol{x}_1, \boldsymbol{b}_1 \in \mathbb{R}^{n-1}$. Muuttuja x_2 voidaan ratkaista yhtälöryhmästä suoraan. Saadaan,

$$x_2 = \frac{b_2}{u_{22}}.$$

Suorittamalla vetorin ja blokkimatriisin tulo saadaa

$$\begin{bmatrix} \tilde{U}\boldsymbol{x}_1 + x_2\boldsymbol{u}_{12} \\ x_2 u_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ b_2 \end{bmatrix}$$

Ensimmäinen yhtälö voidaan kirjoittaa muodossa

$$\tilde{U}\boldsymbol{x}_1 = \boldsymbol{b}_1 - \boldsymbol{u}_{12}x_2$$

Kyseessä on uusi yhtälöryhmä, jonka kerroinmatriisi $\tilde{U} \in \mathbb{R}^{(n-1)\times(n-1)}$ on yläkolmiomatriisi. Tämä yhtälöryhmä voidaan ratkaista toistamalla ylläolevaa menettelyä. Esimerkiksi Matlabissa takaisinsijoitus voidaan toteuttaa muutamalla rivillä,

```
function x = triu_solve(U,b)

N = size(U,2);

x = zeros(N,1);

for i=N:-1:1
    x(i) = b(i)/U(i,i);
    b(1:(i-1)) = b(1:(i-1)) - U(1:(i-1),i)*x(i);
end
```

Tarkastellaan Gaussin eliminaatioa yhtälöryhmän

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2m}x_m &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3m}x_m &= b_3 . \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nm}x_m &= b_n \end{aligned}$$

ratkaisemiseksi. Oletetaan yksinkertaisuuden vuoksi, että $a_{11} \neq 0$. Lukiossa yhtälöryhmän ratkaisemiseksi käytettiin joko sijoitus- tai yhteenlaskumenetelmää. Sijoitusmenetelmässä ratkaisemme ensin yhtälöstä 1. muuttujan x_1 muodossa

$$x_1 = \frac{b_1}{a_{11}} - \sum_{j=2}^{m} \frac{a_{1j}}{a_{11}} x_j.$$

Käyttämällä tätä lauseketta muuttuja x_1 voidaan eliminoida yhtälöistä $2, \ldots, n$. Tällöin saadaan uusi yhtälöryhmä,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1m}x_m &= b_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2m}^{(1)}x_m &= b_2^{(1)} \\ a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 + \ldots + a_{3m}^{(1)}x_m &= b_3^{(1)} \\ &\vdots \\ a_{n2}^{(1)}x_2 + a_{n3}^{(1)}x_3 + \ldots + a_{nm}^{(1)}x_m &= b_n^{(1)}, \end{aligned}$$

jonka kertoimet $a_{ij}^{(1)}$ ovat

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j} \quad i, j > 1.$$

Yllä suoritettu operaatio voidaan myös ymmärtää siten, että yhtälöryhmän yhtälö 1 on kerrottu termillä $-a_{i1}a_{11}^{-1}$ ja laskettu yhteen yhtälön i kanssa. Huomaa, etteivat yhtälöt

$2, \ldots, m$ riipu enää muuttujasta x_1. Muuttujat $x_2, \ldots, x_n$ voidaan ratkaista yhtälöryhmästä

$$\begin{aligned} a_{22}^{(1)} x_2 + a_{23}^{(1)} x_3 + \ldots + a_{2m}^{(1)} x_m &= b_2^{(1)} \\ a_{32}^{(1)} x_2 + a_{33}^{(1)} x_3 + \ldots + a_{3m}^{(1)} x_m &= b_3^{(1)} \\ &\vdots \\ a_{n2}^{(1)} x_2 + a_{n3}^{(1)} x_3 + \ldots + a_{nm}^{(1)} x_m &= b_n^{(1)}, \end{aligned} .$$

Tämän jälkeen muuttuja x_1 ratkaistaan yhtälöstä

$$a_{11} x_1 = b_1 - (a_{12} x_2 + a_{13} x_3 + \ldots + a_{1m} x_m) .$$

Olemme suorittaneet yhden askeleen Gaussin eliminaatioa. Oletetaan, että $a_{22}^{(1)} \neq 0$ ja jatketaan sijoitusmenettelyä yhtälöstä 2. Muuttuja x_2 voidaan ratkaista muodossa

$$x_2 = \frac{b_{22}^{(1)}}{a_{22}^{(1)}} - \sum_{i=3}^{m} \frac{a_{2i}^{(1)}}{a_{22}^{(1)}} x_i.$$

Kuten edellä, muuttuja x_2 voidaan eliminoida yhtälöistä $3, \ldots, n$. Uudet kertoimet lasketaan kuten edellä :

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} a_{2j} \quad i, j > 2$$

Jos oletetaan, että $a_{jj}^{(i)} \neq 0$ toistamalla tätä menettely, yhtälöryhmä voidaan muunta muotoon :

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + a_{13} x_3 + \ldots + a_{1m} x_m &= b_1 \\ a_{22}^{(1)} x_2 + a_{23}^{(1)} x_3 + \ldots + a_{2m}^{(1)} x_m &= b_2^{(1)} \\ a_{33}^{(2)} x_3 + \ldots + a_{3m}^{(2)} x_m &= b_3^{(2)} \\ &\vdots \\ a_{nm}^{(n-1)} x_m &= b_n^{(n-1)} \end{aligned}$$

Ylläolevan yhtälöryhmän diagonaalialkioita kutsutaan seuraavassa tukialkioiksi. Gaussin eliminaatioa suoritettaessa ei yleensä kirjoiteta muuttujia $x_1, \ldots, x_n$ vaan eliminaatio suoritetaan matriisimuodossa.

Esimerkki 5.1. *Tarkastellaan yhtälöryhmää*

$$\begin{cases} x_1 + x_2 + x_3 & = 0 \\ x_1 + 2x_2 + 4x_3 & = 1 \\ x_1 + 3x_2 + 2x_3 & = 7. \end{cases}$$

Tämä yhtälöryhmä voidaan kirjoittaa matriisimuodossa: Etsi $\boldsymbol{x} \in \mathbb{R}^3$ siten, että

$$A\boldsymbol{x} = \boldsymbol{b}, \quad \textit{jossa} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 2 \end{bmatrix} \quad \textit{ja} \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ 1 \\ 7 \end{bmatrix}.$$

Kun Gaussin eliminaatioa suoritetaan käsin, matriisi A ja vektori $\boldsymbol{b}$ kirjoitetaan samaan taulukkoon rinnakkain.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 1 \\ 1 & 3 & 2 & 7 \end{array}\right]$$

Eliminaatiossa suoritetut rivioperaatiot merkitään matriisin vasemmalle puolelle.

$$\begin{array}{c} \\ -Y1 \\ -Y1 \end{array} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 1 \\ 1 & 3 & 2 & 7 \end{array}\right] \rightarrow \quad \begin{array}{c} \\ \\ -2Y2 \end{array} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 2 & 1 & 7 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -5 & 5 \end{array}\right]$$

Gaussin eliminaation tuottama yhtälöryhmä voidaan ratkaista takaisinsijoitusta käyttämällä.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -5 & 5 \end{array}\right] \stackrel{x_3=1}{\rightarrow} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 4 \end{array}\right] \stackrel{x_2=4}{\rightarrow} \left[\begin{array}{c|c} 1 & -3 \end{array}\right] \rightarrow x_1 = -3.$$

Alkuperäisen yhtälöryhmän ratkaisu on näinollen $\boldsymbol{x} = \begin{bmatrix} -3 & 4 & -1 \end{bmatrix}^T$.

Gaussin eliminaatio on helppo toteuttaa esimerkiksi Matlabissa.

```
%
% Gaussin eliminaatio.
%

function [A,b] = gauss(A,b)

for i = 1:size(A,2)

    for j = (i+1):size(A,2)

        % kerroin riville j
        cof = A(j,i)/A(i,i);

        A(j,i:end) = A(j,i:end) - A(i,i:end)*cof
        b(j) = b(j) - b(i)*cof
    end
end
```

5.2 Tuenta

Gaussin eliminaatiossa tukialkio voi saada arvon 0, vaikka yhtälöryhmällä on olemassa ratkaisu. Kun näin tapahtuu, suoritetaan rivien tai sarakkeiden vaihto.

Esimerkki 5.2. *Tarkastellaan yhtälöryhmää*

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

Tämä yhtälöryhmä voidaan kirjoittaa muodossa

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 4 & 3 \\ 1 & 3 & 2 & 7 \end{array}\right]$$

Suoritetaan ensimmäinen askel Gaussin eliminaatioa :

$$\begin{array}{c} \\ -Y1 \\ -Y1 \end{array} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 4 & 3 \\ 1 & 3 & 2 & 7 \end{array}\right] \rightarrow \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 2 & 1 & 7 \end{array}\right]$$

Koska tukialkio on nolla, vaihdetaan rivien 2 *ja* 3 *paikkaa.*

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 7 \\ 0 & 0 & 3 & 3 \end{array}\right]$$

Alkuperäinen kerroinmatriisi on nyt muunnettu yläkolmiomuotoon ja yhtälöryhmä voidaan ratkaista takaisinsijoituksella.

Matriisin $A \in \mathbb{R}^{n\times m}$ rivien tai sarakkeiden vaihto on helppo esittää matriisimuodossa. Tälläistä esitystä käytetään esimerkiksi Matlabissa. Huomataan ensin, että matriisin A i:s sarake ja j:s rivi voidaan laskea operaatiolla

$$A\boldsymbol{e}_i \quad \text{ja} \quad \boldsymbol{e}_j^T A.$$

Vektorin $\boldsymbol{x}$ alkioiden p ja q vaihto on lineaarikuvaus

$$f(\boldsymbol{x})_i = \begin{cases} x_q & i = p \\ x_p & i = q \\ x_i & \text{muulloin} \end{cases}.$$

Tälläiseen kuvaukseen liittyvää matriisia P kutsutaan permutaatiomatriisiksi. Matriisin rivien p ja q paikka voidaan vaihtaa permutaatiomatriisin P avulla. Matriisille AP pätee

$$AP\boldsymbol{e}_i = Af(\boldsymbol{e}_i) = \begin{cases} A\boldsymbol{e}_q & i = p \\ A\boldsymbol{e}_p & i = q \\ A\boldsymbol{e}_i & \text{muulloin} \end{cases}.$$

Samalla tavalla matriisin A sarakkeet voidaan permutoida operaatiolla $P^T A$.

$$\boldsymbol{e}_i^T P^T A = (P\boldsymbol{e}_i)^T A = \begin{cases} \boldsymbol{e}_q^T A & i = p \\ \boldsymbol{e}_q^T A & i = q \\ \boldsymbol{e}_i^T A & \text{muulloin} \end{cases}.$$

Esimerkki 5.3. *Esimerkissä 5.2 vaihdoimme 3×3-matriisin rivien 2 ja 3 paikkaa. Tämä vastaa lineaarikuvausta $f : \mathbb{R}^3 \to \mathbb{R}^3$,*

$$f(\boldsymbol{x}) = \begin{bmatrix} x_1 \\ x_3 \\ x_2 \end{bmatrix}$$

Permutaatiomatriisi P on tällöin,

$$P = \begin{bmatrix} f(\boldsymbol{e}_1) & f(\boldsymbol{e}_2) & f(\boldsymbol{e}_3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Esimerkissä 5.2 Gaussin eliminaatio suoritettiin siis yhtälöryhmälle

$$P^T \boldsymbol{A}\boldsymbol{x} = P^T \boldsymbol{b}.$$

Gaussin eliminaatiossa rivien ja sarakkeiden vaihto voidaan tehdä ennen rivioperaatioden suorittamista tai niiden aikana. Molemmat tavat johtavat samaan lopputulokseen. Yleensä ajatellan, että Gaussin eliminaatio suoritetaan matriisille

$$P^T AQ$$

jonka kaikki tukialkiot ovat nollasta poikkeavia.

Rivien ja sarakkeiden vaihtoa Gaussin eliminaatiossa kutsutaan tuennaksi. Uusi tukialkio voidaan valita monella eri tavalla. Tukialkion valinta vaikuttaa Gaussin eliminaation tuottaman ratkaisun tarkkuuteen, kun algoritmi totutetaan ääreellisellä tarkkuudella esimerkiksi Matlabissa. Kun vaihdetaan ainoastaan matriisin A rivejä tai sarakkeita, kyseessa on rivi- tai saraketuenta. Rivituennassa askeleella k etsitään $j = k, \ldots, n$ siten, että että $|a_{jk}^{(k)}|$ saa suurimman arvonsa. Tämän jälkeen vaihdetaan rivien k ja j paikkaa.

Kun käytetään ääreellistä tarkkuutta, paras ratkaisu saadaan suorittamalla sekä rivien että sarakkeidan vaihtoja. Tällöin puhutaan täystuennasta ja tukialkio valitaan siten, että askeleella k etsitään $i \geq k, j \geq k$ siten että $|a_{ij}^{(k)}|$ saa suurimman arvonsa. Täystuennassa vaihdetaan sekä rivien i ja k että sarakkeiden j ja k paikat.

Käsin laskettaessa voidaan tukialkio valita helpoimalla mahdollisella tavalla.

Esimerkki 5.4. *Tarkastellaan Esimerkin 5.2 yhtälöryhmää*

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 4 & 3 \\ 1 & 3 & 2 & 7 \end{array}\right].$$

Suoritetaan ensimmäinen askel Gaussin eliminaatioa :

$$\begin{array}{c} \\ -Y1 \\ -Y1 \end{array} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 4 & 3 \\ 1 & 3 & 2 & 7 \end{array}\right] \rightarrow \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 2 & 1 & 7 \end{array}\right]$$

Täystuennassa etsitään ylläolevasta matriisista suurin alkio $|a_{ij}|, i, j = 2, 3$. *Suurin alkio on* $a_{23} = 3$ *joten vaihdetaan sarakkeiden 2 ja 3 paikkaa. Saadaan*

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 3 & 0 & 3 \\ 0 & 1 & 2 & 7 \end{array}\right].$$

Suoritetaan nyt viimeinen askel Gaussin eliminaatioa,

$$\begin{array}{c} \\ \\ -\frac{1}{3}Y2 \end{array} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 3 & 0 & 3 \\ 0 & 1 & 2 & 7 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 2 & 6 \end{array}\right]. \tag{15}$$

Suoritettu sarakkeiden permutaatio voidaan kirjoittaa matriisin

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

avulla. Tämän esimerkin tapauksessa Gaussin eliminaatio suoritettiin matriisille AP. *Eli yhtälöryhmä (15) voidaan kirjoittaa muodossa*

$$UP\boldsymbol{x} = \tilde{\boldsymbol{b}}$$

Takaisinsijoituksella ratkaistaan vektori $\boldsymbol{y}$ *siten, että*

$$U\boldsymbol{y} = \tilde{\boldsymbol{b}}.$$

Yhtälöryhmän ratkaisu voidaan lukea yhtälöstä $P\boldsymbol{x} = \boldsymbol{y}$.

5.3 Ratkaisun olemassaolo

Käytännön kannalta on tärkeää tietää, koska yhtälöryhmällä (14) on olemassa ratkaisu ja koska tämä ratkaisu on yksikäsitteinen. Näihin kysymyksiin voidaan vastata käyttämällä lineaariyhdistelyn käsitettä sekä vektorien lineaarista riippumattomuutta. Olkoot $A \in \mathbb{R}^{n \times n}$. Matriisi-vektoritulo $A\boldsymbol{x}$ voidaan kirjoittaa muodossa

$$A\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{a}_i.$$

Kyseessä on matriisin A sarakevektorien $\boldsymbol{a}_i$ lineaariyhdistely. Sarakevektorit muodostavat aliavaruuden, jota kutsutaan matriisin A rangiksi,

$$R(A) = \operatorname{span}\{\boldsymbol{a}_i, i = 1, \dots, n\}.$$

Jotta yhtälöryhmällä (14) voi olla ratkaisu tietyllä $\boldsymbol{b} \in \mathbb{R}^n$, tulee päteä $\boldsymbol{b} \in R(A)$. Tällöin $\boldsymbol{b}$ voidaan muodostaa matriisin A sarakevektorien lineaarikombinaationa. Yleensä ollaan kiinnostuneita tapauksesta, jossa ratkaisu on olemassa jokaisella $\boldsymbol{b} \in \mathbb{R}^n$. Tällöin tulee päteä

$$R(A) = \mathbb{R}^n.$$

Tarvitsemme seuraavassa Lausetta, joka voidaan todistaa käyttämällä apuna Gaussin eliminaatioa.

Lause 5.1. *Olkoot $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in \mathbb{R}^n$ lineaarisesti riippumattomia. Tällöin*

$$\mathbb{R}^n = \mathit{span}\{\boldsymbol{x}_i, i = 1, \dots, n\},$$

Ratkaisun yksikäsitteisyyttä voidaan tutkia homogeeniyhtälön $A\boldsymbol{x} = 0$ avulla. Idea on seuraava : Olkoot $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^n$ tehtävän (14) ratkaisuja. Tällöin,

$$A(\boldsymbol{y_1} - \boldsymbol{y_2}) = 0. \tag{16}$$

Jos matriisiin A liittyvällä homogeeniyhtälöllä on olemassa vain nollaratkaisu, pätee

$$\boldsymbol{y}_1 = \boldsymbol{y}_2.$$

Kyseessä on perustekniikka ratkaisun olemassaolon tutkintaa. Määrittelemme toisen avaruuden A vektoreihin liittyvän aliavaruuden.

$$N(A) := \{\ \boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = 0\ \},$$

jota kutsutaan matriisin A nolla-avaruudeksi. Nolla-avaruuden avulla voidaan esittää kaikki yhtälöryhmän (14) ratkaisut. Jos pätee $N(A) = \{0\}$ sanotaa, että nolla-avaruus on triviaali. Tällöin yhtälölle (14) löytyy yksikäsitteinen ratkaisu.

Lause 5.2. *Olkoot $A \in \mathbb{R}^{n\times n}$ ja $\boldsymbol{b} \in \mathbb{R}^n$. Tällöin yhtälöryhmällä*

$$A\boldsymbol{x} = \boldsymbol{b}$$

on olemassa yksikäsitteinen ratkaisu jokaisella $\boldsymbol{b} \in \mathbb{R}^n$ jos ja vain jos

(i). $R(A) = \mathbb{R}^n$

(ii). $N(A) = \{0\}$

Todistus. Oletetaan, että (i) on totta. Tällöin on olemassa $\boldsymbol{x}_i \in \mathbb{R}^n$ siten, että

$$A\boldsymbol{x}_i = \boldsymbol{e}_i, i = 1, \ldots, n.$$

Olkoot $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ siten että

$$0 = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i.$$

Kerrotaan ylläoleva matriisilla A. Saadaan,

$$0 = \sum_{i=1}^{n} \alpha_i A\boldsymbol{x}_i = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i.$$

Koska luonnollinen kanta on lineaarisesti riippumaton, ainoa ratkaisu on $\alpha_i = 0, i = 1, \ldots, n$. Vektorit $\boldsymbol{x}_i$ ovat siis lineaarisesti riippumattomia ja muodostavat Lauseen 5.1 perusteella $\mathbb{R}^n$ kanna. Tutkitaan matriisin A nolla-avaruutta.

$$A\boldsymbol{x} = 0.$$

Kirjoitetaan $\boldsymbol{x}$ vektorien $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ muodostamassa kannassa,

$$\boldsymbol{x} = \sum_{i=1}^{n} \beta_i \boldsymbol{x}_i.$$

Eli,

$$A\boldsymbol{x} = \sum_{i=1}^{n} \beta_i A\boldsymbol{x}_i = \sum_{i=1}^{n} \beta_i \boldsymbol{e}_i.$$

Koska luonnolliset kantavektorit ovat lineaarisesti riippumattomia $N(A) = \{0\}$ ja tehtävällä on olemassa yksikäsitteinen ratkaisu.

Oletetaan, että (ii) pätee. Tällöin

$$A\boldsymbol{x} = 0 \Rightarrow \boldsymbol{x} = 0,$$

Eli matriisin A sarakkeet ovat lineaarisesti riippumattomia. Lauseen 5.1 perusteella $R(A) = \mathbb{R}^n$ ja voidaan käyttää kohdan (i) tulosta.

Todistetaan seuraavaksi väite toiseen suuntaa. Oletetaan, että yhtälöryhmällä on olemassa yksikäsitteinen ratkaisu jokaisella $\boldsymbol{b} \in \mathbb{R}^n$. Tällöin

$$A\boldsymbol{y} = 0 \Rightarrow \boldsymbol{y} = 0.$$

eli, $N(A) = \{0\}$. Koska A:n sarakkeet ovat lineaarisesti riippumattomia, Lauseen 5.1 perusteella $R(A) = \mathbb{R}^n$. $\square$

Esimerkki 5.5. *Tarkastellaan lineaarikuvausta $f : \mathbb{R}^n \to \mathbb{R}^n$ siten, että*

$$f(\boldsymbol{x})_i = \begin{cases} x_i & i \leq k \\ x_i - \alpha_i x_k & i > k \end{cases},$$

jossa $1 \leq k \leq n$ ja $\alpha_i \in \mathbb{R}^n$. Lineaarikuvaus f liitty Gaussin eliminaatiossa suoritettaviin rivioperaatioihin. Olkoot $A \in \mathbb{R}^{n\times n}$ kuvauksen f matriisimuoto, eli $f(\boldsymbol{x}) = A\boldsymbol{x}$. Matriisin A nolla-avaruudelle pätee,

$$\boldsymbol{x} \in N(A) \Leftrightarrow f(\boldsymbol{x}) = 0.$$

Saadaan,

$$\begin{cases} x_i = 0 & i \leq k \\ x_i - \alpha_i x_k = 0 & i > k \end{cases}.$$

Tiedetään, että $x_i = 0, i = 1, \ldots, k$. Kun $i > k$ pätee

$$x_i - \alpha_i x_k = 0$$

Koska $x_k = 0 \Rightarrow x_i = 0, k < 0 \leq n$. Näinollen, $N(A) = \{0\}$ ja yhtälöryhmällä $A\boldsymbol{x} = \boldsymbol{b}$ on olemassa ratkaisu jokaisella $\boldsymbol{b} \in \mathbb{R}^n$.

Jos matriisin A nolla-avaruus $N(A)$ on ei-triviaali ja $\boldsymbol{b} \in R(A)$ löytyy yhtälöryhmälle (14) ääretön määrä ratkaisuja. Nämä ratkaisut voidaan määrittää, kun tunnetaan matriisin A nolla-avaruus.

Lause 5.3. *Olkoot $A \in \mathbb{R}^{n\times n}$ ja $\boldsymbol{b} \in R(A)$. Tällöin yhtälöryhmällä $A\boldsymbol{x} = \boldsymbol{b}$ on olemassa ratkaisu. Oletetaan lisäksi, että $\boldsymbol{y}$ on yksi yhtälöryhmän ratkaisu. Tällöin kaikki ratkaisut voidaan kirjoittaa muodossa*

$$\{\ \boldsymbol{y} + \boldsymbol{z} | \boldsymbol{z} \in N(A)\ \}. \tag{17}$$

Todistus. Tehdään vasta-oletus. Oletetaan, että $\boldsymbol{b} \in R(A)$, mutta tehtävällä ei ole ratkaisua. Tällöin ei ole olemassa kertoimia $y_1, \ldots, y_n$ siten että

$$\sum_{i=1}^{n} y_i \boldsymbol{a}_i = \boldsymbol{b}.$$

Määritelmän mukaan tällöin pätee $\boldsymbol{b} \notin R(A)$, joka on ristiriita. Eli yhtälöllä on olemassa ratkaisu $\boldsymbol{y}$.

Oletetaan, että yhtälöllä $A\boldsymbol{x} = \boldsymbol{b}$ on ratkaisu $\boldsymbol{w}$, eli

$$A\boldsymbol{w} = \boldsymbol{b}.$$

Oletusten perusteella tiedetään, että $\boldsymbol{y}$ on ratkaisu. Näinollen saadaa

$$A(\boldsymbol{w} - \boldsymbol{y}) = 0$$

Määritelmän mukaan $\boldsymbol{w} - \boldsymbol{y} \in N(A)$, eli $\exists \boldsymbol{z} \in N(A)$ siten että

$$\boldsymbol{w} = \boldsymbol{z} + \boldsymbol{y}.$$

□

5.4 Käänteismatriisi ja käänteiskuvaus

Matriisiin $A \in \mathbb{R}^{n\times n}$ liittyvä käänteismatriisi $B \in \mathbb{R}^{n\times n}$ on määritelty siten että

$$AB = I. \tag{18}$$

Jokaiselle matriisille A ei löydy käänteismatriisia. Jos käänteismatriisi B on olemassa, matriisia A kutsutaan säännölliseksi. Matriisin A säännöllisyys liittyy yhtälöryhmän ratkaisun olemassaoloon. Kirjoitetaan B sarakevektorien avulla muodossa

$$B = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \dots & \boldsymbol{b}_n \end{bmatrix}.$$

Tällöin yhtälön (18) mukaan pätee

$$\begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \dots & A\boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_n \end{bmatrix}.$$

Eli,

$$A\boldsymbol{b}_i = \boldsymbol{e}_i, \quad , i = 1, \dots, n.$$

Lauseen 5.2 perusteella säännölliselle matriisille A tulee päteä $R(A) = \mathbb{R}^n$. Tällöin yllä-oleva yhtälöryhmä määrittää kunkin vektorin $\boldsymbol{b}_i$ yksikäsitteisesti, joten käänteismatriisi B on yksikäsitteinen. Käytetään seuraavassa matriisin A käänteismatriisista merkintää A^{-1}. Käänteismatriisille pätee,

Lemma 5.1. *Olkoot $A, B \in \mathbb{R}^{n\times n}$ säännöllisiä matriiseja. Tällöin pätee*

(i). $A^{-1}A = I$

(ii). $(A^T)^{-1} = (A^{-1})^T$

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

Todistus. (i). Olkoot $C \in \mathbb{R}^{n\times n}$ siten että, $CA = I$. Tällöin määritelmä (18) antaa : $A^{-1} = IA^{-1} = CAA^{-1} = C$.

(ii). Käänteismatriisin määritelmän mukaan $AA^{-1} = I$, joten $(A^{-1})^T A^T = I$. Tällöin kohdan (i) perusteella, $(A^T)^{-1} = (A^{-1})^T$.

(iii). Kotitehtävä. □

Käänteismatriisi on pääasiassa teoreettinen työkalu. Käänteismatriiseja ei käytännössä koskaan muodosteta kun $n > 2$.

5.5 Rivioperaatioiden matriisimuoto

Tässä luvussa todistamme, että Gaussin eliminaation tuottaman yhtälöryhmän ratkaisun on sama kuin alkuperäisen yhtälöryhmän

$$A\boldsymbol{x} = \boldsymbol{b}$$

ratkaisu. Gaussin eliminaatiossa suoritetut rivioperaatiot voidaan esittää vektoreihin $A\boldsymbol{x}$ ja $\boldsymbol{b}$ operoivina lineaarikuvauksina. Tätä näkökulmaa käyttämällä nämä operaatiot voidaan esittää helposti matriisimuodossa. Askeleella k rivi k kerrotaan sopivalla termillä ja lasketaan yhteen rivien $k+1, \ldots, n$ kanssa. Kyseessä on vektoriin $A\boldsymbol{x}$ operoiva lineeaarikuvaus

$$f_k(\boldsymbol{x})_i = \begin{cases} \boldsymbol{x}_i & i \leq k \\ \boldsymbol{x}_i - \frac{a_{ki}^{(k)}}{a_{kk}^{(k)}} \boldsymbol{x}_k & i > k \end{cases}.$$

Kun tukialkiot $a_{kk}^{(k)}$ ovat nollasta poikkeavia, kuvaus f_k on kääntyvä. Käänteiskuvaus voidaan kirjoittaa muodossa

$$f_k^{-1}(\boldsymbol{x})_i = \begin{cases} \boldsymbol{x}_i & i \leq k \\ \boldsymbol{x}_i + \frac{a_{ki}^{(k)}}{a_{kk}^{(k)}} \boldsymbol{x}_k & i > k \end{cases}.$$

Gaussin algoritmin ensimmäinen askel voidaan siis kirjoittaa muodossa

$$f_1(A\boldsymbol{x}) = f_1(\boldsymbol{b}).$$

Olkoot $E_1 \in \mathbb{R}^{n \times n}$ siten että $f_1(\boldsymbol{x}) = E_1\boldsymbol{x}$. Ensimmäinen askel voidaan kirjoittaa myös muodossa

$$E_1 A\boldsymbol{x} = E_1(\boldsymbol{b})$$

Yhtälöryhmän saattaminen alakolmiomuotoon vastaa tällöin operaatioa

$$f(A\boldsymbol{x}) = f(\boldsymbol{b}). \tag{19}$$

jossa f on yhdistetty kuvaus $f = f_{n-1} \circ \ldots f_1$. Käyttämällä lineaarikuvauksen matriisimuotoa saadaan

$$E_{n-1} \ldots E_2 E_1 A\boldsymbol{x} = E_{n-1} \ldots E_2 E_1 \boldsymbol{b}.$$

Kuvauksen f lineaarisuuden perusteella yhtälö (19) voidaan kirjoittaa muodossa

$$f(A\boldsymbol{x} - \boldsymbol{b}) = 0$$

Matriisimuodossa tämä vastaa yhtälöä

$$E_{n-1} \ldots E_2 E_1 (A\boldsymbol{x} - \boldsymbol{b}) = 0$$

Koska kukin matriisi $E_i, i = 1, \ldots, n-1$ on säännöllinen, myös matriisi $E_{n-1} \ldots E_2 E_1$ on säännöllinen. Tällöin, $N(E_{n-1} \ldots E_2 E_1) = \{0\}$, joten Gaussin algoritmi ei siis muuta alkuperäisen tehtävän ratkaisua. Eliminaatiomatriiseja $E_1, \ldots, E_{n-1}$ tai niiden tuloa ei yleensä muodosteta. Ne toimivat Gaussin eliminaatiossa muodostetun muunnoksen kirjoitusasuna ja mahdollistavat menetelmän formaalin tutkimisen.

5.6 LU-hajotelma

Tarkastellaan seuraavassa säännöllisiä matriiseja $A \in \mathbb{R}^{n\times n}$. Olkoot $P, Q \in \mathbb{R}^{n\times n}$ permutaatiomatriiseja, siten että Gaussin eliminaation muokkaa matriisin $P^T AQ$ yläkolmiomuotoon ilman tuentaa. Edellisen luvun merkinnöillä,

$$U = E_{n-1} \cdots E_2 E_1 P^T AQ. \tag{20}$$

Lineaarikuvausten f_k ja f_k^{-1} määritelmistä seuraa, että

$$(E_k)_{ij} = 0, i > j \quad \text{ja} \quad (E_k^{-1})_{ij} = 0, i > j$$

Tälläisiä matriiseja kutsutaan alakolmiomatriiseiksi. Olemme aikasemmin todistaneet, että kahden yläkolmiomatriisin tulo on yläkolmiomatriisi. Vastaava tulos pätee alakolmiomatriiseille. Näinollen $E_{n-1} \cdots E_2 E_1$ ja sen käänteismatriisi $E_1^{-1} E_2^{-1} \cdots E_{n-1}^{-1}$ ovat molemmat alakolmiomatriiseja. Yhtälö (20) voidaan kirjoittaa muodossa

$$(E_{n-1} \cdots E_2 E_1)^{-1} U = P^T AQ.$$

Merkitään seuraavassa $L = (E_k \cdots E_2 E_1)^{-1}$. Ylläolevan perusteella pätee

$$LU = P^T AU.$$

Kyseessä on matriisin A LU-hajotelma. Matriisihajotelmalla tarkoitetaan tapaa, jolla matriisi voidaan kirjoittaa kahden tai useamman "erityis-matriisin tulona. Tieteellisen laskennan algoritmit perustuvat lähes kaikki erilaisiin matriisihajotelmiin. Tällä kurssilla tutustumme LU - hajotelman lisäksi matriisin ominaisarvohajotelmaan, jossa tietyt matriisit A voidaan kirjoittaa muodossa $A = XDX^{-1}$, jossa D on diagonaalimatriisi.

Matriisin LU-hajotelmaa voidaa käyttää, esimerkiksi lineaaristen yhtälöryhmien ratkaisussa. Olkoot $A \in \mathbb{R}^{n\times n}$ siten, että $A = LU$. Tehtävä

$$A\boldsymbol{x} = \boldsymbol{b}$$

voidaan ratkaista kahdessa osassa

$$L\boldsymbol{y} = \boldsymbol{b} \quad \text{ja} \quad U\boldsymbol{x} = \boldsymbol{y}.$$

Tälläistä menettelyä käytetään, kun sama yhtälöryhmä ratkaistaan usealla eri vektorilla $\boldsymbol{b}$.

Matriisi $L \in \mathbb{R}^{n\times n}$ on helppo muodostaa Gaussin eliminaation aikana lasketuista kertoimista.

Lemma 5.2. *Olkoot $\alpha_{ki} \in \mathbb{R}, \alpha_{ki} \neq 0, k = 1, \ldots, n, i = 1, \ldots, n$ ja*

$$f_k(\boldsymbol{x})_i = \begin{cases} \boldsymbol{x}_i & i \leq k \\ \boldsymbol{x}_i + \alpha_{ki}\boldsymbol{x}_k & i > k \end{cases}, \quad k = 1, \ldots, n$$

Olkoot tämän lisäksi $E_k \in \mathbb{R}^{n\times n}$ siten, että $f_k(\boldsymbol{x}) = E_k\boldsymbol{x}$. Tällöin

$$E_1^{-1}E_2^{-1}\cdots E_n^{-1} = \begin{bmatrix} 1 & & & & \\ -\alpha_{12} & 1 & & & \\ -\alpha_{13} & -\alpha_{23} & 1 & & \\ \vdots & \vdots & \dots & \ddots & \\ -\alpha_{1n} & -\alpha_{2n} & \dots & -\alpha_{(n-1)n} & 1 \end{bmatrix}$$

Todistus. Todistetaan lause käyttämällä induktioa parametrin n-suhteen. Lineaarikuvauksen f_k käänteiskuvaus on

$$f_k^{-1}(\boldsymbol{x})_i = \begin{cases} \boldsymbol{x}_i & i \leq k \\ \boldsymbol{x}_i - \alpha_{ki}\boldsymbol{x}_k & i > k \end{cases}, \quad k = 1, \dots, n$$

Perustapaus : $n = 2$. Väite seuraa suoraan lineaarikuvauksen f_1 määritelmästä, eli pätee

$$f_1^{-1} = \begin{bmatrix} 1 & 0 \\ -\alpha_{12} & 0 \end{bmatrix}$$

Induktio-oletus : Väite on totta, kun $n = m$ Eli, matriisitulo

$$\tilde{E}_1^{-1}\tilde{E}_2^{-1}\cdot\tilde{E}_{m-1}^{-1}.$$

on haluttua muotoa.

Induktio-askel : todistataan, väite, kun $n = (m+1)$ Lineaarikuvausten määritelmästä seuraa, että

$$E_k^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{E}_{k-1}^{-1} \end{bmatrix}, \quad k = 2, \dots, (m+1)$$

jossa $\tilde{E}_{k-1} \in \mathbb{R}^{m\times m}$ on eliminaatiomatriisi. Näinollen pätee

$$E_2^{-1}\cdots E_{m+1}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{E}_1^{-1}\cdots\tilde{E}_m^{-1} \end{bmatrix}$$

Lineaarikuvauksen f_1^{-1} matriisimuoto on

$$E_1^{-1} = \begin{bmatrix} 1 & 0 \\ \boldsymbol{a} & I \end{bmatrix}$$

jossa $\boldsymbol{a} \in \mathbb{R}^{m\times 1}$, $\boldsymbol{a} = \begin{bmatrix} \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \end{bmatrix}^T$ ja $I \in \mathbb{R}^{m\times m}$. Suorittamalla blokkimatriisien tulo saadaan

$$E_1^{-1}\cdots E_{m+1}^{-1} = \begin{bmatrix} 1 & 0 \\ \boldsymbol{a} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{E}_1^{-1}\cdots\tilde{E}_m^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \boldsymbol{a} & \tilde{E}_1^{-1}\cdots\tilde{E}_m^{-1} \end{bmatrix}$$

Joka on toivottua muotoa.

□

5.7 Singulaarisen yhtälöryhmän ratkaisu

Täystuettua Gaussin eliminaatioa voidaan käyttää singulaaristen yhtälöryhmien ratkaisuun. Aikaisemmassa luvussa todistimme, etteivät rivioperaatiot muuta matriisiyhtälön ratkaisuja. Matriisin nolla-avaruudelle pätee

$$A\boldsymbol{x} = 0.$$

Täystuetussa Gaussin eliminaatiossa matriisin A rivien ja sarakkeiden järjestysta muutetaan. Voidaan ajatella, että eliminaatio suoritetaan matriisille

$$P^T AQ,$$

jossa P ja Q ovat permutaatiomatriiseja. Nolla-avaruudelle pätee

$$A\boldsymbol{x} = 0 \Leftrightarrow P^T AQ(Q^T \boldsymbol{x}) = 0.$$

Eli,

$$N(A) = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = Q\boldsymbol{y}, \boldsymbol{y} \in N(P^T AQ) \right\}.$$

Koska rivioperaatiot eivät muuta yhtälöryhmän ratkaisua, saadaan

$$E_p \dots E_2 E_1 P^T AQ\boldsymbol{x} = 0 \Rightarrow P^T AQ\boldsymbol{x} = 0.$$

jossa $p < n$. Täystuetussa Gaussin eliminaatiossa singulaarinen matriisiin $P^T AQ$ muuntuu muotoon,

$$E_p \dots E_2 E_1 P^T AQ = \begin{bmatrix} U_{11} & U_{12} \\ & 0 \end{bmatrix}$$

jossa $U_{11} \in \mathbb{R}^{p \times p}$ on yläkolmiomatriisi. Koska U_{11} diagonaalialkiot ovat tukialkiota ja näinolle poikkeavat nollasta, on U_{11} kääntyvä matriisi. Kukin matriisin $P^T AQ$ nolla-avaruuden vektori voidaan kirjoittaa muodossa

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}.$$

Jossa, $\boldsymbol{x}_2 \in \mathbb{R}^p$ ja

$$\boldsymbol{x}_1 = -U_{11}^{-1} U_{12} \boldsymbol{x}_2.$$

Gaussin eliminaation avulla voidaan siis määrittää matriisin A nolla-avaruus. Saadaan

$$N(A) := \left\{ Q \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} \in \mathbb{R}^n \;\middle|\; \boldsymbol{x}_2 \in \mathbb{R}^k \text{ and } \boldsymbol{x}_1 = -U_{11}^{-1} U_{12} \boldsymbol{x}_2 \right\}.$$

Tarkastellaan singulaarisen yhtälöryhmän ratkaisua. Yksi yhtälöryhmän ratkaisu saadaan muodossa

$$\boldsymbol{y} = Q \begin{bmatrix} U_{11}^{-1} P^T \boldsymbol{b}_1 \\ 0 \end{bmatrix}.$$

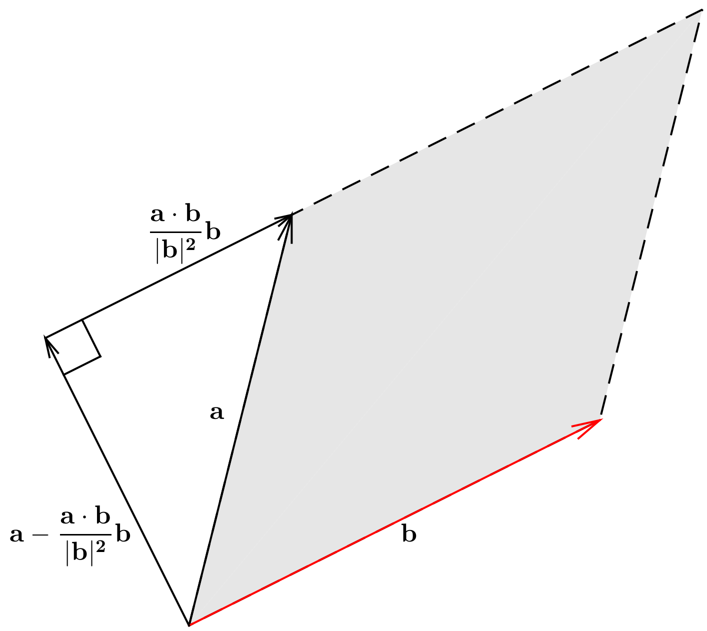
Lauseen 5.3 nojalla kaikki yhtälöryhmän ratkaisut saadaan muodossa

$$\boldsymbol{y} + \boldsymbol{z}, \quad \text{jossa } \boldsymbol{z} \in N(A).$$

Yllä esitettyä menettelyä voidaan käyttää myös kun ratkaistaan yhtälöryhmiä joiden kerroinmatriisi $A \in \mathbb{R}^{m \times n}, m < n$ tai $A \in \mathbb{R}^{n \times n}$, A ei ole säännöllinen.

6 Matriisin determinantti

Matriisin determinantti on kuvaus, joka liittää kuhunkin $\mathbb{R}^{n\times n}$ matriisiin reaaliluvun. Determinantin taustalla on geometrinen ajatus matriisin sarakevektorien muodostaman suunnikkaan tilavuudesta. Kun determinantilta vaaditaan tilavuusmitalle luonnolliset ominaisuudet osoittautuu, että kyseessä on yksikäsitteinen funktion.



Kuva 9: Vektorien $\boldsymbol{a}$ ja $\boldsymbol{b}$ määrittämän suunnikas sekä sen korkeuden laskenta.

Aloitetaan tarkastelemalla 2×2-matriiseja. Sarakevektorien muodostama suunnikas P on esitetty kuvassa 9. Suunnikkaan pinta-ala $|P|$ on sen kanta kertaa korkeus. Suunnikkaan korkeus voidaan laskea käyttämällä apuna pistetuloa.

$$h^2 = \left|\boldsymbol{a} - \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{b}|^2}\boldsymbol{b}\right|^2 = |\boldsymbol{a}|^2 - \frac{(\boldsymbol{a}\cdot\boldsymbol{b})^2}{|\boldsymbol{b}|^2}$$

Näin pinta-alan neliöksi saadaan

$$|P|^2 = |\boldsymbol{b}^2|h^2 = |\boldsymbol{a}|^2|\boldsymbol{b}|^2 - (\boldsymbol{a}\cdot\boldsymbol{b})^2$$

Joka voidaan sieventää muotoon

$$|P|^2 = (a_1b_2 - a_2b_1)^2.$$

Määritellään 2×2-matriisin determinantti ylläolevan lausekkeen mukaisesti,

$$\det \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{b}_1 \\ \boldsymbol{a}_2 & \boldsymbol{b}_2 \end{bmatrix} = a_1 b_2 - a_2 b_1.$$

Determinantti liittyy suunnikkaan pinta-alaan siten, että

$$|P| = |\det A|.$$

Determinantin merkki riippuu vektorien $\boldsymbol{a}$ ja $\boldsymbol{b}$ välisestä kulmasta. Käytetään seuraavassa yksinkertaisuuden vuoksi merkintätapa

$$V(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n) := \det \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots \boldsymbol{a}_n \end{bmatrix}$$

Määrittelemämme 2×2-determinantti toteuttaa seuraavat laskusäännöt :

(i) $V(\boldsymbol{a}_1, \boldsymbol{a}_2) = -V(\boldsymbol{a}_2, \boldsymbol{a}_1)$

(ii) Determinantti saa arvon nolla, jos matriisin A sarakkeet ovat lineaarisesti riippuvia.

(iii) 2×2 - matriisin determinantti on bilineaarinen kuvaus.

(iv) $\det I = 1$

Laajennetaan yllä olevia sääntöjä (i) siten, että :

$$V(\boldsymbol{a}_1, , \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_j, \ldots, \boldsymbol{a}_n) = -V(\boldsymbol{a}_1, , \ldots, \boldsymbol{a}_j, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_n)$$

Tämän lisäksi $n \times n$-matriisin determinantti on multilineaarinen kuvaus.

Määritelmä 6.1. *Olkoot V vektoriavaruus ja $f : V \times V \ldots \times V \to \mathbb{R}$ Tällöin f on multilineaarinen, jos pitämällä kiinitettynä kaikki muut argumentint kuin i kyseessä on lineaarinen kuvaus.*

Voidaan todistaa, että ylläolevat laskusäännöt määrittelevät yksikäsitteisen kuvauksen. Tätä kuvausta kutsutaan matriisin A determinantiksi. Pätee

$$V(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = V(\sum_{i_1=1}^{n} a_{i_1 1} \boldsymbol{e}_{i_1}, \ldots, \sum_{i_n=1}^{n} a_{i_n n} \boldsymbol{e}_{i_1}) = \sum_{i_1=1}^{n} \ldots \sum_{i_n=1}^{n} a_{i_1 1} \ldots a_{i_n n} V(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_n})$$

Käyttämällä determinantin laskusääntöjä, suurin osa ylläolevista termeistä saa arvon nolla. Jäljelle jäävät ainoastaan kantavektorien permutaatiot. Determinantin laskusäännön (ii) mukaan, pätee

$$V(e_{i_1}, \ldots, e_{i_n}) = \begin{cases} -1 & \text{kun paikanvaihtoja on pariton määrä} \\ 1 & \text{kun paikanvaihtoja on parillinen määrä} \end{cases}$$

ylläolevaa kuvausta merkitään yleensä symbolilla $\sigma_{i_1 i_2 \ldots i_n}$. Näinolle determinantille pätee

$$V(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = \sum a_{i_1 1} \ldots a_{i_n n} \sigma_{i_1 i_2 \ldots i_n}$$

Jossa summa lasketaan kaikkien indeksien $1, \ldots, n$ permutaatioiden yli. Koska n - alkion permutaatiota on $n!$ kappaletta, on tämä summan termien lukumäärä. Koska $n!$ kasvaa huomattavan nopeasti n - funktiona ja ylläoleva kaava on vaikea muistaa, ei sitä käytetä yleensä determinanti laskennassa.

6.1 Determinantin laskusääännöjä

Determinantille voidaan todistaa useita hyödyllisiä identiteettejä. Aloitetaan tutkimalla matriisien A ja B tulon determinanttia,

$$\det AB.$$

Tulon determinanttille löytyy geometrinen tulkinta. Aloitetaan tarkastelemalla kaksiulotteista tilannetta. Olkoot $\boldsymbol{e}_1, \boldsymbol{e}_2$ avaruuden $\mathbb{R}^2$:n luonnollinen kanta. Tarkastellaan yksikköneliöä $R = (0,1) \times (0,1)$, joka voidaan kirjoittaa vektorien $\boldsymbol{e}_1$ ja $\boldsymbol{e}_2$ avulla muodossa

$$R = \{\, t\boldsymbol{e}_1 + s\boldsymbol{e}_2 \mid 0 < t, s < 1 \,\}\,.$$

Lineaarisuuden perusteella matriisi AB kuvaa joukon R joukolle

$$ABR = \{\, tAB\boldsymbol{e}_1 + sAB\boldsymbol{e}_2 \mid 0 \le t, s \le 1\}\,.$$

Aikaisemman perusteella determinantti liittyy tämän joukon pinta-alaan siten, että $|ABR| = |\det AB|$. Kuvaus AB voidaan suorittaa kahdessa vaiheessa, kuvataan ensin joukko R joukolle

$$BR = \{\, tB\boldsymbol{e}_1 + sB\boldsymbol{e}_2 \mid 0 \le t, s \le 1 \,\}$$

jonka jälkeen BR kuvataan joukolle ABR. Joukon BR pinta-ala on aikasemman perusteella $|BR| = |\det B|$. Approksimoidaan joukkoa BR neliöillä R_i Kuvan 10 mukaisella tavalla. Kukin neliö R_i voidaan lausua muodossa

$$R_i = \{\, t\boldsymbol{e}_1 + s\boldsymbol{e}_2 + \boldsymbol{b}_i \mid 0 \le t, s \le \epsilon \; \boldsymbol{b}_i \in \mathbb{R} \,\} = \epsilon R + \boldsymbol{b}_i.$$

Determinantin bilineaarisuuden perusteella kunkin joukon AR_i pinta-alalle pätee

$$|AR_i| = |\det \begin{bmatrix} \epsilon\boldsymbol{a}_1 & \epsilon\boldsymbol{a}_2 \end{bmatrix}| = \epsilon^2 |AR|.$$

Ja joukon ABR pinta-alalle

$$|ABR| = \lim_{\epsilon \to 0} \sum_{i}^{N_\epsilon} |AR_i| = |\det A| \lim_{\epsilon \to 0} \sum_{i}^{N_\epsilon} \epsilon^2 = |\det A||\det B|.$$

Joten $|\det AB| = |\det A||\det B|$. Determinantin ominaisuuksista lähtien voimme todistaa seuraavan tuloksen:

Lemma 6.1. *Olkoot $A, B \in \mathbb{R}^{n \times n}$. Tällöin pätee*

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

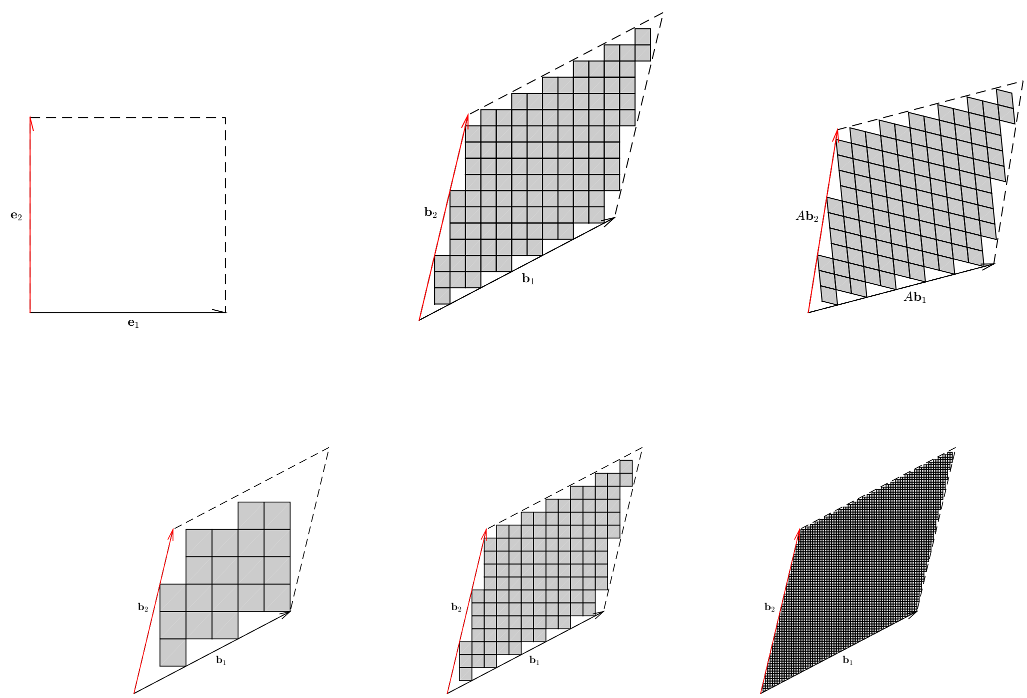
Determinantin tulosäännön avulla voimme todistaa seuraavat tärkeät ominaisuudet,

Lemma 6.2. *Olkoot $A \in \mathbb{R}^{n \times n}$. Tällöin pätee*

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

(ii) $\det A^{-1} = \frac{1}{\det A}$

Todistus. Tarkka todistus jätetään harjoitustehtäväksi. Vihje : molemmissa kohdissa voidaan käyttää determinantin tulosääntöä. Kohdassa (i) matriisin A LU-hajotelma saattaa olla hyödyllinen. □



Kuva 10: Suunnikkaan BR pinta-alan laskenta aproksimoimalla suunnikasta pienillä neliöillä R_i.

6.2 Determinantin laskenta käsin

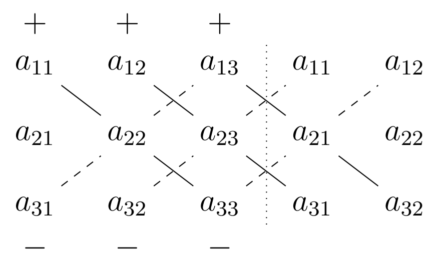
Kun determinatteja lasketaan käsin käytetään joko muistisääntöjä ($n \leq 3$) tai ali - determinattikehitelmää (yleensä kun $n > 3$). Kun $n = 2$ determinantti lasketaan kaavan

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

avulla. Tapauksessa $n = 3$ käytetään Sarruksen sääntöä. Sarruksen säännössä matriisin rivit kiroitetaan uudelleen matriiisn perään ja piirreetään diagonaaliset viivat oheisen kuvan mukaisesti. Viivoilla kohtaavat matriisialkiot kerrotaan keskenään ja lasketaan yhteen toistensa kanssa. Alhaalta ylös piiretyillä viivoilla käytetään negatiivista merkkiä.
Näin laskemalla saadaan

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

Determinantti voidaan aina laskea käyttämällä alideterminanttikehitelmää. Olkoot $A \in \mathbb{R}^{n \times n}$. Seuraavassa merkinnällä $[A]_{i,j}$ tarkoitetaan $\mathbb{R}^{(n-1)\times(n-1)}$ matriisia, joka saadaan kun matriisista A poistetaan rivi i ja sarake j.



Kuva 11: Sarruksen sääntö

.

Esimerkki 6.1. *Olkoot*

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

Tällöin

$$[A]_{1,1} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, [A]_{1,2} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, [A]_{1,3} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

Aloitetaan todistamalla seuraava tulos

Lemma 6.3. *Olkoot* $A \in \mathbb{R}^{n\times n}$ *siten, että*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & A_{22} \end{bmatrix},$$

jossa $A_{22} \in \mathbb{R}^{(n-1)\times(n-1)}$. *Tällöin* $\det A = \det A_{22}$.

Todistus. Seuraa permutaatiokaavasta. □

Käyttämällä Lemma 6.3 voidaan todistaa seuraava tulos

Lemma 6.4. *Olkoot* $A \in \mathbb{R}^{n\times n}$. *Tällöin pätee*

$$\det A = \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \det [A]_{i,1}$$

Todistus. Käytetään determinantin n-linearisuutta. Pätee

$$\det A = V(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n) = V(\sum_{i}^{n} a_{i1}\boldsymbol{e}_i, \dots, \boldsymbol{a}_n)$$

Joten,

$$\det A = \sum_{i}^{n} a_{i1} V(\boldsymbol{e}_i, \dots, \boldsymbol{a}_n)$$

Suoritetaan nyt matriisin A rivien permutaatio siten että permutoidun matriisin rivit ovat järjestyksessä $i, 1, \ldots, i-1, i+1, \ldots, n$. Permutaatioita tarvitaan yhteensä $(i-1)$ kappaletta. Tällöin determinantille pätee

$$V(\boldsymbol{e}_i, \ldots, \boldsymbol{a}_n) = (-1)^{i-1} \det \begin{bmatrix} 1 & \boldsymbol{b}^T \\ 0 & [A]_{i,1} \end{bmatrix}.$$

jossa $\boldsymbol{b}_j = a_{i(j+1)}$. Käyttämällä determinantin n-lineaarisuutta saadaa saadaan

$$V(\boldsymbol{e}_i, \ldots, \boldsymbol{a}_n) = (-1)^{i-1} \det \begin{bmatrix} 1 & 0 \\ 0 & [A]_{i,1} \end{bmatrix}$$

Lemman 6.3 avulla saadaan

$$V(\boldsymbol{e}_i, \ldots, \boldsymbol{a}_n) = (-1)^{i-1} \det [A]_{i,1}.$$

Joka todistaa väitteen. □

Käyttämällä ylläolevaa teoreemaa rekursiivisesti voidaan laskea $n \times n$-matriisien detrminantteja.

Esimerkki 6.2. *Olkoot*

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 5 \\ 0 & 6 & 7 & 0 \\ 8 & 0 & 0 & 10 \end{bmatrix}.$$

Tällöin

$$\det A = 1 \cdot \det \begin{bmatrix} 4 & 0 & 5 \\ 6 & 7 & 0 \\ 0 & 0 & 10 \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 3 & 0 & 5 \\ 0 & 7 & 0 \\ 8 & 0 & 10 \end{bmatrix}.$$

Käyttämällä ali-determinanttikehitelmää saadaan

$$\det \begin{bmatrix} 4 & 0 & 5 \\ 6 & 7 & 0 \\ 0 & 0 & 10 \end{bmatrix} = 4 \det \begin{bmatrix} 7 & 0 \\ 0 & 10 \end{bmatrix} + 5 \det \begin{bmatrix} 6 & 7 \\ 0 & 0 \end{bmatrix} = 280$$

Matriisin $[A]_{1,2}$ *determinantti voidaan laskea käyttämällä samaa menettelyä. Näin ollen saadaan* $\det [A]_{1,2} = -70$ *ja* $\det A = 420$.

Koska $\det A = \det A^T$ ali-determinanttikehitelmä voidaan kehittää rivin tai sarakkeen mukaan. Permutoimalla matriisia A, kehitelmä voidaan laskea minkä vain matriisi rivin tai sarakkeen mukaan. Saadaan,

$$\det A = (-1)^{j-1} \sum_{i=1}^{n} (-1)^{i-1} a_{ij} \det [A]_{i,j}$$

Tai vastaavasti

$$\det A = (-1)^{j-1} \sum_{i=1}^{n} (-1)^{i-1} a_{ji} \det [A]_{j,i}$$

Determinanttia laskettaessa käytettävä kaava valitaan yleensä sen mukaan, mikä matriisin rivi sisältää eniten nollia. Näin käsin laskenta yksinkertaistuu huomattavasti.

Kun käytetään tietokonetta apuna, voidaan determinantti laskea muodostamalla ensi matriisin LU - hajotelma ja käyttämällä tätä hajotelmaan determinantin laskennassa.

Käyttämällä apuna ali-determinanttikehitelmää, voidaan todistaa seuraava lause

Lemma 6.5. *Olkoot U ala- tai yläkolmiomatriisi. Tällöin,*

$$\det U = \prod_{i=1}^{n} u_{ii}.$$

Olkoot $A = LU$. Käyttämällä determinantin tulosääntöä, saadaan

$$\det A = \prod_{i=1}^{n} u_{ii}.$$

6.3 Determinantti ja yhtälöryhmän ratkaisun

Determinanttia voidaan käyttää sekä yhtälöryhmän ratkaisemiseen että yhtälöryhmän ratkaisun olemassaolon tutkimiseen. Yhtälöryhmän ratkaisu olemassaololla ja determinantilla on seuraava yhteys.

Lemma 6.6. *Yhtälöryhmällä (14) on olemassa yksikäsitteinen ratkaisu jos ja vain jos*

$$\det A \neq 0$$

Todistus. Jos $\det A \neq 0$, ovat matriisin A sarakkeet lineaarisesti riippumattomia. Tällöin $R(A) = \mathbb{R}^n$, joten yhtälöryhmällä on Lauseen 5.2 nojalla olemassa yksikäsitteinen ratkaisu. Jos yhtälöryhmälle ei ole yksikäsitteistä ratkaisua, sarakkeet eivät ole lineaarisesti riipppumattomia ja $\det A = 0$. □

Yhtälöryhmiä voidaan myös ratkaista determinantin avulla, käyttämällä ns. Cramerin sääntöä. Cramerin sääntöä ei käytetä kun tehtäviä ratkaistaan tietokoneen avulla, mutta se on hyödyllinen työkalu kun matriisin alkiot ovat symbolisia lausekkeita.

Lemma 6.7. *Olkoot $A \in \mathbb{R}^{n\times n}, \boldsymbol{b} \in \mathbb{R}^n$. Tällöin yhtälöryhmän $A\boldsymbol{x} = \boldsymbol{b}$ ratkaisulle pätee*

$$x_i = \frac{V(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}, \boldsymbol{b}, \boldsymbol{a}_{i+1}, \ldots \boldsymbol{a}_n)}{\det A}.$$

Todistus. Suoraan laskemalla saadaan

$$x_i V(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = V(\boldsymbol{a}_1, \ldots, x_i\boldsymbol{a}_i, \ldots, \boldsymbol{a}_n)$$

Determinantin n-lineaarisuuden ja ominaisuuksien perusteella pätee, että

$$x_i V(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = V(\boldsymbol{a}_1, \ldots, \sum_{j=1}^{n} x_j\boldsymbol{a}_j, \ldots, \boldsymbol{a}_n)$$

Koska $A\boldsymbol{x} = \boldsymbol{b}$ saadaan

$$x_i V(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = V(\boldsymbol{a}_1, \ldots, A\boldsymbol{x}, \ldots, \boldsymbol{a}_n) = V(\boldsymbol{a}_1, \ldots, \boldsymbol{b}, \ldots, \boldsymbol{a}_n).$$

□

6.4 Ristitulo $\mathbb{R}^3$:ssa

Determinanttiin liitty $\mathbb{R}^3$:ssa määritelty kuvaus, ristitulo. Ristitulo on kuvaus $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ joka liittää kahteen vektoriin $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$ niiden määrittämän suunnikkaan pinta-alan sekä normaalin. Ristituloon törmätään tulevilla matematiikan ja fysiikan kursseilla erityisesti, kun lasketaan integraaleja. Historiallisesti ristitulo on määritelty ensimmäistä kertaa kun on tutkittu tetraedreja. Tarkastellaan seuraavassa suuntaissärmiöä,

$$S = \{t\boldsymbol{a} + s\boldsymbol{b} \mid t, s \in [0, 1]\}.$$

Tämän suuntaissärmiön pinta-ala voidaan laskea samalla tavalla kuin aikaisemmin,

$$|S|^2 := |\boldsymbol{a}|^2|\boldsymbol{b}|^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2.$$

Lasketaan seuraavassa vektorien $\boldsymbol{a}, \boldsymbol{b}$ virittämän tason yksikkönormaali $\boldsymbol{n}$. Yksikkönormaalille pätee

$$\boldsymbol{n} \cdot \boldsymbol{a} = 0 \quad \boldsymbol{n} \cdot \boldsymbol{b} = 0 \quad \boldsymbol{n} \cdot \boldsymbol{n} = 1. \tag{21}$$

Kun $\boldsymbol{a}, \boldsymbol{b} \neq 0$ ja $\boldsymbol{a}, \boldsymbol{b}$ eivät ole yhdensuuntaisia, ylläolevalla yhtälöryhmällä on kaksi ratkaisua. Yhtälöt $\boldsymbol{a}^T\boldsymbol{n} = 0$ ja $\boldsymbol{b}^T\boldsymbol{n} = 0$ voidaan ymmärtää ehtona

$$\boldsymbol{n} \in N\left(\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\right).$$

Kun vektorit $\boldsymbol{a}, \boldsymbol{b}$ täyttävät niille asetetut ehdot, kyseessä yksiulotteinen $\mathbb{R}^3$ aliavaruus span$\{\boldsymbol{c}\}$. Näinollen tehtäväksi muodostuu etsiä skalaari α siten että

$$\boldsymbol{n} = \alpha\boldsymbol{c}$$

toteuttaa ehdon $|\boldsymbol{n}| = 1$. Normin määritelmästä seuraa, että ehdon täyttäviä ratkaisuita on kaksi, α ja $-\alpha$.

Kootaan yhtälöt (21) yhtälöryhmäksi

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ n_1 & n_2 & n_3 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Merkitään ylläolevan yhtälörymän kerroinmatriisia A:lla. Vaikka kyseessä ei ole lineaarinen yhtälöryhmä, voidaan se silti ratkaista Cramerin säännöllä. Saadaan

$$n_1 = \frac{\det \begin{bmatrix} 0 & a_2 & a_3 \\ 0 & b_2 & b_3 \\ 1 & n_2 & n_3 \end{bmatrix}}{\det A} = \frac{a_2 b_3 - b_2 a_3}{\det A}$$

sekä

$$n_2 = \frac{a_1 b_3 - b_1 a_3}{\det A} \quad \text{ja} \quad n_3 = \frac{a_1 b_2 - b_2 a_1}{\det A}.$$

Determinantin ominaisuuksien perusteella tiedetään, että $\det A$ on vektorien $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{n}$ virittämän kuution tilavuus. Koska $\boldsymbol{n}$ on yksikkövektori, pätee

$$\det A = \pm |S|,$$

Eli, determinantti on verrannollinen suunnikkaan S pinta-alaan. Määritellään ristitulo siten että

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{bmatrix} a_2 b_3 - b_2 a_3 \\ a_1 b_3 - b_1 a_3 \\ a_1 b_2 - b_2 a_1 \end{bmatrix}.$$

Ristitulo antaa siis suunnikkan S normaalivektorin jonka pituus on sen pinta-alaa. Merkki on valittu oikean käden säännön mukaan.

Ristitulon kaava voidaan muistaa, kun "venytetään"hieman determinantin määritelmää ja kirjoitetaan :

$$\det \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = (a_2 b_3 - b_2 a_3)\boldsymbol{e}_1 + (a_1 b_3 - b_1 a_3)\boldsymbol{e}_2 + (a_1 b_2 - b_2 a_1)\boldsymbol{e}_3.$$

7 Matriisin ominasarvot

Matriisin $A \in \mathbb{R}^{n \times n}$ ominaisarvoilla ja ominaisvektoreilla tarkoitetaan yhtälön : Etsi $(\lambda, \boldsymbol{x}) \in \mathbb{C} \times \mathbb{C}^n, \boldsymbol{x} \neq 0$ siten että

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \tag{22}$$

ratkaisuja. Ominaisarvoja ja vektoreita laskettaessa käytämme kompleksilukuja $\mathbb{C}$. Vektorit, joiden alkiot ovat kompleksilukuja muodostavat avaruuden $\mathbb{C}^n$, jossa laskutoimitukset on määritelty samalla tavalla, kuin avaruudessa $\mathbb{R}^n$. Lähes kaikki aikaisemmin käsittelemämme teoria yleistyy avaruuden $\mathbb{C}^n$ vektoreille. Erotuksena on muutos Eulkidiseen sisätuloon ja normiin,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \boldsymbol{y} \qquad \|\boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^* \boldsymbol{x}}$$

jossa $*$-on konjugaattitranspoosi,

$$\boldsymbol{x}^* = \overline{\boldsymbol{x}}^T.$$

Ominaisarvoja käytetään insinööritieteissä ja matematiikassa monessa eri käyttötarkoituksessa. Esimerkiksi rakenteiden värähtelyt voidaan ratkaista ominaisarvotehtävien avulla. Tällä kurssilla käytämme ominaisarvoja matriisiarvoisten funktioiden, matriisin potenssin ja eksponentin laskentaa. Näitä funktioita voidaan soveltaa rekursio- tai osittaisdifferentiaaliyhtälöiden ratkaisemisessa.

Yhtälö (22) voidaan kirjoittaa muodossa : Etsi $(\lambda, \boldsymbol{x}) \in \mathbb{C} \times \mathbb{C}^n, \boldsymbol{x} \neq 0$ siten että

$$(A - \lambda I)\, \boldsymbol{x} = 0.$$

Kun $\lambda \in \mathbb{C}$ on matriisin A ominaisarvo matriisilla $(A - \lambda I)$ on ei-triviaali nolla-avaruus. Käyttämällä determinantin ja nolla-avaruuden välistä yhteyttä, voidaan ominaisarvo ratkaista yhtälöstä

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

Tämä yhtälö ei riipu ominaisvektorista $\boldsymbol{x}$. Kun ominaisarvot tunnetaan, niihin liittyvät ominaisvektorit ratkaistaan määrittämällä matriisin $(A - \lambda I)$ nolla-avaruus. Yhtälö (22) määrittää ominaisarvon $\lambda \in \mathbb{C}$ yksikäsitteisesti, mutta kutakin ominaisarvoa vastaa äärettömän monta ominaisvektoria, $\boldsymbol{x} \in N(A - \lambda I)$.

Esimerkki 7.1. *Olkoot*

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

Matriisin A ominaisarvot ovat yhtälön

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

ratkaisuja. Eli,

$$\det \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 - 4 = 0$$

Näinollen matriisin A ominaisarvot ovat $\lambda_1 = -1$ ja $\lambda_2 = 3$.

Lasketaan seuraavaksi näihin ominaisarvoihin liittyvät ominaisvektorit, eli määritetään avaruudet $N(A - \lambda_1 I)$ ja $N(A - \lambda_2 I)$. Saadaan

$$A - \lambda_1 I = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \quad ja \quad A - \lambda_2 I = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$$

Tämän esimerkin tapauksessa nolla-avaruus voidaan lukea suoraan matriiseista. Saadaan:

$$N(A - \lambda_1 I) = span\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \quad ja \quad N(A - \lambda_2 I) = span\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} \tag{23}$$

Kuten esimerkissä 7.1, ominaisarvotehtävän (22) ratkaiseminen johtaa polynomi nollakohtien laskemiseen. Olkoon matriisiin $A \in \mathbb{R}^{n\times n}$ karaketristinen polynomi

$$p_A(\lambda) = \det (A - \lambda I).$$

Alideterminanttikehitelmän avulla nähdään, että p_A on enintään n-asteinen polynomi.

Lemma 7.1. *Olkoot $A \in \mathbb{R}^{n\times n}$. Tällöin $p_A(\lambda) = \det (A - \lambda I)$ on n-asteen polynomi.*

Todistus. Todistetaan väite induktiolla matriisin dimension suhteen.

Perus-askel n = 2: Suoraan laskemalla nähdään, että $\det (A - \lambda I) = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21}$.

Induktio-oletus : väite on totta, kun n = k.

Induktio-askel Olkoot $A \in \mathbb{R}^{(k+1)\times(k+1)}$. Kehitetään determinantti matriisin ensimmäisen sarakkeen suhteen käyttämällä ali-determinanttikehitelmää

$$p_A(\lambda) = (a_{11} - \lambda) \det [A - \lambda_I]_{i,1} + \sum_{i=2}^{k} (-1)^{i-1} a_{i1} \det [A - \lambda I]_{i,1}. \tag{24}$$

Kullekkin alideterminantille pätee

$$\det [A - \lambda I]_{i,1} = \det ([A]_{i,1} - \lambda I),$$

Jossa I on $\mathbb{R}^{k\times k}$ identiteettimatriisi. Induktio-oletuksen nojalla $\det ([A]_{i,1} - \lambda I)$ on k-asteinen polynomi, joten yhtälön (24) nojalla p_A on $k+1$ asteen polynomi. □

Koska matriisin ominaisarvot ovat siihen liittyvän karakteristisen polynomin juuria, voidaan ominaisarvoja tutkia tutkimalla karakteristisen polynomin ominaisuuksia.

Lemma 7.2. *Olkoot $A \in \mathbb{R}^{n\times n}$. Tällöin matriisilla A on enintään n erillistä ominaisarvoa.*

Todistus. Matriisin A ominaisarvot ovat sen karakteristisen polynomi p_A juuria. Koska polynomi p_A on astetta n sillä on algebran peruslauseen nojalla korkeintaan n juurta. □

Matriisin ominaisarvo voi olla karteristisen polynomin monikertainen juuri. Olkoot matriisin $A \in \mathbb{R}^{n\times n}$ ominaisarvot $\lambda_1, \ldots, \lambda_k$. Tällöin karakteristinen polynomi voidaan kirjoittaa muodossa

$$p_A(\lambda) = (-1)^n \prod_{i=1}^{k} (\lambda_i - \lambda)^{i_k}.$$

Lukua i_k kutsutaan ominaisarvon algebralliseksi kertaluvuksi. Kuhunkin ominaisarvoon liittyy ominaisvektorien avaruus

$$E_\lambda = N(A - \lambda I).$$

Tämän avaruuden dimensioa (lineaarisesti riippumattomien kantavektorien lukumäärää) kutsutaan ominaisarvon λ geometriseksi kertaluvuksi.

Esimerkki 7.2. *Tarkastellaan matriisia*

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

Lasketaan A:n ominaisarvot.

$$\det A - \lambda I = \det \begin{bmatrix} 3-\lambda & -1 \\ 1 & 1-\lambda \end{bmatrix} = (3-\lambda)(1-\lambda) + 1 = 0$$

Saadaan,

$$(3-\lambda)(1-\lambda) + 1 = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2.$$

Joten $\lambda = 2$. Koska p_A voidaan kirjoittaa muodossa $p_A(\lambda) = (\lambda - 2)^2$, kyseessä on kaksinkertainen ominaisarvo. Lasketaan seuraavaksi ominaisarvoon 2 liittyy ominaisavaruus $N(A - 2I)$.

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

Ominaisavaruus $N(A - \lambda) = span \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$. Ominaisarvon 2 geometrinen kertaluku on siis yksi.

Matriisin eri ominaisarvoihin liittyvät ominaisvektorit ovat lineaarisesti riippumattomia.

Lause 7.1. *Olkoot $A \in \mathbb{R}^{n \times n}$ ja $\lambda_1, \dots, \lambda_k$ matriisin A erilliset ominaisarvot. Tällöin eri ominaisarvoihin liittyvät ominaisvektorit ovat lineaarisesti riippumattomia.*

Todistus. Todistetaan väite induktiolla ominaisarvojen lukumäärän suhteen.

Perusaskel: Yleisyyttä menettämättä voidaan tarkastella ominasarvoja λ_1 ja λ_2. Olkoot $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C}^n$ siten, että

$$A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1 \quad \text{ja} \quad A\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2.$$

Tehdään vasta-oletus: Ominaisvektorit $\boldsymbol{x}_1$ ja $\boldsymbol{x}_2$ ovat lineaarisesti riippuvia, eli on olemassa $\alpha \in \mathbb{C}$, $\alpha \neq 0$ siten että

$$\boldsymbol{x}_1 - \alpha \boldsymbol{x}_2 = 0 \qquad \text{tai} \qquad \boldsymbol{x}_1 = \alpha \boldsymbol{x}_2.$$

$$A\boldsymbol{x}_1 = \alpha A \boldsymbol{x}_2 \qquad \text{siten, että} \qquad \boldsymbol{x}_1 = \frac{\lambda_2}{\lambda_1} \alpha \boldsymbol{x}_2.$$

Josta seuraa, että

$$(1 - \frac{\lambda_2}{\lambda_1}) \alpha \boldsymbol{x}_2.$$

Eli, $\lambda_2 = \lambda_1$, joka on ristiriita.

Induktio-oletus: Oletetaan, että ominaisarvoihin $\lambda_1, \ldots, \lambda_{m-1}$ liittyvät ominaisvektorit ovat lineaarisesti riippumattomia.

Induktio-askel: Olkoot $\boldsymbol{x}_i, i = 1, \ldots, ,$ ominaisvektoreihin $\lambda_1, \ldots, \lambda_m$ liittyviä ominaisvektoreita. Tehdään vastaoletus : $\boldsymbol{x}_i$ ovat lineaarisesti riippuvia. Eli, on olemassa $\boldsymbol{\alpha} \in \mathbb{C}^k$ siten että

$$\sum_i^k \alpha_i \boldsymbol{x}_i = 0. \tag{25}$$

jos $\alpha_i = 0$ jollain i väite seuraa suoraan induktio-oletuksesta. Joten, voimme olettaa, että $\alpha_i \neq 0, i = 1, \ldots, k$. Kerrotaan yhtälö (25) matriisilla A. Tämä antaa

$$\sum_i \boldsymbol{\alpha}_i \lambda_i \boldsymbol{x}_i = 0.$$

Ilman yleisyyden menetystä, voimme olettaa, että $\lambda_1 \neq 0$. Joten,

$$\alpha_1 \boldsymbol{x}_1 = \sum_{i=2}^k \alpha_i \frac{\lambda_i}{\lambda_1} \boldsymbol{x}_i.$$

Yhtälön (25) mukaan $\alpha_1 \boldsymbol{x} = \sum_{i=2}^k \alpha_i \lambda_i \boldsymbol{x}_i$, joten

$$0 = \sum_{i=2}^k \alpha_i (1 - \frac{\lambda_i}{\lambda_1}) \boldsymbol{x}_i.$$

Koska induktio-oletuksen mukaan $\boldsymbol{x}_i, i = 2, \ldots, k$ ovat lineaarisesti riippumattomia ja $\alpha_i \neq 0$ tulee päteä $(1 - \frac{\lambda_i}{\lambda_1}) = 0$, joka on ristiriita. □

Eri ominaisarvoihin liittyvien vektorien lineaarinen riippumattomuus johtaa matriisin A ominaisarvohajotelmaan. Tarkastellaan ensin tilannetta, jossa kunkin ominaisarvon geometrinen ja algbrallinen kertaluku ovat yksi. Kootaan matriisin $A \in \mathbb{R}^{n \times n}$ ominaisvektorit matriisin $X \in \mathbb{C}^{n \times n}$ sarakkeiksi

$$X = \begin{bmatrix} \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_n \end{bmatrix}.$$

Lasketaan matriisitulo AX

$$AX = \begin{bmatrix} A\boldsymbol{x}_1 & \ldots & A\boldsymbol{x}_n \end{bmatrix}$$

saadaan

$$AX = \begin{bmatrix} \lambda_1 \boldsymbol{x}_1 & \ldots & \lambda_n \boldsymbol{x}_n \end{bmatrix}$$

Tämä voidaan kirjoittaa muodossa

$$AX = X\Lambda$$

jossa $\Lambda \in \mathbb{C}^{n\times n}$ on siten, että

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Lauseen 7.1 nojalla matriisin X sarakkeet ovat lineaarisesti riippumattomia, joten matriisi X on säännöllinen. Saadaan

$$A = X\Lambda X^{-1}.$$

Kyseessä on matriisin A ominaisarvohajotelma. Ominaisarvohajotelman avulla voidaan määritellä matriisiarvoisia funktioita, kuten matriisin potenssi tai eksponentti. Esimerkiksi matriisin potenssille

$$A^k = \begin{cases} I & \text{kun } k = 0 \\ AA^{k-1} & \text{kun } k > 0 \end{cases}.$$

pätee

Lemma 7.3. *Olkoot $A \in \mathbb{R}^{n\times n}$ siten että*

$$A = X\Lambda X^{-1},$$

jossa $X \in \mathbb{C}^{n\times n}$ on säännöllinen matriisi ja $\Lambda \in \mathbb{C}^{n\times n}$ on diagonaalimatriisi. Tällöin jokaisella $k \in \mathbb{N}$ pätee

$$A^k = X\Lambda^k X^{-1}.$$

Todistus. Todistetaan väite induktiolla.

Induktio askel, k= 1 : Väite seuraa suoraan oletuksista.

Induktio-oletus : Väite on totta, kun $k = m$

Induktio-askel : Todistetaan väite, kun $k = m + 1$

$$A^{m+1} = AA^m.$$

Induktio-oletuksen ja oletusten perusteella

$$A^{m+1} = X\Lambda X^{-1}X\Lambda^m X^{-1} = X\Lambda^{m+1}X^{-1}.$$

□

Matriisin potenssin avulla voidaan määritellä funktioiden sarjaesityksen kautta matriisiarvioisia funktioita. Esimerkiksi matriisin eksponenttifunktio määritelläämä eksponenttifunktion sarjaesityksen

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

avulla. Olkoot $A \in \mathbb{R}^{n\times n}$. Matriisin A eksponenttifunktio on tällöin

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Sarjan summalle pätee

$$\sum_{n=0}^{\infty} \frac{A^n}{n!} = \lim_{N\to\infty} \sum_{n=0}^{N} \frac{A^n}{n!}.$$

Tällä kurssilla emme ole keskustelleet, siitä miten matriisien suuruutta mitataan. Formaalisti ylläoleva raja-arvo tulee tulkita $\epsilon - \delta$-määritelmän kautta: $\forall \delta > 0$ on olemassa $N_\delta \in \mathbb{N}$ siten, että

$$\left\| \sum_{n=0}^{\infty} \frac{A^n}{n!} - \sum_{n=0}^{N} \frac{A^n}{n!} \right\| \leq \delta \quad \forall \quad N \geq n_\delta,$$

jossa $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$ on jokin matriisille määritelty normi. Matriisinormi on kuvaus, joka liitttä kuhunkin matriisiin reaaliluvun sekä toteuttaa Luvussa 3 vektorinormille valitut ominaisuudet. Jotta matriisin eksponenttifunktio voidaan määritellä formaalisti, tulisi matriisinormista keskustella tarkemmin - jätetään tämä tuleville matematiikan kursseilla.

Diagonalisoituvan matriisin eksponenetti voidaan laskea sarjaesityksen avulla. Olkoot $A \in \mathbb{R}^{n\times n}$ siten, että

$$A = X\Lambda X^{-1}$$

jossa $X \in \mathbb{R}^{n\times n}$ on säännöllinen ja $\Lambda \in \mathbb{R}^{n\times n}$ diagonaalimatriisi. Matriisin A potenssille pätee

$$A^n = X\Lambda^n X^{-1}.$$

Käyttämällä tätä esitystä saadaan

$$\sum_{n=0}^{N} A^n = X \begin{bmatrix} \sum_{k=0}^{N} \frac{\lambda_1^k}{k!} & & & \\ & \sum_{k=0}^{N} \frac{\lambda_2^k}{k!} & & \\ & & \ddots & \\ & & & \sum_{k=0}^{N} \frac{\lambda_n^k}{k!} \end{bmatrix} X^{-1}.$$

Matriisinormin ominaisuuksista seuraa (ei todisteta), että raja arvo voidaan viedä diagonaalimatriisin alkioiden sisään.

$$e^A = \lim_{N\to\infty} \sum_{n=0}^{N} A^n = X \begin{bmatrix} \lim_{N\to\infty} \sum_{k=0}^{N} \frac{\lambda_1^k}{k!} & & & \\ & \lim_{N\to\infty} \sum_{k=0}^{N} \frac{\lambda_2^k}{k!} & & \\ & & \ddots & \\ & & & \lim_{N\to\infty} \sum_{k=0}^{N} \frac{\lambda_n^k}{k!} \end{bmatrix} X^{-1} =$$

Laskemalla raja-arvot saadaan

$$e^A = X \begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix} X^{-1}. \tag{26}$$

Diagonaisoituvan matriisin A eksponenttifunktio on määritelty ylläolevan esityksen avulla.

Esimerkki 7.3. *Lasketaan Esimerkin 7.1 matriisin*

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

eksponentti. Esimerkin 7.1 nojalla pätee,

$$e^A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-1} & 0 \\ 0 & e^3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}.$$

Matriisiarvoisia funktioita käytetään esimerkiksi differentiaaliyhtälöitä ratkaistaessa. Esimerkiksi matriisn eksponenttifunktio antaa tehtävän: Etsi $\boldsymbol{x}(t) : [0, \infty) \to \mathbb{R}^n$ siten että

$$\boldsymbol{x}'(t) = A\boldsymbol{x}(t) \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{27}$$

ratkaisun. Ratkaisuksi saadaan

$$\boldsymbol{x}(t) = e^{tA}\boldsymbol{x}_0.$$

Tarkastetaan, että $\boldsymbol{x}(t)$ on tehtävän (27) ratkaisu. Vektorin $\boldsymbol{x}(t)$ aikaderivaatta voidaan laskea suoraan määritelmän avulla:

$$\boldsymbol{x}'(t) = \lim_{h \to 0} \frac{\boldsymbol{x}(t+h) - \boldsymbol{x}(t)}{h} = \lim_{h \to 0} \frac{e^{(t+h)A}\boldsymbol{x}_0 - e^{tA}\boldsymbol{x}_0}{h}$$

Määritelmästä (26) seuraa, että $e^{(t+h)A} = e^{hA}e^{tA}$. Käyttämällä matriisieksponentin sarjaesitystä saadaan

$$\lim_{h \to 0} \frac{e^{(t+h)A}\boldsymbol{x}_0 - e^{tA}\boldsymbol{x}_0}{h} = \lim_{h \to 0} \frac{(hA + \mathcal{O}(h))e^{tA}\boldsymbol{x}_0}{h} = Ae^{tA}\boldsymbol{x}_0$$

Joten $\boldsymbol{x}(t)$ on tehtävän (27) ratkaisu. Laskuharjoituksissa tehtävä (27) ratkaistaan muuttujanvaihdon avulla.

7.1 Hermiittiset matriisit

Tässä luvussa käsittelemme hermiittisiä matriiseita. Jokainen hermiittinen matriisi $A \in \mathbb{C}^{n \times n}$ voidaan kirjoittaa muodossa

$$A = Q\Lambda Q^*,$$

jossa $Q \in \mathbb{C}^{n \times n}$ on unitaarinen matriisi ja $\Lambda \in \mathbb{C}^{n \times n}$ diagonaalimatriisi.

Määritelmä 7.1. *Olkoot $U \in \mathbb{C}^{n\times n}$ unitaarinen matriisi. Tällöin U on säännöllinen ja*

$$U^{-1} = U^*.$$

Esimerkki 7.4. *Olkoot*

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$$

Tällöin U on unitaarinen matriisi.

Koska hermiittiset matriisit ovat aina unitaaridiagonalisoituvia, ne muodostavat erittäin tärkeän matriisien ryhmän. Hermiittisillä matriiseilla on myös muita tärkeitä ominaisuuksia, esimerkiksi hermiittisen matriisin ominaisarvot ovat reaalisia ja eri ominaisarvoihin liittyvät ominaisvektorit ovat aina ortogonaalisia.

Lemma 7.4. *Olkoot $A \in \mathbb{C}^{n\times n}$ siten, että $A = A^*$. Tällöin kaikki matriisin A ominaisarvot ovat reaalisia.*

Todistus. Olkoot λ ja $\boldsymbol{v}$ siten, että

$$A\boldsymbol{v} = \lambda\boldsymbol{v}.$$

Pätee

$$\boldsymbol{v}^* A\boldsymbol{v} = \lambda|\boldsymbol{v}|^2.$$

Matriisin A symmetrisyyden nojalla pätee,

$$\boldsymbol{v}^* A\boldsymbol{v} = (A\boldsymbol{v})^*\boldsymbol{v} = (\lambda\boldsymbol{v})^*\boldsymbol{v} = \overline{\lambda}|\boldsymbol{v}|^2.$$

Näinollen

$$(\lambda - \overline{\lambda})|\boldsymbol{v}|^2 = 0$$

Eli

$$2\mathrm{Im}\lambda|\boldsymbol{v}|^2 = 0$$

□

.

Lause 7.2. *Olkoot $A \in \mathbb{C}^{n\times n}$ siten, että $A = A^*$ ja $\boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{C}^n$ siten, että*

$$A\boldsymbol{q}_1 = \lambda_1\boldsymbol{q}_1 \quad A\boldsymbol{q}_2 = \lambda_2\boldsymbol{q}_2$$

jossa $\lambda_1 \neq \lambda_2 \in \mathbb{R}$. Tällöin $\boldsymbol{q}_1^*\boldsymbol{q}_2 = 0$.

Todistus.

$$\lambda_1\boldsymbol{q}_1^*\boldsymbol{q}_2 = \lambda_1(\lambda_1\boldsymbol{q}_1)^*\boldsymbol{q}_2 = \boldsymbol{q}_1^* A\boldsymbol{q}_2$$

Toisaalta

$$\lambda_2\boldsymbol{q}_1^*\boldsymbol{q}_2 = \boldsymbol{q}_1^*(\lambda_2\boldsymbol{q}_2) = \boldsymbol{q}_1^* A\boldsymbol{q}_2.$$

Eli

$$(\lambda_2 - \lambda_1)\boldsymbol{q}_1^*\boldsymbol{q}_2 = 0.$$

Koska $\lambda_2 \neq \lambda_1$, tulee päteä $\boldsymbol{q}_1^*\boldsymbol{q}_2 = 0$.

□

Lause 7.3. *Olkoot $A \in \mathbb{C}^{n\times n}$ siten, että $A = A^*$. Tällöin on olemassa unitaarinen $Q \in \mathbb{C}^{n\times n}$ ja diagonaalimatriisi $\Lambda \in \mathbb{R}^{n\times n}$ siten, että*

$$A = Q^* \Lambda Q.$$

Todistus. **Perusaskel :** $n = 1$ Väite on selvästi totta.

Induktio-oletus: väite on totta, kun $n = k$.

Induktio-askel: Jokaisella matriisilla $A \in \mathbb{C}^{(k+1)\times(k+1)}$ on aina vähintään yksi ominaisarvo λ ja sitä vastaava ominaisvektori $\boldsymbol{q}$, $|\boldsymbol{q}| = 1$. Olkoot $Q \in \mathbb{C}^{(k+1)\times k}$ siten, että $Q^*\boldsymbol{q} = 0$ ja $Q^*Q = I$. Tarkastellaan matriisia

$$\begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix}^* A \begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix} = \begin{bmatrix}\boldsymbol{q}^*A\boldsymbol{q} & \boldsymbol{q}^*AQ \\ Q^*A\boldsymbol{q} & Q^*AQ\end{bmatrix} = \begin{bmatrix}\lambda & 0 \\ 0 & Q^*AQ\end{bmatrix}.$$

Koska matriisi Q^*AQ on hermiittinen, Induktio-oletuksen nojalla on olemassa unitaarinen $Q_1 \in \mathbb{C}^{k\times k}$ ja diagonaalimatriisi $\Lambda \in \mathbb{R}^{k\times k}$ siten, että

$$Q^*AQ = Q_1^*\Lambda_1 Q_1.$$

Eli, pätee

$$\begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix}^* A \begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix} = \begin{bmatrix}\lambda & 0 \\ 0 & Q_1^*\Lambda Q_1\end{bmatrix} = \begin{bmatrix}1 & 0 \\ 0 & Q_1^*\end{bmatrix}\begin{bmatrix}\lambda & 0 \\ 0 & \Lambda_1\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & Q_1\end{bmatrix}.$$

Näinollen,

$$A = \begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & Q_1^*\end{bmatrix}\begin{bmatrix}\lambda & 0 \\ 0 & \Lambda_1\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & Q_1\end{bmatrix}\begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix}^*.$$

Koska pätee

$$\begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & Q_1^*\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & Q_1\end{bmatrix}\begin{bmatrix}\boldsymbol{q} & Q\end{bmatrix}^* = I$$

Joka todistaa väitteen. □

7.2 Singulaariarvohajotelma

Jokainen matriisi $A \in C^{n\times m}$ voidaan hajottaa kahden unitaarimatriisin sekä diagonaalimatriisin tuloksi. Kyseessä on matriisin singulaariarvohajotelma,

$$A = U\Sigma V^*.$$

jossa $U \in \mathbb{C}^{n\times n}, V \in \mathbb{C}^{m\times m}$ ovat unitaarisia matriiseja ja $\Sigma \in \mathbb{R}^{n\times m}$ on diagonaalimatriisi. Kun $n < m$

$$\Sigma = \begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_m\end{bmatrix}$$

kun $n > m$,

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \\ & & & \end{bmatrix}$$

Singulaariarvohajotelma ei ole yksikäsitteinen, vaan matriisit U, V ja Σ voidaan valita usealla eri tavalla. Redusoidussa singulaariarvohajotelmassa matriisin U nolla-singulaariarvoja vastaavat sarakkeet poistetaan. Tämä johtaa hajotelmaan

$$A = U_r \Sigma_r V^*,$$

jossa $U_r \in \mathbb{C}^{n\times m}$, $V \in \mathbb{C}^{m\times m}$ ja diagonaalimatriisi $\Sigma_r \in \mathbb{R}^{m\times m}$ ovat siten, että

$$U_r^* U_r = I \quad V^* V = I \quad \Sigma_r = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_m \end{bmatrix}$$

Singulaariarvohajotelma voidaan kirjoittaa muodossa (laskuharjoitustehtävä)

$$A = \sum_{i=1}^{m} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^*.$$

Jokaisella matriisilla A on olemassa redusoitu singulaariarvohajotelma.

Lause 7.4. *Olkoot $A \in \mathbb{C}^{n\times m}$, $m < n$, Tällöin on olemassa $U_r \in \mathbb{C}^{n\times m}, V \in \mathbb{C}^{m\times m}$ ja diagonaalimatriisi $\Sigma_r \in \mathbb{R}^{m\times m}$ siten, että*

$$A = U\Sigma V^*.$$

jossa $U^*U = I$, $V^*V = I$ ja Σ on diagonaalimatriisi, jonka alkiot $\sigma_{ii} \geq 0$.

Todistus. Konstruoidaan matriisi V ja muodostetaan Σ sekä U tämän avulla. Jos matriisilla on olemassa singulaariarvohajotelma pätee,

$$A^*A = V\Sigma^T\Sigma V^* = V \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_m^2 \end{bmatrix} V^*.$$

Koska matriisi A^*A on hermiittinen ja näinollen unitaaridiagonalisoituva, etsitään matriisia V A^*A ominasvektoreista ja singulaariarvoja σ_{ii} vastaavan ominaisarvon neliöjuurina. Olkoot $V \in \mathbb{C}^{m\times m}$ ja diagonaalimatriisi $\Lambda \in \mathbb{R}^{m\times m}$ siten, että

$$A^*A = V\Lambda V^*.$$

Näytetään ensin, että kaikki matriisin A^*A ominaisarvot ovat positiivisia. Pätee,

$$V^*A^*AV = \Lambda$$

Eli,

$$\lambda_{ii} = \boldsymbol{e}_i^* V^*A^*AV\boldsymbol{e}_i = \|AV\boldsymbol{e}_i\| \geq 0$$

.

Oletataan yksinkertaisuuden vuoksi, että $N(A^*A)$ on triviaali. Tällöin pätee lisäksi $\lambda_{ii} \neq 0$ eli $\lambda_{ii} > 0$. Näinollen

$$\Lambda^{-1/2}V^*A^*AV\Lambda^{-1/2} = I.$$

Eli, matriisi $U = AV\Lambda^{-1/2}$ on unitaarinen. Kääntämällä matriisit V ja $\Lambda^{-1/2}$ saadaan

$$A = U\Lambda^{1/2}V^* = U\Sigma V^*.$$

□

Singulaarihajotelma voidaan muodostaa laskemalla matriisien A^*A ominaisarvot ja ominaisvektorit. Matriisi U saadaan muodossa $U = A\Sigma^{-1/2}V$. Käytännössä matriisin singulaariarvohajotelma voidaan lasketaan ainan numeerisesti.

Singulaariarvohajotelma eroaa ominaisarvohajotelmasta siten, että se on olemassa jokaiselle matriisille A. Singulaariarvohajotelmaa ei voi käyttää esimerkiksi matriisin potenssin laskentaa, vaan sillä on erilaisia käyttötarkoituksia. Singulaariarvohajotelmaa käytetään esimerkiksi matriisin pakkaamiseen, matriisin nolla-avaruuden sekä rangin määrittämiseen.