



Vaasan yliopisto
UNIVERSITY OF VAASA

MATTI LAAKSONEN

Lineaarialgebra, math.1040

VAASAN YLIOPISTON JULKAISUJA

OPETUSJULKAISUJA
MATEMATIIKKA

Julkaisija

Vaasan yliopisto

Julkaisuajankohta

Elokuu 2014

Tekijä(t) Matti Laaksonen	Julkaisun tyyppi Opetusmoniste	
	Julkaisusarjan nimi, osan numero Vaasan yliopiston julkaisuja	
Yhteystiedot University of Vaasa Department of Mathematics and Statistics P.O.Box 700 FI-65101 Vaasa, Finland	ISBN	
	ISSN	
	Sivumäärä	Kieli suomi
Julkaisun nimike Lineaarialgebra		
Tiivistelmä Opintojakson ”Lineaarialgebra”, math.1040 luentomateriaalia. Osaamistavoitteet: opintojakson jälkeen opiskelija osaa ratkaista minkä tahansa lineaarisen yhtälöryhmän ja osaa tulkita myös kaikki mahdolliset erikoistapaukset, opiskelija osaa ratkaista matriisin ominaisarvot ja ominaisvektorit, tutkia neliömuodon definiittisyyden, opiskelija osaa tutkia matriisin säännöllisyysasteen, opiskelija osaa käyttää Cramerin kaavoja, opiskelija osaa määrittää kolmiulotteisen vektoriavaruuden suoran ja tason yhtälöt, opiskelija osaa laskea ristitulon ja skalaaritulon ja tuntee niiden tavallisimmat käyttötavat fysiikan laskuissa, opiskelija tuntee tavallisimmat matriisihajotelmat, opiskelija osaa määrittää lineaarisen selitysmallin kertoimet PNS-menetelmällä, opiskelija osaa selittää lineaariavaruuden, lineaarisen aliavaruuden, lineaarikuvauksen, kannan ja dimension käsitteet. Sisältö: vektorit, lineaarinen vektoriavaruus, lineaarikuvaukset, lineaarinen yhtälöryhmä, matriisi, determinantti, ominaisarvo, Singulaariarvo-hajotelma, sisätulo, normi, approksimointi normin mielessä, vektoritulo, suora, taso, pseudoinverssi, PNS-menetelmä, tietokone-ohjelman käyttö vektori- ja matriisilaskuissa. (käytetty ohjelmointikieli ilmoitetaan kurssin alussa; Octave, Python, Java tai C.)		
Asiasanat		

SISÄLTÖ

1	Johdanto	6
1.1	Opintojakson ”Lineaarialgebra” kuvaus 2014	6
1.2	Ohjeita kurssin suorittajalle	7
1.3	Ensimmäiset Octave-ohjelmat	8
2	Trigonometriaa ja fysiikan vektoreita	12
2.1	Trigonometriset funktiot	12
2.1.1	Kulmayksiköistä	12
2.1.2	Suorakulmainen kolmio	14
2.1.3	Kolmiolauseet	16
2.1.4	Trigonometriset funktiot	18
2.1.5	Arcus-funktiot	20
2.2	Pisteet ja vektorit suorakulmaisissa koordinaatistoissa	23
2.3	Fysiikan voima-vektori	31
2.4	Pistetulo ja ristitulo	34
2.5	Skalaari- ja vektoriprojektiot	40
2.5.1	Vektorin projektio vektorin suuntaan	41
2.5.2	Voiman momentti pisteen suhteen tasossa	43
3	Lineaariset yhtälöryhmät	46
3.1	Rivoperaatiot	46
3.2	Lineaarikombinaatiot	54
3.3	Gaussin eliminointikeino	56
3.3.1	Algoritmi, Gaussin eliminointi, helppo tapaus.	57
3.3.2	Algoritmi, Gaussin eliminointi, yleinen tapaus.	61
3.4	Yhtälöryhmän ratkaiseminen Octavella	63
3.5	Homogeeninen yhtälöryhmä	66
4	Matriisit ja matriisilaskut	71
4.1	Matriisi ja matriisilaskut	71
4.2	Transponointi	80
4.3	Käänteismatriisi	82
4.4	Matriisit ja vektorit Octavessa (kesken)	84
4.5	Esimerkki Markovin ketjusta, ’FlyBike’	85
4.6	Matriisin kääntäminen riviopeaatioilla	90
4.7	Yhtälöryhmän ratkaisun herkkyyys	95

5	Determinantit	97
5.1	Determinantti	97
5.1.1	Kaksi- ja kolmirivinen determinantti	97
5.1.2	Isot determinantit	98
5.2	Determinantin laskeminen	103
5.2.1	Determinantin laskeminen rivioperaatioiden avulla	104
5.2.2	Determinantin laskeminen minorikehitelmällä	106
5.3	Matriisin kääntäminen adjungaattilla	108
5.4	Cramerin kaavat	109
5.5	Homogeeniset yhtälöryhmät	113
6	Pistetulon ja ristitulon sovelluksia	116
6.1	Ristitulo determinantin avulla	116
6.2	Pinta-aloja ja tilavuuksia	117
6.3	Skalaarikolmitulo ja vektorikolmitulo $\lll \square$	117
6.4	Vektorikolmikon suunnistus $\lll \square$	117
6.5	Suoran ja tason parametrimuotoinen yhtälö $\lll \square$	117
6.6	Suoran ja tason koordinaattimuotoinen yhtälö $\lll \square$	117
6.7	Etäisyyksiä avaruudessa	117
6.7.1	Pisteen etäisyys suorasta	117
6.7.2	Kahden suoran välinen etäisyys	119
7	Lineaariavaruudet	121
7.1	Rakenteiden hierarkiaa.	121
7.2	$\mathbb{R}$ - ja $\mathbb{C}$ -kertoiminen lineaariavaruus	125
7.3	Sisätulo ja normi	130
7.3.1	Sisätulo E^n :ssä	130
7.3.2	$\mathbb{R}^n$:n normeja	136
7.3.3	Yleinen tarkastelu	139
7.4	Lineaarinen kuvaus	141
7.5	(x, y) -Tason siirrot, venytykset, peilaukset ja kierrot	143
7.6	Lineaarinen aliavaruus ja span $\lll \square$	146
7.7	Ydin ja kuva	146
8	Kannat ja ortogonaalisuus	151
8.1	Lineaarinen riippumattomuus	151
8.2	Kanta ja dimensio	157
8.3	Matriisin säännöllisyysaste	165
8.4	Pseudoinverssi ja ortogonaaliprojektio	168
8.4.1	Ortogonaaliprojektio	170
8.5	Ortonormeerattu kanta, Gram-Schmidt	171
8.6	Ortogonaaliset matriisit, QR-hajoitelma	176
8.7	Lineaarikuvauksen matriisi	181
8.8	Kannan vaihto	185

9	Ominaisarvotehtävä	189
9.1	Kompleksinen lineaariavaruus	189
9.1.1	Kompleksiluvut	189
9.1.2	Kompleksiset vektorit ja matriisit	191
9.2	Matriisin ominaisarvot	192
9.2.1	Perusmääritelmät	192
9.2.2	Geometrinen tulkinta helpossa tapauksessa	202
9.2.3	Ominaisarvohajoitelma	204
9.3	Similaarisuus ja matriisin diagonalisointi	206
9.4	Schurin muoto ja normaalit matriisit	210
9.5	Matriisihajoitelmia	217
9.5.1	Jordanin muoto	220
9.6	Matriisin singulaariarvohajoitelma	221
9.7	Matriisinormit	225
9.8	Matriisin kuntoluku	227
9.9	Neliömuodot ja definiittisyys	229

TODO LIST

Lisää esimerkki	27
Lisää esimerkki.	31
Tarkista toleranssi	65
Kappale kesken	84

1 JOHDANTO

Ch: johdanto

1.1 Opintojakson ”Lineaarialgebra” kuvaus 2014

SEC:ojkuvaus

Koodi ja laajuus: MATH.1040, 5op.

Edellytykset: Lähtötasotestin hyväksytty suoritus.

Osaamistavoitteet: opintojakson jälkeen opiskelija osaa ratkaista minkä tahansa lineaarisen yhtälöryhmän ja osaa tulkita myös kaikki mahdolliset erikoistapaukset, opiskelija osaa ratkaista matriisin ominaisarvot ja ominaisvektorit, tutkia neliömuodon definiittisyyden, opiskelija osaa tutkia matriisin säännöllisyysasteen, opiskelija osaa käyttää Cramerin kaavoja, opiskelija osaa määrittää kolmiulotteisen vektoriavaruuden suoran ja tason yhtälöt, opiskelija osaa laskea ristitulon ja skalaaritulon ja tuntee niiden tavallisimmat käyttötavat fysiikan laskuissa, opiskelija tuntee tavallisimmat matriisihajotelmat, opiskelija osaa määrittää lineaarisen selitysmallin kertoimet PNS-menetelmällä, opiskelija osaa selittää lineaariavaruuden, lineaarisen aliavaruuden, lineaarikuvauksen, kannan ja dimension käsitteet.

Sisältö: vektorit, lineaarinen vektoriavaruus, lineaarikuvaus, lineaarinen yhtälöryhmä, matriisi, determinantti, ominaisarvo, Singulaariarvo-hajotelma, sisätulo, normi, approksimointi normin mielessä, vektoritulo, suora, taso, Kalman-suodin, pseudoinverssi, PNS-menetelmä, tietokone-ohjelman käyttö vektori- ja matriisilaskuissa. (käytetty ohjelmointikieli ilmoitetaan kurssin alussa; Octave, Python, Java tai C.)

Oppimateriaali ja Kirjallisuus:

1. luentomoniste:
2. Kreyszig, E: Advanced Engineering Mathematics, John Wiley & Sons, luvut 6, 7, 8.1–8.3
3. S. K. Kivelä: matriisilasku ja lineaarialgebra, luvut 2, 3, 4 ja 7

Toteutustavat: Luennot 44 h, harjoitukset 20 h.

Suoritustavat:

- a) hyväksytty osallistuminen harjoituksiin ja välikokeet (hyväksytyn osallistumisen kriteeri ilmoitetaan ensimmäisellä luennolla ja opintojakson verkkosivuilla) tai
- b) tentti.

Opetus- ja suorituskielet: suomi

Arvostelu: Asteikolla 1-5 tai hylätty, laskuharjoituksista saa lisäpisteitä.

Vastuuhenkilö: Matti Laaksonen

Opettaja: Matti Laaksonen

Vastuuorganisaatio: Matemaattisten tieteiden yksikkö

Lisätietoja: kurssin materiaali www.uva.fi/~mla/math1040/

1.2 Ohjeita kurssin suorittajalle

Kurssi on Vaasan yliopiston teknillisen tiedekunnan opiskelijoille yksi ensimmäisistä kursseista. Kurssimateriaalia läpikäydessäsi kiinnitä huomiota seuraaviin seikkoihin:

Luvuista olioihin: Koulumatematiikkasi on keskittynyt luvuilla laskemiseen. Jonkin verran olet laskenut myös kirjaimilla. Nyt opimme laskemaan vektoreilla ja matriiseilla. On tärkeätä oppia ajattelemaan uusilla olioilla. Aluksi tämä voi tuntua oudolta, mutta nyt on tarkoitus valloittaa tämä uusi abstraktiotaso.

Algoritmit: Kurssin alussa opetellaan ratkaisemaan yhtälöryhmiä. Koulussa on käsitelty yhtälöparit ja kolmen yhtälön ryhmät. Nyt opimme uuden algoritmin (ratkaisumenettelyn). Ensimmäinen reaktiosi saattaa olla, että tämä on turhaa! Algoritmi on mekaaninen ja se soveltuu kaikkiin tapauksiin. On hyödyllistä osata tämä algoritmi. Opintoissasi tulet kohtaamaan monia algoritmeja. Nyt opimme kohtaamaan algoritmeja! Kurssin kuluessa on moni asia ilmaistu algoritmimuodossa, ja ne on tarkoitettu opeteltaviksi.

Käsitteet: Kurssin aikana opitaan useita käsitteitä, jotka ovat paljon käytössä (esimerkiksi lineaarisuus). Käytännön kokemus on osoittanut, että vaikka sanat tunnettaankin, niin usein niiden todellista merkitystä ei tiedetä. Aina, kun ajattelussa käytetään käsitteitä, joiden täsmällistä merkitystä ei tunneta, ollaan vaarallisilla vesillä. Johtopäätökset voivat mennä rankasti pieleen. Kun materiaalissa on annettu käsitteelle numeroitu määritelmä, niin tätä määritelmää voidaan kysyä tentissä. Tenttitehtävät eivät ole pelkästään laskuja, vaan tehtävistä ainakin yksi on sanallinen ja tyypillisesti siinä kysytään jonkin termin määritelmää.

Todistukset: Kurssimateriaali sisältää melko paljon todistuksia. Todistukset ovat aina sellaisia, että ne on mahdollista ymmärtää aikaisemman avulla. Jos todistus on kohtuuttoman mutkikas, niin toteamme vain, että: "on mahdollista todistaa, että ...".

Jos pystyt seuraamaan todistukseen liittyvän ajatuksenjuoksun, niin käy se läpi huolellisesti ja iloitse. Jos todistus jää sinulta ymmärtämättä, niin älä missään tapauksessa opettele sitä ulkoa. Todistusten opetteleminen ulkoa on hyödytöntä itsensä uuvuttamista.

Kun myöhemmin työelämässä joudut ottamaan selville asioita, joudut perustelemaan asioita itsellesi ja työtovereillesi. Silloin olet vahvemmilla, jos olet tällä kurssilla käynyt todistukset läpi ja tajunnut niiden tyypilliset ajatuskuviot.

Jos haluat varmistaa ylimmän arvosana, niin käy myös todistukset läpi kokeeseen valmistautuessasi. Toisaalta suurin osa koetehtävistä on suorita laskutehtäviä, joten kurssi on mahdollista suorittaa osaamatta todistuksia. Nyt on kysymys siitä, että tunnistat omat lahjasi ja kehität niitä..

Kaavat: Kaavoja ei tarvitse opetella ulkoa. Jos lasket viikkoharjoituksia ahkerasti, niin tärkeimmät kaavat muistat joka tapauksessa. Monisteen lopussa oleva kaavakoelma saa olla tentissä mukana.

Oppimistavoitteet: Kurssin jälkeen opiskelijan tulee osata tietyt asiat. Osattavien asioiden lista on sen verran pitkä, että se on sijoitettu liitteenä monisteen loppuun. Listaa kannattaa seurata kurssin kuluessa.

Tietokoneen käyttö: Tietokoneen käyttö on asetettu tiedekunnan tasolla tärkeäksi oppimisaiheeksi. Jo ensimmäisten kurssien aikana käytämme tietokonetta kurssiin liittyvien esimerkkien ratkaisemiseen. Käytetyt ohjelmat, ympäristöt ja työkalut ovat jatkuvan kokeilu- ja kehittelytyön alla. Siksi kannattaa tarkistaa viimeiset ohjeet kurssin verkkosivulta (www.uva.fi/~mla/math1040/) ennen kuin asentaa mitään ohjelmia omalle tietokoneelle.

1.3 Ensimmäiset Octave-ohjelmat

Sec:EnsOhjelma

Nyt teemme ensimmäiset kaksi yksinkertaista ohjelmaa. Tässä kappaleessa ohjelmat toteutetaan Octavella. Liitteessä ^{laCodes}([laCodes](#)) on useimmille monisteen esimerkki-ongelmille esitetty myös Python-, Java-, ja C-kieliset toteutukset.

Code:es01

Ohjelma-esimerkki 1 Tee ohjelma joka tuottaa Fahrenheit/Celsius -muunnostaulukon. Jos t_F on lämpötila Farenheit asteissa ja t_C on vastaava lämpötila Celsius asteissa, niin

$$t_C = \frac{5}{9}(t_F - 32).$$

(Jotta tuloste ei ole liian pitkä, teemme taulukosta melko karkean. Voit itse muuttaa ohjelmaa siten, että se tekee tarkemman muunnostaulukon.)

Ohjelma tiedostossa: "laEs01.m"

```
# file: "laEs01.m"
printf("Fahrenheit,      Celsius\n");
for n=1:10
    tf = 10*n;
    tc = 5/9*(tf-32);
    printf("t_F = %6.2f,   t_C = %6.2f \n", tf, tc);
endfor
```

Huomautuksia koodiin:

1. Ristikkomerkki # ensimmäisen rivin alussa aloittaa kommentin, jonka tulkki sivuuttaa. Kaikki ristikko-merkin jälkeen rivillä oleva teksti tulkitaan kommentiksi.

Kommentit ovat useimmiten tarkoitettu ohjelmalistausta lukevalle ihmiselle. Tässä tapauksessa kommentti kertoo, minkä nimiseen tiedostoon koodi on talletettu. Joskus kommentteilla ohjataan kääntäjän toimintaa, mutta sellaisia kommentteja on tässä materiaalissa vain muutama.

Huomaa, että ristikkomerkki (risuaita) on eri asia, kuin 'hashtag'. Vasta kun ristikkomerkkiin yhdistetään avainsana, syntyy tunniste, eli 'hashtag'. Esimerkiksi '#linnanjuhlat' on hashtag.

käyttö	Suom.	Engl. (USA)	out of USA	esim.
Yleisnimitys	ristikko	number sign		
Numeron edessä	numero	number		#2
Puhelimessa	ristikko	pound	hash	

Ristikkomerkki # on myös eri asia kuin *sharp* ‡, joka on alunperin nuoteissa esiintyvä ”ylennysmerkki”. Microsoftin ”C sharp”-ohjelmointikielen logo C# (lue: *si: sha:p*) toteutetaan yleensä ristikkomerkin avulla.

2. Rivillä: `printf("Fahrenheit, Celsiusn");` komento `printf("...");` on funktio, joka tulostaa konsolille argumenttinsa, joka on merkkijono. Merkkijono päättyy tällä kertaa rivinvaihtoon `\n`.
3. `for n=1:10 ... endfor` on toistorakenne, jossa samoja komentoja toistetaan kymmenen kertaa `n:n` arvoilla 1,..., 10.
4. Rivi: `tf = 10*n;` päättyy puolipisteeseen, kuten moni komento Octavessa. Kyseessä on sijoituskomento, joka laskee lausekkeen `10*n` arvon ja sijoittaa sen muuttujan `tf` arvoksi. Jos komennon perässä ei ole puolipistettä, kääntäjä kirjoittaa konsolille muuttujan `tf` nimen ja sen saaman uuden arvon. Jos puolipiste on komennon perässä, niin tulkki ei kirjoita konsolille mitään. Ohjelman voisi siis käytännössä toteuttaa seuraavasti:

```
for n=1:10
    tf = 10*n
    tc = 5/9*(tf-32)
endfor
```

mutta tuloste on hieman sekava. (Kokeile!)

5. Toinen `printf` funktio tulostaa taas rivin tekstiä kuitenkin niin, että ensimmäinen `'%6.2f'` korvataan kuudella merkillä, joihin tulee muuttujan `tf` arvo kahdella desimaalilla ja toinen `'%6.2f'` korvataan kuudella merkillä, joihin tulee muuttujan `tc` arvo kahdella desimaalilla. Kirjain `f` kertoo tulkille, että tulostettavan muuttujan arvo on desimaaliluku (engl. `float`).

Ajo:

```
octave:1> laEs01
Fahrenheit,    Celsius
t_F = 10.00,   t_C = -12.22
t_F = 20.00,   t_C = -6.67
t_F = 30.00,   t_C = -1.11
t_F = 40.00,   t_C = 4.44
t_F = 50.00,   t_C = 10.00
t_F = 60.00,   t_C = 15.56
t_F = 70.00,   t_C = 21.11
t_F = 80.00,   t_C = 26.67
t_F = 90.00,   t_C = 32.22
t_F = 100.00,  t_C = 37.78
octave:2>
```

Code:es02

Ohjelma-esimerkki 2 Piirrä funktion $f(x) = x^2 - 2x$ kuvaaja välillä $0 \leq x \leq 3$. Tee me kuvasta hiukan karkean. Voit parantaa kuvaa.

Ohjelma tiedostossa: "laEs02.m"

```
# file: "laEs02.m"
x = [0 0.5 1 1.5 2 2.5 3]
for k=1:7
    y(k) = x(k)^2-2*x(k);
endfor
plot(x, y, 'b-')
```

Fig:Code:es02

Ohjelman ajosta syntyvä graafi on kuvassa II.

Ohjelman "laEs02.m" viimeisellä rivillä olevan `plot`-komennon kolmas argumentti on merkkijono joka ohjaa piirrettävän käyrän ominaisuuksia. Jos merkkijonosta löytyy alla olevan taulukon I mukaisia merkkejä, niin piirrettävän murtoviivan ominaisuudet muuttuvat vastaavasti. Esimerkiksi piirtokomento `plot(x, y, 'r:o')` tuottaa murtoviivan, joka on piirretty punaisella värillä, viivat ovat katkoviivoja ja pisteet merkitty pienillä palloilla.

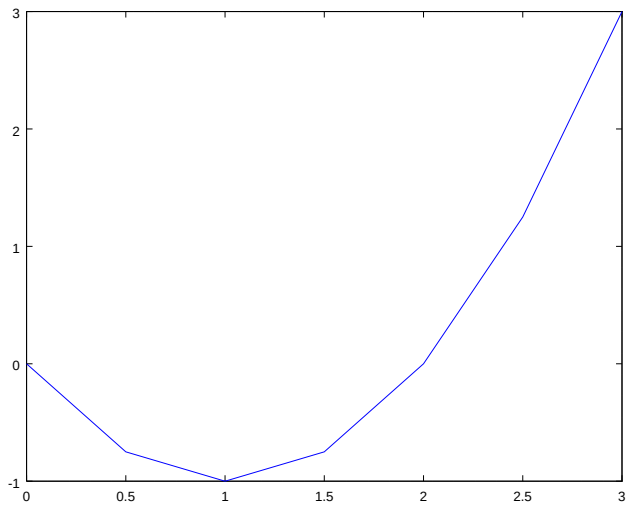


Fig:Code:es02

Kuva 1. Ohjelma-esimerkin 2 tuottama kuva.

Taulukko 1. Octaven plot-komennon viivan ominaisuudet.

ainenOhjelma01

väri		piste	viiva
w	whitew	. point	– dolid
m	magenta	o circle	: dotted
c	cyan	x x-mark	–. dashdot
r	red	+ plus	– dashed
g	green	* star	
b	blue	s square	
y	yellow	d diamond	
k	black	v triangle down	
		^ triangle up	
		< triangle left	
		> triangle right	
		p pentagram	
		h hexagram	

2 TRIGONOMETRIA JA FYSIIKAN VEKTOREITA

Ch:TrigFys

2.1 Trigonometriset funktiot

Sec:TrigFunk

2.1.1 Kulmayksiköistä

Aste, 1° (engl. degree) Kun kellon viisari kiertyy yhden kierroksen, sanomme, että se kääntyy 360° (360 astetta). Ajatus täyden kierroksen jakamisesta 360 asteeseen, juontaa kaldealaiseen astrologiaan. Yhdessä vuorokaudessa aurinko siirtyy noin yhden asteen tähtikuvioden joukossa. Oikeastaan täysi kierros tulisi jakaa 365 osaan, mutta sekään ei karkausvuosien takia anna täysin oikeata jakoa. 360 jako-osien määränä on käytännöllinen, koska se on helppo jakaa pienempiin osiin. Yksi aste jaetaan 60 kulminuuttiin ($1^\circ = 60'$) ja yksi kulminuutti jaetaan 60 kulmasekuntiin ($1' = 60''$). Siis aurinko kiertää yhden kulminuutin kulman 24 minuutissa, ja yhden kulmasekunnin kulman 24 sekunnissa.

Gradiaani, grad, eli gooni (engl. grad, ransk grade, saks. gon, ruots. gon) on 1/400 täydestä kierroksesta.

grad, grade [1], gradian, or gon (1^g or gr or grd) is a unit of angle measurement equal to 1/400 circle, 0.01 right angle, 0.9° , or $54'$. This unit was introduced in France, where it is called the grade, in the early years of the metric system. The grad is the English version, apparently introduced by engineers around 1900. The name gon is used for this unit in German, Swedish, and other northern European languages in which the word grad means degree. Although many calculators will display angle measurements in grads as well as degrees or radians, it is difficult to find actual applications of the grad today. ¹

Radiaani, rad, on kulman ympyrästä erottaman kaaren pituuden suhde ympyrän säteeseen (kun ympyrän keskipiste on kulman kärjessä). Täysi kierros on silloin 2π radiaania. Suorakulma on $\frac{\pi}{2}$ rad.

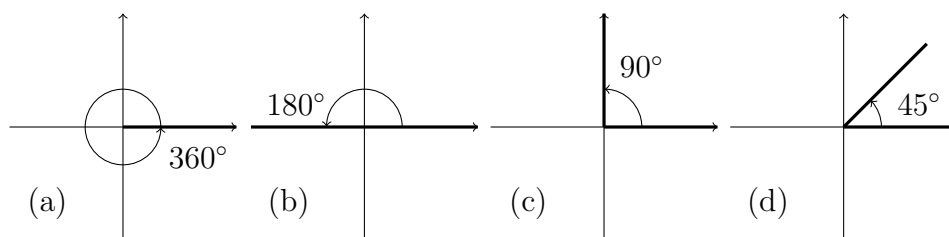
¹ <http://www.unc.edu/~rowlett/units/dictG.html>

Piiru, Kompassipiiru on $1/32$ täydestä kierroksesta. Siis $1 \text{ piiru} = 11,25^\circ$. Kompassipiiru näkyy edelleen merenkulun säädöksissä.

Tykistön piiru on karkeasti se kulma, jossa metrin kokoinen esine näkyy kilometrin päästä katsottuna. Siis täysi kierros on noin $1000 \cdot 2\pi \text{ piirua} \approx 6283,2 \text{ piirua}$. Eri armeijat jakavat täyden kierroksen piiruiksi hiukan eri tavoin. Täysi kierros on 6000 piirua (venäjä), 6300 piirua (Ruotsi) tai 6400 piirua (Nato). Suomessa ollaan siirtymässä venäläisestä piirusta nato-piiruun.

$$360^\circ = 400 \text{ grad} = 2\pi \text{ rad} = 6400 \text{ piirua}$$

Kun tiedämme, mitä täysi kierros on valitussa kulma-asteikossa, niin *oikokulma* on $1/2$ täydestä kierroksesta ($180^\circ = 200 \text{ grad} = \pi \text{ rad} = 3200 \text{ piirua}$), ja *suorakulma* on $1/4$ täydestä kierroksesta ($90^\circ = 100 \text{ grad} = \frac{\pi}{2} \text{ rad} = 1600 \text{ piirua}$).

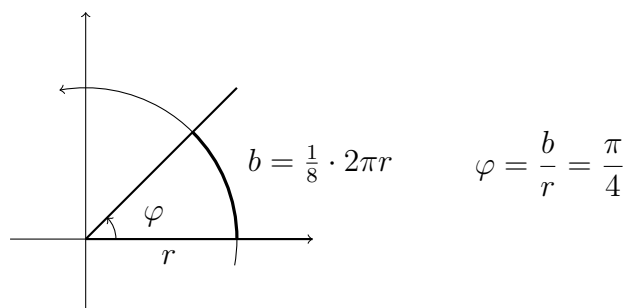


Kuva 2. (a) täysi kierros 2π , (b) oikokulma π , (c) suora kulma $\pi/2$, (d) suoran kulman puolikas $\pi/4$.

Kulma α on *terävä*, jos $0^\circ \leq \alpha < 90^\circ$ ($0 \leq \alpha < \pi/2$). Kulma α on *tylppä*, jos $90^\circ < \alpha < 180^\circ$ ($\pi/2 < \alpha < \pi$).

Radiaani-kulmayksikkö ”rad” kirjoitetaan näkyviin, jos halutaan lukijan varmasti tietävän missä yksiköissä kulma on ilmaistu. Jos käytetty yksikkö on asiayhteydestä selvä, niin silloin ”rad” jätetään pois. Tämä ei ole virhe, sillä kulma radiaaneissa on kaaren pituus jaettuna säteen pituudella, ja tämä laskutoimitus antaa tulokseksi paljaan luvun.

Tällä kurssilla käytetään kulmayksikkönä joko astetta tai radiaania. (Piiru voi jossakin erikoistilanteessa esiintyä, mutta gradia emme käytä koskaan. Opettele omassa laskimessasi tarkistamaan laskimen kulmayksikkö-asetus ja tarvittaessa muuttamaan se. Kokeessa ei nyt väärän vastauksen selitykseksi kelpaa, että ”Tein kyllä kaiken oikein, mutta laskimessa olikin grad-asetus päällä.”)

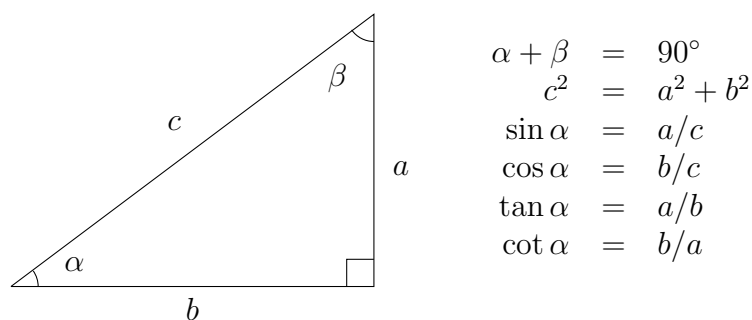


Kuva 3. Täyden kierroksen kahdeksas osa radiaaneina.

2.1.2 Suorakulmainen kolmio

Suorakulmainen kolmio on koulusta tuttu kolmio, jonka yksi kulma on suora (90° , tai $\pi/2$ rad). Käytämme seuraavassa melko laajasti vakiintuneita merkintäperiaatteita:

- Kolmion kulmia merkitsemme kreikkalaisin kirjaimin α , β , γ , jne. Suoraa kulmaa ei aina merkitä kirjaimella, vaan se merkitään kuvaan pienellä neliöllä.
- Kolmion sivuja ja sivujen pituuksia merkitsemme tavallisin kirjaimin a , b , c , jne.
- Ensimmäisessä kuvassa kulmaa α vastassa on sivu a , kulmaa β vastassa on sivu b jne. Kun ensimmäistä kuvaa analysoidaan jakamalla kolmioita pienempiin osiin, joudutaan tästä periaatteesta yleensä luopumaan.
- Suoran kulman kyljet ovat kateetit a ja b , ja suoraa kulmaa vastassa on hypoteenuusa c .



Kuva 4. Suorakulmainen kolmio.

Fig:laSKkolmio

Suorakulmaisen kolmion terävälle kulmalle määritellään seuraavat trigonometriset funktiot (ks kuva 4):

Fig:laSKkolmio

Sini = vastainen kateetti jaettuna hypotenuusalla $\sin \alpha = \frac{a}{c}$

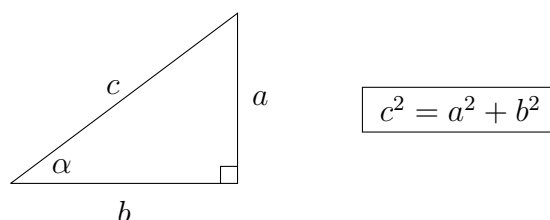
Kosini = viereinen kateetti jaettuna hypotenuusalla $\cos \alpha = \frac{b}{c}$

Tangentti = vastainen kateetti jaettuna viereisellä kateetilla $\tan \alpha = \frac{a}{b}$

Kotangentti = viereinen kateetti jaettuna vastaisella kateetilla $\cot \alpha = \frac{b}{a}$

Funktioiden arvot eivät riipu kolmion suuruudesta. Pienessä ja suuressa suhteet ovat samat.

Lause 1 (Pythagoran lause) Suorakulmaisessa kolmiossa hypotenuusan pituuden neliö on kateettien pituuksien neliöiden summa.



Seuraavassa lauseessa ja tästä eteenpäin muutenkin merkinnällä $\sin^2 \alpha$ tarkoitetaan lauseketta $(\sin \alpha)^2$. Merkintätapa on aiemmin ollut välttämätön, kun on käsitelty pitkiä trigonometrisia sarjakehitelmiä. Muussa yhteydessä merkintä tuntuu aluksi epäloogiselta, mutta siihen tottuu nopeasti, ja sitä on syytä osata käyttää.

Lause 2 (Seuraus Pythagoran lauseesta) Jos α on terävä kulma, niin

$$\sin^2 \alpha + \cos^2 \alpha = 1.$$

Perustelu: Tarkastellaan mitä tahansa suorakulmaista kolmiota, jossa kulma α on terävänä kulmana. Nimetään kolmion sivut ja muut kulmat tavalliseen tapaan. Silloin $a = c \cdot \sin \alpha$ ja $b = c \cdot \cos \alpha$ ja Pythagoran lauseen mukaan

$$\begin{aligned} a^2 + b^2 &= c^2 \\ \Leftrightarrow (c \cdot \sin \alpha)^2 + (c \cdot \cos \alpha)^2 &= c^2 \\ \Leftrightarrow \sin^2 \alpha + \cos^2 \alpha &= 1. \quad \square \end{aligned}$$

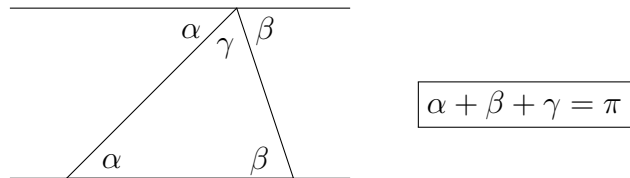
2.1.3 Kolmiolauseet

Tässä kappaleessa kolmion ei tarvitse olla suorakulmainen, ellei erikseen niin sanota. Koska kolmiossa voi olla tylppä kulma, sovimme tylpän kulman sinille ja kosinille seuraavat sopimukset:

$$\text{jos } \pi/2 \leq \alpha \leq \pi, \text{ niin } \sin(\alpha) = \sin(\pi - \alpha), \quad (2.1) \quad \boxed{\text{EqSinTylp}}$$

$$\text{jos } \pi/2 \leq \alpha \leq \pi, \text{ niin } \cos(\alpha) = -\cos(\pi - \alpha). \quad (2.2) \quad \boxed{\text{EqCosTylp}}$$

Lause 3 Kolmion kulmien summa on π (eli 180°).

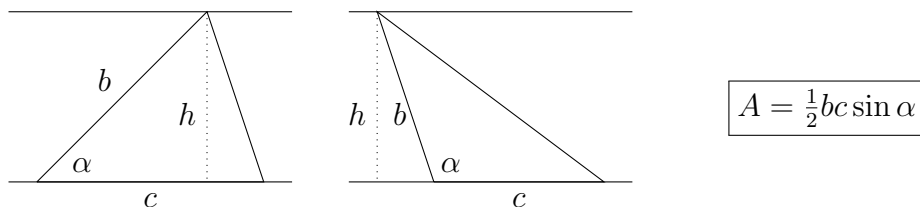


Perustelu: Kuvan mukaisesti kolmion huipun kautta piirretään kannan suuntainen suora. Kun nyt kolmion vasen kylki leikkaa kahta yhdensuuntaista suoraa, ovat syntyvät ristikulmat yhtäsuuret. Kolmion kantakulmaa α vastaava ristikulma on myös suuruudeltaan α . Vastaavasti kantakulmaa β vastaava ristikulma on myös suuruudeltaan β . Huipun kohdalle syntyvä kolmen kulman summa on oikokulma.

Lause 4 (Seuraus edellisestä) Jos kolmiossa on tylppä kulma ($> \pi/2$), niin muut kaksi kulmaa ovat teräviä ($< \pi/2$).

Lause 5 Kolmion ala on puolet kahden sivun pituuksien ja niiden välisen kulman sinin tulosta. (Kaavana: $A = \frac{1}{2} \cdot b \cdot c \cdot \sin \alpha$.)

Th:kolmionala



Perustelu: Kuvan mukaisesti kolmion huipun kautta piirretään kannan suuntainen suora. Piirretyn suoran ja kannan välinen etäisyys on kolmion korkeus h . Olkoon α kannan

c ja kyljen b välinen kulma. Silloin kolmion korkeus on $h = b \sin(\alpha)$. Kolmion pinta-ala on puolet kannan ja korkeuden tulosta, eli

$$A = \frac{1}{2}ch = \frac{1}{2}bc \sin \alpha.$$

□

Lause 6 Sinilause:

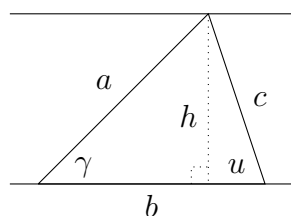
$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma}.$$

Perustelu: Lauseen ^{Th:kolmionala}5 mukaan kolmion kaksinkertainen pinta-ala voidaan laskea seuraavilla tavoilla: $bc \sin \alpha = ac \sin \beta = ab \sin \gamma$. Väite seuraa, kun lausekkeet jaetaan abc :lla ja siirrytään käänteislausekkeisiin. □

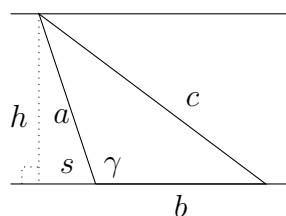
Lause 7 Kosinilause:

$$c^2 = a^2 + b^2 - 2ab \cos \gamma.$$

Perustelu:



Tapaus: "Terävä γ "
 $b - u = a \cos \gamma$



Tapaus: "Tylppä γ "
 $s = -a \cos \gamma$

Yllä olevien kuvien merkinnöin:

Jos γ on terävä, niin

$$\begin{aligned} c^2 &= h^2 + u^2 \\ &= a^2 - (b - u)^2 + u^2 \\ &= a^2 - b^2 + 2bu \\ &= a^2 + b^2 - 2b(b - u) \\ &= a^2 + b^2 - 2ab \cos \gamma \end{aligned}$$

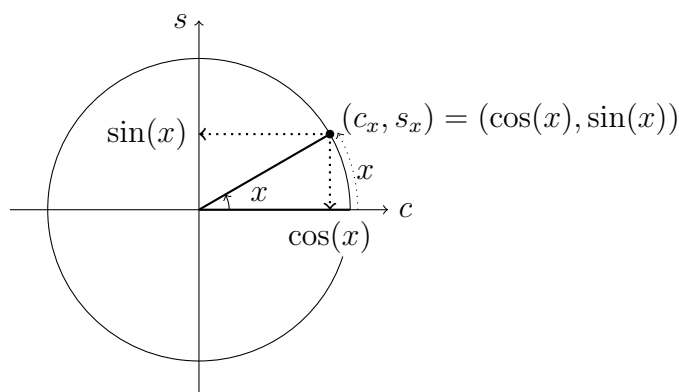
Jos γ on tylppä, niin

$$\begin{aligned} c^2 &= h^2 + (b + s)^2 \\ &= a^2 - s^2 + (b + s)^2 \\ &= a^2 + b^2 + 2bs \\ &= a^2 + b^2 - 2ab \cos \gamma \end{aligned}$$

□

2.1.4 Trigonometriset funktiot

Edellä kulma α oli suorakulmaisen kolmion terävä kulma, joten tähän mennessä $\sin \alpha$, $\cos \alpha$, $\tan \alpha$ ja $\cot \alpha$ on määritelty vain, kun $0 \leq \alpha < \pi/2$. Seuraavaksi määrittelemme, mitä $\sin(x)$, $\cos(x)$, $\tan(x)$ ja $\cot(x)$ ovat, kun x on mikä tahansa reaaliluku ($x \in \mathbb{R}$). Tämän kappaleen merkinnät (x ja sulut sen ympärillä) korostavat trigonometristen funktioiden funktio-luonnetta. Myöhemmin x korvautuu taas muilla kirjaimilla ja sulut jäävät pois.



Kuva 5. Kulman x kehäpisteen koordinaatit.

Piirretään (c, s) -koordinaatisto ja sen päälle origokeskinen yksikköympyrä. c -akseli osoittaa vaakasuoraan oikealle ja s -akseli osoittaa ylöspäin. (Käytämme nyt c - ja s -akseleita, koska muuttuja x on varattu kulmalle.) Piirretään koordinaatistoon yksikköympyrä $c^2 + s^2 = 1$ ja piirretään koordinaatistoon 0-kulma niin, että kulman kärki on origossa ja kulman kumpikin kylki on positiivisella c -akselilla. Avataan kulmaa kääntämällä vasenta kylkeä niin, että vasemman kyljen ja ympyrän kehän leikkauspiste, eli kehäpiste, liikkuu kehää pitkin positiiviseen suuntaan matkan x . Silloin kulman suuruus on x radiaania.

Kulman vasen kylki leikkaa yksikköympyrän kehän kehä-pisteessä (c_x, s_x) . Kehäpisteen s -koordinaattia sanomme kulman x siniksi, ja käytämme sille merkintää

$$\sin x,$$

ja kehäpisteen c -koordinaattia sanomme kulman x kosiniksi, ja käytämme sille merkintää

$$\cos x.$$

Sinin arvo, eli kehäpisteen s -koordinaatti, ja kosinin arvo, eli kehäpisteen c -koordinaatti, riippuvat kulman suuruudesta x . Näin saamme funktiot

$$x \mapsto \sin(x), \quad \text{ja} \quad x \mapsto \cos(x).$$

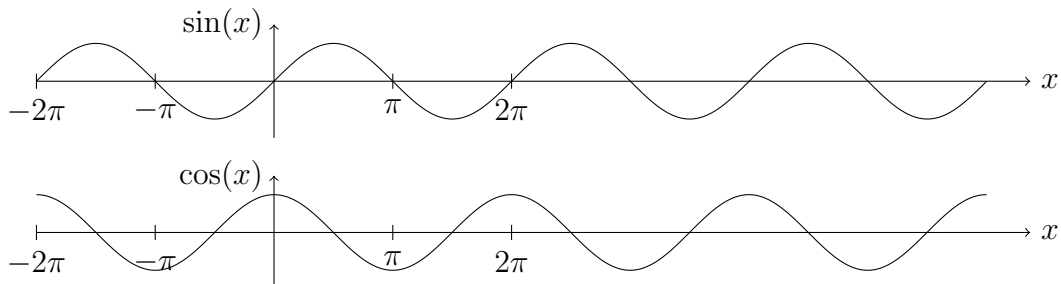
Funktioiden $\sin(x)$ ja $\cos(x)$ arvot on nyt määritelty kahdella erilaisella tavalla: (1) käyttäen suorakulmaista kolmiota ja (2) käyttäen yksikköympyrää. Kun kulma x on terävä, eli $0 \leq x \leq \pi/2$, niin kumpikin määritelmä antaa saman arvon sinille, ja kumpikin määritelmä antaa saman arvon kosinille.

Yksikköympyrän avulla sinille ja kosinille saadaan arvot myös silloin, kun x on suurempi kuin suora kulma. Silloin kehäpiste kulkee kehällä enemmän kuin neljänneksen kierroksen. Kulma x voi olla miten suuri tahansa. Tarvittaessa kehäpiste kiertää kehän vaikka useampaan kertaan kulkiessaan matkan x . Kulma x voi myös olla negatiivinen. Silloin kehäpiste lähtee liikkeelle päinvastaiseen suuntaan, myötäpäivään.

Nyt voimme sanoa, että sini- ja kosinifunktiot on määritelty kaikilla x :n reaaliarvoilla. Kuvassa [laFig:SinCos](#) on funktioiden kuvaajat. Samoin voimme määritellä tangentti- ja kotangenttifunktioiden arvot

$$\begin{aligned}\tan(x) &= \sin(x)/\cos(x), & \text{kun } x \neq \pi/2 + n \cdot \pi, \text{ ja} \\ \cot(x) &= \cos(x)/\sin(x), & \text{kun } x \neq n \cdot \pi.\end{aligned}$$

Tangentin ja kotangentin arvoja emme siis määrittele kaikilla x :n reaaliarvoilla. Kuvassa [laFig:TanCot](#) on näiden funktioiden kuvaajat.



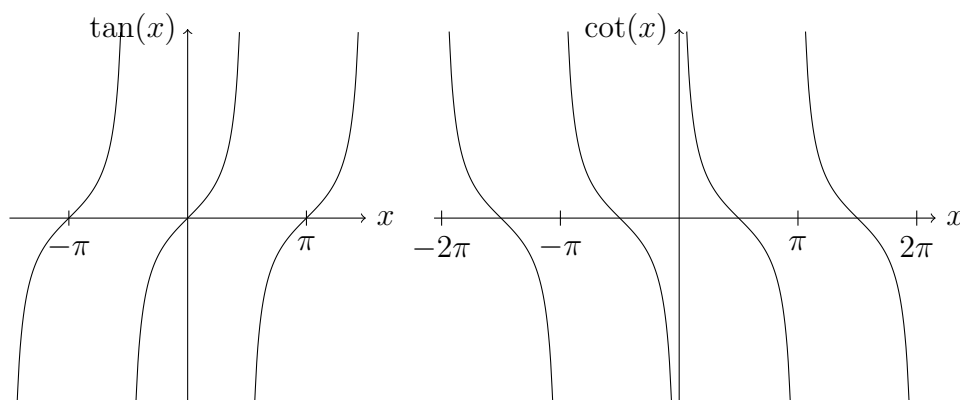
laFig:SinCos

Kuva 6. Sini- ja kosini-funktioiden kuvaajat.

Kun kulmaa kasvatetaan arvosta x arvoon $x + 2\pi$, kulmaa vastaava kehäpiste kiertää yhden täyden kierroksen ja lopussa kehäpiste on samassa paikassa, josta se lähti. Siis

$$\sin(x + 2\pi) = \sin(x), \quad \text{ja} \quad \cos(x + 2\pi) = \cos(x). \quad (2.3)$$

Sanomme, että sini- ja kosinifunktio ovat jaksollisia, perusjaksona 2π . Myös perusjakson kokonaislukumonikerrat ovat jaksoja ($\sin(x + n \cdot 2\pi) = \sin(x)$). Tangentti ja kotangentti ovat jaksollisia perusjaksona π .



laFig:TanCot

Kuva 7. tangentti- ja kotangentti-funktioiden kuvaajat.

Yksikköympyrän avulla on helppo nähdä, että myös seuraavat kaavat ovat tosia kaikilla x :n arvoilla:

$$-1 \leq \sin(x) \leq 1, \quad \text{ja} \quad -1 \leq \cos(x) \leq 1, \quad (2.4) \quad \text{TrigEq02}$$

$$\sin(-x) = -\sin(x), \quad \text{ja} \quad \cos(-x) = \cos(x), \quad (2.5) \quad \text{TrigEq03}$$

$$\sin(\pi - x) = \sin(x), \quad \text{ja} \quad \cos(\pi - x) = -\cos(x). \quad (2.6) \quad \text{TrigEq04}$$

$$\sin(\pi/2 - x) = \cos(x), \quad \text{ja} \quad \cos(\pi/2 - x) = \sin(x). \quad (2.7) \quad \text{TrigEq05}$$

Kaavan (2.5) kohdalla sanomme usein, että sinifunktio on pariton,² ja kosinifunktio on parillinen.³ Kaava (2.6) merkitsee samaa, kuin aiemmin tekemämme sopimukset (2.1) ja (2.2). Sopimukset olivat siis viisaasti valitut (ainoat järkevät). Usein tarvitsemme summakulman sinin ja kosinin kaavoja:

$$\sin(x \pm y) = \sin(x) \cos(y) \pm \cos(x) \sin(y) \quad (2.8)$$

$$\cos(x \pm y) = \cos(x) \cos(y) \mp \sin(x) \sin(y) \quad (2.9)$$

Trigonometrisille funktioille voidaan keksiä paljon kaavoja, joita emme nyt ryhdy luettelemaan. Kaavat löytyvät kaava-liitteestä ?? sivu ??.

2.1.5 Arcus-funktiot

Sini- ja kosinifunktiot ovat jatkuvia, mutta eivät monotonisia. Siksi niillä ei ole käänteisfunktioita (yleisessä merkityksessä). Kun sinifunktion määrittelyjoukko rajoitetaan

²Parittomalle funktiolle on voimassa $f(-x) = -f(x)$. Tämä on tyypillistä parittomalle potenssifunktiolle, kuten $f(x) = x^3$.

³Parilliselle funktiolle on voimassa $f(-x) = f(x)$. Tämä on tyypillistä parilliselle potenssifunktiolle, kuten $f(x) = x^2$.

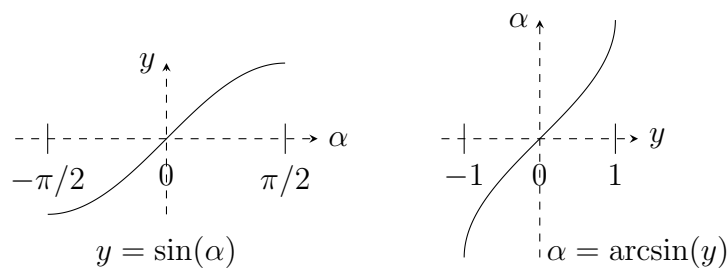
väliksi $I_n = [-\pi/2 + n \cdot \pi, \pi/2 + n \cdot \pi]$, niin sinin käänteisfunktio löytyy. Kun kosinifunktion määrittelyjoukko rajoitetaan väliksi $J_n = [0 + n \cdot \pi, \pi + n \cdot \pi]$, niin kosinin käänteisfunktio löytyy.

Tangenttifunktio on epäjatkuva kohdissa $x = \pi/2 + n \cdot \pi$, mutta se on monotoninen (kasvava) väleillä I_n . Kotangenttifunktio on epäjatkuva kohdissa $x = n \cdot \pi$, mutta se on monotoninen (vähenevä) väleillä J_n .

Määritelmä 8 Tässä luentosarjassa määrittelemme funktiot arcsin, arccos, arctan ja arccot väleille $I_0 = [-\frac{\pi}{2}, \frac{\pi}{2}]$ ja $J_0 = [0, \pi]$ rajoitettujen trigonometrinen funktioiden käänteisfunktioina seuraavasti:

funktio	käänteisfunktio
$\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$	$\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$
$\cos : [0, \pi] \rightarrow [-1, 1]$	$\arccos : [-1, 1] \rightarrow [0, \pi]$
$\tan : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow \mathbb{R}$	$\arctan : \mathbb{R} \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$
$\cot : [0, \pi] \rightarrow \mathbb{R}$	$\operatorname{arccot} : \mathbb{R} \rightarrow [0, \pi]$

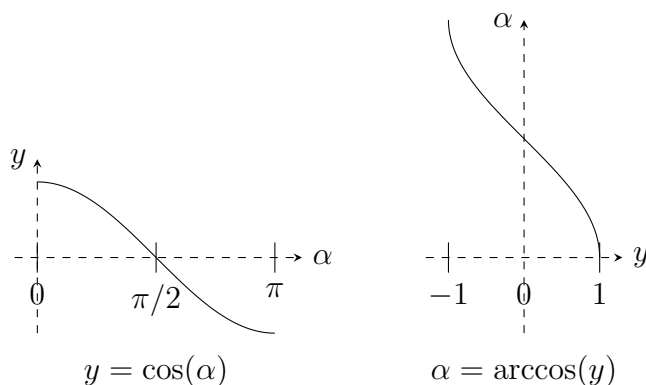
Usein kirjoissa käytetään arcus-funktioiden paikalla käänteisfunktio-merkintää ($\arcsin \sim \sin^{-1}$). Tämä on oikein ja perusteltua, jos kirjassa on trigonometriset funktiot määritetty kompleksimuuttujan kompleksiarvoisina funktioina. Silloin funktioteorian ja Riemannin pintojen avulla käänteisfunktio voidaan määritellä hyvin. Tässä kurssissa ei käsitellä Riemannin pintoja joten sovimme, että tässä materiaalissa arcsin ja arctan saavat arvon väliltä $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ja arccos ja arccot saavat arvon väliltä $[0, \pi]$.



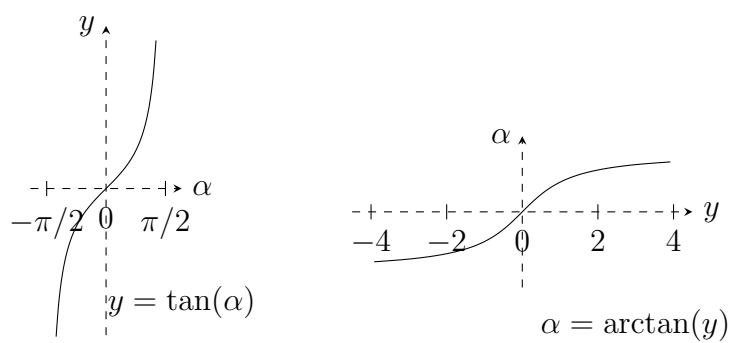
Kuva 8. Sini- ja arcus-sini -funktioiden kuvaajat.

aFig:SinArcsin

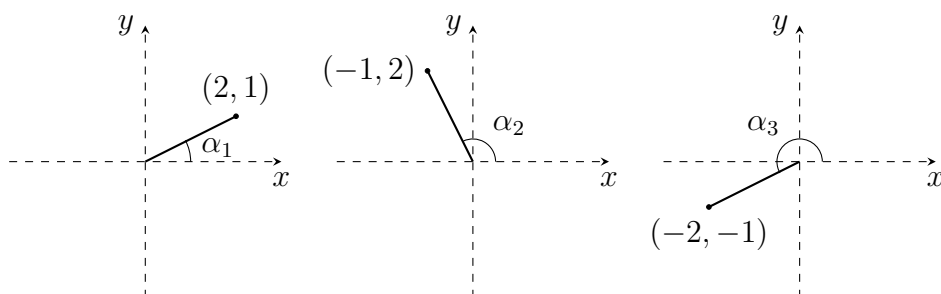
Esimerkki 9 Laske pisteiden $P_1 = (2, 1)$, $P_2 = (-1, 2)$, $P_3 = (-2, -1)$ paikkavektoreiden ja x -akselin väliset kulmat.



Kuva 9. Kosini- ja arcus-kosini -funktioiden kuvaajat.



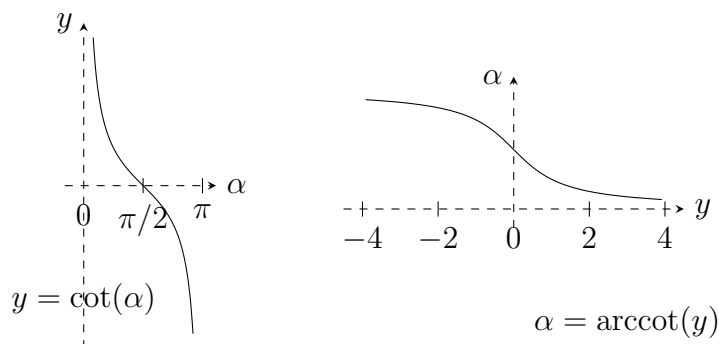
Kuva 10. Tangentti- ja arcus-tangentti -funktioiden kuvaajat.



Periaatteessa tehtävä on helppo, sillä

$$\tan \alpha_k = \frac{y_k}{x_k} \quad \Rightarrow \quad \alpha_k = \arctan\left(\frac{y_k}{x_k}\right) + n \cdot \pi.$$

Nyt tulee laskijan kuitenkin muistaa, että laskin antaa $\arctan$ -funktion arvon aina väliltä $-\pi/2 \leq \alpha \leq \pi/2$, joten laskijan tulee lisätä laskimen antamaan kulmaan π laskiessaan kulmia α_2 ja α_3 mutta ei tule lisätä laskimen antamaan arvoon mitään laskiessaan



Kuva 11. Kotangenti- ja arcus-kotangenti -funktioiden kuvaajat.

KotanArckotan

kulmaa α_1 . Siis

$$\begin{aligned}\alpha_1 &= \arctan\left(\frac{1}{2}\right) = 0.46365(\text{rad}) = 26.565^\circ \\ \alpha_2 &= \arctan\left(\frac{2}{-1}\right) + \pi = 2.03444(\text{rad}) = 116.565^\circ \\ \alpha_1 &= \arctan\left(\frac{-1}{-2}\right) + \pi = 3.60524(\text{rad}) = 206.565^\circ\end{aligned}$$

Johtopäätös esimerkin perusteella

$$\alpha = \begin{cases} \arctan(y/x), & \text{kun } x \geq 0, \\ \arctan(y/x) + \pi, & \text{kun } x < 0. \end{cases}$$

2.2 Pisteet ja vektorit suorakulmaisissa koordinaatistoissa

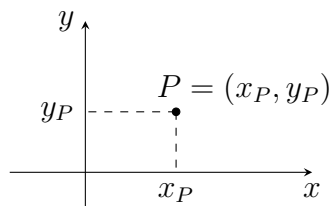
Sec:SKKoordSys

Opiskelija on jo koulussa käsitellyt tason ja avaruuden koordinaatistoja, mutta koulussa asia yleensä yksinkertaistetaan äärimmilleen niin, että syntyy mielikuva siitä, että on olemassa vain yksi koordinaatisto.

Tavallinen tason (x, y) -koordinaatisto

Koordinaatisto asetetaan siten, että x -akseli on vaakasuora, asteikko kasvaa oikealle, ja y -akseli on pystysuora, asteikko kasvaa ylöspäin. Akselit ovat siis kohtisuorassa. Akselit leikkaavat asteikkojen nollakohdissa. akselien leikkauspistettä sanomme origoksi.

Jos tason pisteen P kautta piirretty pystysuora leikka x -akselin kohdassa x_P , ja pisteen P kautta piirretty vaakasuora leikka y -akselin korkeudella y_P niin sanomme, että piste on kohdassa x_P korkeudella y_P ja merkitsemme $P = (x_P, y_P)$. Janaa, joka yhdistää



Kuva 12. Pisteen koordinaatit tasossa.

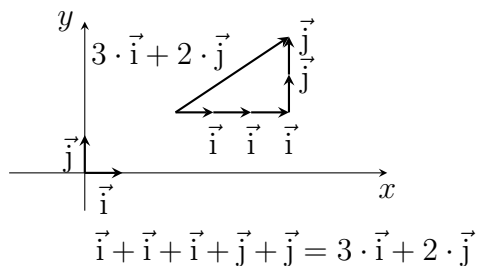
pisteet P ja Q merkitsemme kirjainparilla, jonka päällä on viiva, $\overline{PQ}$. Kun janalle annetaan suunta (P :stä Q :hun), saamme suuntajanan $\overrightarrow{PQ}$. Tarkasti tulkittuna suuntajalla on pituus, suunta ja paikka.

Vektorilla on pituus ja suunta. Kun piirrämme vektorin, niin piirrämme suuntajanan, jolla on sama pituus ja suunta kuin vektorilla. Sanomme, että mikä tahansa suuntajana, jolla on sama pituus ja suunta kuin vektorilla, edustaa vektoria. On tapana sopia nimet koordinaattiakselien suuntaisille yksikkövektoreille:

vektori	pituus	suunta
$\vec{i}$	1	x -akselin suunta
$\vec{j}$	1	y -akselin suunta

Suuntajanan pituus ja suunta on helppo esittää yksikkövektoreiden $\vec{i}$ ja $\vec{j}$ avulla:

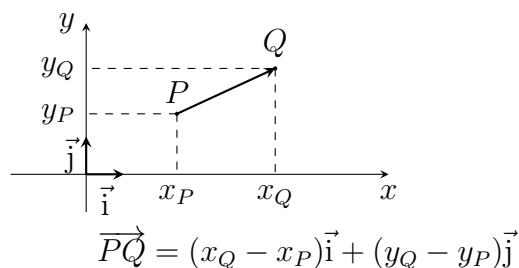
$$\text{”}a \text{ oikealle ja } b \text{ ylös”} = a \cdot \vec{i} + b \cdot \vec{j}$$



Kuva 13. ”Kolme oikealle ja kaksi ylös”.

Kun tiedämme suuntajanan päätepisteiden koordinaatit, niin suuntajanan pituus ja suunta saadaan samalla tavalla

$$\overrightarrow{PQ} = (x_Q - x_P) \cdot \vec{i} + (y_Q - y_P) \cdot \vec{j}$$

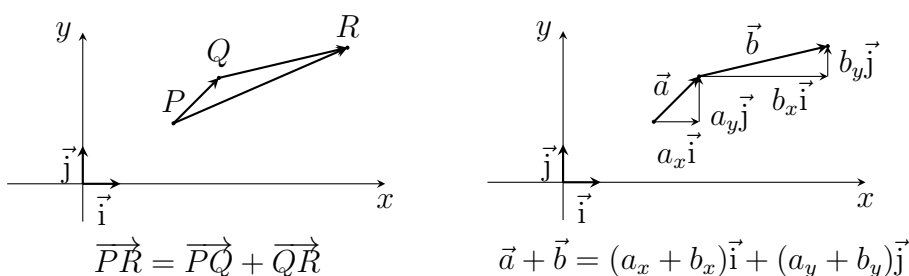


Kuva 14. Suuntajanan pituus ja suunta.

Pisteen $P = (x_P, y_P)$ paikkavektoria $\vec{r}_P$ edustaa suuntajana origosta O pisteeseen P , eli

$$\vec{r}_P = \overrightarrow{OP} = x_P\vec{i} + y_P\vec{j}.$$

Kahden vektorin, $\vec{a}$ ja $\vec{b}$, summa $\vec{a} + \vec{b}$ konstruoidaan piirtäen siten, että ensin piirretään jokin vektoria $\vec{a}$ edustava suuntajana $\overrightarrow{PQ}$, sitten piirretään pisteestä Q alkava vektorin $\vec{b}$ edustaja $\overrightarrow{QR}$. Nyt suuntajana $\overrightarrow{PR}$ edustaa summavektoria $\vec{a} + \vec{b}$.



Lause 10 Tason vektorin $\vec{a} = a_x\vec{i} + a_y\vec{j}$ pituus $|\vec{a}|$ voidaan laskea lausekkeesta

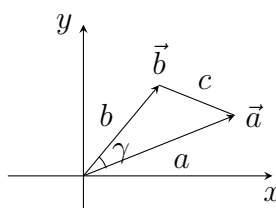
$$|\vec{a}| = \sqrt{a_x^2 + a_y^2}.$$

Perustelu: Kun piirrämme vektorille $\vec{a}$ origosta alkavan edustajan, syntyy luonnollisella tavalla suorakulmainen kolmio, jonka hypotenuusan pituus on sama kuin vektorin pituus ja kateettien pituudet ovat $|a_x|$ ja $|a_y|$. Väite seuraa nyt Pythagoran lauseesta. $\square$

Lause 11 Kaksi tasovektoria $\vec{a} = a_x\vec{i} + a_y\vec{j}$ ja $\vec{b} = b_x\vec{i} + b_y\vec{j}$ ja niiden välinen kulma γ toteuttavat yhtälön

$$\cos \gamma = \frac{a_x b_x + a_y b_y}{|\vec{a}| |\vec{b}|}.$$

Perustelu: Piirretään vektoreille origosta alkavat edustajat. Olkoon edustajien loppupisteitä yhdistävän janan pituus c . Kosinilauseen mukaan:



$$\begin{aligned}
 c^2 &= a^2 + b^2 - 2ab \cos \gamma \\
 (a_x - b_x)^2 + (a_y - b_y)^2 &= a_x^2 + a_y^2 + b_x^2 + b_y^2 - 2ab \cos \gamma \\
 -2a_x b_x - 2a_y b_y &= -2|\vec{a}||\vec{b}| \cos \gamma
 \end{aligned}$$

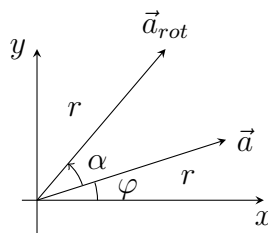
Lause 12 (a) Jos tasovektoria $\vec{a} = a_x \vec{i} + a_y \vec{j}$ kierretään origon ympäri positiiviseen kiertosuuntaan kulman α verran, niin kierretty vektori on

$$\vec{a}_{rot} = (\cos(\alpha) \cdot a_x - \sin(\alpha) \cdot a_y) \vec{i} + (\sin(\alpha) \cdot a_x + \cos(\alpha) \cdot a_y) \vec{j}.$$

(b) Jos tasovektoria $\vec{a} = a_x \vec{i} + a_y \vec{j}$ kierretään origon ympäri positiiviseen kiertosuuntaan kulman $\pi/2$ verran, niin kierretty vektori on

$$\vec{a}_{rot} = -a_y \vec{i} + a_x \vec{j}.$$

Perustelu: (a) Olkoon vektorin $\vec{a}$ pituus r ja vektorin ja x -akselin välinen kulma φ , jolloin $a_x = r \cos \varphi$ ja $a_y = r \sin \varphi$. Nyt kierretty vektori on



$$\begin{aligned}
 \vec{a}_{rot} &= r \cos(\alpha + \varphi) \vec{i} + r \sin(\alpha + \varphi) \vec{j} \\
 &= r(\cos \alpha \cos \varphi - \sin \alpha \sin \varphi) \vec{i} \\
 &\quad + r(\sin \alpha \cos \varphi + \cos \alpha \sin \varphi) \vec{j} \\
 &= (\cos \alpha \cdot a_x - \sin \alpha \cdot a_y) \vec{i} + (\sin \alpha \cdot a_x + \cos \alpha \cdot a_y) \vec{j}
 \end{aligned}$$

(b) Seuraa suoraan a-kohdasta, kun $\alpha = \pi/2$.

Konkreettisen tason (x, y) -koordinaatisto

Jatkossa monen esimerkin ratkaisu alkaa sillä, että asetetaan koordinaatisto, jossa laskut ja analyysit tehdään. Koordinaatisto siis valitaan tehtävään sopivalla tavalla. Edellisessä kappaleessa koordinaattiakselien suunnat olivat ”oikealle” ja ”ylös”. Tällä valinnalla saatiin aikaan se, että kun ensimmäisen koordinaattiakselin suuntaista vektoria käännetään positiiviseen kiertosuuntaan (vastapäivään) kulman $\pi/2$ verran, niin käännetty vektori on toisen koordinaattiakselin suuntainen.

Kun varustamme tason suorakulmaisella koordinaatistolla, valitsemme neljä ominaisuutta

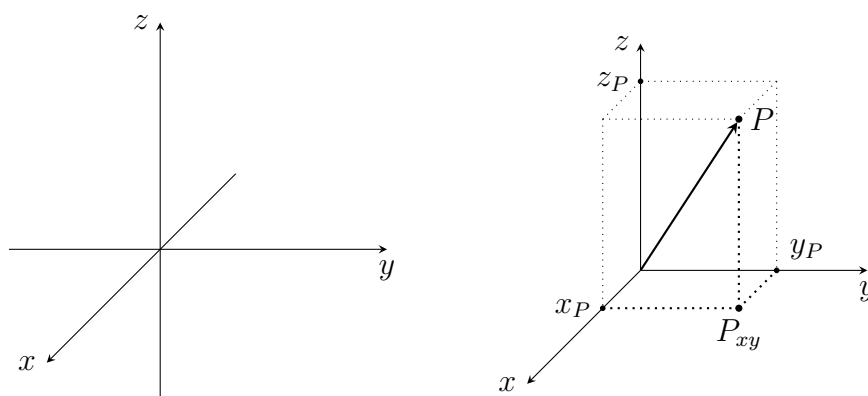
1. Origo. Piste, jossa koordinaattiakselit leikkaavat.
2. x -akseli.
3. y -akseli, joka on kohtisuorassa x -akselia vastaan ja saadaan x -akselista kiertämällä kulman $\pi/2$ verran positiiviseen kiertosuuntaan.
4. Akselien asteikko. Akseleilla tulee olla samat asteikot. Asteikkojen nollakohdat ovat origossa.

Edellä kohta 3 sisältää aidon valinnan, sillä suunnistus riippuu siitä, miltä puolelta katsoimme tasoa. Jos avaruusaseman kaksi astronauttia ovat korjaamassa aurinkopaneelia, ja ovat eri puolilla paneelia, niin he ovat perustellusti eri mieltä suunnistuksesta. Suunnistus riippuu siitä, miten valitsemme tason ylä- ja ala-puolen.

Lisää esimerkki

Tavallinen avaruuden (x, y, z) -koordinaatisto

Aloitetaan koordinaatiston konstruointi siten, että katsoja on kyykyssä ja origo on katsojan edessä. x -akseli on vaakasuora katsojaan päin, y -akseli on vaakasuora oikealle (kohtisuorassa x -akselia vastaan), ja z -akseli osoittaa ylöspäin (kohtisuorassa sekä x -akselia että y -akselia vastaan). Tämän jälkeen katsoja nousee seisomaan ja ottaa askeleen oikealle, ja piirtää kuvan näkemästään. Se, että katsoja muutti hieman asentoaan, saa x -akselinkin näkymään kuvassa hyvin.



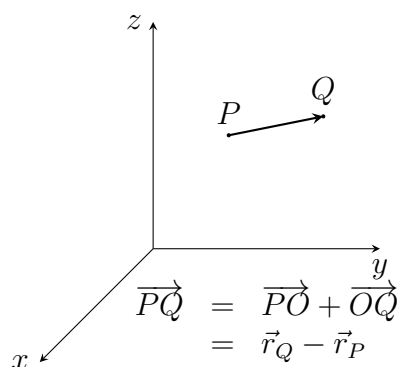
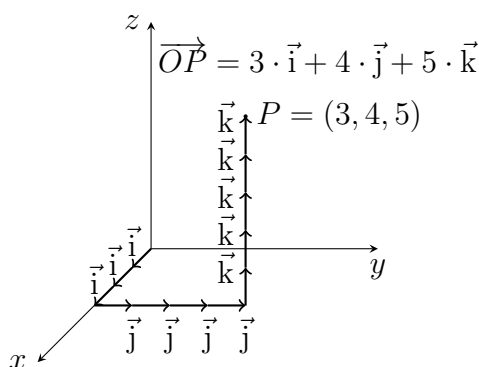
Kuva 15. (x, y, z) -koordinaatisto.

Samoin kuin edellä määritellään koordinaattiakselien suuntaiset yksikkövektorit

vektori	pituus	suunta
$\vec{i}$	1	x -akselin suunta
$\vec{j}$	1	y -akselin suunta
$\vec{k}$	1	z -akselin suunta

Olkoon P avaruuden mikä tahansa piste. Piirretään pisteen P kautta pystysuora (vektorin $\vec{k}$ suuntainen suora), joka leikkaa (x, y) -tason pisteessä $P_{xy} = (x_P, y_P, 0)$. Piste P_{xy} on pisteen P kohtisuora projektio (x, y) -tasoon. z -koordinaatti kertoo, miten korkealla P on (x, y) -tasosta. z_P on plus-merkkinen, jos P on (x, y) -tason yläpuolella, ja miinus-merkkinen, jos P on (x, y) -tason alapuolella. Nyt on selvää, että

$$\begin{aligned}\vec{OP} - z_P \vec{k} &= \vec{OP}_{xy} \\ \Leftrightarrow \vec{OP} - z_P \vec{k} &= x_P \vec{i} + y_P \vec{j} \\ \Leftrightarrow \vec{OP} &= x_P \vec{i} + y_P \vec{j} + z_P \vec{k}\end{aligned}$$



Kun piirrämme pisteestä P vaakasuoran janan z -akselille, niin janan akselin puoleinen päätepiste $P_z = (0, 0, z_P)$ on pisteen P projektio z -akselille.

Edellä sanoimme (x, y) -tasoa vaakasuoraksi ja z -akselia pystysuoraksi. Tämä helpottaa kuvion näkemistä, mutta ei ole mitenkään välttämätöntä. Akselit ovat täysin tasavertaiset. Samalla tavalla kuin edellä piste projisoitiin (x, y) -tasoon ja z -akselille, piste voidaan myös projisoida (y, z) -tasoon ja x -akselille, tai (z, x) -tasoon ja y -akselille.

Edellä konstruoitu koordinaatisto on esimerkki positiivisesti suunnistetusta koordinaatistosta. Suunnistuksen positiivisuuden voi testata ns. ”oikean käden säännöllä”:

Aseta oikean käden peukalo, etusormi ja keskisormi likimain kohtisuoraan toisiinsa nähden. Jos nyt pystyt kääntämään kätesi niin, että peukalo on x -akselin suuntainen, etusormi on y -akselin suuntainen, ja keskisormi on z -akselin suuntainen, niin koordinaatisto on positiivisesti suunnistettu.

Jos koordinaatisto on positiivisesti suunnistettu, niin kahden koordinaattiakselin määräämän tason suunnistus saadaan myös oikean käden avulla:

Olkoon (x, y, z) -koordinaatisto positiivisesti suunnistettu. Jos asetat oikean käden sormet ”peukutus-asentoon” ja asetat peukalon z -akselin suuntaiseksi, niin muut sormet näyttävät (x, y) -tason positiivisen kiertosuunnan. Jos asetat peukalon x -akselin suuntaiseksi, niin muut sormet näyttävät (y, z) -tason positiivisen kiertosuunnan. Jos asetat peukalon y -akselin suuntaiseksi, niin muut sormet näyttävät (z, x) -tason positiivisen kiertosuunnan.

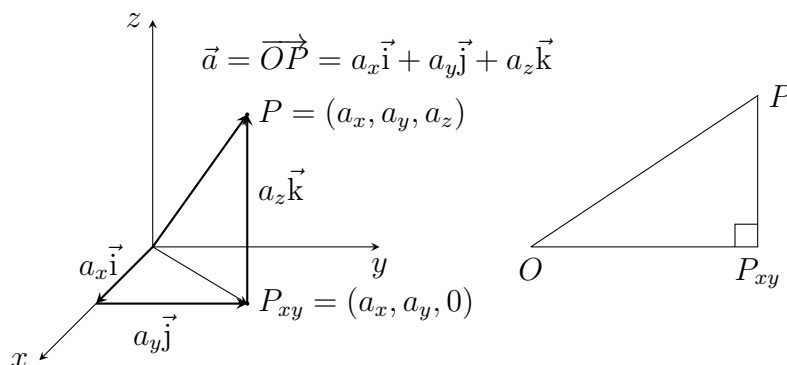
Lause 13 (1) Vektorin $\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k}$ pituus saadaan lausekkeesta

$$|\vec{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}.$$

(2) Olkoon $P = (x_P, y_P, z_P)$ ja $Q = (x_Q, y_Q, z_Q)$ kaksi pistettä. Pisteiden P ja Q etäisyys, eli janan $\overline{PQ}$ pituus, eli suuntajanan $\overrightarrow{PQ}$ pituus saadaan lausekkeesta

$$|\overline{PQ}| = \sqrt{(x_Q - x_P)^2 + (y_Q - y_P)^2 + (z_Q - z_P)^2}.$$

Perustelu: (1)



Vektori $\overrightarrow{OP_{xy}}$ on vaakasuora ja vektori $\overrightarrow{P_{xy}P}$ on pystysuora. Ne ovat siis kohtisuorassa. Kun kopioimme kolmion $\triangle OP_{xy}P$ kuvatasoon, syntyy suorakulmainen kolmio, jolle Pythagoran lauseen nojalla on voimassa

$$\begin{aligned} |\overrightarrow{OP}|^2 &= |\overrightarrow{OP_{xy}}|^2 + |\overrightarrow{P_{xy}P}|^2 \\ \Leftrightarrow |\vec{a}|^2 &= (a_x^2 + a_y^2) + a_z^2. \end{aligned}$$

(2) Seuraa kohdasta (1). □

Lause 14 Kaksi avaruusvektoria $\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k}$ ja $\vec{b} = b_x\vec{i} + b_y\vec{j} + b_z\vec{k}$ ja niiden välinen kulma γ toteuttavat yhtälön

$$\cos \gamma = \frac{a_x b_x + a_y b_y + a_z b_z}{|\vec{a}| |\vec{b}|}.$$

GammaAvaruus

Perustelu: Piirretään vektoreille origosta alkavat edustajat. Olkoon edustajien loppupisteitä yhdistävän janan pituus c . Kosinilauseen mukaan:

$$\begin{aligned} c^2 &= a^2 + b^2 - 2ab \cos \gamma \\ (a_x - b_x)^2 + (a_y - b_y)^2 + (a_z - b_z)^2 &= a_x^2 + a_y^2 + a_z^2 + b_x^2 + b_y^2 + b_z^2 - 2ab \cos \gamma \\ -2a_x b_x - 2a_y b_y - 2a_z b_z &= -2|\vec{a}| |\vec{b}| \cos \gamma \end{aligned}$$

□

Avaruusvektorin kiertäminen yleisesti käsitellään myöhemmin, kun olemme ensin saaneet sopivia työkaluja käyttöön. Jos voimme valita yhden koordinaattiakselin kiertoakselin suuntaiseksi, niin tehtävä helpottuu ja osaamme kuvata kierron helposti

Lause 15 (a) Jos avaruusvektoria $\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k}$ kierretään z -akselin ympäri positiiviseen kiertosuuntaan kulman α verran, niin kierretty vektori on

$$\vec{a}_{rot} = (\cos(\alpha) \cdot a_x - \sin(\alpha) \cdot a_y)\vec{i} + (\sin(\alpha) \cdot a_x + \cos(\alpha) \cdot a_y)\vec{j} + a_z\vec{k}.$$

(b) Jos tasovektoria $\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k}$ kierretään z -akselin ympäri positiiviseen kiertosuuntaan kulman $\pi/2$ verran, niin kierretty vektori on

$$\vec{a}_{rot} = -a_y\vec{i} + a_x\vec{j} + a_z\vec{k}.$$

Konkreettisen avaruuden (x, y, z) -koordinaatisto

Fysiikan laskuissa usein suunnat ”vaakasuoara” ja ”pystysuoara” ovat luonnollisia ja ongelmattomia. Aina näin ei kuitenkaan ole ja joskus laskut helpottuvat, kun akselien suunnat valitaan tilanteeseen sopivasti. jatkossa kuitenkin odotamme aina, että valitut akselit ovat keskenään kohtisuorassa ja muodostavat positiivisesti suunnistetun koordinaatiston.

Kun varustamme tilan suorakulmaisella koordinaatistolla, valitsemme taas neljä ominaisuutta

1. Origo. Piste, jossa koordinaattiakselit leikkaavat.
2. x -akseli.
3. y -akseli, joka on kohtisuorassa x -akselia vastaan. z -akseli määräytyy nyt siitä, että se on kohtisuorassa sekä x - akselia että y akselia vastaan, ja kaikki akselit yhdessä muodostavat positiivisesti suunnistetun koordinaatiston. (Kun siis x - ja y -akselit on valittu, niin z -akseli määräytyy niistä, eikä sen suhteen enää ole valinnan mahdollisuuksia).
4. Akselien asteikko. Akseleilla tulee olla samat asteikot. Asteikkojen nollakohtat ovat origossa.

Edellä kohta 3 sisältää enemmän vaihtoehtoja kuin tason tapauksessa.

Lisää esi-
merkki.

2.3 Fysiikan voima-vektori

Sec:ForceVec

Voiman vaikutus massapisteen liiketilaan ja kiinteiden rakenteiden jännityksiin, on luonnon ilmiö, jota opitaan käsittelemään fysiikan kursseilla. Tässä materiaalissa voimaa käsitellään vain siinä määrin kuin se on tarpeen vektorilaskujen temiseksi sovel-lusesimerkeissä.

Voimalla on kolme ominaisuutta

1. Voimalla on suuruus f joka mitataan Newtonissa.
2. Voimalla on suunta, joka on helpointa ilmaista yksikkövektorilla $\vec{e}$.
3. Voimalla on vaikutuspiste.

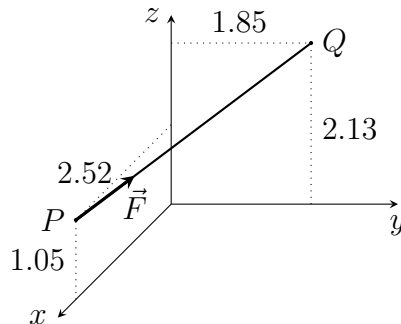
Tavallisesti vaikutuspiste tiedetään ja laskuissa tarvitaan voiman suuruutta ja suuntaa, joka esitetään voima-vektorina

$$\vec{F} = f \cdot \vec{e}.$$

Kun voimavektori $\vec{F}$ tiedetään, niin voiman suuruus on $f = |\vec{F}|$ ja voiman suunta on $\vec{e} = f^{-1}\vec{F}$, mutta vaikutuspiste ei sisälly merkintään.

Esimerkki 16 Huoneen pohjoisseinään kiinnitetään kiinnitysrengas 1,85 m luoteisnurkasta ja 2,13 m lattiasta. Toinen kiinnitysrengas kiinnitetään länsiseinään 2,52 m luoteisnurkasta ja 1,05 m lattiasta. Renkaiisiin kiinnitetään vaijeri, joka kiristetään niin, että kiristysvoima on 135,0 N. Arvioidaan, että vaijerin päätepiste on 8 mm seinän tasosta. Määritä länsiseinään vaikuttavan voiman vektori.

Ratkaisu: Asetetaan koordinaatiston origo lattian luoteisnurkkaan, x -akseli länsiseinän lattialistaa pitkin, y -akseli pohjoisseinän lattialistaa pitkin, z -akseli ylöspäin pitkin nurkkaa.



Vaijerin päätepisteiden koordinaatit ovat

$$P = (2.520; 0.008; 1.050),$$

$$Q = (0.008; 1.850; 2.130).$$

$$\begin{aligned}\vec{e} &= \frac{\vec{PQ}}{|\vec{PQ}|} = \frac{(x_Q - x_P)\vec{i} + (y_Q - y_P)\vec{j} + (z_Q - z_P)\vec{k}}{\sqrt{(x_Q - x_P)^2 + (y_Q - y_P)^2 + (z_Q - z_P)^2}} \\ &= \frac{-2.512\vec{i} + 1.842\vec{j} + 1.080\vec{k}}{\sqrt{(-2.512)^2 + 1.842^2 + 1.080^2}} \\ &= -0.76193\vec{i} + 0.55871\vec{j} + 0.32758\vec{k}\end{aligned}$$

$$\begin{aligned}\vec{F} &= 135.0\text{N} \cdot (-0.76193\vec{i} + 0.55871\vec{j} + 0.32758\vec{k}) \\ &= -102.9\text{N} \cdot \vec{i} + 75.43\text{N} \cdot \vec{j} + 44.22\text{N} \cdot \vec{k}\end{aligned}$$

Esimerkin laskut voidaan tehdä Octavella seuraavasti:

```
octave:1> P = [2.520; 0.008; 1.050];
octave:2> Q = [0.008; 1.850; 2.130];
octave:3> PQ = Q-P
PQ =

    -2.5120
     1.8420
     1.0800

octave:4> e = PQ/norm(PQ)
e =
```



```

-0.76193
 0.55871
 0.32758

octave:5> F = 135.0*e
F =

-102.860
  75.426
  44.223

```

Esimerkki 17 Tarkastellaan kappaletta, jonka massa on 7,50 kg. kappale on vaakasuoralla alustalla. Kappaleen ja alustan välinen täysin kehittynyt lepokitkan kitkakerroin on 0,6 (kumi–märkä asfaltti). Kappaleeseen vaikuttaa voima $\vec{F} = F\vec{e}$, jonka suunta muodostaa 20° kulman vaakatason kanssa. Miten suuri tulee voiman olla, jotta kappale lähtee liikkeelle.

Exa:Vec1

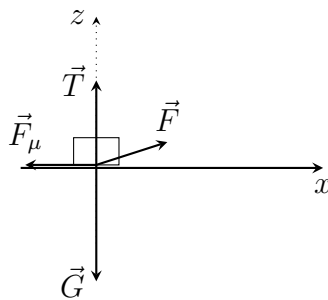


fig:Force01

Kuva 16. Kappale vaakasuoralla tasolla.

Kuvaan (16) on piirretty kappaleeseen vaikuttavia voimia: $\vec{G}$ kappaleen paino, $\vec{T}$ alustan tukivoima, $\vec{F}$ kappaleeseen vaikuttava ulkoinen voima, $\vec{F}_\mu$ kappaleen liikettä vastustava kitkavoima. Koska muut voimat ovat joko pystysuoria tai vaakasuoria, kannattaa myös voima $\vec{F}$ jakaa vaakasuoraan ja pystysuoraan komponenttiin.

$$\vec{F} = \underbrace{F \cos 20^\circ \cdot \vec{i}}_{=\vec{F}_x} + \underbrace{F \sin 20^\circ \cdot \vec{k}}_{=\vec{F}_z}.$$

Paino on

$$\vec{G} = 9,81 \frac{\text{m}}{\text{s}^2} \cdot 7,50 \text{ kg} \cdot (-\vec{k}) = -73,575 \text{ N} \cdot \vec{k}$$

Pystysuorien voimien summa on nolla, joten

$$\begin{aligned} \vec{T} + \vec{G} + \vec{F}_z &= \vec{0} \\ \Leftrightarrow \vec{T} &= -\vec{G} - F \sin 20^\circ \cdot \vec{k} \\ \Leftrightarrow \vec{T} &= (73,575 \text{ N} - F \sin 20^\circ) \vec{k} \end{aligned}$$

Lepokitka-voiman suuruus on enintään $0.6 \cdot |\vec{T}|$, joten kappale lähtee liikkeelle, jos

$$\begin{aligned} |\vec{F}_x| &> 0.6 \cdot |\vec{T}| \\ \Leftrightarrow F \cos 20^\circ &> 0.6 \cdot (73,575 \text{ N} - F \sin 20^\circ) \\ \Leftrightarrow F(\cos 20^\circ + 0.6 \cdot \sin 20^\circ) &> 0.6 \cdot 73,575 \text{ N} \\ \Leftrightarrow F &> \frac{0.6 \cdot 73,575 \text{ N}}{\cos 20^\circ + 0.6 \cdot \sin 20^\circ} = 38,6 \text{ N}. \end{aligned}$$

Vastaus: Koska kitkakerroin tiedetään melko epätarkasti, niin täsmällistä raja-arvoa ei voida ilmoittaa. Annettujen arvojen perusteella voidaan sanoa, että voiman tulee olla noin 39 N, jotta kappale lähtee liikkeelle. $\square$

Edellä voiman jakaminen vaakasuoraan ja pystysuoraan osaan oli oleellista ja myös helppoa. Jatkossa joudumme jakamaan voiman tason suuntaiseen ja tasoa vastaan kohtisuoraan komponenttiin, kun taso ei ole vaakasuora. Tähän on olemassa helpot työkalut, joita ryhdymme seuraavaksi opettelemaan.

2.4 Pistetulo ja ristitulo

Sec:DCIntro

Tässä kappaleessa määrittelemme kaksi vektoreille määriteltyä tuloa; pistetulo ja ristitulo.

Määritelmät ja erilaiset tulosten ominaisuudet saattavat tuntua, sekavalta kokonaisuudelta. Koko tämän kappaleen ajan pidä kirkkaana mielessä se ero, että kahden vektorin pistetulon arvo on luku, ja kahden vektorin ristitulon arvo on vektori.

Ex:DC1

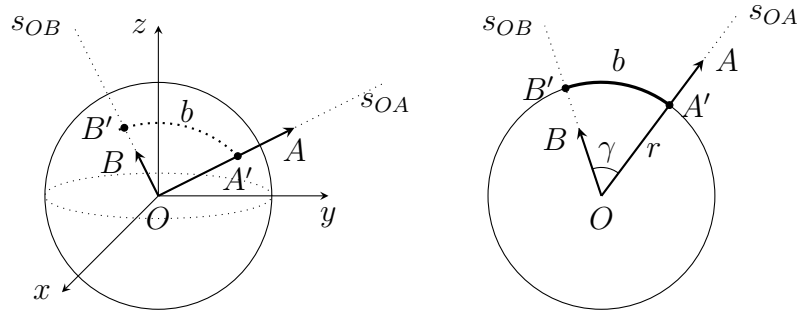
Esimerkki 18 Olkoon $\vec{a} = 2\vec{i} + 3\vec{j} - 4\vec{k}$, ja $\vec{b} = \vec{i} + 2\vec{j} + 3\vec{k}$. Silloin

$$\vec{a} \cdot \vec{b} = -4, \quad (2.10)$$

$$\vec{a} \times \vec{b} = 17\vec{i} - 10\vec{j} + \vec{k}. \quad (2.11)$$

Pian opimme laskemaan pistetulon ja ristitulon. Sitten voit palata tähän esimerkkiin ja tarkistaa laskut. Nyt on tärkeintä ymmärtää, että tulot antavat täysin eri tyyppiset tulokset.

Viimeistään tässä vaiheessa tulee määritellä, mitä tarkoitamme kahden avaruusvektorin välisellä kulmalla. Olkoon $\vec{a}$ ja $\vec{b}$ kaksi avaruusvektoria. Piirrämme vektoreita edustavat origosta alkavat suuntajana $\vec{a} = \overrightarrow{OA}$ ja $\vec{b} = \overrightarrow{OB}$. Piirrämme vielä origosta alkavan ja



Kuva 17. Kahden avaruusvektorin välinen kulma.

laFig:VecAngle

pisteen A kautta kulkevan puolisuoran s_{OA} ja origosta alkavan ja pisteen B kautta kulkevan puolisuoran s_{OB} . Puolisuorat leikkaavat r -säteisen origo-keskisen pallon pisteissä A' ja B' . Leikkaamme kuvion tasolla, joka kulkee pisteiden O , A' , ja B' kautta.

Kuvassa [laFig:VecAngle](#) on piirretty syntyvä leikkauskuvio. Merkitään b :llä pisteitä A' ja B' yhdistävän lyhyemmän kaaren pituutta. Leikkauskuvaan syntyvä kulma

$$\gamma = \angle A'OB' = \angle AOB = \frac{b}{r}(\text{rad})$$

on vektoreiden välinen kulma. Konstruktiosta seuraa, että $0 \leq \gamma \leq \pi$. Jos $\gamma = 0$, niin vektorit ovat samansuuntaiset. Jos $\gamma = \pi$, niin vektorit ovat vastakkaissuuntaiset.

Monessa sovellustilanteessa vektorin avulla ilmaistaan suoran suunta. Siksi on tapana sanoa, että jos pisteiden O ja A kautta piirretty suora, ja pisteiden O ja B kautta piirretty suora ovat yhdensuuntaiset, niin myös vektorit ovat yhdensuuntaiset. Tämä tarkoittaa sitä, että vektorit ovat yhdensuuntaiset, jos ne ovat saman- tai vastakkaissuuntaiset.

Käytämme jatkossa seuraavia merkintöjä:

$$\begin{aligned} \vec{a} \uparrow \vec{b}, & \quad \text{jos vektorit ovat samansuuntaiset,} \\ \vec{a} \updownarrow \vec{b}, & \quad \text{jos vektorit ovat vastakkaissuuntaiset,} \\ \vec{a} \parallel \vec{b}, & \quad \text{jos vektorit ovat yhdensuuntaiset,} \\ \vec{a} \perp \vec{b}, & \quad \text{jos vektorit ovat kohtisuorassa.} \end{aligned}$$

Määritelmä 19 Olkoon $\vec{a}$ ja $\vec{b}$ kaksi tason (tai avaruuden) vektoria ja olkoon $\gamma = \angle(\vec{a}, \vec{b})$ niiden välinen kulma. Vektoreiden $\vec{a}$ ja $\vec{b}$ pistetulo $\vec{a} \cdot \vec{b}$ on reaalityyppinen luku

$$\vec{a} \cdot \vec{b} = |\vec{a}| |\vec{b}| \cos \gamma.$$

Kaksi vektoria ovat siis kohtisuorassa, jos niiden välinen kulma on $\pi/2$ (90°) tai ainakin toinen vektoreista on nollavektori. Pistetulon arvo voidaan laskea koordinaattien perusteelle seuraavasti:

Lause 20 (1) Olkoon $\vec{a} = a_x \vec{i} + a_y \vec{j}$ ja $\vec{b} = b_x \vec{i} + b_y \vec{j}$ kaksi tasovektoria. Silloin

$$\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y.$$

(2) Olkoon $\vec{a} = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}$ ja $\vec{b} = b_x \vec{i} + b_y \vec{j} + b_z \vec{k}$ kaksi avaruusvektoria. Silloin

$$\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z.$$

Perustelu: (1) Suora seuraus määritelmästä ja lauseesta 11 (sivu 25) ja lauseesta 14 (sivu 30). □

Octavessa on pistetulon laskemiseen funktio `dot(vec1, vec2)`. Periaate on se, että ensin esitellään vektorit, ja sitten lasketaan vektoreilla. Esimerkin 18 pistetulo lasketaan Octavella seuraavalla tavalla:

```
octave:1> a = [2; 3; -4];
octave:2> b = [1; 2; 3];
octave:3> ptulo = dot(a,b)
ptulo = -4
```

Pistetulo on siis helppo laskea koordinaattien avulla, vaikka emme tietäisikään vektorien välistä kulmaa. Kaksi pistetulon ominaisuutta tulee esiin varsin usein:

1. Kaksi vektoria, $\vec{a}$ ja $\vec{b}$ ovat kohtisuorassa jos ja vain, jos niiden pistetulo on nolla.

$$\vec{a} \perp \vec{b} \Leftrightarrow \vec{a} \cdot \vec{b} = 0.$$

2. Vektorin pituuden neliö on sen pistetulo itsensä kanssa.

$$|\vec{a}|^2 = \vec{a} \cdot \vec{a}.$$

Määritelmä 21 Olkoon $\vec{a}$ ja $\vec{b}$ kaksi avaruuden vektoria ja olkoon $\gamma = \angle(\vec{a}, \vec{b})$ niiden välinen kulma. Vektoreiden $\vec{a}$ ja $\vec{b}$ ristitulo $\vec{a} \times \vec{b}$ on vektori, jonka pituus on

$$|\vec{a} \times \vec{b}| = |\vec{a}| |\vec{b}| \sin \gamma,$$

ja suunta määräytyy seuraavasta kolmesta ehdosta

1. vektorit $\vec{a}$ ja $\vec{a} \times \vec{b}$ ovat kohtisuorassa toisiaan vastaan,

2. vektorit $\vec{b}$ ja $\vec{a} \times \vec{b}$ ovat kohtisuorassa toisiaan vastaan,
3. vektorit $\vec{a}$, $\vec{b}$ ja $\vec{a} \times \vec{b}$ (tässä järjestyksessä!) muodostavat positiivisesti suunnistettun systeemin.⁴

DCintro_def1

Määritelmässä siis viitataan vektoreiden pituuksiin ja niiden väliseen kulmaan. Jos voimme valita koordinaatiston sopivasti, niin ristitulo voidaan rakentaa kuin lankkuaita. Aina ei ole järkevää käyttää montaa koordinaatistoa samassa laskussa, ja seuraava konstruktio jääkin siksi erikoisuudeksi, jota vähän käytetään. Sen avulla on kuitenkin helppo osoittaa että ristitulo noudattaa osittelulakia, mikä edelleen avaa meille kaavat, joilla ristitulon lasku voidaan tehdä vektoreiden koordinaattien avulla helposti.

Lause 22 Olkoon $\vec{a}$ ja $\vec{b}$ kaksi avaruuden vektoria ja olkoon $\gamma = \angle(\vec{a}, \vec{b})$ niiden välinen kulma. Asetetaan koordinaatisto siten, että z -akseli on vektorin $\vec{a}$ suuntainen. x - ja y -akselit voidaan valita vapaasti, kuitenkin siten että syntyvä koordinaatisto on suora-kulmainen ja positiivisesti suunnistettu. Olkoon vektoreiden esitykset tässä koordinaatistossa

$$\begin{aligned}\vec{a} &= |\vec{a}| \vec{k}, \\ \vec{b} &= b_x \vec{i} + b_y \vec{j} + b_z \vec{k} = \overrightarrow{OP}.\end{aligned}$$

Ristitulo $\vec{a} \times \vec{b}$ voidaan nyt konstruoida kolmella askelella seuraavasti:

1. Projisoidaan P koordinaatiston (x, y) -tasoon. Projektion piste on $P_{xy} = (b_x, b_y, 0)$ ja suuntajanan $\overrightarrow{OP_{xy}}$ pituus on $|\vec{b}| \sin \gamma$.
2. Kierretään suuntajanaa $\overrightarrow{OP_{xy}}$ kulman $\pi/2$ verran positiiviseen kiertosuuntaan (x, y) -tasossa. Kierretty suuntajana on

$$\overrightarrow{OP_{xy, rot}} = -b_y \vec{i} + b_x \vec{j}.$$

3. Ristituluvektori saadaan nyt kertomalla edellä saatu vektori vektorin $\vec{a}$ pituudella:

$$\vec{a} \times \vec{b} = |\vec{a}| \overrightarrow{OP_{xy, rot}} = -b_y |\vec{a}| \vec{i} + b_x |\vec{a}| \vec{j}.$$

Th:Cross1

⁴Ehdot 1–3 merkitsevät sitä, että ristituluvektori on kohtisuorassa $\vec{a}$:n ja $\vec{b}$:n tasoa vastaan ja osoittaa tämän tason yläpuolelle.

Perustelu: Pisteet O , P ja P_{xy} muodostavat $\vec{a}$:n ja $\vec{b}$:n määräämässä tasossa suorakulmaisen kolmion, jonka hypotenuusan pituus on $|\vec{b}|$. Vaakasuora kateetti on pituudeltaan $|\overrightarrow{OP_{xy}}| = |\vec{b}| \sin \gamma$. Kierretty vektori on yhtä pitkä. Lopullisen ristitulovektorin pituus on $|\vec{a}| |\vec{b}| \sin \gamma$, kuten pitääkin.

Se, että kolmannen askeleen lopussa saatu ristitulovektori on kohtisuorassa vektoreita $\vec{a}$ ja $\vec{b}$ vastaan on konstruktion mukaan selvä. Riittävän hyvällä avaruudellisella hahmotuskyvyllä varustettu lukija voi todistaa tämän. Onneksi voimme tarkistaa kohtisuoruuden myös yksinkertaisella laskulla

$$(\vec{a} \times \vec{b}) \cdot \vec{a} = (-b_y |\vec{a}| \vec{i} + b_x |\vec{a}| \vec{j} + 0 \vec{k}) \cdot (-0 \vec{i} + 0 \vec{j} + |\vec{a}| \vec{k}) = 0 \quad (2.12)$$

$$\longrightarrow \vec{a} \times \vec{b} \perp \vec{a}$$

$$(\vec{a} \times \vec{b}) \cdot \vec{b} = (-b_y |\vec{a}| \vec{i} + b_x |\vec{a}| \vec{j} + 0 \vec{k}) \cdot (b_x \vec{i} + b_y \vec{j} + b_z \vec{k}) = 0 \quad (2.13)$$

$$\longrightarrow \vec{a} \times \vec{b} \perp \vec{b}$$

Kolmikon $\vec{a}$, $\vec{b}$, ja $\vec{a} \times \vec{b}$ positiivinen suunnistus nähdään seuraavasti: aseta peukalo, etusormi ja keskisormi yhdensuuntaisina vektorin $\vec{a}$ suuntaisesti. Pidä peukalo $\vec{a}$:n suuntaisena, mutta käännä etu- ja keskisormi $\vec{b}$:n suuntaiseksi. Pidä peukalo ja etusormi paikallaan, mutta jatka keskisormen kääntämistä samaan suuntaan, kunnes se on xy -tasossa kohtisuorassa peukaloa vastaan. Lopuksi käännä keskisormea $\pi/2$ positiiviseen kiertosuuntaan xy -tasossa. Nyt peukalo on $\vec{a}$:n suuntainen, etusormi on $\vec{b}$:n suuntainen ja keskisormi on ristitulon $\vec{a} \times \vec{b}$ suuntainen. $\square$

Edellisen lauseen konstruktiota emme tässä opetusmonisteessa sovelle käytännössä kuin kerran. Se tapahtuu seuraavan lauseen todistuksessa.

Lause 23 *Olko $\vec{a}$, $\vec{b}$ ja $\vec{c}$ kolme avaruusvektoria ja olko r reaalityyppinen luku. Silloin*

- (1) $\vec{b} \times \vec{a} = -\vec{a} \times \vec{b}$,
- (2) $(r\vec{a}) \times \vec{b} = \vec{a} \times (r\vec{b}) = r(\vec{a} \times \vec{b})$
- (3) $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$
- (4) $(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}$

Th:Cross2

Perustelu: (1) Määritelmän mukaan on selvää, että $|\vec{a} \times \vec{b}| = |\vec{b} \times \vec{a}|$. Kun haemme ristitulo-vektorille $\vec{a} \times \vec{b}$ suunnan oikean käden säännöllä, asetamme peukalon vektorin $\vec{a}$ suuntaiseksi ja etusormen vektorin $\vec{b}$ suuntaiseksi. Nyt keskisormi osoittaa ristitulon suunnan. Kun teemme saman ristitulon $\vec{b} \times \vec{a}$ osalta, niin peukalo ja etusormi vaihtavat paikkoja, jolloin keskisormi kääntyy osoittamaan päinvastaiseen suuntaan kuin edellä.

(2) Seuraa suoraan määritelmästä.

(3) Tarkastellaan vektoreita käyttäen suorakulmaista koordinaatistoa, jonka z -akseli on $\vec{a}$ -vektorin suuntainen. Silloin ristitulot voidaan määrittää kuten lauseessa 22. ^{Th:Cross1} Nyt

$$\begin{aligned}\vec{a} \times (\vec{b} + \vec{c}) &= -(b_y + c_y)|\vec{a}|\vec{i} + (b_x + c_x)|\vec{a}|\vec{j} \\ &= (-b_y|\vec{a}|\vec{i} + b_x|\vec{a}|\vec{j}) + (-c_y|\vec{a}|\vec{i} + c_x|\vec{a}|\vec{j}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}\end{aligned}$$

(4) Periaatteessa voisimme tehdä kuten kohdassa (3), mutta helpommalla pääsemme, kun vetoamme kohtiin (1) ja (3) seuraavalla tavalla

$$\begin{aligned}(\vec{a} + \vec{b}) \times \vec{c} &\stackrel{(1)}{=} -(\vec{c} \times (\vec{a} + \vec{b})) \stackrel{(3)}{=} -(\vec{c} \times \vec{a} + \vec{c} \times \vec{b}) \\ &\stackrel{(1)}{=} \vec{a} \times \vec{c} + \vec{b} \times \vec{c}\end{aligned}$$

□

DCintro_lause1

Lause 24 Vektoreiden $\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k}$ ja $\vec{b} = b_x\vec{i} + b_y\vec{j} + b_z\vec{k}$ ristitulo on

$$\vec{a} \times \vec{b} = (a_y b_z - a_z b_y)\vec{i} + (a_z b_x - a_x b_z)\vec{j} + (a_x b_y - a_y b_x)\vec{k}.$$

Perustelu:

Määritelmästä seuraa, että $\vec{i} \times \vec{i} = \vec{j} \times \vec{j} = \vec{k} \times \vec{k} = \vec{0}$ ja

$$\begin{array}{lll}\vec{i} \times \vec{j} = \vec{k} & \vec{j} \times \vec{k} = \vec{i} & \vec{k} \times \vec{i} = \vec{j} \\ \vec{j} \times \vec{i} = -\vec{k} & \vec{k} \times \vec{j} = -\vec{i} & \vec{i} \times \vec{k} = -\vec{j}\end{array}$$

Nämä ja osittelulait (lauseen 23 ^{Th:Cross2} kohdat (3) ja (4)) antavat seuraavan tuloksen

$$\begin{aligned}\vec{a} \times \vec{b} &= (a_x\vec{i} + a_y\vec{j} + a_z\vec{k}) \times (b_x\vec{i} + b_y\vec{j} + b_z\vec{k}) \\ &= a_x b_x \vec{i} \times \vec{i} + a_x b_y \vec{i} \times \vec{j} + a_x b_z \vec{i} \times \vec{k} \\ &\quad + a_y b_x \vec{j} \times \vec{i} + a_y b_y \vec{j} \times \vec{j} + a_y b_z \vec{j} \times \vec{k} \\ &\quad + a_z b_x \vec{k} \times \vec{i} + a_z b_y \vec{k} \times \vec{j} + a_z b_z \vec{k} \times \vec{k} \\ &= +a_x b_y \vec{k} - a_x b_z \vec{j} \\ &\quad - a_y b_x \vec{k} + a_y b_z \vec{i} \\ &\quad + a_z b_x \vec{j} - a_z b_y \vec{i} \\ &= (a_y b_z - a_z b_y)\vec{i} + (a_z b_x - a_x b_z)\vec{j} + (a_x b_y - a_y b_x)\vec{k}.\end{aligned}$$

□

Esimerkki 25 Laskemme tarkistuksen vuoksi esimerkin 18 ^{Ex:DC1} ristitulon ensin käsin, sitten Octavella.

$$\begin{aligned}\vec{a} &= 2\vec{i} + 3\vec{j} - 4\vec{k}, \text{ ja} \\ \vec{b} &= \vec{i} + 2\vec{j} + 3\vec{k}.\end{aligned}$$

$$\begin{aligned}\Rightarrow \vec{a} \times \vec{b} &= (a_y b_z - a_z b_y)\vec{i} + (a_z b_x - a_x b_z)\vec{j} + (a_x b_y - a_y b_x)\vec{k} \\ &= (3 \cdot 3 - (-4) \cdot 2)\vec{i} + ((-4) \cdot 1 - 2 \cdot 3)\vec{j} + (2 \cdot 2 - 3 \cdot 1)\vec{k} \\ &= 17\vec{i} - 10\vec{j} + \vec{k}\end{aligned}$$

Octavella

```
octave:2> a = [2; 3; -4];
octave:3> b = [1; 2; 3];
octave:4> rtulo = cross(a,b)
rtulo =

    17
   -10
     1
```

Edellä todetun perusteella voimme tehdä seuraavan yhteenvedon yhdensuuntaisuuden tutkimisesta pistetulon ja ristitulon avulla:

Lause 26 Olkoon $\vec{a}$ ja $\vec{b}$ kaksi vektoria ja γ niiden välinen kulma. Silloin

$$\begin{array}{lll}(\gamma = 0) & \text{samansuuntaisuus:} & \vec{a} \uparrow \uparrow \vec{b} \Leftrightarrow \vec{a} \cdot \vec{b} = |\vec{a}| |\vec{b}|, \\(\gamma = \pi) & \text{vastakkaissuuntaisuus:} & \vec{a} \uparrow \downarrow \vec{b} \Leftrightarrow \vec{a} \cdot \vec{b} = -|\vec{a}| |\vec{b}|, \\(\gamma = 0, \text{ tai } \gamma = \pi) & \text{yhdensuuntaisuus:} & \vec{a} \parallel \vec{b} \Leftrightarrow \vec{a} \times \vec{b} = \vec{0}, \\(\gamma = \frac{\pi}{2}) & \text{kohtisuoruus:} & \vec{a} \perp \vec{b} \Leftrightarrow \vec{a} \cdot \vec{b} = 0, \\(\gamma < \frac{\pi}{2}) & \text{terävä kulma:} & \Leftrightarrow \vec{a} \cdot \vec{b} > 0, \\(\gamma > \frac{\pi}{2}) & \text{tylppä kulma:} & \Leftrightarrow \vec{a} \cdot \vec{b} < 0.\end{array}$$

2.5 Skalaari- ja vektoriprojektiot

Sec:OProjIntro

Edellä esimerkissä ^{Exa:Vec1}17 voima jaettiin vaakasuoraan ja sitä vastaan kohtisuorassa olevaan pystysuoraan komponenttiin. Jatkossa ”vaakasuoran” sijassa on jokin muu suunta. Edellisessä kappaleessa esitellyt pistetulo ja ristitulo antavat työkalut tähän komponenttien erotteluun.

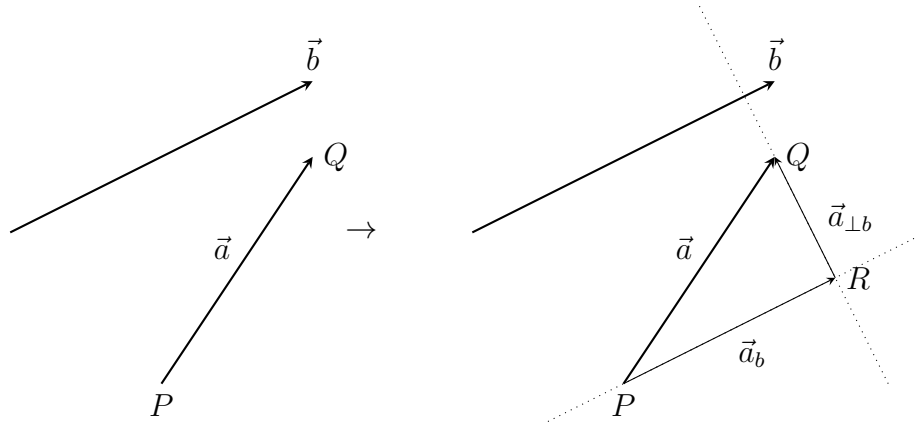
2.5.1 Vektorin projektio vektorin suuntaan

Ssec:VecProj

Tasovektoreiden tapauksessa tämän kappaleen asiat ovat helpommin piirrettävissä, joten aloitamme siitä. Kuvaan 18 on piirretty kaksi tasovektoria $\vec{a} = \overrightarrow{PQ}$ ja $\vec{b}$. Piirretään pisteen P kautta vektorin $\vec{b}$ suuntainen apusuora, ja piirretään pisteen Q kautta apusuora, joka on kohtisuorassa ensin piirrettyä apusuoraa vastaan. Toinen apusuora on myös kohtisuorassa vektoria $\vec{b}$ vastaan. Lisätyt apusuorat leikkaavat toisensa kohtisuorasti pisteessä R . Merkitään

$$\begin{aligned}\overrightarrow{PR} &= \vec{a}_b && \text{”vektorin } b \text{ suuntainen } a\text{:n komponentti”}, \text{ ja} \\ \overrightarrow{RQ} &= \vec{a}_{\perp b} && \text{”vektoria } b \text{ vastaan kohtisuora } a\text{:n komponentti”}.\end{aligned}$$

Nyt $\vec{a} = \vec{a}_b + \vec{a}_{\perp b}$, ja vektori $\vec{a}$ on siis jaettu kahteen osaan, joista ensimmäinen on b :n suuntainen ja toinen on b :tä vastaan kohtisuora.



Kuva 18. Vektorin $\vec{a}$ komponentit: $\vec{b}$:n suuntainen komponentti $\vec{a}_b$, ja $\vec{b}$:tä vastaan kohtisuora komponentti $\vec{a}_{\perp b}$

Fig:VecOComp1

Avaruusvektoreiden tapauksessa edellinen konstruktio vaatii huolellisuutta toisen apusuoran valinnassa. Emme tee konstruktioista avaruusversiota, vaan määrittelemme komponentit helpolla ja yleisellä tavalla:

Määritelmä 27 Olkoon $\vec{a}$ ja $\vec{b}$ kaksi tasovektoria, tai kaksi avaruusvektoria. Jos $\vec{a} = \vec{a}_b + \vec{a}_{\perp b}$, missä $\vec{a}_b$ on vektorin $\vec{b}$ suuntainen ja $\vec{a}_{\perp b}$ on kohtisuorassa vektoria $\vec{b}$ vastaan, niin sanomme, että

$\vec{a}_b$ on vektorin $\vec{b}$ suuntainen vektorin $\vec{a}$ komponentti,

$\vec{a}_{\perp b}$ on vektoria $\vec{b}$ vastaan kohtisuora vektorin $\vec{a}$ komponentti,

Sanomme myös, että $\vec{a}_b$ on vektorin $\vec{a}$ vektoriprojektio vektorin $\vec{b}$ suuntaan.

Vektorin $\vec{a}$ skalaariprojektio vektorin $\vec{b}$ suuntaan, on

$$a_b = \begin{cases} +|\vec{a}_b|, & \text{kun } \vec{a}_b \uparrow\uparrow \vec{b} \\ -|\vec{a}_b|, & \text{kun } \vec{a}_b \uparrow\downarrow \vec{b} \end{cases}.$$

Vektoriprojektio on siis vektori ja skalaari projektio on luku. Projektioille on olemassa helpot kaavat pistetulon avulla.

Lause 28 Olkoon $\vec{b}_0$ vektorin $\vec{b}$ suuntainen yksikkövektori (eli $\vec{b}_0 = \frac{1}{|\vec{b}|}\vec{b}$), ja olkoon γ vektoreiden $\vec{a}$ ja $\vec{b}$ välinen kulma. Silloin

$$\begin{aligned} \vec{a}_b &= |\vec{a}| \cos \gamma \vec{b}_0 \\ &= \frac{(\vec{a} \cdot \vec{b})}{|\vec{b}|} \vec{b}_0 = \frac{(\vec{a} \cdot \vec{b})}{|\vec{b}|^2} \vec{b} = \frac{(\vec{a} \cdot \vec{b})}{(\vec{b} \cdot \vec{b})} \vec{b}, \\ a_b &= |\vec{a}| \cos \gamma = \vec{a} \cdot \vec{b}_0 = \frac{(\vec{a} \cdot \vec{b})}{|\vec{b}|} = \frac{(\vec{a} \cdot \vec{b})}{\sqrt{\vec{b} \cdot \vec{b}}}. \end{aligned}$$

Perustelu: On selvää, että lauseen kaavasta saatu komponentti $\vec{a}_b$ on vektorin $\vec{b}$ suuntainen. On vielä osoitettava, että $(\vec{a} - \vec{a}_b) \perp \vec{b}$. Tämä on suora läpilasku:

$$(\vec{a} - \vec{a}_b) \cdot \vec{b} = \left(\vec{a} - \frac{(\vec{a} \cdot \vec{b})}{(\vec{b} \cdot \vec{b})} \vec{b} \right) \cdot \vec{b} = (\vec{a} \cdot \vec{b} - \frac{(\vec{a} \cdot \vec{b})}{(\vec{b} \cdot \vec{b})} \vec{b} \cdot \vec{b}) = 0.$$

□

Jos $\vec{b}$ on yksikkövektori (jonka pituus on siis yksi), niin vektoriprojektion ja skalaariprojektion kaavat saavat hyvin yksinkertaiset muodot

Th:VecProj1

Lause 29 Jos $\vec{e}$ on yksikkövektori, niin

$$\begin{aligned} \vec{a}_e &= (\vec{a} \cdot \vec{e}) \vec{e}, \\ a_e &= \vec{a} \cdot \vec{e} \end{aligned}$$

Perustelu: HT

□

Lauseen ^{Th:VecProj1}29 kaavat ovat päteviä vain, jos $\vec{e}$ on yksikkövektori. Jos laskija käyttää usein kyseisiä kaavoja, niin on inhimillistä alkaa ajatella, että ”näinhän tämä aina lasketaan”. (Muista: Lauseen ^{Th:VecProj1}29 kaavat eivät ole tosia, jos $|\vec{e}| \neq 1$.)

2.5.2 Voiman momentti pisteen suhteen tasossa

Ssec:FMomPl

Esimerkki 30 Tarkastellaan kuvan [fig:vmP01](#) mukaista tilannetta, jossa tanko, jonka pituus on $L = 50\text{ cm}$ on ripustettu yläpäästään siten, että se pääsee kiertymään ripustuspiisteensä ympäri. Tankoon vaikuttaa paino $G = 20\text{ N}$ alaspäin, jonka vaikutuspiste on tangon keskipiste 25.0 cm ripustuspiisteestä. Lisäksi tankoon vaikuttaa tukivoima $\vec{T}$ ripustuspiisteeseen, ja kaksi vaakasuoraa voimaa: $F_1 = 12.0\text{ N}$ vasemmalle (vaikutuspiste $r_1 = 10.0\text{ cm}$ ripustuspiisteestä) ja $F_2 = 6.0\text{ N}$ oikealle (vaikutuspiste $r_2 = 30.0\text{ cm}$ ripustuspiisteestä). Ensimmäinen voima pyrkii kiertämään tankoa negatiiviseen kiertosuuntaan (myötäpäivään), ja toinen voima pyrkii kiertämään tankoa positiiviseen kiertosuuntaan (vastapäivään). Kuvaan ei ole piirretty ripustuspiisteeseen vaikuttavaa tukivoimaa, eikä massakeskipisteeseen vaikuttavaa painoa – ne otetaan mukaan seuraavassa esimerkissä. Kumpaan suuntaan tanko alkaa kiertyä?

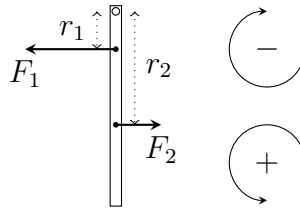


fig:vmP01

Kuva 19. Kilpailevat momentit.

Ratkaisevia ovat nyt voimien momentit (momentti = voima kertaa varsi). Momentit ovat

$$M_1 = -r_1 F_1 = -0.100\text{ m} \cdot 12.0\text{ N} = -1.20\text{ Nm}$$

$$M_2 = +r_2 F_2 = +0.300\text{ m} \cdot 6.0\text{ N} = +1.80\text{ Nm}$$

Momenttien merkit määräytyvät siitä, mihin kiertosuuntaan momentti pyrkii kiertämään kappaletta. Tämä vaatii laskijalta harkintaa. Kokonaismomentti on $M_1 + M_2 = +0.60\text{ Nm}$, joten tanko alkaa kiertyä positiiviseen suuntaan.

Merkit saadaan kätevästi ilman erillistä harkintaakin. Upotetaan taso avaruuden x, y -tasoon, ja korvataan varret vektoreilla kiertopisteestä voiman vaikutuspisteeseen. Momentti lasketaan kaavalla

$$\vec{M} = \vec{r} \times \vec{F}.$$

Nyt

$$\vec{M}_1 = \vec{r}_1 \times \vec{F}_1 = -0.100\text{ m} \vec{j} \times (-12.0\text{ N} \vec{i}) = -1.20\text{ Nm} \vec{k}$$

$$\vec{M}_2 = \vec{r}_2 \times \vec{F}_2 = -0.300\text{ m} \vec{j} \times 6.0\text{ N} \vec{i} = 1.80\text{ Nm} \vec{k}$$

Esimerkki 31 Jatketaan edellisen esimerkin tarkastelua. Tanko kiertyy kulman α verran. Tukivoimalla ei edelleenkään ole momenttia, koska sen varsi on nolla, mutta painolla on nyt momentti

$$M_G = -\frac{L}{2} \sin \alpha G.$$

(Ks. kuva 20) Millä kulman α arvolla kokonaismomentti on nolla?

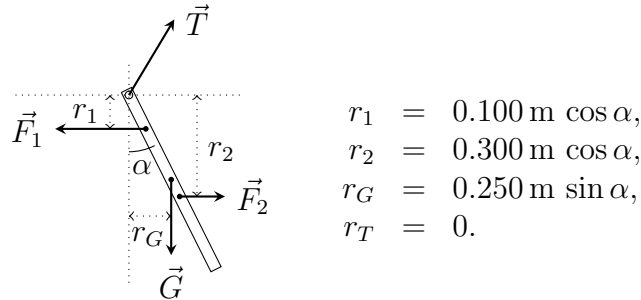


fig:vmP02

Kuva 20. Momenttien tasapaino.

Tasapainon momenttiehto on nyt

$$\begin{aligned}
 -r_1 F_1 + r_2 F_2 - r_G G &= 0 \\
 \Leftrightarrow -0.100 \text{ m} \cos \alpha \cdot 12.0 \text{ N} + 0.300 \text{ m} \cos \alpha \cdot 6.0 \text{ N} - 0.250 \text{ m} \sin \alpha \cdot 20.0 \text{ N} &= 0 \\
 \Leftrightarrow 0.600 \text{ Nm} \cos \alpha - 5.00 \text{ Nm} \sin \alpha &= 0 \\
 \Leftrightarrow \tan \alpha &= 0.120 \\
 \Leftrightarrow \alpha = 0.1194 \text{ (rad)} &= 6.84^\circ. \quad \square
 \end{aligned}$$

Edellä voimat olivat vaaka- tai pystysuoria, mikä aina helpottaa laskemista. Entä jos minkään ei tarvitse olla pystyä tai vaakaa?

Tarkastellaan seuraavaksi kuvan 21 mukaista tilannetta.

Kuvan merkinnöillä on seuraavat tulkinnat:

- $\vec{r}$ = vektori kiertopisteestä voiman vaikutuspisteeseen (varsivektori),
- $r_\perp$ = kiertopisteen etäisyys voiman vaikutussuorasta,
- $\vec{F}$ = voimavektori,
- $\vec{F}_\perp$ = voiman vartta vastaan kohtisuora komponentti,
- α = poikkeama kohtisuoruudesta.

Lause 32 Kuvan 21 merkinnöillä voiman momentti on

$$\begin{aligned}
 M &= (\pm) r F_\perp = (\pm) r_\perp F, \quad \text{ja} \\
 \vec{M} &= \vec{r} \times \vec{F} = M \vec{k}.
 \end{aligned}$$

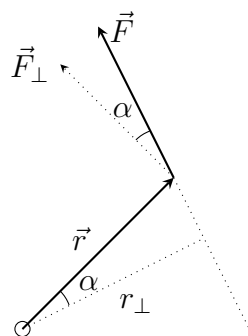


fig:vmP03

Kuva 21. Voiman momentti tasossa.

Perustelu: Tutkimme nyt vain kuvan ^{fig:vmP03}21 tilanteen. (Täydellinen perustelu vaatisi tutkia myös kuviot, joissa momentti on negatiivinen ja $\alpha < 0$)

(1) $M = (\pm)rF_{\perp}$ on momentin perusmääritelmän mukainen lauseke. (Huomaa, että voiman varren suuntaisella komponentilla ei ole momenttia.)

(2)

$$M = (\pm)rF_{\perp} = (\pm)rF \cos \alpha = (\pm)(r \cos \alpha) F = (\pm)r_{\perp} F.$$

(3)

$$|\vec{r} \times \vec{F}| = rF \sin(\pi/2 - \alpha) = rF \cos \alpha = |M|.$$

Merkkien oikeellisuus nähdään oikean käden säännöllä.

□

On syytä korostaa, että edellisen lauseen kaavat koskevat tilannetta, jossa kaikki vektorit ovat samassa tasossa. Kyse on siis taso-ongelman käsittelystä. Erityisesti pyörivien koneiden yhteydessä kappaleen pyöriminen tapahtuu akselin eikä pisteen ympäri. Tämä tuo laskutehtävään yhden uuden vektorin (akselin suunta).

Voiman momenttia akselin suhteen käsitellään myöhemmin luvussa ^{Ssec:FM6SpaceMomSpace}?? (sivu ??).

3 LINEAARISET YHTÄLÖRYHMÄT

Ch:YhtRyh

3.1 Rivioperaatiot

Sec:laRowOper

Sovitaan ensin merkintätavoista. Ratkaisemme ensin yksinkertaisen yhtälöparin

$$\begin{cases} x + 2y = 3 \\ -2x + y = 4 \end{cases} \quad (3.1)$$

Lisäämme ensimmäisen yhtälön kahdella kerrottuna toiseen, jolloin yhtälöryhmä menee muotoon, josta ratkaisu on helppo saada esiin. (Huomaa merkinnät, jotka on tehty yhtälöryhmän oikealle puolelle.)

$$\begin{cases} x + 2y = 3 \\ -2x + y = 4 \end{cases} \quad \begin{array}{l} *2 \\ \leftarrow + \end{array} \quad (3.2)$$

$$\Leftrightarrow \begin{cases} x + 2y = 3 \\ 5y = 10 \end{cases} \quad \begin{array}{l} (1) \\ (2) \end{array} \quad (3.3)$$

Yhtälöstä (2) saadaan arvo $y = 2$ ja sijoittamalla tämä arvo yhtälöön (1) saadaan

$$\begin{aligned} x + 2 \cdot 2 &= 3 \\ x &= -1 \end{aligned}$$

Yhtälöryhmän ratkaisu on siis $x = -1$, $y = 2$.

Edellä käytetty rivioperaatio oli rivin lisääminen luvulla kerrottuna toiseen.

Seuraavassa esimerkissä käytämme samaa rivioperaatiota toistuvasti. Jokaisessa väli-muodossa on yksi termi merkitty hakasuluin. Tätä termiä sanomme *pivot-termiksi*. Rivioperaatioiden kertoimet valitaan niin, että pivot-termin alapuolella olevat samanmuotoiset termit häviävät. Pivot-termin valintaa ja sitä seuraavia rivioperaatioita sanomme *pivotoinniksi*.

Esimerkki 33

$$\left\{ \begin{array}{cccccc} [x] & + & y & - & z & + & 2w & = & 2 \\ -x & & & & + & 3z & - & w & = & 6 \\ & & - & y & + & 2z & + & 3w & = & 2 \\ 2x & + & y & + & 2z & - & w & = & 1 \end{array} \right. \quad \left| \begin{array}{cc} *1 & *2 \\ \leftarrow + & \\ & \leftarrow - \end{array} \right. \quad (3.4)$$

$$\Leftrightarrow \left\{ \begin{array}{cccccc} x & + & y & - & z & + & 2w & = & 2 \\ & & [y] & + & 2z & + & w & = & 8 \\ & & - & y & + & 2z & + & 3w & = & 2 \\ & & - & y & + & 4z & - & 5w & = & -3 \end{array} \right. \quad \left| \begin{array}{cc} *1 & *1 \\ \leftarrow + & \\ & \leftarrow + \end{array} \right. \quad (3.5)$$

$$\Leftrightarrow \left\{ \begin{array}{cccccc} x & + & y & - & z & + & 2w & = & 2 \\ & & y & + & 2z & + & w & = & 8 \\ & & & & [4z] & + & 4w & = & 10 \\ & & & & 6z & - & 4w & = & 5 \end{array} \right. \quad \left| \begin{array}{cc} & *3/2 \\ & \leftarrow - \end{array} \right. \quad (3.6)$$

$$\Leftrightarrow \left\{ \begin{array}{cccccc} x & + & y & - & z & + & 2w & = & 2 \\ & & y & + & 2z & + & w & = & 8 \\ & & & & 4z & + & 4w & = & 10 \\ & & & & & & -10w & = & -10 \end{array} \right. \quad \left| \begin{array}{c} (1) \\ (2) \\ (3) \\ (4) \end{array} \right. \quad (3.7)$$

$$(4) \longrightarrow w = 1$$

$$(3) \longrightarrow 4z + 4 \cdot 1 = 10 \quad \Leftrightarrow z = 1.5$$

$$(2) \longrightarrow y + 2 \cdot 1.5 + 1 = 8 \quad \Leftrightarrow y = 4$$

$$(1) \longrightarrow x + 4 - 1.5 + 2 \cdot 1 = 2 \quad \Leftrightarrow x = -2.5$$

Jatkossa tulemme käyttämään seuraavia rivioperaatioita:

Rivin lisääminen toiseen riviin luvulla kerrottuna. On selvää, että tämä rivioperaatio ei muuta yhtälöryhmän ratkaisua. Esimerkiksi yhtälöryhmässä (3.4) pivot-rivi (rivi 1) lisätään yhdellä kerrottuna riviin 2 ja pivot-rivi vähennetään kahdella kerrottuna rivistä 4. Huomaa, miten pivotoinnissa käytetyt kertoimet nähdään pivot-sarakkeen (x -sarake) vastaavan rivin kertoimista. Varmista edellisen esimerkin avulla, että varmasti ymmärrät rivioperaatiot.

Yhtälöryhmän (3.5) neljäs yhtälö syntyy seuraavasti

termi	$4 \cdot \text{rivi (3.4)}$	$-2 \cdot \text{pivot (3.4)}$	$=$	$4 \cdot \text{rivi (3.5)}$
x -termi	$2x$	$-2 \cdot x$	$=$	
y -termi	y	$-2 \cdot y$	$=$	$-y$
z -termi	$2z$	$-2 \cdot (-z)$	$=$	$4z$
w -termi	$-z$	$-2 \cdot 2w$	$=$	$-5w$
RHS	1	$-2 \cdot 2$	$=$	-3

Nyrkkisääntö: Rivioperaatiota tehtäessä käydään läpi kaikki sarakkeet, myöskin RHS.

Rivin kertominen (tai jakaminen) nolasta eroavalla luvulla.

Kahden rivin vaihtaminen.

On selvää, että rivioperaatiot eivät muuta yhtälöryhmän ratkaisua. Tavoite onkin nyt muuttaa yhtälöryhmän muotoa muuttamatta ratkaisua. Pyrimme siihen, että muoto saadaan sellaiseksi, että viimeisessä yhtälössä on vain yksi muuttuja w , joka saadaan ratkaistua tästä yhtälöstä. Kun toiseksi viimeinen yhtälö viedään muotoon, jossa esiintyy edellisen muuttujan lisäksi vain yksi muu muuttuja (tässä w :n lisäksi z), niin sijoittamalla w :n arvo yhtälöön saadaan uusi muuttuja z ratkaistua.

Useimmissa tapauksissa tavoittelemamme muoto saadaan helposti suorittamalla peräkkäiset pivotoinnit seuraavasti. Ensimmäiseksi pivot-termiksi valitaan ensimmäisen yhtälön x -termi. Toiseen pivotointiin valitsemme pivot-termiksi toisen yhtälön y -termin. Kolmanteen pivotointiin valitsemme pivot-termiksi kolmannen yhtälön z -termin. Jne. Normaali ratkaisu onkin siis valita seuraavaksi pivot-termiksi ”seuraava muuttuja – seuraava rivi”. Näin jatkaen yhtälöryhmä saa tavanomaiselle ratkaisulle tyypillisen kolmiomaisen muodon.

Joskus ”seuraava muuttuja – seuraava rivi” -periaatteen mukaisessa paikassa on tyhjää; haluttu termi puuttuu. Silloin on etsittävä haluttua muuttujaa sisältävää yhtälöä alemmilta riveiltä. Jos tällainen rivi löytyy, vaihdetaan se pivot-rivin paikalle ja jatketaan normaalisti. Jos muuttujaa ei esiinny millään jäljellä olevalla rivillä, niin siirrymme seuraavaan muuttujaan (sarakkeeseen). Jos jouduimme hyppäämään pivotoinneissa yhden sarakkeen yli, on tämä merkille pantava asia. Pivotoimaton sarake kannattaa merkitä käsin laskettaessa, jotta varmasti muistaa jatkossa huomioda sen oikein. Tämän kapaleen esimerkeissä pivot-termit on merkitty hakasuluin. Pivotoinneissa yli-hypätty sarake erottuu siitä, ettei sillä ole viimeisessä esityksessä pivot-merkintää.

Seuraavat esimerkit valaisevat asiaa.

Esimerkki 34

$$\left\{ \begin{array}{cccccc} [x] & + & y & + & z & + & 2w & = & 1 \\ -3x & - & 3y & & & - & w & = & 1 \\ & & + & y & - & 2z & + & 3w & = & -1 \\ 2x & + & 3y & + & 3z & + & 10w & = & 1 \end{array} \right. \left| \begin{array}{cc} *3 & *2 \\ \leftarrow + & \\ & \leftarrow - \end{array} \right. \quad (3.8)$$

$$\Leftrightarrow \left\{ \begin{array}{cccccc} (x) & + & y & + & z & + & 2w & = & 1 \\ & & & & 3z & + & 5w & = & 4 \\ & & y & - & 2z & + & 3w & = & -1 \\ & & y & + & z & + & 6w & = & -1 \end{array} \right. \left| \begin{array}{c} \leftarrow \\ \leftarrow \end{array} \right. \text{vaihdetaan} \quad (3.9)$$

Uudeksi pivot-termiksi ei nyt voida ottaa toisen yhtälön y-termiä. Koska haluamme edetä systemaattisesti, vaihdamme toisen ja kolmannen rivin keskenään ja jatkamme sen jälkeen systemaattisesti.

$$\Leftrightarrow \left\{ \begin{array}{cccccc} (x) & + & y & + & z & + & 2w & = & 1 \\ & & [y] & - & 2z & + & 3w & = & -1 \\ & & & & 3z & + & 5w & = & 4 \\ & & y & + & z & + & 6w & = & -1 \end{array} \right. \left| \begin{array}{c} *1 \\ \leftarrow - \end{array} \right. \quad (3.10)$$

$$\Leftrightarrow \left\{ \begin{array}{cccccc} (x) & + & y & + & z & + & 2w & = & 1 \\ & & (y) & - & 2z & + & 3w & = & -1 \\ & & & & [3z] & + & 5w & = & 4 \\ & & & & 3z & + & 3w & = & 0 \end{array} \right. \left| \begin{array}{c} *1 \\ \leftarrow - \end{array} \right. \quad (3.11)$$

$$\Leftrightarrow \left\{ \begin{array}{cccccc} (x) & + & y & + & z & + & 2w & = & 1 \\ & & (y) & - & 2z & + & 3w & = & -1 \\ & & & & (3z) & + & 5w & = & 4 \\ & & & & [- & 2w] & = & -4 \end{array} \right. \left| \begin{array}{c} (1) \\ (2) \\ (3) \\ (4) \end{array} \right. \quad (3.12)$$

$$(4) \rightarrow w = (-4)/(-2) = 2 \quad (3.13)$$

$$(3) \rightarrow z = (4 - 5 \cdot 2)/3 = -2 \quad (3.14)$$

$$(2) \rightarrow y = -1 + 2 \cdot (-2) - 3 \cdot 2 = -11 \quad (3.15)$$

$$(1) \rightarrow x = 1 - (-11) - (-2) - 2 \cdot 2 = 10 \quad (3.16)$$

Saatu ratkaisu kannattaa aina tarkistaa sijoittamalla saadut arvot **alkuperäiseen** yhtälöryhmään:

$$\left\{ \begin{array}{cccccc} 10 & + & (-11) & + & (-2) & + & 2 \cdot 2 & = & 1 \\ -3 \cdot 10 & - & 3(-11) & & & - & 2 & = & 1 \\ & & + & (-11) & - & 2(-2) & + & 3 \cdot 2 & = & -1 \\ 2 \cdot 10 & + & 3(-11) & + & 3(-2) & + & 10 \cdot 2 & = & 1 \end{array} \right. \left| \begin{array}{c} ok \\ ok \\ ok \\ ok \end{array} \right. \quad (3.17)$$

Joskus käy niin, että jostakin yhtälöstä häviävät kaikki muuttujat. Jos syntyvä yhtälö on tyyppiä $0 = 0$, niin se on (triviaalisti) tosi. Vaihdetaan tämä yhtälö ryhmän viimeiseksi

ja jatketaan normaalisti.

Jos yhtälö on tyyppiä $0 = a$ ($a \neq 0$), niin yhtälö on epätosi ja yhtälöryhmällä ei voi olla yhtään ratkaisua, joten ratkaisun etsiminen voidaan lopettaa.

Jos on syntynyt epätosi yhtälö, tai kaikki pivotoimattomat yhtälöt ovat triviaaleja, päättelemme seuraavasti:

(1) Jos on syntynyt epätosi yhtälö,

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \quad \quad \quad a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad 0 = 5 \quad \text{epätosi} \rightarrow \text{ei ratkaisua} \\ \quad \quad \quad \star + \star + \dots + \star = \star \\ \quad \quad \quad \star + \star + \dots + \star = \star \end{array} \right.$$

niin lopetamme ratkaisun etsimisen ja ilmoitamme, että **ratkaisujoukko on tyhjä**.

(2) Jos ei-triviaaleja yhtälöitä on yhtä monta kuin muuttujaa,

$$\left\{ \begin{array}{l} (a_{11}x_1) + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \quad \quad (a_{22}x_2) + \dots + a_{2n}x_n = b_2 \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \ddots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad (a_{nn}x_n) = b_n \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad 0 = 0 \end{array} \right.$$

niin ratkaisemme muuttujien arvot palautuvien sijoitusten avulla.

(3) Jos ei-triviaaleja yhtälöitä on m kappaletta ja muuttujia on n kpl ja $m < n$, niin siirrämme RHS:iin muuttujat, joita ei esiinny pivot-termeissä. Pivotoinnissa mukana olleet muuttujat ratkaistaan peräkkäisillä sijoituksilla.

$$\left\{ \begin{array}{l} (a_{11}x_1) + a_{12}x_2 + \dots + a_{1m}x_m = a_{1(m+1)}x_{m+1} \dots a_{1n}x_n + b_1 \\ \quad \quad (a_{22}x_2) + \dots + a_{2m}x_m = a_{2(m+1)}x_{m+1} \dots a_{2n}x_n + b_2 \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \ddots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad (a_{mm}x_m) = a_{m(m+1)}x_{m+1} \dots a_{mn}x_n + b_m \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad 0 = 0 \end{array} \right.$$

Ratkaisuja on nyt ääretön määrä. On kuitenkin oltava huolellinen vastausta kirjoittaessa. Yhtälöryhmä ei ole aina tosi. Periaate on se, että muuttujille $x_{m+1}, \dots, x_n$ voidaan antaa mitkä tahansa numeroarvot ja tämän jälkeen muuttujien $x_1, \dots, x_m$ arvot saadaan

palautuvilla sijoituksella viimeisestä yhtälöryhmän muodosta.

Seuraavissa esimerkeissä tulevat vastaan kaikki edellä esitetyt tapaukset. Esimerkeissä yksinkertaistetaan välimuotoja siirtymällä kerroin-kaavio -esitykseen. Tämä merkitsee sitä, että jätämme muuttujat kirjoittamatta. Kertoimen paikka kaaviossa kertoo minkä muuttujan kerroin se on (sarake) ja mihin yhtälöön se kuuluu (rivi).

Esimerkki 35

$$\begin{cases} [x] - y + z = 5 \\ -3x + 3y - 3z = 3 \\ \quad 2y - 2z = -1 \\ 2x + 3y + 3z = 2 \end{cases} \quad (3.18)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} [1] & -1 & 1 & 5 \\ -3 & 3 & -3 & 3 \\ 0 & 2 & -2 & -1 \\ 2 & 3 & 3 & 2 \end{array} \right) \begin{array}{l} *3 \quad *2 \\ \leftarrow + \\ \\ \leftarrow - \end{array} \quad (3.19)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & -1 & 1 & 5 \\ 0 & 0 & 0 & 18 \\ 0 & 2 & -2 & -1 \\ 0 & 5 & 1 & -8 \end{array} \right) \begin{array}{l} \rightarrow \text{epätosi, ei ratkaisua} \\ \\ \\ \end{array} \quad (3.20)$$

Vastaus: ratkaisujoukko on tyhjä.

Esimerkki 36

$$\begin{cases} [x] - y + z = -1 \\ -3x + 3y - 3z = 3 \\ \quad 2y - 2z = 0 \\ 2x + 3y + 3z = 4 \end{cases} \quad (3.21)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} [1] & -1 & 1 & -1 \\ -3 & 3 & -3 & 3 \\ 0 & 2 & -2 & 0 \\ 2 & 3 & 3 & 4 \end{array} \right) \begin{array}{l} *3 \quad *2 \\ \leftarrow + \\ \\ \leftarrow - \end{array} \quad (3.22)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & 5 & 1 & 6 \end{array} \right) \begin{array}{l} \leftarrow | \text{vaihdetaan} \\ \\ \leftarrow | \end{array} \quad (3.23)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & -1 & 1 & -1 \\ 0 & 5 & 1 & 6 \\ 0 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} \leftarrow - \\ *2 \end{array} \quad (3.24)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & -1 & 1 & -1 \\ 0 & [1] & 5 & 6 \\ 0 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} *2 \\ \leftarrow - \end{array} \quad (3.25)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & -1 & 1 & -1 \\ 0 & (1) & 5 & 6 \\ 0 & 0 & (-12) & -12 \\ 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} (1) \\ (2) \\ (3) \\ (4) \end{array} \quad (3.26)$$

$$(3) \rightarrow z = 1 \quad (3.27)$$

$$(2) \rightarrow y = 6 - 5 \cdot 1 = 1 \quad (3.28)$$

$$(1) \rightarrow x = -1 + 1 - 1 = -1 \quad (3.29)$$

Vastaus: $x = -1$, $y = 1$ ja $z = 1$.

Esimerkki 37

$$\begin{cases} x + y + z + 2w = 1 \\ -3x - 3y - z = 1 \\ + y - 2z + 3w = -1 \end{cases} \quad (3.30)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} [1] & 1 & 1 & 2 & 1 \\ -3 & -3 & -1 & 0 & 1 \\ 0 & 1 & -2 & 3 & -1 \end{array} \right) \begin{array}{l} *3 \\ \leftarrow + \end{array} \quad (3.31)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} (1) & 1 & 1 & 2 & 1 \\ 0 & 0 & 2 & 6 & 4 \\ 0 & 1 & -2 & 3 & -1 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \text{vaihdetaan} \quad (3.32)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} (1) & 1 & 1 & 2 & 1 \\ 0 & (1) & -2 & 3 & -1 \\ 0 & 0 & [2] & 6 & 4 \end{array} \right) \quad (3.33)$$

$$\Leftrightarrow \begin{cases} x + y + z = 1 - 2w \\ y - 2z = -1 - 3w \\ 2z = 4 - 6w \end{cases} \begin{array}{l} (1) \\ (2) \\ (3) \end{array} \quad (3.34)$$

$$(3) \rightarrow z = (4 - 6w)/2 = 2 - 3w \quad (3.35)$$

$$(2) \rightarrow y = (-1 - 3w + 2(2 - 3w)) = 3 - 9w \quad (3.36)$$

$$(1) \rightarrow x = (1 - 2w - (3 - 9w) - (2 - 3w)) = -4 + 10w \quad (3.37)$$

Vastaus

$$\begin{cases} x = -4 + 10w \\ y = 3 - 9w \\ z = 2 - 3w \\ w = w \end{cases}$$

Esimerkki 38

$$\begin{cases} x - 2y - z + 2w = 1 \\ -3x + 3y - 5w = 1 \\ y + z = -1 \\ 2x + y + 3z + 6w = 5 \end{cases} \quad (3.38)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} [1] & -2 & -1 & 2 & 1 \\ -3 & 3 & 0 & -5 & 1 \\ 0 & 1 & 1 & 0 & -1 \\ 2 & 1 & 3 & 6 & 5 \end{array} \right) \quad \left| \begin{array}{l} *3 \\ \leftarrow + \\ \leftarrow - \end{array} \right. \quad \begin{array}{l} *2 \\ \\ \end{array} \quad (3.39)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} (1) & -2 & -1 & 2 & 1 \\ 0 & -3 & -3 & 1 & 4 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 5 & 5 & 2 & 3 \end{array} \right) \quad \left| \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \right. \quad \begin{array}{l} \text{vaihdetään} \\ \end{array} \quad (3.40)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} (1) & -2 & -1 & 2 & 1 \\ 0 & [1] & 1 & 0 & -1 \\ 0 & -3 & -3 & 1 & 4 \\ 0 & 5 & 5 & 2 & 3 \end{array} \right) \quad \left| \begin{array}{l} *3 \\ \leftarrow + \\ \leftarrow - \end{array} \right. \quad \begin{array}{l} *5 \\ \\ \end{array} \quad (3.41)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} (1) & -2 & -1 & 2 & 1 \\ 0 & (1) & 1 & 0 & -1 \\ 0 & 0 & 0 & [1] & 4 \\ 0 & 0 & 0 & 2 & 8 \end{array} \right) \quad \left| \begin{array}{l} *2 \\ \leftarrow - \end{array} \right. \quad (3.42)$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} (1) & -2 & -1 & 2 & 1 \\ 0 & (1) & 1 & 0 & -1 \\ 0 & 0 & 0 & (1) & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad (3.43)$$

$$\Leftrightarrow \begin{cases} (x) - 2y + 2w = 1 + z \\ (y) = -1 - z \\ (w) = 4 \\ 0 = 0 \end{cases} \quad \left| \begin{array}{l} (1) \\ (2) \\ (3) \end{array} \right. \quad (3.44)$$

$$(3) \rightarrow w = 4 \quad (3.45)$$

$$(2) \rightarrow y = -1 - z \quad (3.46)$$

$$(1) \rightarrow x = 1 + z + 2(-1 - z) - 2 \cdot 4 = -9 - z \quad (3.47)$$

Vastaus

$$\begin{cases} x = -9 - z \\ y = -1 - z \\ z = z \\ w = 4 \end{cases}$$

3.2 Lineaarikombinaatiot

Sec:laLinComb

Yhtälöryhmää ratkaistaessa laskimme kerroinkaavion riveillä. Jatkossa vastaavia operaatioita tehdään niin paljon, että asiaan kannattaa sopia merkinnät ja laskusäännöt. Sovimme, että n :stä reaalityluvusta muodostuva yksirivinen kaavio

$$\vec{a} = (a_1 \ a_2 \ a_3 \ \dots \ a_n) \quad (3.48)$$

on *rivivektori*, ja vastaava yhdestä sarakkeesta muodostuva kaavio

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (3.49)$$

on *sarakevektori*. Lukuja a_k sanotaan vektorin $\vec{a}$ *koordinaateiksi*. Sarakevektorille sovitetaan myös merkintätapa

$$\vec{b} = (b_1 \ b_2 \ \dots \ b_n)^T \quad (3.50)$$

missä yläindeksi T merkitsee *transponointia*, missä rivit muutetaan sarakkeiksi ja sarakkeet riveiksi.

Määritelmä 39 Olkoon $\vec{a} = (a_1 \ a_2 \ \dots \ a_n)$ ja $\vec{b} = (b_1 \ b_2 \ \dots \ b_n)$ kaksi n -alkioista rivivektoria ja $r \in \mathbb{R}$ reaalityluku. Määritellään reaalityluvulla kertominen ja vektoreiden yhteen- ja vähennyslasku seuraavasti.

$$\begin{aligned} r \cdot \vec{a} &= r \cdot (a_1 \ a_2 \ a_3 \ \dots \ a_n) \\ &= (r \cdot a_1 \ r \cdot a_2 \ r \cdot a_3 \ \dots \ r \cdot a_n) \\ \vec{a} + \vec{b} &= (a_1 \ a_2 \ a_3 \ \dots \ a_n) + (b_1 \ b_2 \ b_3 \ \dots \ b_n) \\ &= (a_1 + b_1 \ a_2 + b_2 \ a_3 + b_3 \ \dots \ a_n + b_n) \\ \vec{a} - \vec{b} &= (a_1 \ a_2 \ a_3 \ \dots \ a_n) - (b_1 \ b_2 \ b_3 \ \dots \ b_n) \\ &= (a_1 - b_1 \ a_2 - b_2 \ a_3 - b_3 \ \dots \ a_n - b_n) \end{aligned}$$

Sarake vektoreiden kertominen reaalityluvulla ja sarakevektoreiden summa määritellään vastaavalla tavalla.

Esimerkki 40

$$\begin{aligned} 5 \cdot (1 \ -2 \ 0) &= (5 \ -10 \ 0) \\ (1 \ -2 \ 0) + (2 \ 4 \ -1) &= (3 \ 2 \ -1) \end{aligned}$$

Määritelmä 41 Vektorit $\vec{a} = (a_1 \ a_2 \ \dots \ a_n)$ ja $\vec{b} = (b_1 \ b_2 \ \dots \ b_n)$ ovat samat (identtiset), jos niillä niiden koordinaatit ovat samat

$$a_k = b_k, \text{ kaikilla } k = 1, \dots, n$$

Määritelmä 42 Vektoria, jonka kaikki koordinaatit ovat nolliä sanomme nollavektoriaksi ja merkitsemme

$$(0 \ 0 \ 0 \ \dots \ 0) = \vec{0}$$

Näillä merkinnöillä voimme nyt tehdä melko tutulta tuntuvaa laskentaa. Vektoreita sisältävien lausekkeiden käsittely muistuttaa hyvin paljon tuttua kirjainlaskentaa. Jos esimerkiksi olemme ratkaisemassa vektoria $\vec{x}$ yhtälöstä $\vec{a} + 3\vec{x} = \vec{x} + \vec{b} - \vec{a}$, niin voimme siirtää termejä ja yhdistellä samaa muotoa olevia termejä seuraavasti

$$\begin{aligned} \vec{a} + 3\vec{x} &= \vec{x} + \vec{b} - \vec{a} \\ \Leftrightarrow 3\vec{x} - \vec{x} &= \vec{b} - \vec{a} - \vec{a} \\ \Leftrightarrow 2\vec{x} &= \vec{b} - 2\vec{a} \\ \Leftrightarrow \vec{x} &= \frac{1}{2}\vec{b} - \vec{a} \end{aligned}$$

Määritelmä 43 Olkoon $\vec{a}$ ja $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ saman mittaisia vektoreita, ja olkoon $c_1, c_2, \dots, c_n$ reaali-lukuja. Jos

$$\vec{a} = c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_n\vec{u}_n$$

niin sanomme, että vektori $\vec{a}$ on edellä lausuttu u -vektoreiden lineaarikombinaationa. Lukuja $c_1, c_2, \dots, c_n$ sanomme lineaarikombinaation kertoimiksi. Jos ainakin yksi lineaarikombinaation kertoimista on nolasta eroava, sanomme lineaarikombinaation olevan ei-triviaalin.

Lause 44 Jos nollavektori $\vec{0}$ voidaan lausua ei-triviaalina lineaarikombinaationa vektoreista $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$, niin jokin u -vektoreista voidaan lausua muiden u -vektoreiden lineaarikombinaationa.

lause12

Todistus: Olkoon

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_n\vec{u}_n = \vec{0}$$

Koska lineaarikombinaatio on ei-triviaali, löytyy kertoimien joukosta nolasta eroavia kertoimia. Olkoon c_k indeksiltään pienin nolasta eroava kerroin. Silloin

$$\vec{u}_k = -\frac{c_{k+1}}{c_k}\vec{u}_{k+1} - \frac{c_{k+2}}{c_k}\vec{u}_{k+2} - \dots - \frac{c_n}{c_k}\vec{u}_n$$

Algoritmi muokkaa kaaviota kierros kierrokselta ja merkitsemme tarvittaessa kaavion järjestysnumeron $k = 1, 2, \dots$ yläindeksinä. (a_{ij}^k on siis muuttujan x_j kerroin k :nnen

kaavion i :nnessä yhtälössä.)

Koska lopullinen algoritmi sisältää lukuisia poikkeustapausten aiheuttamia varmistuksia ja on siksi kummallisen tuntuinen ensi lukemalta, esitämme ensin erikoistapauksen, jolla saa tuntuman perusratkaisuun.

3.3.1 Algoritmi, Gaussin eliminointi, helppo tapaus.

Seuraava algoritmi olettaa, että muuttujia ja yhtälöitä on sama määrä ($m = n$), ja minään yhtälön LHS:ia ei voida lausua muiden yhtälöiden LHS:ien lineaarikombinaationa. Oletus merkitsee, että yhtälöryhmällä on yksikäsitteinen ratkaisu ja monet yleisen algoritmin tarkistukset voidaan jättää pois.

GaussJordanI

Gaussin algoritmi, versio I helpossa tapauksessa.

- (1) Aseta ensimmäiseen kaavioon tutkittavan yhtälöryhmän kertoimet;
- (2) for $k = 1$ to $n - 1$
- (3) % (Etsitään pivot-alkio)
- (4) if $a_{kk}^k = 0$
- (5) then
- (6) etsi samasta sarakkeesta alemmalta riviltä alkio
- (7) $a_{tk}^k \neq 0, t > k$ ja vaihda rivit w_t ja w_k .
- (8) endif
- (9) % (Tehdään uusi kaavio, pivotoidaan)
- (10) for $j = 1$ to k
- (11) $w_j^{k+1} = w_j^k$;
- (12) endfor
- (13) for $j = k + 1$ to m
- (14) $w_j^{k+1} = w_j^k - (a_{jk}^k / a_{kk}^k) \cdot w_k^k$;
- (15) endfor
- (16) endfor
- (17) (% Palautuvat sijoitukset)
- (18) for $p = n$ downto 1
- (19) $x_p = (b_p^k - \sum_{j=p+1}^n a_{pj}^k x_j) / (a_{pp}^k)$
- (20) endfor

GaussJordanII

Gaussin algoritmi, versio II helpossa tapauksessa.

```

(1) Aseta ensimmäiseen kaavioon tutkittavan yhtälöryhmän kertoimet;
(2) for  $k = 1$  to  $n - 1$ 
(3)   % (Etsitään pivot-alkio)
(4)   if  $a_{kk}^k = 0$ 
(5)   then
(6)     etsi samasta sarakkeesta alemmalta riviltä alkio
(7)      $a_{tk}^k \neq 0, t > k$  ja vaihda rivit  $w_t$  ja  $w_k$ .
(8)   endif
(9)   % (Tehdään uusi kaavio, pivotoidaan)
(10)  for  $j = 1$  to  $k - 1$ 
(11)     $w_j^{k+1} = w_j^k - (a_{jk}^k / a_{kk}^k) \cdot w_k^k$ ;
(12)  endfor
(13)   $w_k^{k+1} = w_k^k$ ;
(14)  for  $j = k + 1$  to  $m$ 
(15)     $w_j^{k+1} = w_j^k - (a_{jk}^k / a_{kk}^k) \cdot w_k^k$ ;
(16)  endfor
(17) endfor
(18) (% Muuttujien arvot)
(19) for  $p = 1$  to  $n$ 
(20)    $x_p = b_p^k / a_{pp}^k$ 
(21) endfor

```

Versio II:ssa ei tarvita palautuvia sijoituksia. Sen vuoksi moni pitää toista versiota parempana. Kuitenkin se on herkempi pyöristysvirheille ja versiota I pidetään siksi teoreettisesti parempana.

Esimerkki 45 Olkoon ratkaistavana yhtälöryhmä

$$\begin{cases} x - 2y + 3z = -3 \\ x - 2y + z = -5 \\ 2x + y - 2z = 1 \end{cases} \quad (3.54)$$

Ratkaistaan yhtälöryhmä kummallakin edellä esitetyllä algoritmilla.

Versio I:

$$\left(\begin{array}{ccc|c} [1] & -2 & 3 & -3 \\ 1 & -2 & 1 & -5 \\ 2 & 1 & -2 & 1 \end{array} \right) \quad \left| \begin{array}{cc} *1 & *2 \\ \leftarrow - & \\ & \leftarrow - \end{array} \right. \quad (3.55)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} [1] & -2 & 3 & -3 \\ 0 & 0 & -2 & -2 \\ 0 & 5 & -8 & 7 \end{array} \right) \quad \left| \begin{array}{c} \leftarrow \\ \leftarrow \end{array} \right| \text{vaihdetaan} \quad (3.56)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & -2 & 3 & -3 \\ 0 & (5) & -8 & 7 \\ 0 & 0 & (-2) & -2 \end{array} \right) \quad (3.57)$$

$$z = (-2)/(-2) = 1 \quad (3.58)$$

$$y = (7 - (-8) \cdot 1)/5 = 3 \quad (3.59)$$

$$x = (-3 - (-2) \cdot 3 - 3 \cdot 1)/1 = 0 \quad (3.60)$$

Versio II:

$$\left(\begin{array}{ccc|c} [1] & -2 & 3 & -3 \\ 1 & -2 & 1 & -5 \\ 2 & 1 & -2 & 1 \end{array} \right) \quad \left| \begin{array}{c} *1 \\ \leftarrow - \\ \leftarrow - \end{array} \right. \quad \begin{array}{c} *2 \\ \\ \leftarrow - \end{array} \quad (3.61)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} [1] & -2 & 3 & -3 \\ 0 & 0 & -2 & -2 \\ 0 & 5 & -8 & 7 \end{array} \right) \quad \left| \begin{array}{c} \leftarrow \\ \leftarrow \end{array} \right. \quad \begin{array}{c} \text{vaihdetaan} \\ \end{array} \quad (3.62)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & -2 & 3 & -3 \\ 0 & [5] & -8 & 7 \\ 0 & 0 & -2 & -2 \end{array} \right) \quad \left| \begin{array}{c} \leftarrow + \\ *0.4 \end{array} \right. \quad (3.63)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & 0 & -0.2 & -0.2 \\ 0 & (5) & -8 & 7 \\ 0 & 0 & [-2] & -2 \end{array} \right) \quad \left| \begin{array}{c} \leftarrow \\ -0.1 \end{array} \right. \quad \begin{array}{c} \leftarrow \\ -4 \end{array} \quad (3.64)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (1) & 0 & 0 & 0 \\ 0 & (5) & 0 & 15 \\ 0 & 0 & (-2) & -2 \end{array} \right) \quad (3.65)$$

$$x = 0/1 = 0 \quad (3.66)$$

$$y = 15/5 = 3 \quad (3.67)$$

$$z = (-2)/(-2) = 1 \quad (3.68)$$

Edellä esitetyissä algoritmeissa on vielä yksi merkittävä heikkous. Jos a_{kk} on pieni, mutta nolasta eroava (esim. $a_{kk} = 1.23 \cdot 10^{-10}$ niin version I rivillä (4) ehto on tarkkaan ottaen epätosi, vaikka useimmiten pivot-alkio pitäisi hakea alemmalta riviltä.

Selkein ja helposti ohjelmoitava ratkaisu on hakea itseisarvoltaan suurin kerroin joukosta $\{a_{kk}, a_{(k+1)k}, \dots, a_{nk}\}$. Jos itseisarvoltaan suurin kerroin on a_{qk} rivillä q , niin vaihdetaan rivit w_k ja w_q . Käsin laskettaessa tällainen rivinvaihto on usein tarpeetonta, mutta ohjelmoitaessa algoritmi tietokoneelle, tämä rivinvaihto on ehdottomasti tehtävä versio

III:n mukaisesti.

GaussJordanIII

Gaussin algoritmi, versio III helpossa tapauksessa.

Muuttuneet rivit (verrattuna versioon I) on merkitty tähdellä.

```

(1) Aseta ensimmäiseen kaavioon tutkittavan yhtälöryhmän kertoimet;
(2) for  $k = 1$  to  $n - 1$ 
(3*)   % (Etsitään pivot-rivi)
(4*)    $q = k$ ;
(5*)   for  $i = k + 1$  to  $n$ 
(6*)     if  $|a_{ik}| > |a_{qk}|$ 
(7*)     then
(8*)        $q = i$ ;
(9*)     endif
(10*)  endfor
(11*)  vaihda rivit  $w_k$  ja  $w_q$ 
(12)   % (Tehdään uusi kaavio, pivotoidaan)
(13)   for  $j = 1$  to  $k$ 
(14)      $w_j^{k+1} = w_j^k$ ;
(15)   endfor
(16)   for  $j = k + 1$  to  $m$ 
(17)      $w_j^{k+1} = w_j^k - (a_{jk}^k / a_{kk}^k) \cdot w_k^k$ ;
(18)   endfor
(19) endfor
(20) (% Palautuvat sijoitukset)
(21) for  $p = n$  downto  $1$ 
(22)    $x_p = (b_p^k - \sum_{j=p+1}^n a_{pj}^k x_j) / (a_{pp}^k)$ 
(23) endfor

```

Esimerkki 46 Olkoon ratkaistavana yhtälöryhmä

$$\begin{cases} x - 2y + 3z = -3 \\ x - 2y + z = -5 \\ 2x + y - 2z = 1 \end{cases} \quad (3.69)$$

Ratkaistaan yhtälöryhmä muutetulla algoritmilla.

Versio III:

$$\left(\begin{array}{ccc|c} 1 & -2 & 3 & -3 \\ 1 & -2 & 1 & -5 \\ 2 & 1 & -2 & 1 \end{array} \right) \begin{array}{l} \leftarrow | \text{vaihdetaan} \\ \leftarrow | \end{array} \quad (3.70)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} [2] & 1 & -2 & 1 \\ 1 & -2 & 3 & -3 \\ 1 & -2 & 1 & -5 \end{array} \right) \begin{array}{l} *0.5 \quad *0.5 \\ \leftarrow - \\ \leftarrow - \end{array} \quad (3.71)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (2) & 1 & -2 & 1 \\ 0 & [-2.5] & 4 & -3.5 \\ 0 & -2.5 & 2 & -5.5 \end{array} \right) \begin{array}{l} *1 \\ \leftarrow - \end{array} \quad (3.72)$$

$$\Leftrightarrow \left(\begin{array}{ccc|c} (2) & 1 & -2 & 1 \\ 0 & (-2.5) & 4 & -3.5 \\ 0 & 0 & -2 & -2 \end{array} \right) \quad (3.73)$$

$$z = (-2)/(-2) = 1 \quad (3.74)$$

$$y = (-3.5 - 4 \cdot 1)/(-2.5) = 3 \quad (3.75)$$

$$x = (1 - 1 \cdot 3 + 2 \cdot 1)/2 = 0 \quad (3.76)$$

3.3.2 Algoritmi, Gaussin eliminointi, yleinen tapaus.

Seuraavaksi muotoilemme yleisemmän algoritmin, joka toimii kaikissa tilanteissa. Emme kirjoita täydellistä ohjelmaa, vaan tarkoitus on saada luettava esitys, josta ohjelmointitaitoinen lukija pystyy kirjoittamaan toimivan koodin.

Oletamme, että yhtälöitä on m kappaletta ja muuttujia on n kappaletta. Kun kirjoitamme algoritmin, varaudumme erityisesti seuraaviin erikoistapahtumiin.

- Pivointi epäonnistuu jossakin sarakkeessa.
- Yhtälöt loppuvat ennen kuin kaikki sarakkeet on pivotoitu.
- Syntyy epätosi yhtälö.
- Syntyy triviaali yhtälö $0 = 0$.

Tulemme algoritmissa joskus vaihtamaan sarakkeiden järjestystä. Jotta syntyvät kaaviot olisivat helposti luettavissa lisäämme kaavioon otsikko-rivin, jolta käy ilmi muuttujia, jonka kertoimista on kysymys. Tämä ei ole oikea-oppista, mutta harjoitteluvai-

heessa tarpeellista. Erityisesti tulemme siirtämään kokonaisia sarakkeita RHS:ään. Siirron jälkeen yhtälöryhmän jokaisen yhtälön oikea puoli ei ole enää reaaliluku, vaan lineaarikombinaatio ykkösestä ja siirretyistä muuttujista. Vastaava kaavion rivin RHS-vektori on tämän lineaarikombinaation kerroinvektori. Käytämme tästä rivin k RHS-vektorista merkintää $\vec{r}_k$. Seuraava esimerkki valaisee asiaa. Esimerkki on käyty läpi sekä yhtälöryhmä- että kaaviomuodossa.

Esimerkki 47

$$\begin{aligned} \left\{ \begin{array}{rcl} [x] & + & 2y - z = 6 \\ 2x & + & 6y + 4z = 0 \end{array} \right. & \sim \left(\begin{array}{ccc|c} x & y & z & rhs \\ [1] & 2 & -1 & 6 \\ 2 & 6 & 4 & 0 \end{array} \right) \left| \begin{array}{l} *2 \\ \leftarrow - \end{array} \right. \\ \\ \left\{ \begin{array}{rcl} (x) & + & 2y - z = 6 \\ [2y] & + & 6z = -12 \end{array} \right. & \sim \left(\begin{array}{ccc|c} x & y & z & rhs \\ (1) & 2 & -1 & 6 \\ 0 & [2] & 6 & -12 \end{array} \right) \\ \\ \left\{ \begin{array}{rcl} (x) & + & 2y = 6 + z \\ (2y) & = & -12 - 6z \end{array} \right. & \sim \left(\begin{array}{ccc|cc} x & y & rhs & z & \\ (1) & 2 & 6 & 1 & \\ 0 & (2) & -12 & -6 & \end{array} \right) \left| \begin{array}{l} :2 \end{array} \right. \\ \\ \left\{ \begin{array}{rcl} (x) & + & 2y = 6 + z \\ (y) & = & -6 - 3z \end{array} \right. & \sim \left(\begin{array}{ccc|cc} x & y & rhs & z & \\ (1) & 2 & 6 & 1 & \\ 0 & (1) & -6 & -3 & \end{array} \right) \left| \begin{array}{l} \leftarrow - \\ *2 \end{array} \right. \\ \\ \left\{ \begin{array}{rcl} (x) & & = 18 + 7z \\ (y) & = & -6 - 3z \end{array} \right. & \sim \left(\begin{array}{ccc|cc} x & y & rhs & z & \\ (1) & 0 & 18 & 7 & \\ 0 & (1) & -6 & -3 & \end{array} \right) \end{aligned}$$

Vastaus:

$$\begin{cases} x = 18 + 7z \\ y = -6 - 3z \end{cases}$$

Gaussin algoritmi, versio IV, yleinen tapaus.

- (1) Aseta ensimmäiseen kaavioon tutkittavan yhtälöryhmän kertoimet;
- (2) $k = 1, siirretyt = 0$
- (3) repeat
- (4) % (Etsitään pivot-alkio)
- (5) määritä rivi q siten, että alkio a_{qk}^k on itseisarvoltaan suurin alkioista $\{a_{kk}^k, a_{(k+1)k}^k, \dots, a_{mk}^k\}$
- (6) if $a_{qk}^k = 0$
- (7) then
- (8) siirrä sarake k RHS:iin
- (9) $siirretyt = siirretyt + 1$;

```

(10)     else
(11)         vaida rivit  $w_q$  ja  $w_k$ .
(12)         % (Tehdään uusi kaavio, pivotoidaan)
(13)         for  $j = 1$  to  $k$ 
(14)              $w_j^{k+1} = w_j^k$ ;
(15)         endfor
(16)         for  $j = k+1$  to  $m$ 
(17)              $w_j^{k+1} = w_j^k - (a_{jk}^k/a_{kk}^k) \cdot w_k^k$ ;
(18)         endfor
(19)          $k = k + 1$ ;
(20)     endif
(21) until  $k > m$  or  $k + \text{siirretyt} > n$ 
(22) % (Onko jäljellä pivotoimattomia sarakkeita?)
(23) if  $k + \text{siirretyt} \leq n$ 
(24) then
(25)     siirrä pivotoimattomat LHS:n sarakkeet RHS:iin
(26) endif
(27) % (Onko jäljellä pivotoimattomia rivejä?)
(28) if  $k \leq m$ 
(29) then
(30)     jos jokin pivotoimattomista yhtälöistä on epätosi,
        niin yhtälöryhmällä ei ole ratkaisua. (LOPETA)
(31) endif
(32) % (Ratkaise LHS:n muuttujien arvot.)
(33) for  $q = k - 1$  downto 1
(34)      $w_q = w_q/a_{qq}$ ;
(35)     for  $i = 1$  to  $q - 1$ 
(36)          $w_i = w_i - a_{iq}w_q$ ;
(37)     endfor
(38) endfor
(39) Tulkitse tulos.

```

3.4 Yhtälöryhmän ratkaiseminen Octavella

Sec:laYROctave

Jokaiselle tulee joskus eteen tilanne, jossa ei ole aikaa tehdä rivioperaatioita käsin, vaan yhtälöryhmä tulee ratkaista nopeasti. Silloin Octave (tai Matlab) on ihanteellinen työkalu.

Kun luet tätä kappaletta, älä tyydy vilkaisemaan ensimmäistä esimerkkiä, vaan lue kaikki esimerkit ja myös kaikki tekstit. Octave (Matlab) on hyvä ja helppo työkalu, mutta sitä helposti käyttää väärin, jos ei tiedä sen erikoista tapaa käsitellä yhtälöryhmää, jolla ei ole ratkaisua.

Esimerkki 48 Ratkaise yhtälöryhmä

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

Octavella ryhmän voi ratkaista seuraavin komennoin

```
octave:1> A = [-1 2; 1 -1];
octave:2> b = [1; 2];
octave:3> ratk = A\b
ratk =
     5
     3
```

- Huomaa, että vastaus on vektori, jonka ensimmäinen koordinaatti on ensimmäisen muuttujan arvo (x :n arvo) ja toinen koordinaatti on toisen muuttujan arvo (y :n arvo). Muuttujien järjestys määräytyy niiden järjestyksestä yhtälöryhmässä (ensimmäinen muuttuja ensimmäisessä sarakkeessa, toinen muuttuja toisessa sarakkeessa, jne...). Tällä kertaa vastaus on siis: $x = 5, y = 3$.
- Ensimmäinen komentorivi määrittelee yhtälöryhmän kerroinkaavion. Kaavion kertoimet erotetaan välilyönnein ja uuden rivin alkaminen ilmoitetaan puolipisteellä. (Kaaviossa on siis kaksi riviä, kummallakin kaksi kerrointa.) Komento tallettaa kaavion muuttujalle **A**.
- Toinen komentorivi määrittelee yhtälöryhmän RHS:n (Kaksi riviä, kummassakin yksi luku.) Komento tallettaa RHS:n muuttujalle **b**.
- Kolmas komentorivi ei ole laskutoimitus vaan määrää Octaven ratkaisemaan yhtälöryhmän, jonka kerroinkaavio on **A** ja RHS on **b**. Ratkaisu talletetaan muuttujaan **ratk**. Kahden ensimmäisen komennon lopussa on puolipiste. Tämä saa aikaan sen, että Octave suorittaa laskun mykkänä, eikä kaiuta tulosta konsolille. Kolmannen komennon lopussa ei ole puolipistettä, siksi Octave kaiuttaa tuloksen konsolille. Varsinkin alussa on hyvä antaa komennot ilman puolipistettä lopussa. Silloin näkee onko kaavio juuri sellainen kuin halusi. Jos edeltävät komennot annetaan ilman puolipisteitä, niin sessio näyttää seuraavalta:

```
octave:4> A = [-1 2; 1 -1]
A =

    -1     2
     1    -1

octave:5> b = [1; 2]
```



```

b =

    1
    2

octave:6> ratk = A\b
ratk =

    5
    3

```

sim:YROctave02

Esimerkki 49 Ratkaise yhtälöryhmä

$$\begin{cases} -x + 2y = 1 \\ x - y = 2 \\ 3x - y = 0 \end{cases}.$$

On selvää, että tällä yhtälöryhmällä ei ole ratkaisua. Octavella ryhmän ratkaisuyritys näyttää seuraavalta

```

octave:7> A = [-1 2; 1 -1; 3 -1];
octave:8> b = [1; 2; 0];
octave:9> ratk = A\b
ratk =

    0.20000
    0.20000

```

Octave antaa siis vastauksen, joka ei ole yhtälöryhmän ratkaisu. Octaven antama vektori on ns. ”Pienimmän NeliöSumman” ratkaisu. Pienimmän neliösumman menetelmä käsitellään myöhemmin kurssin kuluessa.

Nyt on tärkeintä muistaa, että kun ratkaistaan yhtälöryhmä Octavella, niin tulos tulee aina tarkistaa. Octave antaa aina parhaan mahdollisen ratkaisun. Jos ratkaisun mukaiset arvot sijoitetaan yhtälöryhmään muuttujille ja jonkin yhtälön LHS ja RHS poikkeavat toisistaan enemmän kuin valittu toleranssi ($10^{-15} \dots 10^{-8}$), niin ratkaisua ei ole olemassa.

Tarkista toleranssi

Onneksi tarkistamisenkin voi tehdä Octavella. Tarkistus perustuu matriisitulon laskemiseen. Matriisitulo opetellaan seuraavassa luvussa. Jos A on yhtälöryhmän kerroinkaavio, b on yhtälöryhmän RHS, ja ratk on ratkaisuvektori, niin $A:n$ ja $\text{ratk}:n$ matriisitulo on b . Kahden edellisen esimerkin yhtälöryhmien ratkaisemiset tarkistuksineen menevät Octavella siis seuraavasti:

```

octave:1> # Ensimmäinen esimerkki
octave:2> A = [-1 2; 1 -1];
octave:3> b = [1; 2];
octave:4> ratk = A\b
ratk =

    5
    3

octave:5> # tarkistus:
octave:5> A*ratk
ans =

    1
    2

octave:6> # sama kuin b. OK ratkaisu x = 5, y = 3
octave:6> # *****
octave:6> # Toinen esimerkki
octave:6> A = [-1 2; 1 -1; 3 -1];
octave:7> b = [1; 2; 0];
octave:8> ratk = A\b
ratk =

    0.20000
    0.20000

octave:9> # tarkistus:
octave:9> A*ratk
ans =

    0.20000
    0.00000
    0.40000

octave:10> # eri kuin b. Ei ratkaisua.

```

3.5 Homogeeninen yhtälöryhmä

Sec:HomogYR1

Määritelmä 50 (1) Yhtälöryhmän ratkaisu on *triviaali*, jos kaikkien muuttujien arvot ovat ratkaisussa nollia.

(2) Ratkaisu on *ei-triviaali*, jos ainakin yksi muuttuja saa ratkaisussa nollasta eroavan arvon.

Määritelmä 51 *Yhtälöryhmä on homogeeninen, jos sen jokaisen yhtälön RHS on nolla.*

On selvää, että homogeenisella yhtälöryhmällä on aina triviaali ratkaisu. Useimmiten tämä triviaali ratkaisu ei kiinnosta laskijaa. Ainakaan sen löytämiseksi ei kannata nähdä suurta vaivaa. Käytännössä aina, kun tutkimme homogeenista yhtälöryhmää, toivomme löytävämme ei-triviaalin ratkaisun.

Lause 52 *Jos homogeenisella yhtälöryhmällä on ei-triviaali ratkaisu, niin jokin yhtälöryhmän kerroinkaavion sarake voidaan lausua muiden sarakkeiden lineaarikombinaationa.*

Perustelu: Koska homogeenisella yhtälöryhmällä on aina triviaali ratkaisu, lauseen ehto merkitsee sitä, että yhtälöryhmällä on nyt enemmän kuin yksi ratkaisu. Silloin lauseen ^{Th:KaksiRat1}?? mukaan jokin kerroinkaavion sarake voidaan lausua muiden sarakkeiden lineaarikombinaationa. □

Edellisessä lauseessa mainittu yhtälöryhmän ominaisuus: ”jokin kerroinkaavion sarake voidaan lausua muiden sarakkeiden lineaarikombinaationa”, opitaan myöhemmin paljastamaan helpolla determinantilaskulla.

Jo tässä vaiheessa on hyvä todeta eräs ongelma, joka nousee esiin, kun ratkaisemme yhtälöryhmän rivioperaatioilla tai laskemme edellä mainitun helpon determinanttilaskun. Kummassakin tapauksessa laskimemme tai tietokoneemme tekee suuren määrän kerto-, yhteen-, ja vähennyslaskuja. Jakolaskua tehdään vähemmän, mutta sitäkään ei voi välttää. Tämän seurauksena tulokseen kertyy väistämättä pyöristysvirheiden vaikutus. Jos laskutuloksen pitäisi olla tasan 0, mutta pyöristysvirheiden takia se onkin $1.3 \cdot 10^{-14}$, niin se ei silloin ole siis tasan nolla! Tässä kurssimateriaalissa palaamme tähän pyöristysvirheiden ongelmaan muutaman kerran. Emme kuitenkaan käsittele sitä kokonaan.

Numeerisen matematiikan kurssilla asiaa pohditaan lisää. Jos työssäsi myöhemmin joudut ohjelmoimaan ohjelmakoodia, joka ratkaisee yhtälöryhmän niin huomaa, että peruskurssin yksinkertaisin algoritmi-versio ei ole professionaalia tasoa.

Seuraavaksi esitämme kolme esimerkkiä homogeenisen yhtälöryhmän ratkaisemisesta. Ensimmäinen esimerkki on helppo ja yksinkertainen. Toinen esimerkki tuo esiin tavan käsitellä yhtälöryhmää niin, että voidaan hyödyntää laskinta tai tietokonetta. Kolmas esimerkki tuo esiin vähän mutkikkaamman tilanteen, jonka tyylikäs ratkaisu onnistuu parhaiten puhtaalla Gauss-Jordan algoritmin läpiviennillä.

Esimerkki 53 Ratkaise yhtälöryhmä

$$\begin{cases} x + 2y + 4z = 0 \\ -x + 3y + z = 0 \\ 3x + y + 7z = 0 \end{cases}.$$

Yhtälöryhmä on homogeeninen. On helppo nähdä, että kerroinkaavion kolmas sarake saadaan, kun toiseen sarakkeeseen lisätään kahdella kerrottuna ensimmäinen sarake ($\vec{a}_{\bullet 3} = \vec{a}_{\bullet 2} + 2 \cdot \vec{a}_{\bullet 1}$.)

Gauss-Jordan algoritmilla

$$\begin{aligned} & \left(\begin{array}{ccc|c} [1] & 2 & 4 & 0 \\ -1 & 3 & 1 & 0 \\ 3 & 1 & 7 & 0 \end{array} \right) \quad \left| \begin{array}{l} *1 \\ \leftarrow + \\ \leftarrow - \end{array} \right. \quad \begin{array}{l} *3 \\ \\ \leftarrow - \end{array} \\ \Leftrightarrow & \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & [5] & 5 & 0 \\ 0 & -5 & -5 & 0 \end{array} \right) \quad \left| \begin{array}{l} *1 \\ \leftarrow + \end{array} \right. \quad :5 \\ \Leftrightarrow & \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & (1) & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \Leftrightarrow \begin{cases} x + 2y + 4z = 0 \\ y + z = 0 \end{cases} \end{aligned}$$

Muuttujan z sarakkeessa ei pivotoitu, siksi muuttujaa z sisältävät termit siirretään RHS:iin. Tämän jälkeen z :lle annetaan jokin arvo. Jos annamme z :lle arvon $z = 1$, niin saamme yhtälöryhmälle ei-triviaalin ratkaisun. Jos annamme z :lle arvon $z = a$, niin saamme ne kaikki.

$$\Leftrightarrow \begin{cases} x + 2y = -4z \\ y = -z \end{cases} \begin{array}{l} (1) \\ (2) \end{array}$$

Valitaan muuttujalle z arvoksi $z = a$. Sen jälkeen x ja y ratkaistaan yhtälöistä tavalliseen tapaan.

$$\Leftrightarrow \begin{cases} z = a \\ y = -a \\ x = -2a \end{cases}$$

Suoraviivainen ratkaisuyritys Octavella jää epätydyttäväksi

```
octave:1> A = [1 2 4; -1 3 1; 3 1 7];
octave:2> b = [0 0 0]';
octave:3> x = A\b
warning: matrix singular to machine precision, rcond =
    9.25186e-18
x =
```

0
0
0

Esimerkki 54 Etsi Octaven avulla ainakin yksi ei-triviaali ratkaisu yhtälöryhmälle

$$\begin{cases} x + 2y + 4z = 0 \\ -x + 3y + z = 0 \\ 3x + y + 7z = 0 \end{cases}.$$

Edellisestä esimerkistä tiedämme, että on olemassa ratkaisuja, joissa kaikki muuttujat saavat nollasta eroavan arvon. Etsimme seuraavassa sen ratkaisun, jossa z saa arvon $z = 1$. Siirretään yhtälöryhmässä jokainen z -termi oikealle ja sijoitetaan z :n paikalle 1. Syntyvä yhtälöryhmä ei ole homogeeninen ja se voidaan helposti ratkaista Octavella

$$\begin{cases} x + 2y + 4z = 0 \\ -x + 3y + z = 0 \\ 3x + y + 7z = 0 \end{cases} \Leftrightarrow \begin{cases} x + 2y = -4 \cdot 1 \\ -x + 3y = -1 \cdot 1 \\ 3x + y = -7 \cdot 1 \end{cases}$$

```
octave:2> A = [1 2; -1 3; 3 1];
octave:3> b = [-4; -1; -7];
octave:4> ratk = A\b
ratk =

    -2.0000
    -1.0000

octave:5> # tarkistus
octave:5> A*ratk
ans =

    -4.0000
    -1.0000
    -7.0000

octave:6> # oikein! OK, siis z = 1, x = -2, y = -1.
```

HomogYR1_esim1

Esimerkki 55 Ratkaise yhtälöryhmä

$$\begin{cases} x + 2y + z = 0 \\ -x - 2y + z - 2w = 0 \\ 3x + 6y + 2z + w = 0 \\ 2x + 4y + z + w = 0 \end{cases}.$$

Octavella saamme jonkin ratkaisun, mutta emme kaikkia. Siksi teemme täydellisen ratkaisun Gauss-Jordan -algoritmilla. Toinen sarake on ensimmäisen lineaarikombinaatio ($\vec{a}_{\bullet 2} = 2 \cdot \vec{a}_{\bullet 1}$) ja neljäs sarake on ensimmäisen ja kolmannen lineaarikombinaatio ($\vec{a}_{\bullet 4} = \vec{a}_{\bullet 1} - \vec{a}_{\bullet 3}$).

$$\begin{aligned}
 & \left(\begin{array}{cccc|c} [1] & 2 & 1 & 0 & 0 \\ -1 & -2 & 1 & -2 & 0 \\ 3 & 6 & 2 & 1 & 0 \\ 2 & 4 & 1 & 1 & 0 \end{array} \right) \left| \begin{array}{c} *1 \\ \leftarrow + \\ \\ \end{array} \right| \left| \begin{array}{c} *3 \\ \\ \leftarrow - \\ \\ \end{array} \right| \left| \begin{array}{c} *2 \\ \\ \\ \leftarrow - \end{array} \right| \\
 & \sim \left(\begin{array}{cccc|c} (1) & 2 & 1 & 0 & 0 \\ 0 & 0 & [2] & -2 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{array} \right) \left| \begin{array}{c} *1/2 \\ \leftarrow + \\ \leftarrow + \end{array} \right| *1/2 \\
 & \sim \left(\begin{array}{cccc|c} (1) & 2 & 1 & 0 & 0 \\ 0 & 0 & (1) & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \left| \begin{array}{c} \leftarrow - \\ *1 \\ \\ \end{array} \right| \\
 & \sim \left(\begin{array}{cccc|c} (1) & 2 & 0 & 1 & 0 \\ 0 & 0 & (1) & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \Leftrightarrow \begin{cases} x + 2y + w = 0 \\ z - w = 0 \end{cases}
 \end{aligned}$$

Muuttujat, joiden sarakeissa ei pivotoitu (y ja w) siirretään RHS:iin ja niille annetaan jotkin sopivat arvot (esim. $y = a$ ja $w = b$. Ainakin toinen ei ole nolla.)

$$\begin{aligned}
 & \Leftrightarrow \begin{cases} x = -2y - w \\ z = w \end{cases} \\
 & \xrightarrow{y=a, w=b} \begin{cases} x = -2a - b \\ y = a \\ z = b \\ w = b \end{cases} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = a \cdot \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}
 \end{aligned}$$

Tässä tapauksessa siis homogeenisen yhtälöryhmän kaikki ratkaisut saadaan kahden ratkaisuvektorin lineaarikombinaatioina. On tapana sanoa, että vektorit $(-2 \ 1 \ 0 \ 0)^T$ ja $(-1 \ 0 \ 1 \ 1)^T$ 'virittävät' homogeenisen yhtälöryhmän ratkaisujoukon.

4 MATRIISIT JA MATRIISILASKUT

Ch:Matrix

4.1 Matriisi ja matriisilaskut

Sec:MatLaskut

Matriisi on suorakulmainen lukukaavio. Matriiseja ovat esimerkiksi:

$$\begin{pmatrix} 2 & 0.4 & 8 \\ 0 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \quad (x_1 \ x_2 \ x_3) \text{ ja } \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Matriiseihin perehtyminen voidaan perustella useilla järkisyillä.

1. Matriisien avulla voidaan suuria lukujoukkoja esittää mahdollisimman pienellä vaivalla ja kokonaisille lukukaavioille voidaan antaa nimiä $\mathbf{A}, \mathbf{B}, \dots$ jne. Tämä helpottaa suuresti asioiden esittämistä. Eräissä tilanteissa matriisi on hyvin onnistunut ja havainnollinen tapa esittää lukuryhmiä. Esimerkiksi yhtälöryhmän

$$\begin{cases} 2x - 5y = -1 \\ 3x + 7y = 13 \end{cases}$$

oleelliset osat tiivistyvät matriiseihin

$$\mathbf{A} = \begin{pmatrix} 2 & -5 \\ 3 & 7 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \text{ ja } \vec{b} = \begin{pmatrix} -1 \\ 13 \end{pmatrix}.$$

2. Matriiseille voidaan määritellä yhteen- ja kertolasku. Matriiseilla voidaan siis laskea. Näiden laskutoimitusten avulla on kehitetty menetelmiä eräiden matriisien avulla kuvattavien ongelmien ratkaisemiseen. Esimerkkinä mainittakoon yhtälöryhmän ratkaiseminen.
3. Vaikka itse et koskaan harrastaisikaan matriisilaskentaa, niin varmaa on, että tietokoneesi harrastaa. Jos tietokone esimerkiksi ei löydä ratkaisua yhtälöryhmälle se yleensä antaa virheilmoituksen, jossa se kertoo mitä vikaa oli kerroinmatriisissa. Virheilmoituksen ymmärtäminen edellyttää matriisilaskennan perusteiden osaamista.
4. Matematiikassa tarkastellaan struktuureja, jotka muodostuvat olioista ja niille määritellyistä laskutoimituksista. Struktuureja luokitellessa on tapana sanoa, että tietyt ehdot täyttävä struktuuri on ryhmä (tarkemmin luvussa III), tietyt ehdot täyttävä ryhmä on rengas ja tietyt ehdot täyttävä rengas on kunta. Reaaliluvut muodostavat kunnan, ja kompleksiluvut muodostavat kunnan. Opiskelija on

peruskoulussa ja lukiossa harjoitellut laskutoimitusten suorittamista reaalitylukukunnassa. Kun sääntönä voidaan pitää, että reaalitylukuja ja kompleksilukuja lukuunottamatta mikään mielenkiintoinen struktuuri ei ole kunta, saatta peruskoulussa hankittu sujuva laskutaito aiheuttaa kitkaa tutustuttaessa uusiin matemaattisiin struktuureihin. Neliömatriisit muodostavat renkaan mutta eivät kuntaa. Niillä laskettaessa tulee muistaa, että matriisilla ei saa jakaa eikä supistaa (sillä kertolaskun käänteisalkiota ei aina ole olemassa) ja kertolasku ei ole vaihdannainen ($\mathbf{AB} \neq \mathbf{BA}$). Tätä kautta matriisilaskenta kehittää teoreettisessa ajattelussa tarvittavaa tarkkuutta ja joustavuutta.

Määritelmä 56 *Suorakulmainen lukukaavio, jossa on m vaakariviä ja n pystysaraketta on $m \times n$ -matriisi. Tyyppeä $m \times 1$ oleva matriisi on pystyvektori ja tyyppeä $1 \times n$ oleva matriisi on vaakavektori. Kaaviossa esiintyviä lukuja sanomme matriisialkioiksi.*

Merkintäsopimuksia:

(1) Tulemme kirjoittamaan lukukaavion aina isojen sulkeiden sisään. On myös luvalista käyttää hakasulkumerkintää, mutta tässä monisteessa käytämme systemaattisesti kaarisulkeita.

(2) Matriisia merkitsemme kirjaimella (iso, vahvennettu ja kursiivi) ja matriisissa olevia alkioita merkitsemme samalla kirjaimella (pieni, vahventamaton ja kursiivi) varustettuna alaindeksein, jotka ilmaisevat millä rivillä ja missä sarakkeessa alkio on. Esimerkiksi a_{12} on matriisin $\mathbf{A}$ rivillä 1 sarakkeessa 2 oleva alkio.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

(3) Vaakavektoria merkitään samoin kuin matriisia, mutta usein valitaan pieni kirjain, ja rivi-indeksi voidaan jättää pois. Pystyvektorin tapauksessa sarakeindeksi voidaan jättää pois.

$$\mathbf{r} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \end{pmatrix} = \begin{pmatrix} r_1 & r_2 & \cdots & r_n \end{pmatrix}, \quad \mathbf{s} = \begin{pmatrix} s_{11} \\ s_{21} \\ \vdots \\ s_{m1} \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{pmatrix}$$

(4) Fysiikassa ja tekniikassa on tapana merkitä vektoreita vahventamattomalla, kursivilla vektorilla, jonka päällä on nuoli.

$$\vec{r} = \begin{pmatrix} r_1 & r_2 & \cdots & r_n \end{pmatrix}, \quad \vec{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{pmatrix}$$

(5) Tulemme käyttämään kumpaakin merkintää (3 ja 4) vektorille. On tärkeätä, että opiskelija tottuu siihen, että eri tilanteissa käytetään hieman erilaisia merkintöjä.

Määritelmä 57 Usein käytämme matriisille merkintää $\mathbf{A} = (a_{ij})$. Vastaavasti matriisin $\mathbf{A}$ alkion a_{ij} käytetään joskus merkintää $(\mathbf{A})_{ij}$.

Määritelmä 58 (a) Matriisit $\mathbf{A} = (a_{ij})$ ja $\mathbf{B} = (b_{ij})$ ovat samat (identtiset, engl. equal), jos ne ovat saman kokoiset ja, $a_{ij} = b_{ij}$ kaikilla i, j . Tällöin merkitään $\mathbf{A} = \mathbf{B}$.

(b) Kahden $m \times n$ -matriisin $\mathbf{A} = (a_{ij})$ ja $\mathbf{B} = (b_{ij})$ summa on matriisi

$$\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$$

(c) Reaaliluvun λ ja $m \times n$ -matriisin $\mathbf{A} = (a_{ij})$ tulo $\lambda \mathbf{A}$ on $m \times n$ -matriisi $(\lambda \cdot a_{ij})$. (Kaikki alkion a_{ij} kerrotaan λ :lla.)

(d) Sovimme lisäksi merkintätavoista

$$\begin{aligned} (-1)\mathbf{A} &= -\mathbf{A}, \\ \mathbf{A} - \mathbf{B} &= \mathbf{A} + (-1)\mathbf{B} = \mathbf{A} + (-\mathbf{B}) \end{aligned}$$

Esimerkki 59

$$\begin{aligned} 2 \cdot \begin{pmatrix} 0 & -1 & 2 \\ 1 & 4 & 0 \end{pmatrix} - 3 \cdot \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 3 \end{pmatrix} &= \begin{pmatrix} 0 & -2 & 4 \\ 2 & 8 & 0 \end{pmatrix} - \begin{pmatrix} 3 & -6 & 0 \\ 0 & 3 & 9 \end{pmatrix} \\ &= \begin{pmatrix} -3 & 4 & 4 \\ 2 & 5 & -9 \end{pmatrix} \end{aligned}$$

Samankokoisten matriisien yhteen- ja vähennyslasku ja vähennyslasku ovat melko selviä asioita. Seuraavaksi määrittelemme matriisin kertomisen matriisilla.

Koulussa olet todennäköisesti oppinut kahden vektorin pistetulon $\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$. Kertaamme asian ensin

Määritelmä 60 Kahden samanmittaisen vektorin $\vec{u} = (u_i)$ ja $\vec{v} = (v_i)$ pistetulo on

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Määritelmä 61 $m \times n$ -matriisin $\mathbf{A} = (a_{ij})$ ja $n \times p$ -matriisin $\mathbf{B} = (b_{jk})$ matriisitulo $\mathbf{AB}$ on $m \times p$ -matriisi

$$(\mathbf{AB})_{ij} = \sum_{k=1}^n a_{ik} b_{kj} = \vec{a}_{i\bullet} \cdot \vec{b}_{\bullet j}$$

(Siis: $\mathbf{AB}$:n rivillä i sarakkeessa j oleva luku on $\mathbf{A}$:n i :nnen rivin ja $\mathbf{B}$:n j :nnen sarakkeen pistetulo.)

Esimerkki 62 Seuraavassa esimerkissä on lisätty sarakkeita ja rivejä erottavat viivat joihinkin kaavioihin. Tämä ei ole kovin tavallista, mutta tässä tapauksessa saattaa helpottaa kaavion lukemista.

$$\begin{aligned} & \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & -2 & 3 \end{pmatrix} = \\ & = \begin{pmatrix} (2 \ -1) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} & (2 \ -1) \cdot \begin{pmatrix} -1 \\ -2 \end{pmatrix} & (2 \ -1) \cdot \begin{pmatrix} 0 \\ 3 \end{pmatrix} \\ (1 \ 0) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} & (1 \ 0) \cdot \begin{pmatrix} -1 \\ -2 \end{pmatrix} & (1 \ 0) \cdot \begin{pmatrix} 0 \\ 3 \end{pmatrix} \\ (0 \ 3) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} & (0 \ 3) \cdot \begin{pmatrix} -1 \\ -2 \end{pmatrix} & (0 \ 3) \cdot \begin{pmatrix} 0 \\ 3 \end{pmatrix} \end{pmatrix} \\ & = \begin{pmatrix} \frac{2 \cdot 1 + (-1) \cdot 0}{1 \cdot 1 + 0 \cdot 0} & \frac{2 \cdot (-1) + (-1) \cdot (-2)}{1 \cdot (-1) + 0 \cdot (-2)} & \frac{2 \cdot 0 + (-1) \cdot 3}{1 \cdot 0 + 0 \cdot 3} \\ \frac{0 \cdot 1 + 3 \cdot 0}{0 \cdot (-1) + 3 \cdot (-2)} & \frac{0 \cdot (-1) + 3 \cdot (-2)}{0 \cdot 0 + 3 \cdot 3} \end{pmatrix} \\ & = \begin{pmatrix} 2 & 0 & -3 \\ 1 & -1 & 0 \\ 0 & -6 & 9 \end{pmatrix} \end{aligned}$$

c3s1esim1

Huomaa tyyppien yhteensopivuus $(3 \times 2\text{-matriisi})(2 \times 3\text{-matriisi}) = (3 \times 3\text{-matriisi})!$

Edeltävä määritelmä ja esimerkki kannattaa tiivistää seuraavalla tavalla: ”Kun ollaan laskemassa matriisia $\mathbf{C} = \mathbf{AB}$, niin laskettaessa lukua riville i sarakkeeseen j , tarvitaan ensimmäisestä matriisista i :s rivi ja toisesta matriisista j :s sarake. Rivin ja sarakkeen pistetulo antaa alkion tulo-kaavion vastaavaan paikkaan.”

Alla taulukossa (2) (sivu 75) on esitetty edeltävän esimerkin tulokaavion kaikkien alkoiden lasku ja erityisesti ”mistä luvut tulevat?”.

Taulukko 2. Esimerkissä (62) (sivu 74) laskettujen lausekkeiden sisältämien lukujen sijainti alkuperäisissä matriiseissa.

c3s1taulu1

mikä = mitä?	mistä luvut?
$c_{11} = 2 \cdot 1 + (-1) \cdot 0$	$\begin{pmatrix} \boxed{2} & \boxed{-1} \\ 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \boxed{1} & -1 & 0 \\ \boxed{0} & -2 & 3 \end{pmatrix}$
$c_{12} = 2 \cdot (-1) + (-1) \cdot (-2)$	$\begin{pmatrix} \boxed{2} & \boxed{-1} \\ 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & \boxed{-1} & 0 \\ 0 & \boxed{-2} & 3 \end{pmatrix}$
$c_{13} = 2 \cdot 0 + (-1) \cdot 3$	$\begin{pmatrix} \boxed{2} & \boxed{-1} \\ 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 & \boxed{0} \\ 0 & -2 & \boxed{3} \end{pmatrix}$
$c_{21} = 1 \cdot 1 + 0 \cdot 0$	$\begin{pmatrix} 2 & -1 \\ \boxed{1} & \boxed{0} \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \boxed{1} & -1 & 0 \\ \boxed{0} & -2 & 3 \end{pmatrix}$
$c_{22} = 1 \cdot (-1) + 0 \cdot (-2)$	$\begin{pmatrix} 2 & -1 \\ \boxed{1} & \boxed{0} \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & \boxed{-1} & 0 \\ 0 & \boxed{-2} & 3 \end{pmatrix}$
$c_{23} = 1 \cdot 0 + 0 \cdot 3$	$\begin{pmatrix} 2 & -1 \\ \boxed{1} & \boxed{0} \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 & \boxed{0} \\ 0 & -2 & \boxed{3} \end{pmatrix}$
$c_{31} = 0 \cdot 1 + 3 \cdot 0$	$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ \boxed{0} & \boxed{3} \end{pmatrix} \begin{pmatrix} \boxed{1} & -1 & 0 \\ \boxed{0} & -2 & 3 \end{pmatrix}$
$c_{32} = 0 \cdot (-1) + 3 \cdot (-2)$	$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ \boxed{0} & \boxed{3} \end{pmatrix} \begin{pmatrix} 1 & \boxed{-1} & 0 \\ 0 & \boxed{-2} & 3 \end{pmatrix}$
$c_{33} = 0 \cdot 0 + 3 \cdot 3$	$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ \boxed{0} & \boxed{3} \end{pmatrix} \begin{pmatrix} 1 & -1 & \boxed{0} \\ 0 & -2 & \boxed{3} \end{pmatrix}$

Esimerkki 63 Seuraavassa välimuotoja on vähemmän. Saatko saman tuloksen?

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 2-1+0 & -1+0+0 \\ 0-2+0 & 0+0+9 \\ 2+1+0 & -1+0+3 \\ 2+0+0 & -1+0+0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -2 & 9 \\ 3 & 2 \\ 2 & -1 \end{pmatrix}$$

Huomaa taas tyyppien yhteensopivuus $(4 \times 3\text{-matriisi})(3 \times 2\text{-matriisi}) = (4 \times 2\text{-matriisi})!$

Esimerkki 64 *Olkoon*

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

Laskemalla tulot nähdään, että a) tulo $\mathbf{AB}$ on hyvin määritelty ja laskettavissa, mutta tuloa $\mathbf{BA}$ ei voi laskea. b) Tulot $\mathbf{BC}$ ja $\mathbf{CB}$ voidaan laskea, mutta niistä saadaan eri tulokset. Siis $\mathbf{BC} \neq \mathbf{CB}$.

Esimerkki 65 *Huomaa, että vektoreiden*

$$\mathbf{u} = \begin{pmatrix} 2 & 3 \end{pmatrix} \text{ ja } \mathbf{v} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

Matriisitulo ja pistetulo ovat tarkasti tulkittuina eri asiat

$$\begin{aligned} \mathbf{uv} &= \begin{pmatrix} 2 \cdot 5 + 3 \cdot 6 \end{pmatrix} = \begin{pmatrix} 28 \end{pmatrix} \\ \mathbf{u} \cdot \mathbf{v} &= 2 \cdot 5 + 3 \cdot 6 = 28 \end{aligned}$$

Esimerkki 66 *Luvun alussa esiintyi yhtälöryhmä ja siihen liittyvät matriisit*

$$\begin{cases} 2x - 5y = -1 \\ 3x + 7y = 13 \end{cases}$$

$$\mathbf{A} = \begin{pmatrix} 2 & -5 \\ 3 & 7 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \text{ ja } \mathbf{b} = \begin{pmatrix} -1 \\ 13 \end{pmatrix}.$$

Matriisikertolaskun määritelmästä ja matriisien yhtäsuuruuden määritelmästä seuraa, että seuraavat kolme riviä sanovat saman asian:

$$\mathbf{Ax} = \mathbf{b} \tag{4.1}$$

$$\Leftrightarrow \begin{pmatrix} 2 & -5 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 13 \end{pmatrix} \tag{4.2}$$

$$\Leftrightarrow \begin{cases} 2x - 5y = -1 \\ 3x + 7y = 13 \end{cases} \tag{4.3}$$

Määritelmä 67 Nollamatriisi $\mathbf{O}$ on matriisi, jonka kaikki alkiot ovat nollia. Jos matriisissa on sarakkeita yhtä monta kuin riviä, se on neliömatrissi. $n \times n$ -neliömatrissin (a_{ij}) päälävistäjä muodostuu alkiosta a_{kk} , $1 < k < n$. Neliömatrissi on diagonaalinen, jos sen ainoat nollasta eroavat alkiot ovat päälävistäjällä. Neliömatrissi $\mathbf{I}$ on yksikkömatrissi, jonka päälävistäjän alkiot ovat ykkösiä ja muut alkiot ovat nollia.

$$\mathbf{O} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

Seuraava määritelmä saattaa tuntua keinotekoiselta, mutta määritelmässä esitelty Kroneckerin delta on paljon käytössä teknisessä laskennassa. Asia on ulkoa opettelu arvoisen.

Määritelmä 68 *Kroneckerin delta on*

$$\delta_{ij} = \begin{cases} 1, & \text{kun } i = j \\ 0, & \text{kun } i \neq j \end{cases}$$

Siiis $\mathbf{I} = (\delta_{ij})$.

Seuraavilla lauseilla on täsmällinen matemaattiseen käsitteistöön liittyvä merkitys, johon palataan opetuksessa myöhemmin. Lauseet (69) ja (70) luettelevat matriiseille voimassa olevia tavallisesta laskemisesta tuttuja ominaisuuksia. Esimerkki (71) kertoo missä suhteessa matriiseilla laskeminen poikkeaa reaalityyppien laskemisesta. (Tässä vaiheessa on tärkeää muistaa esimerkin (71) antama viesti.) Lauseet on koottu peräkkäin, jotta kaavat olisivat helposti silmäiltävissä. Todistukset on esitetty esimerkin jälkeen. Seuraavassa merkinnät $+_m$ ja $\cdot_m$ tarkoittavat matriisien laskutoimituksia.

h ominaisuuudet

Lause 69 *Olko $\mathbf{A}, \mathbf{B}$ ja $\mathbf{C}$ $m \times n$ -matriiseja ja $\alpha, \beta \in \mathbb{R}$ reaalityypin lukuja. Silloin*

- (1) $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
- (2) $\mathbf{A} + \mathbf{O} = \mathbf{A}$
- (3) $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$
- (4) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- (5) $1 \cdot \mathbf{A} = \mathbf{A}$
- (6) $\alpha(\beta \mathbf{A}) = (\alpha\beta)\mathbf{A}$
- (7) $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$
- (8) $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$

Kommentteja: (1) $+_m$ on liitännäinen (eli assosiatiivinen), (2) nollamatriisi on $+_m$:n neutraalialkio, (3) $-\mathbf{A}$ on $\mathbf{A}$:n käänteisalkio, (4) $+_m$ on vaihdannainen (eli kommutatiivinen)

n ominaisuuudet

Lause 70 Olkoon $\mathbf{A}, \mathbf{B}$ ja $\mathbf{C}$ matriiseja ja $\alpha \in \mathbb{R}$ reaaliluku. Jos matriisien tyypit ovat sellaiset, että seuraavat tulot ovat olemassa, niin silloin

- (1) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- (2) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- (3) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
- (4) $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$
- (5) $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$

Kommentteja: (1) $\cdot_m$ on liitännäinen (assosiatiivinen), (2,3) $+_m$ ja $\cdot_m$ noudattavat osittelulakeja, (5) $\mathbf{I}$ on matriisikertolaskun $\cdot_m$ neutraalialkio.

kussa on outoa

Esimerkki 71 Olkoon $\mathbf{A}, \mathbf{B}$ ja $\mathbf{C}$ matriiseja. Jos matriisien tyypit ovat sellaiset, että seuraavat tulot ovat olemassa, niin on mahdollista, että

- (1) $\mathbf{AB} \neq \mathbf{BA}$
- (2) $(\mathbf{AC} = \mathbf{BC})$ vaikka $\mathbf{A} \neq \mathbf{B}$
- (3) $(\mathbf{AB} = \mathbf{O})$ vaikka $(\mathbf{A} \neq \mathbf{O}$ ja $\mathbf{B} \neq \mathbf{O})$

Lauseiden (69) ja (70) osalta todistamme vain muutaman kohdan ja esimerkin (71) väitteet osoitetaan oikeiksi antamalla esimerkki jokaiseen kohtaan.

Todistuksia: (Lause (69) väite (I))

$$\begin{aligned} \forall i, j: (\mathbf{A} + (\mathbf{B} + \mathbf{C}))_{ij} &= a_{ij} + (\mathbf{B} + \mathbf{C})_{ij} = a_{ij} + b_{ij} + c_{ij} \\ &= (\mathbf{A} + \mathbf{B})_{ij} + c_{ij} = ((\mathbf{A} + \mathbf{B}) + \mathbf{C})_{ij} \end{aligned}$$

(Lause (70) väite (I)) Olkoot $\mathbf{A}, \mathbf{B}$ ja $\mathbf{C}$ tyyppisiä $m \times n$, $n \times p$ ja $p \times r$ olevia matriiseja. Silloin matriisit $(\mathbf{AB})\mathbf{C}$ ja $\mathbf{A}(\mathbf{BC})$ ovat kumpikin tyyppiä $m \times r$ ja

$$(\mathbf{AB})_{ij} = \sum_{k=1}^n a_{ik}b_{kj} \text{ ja } (\mathbf{BC})_{st} = \sum_{q=1}^p a_{sq}b_{qt}$$

joten

$$\begin{aligned}
 ((\mathbf{AB})\mathbf{C})_{ij} &= \sum_{q=1}^p (\mathbf{AB})_{iq} c_{qj} = \sum_{q=1}^p \left(\sum_{k=1}^n a_{ik} b_{kq} \right) c_{qj} \\
 &= (a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1})c_{1j} \\
 &\quad + (a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2})c_{2j} \\
 &\quad + \cdots \\
 &\quad + (a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np})c_{pj} \\
 &= a_{i1}(b_{11}c_{1j} + b_{12}c_{2j} + \cdots + b_{1p}c_{pj}) \\
 &\quad + a_{i2}(b_{21}c_{1j} + b_{22}c_{2j} + \cdots + b_{2p}c_{pj}) \\
 &\quad + \cdots \\
 &\quad + a_{in}(b_{n1}c_{1j} + b_{n2}c_{2j} + \cdots + b_{np}c_{pj}) \\
 &= \sum_{k=1}^n a_{ik} \left(\sum_{q=1}^p b_{kq} c_{qj} \right) = \sum_{k=1}^n a_{ik} (\mathbf{BC})_{kj} \\
 &= (\mathbf{A}(\mathbf{BC}))_{ij}
 \end{aligned}$$

th_matkerto:kertolaskun ominaisuudet
(Lause (70) väite (2))

$$\begin{aligned}
 (\mathbf{A}(\mathbf{B} + \mathbf{C}))_{ij} &= \sum_{k=1}^n a_{ik} (\mathbf{B} + \mathbf{C})_{kj} = \sum_{k=1}^n a_{ik} (b_{kj} + c_{kj}) \\
 &= a_{i1}(b_{1j} + c_{1j}) + a_{i2}(b_{2j} + c_{2j}) + \cdots + a_{in}(b_{nj} + c_{nj}) \\
 &= (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}) + (a_{i1}c_{1j} + a_{i2}c_{2j} + \cdots + a_{in}c_{nj}) \\
 &= \sum_{k=1}^n a_{ik} b_{kj} + \sum_{k=1}^n a_{ik} c_{kj} = (\mathbf{AB} + \mathbf{AC})_{ij}
 \end{aligned}$$

th_matkerto:kertolaskun ominaisuudet
(Lause (70) väite (5)) Tutkitaan $m \times n$ -matriisin $\mathbf{A} = (a_{ij})$, $m \times m$ -yksikkömatriisin $\mathbf{I} = (\delta_{ij})$ ja $n \times n$ -yksikkömatriisin $\mathbf{J} = (\delta_{ij})$ tuloja.

$$\begin{aligned}
 (\mathbf{IA})_{ij} &= \sum_{k=1}^m \delta_{ik} a_{kj} = \delta_{i1}a_{1j} + \delta_{i2}a_{2j} + \cdots + \delta_{im}a_{mj} = a_{ij} = (\mathbf{A})_{ij} \\
 (\mathbf{AJ})_{ij} &= \sum_{k=1}^n a_{ik} \delta_{kj} = a_{i1}\delta_{1j} + a_{i2}\delta_{2j} + \cdots + a_{in}\delta_{nj} = a_{ij} = (\mathbf{A})_{ij}
 \end{aligned}$$

c3s1esim2:mik\IeC {\\"a} matlaskussa on outoa
(Esimerkki (71) Olkoon

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 1 & 5 \\ 3 & 6 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 7 & 8 \\ 0 & 0 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{E} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

Silloin suora lasku osoittaa, että:

- (1) $\mathbf{AB} \neq \mathbf{BA}$,
- (2) $\mathbf{AC} = \mathbf{BC}$, vaikka $\mathbf{A} \neq \mathbf{B}$,
- (3) $\mathbf{DE} = \mathbf{O}$, vaikka $\mathbf{D} \neq \mathbf{O}$ ja $\mathbf{E} \neq \mathbf{O}$.

4.2 Transponointi

sec:MatTransp

Seuraavassa määriteltävää transpoosimerkintää käytetään paljon. Vaikka asia on yksinkertainen, sen pitää tulla täydellisen tutuksi.

Määritelmä 72 $m \times n$ -matriisin $\mathbf{A} = (a_{ij})$ transponoitu matriisi eli transpoosi on $n \times m$ -matriisi $\mathbf{A}^T$ jolle

$$(\mathbf{A}^T)_{ij} = a_{ji}.$$

(Matriisi $\mathbf{A}^T$ saadaan siis $\mathbf{A}$:sta vaihtamalla vaakarivit pystysarakkeiksi.)

Esimerkki 73

$$\begin{pmatrix} 0 & 2 & 3 \\ 5 & -1 & 6 \end{pmatrix}^T = \begin{pmatrix} 0 & 5 \\ 2 & -1 \\ 3 & 6 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1 \ x_2 \ \cdots \ x_n)^T$$

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

(Huomaa: vasen ylänurkka pysyy paikallaan, ensimmäisestä rivistä tulee ensimmäinen sarake, toisesta rivistä tulee toinen sarake, ...)

Lause 74 Jos matriisien $\mathbf{A}$ ja $\mathbf{B}$ väliset laskutoimitukset ovat määriteltyjä ja $\lambda \in \mathbb{R}$, niin

- (1) $(\mathbf{A}^T)^T = \mathbf{A}$
- (2) $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- (3) $(\lambda \mathbf{A})^T = \lambda \mathbf{A}^T$
- (4) $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T \quad \lll \text{TÄRKEÄ!}$

Todistus: Kohdat (1) - (3) nähdään välittömästi. (4):

$$((\mathbf{AB})^T)_{ij} = (\mathbf{AB})_{ji} = \sum_{k=1}^n a_{jk}b_{ki} = \sum_{k=1}^n (\mathbf{B}^T)_{ik}(\mathbf{A}^T)_{kj} = (\mathbf{B}^T\mathbf{A}^T)_{ij}$$

Esimerkki 75 Olkoon

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \text{ ja } \mathbf{B} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

Lasketaan ensin väitteen (4) mukaiset lausekkeet

$$\begin{aligned} (\mathbf{AB})^T &= \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \right)^T = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}^T = \begin{pmatrix} 19 & 43 \\ 22 & 50 \end{pmatrix} \\ \mathbf{B}^T\mathbf{A}^T &= \begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 19 & 43 \\ 22 & 50 \end{pmatrix} \end{aligned}$$

Seuraavaksi kirjaamme uudelleen samat laskut, mutta nyt merkitsemme erityisesti mistä tulevat ne luvut joista laskemme tulomatriisin rivillä yksi sarakkeessa kaksi olevan alkion 43.

$$\begin{aligned} (\mathbf{AB})^T &= \left(\left(\begin{pmatrix} 1 & 2 \\ \boxed{3} & \boxed{4} \end{pmatrix} \begin{pmatrix} \boxed{5} & 6 \\ \boxed{7} & 8 \end{pmatrix} \right) \right)^T = \begin{pmatrix} 19 & 22 \\ \boxed{43} & 50 \end{pmatrix}^T = \begin{pmatrix} 19 & \boxed{43} \\ 22 & 50 \end{pmatrix} \\ \mathbf{B}^T\mathbf{A}^T &= \begin{pmatrix} \boxed{5} & \boxed{7} \\ 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & \boxed{3} \\ 2 & \boxed{4} \end{pmatrix} = \begin{pmatrix} 19 & \boxed{43} \\ 22 & 50 \end{pmatrix} \end{aligned}$$

Määritelmä 76 Neliömatriisi $\mathbf{A}$ on symmetrinen, jos $\mathbf{A}^T = \mathbf{A}$ (eli $a_{ij} = a_{ji}, \forall i, j$) ja antisymmetrinen, jos $\mathbf{A}^T = -\mathbf{A}$ (eli $a_{ij} = -a_{ji}, \forall i, j$)

Jokainen matriisi voidaan esittää symmetrisen ja antisymmetrisen matriisin summana

$$\mathbf{A} = \mathbf{A}^+ + \mathbf{A}^- \quad \text{missä}$$

$$\mathbf{A}^+ = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) \text{ on symmetrinen, ja}$$

$$\mathbf{A}^- = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T) \text{ on antisymmetrinen.}$$

Jos esimerkiksi

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \text{niin} \\ \mathbf{A}^+ &= \frac{1}{2} \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2.5 \\ 2.5 & 4 \end{pmatrix} \\ \mathbf{A}^- &= \frac{1}{2} \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \right) = \begin{pmatrix} 0 & -0.5 \\ 0.5 & 0 \end{pmatrix}, \quad \text{ja} \\ \mathbf{A}^+ + \mathbf{A}^- &= \begin{pmatrix} 1 & 2.5 \\ 2.5 & 4 \end{pmatrix} + \begin{pmatrix} 0 & -0.5 \\ 0.5 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \mathbf{A}. \end{aligned}$$

4.3 Käänteismatriisi

Sec:MatInv

Reaalilukujen tapauksessa luvun a käänteisluku a^{-1} on sellainen luku, että $a^{-1} \cdot a = a \cdot a^{-1} = 1$ (missä $1 =$ kertolaskun neutraalialkio). Matriisin käänteismatriisilta vaaditaan täsmälleen sama ominaisuus. Asioita mutkistaa se, että kaikilla $\mathbf{O}$:sta poikkeavilla matriiseilla ei ole käänteismatriisia.

Määritelmä 77 Neliömatriisi $\mathbf{A}$ on säännöllinen, jos on olemassa käänteismatriisi $\mathbf{A}^{-1}$ siten, että $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

sikäisitteisyys

Lause 78 Jos neliömatriisi $\mathbf{A}$ on säännöllinen, niin

(1) käänteismatriisi $\mathbf{A}^{-1}$ on yksikäsitteinen ja

(2) matriisiyhtälöllä $\mathbf{A}\vec{x} = \vec{b}$ on yksikäsitteinen ratkaisu $\vec{x} = \mathbf{A}^{-1}\vec{b}$.

Todistus: Todistus perustuu käänteismatriisin määritelmään, matriisikertolaskun liitännäisyyteen ja matriisilaskujen osittelulakeihin (lause (70)(1,2)). th matkerto:kertolaskun ominaisuudet

(1) Olkoot $\mathbf{V}$ ja $\mathbf{W}$ kaksi $\mathbf{A}$:n käänteismatriisia. Silloin

$$\mathbf{V} = \mathbf{V}\mathbf{I} = \mathbf{V}(\mathbf{A}\mathbf{W}) = (\mathbf{V}\mathbf{A})\mathbf{W} = \mathbf{I}\mathbf{W} = \mathbf{W}$$

(2) $\mathbf{A}^{-1}\vec{b}$ on matriisiyhtälön $\mathbf{A}\vec{x} = \vec{b}$ ratkaisu, sillä

$$\mathbf{A}(\mathbf{A}^{-1}\vec{b}) = (\mathbf{A}\mathbf{A}^{-1})\vec{b} = \vec{b}$$

Olkoot $\vec{v}$ ja $\vec{w}$ kaksi matriisiyhtälön $\mathbf{A}\vec{x} = \vec{b}$ ratkaisua. Silloin

$$\begin{aligned} \vec{v} - \vec{w} &= \mathbf{A}^{-1}(\mathbf{A}(\vec{v} - \vec{w})) = \mathbf{A}^{-1}(\vec{b} - \vec{b}) = \mathbf{A}^{-1}\vec{0} = \vec{0} \\ \Leftrightarrow \vec{v} &= \vec{w} \end{aligned}$$

□

Edellisen lauseen väite (2) merkitsee sitä, että kun yhtälöryhmässä on yhtä monta yhtälöä kuin muuttujaa, niin ratkaisu saadaan helposti, jos kerroinmatriisin käänteismatriisi on tiedossa. Tulemme seuraavassa kappaleessa esittelemään keinon, jolla käänteismatriisi saadaan määritettyä. Myöhemmin tässä luvussa esitämme samaan tarkoitukseen vielä toisen keinon.

Kysymys siitä, milloin neliömatriisi on säännöllinen saa seuraavissa kapaleissa yksinkertaisen vastauksen. Seuraavassa luvussa yleistämme säännöllisyyden käsitteen matriiseille, joilla on rivejä enemmän kuin sarakkeita. Yleistys auttaa käsittelemään yhtälöryhmiä, joissa on yhtälöitä enemmän kuin muuttujia.

Kaikesta edellä sanotusta voimme todeta, että olemme nyt saavuttaneet erään tämän kurssin ydin-asian: ”matriisin säännöllisyys ja käänteismatriisin määrittäminen”.

Esimerkki 79 Tarkastellaan esimerkkinä lauseen ^{th_c3s3a:kmatr_yksikasitteisyys} (78) väitteeseen (2) yhtälöryhmää

$$\begin{cases} x + y = 4 \\ -2x + y = 1 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

Kerroinmatriisi ja sen käänteismatriisi ovat

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}, \quad \mathbf{A}^{-1} = \begin{pmatrix} 1/3 & -1/3 \\ 2/3 & 1/3 \end{pmatrix},$$

sillä

$$\begin{aligned} \mathbf{A}\mathbf{A}^{-1} &= \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1/3 & -1/3 \\ 2/3 & 1/3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{ja} \\ \mathbf{A}^{-1}\mathbf{A} &= \begin{pmatrix} 1/3 & -1/3 \\ 2/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Lauseen mukaan yhtälöryhmän ratkaisu saadaan kaavalla

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A}^{-1}\vec{b} = \begin{pmatrix} 1/3 & -1/3 \\ 2/3 & 1/3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Sijoittamalla alkuperäisen yhtälöryhmään huomaa, että ratkaisu on oikea.

Seuraavaksi listaamme joukon käänteismatriisin ominaisuuksia.

Lause 80 Olkoot $\mathbf{A}$ ja $\mathbf{B}$ säännöllisiä $n \times n$ -matriiseja. Silloin myös $\mathbf{A}^{-1}$, $\mathbf{A}^T$ ja $\mathbf{AB}$ ovat säännöllisiä ja

- (1) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- (2) $(\lambda\mathbf{A})^{-1} = \frac{1}{\lambda}\mathbf{A}^{-1}$
- (3) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ «« TÄRKEÄ!
- (4) $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$

Todistus:

$$\begin{aligned}
(1) \quad & \begin{cases} (\mathbf{A}^{-1})(\mathbf{A}) = \mathbf{I} \\ (\mathbf{A})(\mathbf{A}^{-1}) = \mathbf{I} \end{cases} \\
(2) \quad & \begin{cases} (\lambda \mathbf{A})(\lambda^{-1} \mathbf{A}^{-1}) = \lambda \cdot \lambda^{-1} \mathbf{A} \mathbf{A}^{-1} = \mathbf{I} \\ (\lambda^{-1} \mathbf{A}^{-1})(\lambda \mathbf{A}) = \lambda^{-1} \cdot \lambda \mathbf{A}^{-1} \mathbf{A} = \mathbf{I} \end{cases} \\
(3) \quad & \begin{cases} (\mathbf{A}\mathbf{B})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{B}\mathbf{B}^{-1})\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \\ (\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{A}\mathbf{B}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I} \end{cases} \\
(3) \quad & \begin{cases} (\mathbf{A}^T)(\mathbf{A}^{-1})^T = (\mathbf{A}^{-1}\mathbf{A})^T = \mathbf{I} \\ (\mathbf{A}^{-1})^T(\mathbf{A}^T) = (\mathbf{A}\mathbf{A}^{-1})^T = \mathbf{I} \end{cases}
\end{aligned}$$

□

Lause 81 Jos matriisi $\mathbf{A}$ on säännöllinen, niin

$$\begin{aligned}
(1) \quad & (\mathbf{AB} = \mathbf{AC}) \Rightarrow (\mathbf{B} = \mathbf{C}) \\
(2) \quad & (\mathbf{BA} = \mathbf{CA}) \Rightarrow (\mathbf{B} = \mathbf{C}) \\
(3) \quad & (\mathbf{AB} = \mathbf{O}) \Rightarrow (\mathbf{B} = \mathbf{O}) \\
(4) \quad & (\mathbf{BA} = \mathbf{O}) \Rightarrow (\mathbf{B} = \mathbf{O})
\end{aligned}$$

Todistus: (1) $\mathbf{B} = \mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{AC} = \mathbf{C}$, (2) vastaavasti ja (3) seuraa (1):stä, kun $\mathbf{C} = \mathbf{O}$. (4) vastaavasti. □

Siis esimerkiksi (71) sivulla 78 mainitut ongelmat (2) ja (3) eivät koske säännöllisiä neliömatriiseja. Esimerkin (71) yhteydessä on tapana sanoa, että ”matriisilla ei saa jakaa, mutta käänteismatriisilla saa kertoa (jos käänteismatriisi on olemassa).”

4.4 Matriisit ja vektorit Octavessa (kesken)

Sec:MVOct

Kappale
kesken

Octavea on jo käytetty lineaaristen yhtälöryhmien ratkaisemisessa. Nyt käytämme samaa ohjelmaa matriisi- ja vektori-laskujen laskemiseen.

Matriisi

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

voidaan esitellä Octavelle kahdella tavalla:

Tapa 1:

```
octave:2> A = [2 -4; 1 3]
A =

     2    -4
     1     3
```

Muuttujalle A sijoitetaan matriisi. Matriisimerkintä alkaa (alku)hahasululla ja päättyy (loppu)hahasulkuun. Matriisin alkiot luetellaan välilyönnein (tai pilkuin) erotettuina ja rivit erotetaan puolipisteellä. Jos komento päättyy puolipisteeseen, niin Octave ei kaiuta matriisia näytölle, ja jos puolipiste puuttuu komennon perästä, niin Octave nielaisee komenon kaiuttamatta mitään. (Kokeile.) Kummassakin tapauksessa muuttuja A viittaa komennon jälkeen siihen edellä sijoitettuun matriisiin.

Tapa 2:

```
octave:3> A = [2 -4;
>              1  3]
A =

     2    -4
     1     3
```

Toisessa tavassa matriisi kirjoitetaan rivi kerrallaan. Octave-tulkki tulkitsee ensimmäisen komentorivin tarkoittavan matriisin määrittelyä, mutta toteaa komennon olevan vielä kesken, ja odottaa seuraavan rivin täydentävän komennon. Rivejä voi olla enemmänkin kuin kaksi. Komento tulee valmiiksi, kun loppu-hahasulku lopettaa matriisin.

4.5 Esimerkki Markovin ketjusta, 'FlyBike'

Sec:FlyBike

laFlyBike

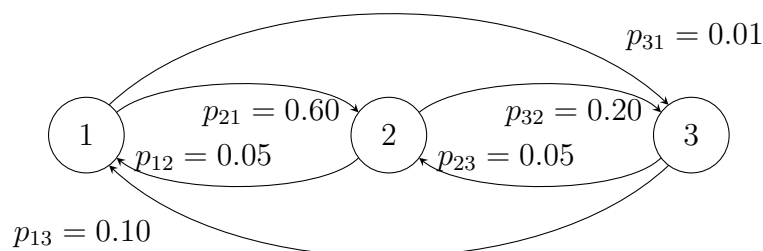
Esimerkki 82 *Suomalainen 'FlyBike'-talli ottaa vuosittain osaa Trans-Silvanian ympäriajoon kymmenellä pyöräilijällä. Tallin manageri haluaa arvion tulevien vuosien tuotoista taloussuunnittelua varten. Manageri luokittelee tallinsa urheilijat kolmeen ryhmään "aloittelijat", "kokeneet" ja "tähdet". Talli tekee yhdessä jokaisen urheilijan kanssa mainos ja pr-sopimuksen. Keskimääräiset tallin tuotot eri ryhmissä ovat seuraavan taulukon mukaiset. Talukossa on myös eri ryhmiin kuuluvien urheilijoiden lukumäärät tällä kaudella.*

ryhmä	tuotto	lkm nyt
(1) aloittelijat	$r_1 = 300$	$n_1 = 6$
(2) kokeneet	$r_2 = 2000$	$n_2 = 3$
(3) tähdet	$r_3 = 10000$	$n_3 = 1$

Muodostetaan taulukon sarakkeista vektorit $\vec{r} = (300, 2000, 10000)^T$ ja $\vec{n} = (6, 3, 1)^T$. Tallin tämän kauden tuotto on

$$R_0 = \vec{n} \cdot \vec{r} = 6 \cdot 300 + 3 \cdot 2000 + 1 \cdot 10000 = 17\,800.$$

Tallin harjoitusohjelma sekä valmennus- ja huoltotuki takaa sen, että urheilijat kehittyvät uransa aikana. Siirtymätodennäköisyydet urheilijan siirtymiselle ryhmästä toiseen yhden kauden kuluessa ovat seuraavat:



Tee managerille arvio lähikausien tuotoista.

Ratkaisu:

Kaaviossa kirjaimella p_{ij} merkitään todennäköisyyttä sille, että tällä kaudella ryhmässä j polkeva urheilija, kilpailee seuraavalla kaudella ryhmässä i . Siirtymät $(1 \rightarrow 2)$ ja $(2 \rightarrow 3)$ ovat luonnollisia valmennuksesta seuraavaa urakehitystä. Siirtymät $(2 \rightarrow 1)$, $(3 \rightarrow 2)$ ja $(3 \rightarrow 1)$ ovat joko loukkaantumisesta johtuvia taantumia, tai seurausta siitä, että kokenut urheilija tai tähti lopettaa uransa ja hänet korvataan kokemattommalla.

Saattaa tuntua oudolta, että siirtymään $(1 \rightarrow 2)$ liitetään siirtymätodennäköisyys p_{21} . Tuntuu kuin indeksit olisivat väärinpäin. Pian huomaat, että indeksit kannattaa merkitä juuri näin!

p_{kk} on todennäköisyys sille, että urheilija joka on nyt ryhmässä k , jatkaa seuraavankin kauden samassa ryhmässä. Alkeistodennäköisyyksien summa on yksi, joten

$$p_{11} = 1 - p_{21} - p_{31} = 1 - 0.60 - 0.01 = 0.39$$

$$p_{22} = 1 - p_{12} - p_{32} = 1 - 0.05 - 0.20 = 0.75$$

$$p_{33} = 1 - p_{13} - p_{23} = 1 - 0.10 - 0.05 = 0.85$$

Tiedämme siis urheilijoiden lukumäärät eri ryhmissä nyt ($\vec{n}(0)$) sekä siirtymätodennäköisyydet. Näiden avulla voimme laskea odotusarvot urheilijoiden lukumäärille eri ryhmissä ensi kaudella ($\vec{n}(1)$) kaavoilla

$$n_1(1) = p_{11}n_1(0) + p_{12}n_2(0) + p_{13}n_3(0)$$

$$n_2(1) = p_{21}n_1(0) + p_{22}n_2(0) + p_{23}n_3(0)$$

$$n_3(1) = p_{31}n_1(0) + p_{32}n_2(0) + p_{33}n_3(0)$$

eli

$$\vec{n}(1) = \mathbf{P}\vec{n}(0),$$

missä $\mathbf{P}$ on matriisi

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} 0.39 & 0.05 & 0.10 \\ 0.60 & 0.75 & 0.05 \\ 0.01 & 0.20 & 0.85 \end{pmatrix}.$$

Odotusarvo k :nnen kauden ryhmien suuruuksille on

$$\vec{n}(k) = \mathbf{P}\vec{n}(k-1) = \mathbf{P}^k\vec{n}(0).$$

Teemme pienen python-ohjelman nähdäksemme seuraavan kymmenen kauden odotusarvot. Kirjoitamme seuraavat komentorivit tekstitiedostoon "laFlyBike.py":

```
from numpy import *
P = matrix([[0.39, 0.10, 0.20],
            [0.60, 0.75, 0.10],
            [0.01, 0.10, 0.85]],float);
n = matrix([[6],[3],[1]],float)
#
print('\n k    n1(k)    n2(k)    n3(k)')
for k in range(11):
    nk = (P**k)*n
    print('{0:3d}    {1:5.3f}    {2:5.3f}    {3:5.3f}'.
          format(k,nk[0,0],nk[1,0],nk[2,0]))
```

Komento 'python laFlyBike.py' tuottaa tulosteen:

k	n1(k)	n2(k)	n3(k)
0	6.000	3.000	1.000
1	2.840	5.950	1.210
2	1.945	6.287	1.652
3	1.718	6.048	2.052
4	1.685	5.771	2.366
5	1.708	5.576	2.605
6	1.745	5.467	2.789
7	1.785	5.426	2.935
8	1.826	5.434	3.055
9	1.867	5.477	3.159
10	1.907	5.543	3.251

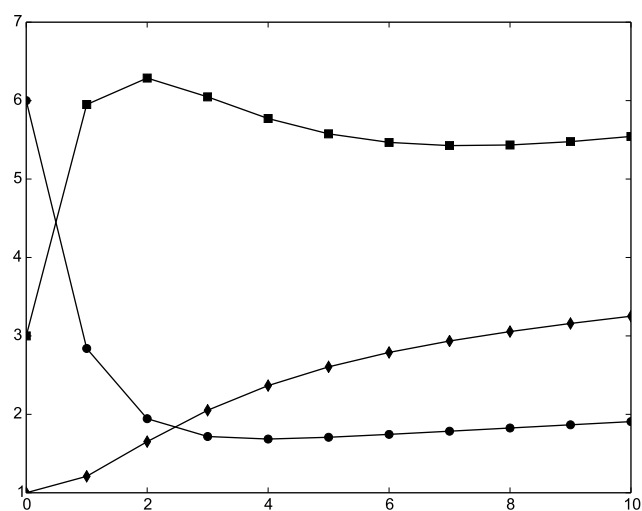
Pythonilla saamme myös helposti graafisen esityksen eri ryhmien kokojen kehitykselle. Tiedostoon ”laFlyBike2.py” kirjoitamme ohjelman:

```
from numpy import *
import matplotlib.pyplot as plt
#
P = matrix([[0.39, 0.10, 0.20],
            [0.60, 0.75, 0.10],
            [0.01, 0.10, 0.85]],float);
n = matrix([[6],[3],[1]],float)
#
x = zeros((11,1));
n1 = zeros((11,1));
n2 = zeros((11,1));
n3 = zeros((11,1));
for k in range(11):
    nk = (P**k)*n;
    n1[k] = nk[0];
    n2[k] = nk[1];
    n3[k] = nk[2];
    x[k] =k;
#
plt.plot(x,n1,'k-o',x,n2,'k-s',x,n3,'k-d');
plt.show()
```

Ohjelman piirtämä graafi on kuvassa [LAPc3s1fig1](#) 22.

Jos mallin siirtymätodennäköisyydet on arvioitu oikein ja ne eivät myöskään muutu tulevien kausien aikana, niin joukkueen urheilijoista tulee keskimäärin kokeneempia ja heidän tulonsakin keskimäärin kasvavat. Tallin tulot ovat siis kasvussa, mutta miten paljon?

Lisätään vielä samaan kuvaan tallin saama tulo kausittain. Piirrämme kaksi erillistä



LAPc3s1fig1

Kuva 22. Ryhmien koko tulevina kausina.

koordinaatistoa samaan kuvaan ja lisäämme joitakin otsikkotekstejä. Seuraavassa tiedosto ”laFlyBike3.py” ja sen piirtämä kuva:

```
from numpy import *
import matplotlib.pyplot as plt
#
P = matrix([[0.39, 0.10, 0.20],[0.60, 0.75, 0.10],[0.01,
0.10, 0.85]],float);
r = matrix([[300,2000,10000]]);
n = matrix([[6],[3],[1]],float)
#
x = zeros((11,1));
n1 = zeros((11,1));
n2 = zeros((11,1));
n3 = zeros((11,1));
tulo = zeros((11,1));
#
for k in range(11):
    nk = (P**k)*n;
    n1[k] = nk[0];
    n2[k] = nk[1];
    n3[k] = nk[2];
    x[k] =k;
    tulo[k] = r*nk;
#
fig = plt.figure();

kuval = fig.add_subplot(211);
```

```

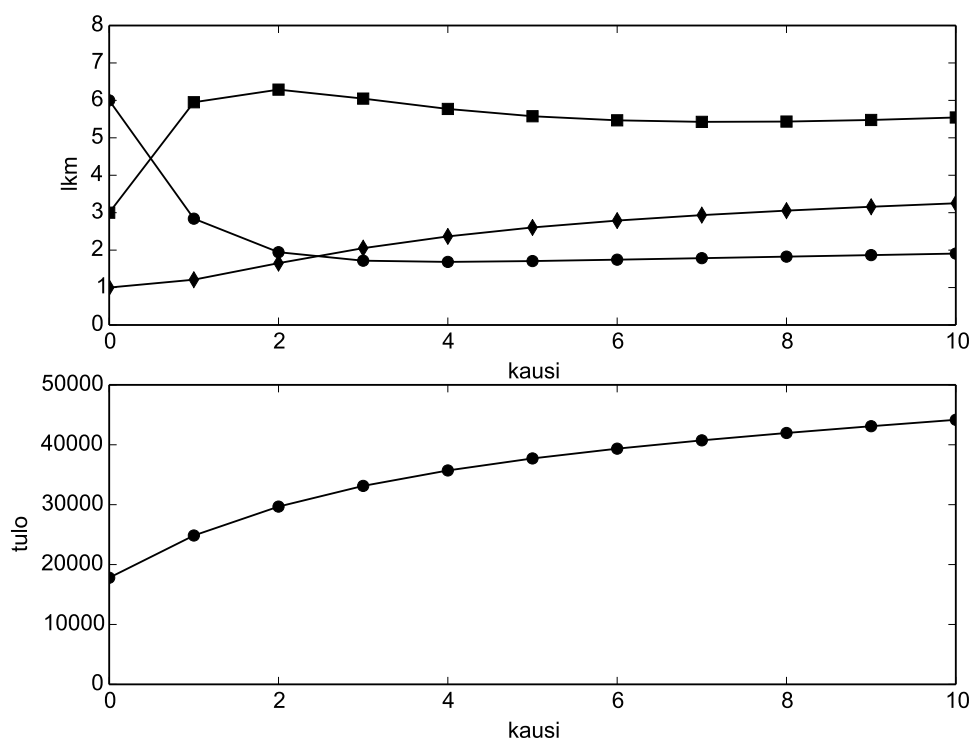
kuva1.plot(x,n1,'k-o',x,n2,'k-s',x,n3,'k-d');
kuva1.set_xlabel('kausi');
kuva1.set_ylabel('lkm');
kuva1.axis([0,10,0,8]);

kuva2 = fig.add_subplot(212);
kuva2.plot(x,tulo,'k-o');
kuva2.set_xlabel('kausi');
kuva2.set_ylabel('tulo');
kuva2.axis([0,10,0,50000]);

plt.show()

```

Ohjelman piirtämä graafi on kuvassa [LAPc3s1fig2](#) 23.



Kuva 23. Ryhmien koko ja tallin tulo tulevina kausina.

4.6 Matriisin kääntäminen rivioperaatioilla

InvGaussJordan

Seuraavassa esitetty Gaussin algoritmi perustuu tuttuihin rivioperaatioihin. Teoria on esitetty mahdollisimman lyhyesti ja esimerkit saavat selvittää asian. Algoritmi esite-

tään vasta lopuksi, koska se tiivistää esimerkit.

Olkoon $\mathbf{A}$ seuraavassa $n \times n$ -matriisi. Rivioperaatiot voidaan muuntaa matriisilla (rivioperaattorilla) kertomiseksi seuraavalla tavalla:

1) kahden rivin vaihto

Kun matriisi $\mathbf{A}$ kerrotaan matriisilla (rivioperaattorilla)

$$\mathbf{E}_{p;q}^{\text{vaihda}} = (e_{ij}), \text{ missä } \begin{cases} e_{pq} = e_{qp} = 1 \\ e_{ii} = 1, \text{ kun } i \neq p \text{ ja } i \neq q \\ e_{ij} = 0, \text{ muulloin} \end{cases},$$

niin rivit p ja q vaihtavat paikkaa.

Esim:

$$\mathbf{E}_{1;2}^{\text{vaihda}} \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} = \begin{pmatrix} 21 & 22 & 23 \\ 11 & 12 & 13 \\ 31 & 32 & 33 \end{pmatrix}$$

2) rivin kertominen luvulla λ

Kun matriisi $\mathbf{A}$ kerrotaan matriisilla (rivioperaattorilla)

$$\mathbf{E}_{p;\lambda}^{\text{kerro}} = (e_{ij}), \text{ missä } \begin{cases} e_{pp} = \lambda \\ e_{ii} = 1, \text{ kun } i \neq p \\ e_{ij} = 0, \text{ muulloin} \end{cases},$$

niin rivi p tulee kerrotuksi luvulla λ .

Esim:

$$\mathbf{E}_{2;10}^{\text{kerro}} \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} = \begin{pmatrix} 11 & 12 & 13 \\ 210 & 220 & 230 \\ 31 & 32 & 33 \end{pmatrix}$$

3) rivin p lisääminen luvulla λ kerrottuna riviin q

Kun matriisi $\mathbf{A}$ kerrotaan matriisilla (rivioperaattorilla)

$$\mathbf{E}_{p,q;\lambda}^{\text{lisaa}} = (e_{ij}), \text{ missä } \begin{cases} e_{ii} = 1, \forall i \\ e_{qp} = \lambda \\ e_{ij} = 0, \text{ muulloin} \end{cases},$$

niin rivi p tulee lisätyksi luvulla λ kerrottuna riviin q .

Esim:

$$\mathbf{E}_{1,2;100}^{\text{lisaa}} \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 100 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} = \begin{pmatrix} 11 & 12 & 13 \\ 1121 & 1222 & 1323 \\ 31 & 32 & 33 \end{pmatrix}$$

Kaikki edellä esiintyneet rivioperaattorit ovat säännöllisiä sillä helposti nähdään että

$$\left(\mathbf{E}_{p,q}^{\text{vaihto}}\right)^{-1} = \mathbf{E}_{p,q}^{\text{vaihto}}, \quad \left(\mathbf{E}_{p;\lambda}^{\text{kerro}}\right)^{-1} = \mathbf{E}_{p;1/\lambda}^{\text{kerro}} \quad \text{ja} \quad \left(\mathbf{E}_{p,q;\lambda}^{\text{lisaa}}\right)^{-1} = \mathbf{E}_{p,q;-\lambda}^{\text{lisaa}},$$

Lause 83 *Olkoon $\mathbf{A}$ säännöllinen $n \times n$ -matriisi. Jos on olemassa $n \times n$ -matriisi $\mathbf{B}$, jolle $\mathbf{BA} = \mathbf{I}$, niin $\mathbf{A}^{-1} = \mathbf{B}$.*

Todistus: $\mathbf{AB} = \mathbf{ABAA}^{-1} = \mathbf{AIA}^{-1} = \mathbf{I}$ □

Lause 84 *Olkoon $\mathbf{A}$ säännöllinen $n \times n$ -matriisi. Jos $\mathbf{A}$ voidaan muuttaa jonolla rivio-
peraatioita yksikkömatriisiksi $\mathbf{I}$, niin kääntematriisi $\mathbf{A}^{-1}$ saadaan tekemällä samat
rivioperaatiot yksikkömatriisille $\mathbf{I}$.*

Todistus: Olkoon $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$ jono rivioperaattoreita siten, että $\mathbf{E}_n \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{I}$.
Silloin edellisen lauseen mukaan

$$\mathbf{A}^{-1} = \mathbf{E}_n \cdots \mathbf{E}_2 \mathbf{E}_1 = \mathbf{E}_n \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{I}$$

□

Esimerkki 85 *Määritetään kääntematriisi matriisille*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -2 \\ -1 & 0 & 4 \\ -2 & -8 & 1 \end{pmatrix}.$$

Kirjoitetaan ensin kaavio $(\mathbf{A}|\mathbf{I})$, jolle ryhdymme tekemään rivioperaatioita. Tällä var-

mistamme, että $\mathbf{I}$:hin kohdistuu täsmälleen samat rivioperaatiot kuin $\mathbf{A}$:han.

$$\begin{aligned}
 & (\mathbf{A}|\mathbf{I}) \\
 & \sim \left(\begin{array}{ccc|ccc} [1] & 2 & -2 & 1 & 0 & 0 \\ -1 & 0 & 4 & 0 & 1 & 0 \\ -2 & -8 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} +1 \\ \leftarrow \\ \leftarrow \end{array} \\
 & \sim \left(\begin{array}{ccc|ccc} (1) & 2 & -2 & 1 & 0 & 0 \\ 0 & [2] & 2 & 1 & 1 & 0 \\ 0 & -4 & -3 & 2 & 0 & 1 \end{array} \right) \begin{array}{l} \leftarrow \\ -1 \\ \leftarrow \end{array} \\
 & \sim \left(\begin{array}{ccc|ccc} (1) & 0 & -4 & 0 & -1 & 0 \\ 0 & (2) & 2 & 1 & 1 & 0 \\ 0 & 0 & [1] & 4 & 2 & 1 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow \\ -2 \end{array} \\
 & \sim \left(\begin{array}{ccc|ccc} (1) & 0 & 0 & 16 & 7 & 4 \\ 0 & (2) & 0 & -7 & -3 & -2 \\ 0 & 0 & [1] & 4 & 2 & 1 \end{array} \right) :2 \\
 & \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 16 & 7 & 4 \\ 0 & 1 & 0 & -7/2 & -3/2 & -1 \\ 0 & 0 & 1 & 4 & 2 & 1 \end{array} \right) \\
 & \Rightarrow \mathbf{A}^{-1} = \begin{pmatrix} 16 & 7 & 4 \\ -7/2 & -3/2 & -1 \\ 4 & 2 & 1 \end{pmatrix}
 \end{aligned}$$

Tarkistus

$$\begin{aligned}
 \mathbf{A}\mathbf{A}^{-1} &= \begin{pmatrix} 1 & 2 & -2 \\ -1 & 0 & 4 \\ -2 & -8 & 1 \end{pmatrix} \begin{pmatrix} 16 & 7 & 4 \\ -7/2 & -3/2 & -1 \\ 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \mathbf{A}^{-1}\mathbf{A} &= \begin{pmatrix} 16 & 7 & 4 \\ -7/2 & -3/2 & -1 \\ 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -2 \\ -1 & 0 & 4 \\ -2 & -8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

Esimerkki 86 Määritetään käänteismatriisi matriisille

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

$$\begin{aligned}
& (\mathbf{A}|\mathbf{I}) \\
& \sim \left(\begin{array}{ccc|ccc} [1] & -1 & -2 & 1 & 0 & 0 \\ -1 & 3 & 4 & 0 & 1 & 0 \\ -2 & -1 & 1 & 0 & 0 & 1 \end{array} \right) \left| \begin{array}{l} +1 \\ \leftarrow \\ \leftarrow \end{array} \right| \begin{array}{l} +2 \\ \\ \leftarrow \end{array} \\
& \sim \left(\begin{array}{ccc|ccc} (1) & -1 & -2 & 1 & 0 & 0 \\ 0 & [2] & 2 & 1 & 1 & 0 \\ 0 & -3 & -3 & 2 & 0 & 1 \end{array} \right) \left| \begin{array}{l} \leftarrow \\ +1/2 \\ \leftarrow \end{array} \right| \begin{array}{l} \\ +3/2 \\ \leftarrow \end{array} \\
& \sim \left(\begin{array}{ccc|ccc} (1) & 0 & -1 & 3/2 & 1/2 & 0 \\ 0 & (2) & 2 & 1 & 1 & 0 \\ 0 & 0 & [0] & 7/2 & 3/2 & 1 \end{array} \right)
\end{aligned}$$

Viimeinen pivotointi epäonnistuu, jonka tulkitsemme merkiksi siitä, että matriisilla ei ole käänteismatriisia.

Edellä kuvattu menetelmä johtaa helposti hankaliin murtolukulaskuihin, mutta se on erittäin suoraviivainen ja sellaisena helppo omaksua.

Gaussin algoritmi (käänteismatriisin määrittämiseksi):

- (0) Aseta $k = 1$, $\mathbf{B}(1) = (\mathbf{A}|\mathbf{I})$.
- (1) Otetaan tukialkioksi (pivot element) alkio $b(k)_{kk}$.
- (2) Jos tukialkio on 0 ja tukialkion sarakkeessa tukialkion alapuolella on nollasta poikkeava alkio $b(k)_{rk}$, $r > k$, vaihdetaan rivit r ja k keskenään. Jos tukialkio on 0 ja sen alapuolelta ei löydy nollasta poikkeavaa alkioita, lopetetaan ja todetaan ettei käänteismatriisia ole olemassa.
- (3) Jaetaan tukialkion rivi tukialkiolla. (Eli tehdään tukialkio 1:ksi.)
- (4) Lisätään tukialkion rivi muihin riveihin sopivalla luvulla kerrottuna siten, että tukialkion ylä- ja alapuolella olevat alkion muuttuvat nolliksi.
- (5) Jos $k < n$, niin kasvatetaan k :ta yhdellä ja siirrytään kohtaan (1).
- (6) $\mathbf{A}^{-1}$ on saadun $n \times 2n$ -matriisin $\mathbf{B}(k)$ oikea puolikas.

4.7 Yhtälöryhmän ratkaisun herkkyyys

Yhtälöryhmän

$$\begin{cases} 3x + y - z = 2 \\ x + 2y + z = 5 \\ -2x + 2y + 2z = 4 \end{cases} \\ \Leftrightarrow \underbrace{\begin{pmatrix} 3 & 1 & -1 \\ 1 & 2 & 1 \\ -2 & 2 & 2 \end{pmatrix}}_{=\mathbf{A}} \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{=\vec{x}_1} = \underbrace{\begin{pmatrix} 2 \\ 5 \\ 4 \end{pmatrix}}_{=\vec{b}_1}$$

ratkaisu on

$$\vec{x}_1 = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{A}^{-1} \vec{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Jos jonkun yhtälön RHS muuttuu, niin myös ratkaisu muuttuu.

Jos kolmannen yhtälön RHS kasvaa arvosta 4 arvoon 5, niin yhtälöryhmän ratkaisu muuttuu. Uusi ratkaisu on

$$\begin{cases} 3x + y - z = 2 \\ x + 2y + z = 5 \\ -2x + 2y + 2z = 5 \end{cases} \\ \Leftrightarrow \underbrace{\begin{pmatrix} 3 & 1 & -1 \\ 1 & 2 & 1 \\ -2 & 2 & 2 \end{pmatrix}}_{=\mathbf{A}} \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{=\vec{x}_2} = \underbrace{\begin{pmatrix} 2 \\ 5 \\ 5 \end{pmatrix}}_{=\vec{b}_2}$$

ratkaisu on

$$\vec{x}_2 = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{A}^{-1} \vec{b}_2 = \begin{pmatrix} 0.25 \\ 2 \\ 0.75 \end{pmatrix}$$

Muutos ratkaisussa on

$$\Delta \vec{x} = \vec{x}_2 - \vec{x}_1 = \begin{pmatrix} 0.25 \\ 2 \\ 0.75 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -0.75 \\ 1 \\ -1.25 \end{pmatrix} = \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}$$

Kun siis yhdenkin yhtälön RHS muuttui, niin jokaisen muuttujan arvo ratkaisussa muuttui!

Jos pienikin muutos yhtälöryhmän RHS-luvuissa saa aikaan suuria muutoksia muuttujien ratkaisu-arvoihin sanomme, että yhtälöryhmä on herkkä. Herkkyydelle on olemassa mitta (kerroinmatriisin kuntoluku), mutta emme opettele sitä vielä nyt.

Toinen esimerkki

Edellisessä esimerkissä muutos RHS:ssä oli melko suuri. Seuraavaksi katsomme vastaavan esimerkin, jossa RHS:n muutos on pienempi (realistisempi, koska pieniä muutoksia on aina).

Olkoon yhtälöryhmän kertoimet samat kuin edellisessä esimerkissä. Alkutilanteen RHS on

$$\vec{b}_1 = \begin{pmatrix} 2 \\ 5 \\ 4 \end{pmatrix}$$

ja lopputilanteen RHS olkoon

$$\vec{b}_3 = \begin{pmatrix} 2 \\ 5 \\ 4.1 \end{pmatrix} = \vec{b}_1 + \begin{pmatrix} 0 \\ 0 \\ 0.1 \end{pmatrix}$$

$$\vec{b}_1 = \begin{pmatrix} 2 \\ 5 \\ 4 \end{pmatrix}, \quad \Delta\vec{b} = \begin{pmatrix} 0 \\ 0 \\ 0.1 \end{pmatrix} = 0.1 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{b}_3 = \vec{b}_1 + \Delta\vec{b} = \begin{pmatrix} 2 \\ 5 \\ 4.1 \end{pmatrix}$$

Laskemme vastaavan muutoksen ratkaisuvektorissa $\vec{x}$:

$$\begin{aligned} \Delta\vec{x} &= \vec{x}_3 - \vec{x}_1 = \mathbf{A}^{-1}\vec{b}_3 - \mathbf{A}^{-1}\vec{b}_1 \\ &= \mathbf{A}^{-1}(\vec{b}_3 - \vec{b}_1) = \mathbf{A}^{-1}\Delta\vec{b} \quad \longrightarrow \boxed{\Delta\vec{x} = \mathbf{A}^{-1}\Delta\vec{b}} \\ &= 0.1 \cdot \begin{pmatrix} -0.5 & 1 & -0.75 \\ 1 & -1 & 1 \\ -1.5 & 2 & -1.25 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.075 \\ 0.1 \\ -0.125 \end{pmatrix} \end{aligned}$$

Esimerkkien yhteenveto

Yksinkertainen periaate: Tarkastellaan yhtälöryhmää $\mathbf{A}\vec{x} = \vec{b}$. Jos k :nnen yhtälön RHS kasvaa yhdellä, niin ratkaisuvektori muuttuu määrällä, joka on kerroinmatriisin käänteismatriisin k :nnes sarake.

Merkitään $\mathbf{A}$:n käänteismatriisin k :tta saraketta vektorilla $\vec{u}_k = (u_{1k} \ u_{2k} \ \dots \ u_{nk})^T$

Täsmällisempi periaate: Tarkastellaan yhtälöryhmää $\mathbf{A}\vec{x} = \vec{b}$. Jos yhtälöryhmän RHS kasvaa määrällä $\Delta\vec{b} = (\Delta b_1 \ \Delta b_2 \ \dots \ \Delta b_n)^T$, niin ratkaisuvektori muuttuu määrällä

$$\Delta\vec{x} = \mathbf{A}^{-1}\Delta\vec{b} = \Delta b_1 \vec{u}_1 + \Delta b_2 \vec{u}_2 + \dots \Delta b_n \vec{u}_n$$

5 DETERMINANTIT

Ch: Determ

5.1 Determinantti

Sec: DetDef

Tämä kappale jakautuu kolmeen alakappaleeseen. Ensimmäisessä alakappaleessa määrittelemme kaksi- ja kolmiriviset determinantit. Nämä kaavat löytyvät taulukkokirjoista. Toisessa alakappaleessa yleistämme determinantin isommille kaavioille ja toteamme koko joukon ominaisuuksia determinantille. Ominaisuuksien perustelut löytyvät lisämateriaalista.

Tekniikan opiskelussa usein ajatellaan, että kaksi- ja kolmiriviset determinantit riittävät. Tämä ei kuitenkaan nyt riitä. Esimerkiksi lineaarisen optimoinnin tarpeita varten opiskelijan tulee opetella myös käsittelemään isojakin lukukaavioita.

5.1.1 Kaksi- ja kolmirivinen determinantti

c_kaksijakolme

def_kaksi_det

Määritelmä 87 (*Kaksirivinen determinantti, [määritelmä löytyy taulukkokirjoista]*) 2×2 -matriisin $\mathbf{A} = (a_{ij})$ determinantti on $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$.

Kaksirivisen matriisin determinantille käytämme myös merkintää

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (5.1) \quad \text{kaksirivinen_det}$$

Huomaa, että merkintä kaksirivinen_det (5.1) tarkoittaa siis reaalilukua $a_{11}a_{22} - a_{12}a_{21}$ eikä mitään lukukaaviota.

Esimerkki 88

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2 \quad (5.2)$$

$$\begin{vmatrix} 3 & -4 \\ 5 & -6 \end{vmatrix} = 3 \cdot (-6) - (-4) \cdot 5 = -18 + 20 = 2 \quad (5.3)$$

Huomaa, että lausekkeessa $3 \cdot (-6) - (-4) \cdot 5$ keskimäinen miinusmerkki tulee määritelmästä ja toiset kaksi miinus-merkkiä ovat matriisialkioiden merkkejä. Merkkivirheitä voi estää vain olemalla huolellinen.

Esimerkki 89 Tarkastellaan esimerkkinä yhtälöparia

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

Ratkaistaan yhtälöpari tavalliseen tapaan

$$\left(\begin{array}{ccc|c} a_{11} & a_{12} & b_1 & -a_{21}/a_{11} \\ a_{21} & a_{22} & b_2 & \leftarrow \end{array} \right) \cdot a_{11} \sim \left(\begin{array}{ccc|c} a_{11} & a_{12} & b_1 & \\ & \det(\mathbf{A}) & a_{11}b_2 - a_{21}b_1 & \end{array} \right)$$

Siis yhtälöryhmällä on yksikäsitteinen ratkaisu, jos ja vain jos $\det(\mathbf{A}) \neq 0$.

def_kolme_det

Määritelmä 90 (Kolmirivinen determinantti, [määritelmä löytyy taulukkokirjoista])
 3×3 -matriisin $\mathbf{A} = (a_{ij})$ determinantti on

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

Esimerkki 91

$$\begin{aligned} \begin{vmatrix} 1 & -2 & -1 \\ 3 & 4 & -2 \\ 0 & 1 & 5 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 4 & -2 \\ 1 & 5 \end{vmatrix} - (-2) \cdot \begin{vmatrix} 3 & -2 \\ 0 & 5 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 3 & 4 \\ 0 & 1 \end{vmatrix} \\ &= 1 \cdot (4 \cdot 5 - (-2) \cdot 1) + 2 \cdot (3 \cdot 5 - (-2) \cdot 0) - 1 \cdot (3 \cdot 1 - 4 \cdot 0) \\ &= (20 + 2) + 2 \cdot (15 - 0) - (3 - 0) = 49 \end{aligned}$$

5.1.2 Isot determinantit

det_isot

Määritelmä 92 $n \times n$ -matriisin $\mathbf{A} = (a_{ij})$ determinantti on luku

$$\det \mathbf{A} = \sum_{k=1}^n (-1)^{1+k} a_{1k} \text{Det} \mathbf{A}_{1k},$$

missä $\mathbf{A}_{1k}$ on matriisin $\mathbf{A}$ alimatriisi, joka saadaan, kun $\mathbf{A}$:n ensimmäinen rivi ja k :s sarake poistetaan. Käytämme determinantille myös merkintää

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Ensimmäiseksi tarkistamme, ettei uusi määritelmä ole ristiriidassa aiemmin antamiemme kaksi- ja kolmirivisten determinantin kaavojen kanssa.

Esimerkki 93 (1) Sovelletaan ensin uutta määritelmää 3-riviseen kaavioon.

$$\begin{aligned} & \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= (-1)^{1+1}a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{1+2}a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{1+3}a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \end{aligned}$$

Uusi määritelmä antaa siis täsmälleen saman lausekkeen kuin aiempi määritelmä (90). def_kolme_det

(2) Toiseksi tutkimme kaavaa 2-rivisen kaavion yhteydessä. Sovimme, että yksirivinen determinantti on sama kuin kaavion ainoa alkio. (Siis $\det(a) = |a| = a$. Huomaa, että merkintä $|a|$ ei nyt tarkoita luvun itseisarvoa, vaan 1×1 matriisin (a) determinanttia.) Kun näin on sovittu, voidaan uutta määritelmää soveltaa myös kaksiriviseen determinanttiin.

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= (-1)^{1+1}a_{11}|a_{22}| + (-1)^{1+2}a_{12}|a_{21}| \\ &= a_{11}a_{22} - a_{21}a_{12} \end{aligned}$$

Taas oikein. Siis kaksirivisen determinantin tapauksessa saadaan sama lauseke kuin määritelmässä (87) sivulla 97. def_kaksidetkaksi_det

Seuraavassa esitellyt determinantin ominaisuudet ovat aiheen ymmärtämisen kannalta hyvin tärkeitä. Ne tulee ymmärtää hyvin. Siksi niistä esitetään sekä sanallinen (mahdollisimman havainnollinen), että kaavanomainen (mahdollisimman täsmällinen) muoto. Ominaisuudet on esitetty lauseen muodossa. Lauseiden todistukset on esitetty lisämateriaalissa.

Jotta kaavat olisivat helpot muodostaa, sovimme epästandardin, vain tässä yhteydessä käytettävän merkinnän.

Olkoon tarkasteltavina $n \times n$ -matriisi $\mathbf{A} = (a_{ij})$ ja n -alkioinen pystyvektori $\vec{u} = (u_1 \ u_2 \ \dots \ u_n)^T$. Sovimme, että tässä kappaleessa (ja myöhemmin Cramerin kaavojen yhteydessä) merkintä

$$\mathbf{A}[k : \vec{u}]$$

tarkoittaa matriisia, joka saadaan, kun $\mathbf{A}$:n k :s sarake korvataan pystyvektorilla $\vec{u}$. Jos siis esimerkiksi

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & 4 \\ 5 & -3 & 1 \end{pmatrix}, \quad \text{ja} \quad \vec{u} = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix},$$

niin

$$\mathbf{A}[2 : \vec{u}] = \begin{pmatrix} 2 & 7 & 3 \\ 1 & 8 & 4 \\ 5 & 9 & 1 \end{pmatrix}$$

Lisäksi käytämme paljon käytettyä merkintää, jossa matriisin $\mathbf{A} = (a_{ij})$ k :tta saraketta merkitään pystyvektorilla

$$\vec{a}_{\bullet k} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix}$$

th_det1 **Lause 94** Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.

(Sanallinen muoto:) Jos neliömatriisin $\mathbf{A}$ jonkin sarakkeen kaikki alkiot ovat nollia, niin matriisin determinanti on nolla.

(Kaavana:)

$$\vec{a}_{\bullet k} = \vec{0} \Rightarrow \det(\mathbf{A}) = 0$$

(esimerkki:)

$$\begin{vmatrix} 2 & 0 & 1 \\ 7 & 0 & -2 \\ 3 & 0 & 5 \end{vmatrix} = 0$$

th_det2 **Lause 95** Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.

(Sanallinen muoto:) Jos neliömatriisin $\mathbf{A}$ kaksi saraketta vaihtavat paikkaa, niin syntyvän matriisin determinanti on alkuperäisen vastaluku.

(Kaavana:)

$$\mathbf{B} = \mathbf{A}[j : \vec{a}_{\bullet k}][k : \vec{a}_{\bullet j}] \Rightarrow \det(\mathbf{B}) = -\det(\mathbf{A})$$

(Esimerkki:)

$$\begin{vmatrix} 2 & 1 & 3 \\ -5 & 0 & 4 \\ 0 & 0 & 7 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 3 \\ 0 & -5 & 4 \\ 0 & 0 & 7 \end{vmatrix}$$

th_det3 **Lause 96** Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.

(Sanallinen muoto:) Jos neliömatriisissa $\mathbf{A}$ on kaksi identtistä saraketta, niin matriisin determinanti on 0.

(Kaavana:)

$$j \neq k, \text{ ja } \vec{a}_{\bullet j} = \vec{a}_{\bullet k} \Rightarrow \det(\mathbf{A}) = 0$$

(Esimerkki:)

$$\begin{vmatrix} 2 & 2 & 3 \\ -5 & -5 & 4 \\ 0 & 0 & 7 \end{vmatrix} = 0$$

th_det4

Lause 97 Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.(Sanallinen muoto:) Neliömatriisin $\mathbf{A}$ determinantin lasku on k :nnen sarakkeen suhteen lineaarinen operaatio.

(Kaavana:)

$$\begin{aligned} \det(\mathbf{A}[k : \vec{u} + \vec{v}]) &= \det(\mathbf{A}[k : \vec{u}]) + \det(\mathbf{A}[k : \vec{v}]), \quad \text{ja} \\ \det(\mathbf{A}[k : \lambda \cdot \vec{u}]) &= \lambda \cdot \det(\mathbf{A}[k : \vec{u}]) \end{aligned}$$

(Esimerkki:)

$$\begin{aligned} \begin{vmatrix} 2 & 5+1 & 3 \\ -5 & 2+9 & 4 \\ 0 & 3-2 & 7 \end{vmatrix} &= \begin{vmatrix} 2 & 5 & 3 \\ -5 & 2 & 4 \\ 0 & 3 & 7 \end{vmatrix} + \begin{vmatrix} 2 & 1 & 3 \\ -5 & 9 & 4 \\ 0 & -2 & 7 \end{vmatrix} \\ \begin{vmatrix} 2 & 100 \cdot 5 & 3 \\ -5 & 100 \cdot 2 & 4 \\ 0 & 100 \cdot 3 & 7 \end{vmatrix} &= 100 \cdot \begin{vmatrix} 2 & 5 & 3 \\ -5 & 2 & 4 \\ 0 & 3 & 7 \end{vmatrix} \end{aligned}$$

th_det5

Lause 98 Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.(Sanallinen muoto:) Jos neliömatriisin $\mathbf{A}$ sarake kerrotaan reaalityyppisellä λ ja näin saatu sarake lisätään toiseen $\mathbf{A}$:n sarakkeeseen, niin syntyvän uuden matriisin determinantti on sama, kuin alkuperäisen matriisin determinantti.

(Kaavana:)

$$j \neq k \quad \Rightarrow \quad \det(\mathbf{A}[k : \vec{a}_{\bullet k} + \lambda \vec{a}_{\bullet j}]) = \det(\mathbf{A})$$

(Esimerkki:)

$$\begin{vmatrix} 2 & 1 & 3 \\ -5 & 1 & 4 \\ 0 & 0 & 7 \end{vmatrix} = \begin{vmatrix} 7 & 1 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 7 \end{vmatrix} = \det \mathbf{A}[1 : \vec{a}_{\bullet 1} + 5 \cdot \vec{a}_{\bullet 2}]$$

th_det6

Lause 99 Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.(Sanallinen muoto:) Jos jokin neliömatriisin $\mathbf{A}$ sarake voidaan lausua muiden sarakkeiden lineaarikombinaationa, niin matriisin determinantti on nolla.

(Esimerkki:)

$$\begin{vmatrix} 2 & 1 & 4 \\ -5 & 1 & -3 \\ 0 & 1 & 2 \end{vmatrix} = 0, \quad \text{sillä } \vec{a}_{\bullet 3} = \vec{a}_{\bullet 1} + 2 \cdot \vec{a}_{\bullet 2}$$

th_det6b

Lause 100 Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.(Sanallinen muoto:) Neliömatriisilla $\mathbf{A}$ ja sen transpoosilla $\mathbf{A}^T$ on samat determinantit.

(Kaavana:)

$$\det(\mathbf{A}^T) = \det(\mathbf{A})$$

(Esimerkki:)

$$\begin{vmatrix} 2 & -5 & 0 \\ 1 & 1 & 0 \\ 3 & 4 & 7 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 \\ -5 & 1 & 4 \\ 0 & 0 & 7 \end{vmatrix}$$

th_det7

Lause 101 Lauseissa (94), (95), (96), (97), (98) ja (99) matriisin sarakkeita koskevat tulokset ovat voimassa myös matriisin riveille.

th_det8

Lause 102 Tarkastellaan $n \times n$ -matriiseja $\mathbf{A}$ ja $\mathbf{B}$.

(Sanallinen muoto:) Tulon determinantti on determinanttien tulo.

(Kaavana:)

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$$

th_det9

Lause 103 (Sanallinen muoto:) Yksikkömatriisin determinantti on 1.

(Kaavana:)

$$\det(\mathbf{I}) = 1$$

th_det10

Lause 104 Tarkastellaan $n \times n$ -matriisia $\mathbf{A}$.

(Sanallinen muoto:) Matriisi on säännöllinen (eli sille on olemassa käänteismatriisi), jos ja vain jos matriisin determinantti ei ole nolla.

(Kaavana:)

$$\exists \mathbf{A}^{-1} \Leftrightarrow \det(\mathbf{A}) \neq 0$$

Todistus: $\Rightarrow$ (eli $\exists \mathbf{A}^{-1} \Rightarrow \det(\mathbf{A}) \neq 0$)

Oletamme nyt, että käänteismatriisi on olemassa. Määritetään käänteismatriisi rivioperaatioiden avulla Gaussin algoritmeilla. Aluksi laadimme kaavion, jonka vasen puoli on $\mathbf{A}$. Jokaisessa vaiheessa algoritmin aikana kaavion vasemman osan determinantti joko, säilyy muuttumattomana, vaihtaa merkkiä tai tulee kerrotuksi nolasta eroavalla luvulla. Viimeisen kaavion vasemman puolen determinantti on 1. Siis ensimmäisen kaavion vasemman osan determinantti ei voi olla nolla.

 $\Leftarrow$ (eli $\nexists \mathbf{A}^{-1} \Rightarrow \det(\mathbf{A}) = 0$)

Oletamme nyt, että käänteismatriisia ei ole olemassa. Aloitetaan käänteismatriisin määrittäminen kuten edellä. Oletuksen mukaan Gaussin algoritmi ei onnistu, vaan pivotointi epäonnistuu jossakin vaiheessa. Tämä merkitsee, että viimeisessä kaaviossa on rivi, jonka vasen osa on täynnä nollia. Siis jokin $\mathbf{A}$:n rivi voidaan lausua muiden lineaarikombinaationa ja lauseiden (99) ja (101) mukaan determinantti on nolla. $\square$

th_det11

Lause 105 Jos $n \times n$ -matriisia $\mathbf{A}$ on säännöllinen, niin

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

Edellä oli suuri joukko lauseita. Tehdään vielä yhteenveto:

(lause 94 sivu 100)	th_det1	Jos neliömatriisin $\mathbf{A}$ jonkin sarakkeen kaikki alkiot ovat nollia, niin matriisin determinanti on nolla.
(lause 95 sivu 100)	th_det2	Jos neliömatriisin $\mathbf{A}$ kaksi saraketta vaihtavat paikkaa, niin syntyvän matriisin determinanti on alkuperäisen vastaluku.
(lause 96 sivu 100)	th_det3	Jos neliömatriisissa $\mathbf{A}$ on kaksi identtistä saraketta, niin matriisin determinanti on 0.
(lause 97 sivu 101)	th_det4	Neliömatriisin $\mathbf{A}$ determinantin lasku on k :nnen sarakkeen suhteen lineaarinen operaatio.
(lause 98 sivu 101)	th_det5	Jos neliömatriisin $\mathbf{A}$ sarake kerrotaan reaaliluvulla λ ja näin saatu sarake lisätään toiseen $\mathbf{A}$:n sarakkeeseen, niin syntyvän uuden matriisin determinanti on sama, kuin alkuperäisen matriisin determinanti.
(lause 99 sivu 101)	th_det6	Jos jokin neliömatriisin $\mathbf{A}$ sarake voidaan lausua muiden sarakkeiden lineaarikombinaationa, niin matriisin determinanti on nolla.
(lause 100 sivu 102)	th_det6b	Neliömatriisilla $\mathbf{A}$ ja sen transpoosilla $\mathbf{A}^T$ on samat determinantit.
(lause 101 sivu 102)	th_det7	Lauseissa (94), (95), (96), (97), (98) ja (99) matriisin sarakkeita koskevat tulokset ovat voimassa myös matriisin riveille.
(lause 102 sivu 102)	th_det8	$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$
(lause 103 sivu 102)	th_det9	$\det(\mathbf{I}) = 1$
(lause 104 sivu 102)	th_det10	Matriisi on säännöllinen (eli sille on olemassa käänteismatriisi), jos ja vain jos matriisin determinanti ei ole nolla.
(lause 105 sivu 103)	th_det11	Jos $n \times n$ -matriisia $\mathbf{A}$ on säännöllinen, niin
		$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$

5.2 Determinantin laskeminen

Sec:DetLasku

Tässä kappaleessa opimme kaksi tapaa laskea determinantin arvo. Ensimmäinen tapa, Gaussin algoritmi, perustuu rivioperaatioihin. Rivioperaatiot osataan jo, joten menettely on varsin helppo omaksua. Toinen tapa perustuu determinantin lausumiseen pienempien alideterminanttien lineaarikombinaationa. Menettely johtaa selvään kaavaan,

mikä usein koetaan helpotuksena.

Nämä kaksi tapaa näyttävät aluksi hyvin erilaisilta. Kumpikin tapa kuitenkin perustuu determinantin määritelmään ja determinantin ominaisuuksiin. Kummallakin tavalla on myös jatkossa oma roolinsa, joten kumpikin laskutapa tulee osata.

5.2.1 Determinantin laskeminen rivioperaatioiden avulla

detGauss

Ensin käsittelemme Gaussin algoritmin, jossa determinantti ensin muutetaan yläkolmiomuotoon. Yläkolmiomuoto (alacolmiomuoto) on muoto, jossa päälävistäjän alla (yllä) on vain nollia. Esimerkiksi determinantit

$$\begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} \quad \text{ja} \quad \begin{vmatrix} 1 & 3 & -1 & 0 \\ 0 & 2 & 7 & 2 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & -2 \end{vmatrix}$$

ovat yläkolmiomuodossa.

Lause 106 Jos $n \times n$ -matriisi $\mathbf{A}$ on yläkolmiomuodossa (tai alacolmiomuodossa) niin determinantin arvo on päälävistäjällä olevien alkioden tulo, $\det(\mathbf{A}) = a_{11}a_{22} \cdots a_{nn}$.

Todistus: HT

□

Lauseiden ^{1th det5} (98) ja ^{det6b} (100) mukaan determinantin arvo ei muutu, jos sen riviin lisätää skaalarilla kerrottu toinen rivi. Tällaisten rivioperaatioiden avulla (ja mahdollisesti vaihtamalla rivejä keskenään) voidaan determinantti aina muuttaa yläkolmiomuotoon. Ennen Gaussin algoritmia esitämme muutaman esimerkin.

Esimerkki 107

$$\begin{aligned} \begin{vmatrix} [1] & 3 & -1 \\ 1 & 5 & 0 \\ -2 & 2 & -3 \end{vmatrix} \begin{array}{cc} -1 & +2 \\ \leftarrow & \\ & \leftarrow \end{array} &= \begin{vmatrix} 1 & 3 & -1 \\ 0 & [2] & 1 \\ 0 & 8 & -5 \end{vmatrix} \begin{array}{c} \\ -4 \\ \leftarrow \end{array} \\ &= \begin{vmatrix} 1 & 3 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & -9 \end{vmatrix} \\ &= 1 \cdot 2 \cdot (-9) = -18 \end{aligned}$$

Esimerkki 108

$$\begin{aligned}
& \left| \begin{array}{cccc|ccc} [2] & 1 & 0 & -1 & -1 & -2 & +1 \\ 2 & 1 & 1 & 3 & \leftarrow & & \\ 4 & 0 & 1 & 3 & & \leftarrow & \\ -2 & 1 & 0 & 5 & & & \leftarrow \end{array} \right| &= \left| \begin{array}{cccc|ccc} [2] & 1 & 0 & -1 & & & \\ 0 & 0 & 1 & 4 & & & \leftarrow \text{vaihda} \\ 0 & -2 & 1 & 5 & & & \leftarrow \text{vaihda} \\ 0 & 2 & 0 & 4 & & & \end{array} \right| \\
&= (-1) \cdot \left| \begin{array}{cccc|ccc} 2 & 1 & 0 & -1 & & & \\ 0 & [-2] & 1 & 5 & & & +1 \\ 0 & 0 & 1 & 4 & & & \\ 0 & 2 & 0 & 4 & & & \leftarrow \end{array} \right| \\
&= (-1) \cdot \left| \begin{array}{cccc|ccc} 2 & 1 & 0 & -1 & & & \\ 0 & -2 & 1 & 5 & & & \\ 0 & 0 & [1] & 4 & & & -1 \\ 0 & 0 & 1 & 9 & & & \leftarrow \end{array} \right| \\
&= (-1) \cdot \left| \begin{array}{cccc|ccc} 2 & 1 & 0 & -1 & & & \\ 0 & -2 & 1 & 5 & & & \\ 0 & 0 & 1 & 4 & & & \\ 0 & 0 & 0 & 5 & & & \end{array} \right| \\
&= (-1) \cdot 2 \cdot (-2) \cdot 1 \cdot 5 = 20
\end{aligned}$$

Olkoon $\mathbf{A}$ seuraavassa $n \times n$ -matriisi. Seuraavan prosessin aikana determinanttikaavion ulkoasu muutetaan vaiheittain yläkolmiomuotoon siten, että lausekkeen arvo koko ajan säilyy samana. Merkitään peräkkäisiä muotoja: $|\mathbf{A}| = |\mathbf{A}(1)| = |\mathbf{A}(2)| = |\mathbf{A}(3)| = \dots = |\text{yläkolmiomuoto}|$

Gaussin algoritmi:

(0) Aseta $k = 1$, $|\mathbf{A}(1)| = |\mathbf{A}|$

(1) Otetaan tukialkioksi (pivot element) alkio $a(k)_{kk}$.

(2) Jos tukialkio on 0 ja tukialkion sarakkeessa tukialkion alapuolella on nollasta poikkeava alkio $a(k)_{rk}$, $r > k$, vaihdetaan rivit r ja k keskenään ja saatu uusi determinantti kerrotaan (-1) :llä.

(3) Lisätään tukialkion rivin alapuolella oleviin riveihin tukialkion rivi sopivalla luvulla kerrottuna siten, että tukialkion alapuolella olevat alkion muuttuvat nolliksi.

(4) Jos $k < n - 1$, niin kasvatetaan k :ta yhdellä ja siirrytään kohtaan (1), muuten kohtaan (5).

(5) $\det(\mathbf{A}) = a(k)_{11}a(k)_{22} \cdots a(k)_{nn}$

Huomautus: Kohdan (3) rivioperaatiot eivät onnistu, jos tukialkio on 0. Tämän takia tehdään rivinvaihto kohdassa (2), jotta tukialkio olisi nolasta poikkeava. Jos tukialkio ja lisäksi kaikki sen alla olevat alkio ovat nollia, ei kohdassa (3) tarvitse tehdä mitään. Poiketen edeltävistä Gaussin nimeä kantavista algoritmeista, tämä algoritmi toimii myös, kun $\det(\mathbf{A}) = 0$. Silloin diagonaalille jää 0 ja determinantin arvoksi tulee 0, niin kuin pitääkin!

On hyvä huomata, että edellä kuvattu menettely toimii aina. Determinantin laskeminen rivioperaatioiden avulla voidaan tehdä mekaanisesti, mutta hyvä ”pelisilmä” auttaa usein löytämään nopeita reittejä kolmiokaavioon. Kaavioiden käsittely determinantin laskun aikana sallii ”sarakeoperaatiot”, mutta näitä kannattaa nyt välttää. Oppimisen kannalta on nyt hyvä keskittyä rivioperaatioiden oppimiseen. Jos siis havaitset mahdollisuuden ”sarakeoperaatioihin”, niin transponoi kaavio ja jatka rivioperaatioilla.

detMinor*i*

5.2.2 Determinantin laskeminen minorikehitelmällä

Määritelmä 109 Matriisin $\mathbf{A}$ alimatriisi $\mathbf{A}_{ij}$ saadaan poistamalla $\mathbf{A}$:sta i :s rivi ja j :s sarake. Alimatriisin determinanti $\det(\mathbf{A}_{ij})$ on matriisin $\mathbf{A}$ alkioon a_{ij} liittyvä minori (alideterminantti). Lukua $(-1)^{i+j} \det(\mathbf{A}_{ij})$ sanotaan matriisin $\mathbf{A}$ alkioon a_{ij} liittyväksi kofaktoriksi.

Esimerkki 110 Jos

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 & 2 \\ 3 & 4 & 7 & 1 \\ -2 & 4 & 5 & 1 \\ 0 & 1 & 8 & 2 \end{pmatrix}$$

niin

$$\mathbf{A}_{23} = \left(\begin{array}{ccc|c} 1 & 0 & 3 & 2 \\ 3 & 4 & 7 & 1 \\ -2 & 4 & 5 & 1 \\ 0 & 1 & 8 & 2 \end{array} \right) = \begin{pmatrix} 1 & 0 & 2 \\ -2 & 4 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

Vastaavat alkioon a_{23} liittyvät minori ja kofaktori ovat

$$\begin{aligned} \text{minori} &= \det(\mathbf{A}_{23}) = 3 \\ \text{kofaktori} &= (-1)^{2+3} \det(\mathbf{A}_{23}) = -3 \end{aligned}$$

Lause 111 $n \times n$ -neliömatriisin $\mathbf{A}$ determinantti saadaan laskettua kaavoilla

$$\det(\mathbf{A}) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}) \quad \text{kehitemä } i\text{:nnen rivin suhteen}$$

$$\det(\mathbf{A}) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}) \quad \text{kehitemä } j\text{:nnen sarakkeen suhteen}$$

Determinantin kehittäminen kannattaa perustaa riviin tai sarakkeeseen, jolla on mahdollisimman monta nollaa.

Esimerkki 112 (1) Kehitetään determinantti ensimmäisen rivin suhteen

$$\begin{aligned} & \begin{vmatrix} \mathbf{1} & \mathbf{3} & -\mathbf{1} \\ \mathbf{1} & \mathbf{5} & \mathbf{0} \\ -\mathbf{2} & \mathbf{2} & -\mathbf{3} \end{vmatrix} = \sum_{j=1}^3 (-1)^{1+j} a_{1j} \det(\mathbf{A}_{1j}) \\ &= (-1)^{1+1} \cdot 1 \cdot \begin{vmatrix} \mathbf{5} & \mathbf{0} \\ \mathbf{2} & -\mathbf{3} \end{vmatrix} + (-1)^{1+2} \cdot 3 \cdot \begin{vmatrix} \mathbf{1} & \mathbf{0} \\ -\mathbf{2} & -\mathbf{3} \end{vmatrix} + (-1)^{1+3} \cdot (-1) \cdot \begin{vmatrix} \mathbf{1} & \mathbf{5} \\ -\mathbf{2} & \mathbf{2} \end{vmatrix} \\ &= 1 \cdot (-15 - 0) + (-1) \cdot 3 \cdot (-3 - 0) + (-1) \cdot (2 + 10) \\ &= -18 \end{aligned}$$

(2) Kehitetään sama determinantti toisen rivin suhteen

$$\begin{aligned} & \begin{vmatrix} \mathbf{1} & \mathbf{3} & -\mathbf{1} \\ \mathbf{1} & \mathbf{5} & \mathbf{0} \\ -\mathbf{2} & \mathbf{2} & -\mathbf{3} \end{vmatrix} = \sum_{j=1}^3 (-1)^{2+j} a_{2j} \det(\mathbf{A}_{2j}) \\ &= (-1)^{2+1} \cdot 1 \cdot \begin{vmatrix} \mathbf{3} & -\mathbf{1} \\ \mathbf{2} & -\mathbf{3} \end{vmatrix} + (-1)^{2+2} \cdot 5 \cdot \begin{vmatrix} \mathbf{1} & -\mathbf{1} \\ -\mathbf{2} & -\mathbf{3} \end{vmatrix} + (-1)^{2+3} \cdot 0 \cdot \begin{vmatrix} \mathbf{1} & \mathbf{3} \\ -\mathbf{2} & \mathbf{2} \end{vmatrix} \\ &= (-1) \cdot (-9 + 2) + 5 \cdot (-3 - 2) - 0 \\ &= -18 \end{aligned}$$

Käytännön laskemista helpottaa se, että kofaktoreiden merkkikaavio on varsin yksinkertainen:

$$\begin{pmatrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

5.3 Matriisin kääntäminen adjungaattilla

Sec:MatInvAdj

Määritelmä 113 $n \times n$ -matriisin $\mathbf{A}$ adjungaatti on $n \times n$ -matriisi $\text{adj}(\mathbf{A}) = (\alpha_{ij})$, missä $\alpha_{ij} = (-1)^{i+j} \det(\mathbf{A}_{ji})$ (, joka on siis alkioon a_{ji} liittyvä kofaktori).

Siis adjungaatti on kofaktorikaavion transpoosi.

Lause 114 (1)

$$\text{adj}(\mathbf{A})\mathbf{A} = \mathbf{A}\text{adj}(\mathbf{A}) = \det(\mathbf{A})\mathbf{I}$$

(2) Jos $\det(\mathbf{A}) \neq 0$, niin

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \cdot \text{adj}(\mathbf{A})$$

Todistus: (1)

$$\begin{aligned} (\text{adj}(\mathbf{A})\mathbf{A})_{ij} &= \sum_{k=1}^n \alpha_{ik} a_{kj} = \sum_{k=1}^n a_{kj} (-1)^{i+k} \det(\mathbf{A}_{ki}) \\ &= \det(\mathbf{A}[i : \sum_{k=1}^n a_{kj} \vec{e}_k]) = \det(\mathbf{A}[i : \vec{a}_{\bullet j}]) \\ &= \begin{cases} 0, & \text{kun } i \neq j \\ \det(\mathbf{A}), & \text{kun } i = j \end{cases} \\ &= \det(\mathbf{A}) \delta_{ij} \end{aligned}$$

Siis $\text{adj}(\mathbf{A})\mathbf{A} = \det(\mathbf{A})\mathbf{I}$. Toisaalta voimme päätellä, että adjungaatin transpoosi on transpoosin adjungaatti seuraavasti:

$$(\text{adj}(\mathbf{A})^T)_{ij} = \text{adj}(\mathbf{A})_{ji} = (-1)^{j+i} \det(\mathbf{A}_{ij}) = (-1)^{i+j} \det((\mathbf{A}^T)_{ji}) = (\text{adj}(\mathbf{A}^T))_{ij}$$

Tästä seuraa, että

$$\mathbf{A}\text{adj}(\mathbf{A}) = (\text{adj}(\mathbf{A})^T \mathbf{A}^T)^T = (\text{adj}(\mathbf{A}^T) \mathbf{A}^T)^T = (\det(\mathbf{A}^T) \mathbf{I})^T = \det(\mathbf{A}) \mathbf{I}$$

(2) Seuraa välittömästi kohdasta (1). □

Esimerkki 115 Olkoon

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

Silloin

$$\begin{aligned} \det(\mathbf{A}) &= (-1)^{1+1} \cdot 3 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} + (-1)^{1+2} \cdot 3 \cdot \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} + (-1)^{1+3} \cdot 1 \cdot \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} \\ &= 3 \cdot (3 - 1) - 3 \cdot (2 - 0) + (2 - 0) = 2 \end{aligned}$$

$$|\mathbf{A}_{11}| = \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} = 2; \quad |\mathbf{A}_{12}| = \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = 2; \quad |\mathbf{A}_{13}| = \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} = 2$$

$$|\mathbf{A}_{21}| = \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} = 2; \quad |\mathbf{A}_{22}| = \begin{vmatrix} 3 & 1 \\ 0 & 1 \end{vmatrix} = 3; \quad |\mathbf{A}_{23}| = \begin{vmatrix} 3 & 3 \\ 0 & 1 \end{vmatrix} = 3$$

$$|\mathbf{A}_{31}| = \begin{vmatrix} 3 & 1 \\ 3 & 1 \end{vmatrix} = 0; \quad |\mathbf{A}_{32}| = \begin{vmatrix} 3 & 1 \\ 2 & 1 \end{vmatrix} = 1; \quad |\mathbf{A}_{33}| = \begin{vmatrix} 3 & 3 \\ 2 & 3 \end{vmatrix} = 3$$

ja

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{\det(\mathbf{A})} \cdot \begin{pmatrix} +|\mathbf{A}_{11}| & -|\mathbf{A}_{21}| & +|\mathbf{A}_{31}| \\ -|\mathbf{A}_{12}| & +|\mathbf{A}_{22}| & -|\mathbf{A}_{32}| \\ +|\mathbf{A}_{13}| & -|\mathbf{A}_{23}| & +|\mathbf{A}_{33}| \end{pmatrix} \\ &= \frac{1}{2} \cdot \begin{pmatrix} 2 & -2 & 0 \\ -2 & 3 & -1 \\ 2 & -3 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1.5 & -0.5 \\ 1 & -1.5 & 1.5 \end{pmatrix} \end{aligned}$$

Tarkistus:

$$\begin{aligned} \begin{pmatrix} 3 & 3 & 1 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1.5 & -0.5 \\ 1 & -1.5 & 1.5 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1.5 & -0.5 \\ 1 & -1.5 & 1.5 \end{pmatrix} \begin{pmatrix} 3 & 3 & 1 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

5.4 Cramerin kaavat

Sec:Cramer

Edellä käydyt tarkastelut antavat nyt mahdollisuuden johtaa kaava myös yhtälöryhmän ratkaisulle silloin, kun ratkaisu on yksikäsitteinen. Kysymys ei kuitenkaan ole vain keinosta saada kaava, vaan esimerkit tuovat esiin, miksi Cramerin kaavat ovat korvaamaton apu eräissä ongelmissa.

Lause 116 (Cramerin kaavat) (Oletus:) Olkoon $\mathbf{A}\vec{x} = \vec{b}$, jossa muuttujia on yhtä monta kuin yhtälöitä, ja yhtälöryhmän kerroinmatriisi on säännöllinen. Auki kirjoitettuna:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad \text{ja} \quad \det(\mathbf{A}) \neq 0.$$

(Väite sanallisesti:) Yhtälöryhmän k :s tuntematon saadaan lausekkeesta

$$x_k = \frac{D_k}{\det(\mathbf{A})},$$

missä D_k on determinantti, joka saadaan, kun kerroinmatriisin k :s sarake korvataan $\vec{b}$:llä (eli yhtälöryhmän RHS:llä).

(Väite kaavana:)

$$x_k = \frac{\det(\mathbf{A}[k : \vec{b}])}{\det(\mathbf{A})}.$$

(Kaava aukikirjoitettuna:)

$$x_k = \frac{1}{\det(\mathbf{A})} \cdot \begin{vmatrix} a_{11} & \cdots & a_{1(k-1)} & b_1 & a_{1(k+1)} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2(k-1)} & b_2 & a_{2(k+1)} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(k-1)} & b_1 & a_{n(k+1)} & \cdots & a_{nn} \end{vmatrix}$$

Todistus: Lausekkeiden yksinkertaistamiseksi merkitsemme tämän todistuksen ajan matriisiin $\mathbf{A}$ alkioon a_{ij} liittyvää kofaktoria symbolilla c_{ij} . Koska $\mathbf{A}$ on säännöllinen, on sillä käänteismatriisi, jonka voimme lausua adjungaatin avulla. Siis

$$\begin{aligned} \vec{x} &= \mathbf{A}^{-1}\vec{b} = \frac{1}{\det(\mathbf{A})} \cdot \text{adj}(\mathbf{A})\vec{b} \\ &= \frac{1}{\det(\mathbf{A})} \cdot \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & \cdots & c_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \end{aligned}$$

Ratkaisuvektorin k :s koordinaatti on siis

$$\begin{aligned} x_k &= \frac{1}{\det(\mathbf{A})} \cdot \sum_{j=1}^n c_{jk} b_j \\ &= \frac{1}{\det(\mathbf{A})} \cdot \sum_{j=1}^n (-1)^{j+k} b_j \det(\mathbf{A}_{jk}) \\ &= \frac{1}{\det(\mathbf{A})} \cdot \det(\mathbf{A}[k : \vec{b}]) \end{aligned}$$

□

Esimerkki 117 Ratkaistaksemme yhtälöryhmän

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

laskemme determinantit

$$D = \begin{vmatrix} 2 & 9 & -2 \\ 7 & -3 & -6 \\ -1 & 2 & 3 \end{vmatrix} \\ = 2 \cdot \begin{vmatrix} -3 & -6 \\ 2 & 3 \end{vmatrix} - 9 \cdot \begin{vmatrix} 7 & -6 \\ -1 & 3 \end{vmatrix} + (-2) \cdot \begin{vmatrix} 7 & -3 \\ -1 & 2 \end{vmatrix} = -151$$

$$D_1 = \begin{vmatrix} \mathbf{2} & 9 & -2 \\ \mathbf{8} & -3 & -6 \\ \mathbf{1} & 2 & 3 \end{vmatrix} \\ = \mathbf{2} \cdot \begin{vmatrix} -3 & -6 \\ 2 & 3 \end{vmatrix} - 9 \cdot \begin{vmatrix} \mathbf{8} & -6 \\ \mathbf{1} & 3 \end{vmatrix} + (-2) \cdot \begin{vmatrix} \mathbf{8} & -3 \\ \mathbf{1} & 2 \end{vmatrix} = -302$$

$$D_2 = \begin{vmatrix} 2 & \mathbf{2} & -2 \\ 7 & \mathbf{8} & -6 \\ -1 & \mathbf{1} & 3 \end{vmatrix} \\ = 2 \cdot \begin{vmatrix} \mathbf{8} & -6 \\ \mathbf{1} & 3 \end{vmatrix} - \mathbf{2} \cdot \begin{vmatrix} 7 & -6 \\ -1 & 3 \end{vmatrix} + (-2) \cdot \begin{vmatrix} 7 & \mathbf{8} \\ -1 & \mathbf{1} \end{vmatrix} = 0$$

$$D_3 = \begin{vmatrix} 2 & 9 & \mathbf{2} \\ 7 & -3 & \mathbf{8} \\ -1 & 2 & \mathbf{1} \end{vmatrix} \\ = 2 \cdot \begin{vmatrix} -3 & \mathbf{8} \\ 2 & \mathbf{1} \end{vmatrix} - 9 \cdot \begin{vmatrix} 7 & \mathbf{8} \\ -1 & \mathbf{1} \end{vmatrix} + \mathbf{2} \cdot \begin{vmatrix} 7 & -3 \\ -1 & 2 \end{vmatrix} = -151$$

Nyt $x = D_1/D = 2$, $y = D_2/D = 0$ ja $z = D_3/D = 1$.

Esimerkki 118 Seuraavaksi palaamme esimerkkiin ^{|esim1_c2s4}(??), joka oli jo sivulla ^{|esim1_c2s4}?? Rat-
kaistavana oli yhtälöryhmä

$$\begin{cases} -x + y - z = 0 \\ x - \alpha\beta y = 0 \\ -ty + z = I_0 \end{cases},$$

missä

x = kulutus tasapainotilassa ,

y = kansantulo tasapainotilassa ,

z = investoinnit tasapainotilassa ,

I_0 = keskimääräinen investointitaso (oletetaan vakioksi),

s = säästämisaste ($0 \leq s \leq 1$, oletetaan vakioksi),

t = veroaste ($0 \leq t \leq 1$, oletetaan vakioksi),

k = investointien herkyys kulutuksen muutokselle, (oletetaan vakioksi)

$\alpha = 1 - s$,

$\beta = 1 - t$

Yhtälöryhmässä esiintyy nyt kirjainkertoimia, joiden numeroarvot eivät ole tiedossa. Silti haluamme voida ratkaista yhtälöryhmän. Toisessa pivotoinnissa voi tulla ongelmia, jos käytämme rivioperaatioita. Cramerin kaavoja käytettäessä, ei tarvitse pivotoida, vaan kaavat perustuvat determinantteihin, joiden lasku on tässä tilanteessa suoraviivaista. Kerroinmatriisi ja RHS ovat nyt:

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & -1 \\ 1 & -\alpha\beta & 0 \\ 0 & -t & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0 \\ 0 \\ I_0 \end{pmatrix}$$

ja tarvittavat determinantit ovat:

$$\begin{aligned} D &= \begin{vmatrix} -1 & 1 & -1 \\ 1 & -\alpha\beta & 0 \\ 0 & -t & 1 \end{vmatrix} \quad (\text{kehitetään 1. sarakkeen suhteen}) \\ &= +(-1) \cdot \begin{vmatrix} -\alpha\beta & 0 \\ -t & 1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & -1 \\ -t & 1 \end{vmatrix} + 0 \\ &= \alpha\beta - (1-t) = -s(1-t) \end{aligned}$$

$$\begin{aligned} D_1 &= \begin{vmatrix} 0 & 1 & -1 \\ 0 & -\alpha\beta & 0 \\ I_0 & -t & 1 \end{vmatrix} \quad (\text{kehitetään 1. sarakkeen suhteen}) \\ &= +0 - 0 + I_0 \cdot \begin{vmatrix} 1 & -1 \\ -\alpha\beta & 0 \end{vmatrix} \\ &= -I_0\alpha\beta = -I_0(1-s)(1-t) \end{aligned}$$

$$\begin{aligned} D_2 &= \begin{vmatrix} -1 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & I_0 & 1 \end{vmatrix} \quad (\text{kehitetään 2. sarakkeen suhteen}) \\ &= -0 + 0 - I_0 \cdot \begin{vmatrix} -1 & -1 \\ 1 & 0 \end{vmatrix} = -I_0 \end{aligned}$$

$$\begin{aligned} D_3 &= \begin{vmatrix} -1 & 1 & 0 \\ 1 & -\alpha\beta & 0 \\ 0 & -t & I_0 \end{vmatrix} \quad (\text{kehitetään 3. sarakkeen suhteen}) \\ &= +0 - 0 + I_0 \cdot \begin{vmatrix} -1 & 1 \\ 1 & -\alpha\beta \end{vmatrix} = I_0(\alpha\beta - 1) = I_0((1-s)(1-t) - 1) \end{aligned}$$

Siis ratkaisu on olemassa, kun $s \neq 0$ ja $t \neq 1$ ja silloin tasapainossa

$$\text{kulutus} = x = \frac{D_1}{D} = \frac{-I_0(1-s)(1-t)}{-s(1-t)} = \frac{I_0(1-s)}{s}$$

$$\text{kansantulo} = y = \frac{D_2}{D} = \frac{-I_0}{-s(1-t)} = \frac{I_0}{s(1-t)}$$

$$\text{investoinnit} = z = \frac{D_3}{D} = \frac{I_0((1-s)(1-t)-1)}{-s(1-t)} = I_0 \left(1 + \frac{t}{s(1-t)} \right)$$

5.5 Homogeeniset yhtälöryhmät

Sec:YRHomog

Tässä kappaleessa tutkitaan yhtälöryhmää, jossa on n muuttujaa ja n yhtälöä, ja yksi yhtälöryhmän kerroinmatriisin sarake voidaan lausua muiden sarakkeiden lineaarikombinaationa. Voisimme toki tutkia jo tapausta, jossa ”yksi tai useampi” sarake voidaan lausua muiden lineaarikombinaationa, mutta lykkäämme tämän tapauksen käsittelyn lukuun 5.

Yksinkertaistamalla ongelmanasettelua saamme helpommin luettavan esityksen. Näillä eväillä tulemme luvussa 4 laskemaan matriisien ominaisvektoreita.

Määritelmä 119 *Yhtälöryhmä on homogeeninen, jos jokaisen yhtälön RHS on nolla.*

Homogeeninen yhtälöryhmä on siis muotoa

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

Homogeenisella yhtälöryhmällä on aina ratkaisu $\vec{x} = \vec{0}$, jota sanomme triviaaliksi ratkaisuksi. Jos yhtälöryhmän kerroinmatriisi on säännöllinen ($\det(\mathbf{A}) \neq 0$), niin triviaali ratkaisu on ainoa ratkaisu.

Jotta homogeenisella yhtälöryhmällä (n muuttujaa ja n yhtälöä) olisi ei-triviaali ratkaisu, tulee kerroinmatriisin determinantin olla 0. Silloin yhtälöryhmää ei voida ratkaista

käänteismatriisin tai Cramerin kaavojen avulla, vaan tulee suorittaa yhtälöryhmän ratkaisu Gaussin algoritmin yleistä muotoa noudattaen.

esim01_Markov

Esimerkki 120 Firmassa on käytössä 70 autoa. Autojen kunto vaihtelee niin, että autot luokitellaan kolmeen kuntoluokkaan (1) hyvä kunto, (2) kohtalainen kunto ja (3) heikko kunto. Autoja huolletaan määräajoin ja tarvittaessa korjataan sopimushuollossa. Tarvittaessa loppuunajettu auto korvataan uudella, jolloin kuntoluokka nousee (3):sta (1):een. Huoltovälit ja huollon taso ovat sellaiset, että siirtymistodennäköisyydet eri luokkien välillä puolen vuoden jaksossa arvioidaan seuraaviksi:

	luokasta (1)	luokasta (2)	luokasta (3)
luokkaan (1)	$p_{11} = 0.80$	$p_{12} = 0.05$	$p_{13} = 0.10$
luokkaan (2)	$p_{21} = 0.15$	$p_{22} = 0.70$	$p_{23} = 0.10$
luokkaan (3)	$p_{31} = 0.05$	$p_{32} = 0.25$	$p_{33} = 0.80$

Jos hetkellä t eri luokissa olevien autojen lukumäärät ovat $x_1(t)$ autoa luokassa (1), $x_2(t)$ autoa luokassa (2) ja $x_3(t)$ autoa luokassa (3), niin hetkellä $t + 1$ lukumäärät ovat

$$\begin{aligned}x_1(t+1) &= p_{11}x_1(t) + p_{12}x_2(t) + p_{13}x_3(t) \\x_2(t+1) &= p_{21}x_1(t) + p_{22}x_2(t) + p_{23}x_3(t) \\x_3(t+1) &= p_{31}x_1(t) + p_{32}x_2(t) + p_{33}x_3(t)\end{aligned}$$

eli

$$\vec{x}(t+1) = \mathbf{P}\vec{x}(t)$$

Jos huoltosysteemi ei muutu ja siirtymätodennäköisyydet pysyvät samoina pitkän ajan, lähestytään tasapainotilaa, jossa tasapainoarvot $x_k(t+1) = x_k(t) = x_k$ toteuttavat yhtälöt

$$\begin{aligned}&\begin{cases} p_{11}x_1 + p_{12}x_2 + p_{13}x_3 = x_1 \\ p_{21}x_1 + p_{22}x_2 + p_{23}x_3 = x_2 \\ p_{31}x_1 + p_{32}x_2 + p_{33}x_3 = x_3 \end{cases} \\ \Leftrightarrow &\begin{cases} (p_{11}-1)x_1 + p_{12}x_2 + p_{13}x_3 = 0 \\ p_{21}x_1 + (p_{22}-1)x_2 + p_{23}x_3 = 0 \\ p_{31}x_1 + p_{32}x_2 + (p_{33}-1)x_3 = 0 \end{cases} \\ \Leftrightarrow &\begin{cases} -0.20x_1 + 0.05x_2 + 0.10x_3 = 0 \\ 0.15x_1 - 0.30x_2 + 0.10x_3 = 0 \\ 0.05x_1 + 0.25x_2 - 0.20x_3 = 0 \end{cases}\end{aligned}$$

Kun kerroinmatriisin rivit lasketaan yhteen, saadaan nollarivi. Siis yhtälöryhmän kerroinmatriisin determinantti on nolla, ja homogeenisella yhtälöryhmällä on ei-triviaaleja ratkaisuja. Haetaan nämä ratkaisut Gaussin algoritmilla

$$\begin{aligned}
& \left(\begin{array}{ccc|c} -0.2000 & 0.0500 & 0.1000 & 0 \\ 0.1500 & -0.3000 & 0.1000 & 0 \\ 0.0500 & 0.2500 & -0.2000 & 0 \end{array} \right) \quad \left| \begin{array}{l} \cdot(-5) \\ \\ \end{array} \right. \\
& \sim \left(\begin{array}{ccc|c} [1] & -0.2500 & -0.5000 & 0 \\ 0.1500 & -0.3000 & 0.1000 & 0 \\ 0.0500 & 0.2500 & -0.2000 & 0 \end{array} \right) \quad \left| \begin{array}{l} 0.15 \\ \leftarrow \\ \\ \end{array} \right. \begin{array}{l} -0.05 \\ \\ \leftarrow \end{array} \\
& \sim \left(\begin{array}{ccc|c} (1) & -0.2500 & -0.5000 & 0 \\ 0 & [-0.2625] & 0.1750 & 0 \\ 0 & 0.2625 & -0.1750 & 0 \end{array} \right) \quad \left| \begin{array}{l} \\ +1 \\ \leftarrow \end{array} \right. \\
& \sim \left(\begin{array}{ccc|c} (1) & -0.2500 & -0.5000 & 0 \\ 0 & (-0.2625) & 0.1750 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)
\end{aligned}$$

Kirjoitamme kaavion takaisin yhtälöiksi, ja koska x_3 :n sarakkeessa emme pivotoineet, siirrämme x_3 :a sisältävät termit RHS:iin.

$$\begin{cases} x_1 - 0.2500x_2 = 0.5000x_3 \\ -0.2625x_2 = -0.1750x_3 \end{cases}$$

$$\text{Siis} \quad x_2 = \frac{-0.1750}{-0.2625}x_3 = \frac{2}{3}x_3,$$

$$\text{ja} \quad x_1 = \frac{1}{2}x_3 + \frac{1}{4} \cdot \frac{2}{3}x_3 = \frac{2}{3}x_3.$$

Homogeenisen yhtälöryhmän ei-triviaali ratkaisu on siis

$$\begin{cases} x_1 = 2a/3 \\ x_2 = 2a/3 \\ x_3 = a \end{cases}, \quad \text{missä } a \neq 0$$

Toisaalta tässä esimerkissä $x_1 + x_2 + x_3 = 70$, joten

$$\begin{aligned}
\frac{2a}{3} + \frac{2a}{3} + a &= 70 \\
\Leftrightarrow \frac{2a}{3} + \frac{2a}{3} + \frac{3a}{3} &= 70 \\
\Leftrightarrow \frac{7a}{3} &= 70 \\
\Leftrightarrow a &= 30
\end{aligned}$$

Siis

$$\begin{cases} x_1 = 20 \\ x_2 = 20 \\ x_3 = 30 \end{cases}.$$

6 PISTETULON JA RISTITULON SOVELLUKSIA

Ch:DotCross

6.1 Ristitulo determinantin avulla

Sec:CrosDet

Vektoreiden $\vec{a}$ ja $\vec{b}$ ristitulo määriteltiin sivulla ^{|DCintro_def1|}37 seuraavasti:

Olkoon $\vec{a}$ ja $\vec{b}$ kaksi avaruuden vektoria ja olkoon $\gamma = \angle(\vec{a}, \vec{b})$ niiden välinen kulma. Vektoreiden $\vec{a}$ ja $\vec{b}$ ristitulo $\vec{a} \times \vec{b}$ on vektori, jonka pituus on

$$|\vec{a} \times \vec{b}| = |\vec{a}| |\vec{b}| \sin \gamma,$$

ja suunta määräytyy seuraavasta kolmesta ehdosta

1. vektorit $\vec{a}$ ja $\vec{a} \times \vec{b}$ ovat kohtisuorassa toisiaan vastaan,
2. vektorit $\vec{b}$ ja $\vec{a} \times \vec{b}$ ovat kohtisuorassa toisiaan vastaan,
3. vektorit $\vec{a}$, $\vec{b}$ ja $\vec{a} \times \vec{b}$ (tässä järjestyksessä!) muodostavat positiivisesti suunnistettun systeemin.

Lauseessa ^{|laDCintro_lausel|}(24) sivulla ^{|laDCintro_lausel|}39 annettiin vektoreiden $\vec{a} = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}$ ja $\vec{b} = b_x \vec{i} + b_y \vec{j} + b_z \vec{k}$ ristitulolle lauseke, jonka saamme determinantti-merkinnän avulla helposti muistettavaan muotoon

$$\vec{a} \times \vec{b} = (a_y b_z - a_z b_y) \vec{i} + (a_z b_x - a_x b_z) \vec{j} + (a_x b_y - a_y b_x) \vec{k} \quad (6.1)$$

$$\begin{aligned} &= \vec{i} \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} - \vec{j} \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} + \vec{k} \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \\ &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}. \end{aligned} \quad (6.2)$$

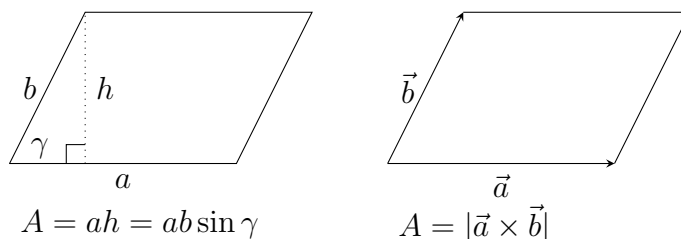
Huomaa, että lauseke ^{|CrosDet_eq2|}(6.2) ei nyt ole tavallinen lukukaavion determinantti, vaan se on determinantti-merkintä. Kun merkintä kehitetään ensimmäisen rivinsä suhteen, syntyy täsmälleen oikea määritelmän mukainen vektorilauseke ^{|CrosDet_eq1|}(6.1). Merkinnän voi itse asiassa kehittää minkä tahansa rivin tai sarakkeen suhteen, ja aina saadaan täsmälleen oikea lauseke.

6.2 Pinta-aloja ja tilavuuksia

c:DCpinta_alat

Tason suunnikas.

Suunnikas on nelikulmio, jonka sivut ovat pareittain yhtä pitkät ja yhdensuuntaiset. Jos mahdollista suunnikas on tapana piirtää niin, että yksi sivu a on vaakasuora. Sanomme sitä kannaksi. Toinen sivu b on vino. Sanomme sitä kyljeksi. Kannan ja kyljen välinen kulma on γ . Kun kuvaan lisätään vielä vaakasuorien sivujen kohtisuora etäisyys, eli suunnikkaan korkeus h , niin suorakulmaisten kolmioiden trigonometriasta saamme helposti suunnikkaan pinta-alan.



Kaava $A = |\vec{a} \times \vec{b}|$ tuntuu nyt oudolta. Olemme nyt tasossa ja ristitulo kuuluu kolmiulotteisen avaruuden laskuihin. Ihminen kuitenkin aina elää ja hengittää kolmiulotteisessa maailmassa. Jos tason koordinaattiakselien yksikkö-suuntavektorit ovat $\vec{i}$ ja $\vec{j}$ ($\vec{i} \perp \vec{j}$), niin vektori $\vec{k} = \vec{i} \times \vec{j}$ on kohtisuorassa kumpaakin akselia vastaan ja kolmikko $\{\vec{i}, \vec{j}, \vec{k}\}$ on positiivisesti suunnistettu. Jos alkuperäiset vektorit ovat $\vec{a} = a_x \vec{i} + a_y \vec{j}$ ja $\vec{b} = b_x \vec{i} + b_y \vec{j}$, niin täydennetään vektorit muotoon $\vec{a} = a_x \vec{i} + a_y \vec{j} + 0 \cdot \vec{k}$ ja $\vec{b} = b_x \vec{i} + b_y \vec{j} + 0 \cdot \vec{k}$. Täydennyksen jälkeen ristitulossa ei enää ole mitään ongelmaa.

6.3 Skalaarikolmitulo ja vektorikolmitulo

6.4 Vektorikolmikon suunnistus

6.5 Suoran ja tason parametrimuotoinen yhtälö

6.6 Suoran ja tason koordinaattimuotoinen yhtälö

6.7 Etäisyyksiä avaruudessa

Sec:DistSpace

6.7.1 Pisteen etäisyys suorasta

Ssec:distPS

Tasossa pisteen etäisyys suorasta on helppo selvittää. Piirrämme taso-ongelmasta kuvan ja teemme konstruktion, joka helpottaa asian hahmottamista. Tämän jälkeen yleis-
tämme saamamme kaavat kolmiulotteiseen tilanteeseen.

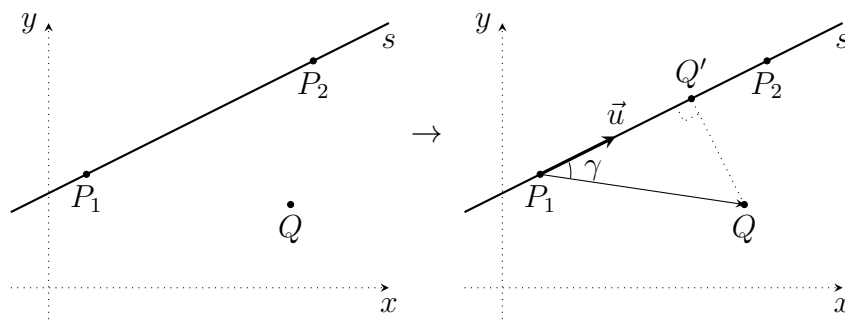


fig:dPs01

Kuva 24. Tason pisteen etäisyys suorasta.

Olkoon s suora, ja olkoon $Q = (x_0, y_0)$ piste suoran ulkopuolella. Etsitään suoralta piste Q' , joka on lähinnä pistettä $Q = (x_0, y_0)$. Pisteen etäisyys suorasta on sen etäisyys suoran lähimmästä pisteestä, $d(Q, s) = d(Q, Q')$.

Lause 121 (Tason pisteen etäisyys suorasta.) Jos suoran s yhtälö on

$$ax + by + c = 0,$$

niin pisteen $Q = (x_0, y_0)$ etäisyys suorasta on

$$d = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

Perustelu: (Tapaus $b = 0$.) Suora on pystysuora $x = -c/a$, jolloin pisteen etäisyys suorasta on

$$d(Q, s) = \left| x_0 + \frac{c}{a} \right| = \frac{|ax_0 + c|}{|a|} = \frac{|ax_0 + c|}{\sqrt{a^2}}.$$

(Tapaus $b \neq 0$.) Valitaan suoralta pisteet $P_1 = (0, -c/b)$ ja $P_2 = (b, -c/b - a)$. Suoran suuntavektoriksi voimme valita vektorin $\vec{u} = \overrightarrow{P_1P_2} = b\vec{i} - a\vec{j}$. Seuraavaksi upotamme vektorit ja pisteet x, y, z -koordinaatistoon asettamalla z -koordinaatit nolliksi. Kuvan 24 merkinnöin

$$\begin{aligned} d(Q, s) &= d(Q, Q') = |\overrightarrow{P_1Q}| \sin \gamma = \frac{|\overrightarrow{P_1Q} \times \vec{u}|}{|\vec{u}|} \\ &= \frac{|(x_0\vec{i} + (y_0 + \frac{c}{b})\vec{j} + 0\vec{k}) \times (b\vec{i} - a\vec{j} + 0\vec{k})|}{\sqrt{b^2 + (-a)^2}} = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}. \end{aligned}$$

□

Lause 122 (Avaruuden pisteen etäisyys suorasta.) Jos suora s kulkee pisteiden P_1 ja P_2 kautta niin pisteen Q etäisyys suorasta on

$$d = \frac{|\overrightarrow{P_1Q} \times \overrightarrow{P_1P_2}|}{|\overrightarrow{P_1P_2}|}.$$

Perustelu: Kuvan [fig:dPs02](#) 25 merkinnöin valitaan $\vec{u} = \overrightarrow{P_1P_2}$.

$$d(Q, s) = d(Q, Q') = |\overrightarrow{P_1Q}| \sin \gamma = \frac{|\overrightarrow{P_1Q} \times \overrightarrow{P_1P_2}|}{|\overrightarrow{P_1P_2}|}.$$

□

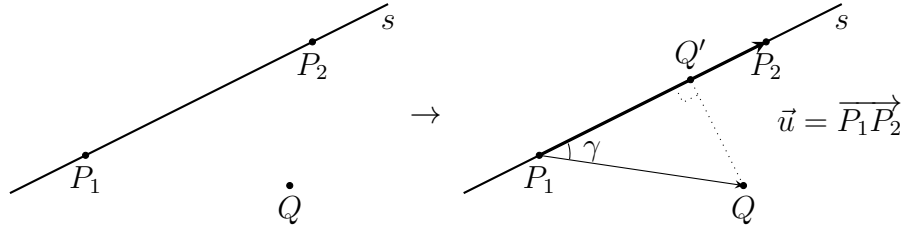


fig:dPs02

Kuva 25. Avaruuden pisteen etäisyys suorasta.

Lauseen [Th:dPs02](#) 122 antama etäisyyden lauseke voidaan tarvittaessa kirjoittaa auki käyttäen pisteiden P_1 , P_2 ja Q koordinaatteja, mutta syntyvä lauseke on melko mutkikas, joten emme tee sitä tässä.

6.7.2 Kahden suoran välinen etäisyys

Ssec:distSS

Seuraavaksi selvitämme kahden avaruudessa olevan suoran etäisyyden. Oletamme, että suora s_1 kulkee pisteen P_1 kautta ja on vektorin $\vec{u}_1$ suuntainen, ja että suora s_2 kulkee pisteen P_2 kautta ja on vektorin $\vec{u}_2$ suuntainen.

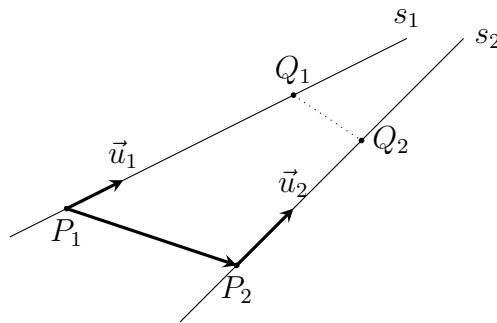


fig:dSS01

Kuva 26. Suoran etäisyys suorasta.

Olkoot $Q_1 \in s_1$ ja $Q_2 \in s_2$ ne suorien pisteet, jotka ovat lähinnä toisiaan. On helppo perustella, että jana $\overline{Q_1Q_2}$ on kohtisuorassa suoria vastaan. Siis $\overline{Q_1Q_2}$ on vektorin $\vec{u}_1 \times \vec{u}_2$ suuntainen. Suuntajana $\overrightarrow{P_1P_2}$ voidaan jakaa komponentteihin seuraavasti:

$$\overrightarrow{P_1P_2} = \overrightarrow{P_1Q_1} + \overrightarrow{Q_1Q_2} + \overrightarrow{Q_2P_2} = \underbrace{\overrightarrow{P_1Q_1} + \overrightarrow{Q_2P_2}}_{\perp \vec{u}_1 \times \vec{u}_2} + \underbrace{\overrightarrow{Q_1Q_2}}_{\parallel \vec{u}_1 \times \vec{u}_2}.$$

Siis suuntajana $\overrightarrow{Q_1Q_2}$ on suuntajanan $\overrightarrow{P_1P_2}$ vektoriprojektion vektorin $\vec{u}_1 \times \vec{u}_2$ suuntaan.

$$\begin{aligned} d(S_1, S_2) &= d(Q_1, Q_2) = |\overrightarrow{Q_1Q_2}| = |\overrightarrow{P_1P_2}_{\vec{u}_1 \times \vec{u}_2}| \\ &= \frac{|\overrightarrow{P_1P_2} \cdot \vec{u}_1 \times \vec{u}_2|}{|\vec{u}_1 \times \vec{u}_2|} \end{aligned} \quad (6.3)$$

7 LINEAARIAVARUDET

Ch:LinSpaces

7.1 Rakenteiden hierarkiaa.

Sec:AlgStruct

Matematiikan historian suurimmat saavutuksen on saatu, kun on tutkittu rakenteita. Seuraavassa luetellaan ne rakenteet, jotka ovat parhaiten kestäneet ajan kulutusta. Seuraavan parin sivun sisältämä luettelo kannattaa ensin lukea kursoorisesti (pysähtymättä yksityiskohtiin). Annettuja määritelmiä ei kaikkia ole nyt numeroitu, ja numeroimattomia ei kysytä tentissä. Jotkut luettelossa olevat termit esiintyvät usein myöhemmillä kursseilla. Nämä määritelmät on numeroitu, ja ne pitää muistaa.

Joukossa A määritelty laskutoimitus $\hat{+}$ on kuvaus $\hat{+} : A \times A \rightarrow A, (a, b) \mapsto a \hat{+} b$. Joskus joukon A alkio kerrotaan joukon K alkiolla. Tälle kertolaskulle käytämme laskutoimituksen kaltaista merkintää $\circ : K \times A \rightarrow A, (k, a) \mapsto k \circ a$.

ath1040c4s1es1

Esimerkki 123 Olkoon $A = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ ja määritellään laskutoimitukset

$$a \hat{+} b = \text{jakojäännös, kun } a + b \text{ jaetaan } 5:\text{llä, ja} \quad (7.1)$$

$$a \hat{*} b = \text{jakojäännös, kun } a \cdot b \text{ jaetaan } 5:\text{llä.} \quad (7.2)$$

Silloin laskutoimitusten laskutaulut ovat seuraavat:

$\hat{+}$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\hat{*}$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Näillä laskutoimituksilla on monia hyvin tuttuja ominaisuuksia. Esim:

$$a \hat{+} (b \hat{+} c) = (a \hat{+} b) \hat{+} c \quad (7.3)$$

$$a \hat{*} (b \hat{+} c) = (a \hat{*} b) \hat{+} (a \hat{*} c) \quad (7.4)$$

Jos tämä oli ensimmäinen kerta, kun luit tästä joukosta ja näistä laskutoimituksista, niin on luonnollista, jos asia tuntuu keinotekoiselta. Myöhemmin tulet tapaamaan tämän tai vastaavan rakenteen uudelleen.

On syntynyt tapa sanoa, että kun oliojoukossa A on määritelty laskutoimitus $\hat{+}$ ja pari $(A, \hat{+})$ toteuttaa joukon aksioomia, niin parilla $(A, \hat{+})$ on tähän aksioomajoukkoon liittyvä strukturi.

Parhaiten ajan koetusta ovat kestäneet seuraavat struktuurit:

Ryhmä. Pari $(A, \hat{+})$ on ryhmä (group), jos se toteuttaa seuraavat aksioomat:
 (1) $a\hat{+}(b\hat{+}c) = (a\hat{+}b)\hat{+}c$, kaikille $a, b, c \in A$, (eli laskutoimitus on liitännäinen)
 (2) On olemassa neutraalialkio $o \in A$ siten, että $o\hat{+}a = a\hat{+}o = a$ kaikille $a \in A$
 (3) jokaiselle $a \in A$ on olemassa vasta-alkio $a' \in A$ siten, että $a'\hat{+}a = a\hat{+}a' = o$.

Abelin ryhmä. Ryhmä $(A, \hat{+})$ on Abelin ryhmä (abelian group), jos se edellisen lisäksi toteuttaa aksiooman:
 (4) $a\hat{+}b = b\hat{+}a$, kaikilla $a, b \in A$, (eli laskutoimitus on vaihdannainen).

Rengas. Kolmikko $(A, \hat{+}, \hat{*})$ on rengas (ring), jos
 (1) $(A, \hat{+})$ on Abelin ryhmä,
 (2) $a\hat{*}(b\hat{*}c) = (a\hat{*}b)\hat{*}c$, kaikilla $a, b, c \in A$, (eli kertolasku on liitännäinen),
 (3) on voimassa osittelulait:

$$a\hat{*}(b\hat{+}c) = (a\hat{*}b)\hat{+}(a\hat{*}c) \quad (7.5)$$

$$(b\hat{+}c)\hat{*}a = (b\hat{*}a)\hat{+}(c\hat{*}a) \quad (7.6)$$

kaikilla $a, b, c \in A$.

Vaihdannainen (kommutatiivinen) rengas. Rengas $(A, \hat{+}, \hat{*})$ on vaihdannainen rengas (commutative ring), jos edellisen lisäksi
 (4) $a\hat{*}b = b\hat{*}a$, kaikilla $a, b \in A$, (eli kertolasku on vaihdannainen).

Ykkösellinen rengas. Rengas $(A, \hat{+}, \hat{*})$ on ykkösellinen rengas (ring with unit), jos
 (4) on olemassa yksikkö-alkio e , jolle $e\hat{*}a = a\hat{*}e = a$ kaikille $a \in A$.

Jakorengas. Ykkösellinen rengas $(A, \hat{+}, \hat{*})$ on jakorengas (division ring), jos
 (5) $e \neq o$,
 (6) jokaiselle $a \in A, a \neq o$ on olemassa kertolaskun käänteisalkio a^{-1} siten, että $(a^{-1})\hat{*}a = a\hat{*}(a^{-1}) = e$.

Kunta. Kommutatiivinen jakorengas $(A, \hat{+}, \hat{*})$ on kunta (field).

Kunnan määritelmä on tässä kumulatiivisessa määritelmäketjussa hämäävän yksinkertainen. On syytä koota yhteenveto kunnan laskutoimitusten ominaisuuksista.

(1) Yhteenlasku on vaihdannainen ja liitännäinen, sillä on neutraalialkio 0 ja jokaisella alkiolla a on yhteenlaskun vasta-alkio $-a$.

(2) Kertolasku on vaihdannainen ja liitännäinen, sillä on ykkösalkio e ja jokaisella yhteenlaskun neutraalialkiosta poikkeavalla alkiolla a on kertolaskun käänteisalkio a^{-1} . Tästä seuraa, että jakolasku on mahdollista.

(3) Yhteen- ja kertolasku noudattavat osittelulakeja, joten lausekkeiden sieventämisessä voidaan käyttää tavanomaisia numerolaskennasta tuttuja keinoja.

Siis kaikki mahdolliset numerolaskennasta tutut laskulait ovat voimassa.

Esimerkki 124 Reaalilukujen joukko varustettuna tavallisella yhteen- ja kertolaskulla, $(\mathbb{R}, +, \cdot)$ on kunta. $(\mathbb{Z}, +, \cdot)$ ei ole kunta. Esimerkissä (123) esiintynyt struktuuri $(\mathbb{Z}_5, \hat{+}, \hat{*})$ on kunta.

Olkoon $(A, \hat{+})$ Abelin ryhmä ja $(K, \bar{+}, \bar{*})$ ykkösellinen rengas. Olkoon lisäksi määritelty skalaarilla kertominen $\circ : K \times A \rightarrow A$.

Moduli. Nelikko $(A, (K, \bar{+}, \bar{*}), \hat{+}, \circ)$ on K-moduli (K -module), jos

- (1) $(A, \hat{+})$ on Abelin ryhmä,
- (2) $(K, \bar{+}, \bar{*})$ on ykkösellinen rengas (ns. kerroinrengas),
- (3) skalaarilla kertominen toteuttaa aksioomat:

$$\lambda \circ (\mu \circ a) = (\lambda \bar{*} \mu) \circ a, \quad \text{kaikilla } \lambda, \mu \in K, a \in A$$

$$(\lambda \bar{+} \mu) \circ a = \lambda \circ a \hat{+} \mu \circ a, \quad \text{kaikilla } \lambda, \mu \in K, a \in A$$

$$\lambda \circ (a \hat{+} b) = \lambda \circ a \hat{+} \lambda \circ b, \quad \text{kaikilla } \lambda \in K, a, b \in A$$

$$e \circ a = a, \quad \text{kaikilla } a \in A. \quad (\text{Tässä } e \text{ on kerroinrenkaan } K \text{ yksikköalkio.})$$

Lineaariavaruus. K -moduli $(A, K, \hat{+}, \circ)$ on K-kertoiminen lineaariavaruus (linear space over K), jos kerroinrengas $(K, \bar{+}, \bar{*})$ on kunta;

Algebra. Viisikko $(A, K, \hat{+}, \hat{*}, \circ)$ on K -algebra, jos

(1) $(A, K, \hat{+}, \circ)$ on K -kertoiminen lineaariavaruus,

(2) $(A, \hat{+}, \hat{*})$ on rengas ja

(3) skalaarilla kertominen ja kertolasku toteuttavat aksiooman:

$$\lambda \circ (a \hat{*} b) = (\lambda \circ a) \hat{*} b = a \hat{*} (\lambda \circ b), \quad \text{kaikille } \lambda \in K, a, b \in A$$

Jakoalgebra. K -algebra $(A, K, \hat{+}, \hat{*}, \circ)$ on K -jakoalgebra, jos

(2) rengas $(A, \hat{+}, \hat{*})$ on jakorengas.

Edellä esiintyvät aksioomat ovat niitä laskulakeja, joita numeroilla laskemaan tottunut ihminen pitää itsestäänselvyyksinä. Matriisien tapauksessa havaitsimme, että eräs numerolaskennan itsestäänselvyys - kertolaskun vaihdantalaki ole voimassa. Tämä on tärkeä havainto, kun nyt ryhdymme laskemaan numeroiden sijasta yleisemmillä olioilla (vektoreilla, matriiseilla, funktioilla jne.)

Kun tarkastelemme jotakin oliojoukkoa ja siinä määriteltyjä laskutoimituksia, niin mitä vahvempi struktuuri sillä on asteikolla "ryhmä < rengas < jakorengas < kunta" tai asteikolla " K -moduli < K -kertoiminen lineaariavaruus < K -algebra < K -jakoalgebra", sitä helpompaa on olioilla laskeminen. Mitä useampi aksiooma on voimassa, sitä useampi numerolaskennan itsestäänselvyys on voimassa.

Kahdessa edellisessä kappaleessa perehdyimme matriiseihin. Neliömatriisit muodostavat renkaan, mutta eivät kommutoivaa rengasta eivätkä jakorengasta. Kun neliömatriisien joukko varustetaan kerroinkunnalla R , ne muodostavat R -algebran, mutta eivät R -jakoalgebraa.

Jatkossa keskitymme alkeellisempaan struktuuriin, nimittäin lineaariavaruuteen, jonka kerroinkunta on reaalilukujen kunta. Tämä merkitsee siis sitä, että olioiden tuloa ei ole määritelty, mutta niiden summa ja olioiden kertominen reaaliluvulla noudattavat kaikkia vanhastaan tuttuja laskulakeja.

7.2 $\mathbb{R}$ - ja $\mathbb{C}$ -kertoiminen lineaariavaruus

Sec:RCLinSpace

Tässä esityksessä tärkein ja eniten käsitelty lineaariavaruus on Euklidinen vektoriavaruus $(\mathbb{R}^n, \mathbb{R}, +, \circ) = \mathbb{R}^n = E^n$ (engl. Euclidean Vector Space), jonka kerroinkunta on $\mathbb{R}$.

Merkintä $\mathbb{R}^n$ on loogisempi, mutta merkintää E^n käytetään paljon. Merkintä korostaa, sitä, että vektoriavaruus ei ole pelkkä karteesinen tulo-joukko, vaan sillä on vektoriavaruuden rakenne, eli yhteenlasku ja kerroinkunta. Kumpikin merkintä esiintyy jatkossa.

Kompleksinen lineaariavaruus $(\mathbb{C}^n, \mathbb{C}, +, \circ) = \mathbb{C}^n$ (engl. Complex Vector Space) on tärkeä, mutta se jätetään nyt hetkeksi aikaa taka-alalle. Asiaan palataan kappaleessa ^{$\mathbb{C}^n$} (??) sivulla ^{$\mathbb{C}^n$} ??.

Usein virheellisesti lineaariavaruus käsitetään samaksi, kuin E^n , mutta tämä ei ole oikein. Monet funktiojoukot voidaan helposti varustaa lineaariavaruuden struktuurilla. Tällaiset lineaariset funktioavaruudet ovat erittäin hyödyllisiä olioita ja monessa suhteessa E^n :stä poikkeavia. Seuraavassa yleiset lineaariavaruuksiin liittyvät määritelmät ja tulokset muotoillaan koskemaan reaalikertoimista lineaariavaruutta $L = (L, \mathbb{R}, +, \circ)$. Tällainen tulos on sellaisenaan voimassa myös, kun $L = E^n$. Jos väite esitetään E^n :lle, niin se ei välttämättä ole voimassa kaikille lineaariavaruuksille.

Puramme ensin edellisen kappaleen hierarkiseen määritelmäketjuun piiloutuneet ominaisuudet seuraavaan määritelmään. Huomaa, että tässä oikeastaan toistetaan jo todettua.

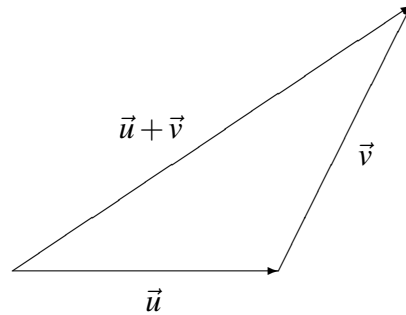
Määritelmä 125 Nelikko $(L, \mathbb{R}, +, \circ)$ on reaalinen lineaariavaruus (eli reaalinen vektoriavaruus), jos yhteenlasku $L \times L \rightarrow L, (\vec{u}, \vec{v}) \mapsto \vec{a} + \vec{b}$ ja reaaliluvulla kertominen

$\mathbb{R} \times L \rightarrow L, (\lambda, \vec{u}) \mapsto \lambda \circ \vec{u}$ toteuttavat seuraavat ehdot:

- (A1) $\vec{u} + \vec{v} = \vec{v} + \vec{u},$ kaikilla $\vec{u}, \vec{v} \in L,$
- (A2) $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ kaikilla $\vec{u}, \vec{v}, \vec{w} \in L,$
- (A3) on olemassa (yksikäsitteinen) nolla-alkio $\vec{0} \in L$ siten, että $\vec{u} + \vec{0} = \vec{u}$ kaikilla $\vec{u} \in L,$
- (A4) jokaista $\vec{u} \in L$ vastaa (yksikäsitteinen) vasta-alkio $-\vec{u}$ siten, että $\vec{u} + (-\vec{u}) = \vec{0},$ (merkitään $\vec{v} - \vec{u} = \vec{v} + (-\vec{u})$)
- (A5) $\lambda(\mu \circ \vec{u}) = (\lambda \cdot \mu) \circ \vec{u},$ kaikilla $\lambda, \mu \in \mathbb{R}, \vec{u} \in L,$
- (A6) $(\lambda + \mu) \circ \vec{u} = \lambda \circ \vec{u} + \mu \circ \vec{u},$ kaikilla $\lambda, \mu \in \mathbb{R}, \vec{u} \in L,$
- (A7) $\lambda \circ (\vec{u} + \vec{v}) = \lambda \circ \vec{u} + \lambda \circ \vec{v},$ kaikilla $\lambda \in \mathbb{R}, \vec{u}, \vec{v} \in L,$
- (A8) $1 \circ \vec{u} = \vec{u}$ kaikilla $\vec{u} \in L,$

Avaruuden L alkioita sanomme vektoreiksi. Jos laskutoimitukset ovat asiayhteydestä selvät sanomme, että ” L on reaalinen lineaariavaruus”.

Esimerkki 126 Tason vektoreiden joukko on reaalinen lineaariavaruus



Tason vektori voidaan esittää lukuparina $\vec{u} = (x, y)$, jolloin $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ ja $\lambda \circ (x, y) = (\lambda \cdot x, \lambda \cdot y)$. Suora lasku osoittaa, että nolla-alkio on $(0, 0)$ ja aksioomat (A1 - A8) ovat voimassa.

Esimerkki 127 Joukko $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n)^T \mid x_i \in \mathbb{R}, \text{ kaikille } 1 \leq i \leq n\}$ varustettuna pystyvektorien yhteenlaskulla ja reaalityluvulla kertomisella

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad \lambda \circ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot a_1 \\ \lambda \cdot a_2 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix}$$

on reaalinen lineaariavaruus $E^n = (\mathbb{R}^n, \mathbb{R}, +, \circ)$.

Esimerkki 128 $m \times n$ -matriisien joukko on reaalinen lineaariavaruus.

Esimerkki 129 Olkoon $A \neq \emptyset$ ja olkoon $\mathcal{F}(A, \mathbb{R})$ kaikkien kuvausten $f: A \rightarrow \mathbb{R}$ joukko. Joukko $\mathcal{F}(A, \mathbb{R})$ on reaalinen lineaariavaruus, kun määrittelemme

$$\begin{aligned}(f+g)(x) &= f(x) + g(x), \quad \forall x \in A \\ (\lambda \circ f)(x) &= \lambda \cdot f(x), \quad \forall x \in A\end{aligned}$$

Lause 130 Olkoon L reaalinen lineaariavaruus. Silloin

$$\begin{aligned}(1) \quad & 0 \circ \vec{u} = \vec{0}, \quad \forall \vec{u} \in L, \\ (2) \quad & \lambda \circ \vec{0} = \vec{0}, \quad \forall \lambda \in \mathbb{R}, \\ (3) \quad & [\lambda \circ \vec{u} = \vec{0}] \Leftrightarrow [\lambda = 0 \text{ tai } \vec{u} = \vec{0}], \\ (4) \quad & (-1) \circ \vec{u} = -\vec{u}, \quad \forall \vec{u} \in L.\end{aligned}$$

Todistus.

$$\begin{aligned}(1) \quad & 0 \circ \vec{u} \stackrel{(A3)}{=} 0 \circ \vec{u} + \vec{0} \stackrel{(A4)}{=} 0 \circ \vec{u} + (0 \circ \vec{u} + (- (0 \circ \vec{u}))) \stackrel{(A3)}{=} (0 \circ \vec{u} + 0 \circ \vec{u}) - (0 \circ \vec{u}) \\ & \stackrel{(A6)}{=} (0 + 0) \circ \vec{u} - (0 \circ \vec{u}) = \vec{0}\end{aligned}$$

$$(2) \quad \lambda \circ \vec{0} \stackrel{(A3)}{=} \lambda \circ \vec{0} + \vec{0} \stackrel{(A4)}{=} \lambda \circ \vec{0} + \lambda \circ \vec{0} + (- (\lambda \circ \vec{0})) \stackrel{(A7)}{=} \lambda \circ (\vec{0} + \vec{0}) - (\lambda \circ \vec{0}) = \vec{0}$$

$$(3) \quad \boxed{\Leftarrow} \text{ seuraa kohdista (1) ja (2).}$$

$$\boxed{\Rightarrow} \text{ Olkoon } \lambda \circ \vec{u} = \vec{0}. \text{ Jos } \lambda \neq 0, \text{ niin } \lambda^{-1} \text{ on olemassa ja}$$

$$\vec{u} \stackrel{(A8)}{=} 1 \circ \vec{u} = (\lambda^{-1} \lambda) \circ \vec{u} \stackrel{(A5)}{=} \lambda^{-1} (\lambda \circ \vec{u}) = \lambda^{-1} \circ \vec{0} \stackrel{(1)}{=} \vec{0}$$

$$(4) \quad \vec{u} + (-1) \circ \vec{u} \stackrel{(A8)}{=} 1 \circ \vec{u} - 1 \circ \vec{u} \stackrel{(A6)}{=} (1 - 1) \circ \vec{u} = 0 \circ \vec{u} \stackrel{(1)}{=} \vec{0}$$

□

Määritelmä 131 Lineaariavaruuden L osajoukko M on lineaarinen aliavaruus, jos

$$\begin{aligned}(1) \quad & \vec{0} \in M \\ (2) \quad & \vec{u}, \vec{v} \in M \Rightarrow \vec{u} + \vec{v} \in M \\ (3) \quad & \lambda \in \mathbb{R}, \vec{u} \in M \Rightarrow \lambda \circ \vec{u} \in M\end{aligned}$$

(Engl. subspace)

Lineaarinen aliavaruus siis sisältää nollavektorin ja se on yhteenlaskun ja skalaarilla kertomisen suhteen suljettu osajoukko.

Lause 132 Olkoon M_1 ja M_2 kaksi lineaariavaruuden L aliavaruutta. Silloin:

(1) Aliavaruuksien leikkaus

$$M_1 \cap M_2 = \{\vec{u} \in L \mid \vec{u} \in M_1 \text{ ja } \vec{u} \in M_2\}$$

on L :n lineaarinen aliavaruus.

(1) Aliavaruuksien summa

$$M_1 + M_2 = \{\vec{u} + \vec{v} \in L \mid \vec{u} \in M_1 \text{ ja } \vec{v} \in M_2\}$$

on L :n lineaarinen aliavaruus.

Todistus. HT

□

Tyhjä joukko ei ole aliavaruus, mutta pelkän nollavektorin sisältävä joukko $\{\vec{0}\}$ on aliavaruus (triviaali aliavaruus). Koska jokainen aliavaruus sisältää nolavektorin, sisältää kahden aliavaruuden leikkaus aina ainakin nollavektorin.

Määritelmä 133 Jos M_1 ja M_2 ovat kaksi lineaariavaruuden L aliavaruutta, joiden leikkaus on triviaali ($M_1 \cap M_2 = \{\vec{0}\}$), niin niiden summaa sanotaan suoraksi summaksi ja sitä merkitään

$$M_1 \oplus M_2.$$

c4s2th1

Lause 134 Jos lineaariavaruus L on kahden aliavaruutensa M_1 ja M_2 suora summa, eli $L = M_1 \oplus M_2$, niin jokaista vektoria $\vec{u} \in L$ vastaa yhdet ja vain yhdet komponenttivetektorit $\vec{u}_1 \in M_1$ ja $\vec{u}_2 \in M_2$ siten, että

$$\vec{u} = \vec{u}_1 + \vec{u}_2.$$

Todistus. Koska $L = M_1 \oplus M_2$ on olemassa ainakin yhdet komponenttivetektorit $\vec{u}_1 \in M_1$ ja $\vec{u}_2 \in M_2$, joille $\vec{u} = \vec{u}_1 + \vec{u}_2$. On siis osoitettava, että muita ei löydy.

Olkoon $\vec{w}_1 \in M_1$ ja $\vec{w}_2 \in M_2$ toiset komponentit siten, että $\vec{u} = \vec{w}_1 + \vec{w}_2$. Nyt

$$\begin{aligned} \vec{u}_1 + \vec{u}_2 &= \vec{w}_1 + \vec{w}_2 \\ \Leftrightarrow \underbrace{\vec{u}_1 - \vec{w}_1}_{\in M_1} &= \underbrace{\vec{w}_2 - \vec{u}_2}_{\in M_2} \end{aligned}$$

Koska $M_1 \cap M_2 = \{\vec{0}\}$, niin viimeinen yhtälö voi olla tosi vain silloin, kun LHS ja RHS ovat nollavektoreita. Siis $\vec{w}_1 = \vec{u}_1$ ja $\vec{w}_2 = \vec{u}_2$. $\square$

Seuraavassa määritelty joukko on hyvin tärkeä ja usein esiintyvä aliavaruus.

Kerrataan ensin jo luvussa 2 esiintynyt lineaarikombinaation käsite.

Määritelmä 135 Olkoot $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n \in L$ lineaariavaruuden L vektoreita, ja olkoot $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ reaalityyppisiä lukuja. Silloin vektori $\lambda_1 \circ \vec{u}_1 + \lambda_2 \circ \vec{u}_2 + \dots + \lambda_n \circ \vec{u}_n \in L$ on lineaarikombinaatio vektoreista $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ kertoimin $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. (Engl. linear combination.)

Määritelmä 136 Olkoon L reaalinen lineaariavaruus ja olkoon M sen osajoukko. Joukon M virittämä aliavaruus, $\text{span}(M)$, on kaikkien M :n alkioiden äärellisten lineaarikombinaatioiden joukko. Sovimme edellisen lisäksi, että tyhjä joukko virittää triviaalin aliavaruuden, $\text{span}(\emptyset) = \{\vec{0}\}$. (Engl. subspace spanned by M .)

Lause 137 Olkoon L reaalinen lineaariavaruus ja olkoon M sen osajoukko. Silloin $\text{span}(M)$ on L :n lineaarinen aliavaruus.

Todistus. Jos $M = \emptyset$, niin $\text{span}(M) = \{\vec{0}\}$ on aliavaruus. Jos $M \neq \emptyset$, niin on olemassa $\vec{u} \in M$ joten $\vec{0} = 0 \circ \vec{u} \in \text{span}(M)$. Siis aliavaruuden määritelmän ehto (1) on voimassa.

(2) Jos $\vec{v}, \vec{w} \in \text{span}(M)$, niin ne voidaan esittää äärellisinä lineaarikombinaatioina

$$\begin{aligned}\vec{v} &= \lambda_1 \circ \vec{u}_1 + \dots + \lambda_m \circ \vec{u}_m, \text{ ja} \\ \vec{w} &= \lambda_{m+1} \circ \vec{u}_{m+1} + \dots + \lambda_n \circ \vec{u}_n, \text{ joten} \\ \vec{v} + \vec{w} &= \lambda_1 \circ \vec{u}_1 + \dots + \lambda_n \circ \vec{u}_n \in \text{span}(M)\end{aligned}$$

(3) Jos $\lambda \in \mathbb{R}$ ja $\vec{v} = \lambda_1 \circ \vec{u}_1 + \dots + \lambda_m \circ \vec{u}_m$, niin

$$\lambda \circ \vec{v} = (\lambda \cdot \lambda_1) \circ \vec{u}_1 + \dots + (\lambda \cdot \lambda_m) \circ \vec{u}_m \in \text{span}(M)$$

$\square$

Esimerkki 138 Olkoon L tasovektoreiden muodostama vektoriavaruus. Olkoon $\vec{i}$ yksikön mittainen vektori x -akselin suuntaan, ja olkoon $\vec{j}$ yksikön mittainen vektori y -akselin suuntaan. Silloin $L = \text{span}(\vec{i}, \vec{j})$.

7.3 Sisätulo ja normi

Sec:InnerPNorm

Etenemme asteittain niin, että ensin esittelemme E^n :ssä sisätulon ja kolme erilaista normia. Näiden käsitteiden avulla määrittelemme kohtisuoruuden ja pituuden käsitteet. Lopuksi vielä annamme määritelmät yleisen lineaariavaruuden tapauksessa.

7.3.1 Sisätulo E^n :ssä

Määritelmä 139 Kahden vektorin $\vec{a} = (a_1, a_2, \dots, a_n)^T \in \mathbb{R}^n$ ja $\vec{b} = (b_1, b_2, \dots, b_n)^T \in \mathbb{R}^n$ sisätulo $\langle \vec{a} | \vec{b} \rangle$ määritellään kaavalla

$$\langle \vec{a} | \vec{b} \rangle = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \vec{a} \cdot \vec{b}$$

(Kyseessä on siis jo ennemmin määritelty pistetulo, mutta nyt otamme käyttöön kulmasulkumerkinnän.)

Määritelmästä seuraa välittömästi, että sisätulo on vaihdannainen ja kummankin tekijänsä suhteen lineaarinen. Lisäksi nollavektori on ainoa vektori, jonka sisätulo itsensä kanssa on nolla. Kokoamme nämä selvät ominaisuudet lauseeksi, jotta ne myöhemmin löytyvät helposti.

c4s3th1 **Lause 140** Olkoot $\vec{u}, \vec{v} \in \mathbb{R}^n$ ja $\lambda \in \mathbb{R}$. Silloin

- (S1) $\langle \vec{u} | \vec{v} \rangle = \langle \vec{v} | \vec{u} \rangle,$
- (S2) $\langle \vec{u} | \vec{v} + \vec{w} \rangle = \langle \vec{u} | \vec{v} \rangle + \langle \vec{u} | \vec{w} \rangle,$
- (S3) $\lambda \cdot \langle \vec{u} | \vec{v} \rangle = \langle \lambda \vec{u} | \vec{v} \rangle = \langle \vec{u} | \lambda \vec{v} \rangle,$
- (S4) $\langle \vec{u} | \vec{u} \rangle \geq 0$ ja $\langle \vec{u} | \vec{u} \rangle = 0$ jos ja vain jos $\vec{u} = \vec{0}$

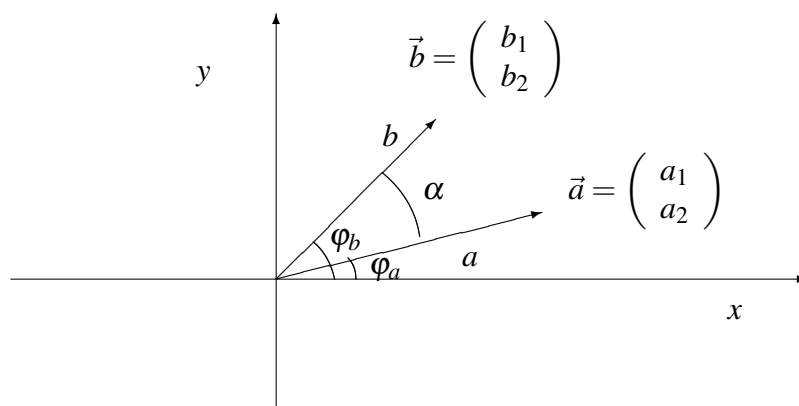
Todistus. HT.

Kaksi- ja kolmiulotteisissa Euklididissa vektoriavaruuksissa E^2 ja E^3 sisätulolla on seuraava koulukurssista tuttu geometrinen tulkinta.

Lause 141 Tarkastellaan E^2 :n vektoreita $\vec{a} = (a_1, a_2)^T$ ja $\vec{b} = (b_1, b_2)^T$. Vektoreiden pituudet ovat $a = \sqrt{a_1^2 + a_2^2}$ ja $b = \sqrt{b_1^2 + b_2^2}$. Jos vektoreiden välinen kulma on α , niin

$$\langle \vec{a} | \vec{b} \rangle = a \cdot b \cdot \cos \alpha$$

c4s3fig1



Kuva 27. vektoreiden $\vec{a}$ ja $\vec{b}$ sisätulo.

Todistus. Olkoon vektorin $\vec{a}$ ja vaaka-akselin välinen kulma φ_a ja vastaavasti vektorin $\vec{b}$ ja vaaka-akselin välinen kulma φ_b . Silloin (Ks. kuva (c4s3fig1))

$$\begin{cases} a_1 = a \cos \varphi_a \\ a_2 = a \sin \varphi_a \end{cases} \quad \text{ja} \quad \begin{cases} b_1 = b \cos \varphi_b \\ b_2 = b \sin \varphi_b \end{cases}.$$

Trigonometristen funktioiden kaavojen mukaan $\cos(x-y) = \cos x \cos y + \sin x \sin y$. Siis

$$\begin{aligned} \langle \vec{a} | \vec{b} \rangle &= a_1 b_1 + a_2 b_2 = a \cos(\varphi_a) \cdot b \cos(\varphi_b) + a \sin(\varphi_a) \cdot b \sin(\varphi_b) \\ &= ab \cos(\varphi_b - \varphi_a) = ab \cos \alpha \end{aligned}$$

□

Edeltävä tulos pitää paikkansa myös E^3 :ssa. Emme nyt todista tätä, vaan myöhemmin, palaamme asiaan. Joka tapauksessa teemme nyt seuraavat yleistykset.

Määritelmä 142 Vektorit $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$ ovat

(1) kohtisuorassa toisiaan vastaan ($\vec{a} \perp \vec{b}$) eli ortogonaaliset (engl. *orthogonal*), jos

$$\langle \vec{a} | \vec{b} \rangle = 0.$$

(2) yhdensuuntaiset ($\vec{a} \parallel \vec{b}$), jos on olemassa reaaliluku $\lambda \in \mathbb{R}$ siten, että

$$\vec{a} = \lambda \vec{b}$$

(3) samansuuntaiset ($\vec{a} \uparrow \vec{b}$), jos on olemassa positiivinen reaaliluku $\lambda \in \mathbb{R}$ siten, että

$$\vec{a} = \lambda \vec{b}, \lambda > 0$$

(4) vastakkaissuuntaiset ($\vec{a} \updownarrow \vec{b}$), jos on olemassa negatiivinen reaaliluku $\lambda \in \mathbb{R}$ siten, että

$$\vec{a} = \lambda \vec{b}, \lambda < 0$$

Olkon $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$ mitkä tahansa kaksi vektoria. Vektori $\vec{a}$ voidaan aina lausua kahden vektorin summana $\vec{a} = \vec{a}_b + \vec{a}_\perp$ niin, että ensimmäinen vektoreista on $\vec{b}$:n suuntainen ja toinen on $\vec{b}$:tä vastaan kohtisuorassa. Jos nimittäin asetamme

$$\vec{a}_b = \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} \quad (7.7)$$

$$\vec{a}_\perp = \vec{a} - \vec{a}_b = \vec{a} - \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} \quad (7.8)$$

niin selvästi vektoreiden summa on $\vec{a}$, ensimmäinen vektori on yhdensuuntainen $\vec{b}$:n kanssa ja lineaarisuuden perusteella (ks. (S2) ja (S3) lauseessa (I40))

$$\langle \vec{a}_\perp | \vec{b} \rangle = \left\langle \vec{a} - \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} \middle| \vec{b} \right\rangle \quad (7.9)$$

$$= \langle \vec{a} | \vec{b} \rangle - \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \langle \vec{b} | \vec{b} \rangle = 0 \quad (7.10)$$

c4s3esim1

Esimerkki 143 Tarkastellaan esimerkkiä, jonka vektorit voimme piirtää (ks. kuva (??) sivulla I33). Olkoon $\vec{a} = (2, 6)^T \in \mathbb{R}^2$ ja $\vec{b} = (8, 4)^T \in \mathbb{R}^2$. Silloin

$$\vec{a}_b = \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b} = \left(\frac{2 \cdot 8 + 6 \cdot 4}{8 \cdot 8 + 4 \cdot 4} \right) \vec{b} = 0.5 \begin{pmatrix} 8 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad (7.11)$$

$$\vec{a}_\perp = \vec{a} - \vec{a}_b = \begin{pmatrix} 2 \\ 6 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix} \quad (7.12)$$

Määritelmä 144 Olkoon $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$. Vektori

$$\vec{a}_b = \frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle} \vec{b}$$

on vektorin $\vec{a}$ projektio vektorin $\vec{b}$ suuntaan (vektoriprojektio).

Määritelmä 145 Olkoon $M \subseteq \mathbb{R}^n$ lineaarinen aliavaruus.

(1) Olkoon $\vec{u} \in \mathbb{R}^n$ vektori. Jos

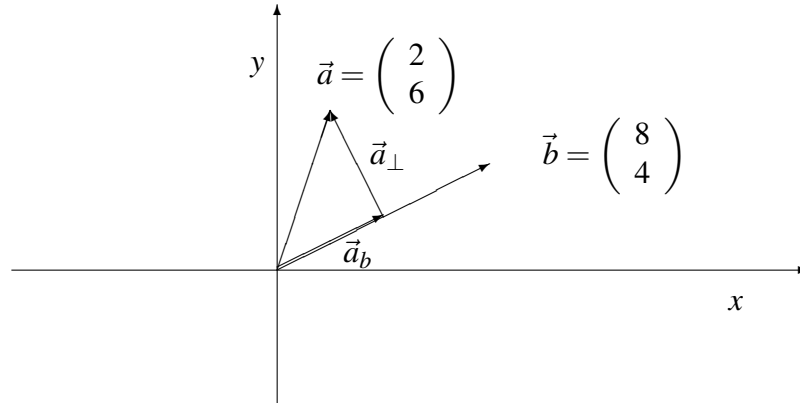
$$\vec{u} \perp \vec{v}, \quad \forall \vec{v} \in M,$$

niin sanomme, että $\vec{u}$ on kohtisuorassa aliavaruutta M vastaan ja merkitsemme

$$\vec{u} \perp M.$$

(2) Aliavaruuden M ortogonaalikomplementti on

$$M^\perp = \{ \vec{u} \in \mathbb{R}^n \mid \vec{u} \perp M \}.$$



c4s3fig2

Kuva 28. Vektorin $\vec{a}$ projektio vektorin $\vec{b}$ suuntaan.

Lause 146 Olkoon $M \subseteq \mathbb{R}^n$ lineaarinen aliavaruus. Silloin myös $M^\perp$ on lineaarinen aliavaruus.

Todistus. On siis osoitettava, että kaikilla $\vec{a}, \vec{b} \in M^\perp$ ja $\lambda \in \mathbb{R}$

$$\begin{aligned}\vec{a} + \vec{b} &\in M^\perp \text{ ja} \\ \lambda \vec{a} &\in M^\perp\end{aligned}$$

Olkoon nyt $\vec{a}, \vec{b} \in M^\perp$ ja $\lambda \in \mathbb{R}$ kuten edellä. Silloin

$$\begin{aligned}\forall \vec{c} \in M: \quad \langle \vec{a} + \vec{b} | \vec{c} \rangle &= \langle \vec{a} | \vec{c} \rangle + \langle \vec{b} | \vec{c} \rangle = 0 \Rightarrow \vec{a} + \vec{b} \in M^\perp \\ \forall \vec{c} \in M: \quad \langle \lambda \vec{a} | \vec{c} \rangle &= \lambda \langle \vec{a} | \vec{c} \rangle = 0 \Rightarrow \lambda \vec{a} \in M^\perp\end{aligned}$$

□

Käytännössä suurin osa $\mathbb{R}^n$:n aliavaruuksista on viritettyjä, joten seuraava lause on lähes aina käytettävissä, kun tutkimme kohtisuoruutta.

c4s3th2

Lause 147 Olkoon $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q) \subseteq \mathbb{R}^n$ vektoreiden $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ virittämä aliavaruus ja olkoon $\mathbf{V}$ matriisi, jonka k :s sarake on $\vec{v}_k$. Olkoon $\vec{u} \in \mathbb{R}^n$. Silloin

$$\vec{u} \perp M. \Leftrightarrow \vec{u} \perp \vec{v}_k, \forall k = 1, 2, \dots, q \quad (7.13)$$

$$\Leftrightarrow \mathbf{V}^T \vec{u} = \vec{0} \quad (7.14)$$

Todistus. Helppo HT.

□

Jos siis vektori on kohtisuorassa jokaista virittäjää vastaan, se on kohtisuorassa koko aliavaruutta vastaan.

Seuraava lause on ”tekninen temppu”, mutta se tekee monet tarkastelut helpoiksi.

Lause 148 Olkoon $\vec{b} \in \mathbb{R}^n$, $\vec{c} \in \mathbb{R}^m$ ja olkoon $\mathbf{A}$ $m \times n$ -matriisi. Silloin

$$\langle \mathbf{A}\vec{b} | \vec{c} \rangle = \langle \vec{b} | \mathbf{A}^T \vec{c} \rangle$$

Todistus.

$$\begin{aligned} \langle \mathbf{A}\vec{b} | \vec{c} \rangle &= \sum_{k=1}^m (\mathbf{A}\vec{b})_k c_k = \sum_{k=1}^m \left(\sum_{j=1}^n a_{kj} b_j \right) c_k \\ &= (a_{11}b_1 + a_{12}b_2 + \cdots + a_{1n}b_n)c_1 \\ &\quad + (a_{21}b_1 + a_{22}b_2 + \cdots + a_{2n}b_n)c_2 \\ &\quad \cdots + (a_{m1}b_1 + a_{m2}b_2 + \cdots + a_{mn}b_n)c_m \\ &= b_1(a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m) \\ &\quad + b_2(a_{12}c_1 + a_{22}c_2 + \cdots + a_{m2}c_m) \\ &\quad \cdots + b_n(a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m) \\ &= \sum_{j=1}^n b_j \left(\sum_{k=1}^m a_{kj} c_k \right) = \sum_{j=1}^n b_j (\mathbf{A}^T \vec{c})_j = \langle \vec{b} | \mathbf{A}^T \vec{c} \rangle \end{aligned}$$

□

Edellisen lauseen yhteydessä on tärkeätä muistaa transpoosin peruskaavat

$$(\mathbf{A}^T)^T = \mathbf{A} \quad (7.15)$$

$$(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T \quad (7.16)$$

$$(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} \quad (7.17)$$

Edellä esimerkissä ^{lc4s3esim1}(143) jaoinme vektorin $\vec{a} \in \mathbb{R}^n$ kahteen osaan $\vec{a} = \vec{a}_b + \vec{a}^\perp$ siten, että ensimmäinen komponentti on vektorin $\vec{b}$ suuntainen ja toinen on vektoria $\vec{b}$ vastaan kohtisuora. Lineaarisen aliavaruuden $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q)$ tapauksessa voimme erään ehdon vallitessa jakaa vektorin $\vec{a} \in \mathbb{R}^n$ kahteen komponenttiin $\vec{a} = \vec{a}_M + \vec{a}_{M^\perp}$ siten että $\vec{a}_M \in M$ ja $\vec{a}_{M^\perp} \in M^\perp$. Komponentit osoittautuvat yksikäsitteisiksi.

Ehto, joka nyt joudutaan olettamaan, voidaan myöhemmin kokonaan poistaa pienellä lauseen uudelleenmuotoilulla. Voisimme lykätä tarkastelun myöhemmäksi, mutta otamme tämän nyt esimerkkinä sisätulon ominaisuuksista.

c4s3th6

Lause 149 Olkoon $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q) \subseteq \mathbb{R}^n$ vektoreiden $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ virittämä lineaarinen aliavaruus ja olkoon $\mathbf{V}$ matriisi, jonka k :s sarake on $\vec{v}_k$.

Teemme nyt erikoisen oletuksen olettamalla, että matriisi $\mathbf{V}^T \mathbf{V}$ on säännöllinen (eli on olemassa $(\mathbf{V}^T \mathbf{V})^{-1}$).

Silloin

$$\mathbb{R}^n = M \oplus M^\perp.$$

Todistus. Ensin osoitamme, että jokaisella $\vec{a} \in \mathbb{R}^n$ on olemassa vektorit $\vec{a}_M \in M$ ja $\vec{a}_{M^\perp} \in M^\perp$ siten, että $\vec{a} = \vec{a}_M + \vec{a}_{M^\perp}$.

Tekemämme oletuksen perusteella voimme asettaa

$$\vec{a}_M = \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a}, \quad \text{ja} \quad \vec{a}_{M^\perp} = \vec{a} - \vec{a}_M$$

(1) Jos merkitsemme $\vec{c} = (c_1, c_2, \dots, c_q)^T = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a}$, niin

$$\vec{a}_M = \mathbf{V} \vec{c} = c_1 \vec{v}_1 + \dots + c_q \vec{v}_q \in M.$$

Siis ensimmäinen komponentti on aliavaruuden vektori.

(2) Toiseksi tutkimme toisen komponentin kohtisuoruutta lauseen (I47) mukaisesti

$$\mathbf{V}^T \vec{a}_{M^\perp} = \mathbf{V}^T \vec{a} - \mathbf{V}^T \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} = \vec{0}$$

Siis toinen komponentti on kohtisuorassa aliavaruutta M vastaan.

(3) Kolmanneksi toteamme summan suoraksi osoittamalla, että $M \cap M^\perp = \{\vec{0}\}$. Olkoon siis $\vec{u} \in M \cap M^\perp$. Silloin erityisesti

$$\begin{aligned} \vec{u} \perp \vec{u} &\Leftrightarrow \langle \vec{u} | \vec{u} \rangle = u_1^2 + u_2^2 + \dots + u_n^2 = 0 \\ &\Leftrightarrow \vec{u} = \vec{0} \end{aligned}$$

□

Lause sanoo sen, että mikä tahansa vektori voidaan lausua kahden vektorin summan niin, että ensimmäinen yhteenlaskettava kuuluu aliavaruuteen M ja toinen yhteenlaskettava on kohtisuorassa aliavaruutta M vastaan. Tämä on jatkossa erittäin hyödyllinen tulos. Lauseessa oleva lisäehto $\det(\mathbf{V}^T \mathbf{V}) = 0$ tulee jatkossa hoidettua pois.

Esimerkki 150 Edellisen lauseen todistus on hämäävän helppo. Testaamme vielä fundamentaalilla tavalla, ovatko vektorit $\vec{a}_{M^\perp}$ ja $\vec{v}_k$, $1 \leq k \leq q$ kohtisuorassa:

$$\langle \vec{a}_{M^\perp} | \vec{v}_k \rangle = \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} | \vec{v}_k \rangle \quad (7.18)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}(\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} | \mathbf{V} \vec{e}_k \rangle \quad (7.19)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a} | \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.20)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | ((\mathbf{V}^T \mathbf{V})^{-1})^T \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.21)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | ((\mathbf{V}^T \mathbf{V})^T)^{-1} \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.22)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \mathbf{V} \vec{e}_k \rangle \quad (7.23)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \mathbf{V}^T \vec{a} | \vec{e}_k \rangle \quad (7.24)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \vec{a} | \mathbf{V} \vec{e}_k \rangle \quad (7.25)$$

$$= \langle \vec{a} | \vec{v}_k \rangle - \langle \vec{a} | \vec{v}_k \rangle = 0 \quad (7.26)$$

7.3.2 $\mathbb{R}^n$:n normeja

Vektoreiden $\vec{a} = (a_1, a_2)^T \in \mathbb{R}^2$ ja $\vec{b} = (b_1, b_2, b_3)^T \in \mathbb{R}^3$ pituudet on koulukurssissa laskettu kaavoilla

$$\begin{aligned} |\vec{a}| &= \sqrt{a_1^2 + a_2^2}, \text{ ja} \\ |\vec{b}| &= \sqrt{b_1^2 + b_2^2 + b_3^2}. \end{aligned}$$

Normi yleistää pituuden idean $\mathbb{R}^n$:ään. Perusidea on se, että pienellä vektorilla on pieni normi ja suurella vektorilla on suuri normi. Kaksi vektoria ovat lähellä toisiaan, jos niiden erotuksen normi on pieni.

Määrittelemme kolme erilaista normia.

c4s3defnorm

Määritelmä 151 Vektorin $\vec{a} = (a_1, a_2, \dots, a_n)^T \in \mathbb{R}^n$ 1-normi $\|\vec{a}\|_1$, 2-normi $\|\vec{a}\|_2$ ja ääretön-normi $\|\vec{a}\|_\infty$ määritellään kaavoilla

$$(1) \quad \|\vec{a}\|_1 = |a_1| + |a_2| + \dots + |a_n| = \sum_{k=1}^n |a_k|$$

$$(2) \quad \|\vec{a}\|_2 = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2} = \sqrt{\langle \vec{a} | \vec{a} \rangle}$$

$$(3) \quad \|\vec{a}\|_\infty = \max\{|a_1|, |a_2|, \dots, |a_n|\} = \max_{1 \leq k \leq n} |a_k|$$

CauchySchwarz

Lause 152 (Cauchy-Schwarz) Kaikille $\vec{a} \in \mathbb{R}^n$ ja $\vec{b} \in \mathbb{R}^n$ on voimassa epäyhtälö

$$|\langle \vec{a} | \vec{b} \rangle| \leq \|\vec{a}\|_2 \|\vec{b}\|_2$$

Todistus. Seuraavassa todistuksessa normi $\|\cdot\|$ tarkoittaa 2-normia. Sisätulon ominaisuuksista seuraa, että kaikilla $\lambda \in \mathbb{R}$ on voimassa

$$0 \leq \langle \vec{a} + \lambda \vec{b} | \vec{a} + \lambda \vec{b} \rangle = \|\vec{a}\|^2 + \lambda^2 \|\vec{b}\|^2 + 2\lambda \langle \vec{a} | \vec{b} \rangle$$

Kun tähän sijoitetaan

$$\lambda = -\frac{\langle \vec{a} | \vec{b} \rangle}{\langle \vec{b} | \vec{b} \rangle},$$

niin saadaan

$$0 \leq \|\vec{a}\|^2 - \frac{\langle \vec{a} | \vec{b} \rangle^2}{\|\vec{b}\|^2},$$

mistä seuraa lauseen väite. □

Kaikki kolme normia toteuttavat kolme yksinkertaista ominaisuutta.

Lause 153 Kaikilla $\vec{u}, \vec{v} \in \mathbb{R}^n$ ja $\lambda \in \mathbb{R}$ edellä määritellyt 1-normi, 2-normi ja ääretön-normi toteuttavat seuraavat ehdot

$$(N1) \quad \|\vec{u}\| \geq 0 \text{ ja } \|\vec{u}\| = 0 \text{ jos ja vain jos } \vec{u} = \vec{0}$$

$$(N2) \quad \|\lambda \vec{u}\| = |\lambda| \cdot \|\vec{u}\|,$$

$$(N3) \quad \|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|, \quad (\text{kolmioepäyhtälö})$$

Todistus. Kohdat (N1) ja (N2) seuraavat suoraan määritelmistä, mutta kolmioepäyhtälö kannattaa käydä tarkemmin läpi: Merkitään todistuksen ajan $\vec{u} + \vec{v} = \vec{w}$ **1-normi:** Olkoon x ja y reaalilukuja. Tarvittaessa nimet vaihtaen voimme olettaa, että $|x| \geq |y|$. Nyt

$$|x+y| = \begin{cases} |x|+|y|, & \text{kun } x \text{ ja } y \text{ ovat saman merkkiset, ja} \\ |x|-|y| \leq |x|+|y|, & \text{kun } x \text{ ja } y \text{ ovat erimerkkiset.} \end{cases}$$

Aina on siis kahdelle reaaliluvulle totta, että $|x+y| \leq |x|+|y|$. Nyt

$$\begin{aligned} \|\vec{w}\|_1 &= |u_1 + v_1| + |u_2 + v_2| + \cdots + |u_n + v_n| \\ &\leq |u_1| + |v_1| + |u_2| + |v_2| + \cdots + |u_n| + |v_n| = \|\vec{u}\|_1 + \|\vec{v}\|_1 \end{aligned}$$

2-normi: Cauchy-Schwarzin epäyhtälöä käyttäen saamme

$$\begin{aligned}\|\vec{w}\|_2^2 &= \langle \vec{u} + \vec{v} | \vec{u} + \vec{v} \rangle = \|\vec{u}\|_2^2 + 2\langle \vec{u} | \vec{v} \rangle + \|\vec{v}\|_2^2 \\ &\leq \|\vec{u}\|_2^2 + 2|\langle \vec{u} | \vec{v} \rangle| + \|\vec{v}\|_2^2 \\ &\leq \|\vec{u}\|_2^2 + 2\|\vec{u}\|_2\|\vec{v}\|_2 + \|\vec{v}\|_2^2 = (\|\vec{u}\|_2 + \|\vec{v}\|_2)^2\end{aligned}$$

∞ -normi: Olkoon $\|\vec{w}\|_\infty = |w_i|$, $\|\vec{u}\|_\infty = |u_j|$ ja $\|\vec{v}\|_\infty = |v_k|$. Silloin

$$\|\vec{w}\|_\infty = |w_i| = |u_i + v_i| \leq |u_i| + |v_i| \leq |u_j| + |v_k| = \|\vec{u}\|_\infty + \|\vec{v}\|_\infty$$

□

Kolmen erilaisen normin käyttö ei ole tässä vaiheessa kovin motivoivaa. Jos laadimme numeerisia algoritmeja, niin 1-normin ja ∞ -normin laskeminen on helpompaa kuin 2-normin laskeminen. Ongelma voidaan useimmiten helposti kiertää.

Kun myöhemmin määrittelemme vektorinormiin liittyvän vastaavan matriisinormin, alkavat syntyvät kolme vaihtoehtoa erota oleellisesti toisistaan. Palaamme tähän myöhemmin.

Normia tarvitaan yleensä, kun halutaan tutkia, ovatko kaksi vektoria ”melkein samat”, eli ”lähellä toisiaan”. Jos esimerkiksi $\|\vec{a} - \vec{b}\|_\infty = 10^{-15}$, niin vektorit ovat käytännössä melkein identtiset, sillä kaikilla k on $|a_k - b_k| \leq 10^{-15}$. ”Lähellä oleminen” ei nyt riipu käytetystä normista, sillä helposti nähdään, että

$$\frac{1}{\sqrt{n}}\|\vec{u}\|_2 \leq \|\vec{u}\|_1 \leq \|\vec{u}\|_2 \quad (7.27)$$

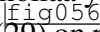
$$\frac{1}{n}\|\vec{u}\|_\infty \leq \|\vec{u}\|_1 \leq \|\vec{u}\|_\infty \quad (7.28)$$

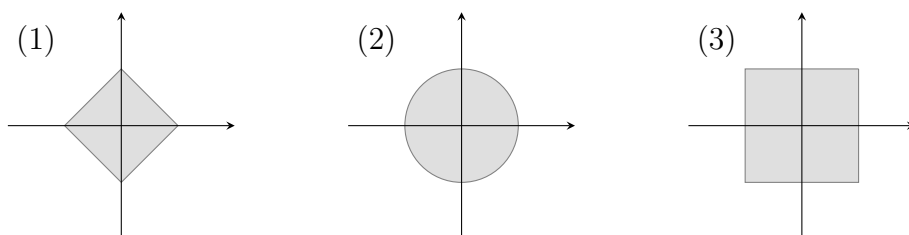
$$\|\vec{u}\|_1 \leq \|\vec{u}\|_2 \leq \sqrt{n}\|\vec{u}\|_1 \quad (7.29)$$

$$\frac{1}{\sqrt{n}}\|\vec{u}\|_\infty \leq \|\vec{u}\|_2 \leq \|\vec{u}\|_\infty \quad (7.30)$$

$$\|\vec{u}\|_1 \leq \|\vec{u}\|_\infty \leq n\|\vec{u}\|_1 \quad (7.31)$$

$$\|\vec{u}\|_2 \leq \|\vec{u}\|_\infty \leq \sqrt{n}\|\vec{u}\|_2 \quad (7.32)$$

Erityisesti, jos rajankäynnissä normi lähestyy nollaa yhden normin mielessä, niin myös muut kaksi normia lähestyvät nollaa. Kuvaan  (29) on piirretty kaikki pisteet $\vec{z} = (x, y)^T$, joille on totta, että (1) $\|\vec{z}\|_1 \leq 1$, (2) $\|\vec{z}\|_2 \leq 1$, (3) $\|\vec{z}\|_\infty \leq 1$.



Kuva 29. Pisteet, jotka ovat enintään etäisyydellä 1 origosta: (1) $\|\vec{z}\|_1 \leq 1$, (2) $\|\vec{z}\|_2 \leq 1$,
(3) $\|\vec{z}\|_\infty \leq 1$

fig056

7.3.3 Yleinen tarkastelu

Määritelmä 154 Olkoon L reaallinen lineaariavaruus. Jos on olemassa kuvaus

$$L \times L \rightarrow \mathbb{R}, (\mathbf{a}, \mathbf{b}) \mapsto \langle \mathbf{a} | \mathbf{b} \rangle,$$

joka toteuttaa seuraavat aksioomat

- (S1) $\langle \mathbf{u} | \mathbf{v} \rangle = \langle \mathbf{v} | \mathbf{u} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in L$
- (S2) $\langle \mathbf{u} | \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} | \mathbf{w} \rangle, \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in L$
- (S3) $\lambda \cdot \langle \mathbf{u} | \mathbf{v} \rangle = \langle \lambda \mathbf{u} | \mathbf{v} \rangle = \langle \mathbf{u} | \lambda \mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in L, \forall \lambda \in \mathbb{R}$
- (S4) $\langle \mathbf{u} | \mathbf{u} \rangle \geq 0$ ja $\langle \mathbf{u} | \mathbf{u} \rangle = 0$ jos ja vain jos $\mathbf{u} = \mathbf{0}$

niin kuvausta sanotaan sisätuloksi ja lukua $\langle \mathbf{u} | \mathbf{v} \rangle$ sanotaan vektoreiden $\mathbf{u}$ ja $\mathbf{v}$ sisätuloksi (engl. *inner product*).

Määritelmä 155 Olkoon L reaallinen lineaariavaruus. Jos on olemassa kuvaus

$$L \rightarrow \mathbb{R}, \mathbf{u} \mapsto \|\mathbf{u}\|,$$

joka toteuttaa seuraavat aksioomat

- (N1) $\|\mathbf{u}\| \geq 0$ ja $\|\mathbf{u}\| = 0$ jos ja vain jos $\mathbf{u} = \vec{0}$
- (N2) $\|\lambda \mathbf{u}\| = |\lambda| \cdot \|\mathbf{u}\|, \quad \forall \mathbf{u} \in L, \forall \lambda \in \mathbb{R}$
- (N3) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|, \quad \forall \mathbf{u}, \mathbf{v} \in L,$

niin kuvausta sanotaan normiksi ja lukua $\|\mathbf{u}\|$ sanotaan vektorin $\mathbf{u}$ normiksi (engl. *norm*).

Yleisessä tapauksessa sisätulo ja normi voidaan määritellä monella tavalla. (Huomaa miten tämä näkyy edellä esitettyjen määritelmien muotoilussa. Määritelmä ei sano mikä normi on, vaan se kertoo millä ehdolla kuvauksesta saa sanoa, että se on normi.)

Jokaiseen sisätuloon $\langle \mathbf{u} | \mathbf{v} \rangle$ liittyy sen avulla määritelty normi $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u} | \mathbf{u} \rangle}$. (Usein sanotaan, että ”sisätulo indusoi normin”). Toisaalta normiin ei tarvitse liittyä mitään sisätuloa.

$\mathbb{R}^n$:ssä sisätulo indusoi 2-normin, mutta 1-normiin ja ∞ -normiin ei liity sisätuloa, eikä siis myöskään kohtisuoruutta.

Esimerkki 156 Olkoon $L = C([-1, 1]) = \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ on jatkuva kuvaus}\}$. L on reaalinen lineaariavaruus, ja voimme määritellä sisätulon ja normin kaavoilla

$$\langle f | g \rangle = \int_{-1}^1 f(x)g(x) \, dx \quad (7.33)$$

$$\|f\| = \sqrt{\langle f | f \rangle} = \left(\int_{-1}^1 (f(x))^2 \, dx \right)^{1/2} \quad (7.34)$$

Lasketaan seuraavien funktioiden normit ja niiden väliset sisätulot:

$$e(x) = 1, \quad f(x) = x, \quad g(x) = x^2.$$

$$\|e\| = \left(\int_{-1}^1 (1)^2 \, dx \right)^{1/2} = \sqrt{2}$$

$$\|f\| = \left(\int_{-1}^1 x^2 \, dx \right)^{1/2} = \left(\left[\frac{1}{3} x^3 \right]_{-1}^1 \right)^{1/2} = \sqrt{2/3}$$

$$\|g\| = \left(\int_{-1}^1 (x^2)^2 \, dx \right)^{1/2} = \left(\left[\frac{1}{5} x^5 \right]_{-1}^1 \right)^{1/2} = \sqrt{2/5}$$

$$\langle e | f \rangle = \int_{-1}^1 1 \cdot x \, dx = 0$$

$$\langle e | g \rangle = \int_{-1}^1 1 \cdot x^2 \, dx = \frac{2}{3}$$

$$\langle f | g \rangle = \int_{-1}^1 x \cdot x^2 \, dx = 0$$

Siis $e \perp f$ ja $f \perp g$, mutta $e \not\perp g$. (Kohtisuoruus ei siis ole transitiivinen relaatio.)

Cauchy-Schwarz -epäyhtälön todistus (152) sivulla 137 perustuu sisätulon ja siihen liittyvän normin ominaisuuksiin. Epäyhtälö on siis tässäkin tapauksessa voimassa.

$$\begin{aligned} |\langle f | g \rangle| &\leq \|f\| \cdot \|g\| \\ \Leftrightarrow \left| \int_{-1}^1 f(x)g(x) \, dx \right| &\leq \left(\int_{-1}^1 (f(x))^2 \, dx \right)^{1/2} \left(\int_{-1}^1 (g(x))^2 \, dx \right)^{1/2} \end{aligned}$$

7.4 Lineaarinen kuvaus

Sec:LinearMap

Tämä on lyhyt kappale, jossa on vain muutama määritelmä. Jokainen määritelmä ja lause on tärkeä. Sanaa ”lineaarinen” käytetään paljon eri yhteyksissä. Matematiikan ulkopuolella sana merkitsee yleensä jotakin suoraa tai mutkatonta.

Kerran kirjoittaja joutui kansainvälisessä konferenssissa kuuntelemaan esitelmää, jossa tutkija määritteli amerikkalaisen tavan käydä liikeneuvottelua ”lineaariseksi” ja kiinalaisen tavan käydä liikeneuvottelua ”epälineaariseksi”. Esitelmässä vertailtiin näitä kahta neuvottelukulttuuria. Koska esitelmän otsikossa oli sana ”lineaarinen”, sai tutkija esittää esitelmänsä matemaattisessa sessiossa. Esitelmä oli hauska, mutta täysin väärässä paikassa.

Matematiikassa sana ”lineaarinen” liitetään aina kahden lineaariavaruuden väliseen kuvaukseen.

Määritelmä 157 Olkoon $(L, +, \circ)$ ja $(M, \hat{+}, \hat{\circ})$ reaalisia lineaariavaruuksia, ja olkoon $f : L \rightarrow M, \vec{u} \mapsto f(\vec{u})$ kuvaus. f on lineaarinen eli lineaarikuvaus, jos

$$\begin{aligned} f(\vec{a} + \vec{b}) &= f(\vec{a}) \hat{+} f(\vec{b}), \quad \forall \vec{a}, \vec{b} \in L, \text{ ja} \\ f(\lambda \circ \vec{a}) &= \lambda \hat{\circ} f(\vec{a}), \quad \forall \lambda \in \mathbb{R}, \forall \vec{a} \in L. \end{aligned}$$

Lineaarikuvaus on siis kuvaus, joka kuvaa vektorin vektorille.

kuvaus_c4s3es1

Esimerkki 158 Jos f on määritelty seuraavasti: $f : E^3 \rightarrow E^3, \vec{u} \mapsto f(\vec{u}) = 2\vec{u}$, niin esimerkiksi

$$f\left(\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}\right) = 2 \circ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 6 \end{pmatrix}.$$

Esimerkki 159 Jokainen $m \times n$ -matriisi $\mathbf{A}$ määrittelee kuvauksen $f_{\mathbf{A}} : E^n \rightarrow E^m, \vec{u} \mapsto \mathbf{A}\vec{u}$. Tämä kuvaus on lineaarinen, sillä

$$\begin{aligned} f_{\mathbf{A}}(\vec{a} + \vec{b}) &= \mathbf{A}(\vec{a} + \vec{b}) = \mathbf{A}\vec{a} + \mathbf{A}\vec{b} = f_{\mathbf{A}}(\vec{a}) + f_{\mathbf{A}}(\vec{b}), \text{ ja} \\ f_{\mathbf{A}}(\lambda \circ \vec{a}) &= \mathbf{A}(\lambda \vec{a}) = \lambda(\mathbf{A}\vec{a}) = \lambda \circ f_{\mathbf{A}}(\vec{a}). \end{aligned}$$

kuvmat_c4s3th1

Lause 160 Olkoon $f : E^n \rightarrow E^m, \vec{u} \mapsto f(\vec{u})$ lineaarinen kuvaus. Silloin on olemassa $m \times n$ -matriisi $\mathbf{A}$, jolle

$$f(\vec{v}) = \mathbf{A}\vec{v}, \quad \forall \vec{v} \in E^n.$$

Todistus. Käytetään seuraavassa konstruktiossa vanhoja tuttuja vektoreita (ks. sivu) $\vec{e}_k = (\delta_{1k}, \delta_{2k}, \dots, \delta_{nk})^T$. Jos siis esimerkiksi $n = 4$, niin

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \text{ ja } \vec{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Muodostamme matriisin $\mathbf{A}$ siten, että sen k :s sarake on $f(\vec{e}_k)$. Olkoon nyt $\vec{v} = (v_1, v_2, \dots, v_n)^T \in E^n$ mikä tahansa vektori. Suora lasku osoittaa, että

$$\begin{aligned} f(\vec{v}) &= f(v_1\vec{e}_1 + v_2\vec{e}_2 + \dots + v_n\vec{e}_n) \\ &= v_1f(\vec{e}_1) + v_2f(\vec{e}_2) + \dots + v_nf(\vec{e}_n) \\ &= v_1\vec{a}_{\bullet 1} + v_2\vec{a}_{\bullet 2} + \dots + v_n\vec{a}_{\bullet n} \\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \mathbf{A}\vec{v} \end{aligned}$$

□

Esimerkki 161 Esimerkin [linkuvaus_c4s3es1](#) (I58) lineaarikuvauksen matriisi on

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Huomaa, että lause [linkuvmat_c4s3th1](#) (I60) koskee vain Euklidisia vektoriavaruuksia E^n . Syy on tietenkin se, ettei esimerkiksi funktioavaruuksien tapauksessa funktion kertominen matriisillä ole hyvin määritelty. Sen jälkeen, kun kanta ja dimensio on määritelty, voidaan väite yleistää äärellisulotteisille lineaariavaruuksille. Mutta senkin jälkeen on paljon tärkeitä lineaariavaruuksia, joiden välisiä lineaarikuvauksia ei voi esittää matriisien avulla. Asiaan palataan tämän luvun viimeisessä kappaleessa.

Lause 162 Kahdesta lineaarikuvauksesta yhdistämällä saatu yhdistetty kuvaus on myös lineaarikuvaus.

Todistus. Olkoon $g : L \rightarrow M$, ja $f : M \rightarrow N$ lineaarikuvauksia. Silloin

$$\begin{aligned}(f \circ g)(\vec{a} + \vec{b}) &= f(g(\vec{a} + \vec{b})) = f(g(\vec{a}) + g(\vec{b})) = f(g(\vec{a})) + f(g(\vec{b})) \\ &= (f \circ g)(\vec{a}) + (f \circ g)(\vec{b}), \text{ ja} \\ (f \circ g)(\lambda \circ \vec{a}) &= f(g(\lambda \vec{a})) = f(\lambda g(\vec{a})) = \lambda f(g(\vec{a})) \\ &= \lambda (f \circ g)(\vec{a}).\end{aligned}$$

□

Lause 163 Jos $f : L \rightarrow M$ on lineaarikuvaus, niin $f(\vec{0}) = \vec{0}$.

Todistus. $f(\vec{0}) = f(0 \circ \vec{0}) = 0 \circ f(\vec{0}) = \vec{0}$.

□

7.5 (x, y) -Tason siirrot, venytykset, peilaukset ja kierrot

Sec:R2Transf

Tässä kappaleessa tutkimme yksinkertaisia kuvauksia $E^2 \rightarrow E^2$. Otamme käyttöön koulusta tutun tavan piirtää (x, y) -koordinaatisto, jonka piste kohdassa x korkeudella y vastaa vektoria $(x, y)^T$.

Käytämme seuraavassa kuvapisteele $f((x, y)^T)$ merkintää $(u, v)^T$. Piirrämme samaan kuvaan (x, y) -koordinaston päälle (u, v) -koordinaatiston. Näin voimme havainnollisesti esittää kuvapisteen ja alkukuvapisteen samassa kuvassa. (Jos haluamme tiivistää esityksemme äärimmilleen, kirjoitamme: $\vec{z} = (x, y)^T$, $\vec{w} = (u, v)^T$ ja $f(\vec{z}) = \vec{w}$.)

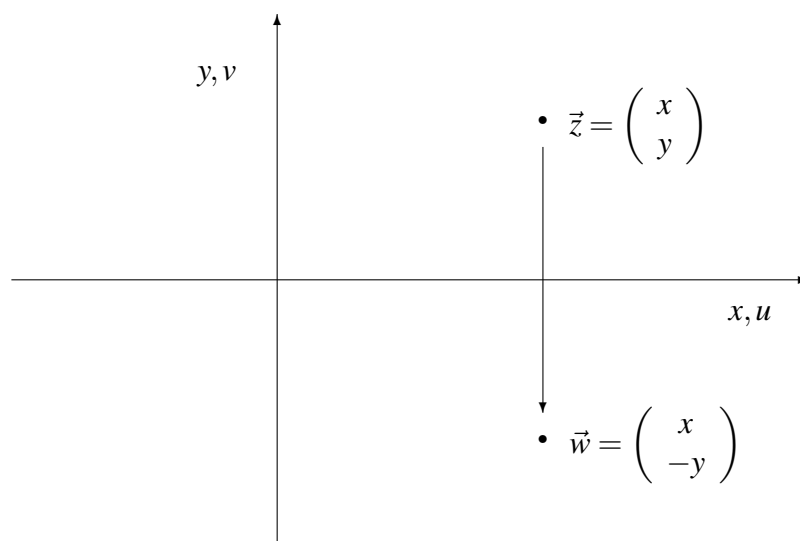
Peilaus x -akselin suhteen. Kuvaus $p_x : E^2 \rightarrow E^2$, $(x, y)^T \mapsto (x, -y)^T$ peilaa pisteen x -akselin toiselle puolelle. (Ks. kuva (??)).

Kuvaus p_x voidaan esittää myös matriisin avulla

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

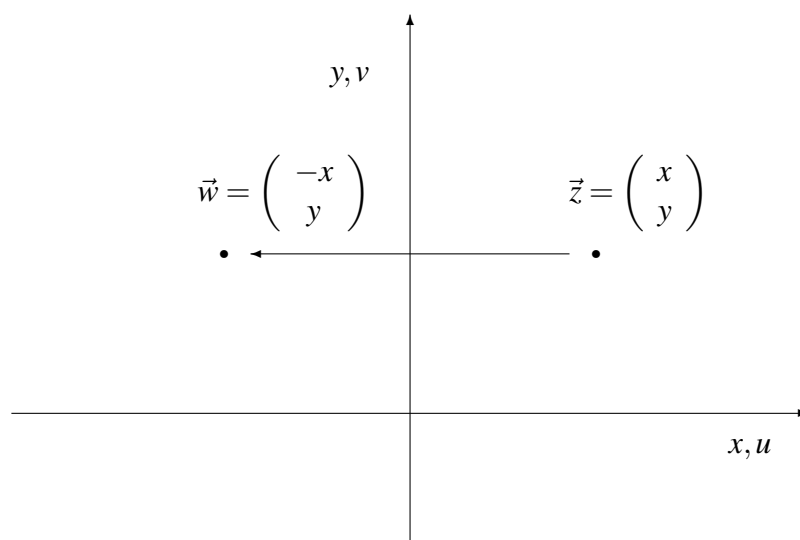
Koska kuvaus voidaan määritellä näin matriisin avulla, se on lineaarikuvaus.

c4s4fig1



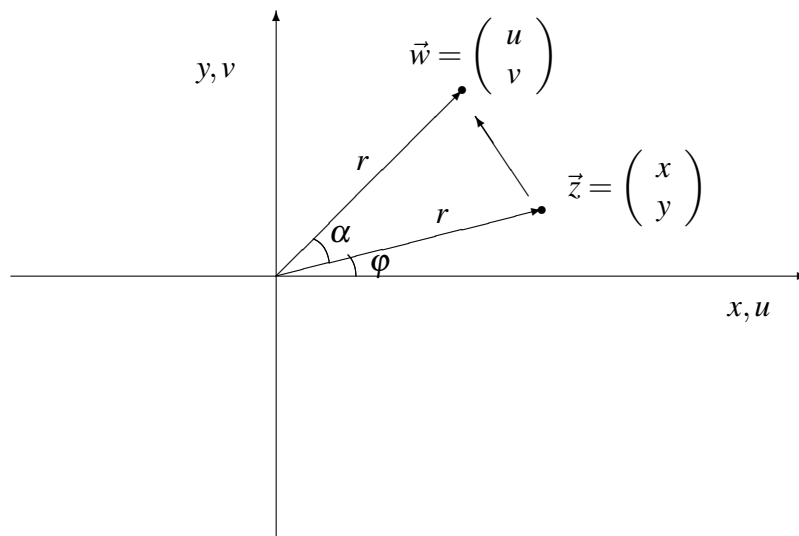
Kuva 30. Peilaus x -akselin suhteen.

c4s4fig2



Kuva 31. Peilaus y -akselin suhteen.

c4s4fig3



Kuva 32. Kierto kulman α verran origon ympäri.

Peilaus y-akselin suhteen. Kuvaus $p_x: E^2 \rightarrow E^2, (x, y)^T \mapsto (-x, y)^T$ peilaa pisteen x -akselin toiselle puolelle. (Ks. kuva (??)).

Kuvaus p_y voidaan esittää myös matriisin avulla

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

Koska kuvaus voidaan määritellä näin matriisin avulla, se on lineaarikuvaus.

Kierto kulman α verran origon ympäri. Määritellään, että piste $\vec{w} = r_\alpha(\vec{z})$ on yhtä kaukana origosta kuin $\vec{z}$, mutta sen paikkavektorin ja x -akselin välinen kulma on α :n verran suurempi kuin vastaava pisteeseen $\vec{z}$ liittyvä kulma. (Ks. kuva (??)).

Jos alkukuva on piste

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

niin trigonometristen kaavojen avulla saamme välittömästi

$$\begin{aligned} \begin{pmatrix} u \\ v \end{pmatrix} &= \begin{pmatrix} r \cos(\alpha + \varphi) \\ r \sin(\alpha + \varphi) \end{pmatrix} = \begin{pmatrix} r \cos \alpha \cos \varphi - r \sin \alpha \sin \varphi \\ r \sin \alpha \cos \varphi + r \cos \alpha \sin \varphi \end{pmatrix} \\ &= \begin{pmatrix} x \cos \alpha - y \sin \alpha \\ x \sin \alpha + y \cos \alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Kierto origon ympäri on siis lineaarinen kuvaus!

Siirto matkan $\vec{d}$ verran. Siirto $s_{\vec{d}}: E^2 \rightarrow E^2, \vec{z} \mapsto \vec{z} + \vec{d}$ ei ole lineaarikuvaus, sillä $s_{\vec{d}}(\vec{0}) \neq \vec{0}$.

Siis: kierto origon ympäri on lineaarinen kuvaus, mutta siirto ei ole lineaarinen kuvaus.

Siirto on käytännössä tärkeä kuvaus ja sitä paljon tarvitaan. On muistettava, että kun tässä kappaleessa mainittuja kuvauksia yhdistetään, niin jos yksikin osakuvaus on siirto, niin yhdistetty kuvaus ei ole lineaarinen, eikä sitä voida esittää yhden matriisin avulla.

Peilaus suoran $y = x$ suhteen. Peilaus suoran $y = x$ suhteen saadaan yhdistämällä kolme edellä kuvattua kuvausta: (1) kierto kulman $-\pi/4$ verran (2) peilaus x -akselin suhteen (3) kierto kulman $\pi/4$ verran. Kuvaukseen liittyvä matriisi on siis

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(-\pi/4) & -\sin(-\pi/4) \\ \sin(-\pi/4) & \cos(-\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(\pi/4) & \sin(\pi/4) \\ -\sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix} \begin{pmatrix} \cos(\pi/4) & \sin(\pi/4) \\ \sin(\pi/4) & -\cos(\pi/4) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

7.6 Lineaarinen aliavaruus ja span

7.7 Ydin ja kuva

Sec:MapKer

Teemme ensin yleiset tarkastelut ja sitten keskitymme matriiseihin.

Määritelmä 164 Olkoon $f: L_1 \rightarrow L_2$ lineaarikuvaus. Sanomme, että L_1 on kuvauksen lähtöavaruus (engl. domain) ja L_2 on kuvauksen maaliavaruus (engl. codomain). Lineaarikuvauksen ydin (engl. kernel) on joukko

$$\ker(f) = \{\vec{u} \in L_1 \mid f(\vec{u}) = \vec{0}\} \subseteq L_1.$$

Lineaarikuvauksen kuva (engl. range) on joukko

$$\text{ran}(f) = \{\vec{w} \in L_2 \mid f(\vec{u}) = \vec{w} \text{ jollakin } \vec{u} \in L_1\} \subseteq L_2.$$

Ydin on siis eräänlainen ”nollakohtien” joukko ja kuva on täsmälleen se mitä peruskurssilla jo on määritelty kuvauksen kuvaksi.

c4s6th2

Lause 165 Olkoon $f : L_1 \rightarrow L_2$ lineaarikuvaus. Lineaarikuvauksen ydin $\ker(f)$ on L_1 :n lineaarinen aliavaruus, ja kuva $\text{ran}(f)$ on L_2 :n lineaarinen aliavaruus.

Todistus. HT. □

Lause 166 Olkoon $f : L_1 \rightarrow L_2$ lineaarikuvaus.

(1) $\vec{0}_1 \in \ker(f)$ ja $\vec{0}_2 \in \text{ran}(f)$.

(2) Lineaarikuvaus f on injektio, jos ja vain jos $\ker(f) = \{\vec{0}_1\}$.

(3) Lineaarikuvaus on surjektio, jos ja vain jos $\text{ran}(f) = L_2$.

Todistus. (1) $f(\vec{0}_1) = f(0 \circ \vec{0}_1) = 0 \circ f(\vec{0}_1) = \vec{0}_2$.

(2) $\Rightarrow$ Olkoon f injektio. Silloin $\vec{0}_2$:n alkukuvia voi olla vain yksi. Siis $f^{-1}(\{\vec{0}_2\}) = \ker(f) = \{\vec{0}_1\}$.

$\Leftarrow$ Olkoon $\ker(f) = \{\vec{0}_1\}$ ja olkoon $\vec{u} \in L_1$ ja $\vec{v} \in L_1$ sellaiset, että $f(\vec{u}) = f(\vec{v})$. Kuvauksen f lineaarisuuden perusteella

$$f(\vec{u} - \vec{v}) = f(\vec{u}) - f(\vec{v}) = \vec{0}_2$$

Siis $\vec{u} - \vec{v} \in \ker(f)$. Oletuksen mukaan ydin on triviaali, joten

$$\begin{aligned}\vec{u} - \vec{v} &= \vec{0} \\ \Leftrightarrow \vec{u} &= \vec{v}\end{aligned}$$

(3) Seuraa suoraan surjektion ja kuvan määritelmästä. □

Samoista asioista käytetään hieman erilaisia merkintöjä, kun lineaarikuvaus on matriisin määräämä Euklidisten lineaariavaruuksien välinen kuvaus.

Määritelmä 167 Olkoon $f : E^n \rightarrow E^m, f(\vec{u}) = \mathbf{A}\vec{u}$ lineaarikuvaus. Tämän kuvauksen ydin on matriisin $\mathbf{A}$ ydin (tai nolla-avaruus) (engl. null space)

$$\text{null}(\mathbf{A}) = \{\vec{u} \in E^n \mid \mathbf{A}\vec{u} = \vec{0}\}$$

Kuvauksen kuva on matriisin $\mathbf{A}$ kuva (tai sarake-avaruus) (engl. column space)

$$\text{col}(\mathbf{A}) = \{\vec{v} \in E^m \mid \mathbf{A}\vec{u} = \vec{v}, \text{ jollakin } \vec{u} \in E^n\}$$

Matriisin sarake-avaruus on helppo luonnehtia sanomalla, että:

$$\begin{aligned}
 \text{col}(\mathbf{A}) &= \{\mathbf{A}\vec{u} \mid \vec{u} \in E^n\} \\
 &= \{u_1\vec{a}_{\bullet 1} + u_2\vec{a}_{\bullet 2} + \cdots + u_n\vec{a}_{\bullet n} \mid u_1, u_2, \dots, u_n \in \mathbb{R}\} \\
 &= \text{span}\{\vec{a}_{\bullet 1}, \vec{a}_{\bullet 2}, \dots, \vec{a}_{\bullet n}\} \\
 &= \text{matriisin } \mathbf{A} \text{ sarakkeiden virittämä } E^n\text{:n aliavaruus}
 \end{aligned}$$

Matriisin ytimen määrittäminen on hieman mutkikas tehtävä, mutta se perustuu tuttuun tekniikkaan, joten esitämme esimerkin.

c4s6es1

Esimerkki 168 Olkoon tutkittavana 4×6 -matriisi

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & 3 & 0 & 3 & 0 \\ 0 & 1 & 1 & 2 & 3 & 0 \\ 1 & 4 & 5 & -1 & 4 & 1 \\ 2 & 0 & 2 & -3 & -1 & 2 \end{pmatrix}$$

Määritämme matriisin ytimen. On siis löydettävä kaikki yhtälön $\mathbf{M}\vec{x} = \vec{0}$ ratkaisut. Sitä varten alamme ratkaista tätä yhtälöryhmää. Käytämme rivioperaatioita algortmiversion II mukaisesti, ja aina kun pivotointi epäonistuu, jätämme sarakkeen paikalleen ja jatkamme pivotointia seuraavasta sarakkeesta.

$$\begin{aligned}
 &\left(\begin{array}{cccccc|c} [1] & 2 & 3 & 0 & 3 & 0 & 0 \\ 0 & 1 & 1 & 2 & 3 & 0 & 0 \\ 1 & 4 & 5 & 2 & 7 & 1 & 0 \\ 2 & 0 & 2 & -4 & -2 & 2 & 0 \end{array} \right) \left| \begin{array}{c} -1 \\ \leftarrow \\ \leftarrow \end{array} \right| \begin{array}{c} -2 \\ \\ \leftarrow \end{array} \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 2 & 3 & 0 & 3 & 0 & 0 \\ 0 & [1] & 1 & 2 & 3 & 0 & 0 \\ 0 & 2 & 2 & 2 & 4 & 1 & 0 \\ 0 & -4 & -4 & -4 & -8 & 2 & 0 \end{array} \right) \left| \begin{array}{c} \leftarrow \\ -2 \\ \leftarrow \end{array} \right| \begin{array}{c} -2 \\ +4 \\ \leftarrow \end{array} \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 0 & 1 & -4 & -3 & 0 & 0 \\ 0 & (1) & 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & [-2] & -2 & 1 & 0 \\ 0 & 0 & 0 & 4 & 4 & 2 & 0 \end{array} \right) \left| \begin{array}{c} \leftarrow \\ \leftarrow \\ -2 \end{array} \right| \begin{array}{c} \leftarrow \\ +1 \\ +2 \end{array} \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 0 & 1 & 0 & 1 & -2 & 0 \\ 0 & (1) & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & (-2) & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & [4] & 0 \end{array} \right) \left| \begin{array}{c} \leftarrow \\ \leftarrow \\ +1/2 \end{array} \right| \begin{array}{c} \leftarrow \\ \leftarrow \\ -1/4 \end{array} \left| \begin{array}{c} \leftarrow \\ \leftarrow \\ -1/4 \end{array} \right| \begin{array}{c} \leftarrow \\ :4 \\ \end{array} : (-2) \\
 &\sim \left(\begin{array}{cccccc|c} (1) & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & (1) & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & (1) & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1) & 0 \end{array} \right)
 \end{aligned}$$

Siirretään kolmas ja viides sarake *RHS*:iin ja kirjoitetaan yhtälöt yhtälöiksi. Lisätään yhtälöihin kaksi triviaalia yhtälöä $x_3 = x_3$ ja $x_5 = x_5$.

$$\begin{cases} x_1 = -x_3 - x_5 \\ x_2 = -x_3 - x_5 \\ x_3 = x_3 \\ x_4 = -x_5 \\ x_5 = x_5 \\ x_6 = 0 \end{cases}$$

Jos valitsemme arvot $x_3 = a \in \mathbb{R}$ ja $x_5 = b \in \mathbb{R}$, niin ytimeen kuuluvalla vektorille $\vec{x}$ saadaan esitys

$$\vec{x} = a \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} = a\vec{u}_1 + b\vec{u}_2$$

Koska a ja b voidaan valita vapaasti, kuuluvat kaikki mahdolliset vektoreiden $\vec{u}_1$ ja $\vec{u}_2$ lineaarikombinaatiot matriisin $\mathbf{M}$ ytimeen. Siis

$$\text{null}(\mathbf{M}) = \text{span}\{\vec{u}_1, \vec{u}_2\}.$$

c4s6th3

Lause 169 Olkoon $\mathbf{A}$ $m \times n$ -matriisi ja olkoon $f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \vec{x} \mapsto \mathbf{A}\vec{x}$ matriisin generoitu lineaarikuvaus. Silloin

(1) jos f on injektio, niin $m \geq n$.

(2) jos f on surjektio, niin $m \leq n$.

(3) jos f on bijektio, niin $m = n$.

Todistus. (1) Jos f on injektio, niin lauseen ^(c4s6th2)(I65) mukaan $\ker(\mathbf{A}) = \{\vec{0}\}$. Silloin tulee esimerkin ^(c4s6es1)(I68) mukaisessa konstruktiossa pivotointien onnistua jokaisessa sarakkeessa, joten on mahdotonta, että rivejä olisi vähemmän kuin sarakkeita.

(2) Osoitamme, että jos $m > n$, niin f ei voi olla surjektio. Olkoon siis $m > n$. Ratkaisemme Rivioperaatioiden avulla yhtälöryhmän $\mathbf{A}\vec{x} = \vec{0}$. Koska kerroinmatriisissa rivejä on enemmän kuin sarakkeita, on viimeisen kaavion viimeinen rivi muotoa $0 = 0$. Jos ratkaisun aikana tehtyjen rivioperaatioiden jono vastaa matriiseilla $\mathbf{E}_p \cdot \mathbf{E}_2 \mathbf{E}_1$ kertomista, niin asetamme $\vec{b} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdot \mathbf{E}_p^{-1} \vec{e}_n$. Jos nyt ratkaisemme yhtälöryhmän $\mathbf{A}\vec{x} = \vec{b}$ rivioperaatioilla, niin viimeisen kaavion viimeinen rivi on tyyppiä $0 = 1$. Siis vektorilla $\vec{b} \in \mathbb{R}^n$ ei ole alkukuvaa kuvauksessa f .

(3) Seuraa kohdista (1) ja (2). □

c4s6th1

Lause 170 $n \times n$ -neliömatriisin $\mathbf{A}$ ydin on triviaali, jos ja vain jos matriisin determinantti eroaa nolasta, $\det(\mathbf{A}) \neq 0$.

Todistus. Edellisen esimerkin mukainen konstruktio ei löydä yhtään ytimeen kuuluvaa vektoria, jos ja vain jos pivotointi onnistuu joka sarakkeessa ratkaistaessa yhtälöryhmää $\mathbf{A}\vec{x} = \vec{0}$ Gaussin algoritmillä. Tämä on yhtäpitävää sen kanssa, että $\det(\mathbf{A}) \neq 0$. $\square$

8 KANNAT JA ORTOGONAALISUUS

KantaOrthogon

8.1 Lineaarinen riippumattomuus

Sec:LinIndep

Lineaarinen riippumattomuus on oikeastaan jo määritelty, mutta kirjoitamme määritelmät uudestaan tähän kappaleeseen, jotta esitys olisi helpommin luettavissa. Lisäksi otamme käyttöön uuden tavan erotella 'joukot', 'jonot' ja 'systeemit'.

Vektorijoukolle, jossa on q vektoria käytetään merkintää $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$. Alkioiden järjestyksellä ei nyt ole väliä, vaikka käytännössä ne joudutaankin aina kirjoittamaan johonkin järjestykseen. Sama alkio ei toistu siistityssä vektori-luettelossa.

Vektorijonossa $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ vektoreiden järjestys on listan mukainen ja sama vektori voi esiintyä useamman kuin yhden kerran. Matematiikassa tyypillinen tapa merkitä jono kaarisulkeita käyttäen voisi nyt aiheuttaa väärinkäsityksiä. Siksi käytämme vektorijonolle samaa merkintää kuin vektorijoukolle!

Vektorisysteemissä $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ voi sama vektori esiintyä useamman kuin yhden kerran, mutta vektoreiden järjestyksellä ei ole väliä. (Voidaan sanoa, että kaksi vektorisysteemiä ovat samat, jos niiden luetteloista saadaan permutoimalla samat.)

Esimerkki 171 (1) Olkoon $L = \{\mathbf{a}, \mathbf{b}\}$, $M = \{\mathbf{a}, \mathbf{a}, \mathbf{b}\}$ ja $N = \{\mathbf{a}, \mathbf{b}, \mathbf{a}\}$. Silloin vektorijoukkoina $L = M = N$; vektorijonoina $M \neq N$; mutta vektorisysteeminä $M = N$.

(2) a) Jos viisipäiväisen loman aikana tiskarijoukko on $\{\text{Matti, Heikki, Ville}\}$, niin nämä kolme henkilöä siis tiskaavat.

b) Jos loman aikana tiskarijono on $\{\text{Matti, Ville, Heikki, Ville, Ville}\}$, niin luettelo kertoo täsmällisesti tiskivuorojen järjestyksen.

c) Jos tiskarisysteemi on $\{\text{Matti, Heikki, Ville, Ville, Ville}\}$, niin kunkin osalta on tiedossa miten monena päivänä joutuu tiskaamaan, mutta järjestys on jätetty auki.

Seuraavassa lähes aina, kun puhumme vektorisysteemistä, joudumme käytännössä käsittelemään vektorijonoa. Kun sanomme, että vektorisysteemillä on tietty ominaisuus,

niin tarkoitamme, että kyseinen ominaisuus on vastaavan vektorijonon kaikilla permutaatioilla. (Siis ”järjestyksellä ei ole väliä”).

Määritelmät annetaan taas yleisessä muodossa, mutta tärkeimmät sovellukset koskevat $\mathbb{R}^n$:ää.

Määritelmä 172 L :n vektorisysteemi $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippuva, eli sidottu (engl. linearly dependent), jos on olemassa reaalityyppiset $\lambda_1, \lambda_2, \dots, \lambda_q \in \mathbb{R}$ siten, että

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \dots + \lambda_q \mathbf{u}_q = \mathbf{0}$$

ja ainakin yksi kertoimista eroaa nolasta ($\lambda_k \neq 0$), jollakin $k = 1, 2, \dots, q$.

Vektorisysteemi $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippumaton, eli vapaa (engl. linearly independent), jos se ei ole sidottu.

Määritelmästä seuraa välittömästi seuraavat tulokset:

Lause 173 Tarkastellaan vektorisysteemiä $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$.

- (1) Jos jokin systeemin vektoreista on nollavektori, niin systeemi on sidottu.
- (1b) Jos systeemi on vapaa, niin mikään vektoreista ei ole nollavektori.
- (2) Jos sama vektori esiintyy luettelossa kaksi kertaa, niin systeemi on sidottu.
- (2b) Jos systeemi on vapaa, niin $\mathbf{u}_i \neq \mathbf{u}_j$ aina, kun $i \neq j$.
- (3) Jos jokin tarkasteltavan systeemin osasysteemi on sidottu, niin systeemi on sidottu.
- (3b) Jos systeemi on vapaa, niin jokainen sen osasysteemi on vapaa.

Todistus.HT.

□

Määritelmä 174 Olkoon $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ vektorijono ja olkoot $\lambda_1, \lambda_2, \dots, \lambda_q \in \mathbb{R}^n$ reaalityyppiset luvut. Lauseketta

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \dots + \lambda_q \mathbf{u}_q$$

sanomme linearikombinaatioksi vektoreista $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ ja lukuja $\lambda_1, \lambda_2, \dots, \lambda_q$ sanomme linearikombinaation kertoimiksi. Jos kaikki kertoimet ovat nollia, $\lambda_1 = \lambda_2 = \dots = \lambda_q = 0$, niin sanomme, että lineaarikombinaatio on triviaali. Jos jokin kertoimista eroaa nolasta, $\lambda_k \neq 0$, niin sanomme, että lineaarikombinaatio on ei-triviaali.

Jos vektori $\mathbf{a}$ voidaan lausua summana

$$\mathbf{a} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_q \mathbf{u}_q,$$

niin sanomme, että $\mathbf{a}$ voidaan lausua lineaarikombinaationa vektoreista $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$. Jos edellä ainakin yksi kerroin eroaa nollasta, niin sanomme, että vektori $\mathbf{a}$ voidaan lausua ei-triviaalina lineaarikombinaationa vektoreista $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$.

Lause 175 Jos vektorisysteemi $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippuva, niin jokin systeemin vektoreista voidaan lausua ei-triviaalina lineaarikombinaationa muista systeemin vektoreista.

Todistus. HT

□

$\mathbb{R}^n$:ssä liitämme vektorijonoon (-systeemiin) $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ matriisin $\mathbf{V}$, jonka k :s sarakke on jonon k :s vektori. Siis

$$(\mathbf{V})_{\bullet k} = \vec{v}_k.$$

Jos $\vec{a} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_q \vec{v}_q$, niin

$$\mathbf{V}\vec{\lambda} = \vec{a}, \text{ missä } \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_q \end{pmatrix}.$$

Näillä merkinnöillä vektorisysteemin lineaarinen riippumattomuus voidaan luonnehtia seuraavasti:

$$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\} \text{ on vapaa} \quad (8.1)$$

$$\Leftrightarrow \mathbf{V}\vec{\lambda} = \vec{0}, \text{ jos ja vain jos } \vec{\lambda} = \vec{0} \quad (8.2)$$

$$\Leftrightarrow \ker(\mathbf{V}) = \{\vec{0}\} \quad (8.3)$$

Seuraavassa lauseessa tätä ideaa kehitetään edelleen.

c4s7th1

Lause 176 Tarkastellaan $\mathbb{R}^n$:n vektorisysteemiä $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ ja siihen liittyvää matriisia $\mathbf{V}$. Silloin seuraavat ehdot ovat yhtäpitävät

$$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\} \text{ on vapaa} \quad (8.4)$$

$$\Leftrightarrow \ker(\mathbf{V}) = \{\vec{0}\} \quad (8.5)$$

$$\Leftrightarrow \ker(\mathbf{V}^T \mathbf{V}) = \{\vec{0}\} \quad (8.6)$$

$$\Leftrightarrow \det(\mathbf{V}^T \mathbf{V}) \neq 0 \quad (8.7)$$

$$\Leftrightarrow \mathbf{V}^T \mathbf{V} \text{ on säännöllinen} \quad (8.8)$$

Todistus. $(8.4) \Leftrightarrow (8.5)$ seuraa suoraan määritelmistä.

$(8.5) \Leftrightarrow (8.6)$ Jos $\vec{u} \in \ker(V)$, niin $V^T V \vec{u} = V^T \vec{0} = \vec{0}$. Siis $\ker(V) \subseteq \ker(V^T V)$.
Toisaalta olkoon $\vec{u} \in \ker(V^T V)$. Silloin

$$0 = \langle V^T V \vec{u} | \vec{u} \rangle = \langle V \vec{u} | V \vec{u} \rangle \Rightarrow V \vec{u} = \vec{0}$$

Siis $\ker(V^T V) \subseteq \ker(V)$.

$(8.6) \Leftrightarrow (8.7)$ Seuraa siitä, että $V^T V$ on neliömatriisi ja lauseesta (I70) sivulla 150.

$(8.7) \Leftrightarrow (8.8)$ Seuraa lauseesta (I04) sivulla 102. □

Esimerkki 177 Tutkitaan onko vektorijono $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ lineaarisesti riippumaton, kun

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -3 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} 4 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Tutkitaan asia (1) ”kynällä ja paperilla” ja (2) ”tietokonetta käyttäen”.

(1) Etsimme ei-triviaaleja λ -kertoimia, joilla yhtälö $\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0}$ on tosi. Jos löytyy ei-triviaalit kertoimet, jono on sidottu, ja jos ei löydy, niin jono on vapaa. Ryhdymme työhön:

$$\begin{aligned} \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0} &\Leftrightarrow \begin{cases} \lambda_1 + 2\lambda_2 + 4\lambda_3 = 0 \\ \lambda_2 - \lambda_3 = 0 \\ 2\lambda_1 - \lambda_2 = 0 \\ -3\lambda_1 + \lambda_3 = 0 \end{cases} \\ &\sim \left(\begin{array}{ccc|c} [1] & 2 & 4 & 0 \\ 0 & 1 & -1 & 0 \\ 2 & -1 & 0 & 0 \\ -3 & 0 & 1 & 0 \end{array} \right) \begin{array}{l} -2 \\ \leftarrow \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & [1] & -1 & 0 \\ 0 & -5 & -8 & 0 \\ 0 & 6 & 12 & 0 \end{array} \right) \begin{array}{l} +5 \\ \leftarrow \\ \leftarrow \end{array} \\ &\sim \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & (1) & -1 & 0 \\ 0 & 0 & [-13] & 0 \\ 0 & 0 & 18 & 0 \end{array} \right) \begin{array}{l} +18/13 \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|c} (1) & 2 & 4 & 0 \\ 0 & (1) & -1 & 0 \\ 0 & 0 & (-13) & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \\ &\Leftrightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0 \end{aligned}$$

Siis vektorisysteemi on vapaa, eli lineaarisesti riippumaton.

(2) Matlabilla saman asian voi todeta lauseen (I76) perusteella:

```
>> U = [1 0 2 -3 ; 2 1 -1 0 ; 4 -1 0 1]'
U =
     1     2     4
     0     1    -1
     2    -1     0
    -3     0     1

>> det (U' * U)
ans =
     820
>>
```

Koska siis $\det(\mathbf{U}^T \mathbf{U}) \neq 0$, niin vektorisysteemi on vapaa.

Esimerkki 178 Otetaan vielä esimerkki, jossa johtopäätös on erilainen. Tutkitaan onko vektorijono $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ lineaarisesti riippumaton, kun

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -3 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} -1 \\ -2 \\ 8 \\ -9 \end{pmatrix}.$$

(1) Etsimme taas ei-triviaaleja λ -kertoimia, joilla yhtälö $\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0}$ on tosi. Jos löytyy ei-triviaalit kertoimet, jono on sidottu, ja jos ei löydy, niin jono on vapaa. Ryhdymme työhön:

$$\begin{aligned} \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0} &\Leftrightarrow \begin{cases} \lambda_1 + 2\lambda_2 - \lambda_3 = 0 \\ \lambda_2 - 2\lambda_3 = 0 \\ 2\lambda_1 - \lambda_2 + 8\lambda_3 = 0 \\ -3\lambda_1 - 9\lambda_3 = 0 \end{cases} \\ &\sim \left(\begin{array}{ccc|c} [1] & 2 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 2 & -1 & 8 & 0 \\ -3 & 0 & -9 & 0 \end{array} \right) \begin{array}{l} -2 \\ \leftarrow \\ \leftarrow \end{array} \sim \left(\begin{array}{ccc|c} (1) & 2 & 1 & 0 \\ 0 & [1] & -2 & 0 \\ 0 & -5 & 10 & 0 \\ 0 & 6 & -12 & 0 \end{array} \right) \begin{array}{l} +3 \\ \leftarrow \\ \leftarrow \end{array} \\ &\sim \left(\begin{array}{ccc|c} (1) & 2 & 1 & 0 \\ 0 & (1) & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \Leftrightarrow \begin{cases} \lambda_1 = 3a \\ \lambda_2 = 2a \\ \lambda_3 = a \end{cases}, \quad a \in \mathbb{R} \end{aligned}$$

Koska löytyy ei-triviaalit kertoimet, vektorisysteemi on sidottu, eli lineaarisesti riippuva.

(2) Matlabilla saman asian voi todeta lauseen ^{c4s7th1}(I76) perusteella:

```
>> U = [1 0 2 -3 ; 2 1 -1 0 ; -1 -2 8 -9]'
```

```
U =
```

```
    1    2   -1
    0    1   -2
    2   -1    8
   -3    0   -9
```

```
>> det (U' * U)
```

```
ans =
```

```
    0
```

```
>>
```

Koska siis $\det(\mathbf{U}^T \mathbf{U}) = 0$, niin vektorisysteemi on sidottu.

Lause 179 Olkoon $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ lineaarisesti riippumaton $\mathbb{R}^n$:n vektorisysteemi. Jos $\vec{a} \in \mathbb{R}^n$, $\vec{a} \neq \vec{0}$ ja $\vec{a} \perp \vec{v}_k$, $\forall k = 1, \dots, q$, niin vektorisysteemi

$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{a}\}$ on vapaa.

Todistus. Olkoon $\mathbf{V}$ se $n \times q$ -matriisi, jonka sarakkeina ovat vektorit $\vec{v}_k, k = 1, \dots, q$ ja olkoon $\mathbf{W} = (\mathbf{V} | \vec{a})$ vastaava $n \times (q+1)$ -matriisi, jonka viimeinen sarake on vektori $\vec{a}$. Oletuksesta $\vec{a} \neq \vec{0}$ seuraa, että $\|\vec{a}\| > 0$ ja oletuksesta $\vec{a} \perp \vec{v}_k, \forall \vec{v}_k$ seuraa, että $\mathbf{V}^T \vec{a} = \vec{0}$. Nyt voimme laskea suoraan

$$\begin{aligned} \mathbf{W}^T \mathbf{W} &= (\mathbf{V} | \vec{a})^T (\mathbf{V} | \vec{a}) = \begin{pmatrix} \mathbf{V}^T \\ \vec{a}^T \end{pmatrix} (\mathbf{V} | \vec{a}) \\ &= \left(\begin{array}{c|c} \mathbf{V}^T \mathbf{V} & \mathbf{V}^T \vec{a} \\ \hline \vec{a}^T \mathbf{V} & \vec{a}^T \vec{a} \end{array} \right) = \left(\begin{array}{c|c} \mathbf{V}^T \mathbf{V} & \vec{0} \\ \hline \vec{0}^T & \|\vec{a}\|^2 \end{array} \right) \end{aligned}$$

$$\Rightarrow \det(\mathbf{W}^T \mathbf{W}) = \|\vec{a}\|^2 \det(\mathbf{V}^T \mathbf{V}) \neq 0$$

Siis lauseen ^{|c4s7th1}(I76) mukaan vektorijono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{a}\}$ on vapaa. □

Yhdistämällä lause ^{|c4s7th1}(I76) lauseeseen ^{|c4s3th6}(I49) saadaan seuraava tulos, jota myöhemmin pystymme vielä parantamaan:

Lause 180 Jos $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_q\}$ on lineaarisesti riippumaton $\mathbb{R}^n$:n vektorisysteemi ja $M = \text{span}(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_q)$, niin

$$\mathbb{R}^n = M \oplus M^\perp$$

Seuraa suoraan lauseesta ^{|c4s7th1}(I76) sivulla ^{|c4s7th1}I53 ja lauseesta ^{|c4s3th6}(I49) sivulla ^{|c4s3th6}I35. □

8.2 Kanta ja dimensio

Sec:BaseDim

c4s8def1

Määritelmä 181 Lineaariavaruuden L vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaariavaruuden L kanta, jos

1. Jokainen L :n vektori $\mathbf{a} \in L$ voidaan lausua lineaarikombinaationa jonon F vektoreista

$$\mathbf{a} = \mu_1 \mathbf{u}_1 + \mu_2 \mathbf{u}_2 + \dots + \mu_q \mathbf{u}_q.$$

2. Jono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on vapaa.

Jos $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on kanta ja $\mathbf{a} = \mu_1 \mathbf{u}_1 + \mu_2 \mathbf{u}_2 + \dots + \mu_q \mathbf{u}_q$, niin kertoimet $\mu_1, \mu_2, \dots, \mu_q$ ovat vektorin $\mathbf{a}$ koordinaatit kannassa F .

Koordinaateista muodostettua vektoria $\vec{\mu} = (\mu_1, \mu_2, \dots, \mu_q)^T \in \mathbb{R}^q$ sanomme vektorin $\mathbf{a}$ koordinaattivektoriksi (kannassa F).

Esimerkki 182 (1) Vektorijono $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ eli vektorit

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

on $\mathbb{R}^n$:n kanoninen kanta. Kantaa sanotaan kanoniseksi, koska sillä on monta yksinkertaista ominaisuutta: jokaisen kantavektorin normi on 1, kantavektorit ovat keskenään kohtisuorat.

(2) Jos lineaariavaruuden L vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippumaton, niin se on virittämänsä aliavaruuden $M = \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q)$ kanta.

Jos vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaariavaruuden L kanta ja vektori $\mathbf{a} \in L$ on mikä tahansa L :n vektori, niin vektorin $\mathbf{a}$ koordinaatit kannassa F ovat yksikäsitteiset. Siis on olemassa yhdet ja vain yhdet koordinaatit. Jos nimittäin olisi olemassa kahdet erilaiset koordinaatit

$$\begin{aligned} \mathbf{a} &= a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_q \mathbf{u}_q, \\ \mathbf{a} &= \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_q \mathbf{u}_q, \quad \text{ja } a_k \neq \alpha_k \text{ jollakin } k, \end{aligned}$$

niin silloin olisi

$$(a_1 - \alpha_1) \mathbf{u}_1 + (a_2 - \alpha_2) \mathbf{u}_2 + \dots + (a_q - \alpha_q) \mathbf{u}_q = \mathbf{0}, \quad (a_k - \alpha_k) \neq 0$$

mikä on mahdotonta, koska kanta on vapaa.

Lause 183 Vektorin koordinaatit kannassa ovat yksikäsitteiset.

Todistus. Perustelu edellä. □

Lause 184 Olkoon L lineaariavaruus ja olkoon vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ sen kanta. Silloin identtinen kuvaus

$$\text{id} : L \rightarrow L, \mathbf{a} \mapsto \mathbf{a}$$

ja kuvaus, joka kuvaa vektorin sen koordinaattivektorille kannassa F

$$v_F : L \rightarrow \mathbb{R}^q, \mu_1 \mathbf{u}_1 + \mu_2 \mathbf{u}_2 + \dots + \mu_q \mathbf{u}_q \mapsto (\mu_1, \mu_2, \dots, \mu_q)^T$$

ovat bijektiivisiä lineaarikuvauksia.

Todistus. Olkoon $\mathbf{a} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_q \mathbf{u}_q$, $\mathbf{b} = \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \dots + \beta_q \mathbf{u}_q$ ja $\lambda \in \mathbb{R}$. Silloin

$$\begin{aligned} \text{id}(\mathbf{a} + \mathbf{b}) &= \mathbf{a} + \mathbf{b} = \text{id}(\mathbf{a}) + \text{id}(\mathbf{b}) \\ \text{id}(\lambda \mathbf{a}) &= \lambda \mathbf{a} = \lambda \text{id}(\mathbf{a}) \\ v_F(\mathbf{a} + \mathbf{b}) &= \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_q + \beta_q \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_q \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_q \end{pmatrix} = v_F(\mathbf{a}) + v_F(\mathbf{b}) \\ v_F(\lambda \mathbf{a}) &= \begin{pmatrix} \lambda \alpha_1 \\ \lambda \alpha_2 \\ \vdots \\ \lambda \alpha_q \end{pmatrix} = \lambda v_F(\mathbf{a}) \end{aligned}$$

Siis kumpikin kuvaus on lineaarinen. Kumpikin kuvaus on selvästi bijektio. □

Lause 185 Olkoon vektorijono $F = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ lineaariavaruuden L kanta. Jos $G = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ on toinen lineaariavaruuden L kanta, niin

$$k = q.$$

Toisin sanoen: Kaikissa mahdollisissa L :n kannoissa on sama määrä kantavektoreita.

Todistus. Seuraavassa kaaviossa kaikki kuvaukset ovat bijektiivisiä lineaarikuvauksia.

$$\begin{array}{ccc} & \text{id} & \\ L & \longrightarrow & L \\ v_F \downarrow & & \downarrow v_G \\ \mathbb{R}^q & & \mathbb{R}^k \end{array}$$

Yhdistämällä nämä kuvaukset saamme bijektiivisen lineaarikuvauksen

$$v_G \circ \text{id} \circ v_F^{-1} : \mathbb{R}^q \rightarrow \mathbb{R}^k$$

Tämä on lauseen (I69) ^{c4s6th3} mukaan mahdollista vain, jos $k = q$. □

Koska lineaariavaruuden kannan kantavektoreiden lukumäärä on sama kaikilla kannoilla, voimme määritellä seuraavasti.

Määritelmä 186 Jos lineaariavaruudella L on kanta, jossa on k kantavektoria, niin sanomme, että lineaariavaruuden L dimensio, $\dim(L)$, on k

Esimerkki 187 (1)

$$\dim(\mathbb{R}^n) = n$$

(2) Jos $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q\}$ on lineaarisesti riippumaton vektorijono, niin

$$\dim(\text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_q)) = q$$

Siis vapaan vektorijonon virittämän aliavaruuden dimensio on sama kuin virittäjiensä lukumäärä.

Kaikki tason vektorit voidaan lausua kahden $\{\vec{i}, \vec{j}\}$ vektorin lineaarikombinaatioina, joten tason dimensio on 2. Toisaalta kolme tasovektoria muodostaa aina sidotun systeemin. Yleisemmin: jos vektorijonossa on liikaa vektoreita, se ei voi olla vapaa.

c4s8th5

Lause 188 Olkoon L lineaariavaruus, ja $\dim(L) = n$. Tarkastellaan L :n vektorijonoa $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$. Jos $q > n = \dim(L)$, niin vektorijono F on sidottu.

Todistus. (1) Todistetaan ensin se erikoistapaus, jossa $L = \mathbb{R}^n$ ja $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$. Olkoon $\mathbf{V}$ se $n \times q$ -matriisi, jonka sarakkeina ovat jonon F vektorit. Lähdemme nyt etsimään ei-triviaaleja kertoimia λ_i niin, että

$$\begin{aligned} \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_q \vec{v}_q &= \vec{0} \\ \Leftrightarrow \mathbf{V} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_q \end{pmatrix} &= \vec{0} \end{aligned}$$

Tämä on homogeeninen yhtälöryhmä, jonka kerroinmatriisissa $\mathbf{V}$ on n riviä ja q saraketta. Koska sarakkeita on enemmän kuin rivejä, jää Gaussin algoritmin mukaisessa ratkaisussa aina osa sarakkeista pivotoimatta, joten yhtälöryhmälle löytyy ei-triviaaleja ratkaisuja. Siis jono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ on sidottu.

(2) Nyt palaamme alkuperäiseen väitteeseen. Koska L :n dimensio on n , on olemassa kanta $G = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$, jossa on n kantavektoria. Olkoon nyt $\vec{v}_k = v_G(\mathbf{v}_k)$ vektorin $\mathbf{v}_k$ koordinaattivektori kannassa G . Todistuksen alkuosan mukaan nyt on olemassa ei-triviaalit kertoimet λ_i siten, että

$$\mathbf{V}\vec{\lambda} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \dots + \lambda_q\vec{v}_q = \vec{0}.$$

Silloin

$$\begin{aligned} \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \dots + \lambda_q\mathbf{v}_q &= \sum_{i=1}^q \lambda_i \left(\sum_{k=1}^n v_{ki}\mathbf{u}_k \right) \\ &= \lambda_1(v_{11}\mathbf{u}_1 + v_{21}\mathbf{u}_2 + \dots + v_{n1}\mathbf{u}_n) \\ &\quad + \lambda_2(v_{12}\mathbf{u}_1 + v_{22}\mathbf{u}_2 + \dots + v_{n2}\mathbf{u}_n) \\ &\quad + \dots \\ &\quad + \lambda_q(v_{1q}\mathbf{u}_1 + v_{2q}\mathbf{u}_2 + \dots + v_{nq}\mathbf{u}_n) \\ &= \mathbf{u}_1(v_{11}\lambda_1 + v_{12}\lambda_2 + \dots + v_{1q}\lambda_q) \\ &\quad + \mathbf{u}_2(v_{21}\lambda_1 + v_{22}\lambda_2 + \dots + v_{2q}\lambda_q) \\ &\quad + \dots \\ &\quad + \mathbf{u}_n(v_{n1}\lambda_1 + v_{n2}\lambda_2 + \dots + v_{nq}\lambda_q) \\ &= \sum_{k=1}^n \mathbf{u}_k(\mathbf{V}\vec{\lambda})_k = \vec{0} \end{aligned}$$

□

c4s8th6

Lause 189 *Olkoon L lineaariavaruus, ja $\dim(L) = n$. Jos L :n vektorijono $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ on vapaa, niin se on L :n kanta.*

Todistus. Kannan määritelmän mukaan nyt riittää osoittaa, että vektorijono virittää L :n. Olkoon sitä varten $\mathbf{a}$ mikä tahansa L :n vektori. Silloin jonossa $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{a}\}$ on ”liian monta vektoria ollakseen vapaa”. Edellisen lauseen mukaan $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{a}\}$ on sidottu. Siis joku vektoreista voidaan lausua muiden lineaarikombinaationa. Siinä yhtälössä $\mathbf{a}$ on mukana ja voidaan ratkaista F :n vektoreiden lineaarikombinaationa. □

c4s8th7

Lause 190 Olkoon L lineaariavaruus, ja $\dim(L) = n$. Tarkastellaan L :n vektorijonoa $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$. Jos $q < n$, ja vektorijono F on vapaa, niin on olemassa $n - q$ vektoria $\mathbf{a}_j \in L, j = 1, \dots, n - q$ siten, että jono

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q, \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-q}\}$$

on L :n kanta. (Sanomme, että vapaa jono $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ on täydennetty kannaksi vektoreilla $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n-q}\}$.)

Todistus. Koska $\dim(L) = n$, on lineaariavaruudella kanta $G = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$. Jokin kannan G kantavektoreista $\mathbf{u}_k$ ei ole lausuttavissa jonon $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ vektoreiden lineaarikombinaationa, muuten jono $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ olisi L :n kanta, mikä on mahdottomaa, koska siinä ei ole tarpeeksi vektoreita. Lisätään kantavektori $\mathbf{u}_k$ vektorijonoon viimeiseksi. Uusi täydennetty vektorijono $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q, \mathbf{u}_k\}$ on vapaa ja siinä on $q + 1$ vektoria.

Toistamalla äskeistä menettelyä saamme lopulta täydennetyn vektorijonon, joka on vapaa ja jossa on n vektoria. Se on kanta. $\square$

c4s8th7b

Lause 191 Olkoon $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ vapaa jono lineaariavaruuden $\mathbb{R}^n$ vektoreita. Jos $\vec{b} \in L$ on kohtisuorassa kaikkia jonon vektoreita vastaan, niin täydennetty jono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{b}\}$ on vapaa.

Todistus. Kysymys on nyt siitä voidaanko $\vec{b}$ lausua ei-triviaalina lineaarikombinaationa jonon F vektoreista. Olkoon nyt

$$\vec{b} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_q \vec{v}_q = \mathbf{V} \vec{\lambda}$$

Koska $\vec{b}$ on kohtisuorassa jonon vektoreita vastaan, on

$$\vec{0} = \mathbf{V}^T \vec{b} = \mathbf{V}^T \mathbf{V} \vec{\lambda}$$

Lauseen (I 76) mukaan tämä on mahdollista vain, kun $\lambda_1 = \lambda_2 = \dots = \lambda_q = 0$. Siis täydennetty jono on vapaa. $\square$

c4s7th1

c4s8th8

Lause 192 Tarkastellaan n -ulotteisen lineaariavaruuden $\mathbb{R}^n$ vapaata vektorijonoa $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ ja sen virittämää aliavaruutta $M = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$. Jos $q < n$, niin on olemassa $n - q$ vektoria $\vec{b}_j \in M^\perp, j = 1, \dots, n - q$ siten, että jono

$$G = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{b}_1, \vec{b}_2, \dots, \vec{b}_{n-q}\}$$

on L :n kanta ja jono

$$H = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_{n-q}\}$$

on aliavaruuden ortogonaalikomplementin $M^\perp$ kanta.

Erityisesti

$$\dim(M) + \dim(M^\perp) = n$$

Todistus. Lauseen (I90) todistuksen konstruktio tapahtui vaiheittain siten, että joka askeleessa jonoon lisätään yksi vektori $\vec{a}_k$. Koska $\mathbb{R}^n = M \oplus M^\perp$ on suora summa, joten lauseen (I34) mukaan vektori $\vec{a}_k$ voidaan yksikäsitteisellä tavalla jakaa kahden vektorin summaksi $\vec{a}_k = \vec{a}_k^M + \vec{a}_k^\perp$ siten, että $\vec{a}_k^M \in M$ ja $\vec{a}_k^\perp \in M^\perp$. Asetetaan $\vec{b}_k = \vec{a}_k^\perp$.

Lauseen (I91) mukaan täydennetty jono on vapaa, ja toistamalla menettelyä saamme täydennetyn jonon

$$G = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q, \vec{a}_1^\perp, \vec{a}_2^\perp, \dots, \vec{a}_{n-q}^\perp\},$$

joka on vapaa ja siis $\mathbb{R}^n$:n kanta.

Jono $H = \{\vec{a}_1^\perp, \vec{a}_2^\perp, \dots, \vec{a}_{n-q}^\perp\}$ on vapaan jonon osajonona myös vapaa. Jäljellä on enää sen osoittaminen, että H virittää $M^\perp$:n.

Olkoon $\vec{u} \in M^\perp$ mikä tahansa ortogonaalikomplementin vektori. Myös vektori $\vec{u}$ voidaan jakaa kahden komponentin summaksi

$$\vec{u} = \vec{u}^M + \vec{u}^\perp = \vec{0} + \vec{u}$$

Jos siis vektori $\vec{u}$ kirjoitetaan kannassa G , saadaan

$$\vec{u} = \underbrace{\mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \dots + \mu_q \vec{v}_q}_{=\vec{u}^M=\vec{0}} + \underbrace{\mu_{q+1} \vec{a}_1^\perp + \mu_{q+2} \vec{a}_2^\perp + \dots + \mu_n \vec{a}_{n-q}^\perp}_{=\vec{u}^\perp=\vec{u}}$$

Koska jono $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ on vapaa, on ensimmäisten koordinaattien oltava nollia, $\mu_1 = \mu_2 = \dots = \mu_q = 0$. Siis

$$\vec{u} \in \text{span}(\vec{a}_1^\perp, \vec{a}_2^\perp, \dots, \vec{a}_{n-q}^\perp)$$

□

Kootaan joitakin aikasemmin todettuja tuloksia yhteen seuraavasti:

Lause 193 Tarkastellaan $\mathbb{R}^n$:n vektorisysteemiä $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ ja siihen liittyvää matriisia $\mathbf{V}$.

(1) Jos $q < n$, niin systeemi on vapaa, jos ja vain jos $\det(\mathbf{V}^T \mathbf{V}) \neq 0$.

(2) Jos $q = n$, niin systeemi on vapaa, jos ja vain jos $\det(\mathbf{V}) \neq 0$.

(3) Jos $q > n$, niin systeemi on sidottu.

Todistus. (1) Todistettiin lauseessa (I76 sivu I53). (2) Kun $q = n$, matriisi $\mathbf{V}$ on neliö-matriisi, ja $\det(\mathbf{V}^T \mathbf{V}) = (\det(\mathbf{V}))^2$. (3) Todettiin lauseessa (I88). $\square$

Lopuksi tarkastelemme tärkeitä kysymystä siitä, miten vektorijonon virittämälle aliavaruudelle saadaan kanta, jos virittäjäjono ei ole vapaa.

Olkoon lähtökohtanamme $\mathbb{R}^n$:n vektorijono

$$F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$$

ja uusi vektorijono, jossa alkutilanteessa on vain yksi vektori

$$G = \{\vec{v}_1\}.$$

Käytään alkuperäisen vektorijonon F vektorit läpi alkaen toisesta ja päättyen viimeiseen siten, että jokaisen vektorin $\vec{v}_k$ kohdalla tehdään seuraava konstruktio:

Jos tutkittavaa vektoria $\vec{v}_k$ ei voi lausua sen edellä alkuperäisessä vektorijonossa olevien vektorien lineaarikombinaationa ($\vec{v}_k \notin \text{span}(\vec{v}_1, \dots, \vec{v}_{k-1})$), niin lisätään vektori $\vec{v}_k$ uuteen jonoon G .

Kun kaikki alkuperäisen jonon vektorit on näin käyty läpi, on uudessa jonossa G alkuperäisen jonon osajono. Tästä uudesta vektorijonosta sanomme, että se on saatu karsimalla alkuperäinen jono vapaaksi.

Helpoin tapa tehdä karsiminen, on muodostaa vektorijonoa $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ vastaava matriisi $\mathbf{V}$, ja tehdä matriisille Gaussin algoritmi mukaiset rivioperaatiot. Aina, kun pivotointi jossakin sarakkeessa epäonnistuu, se johtuu siitä, että kyseinen sarake on lausuttavissa edellisten sarakkeiden lineaarikombinaationa. Teemme siis siten, että Vektorijonosta poistetaan kaikki vektorit, joita vastaavassa $\mathbf{V}$:n sarakkeessa pivotointi epäonnistui tai jäi tekemättä.

Otamme ensin esimerkin, jotta menettely tulee varmasti ymmärrettyä oikein.

Esimerkki 194 Karsimme vapaaksi vektorijonon

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 0 \\ 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -4 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ -1 \end{pmatrix} \right\}$$

Muodostamme asiaankuuluvan matriisin ja ryhdymme pivotoimaan

$$\begin{aligned} & \left(\begin{array}{ccccc|c} [1] & 2 & 1 & 1 & 0 & -2 \\ 2 & 5 & 4 & 1 & 0 & \leftarrow \\ 0 & 0 & 0 & 0 & 1 & \\ -1 & 1 & 5 & -4 & 2 & \\ 0 & 2 & 4 & -2 & -1 & \end{array} \right) \begin{array}{l} \\ \\ \leftarrow \\ \\ \end{array} \begin{array}{l} +1 \\ \\ \\ \\ \end{array} \sim \left(\begin{array}{ccccc|c} (1) & 2 & 1 & 1 & 0 & \\ 0 & [1] & 2 & -1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 & \\ 0 & 3 & 6 & -3 & 2 & \leftarrow \\ 0 & 2 & 4 & -2 & -1 & \leftarrow \end{array} \right) \begin{array}{l} \\ \\ \\ \leftarrow \\ \leftarrow \end{array} \\ & \sim \left(\begin{array}{ccccc|c} (1) & 2 & 1 & 1 & 0 & \\ 0 & (1) & 2 & -1 & 0 & \\ 0 & 0 & 0 & 0 & [1] & -2 \\ 0 & 0 & 0 & 0 & 2 & \leftarrow \\ 0 & 0 & 0 & 0 & -1 & \leftarrow \end{array} \right) \begin{array}{l} \\ \\ \\ \leftarrow \\ \leftarrow \end{array} \sim \left(\begin{array}{ccccc|c} (1) & 2 & 1 & 1 & 0 & \\ 0 & (1) & 2 & -1 & 0 & \\ 0 & 0 & 0 & 0 & (1) & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \end{array} \right) \begin{array}{l} \\ \\ \\ \\ \leftarrow \end{array} \end{aligned}$$

Siis kolmas ja neljäs vektori karsitaan pois; jäljelle jäävät ensimmäinen, toinen ja viides vektori. Vapaaksi karsittu jono on

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ -1 \end{pmatrix} \right\}.$$

Lause 195 Olkoon

$$F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$$

$\mathbb{R}^n$:n vektorijono ja olkoon

$$G = \{\vec{v}_{k_1}, \vec{v}_{k_2}, \dots, \vec{v}_{k_r}\}$$

jonosta F vapaaksi karsittu jono. Silloin G on vapaa ja virittää saman aliavaruuden kuin F . Toisin sanoen G on aliavaruuden $M = \text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q)$ kanta.

Todistus. Jos G olisi sidottu, niin jokin G :n vektoreista voitaisiin lausua sitä edeltävien vektoreiden lineaarikombinaationa. Tämä on mahdotonta, koska tällaiset vektorit on

karsittu pois. Siis G on vapaa.

Olkoon toiseksi

$$\vec{a} = \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \cdots + \mu_q \vec{v}_q$$

mikä tahansa aliavaruuden M vektori. Käymme termit läpi alkaen viimeisestä. Jos $\vec{v}_k$ karsittiin jonosta pois, niin se voidaan korvata lineaarikombinaationa edeltävistä vektoreista. Sijoitetaan tämä lineaarikombinaatio vektorin $\vec{v}_k$ paikalle ja sievennetään lauseke. Jatketaan seuraavasta termistä. Näin jatkaen löytyy lineaarikombinaatio, jossa on vain karsitun jonon vektoreita. $\square$

Esimerkki 196 Edellisessä esimerkissä

$$\vec{v}_3 = -3\vec{v}_1 + 2\vec{v}_2, \text{ ja}$$

$$\vec{v}_4 = \vec{v}_2 - \vec{v}_3$$

Siis

$$\begin{aligned} & \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \mu_3 \vec{v}_3 + \mu_4 \vec{v}_4 + \mu_5 \vec{v}_5 \\ = & \mu_1 \vec{v}_1 + \mu_2 \vec{v}_2 + \mu_3 \vec{v}_3 + \mu_4 (\vec{v}_2 - \vec{v}_3) + \mu_5 \vec{v}_5 \\ = & \mu_1 \vec{v}_1 + (\mu_2 + \mu_4) \vec{v}_2 + (\mu_3 - \mu_4) \vec{v}_3 + \mu_5 \vec{v}_5 \\ = & \mu_1 \vec{v}_1 + (\mu_2 + \mu_4) \vec{v}_2 + (\mu_3 - \mu_4) (-3\vec{v}_1 + 2\vec{v}_2) + \mu_5 \vec{v}_5 \\ = & (\mu_1 - 3\mu_3 + 3\mu_4) \vec{v}_1 + (\mu_2 + 2\mu_3 - \mu_4) \vec{v}_2 + \mu_5 \vec{v}_5 \in \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_5\} \end{aligned}$$

8.3 Matriisin säännöllisyysaste

Sec:MatRank

Tässä kappaleessa tarkastelemme $m \times n$ -matriisia $\mathbf{A}$. Merkitään matriisin k :ttä saraketta vektorina $\vec{a}_k$. Kappaleessa (??) määrittelimme matriisin sarakeavaruuden

$$\text{coll}(\mathbf{A}) = \text{span}\{\vec{a}_1, \dots, \vec{a}_n\},$$

joka on siis matriisin sarakkeiden virittämä $\mathbb{R}^m$:n aliavaruus, ja ytimen

$$\text{null}(\mathbf{A}) = \{\vec{x} \in \mathbb{R}^n \mid \mathbf{A}\vec{x} = \vec{0}\},$$

joka on $\mathbb{R}^n$:n aliavaruus.

Määritelmä 197 $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste, eli rangi, $\text{rank}(\mathbf{A})$, on matriisin sarakeavaruuden dimensio

$$\text{rank}(\mathbf{A}) = \dim(\text{coll}(\mathbf{A}))$$

ja matriisin nulliteetti on sen ytimen dimensio

$$\dim(\text{null}(\mathbf{A})).$$

Nulliteetille emme ota käyttöön mitään erityistä merkintää. Kun käsitettä tarvitaan, se on luonnollista merkitä ytimen avulla. Rangia käytetään paljon, sillä aiemmasta tiedämme, että

$$\begin{aligned} n \times n - \text{neliömatriisi } \mathbf{A} \text{ on säännöllinen} \\ \Leftrightarrow \text{rank}(\mathbf{A}) = n \end{aligned}$$

Tässä tapauksessa englanninkielinen kirjallisuus ei yleensä sano, että matriisi olisi ”regular”, vaan se sanoo, että matriisi on ”a matrix of full rank”.

Aiemmasta tiedämme, että rangi ja nulliteetti saadaan määritettyä tekemällä matriisille Gaussin algoritmin mukaiset rivioperaatiot. Onnistuneiden pivotointien lukumäärä on rangi ja epäonnistuneiden tai tekemättä jääneiden pivotointien lukumäärä on nulliteetti (vrt. esimerkki (I68) sivulla I48). Siis

Lause 198 $m \times n$ -matriisille $\mathbf{A}$ on voimassa

$$\text{rank}(\mathbf{A}) + \dim(\text{null}(\mathbf{A})) = n$$

Todistus. Jos ydin on triviaali, niin lauseen (I76) mukaan matriisin sarakkeet muodostavat vapaan systeemin ja lause on tosi.

Jos matriisin ydin ei ole triviaali, niin

$$\begin{aligned} \text{col}(\mathbf{A}^T)^\perp &= \{\vec{y} \in \mathbb{R}^m \mid \langle \vec{y} | \vec{z} \rangle = 0, \forall \vec{z} \in \text{col}(\mathbf{A}^T)\} \\ &= \{\vec{y} \in \mathbb{R}^m \mid \langle \vec{y} | \mathbf{A}^T \vec{x} \rangle = 0, \forall \vec{x} \in \mathbb{R}^n\} \\ &= \{\vec{y} \in \mathbb{R}^m \mid \langle \mathbf{A} \vec{y} | \vec{x} \rangle = 0, \forall \vec{x} \in \mathbb{R}^n\} \\ &= \text{null}(\mathbf{A}) \end{aligned} \tag{8.9}$$

Kaava tarkoittaa nyt sitä, että ytimen kanta voidaan aina konstruoida. Transpoosin sarakkeavaruuden kanta saadaan karsimalla sarakkeista ja ortogonaalikomplementin (eli ytimen) kanta saadaan lauseen (I92) konstruktiolla.

Olkoon nyt vektorijono $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_q\}$ ytimen kanta. Täydennetään tämä jono $\mathbb{R}^n$:n kannaksi $\{\vec{b}_1, \dots, \vec{b}_q, \vec{c}_1, \dots, \vec{c}_{n-q}\}$. Käytetään tässäkin lauseen (I92) konstruktiota.

Seuraavaksi osoitamme, että vektorijono $\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\}$ on aliavaruuden $\text{col}(\mathbf{A})$ kanta. Jos näin on, niin $\dim(\text{null}(\mathbf{A})) = q$ ja $\text{rank}(\mathbf{A}) = \dim(\text{col}(\mathbf{A})) = n - q$, joten

lauseen väite on tosi.

Ensin osoitamme, että $\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\}$ virittää aliavaruuden $\text{col}(\mathbf{A})$. Olkoon sitä varten $\vec{y} \in \mathbb{R}^m$ mikä tahansa $\text{col}(\mathbf{A})$:n vektori. Silloin on olemassa $\vec{x} \in \mathbb{R}^n$ siten, että

$$\vec{y} = x_1 \vec{a}_1 + \dots + x_n \vec{a}_n = \mathbf{A}\vec{x}.$$

Vektorilla $\vec{x}$ on yksikäsitteinen esitys $\mathbb{R}^n$:n kannassa $\{\vec{b}_1, \dots, \vec{b}_q, \vec{c}_1, \dots, \vec{c}_{n-q}\}$

$$\vec{x} = \mu_1 \vec{b}_1 + \dots + \mu_q \vec{b}_q + \mu_{q+1} \vec{c}_1 + \dots + \mu_n \vec{c}_{n-q}.$$

Silloin koska $\mathbf{A}\vec{b}_k = \vec{0}, \forall k$, saamme

$$\begin{aligned} \vec{y} &= \mathbf{A}(\mu_1 \vec{b}_1 + \dots + \mu_q \vec{b}_q + \mu_{q+1} \vec{c}_1 + \dots + \mu_n \vec{c}_{n-q}) \\ &= \mu_{q+1} \mathbf{A}\vec{c}_1 + \dots + \mu_n \mathbf{A}\vec{c}_{n-q} \\ &\in \text{span}\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\} \end{aligned}$$

Seuraavaksi osoitamme, että $\{\mathbf{A}\vec{c}_1, \mathbf{A}\vec{c}_2, \dots, \mathbf{A}\vec{c}_{n-q}\}$ on vapaa.

$$\begin{aligned} \lambda_1 \mathbf{A}\vec{c}_1 + \lambda_2 \mathbf{A}\vec{c}_2 + \dots + \lambda_{n-q} \mathbf{A}\vec{c}_{n-q} &= \vec{0} \\ \Leftrightarrow \mathbf{A}(\lambda_1 \vec{c}_1 + \lambda_2 \vec{c}_2 + \dots + \lambda_{n-q} \vec{c}_{n-q}) &= \vec{0} \\ \Leftrightarrow \lambda_1 \vec{c}_1 + \lambda_2 \vec{c}_2 + \dots + \lambda_{n-q} \vec{c}_{n-q} &\in \text{null}(\mathbf{A}) \end{aligned}$$

Nyt löytyy sellaiset kertoimet $\mu_1, \dots, \mu_q$, että

$$\begin{aligned} \lambda_1 \vec{c}_1 + \lambda_2 \vec{c}_2 + \dots + \lambda_{n-q} \vec{c}_{n-q} - \mu_1 \vec{b}_1 - \dots - \mu_q \vec{b}_q &= \vec{0} \\ \Leftrightarrow \lambda_1 = \lambda_2 = \dots = \lambda_{n-q} = \mu_1 = \dots = \mu_q &= 0 \end{aligned}$$

(Viimeinen ekvivalenssi johtuu siitä, että jono $\{\vec{b}_1, \dots, \vec{b}_q, \vec{c}_1, \dots, \vec{c}_{n-q}\}$ on kantana vapaa.) $\square$

Lause 199 $m \times n$ -matriisille $\mathbf{A}$ on voimassa

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T)$$

Todistus. Kaavan [\(8.9\)](#) mukaan $\text{col}(\mathbf{A}^T)^\perp = \text{null}(\mathbf{A})$, joten lauseen [\(I92\)](#) mukaan

$$\text{rank}(\mathbf{A}^T) + \dim(\text{null}(\mathbf{A})) = n$$

Edellisen lauseen mukaan

$$\text{rank}(\mathbf{A}) + \dim(\text{null}(\mathbf{A})) = n.$$

Ratkaisemalla rangit kummastakin yhtälöstä saadaan

$$\text{rank}(\mathbf{A}) = n - \dim(\text{null}(\mathbf{A})) = \text{rank}(\mathbf{A}^T)$$

□

Siis $m \times n$ -matriisin säännöllisyysaste on korkeintaan pienempi luvuista m ja n . Rangin laskeminen ei ole aina ongelmatonta. Ongelmat ovat numeerisia, eikä niihin puututa tässä kappaleessa. Singulaariarvo-hajoitelman avulla saamme numeerisesti hyvin käyttökelpoisen tavan laskea säännöllisyysaste. Toistaiseksi oletamme, että matriisin säännöllisyysaste tunnetaan tarvittaessa tarkasti.

8.4 Pseudoinverssi ja ortogonaaliprojektio

Sec:MatPInv

Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n ja $m > n$. ($\text{rank}(\mathbf{A}) = n$) Pohdimme (aina vain) yhtälöryhmän

$$\mathbf{A}\vec{x} = \vec{b}$$

ratkaisua. Ratkaisu on olemassa, jos

$$\begin{aligned} \vec{b} &\in \text{col}(\mathbf{A}) \\ \Leftrightarrow \quad &\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n, \vec{b}\} \text{ on sidottu} \\ \Leftrightarrow \quad &\det((\mathbf{A} | \vec{b})^T (\mathbf{A} | \vec{b})) = 0 \end{aligned}$$

Ratkaisun olemassaolo on siis helppo tarkistaa.

Jos kerroinmatriisin QR-hajotelma on tiedossa $\mathbf{A} = \mathbf{Q}\mathbf{R}$, niin yhtälöryhmällä $\mathbf{A}\vec{x} = \vec{b}$ on ratkaisu, jos ja vain jos yhtälöryhmällä

$$\mathbf{R}\vec{x} = \mathbf{Q}^T \vec{b} = \vec{c} \tag{8.10}$$

on ratkaisu. Jos RHS:n $m - n$ viimeistä koordinaattia ovat nollija, $c_{n+1} = \dots = c_m = 0$, niin yhtälöryhmällä on yksikäsitteinen ratkaisu, joka saadaan yksinkertaisten palautuvien sijoitusten avulla.

Jos $\mathbf{A}$ olisi neliömatriisi, niin saisimme ratkaisun laskettua käänteismatriisin avulla $\vec{x} = \mathbf{A}^{-1}\vec{b}$. Nyt kuitenkin käänteismatriisia ei ole olemassa.

Lause 200 Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n ja $m > n$. Jos $\vec{b} \in \text{col}(\mathbf{A})$, niin yhtälöryhmän

$$\mathbf{A}\vec{x} = \vec{b}$$

ratkaisu saadaan kertomalla RSH kerroinmatriisin pseudoinverssillä $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$. Siis

$$\vec{x} = \mathbf{A}^\dagger \vec{b}.$$

Todistus. Oletuksen mukaan on olemassa $\vec{x}$ siten, että $\mathbf{A}\vec{x} = \vec{b}$. Nyt

$$\mathbf{A}^\dagger \vec{b} = [(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T] \mathbf{A} \vec{x} = \vec{x}$$

□

Pseudoinverssi QR-hajotelman avulla.

Jos $m \times n$ -matriisin $\mathbf{A}$, jolle $\text{rank}(\mathbf{A}) = n$, QR-hajotelma on

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

niin

$$\begin{aligned} \mathbf{A}^\dagger &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \\ &= ((\mathbf{Q}\mathbf{R})^T \mathbf{Q}\mathbf{R})^{-1} (\mathbf{Q}\mathbf{R})^T \\ &= ((\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{Q}^T \\ &= \mathbf{R}^\dagger \mathbf{Q}^T \end{aligned}$$

Pseudoinverssi singulaariarvohajotelman avulla.

Jos $m \times n$ -matriisin $\mathbf{A}$, jolle $\text{rank}(\mathbf{A}) = n$, singulaariarvohajotelma on

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T,$$

niin

$$\begin{aligned} \mathbf{A}^\dagger &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \\ &= \left[\left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right]^{-1} \left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \\ &= \left[\mathbf{V} \begin{pmatrix} \mathbf{\Sigma}^T & \mathbf{0} \end{pmatrix} \mathbf{U}^T \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right]^{-1} \left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \\ &= [\mathbf{V} \mathbf{\Sigma}^2 \mathbf{V}^T]^{-1} \left(\mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V}^T \right)^T \\ &= [\mathbf{V} \mathbf{\Sigma}^{-2} \mathbf{V}^T] (\mathbf{V} \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \end{pmatrix} \mathbf{U}^T) \\ &= \mathbf{V} \begin{pmatrix} \mathbf{\Sigma}^{-1} & \mathbf{0} \end{pmatrix} \mathbf{U}^T \end{aligned}$$

8.4.1 Ortogonaaliprojektio

Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n . Oletamme, että vektori $\vec{y} = (y_1, y_2, \dots, y_n)^T$ ei kuulu $\mathbf{A}$:n sarakeavaruuteen

$$\vec{y} \notin \text{col}(\mathbf{A}) = \text{span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}.$$

Asetamme tehtäväksi nyt etsiä se sarakeavaruuden vektori $\vec{x} \in \text{col}(\mathbf{A})$, jolla

$$\|\vec{x} - \vec{y}\|_2$$

saa pienimmän mahdollisen arvon. Muutaman välivaiheen kautta osoitamme, että tehtävän ratkaisu on

$$\vec{x} = \mathbf{A}\mathbf{A}^\dagger \vec{y}.$$

Lause 201 *Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n . Jos $\vec{x} = \mathbf{A}\mathbf{A}^\dagger \vec{y}$, niin*

$$(\vec{x} - \vec{y}) \perp \text{col}(\mathbf{A}).$$

Todistus.

$$\mathbf{A}^T(\vec{x} - \vec{y}) = \mathbf{A}^T \mathbf{A} \mathbf{A}^\dagger \vec{y} - \mathbf{A}^T \vec{y} = 0 = \mathbf{A}^T \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \vec{y} - \mathbf{A}^T \vec{y} = 0$$

□

c4s21thortproj

Lause 202 *Olkoon $m \times n$ -matriisin $\mathbf{A}$ säännöllisyysaste n . Jos $\vec{x} = \mathbf{A}\mathbf{A}^\dagger \vec{y}$, niin $\vec{x}$ on lähinnä vektoria $\vec{y}$ oleva sarakeavaruuden $\text{col}(\mathbf{A})$ vektori. Siis*

$$\|\vec{x} - \vec{y}\|_2 = \min_{\vec{w} \in \text{col}(\mathbf{A})} \|\vec{w} - \vec{y}\|_2.$$

Todistus. Olkoon $\vec{w} \in \text{col}(\mathbf{A})$ mikä tahansa sarakeavaruuden vektori. Silloin

$$\vec{w} - \vec{y} = (\vec{w} - \vec{x}) + (\vec{x} - \vec{y})$$

ja $(\vec{w} - \vec{x}) \perp (\vec{x} - \vec{y})$. Silloin Pythagoran lauseen nojalla

$$\begin{aligned} \|\vec{w} - \vec{y}\|_2^2 &= \|\vec{w} - \vec{x}\|_2^2 + \|\vec{x} - \vec{y}\|_2^2 \\ \Rightarrow \|\vec{w} - \vec{y}\|_2 &\geq \|\vec{x} - \vec{y}\|_2 \end{aligned}$$

□

8.5 Ortonormeerattu kanta, Gram-Schmidt

Sec:BaseOrt

Tässä vaiheessa selvä perusongelma on vektorin koordinaattien laskeminen annetussa kannassa. Olkoon $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ lineaariavaruuden $L \subseteq \mathbb{R}^m$ kanta ja $\mathbf{V}$ $m \times n$ -matriisi, jonka sarakkeina ovat kantavektorit. Oletamme seuraavassa, että vektori $\vec{a}$ kuuluu kannan virittämään aliavaruuteen. Vektorin $\vec{a}$ koordinaatit $\alpha_1, \dots, \alpha_m$ ja niistä muodostettu koordinaattivektori $\vec{\alpha} = (\alpha_1 \dots \alpha_m)^T$ toteuttavat yhtälöt

$$\begin{aligned} \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_m \vec{v}_m &= \vec{a} \\ \Leftrightarrow \mathbf{V} \vec{\alpha} &= \vec{a}. \end{aligned} \quad (8.11)$$

Koska kanta on vapaa, on $\text{rank}(\mathbf{V}) = n$, ja koordinaattivektori voidaan kätevästi ratkaista kaavalla

$$\vec{\alpha} = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \vec{a}. \quad (8.12) \quad \text{kooreq}$$

Huomaa, että tämä kaava antaa väärän vastauksen, jos $\vec{a} \notin \text{col}(\mathbf{V})$. Silloin ei nimittäin koordinaatteja ole olemassa, mutta kaavasta sellaiset saadaan laskettua. Ne ovat $\text{coll}(\mathbf{V})$:n lähimpänä $\vec{a}$:ta olevan vektorin koordinaatit.

Tarkastellaan ongelmaa hieman eri näkökulmasta. Lasketaan vektorin $\vec{a}$ sisätulot kantavektoreiden kanssa.

$$\begin{aligned} \langle \vec{a} | \vec{v}_i \rangle &= \langle \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_m \vec{v}_m | \vec{v}_i \rangle \\ &= \alpha_1 \langle \vec{v}_1 | \vec{v}_i \rangle + \alpha_2 \langle \vec{v}_2 | \vec{v}_i \rangle + \dots + \alpha_m \langle \vec{v}_m | \vec{v}_i \rangle \end{aligned}$$

Jos nyt $\langle \vec{v}_i | \vec{v}_j \rangle = \delta_{ij}$ eli

$$\vec{v}_i \perp \vec{v}_j, \quad \forall i \neq j \quad \text{ja} \quad (8.13)$$

$$\|\vec{v}_i\| = 1, \quad \forall i, \quad (8.14)$$

niin

$$\alpha_i = \langle \vec{a} | \vec{v}_i \rangle$$

eli

$$\vec{a} = \langle \vec{a} | \vec{v}_1 \rangle \vec{v}_1 + \langle \vec{a} | \vec{v}_2 \rangle \vec{v}_2 + \dots + \langle \vec{a} | \vec{v}_m \rangle \vec{v}_m \quad (8.15) \quad \text{kooreq2}$$

Esitys (8.15) on yksinkertainen ja helppo muistaa. Esitys on käytettävissä vain jos käytetty kanta toteuttaa ehdot (8.13, 8.14). Annamme ehdot toteuttaville kannoille nimen, jotta voimme jatkossa puhua niistä sujuvasti

Määritelmä 203 Lineaariavaruuden L vektorisysteemi $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ on ortonormaalinen, jos systeemin vektorit ovat pareittain kohtisuorassa keskenään,

$$i \neq j \Rightarrow \mathbf{v}_i \perp \mathbf{v}_j.$$

Jos lisäksi

$$\|\mathbf{v}_i\| = \sqrt{\langle \mathbf{v}_i | \mathbf{v}_i \rangle} = 1, \quad \forall i = 1, \dots, n,$$

niin systeemi on ortonormaali (ortonormeerattu).

Laksut usein yksinkertaistuvat, kun käytetty kanta on ortonormitettu. Koulukurssista tutut vektorit $\vec{i} = (1, 0, 0)^T$, $\vec{j} = (0, 1, 0)^T$ ja $\vec{k} = (0, 0, 1)^T$ muodostavat $\mathbb{R}^3$:n kannan $F = \{\vec{i}, \vec{j}, \vec{k}\}$. Toisaalta myös jono $G = \{2\vec{i}, 2\vec{j}, 2\vec{k}\}$ on $\mathbb{R}^3$:n kanta. Jos vektorin $\vec{a}$ koordinaatit kannassa F ovat $(1 \ 1 \ 1)^T$, niin sen koordinaatit kannassa G ovat $(0.5 \ 0.5 \ 0.5)^T$. Nyt on selvää, ettei vektorin pituutta voi laskea suoraan koordinaateista ilman jonkinlaista tietoa koordinaatteihin liittyvästä kannasta. Ortonormalisuus riittää.

Lause 204 (Pythagoraan lauseen yleistys) Olkoon L reaalinen lineaariavaruus, jossa on määritelty sisätulo, ja olkoon vektorijono $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q\}$ sen ortonormaali kanta. Jos

$$\begin{aligned} \mathbf{a} &= a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_q \mathbf{v}_q, \quad \text{ja} \\ \mathbf{b} &= b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \dots + b_q \mathbf{v}_q, \end{aligned}$$

niin

$$\begin{aligned} \langle \mathbf{a} | \mathbf{b} \rangle &= a_1 b_1 + a_2 b_2 + \dots + a_q b_q \\ \|\mathbf{a}\|^2 &= \langle \mathbf{a} | \mathbf{a} \rangle = a_1^2 + a_2^2 + \dots + a_q^2. \end{aligned}$$

Todistus. Suora lasku. □

Siis vektorin pituus ja kahden vektorin kohtisuoruus voidaan tutkia suoraan koordinaattien avulla, JOS KANTA ON ORTONORMITETTU.

Seuraavaksi kuvailemme Gram-Schmidt -algoritmin, jonka avulla saamme aina muutettua $\mathbb{R}^n$:n aliavaruuden kannan ortonormeeratuksi. Oletamme, että vektorijono $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ on vapaa eli aliavaruuden $M = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ kanta. Muodostamme M :lle uuden kannan $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_q\}$ seuraavasti.

Gram-Schmidt:

(1) Ensimmäiseksi asetamme

$$\vec{b}_1 = \frac{1}{\|\vec{v}_1\|} \vec{v}_1. \quad (8.16) \quad \text{GrSm1}$$

(2) Toiseksi asetamme

$$\vec{a}_2 = \vec{v}_2 - \langle \vec{v}_2 | \vec{b}_1 \rangle \vec{b}_1 \quad (8.17)$$

$$\vec{b}_2 = \frac{1}{\|\vec{a}_2\|} \vec{a}_2. \quad (8.18)$$

(3) Kolmanneksi asetamme

$$\vec{a}_3 = \vec{v}_3 - \langle \vec{v}_3 | \vec{b}_1 \rangle \vec{b}_1 - \langle \vec{v}_3 | \vec{b}_2 \rangle \vec{b}_2 \quad (8.19)$$

$$\vec{b}_3 = \frac{1}{\|\vec{a}_3\|} \vec{a}_3. \quad (8.20)$$

(4...q) k:s kantavektori saadaan asettamalla

$$\vec{a}_k = \vec{v}_k - \sum_{j=1}^{k-1} \langle \vec{v}_k | \vec{b}_j \rangle \vec{b}_j \quad (8.21)$$

$$\vec{b}_k = \frac{1}{\|\vec{a}_k\|} \vec{a}_k. \quad (8.22)$$

Lause 205 Olkoon $F = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ vapaa jono $\mathbb{R}^n$:n vektoreita, eli virittämässä aliavaruuden $M = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ kanta. Gram-Schmidt -menetelmällä (8.16-8.22) muodostettu vektorijono $G = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_q\}$ on M :n ortonormaali kanta. GrSm1

Todistus. Vektorit $\vec{v}_1, \dots, \vec{v}_q$ ovat selvästi yksikkövektoreita, kunhan mikään $\vec{a}_k$ ei ole nollavektori. Riittää siis osoittaa, että jono $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_q\}$ on vapaa ja ortogonaalinen.

Tehdään induktiotarkastelu:

Induktion lähtökohta: jono $\{\vec{a}_1\}$ on vapaa ja ortogonaalinen.

Induktio-oletus: jono $\{\vec{a}_1, \dots, \vec{a}_{k-1}\}$ on vapaa ja ortogonaalinen.

Induktioväite: jono $\{\vec{a}_1, \dots, \vec{a}_k\}$ on vapaa ja ortogonaalinen.

Induktiotodistus: Jos pystymme osoittamaan, että $\vec{a}_k \perp \vec{a}_s$ kaikilla $s = 1, \dots, k-1$, niin jono $\{\vec{a}_1, \dots, \vec{a}_k\}$ on ortogonaalinen ja lauseen (I91) mukaan vapaa. c4s8th7b

Yhtälöt ^{GrSm6}(8.21) ja ^{GrSm7}(8.22) voidaan tiivistää yhdeksi yhtälöksi

$$\vec{a}_k = \vec{v}_k - \sum_{j=1}^{k-1} \rho_{kj} \vec{a}_j, \quad \rho_{kj} = \frac{\langle \vec{v}_k | \vec{a}_j \rangle}{\langle \vec{a}_j | \vec{a}_j \rangle}$$

Induktio-oletuksen mukaan kaikilla $s < k$ on

$$\begin{aligned} \langle \vec{a}_k | \vec{a}_s \rangle &= \left\langle \vec{v}_k - \sum_{j=1}^{k-1} \rho_{kj} \vec{a}_j \middle| \vec{a}_s \right\rangle \\ &= \langle \vec{v}_k | \vec{a}_s \rangle - \sum_{j=1}^{k-1} \rho_{kj} \underbrace{\langle \vec{a}_j | \vec{a}_s \rangle}_{=0, \text{ kun } j \neq s} \\ &= \langle \vec{v}_k | \vec{a}_s \rangle - \rho_{ks} \langle \vec{a}_s | \vec{a}_s \rangle \\ &= 0 \end{aligned}$$

□

Gram-Schmidt -menettely on käsin laskettaessa raskas. Se onkin enemmän teoreettinen asia, kuin käytännöllinen työkalu.

Lause 206 Jos $\mathbb{R}^n$:n aliavaruudessa M on nollavektorin lisäksi muitakin vektoreita, niin aliavaruudella on ortonormeerattu kanta.

Todistus. Olkoon $\vec{a} \in \mathbb{R}^n$ ja $\vec{a} \neq \vec{0}$. Silloin jono $\{\vec{a}\}$ on vapaa jono, jonka voimme täydentää $\mathbb{R}^n$:n kannaksi. Tämä kanta voidaan edelleen ortonormeerata Gram-Schmidt -menetelmällä. □

Esimerkki 207 Tarkastellaan $\mathbb{R}^5$:n vektoreita $\vec{v}_1 = (1, 0, 2, -1, 3)^T$, $\vec{v}_2 = (3, 1, -1, 2, 0)^T$ ja $\vec{v}_3 = (0, 1, 3, 5, -1)^T$. Muodostetaan näiden vektoreiden virittämän aliavaruuden ortonormeerattu kanta Matlabilla seuraavasti.

```
>> v1 = [1 0 2 -1 3]';
>> v2 = [3 1 -1 2 0]';
>> v3 = [0 1 3 5 -1]';
```

```
>> b1 = v1/norm(v1)
b1 =
    0.2582
         0
```

```
0.5164
-0.2582
0.7746

>> a2 = v2 - (v2'*b1)*b1;
>> b2 = a2/norm(a2)
b2 =
    0.7936
    0.2588
   -0.2243
    0.5003
    0.0518

>> a3 = v3 - (v3'*b1)*b1 - (v3'*b2)*b2;
>> b3 = a3/norm(a3)
b3 =
   -0.2637
    0.0842
    0.6624
    0.6847
   -0.1255

>> %Tarkistuksia
>> b1'*b2
ans =
   -2.7756e-017

>> b1'*b3
ans =
   -8.3267e-017

>> b2'*b3
ans =
    1.1102e-016

>> norm(b1)
ans =
    1

>> norm(b2)
ans =
```

```

1

>> norm(b3)
ans =
    1
>>

```

8.6 Ortogonaaliset matriisit, QR-hajoitelma

Sec:MatOrt

Määritelmä 208 Neliömatriisi $\mathbf{A}$ on ortogonaalinen, jos

$$\mathbf{A}^{-1} = \mathbf{A}^T.$$

Määritelmä on hämäävän yksinkertainen. Seuraava lause kertoo paremmin mistä oikein on kysymys.

Lause 209 Olkoon $\mathbf{A}$ neliömatriisi. Silloin seuraavat kolme ominaisuutta ovat yhtäpitävät:

- (1) $\mathbf{A}$ on ortogonaalinen.
- (2) $\mathbf{A}$:n sarakkeet muodostavat ortonormeeratun systeemin.
- (3) $\mathbf{A}$:n rivit muodostavat ortonormeeratun systeemin.

Todistus. Ensin osoitetaan, että (1):stä seuraa (2) ja (3).

$$\begin{aligned}
 (1) \quad &\Leftrightarrow \mathbf{A}^{-1} = \mathbf{A}^T \\
 &\Leftrightarrow \begin{cases} \mathbf{A}^T \mathbf{A} = \mathbf{I} & \Rightarrow (2) \\ \mathbf{A} \mathbf{A}^T = \mathbf{I} & \Rightarrow (3) \end{cases}
 \end{aligned}$$

Toiseksi osoitamme, että (2):sta seuraa (1). Oletamme siis, että $\mathbf{A}$ on neliömatriisi, jonka sarakkeet muodostavat ortonormeeratun systeemin. Silloin $\mathbf{A}^T \mathbf{A} = \mathbf{I}$, joten $\det(\mathbf{A}) \neq 0$. Siis $\mathbf{A}$ on säännöllinen eli sillä on käänteismatriisi. Nyt

$$\mathbf{A} \mathbf{A}^T = \mathbf{A} \underbrace{\mathbf{A}^T \mathbf{A}}_{=\mathbf{I}} \mathbf{A}^{-1} = \mathbf{I}$$

Kolmanneksi osoitamme, että (3):sta seuraa (1)

$$\begin{aligned}
 (3) \mathbf{A}: \text{lle} &\Leftrightarrow (2) \mathbf{A}^T: \text{lle} \\
 &\Rightarrow (1) \mathbf{A}^T: \text{lle} \\
 &\Leftrightarrow (\mathbf{A}^T)^{-1} = (\mathbf{A}^T)^T \\
 &\Leftrightarrow (\mathbf{A}^{-1})^T = \mathbf{A} \\
 &\Leftrightarrow \mathbf{A}^{-1} = \mathbf{A}^T \\
 &\Leftrightarrow (1) \mathbf{A}: \text{lle}
 \end{aligned}$$

□

Lause 210 Olkoon $n \times n$ -matriisi $\mathbf{A}$ ortogonaalinen ja $\vec{b}, \vec{c} \in \mathbb{R}^n$.

(1) Jos $\vec{b} \perp \vec{c}$, niin $\mathbf{A}\vec{b} \perp \mathbf{A}\vec{c}$.

(2) $\|\mathbf{A}\vec{b}\|_2 = \|\vec{b}\|_2$

(3) Jos $\mathbb{R}^n$:n vektorijono $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ on $\mathbb{R}^n$:n ortonormitettu kanta, niin vektorijono $\{\mathbf{A}\vec{u}_1, \mathbf{A}\vec{u}_2, \dots, \mathbf{A}\vec{u}_n\}$ on myös ortonormitettu kanta.

(Verbaalisti: Ortonormaalilla matriisilla kertominen säilyttää kohtisuoruuden ja 2-normin.)

Todistus. HT.

□

Lause 211 $n \times n$ -matriisi $\mathbf{A}$, jonka sarakkeet ovat lineaarisesti riippumattomat (siis $\text{rank}(\mathbf{A}) = n$), voidaan aina kirjoittaa muotoon

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

missä $\mathbf{Q}$ on ortogonaalinen $n \times n$ -matriisi ja $\mathbf{R}$ on $n \times n$ -yläkolmiomatriisi ($r_{ij} = 0$, kun $i > j$), jonka lävistäjäalkiot ovat nollasta eroavat.

Todistus. Matriisin $\mathbf{A}$ sarakevektorit $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ muodostavat vapaan jonon, joten voimme ortonormeerata sen Gram-Schmidt -menetelmällä jonoksi $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$. Kirjoitamme uudelleen Gram-Schmidt -menetelmän yhtälöt (8.21) ja (8.22) tähän tilanteeseen kuuluvien merkinnöin.

$$\vec{c}_k = \vec{a}_k - \sum_{j=1}^{k-1} \langle \vec{a}_k | \vec{q}_j \rangle \vec{q}_j \quad (8.23)$$

$$\vec{q}_k = \frac{1}{\|\vec{c}_k\|} \vec{c}_k. \quad (8.24)$$

Yhtälöt voidaan yhdistää seuraavasti

$$\vec{a}_k = \sum_{j=1}^k r_{jk} \vec{q}_j \quad (8.25) \quad \boxed{\text{c4s11GrSmeq}}$$

ja $r_{kk} \neq 0, \forall k$. Voimme lisäksi asettaa $r_{jk} = 0$, kun $j > k$. Edellinen yhtälö pitää kirjoittaa jokaiselle k :n arvolle. Siis n yhtälöä. Nämä kaikki yhtälöt saadaan koottua siistiin pakettiin kirjoittamalla seuraava matriisiyhtälö

$$\begin{pmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{pmatrix} = \begin{pmatrix} \vec{q}_1 & \dots & \vec{q}_n \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix} \quad (8.26)$$

$$\Leftrightarrow \mathbf{A} = \mathbf{QR}$$

□

Todistus on valmis, mutta koska on olemassa vaara, että yhtälö (8.26) jää ymmärtämättä, kirjoitamme sen auki poikkeuksellisen perusteellisesti

$$\begin{pmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{pmatrix} = \begin{pmatrix} \vec{q}_1 & \dots & \vec{q}_n \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \dots & q_{nn} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix}$$

Laskemalla matriisikertolasku RHS:ssa saadaan

$$\vec{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = \begin{pmatrix} r_{11}q_{11} \\ r_{11}q_{21} \\ \vdots \\ r_{11}q_{n1} \end{pmatrix} = r_{11}\vec{q}_1$$

$$\vec{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} = \begin{pmatrix} r_{12}q_{11} + r_{22}q_{12} \\ r_{12}q_{21} + r_{22}q_{22} \\ \vdots \\ r_{12}q_{n1} + r_{22}q_{n2} \end{pmatrix} = r_{12}\vec{q}_1 + r_{22}\vec{q}_2$$

ja niin edelleen yhtälön (8.25) mukaisesti.

QR-hajoitelman olemassaolo perusteltiin edellä Gram-Schmidt algoritmin avulla. QR-hajoitelmaa ei käytännössä kannata laskea Gram-Schmidt -algoritmillä, vaan siihen tarkoitukseen on olemassa numeerisesti luotettavampia menetelmiä, joita emme nyt käsittele. (Ns. Housholderin transformaatiot. ks Kivelä: 'Matriisilasku ja Lineaarialgebra' ss. 123–127.)

Esimerkki 212 Jos käytettävissä on Matlab-ohjelma, niin voimme pyytää QR-hajoitelman yksinkertaisella komennolla. Matriisiin

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

QR-hajoitelma saadaan seuraavasti:

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

```
A =
```

```

1      2      0      1
3      2      1      0
0      1      2      1
4     -1      3      1
```

```
>> [Q,R]=qr(A)
```

```
Q =
```

```

-0.1961    -0.6026     0.1407    -0.7606
-0.5883    -0.5022     0.2272     0.5916
         0    -0.3264    -0.9414     0.0845
-0.7845     0.5273    -0.2056    -0.2535
```

```
R =
```

```

-5.0990    -0.7845    -2.9417    -0.9806
         0    -3.0634     0.4269    -0.4018
         0         0    -2.2724    -1.0064
         0         0         0    -0.9297
```

```
>>
```

QR-hajoitelma voidaan muodostaa myös $m \times n$ -matriisille, jos $m > n$ ja $\text{rank}(A) = n$. Sama konstruktio kuin edellä, johtaa $m \times n$ matriisiin Q_1 ja $n \times n$ -yläkolmiomatriisiin

$\mathbf{R}_1$ siten, että matriisin $\mathbf{Q}_1$ sarakkeet muodostavat ortonormeeratun systeemin, $\mathbf{R}_1$:n diagonaali-alkiot eroavat nolasta ja

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R}_1.$$

Koska matriisin $\mathbf{Q}_1$ sarakevektorit $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$ muodostavat vapaan jonon, se voidaan täydentää $\mathbb{R}^m$:n kannaksi ja sen jälkeen ortonormeerata. Näin saadaan uusi vektorijono $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n, \vec{w}_1, \dots, \vec{w}_{m-n}\}$, jossa ensimmäiset vektorit ovat $\mathbf{Q}_1$:n sarakkeet ja loput vektorit ovat uusia. Näistä koottu matriisi

$$\mathbf{Q} = (\mathbf{Q}_1 | \mathbf{W})$$

on ortogonaalinen $m \times m$ -matriisi. Täydennetään edellä saatua $\mathbf{R}_1$ matriisia lisäämällä siihen $m - n$ nollariviä alimmiksi riveiksi. Saatua matriisi

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{pmatrix}$$

on $m \times n$ matriisi, jossa diagonaalin alapuolella on vain nollia. Nyt

$$\mathbf{QR} = (\mathbf{Q}_1 | \mathbf{W}) \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{pmatrix} = \mathbf{Q}_1 \mathbf{R}_1 = \mathbf{A}$$

Esimerkki 213 *Matriisin*

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \\ 0 & 1 \\ 4 & -1 \end{pmatrix}$$

QR-hajoitelma saadaan seuraavasti:

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

```
A =
```

```
1      2
3      2
0      1
4     -1
```

```
>> [Q,R] = qr(A)
```

```
Q =
```

```
-0.1961    -0.6026    -0.0039    -0.7735
-0.5883    -0.5022    -0.3283     0.5421
0          -0.3264     0.9116     0.2498
-0.7845     0.5273     0.2472    -0.2132
```

```

R =
    -5.0990    -0.7845
         0    -3.0634
         0         0
         0         0
>>

```

Matlabin komento on myös mahdollista antaa muodossa, jossa lisäykset jää pois

```

>> [Q,R] = qr(A,0)
Q =
    -0.1961    -0.6026
    -0.5883    -0.5022
         0    -0.3264
    -0.7845     0.5273

R =
    -5.0990    -0.7845
         0    -3.0634
>>

```

Nyt siis saimme $\mathbf{Q}_{1:n}$ ja $\mathbf{R}_{1:n}$

8.7 Lineaarikuvauksen matriisi

Sec:LatLinMap

Tässä kappaleessa oletamme, että L ja M ovat äärellisulotteisia lineaariavaruuksia. Olkoon $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ lineaariavaruuden L kanta ja olkoon $G = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ lineaariavaruuden M kanta. (Toisin sanoen $\dim(L) = n$ ja $\dim(M) = m$.)

Jokaisella L :n vektorilla $\mathbf{x}$ on yksikäsitteiset koordinaatit $(x_1, x_2, \dots, x_n)$ kannassa F siten, että

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n$$

Vastaavasti jokaisella M :n vektorilla $\mathbf{y}$ on yksikäsitteiset koordinaatit $(y_1, y_2, \dots, y_m)$ kannassa G siten, että

$$\mathbf{y} = y_1 \mathbf{w}_1 + y_2 \mathbf{w}_2 + \dots + y_m \mathbf{w}_m$$

Olkoon $f : L \rightarrow M$ on lineaarikuvaus. Jos tunnemme kantavektoreiden $\mathbf{v}_i$ kuvat, niin koko kuvaus tunnetaan täydellisesti. Olkoon nimittäin

$$f(\mathbf{v}_i) = a_{i1}\mathbf{w}_1 + a_{i2}\mathbf{w}_2 + \cdots + a_{im}\mathbf{w}_m, \quad \text{kaikilla } i = 1, 2, \dots, n$$

Jos nyt $\mathbf{y} = f(\mathbf{x})$, niin

$$\begin{aligned} \mathbf{y} &= y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_n\mathbf{w}_m \\ &= f(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n) \\ &= x_1f(\mathbf{v}_1) + x_2f(\mathbf{v}_2) + \cdots + x_nf(\mathbf{v}_n) \\ &= x_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m) \\ &\quad + x_2(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m) \\ &\quad + \cdots + \\ &\quad x_n(a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m) \end{aligned}$$

Vertaamalla ensimmäisen ja viimeisen lausekkeen kertoimia keskenään, näemme, että

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\Leftrightarrow \quad \vec{y}_G = \mathbf{A}\vec{x}_F$$

Muotoilemme edellisen lauseeksi:

Lause 214 *Olkoon $f : L \rightarrow M$ lineaarikuvaus. Jos $F = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ on lineaariavaruuden L kanta, ja $G = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ on lineaariavaruuden M kanta, niin on olemassa kantoihin F ja G ja lineaarikuvaukseen f liittyvä $m \times n$ -matriisi $\mathbf{A}$ siten, että seuraava pitää paikkansa:*

Jos $\mathbf{x} \in L$ ja $f(\mathbf{x}) = \mathbf{y} \in M$, niin

$$\vec{y}_G = \mathbf{A}\vec{x}_F,$$

missä $\vec{y}_G$ on $\mathbf{y}$:n koordinaattivektori kannassa G ja $\vec{x}_F$ on $\mathbf{x}$:n koordinaattivektori kannassa F . (Matriisin $\mathbf{A}$ k :s sarake on vektorin $f(\mathbf{v}_k)$ koordinaattivektori kannassa G .) $\square$

Esimerkki 215 Olkoon P_4 korkeintaan astetta 4 olevien polynomifunktioiden lineaariavaruus. Kantana voimme käyttää viisikkoa

$$\{1, x, x^2, x^3, x^4\}$$

Derivaattakuvauksen

$$D : P_4 \rightarrow P_4, p(x) \mapsto p'(x)$$

matriisi on nyt

$$\mathbf{A}_D = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Esimerkki 216 Tarkastellaan toiseksi matriisin

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

määrittää lineaarikuvausta $f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f_A(\vec{x}) = \mathbf{A}\vec{x}$. Jos varustamme lähtö- ja maaliavaruudet luonnollisilla kannoilla

$$F = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad \text{ja} \quad G = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

niin kuvauksen f_A matriisi on $\mathbf{A}$ itse.

Toisaalta on hyvin luonnollista rajoittaa kuvauksen maaliavaruus aliavaruuteen $M = \text{col}(\mathbf{A}) = \text{span}(\vec{a}_1, \vec{a}_2)$. Nyt kuvauksen $f_A : \mathbb{R}^2 \rightarrow M$, $f_A(\vec{x}) = \mathbf{A}\vec{x}$ maaliavaruuden kannaksi on luonnollista valita vapaa vektorijono

$$H = \left\{ \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} \right\}.$$

Tällä kanta-valinnalla kuvauksen f_A matriisi on

$$\mathbf{A}_f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Kolmanneksi siirrymme ortonormitettuihin koordinaatteihin. F on ortonormitettu, mutta H ei ole. Ortonormitamme vektorijonon H . Muodostetaan vastaava matriisi $\mathbf{H}$ ja sen QR-hajoitelma

```
>> H=[1 2 ; 2 -1 ; 2 2]
```

```
H =
```

```
    1    2
    2   -1
    2    2
```

```
>> [Q1,R1] = qr(H,0)
```

```
Q1 =
```

```
-0.3333    0.5788
-0.6667   -0.7029
-0.6667    0.4134
```

```
R1 =
```

```
-3.0000   -1.3333
         0    2.6874
```

```
>>
```

Siis uusi ortonormitettu maaliavaruuden kanta on

$$H_2 = \left\{ \begin{pmatrix} -0.3333 \\ -0.6667 \\ -0.6667 \end{pmatrix}, \begin{pmatrix} 0.5788 \\ -0.7029 \\ 0.4134 \end{pmatrix} \right\}$$

Määritämme nyt kuvauksen matriisin tällä kanta valinnalla.

$$f(\vec{e}_1) = \mathbf{A}\vec{e}_1 = \mathbf{Q}_1\mathbf{R}_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mathbf{Q}_1\vec{r}_1$$

$$f(\vec{e}_2) = \mathbf{A}\vec{e}_2 = \mathbf{Q}_1\mathbf{R}_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{Q}_1\vec{r}_2$$

ja kuvauksen matriisi on tällä kanta-valinnalla

$$\mathbf{A}_f = \mathbf{R}_1 = \begin{pmatrix} -3 & -1.3333 \\ 0 & 2.6874 \end{pmatrix}.$$

Tärkeintä on nyt muistaa, että kuvauksen matriisi riippuu valituista lähtö ja maaliavaruuden kannoista!

8.8 Kannan vaihto

Sec:BaseChange

Tarkastellaan kahta n -ulotteisen lineaariavaruuden L kantaa $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ ja $\hat{E} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_n\}$. Kannan E kantavektorit voidaan lausua yksikäsitteisellä tavalla kannassa $\hat{E}$

$$\begin{cases} \mathbf{e}_1 = a_{11}\hat{\mathbf{e}}_1 + a_{21}\hat{\mathbf{e}}_2 + \dots + a_{n1}\hat{\mathbf{e}}_n \\ \mathbf{e}_2 = a_{12}\hat{\mathbf{e}}_1 + a_{22}\hat{\mathbf{e}}_2 + \dots + a_{n2}\hat{\mathbf{e}}_n \\ \vdots \\ \mathbf{e}_n = a_{1n}\hat{\mathbf{e}}_1 + a_{2n}\hat{\mathbf{e}}_2 + \dots + a_{nn}\hat{\mathbf{e}}_n \end{cases}$$

Huomaa, että indeksointi on nyt tehty erikoisesti. Edelläoleva on helpointa muistaa seuraavasti

$$E^T = \mathbf{A}^T \hat{E}^T \Leftrightarrow E = \hat{E} \mathbf{A}.$$

Olkoon nyt $\mathbf{u}$ lineaariavaruuden L vektori, jolla on eri kannoissa esitykset

$$\begin{aligned} \mathbf{u} &= x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n, \quad \text{ja} \\ \mathbf{u} &= \hat{x}_1 \hat{\mathbf{e}}_1 + \hat{x}_2 \hat{\mathbf{e}}_2 + \dots + \hat{x}_n \hat{\mathbf{e}}_n. \end{aligned}$$

Silloin

$$\begin{aligned} \mathbf{u} &= \sum_{j=1}^n x_j \mathbf{e}_j = \sum_{j=1}^n x_j \left(\sum_{k=1}^n a_{kj} \hat{\mathbf{e}}_k \right) \\ &= x_1 (a_{11} \hat{\mathbf{e}}_1 + a_{21} \hat{\mathbf{e}}_2 + \dots + a_{n1} \hat{\mathbf{e}}_n) \\ &\quad + x_2 (a_{12} \hat{\mathbf{e}}_1 + a_{22} \hat{\mathbf{e}}_2 + \dots + a_{n2} \hat{\mathbf{e}}_n) \\ &\quad + \dots + \\ &\quad + x_n (a_{1n} \hat{\mathbf{e}}_1 + a_{2n} \hat{\mathbf{e}}_2 + \dots + a_{nn} \hat{\mathbf{e}}_n) \\ &= (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) \hat{\mathbf{e}}_1 \\ &\quad + (a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n) \hat{\mathbf{e}}_2 \\ &\quad + \dots \\ &\quad + (a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n) \hat{\mathbf{e}}_n \end{aligned}$$

Viimeisestä muodosta voimme lukea koordinaatit hatullisessa kannassa. Siis

$$\begin{cases} \hat{x}_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \hat{x}_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ \hat{x}_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{cases}$$

Merkitsemme seuraavasti.

Kannat:

$$E = (\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_n), \quad \text{ja} \quad \hat{E} = (\hat{\mathbf{e}}_1 \quad \hat{\mathbf{e}}_2 \quad \dots \quad \hat{\mathbf{e}}_n)$$

Vektorin $\mathbf{u}$ koordinaattivektorit kannoissa:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \text{ missä } \mathbf{u} = E\vec{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$$

$$\vec{\hat{x}} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_n \end{pmatrix}, \text{ missä } \mathbf{u} = \hat{E}\vec{\hat{x}} = \hat{x}_1\hat{\mathbf{e}}_1 + \cdots + \hat{x}_n\hat{\mathbf{e}}_n$$

Näillä merkinnöillä saamme tuloksen.

Lause 217 Jos $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ ja $\hat{E} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_n\}$ ovat kaksi lineaariavaruuden L kantaa ja jos $E^T = \mathbf{A}^T \hat{E}^T$ (eli $E = \hat{E}\mathbf{A}$), niin vektorin koordinaatit toteuttavat yhtälöt

$$\vec{\hat{x}} = \mathbf{A}\vec{x}$$

$$\vec{x} = \mathbf{A}^{-1}\vec{\hat{x}}$$

□

Esimerkki 218 Olkoon $E = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ ja $\hat{E} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ kaksi lineaariavaruuden kantaa siten, että

$$\begin{cases} \mathbf{e}_1 = \hat{\mathbf{e}}_1 + 2\hat{\mathbf{e}}_2 - 3\hat{\mathbf{e}}_3 \\ \mathbf{e}_2 = \hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_2 \\ \mathbf{e}_3 = \hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3 \end{cases} \quad \text{eli}$$

$$\Leftrightarrow \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -3 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \\ \hat{\mathbf{e}}_3 \end{pmatrix}$$

a) Laskemme vektorin $\mathbf{u} = 2\hat{\mathbf{e}}_1 + 3\hat{\mathbf{e}}_2 - \hat{\mathbf{e}}_3$ koordinaatit hatuttomassa kannassa.

b) Laskemme vektorin $\mathbf{v} = 4\mathbf{e}_1 - 2\mathbf{e}_2 + \mathbf{e}_3$ koordinaatit hatullisessa kannassa.

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

```
A =
```

```
    1    1    0
    2    1    1
   -3    0    1
```

```
>> uhat = [2; 3; -1]
```

```
uhat =
```

```
2
3
-1
```

```
>> v = [4; -2; 1]
```

```
v =
```

```
4
-2
1
```

```
>> u = inv(A)*uhat
```

```
u =
```

```
0.5000
1.5000
0.5000
```

```
>> vhat = A*v
```

```
vhat =
```

```
2
7
-11
```

Siiis

$$\mathbf{u} = 2\hat{\mathbf{e}}_1 + 3\hat{\mathbf{e}}_2 - \hat{\mathbf{e}}_3 = 0.5\mathbf{e}_1 + 1.5\mathbf{e}_2 + 0.5\mathbf{e}_3$$

$$\mathbf{v} = 4\mathbf{e}_1 - 2\mathbf{e}_2 + \mathbf{e}_3 = 2\hat{\mathbf{e}}_1 + 7\hat{\mathbf{e}}_2 - 11\hat{\mathbf{e}}_3$$

Lause 219 Kahden ortonormitetun kannan välinen kannanvaihto-matriisi on ortogonaalinen.

Todistus. Jos

$$\mathbf{e}_j = a_{1j}\hat{\mathbf{e}}_1 + a_{2j}\hat{\mathbf{e}}_2 + \cdots + a_{nj}\hat{\mathbf{e}}_n, \quad \forall j,$$

niin ortonormitusten perusteella

$$\delta_{ij} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = \vec{a}_i \cdot \vec{a}_j$$

□

Seuraavaksi mietimme, miten kannan vaihto vaikuttaa lineaarikuvauksen matriisiin. Olkoon $f : L \rightarrow L$ lineaarikuvaus, jossa lähtö- ja maaliavaruuksien ollessa sama lineaariavaruus L . Olkoon kannassa E kuvauksen matriisi $\mathbf{M}$. Jos $f(\mathbf{x}) = \mathbf{y}$, niin

$$\vec{y} = \mathbf{M}\vec{x}.$$

Jos $\hat{E}$ on toinen kanta siten, että kantojen välinen kannanvaihto-matriisi on $\mathbf{A}^T$, niin $E = \hat{E}\mathbf{A}$ ja

$$\vec{y} = \mathbf{A}\mathbf{M}\mathbf{A}^{-1}\vec{\hat{x}}$$

9 OMINAISARVOTEHTÄVÄ

Ch:EigSystem

9.1 Kompleksinen lineaariavaruus

Sec:CMatrix

9.1.1 Kompleksiluvut

Pian tulemme tarvitsemaan kompleksisen lineaariavaruuden alkeita. Tätä varten käymme läpi kompleksilukujen alkeet. Asiasta on oma kurssi, joten nyt toteamme asioita vain sen verran, että saamme tässä kurssissa tarvittavat asiat esitettyä.

Kompleksilukujen lukupariesitys. Kompleksiluvut voidaan esittää lukupareina

$$\mathbb{C} = \{(x, y) \mid x, y \in \mathbb{R}\}$$

joiden joukossa on määritelty yhteen- ja kertolasku seuraavasti

$$(x_1, y_1) \hat{+} (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad (9.1)$$

$$(x_1, y_1) \hat{*} (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) \quad (9.2)$$

Kolmikko $(\mathbb{C}, \hat{+}, \hat{*})$ on kunta (ks. sivu 102). Kunnassa laskeminen on helppoa, kunhan kompleksilukujen esittäminen saadaan helpoksi seuraavalla merkinnnällä.

Kompleksilukujen esitys imaginääriyksikön avulla. Sovimme seuraavan lyhennysmerkinnän

$$\begin{aligned} (0, 1) &=_{def} i = \text{imaginääriyksikkö} \\ (1, 0) &=_{def} 1 \\ \longrightarrow (x, y) &= x + iy \end{aligned}$$

Imaginääriyksikön tärkein ominaisuus, on

$$i^2 = -1$$

Tämän jälkeen kompleksiluvuilla laskeminen sujuu automaattisesti oikein. Laskusäännöt (9.1) ja (9.2) saavat nyt esityksen

$$\begin{aligned} (x_1 + iy_1) + (x_2 + iy_2) &= x_1 + x_2 + i(y_1 + y_2) \\ (x_1 + iy_1)(x_2 + iy_2) &= x_1 x_2 + ix_1 y_2 + ix_2 y_1 + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1) \end{aligned}$$

Määritelmä 220 Kompleksiluvun $z = x + iy \in \mathbb{C}$ reaaliosa on

$$\operatorname{Re} z = x.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ imaginaariosa on

$$\operatorname{Im} z = y.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ liittoluku on

$$z^* = x - iy.$$

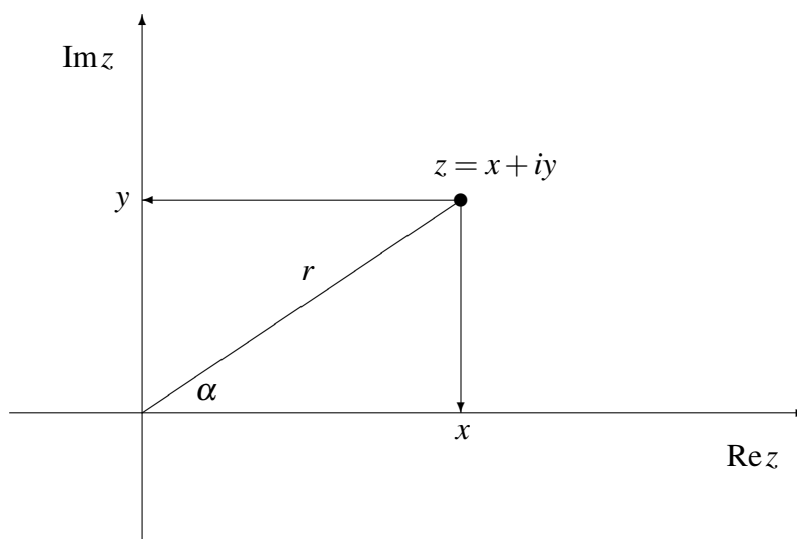
Kompleksiluvun $z = x + iy \in \mathbb{C}$ moduli on

$$|z| = \sqrt{x^2 + y^2} = \sqrt{zz^*}.$$

Kompleksiluvun $z = x + iy \in \mathbb{C}$ argumentti on

$$\arg z = \arctan\left(\frac{y}{x}\right).$$

Kompleksitaso. Kompleksiluku $z = x + iy$ "sijoitetaan" xy -tasoon pisteeseen (x, y) . Jos nyt $r = |z|$ ja $\alpha = \arg z$, niin suoraan kuvasta nähdään, että r on pisteen etäisyys origosta ja α on pisteen paikkavektorin ja x -akselin välinen kulma. Silloin



$$x = r \cos \alpha \quad \text{ja} \quad y = r \sin \alpha$$

9.1.2 Kompleksiset vektorit ja matriisit

Kompleksiset vektorit $\mathbb{C}^n$ ja matriisit määritellään nyt kaavioiksi, joihin on sijoitettu kompleksilukuja. Suora lasku osoittaa, että viisikko $(\mathbb{C}^n, \mathbb{C}, +, *, \circ)$ on kompleksiker-toiminen lineaariavaruus.

Muutama asia pitää määritellä uudelleen.

Määritelmä 221 (1) Matriisin $\mathbf{A} = (a_{ij})$ kompleksikonjugaatti on matriisi

$$\mathbf{A}^* = (a_{ij}^*).$$

(Jokainen kaavion luku siis konjugoidaan.)

(2) $m \times n$ -matriisin $\mathbf{A}$ hermitoitu matriisi on

$$\mathbf{A}^H = (\mathbf{A}^*)^T$$

(Matriisi siis konjugoidaan ja transponoidaan.)

(3) Kahden kompleksisen vektorin $\vec{z} \in \mathbb{C}^n$ ja $\vec{w} \in \mathbb{C}^n$ sisätulo määritellään kaavalla

$$\langle \vec{z} | \vec{w} \rangle = z_1^* w_1 + z_2^* w_2 + \cdots + z_n^* w_n = \vec{z}^H \vec{w}$$

Tehdyt määritelmät eivät muuta aiemmin tehtyjä reaalisten vektoreiden ja matriisien vastaavia määritelmiä. Jos nimittäin kaaviossa on vain reaalilukuja, niin herminointi palautuu transponoinniksi.

Kompleksisten vektorien ja matriisien tapauksessa erään aiemmat kaavat saavat uuden asun. Seuraavat kaavat ovat suoria seuraksia tehdyistä määritelmistä, joten emme esitä niille erillisiä todistuksia.

$$(\mathbf{A}^H)^H = \mathbf{A} \quad (9.3)$$

$$(\mathbf{A} + \mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H \quad (9.4)$$

$$(\lambda \mathbf{A})^H = \lambda^* \mathbf{A}^H \quad (9.5)$$

$$(\mathbf{A}\mathbf{B})^H = \mathbf{B}^H \mathbf{A}^H \quad (9.6)$$

$$\langle \vec{z} | \vec{w} \rangle = \langle \vec{w} | \vec{z} \rangle^* \quad (9.7)$$

$$\langle \lambda \vec{z} | \vec{w} \rangle = \lambda^* \langle \vec{z} | \vec{w} \rangle \quad (9.8)$$

$$\langle \vec{z} | \lambda \vec{w} \rangle = \lambda \langle \vec{z} | \vec{w} \rangle \quad (9.9)$$

$$\langle \lambda \vec{z} | \vec{w} \rangle = \langle \vec{z} | \lambda^* \vec{w} \rangle \quad (9.10)$$

$$\langle \mathbf{A}\vec{z} | \vec{w} \rangle = \langle \vec{z} | \mathbf{A}^H \vec{w} \rangle \quad (9.11)$$

$$\|\vec{z}\| = \sqrt{\langle \vec{z} | \vec{z} \rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2} \in \mathbb{R} \quad (9.12)$$

Määritelmä 222 Neliömatriisi $\mathbf{A}$ on

$$\begin{aligned} \text{symmetrinen, jos} \quad & \mathbf{A} = \mathbf{A}^T, \\ \text{hermiittinen, jos} \quad & \mathbf{A} = \mathbf{A}^H, \\ \text{ortogonaalinen, jos} \quad & \mathbf{A}^{-1} = \mathbf{A}^T, \\ \text{unitaarinen, jos} \quad & \mathbf{A}^{-1} = \mathbf{A}^H. \end{aligned}$$

9.2 Matriisin ominaisarvot

Sec:MatEig

9.2.1 Perusmääritelmät

Tässä kappaleessa oletamme skalaarit, vektorit ja matriisit kompleksisiksi. Vaikka lähtökohta olisikin reaalinen matriisi, niin laskujen aikana saattaa tulla esiin kompleksilukuja ja me nyt hyväksymme sen. Tämä helpottaa asian tarkastelua suuresti.

Seuraavaksi tarkastelemme teknistä ongelmaa, joka on yksinkertaisissa tapauksissaan mahdollista ratkaista suoraan laskemalla. Ensin opimme ratkaisemaan ominaisarvotävän helpossa tapauksessa. Myöhemmin opimme lisää ratkaisun ominaisuuksista, ja niistä ehdoista joilla ratkaisu on olemassa.

Määritelmä 223 Olkoot $\mathbf{A}$ $n \times n$ -matriisi, $\vec{x} \in \mathbb{R}^n, \vec{x} \neq \vec{0}$ vektori ja $\lambda \in \mathbb{R}$ reaaliluku. Jos

$$\mathbf{A}\vec{x} = \lambda\vec{x}, \tag{9.13} \quad \text{eigveceq}$$

niin sanomme, että λ on matriisin $\mathbf{A}$ ominaisarvo (engl. *eigenvalue*) ja vektori $\vec{x}$ on ominaisarvoon λ liittyvä matriisin $\mathbf{A}$ ominaisvektori (engl. *eigenvector*).

Jos λ on matriisin $\mathbf{A}$ ominaisarvo, niin homogeenisella yhtälöryhmällä

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

on ei-triviaali ratkaisu. Tämä on mahdollista vain jos

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \tag{9.14}$$

$$\Leftrightarrow \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{9.15}$$

Ominaisarvotehtävän ratkaisu tehdään $(n + 1)$:ssä vaiheessa:

(1) Ensin ratkaistaan ominaisarvot karakteristisesta yhtälöstä

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0. \quad (9.16)$$

Karakteristinen yhtälö on λ :n suhteen astetta n oleva yhtälö, joten sillä on korkeintaan n juurta (ominaisarvoa).

(2) Jokaista karakteristisen yhtälön juurta (ominaisarvoa) λ kohti ratkaistaan vastaava ominaisvektori homogeenisesta yhtälöryhmästä

$$\mathbf{A}\vec{x} = \lambda\vec{x}. \quad (9.17)$$

c4s15es1

Esimerkki 224 Ratkaise matriisin

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

ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö:

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

$$\Leftrightarrow \begin{vmatrix} 2-\lambda & 2 & 0 \\ -1 & 3-\lambda & 2 \\ 0 & 2 & 2-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow (2-\lambda) \begin{vmatrix} 3-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} - 2 \begin{vmatrix} -1 & 2 \\ 0 & 2-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow (2-\lambda)[(3-\lambda)(2-\lambda) - 2 \cdot 2] - 2[-1 \cdot (2-\lambda) - 0 \cdot 2] = 0$$

$$\Leftrightarrow (2-\lambda)[\lambda^2 - 5\lambda + 2] + 2(2-\lambda) = 0$$

$$\Leftrightarrow (2-\lambda)(\lambda^2 - 5\lambda + 4) = 0$$

$$\Leftrightarrow (2-\lambda) = 0 \text{ tai } (\lambda^2 - 5\lambda + 4) = 0$$

$$\Leftrightarrow \lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 4$$

(2) Ominaisarvoa $\lambda_1 = 1$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned}
 & \mathbf{A}\vec{x} = \lambda_1 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 1 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + 2x_2 = x_1 \\ -x_1 + 3x_2 + 2x_3 = x_2 \\ 2x_2 + 2x_3 = x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} [x_1] + 2x_2 = 0 \\ -x_1 + 2x_2 + 2x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \\ \end{array} \\
 \Leftrightarrow & \begin{cases} (x_1) + 2x_2 = 0 \\ + [4x_2] + 2x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \begin{array}{l} \\ -1/2 \\ \leftarrow \end{array} \\
 \Leftrightarrow & \begin{cases} (x_1) + 2x_2 = 0 \\ + (4x_2) + 2x_3 = 0 \\ 0 = 0 \end{cases} \begin{array}{l} \rightarrow x_1 = -2x_2 \\ \rightarrow x_3 = -2x_2 \\ \end{array}
 \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} -2a \\ a \\ -2a \end{pmatrix}, a \in \mathbb{R}.$$

Yleensä vektori valitaan niin, että sen pituus on yksi. Siis $(-2a)^2 + a^2 + (-2a)^2 = 1$, josta ratkeaa $a = 1/3$. Siis lopullinen ominaisarvoon $\lambda_1 = 1$ liittyvä ominaisvektori on

$$\vec{x}_1 = \begin{pmatrix} -2/3 \\ 1/3 \\ -2/3 \end{pmatrix}.$$

(3) Ominaisarvoa $\lambda_2 = 2$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned}
 & \mathbf{A}\vec{x} = \lambda_2 \vec{x} \\
 \Leftrightarrow & \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow & \begin{cases} 2x_1 + 2x_2 = 2x_1 \\ -x_1 + 3x_2 + 2x_3 = 2x_2 \\ 2x_2 + 2x_3 = 2x_3 \end{cases} \\
 \Leftrightarrow & \begin{cases} 2x_2 = 0 \\ -x_1 + x_2 + 2x_3 = 0 \\ 2x_2 = 0 \end{cases} \begin{array}{l} \rightarrow x_2 = 0 \\ \rightarrow x_1 = 2x_3 \\ \end{array}
 \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} 2b \\ 0 \\ b \end{pmatrix}, b \in \mathbb{R}.$$

Vektori valitaan taas niin, että sen pituus on yksi. Siis $(2b)^2 + 0^2 + (b)^2 = 1$, josta ratkeaa $b = 1/\sqrt{5}$. Siis lopullinen ominaisarvoon $\lambda_2 = 2$ liittyvä ominaisvektori on

$$\vec{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}.$$

(4) Ominaisarvoa $\lambda_3 = 4$ vastaava ominaisvektori ratkaistaan yhtälöryhmästä

$$\begin{aligned} \mathbf{A}\vec{x} &= \lambda_3\vec{x} \\ \Leftrightarrow \begin{pmatrix} 2 & 2 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 4 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ \Leftrightarrow \begin{cases} 2x_1 + 2x_2 &= 4x_1 \\ -x_1 + 3x_2 + 2x_3 &= 4x_2 \\ 2x_2 + 2x_3 &= 4x_3 \end{cases} \\ \Leftrightarrow \begin{cases} [-2x_1] + 2x_2 &= 0 \\ -x_1 - x_2 + 2x_3 &= 0 \\ 2x_2 - 2x_3 &= 0 \end{cases} \begin{array}{l} -1/2 \\ \leftarrow \\ \end{array} \\ \Leftrightarrow \begin{cases} (-2x_1) + 2x_2 &= 0 \\ \quad [-2x_2] + 2x_3 &= 0 \\ \quad 2x_2 - 2x_3 &= 0 \end{cases} \begin{array}{l} \\ +1 \\ \leftarrow \end{array} \\ \Leftrightarrow \begin{cases} (-2x_1) + 2x_2 &= 0 \\ \quad (-2x_2) + 2x_3 &= 0 \\ \quad 0 &= 0 \end{cases} \begin{array}{l} \rightarrow x_1 = x_2 \\ \rightarrow x_3 = x_2 \\ \end{array} \end{aligned}$$

Siis

$$\vec{x} = \begin{pmatrix} c \\ c \\ c \end{pmatrix}, c \in \mathbb{R}.$$

Vektori valitaan taas niin, että sen pituus on yksi. Siis $c^2 + c^2 + c^2 = 1$, josta ratkeaa $c = 1/\sqrt{3}$. Siis lopullinen ominaisarvoon $\lambda_3 = 4$ liittyvä ominaisvektori on

$$\vec{x}_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

Vastauksena ongelmaan voimme siis luetella ominaisarvot ja niihin liittyvät ominaisvektorit

$$\lambda_1 = 1, \vec{x}_1 = \begin{pmatrix} -2/3 \\ 1/3 \\ -2/3 \end{pmatrix}, \text{ ja } \lambda_2 = 2, \vec{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}, \text{ ja } \lambda_3 = 4, \vec{x}_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$$

Saman matriisin ominaisarvot ja ominaisvektorit saadaan Matlabin komennolla

[V,D] = eigs(A) seuraavasti:

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

```
A =
```

```
    2    2    0
   -1    3    2
    0    2    2
```

```
>> [V,D] = eigs(A)
```

```
V =
```

```
    0.5774   -0.8944    0.6667
    0.5774   -0.0000   -0.3333
    0.5774   -0.4472    0.6667
```

```
D =
```

```
    4.0000         0         0
         0    2.0000         0
         0         0    1.0000
```

```
>>
```

*Siis matriisin **V** sarakkeet ovat ominaisvektoreita ja diagonaalimatriisin **D** diagonaalialkiot ovat vastaavat ominaisarvot.*

On melko tavallista, että karakteristisella yhtälöllä on kompleksisia juuria. Nämä juuret ovat myös ominaisarvoja ja niitä vastaavat ominaisvektorit ovat yleensä myös kompleksisia. Kompleksisen esimerkin laskeminen käsin on melko työläs tehtävä, joten tyydymme nyt tietokoneella laskettuun esimerkkiin.

c4s15es2

Esimerkki 225 >> A = [2 2 0 ; -1 3 2 ; 0 2 -1]

```
A =
```

```
    2    2    0
   -1    3    2
    0    2   -1
```

```
>> [V,D] = eigs(A)
```

```
V =
```

```
    0.8165         0.8165         0.1851
    0.3585 - 0.3711i    0.3585 + 0.3711i   -0.3476
    0.2178 - 0.1403i    0.2178 + 0.1403i    0.9192
```

```

D =
    2.8782 - 0.9089i    0    0
         0    2.8782 + 0.9089i    0
         0    0    -1.7563
>>

```

Edellisen esimerkin matriisi ei näyttänyt hankalalta, mutta silti ominaisarvot ja ominaisvektorit ovat kompleksiset. Tässä vaiheessa helposti tulee epätoivoinen tunne siitä, että asia on liian mutkikas.

Jotta epätoivo ei ahdistaisi toteamme kaksi tulosta, jotka perustelemme pian:

(1) Symmetrisen matriisin ($\mathbf{A}^T = \mathbf{A}$) ominaisarvot ovat reaaliset.

(2) Ns. normaalin matriisin ($\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T$) ominaisvektoreista voidaan muodostaa $\mathbb{R}^n$:n kanta.

Useissa sovelluksissa tarkasteltava matriisi on symmetrinen, ja tiedämme siis, että ominaisarvot eivät ole kompleksisia.

Kun karakteristisen yhtälön, $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$, kaikki termit viedään vasemmalle puolelle yhtälöä, niin sinne syntyy karakteristinen polynomi $P(\lambda)$. Karakteristinen polynomi voidaan aina myös saattaa tulomuotoon

$$\begin{aligned}
 P(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) \\
 &= (-1)^n \lambda^n + \alpha_{n-1} \lambda^{n-1} + \cdots + \alpha_0
 \end{aligned} \tag{9.18}$$

$$= (-1)^n (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_q)^{k_q} \tag{9.19}$$

Jos matriisin karakteristinen polynomi on muotoa ^{karpoltulo} (9.19), niin sanomme, että ominaisarvon λ_j algebraalinen kertaluku on k_j . Kompleksianalyysin kurssilla opimme, että kertalukujen summa on $k_1 + k_2 + \cdots + k_q = n$.

$\mathbb{C}^n$:n lineaarinen aliavaruus $\ker(\mathbf{A} - \lambda_j \mathbf{I})$ sisältää kaikki ominaisarvoon λ_j liittyvät ominaisvektorit ja nollavektorin (joka ei ole ominaisvektori). Tämän aliavaruuden dimensiota

$$\rho(\lambda_j) = \dim \ker(\mathbf{A} - \lambda_j \mathbf{I}) \geq 1$$

sanomme ominaisarvon λ_j geometriseksi kertaluvuksi. Geometrinen kertaluku kertoo, miten monta lineaarisesti riippumatonta ominaisvektoria voidaan liittää ominaisarvoon λ_j .

Normaalisti matriisin ominaisarvojen algebralliset- ja geometriset kertaluvut yhtyvät. Jos näin ei tapahdu, niin sanomme, että matriisi on defektiivinen. Defektiivisten matriisien tutkiminen ylittää tämän peruskurssin tavoitteet. Termi mainitaan tässä, jotta opiskelija osaa alkaa etsiä lisätietoa kirjallisuudesta, jos laskentaohjelma antaa termin osana virheilmoitusta.

c4s15es3

Esimerkki 226 Tutkitaan matriisin

$$\mathbf{M} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

Ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö saa muodon

$$\begin{aligned} \det(\mathbf{M} - \lambda \mathbf{I}) = 0 &\Leftrightarrow \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & 1 & 3-\lambda \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda) \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda)((3-\lambda)(3-\lambda) - 1) = 0 \\ &\Leftrightarrow (\lambda - 2)^2(\lambda - 4) = 0 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ algebrallinen kertaluku on 2, ja ominaisarvon $\lambda_2 = 4$ algebrallinen kertaluku on 1.

(2) Seuraavaksi määritämme ominaisarvoon $\lambda_1 = 2$ liittyvät ominaisvektorit.

$$\begin{aligned}
 \mathbf{M}\vec{x} &= \lambda_1 \vec{x} \\
 \Leftrightarrow \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow \begin{cases} 2x_1 & = 2x_1 \\ 3x_2 + x_3 & = 2x_2 \\ x_2 + 3x_3 & = 2x_3 \end{cases} \\
 \Leftrightarrow \begin{cases} 0 & = 0 & \rightarrow x_1 = \text{vapaa} \\ x_2 + x_3 & = 0 & \rightarrow x_2 = -x_3 \\ x_2 + x_3 & = 0 \end{cases} \\
 \Leftrightarrow \begin{cases} x_1 & = x_1 \\ x_2 & = -x_3 \\ x_3 & = x_3 \end{cases} \\
 \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}
 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ geometrinen kertaluku on 2 ja ominaisvektorien virittäjiksi voidaan valita

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{ja} \quad \vec{x}_2 = \begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

(3) Seuraavaksi määritämme ominaisarvoon $\lambda_2 = 4$ liittyvät ominaisvektorit.

$$\begin{aligned}
 \mathbf{M}\vec{x} &= \lambda_2 \vec{x} \\
 \Leftrightarrow \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 4 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow \begin{cases} 2x_1 & = 4x_1 \\ 3x_2 + x_3 & = 4x_2 \\ x_2 + 3x_3 & = 4x_3 \end{cases} \\
 \Leftrightarrow \begin{cases} -2x_1 & = 0 & \rightarrow x_1 = 0 \\ -x_2 + x_3 & = 0 & \rightarrow x_2 = x_3 \\ x_2 + -x_3 & = 0 \end{cases}
 \end{aligned}$$

Ominaisvektoriksi voidaan valita

$$\vec{x}_3 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Sama matlabilla:

```
>> M = [2 0 0 ; 0 3 1 ; 0 1 3]
```

```
M =
```

```
      2      0      0
      0      3      1
      0      1      3
```

```
>> [V,D] = eigs(M)
```

```
V =
```

```
      0      0      1.0000
  0.7071  -0.7071      0
  0.7071   0.7071      0
```

```
D =
```

```
      4      0      0
      0      2      0
      0      0      2
```

```
>>
```

c4s15es4

Esimerkki 227 Tutkitaan matriisin

$$\mathbf{Q} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

Ominaisarvot ja ominaisvektorit.

(1) Karakteristinen yhtälö saa muodon

$$\begin{aligned} \det(\mathbf{Q} - \lambda \mathbf{I}) = 0 &\Leftrightarrow \begin{vmatrix} 2-\lambda & 1 & 0 \\ 0 & 1-\lambda & -1 \\ 0 & 2 & 4-\lambda \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda) \begin{vmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{vmatrix} - 1 \cdot \begin{vmatrix} 0 & 1-\lambda \\ 0 & 2 \end{vmatrix} = 0 \\ &\Leftrightarrow (2-\lambda)((1-\lambda)(4-\lambda) + 2) = 0 \\ &\Leftrightarrow (\lambda-2)^2(\lambda-3) = 0 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ algebrallinen kertaluku on 2, ja ominaisarvon $\lambda_2 = 3$ geometrinen kertaluku on 1.

(2) Seuraavaksi määritämme ominaisarvoon $\lambda_1 = 2$ liittyvät ominaisvektorit.

$$\begin{aligned}
 Q\vec{x} &= \lambda_1 \vec{x} \\
 \Leftrightarrow \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 2 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow \begin{cases} 2x_1 + x_2 &= 2x_1 \\ x_2 - x_3 &= 2x_2 \\ 2x_2 + 4x_3 &= 2x_3 \end{cases} \\
 \Leftrightarrow \begin{cases} x_2 &= 0 \\ -x_2 - x_3 &= 0 \\ 2x_2 + 2x_3 &= 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \end{array} \\
 \Leftrightarrow \begin{cases} x_2 &= 0 \\ -x_2 - x_3 &= 0 \\ 2x_2 + 2x_3 &= 0 \end{cases} \begin{array}{l} \rightarrow x_1 = \text{vapaa} \\ \rightarrow x_2 = x_3 = 0 \end{array} \\
 \Leftrightarrow \vec{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
 \end{aligned}$$

Siis ominaisarvon $\lambda_1 = 2$ geometrinen kertaluku on 1.

(3) Seuraavaksi määritämme ominaisarvoon $\lambda_2 = 3$ liittyvät ominaisvektorit.

$$\begin{aligned}
 Q\vec{x} &= \lambda_2 \vec{x} \\
 \Leftrightarrow \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= 3 \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
 \Leftrightarrow \begin{cases} 2x_1 + x_2 &= 3x_1 \\ x_2 - x_3 &= 3x_2 \\ 2x_2 + 4x_3 &= 3x_3 \end{cases} \\
 \Leftrightarrow \begin{cases} -x_1 + x_2 &= 0 \\ -2x_2 - x_3 &= 0 \\ 2x_2 + x_3 &= 0 \end{cases} \begin{array}{l} +1 \\ \leftarrow \end{array} \\
 \Leftrightarrow \begin{cases} -x_1 + x_2 &= 0 \\ -2x_2 - x_3 &= 0 \\ 0 &= 0 \end{cases} \begin{array}{l} \rightarrow x_1 = x_2 \\ \rightarrow x_3 = -2x_2 \end{array}
 \end{aligned}$$

Ominaisvektoriksi voidaan valita

$$\vec{x}_2 = \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}$$

Koska ensimmäisen ominaisarvon algebrallinen kertaluku on 2 ja geometrinen kertaluku on 1, on matriisi defektiivinen.

Sama matlabilla:

```
>> Q = [2 1 0 ; 0 1 -1 ; 0 2 4]
```

```
Q =
```

```
    2    1    0
    0    1   -1
    0    2    4
```

```
>> [V,D] = eigs(Q)
```

```
V =
```

```
    0.4082    1.0000    1.0000
    0.4082   -0.0000         0
   -0.8165    0.0000         0
```

```
D =
```

```
    3    0    0
    0    2    0
    0    0    2
```

```
>>
```

Defektiivisyys paljastuu nyt siitä, että tuloksena saadut ominisvektorit eivät muodosta vapaata systeemiä, sillä

```
>> det(V'*V)
```

```
ans =
```

```
    4.6259e-017
```

```
>>
```

Ominaisarvoihin ja ominaisvektoreihin palaamme takaisin, mutta ensin toteamme joi-takin uusia asioita.

9.2.2 Geometrinen tulkinta helpossa tapauksessa

Edellä on annettu vain tekninen määritelmä, joka saattaa joskus johtaa kompleksisiin ominaisarvoihin ja ominaisvektoreihin. Jotta asiasta saisi nyt konkreettisen hyödyllisen mielikuvan, oletamme tässä alakappaleessa, että $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa $\lambda_1, \lambda_2, \dots, \lambda_n$. Tämä yksinkertaistaa kovasti tarkastelua ja antaa mahdollisuuden nähdä, mitä merkitystä ominaisarvoilla on.

Lause 228 Olkoon $\mathbf{A}$ $n \times n$ -matriisi, jonka ominaisarvot ovat reaaliset. Matriisin $\mathbf{A}$ ominaisvektorit muodostavat $\mathbb{R}^n$:n kannan, jos ja vain jos $\mathbf{A}$ ei ole defektiivinen.

Todistus. $\Rightarrow$ Jos $\mathbf{A}$ on defektiivinen, niin lineaarisesti riippumattomien ominaisvektorien lukumäärä on pienempi kuin n , joten ominaisvektorit eivät voi virittää $\mathbb{R}^n$:ää.

$\Leftarrow$ Oletetaan seuraavaksi, että $\mathbf{A}$ ei ole defektiivinen. Luetellaan matriisin ominaisarvot suuruusjärjestyksessä $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ siten, että jokainen ominaisarvo toistuu luettelossa kertalukunsa mukaisesti. Olkoot $\{\vec{x}_1, \dots, \vec{x}_n\}$, vastaavat ominaisvektorit. Koska $\mathbf{A}$ ei ole defektiivinen, voidaan samaan ominaisarvoon liittyvät ominaisvektorit valita vapaiksi.

Ominaisarvoon λ_1 liittyvä ominaisvektori $\{\vec{x}_1\}$ muodostaa vapaan systeemin. Oletamme seuraavaksi, että ominaisarvoihin $(\lambda_1, \lambda_2, \dots, \lambda_{p-1})$ liittyvät ominaisvektorit $\{\vec{x}_1, \dots, \vec{x}_{p-1}\}$ muodostavat vapaan systeemin. Osoitamme seuraavaksi, että ominaisarvoihin $(\lambda_1, \lambda_2, \dots, \lambda_p)$ liittyvät ominaisvektorit $\{\vec{x}_1, \dots, \vec{x}_{p-1}, \vec{x}_p\}$ muodostavat vapaan systeemin. oletamme sitä varten, että

$$\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_p \vec{x}_p = \vec{0}. \quad (9.20) \quad \text{eigbase0}$$

Oletamme, että k viimeistä ominaisvektoria liittyvät samaan ominaisarvoon ($\lambda_{p-k+1} = \lambda_{p-k+2} = \dots = \lambda_p$). Nyt

$$\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_{p-k} \vec{x}_{p-k} = -\mu_{p-k+1} \vec{x}_{p-k+1} - \dots - \mu_p \vec{x}_p = \vec{a}. \quad (9.21) \quad \text{eigbase1}$$

missä vektori $\vec{a}$ on ominaisarvoon λ_p liittyvä ominaisvektori. Kertomalla yhtälö (9.21) puolittain vasemmalta $\mathbf{A}$:lla saadaan

$$\mu_1 \lambda_1 \vec{x}_1 + \mu_2 \lambda_2 \vec{x}_2 + \dots + \mu_{p-k} \lambda_{p-k} \vec{x}_{p-k} = \lambda_p \vec{a}, \quad (9.22) \quad \text{eigbase2}$$

Yhdistämällä (9.21) ja (9.22) saadaan

$$\mu_1 \left(1 - \frac{\lambda_1}{\lambda_p}\right) \vec{x}_1 + \mu_2 \left(1 - \frac{\lambda_2}{\lambda_p}\right) \vec{x}_2 + \dots + \mu_{p-k} \left(1 - \frac{\lambda_{p-k}}{\lambda_p}\right) \vec{x}_{p-k} = \vec{0}.$$

Tämä on mahdollista vain, jos $\mu_1 = \mu_2 = \dots = \mu_{p-k} = 0$. Siis $\vec{a} = \vec{0}$, mikä edelleen on mahdollista vain, jos $\mu_{p-k+1} = \mu_{p-k+2} = \dots = \mu_p = 0$. Olemme siis osoittaneet, että

$$(\mu_1 \vec{x}_1 + \mu_2 \vec{x}_2 + \dots + \mu_p \vec{x}_p = \vec{0}) \Rightarrow (\mu_1 = \mu_2 = \dots = \mu_p = 0).$$

Induktioaskel on valmis, ja lauseen todistus saadaan induktiolla p :n suhteen. $\square$

Lause 229 Jos $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa $\lambda_1, \lambda_2, \dots, \lambda_n$, niin vastaavista ominaisvektoreista muodostettu jono $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ on vapaa ja siis $\mathbb{R}^n$:n kanta.

Todistus. Suora seuraus edellisestä. □

Tarkastellaan kuvausta $f: \mathbb{R}^n \rightarrow \mathbb{R}^n, \vec{u} \mapsto \mathbf{A}\vec{u}$. Kun käytämme $\mathbb{R}^n$:n kantana ominaisvektoreista muodostettua kantaa, niin jokainen vektori voidaan lausua kannassa

$$\vec{u} = u_1 \vec{x}_1 + u_2 \vec{x}_2 + \dots + u_n \vec{x}_n$$

ja vektorin kuva on vastaavasti

$$\begin{aligned} f(\vec{u}) &= \mathbf{A}\vec{u} = \mathbf{A}(u_1 \vec{x}_1 + u_2 \vec{x}_2 + \dots + u_n \vec{x}_n) \\ &= u_1 \lambda_1 \vec{x}_1 + u_2 \lambda_2 \vec{x}_2 + \dots + u_n \lambda_n \vec{x}_n \end{aligned}$$

joten kuvauksen matriisi on

$$\mathbf{M}_f = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Geometrinen tulkinta on nyt siis se, että kun vektori lausutaan kantavektoreiden lineaarikombinaationa, se jaetaan kantavektoreiden suuntaisiin komponentteihin. Kerrotaessa matriisilla $\mathbf{A}$ jokainen komponentti tulee kerrottua vastaavalla ominaisarvolla. Jos $|\lambda_j| > 1$, niin komponentti venyy, jos $|\lambda_j| < 1$, niin komponentti kutistuu.

9.2.3 Ominaisarvohajoitelma

Merkitsemme $\mathbf{X}$:llä matriisia, jonka sarakkeina ovat matriisin $\mathbf{A}$ ominaisvektorit ja $\mathbf{D}$:llä disgonaalimatriisia, jonka diagonaalilla ovat matriisin $\mathbf{A}$ ominaisarvot. Jos nyt matriisin $\mathbf{A}$ ominaisvektorit muodostavat ortonormitetun systeemin, niin

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^T. \tag{9.23} \quad \text{eigsys haj}$$

Kaavan ^{eigsvshaj}(9.23) perusteluksi riittää osoittaa, että $\mathbf{XDX}^T \vec{x}_j = \lambda_j \vec{x}_j = \mathbf{A} \vec{x}_j$ kaikilla $j = 1, 2, \dots, n$. Tämä taas on suora lasku, sillä

$$\begin{aligned}\mathbf{X}^T \vec{x}_j &= \vec{e}_j \\ \Rightarrow \mathbf{DX}^T \vec{x}_j &= \mathbf{D} \vec{e}_j = \lambda_j \vec{e}_j \\ \Rightarrow \mathbf{XDX}^T \vec{x}_j &= \lambda_j \mathbf{X} \vec{e}_j = \lambda_j \vec{x}_j\end{aligned}$$

Ominaisarvohajoittelman avulla on helppo perustella joitakin tuloksia.

Lause 230 Jos neliömatriiseilla $\mathbf{A}$ ja $\mathbf{B}$ ovat samat ominaisvektorit ja ominaisvektorit muodostavat ortonormitetun jonon, niin

$$\mathbf{AB} = \mathbf{BA}$$

eli matriisit kommutoivat.

Todistus. Merkitään $\mathbf{A}$:n ominaisarvoista muodostettua diagonaalimatriisia $\mathbf{D}_A$:lla ja vastaavasti $\mathbf{B}$:n ominaisarvoista muodostettua diagonaalimatriisia $\mathbf{D}_B$:lla. Diagonaalimatriisit kommutoivat, joten

$$\begin{aligned}\mathbf{AB} &= \mathbf{XD}_A \mathbf{X}^T \mathbf{XD}_B \mathbf{X}^T = \mathbf{XD}_A \mathbf{D}_B \mathbf{X}^T = \mathbf{XD}_B \mathbf{D}_A \mathbf{X}^T \\ &= \mathbf{XD}_B \mathbf{X}^T \mathbf{XD}_A \mathbf{X}^T = \mathbf{BA}\end{aligned}$$

□

Matriisin potenssit on nyt helppo laskea käyttäen hajoitelmaa ^{eigsvshaj}(9.23):

$$\begin{aligned}\mathbf{A}^2 &= \mathbf{AA} = \mathbf{XDX}^T \mathbf{XDX}^T \\ &= \mathbf{XD}^2 \mathbf{X}^T = \mathbf{X} \begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix} \mathbf{X}^T \\ \mathbf{A}^n &= \mathbf{XD}^n \mathbf{X}^T = \mathbf{X} \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^n \end{pmatrix} \mathbf{X}^T\end{aligned}$$

Lause 231 (Cayley-Hamilton) Olkoon neliömatriisilla $\mathbf{A}$ ominaisarvohajoitelma $\mathbf{A} = \mathbf{XDX}^T$, missä ominaisvektoreista muodostettu matriisi on ortogonaalinen. Matriisi $\mathbf{A}$ toteuttaa oman karakteristisen yhtälönsä.

Todistus.

$$\begin{aligned} P(\mathbf{A}) &= P(\mathbf{XDX}^T) = \mathbf{X}P(\mathbf{D})\mathbf{X}^T \\ &= \mathbf{X} \begin{pmatrix} P(\lambda_1) & 0 & \cdots & 0 \\ 0 & P(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P(\lambda_n) \end{pmatrix} \mathbf{X}^T = \mathbf{0} \end{aligned}$$

□

9.3 Similaarisuus ja matriisin diagonalisointi

Sec:MatSim

Määritelmä 232 Neliömatriisit $\mathbf{A}$ ja $\mathbf{B}$ ovat similaariset, jos on olemassa säännöllinen matriisi $\mathbf{T}$ siten, että

$$\mathbf{A} = \mathbf{T}^{-1}\mathbf{B}\mathbf{T}.$$

Matriisia $\mathbf{T}$ sanomme similariteettimuunnosmatriisiksi ja muunnosta $\mathbf{B} \mapsto \mathbf{T}^{-1}\mathbf{B}\mathbf{T}$ sanomme similariteettimuunnokseksi.

Jos similariteettimuunnosmatriisi on unitaarinen, sanomme että matriisit ovat unitaarisimilaariset, eli

$$\mathbf{A} = \mathbf{T}^H\mathbf{B}\mathbf{T}$$

Lause 233 Jos matriisit $\mathbf{A}$ ja $\mathbf{B}$ ovat similaariset, $\mathbf{A} = \mathbf{T}^{-1}\mathbf{B}\mathbf{T}$, niin

- (1) matriiseilla on sama karakteristinen polynomi,
- (2) matriiseilla on samat ominaisarvot
- (3) jos $\vec{x}$ on $\mathbf{A}$:n ominaisvektori, niin $\mathbf{T}\vec{x}$ on vastaava $\mathbf{B}$:n ominaisvektori.

Todistus. (1)

$$\begin{aligned} P_A(\lambda) &= \det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{T}^{-1}(\mathbf{B} - \lambda\mathbf{I})\mathbf{T}) \\ &= \det(\mathbf{T}^{-1})\det(\mathbf{B} - \lambda\mathbf{I})\det(\mathbf{T}) \\ &= \det(\mathbf{B} - \lambda\mathbf{I}) = P_B(\lambda) \end{aligned}$$

(2) Seuraa suoraan edellisestä.

(3)

$$\mathbf{B}(\mathbf{T}\vec{x}) = \mathbf{T}(\mathbf{T}^{-1}\mathbf{B}\mathbf{T})\vec{x} = \mathbf{T}\mathbf{A}\vec{x} = \mathbf{T}\lambda\vec{x} = \lambda(\mathbf{T}\vec{x})$$

□

Määritelmä 234 $n \times n$ -matriisin $\mathbf{A}$ jälki (trace) on sen diagonaalialkioiden summa

$$\text{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}$$

Lause 235

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$$

Todistus. Suora lasku osoittaa, että

$$\begin{aligned} \text{tr}(\mathbf{AB}) &= (\mathbf{AB})_{11} + (\mathbf{AB})_{22} + \cdots + (\mathbf{AB})_{nn} \\ &= (a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}) \\ &\quad + (a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2}) \\ &\quad + \cdots + (a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn}) \\ &= (b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1n}a_{n1}) \\ &\quad + (b_{21}a_{12} + b_{22}a_{22} + \cdots + b_{2n}a_{n2}) \\ &\quad + \cdots + (b_{n1}a_{1n} + b_{n2}a_{2n} + \cdots + b_{nn}a_{nn}) \\ &= (\mathbf{BA})_{11} + (\mathbf{BA})_{22} + \cdots + (\mathbf{BA})_{nn} = \text{tr}(\mathbf{BA}) \end{aligned}$$

□

Lause 236 Similaarisilla matriiseilla on samat jäljet.

Todistus. Olkoon $\mathbf{A} = \mathbf{T}^{-1}\mathbf{BT}$. Silloin

$$\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{T}^{-1}\mathbf{BT}) = \text{tr}(\mathbf{BTT}^{-1}) = \text{tr}(\mathbf{B})$$

□

Jos herää kysymys siitä, ovatko kaksi annettua matriisia similaariset keskenään, niin ensin kannattaa laskea matriisien jäljet ja verrata niitä. Jos jäljet eivät ole samat, niin matriisit eivät voi olla similaariset.

Merkitään jatkossa diagonaalista matriisia merkinnällä

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

Määritelmä 237 Neliömatriisi $\mathbf{A}$ on diagonalisoituva, jos se on similaarinen jonkin diagonaalimatriisiin $\mathbf{D}$ kanssa, eli

$$\mathbf{D} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}.$$

Jos kuvauksen $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ matriisi on diagonalisoituva, niin sopivalla kannan vaihdolla saadaan kuvauksen matriisi diagonaaliseksi. Tämä saattaa ratkaisevalla tavalla helpottaa kuvauksen tarkastelua. Seuraavassa esimerkki.

Esimerkki 238 Tarkastellaan differentiaaliyhtälöryhmää

$$\begin{cases} x_1'(t) = 3x_1(t) - 0.5x_2(t) - 0.5x_3(t) \\ x_2'(t) = 0.5x_2(t) + 1.5x_3(t) \\ x_3'(t) = 1.5x_2(t) + 0.5x_3(t) \end{cases}$$

Jätetään jatkossa monessa kohdassa ”(t)” pois. Tämä yksinkertaistaa lausekkeita ja saa ne helpommin luettaviksi. Nyt siis

$$\begin{pmatrix} dx_1/dt \\ dx_2/dt \\ dx_3/dt \end{pmatrix} = \begin{pmatrix} 3 & -0.5 & -0.5 \\ 0 & 0.5 & 1.5 \\ 0 & 1.5 & 0.5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (9.24) \quad \boxed{\text{dyryhma}}$$

Merkitsemme nyt

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \frac{d\vec{x}}{dt} = \begin{pmatrix} dx_1/dt \\ dx_2/dt \\ dx_3/dt \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 3 & -0.5 & -0.5 \\ 0 & 0.5 & 1.5 \\ 0 & 1.5 & 0.5 \end{pmatrix}$$

Voimme siis nyt kirjoittaa yhtälön ^{dyryhma}(9.24) muotoon

$$\frac{d\vec{x}}{dt} = \mathbf{A}\vec{x}$$

Kerroinmatriisi on diagonalisoituva, sillä

$$\mathbf{D} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T},$$

kun

$$\mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \mathbf{T}^{-1} = \begin{pmatrix} 1 & -0.5 & -0.5 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & -0.5 \end{pmatrix}$$

Nyt muodostamme uudet funktiot, jotka ovat lineaarikombinaatioita alkuperäisistä:

$$\begin{cases} w_1(t) = x_1(t) - 0.5x_2(t) - 0.5x_3(t) \\ w_2(t) = 0.5x_2(t) + 0.5x_3(t) \\ w_3(t) = 0.5x_2(t) - 0.5x_3(t) \end{cases} \\ \Leftrightarrow \vec{w} = \mathbf{T}^{-1}\vec{x}$$

Nyt suora lasku osoittaa, että w -funktiot toteuttavat hyvin yksinkertaiset differentiaaliyhtälöt

$$\begin{aligned} & \begin{cases} dw_1/dt = dx_1/dt - 0.5 dx_2/dt - 0.5 dx_3/dt \\ dw_2/dt = 0.5 dx_2/dt + 0.5 dx_3/dt \\ dw_3/dt = 0.5 dx_2/dt - 0.5 dx_3/dt \end{cases} \\ \Leftrightarrow & \begin{pmatrix} dw_1/dt \\ dw_2/dt \\ dw_3/dt \end{pmatrix} = \mathbf{T}^{-1} \frac{d\vec{x}}{dt} = \mathbf{T}^{-1} \mathbf{A} \vec{x} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T} \vec{w} = \mathbf{D} \vec{w} \\ \Leftrightarrow & \begin{pmatrix} dw_1/dt \\ dw_2/dt \\ dw_3/dt \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \\ \Leftrightarrow & \begin{cases} dw_1/dt = 3w_1(t) \\ dw_2/dt = 2w_2(t) \\ dw_3/dt = -w_3(t) \end{cases} \end{aligned}$$

Tämän ryhmän ratkaisufunktiot voidaan kirjoittaa suoraan

$$\vec{w} = \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{pmatrix} = \begin{pmatrix} C_1 e^{3t} \\ C_2 e^{2t} \\ C_3 e^{-t} \end{pmatrix},$$

ja lopulta

$$\vec{x} = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \mathbf{T} \vec{w} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} C_1 e^{3t} \\ C_2 e^{2t} \\ C_3 e^{-t} \end{pmatrix} = \begin{pmatrix} C_1 e^{3t} + C_2 e^{2t} \\ C_2 e^{2t} + C_3 e^{-t} \\ C_2 e^{2t} - C_3 e^{-t} \end{pmatrix}.$$

Lause 239 Olkoon $\mathbf{A}$ $n \times n$ -matriisi, jonka ominaisarvot ovat reaaliset. Matriisi $\mathbf{A}$ on diagonalisoituva, jos ja vain jos $\mathbf{A}$ ei ole defektiivinen.

Todistus. $\Rightarrow$ Olkoon

$$\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n) = \mathbf{T}^{-1} \mathbf{A} \mathbf{T}.$$

Nyt

$$\mathbf{T} \mathbf{D} = \mathbf{A} \mathbf{T}.$$

Edellisen yhtälön LHS:n k :s sarake on $\lambda_k \vec{t}_{\bullet,k}$ ja RHS:n k :s sarake on $\mathbf{A} \vec{t}_{\bullet,k}$. Siis $\mathbf{T}$:n sarakkeet ovat matriisin $\mathbf{A}$ ominaisvektorit. Koska $\mathbf{T}$ on säännöllinen sen rangi on n , joten $\mathbf{A}$:n ominaisvektorit muodostavat vapaan systeemin. Siis $\mathbf{A}$ ei ole defektiivinen.

$\Leftarrow$ Olkoon $\mathbf{A}$ ei-defektiivinen. Silloin lauseen (228) ^{|c4s15theigbase} mukaan matriisilla $\mathbf{A}$ on n ominaisvektoria, jotka muodostavat vapaan systeemin. Tehdään ominaisvektoreista matriisi $\mathbf{T}$ ja ominaisarvoista diagonaalimatriisi $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Silloin vastaavasti kuin edellä

$$\begin{aligned}\mathbf{T}\mathbf{D} &= \mathbf{A}\mathbf{T} \\ \Rightarrow \mathbf{D} &= \mathbf{T}^{-1}\mathbf{A}\mathbf{T}.\end{aligned}$$

□

Lause 240 Jos $n \times n$ -matriisilla $\mathbf{A}$ on n erisuurta reaalista ominaisarvoa, niin matriisi on diagonalisoituva

Todistus. Lauseen (229) ^{|c4s15threalbase} mukaan matriisi ei ole defektiivinen jolloin lauseen (239) ^{|c4s16thdiag1} mukaan se on diagonalisoituva. □

9.4 Schurin muoto ja normaalit matriisit

Sec:MatSchur

Tässä kappaleessa esitämme todistukset vain tapauksessa, jossa matriisin ominaisarvot ovat reaaliset. Todistukset on mahdollista yleistää kompleksisten ominaisarvojen tapaukseen (Ks. Kivelä: ”Matriisilasku ja Lineaarialgebra” sivut 176–178).

Lause 241 Olkoot $\lambda_1, \lambda_2, \dots, \lambda_n$ reaalisen $n \times n$ -matriisin $\mathbf{A}$ ominaisarvot. On olemassa unitaarimatriisi $\mathbf{M}$ siten, että

$$\mathbf{M}^H \mathbf{A} \mathbf{M} = \begin{pmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & \mathbf{C} \end{pmatrix},$$

^{thschur1} missä matriisin $\mathbf{C}$ ominaisarvot ovat $\lambda_2, \dots, \lambda_n$.

Todistus. Olkoon $\vec{x}$ ominaisarvoon λ_1 liittyvä normitettu ominaisvektori (ts. $\vec{x}^T \vec{x} = 1$). Oletamme nyt, että λ_1 on reaalinen. (Kompleksinen tapaus ks. Kivelä: ”Matriisilasku ja Lineaarialgebra” sivut 176–178.) Täydennetään yhden vektorin systeemi $\{\vec{x}\}$ ortonormeeratuksi kannaksi $\{\vec{x}, \vec{v}_1, \vec{v}_2, \dots, \vec{v}_{n-1}\}$. Muodostetaan näistä vektoreista (sarakeisiin sijoittamalla) matriisi

$$\mathbf{M} = \begin{pmatrix} \vec{x} & \mathbf{V} \end{pmatrix},$$

joka on sarakkeiden ortonormeerauksen perusteella unitaarinen. Edelleen ortonormeerauksen takia $\mathbf{V}^T \vec{x} = \vec{0}$. Siis

$$\begin{aligned} \mathbf{M}^H \mathbf{A} \mathbf{M} &= \begin{pmatrix} \vec{x}^T \\ \mathbf{V}^T \end{pmatrix} \mathbf{A} \begin{pmatrix} \vec{x} & \mathbf{V} \end{pmatrix} = \begin{pmatrix} \vec{x}^T \mathbf{A} \vec{x} & \vec{x}^T \mathbf{A} \mathbf{V} \\ \mathbf{V}^T \mathbf{A} \vec{x} & \mathbf{V}^T \mathbf{A} \mathbf{V} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & \vec{x}^T \mathbf{A} \mathbf{V} \\ \vec{0} & \mathbf{C} \end{pmatrix} = \mathbf{G} \end{aligned}$$

Koska matriisit $\mathbf{A}$ ja $\mathbf{G}$ ovat similaariset, niillä on samat karakteristiset polynomit. Siis

$$P_A(\lambda) = P_G(\lambda) = (\lambda_1 - \lambda)P_C(\lambda)$$

ja $\mathbf{C}$:n ominaisarvot ovat $\lambda_2, \dots, \lambda_n$. □

Lause 242 $n \times n$ -matriisi $\mathbf{A}$ on unitaarisimilaarinen yläkolmiomatriisin kanssa.

$$\mathbf{M}^H \mathbf{A} \mathbf{M} = \begin{pmatrix} \lambda_1 & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & \lambda_2 & u_{23} & \cdots & u_{2n} \\ 0 & 0 & \lambda_3 & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Yläkolmiomatriisin lävistäjäalkiot ovat $\mathbf{A}$:n ominaisarvot.

Todistus. Tehdään induktio n :n suhteen.

Kun $n = 1$, väite on tosi, koska 1×1 -matriisi on valmiiksi yläkolmiomatriisi.

Oletamme seuraavaksi, että väite on tosi, kun $n = k - 1$ ja osoitamme sen silloin olevan tosi myös, kun $n = n$. Lauseen (241) mukaan on olemassa unitaarinen $k \times k$ -matriisi $\mathbf{M}_1$ siten, että

$$\mathbf{M}_1^H \mathbf{A} \mathbf{M}_1 = \begin{pmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & \mathbf{C}_1 \end{pmatrix},$$

missä λ_1 on matriisin $\mathbf{A}$ ominaisarvo ja matriisin $\mathbf{C}_1$ ominaisarvot ovat $\lambda_2, \lambda_3, \dots, \lambda_k$.

Induktio-oletuksen mukaan on olemassa unitaarinen $(k - 1) \times (k - 1)$ -matriisi $\mathbf{M}_2$ siten, että

$$\mathbf{C} = \mathbf{M}_2^H \mathbf{C}_1 \mathbf{M}_2$$

on yläkolmiomatriisi, jonka lävistäjäalkiot ovat $\lambda_2, \lambda_3, \dots, \lambda_k$. Nyt matriisi

$$\mathbf{M} = \mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix}$$

on unitaarinen ja

$$\begin{aligned}
 \mathbf{M}^H \mathbf{A} \mathbf{M} &= \left(\mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \right)^H \mathbf{A} \mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix}^H \mathbf{M}_1^H \mathbf{A} \mathbf{M}_1 \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix}^H \begin{pmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & \mathbf{C}_1 \end{pmatrix} \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{M}_2^H \end{pmatrix} \begin{pmatrix} \lambda_1 & \vec{b}^T \mathbf{M}_2 \\ \vec{0} & \mathbf{C}_1 \mathbf{M}_2 \end{pmatrix} \\
 &= \begin{pmatrix} \lambda_1 & \vec{b}^T \mathbf{M}_2 \\ \vec{0} & \mathbf{C} \end{pmatrix}
 \end{aligned}$$

□

Esimerkki 243 Esimerkkien (224 sivu 193), (225 sivu 196), (226 sivu 198) ja (227 sivu 200) matriisien Schurin muodot ovat

```
>> help schur
```

```
SCHUR Schur decomposition.
```

```
[U,T] = SCHUR(X) produces a quasitriangular Schur matrix T and
a unitary matrix U so that X = U*T*U' and U'*U = EYE(SIZE(U)).
X must be square.
```

```
T = SCHUR(X) returns just the Schur matrix T.
```

```
If X is complex, the complex Schur form is returned in matrix T.
The complex Schur form is upper triangular with the eigenvalues
of X on the diagonal.
```

```
If X is real, two different decompositions are available.
```

```
SCHUR(X,'real') has the real eigenvalues on the diagonal and the
complex eigenvalues in 2-by-2 blocks on the diagonal.
```

```
SCHUR(X,'complex') is triangular and is complex if X has complex
eigenvalues. SCHUR(X,'real') is the default.
```

```
See RSF2CSF to convert from Real to Complex Schur form.
```

Esimerkki (224 sivu 193):

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

```
A =
```

```
    2    2    0
   -1    3    2
    0    2    2
```

```
>> [U,T] = schur(A)
```

```
U =
```

```
    0.5774   -0.7071   -0.4082
    0.5774    0.7071   -0.4082
    0.5774    0.0000    0.8165
```

```
T =
```

```
    4.0000    2.4495   -0.0000
         0    2.0000    1.7321
         0         0    1.0000
```

```
>>
```

Esimerkki (225 sivu 196):

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

```
A =
```

```
    2    2    0
   -1    3    2
    0    2   -1
```

```
>> [U,T] = schur(A)
```

```
U =
```

```
    0.8490   -0.5199   -0.0938
    0.4602    0.8150   -0.3522
    0.2595    0.2559    0.9312
```

```
T =
```

```
    2.8782    2.4575   -0.7676
   -0.3362    2.8782    0.7785
         0         0   -1.7563
```

```
>> [U,T] = schur(A,'complex')
```

```
U =
```

```
    0.8165 + 0.0046i    0.4821 + 0.3035i   -0.0938
    0.3564 + 0.3731i   -0.2006 - 0.7547i   -0.3522
```

```

0.2170 + 0.1415i -0.0273 - 0.2548i 0.9312

T =
2.8782 + 0.9089i -1.3569 - 1.6306i -0.7623 - 0.1001i
0 2.8782 - 0.9089i -0.5658 + 0.5329i
0 0 -1.7563

>>
\end{verbatim}

```

Esimerkki (\ref{c4s15es3} sivu \pageref{c4s15es3}):

```

\begin{verbatim}
>> M = [2 0 0 ; 0 3 1 ; 0 1 3]
M =
2      0      0
0      3      1
0      1      3

```

```

>> [U,T] = schur(M)
U =
1.0000      0      0
0 -0.7071 0.7071
0 0.7071 0.7071

```

```

T =
2      0      0
0      2      0
0      0      4

```

```

>>

```

Esimerkki (\ref{c4s15es4} sivu 200):

```

>> Q = [2 1 0 ; 0 1 -1 ; 0 2 4]
Q =
2      1      0
0      1     -1
0      2      4

```

```

>> [U,T] = schur(Q)
U =
1.0000      0      0

```

$$\begin{array}{ccc} 0 & -0.7071 & -0.7071 \\ 0 & 0.7071 & -0.7071 \end{array}$$

$$\mathbf{T} = \begin{array}{ccc} 2.0000 & -0.7071 & -0.7071 \\ 0 & 2.0000 & -3.0000 \\ 0 & 0 & 3.0000 \end{array}$$

>>

Määritelmä 244 Neliömatriisi $\mathbf{A}$ on normaali, jos

$$\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H.$$

Reaalisen matriisin tapauksessa normaalisuus merkitsee siis, että

$$\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T.$$

Erityisesti hermiittinen ($\mathbf{A}^H = \mathbf{A}$) ja unitaarinen ($\mathbf{A}^H = \mathbf{A}^{-1}$) sekä reaalisten matriisien tapauksessa symmetrinen ($\mathbf{A}^T = \mathbf{A}$) ja ortogonaalinen ($\mathbf{A}^T = \mathbf{A}^{-1}$) ovat normaalin matriisin erikoistapauksia.

Lause 245 Olkoon neliömatriisi $\mathbf{A}$ unitaarisimilaarinen matriisin $\mathbf{B}$ kanssa ($\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{B}$). Silloin $\mathbf{A}$ on normaali, jos ja vain jos $\mathbf{B}$ on normaali.

Todistus. $\Rightarrow$ Oletamme siis, että $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H$. Nyt

$$\begin{aligned} \mathbf{B}^H \mathbf{B} &= (\mathbf{U}^H \mathbf{A} \mathbf{U})^H (\mathbf{U}^H \mathbf{A} \mathbf{U}) = (\mathbf{U}^H \mathbf{A}^H \mathbf{U}) (\mathbf{U}^H \mathbf{A} \mathbf{U}) = \mathbf{U}^H \mathbf{A}^H \mathbf{A} \mathbf{U} \\ &= \mathbf{U}^H \mathbf{A} \mathbf{A}^H \mathbf{U} = (\mathbf{U}^H \mathbf{A} \mathbf{U}) (\mathbf{U}^H \mathbf{A}^H \mathbf{U}) = \mathbf{B} \mathbf{B}^H \end{aligned}$$

$\Leftarrow$ Toisinpäin $\mathbf{A} = \mathbf{U} \mathbf{B} \mathbf{U}^H$ ja oletamme, että $\mathbf{B}^H \mathbf{B} = \mathbf{B} \mathbf{B}^H$. Nyt

$$\begin{aligned} \mathbf{A}^H \mathbf{A} &= (\mathbf{U} \mathbf{B} \mathbf{U}^H)^H (\mathbf{U} \mathbf{B} \mathbf{U}^H) = (\mathbf{U} \mathbf{B}^H \mathbf{U}^H) (\mathbf{U} \mathbf{B} \mathbf{U}^H) = \mathbf{U} \mathbf{B}^H \mathbf{B} \mathbf{U}^H \\ &= \mathbf{U} \mathbf{B} \mathbf{B}^H \mathbf{U}^H = (\mathbf{U} \mathbf{B} \mathbf{U}^H) (\mathbf{U} \mathbf{B}^H \mathbf{U}^H) = \mathbf{A} \mathbf{A}^H \end{aligned}$$

□

Lause 246 Matriisi $\mathbf{A}$ on unitaarisimilaarinen diagonaalimatriisin kanssa, jos ja vain jos se on normaali.

Todistus. $\Rightarrow$ Olkoon $U^H A U = D$, missä U on unitaarinen ja D on diagonaalinen. Koska diagonaalimatriisi on normaali, on edellisen lauseen perusteella myös A normaali.

$\Leftarrow$ Seuraavaksi oletamme, että A on normaali. Koska A :lla on Schurin muoto, se on unitaarisimilaarinen yläkolmiomatriisin kanssa. Koska A on normaali, myös yläkolmiomatriisi on normaali. On helppo nähdä, että normaali yläkolmiomatriisi on välttämättä diagonaalinen. $\square$

Lause 247 Normaalin $n \times n$ -matriisin ominaisvektoreista voidaan muodostaa ortonormeerattu jono.

Todistus. Edellisen lauseen mukaan normaali matriisi A on unitaarisimilaarinen diagonaalimatriisin kanssa. Siis

$$U^H A U = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

missä U on unitaarinen ja D on diagonaalinen. Silloin

$$A U = U D$$

$$\Leftrightarrow A \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \vec{u}_1 & \lambda_2 \vec{u}_2 & \cdots & \lambda_n \vec{u}_n \end{pmatrix}.$$

Siis U :n sarakkeet ovat ominaisvektoreita, ja U :n unitaarisuudesta seuraa, että vektori-systeemi $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ on vapaa ja ortonormeerattu. $\square$

Lause 248 Hermiittisen matriisin ($A = A^H$) ja reaalisen symmetrisen matriisin ($A = A^T$) Schurin muoto on diagonaalinen ja ominaisarvot ovat reaaliset.

Todistus. Lauseen ^{lc4s17thnormaali} (246) mukaan matriisin Schurin muoto on diagonaalinen. Olkoon $A = A^H$ ja $U^H A U = D$, missä U on unitaarinen ja D on diagonaalinen Schurin muoto. Silloin

$$D^H = (U^H A U)^H = U^H A^H U = U^H A U = D.$$

Tästä seuraa, että D :n lävistäjäalkiot, eli A :n ominaisarvot ovat reaaliset. $\square$

YHTEENVETO: Symmetrisen $n \times n$ -matriisin ominaisarvotekijä ei johda ongelmiin, vaan ominaisarvot ovat reaaliset ja ominaisvektoreista voidaan muodostaa $\mathbb{R}^n$:n ortonormitettu kanta.

Lause 249 Olkoon $n \times n$ -neliömatriisin $\mathbf{A}$ ominaisarvot $\lambda_1, \lambda_2, \dots, \lambda_n$. Silloin

$$\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$$

ja

$$\operatorname{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Todistus. Matriisi on unitaarisimilaarinen Schurin muotonsa kanssa

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{S},$$

missä $\mathbf{U}$ on unitaarinen ja $\mathbf{S}$ on yläkolmiomatriisi, jonka diagonaali-alkiot ovat $\lambda_1, \lambda_2, \dots, \lambda_n$. Silloin koska similaarisilla matriiseilla on samat determinantit ja jäljet, niin

$$\begin{aligned} \det(\mathbf{A}) &= \det(\mathbf{S}) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n, & \text{ja} \\ \operatorname{tr}(\mathbf{A}) &= \operatorname{tr}(\mathbf{S}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n \end{aligned}$$

□

9.5 Matriisihajoitelmia

Sec:MatDecomp

Aiemmissa kappaleissa on jo esiintynyt kolme matriisi-hajoitelmaa. Seuraavassa on tiivistelmä näistä kolmesta.

QR-hajoitelma.

(1) Jos $\mathbf{A}$ on $m \times n$ -matriisi, jolle $\operatorname{rank}(\mathbf{A}) = n$ (eli $\mathbf{A}$ on ”a matrix of full rank”), niin matriisilla on QR-hajoitelma

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \tag{9.25}$$

missä $\mathbf{Q}$ on ortogonaalinen $m \times m$ matriisi, ja $\mathbf{R}$ on $m \times n$ yläkolmiomatriisi, jonka diagonaali-alkiot eroavat nolasta.

(2) Jos $\mathbf{A}$ on $m \times n$ -matriisi, jolle $\operatorname{rank}(\mathbf{A}) < n$, niin matriisilla on QR-hajoitelma

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \tag{9.26}$$

missä $\mathbf{Q}$ on ortogonaalinen $m \times m$ matriisi, ja $\mathbf{R}$ on $m \times n$ yläkolmiomatriisi, jonka diagonaalilla on nollia.

(3) Jos $\mathbf{A}$ on $m \times n$ -matriisi, jolle $\text{rank}(\mathbf{A}) = n$, niin matriisilla on QR-hajoitelma

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R}_1, \quad (9.27)$$

missä $\mathbf{Q}_1$ on $m \times n$ matriisi jonka sarakevektorit muodostavat ortonormitetun jonon (siis $\mathbf{Q}_1^T \mathbf{Q}_1 = \mathbf{I}$, mutta $\mathbf{Q}_1 \mathbf{Q}_1^T \neq \mathbf{I}$), ja $\mathbf{R}_1$ on $n \times n$ yläkolmiomatriisi, jonka diagonaalialkiot eroavat nolasta.

c4s18qresim1

Esimerkki 250

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

```
A =
```

```
    1    2    3
    0    1    1
    1    1    1
    3   -1    2
```

```
>> [Q,R] = qr(A)
```

```
Q =
```

```
-0.3015   -0.7559    0.4303   -0.3906
         0   -0.3780    0.1632    0.9113
-0.3015   -0.3780   -0.8753   -0.0000
-0.9045    0.3780    0.1484    0.1302
```

```
R =
```

```
-3.3166         0   -3.0151
         0   -2.6458   -2.2678
         0         0    0.8753
         0         0         0
```

```
>> [Q1,R1] = qr(A,0)
```

```
Q1 =
```

```
-0.3015   -0.7559    0.4303
         0   -0.3780    0.1632
-0.3015   -0.3780   -0.8753
-0.9045    0.3780    0.1484
```

```
R1 =
```

```
-3.3166         0   -3.0151
         0   -2.6458   -2.2678
         0         0    0.8753
```

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

```
B =
```

```
    1    2    3
    0    1    1
    1    1    2
```

```

      3      -1      2
>> [Q2,R2] = qr(B)
Q2 =
-0.3015 -0.7559 -0.4128 -0.4090
      0 -0.3780 -0.2028  0.9033
-0.3015 -0.3780  0.8745  0.0382
-0.9045  0.3780 -0.1539  0.1236

R2 =
-3.3166      0 -3.3166
      0 -2.6458 -2.6458
      0      0  0.0000
      0      0      0
>>

```

Ominaisarvohajoitelma.

Jos reaalinen $n \times n$ -neliömatriisi $\mathbf{A}$ on symmetrinen ($\mathbf{A}^T = \mathbf{A}$), niin ominaisarvot ovat reaaliset ja ominaisvektoreista voidaan muodostaa ortonormitettu kanta. Lisäksi matriisilla on ominaisarvohajoitelma

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^T, \quad (9.28)$$

missä diagonaalimatriisin $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ lävistäjällä ovat matriisin ominaisarvot ja matriisin $\mathbf{X}$ sarakkeissa ovat vastaavat normitetut ominaisvektorit.

Matriisin diagonalisointi.

Jos reaalinen $n \times n$ -neliömatriisi $\mathbf{A}$ ei ole defektiivinen (eli se on normaali $\mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T$, esimerkiksi symmetrinen tai ortogonaalinen), niin se on similaarinen diagonaalimatriisin kanssa

$$\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}, \quad (9.29)$$

missä similariteetti-muunnosmatriisi $\mathbf{X}$ voidaan muodostaa lineaarisesti riippumattomista ominaisvektoreista ja matriisin $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ lävistäjällä ovat vastaavat matriisin ominaisarvot. (Huomaa, että nyt $\mathbf{X}$:n ei tarvitse olla ortonormeerattu.)

Schurin-muoto.

Jokainen reaalinen $n \times n$ -matriisi $\mathbf{A}$ on unitaarisimilaarinen Schurin-muotonsa $\mathbf{S}$ kanssa

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{U}^H, \quad (9.30)$$

missä $\mathbf{U}$ on unitaarinen ($\mathbf{U}^H = \mathbf{U}^{-1}$) ja $\mathbf{S}$ on yläkolmiomatriisi, jonka diagonaalilla on matriisin $\mathbf{A}$ ominaisarvot.

9.5.1 Jordanin muoto

Jordanin lohkoksi sanotaan $r \times r$ matriisia

$$\mathbf{J}(\lambda, r) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

$n \times n$ Jordanin matriisi on lohkomatriisi, jonka lavistaja-lohkot ovat Jordanin lohkoja, ja muut lohkot ovat nolla-lohkoja.

$$\left(\begin{array}{c|c|c|c} \mathbf{J}(\lambda_1, r_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{J}(\lambda_2, r_2) & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}(\lambda_q, r_q) \end{array} \right)$$

Esimerkki 251

$$\begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & 0 & & & \\ & & & 3 & 0 & & \\ & & & & 3 & 0 & \\ & & & & & 5 & 1 \\ & & & & & & 5 \end{pmatrix} = \begin{pmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & & & & \\ \hline & & & 3 & & & \\ \hline & & & & 3 & & \\ \hline & & & & & 5 & 1 \\ & & & & & & 5 \end{pmatrix}$$

Matriisin Jordanin muoto.

Jokainen reaalinen $n \times n$ -neliömatriisi $\mathbf{A}$ on similaarinen Jordanin-matriisin kanssa

$$\mathbf{A} = \mathbf{X}\mathbf{J}\mathbf{X}^{-1}, \quad (9.31) \quad \boxed{\text{Jordan}}$$

missä similariteetti-muunnosmatriisin $\mathbf{X}$ sarakkeita sanotaan yleistetyiksi ominaisvektoreiksi ja $\mathbf{J} = \text{diag}(\mathbf{J}(\lambda_1, r_1), \mathbf{J}(\lambda_2, r_2), \dots, \mathbf{J}(\lambda_q, r_q))$, $r_1 + r_2 + \cdots + r_q = n$.

Perustelu Jordanin muodon olemassaololle on melko mutkikas ja sivuutamme sen tässä. (Ks. Simo Kivel: ”Matriisilasku ja Lineaarialgebra”, sivut 183–190.)

Yhtälö (9.31) voidaan myös kirjoittaa muotoon

$$\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{J},$$

mistä seuraa, että yleistetty ominaisvektori toteuttaa toisen seuraavista yhtälöistä

$$(1) \quad \mathbf{A}\vec{x}_k = \lambda_k \vec{x}_k, \quad \text{tai}$$

$$(2) \quad \mathbf{A}\vec{x}_k = \lambda_k \vec{x}_k + \vec{x}_{k-1}.$$

9.6 Matriisin singulaariarvohajoitelma

Sec:MatSVD

Kun halutaan saada mielikuva lineaarikuvauksesta

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^n, \mathbf{x} \mapsto \mathbf{A}\mathbf{x}, \quad (9.32)$$

niin eräs hyvä tapa on laskea kuvavektorin pituus, kun $\|\mathbf{x}\| = 1$. Jos λ on matriisin ominaisarvo ja $\mathbf{x}$ on sitä vastaava normeerattu ominaisvektori, niin

$$\|\mathbf{A}\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda| \quad (9.33)$$

Jos ominaisvektorit virittävät $\mathbb{R}^n$:n, niin saamme varsin hyvän käsityksen kuvauksesta. Valitettavasti kaikkien matriisien ominaisvektorit eivät ole reaalisia. Onneksi tähän löytyy seuraava ratkaisu.

Olkoon $\mathbf{A}$ mikä tahansa (!) $(m \times n)$ -matriisi. Tutkitaan kuvavektorin normin sijasta normin neliötä

$$\|\mathbf{A}\mathbf{x}\|^2 = (\mathbf{A}\mathbf{x})^T \mathbf{A}\mathbf{x} = \mathbf{x}^T \mathbf{A}^T \mathbf{A}\mathbf{x} \quad (9.34)$$

Matriisi $\mathbf{A}^T \mathbf{A}$ on symmetrinen, sen ominaisarvot ovat reaaliset. Jos λ_j on matriisin $\mathbf{A}^T \mathbf{A}$ ominaisarvo ja $\mathbf{v}_j$ on sitä vastaava normeerattu ominaisvektori, niin

$$\|\mathbf{A}\mathbf{v}_j\|^2 = \mathbf{v}_j^T \mathbf{A}^T \mathbf{A}\mathbf{v}_j = \mathbf{v}_j^T \lambda_j \mathbf{v}_j = \lambda_j \quad (9.35)$$

Matriisin $\mathbf{A}^T \mathbf{A}$ ominaisarvot ovat siis ei-negatiiviset ja ne yleensä järjestetään suuruusjärjestykseen $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0$. Voimme määritellä

Määritelmä 252 $(m \times n)$ -matriisin $\mathbf{A}$ singulaariarvot (*singular values*) ovat

$$\sigma_j = \sqrt{\lambda_j}, j = 1, \dots, n$$

missä $\lambda_1, \dots, \lambda_n$ ovat matriisin $\mathbf{A}^T \mathbf{A}$ ominaisarvot.

Siis $\|\mathbf{A}\mathbf{v}_j\| = \sigma_j$.

Lause 253 Jos $(m \times n)$ -matriisin $\mathbf{A}$ säännöllisyysaste on $\text{Rank}(\mathbf{A}) = r$, niin on olemassa hajoitelma

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T,$$

missä $\mathbf{U}$ on ortogonaalinen $(m \times m)$ -matriisi ja $\mathbf{V}$ on ortogonaalinen $(n \times n)$ -matriisi ja

$$\mathbf{D} = \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$$

Todistus. Matriisi $\mathbf{A}^H \mathbf{A}$ on hermiittinen $n \times n$ -matriisi ja sen ominaisarvot ovat ≥ 0 . Luetellaan ominaisarvot $\lambda_j = \sigma_j^2$ suuruusjärjestyksessä

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0.$$

(Huomaa, että r ensimmäistä ominaisarvoa ovat positiivisia ja loput nollia.) Vastaavat ortonormitetut ominaisvektorit ovat $\vec{v}_1, \dots, \vec{v}_n$. Merkitään

$$\mathbf{V} = (\mathbf{V}_1 \mid \mathbf{V}_2) = (\vec{v}_1 \quad \dots \quad \vec{v}_r \mid \vec{v}_{r+1} \quad \dots \quad \vec{v}_n)$$

Matriisin $\mathbf{A}^H \mathbf{A}$ ominaisarvohajoitelma on

$$\begin{aligned} \mathbf{A}^H \mathbf{A} &= \mathbf{V} \begin{pmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^H, \quad \mathbf{\Sigma}^2 = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2) \\ \Rightarrow \begin{pmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} &= \begin{pmatrix} \mathbf{V}_1^H \\ \mathbf{V}_2^H \end{pmatrix} \mathbf{A}^H \mathbf{A} (\mathbf{V}_1 \mid \mathbf{V}_2) \\ &\Rightarrow \begin{cases} \mathbf{V}_1^H \mathbf{A}^H \mathbf{A} \mathbf{V}_1 = \mathbf{\Sigma}^2 \\ \mathbf{A} \mathbf{V}_2 = \mathbf{0} \end{cases} \end{aligned}$$

Asetetaan

$$\mathbf{U}_1 = \mathbf{A} \mathbf{V}_1 \mathbf{\Sigma}^{-1},$$

jolloin edeltävän mukaan $\mathbf{U}_1^H \mathbf{U}_1 = \mathbf{I}$, ts. $\mathbf{U}_1$:n sarakkeet muodostavat ortonormeeratun jonon. $\mathbf{U}_1$ voidaan täydentää unitaariseksi matriisiksi $\mathbf{U} = (\mathbf{U}_1 \mid \mathbf{U}_2)$. Unitarisuudesta seuraa, että $\mathbf{U}_2^H \mathbf{U}_1 = \mathbf{0}$. Nyt

$$\begin{aligned} \mathbf{U}^H \mathbf{A} \mathbf{V} &= \begin{pmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{pmatrix} \mathbf{A} (\mathbf{V}_1 \mid \mathbf{V}_2) = \begin{pmatrix} \mathbf{U}_1^H \mathbf{A} \mathbf{V}_1 & \mathbf{U}_1^H \mathbf{A} \mathbf{V}_2 \\ \mathbf{U}_2^H \mathbf{A} \mathbf{V}_1 & \mathbf{U}_2^H \mathbf{A} \mathbf{V}_2 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{\Sigma}^{-1} \mathbf{V}_1^H \mathbf{A}^H \mathbf{A} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{U}_2^H \mathbf{U}_1 \mathbf{\Sigma} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \end{aligned}$$

□

Singulaariarvohajoitelman suuri etu on siinä, että **reaalisen $m \times n$ -matriisin $\mathbf{A}$ singulaariarvohajoitelma on aina olemassa ja se on reaalinen!** Singulaariarvohajoitelman laskemiseen on olemassa hyvin luotettavia ja hyvin toimivia algoritmeja, joihin emme nyt syvenny.

Lause 254 Jos $n \times n$ -matriisilla $\mathbf{A}$ on n positiivista singulaariarvoa

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad \mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n),$$

niin $\mathbf{A}$ on säännöllinen ja sen käänteismatriisi on

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T, \quad \mathbf{\Sigma}^{-1} = \text{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}).$$

Todistus. HT

□

Esimerkki 255 Esimerkin ^{|c4s18qresim1}(250) matriisien singulaariarvohajotelmat saadaan seuraavasti.

```
>> help svd
```

```
SVD    Singular value decomposition.
```

```
[U,S,V] = SVD(X) produces a diagonal matrix S, of the same
dimension as X and with nonnegative diagonal elements in
decreasing order, and unitary matrices U and V so that
X = U*S*V' .
```

```
S = SVD(X) returns a vector containing the singular values.
```

```
[U,S,V] = SVD(X,0) produces the "economy size"
decomposition. If X is m-by-n with m > n, then only the
first n columns of U are computed and S is n-by-n.
```

```
See also SVDS, GSVD.
```

```
Overloaded methods
```

```
help sym/svd.m
```

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

```
A =
```

```
    1    2    3
    0    1    1
    1    1    1
    3   -1    2
```

```
>> [U,S,V] = svd(A)
```

```
U =
```

```
-0.6930    0.5189    0.3129   -0.3906
-0.2102    0.3333    0.1192    0.9113
-0.3263    0.1290   -0.9364   -0.0000
-0.6075   -0.7765    0.1048    0.1302
```

```

S =
    4.9476         0         0
         0    2.8685         0
         0         0    0.5412
         0         0         0

V =
   -0.5744   -0.5862   -0.5713
   -0.2658    0.7937   -0.5472
   -0.7742    0.1625    0.6117

>> B = [1 2 3 ; 0 1 1 ; 1 1 2 ; 3 -1 2]
B =
     1     2     3
     0     1     1
     1     1     2
     3    -1     2

>> [U2,S2,V2] = svd(B)
U2 =
   -0.6625    0.4727   -0.4128   -0.4090
   -0.2045    0.3179   -0.2028    0.9033
   -0.4580    0.1548    0.8745    0.0382
   -0.5563   -0.8072   -0.1539    0.1236

S2 =
    5.2577         0         0
         0    2.8907         0
         0         0    0.0000
         0         0         0

V2 =
   -0.5306   -0.6206   -0.5774
   -0.2722    0.7698   -0.5774
   -0.8028    0.1492    0.5774

>>

```

Esimerkin (227) defektiivisen matriisin singulaariarvohajoitelma on

```

>> Q = [2 1 0 ; 0 1 -1 ; 0 2 4]
Q =
     2     1     0
     0     1    -1
     0     2     4

>> [U,S,V] = svd(Q)

```



```

U =
    0.1213    -0.9359   -0.3307
   -0.1005   -0.3430    0.9339
    0.9875     0.0801    0.1357

S =
    4.5221         0         0
         0     2.2793         0
         0         0     1.1642

V =
    0.0537   -0.8212   -0.5681
    0.4414   -0.4908    0.7512
    0.8957    0.2911   -0.3361

>>

```

Ei siis lainkaan ongelmia defektiivisyydestä.

9.7 Matriisinormit

Sec:MatNorms

Lineaarinen yhtälöryhmä $A\vec{x} = \vec{b}$ on hyväkuntoinen, jos pienet muutokset kerroinmatriisissa A tai RHS-vektorissa $\vec{b}$ aiheuttavat vain pieniä muutoksia ratkaisuvektorissa $\vec{x}$. Muussa tapauksessa ryhmä on huonokuntoinen.

c4s19es1

Esimerkki 256 Ryhmän

$$\begin{cases} 0.9999x_1 - 1.0001x_2 = 1 \\ x_1 - x_2 = 1 \end{cases}$$

yksikäsitteinen ratkaisu on $x_1 = 0.5$ ja $x_2 = -0.5$. Jos b_2 :ta hiukan muutetaan, saadaan

$$\begin{cases} 0.9999x_1 - 1.0001x_2 = 1 \\ x_1 - x_2 = 1 + \varepsilon \end{cases} \Leftrightarrow \begin{cases} x_1 = 0.5 + 5000.5\varepsilon \\ x_2 = -0.5 + 4999.5\varepsilon \end{cases}$$

Ratkaisu on hyvin herkkä lähtötietojen muutoksille, joten yhtälöryhmä on huonokuntoinen.

Tämän ilmiön ymmärtämiseksi tutustumme seuraavaksi matriisin normiin. Vektorinormit $\|\cdot\|_1$, $\|\cdot\|_2$ ja $\|\cdot\|_\infty$ esiteltiin määritelmässä (151) sivulla 136.

$$(1) \quad \|\vec{a}\|_1 = |a_1| + |a_2| + \cdots + |a_n| = \sum_{k=1}^n |a_k|$$

$$(2) \quad \|\vec{a}\|_2 = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} = \sqrt{\langle \vec{a} | \vec{a} \rangle}$$

$$(3) \quad \|\vec{a}\|_\infty = \max\{|a_1|, |a_2|, \dots, |a_n|\} = \max_{1 \leq k \leq n} |a_k|$$

Lineaariavaruuden L normi on funktio $N : L \rightarrow \mathbb{R}$ siten, että

$$(N1) \quad N(x) \geq 0, \forall x \in L \text{ ja } N(x) = 0 \Leftrightarrow x = \theta \in L$$

$$(N2) \quad N(\alpha x) = |\alpha|N(x), \forall x \in L, \alpha \in \mathbb{R}$$

$$(N3) \quad N(x+y) \leq N(x) + N(y), \forall x, y \in L$$

Normia on tapana merkitä $N(x) = \|\vec{x}\|$. Ja kahden vektorin $\vec{x}$ ja $\vec{y}$ väliseksi etäisyydeksi sanomme niiden erotuksen normia $\|\vec{x} - \vec{y}\|$.

Valittu vektorinormi indusoi matriiseille vastaavan operaattorinormin.

$$\|\mathbf{A}\|_p = \max_{\vec{x} \neq \vec{0}} \frac{\|\mathbf{A}\vec{x}\|_p}{\|\vec{x}\|_p} = \max_{\|\vec{x}\|_p=1} \|\mathbf{A}\vec{x}\|_p \quad (9.36) \quad \boxed{\text{matnormi}}$$

Erityisesti, kun $p = 1$, $p = 2$ tai $p \rightarrow \infty$, saadaan $m \times n$ -matriisin $\mathbf{A}$ normeille kaavat

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^m |a_{ij}| \quad (= \text{itseisarvojen sarakesummien maksimi})$$

$$\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}| \quad (= \text{itseisarvojen rivisummien maksimi})$$

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A})} = \sigma_{\max}(\mathbf{A})$$

missä $\lambda_{\max}(\mathbf{A}^T \mathbf{A})$ on matriisin $\mathbf{A}^T \mathbf{A}$ itseisarvoltaan suurin ominaisarvo eli matriisin $\mathbf{A}$ suurin singulaariarvo .

Unitaarimatriisin kaikki singulaariarvot ovat ykkösiä ($\mathbf{U} = \mathbf{U}\mathbf{I}\mathbf{I}^T$), joten unitaarimatriisin 2-normi on $\|\mathbf{U}\|_2 = 1$.

Edellä määriteltyjen operaattorinormien lisäksi $m \times n$ -matriisille $\mathbf{A}$ käytetään varsin paljon myös Frobenius normia

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (9.37)$$

Määritelmän seuraa, että matriisin operaattorinormeilla on ominaisuudet

$$\|\mathbf{A}\vec{x}\|_p \leq \|\mathbf{A}\|_p \|\vec{x}\|_p \quad (9.38)$$

$$\|\mathbf{A}\mathbf{B}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p \quad (9.39)$$

$$\|\mathbf{A}^n\|_p \leq \|\mathbf{A}\|_p^n \quad (9.40)$$

9.8 Matriisin kuntoluku

Tarkastellaan nyt yhtälöryhmää, jossa kerroinmatriisi on $n \times n$ -matriisi

$$\mathbf{A}\vec{x} = \vec{b}.$$

Jos RHS muuttuu arvoon $\vec{b} + \Delta\vec{b}$, niin vastaava uusi ratkaisu on $\vec{x} + \Delta\vec{x}$. Arvioimme seuraavassa suhteellista virhettä normin $\|\cdot\|_p$ mielessä.

$$\begin{aligned} & \begin{cases} \mathbf{A}(\vec{x} + \Delta\vec{x}) = \vec{b} + \Delta\vec{b} \\ \mathbf{A}\vec{x} = \vec{b} \end{cases} \\ \Rightarrow & \Delta\vec{x} = \mathbf{A}^{-1}\Delta\vec{b} \\ \Rightarrow & \|\Delta\vec{x}\|_p \leq \|\mathbf{A}^{-1}\|_p \|\Delta\vec{b}\|_p \end{aligned} \quad (9.41)$$

Toisaalta vastaavasti

$$\begin{aligned} \|\vec{b}\|_p &= \|\mathbf{A}\vec{x}\|_p \leq \|\mathbf{A}\|_p \|\vec{x}\|_p \\ \Rightarrow & \frac{1}{\|\vec{x}\|_p} \leq \frac{\|\mathbf{A}\|_p}{\|\vec{b}\|_p} \end{aligned} \quad (9.42)$$

Yhdistämällä epäyhtälöt ^{mev1}(9.41) ja ^{mev2}(9.42) saadaan

$$\frac{\|\Delta\vec{x}\|_p}{\|\vec{x}\|_p} \leq \underbrace{\|\mathbf{A}\|_p \|\mathbf{A}^{-1}\|_p}_{=\text{Cond}_p(\mathbf{A})} \frac{\|\Delta\vec{b}\|_p}{\|\vec{b}\|_p} \quad (9.43) \quad \boxed{\text{cond1}}$$

Kaavassa ^{cond1}(9.43) olevaa kerrointa $\text{Cond}_p(\mathbf{A}) = \|\mathbf{A}\|_p \|\mathbf{A}^{-1}\|_p$ sanotaan matriisin kuntoluvuksi. Hyväkuntoisen yhtälöryhmän kerroinmatriisin kuntoluku on pieni (kuitenkin aina ≥ 1).

Tarkastellaan edelleen yhtälöryhmää

$$\mathbf{A}\vec{x} = \vec{b}$$

Jos kerroinmatriisi muuttuu arvoon $\mathbf{A} + \Delta\mathbf{A}$, niin vastaava uusi ratkaisu on $\vec{x} + \Delta\vec{x}$. Arvioimme seuraavassa suhteellista virhettä normin $\|\cdot\|_p$ mielessä.

$$\begin{aligned} & (\mathbf{A} + \Delta\mathbf{A})(\vec{x} + \Delta\vec{x}) = \vec{b} \\ \Leftrightarrow & \mathbf{A}\Delta\vec{x} + \Delta\mathbf{A}(\vec{x} + \Delta\vec{x}) = \vec{0} \\ \Leftrightarrow & \Delta\vec{x} = -\mathbf{A}^{-1}\Delta\mathbf{A}(\vec{x} + \Delta\vec{x}) \\ \Rightarrow & \|\Delta\vec{x}\|_p \leq \|\mathbf{A}^{-1}\|_p \|\Delta\mathbf{A}\|_p \|\vec{x} + \Delta\vec{x}\|_p \\ & = \text{Cond}(\mathbf{A}) \cdot \frac{\|\Delta\mathbf{A}\|_p}{\|\mathbf{A}\|_p} \cdot \|\vec{x} + \Delta\vec{x}\|_p \end{aligned}$$

Voimme siis arvioida

$$\frac{\|\Delta \vec{x}\|_p}{\|\vec{x}\|_p} \approx \frac{\|\Delta \vec{x}\|_p}{\|\vec{x} + \Delta \vec{x}\|_p} \leq \text{Cond}(\mathbf{A}) \cdot \frac{\|\Delta \mathbf{A}\|_p}{\|\mathbf{A}\|_p} \quad (9.44)$$

Jos matriisin kuntoluku on suuri, on odotettavissa ongelmia laskettaessa matriisilla.

Pieni kuntoluku on hyvä kuntoluku.

Kahden matriisin tulon kuntolukua voimme arvioida seuraavasti

$$\begin{aligned} \text{Cond}_p(\mathbf{AB}) &= \|\mathbf{AB}\|_p \|(\mathbf{AB})^{-1}\|_p = \|\mathbf{AB}\|_p \|\mathbf{B}^{-1} \mathbf{A}^{-1}\|_p \\ &\leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p \|\mathbf{B}^{-1}\|_p \|\mathbf{A}^{-1}\|_p = \text{Cond}(\mathbf{A}) \text{Cond}(\mathbf{B}) \end{aligned}$$

Koska unitaarinen matriisi säilyttää 2-normin, niin unitaarille matriisille $\mathbf{U}$ pätee

$$\begin{aligned} \|\mathbf{U}\|_2 &= 1, \quad \text{ja} \\ \text{Cond}_2(\mathbf{U}) &= 1. \end{aligned}$$

Vaikka unitaarinen (ortogonaalinen) matriisi onkin käsin laskettaessa työläs olio, niin laskettaessa tietokoneella numeerisesti, se siisti ja ongelmaton olio. Unitaarisella matriisilla kertominen ei kasvata lausekkeiden kuntolukuja.

Esimerkki 257 Pyydämme vielä Matlabilta arvion kuntoluvusta esimerkissä ^{|c4s19es1}(256) esiintyneelle kerroinmatriisille

$$\begin{pmatrix} 0.9999 & -1.0001 \\ 1 & -1 \end{pmatrix}.$$

Kun katsoo esimerkkiä ^{|c4s19es1}(256), niin ensimmäinen ajatus on, että kuntoluku olisi noin 5000, mutta asia täytyy tutkia tarkemmin:

Esimerkissä suhteellinen virhe RHS:ssä oli

$$\frac{\|\Delta \vec{b}\|_1}{\|\vec{b}\|_1} = \frac{\varepsilon}{2}$$

ja ratkaisun suhteellinen virhe oli

$$\frac{\|\Delta \vec{x}\|_1}{\|\vec{x}\|_1} = \frac{10000\varepsilon}{1}$$

ja näiden suhde on siis 20000.

```

>> help cond
COND    Condition number with respect to inversion.
COND(X) returns the 2-norm condition number (the ratio of the
largest singular value of X to the smallest). Large condition
numbers indicate a nearly singular matrix.

COND(X,P) returns the condition number of X in P-norm:

      NORM(X,P) * NORM(INV(X),P).

where P = 1, 2, inf, or 'fro'.

Class support for input X:
      float: double, single

See also rcond, condest, condeig, norm, normest.

>> A = [0.9999 -1.0001 ; 1 -1]
A =
    0.9999    -1.0001
    1.0000    -1.0000

>> cond(A,1)
ans =
    2.0001e+004

>> cond(A,2)
ans =
    2.0000e+004

>> cond(A,inf)
ans =
    2.0001e+004

>> cond(A,'fro')
ans =
    2.0000e+004

>>

```

9.9 Neliömuodot ja definiittisyys

Sec:FormDef

Seuraavassa määriteltävä neliömuoto ei ole lineaarikuvaus, vaikka sen ominaisuuksia tutkitaan matriisien avulla. Tässä kurssissa käsiteltävät sovellukset liittyvät optimoin-

tiin.

Määritelmä 258 (Epälineaarinen) kuvaus $\mathbb{R}^n \rightarrow \mathbb{R}$

$$Q(\mathbf{x}) = \sum q_{ij}x_i x_j$$

on $\mathbb{R}^n$:n neliömuoto.

Neliömuoto voidaan aina esittää symmetrisen matriisin avulla.

Esimerkki 259

$$\begin{aligned} & x_1^2 + x_2^2 + 4x_1x_2 - 2x_2x_3 \\ = & x_1x_1 + x_2x_2 + 2x_1x_2 + 2x_2x_1 - x_2x_3 - x_3x_2 \\ = & x_1(x_1 + 2x_2) + x_2(x_2 + 2x_1 - x_3) - x_3(x_2) \\ = & (x_1 \ x_2 \ x_3) \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 - x_3 \\ -x_2 \end{pmatrix} \\ = & (x_1 \ x_2 \ x_3) \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned} \quad (9.45)$$

□

Voimme siis aina olettaa, että neliömuoto voidaan esittää symmetrisen matriisin $\mathbf{A}$ avulla

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \quad (9.47)$$

Määritelmä 260 Symmetrinen $(n \times n)$ -neliömatriisi $\mathbf{A}$ on positiivisesti definiitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} > 0, \text{ kaikilla } \mathbf{x} \neq \mathbf{0}.$$

negatiivisesti definiitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} < 0, \text{ kaikilla } \mathbf{x} \neq \mathbf{0},$$

positiivisesti semidefiniitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0, \text{ kaikilla } \mathbf{x} \in \mathbb{R}^n,$$

negatiivisesti semidefiniitti, jos

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0, \text{ kaikilla } \mathbf{x} \in \mathbb{R}^n,$$

indefiniitti, jos se ei ole mitään edeltävistä.

Lause 261 Olkoon $\mathbf{A}$ symmetrinen $(n \times n)$ -neliömatriisi, jonka reaaliset ominaisarvot ovat $\lambda_1, \lambda_2, \dots, \lambda_n$.

(1) Jos kaikki ominaisarvot ovat positiivisia ($\lambda_j > 0$), niin $\mathbf{A}$ on positiivisesti definiitti.

(2) Jos kaikki ominaisarvot ovat negatiivisia ($\lambda_j < 0$), niin $\mathbf{A}$ on negatiivisesti definiitti.

(3) Jos kaikki ominaisarvot ovat ei-negatiivisia ($\lambda_j \geq 0$), niin $\mathbf{A}$ on positiivisesti semi-definiitti.

(4) Jos kaikki ominaisarvot ovat ei-positiivisia ($\lambda_j \leq 0$), niin $\mathbf{A}$ on negatiivisesti semi-definiitti.

(5) Jos on olemassa sekä positiivinen ominaisarvo, että negatiivinen ominaisarvo, niin $\mathbf{A}$ on indefiniitti.

Esimerkki 262 Tutkitaan definiittisyys matriisille

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad (9.48)$$

Lasketaan ensin ominaisarvot karakteristisesta yhtälöstä

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \quad (9.49)$$

$$\Leftrightarrow \begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 1-\lambda & -1 \\ 0 & -1 & -\lambda \end{vmatrix} = 0 \quad (9.50)$$

$$\Leftrightarrow (1-\lambda)[(1-\lambda)(-\lambda) - 1] - 2 \cdot 2 \cdot (-\lambda) = 0$$

$$\Leftrightarrow (1-\lambda)(\lambda^2 - \lambda - 1) + 4\lambda = 0$$

$$\Leftrightarrow -\lambda^3 + 2\lambda^2 + 2\lambda - 1 = 0$$

$$\Leftrightarrow (\lambda + 1)(-\lambda^2 + 3\lambda - 1) = 0$$

$$\Leftrightarrow -(\lambda + 1)\left(\lambda - \frac{3+\sqrt{5}}{2}\right)\left(\lambda - \frac{3-\sqrt{5}}{2}\right) = 0$$

Ominaisarvot ovat siis

$$\lambda_1 = \frac{3+\sqrt{5}}{2} \approx 2.618 \quad \lambda_2 = \frac{3-\sqrt{5}}{2} \approx 0.382 \quad \lambda_3 = -1$$

Siis matriisi on indefiniitti.

Pienten yksinkertaisten matriisien tapauksessa on usein helpompaa soveltaa seuraavaa lausetta

Lause 263 Olkoon $\mathbf{A}$ ($n \times n$) neliömuoto. Lasketaan apudeterminantit (vasemmasta ylänurkasta) $|D_1|, |D_2|, \dots, |D_n|$ asettamalla

$$|D_1| = a_{11} \quad \text{ja} \quad |D_k| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}$$

Jos kaikki apudeterminantit ovat positiivisia, $|D_k| > 0$, niin matriisi on positiivisesti definiitti.

Jos parittomat apudeterminantit ovat negatiivisia, $|D_k| < 0$, kun $k = 1, 3, \dots$, ja parilliset ovat positiivisia, $|D_k| > 0$, kun $k = 2, 4, \dots$, niin matriisi on negatiivisesti definiitti.

Todistus. Siuvuutetaan. □

Esimerkki 264 Tutkitaan vielä kerran definiittisyys matriisille

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad (9.51)$$

Apudeterminantit ovat

$$|D_1| = 1 > 0 \quad (9.52)$$

$$|D_2| = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -1 < 0 \quad (9.53)$$

$$|D_3| = \begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 0 \end{vmatrix} = -1 < 0 \quad (9.54)$$

Siis matriisi on indefiniitti.

Hakemisto

antisymmetrinen matriisi, 79

arccos, 19

arccot, 19

arcsin, 19

arctan, 19

arcus-funktiot, 18

aste, 10

cos, 16

Gauss-Jordan algoritmi, 54

Gaussin eliminointikeino, 54

gooni, 10

grad, 10

homogeeninen yhtälöryhmä, 65

hypotenuusa, 12

kateetti, 12

kosinifunktio, 16

kosinilause, 15

matriisi, 69

neliömatrissi, 74

nollamatriisi, 74

piiru, 11

pistetulo, 32, 33

pystyvektori, 70

Pythagoran lause, 13

rad, 10

radiaani, 10

ristitulo, 32, 35

sin, 16

sinifunktio, 16

sinilause, 15

skalaariprojektio, 38, 40

suorakulmainen kolmio, 12

symmetrinen matriisi, 79

terävä kulma, 11

transpoosi, 78

tylppä kulma, 11

vaakavektori, 70

vektoriprojektio, 38, 39

voima, 29

voiman momentti, 41, 43

yksikkömatrissi, 74