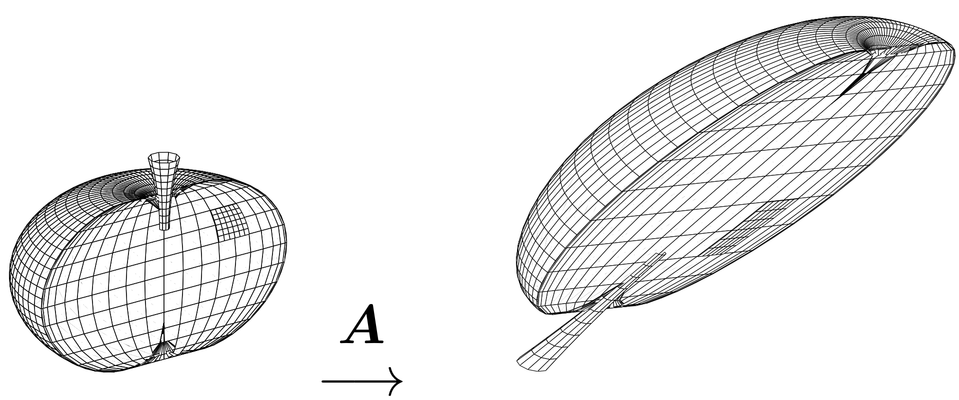
Alkuosa kurssista MS-C1340

LINEAARIALGEBRAN PERUSTEET

TIMO EIROLA



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

Aalto Yliopisto, Perustieteiden Korkeakoulu
Matematiikan ja Systeemianalyysin laitos
Syksy 2013

Tämä moniste on kirjoitettu uudelleen syksyllä 2013. Mukana on jonkin verran Ulla Miekkalan ja Jukka Tuomelan aiemman monisteen tekstiä ja esimerkkejä.

Sisältö

Johdanto, merkintöjä ii
1. Vektoriavaruudet 1
1.1. Lineaarinen riippumattomuus, kanta ja dimensio 3
2. Lineaarikuvaukset 12
2.1. Lineaarikuvauksen matriisiesitys 13
2.2. Kuva-avaruus ja nolla-avaruus 17
3. Normi ja sisätulo 21
3.1. Ortogonaalisuus 24
3.2. QR-hajotelma, pienimmän neliösumman tehtävä 28
3.3. Matriisinormi ja häiriöalttius 34
4. Ominaisarvoteoriaa 38
4.1. Similaarisuus ja diagonalisointi 43
4.2. Schurin hajotelma 44
4.3. Jordanin muoto 50
4.4. Ominaisarvojen ja -vektoreiden numeerinen laskeminen 52
4.5. Positiividefiniitit matriisit 54
4.6. Matriisin sarjat ja eksponenttifunktio 58

Johdanto, merkintöjä

Lineaarialgebra tutkii lineaarisia objekteja: suoria, tasoja, ja näiden yleistyksiä: *vektoriavaruuksia*, sekä tällaisten objektien välisiä rakenteen säilyttäviä kuvauksia *lineaarikuvauksia*. Eräät käsiteltävät asiat ovat esiintyneet matriisilaskun yhteydessä peruskursseilla 1 ja 2. Tällä kurssilla näitä osin tuttuja asioita käsitellään hieman abstraktimmin, jolloin huomaamme, että samanlaisia rakenteita esiintyy myös tarkasteltaessa vaikkapa polynomeja tai muita funktiojoukkoja. Seuraavassa muutamia esimerkkejä lineaarialgebran sovelluksista.

- Differentiaali- ja osittaisdifferentiaaliyhtälöiden teoriassa lineaarialgebra on keskeisessä asemassa. Tällöin tyypillisesti tarkastellaan matriisien *ominaisarvoja*. Näiden yhtälöiden numeerisessa ratkaisussa tarvitaan tehokkaita matriisioperaatioiden ja lineaaristen yhtälöryhmien algoritmeja. Näitä asioita käsitellään tämän kurssin lisäksi edelleen kurssilla L4.

- Signaalinkäsittelyssä monet suodatus- ja kompressointimenetelmät perustuvat tehokkaisiin matriisioperaatioihin. Esimerkiksi keskeinen työkalu – Fourier-muunnos – on unitaarinen lineaarikuvaus.

- Systeemi- ja säätöteorian keskeisten tehtävien, kuten stabilointi, optimisäätö, estimointi ja mallin reduktio, ratkaisumenetelmät perustuvat lineaarialgebran ja matriisilaskennan algoritmeihin.

- Lähes kaikki suuret laskentatehtävät käyttävät apuna lineaarialgebran ja numeerisen matriisilaskennan yhä nopeasti kehittyviä välineitä.

- Kvanttimekaniikka muotoillaan tavallisesti lineaarialgebran käsitteiden avulla, siellä tulevat vastaan itseadjungoidut lineaarikuvaukset, ominaisarvot, ominaisvektorit, kommutoivat operaattorit (ääretöndimensioisessa avaruudessa).

Tällä kurssilla lineaarialgebran numeerisia menetelmiä käsitellään varsin lyhyesti. Jatkokurssina suositellaan: Mat-1.175 Numeerinen matriisilaskenta.

Tähän on kerätty eräitä matriisioperaatioiden ominaisuuksia peruskursseilta 1 ja 2. Suurimmassa osassa eteemme tulevista tilanteista ei ole mitään eroa, onko käytettävä skalaarikunta reaaliluvut $\mathbb{R}$ vai kompleksiluvut $\mathbb{C}$. Tällöin käytetään merkintää $\mathbb{K}$, jonka tilalle voi sijoittaa kumman tahansa[1], kunhan tämä sijoitus on koko asiayhteyden ajan kiinnitetty samaksi. Merkinnällä $a \in \mathbb{K}$, tarkoitetaan siis “a on skalaari”. Vastaavasti $\boldsymbol{x} \in \mathbb{K}^n$ tarkoittaa, että $\boldsymbol{x} \in \mathbb{R}^n$ tai $\boldsymbol{x} \in \mathbb{C}^n$.

Vektoreille ja matriiseille käytämme tässä monisteessa lihavoituja kirjaimia. Matriisin $\boldsymbol{A} \in \mathbb{K}^{m\times n}$ alkioita merkitsemme joko $(\boldsymbol{A})_{ij}$ tai vastaavalla pienellä kirjaimella a_{ij}.

1. Matriisi $\boldsymbol{A} \in \mathbb{K}^{m\times n}$ on *neliömatriisi*, jos $m = n$.

2. Neliömatriisi $\boldsymbol{A}$ on *ylä(ala)kolmiomatriisi*, jos $a_{ij} = 0$ kun $i > j$ ($i < j$).

3. Neliömatriisi $\boldsymbol{A}$ on *lävistäjämatriisi* (tai *diagonaalimatriisi*), jos se on sekä ylä- että alakolmiomatriisi eli jos $a_{ij} = 0$, kun $i \neq j$. Merkinnällä $\operatorname{diag}(d_1, \ldots, d_n)$ tarkoitetaan diagonaalimatriisia, jonka diagonaalialkiot ovat $d_1, \ldots, d_n$.

4. *Nollamatriisia* merkitään yksinkertaisesti nollalla.

5. Matriisien $\boldsymbol{A} \in \mathbb{K}^{m\times p}$ ja $\boldsymbol{B} \in \mathbb{K}^{p\times n}$ *tulo* $\boldsymbol{AB} \in \mathbb{K}^{m\times n}$ määritellään
$$(\boldsymbol{AB})_{ij} = \textstyle\sum_{k=1}^{p} a_{ik}\, b_{kj} \ .$$
Matriisitulo on liitännäinen: $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$, mutta ei vaihdannainen: yleensä $\boldsymbol{AB} \neq \boldsymbol{BA}$.

6. Neliömatriisin determinanttia merkitään $\det(\boldsymbol{A})$. Sen määritelmähän oli aika hankala, mutta aina muistetaan, että kolmiomatriisin determinantti on sen lävistäjäalkioiden tulo ja että pätee:
$$\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B}) \ .$$
Determinantin laskeminen suoritetaan tavallisimmin saattamalla matriisi Gaussin eliminointiaskelin (jotka eivät muuta determinanttia) yläkolmiomuotoon.

7. *Yksikkömatriisi* $\boldsymbol{I} = \operatorname{diag}(1, 1, \ldots, 1)$ on diagonaalimatriisi, jonka diagonaalialkiot ovat kaikki ykkösiä eli $(\boldsymbol{I})_{ij} = \delta_{ij}$. Sille pätee $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$ kaikilla matriiseilla $\boldsymbol{A}$.

8. Jos $\det(\boldsymbol{A}) \neq 0$, niin $\boldsymbol{A}$:lla on olemassa *käänteismatriisi* $\boldsymbol{A}^{-1}$, jolle pätee
$$\boldsymbol{AA}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I} \ .$$

9. Matriisin $\boldsymbol{A} \in \mathbb{K}^{m\times n}$:n *transpoosia* merkitään $\boldsymbol{A}^T \in \mathbb{K}^{n\times m}$, jolloin siis $(\boldsymbol{A}^T)_{ij} = (\boldsymbol{A})_{ji}$. Operaatio $\boldsymbol{A}^* = \overline{\boldsymbol{A}^T}$ on $\boldsymbol{A}$:n *hermitointi*. Matriisia $\boldsymbol{A}^*$ (reaalisessa tapauksessa $\boldsymbol{A}^T$) kutsutaan $\boldsymbol{A}$:n *liittomatriisiksi*.

10. $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ on *symmetrinen*, jos $\boldsymbol{A}^T = \boldsymbol{A}$.

11. $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ on *vinosymmetrinen*, jos $\boldsymbol{A}^T = -\boldsymbol{A}$.

[1] Joissakin tilanteissa myös rationaalilukujen joukko $\mathbb{Q}$ voi olla kätevä.

12. $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ on *hermiittinen*, jos $\boldsymbol{A}^* = \boldsymbol{A}$.

13. $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ on *ortogonaalinen*, jos $\boldsymbol{A}^T = \boldsymbol{A}^{-1}$, eli $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$.

14. $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ on *unitaarinen*, jos $\boldsymbol{A}^* = \boldsymbol{A}^{-1}$, eli $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}$.

15. Matriisitulolle pätee $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$ ja invertoituville matriiseille $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

16. Olkoon $\boldsymbol{A} \in \mathbb{K}^{m\times n}$. $\boldsymbol{A}$:n *nolla-avaruus* (eli *ydin*) on
$$N(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{K}^n \mid \boldsymbol{A}\boldsymbol{x} = 0\} ,$$
ja $\boldsymbol{A}$:n *kuva-avaruus* on
$$R(\boldsymbol{A}) = \{\boldsymbol{A}\boldsymbol{x} \in \mathbb{K}^m \mid \boldsymbol{x} \in \mathbb{K}^n\} .$$

17. Reaalisille vektoreille $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^n$ määritellään *sisätulo*[2]: $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^T\boldsymbol{y} = \sum_{i=1}^n x_i\, y_i$.

18. Kompleksisille vektoreille $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{C}^n$ se määritellään: $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^T\overline{\boldsymbol{y}} = \sum_{i=1}^n x_i\, \overline{y}_i$.

19. Molemmissa tapauksissa määritellään *normi*: $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$. Tarkemmin: tätä normia merkitään $\|\cdot\|_2$, sillä jatkossa näemme muitakin normeja.

20. Kun $\boldsymbol{x} \in \mathbb{C}^n$, $\boldsymbol{y} \in \mathbb{C}^m$ ja $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, niin
$$\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{y}\rangle .$$
Reaalisessa tapauksessa: $\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y}\rangle = \left\langle \boldsymbol{x}, \boldsymbol{A}^T\boldsymbol{y}\right\rangle$.

Seuraavaan lauseeseen on koottu säännöllisiä (eli invertoituvia) matriiseja koskevat tärkeimmät ominaisuudet:

Lause 0.1. *Olkoon* $\boldsymbol{A} \in \mathbb{K}^{n\times n}$. *Seuraavat väitteet ovat yhtäpitäviä:*

1. $\boldsymbol{A}^{-1}$ *on olemassa.*
2. $\det(\boldsymbol{A}) \neq 0$.
3. *Systeemillä* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *on yksikäsitteinen ratkaisu kaikilla* $\boldsymbol{b}$.
4. $\boldsymbol{A}$ *:n sarakkeet ovat lineaarisesti riippumattomat.*
5. $\boldsymbol{A}$ *:n rivit ovat lineaarisesti riippumattomat.*
6. $N(\boldsymbol{A}) = \{0\}$.
7. $R(\boldsymbol{A}) = \mathbb{K}^n$.

[2] Tässä vektori $\boldsymbol{x} \in \mathbb{K}^n$ ajatellaan $n \times 1$-matriisiksi, jolloin $\boldsymbol{x}^T$ on $1 \times n$-matriisi

1. Vektoriavaruudet

Kaikkien tason (avaruuden) vektorit muodostavat joukon, jonka alkioita voidaan laskea yhteen ja kertoa skalaarilla. Nämä laskutoimitukset toteuttavat joukon laskusääntöjä kuten vaihdannaisuus, liitännäisyys ja osittelulait, jolloin joukkoa kutsutaan vektoriavaruudeksi. Tämä yleistyy heti n–pituisille vektoreille, jotka muodostavat vektoriavaruuden $\mathbb{R}^n$. Samoin kompleksisista vektoreista voidaan muodostaa vektoriavaruus $\mathbb{C}^n$. Vektoriavaruus (eli lineaariavaruus) on kuitenkin vielä yleisempi käsite:

Määritelmä 1.1. Joukkoa V sanotaan *vektoriavaruudeksi*, jos V:n alkioille on määritelty yhteenlasku $+\ :\ V \times V \to V$ ja skalaarilla kertominen $\cdot\ :\ \mathbb{K} \times V \to V$ siten, että[3]

(1)	$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} = \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w}) \quad \forall\, \boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in V$	(liitäntälaki)
(2)	On olemassa alkio $0 \in V$, s.e. $\boldsymbol{v}+0=\boldsymbol{v} \quad \forall\, \boldsymbol{v} \in V$	(nolla-alkio)
(3)	jokaisella $\boldsymbol{v} \in V$ on olemassa $-\boldsymbol{v} \in V$ s.e. $\boldsymbol{v}+(-\boldsymbol{v})=0$	(vasta-alkio)
(4)	$\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u} \quad \forall\, \boldsymbol{u},\boldsymbol{v} \in V$	(vaihdantalaki)
(5)	$\alpha\,(\boldsymbol{u}+\boldsymbol{v}) = \alpha\,\boldsymbol{u}+\alpha\,\boldsymbol{v} \quad \forall\, \alpha \in \mathbb{K},\ \boldsymbol{u},\boldsymbol{v} \in V$	(osittelulaki)
(6)	$(\alpha+\beta)\,\boldsymbol{v} = \alpha\,\boldsymbol{v}+\beta\,\boldsymbol{v} \quad \forall\, \alpha,\beta \in \mathbb{K},\ \boldsymbol{v} \in V$	(osittelulaki)
(7)	$\alpha\,(\beta\,\boldsymbol{v}) = (\alpha\,\beta)\,\boldsymbol{v} \quad \forall\, \alpha,\beta \in \mathbb{K},\ \boldsymbol{v} \in V$	(liitäntälaki)
(8)	$1\,\boldsymbol{v}=\boldsymbol{v} \quad \forall\, \boldsymbol{v} \in V$.	

Jos $\mathbb{K}=\mathbb{R}$, niin sanotaan, että V on reaalikertoiminen vektoriavaruus. Kun $\mathbb{K}=\mathbb{C}$, se on kompleksikertoiminen. Vektoriavaruuden alkioita kutsutaan vektoreiksi[4].

Esimerkki 1.1. $\mathbb{R}^2$ varustettuna tutuilla yhteenlasku-ja kertomissäännöillä

$$(u_1,u_2)+(v_1,v_2)=(u_1+v_1,u_2+v_2)\ , \qquad \alpha(v_1,v_2)=(\alpha v_1,\alpha v_2)$$

on vektoriavaruus, kuten on helposti tarkistettavissa. Samoin $\mathbb{R}^n$ on vektoriavaruus.

Jos $\mathbb{R}^2$:ssa määriteltäisiin yhteenlasku säännöllä

$$(u_1,u_2)+(v_1,v_2)=(u_1+v_2,u_2+v_1)\ ,$$

huomataan, että vaihdantalaki ei olisi voimassa:

$$\boldsymbol{u}+\boldsymbol{v}=(u_1+v_2,u_2+v_1) \neq (v_1+u_2,v_2+u_1)=\boldsymbol{v}+\boldsymbol{u}\ .$$

Siten $\mathbb{R}^2$ varustettuna tällä yhteenlaskulla (ja kertosäännöllä $\alpha(v_1,v_2)=(\alpha v_1,\alpha v_2)$) ei ole vektoriavaruus.

Esimerkki 1.2. $V=\{\boldsymbol{v} \in \mathbb{R}^2 \mid v_2 = m v_1\}$, jossa m on annettu kiinteä reaaliluku. Kyseessä on siis tason vektorit, joiden kärki on origon kautta kulkevalla suoralla $y=mx$. Nämä muodostavat vektoriavaruuden, kun laskutoimitukset ovat tavalliset tason vektorien

[3] Huomaa, että kertolaskun piste jätetään tavallisesti kirjoittamatta.

[4] vaikka ne monissa tilanteissa todellisuudessa ovat matriiseja, funktioita, operaattoreita tms.

yhteenlasku ja reaaliluvulla kertominen:

$$\begin{aligned} &\text{a)} \qquad \boldsymbol{u},\ \boldsymbol{v} \in V \Rightarrow\ \boldsymbol{u}+\boldsymbol{v} = (u_1+v_1, u_2+v_2) = \\ &\qquad\quad (u_1+v_1, mu_1+mv_1) = (u_1+v_1, m(u_1+v_1)) \in V \\ &\text{b)} \qquad \boldsymbol{v} \in V, \alpha \in \mathbb{R}\ \Rightarrow\ \alpha\boldsymbol{v} = \alpha(v_1,v_2) = (\alpha v_1, \alpha v_2) = \\ &\qquad\quad (\alpha v_1, \alpha(mv_1)) = (\alpha v_1, m(\alpha v_1)) \in V \end{aligned}$$

V sisältää nollavektorin $0 = (0,0) \in V$ ja kunkin V:n vektorin $\boldsymbol{v}$ vastavektori on $-\boldsymbol{v} \in V$. Muutkin aksioomat toteutuvat, koska ne pätevät $\mathbb{R}^2$:ssa.

Esimerkki 1.3. $V = \{\boldsymbol{v} \in \mathbb{R}^2 \mid v_2 = mv_1 - 1\}$, eli ne tason vektorit, jotka muodostavat suoran $y = mx - 1$. Koska $0 \notin V$, V ei voi olla vektoriavaruus.

Pienin mahdollinen vektoriavaruus on nk. triviaaliavaruus $V = \{0\}$.

Esimerkki 1.4. Kaikkien reaalisten $m \times n$-matriisien muodostama joukko $\mathbb{R}^{m\times n}$ on vektoriavaruus, kun yhteenlasku ja reaaliluvulla kertominen on määritelty alkioittain kuten matriiseille on tapana. Nolla-alkio on tällöin $m \times n$-nollamatriisi ja määritelmän ehdot on helppo tarkistaa. Samoin kompleksisten $m \times n$-matriisien joukko $\mathbb{C}^{m\times n}$ on kompleksikertoiminen vektoriavaruus.

Esimerkki 1.5. Olkoon ℓ_∞ kaikkien rajoitettujen reaalilukujonojen joukko:

$$\ell_\infty = \left\{\, \boldsymbol{\xi} = (\xi_1, \xi_2, \dots) \,\middle|\, \sup_j |\xi_j| < \infty \right\} .$$

Tämä on $\mathbb{R}$-kertoiminen vektoriavaruus, kun määritellään:

$$(\xi_1,\xi_2,\dots)+(\eta_1,\eta_2,\dots) = (\xi_1+\eta_1,\ \xi_2+\eta_2,\ \dots) \qquad \text{ja} \qquad \alpha\,(\xi_1,\xi_2,\dots) = (\alpha\,\xi_1,\ \alpha\,\xi_2,\ \dots)\,.$$

Vastaavasti rajoitettujen kompleksilukujonojen joukko on $\mathbb{C}$-kertoiminen vektoriavaruus.

Esimerkki 1.6. Korkeintaan toista astetta olevien x:n polynomien joukko

$$\mathbb{P}_2 = \{p \mid p(x) = a_0 + a_1 x + a_2 x^2, a_0,\ a_1,\ a_2 \in \mathbb{C}\}$$

on vektoriavaruus, kun polynomien yhteenlasku ja skalaarilla kertominen määritellään normaaliin tapaan. Samoin korkeintaan astetta n olevien polynomien joukko $\mathbb{P}_n$ on vektoriavaruus. Edelleen kaikkien polynomien joukko $\mathbb{P}_\infty$ on myös vektoriavaruus.

Edellä olevan esimerkin vektoriavaruuksien alkiot ovat funktioita, joten niitä kutsutaan *funktioavaruuksiksi.* Laskusäännöt funktioille määritellään luonnollisesti

$$(f+g)(x) = f(x) + g(x) \qquad \text{ja} \qquad (\alpha f)(x) = \alpha(f(x))\ .$$

Toinen tärkeä esimerkki funktioavaruuksista on seuraava.

Esimerkki 1.7. Välillä $[a,b]$ jatkuvien reaaliarvoisten funktioiden joukko $C[a,b]$ on vektoriavaruus, sillä kahden jatkuvan funktion summa on jatkuva ja skalaarilla kertominen kuvaa jatkuvat funktiot jatkuville. Vektoriavaruuden aksioomat on helppo tarkistaa.

Määritelmä 1.2. $\mathbb{K}$-kertoimisen vektoriavaruuden V ei-tyhjä osajoukko $S \subset V$ on V:n *aliavaruus*, jos pätee

a) $\boldsymbol{u}, \boldsymbol{v} \in S\ \Rightarrow\ \boldsymbol{u}+\boldsymbol{v} \in S$,

b) $\boldsymbol{v} \in S\,,\ \alpha \in \mathbb{K} \ \Rightarrow\ \alpha\boldsymbol{v} \in S\,,$

kun käytetään V:ssä määriteltyjä laskutoimituksia. Ehdot a) ja b) voidaan kirjoittaa yhteen ekvivalentiksi ehdoksi:

$$\boldsymbol{u}, \boldsymbol{v} \in S\,,\ \alpha, \beta \in \mathbb{K} \ \Rightarrow\ \alpha\,\boldsymbol{u} + \beta\,\boldsymbol{v} \in S\ .$$

Aliavaruus on myös vektoriavaruus. Vektoriavaruuden määritelmässä vaadittavat aksioomat pätevät aliavaruuden S alkioille, sillä ne ovat myös V:n alkioita. S sisältää nolla-alkion ja vasta-alkiot. Tämä nähdään, kun valitaan $\alpha = 0$ ja $\alpha = -1$ ehdossa b).

Esimerkin 1.2 vektoriavaruutta $V = \{\boldsymbol{v} \in \mathbb{R}^2 \mid v_2 = mv_1\}$ voidaan tarkastella myös tason $\mathbb{R}^2$ aliavaruutena. Samoin avaruus $\mathbb{P}_2$ voidaan tulkita avaruuksien $\mathbb{P}_n$, $\mathbb{P}_\infty$ tai $C[0,1]$ aliavaruudeksi.

Esimerkki 1.8. Avaruuden $\mathbb{R}^3$ kaikki aliavaruudet ovat

i) Triviaaliavaruus $\{(0,0,0)\}$.
ii) Kaikki muotoa $\{a\boldsymbol{w} \mid a \in \mathbb{R}\}$ olevat joukot, missä $\boldsymbol{w}$ on kiinteä, eli origon kautta kulkevat suorat.
iii) Kaikki muotoa $\{a\boldsymbol{w}^1 + b\boldsymbol{w}^2 \mid a, b \in \mathbb{R}\}$ olevat joukot, missä $\boldsymbol{w}^1$ ja $\boldsymbol{w}^2 \in \mathbb{R}^3$ ovat kiinteitä, eli origon kautta kulkevat tasot.
iv) Avaruus $\mathbb{R}^3$ itse.

Kaikki vektoriavaruudet sisältävät ainakin kaksi aliavaruutta: avaruuden itsensä ja triviaaliavaruuden.

Esimerkki 1.9. Onko $\mathbb{R}^2$:n osajoukko $S = \{\boldsymbol{v} \in \mathbb{R}^2 \mid v_1 \geq 0\}$ aliavaruus? Selvästi määritelmän ehto a) on voimassa. Jos ehdossa b) valitaan negatiivinen skalaarikerroin, huomataan, ettei ehto toteudu, joten S ei ole aliavaruus.

Esimerkki 1.10. Kaikkien symmetristen 3×3-matriisien joukko on vektoriavaruuden $\mathbb{R}^{3\times 3}$ (kaikki 3×3- matriisit) aliavaruus, sillä matriisien yhteenlasku ja skalaarilla kertominen säilyttävät symmetrian.

1.1. **Lineaarinen riippumattomuus, kanta ja dimensio.** Olkoot $\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^n$ $\mathbb{K}$-kertoimisen vektoriavaruuden V alkioita ja $c_1, \ldots, c_n \in \mathbb{K}$. Muotoa

$$c_1\,\boldsymbol{v}_1 + c_2\,\boldsymbol{v}_2 + \cdots + c_n\,\boldsymbol{v}_n$$

olevaa lauseketta kutsutaan vektoreiden $\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n$ *lineaarikombinaatioksi.*

Esimerkki 1.11. Onko $q(x) = x^2 + x + 2$ polynomien $p_1(x) = x^2 + 5$ ja $p_2(x) = x^2 + 2x - 1$ lineaarikombinaatio? Toisin sanoen: onko olemassa c_1, $c_2 \in \mathbb{R}$ siten, että $q(x) = c_1p_1(x) + c_2p_2(x)$? Saadaan yhtälöt

$$\begin{aligned} x^2 + x + 2 &= c_1(x^2 + 5) + c_2(x^2 + 2x - 1) \\ \Leftrightarrow\ x^2 + x + 2 &= (c_1 + c_2)x^2 + 2c_2x + 5c_1 - c_2\ , \end{aligned}$$

joten päädytään lineaarisen yhtälöryhmään

$$\begin{array}{rcrcl} c_1 & + & c_2 & = & 1 \\ & & 2c_2 & = & 1 \\ 5c_1 & - & c_2 & = & 2 \end{array}$$

Koska tällä yhtälöryhmällä on ratkaisu: $c_1 = c_2 = 1/2$, polynomi q voidaan esittää p_1:n ja p_2:n lineaarikombinaationa.

Määritelmä 1.3. Ei-tyhjän joukon $S \subset V$ *viritelmä* (engl. span) $\mathrm{sp}(S)$ on kaikkien S:n alkioista muodostettujen lineaarikombinaatioiden joukko eli

$$\mathrm{sp}(S) = \left\{ \boldsymbol{u} \in V \mid \boldsymbol{u} = \textstyle\sum_{i=1}^n a_i\, \boldsymbol{s}^i,\ a_i \in \mathbb{K},\ \boldsymbol{s}^i \in S,\ 1 \le n < \infty \right\} .$$

Viritelmästä käytetään myös nimitystä joukon S *virittämä aliavaruus*, sillä se on aina aliavaruus:

Tod. Olkoot $\boldsymbol{u}, \boldsymbol{v} \in \mathrm{sp}(S)$ ja $\alpha, \beta \in \mathbb{K}$. Tällöin on olemassa $a_1, \ldots, a_n \in \mathbb{K}$ ja $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n \in S$ siten, että $\boldsymbol{u} = a_1\, \boldsymbol{s}_1 + \cdots + a_n\, \boldsymbol{s}_n$ sekä $b_1, \ldots, b_m \in \mathbb{K}$ ja $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m \in S$ siten, että $\boldsymbol{v} = b_1\, \boldsymbol{t}_1 + \cdots + b_m\, \boldsymbol{t}_m$. Saadaan

$$\alpha\, \boldsymbol{u} + \beta\, \boldsymbol{v} = \alpha a_1\, \boldsymbol{s}_1 + \cdots + \alpha a_n\, \boldsymbol{s}_n + \beta b_1\, \boldsymbol{t}_1 + \cdots + \beta b_m\, \boldsymbol{t}_m \ .$$

Tämähän on S:n vektoreiden $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n, \boldsymbol{t}_1, \ldots, \boldsymbol{t}_m$ lineaarikombinaatio, joten $\alpha\, \boldsymbol{u} + \beta\, \boldsymbol{v} \in \mathrm{sp}(S)$. □

Jos S sisältää vain äärellisen määrän vektoreita on tapana kirjoittaa $\mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k)$. Jos $U = \mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k)$, sanotaan myös, että vektorijoukko $\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k\}$ *virittää* U:n.

Esimerkki 1.12. Olkoot $\boldsymbol{v}^1 = (1, 1, 1)$ ja $\boldsymbol{v}^2 = (-2, 0, 1)$. Näiden viritelmä

$$\begin{aligned} \mathrm{sp}(\boldsymbol{v}^1, \boldsymbol{v}^2) &= \left\{ \boldsymbol{w} \in \mathbb{R}^3 \mid \boldsymbol{w} = c_1 \boldsymbol{v}^1 + c_2 \boldsymbol{v}^2,\ c_1, c_2 \in \mathbb{R} \right\} \\ &= \left\{ (c_1 - 2c_2,\ c_1,\ c_1 + c_2) \mid c_1,\ c_2 \in \mathbb{R} \right\} \end{aligned}$$

sisältää kaikki avaruuden vektorit, joiden pätepiste sijaitsee vektoreiden $\boldsymbol{v}^1$ ja $\boldsymbol{v}^2$ määräämällä, origon kautta kulkevalla tasolla.

Esimerkki 1.13. $\mathrm{sp}(1, x, x^2) = \mathbb{P}_2$, sillä jokainen toisen asteen polynomi voidaan kirjoittaa muodossa $c_0 + c_1\, x + c_2\, x^2$. Toisaalta $\mathrm{sp}(1, x^2)$ on niiden toisen asteen polynomien joukko (ja $\mathbb{P}_2$:n aliavaruus), jotka eivät sisällä ensimmäisen asteen termiä.

Määritelmä 1.4. Vektoriavaruuden osajoukko $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^n\}$ on *lineaarisesti riippumaton,* jos nollavektori voidaan esittää näiden lineaarikombinaationa *vain* siten, että kaikki kertoimet ovat nollia, eli kun $c_1 = c_2 = \cdots = c_n = 0$ on yhtälön

$$c_1\, \boldsymbol{v}^1 + c_2\, \boldsymbol{v}^2 + \ldots + c_n\, \boldsymbol{v}^n = 0 \tag{1.1}$$

ainoa ratkaisu. Jos muitakin ratkaisuja on, niin joukkoa $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^n\}$ sanotaan *lineaarisesti riippuvaksi*

Lineaarisen riippumattomuuden tutkiminen vektoriavaruuksissa, jotka voidaan samaistaa $\mathbb{R}^n$:n tai jonkin sen aliavaruuden kanssa on samankaltaista kuin $\mathbb{R}^n$:n vektoreille. Määritelmässä esiintyvä yhtälö (1.1), jossa kertoimet c_k ovat tuntemattomia ja vektorit $\boldsymbol{v}^k$ annettuja, voidaan kirjoittaa matriisimuotoon

$$\boldsymbol{A}\,\boldsymbol{c} = 0 \,, \tag{1.2}$$

missä $\boldsymbol{A} = [\boldsymbol{v}_1\ \boldsymbol{v}_2\ \dots\ \boldsymbol{v}_n]$ eli vektorit $\boldsymbol{v}^k$ muodostavat kerroinmatriisin $\boldsymbol{A}$ sarakkeet ja $\boldsymbol{c} = (c_1, \dots, c_n)$. Syntyneestä homogeenisesta lineaarisesta yhtälöryhmästä voidaan ratkaista Gaussin eliminaatiolla kertoimet c_k. (Itse asiassa riittää vain selvittää, onko yhtälöryhmällä ei-triviaali ratkaisu.)

Esimerkki 1.14. Ovatko vektoriavaruuden $\mathbb{R}^{2\times 2}$ alkiot $\left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}1&2\\0&0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}4&-1\\3&0\end{smallmatrix}\right]$ ja $\left[\begin{smallmatrix}-2&5\\6&1\end{smallmatrix}\right]$ lineaarisesti riippumattomia? Saamme yhtälön

$$c_1 \left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right] + c_2 \left[\begin{smallmatrix}1&2\\0&0\end{smallmatrix}\right] + c_3 \left[\begin{smallmatrix}4&-1\\3&0\end{smallmatrix}\right] + c_4 \left[\begin{smallmatrix}-2&5\\6&1\end{smallmatrix}\right] = 0$$

eli

$$\begin{bmatrix} c_1 + c_2 + 4c_3 - 2c_4 & 2c_2 - c_3 + 5c_4 \\ c_1 + 3c_3 + 6c_4 & c_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} .$$

Matriisimuodossa tämä on

$$\begin{bmatrix} 1 & 1 & 4 & -2 \\ 0 & 2 & -1 & 5 \\ 1 & 0 & 3 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} .$$

Tämän kerroinmatriisi on säännöllinen: sen determinantti on -3. Siten homogeenisella yhtälöryhmällä (1.2) on vain nollaratkaisu, ja matriisijoukko on lineaarisesti riippumaton.

Funktioavaruuksien kohdalla tarkastelu voi olla hankalampaa. Korkeintaan astetta n olevat polynomit voidaan samaistaa $\mathbb{K}^{n+1}$:n vektoreihin, mutta yleisemmin tarvitaan muita keinoja.

Esimerkki 1.15. Ovatko funktiot $\sin^2 x$, $\cos 2x$ ja 1 lineaarisesti riippumattomia? Yhtälö (1.1) saa nyt muodon

$$c_1 \sin^2 x + c_2 \cos 2x + c_3 = 0 \tag{1.3}$$

Sijoittamalla tähän $\cos 2x = \cos^2 x - \sin^2 x = 1 - 2\sin^2 x$ saadaan

$$(c_1 - 2c_2)\sin^2 x + c_2 + c_3 = 0,$$

josta ratkaistaan $c_3 = -c_2$ ja $c_1 = 2c_2$. Koska c_2 voidaan valita vapaasti, muodostavat funktiot lineaarisesti riippuvan joukon.

Esimerkki 1.16. Jos funktiojoukko on lineaarisesti riippumaton, sen voi usein näyttää tarkastelemalla näitä funktioita riittävän monessa pisteessä.
Esimerkiksi $C[-1,1]$:n funktiot 1, e^x, e^{-x} ovat lineaarisesti riippumattomat, sillä kirjoittamalla yhtälö $c_1 + c_2\, e^x + c_3\, e^{-x} = 0$ pisteissä $x \in \{-1, 0, 1\}$ saadaan yhtälöryhmä

$$\begin{bmatrix} 1 & 1/e & e \\ 1 & 1 & 1 \\ 1 & e & 1/e \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0 \,,$$

jonka kerroinmatriisin determinantti on $-(1+e)(1-e)^3/e^2$. Tämä ei ole nolla, joten jo tämän kolmessa pisteessä kirjoitetun yhtälön ainoa ratkaisu on $c_1 = c_2 = c_3 = 0$. Siispä funktiot ovat lineaarisesti riippumattomat.

Lause 1.1. *Vektorijoukko* $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^n\}$ *on lineaarisesti riippuva jos ja vain jos jokin vektoreista voidaan esittää muiden vektoreiden lineaarikombinaationa.*

Tod. "$\Rightarrow$" Jos vektorijoukko on lineaarisesti riippuva, niin yhtälölle (1.1) löytyy ratkaisu, jossa ainakin yksi kerroin $c_l \neq 0$. Jakamalla yhtälö c_l:llä, saadaan

$$\boldsymbol{v}^l = -\tfrac{c_1}{c_l}\,\boldsymbol{v}^1 - \cdots - \tfrac{c_{l-1}}{c_l}\,\boldsymbol{v}^{l-1} - \tfrac{c_{l+1}}{c_l}\,\boldsymbol{v}^{l+1} - \cdots - \tfrac{c_n}{c_l}\,\boldsymbol{v}^n ,$$

mistä nähdään, että $\boldsymbol{v}^l$ on muiden vektoreiden lineaarikombinaatio.

"$\Leftarrow$" Oletetaan, että $\boldsymbol{v}^k$ voidaan lausua muiden vektoreiden lineaarikombinaationa

$$\boldsymbol{v}^k = d_1\,\boldsymbol{v}^1 + \cdots + d_{k-1}\,\boldsymbol{v}^{k-1} + d_{k+1}\,\boldsymbol{v}^{k+1} + \cdots + d_n\,\boldsymbol{v}^n .$$

Siirretään $\boldsymbol{v}^k$ tässä yhtälössä toiselle puolelle, jolloin saadaan

$$d_1\,\boldsymbol{v}^1 + \cdots + d_{k-1}\,\boldsymbol{v}^{k-1} - \boldsymbol{v}^k + d_{k+1}\,\boldsymbol{v}^{k+1} + \cdots + d_n\,\boldsymbol{v}^n = 0 .$$

Tämä on sama kuin yhtälö (1.1) määritelmässä, ja ainakin yksi sen kerroin on $\neq 0$. Siten vektorit ovat lineaarisesti riippuvia. $\square$

Vektoriavaruuden V äärellistä osajoukkoa $B = \{\boldsymbol{b}^1, \boldsymbol{b}^2, \ldots, \boldsymbol{b}^n\}$ sanotaan V:n *kannaksi*, jos se on lineaarisesti riippumaton ja virittää V:n (eli $\mathrm{sp}(B) = V$).

$\mathbb{R}^n$:n *luonnolliseksi kannaksi* kutsutaan vektorijoukkoa $\{\boldsymbol{e}^1, \boldsymbol{e}^2, \ldots, \boldsymbol{e}^n\}$, jossa $\boldsymbol{e}^i \in \mathbb{R}^n$ sisältää muuten nollia paitsi sen i:s komponentti on 1. Siten vektorianalyysistä tuttuja kantavektoreita $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ merkitään $\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\}$.

Esimerkki 1.17. $\mathbb{P}_2$:ssa "luonnolliseksi kannaksi" voisi kutsua joukkoa $\{1, x, x^2\}$. Toinen kanta $\mathbb{P}_2$:lle on esim. $Q = \{x^2+1, 2x^2+x-1, x^2-x\}$. Joukko on lineaarisesti riippumaton, sillä

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

ja asettamalla yhtälön kertoimet nolliksi saadaan ainoaksi ratkaisuksi $c_1 = c_2 = c_3 = 0$. Ja jokainen toisen asteen polynomi voidaan lausua näiden avulla, sillä

$$a + b\,x + c\,x^2 = \tfrac{1}{4}(3a+b+c)\,(x^2+1) + \tfrac{1}{4}(-a+b+c)\,(2x^2+x-1) + \tfrac{1}{4}(-a-3b+c)\,(x^2-x) .$$

Kertoimet on ratkaistu lineaarisesta yhtälöryhmästä.

Vektoriavaruuden kanta voidaan valita monella tavalla, mutta jokainen kanta sisältää yhtä monta vektoria. Tämä saadaan seuraavasta lauseesta.

Lause 1.2. *Olkoon* V *vektoriavaruus. Jos joukko* $\{\boldsymbol{w}^1, \boldsymbol{w}^2, \ldots, \boldsymbol{w}^n\}$ *virittää* V*:n ja* $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^k\} \subset V$ *on lineaarisesti riippumaton, niin pätee* $n \geq k$.

Tod. Käytetään epäsuoraa todistusta, eli oletetaan että $n < k$. Koska $\{\boldsymbol{w}^1, \boldsymbol{w}^2, \ldots, \boldsymbol{w}^n\}$ virittää V:n, niin vektorit $\boldsymbol{v}^i$ voidaan kirjoittaa näiden lineaarikombinaationa:

$$\begin{aligned}
\boldsymbol{v}^1 &= a_{11}\,\boldsymbol{w}^1 + a_{12}\,\boldsymbol{w}^2 + \cdots + a_{1n}\,\boldsymbol{w}^n \ , \\
\boldsymbol{v}^2 &= a_{21}\,\boldsymbol{w}^1 + a_{22}\,\boldsymbol{w}^2 + \cdots + a_{2n}\,\boldsymbol{w}^n \ , \\
\vdots \quad & \quad \vdots \qquad\quad \vdots \qquad\qquad\qquad \vdots \\
\boldsymbol{v}^k &= a_{k1}\,\boldsymbol{w}^1 + a_{k2}\,\boldsymbol{w}^2 + \cdots + a_{kn}\,\boldsymbol{w}^n \ .
\end{aligned}$$

Olkoon

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{bmatrix}$$

ja $\boldsymbol{c} \in \mathbb{K}^k$. Lineaaristen yhtälöryhmien teoriasta tiedetään, että homogeenisella systeemillä $\boldsymbol{A}^T\boldsymbol{c} = 0$ on ei-triviaali ratkaisu, kun yhtälöitä on vähemmän kuin tuntemattomia ($n < k$). Olkoon $\boldsymbol{c} = (c_1, c_2, \ldots, c_k) \neq 0$ eräs ratkaisu. Lausumalla vektorit $\boldsymbol{v}^i$ vektoreiden $\boldsymbol{w}^i$ avulla ja keräämällä $\boldsymbol{w}^i$-vektorien kertoimet yhteen saadaan

$$\begin{aligned}
& c_1\,\boldsymbol{v}^1 + c_2\,\boldsymbol{v}^2 + \cdots + c_k\,\boldsymbol{v}^k \\
&\qquad = c_1\,(a_{11}\,\boldsymbol{w}^1 + \cdots + a_{1n}\,\boldsymbol{w}^n) + c_2\,(a_{21}\,\boldsymbol{w}^1 + \cdots + a_{2n}\,\boldsymbol{w}^n) + \cdots + c_k\,(a_{k1}\,\boldsymbol{w}^1 + \cdots + a_{kn}\,\boldsymbol{w}^n) \\
&\qquad = (c_1a_{11} + c_2a_{21} + \cdots + c_ka_{k1})\,\boldsymbol{w}^1 + \cdots + (c_1a_{1n} + c_2a_{2n} + \cdots + c_ka_{kn})\,\boldsymbol{w}^n \\
&\qquad = 0\ \ \boldsymbol{w}^1 + \cdots + 0\,\boldsymbol{w}^n = 0\ .
\end{aligned}$$

Siis vektoreiden $\boldsymbol{v}^i$ lineaarikombinaatio, jossa ainakin jokin kerroin $\neq 0$ tuottaa nollavektorin. Tämähän on ristiriidassa sen oletuksen kanssa, että $\{\boldsymbol{v}^1,\ \boldsymbol{v}^2, \ldots,\ \boldsymbol{v}^k\}$ on lineaarisesti riippumaton. Siis on oltava $n \geq k$. □

Tarkastellaan V:n kahta kantaa $\{\boldsymbol{b}^1, \boldsymbol{b}^2, \ldots, \boldsymbol{b}^n\}$ ja $\{\widehat{\boldsymbol{b}}^1, \widehat{\boldsymbol{b}}^2, \ldots, \widehat{\boldsymbol{b}}^{\widehat{n}}\}$, jotka siis molemmat virittävät V:n ja ovat lineaarisesti riippumattomia. Soveltamalla edellistä lausetta kahteen kertaan saadaan $n \leq \widehat{n}$ ja $\widehat{n} \leq n$. Siispä näissä on yhtä monta vektoria.

Vektoriavaruuden V *dimensio*, $\dim(V)$, on V:n kantavektoreiden lukumäärä.

Esimerkki 1.18. Joukko $\{1, x, x^2, \ldots, x^n\}$ on polynomiavaruuden $\mathbb{P}_n$ kanta, joten $\dim(\mathbb{P}_n) = n+1$. Vastaavasti $m \times n$-matriisien muodostaman vektoriavaruuden dimensio (harjoitustehtävä) on $\dim(\mathbb{R}^{m\times n}) = m\,n$. Mikä olisi tämän matriisien avaruuden “luonnollinen kanta” ?

Esimerkki 1.19. Avaruuden $\mathbb{C}^n$ dimensio riippuu siitä, tarkastellaanko sitä reaali- vai kompleksikertoimisena avaruutena. Jos $\mathbb{K} = \mathbb{C}$, niin $\dim(\mathbb{C}^n) = n$. Jos $\mathbb{K} = \mathbb{R}$, niin $\dim(\mathbb{C}^n) = 2n$. Mikä olisi jälkimmäisessä tapauksessa “luonnollinen kanta” ? Ellei erikseen muuta mainita, niin $\mathbb{C}^n$ ja $\mathbb{C}^{m\times n}$ tulkitaan aina kompleksikertoimisiksi avaruuksiksi.

Näissä esimerkeissä avaruuden dimensio on äärellinen luku; ne ovat äärellisulotteisia. Tyypillisesti monet funktioavaruudet eivät ole äärellisulotteisia. Näistä löytyy mielivaltaisen monta lineaarisesti riippumatonta alkiota.

Sanotaan, että vektoriavaruus V on *ääretönulotteinen*, $\dim(V) = \infty$, jos on olemassa vektorijono $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots\} \subset V$ siten, että joukko $\{\boldsymbol{v}^1, \dots, \boldsymbol{v}^n\}$ on lineaarisesti riippumaton kaikilla $n \in \mathbb{N}$.

Esimerkki 1.20. Esimerkin 1.5 vektoriavaruus eli kaikkien rajoitettujen lukujonojen joukko ℓ_∞ on ääretönulotteinen: vektorijono $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots\}$, jossa $\boldsymbol{v}^j = (\delta_{j1}, \delta_{j2}, \dots)$, toteuttaa selvästi ehdon: $\{\boldsymbol{v}^1, \dots, \boldsymbol{v}^n\}$ on lineaarisesti riippumaton kaikilla $n \in \mathbb{N}$.

Kaikkien polynomien muodostama avaruus $\mathbb{P}_\infty$ ja kaikkien välillä $[a, b]$ jatkuvien funktioiden vektoriavaruus $C[a, b]$ ovat myös ääretönulotteisia.

Lemma 1.3. *Olkoon $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k\}$ vektoriavaruuden V lineaarisesti riippumaton osajoukko. Jos V:n vektori $\boldsymbol{v}^{k+1} \notin \mathrm{sp}(\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k)$, niin joukko $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k, \boldsymbol{v}^{k+1}\}$ on lineaarisesti riippumaton.*

Tod. Osoitetaan, että yhtälöllä

$$c_1 \boldsymbol{v}^1 + c_2 \boldsymbol{v}^2 + \dots + c_k \boldsymbol{v}^k + c_{k+1} \boldsymbol{v}^{k+1} = 0$$

on vain triviaaliratkaisu. Koska $\boldsymbol{v}^{k+1} \notin \mathrm{sp}(\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k)$, niin on oltava $c_{k+1} = 0$, sillä muuten $\boldsymbol{v}^{k+1}$ voitaisiin lausua muiden vektorien lineaarikombinaationa. Jäljelle jää

$$c_1 \boldsymbol{v}^1 + c_2 \boldsymbol{v}^2 + \dots + c_k \boldsymbol{v}^k = 0 \ .$$

Koska $\{\boldsymbol{v}^1, \dots, \boldsymbol{v}^k\}$ on lineaarisesti riippumaton, saadaan myös $c_1 = \dots = c_k = 0$. □

Lause 1.4. *Olkoon $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k\}$ äärellisulotteisen vektoriavaruuden V lineaarisesti riippumaton osajoukko. Tällöin on olemassa vektorit $\boldsymbol{v}^{k+1}, \dots, \boldsymbol{v}^{k+m} \in V$ siten, että $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k, \boldsymbol{v}^{k+1}, \dots, \boldsymbol{v}^{k+m}\}$ on V:n kanta.*

Todistus. Jos $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k\}$ virittää V:n, niin se on kanta ja lisävektoreita ei tarvita. Muuten löytyy $\boldsymbol{v}^{k+1} \in V$ siten, että sitä ei voi lausua vektoreiden $\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k$ lineaarikombinaationa. Edellisen lemman mukaan $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^k, \boldsymbol{v}^{k+1}\}$ on lineaarisesti riippumaton. Jos tämä joukko virittää koko V:n, niin on saatu kanta. Jos ei, jatketaan lisäämällä vektoreita samaan tapaan yksi kerrallaan, kunnes niitä on $\dim(V)$:n verran. Tällöin V:lle on saatu kanta. □

Seuraus 1.5. *Jos $\dim(V) = n$ ja $\{\boldsymbol{v}^1, \dots, \boldsymbol{v}^n\} \subset V$ on lineaarisesti riippumaton, niin se on V:n kanta.*

Tod. Jos se ei olisi kanta, niin lauseen 1.4 se voitaisiin täydentää kannaksi, jolloin V:hen saataisiin yli n kpl lineaarisesti riippumattomia vektoreita. Tämä olisi ristiriidassa lauseen 1.2 kanssa. □

Edellisen lauseen todistus tarjoaa menetelmän, jolla lineaarisesti riippumaton joukko voidaan täydentää kannaksi. Käänteinen tilanne on, jos tunnetaan V:n virittävä vektorijoukko, joka voi olla lineaarisesti riippuva.

Lause 1.6. *Olkoon $V = \mathrm{sp}(\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^n)$. Tällöin jokin joukon $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \dots, \boldsymbol{v}^n\}$ osajoukko on V:n kanta.*

Tod. Jos $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^n\}$ on lineaarisesti riippuva, niin lauseen 1.1 perusteella jokin vektori $\boldsymbol{v}^k$ voidaan lausua muiden lineaarikombinaationa. Tällöin (näytä)

$$\mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^{k-1}, \boldsymbol{v}^{k+1}, \ldots, \boldsymbol{v}^n) = \mathrm{sp}(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^n) \ .$$

Siis $\boldsymbol{v}^k$ voidaan pudottaa pois joukosta, ja jäljelle jäävä vektorijoukko virittää edelleen V:n. Näin jatketaan, kunnes jäljellä on lineaarisesti riippumaton joukko. Tämä on V:n kanta. $\square$

Kannan tärkein merkitys on seuraava.

Lause 1.7. *Jos $B = \{\boldsymbol{b}^1, \boldsymbol{b}^2, \ldots, \boldsymbol{b}^n\}$ on vektoriavaruuden V kanta, niin jokainen vektori $\boldsymbol{v} \in V$ voidaan esittää muodossa $\boldsymbol{v} = c_1\, \boldsymbol{b}^1 + c_2\, \boldsymbol{b}^2 + \cdots + c_n\, \boldsymbol{b}^n$ täsmälleen yhdellä tavalla.*

Tod. Koska kanta virittää V:n, jokainen $\boldsymbol{v} \in V$ voidaan esittää kantavektorien lineaarikombinaationa. Tarvitsee siis vain todistaa, että tämä esitys on yksikäsitteinen. Oletetaan, että $\boldsymbol{v}$:lle on esitykset:

$$\begin{aligned} \boldsymbol{v} &= c_1\, \boldsymbol{b}^1 + c_2\, \boldsymbol{b}^2 + \cdots + c_n\, \boldsymbol{b}^n \quad \text{ja} \\ \boldsymbol{v} &= d_1\, \boldsymbol{b}^1 + d_2\, \boldsymbol{b}^2 + \cdots + d_n\, \boldsymbol{b}^n \end{aligned}$$

Vähentämällä nämä yhtälöt puolittain saadaan

$$(c_1 - d_1)\, \boldsymbol{b}^1 + (c_2 - d_2)\, \boldsymbol{b}^2 + \cdots + (c_n - d_n)\, \boldsymbol{b}^n = 0 \ .$$

Koska $\boldsymbol{b}^i$:t ovat lineaarisesti riippumattomat, on oltava

$$c_1 = d_1,\ c_2 = d_2, \ldots, c_n = d_n \ ,$$

eli esitykset ovat samat. $\square$

Kun vektoriavaruuteen on valittu jokin kanta, niin lauseen mukaan jokaisen vektorin esitys kantavektoreiden lineaarikombinaationa on yksikäsitteinen. Siten vektoriin voidaan yksikäsitteisellä tavalla viitata käyttämällä sen esittämiseen vain lineaarikombinaation c_i-kertoimia. Näitä kutsutaan sen *koordinaateiksi* ko. kannan suhteen.

Olkoon $B = \{\boldsymbol{b}^1, \boldsymbol{b}^2, \ldots, \boldsymbol{b}^n\}$ vektoriavaruuden V (järjestetty) kanta. Jos vektorin $\boldsymbol{v} \in V$ esitys kannassa B on $\boldsymbol{v} = c_1\, \boldsymbol{b}^1 + c_2\, \boldsymbol{b}^2 + \cdots + c_n\, \boldsymbol{b}^n$, niin $[\boldsymbol{v}]_B = (c_1, c_2, \ldots, c_n)$ on $\boldsymbol{v}$:n *koordinaattivektori kannan B suhteen.*

Esimerkki 1.21. Olkoon vektorin $\boldsymbol{v} \in \mathbb{R}^3$ koordinaatit luonnollisessa kannassa $(5, -1, 9)$. Vektorit $\boldsymbol{b}^1 = (1, 2, 1)$, $\boldsymbol{b}^2 = (2, 9, 0)$ ja $\boldsymbol{b}^3 = (3, 3, 4)$ muodostavat myös $\mathbb{R}^3$:n kannan. Ne ovat lineaarisesti riippumattomia, sillä yhtälössä (1.2) esiintyvä kerroinmatriisi saa muodon

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix} \ .$$

Tämän determinantti on -1, joten yhtälöllä (1.2) on vain triviaaliratkaisu. Vektorijoukko myös virittää koko $\mathbb{R}^3$:n, sillä $\dim(\mathbb{R}^3) = 3$. Haluttaisiin laskea $\boldsymbol{v}$:n koordinaatit

(c_1, c_2, c_3) kannassa $B = \{\boldsymbol{b}^1, \boldsymbol{b}^2, \boldsymbol{b}^3\}$. Saadaan yhtälö

$$\begin{bmatrix} 5 \\ -1 \\ 9 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 9 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 3 \\ 3 \\ 4 \end{bmatrix} ,$$

matriisimuodossa

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 9 \end{bmatrix} .$$

Tämän ratkaisusta saadaan koordinaatit $[\boldsymbol{v}]_B = (1, -1, 2)$.

Esimerkki 1.22. Polynomijoukko $Q = \{x^2+1, 2x^2+x-1, x^2-x\}$ osoitettiin esimerkissä 1.17 $\mathbb{P}_2$:n kannaksi. Polynomin $p(x) = a+bx+cx^2$ koordinaatit tässä kannassa tuli myös ratkaistua: $[p]_Q = (\frac{3a+b+c}{4}, \frac{-a+b+c}{4}, \frac{-a-3b+c}{4})$.

Polynomiavaruudet ovat tärkeitä funktioiden approksimoinnissa ja numeerisessa analyysissä. Luonnollinen kanta ei välttämättä sovelluksen kannalta ole hyvä, vaan tehtävä kannattaa usein formuloida jossain muussa polynomikannassa laskennan yksinkertaistamiseksi.

Kannanvaihto. Tarkastellaan tilannetta, jossa tunnetaan vektorin esitys kannassa $B = \{\boldsymbol{b}^1, \boldsymbol{b}^2, \ldots, \boldsymbol{b}^n\}$ ja halutaan vaihtaa toiseen kantaan $U = \{\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^n\}$. Lasketaan, miten uudet koordinaatit saadaan lausuttua vanhojen avulla.

Merkitään vektorin $\boldsymbol{v}$ koordinaatteja näissä kannoissa

$$[\boldsymbol{v}]_B = (\beta_1, \ldots, \beta_n) \qquad \text{ja} \qquad [\boldsymbol{v}]_U = (\eta_1, \ldots, \eta_n) .$$

Oletetaan, että vanhat kantavektorit $\boldsymbol{b}^j$ on lausuttu uusien kantavektoreiden $\boldsymbol{u}^i$ avulla

$$\boldsymbol{b}^j = \sum_{i=1}^{n} s_{ij}\, \boldsymbol{u}^i , \qquad j = 1, \ldots, n . \tag{1.4}$$

Tällöin saadaan

$$\boldsymbol{v} = \sum_{i=1}^{n} \eta_i\, \boldsymbol{u}^i = \sum_{j=1}^{n} \beta_j\, \boldsymbol{b}^j = \sum_{j=1}^{n} \beta_j \sum_{i=1}^{n} s_{ij}\, \boldsymbol{u}^i = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} s_{ij}\, \beta_j \Big) \boldsymbol{u}^i .$$

Koska vektorin koordinaatit (kannassa U) ovat yksikäsitteiset, on oltava

$$\eta_i = \sum_{j=1}^{n} s_{ij}\, \beta_j , \qquad i = 1, \ldots, n . \tag{1.5}$$

Merkitään $\boldsymbol{S} = \begin{bmatrix} s_{11} & \ldots & s_{1n} \\ \vdots & & \vdots \\ s_{n1} & \ldots & s_{nn} \end{bmatrix}$. Tällöin koordinaattien välinen yhtälö (1.5) voidaan kirjoittaa

$$[\boldsymbol{v}]_U = \boldsymbol{S}\, [\boldsymbol{v}]_B . \tag{1.6}$$

Siis: uudet koordinaatit saadaan matriisilla $\boldsymbol{S}$ kertomalla vanhoista, kun $\boldsymbol{S}$:n sarakkeina on vanhojen kantavektoreiden koordinaattivektorit uudessa kannassa.

Matriisia $\boldsymbol{S}$ kutsutaan *kannanvaihtomatriisiksi* ja se siis välittää yhtälön (1.6) mukaisesti koordinaattimuunnoksen. Kannanvaihtomatriisi on aina invertoituva (harjoitustehtävä) ja vanhat koordinaatit saadaan uusista kaavalla

$$[\boldsymbol{v}]_B = \boldsymbol{S}^{-1}\,[\boldsymbol{v}]_U \ .$$

Esimerkki 1.23. Esimerkissä 1.21 ...

Harjoitustehtäviä

1. Osoita, että $\dim(\mathbb{R}^{k\times n}) = kn$.
2. Näytä, että $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$ ja $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$.
3. Osoita, että $C[a,b]$ on ääretönulotteinen.
4. Mikä kannanvaihtomatriisi on esimerkissä 1.17 siirryttäessä luonnollisesta kannasta kantaan Q?

2. Lineaarikuvaukset

Monet tärkeät operaatiot, kuten derivointi, integrointi, Fourier-muunnos, avaruuden kierto ja peilaus, ovat kuvauksia, jotka säilyttävät vektoriavaruuden struktuurin. Tässä luvussa alamme tarkastella näiden nk. lineaarikuvausten ominaisuuksia. Myös koko monisteen loppuosa painottuu melkoisesti näiden tarkastelemiseen.

Olkoot U ja V $\mathbb{K}$-kertoimisia vektoriavaruuksia. Kuvaus $T:\ U \mapsto V$ on *lineaarikuvaus*, jos kaikilla $\boldsymbol{x}, \boldsymbol{y} \in U$ ja $\alpha, \beta \in \mathbb{K}$ pätee

$$T(\alpha\,\boldsymbol{x} + \beta\,\boldsymbol{y}) = \alpha\,T(\boldsymbol{x}) + \beta\,T(\boldsymbol{y})\ . \tag{2.1}$$

Tällöin käytetään myös merkintää $T\boldsymbol{x} = T(\boldsymbol{x})$ eli sulut jätetään pois, mikäli ei ole sekaannuksen vaaraa.

Esimerkki 2.1. Olkoon $T:\mathbb{R}^3 \mapsto \mathbb{R}^2$ määritelty kaavalla $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ eli T projisoi avaruuden vektorin x_1, x_2 -tasoon. T on lineaarikuvaus, sillä:

$$T(\alpha\,\boldsymbol{x} + \beta\,\boldsymbol{y}) = T\begin{bmatrix} \alpha\,x_1 + \beta\,y_1 \\ \alpha\,x_2 + \beta\,y_2 \\ \alpha\,x_3 + \beta\,y_3 \end{bmatrix} = \begin{bmatrix} \alpha\,x_1 + \beta\,y_1 \\ \alpha\,x_2 + \beta\,y_2 \end{bmatrix} = \alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \alpha\,T\boldsymbol{x} + \beta\,T\boldsymbol{y}\ .$$

Kuvaus $S: C[a,b] \mapsto \mathbb{R},\quad S(f) = \int_a^b f(x)dx$ eli funktion integrointi on myös lineaarikuvaus, sillä

$$S(\alpha\,f + \beta\,g) = \int_a^b (\alpha\,f(x) + \beta\,g(x))\,dx = \alpha\int_a^b f(x)\,dx + \beta\int_a^b g(x)\,dx = \alpha\,S(f) + \beta\,S(g)\ .$$

Derivointi $D(f) = f'$ on lineaarikuvaus $C^1[a,b] \mapsto C[a,b]$, missä $C^1[a,b]$:llä tarkoitetaan välillä $[a,b]$ jatkuvasti derivoituvien funktioiden vektoriavaruutta.
Samoin D määrittelee lineaarikuvauksen $\mathbb{P}_\infty \mapsto \mathbb{P}_\infty$ ja $\mathbb{P}_n \mapsto \mathbb{P}_{n-1}$.

Matriisille $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ määritellään kuvaus $L_{\boldsymbol{A}}: \mathbb{R}^n \mapsto \mathbb{R}^m$ (tai $L_{\boldsymbol{A}}: \mathbb{C}^n \mapsto \mathbb{C}^m$) kaavalla

$$L_{\boldsymbol{A}}(\boldsymbol{x}) = \boldsymbol{A}\,\boldsymbol{x}\,,\ \boldsymbol{x} \in \mathbb{R}^n\ .$$

Toisin sanoen kuvaus $L_{\boldsymbol{A}}$ on matriisilla $\boldsymbol{A}$ kertominen. Selvästi tämä on lineaarikuvaus.

Jatkossa tullaan todistamaan, että myös käänteinen pätee, eli jokainen äärellisdimensioisten avaruuksien välinen lineaarikuvaus voidaan esittää matriisin avulla, kun avaruuksiin U ja V on kiinnitetty kannat.

Käyttäen kaavaa (2.1) useaan kertaan saadaan, että lineaarikuvaukselle pätee

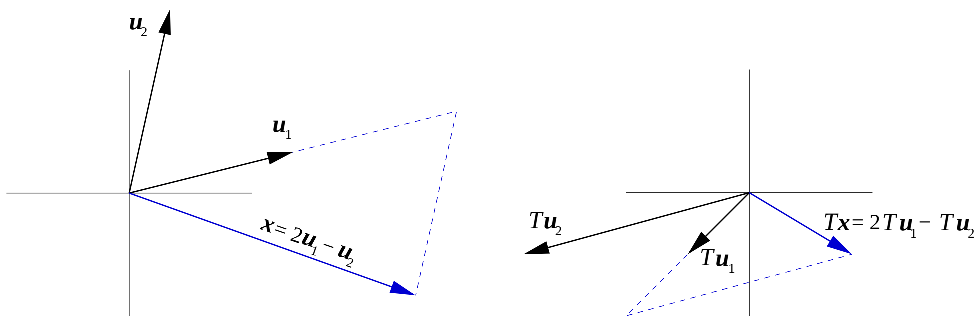
$$T\Big(\sum_{j=1}^{k} \alpha_j\,\boldsymbol{u}^j\Big) = \sum_{j=1}^{k} \alpha_j\,T\boldsymbol{u}^j\ .$$

Lause 2.1. *Oletetaan, että joukko* $\{\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n\}$ *virittää avaruuden* U *. Olkoot* T *ja* $\tilde{T}$ *kaksi lineaarikuvausta* $U \mapsto V$ *siten, että* $T\boldsymbol{u}^i = \tilde{T}\boldsymbol{u}^i$ *kaikilla* $i = 1, \ldots, n$ *. Tällöin* $T = \tilde{T}$ *.*

Tod. Koska $\{\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n\}$ virittää U:n, mielivaltainen vektori $\boldsymbol{v} \in U$ voidaan esittää lineaarikombinaationa $\boldsymbol{v} = \alpha_1 \boldsymbol{u}^1 + \cdots + \alpha_n \boldsymbol{u}^n$. Täten

$$\begin{aligned} T(\boldsymbol{v}) &= T(\alpha_1 \boldsymbol{u}^1 + \cdots + \alpha_n \boldsymbol{u}^n) = \alpha_1 T\boldsymbol{u}^1 + \cdots + \alpha_n T\boldsymbol{u}^n \\ &= \alpha_1 \tilde{T}\boldsymbol{u}^1 + \cdots + \alpha_n \tilde{T}\boldsymbol{u}^n = \tilde{T}(\alpha_1 \boldsymbol{u}^1 + \cdots + \alpha_n \boldsymbol{u}^n) = \tilde{T}(\boldsymbol{v}) \ . \end{aligned}$$

□



Lause siis sanoo, että jos tunnetaan miten T kuvaa jonkin U:n virittävän joukon vektorit, niin T on yksikäsitteisesti määrätty. Riittää siis tietää jonkin kannan vektoreiden kuvat.

Esimerkki 2.2. Olkoon $T:\ \mathbb{R}^2 \mapsto \mathbb{R}^2$ lineaarikuvaus siten, että $T\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}3\\2\end{smallmatrix}\right]$ ja $T\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right] = \left[\begin{smallmatrix}3\\0\end{smallmatrix}\right]$. Halutaan laskea $T\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$. Riittää lausua vektori $\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$ vektoreiden $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ ja $\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]$ avulla ja käyttää lineaarikuvauksen ominaisuuksia. Yhtälöllä

$$\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right] = c_1 \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] + c_2 \left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]$$

on ratkaisu $c_1 = \frac{3}{2}$ ja $c_2 = \frac{1}{2}$. Siten

$$T\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right] = \tfrac{3}{2} T\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] + \tfrac{1}{2} T\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right] = \tfrac{3}{2}\left[\begin{smallmatrix}3\\2\end{smallmatrix}\right] + \tfrac{1}{2}\left[\begin{smallmatrix}3\\0\end{smallmatrix}\right] = \left[\begin{smallmatrix}6\\3\end{smallmatrix}\right] \ .$$

2.1. **Lineaarikuvauksen matriisiesitys.**

Tarkastellaan aluksi lineaarikuvausta $T:\ \mathbb{R}^n \mapsto \mathbb{R}^m$. Olkoon $E_n = \{\boldsymbol{e}^1, \ldots, \boldsymbol{e}^n\}$ $\mathbb{R}^n$:n luonnollinen kanta ja $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ matriisi, jonka sarakkeina ovat vektorit $T\boldsymbol{e}^j$, $j = 1, \ldots, n$. Tällöin mielivaltaiselle vektorille $\boldsymbol{x} \in \mathbb{R}^n$ saadaan

$$\begin{aligned} \boldsymbol{A}\,\boldsymbol{x} &= [T\boldsymbol{e}^1\ T\boldsymbol{e}^2\ \ldots\ T\boldsymbol{e}^n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= x_1 T\boldsymbol{e}^1 + x_2 T\boldsymbol{e}^2 + \cdots + x_n T\boldsymbol{e}^n \\ &= T\,(x_1 \boldsymbol{e}^1 + x_2 \boldsymbol{e}^2 + \cdots + x_n \boldsymbol{e}^n) = T(\boldsymbol{x}) \ . \end{aligned}$$

Siten $T = L_{\boldsymbol{A}}$ eli T on lineaarikuvaus: “matriisilla $\boldsymbol{A}$ kertominen”.

Olkoon nyt U ja V kaksi äärellisulotteista vektoriavaruutta ja T lineaarikuvaus $U \mapsto V$. Olkoon $B_U = \{\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n\}$ U:n ja $B_V = \{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^m\}$ V:n kanta. Lausutaan jokaisen kantavektorin $\boldsymbol{u}^j$ kuva $T\boldsymbol{u}^j$ V:n kantavektorien avulla

$$T\boldsymbol{u}^j = \textstyle\sum_{i=1}^m a_{ij}\,\boldsymbol{v}^i\ , \quad j = 1, \ldots, n\,.$$

Olkoon vektorin $\boldsymbol{u} \in U$ koordinaatit $[\boldsymbol{u}]_{B_U} = (c_1, \ldots, c_n)$. Tällöin

$$T(\boldsymbol{u}) = T\Big(\sum_{j=1}^n c_j\,\boldsymbol{u}^j\Big) = \sum_{j=1}^n c_j\,T\boldsymbol{u}^j = \sum_{j=1}^n c_j \sum_{i=1}^m a_{ij}\,\boldsymbol{v}^i = \sum_{i=1}^m\Big(\sum_{j=1}^n a_{ij}\,c_j\Big)\,\boldsymbol{v}^i\ .$$

Siispä $T(\boldsymbol{u})$:n koordinaatit kannan B_V suhteen ovat luvut $\sum_{j=1}^n a_{ij}\,c_j$, $i = 1, \ldots, m$ eli

$$[T(\boldsymbol{u})]_{B_V} = \boldsymbol{A}\,[\boldsymbol{u}]_{B_U}\ .$$

Toisin sanoen: vektorin kuvan koordinaatit saadaan vektorin itsensä koordinaateista kertomalla matriisilla

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}\ .$$

Tämän matriisin sarakkeet muodostuvat lähtöavaruuden kantavektorien kuvien koordinaattivektoreista. Tätä matriisia kutsutaan *lineaarikuvauksen T matriisi*ksi kantojen B_U ja B_V suhteen.

On siis muistettava, että *lineaarikuvauksen matriisiesitys riippuu valituista kannoista.* Kun haluamme ilmaista eksplisiittisesti, minkä kantojen suhteen matriisi on määritelty, merkitään: $\boldsymbol{A} = [T]_{B_U,B_V}$. Kohdan 2.1 alussa esiintyneen lineaarikuvauksen $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ matriisi on siis tarkemmalla merkinnällä $\boldsymbol{A} = [T]_{E_n,E_m}$.

Esimerkki 2.3. Tarkastellaan lineaarikuvausta $T : \mathbb{R}^2 \mapsto \mathbb{R}^3$:

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -5x_1 + 13x_2 \\ -7x_1 + 16x_2 \end{bmatrix} \qquad \text{eli} \qquad [T]_{E_2,E_3} = \begin{bmatrix} 0 & 1 \\ -5 & 13 \\ -7 & 16 \end{bmatrix}\ .$$

Olkoon nyt $\mathbb{R}^2$:ssa valittu kanta $B_U = \{\boldsymbol{u}^1, \boldsymbol{u}^2\}$ ja $\mathbb{R}^3$:ssa kanta $B_V = \{\boldsymbol{v}^1, \boldsymbol{v}^2, \boldsymbol{v}^3\}$, missä

$$\boldsymbol{u}^1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix},\ \boldsymbol{u}^2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix},\ \boldsymbol{v}^1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix},\ \boldsymbol{v}^2 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix},\ \boldsymbol{v}^3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\ .$$

Lasketaan T:n matriisi näiden kantojen suhteen. Ensin lasketaan kantavektorien kuvat

$$T(\boldsymbol{u}^1) = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix},\quad T(\boldsymbol{u}^2) = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}\ ,$$

jotka lausutaan kannassa B_V. Lineaarisen yhtälöryhmän ratkaisemisen jälkeen saadaan

$$T(\boldsymbol{u}^1) = \boldsymbol{v}^1 - 2\boldsymbol{v}^3,\quad T(\boldsymbol{u}^2) = 3\boldsymbol{v}^1 + \boldsymbol{v}^2 - \boldsymbol{v}^3\ .$$

Näissä esiintyvät koordinaattivektorit $[T(\boldsymbol{u}^1)]_{B_V}$ ja $[T(\boldsymbol{u}^2)]_{B_V}$, asetetaan matriisin $\boldsymbol{A}$ sarakkeiksi, jolloin saadaan

$$\boldsymbol{A} = [T]_{B_U,B_V} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ -2 & -1 \end{bmatrix} .$$

Tässä uudet kannat eivät juuri antaneet lisäinfoa T:stä, olipahan vain tuollainen sormiharjoitus. Seuraava esimerkki on kiinnostavampi

Esimerkki 2.4. Olkoon $T: \mathbb{P}_3 \mapsto \mathbb{P}_3$ lineaarikuvaus $T(p)(x) = p(2-x)$. Olkoon $\mathbb{P}_3$:ssa kanta $B = \{p_1, p_2, p_3, p_4\} = \{1, x, x^2, x^3\}$ (nyt vain yksi kanta, koska $V = U$). Tässä siis kantavektorit ovat polynomeja $p_i(x)$. Saadaan

$$\begin{aligned} T(p_1)(x) &= 1 = p_1(x) \\ T(p_2)(x) &= 2 - x = (2p_1 - p_2)(x) \\ T(p_3)(x) &= (2-x)^2 = 4 - 4x + x^2 = (4p_1 - 4p_2 + p_3)(x) \\ T(p_4)(x) &= (2-x)^3 = 8 - 12x + 6x^2 - x^3 = (8p_1 - 12p_2 + 6p_3 - p_4)(x) \end{aligned}$$

$\mathbb{P}_3$:n polynomin, jonka kertoimet ovat $[p]_B = (c1, c2, c3, c4)$ kuva saadaan siten

$$\begin{aligned} &T(c_1\,p_1 + c_2\,p_2 + c_3\,p_3 + c_4\,p_4) \\ &\quad = c_1\,p_1 + c_2\,(2p_1 - p_2) + c_3\,(4p_1 - 4p_2 + p_3) + c_4\,(8p_1 - 12p_2 + 6p_3 - p_4) \\ &\quad = (c_1 + 2c_2 + 4c_3 + 8c_4)\,p_1 + (-c_2 - 4c_3 - 12c_4)\,p_2 + (c_3 + 6c_4)\,p_3 - c_4\,p_4 \ . \end{aligned}$$

Siispä T:n matriisiksi saadaan

$$[T]_B = \begin{bmatrix} 1 & 2 & 4 & 8 \\ 0 & -1 & -4 & -12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & -1 \end{bmatrix} .$$

Kun kuvaus on avaruudelta itselleen merkitään $[T]_{B,B}$:n sijasta lyhyemmin: $[T]_B$.

Esimerkki 2.5. Lasketaan vielä edellisen esimerkin lineaarikuvauksen matriisi kannan $\widehat{B} = \{\widehat{p}_1, \widehat{p}_2, \widehat{p}_3, \widehat{p}_4\} = \{1, 1-x, (1-x)^2, (1-x)^3\}$ suhteen. Tässä kannassa saadaan

$$\begin{aligned} T(\widehat{p}_1)(x) &= 1 = \widehat{p}_1(x) \\ T(\widehat{p}_2)(x) &= 1 - (2-x) = x - 1 = -\widehat{p}_2(x) \\ T(\widehat{p}_3)(x) &= (1-(2-x))^2 = (x-1)^2 = \widehat{p}_3(x) \\ T(\widehat{p}_4)(x) &= (1-(2-x))^3 = (x-1)^3 = -\widehat{p}_4(x) \ . \end{aligned}$$

Huomataan, että T:n matriisi tämän kannan suhteen on lävistäjämatriisi[5]

$$[T]_{\widehat{B}} = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} .$$

Tällaisia *diagonalisoivia kantoja* näemme lisää ominaisarvoteorian yhteydessä.

[5]Tässä, kuten usein myöhemminkin, matriisissa esiintyvät nollat jätetään kirjoittamatta.

Huomautus 2.1. Kaavan (1.6) kannanvaihtomatriisi voidaan tulkita identiteettikuvauksen[6] $I : V \mapsto \widehat{V}$ matriisiksi, kun V:ssä on kanta B ja $\widehat{V} = V$, mutta kantana on U. Lyhyesti sanottuna: $\boldsymbol{S} = [I]_{B,U}$.

Yhdistetty kuvaus.

Olkoot $T : U \mapsto V$ ja $S : V \mapsto W$ lineaarikuvauksia. Tällöin yhdistetty kuvaus $ST : U \mapsto W$ määritellään $(ST)(\boldsymbol{u}) = S(T(\boldsymbol{u}))$. Se on myös lineaarikuvaus.

Olkoon avaruuksissa U, V ja W kannat B_U, B_V ja B_W ja olkoot $\boldsymbol{A} = [T]_{B_U,B_V}$ ja $\boldsymbol{B} = [S]_{B_V,B_W}$ kuvausten T ja S matriisit näiden kantojen suhteen. Tämä tarkoittaa

$$[T(\boldsymbol{u})]_{B_V} = \boldsymbol{A}\,[\boldsymbol{u}]_{B_U} \qquad \text{ja} \qquad [S(\boldsymbol{v})]_{B_W} = \boldsymbol{B}\,[\boldsymbol{v}]_{B_V} \;,$$

missä $[\boldsymbol{u}]_{B_U}$ on $\boldsymbol{u}$:n koordinaattivektori kannassa B_U jne. Kun $\boldsymbol{u} \in U$, niin

$$[(ST)(\boldsymbol{u})]_{B_W} = [S(T(\boldsymbol{u}))]_{B_W} = \boldsymbol{B}\,[T(\boldsymbol{u})]_{B_V} = \boldsymbol{B}\,(\boldsymbol{A}\,[\boldsymbol{u}]_{B_U}) = (\boldsymbol{BA})[\boldsymbol{u}]_{B_U} \;.$$

Siispä lineaarikuvauksen ST matriisi (kantojen B_U ja B_W suhteen) on $\boldsymbol{BA}$. Näin saatiin: *yhdistetun kuvauksen matriisi saadaan kertomalla kuvausten matriisit keskenään,* eli kapulakielellä

$$[ST]_{B_U,B_W} = [S]_{B_V,B_W}[T]_{B_U,B_V} \;.$$

Kannanvaihto.

Olkoon avaruuksissa U ja V kannat B_U ja B_V ja olkoon $\boldsymbol{A} = [T]_{B_U,B_V}$ lineaarikuvauksen $T : U \mapsto V$ matriisi näiden suhteen. Olkoon U:ssa ja V:ssä myös uudet kannat $\widehat{B}_U$ ja $\widehat{B}_V$ ja olkoot $\boldsymbol{S}$ ja $\boldsymbol{R}$ kannanvaihtomatriisit eli kaikille $\boldsymbol{u} \in U$ ja $\boldsymbol{v} \in V$ pätee

$$[\boldsymbol{u}]_{\widehat{B}_U} = \boldsymbol{S}\,[\boldsymbol{u}]_{B_U} \qquad \text{ja} \qquad [\boldsymbol{v}]_{\widehat{B}_V} = \boldsymbol{R}\,[\boldsymbol{v}]_{B_V} \;. \tag{2.2}$$

Tällöin $[\boldsymbol{u}]_{B_U} = \boldsymbol{S}^{-1}\,[\boldsymbol{u}]_{\widehat{B}_U}$. Se, että $\boldsymbol{A}$ on lineaarikuvauksen T matriisi kantojen B_U ja B_V suhteen tarkoittaa:

$$[T(\boldsymbol{u})]_{B_V} = \boldsymbol{A}\,[\boldsymbol{u}]_{B_U} \qquad \text{kaikilla } \boldsymbol{u} \in U \;.$$

Yhdistetään tämä kaavoihin (2.2), jolloin saadaan

$$[T(\boldsymbol{u})]_{\widehat{B}_V} = \boldsymbol{R}\,[T(\boldsymbol{u})]_{B_V} = \boldsymbol{R}\,(\boldsymbol{A}\,[\boldsymbol{u}]_{B_U}) = \boldsymbol{R}\,(\boldsymbol{A}\,(\boldsymbol{S}^{-1}\,[\boldsymbol{u}]_{\widehat{B}_U})) = (\boldsymbol{RAS}^{-1})\,[\boldsymbol{u}]_{\widehat{B}_U}$$

Siispä T:n matriisi uusissa kannoissa saadaan:

$$\widehat{\boldsymbol{A}} = [T]_{\widehat{B}_U,\widehat{B}_V} = \boldsymbol{RAS}^{-1} \;.$$

Erityisesti, jos $V = U$ ja $B_V = B_U$, $\widehat{B}_V = \widehat{B}_U$, niin $\boldsymbol{R} = \boldsymbol{S}$ ja T:n matriisi uudessa kannassa saadaan

$$\widehat{\boldsymbol{A}} = \boldsymbol{SAS}^{-1} \;. \tag{2.3}$$

Tällaista kutsutaan matriisin $\boldsymbol{A}$ *similaarimuunnokseksi.* Tästä paljon lisää tuonnempana.

[6] $I(\boldsymbol{x}) = \boldsymbol{x}$ kaikilla $\boldsymbol{x}$

Esimerkki 2.6. Esimerkissä 2.4 lineaarikuvauksen $T:\ \mathbb{P}_3 \mapsto \mathbb{P}_3\,,\ T(p)(x) = p(2-x)$ matriisiksi kannan $B = \{p_1, p_2, p_3, p_4\} = \{1, x, x^2, x^3\}$ suhteen saatiin

$$\boldsymbol{A} = [T]_B = \begin{bmatrix} 1 & 2 & 4 & 8 \\ 0 & -1 & -4 & -12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & -1 \end{bmatrix} .$$

Olkoon $\mathbb{P}_3$:n uusi kanta $\widehat{B} = \{\widehat{p}_1, \widehat{p}_2, \widehat{p}_3, \widehat{p}_4\} = \{1, 1-x, (1-x)^2, (1-x)^3\}$. Pätee $\widehat{p}_j = \sum_{i=1}^4 r_{ij} p_i$, missä

$$\boldsymbol{R} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{bmatrix} .$$

Niinpä vanhat kantapolynomit saadaan uusista $p_j = \sum_{i=1}^4 s_{ij}\,\widehat{p}_i$, (ks. (1.4)), joten yhtälön (1.6):n mukaiseksi kannanvaihtomatriisiksi saadaan $\boldsymbol{S} = \boldsymbol{R}^{-1}$. Näin T:n matriisiksi uudessa kannassa tulee

$$\widehat{\boldsymbol{A}} = [T]_{\widehat{B}} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{S}^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 & 4 & 8 \\ 0 & -1 & -4 & -12 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} .$$

Tämähän on sama kuin esimerkissä 2.5 saatu matriisi, kuten pitääkin.

2.2. **Kuva-avaruus ja nolla-avaruus.**

Jokaiseen lineaarikuvaukseen liittyy kaksi aliavaruutta: kuva-avaruus ja nolla-avaruus, joiden ominaisuuksia ja yhteyksiä ryhdytään nyt tarkastelemaan.

Määritelmä 2.1. Olkoon T lineaarikuvaus $U \mapsto V$.

- T:n *nolla-avaruus* (eli ydin) on $N(T) = \{\boldsymbol{u} \in U \mid T\,\boldsymbol{u} = 0\} \subset U$.
- T:n *kuva-avaruus* on $R(T) = \{T\,\boldsymbol{u} \mid \boldsymbol{u} \in U\} \subset V$.

Kuva-avaruus (lyh. kuva) määritellään siis kuten funktion arvojoukko yleensäkin. Sitä merkitään myös TU:lla. Lineaarikuvauksen tapauksessa arvojoukko osoittautuu aliavaruudeksi.

Lause 2.2. *Olkoon $T : U \mapsto V$ lineaarikuvaus. Tällöin*

a) *$N(T)$ on U:n aliavaruus ja*
b) *$R(T)$ on V:n aliavaruus.*

Tod. a) Jos $\boldsymbol{u}^1, \boldsymbol{u}^2 \in N(T)$ ja $\alpha_1, \alpha_2 \in \mathbb{K}$, niin

$$T(\alpha_1\,\boldsymbol{u}^1 + \alpha_2\,\boldsymbol{u}^2) = \alpha_1\,T\boldsymbol{u}^1 + \alpha_2\,T\boldsymbol{u}^2 = 0\ ,$$

joten $\alpha_1\,\boldsymbol{u}^1 + \alpha_2\,\boldsymbol{u}^2 \in N(T)$.

b) Jos $\boldsymbol{v}^1, \boldsymbol{v}^2 \in R(T)$ ja $\beta_1, \beta_2 \in \mathbb{K}$, niin on olemassa $\boldsymbol{u}^1, \boldsymbol{u}^2 \in U$ siten, että $\boldsymbol{v}^1 = T\boldsymbol{u}^1$ ja $\boldsymbol{v}^2 = T\boldsymbol{u}^2$. Tällöin

$$T(\beta_1\,\boldsymbol{u}^1 + \beta_2\,\boldsymbol{u}^2) = \beta_1\,T\boldsymbol{u}^1 + \beta_2\,T\boldsymbol{u}^2 = \beta_1\,\boldsymbol{v}^1 + \beta_2\,\boldsymbol{v}^2\ ,$$

joten $\beta_1\,\boldsymbol{v}^1 + \beta_2\,\boldsymbol{v}^2 \in R(T)$. □

Lineaarikuvauksen ydin voidaan laskea ratkaisemalla yhtälö $T\boldsymbol{u} = 0$.

Jos $N(T) = \{0\}$, niin T on injektio, sillä

$$T\boldsymbol{u} = T\boldsymbol{v} \implies T(\boldsymbol{u} - \boldsymbol{v}) = 0 \implies \boldsymbol{u} - \boldsymbol{v} \in N(T) \implies \boldsymbol{u} = \boldsymbol{v} .$$

Jos $R(T) = V$, niin T on (määritelmän mukaan) surjektio. Jos kuvaus on sekä injektio että surjektio, niin se on bijektio ja sillä on olemassa käänteiskuvaus.

Tarkastellaan matriisiin $\boldsymbol{A} = [\boldsymbol{a}^1 \ \dots \ \boldsymbol{a}^n] \in \mathbb{R}^{m\times n}$ liittyvää lineaarikuvausta $L_{\boldsymbol{A}} :\ \mathbb{R}^n \mapsto \mathbb{R}^m :\ L_{\boldsymbol{A}}(\boldsymbol{x}) = \boldsymbol{A}\,\boldsymbol{x}$. Tämän kuva-avaruus on sarakkeiden viritelmä:

Lause 2.3. $R(L_{\boldsymbol{A}}) = \mathrm{sp}(\boldsymbol{a}^1, \dots, \boldsymbol{a}^n)$.

Tod. Tämä on selvä, kun kirjoitetaan auki:

$$\boldsymbol{A}\,\boldsymbol{x} = [\boldsymbol{a}^1 \ \dots \ \boldsymbol{a}^n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\,\boldsymbol{a}^1 + \cdots + x_n\,\boldsymbol{a}^n \ .$$

□

Esimerkki 2.7. Projektiokuvauksen $T : \mathbb{R}^3 \mapsto \mathbb{R}^2,\ T(\boldsymbol{x}) = (x_1, x_2)$ nolla-avaruus on selvästi

$$N(T) = \{\boldsymbol{x} = (0, 0, x_3) \mid x_3 \in \mathbb{R}\}$$

eli kantavektorin $\boldsymbol{e}^3$ suuntaiset vektorit. T:n kuva $R(T)$ taas on koko $\mathbb{R}^2$.

Määritelmä 2.2. Olkoon $T :\ U \mapsto V$ lineaarikuvaus. Määritellään

- T:n *nulliteetti*: $\nu(T) = \dim(N(T))$.
- T:n *rangi* $r(T) = \dim(R(T))$.

Nämä voivat olla äärellisiä tai äärettömiä.

Matriisilaskussa matriisin rangi määritellään $\boldsymbol{A}$:n sarakeavaruuden dimensioksi. Edellä olevan lauseen mukaan se on myös $\boldsymbol{A}$:han liittyvän lineaarikuvauksen $L_{\boldsymbol{A}}$ rangi.

Käsitteille melkoista väkivaltaa tehden on tullut tavaksi puhua matriisin $\boldsymbol{A}$ nolla-avaruudesta (ytimestä), kuva-avaruudesta, rangista, nulliteetista (ja myöhemmin ominaisarvoista ja -vektoreista), kun itse asiassa tarkoitetaan vastaavia $L_{\boldsymbol{A}}$:n juttuja. Käytäntö on osoittanut, että tästä ei pitemmän päälle ole suurta vaaraa, kunhan alussa ollaan hetki tarkkoja.

Vektoriavaruuksien ja lineaarikuvausten teorian keskeinen tulos on seuraava

Lause 2.4 (Lineaarialgebran peruslause)**.** *Olkoon U äärellisulotteinen ja $T :\ U \mapsto V$ lineaarikuvaus. Tällöin*

$$r(T) + \nu(T) = \dim(U) \ .$$

Tod. Merkitään $n = \dim(U)$, $m = \nu(T)$. Olkoon $\{\boldsymbol{u}^1, \dots, \boldsymbol{u}^m\}$ T:n ytimen kanta. Lauseen 1.4 mukaan se voidaan täydentää koko U:n kannaksi $\{\boldsymbol{u}^1, \dots, \boldsymbol{u}^m, \boldsymbol{u}^{m+1}, \dots, \boldsymbol{u}^n\}$. Väitetään, että $\{T(\boldsymbol{u}^{m+1}), \dots, T(\boldsymbol{u}^n)\}$ on $R(T)$:n kanta. Tällöin $r(T)$ olisi $n - m$ ja lause olisi todistettu.

Kannalle pitää näyttää viritys ja lineaarinen riippumattomuus.

Viritys:
Olkoon $\boldsymbol{v} \in R(T)$ mielivaltainen. Tälle on olemassa $\boldsymbol{u} \in U$ siten, että $\boldsymbol{v} = T\,\boldsymbol{u}$. Esitetään se U:n kannassa: $\boldsymbol{u} = \alpha_1\,\boldsymbol{u}^1 + \cdots + \alpha_n\,\boldsymbol{u}^n$. Tällöin

$$\begin{aligned}\boldsymbol{v} &= T(\alpha_1\,\boldsymbol{u}^1 + \cdots + \alpha_n\,\boldsymbol{u}^n)\\ &= \alpha_1\,T\boldsymbol{u}^1 + \cdots + \alpha_m\,T\boldsymbol{u}^m + \alpha_{m+1}\,T\boldsymbol{u}^{m+1} + \cdots + \alpha_n\,T\boldsymbol{u}^n\\ &= \alpha_{m+1}\,T\boldsymbol{u}^{m+1} + \cdots + \alpha_n\,T\boldsymbol{u}^n\ .\end{aligned}$$

Siispä $R(T) = \mathrm{sp}(T(\boldsymbol{u}^{m+1}), \ldots, T(\boldsymbol{u}^n))$.

Lineaarinen riippumattomuus:
Lineaarisuudella ja koska $\{\boldsymbol{u}^1, \ldots, \boldsymbol{u}^m\}$ on $N(T)$:n kanta saadaan

$$\begin{aligned}&\alpha_{m+1}\,T\boldsymbol{u}^{m+1} + \cdots + \alpha_n\,T\boldsymbol{u}^n = 0\\ &\quad\Rightarrow\quad T(\alpha_{m+1}\,\boldsymbol{u}^{m+1} + \cdots + \alpha_n\,\boldsymbol{u}^n) = 0\\ &\quad\Rightarrow\quad \alpha_{m+1}\,\boldsymbol{u}^{m+1} + \cdots + \alpha_n\,\boldsymbol{u}^n \in N(T)\\ &\quad\Rightarrow\quad \exists\ \alpha_1, \ldots, \alpha_m \quad \text{s.e.} \quad \alpha_1\,\boldsymbol{u}^1 + \cdots + \alpha_m\,\boldsymbol{u}^m = \alpha_{m+1}\,\boldsymbol{u}^{m+1} + \cdots + \alpha_n\,\boldsymbol{u}^n\\ &\quad\Rightarrow\quad \alpha_1 = \cdots = \alpha_m = \alpha_{m+1} = \cdots = \alpha_n = 0\ ,\end{aligned}$$

sillä $\{\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n\}$ on lineaarisesti riippumaton. Siispä $\{T(\boldsymbol{u}^{m+1}), \ldots, T(\boldsymbol{u}^n)\}$ on lineaarisesti riippumaton.

□

Edellistä lausetta kutsutaan myös *dimensiolauseeksi.*

Esimerkki 2.8. Tarkastellaan derivointia $D(p) = p'$ lineaarikuvauksena $\mathbb{P}_n \mapsto \mathbb{P}_n$. Selvästi $N(D) = \mathrm{sp}(1)$ (vakiot) ja $R(D) = \mathbb{P}_{n-1}$. Näiden dimensioiden summa on $n+1$ niin kuin pitääkin.

Esimerkki 2.9. Olkoon $\boldsymbol{A} = \begin{bmatrix} 1 & 3 & -1\\ -1 & 2 & 0\\ 1 & 8 & -2\end{bmatrix}$. Lasketaan $L_{\boldsymbol{A}}$:n ($\sim \boldsymbol{A}$:n) ydin ja kuva-avaruus sekä niille kannat. Gaussin eliminointi kulkee askelet:

$$\begin{bmatrix} 1 & 3 & -1\\ -1 & 2 & 0\\ -1 & 7 & -1\end{bmatrix} \to \begin{bmatrix} 1 & 3 & -1\\ 0 & 5 & -1\\ 0 & 10 & -2\end{bmatrix} \to \begin{bmatrix} 1 & 3 & -1\\ 0 & 5 & -1\\ 0 & 0 & 0\end{bmatrix}\ ,$$

josta nähdään: $\boldsymbol{A}\,\boldsymbol{x} = 0 \ \Leftrightarrow\ \boldsymbol{x} = \alpha \left[\begin{smallmatrix} 2\\ 1\\ 5\end{smallmatrix}\right]$ eli $N(L_{\boldsymbol{A}}) = \mathrm{sp}(\left[\begin{smallmatrix} 2\\ 1\\ 5\end{smallmatrix}\right])$. Siis $\nu(L_{\boldsymbol{A}}) = 1$.

Lauseen 2.3 perusteella

$$R(L_{\boldsymbol{A}}) = \mathrm{sp}\left(\left[\begin{smallmatrix} 1\\ -1\\ 1\end{smallmatrix}\right], \left[\begin{smallmatrix} 3\\ 2\\ 8\end{smallmatrix}\right], \left[\begin{smallmatrix} -1\\ 0\\ -2\end{smallmatrix}\right]\right).$$

Nämä vektorit eivät kuitenkaan muodosta kantaa, sillä $\dim(R(L_{\boldsymbol{A}})) = r(L_{\boldsymbol{A}}) = 3-1 = 2$, joten ne ovat lineaarisesti riippuvat. Kanta tästä saadaan, kun poimitaan näistä mitkä tahansa kaksi lineaarisesti riippumatonta, esim. $\left[\begin{smallmatrix} 1\\ -1\\ 1\end{smallmatrix}\right], \left[\begin{smallmatrix} -1\\ 0\\ -2\end{smallmatrix}\right]$.

Lopuksi vielä yksinkertainen, mutta tärkeä tulos

Lause 2.5. *Olkoon* $T : U \mapsto V$ *lineaarikuvaus. Tällöin*

a) *Jos vektorit* $\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n$ *ovat lineaarisesti riippuvat, niin* $T\boldsymbol{u}^1, \ldots, T\boldsymbol{u}^n$ *ovat myös.*

b) *Jos* $\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n$ *ovat lineaarisesti riippumattomat ja* $N(T) = \{0\}$ *(eli* T *on injektio), niin* $T\boldsymbol{u}^1, \ldots, T\boldsymbol{u}^n$ *ovat lineaarisesti riippumattomat.*

Tod. a) Jos $c_1\,\boldsymbol{u}^1 + \cdots + c_n\,\boldsymbol{u}^n = 0$ ja jokin kerroin on $\neq 0$, niin samoin

$$c_1\,T\boldsymbol{u}^1 + \cdots + c_n\,T\boldsymbol{u}^n = T(c_1\,\boldsymbol{u}^1 + \cdots + c_n\,\boldsymbol{u}^n) = 0 \ .$$

b) Olkoot $\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n$ lineaarisesti riippumattomat ja $N(T) = \{0\}$. Jos $c_1\,T\boldsymbol{u}^1 + \cdots + c_n\,T\boldsymbol{u}^n = 0$, niin $c_1\,\boldsymbol{u}^1 + \cdots + c_n\,\boldsymbol{u}^n \in N(T)$, joten $c_1\,\boldsymbol{u}^1 + \cdots + c_n\,\boldsymbol{u}^n = 0$ ja lineaarisesta riippumattomuudesta $c_1 = \cdots = c_n = 0$. □

Tehtävä 2.1. Olkoon $\dim(U) = \dim(V) < \infty$ ja $T : U \mapsto V$ lineaarinen injektio ($N(T) = \{0\}$). Näytä, että on olemassa lineaarinen $S : V \mapsto U$ siten, että

$$S(T(\boldsymbol{u})) = \boldsymbol{u} \qquad \forall\ \boldsymbol{u} \in U \qquad \text{ja} \qquad T(S(\boldsymbol{v})) = \boldsymbol{v} \qquad \forall\ \boldsymbol{v} \in V \ .$$

Tämä on T:n käänteiskuvaus ja merkitään $T^{-1} = S$.

3. Normi ja sisätulo

Vektoriavaruuden määritelmässä riitti olettaa, että joukon alkioille on määritelty aksioomat toteuttavat yhteenlasku ja skalaarilla kertominen. Kuitenkin monissa vektoriavaruuksissa voidaan tunnetusti tehdä muitakin laskutoimituksia. Esim. $\mathbb{R}^2$:ssa tai $\mathbb{R}^3$:ssa voidaan laskea vektoreiden pituuksia, välisiä kulmia ja pistetuloja. Jatkuvia funktioita voidaan kertoa keskenään, integroida, niiden maksimeja voi etsiä jne.

Normiavaruus on sellainen vektoriavaruus, jossa vektoreille on määritelty pituusfunktio, jota kutsutaan normiksi. Sisätuloavaruus on puolestaan normiavaruus, jossa lisäksi kulmien mittaaminen on mahdollista ja erityisesti kohtisuoruus eli ortogonaalisuus on määritelty. Seuraavassa tarkastellaan lähemmin, miten tällaisia pituus- ja kulmafunktioita voidaan määritellä.

Määritelmä 3.1. Olkoon V $\mathbb{K}$-kertoiminen vektoriavaruus. Kuvaus $\|\cdot\| : V \mapsto \mathbb{R}$ on *normi*, jos se toteuttaa

$$\text{(1)} \qquad \|\boldsymbol{v}\| \geq 0 \quad \forall\, \boldsymbol{v} \in V\,.$$

$$\text{(2)} \qquad \|\boldsymbol{v}\| = 0 \implies \boldsymbol{v} = 0\,.$$

$$\text{(3)} \qquad \|\boldsymbol{u}+\boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\| \quad \forall\, \boldsymbol{u}, \boldsymbol{v} \in V\,.$$

$$\text{(4)} \qquad \|\alpha\, \boldsymbol{v}\| = |\alpha|\; \|\boldsymbol{v}\| \quad \forall\, \alpha \in \mathbb{K},\; \boldsymbol{v} \in V\,.$$

Vektoriavaruutta, jossa on määritelty jokin normi kutsutaan normiavaruudeksi.

Esimerkki 3.1. Vektoriavaruudessa $\mathbb{R}^n$ tavallisin normi on nk. euklidinen normi

$$\|\boldsymbol{x}\|_2 = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}}\,.$$

Selvästi tämä toteuttaa ehdot (1), (2) ja (4). Ominaisuuden (3) eli *kolmioepäyhtälön* näytämme hieman myöhemmin. Muita usein käytettyjä normeja $\mathbb{R}^n$:ssä ovat[7]

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i| \qquad \text{ja} \qquad \|\boldsymbol{x}\|_\infty = \max_{1\leq i\leq n} |x_i|\,.$$

Näistä on helppo näyttää ominaisuudet (1)-(4). Ellei toisin mainita, käytetään $\mathbb{R}^n$:ssä normia $\|\cdot\| = \|\cdot\|_2$.

Avaruudessa $\mathbb{C}^n$ käytetään myös aivan samalla tavalla määriteltyjä normeja.

Esimerkki 3.2. Avaruuteen ℓ_∞ (rajoitetut reaalilukujonot) normiksi sopii hyvin[8]

$$\|\boldsymbol{x}\|_\infty = \sup_{1\leq i<\infty} |x_i|\,.$$

[7] Normia $\|\cdot\|_1$ kutsutaan taksikuskin normiksi. Miksiköhän?

[8] Huomaa, että tässä tarvitaan sup, sillä jonon alkioiden itseisarvoilla ei välttämättä ole maksimia.

Esimerkki 3.3. Avaruudessa $C[a,b]$ voidaan määritellä esimerkiksi

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)| \ , \qquad \|f\|_1 = \int_a^b |f(x)|\, dx \qquad \text{ja} \qquad \|f\|_2 = \left(\int_a^b |f(x)|^2\, dx\right)^{\frac{1}{2}} .$$

Näistä jälleen kaksi ensimmäistä on helppo osoittaa toteuttavan normin ehdot. Näin on myös $\|f\|_2$:n kohdalla muut paitsi kolmioepäyhtälö, johon palaamme pian.

Normiavaruudessa voidaan mitata vektoreiden pituuksia. Normin avulla myös määritellään alkioiden välinen *etäisyys*: $d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\|$. Kun se on määritelty, voidaan puhua vektorijonojen *suppenemisesta*: sanotaan, että vektorijono $\{\boldsymbol{x}^k\}_{k=1}^\infty \subset V$ suppenee kohti vektoria $\boldsymbol{x} \in V$, jos reaalilukujono $\{\|\boldsymbol{x}^k - \boldsymbol{x}\|\}_{k=1}^\infty$ menee nollaan eli

$$\lim_{k\to\infty} \|\boldsymbol{x}^k - \boldsymbol{x}\| = 0 \ .$$

Edelleen normiavaruuksien välisten kuvausten *jatkuvuus* määritellään aivan kuten aiemmin: $F: U \mapsto V$ on jatkuva pisteessä $\boldsymbol{u}^0 \in U$, jos

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \ \text{ s.e. } \ \|\boldsymbol{u} - \boldsymbol{u}^0\|_U < \delta \implies \|F(\boldsymbol{u}) - F(\boldsymbol{u}^0)\|_V < \varepsilon \ .$$

Näin jatkaen, kun on määritelty normi, koko analyysin käsitteistö putoaa syliin ja moni tuttu asia saa yhä vahvemman ja yleisemmän merkityksen.

Tässä on kuitenkin tarkoitus jatkaa vektoriavaruuden erilaisten yleisten rakenteiden selvittelyä. Annetaan analyysin odottaa monisteen loppupuolelle ja siirrytään tarkastelemaan sisätuloa, joka myös määrittelee normin.

Määritelmä 3.2. Olkoon V $\mathbb{K}$-kertoiminen vektoriavaruus. Kuvaus $\langle\,\cdot\,,\,\cdot\,\rangle : V\times V \mapsto \mathbb{K}$ on *sisätulo*, jos se toteuttaa ehdot

(1) $\langle \boldsymbol{v},\boldsymbol{v}\rangle \ge 0$ kaikilla $\boldsymbol{v} \in V$.
(2) $\langle \boldsymbol{v},\boldsymbol{v}\rangle = 0 \implies \boldsymbol{v} = 0$.
(3) $\langle \boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle = \langle \boldsymbol{u},\boldsymbol{w}\rangle + \langle \boldsymbol{v},\boldsymbol{w}\rangle$ kaikilla $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in V$.
(4) $\langle \alpha\boldsymbol{u},\boldsymbol{v}\rangle = \alpha\,\langle \boldsymbol{u},\boldsymbol{v}\rangle$ kaikilla $\alpha \in \mathbb{K}$, $\boldsymbol{u},\boldsymbol{v} \in V$.
(5) $\langle \boldsymbol{v},\boldsymbol{u}\rangle = \overline{\langle \boldsymbol{u},\boldsymbol{v}\rangle}$ kaikilla $\boldsymbol{u},\boldsymbol{v} \in V$.

Sisätulolla varustettua vektoriavaruutta sanotaan sisätuloavaruudeksi.

Reaalisessa tapauksessa (5) saa muodon $\langle \boldsymbol{v},\boldsymbol{u}\rangle = \langle \boldsymbol{u},\boldsymbol{v}\rangle$ eli reaalinen sisätulo on symmetrinen. Ominaisuudet (3) ja (4) sanovat, että sisätulo on lineaarinen ensimmäisen argumentin suhteen. Toisen argumentin suhteen saadaan:

$$\begin{aligned}\langle \boldsymbol{u}\,,\,\alpha\,\boldsymbol{v} + \beta\,\boldsymbol{w}\rangle &\overset{(5)}{=} \overline{\langle \alpha\,\boldsymbol{v} + \beta\,\boldsymbol{w}\,,\,\boldsymbol{u}\rangle} \overset{(3),(4)}{=} \overline{\alpha\,\langle \boldsymbol{v}\,,\,\boldsymbol{u}\rangle + \beta\,\langle \boldsymbol{w}\,,\,\boldsymbol{u}\rangle} \\ &= \overline{\alpha}\,\overline{\langle \boldsymbol{v}\,,\,\boldsymbol{u}\rangle} + \overline{\beta}\,\overline{\langle \boldsymbol{w}\,,\,\boldsymbol{u}\rangle} \overset{(5)}{=} \overline{\alpha}\,\langle \boldsymbol{u}\,,\,\boldsymbol{v}\rangle + \overline{\beta}\,\langle \boldsymbol{u}\,,\,\boldsymbol{w}\rangle \ .\end{aligned} \tag{3.1}$$

Täten sisätulo on *konjugoidusti lineaarinen* toisen argumentin suhteen: skalaarit saadaan ulos kompleksikonjugaatteina. Reaalisessa tapauksessa sisätulo on siten lineaarinen myös toisen argumentin suhteen.

Vektoriavaruudesta $\mathbb{R}^n$ tuttu vektoreiden välinen pistetulo[9]: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \boldsymbol{y} = \sum_{i=1}^n x_i \, y_i$ toteuttaa sisätulon ehdot.
Vastaavasti $\mathbb{C}^n$:n vektoreille määritellään $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \, \overline{\boldsymbol{y}} = \sum_{i=1}^n x_i \, \overline{y_i}$.

Esimerkki 3.4. Avaruudessa $C[a,b]$ voidaan määritellä

$$\langle f, g \rangle = \int_a^b f(x) g(x) \, dx \; .$$

Ehdot (1)-(5) seuraavat suoraan integraalin ominaisuuksista.

Esimerkiksi $C[-\pi, \pi]$:ssä funktioiden $f(x) = \sin x$ ja $g(x) = \cos x$ väliset sisätulot ovat

$$\langle f, g \rangle = \int_{-\pi}^{\pi} \sin x \cos x \, dx = \int_{-\pi}^{\pi} \tfrac{1}{2} \sin 2x \, dx = 0$$

$$\langle f, f \rangle = \int_{-\pi}^{\pi} \sin^2 x \, dx = \int_{-\pi}^{\pi} \tfrac{1}{2}(1 - \cos 2x) \, dx = \pi \; .$$

Samoin $\langle g, g \rangle = \pi$.

Kuten normeja, avaruudessa voidaan määritellä myös useita sisätuloja.

Esimerkki 3.5. Olkoon $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ säännöllinen. Määritellään

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{A}} = (\boldsymbol{A}\boldsymbol{x})^T \boldsymbol{A}\boldsymbol{y} \; .$$

Toteamalla, että $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{A}} = \langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} \rangle$, ja käyttäen $\boldsymbol{A}$:n säännöllisyyttä nähdään, että ehdot (1)-(5) toteutuvat.

Sisätulon tärkeä ominaisuus on, että se määrittelee heti myös normin: jos V on sisätuloavaruus, asetetaan

$$\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \; . \tag{3.2}$$

Sisätulon ehdoista saadaan normin ehdot (1),(2) ja (4) helposti. (3) eli kolmioepäyhtälö vaatii hieman laskemista. Todistetaan ensin *Schwarzin epäyhtälö*[10]: sisätulo ja sen avulla kaavalla (3.2) määritelty $\|\cdot\|$ (jota vielä ei tiedetä normiksi) toteuttavat:

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \; . \tag{3.3}$$

Tod. Jos $\boldsymbol{v} = 0$, niin väite on triviaalisti totta. Oletetaan siis, että $\boldsymbol{v} \neq 0$. Mielivaltaisella $\alpha \in \mathbb{K}$ pätee

$$0 \leq \langle \boldsymbol{u} - \alpha \, \boldsymbol{v} \, , \, \boldsymbol{u} - \alpha \, \boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{u} \rangle - \overline{\alpha} \, \langle \boldsymbol{u}, \boldsymbol{v} \rangle - \alpha \, \langle \boldsymbol{v}, \boldsymbol{u} \rangle + \alpha \, \overline{\alpha} \, \langle \boldsymbol{v}, \boldsymbol{v} \rangle \; .$$

Valitsemalla $\alpha = \langle \boldsymbol{u}, \boldsymbol{v} \rangle \, / \, \langle \boldsymbol{v}, \boldsymbol{v} \rangle$ saadaan

$$0 \leq \langle \boldsymbol{u}, \boldsymbol{u} \rangle - \frac{\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle} \, \langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} - \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle \, \overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} + \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle \, \overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle^2} \, \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{u} \rangle - \frac{|\langle \boldsymbol{u}, \boldsymbol{v} \rangle|^2}{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \; ,$$

mistä väite seuraa. □

---

[9]Lausekkeessa $\boldsymbol{x}^T \boldsymbol{y}$ vektorit on ajateltu $n \times 1$-matriiseiksi, jolloin $\boldsymbol{x}^T$ on $1 \times n$-matriisi ja $\boldsymbol{x}^T \boldsymbol{y}$ on 1×1-matriisi eli skalaari.

[10]Täydellisemmin: Cauchy-Schwarz-Bunjakovskin epäyhtälö.

Edellisestä todistuksesta nähdään, että epäyhtälö on aito, ellei $\boldsymbol{u} = \alpha\,\boldsymbol{v}$ eli ellei $\boldsymbol{u}$ ja $\boldsymbol{v}$ ole lineaarisesti riippuvia. Huomaa, että todistuksessa valittu α minimoi oikean puolen.

Näytetään nyt, että kaavalla (3.2) määritelty $\|\cdot\|$ toteuttaa normin ehdon (3) eli kolmioepäyhtälön

$$\|\boldsymbol{u}+\boldsymbol{v}\| \le \|\boldsymbol{u}\| + \|\boldsymbol{v}\| \ .$$

Tod. Käyttäen sisätulon ominaisuuksia ja Schwarzin epäyhtälöä saadaan

$$\begin{aligned}\|\boldsymbol{u}+\boldsymbol{v}\|^2 &= \langle \boldsymbol{u}+\boldsymbol{v}\,,\,\boldsymbol{u}+\boldsymbol{v}\rangle = \langle \boldsymbol{u},\boldsymbol{u}\rangle + \langle \boldsymbol{u},\boldsymbol{v}\rangle + \langle \boldsymbol{v},\boldsymbol{u}\rangle + \langle \boldsymbol{v},\boldsymbol{v}\rangle \\ &\le \|\boldsymbol{u}\|^2 + 2\,|\langle \boldsymbol{u},\boldsymbol{v}\rangle| + \|\boldsymbol{v}\|^2 \\ &\le \|\boldsymbol{u}\|^2 + 2\,\|\boldsymbol{u}\|\,\|\boldsymbol{v}\| + \|\boldsymbol{v}\|^2 = (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)^2 \ ,\end{aligned}$$

josta väite seuraa. □

Esimerkki 3.6. Kolmioepäyhtälö ja Schwarzin epäyhtälö ovat ei-triviaaleja. Esimerkiksi edellä $C[a,b]$:hen määriteltyyn sisätuloon sovellettuna Schwarz sanoo:

$$\left|\int_a^b f(x)g(x)\,dx\right| \le \left(\int_a^b |f(x)|^2\,dx\,\int_a^b |g(x)|^2\,dx\right)^{\frac{1}{2}}$$

ja asettamalla $f \leftarrow |f|$ ja $g(x) = 1$

$$\int_a^b |f(x)|\,dx \le \sqrt{b-a}\,\left(\int_a^b |f(x)|^2\,dx\right)^{\frac{1}{2}} \ .$$

Näitä olisi hankala todistaa muilla keinoin.

Tehtävä 3.1. Todista seuraavat epäyhtälöt mielivaltaisille $\boldsymbol{v}, \boldsymbol{u} \in \mathbb{C}^n$

1. $\|\boldsymbol{v}\|_\infty \le \|\boldsymbol{v}\|_2 \le \|\boldsymbol{v}\|_1$.
2. $\|\boldsymbol{v}\|_1 \le \sqrt{n}\,\|\boldsymbol{v}\|_2 \le n\,\|\boldsymbol{v}\|_\infty$.
3. $|\langle \boldsymbol{u},\boldsymbol{v}\rangle| \le \|\boldsymbol{u}\|_\infty \|\boldsymbol{v}\|_1$.

3.1. **Ortogonaalisuus.**

Sisätulon avulla voidaan reaalikertoimisessa avaruudessa määritellä vektoreiden väliset kulmat. Schwarzin epäyhtälön mukaan

$$-1 \le \frac{\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|} \le 1 \ ,$$

joten voidaan määritellä vektoreiden $\boldsymbol{u} \ne 0$ ja $\boldsymbol{v} \ne 0$ välinen kulma

$$\sphericalangle(\boldsymbol{u},\boldsymbol{v}) = \arccos\left(\frac{\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}\right) \ .$$

Vektorit $\boldsymbol{u}$ ja $\boldsymbol{v}$ ovat *ortogonaaliset*, kun $\sphericalangle(\boldsymbol{u},\boldsymbol{v}) = \pi/2$ eli kun $\langle \boldsymbol{u},\boldsymbol{v}\rangle = 0$. Ortogonaalisuus määritellään samoin kompleksikertoimisissa vektoriavaruuksissa. Täten $\left[\begin{smallmatrix}1\\ i\end{smallmatrix}\right]$ ja $\left[\begin{smallmatrix}i\\ 1\end{smallmatrix}\right]$ ovat ortogonaaliset $\mathbb{C}^2$:ssa.

Kun $\boldsymbol{u}$ ja $\boldsymbol{v}$ ovat ortogonaaliset, pätee Pythagoraan kaava:

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \langle \boldsymbol{u},\boldsymbol{u}\rangle + \langle \boldsymbol{u},\boldsymbol{v}\rangle + \langle \boldsymbol{v},\boldsymbol{u}\rangle + \langle \boldsymbol{v},\boldsymbol{v}\rangle = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 \ .$$

Sisätuloavaruuden vektorijoukkoa $S = \{\boldsymbol{v}^1, \dots, \boldsymbol{v}^k\}$ sanotaan ortogonaaliseksi, jos kaikki sen vektorit ovat keskenään ortogonaaliset: $\langle \boldsymbol{v}^i, \boldsymbol{v}^j\rangle = 0$, kun $i \neq j$.

Esimerkki 3.7. Tarkastellaan avaruutta $\mathbb{P}_2$ sisätulolla

$$\langle p,q\rangle = \int_{-1}^{1} p(x)q(x)\,dx \ .$$

Polynomit $p_1(x) = x$ ja $p_2(x) = x^2$ ovat ortogonaaliset, sillä

$$\langle p_1,p_2\rangle = \int_{-1}^{1} x\,x^2\,dx = 0 \ .$$

Samoin $p_0(x) = 1$ on ortogonaalinen p_1 :n kanssa. Sen sijaan $\langle p_0,p_2\rangle = \int_{-1}^{1} x^2\,dx = 2/3$. Olkoon $\widehat{p}_2 = p_2 + \tau\, p_0$. Selvästi $\langle p_1,\widehat{p}_2\rangle = 0$. Etsitään τ siten, että myös $\langle p_0,\widehat{p}_2\rangle = 0$:

$$\langle p_0,\widehat{p}_2\rangle = \int_{-1}^{1}(x^2+\tau)\,dx = 2/3 + 2\tau = 0 \qquad \Longrightarrow \qquad \tau = -1/3 \ .$$

Näin, kun $\widehat{p}_2(x) = x^2 - 1/3$, saadaan ortogonaalinen joukko $\{p_0, p_1, \widehat{p}_2\}$.

Ortogonaalinen vektorijoukko $\{\boldsymbol{v}^1, \dots, \boldsymbol{v}^n\}$ on myös lineaarisesti riippumaton edellyttäen, että se ei sisällä nollavektoria. Tämä nähdään seuraavasti. Jos $c_1\,\boldsymbol{v}^1 + \cdots + c_n\,\boldsymbol{v}^n = 0$, otetaan tämän sisätulo $\boldsymbol{v}^k$:n kanssa, jolloin

$$0 = \left\langle c_1\,\boldsymbol{v}^1 + \cdots + c_n\,\boldsymbol{v}^n\ ,\ \boldsymbol{v}^k\right\rangle = c_1\left\langle \boldsymbol{v}^1, \boldsymbol{v}^k\right\rangle + \cdots + c_k\left\langle \boldsymbol{v}^k, \boldsymbol{v}^k\right\rangle + \cdots + c_n\left\langle \boldsymbol{v}^n, \boldsymbol{v}^k\right\rangle = c_k\,\|\boldsymbol{v}^k\|^2$$

ja koska $\boldsymbol{v}^k \neq 0$, saadaan $c_k = 0$. Näin kaikki kertoimet saadaan yksitellen nolliksi, joten $\{\boldsymbol{v}^1, \dots, \boldsymbol{v}^n\}$ on lineaarisesti riippumaton.

Jos ortogonaalisen joukon vektorit ovat lisäksi pituudeltaan ykkösiä kutsutaan joukkoa *ortonormaaliksi.*

Esimerkki 3.8. Lasketaan edellisen esimerkin ortogonaalisten polynomien normit:

$$\begin{aligned} \|p_0\|^2 &= \int_{-1}^{1} 1\,dx = 2 \\ \|p_1\|^2 &= \int_{-1}^{1} x^2\,dx = \tfrac{2}{3} \\ \|\widehat{p}_2\|^2 &= \int_{-1}^{1} (x^2 - \tfrac{1}{3})^2\,dx = \tfrac{8}{45} \ . \end{aligned}$$

Täten $\{\frac{1}{\sqrt{2}},\ \sqrt{\frac{3}{2}}\,x,\ \frac{3}{2}\sqrt{\frac{5}{2}}\,(x^2 - \frac{1}{3})\}$ on ortonormaali joukko. Se on lineaarisesti riippumaton ja, koska $\dim(\mathbb{P}_2) = 3$, se on $\mathbb{P}_2$:n *ortonormaali kanta.*

Olkoon $\{\boldsymbol{b}^1, \dots, \boldsymbol{b}^n\}$ jokin $\mathbb{R}^n$:n ortonormaali kanta. Asetetaan $\boldsymbol{B} = [\boldsymbol{b}^1\ \dots\ \boldsymbol{b}^n]$, jolloin

$$\boldsymbol{B}^T\boldsymbol{B} = \begin{bmatrix} (\boldsymbol{b}^1)^T \\ \vdots \\ (\boldsymbol{b}^n)^T \end{bmatrix} [\boldsymbol{b}^1\ \dots\ \boldsymbol{b}^n] = \begin{bmatrix} \left\langle \boldsymbol{b}^1, \boldsymbol{b}^1\right\rangle & \dots & \left\langle \boldsymbol{b}^1, \boldsymbol{b}^n\right\rangle \\ \vdots & & \vdots \\ \left\langle \boldsymbol{b}^n, \boldsymbol{b}^1\right\rangle & \dots & \left\langle \boldsymbol{b}^n, \boldsymbol{b}^n\right\rangle \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} = \boldsymbol{I} \ ,$$

eli $\boldsymbol{B}$ on ortogonaalinen matriisi: $\boldsymbol{B}^T = \boldsymbol{B}^{-1}$. Kääntäen: ortogonaalisen matriisin sarakkeet muodostavat $\mathbb{R}^n$:n ortonormaalin kannan.

Samoin, jos matriisin $\boldsymbol{Q} \in \mathbb{R}^{m \times n}$ sarakkeet ovat ortonormaalit (jolloin välttämättä $m \geq n$), saadaan $\boldsymbol{Q}^T \boldsymbol{Q} = \boldsymbol{I}$. Jos $m > n$, niin $\boldsymbol{Q}$ ei kuitenkaan ole invertoituva; sillä on vain *vasemmanpuoleinen inverssi.*

Olkoon $\boldsymbol{U}$ reaalinen tai kompleksinen matriisi, jonka sarakkeet ovat ortonormaalit Tällöin[11] $\boldsymbol{U}^* \boldsymbol{U} = \boldsymbol{I}$ ja

$$\langle \boldsymbol{U x}\,,\, \boldsymbol{U y} \rangle = (\boldsymbol{U y})^* \boldsymbol{U x} = \boldsymbol{y}^* \boldsymbol{U}^* \boldsymbol{U x} = \boldsymbol{y}^* \boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \ .$$

Erityisesti: unitaarisella (reaalisessa tapauksessa ortogonaalisella) matriisilla kerrottaessa vektoreiden pituudet ja niiden väliset sisätulot säilyvät.

Annetun vektorin koordinaatit ortonormaalin kannan suhteen on helppo laskea:

Olkoon $B = \{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^n\}$ sisätuloavaruuden V ortonormaali kanta. Jos $\boldsymbol{v} = c_1 \boldsymbol{b}^1 + \cdots + c_n \boldsymbol{b}^n$, otetaan tämän sisätulo $\boldsymbol{b}^k$:n kanssa, jolloin $\left\langle \boldsymbol{v}, \boldsymbol{b}^k \right\rangle = c_k \left\langle \boldsymbol{b}^k, \boldsymbol{b}^k \right\rangle = c_k$. Näin saadaan kaikki kertoimet. Siis esitys ortonormaalissa kannassa saadaan:

$$\boldsymbol{v} = \sum_{k=1}^{n} \left\langle \boldsymbol{v}, \boldsymbol{b}^k \right\rangle \boldsymbol{b}^k \ , \qquad \text{kaikilla } \boldsymbol{v} \in V \, .$$

Ortonormaaleja kantoja voidaan muodostaa nk. Gram–Schmidtin prosessilla. Olkoon $(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots)$ (äärellinen tai ääretön) jono lineaarisesti riippumattomia sisätuloavaruuden vektoreita. Muodostetaan yhtä pitkä jono $(\boldsymbol{q}^1, \boldsymbol{q}^2, \ldots)$ ortonormaaleja vektoreita seuraavasti:

$$\begin{aligned} \boldsymbol{q}^1 &= \boldsymbol{v}^1 / \|\boldsymbol{v}^1\| \ , \\ \left.\begin{aligned} \boldsymbol{w}^k &= \boldsymbol{v}^k - \textstyle\sum_{j=1}^{k-1} \left\langle \boldsymbol{v}^k, \boldsymbol{q}^j \right\rangle \boldsymbol{q}^j \ , \\ \boldsymbol{q}^k &= \boldsymbol{w}^k / \|\boldsymbol{w}^k\| \ . \end{aligned}\right\} &\quad k = 2, 3, \ldots \end{aligned} \tag{3.4}$$

Tässä keskimmäisellä rivillä $\boldsymbol{v}^k$:sta poistetaan sen komponentit jo muodostetuilla suunnilla $\boldsymbol{q}^1, \ldots, \boldsymbol{q}^{k-1}$. Viimeisellä rivillä jäljelle jäävä osa normeerataan ykkösen pituiseksi.

Lause 3.1. *Edellä esitetylle Gram-Schmidtin prosessille pätee:*

a) $(\boldsymbol{q}^1, \boldsymbol{q}^2, \ldots)$ *on ortonormaali.*

b) $\mathrm{sp}(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^k) = \mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k)$ *kaikilla* $k \geq 1$.

Erityisesti, jos V *on äärellisdimensioinen ja* $\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n\}$ *on sen kanta, niin* $\{\boldsymbol{q}^1, \ldots, \boldsymbol{q}^n\}$ *on* V *:n ortonormaali kanta.*

Tod. Prosessi pyörii niin kauan, kun $\boldsymbol{w}^k \neq 0$ (tai $\boldsymbol{v}^j$-vektorit loppuvat). Näytetään aluksi, että b) on voimassa tähän asti. Koska

$$\boldsymbol{v}^k = \|\boldsymbol{w}^k\| \, \boldsymbol{q}^k + \sum_{j=1}^{k-1} \left\langle \boldsymbol{v}^k, \boldsymbol{q}^j \right\rangle \boldsymbol{q}^j \ ,$$

[11]Kompleksiselle matriisille $\boldsymbol{M}^* = \overline{\boldsymbol{M}}^T$ ja reaaliselle $\boldsymbol{M}^* = \boldsymbol{M}^T$.

saadaan kaikilla k : $\boldsymbol{v}^k \in \mathrm{sp}(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^k)$, josta $\mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k) \subset \mathrm{sp}(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^k)$. Toisaalta, jokaiselle $\boldsymbol{q}^k$ selvästi pätee $\boldsymbol{q}^k \in \mathrm{sp}(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^{k-1}, \boldsymbol{v}^k)$. Täten induktiivisesti

$$\boldsymbol{q}^k \in \mathrm{sp}(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^{k-1}, \boldsymbol{v}^k) \subset \mathrm{sp}(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^{k-2}, \boldsymbol{v}^{k-1}, \boldsymbol{v}^k) \subset \cdots \subset \mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k) \ .$$

Näin kaikilla k , joten $\mathrm{sp}(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^k) \subset \mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k)$ ja b) on voimassa.

Jos olisi $\boldsymbol{w}^k = 0$ jollakin k , tämä tarkoittaisi, että

$$\boldsymbol{v}^k = \sum_{j=1}^{k-1} \left\langle \boldsymbol{v}^k, \boldsymbol{q}^j \right\rangle \boldsymbol{q}^j \in \mathrm{sp}(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^{k-1})$$

(sillä b) on voimassa vielä edellisellä kierroksella). Mutta tämä on mahdotonta, koska $\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k$ ovat lineaarisesti riippumattomat. Siispä $\boldsymbol{w}^k$:t eivät koskaan tule nolliksi.

Todistetaan a) induktiolla:

Selvästi $\{\boldsymbol{q}^1\}$ on ortonormaali.

Oletetaan, että $\{\boldsymbol{q}^1, \ldots, \boldsymbol{q}^k\}$ on ortonormaali. Tällöin, kun $i \le k$, saadaan

$$\begin{aligned}
\left\langle \boldsymbol{q}^{k+1}, \boldsymbol{q}^i \right\rangle &= \left\langle \tfrac{1}{\|\boldsymbol{w}^{k+1}\|} \big(\boldsymbol{v}^{k+1} - \textstyle\sum_{j=1}^k \left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^j \right\rangle \boldsymbol{q}^j\big) \, , \, \boldsymbol{q}^i \right\rangle \\
&= \tfrac{1}{\|\boldsymbol{w}^{k+1}\|} \big(\left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^i \right\rangle - \textstyle\sum_{j=1}^k \left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^j \right\rangle \left\langle \boldsymbol{q}^j, \boldsymbol{q}^i \right\rangle \big) \\
&= \tfrac{1}{\|\boldsymbol{w}^{k+1}\|} \big(\left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^i \right\rangle - \left\langle \boldsymbol{v}^{k+1}, \boldsymbol{q}^i \right\rangle \big) = 0 \ .
\end{aligned}$$

Näin $\boldsymbol{q}^{k+1}$ on kohtisuorassa kaikkia $\boldsymbol{q}^i$, $i \le k$ vastaan. Selvästi $\|\boldsymbol{q}^{k+1}\| = 1$. Ja kun muutkin ovat keskenään ortonormaalit, $\{\boldsymbol{q}^1, \ldots, \boldsymbol{q}^{k+1}\}$ on ortonormaali. □

Huomaa, että saatava ortonormaali joukko riippuu paitsi vektoreista $\boldsymbol{v}^j$ myös niiden järjestyksestä.

Tehtävä 3.2. Näytä, että äärellisdimensioisen sisätuloavaruuden mielivaltainen ortonormaali joukko voidaan täydentää ortonormaaliksi kannaksi.

Esimerkki 3.9. Lähdetään liikkeelle $\mathbb{R}^3$:n kannasta $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \boldsymbol{v}^3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$.
Saadaan:

$$\begin{aligned}
\boldsymbol{q}^1 &= \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\
\boldsymbol{w}^2 &= \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \tfrac{3}{\sqrt{2}} \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} \\
\boldsymbol{q}^2 &= \tfrac{1}{\sqrt{\frac{1}{2}}} \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\
\boldsymbol{w}^3 &= \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \tfrac{2}{\sqrt{2}} \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0 \cdot \tfrac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \\
\boldsymbol{q}^3 &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \ .
\end{aligned}$$

Näin saatiin ortonormaali kanta $\left\{ \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} , \tfrac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} , \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

3.2. **QR-hajotelma, pienimmän neliösumman tehtävä.**

Olkoon V sisätuloavaruus, A sen äärellisdimensioinen aliavaruus ja $\boldsymbol{v} \in V$. Tarkastellaan seuraavaa approksimointitehtävää[12]:

Etsi $\boldsymbol{a} \in A$ siten, että $\|\boldsymbol{v}-\boldsymbol{a}\|$ on pienin mahdollinen.

Lemma 3.2. *Tällä tehtävällä on olemassa yksikäsitteinen ratkaisu $\boldsymbol{a}^*$ ja sille pätee:*

$$\langle \boldsymbol{v}-\boldsymbol{a}^*, \boldsymbol{u}\rangle = 0 \qquad \textit{kaikilla } \boldsymbol{u} \in A . \tag{3.5}$$

Tod. Olkoon $\boldsymbol{a}^* \in A$ siten, että (3.5) on voimassa. Tällöin mielivaltaiselle $\boldsymbol{a} \in A$ pätee

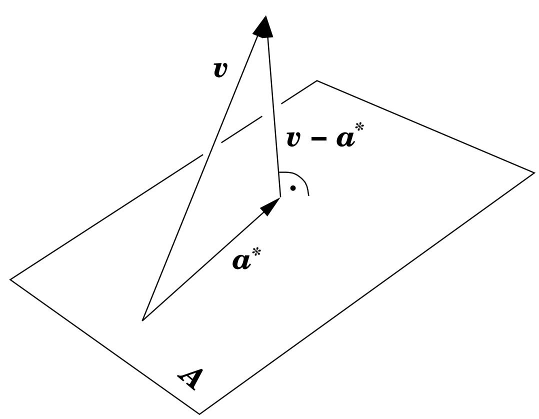
$$\begin{aligned}\|\boldsymbol{v}-\boldsymbol{a}\|^2 &= \|\boldsymbol{v}-\boldsymbol{a}^*+\boldsymbol{a}^*-\boldsymbol{a}\|^2 \\ &= \|\boldsymbol{v}-\boldsymbol{a}^*\|^2 + \langle \boldsymbol{v}-\boldsymbol{a}^*, \boldsymbol{a}^*-\boldsymbol{a}\rangle + \langle \boldsymbol{a}^*-\boldsymbol{a}, \boldsymbol{v}-\boldsymbol{a}^*\rangle + \|\boldsymbol{a}^*-\boldsymbol{a}\|^2 \\ &= \|\boldsymbol{v}-\boldsymbol{a}^*\|^2 + \|\boldsymbol{a}^*-\boldsymbol{a}\|^2 ,\end{aligned}$$

sillä $\boldsymbol{u} = \boldsymbol{a}^* - \boldsymbol{a} \in A$. Nähdään, että $\|\boldsymbol{v}-\boldsymbol{a}\| > \|\boldsymbol{v}-\boldsymbol{a}^*\|$ heti, kun $\boldsymbol{a} \neq \boldsymbol{a}^*$. Näin ollen ratkaisu, mikäli olemassa, on yksikäsitteinen.

Näytetään nyt, että ehdon (3.5) toteuttava ratkaisu löytyy. Olkoon $\{\boldsymbol{q}^1, \ldots, \boldsymbol{q}^m\}$ aliavaruuden A ortonormaali kanta. Asetetaan $\boldsymbol{a}^* = \sum_{j=1}^m \langle \boldsymbol{v}, \boldsymbol{q}^j\rangle\, \boldsymbol{q}^j$. Tällöin jokaisella $\boldsymbol{q}^i$ saadaan

$$\begin{aligned}\left\langle \boldsymbol{v}-\boldsymbol{a}^*, \boldsymbol{q}^i\right\rangle &= \left\langle \boldsymbol{v} - \textstyle\sum_{j=1}^m \left\langle \boldsymbol{v}, \boldsymbol{q}^j\right\rangle \boldsymbol{q}^j , \boldsymbol{q}^i\right\rangle \\ &= \left\langle \boldsymbol{v}, \boldsymbol{q}^i\right\rangle - \textstyle\sum_{j=1}^m \left\langle \boldsymbol{v}, \boldsymbol{q}^j\right\rangle \left\langle \boldsymbol{q}^j, \boldsymbol{q}^i\right\rangle \\ &= \left\langle \boldsymbol{v}, \boldsymbol{q}^i\right\rangle - \left\langle \boldsymbol{v}, \boldsymbol{q}^i\right\rangle = 0 .\end{aligned}$$

Siten $\boldsymbol{v}-\boldsymbol{a}^*$ on kohtisuorassa kaikkia kantavektoreita vastaan, joten se on kohtisuorassa koko alivaruutta vastaan. □



Edellisen lemman vektoria $\boldsymbol{a}^*$ kutsutaan $\boldsymbol{v}$:n *kohtisuoraksi projektioksi* aliavaruudelle A ja merkitään $\boldsymbol{a}^* = P_A^\perp \boldsymbol{v}$. Kaavasta

$$P_A^\perp \boldsymbol{v} = \textstyle\sum_{j=1}^m \left\langle \boldsymbol{v}, \boldsymbol{q}^j\right\rangle \boldsymbol{q}^j \tag{3.6}$$

ja sisätulon lineaarisuudesta nähdään myös, että $P_A^\perp$ on lineaarikuvaus $V \mapsto A$.

[12] Koska $\mathbb{R}^n$:ssä $\|\boldsymbol{v}-\boldsymbol{a}\|^2 = \sum_{j=1}^n (v_j - a_j)^2$, tämän minimointia kutsutaan usein *pienimmän neliösumman tehtäväksi.*

Olkoon edellisen lemman tilanteessa $V = \mathbb{R}^n$ ja aliavaruus A määritelty muodossa $A = \mathrm{sp}(\boldsymbol{a}^1, \dots, \boldsymbol{a}^m)$. Merkitään $\boldsymbol{A} = [\boldsymbol{a}^1 \ \dots \ \boldsymbol{a}^m] \in \mathbb{R}^{n\times m}$. Tällöin

$$A = R(\boldsymbol{A}) = \left\{ \boldsymbol{A}\boldsymbol{x} \,\middle|\, \boldsymbol{x} \in \mathbb{R}^m \right\} \ .$$

Etsitään ratkaisua muodossa $\boldsymbol{a}^* = \boldsymbol{A}\boldsymbol{c}$, missä $\boldsymbol{c} \in \mathbb{R}^m$. Lemman mukaan ratkaisun on toteutettava

$$\langle \boldsymbol{v} - \boldsymbol{A}\boldsymbol{c}\,,\, \boldsymbol{u}\rangle = 0 \qquad \text{kaikilla } \boldsymbol{u} \in R(\boldsymbol{A})\,,$$

eli

$$(\boldsymbol{v} - \boldsymbol{A}\boldsymbol{c})^T\, \boldsymbol{A}\boldsymbol{x} = 0 \qquad \text{kaikilla } \boldsymbol{x} \in \mathbb{R}^m\,.$$

Tämä on mahdollista vain, jos $(\boldsymbol{v} - \boldsymbol{A}\boldsymbol{c})^T\,\boldsymbol{A} = 0$ eli $\boldsymbol{A}^T\,(\boldsymbol{v} - \boldsymbol{A}\boldsymbol{c}) = 0$. Näin $\boldsymbol{c}$:n on toteutettava

$$\boldsymbol{A}^T \boldsymbol{A}\,\boldsymbol{c} = \boldsymbol{A}^T\,\boldsymbol{v} \ . \tag{3.7}$$

Jos $\boldsymbol{A}^T\boldsymbol{A}$ on invertoituva ($\boldsymbol{A}$:n sarakkeet lineaarisesti riippumattomat) tästä voidaan ratkaista $\boldsymbol{c} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\,\boldsymbol{v}$ ja

$$P_A^{\perp}\,\boldsymbol{v} = \boldsymbol{A}\,(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\,\boldsymbol{v} \ . \tag{3.8}$$

Kun aliavaruuteen on asetettu ortonormaali kanta, kohtisuora projektio on helppo laskea kaavasta (3.6). Mielivaltainen sisätuloavaruuden kanta voidaan ortonormalisoida Gram-Schmidt -prosessilla.

Olkoon matriisin $\boldsymbol{A} = [\boldsymbol{a}^1 \ \dots \ \boldsymbol{a}^m] \in \mathbb{K}^{n\times m}$ sarakkeet lineaarisesti riippumattomat. Tehdään $\boldsymbol{A}$:n sarakkeille ortonormalisointi. Kaavoista (3.4) saadaan (kun $\boldsymbol{v}^j = \boldsymbol{a}^j$)

$$\boldsymbol{A} = [\boldsymbol{a}^1\ \boldsymbol{a}^2\ \dots\ \boldsymbol{a}^m] = [\boldsymbol{q}^1\ \boldsymbol{q}^2\ \dots\ \boldsymbol{q}^m] \begin{bmatrix} \|\boldsymbol{a}^1\| & \langle \boldsymbol{a}^2, \boldsymbol{q}^1\rangle & \dots & \langle \boldsymbol{a}^m, \boldsymbol{q}^1\rangle \\ & \|\boldsymbol{w}^2\| & \dots & \langle \boldsymbol{a}^m, \boldsymbol{q}^2\rangle \\ & & \ddots & \vdots \\ & & & \|\boldsymbol{w}^m\| \end{bmatrix} = \boldsymbol{Q}\,\boldsymbol{R}\,,$$

missä yläkolmiomatriisin $\boldsymbol{R} \in \mathbb{K}^{m\times m}$ diagonaalilla on skaalaustekijät ja yläpuolella sisätulot $r_{ij} = \langle \boldsymbol{a}^j, \boldsymbol{q}^i\rangle$, $i < j$. Matriisin $\boldsymbol{Q}$ sarakkeet ovat ortonormaalit, joten $\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{I}$.

Tätä esitystä $\boldsymbol{A} = \boldsymbol{Q}\,\boldsymbol{R}$ kutsutaan $\boldsymbol{A}$:n (suppeaksi) QR-hajotelmaksi.

Sijoittamalla $\boldsymbol{A}$:n QR-hajotelma kaavaan (3.7) saadaan

$$\boldsymbol{R}^*\boldsymbol{Q}^*\boldsymbol{Q}\boldsymbol{R}\,\boldsymbol{c} = \boldsymbol{R}^*\boldsymbol{Q}^*\,\boldsymbol{v} \qquad \text{eli} \qquad \boldsymbol{R}\,\boldsymbol{c} = \boldsymbol{Q}^*\,\boldsymbol{v}\ ,$$

josta $\boldsymbol{c}$ on helppo ratkaista, koska $\boldsymbol{R}$ on yläkolmiomatriisi. Kohtisuora projektio $R(\boldsymbol{A})$:lle saadaan nyt yksinkertaisesti

$$P_{R(\boldsymbol{A})}^{\perp}\,\boldsymbol{v} = \boldsymbol{Q}\boldsymbol{Q}^*\,\boldsymbol{v} \ .$$

Täydentämällä $\{\boldsymbol{q}^1, \dots, \boldsymbol{q}^m\}$ koko $\mathbb{K}^n$:n ortonormaaliksi kannaksi saadaan unitaarinen (reaalisessa tapauksessa ortogonaalinen) neliömatriisi $\widehat{\boldsymbol{Q}} = [\boldsymbol{q}^1 \ \dots \ \boldsymbol{q}^m\ \boldsymbol{q}^{m+1} \ \dots \ \boldsymbol{q}^n] = [\boldsymbol{Q}\ \boldsymbol{Q}_2]$ ja $\boldsymbol{A}$:lle laajempi hajotelma

$$\boldsymbol{A} = [\boldsymbol{Q}\ \boldsymbol{Q}_2]\begin{bmatrix}\boldsymbol{R}\\ 0\end{bmatrix} = \widehat{\boldsymbol{Q}}\widehat{\boldsymbol{R}}\ .$$

Tätä kutsutaan $\boldsymbol{A}$:n (varsinaiseksi) QR-hajotelmaksi ja ylempänä esiintynyttä $\boldsymbol{A}$:n suppeaksi QR-hajotelmaksi. Jatkossa käytämme vapaasti kumpaa haluamme ja myös laajemmasta (varsinaisesta) hajotelmasta jätämme usein hatut pois.

Pienimmän neliösumman tehtävä on tavallisimmillaan seuraava: mitattavan suureen y oletetaan noudattavan lineaarista mallia

$$y = c_1\, x_1 + \cdots + c_n\, x_n \ .$$

Olkoon muuttujien arvoilla $(x_{i1}, x_{i2}, \ldots, x_{in})$, mitattu arvot y_i, $i = 1, \ldots, m$. Millä kertoimilla c_j malli kuvaisi parhaiten mittausaineistoa? Järjestetään mittauspisteistä saadut yhtälöt matriisiyhtälöksi $\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{c}$, missä vektori $\boldsymbol{y}$ sisältää mitatut arvot y_i, $\boldsymbol{c}$ tuntemattomat kertoimet ja matriisi $\boldsymbol{A} = (x_{ij}) \in \mathbb{R}^{m\times n}$. Koska yleensä $m > n$, niin tehtävällä ei välttämättä ole ratkaisua, joten etsitään kertoimia, jotka minimoivat virheen $\|\boldsymbol{y} - \boldsymbol{A}\,\boldsymbol{c}\|$.

Esimerkki 3.10. Sovitetaan mallia $y = c_0 + c_1\, x_1 + c_2\, x_2$ dataan:

$$\left\{ \begin{bmatrix} x_{1,j} \\ x_{2,j} \\ y_j \end{bmatrix} \right\}_{j=1}^{50} = \left\{ \begin{bmatrix} 3.7 \\ 7.8 \\ 12.0 \end{bmatrix} \begin{bmatrix} 2.4 \\ 8.3 \\ 16.0 \end{bmatrix} \begin{bmatrix} 2.5 \\ 7.8 \\ 15.7 \end{bmatrix} \begin{bmatrix} 9.8 \\ 7.0 \\ 7.1 \end{bmatrix} \cdots \begin{bmatrix} 0.5 \\ 2.6 \\ 8.4 \end{bmatrix} \begin{bmatrix} 2.8 \\ 6.6 \\ 11.9 \end{bmatrix} \begin{bmatrix} 9.0 \\ 7.3 \\ 9.0 \end{bmatrix} \begin{bmatrix} 3.3 \\ 0.3 \\ -2.6 \end{bmatrix} \right\} .$$

Saadaan yhtälöryhmä:

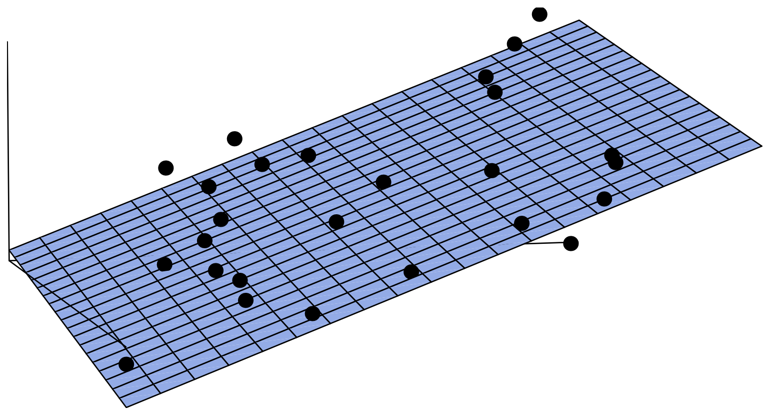
$$\begin{bmatrix} 1 & 3.7 & 7.8 \\ 1 & 2.4 & 8.3 \\ 1 & 2.5 & 7.8 \\ 1 & 9.8 & 7.0 \\ \vdots & \vdots & \vdots \\ 1 & 0.5 & 2.6 \\ 1 & 2.8 & 6.6 \\ 1 & 9.0 & 7.3 \\ 1 & 3.3 & 0.3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 12.0 \\ 16.0 \\ 15.7 \\ 7.1 \\ \vdots \\ 8.4 \\ 11.9 \\ 9.0 \\ -2.6 \end{bmatrix}$$

Tekemällä tälle 50×3-matriisille QR-hajotelma, saadaan pienimmän neliösumman ratkaisuksi[13] (pyöristettynä)

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -7.07 & -35.25 & -34.03 \\ 0 & 22.41 & -1.52 \\ 0 & 0 & -18.87 \end{bmatrix}^{-1} \begin{bmatrix} -0.141 & -0.141 & -0.141 & \ldots & -0.141 & -0.141 & -0.141 \\ -0.057 & -0.115 & -0.119 & \ldots & -0.097 & 0.179 & -0.075 \\ -0.153 & -0.175 & -0.149 & \ldots & -0.0868 & -0.146 & 0.245 \end{bmatrix} \begin{bmatrix} 12.0 \\ 16.0 \\ 15.7 \\ \vdots \\ 11.9 \\ 9.0 \\ -2.6 \end{bmatrix} = \begin{bmatrix} 0.97 \\ -0.59 \\ 1.99 \end{bmatrix} .$$

Alla on piirretty mittauspisteet sekä taso $y = 0.97 - 0.59\, x_1 + 1.99\, x_2$, $(x_1, x_2) \in [0, 10]^2$.

[13] Kun kerroinmatriisi on `A` ja oikea puoli `y`, saadaan tämä `matlab`illa seuraavasti:
`[Q,R]=qr(A); c=inv(R)*Q'*y;` tai vielä yksinkertaisemmin: `c=A\y;`.



Esimerkki 3.11. Olkoon $V = C[-\pi, \pi]$ varustettu sisätulolla $\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx$ ja olkoon T_n funktioiden $1, \cos(x), \sin(x), \ldots, \cos(nx), \sin(nx)$ virittämä aliavaruus. Nämä ovat (harj. teht) keskenään ortogonaaliset ja

$$\int_{-\pi}^{\pi} \sin^2(kx)\,dx = \int_{-\pi}^{\pi} \cos^2(kx)\,dx = \pi \ ,$$

kun $k = 1, 2, \ldots$. Normeerauksen jälkeen saadaan ortonormaaliksi kannaksi

$$\left\{q^0, q^1, \ldots, q^{2n}\right\} = \left\{\tfrac{1}{\sqrt{2\pi}}, \tfrac{1}{\sqrt{\pi}}\ \sin(x), \tfrac{1}{\sqrt{\pi}}\ \cos(x), \ldots, \tfrac{1}{\sqrt{\pi}}\ \sin(nx), \tfrac{1}{\sqrt{\pi}}\ \cos(nx)\right\} \ .$$

Täten (3.6):n mukaan annetun funktion f paras approksimaatio tässä aliavaruudessa saadaan funktioiden q^j lineaarikombinaationa kertoimilla $\langle f, q^j\rangle$. Palaten takaisin trigonometrisiin funktioihin saadaan f :n paras approksimaatio trigonometrisellä polynomilla:

$$P_{T_n}^{\perp} f(x) = a_0 + \sum_{k=1}^{n} \left[\, a_k \cos(kx) + b_k \sin(kx) \,\right] ,$$

missä

$$\begin{aligned} a_0 &= \tfrac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\,dx \ , \\ a_k &= \tfrac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \cos(kx)\,dx \ , \qquad k = 1, \ldots, n \ , \\ b_k &= \tfrac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \sin(kx)\,dx \ , \qquad k = 1, \ldots, n \ . \end{aligned}$$

Tämä on f :n *Fourier-sarjan* alkupää.

Paitsi pienimmän neliösumman tehtävien ratkaisemiseen, QR-hajotelmaa käytetään mm. ominaisarvojen ja -vektoreiden numeeriseen laskemiseen. Se on yhtä tärkeä numeerisen matriisilaskennan perustyökalu kuin Gaussin eliminointi.

Edellä oletettiin matriisin $\boldsymbol{A}$ sarakkeet lineaarisesti riippumattomiksi. Tämä tarvittiin, jotta voitiin käyttää Gram-Schmidtin ortonormeerausprosessia. QR-hajotelman laskemiseksi tämä ei kuitenkaan ole välttämätöntä. On vain löydettävä unitaarinen (reaalisessa tapauksessa ortogonaalinen) matriisi $\boldsymbol{Q}$ siten, että $\boldsymbol{Q}^*\boldsymbol{A}$ on yläkolmiomatriisi. Lähdetään tekemään tätä nk. *Householderin muunnoksilla.*

Annetulle $\boldsymbol{v} \in \mathbb{C}^n \setminus \{0\}$ määritellään Householderin matriisi (muunnos)

$$\boldsymbol{H} = \boldsymbol{I} - 2\,\frac{\boldsymbol{v}\,\boldsymbol{v}^*}{\boldsymbol{v}^*\boldsymbol{v}} \ .$$

Tämä on hermiittinen $\boldsymbol{H}^* = \boldsymbol{H}$ ja unitaarinen

$$\boldsymbol{H}^*\boldsymbol{H} = \boldsymbol{H}^2 = \boldsymbol{I} - 4\,\frac{\boldsymbol{v}\,\boldsymbol{v}^*}{\boldsymbol{v}^*\boldsymbol{v}} + 4\,\frac{\boldsymbol{v}\,\boldsymbol{v}^*\boldsymbol{v}\,\boldsymbol{v}^*}{(\boldsymbol{v}^*\boldsymbol{v})^2} = \boldsymbol{I} \ .$$

Usein tarvitaan sellaista matriisia $\boldsymbol{H}$, että annettu vektori $\boldsymbol{x}$ tulisi $\boldsymbol{H}$:lla kerrottuna annetun yksikkövektorin $\boldsymbol{e}$ (tyypillisesti $\boldsymbol{e}^1$) suuntaiseksi. Koska

$$\boldsymbol{H}\,\boldsymbol{x} = \boldsymbol{x} - 2\,\frac{\langle \boldsymbol{x}, \boldsymbol{v}\rangle}{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}\,\boldsymbol{v} \ ,$$

nähdään, että $\boldsymbol{H}\boldsymbol{x} = \alpha\,\boldsymbol{e} \implies \boldsymbol{v} \in \mathrm{sp}(\boldsymbol{x}, \boldsymbol{e})$, joten yritetään: $\boldsymbol{v} = \boldsymbol{x} + \alpha\,\boldsymbol{e}$. Saadaan

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{v}\rangle &= \|\boldsymbol{x}\|^2 + \bar{\alpha}\,\langle \boldsymbol{x}, \boldsymbol{e}\rangle \qquad \text{ja} \\ \langle \boldsymbol{v}, \boldsymbol{v}\rangle &= \|\boldsymbol{x}\|^2 + \bar{\alpha}\,\langle \boldsymbol{x}, \boldsymbol{e}\rangle + \alpha\overline{\langle \boldsymbol{x}, \boldsymbol{e}\rangle} + |\alpha|^2 \ , \end{aligned}$$

joten

$$\boldsymbol{H}\,\boldsymbol{x} = \Big(1 - 2\,\frac{\|\boldsymbol{x}\|^2 + \bar{\alpha}\,\langle \boldsymbol{x}, \boldsymbol{e}\rangle}{\|\boldsymbol{x}\|^2 + \bar{\alpha}\,\langle \boldsymbol{x}, \boldsymbol{e}\rangle + \alpha\overline{\langle \boldsymbol{x}, \boldsymbol{e}\rangle} + |\alpha|^2}\Big)\,\boldsymbol{x} - 2\,\alpha\,\frac{\langle \boldsymbol{x}, \boldsymbol{v}\rangle}{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}\,\boldsymbol{e} \ .$$

Tämä on $\boldsymbol{e}$:n suuntainen, jos $\boldsymbol{x}$:n kerroin on nolla. Näin on, kun

$$2\,\mathrm{Im}\left(\alpha\,\overline{\langle \boldsymbol{x}, \boldsymbol{e}\rangle}\right) + |\alpha|^2 = \|\boldsymbol{x}\|^2 \ .$$

Valitaan siis α siten, että $\alpha\,\overline{\langle \boldsymbol{x}, \boldsymbol{e}\rangle} \in \mathbb{R}$ ja $|\alpha| = \|\boldsymbol{x}\|$ eli $\alpha = \frac{\langle \boldsymbol{x}, \boldsymbol{e}\rangle}{|\langle \boldsymbol{x}, \boldsymbol{e}\rangle|}\,\|\boldsymbol{x}\|$ ($\alpha = \|\boldsymbol{x}\|$, jos $\langle \boldsymbol{x}, \boldsymbol{e}\rangle = 0$). Saadaan: $\langle \boldsymbol{x}, \boldsymbol{v}\rangle = \frac{1}{2}\,\langle \boldsymbol{v}, \boldsymbol{v}\rangle$ ja

$$\boldsymbol{H}\,\boldsymbol{x} = \boldsymbol{x} - 2\,\frac{\langle \boldsymbol{x}, \boldsymbol{v}\rangle}{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}\,(\boldsymbol{x} + \alpha\,\boldsymbol{e}) = -\alpha\,\boldsymbol{e} \ .$$

Käytännössä matriisia $\boldsymbol{H}$ ei koskaan muodosteta, ainoastaan vektori $\boldsymbol{v}$ talletetaan ja esimerkiksi kertolasku $\boldsymbol{H}\boldsymbol{A}$ suoritetaan:

$$\boldsymbol{H}\boldsymbol{A} = \boldsymbol{A} - \tfrac{2}{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}\,\boldsymbol{v}\,\boldsymbol{v}^*\boldsymbol{A} \ ,$$

eli $\boldsymbol{A}$:han lisätään matriisi, jonka rangi on yksi (kaikki sarakkeet ovat $\boldsymbol{v}$:n suuntaiset).

QR-hajotelma Householderin muunnoksilla.
Olkoon $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ ja $\boldsymbol{a}^1$ sen ensimmäinen sarake. Olkoon $\boldsymbol{H}_1 = \boldsymbol{I} - 2\,\frac{\boldsymbol{v}^1(\boldsymbol{v}^1)^*}{(\boldsymbol{v}^1)^*\boldsymbol{v}^1}$ Householderin muunnos siten, että $\boldsymbol{H}_1\boldsymbol{a}^1 = r_{1,1}\,\boldsymbol{e}^1$. Tällöin

$$\boldsymbol{H}_1\,\boldsymbol{A} = \begin{bmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,n} \\ 0 & & & \\ \vdots & & \widetilde{\boldsymbol{A}}_1 & \\ 0 & & & \end{bmatrix} \ .$$

Olkoon $\widetilde{\boldsymbol{a}}^1$ matriisin $\widetilde{\boldsymbol{A}}_1$ ensimmäinen sarake ja $\widetilde{\boldsymbol{H}}_2 = I - 2\,\frac{\widetilde{\boldsymbol{v}}^2(\widetilde{\boldsymbol{v}}^2)^*}{(\widetilde{\boldsymbol{v}}^2)^*\widetilde{\boldsymbol{v}}^2}$ siten, että $\widetilde{\boldsymbol{H}}_2\widetilde{\boldsymbol{a}}^1 = r_{2,2}\,\boldsymbol{e}^1 \in \mathbb{C}^{m-1}$. Asetetaan $\boldsymbol{v}^2 = \begin{bmatrix} 0 \\ \widetilde{\boldsymbol{v}}^2 \end{bmatrix}$ ja $\boldsymbol{H}_2 = I - 2\,\dfrac{\boldsymbol{v}^2(\boldsymbol{v}^2)^*}{(\boldsymbol{v}^2)^*\boldsymbol{v}^2} = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{\boldsymbol{H}}_2 \end{bmatrix}$. Tällöin

$$\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A} = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & \dots & r_{1,n} \\ 0 & r_{2,2} & r_{2,3} & \dots & r_{2,n} \\ 0 & 0 & & & \\ \vdots & \vdots & & \tilde{\boldsymbol{A}}_2 & \\ 0 & 0 & & & \end{bmatrix}.$$

Näin jatkaen saadaan tapauksessa $m \geq n$:

$$\boldsymbol{H}_n \cdots \boldsymbol{H}_2\boldsymbol{H}_1\,\boldsymbol{A} = \begin{bmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,n} \\ 0 & r_{2,2} & \dots & r_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & r_{n,n} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R} \\ 0 \end{bmatrix},$$

ja $\boldsymbol{A}$:lle QR-hajotelma, jossa $\boldsymbol{Q} = \boldsymbol{H}_1\boldsymbol{H}_2\cdots\boldsymbol{H}_n$ (sillä $\boldsymbol{H}_j^* = \boldsymbol{H}_j = \boldsymbol{H}_j^{-1}$).

Vastaavasti, jos $m < n$, päädytään muotoon

$$\boldsymbol{A} = \boldsymbol{H}_1\boldsymbol{H}_2\cdots\boldsymbol{H}_m \begin{bmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,m} & \dots & r_{1,n} \\ 0 & r_{2,2} & \dots & r_{2,n} & \dots & r_{2,n} \\ \vdots & & \ddots & \vdots & & \vdots \\ 0 & 0 & \dots & r_{m,m} & \dots & r_{m,n} \end{bmatrix}.$$

Ratkaistaessa pienimmän neliösumman tehtävää $\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|$, missä $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, $m \geq n$, matriisia $\boldsymbol{Q}$ ei tarvitse muodostaa, eikä vektoreita $\boldsymbol{v}^k$ tallettaa, kun jokaisella askelella myös oikea puoli kerrotaan matriisilla $\boldsymbol{H}_k$, sillä näiden unitaarisuuden perusteella

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\| = \|\boldsymbol{H}_n \dots \boldsymbol{H}_2\boldsymbol{H}_1(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})\| = \left\| \begin{bmatrix} \boldsymbol{R}\boldsymbol{x} - \tilde{\boldsymbol{b}}_1 \\ \tilde{\boldsymbol{b}}_2 \end{bmatrix} \right\|,$$

joten paras $\boldsymbol{x}$ saadaan systeemin $\boldsymbol{R}\boldsymbol{x} = \tilde{\boldsymbol{b}}_1$ ratkaisuna.

Jos tulo

$$\boldsymbol{Q}_n = \boldsymbol{H}_1\boldsymbol{H}_2 \dots \boldsymbol{H}_n = \Big(\boldsymbol{I} - 2\,\frac{\boldsymbol{v}^1(\boldsymbol{v}^1)^*}{\|\boldsymbol{v}^1\|^2}\Big)\Big(\boldsymbol{I} - 2\,\frac{\boldsymbol{v}^2(\boldsymbol{v}^2)^*}{\|\boldsymbol{v}^2\|^2}\Big)\dots\Big(\boldsymbol{I} - 2\,\frac{\boldsymbol{v}^n(\boldsymbol{v}^n)^*}{\|\boldsymbol{v}^n\|^2}\Big)$$

halutaan muodostaa, se on parempi tehdä muodossa $\boldsymbol{Q}_k = \boldsymbol{I} - \boldsymbol{W}_k\boldsymbol{V}_k^*$, $k = 1, \dots, n$ seuraavasti:

asetetaan: $\boldsymbol{W}_1 = \boldsymbol{w}_1 = \frac{2}{\|\boldsymbol{v}^1\|}\, \boldsymbol{v}^1$, $\boldsymbol{V}_1 = \frac{1}{\|\boldsymbol{v}^1\|}\boldsymbol{v}^1$. Tällöin $\boldsymbol{Q}_1 = \boldsymbol{H}_1 = I - \boldsymbol{W}_1\boldsymbol{V}_1^*$ ja

$$\begin{aligned}\boldsymbol{Q}_k = \boldsymbol{Q}_{k-1}\boldsymbol{H}_k &= \Big(\boldsymbol{I} - \boldsymbol{W}_{k-1}\boldsymbol{V}_{k-1}^*\Big)\Big(I - 2\,\frac{\boldsymbol{v}^k(\boldsymbol{v}^k)^*}{\|\boldsymbol{v}^k\|^2}\Big) = \\ &= \boldsymbol{I} - \boldsymbol{W}_{k-1}\boldsymbol{V}_{k-1}^* - \boldsymbol{w}^k\,\frac{(\boldsymbol{v}^k)^*}{\|\boldsymbol{v}^k\|} = \boldsymbol{I} - \boldsymbol{W}_k\boldsymbol{V}_k^*\ ,\end{aligned}$$

missä $\boldsymbol{w}^k = \frac{2}{\|\boldsymbol{v}^k\|}\,\boldsymbol{Q}_{k-1}\boldsymbol{v}^k$ ja $\boldsymbol{W}_k = [\boldsymbol{W}_{k-1}\ \boldsymbol{w}^k]$, $\boldsymbol{V}_k = [\boldsymbol{V}_{k-1}\ \frac{1}{\|\boldsymbol{v}^k\|}\boldsymbol{v}^k]$.

Tämä esitystapa on varsin ekonominen erityisesti, jos $n \ll m$.

Joukon $S \subset V$ *ortogonaalikomplementti* $S^\perp$ on niiden vektoreiden joukko, jotka ovat ortogonaalisia kaikkia joukon S vektoreita vastaan eli

$$S^\perp = \{\boldsymbol{w} \in V \mid \langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0 \ \ \forall \boldsymbol{v} \in S\}\ .$$

Tehtävä 3.3. Näytä: Joukon S ortogonaalikomplementti on aina aliavaruus ja äärellisdimensioisessa avaruudessa $(S^\perp)^\perp = \mathrm{sp}(S)$.

Lemma 3.2 voidaan kirjoittaa muotoon:

Lause 3.3. *Olkoon A sisätuloavaruuden V äärellisdimensioinen aliavaruus ja $\boldsymbol{v} \in V$. Tällöin on olemassa yksikäsitteinen vektori $\boldsymbol{a} \in A$ siten, että $\boldsymbol{v} - \boldsymbol{a} \in A^\perp$.*

Tämä vektorihan oli $\boldsymbol{v}$:n paras approksimaatio A :n vektoreilla ja se saatiin kohtisuorana projektiona $\boldsymbol{a} = P_A^\perp\,\boldsymbol{v}$.

Matriisiin liitetyille avaruuksille pätee:

Lause 3.4. *Olkoon $\boldsymbol{A} \in \mathbb{K}^{m\times n}$. Tällöin $R(\boldsymbol{A}^*)^\perp = N(\boldsymbol{A})$, eli $\boldsymbol{A}$:n nolla-avaruus ja $\boldsymbol{A}^*$:n kuva-avaruus ovat toistensa ortogonaalikomplementteja $\mathbb{K}^n$:ssä.*

Tod.

$$\begin{array}{cccc} & \boldsymbol{x} \in (R(\boldsymbol{A}^*))^\perp & \Leftrightarrow & \langle \boldsymbol{w}, \boldsymbol{x}\rangle = 0 \ \ \forall \boldsymbol{w} \in R(\boldsymbol{A}^*) \\ \Leftrightarrow & \langle \boldsymbol{A}^*\boldsymbol{y}\,,\, \boldsymbol{x}\rangle = 0 \ \ \forall \boldsymbol{y} & \Leftrightarrow & \langle \boldsymbol{y}\,,\, \boldsymbol{A}\boldsymbol{x}\rangle = 0 \ \ \forall \boldsymbol{y} \\ \Leftrightarrow & \boldsymbol{A}\boldsymbol{x} = 0 & \Leftrightarrow & \boldsymbol{x} \in N(\boldsymbol{A})\ . \end{array}$$

□

3.3. **Matriisinormi ja häiriöalttius.**

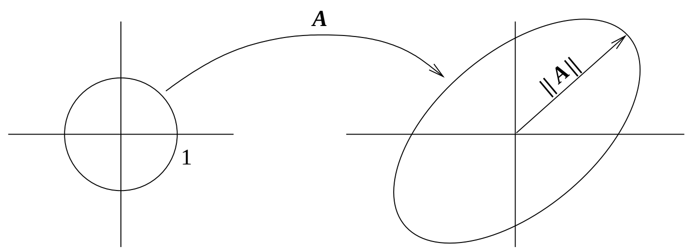
Vektorin normi mittaa vektorin pituutta. Matriiseille ja lineaarikuvauksille voidaan myös määritellä normeja. Erityisen hyödyllisiksi osoittautuvat sellaiset normit, jotka on määritelty vektorinormien avulla. Rajoitumme tässä tarkastelemaan vain matriisien normeja, normiavaruuksien välisten lineaarikuvausten normit määritellään samalla tavalla.

Olkoon $\|\cdot\|$ jokin vektorinormi (esim. $\|\cdot\|_2$ tai $\|\cdot\|_\infty$). Mitataan matriisin kokoa sillä, kuinka pitkiksi vektoreiksi matriisilla kerrottaessa yksikkövektorit saattavat kuvautua.

Niinpä matriisille $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ asetetaan

$$\|\boldsymbol{A}\| = \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\| \,. \tag{3.9}$$

Tässä siis oikealla puolella esiintyy vektoreiden $\boldsymbol{x} \in \mathbb{C}^n$ ja $\boldsymbol{A}\boldsymbol{x} \in \mathbb{C}^m$ normeja.



Näin määritelty $\|\boldsymbol{A}\|$ toteuttaa määritelmän 3.1 neljä ehtoa:

(1) (3.9):n oikealla puolella esiintyy vain ei-negatiivisia lukuja, joten $\|\boldsymbol{A}\| \geq 0$.

(2) Jos $\boldsymbol{A} \neq 0$, niin sillä on olemassa ei-nolla elementti a_{ij}. Valitaan $\boldsymbol{x} = \boldsymbol{e}^j$, jolloin $\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \neq 0$ ja $\|\boldsymbol{A}\| \geq \|\boldsymbol{A}\boldsymbol{x}\| > 0$.

(3)
$$\begin{aligned} \|\boldsymbol{A}+\boldsymbol{B}\| &= \max_{\|\boldsymbol{x}\|=1} \|(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}\| \leq \max_{\|\boldsymbol{x}\|=1} (\|\boldsymbol{A}\boldsymbol{x}\| + \|\boldsymbol{B}\boldsymbol{x}\|) \\ &\leq \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\| + \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{B}\boldsymbol{x}\| = \|\boldsymbol{A}\| + \|\boldsymbol{B}\| \,. \end{aligned}$$

Tässä käytettiin aluksi vektorinormin kolmioepäyhtälöä.

(4) $\|\alpha\boldsymbol{A}\| = \max_{\|\boldsymbol{x}\|=1} \|\alpha\boldsymbol{A}\boldsymbol{x}\| = \max_{\|\boldsymbol{x}\|=1} |\alpha|\,\|\boldsymbol{A}\boldsymbol{x}\| = |\alpha|\,\|\boldsymbol{A}\|$
jälleen vektorinormin vastaavan ominaisuuden perusteella.

Matriisinormilla ja vastaavalla vektorinormilla on lisäksi ominaisuudet (harjoitustehtävä)

$$\|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\| \,, \tag{3.10}$$

$$\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{B}\| \,, \tag{3.11}$$

$$\|\boldsymbol{A}^k\| \leq \|\boldsymbol{A}\|^k, \; k = 1, 2, \ldots \,. \tag{3.12}$$

Kun halutaan korostaa, minkä vektorinormin avulla matriisinormi on määritelty käytetään vastaavaa merkkiä. Esimerkiksi

$$\|\boldsymbol{A}\|_1 = \max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1 \qquad \text{ja} \qquad \|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2 \,.$$

Riippuen valitusta vektorinormista matriisin normin laskeminen voi olla hankalaa tai helpompaa. 1- ja ∞-normit ovat laskuissa monesti käteviä:

Lause 3.5. *Olkoon $A \in \mathbb{C}^{m \times n}$. Tällöin*

$$\|\boldsymbol{A}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}| \qquad \textit{ja} \qquad \|\boldsymbol{A}\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|$$

Tod. Jos $\|\boldsymbol{x}\|_1 = \sum_{k=1}^n |x_k| = 1$, niin

$$\|\boldsymbol{A}\boldsymbol{x}\|_1 = \sum_{i=1}^m |(\boldsymbol{A}\boldsymbol{x})_i| = \sum_{i=1}^m \left| \sum_{k=1}^n a_{ik} x_k \right| \le \sum_{i=1}^m \sum_{k=1}^n |a_{ik} x_k|$$
$$= \sum_{k=1}^n |x_k| \sum_{i=1}^m |a_{ik}| \le \sum_{k=1}^n |x_k| \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}| = \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}| \ .$$

Siten $\|\boldsymbol{A}\|_1 \le \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}|$. Toisaalta, jos l on siten, että

$$\sum_{i=1}^m |a_{i\,l}| = \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}| \ ,$$

niin

$$\|\boldsymbol{A}\boldsymbol{e}^l\|_1 = \left\| \begin{bmatrix} a_{1\,l} \\ \vdots \\ a_{m\,l} \end{bmatrix} \right\|_1 = \sum_{i=1}^m |a_{i\,l}| \ ,$$

joten $\|\boldsymbol{A}\|_1 \ge \max_{1\le j\le n} \sum_{i=1}^m |a_{ij}|$.

$\|\cdot\|_\infty$-normia koskeva väite jätetään harjoitustehtäväksi. □

Tehtävä 3.4. Millaisia yleisesti päteviä epäyhtälöitä saat matriisin $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ normien $\|\boldsymbol{A}\|_1$, $\|\boldsymbol{A}\|_2$ ja $\|\boldsymbol{A}\|_\infty$ välille? Katso vastaavien vektorinormien välisiä epäyhtälöitä (tehtävä 3.1).

Seuraava tärkeä tulos tulee käyttöön vielä useasti. Loppupuolella esitämme sille myös toisen todistuksen.

Lause 3.6. *Olkoon* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *siten, että* $\|\boldsymbol{A}\| < 1$. *Tällöin* $\boldsymbol{I} - \boldsymbol{A}$ *on invertoituva ja* $\|(\boldsymbol{I} - \boldsymbol{A})^{-1}\| \le \frac{1}{1-\|\boldsymbol{A}\|}$.

Tod. Jos $\boldsymbol{I} - \boldsymbol{A}$ ei ole invertoituva, niin on olemassa $\boldsymbol{x} \in \mathbb{C}^n$ siten, että $\|\boldsymbol{x}\| = 1$ ja $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0$. Tällöin $\|\boldsymbol{A}\| \ge \|\boldsymbol{A}\boldsymbol{x}\| = \|\boldsymbol{x}\| = 1$, mikä on ristiriita.

Jos $\|\boldsymbol{x}\| = 1$ ja $\boldsymbol{v} = (\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{x}$, niin

$$1 = \|(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{v}\| \ge \|\boldsymbol{v}\| - \|\boldsymbol{A}\boldsymbol{v}\| \ge \|\boldsymbol{v}\| - \|\boldsymbol{A}\|\|\boldsymbol{v}\| = (1 - \|\boldsymbol{A}\|)\|\boldsymbol{v}\| \ .$$

Siten $\|\boldsymbol{v}\| \le \frac{1}{1-\|\boldsymbol{A}\|}$. □

Häiriöalttius. Kun käytännön tehtävissä päädytään lineaariseen malliin $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, niin usein yhtälöiden kertoimissa ja datassa eli matriisin $\boldsymbol{A}$ tai vektorin $\boldsymbol{b}$ alkioissa, on epävarmuutta. Kertoimet on voitu saada esimerkiksi mittausten tuloksena. Halutaan tietää, miten suuri virhe tästä voi aiheutua ratkaisuun $\boldsymbol{x}$. Tarkastellaan ensin, miten $\delta\boldsymbol{b}$:n suuruinen häiriövektori oikean puolen vektorissa vaikuttaa ratkaisuun. Merkitään $\delta\boldsymbol{x}$:llä ratkaisuvektorin muutosta. Vähentämällä yhtälöt

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \qquad \text{ja} \qquad \boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$$

puolittain, saadaan $\delta \boldsymbol{x} = \boldsymbol{A}^{-1}\delta \boldsymbol{b}$. Siten *absoluuttisen virheen* normille saadaan yläraja

$$\|\delta \boldsymbol{x}\| \leq \|\boldsymbol{A}^{-1}\|\|\delta \boldsymbol{b}\| \ . \tag{3.13}$$

Lineaarisen yhtälöryhmän ratkaisun voi kerroinmatriisia skaalaamalla saada pienemmäksi, jolloin myös absoluuttinen virhe pienenee. Paremmin ratkaisun virhettä kuvaakin *suhteellinen virhe* $\|\delta \boldsymbol{x}\|/\|\boldsymbol{x}\|$. Koska

$$\|\boldsymbol{b}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\| \tag{3.14}$$

niin epäyhtälöistä (3.13) ja (3.14) saadaan suhteelliselle virheelle yläraja-arvio

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\| \frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} \ .$$

Tämän perusteella asetetaan

Määritelmä 3.3. Matriisin *häiriöalttius* on

$$\kappa(\boldsymbol{A}) = \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\| \ .$$

Suuri häiriöalttius merkitsee siten, että pienikin suhteellinen virhe $\boldsymbol{b}$:ssä voi aiheuttaa ratkaisuun $\boldsymbol{x}$ suuren epävarmuuden. Aivan vastaavasti voidaan tarkastella matriisin $\boldsymbol{A}$ häiriön $\delta \boldsymbol{A}$ aiheuttamaa virhettä ratkaisuun, ja saadaan

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x} + \delta \boldsymbol{x}\|} \leq \kappa(\boldsymbol{A}) \frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|} \ .$$

Häiriöalttius riippuu (hieman) siitä, missä matriisinormissa (ja vastaavassa vektorinormissa) asioita mitataan. Koska $1 = \|\boldsymbol{I}\| = \|\boldsymbol{A}\boldsymbol{A}^{-1}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|$, saadaan $\kappa(\boldsymbol{A}) \geq 1$ jokaiselle (invertoituvalle) matriisille normista riippumatta.

Huomaa, että (toisin kuin determinantti) häiriöalttius ei riipu matriisin skaalauksesta:

$$\kappa(\alpha \boldsymbol{A}) = \|\alpha \boldsymbol{A}\|\|(\alpha \boldsymbol{A})^{-1}\| = |\alpha| \, \|\boldsymbol{A}\| \, \|\tfrac{1}{\alpha} \ \boldsymbol{A}^{-1}\| = \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\| = \kappa(\boldsymbol{A}) \ .$$

Unitaariselle matriisille $\boldsymbol{U}$ pätee $\|\boldsymbol{U}\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$, joten $\|\boldsymbol{U}\|_2 = 1$ ja samoin $\|\boldsymbol{U}^{-1}\|_2 = \|\boldsymbol{U}^*\|_2 = 1$, joten $\kappa_2(\boldsymbol{U}) = 1$. Siten unitaarisen matriisin häiriöalttius (2-normissa mitattuna) on pienin mahdollinen.

Esimerkki 3.12. Lasketaan $\kappa_1(\boldsymbol{A})$, kun $\boldsymbol{A} = \begin{bmatrix} 1 & \varepsilon \\ 1 & -\varepsilon \end{bmatrix}$. Nyt $\boldsymbol{A}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1/\varepsilon & -1/\varepsilon \end{bmatrix}$ joten häiriöalttiudeksi saadaan lauseella 3.5 (kun $\varepsilon \in (0,1)$)

$$\kappa_1(\boldsymbol{A}) = \|\boldsymbol{A}\|_1\|\boldsymbol{A}^{-1}\|_1 = 2\,\tfrac{1}{2}\,(1 + 1/\varepsilon) = 1 + 1/\varepsilon \ ,$$

joka on suuri ε:n ollessa pieni.

4. Ominaisarvoteoriaa

Tässä luvussa käsitellään lineaarikuvausten ja matriisien ominaisarvoja ja -vektoreita. Näillä on sovelluksia lähes kaikilla tieteen ja tekniikan aloilla: rakenteen tai sähköpiirin värähtelytaajuudet vastaavat tiettyjä ominaisarvoja jne.

Tarkastelemme ominaisarvojen ja -vektoreiden ominaisuuksia sekä lineaarioperaattoreille mukavia kantoja, joiden avulla voidaan ratkaista erilaisia dynaamisia tehtäviä, kuten lineaarisia differenssi- ja differentiaaliyhtälöitä. Tarkastelemme myös hieman ominaisarvojen ja -vektoreiden numeerista laskemista.

Olkoon V $\mathbb{K}$-kertoiminen lineaariavaruus. Lineaarikuvausta avaruudelta itselleen: $T : V \mapsto V$ kutsutaan tavallisesti (lineaari-)operaattoriksi. Jos on olemassa $\lambda \in \mathbb{K}$ ja vektori $\boldsymbol{v} \in V \setminus \{0\}$ siten, että

$$T\,\boldsymbol{v} = \lambda\,\boldsymbol{v}\ , \tag{4.1}$$

niin sanotaan, että λ on T:n *ominaisarvo* ja vektori $\boldsymbol{v}$ on tähän liittyvä T:n *ominaisvektori*. Ominaisvektori $\boldsymbol{v}$ on siis sellainen, että T kuvaa sen samansuuntaiseksi vektoriksi. Yhtälöä (4.1) sanotaan *ominaisarvoyhtälöksi*. Ominaisarvo ja vastaava ominaisvektori -paria kutsutaan lyhyesti ominaispariksi $(\lambda, \boldsymbol{v})$.

Esimerkki 4.1. Olkoon $\boldsymbol{A} = \left[\begin{smallmatrix}1&3\\2&2\end{smallmatrix}\right]$. Tällöin ($\boldsymbol{A}$:lla kertomis-) lineaarikuvaukselle $L_{\boldsymbol{A}} : \mathbb{R}^2 \mapsto \mathbb{R}^2$: $L_{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{A}\,\boldsymbol{x}$ vektorilla $\boldsymbol{v} = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ saadaan:

$$L_{\boldsymbol{A}}\boldsymbol{v} = \left[\begin{smallmatrix}1&3\\2&2\end{smallmatrix}\right]\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}4\\4\end{smallmatrix}\right] = 4\,\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] = 4\,\boldsymbol{v}\ ,$$

joten $\boldsymbol{v}$ on $L_{\boldsymbol{A}}$:n ominaisvektori ja 4 vastaava ominaisarvo.

Esimerkki 4.2. Esimerkin 2.5 polynomiavaruudessa $\mathbb{P}_3$ määritellylle lineaarikuvaukselle $Tp(x) = p(2-x)$ saatiin polynomeille $\{\widehat{p}_1, \widehat{p}_2, \widehat{p}_3, \widehat{p}_4\} = \{1, 1-x, (1-x)^2, (1-x)^3\}$

$$T\widehat{p}_1 = \widehat{p}_1\ , \quad T\widehat{p}_2 = -\widehat{p}_2\ , \quad T\widehat{p}_3 = \widehat{p}_3 \quad \text{ja} \quad T\widehat{p}_4 = -\widehat{p}_4\ .$$

Täten T:llä on ominaisarvot 1 ja -1. Polynomit 1 ja $(1-x)^2$ ovat ominaisarvoa 1 vastaavia ominaisvektoreita ja $1-x$ ja $(1-x)^3$ vastaavat ominaisarvoa -1.

Esimerkki 4.3. Olkoon $C^\infty[0,1]$ välillä $[0,1]$ äärettömän monta kertaa derivoituvien funktioiden joukko. Selvästi tämä on vektoriavaruus. Tarkastellaan derivointioperaattoria $D : C^\infty[0,1] \mapsto C^\infty[0,1]$: $Df = f'$. Koska $\frac{d}{dx}e^{\alpha\,x} = \alpha\,e^{\alpha\,x}$, nähdään, että D:llä on ominaisvektorit (eli ominaisfunktiot) $e^{\alpha\,x}$ vastaten ominaisarvoja $\alpha \in \mathbb{R}$ (tai $\alpha \in \mathbb{C}$, kun tarkastellaan kompleksiarvoisia funktioita). Täten kaikki luvut ovat D:n ominaisarvoja tässä avaruudessa.

Jos tarkastellaan D:tä polynomien muodostamassa avaruudessa $\mathbb{P}_n$, nähdään, että D:llä ei ole muita ominaisfunktioita kuin vakiot, jotka vastaavat ominaisarvoa nolla.

Jos $\boldsymbol{u}$ ja $\boldsymbol{v}$ ovat lineaarikuvauksen T *samaan* ominaisarvoon λ liittyviä ominaisvektoreita, niin

$$T(\alpha\,\boldsymbol{u} + \beta\,\boldsymbol{v}) = \alpha\,T\boldsymbol{u} + \beta\,T\boldsymbol{v} = \alpha\,\lambda\,\boldsymbol{u} + \beta\,\lambda\,\boldsymbol{v} = \lambda\,(\alpha\,\boldsymbol{u} + \beta\,\boldsymbol{v})\ ,$$

joten niiden lineaarikombinaatiokin on λ:aan liittyvä ominaisvektori (kunhan ei ole nolla). Täten yhteen ominaisarvoon λ liittyvät ominaisvektorit muodostavat nollan kanssa *ominaisavaruuden*[14]

$$E_T(\lambda) = \{\boldsymbol{v} \in V \,|\, T\,\boldsymbol{v} = \lambda\,\boldsymbol{v}\} = N(T - \lambda I) \ .$$

Jatkossa pyrimme muodostamaan ominaisvektoreista kantoja. Tällöin seuraava tulos osoittautuu tärkeäksi. Se sanoo, että erisuuriin ominaisarvoihin liittyvät ominaisvektorit ovat aina lineaarisesti riippumattomat.

Lause 4.1. *Olkoot $\lambda_1, \ldots, \lambda_n$ lineaarikuvauksen $T : V \mapsto V$ erisuuria ominaisarvoja ja $\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n$ näitä vastaavia ominaisvektoreita. Tällöin $\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n$ ovat lineaarisesti riippumattomat.*

Tod. Olkoon

$$c_1\,\boldsymbol{v}^1 + c_2\,\boldsymbol{v}^2 + \cdots + c_n\,\boldsymbol{v}^n = 0 \ . \tag{4.2}$$

Kuvataan tämän sekä vasen että oikea puoli T:llä, jolloin saadaan

$$\begin{aligned} T(c_1\,\boldsymbol{v}^1 + c_2\,\boldsymbol{v}^2 + \cdots + c_n\,\boldsymbol{v}^n) &= c_1\,T\boldsymbol{v}^1 + c_2\,T\boldsymbol{v}^2 + \cdots + c_n\,T\boldsymbol{v}^n \\ &= c_1\,\lambda_1\boldsymbol{v}^1 + c_2\,\lambda_2\boldsymbol{v}^2 + \cdots + c_n\,\lambda_n\boldsymbol{v}^n = 0 \ . \end{aligned}$$

Vähennetään tästä (4.2) λ_1:llä kerrottuna, jolloin saadaan:

$$c_2\,(\lambda_2 - \lambda_1)\,\boldsymbol{v}^2 + c_3\,(\lambda_3 - \lambda_1)\,\boldsymbol{v}^3 + \cdots + c_n\,(\lambda_n - \lambda_1)\,\boldsymbol{v}^n = 0 \ . \tag{4.3}$$

Tämä olisi saatu myös soveltamalla (4.2):een suoraan kuvausta $T - \lambda_1 I$. Sovelletaan nyt (4.3):een kuvausta $T - \lambda_2 I$, jolloin saadaan

$$c_2\,(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)\,\boldsymbol{v}^3 + \cdots + c_n\,(\lambda_n - \lambda_1)(\lambda_n - \lambda_2)\,\boldsymbol{v}^n = 0 \ . \tag{4.4}$$

Näin jatkaen saadaan lopulta

$$c_{n-1}\,(\lambda_{n-1} - \lambda_1)\cdots(\lambda_{n-1} - \lambda_{n-2})\,\boldsymbol{v}^{n-1} + c_n\,(\lambda_n - \lambda_1)\cdots(\lambda_n - \lambda_{n-2})\,\boldsymbol{v}^n = 0$$

ja

$$c_n\,(\lambda_n - \lambda_1)\cdots(\lambda_n - \lambda_{n-2})(\lambda_n - \lambda_{n-1})\,\boldsymbol{v}^n = 0 \ .$$

Koska $\boldsymbol{v}^n \neq 0$ ja $\lambda_i \neq \lambda_j$, $i \neq j$, on oltava $c_n = 0$. Sijoitetaan tämä edelliseen yhtälöön jolloin saadaan $c_{n-1} = 0$ ja näin edelleen $0 = c_{n-2} = \cdots = c_2 = c_1$. Siispä $\boldsymbol{v}^1, \ldots, \boldsymbol{v}^n$ ovat lineaarisesti riippumattomat. □

Neliömatriisin $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ ominaisarvoilla ja ominaisvektoreilla tarkoitetaan $\boldsymbol{A}$:lla kertomiskuvauksen $L_{\boldsymbol{A}} : \mathbb{C}^n \mapsto \mathbb{C}^n \ : L_{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{A}\,\boldsymbol{x}$ ominaisarvoja ja -vektoreita. Huomaa, että vaikka $\boldsymbol{A}$ olisi reaalinen, sillä kertominen määrittelee lineaarikuvauksen myös $\mathbb{C}^n \mapsto \mathbb{C}^n$. Ominaisarvojen yhteydessä $L_{\boldsymbol{A}}$ tulkitaan täksi kuvaukseksi. Tämä osoittautuu mukavaksi, koska $L_{\boldsymbol{A}}$:lla voi olla kompleksisia ominaisarvoja, vaikka $\boldsymbol{A}$ olisi reaalinen:

$$\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i - 2 \\ 2i + 1 \end{bmatrix} = (1 + 2i) \begin{bmatrix} i \\ 1 \end{bmatrix} \ .$$

Siis $1 + 2i$ on matriisin $\boldsymbol{A} = \left[\begin{smallmatrix} 1 & -2 \\ 2 & 1 \end{smallmatrix}\right]$ (tai $L_{\boldsymbol{A}}$:n) ominaisarvo ja $\left[\begin{smallmatrix} i \\ 1 \end{smallmatrix}\right]$ on vastaava ominaisvektori.

[14] I on identiteettikuvaus: $I\boldsymbol{v} = \boldsymbol{v}$ kaikilla $\boldsymbol{v}$.

Lineaarikuvaukselle $L_{\boldsymbol{A}}$ ominaisarvoyhtälö (4.1) saa muodon

$$\boldsymbol{A}\,\boldsymbol{v} = \lambda\,\boldsymbol{v} \qquad \text{eli} \qquad (\lambda\,\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{v} = 0\ .$$

Tällä on ei-triviaaleja ratkaisuja ($\boldsymbol{v} \neq 0$) täsmälleen silloin kun

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0\ . \tag{4.5}$$

Matriisin $\boldsymbol{A}$ ominaisarvot ovat siis $\boldsymbol{A}$:n *karakteristisen polynomin* $p_{\boldsymbol{A}}(z) = \det(z\,\boldsymbol{I} - \boldsymbol{A})$ nollakohtia.

Tehtävä 4.1. Näytä, että $\quad p_{\boldsymbol{A}}(z) = z^n + \alpha_{n-1}\,z^{n-1} + \cdots + \alpha_1\,z + \alpha_0\,,\quad$ missä

$$\alpha_{n-1} = -(a_{1,1} + a_{2,2} + \cdots + a_{n,n}) \qquad \text{ja} \qquad \alpha_0 = (-1)^n \det(\boldsymbol{A})\ .$$

Karakteristinen polynomi on siten n-asteinen. Algebran peruslauseen mukaan n-asteisella polynomilla on kertaluvut mukaan lukien täsmälleen n juurta, joten $n \times n$-matriisilla on n ominaisarvoa, joista osa voi olla moninkertaisia. Jos $\boldsymbol{A}$ on reaalinen, sen karakteristinen polynomi on reaalikertoiminen. Tällöin, jos λ on $\boldsymbol{A}$:n ominaisarvo, niin $\overline{\lambda}$ on myös.

Tehtävä 4.2. Olkoot $\boldsymbol{A}$:n ominaisarvot $\lambda_1, \ldots, \lambda_n$. Näytä edelleen, että

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr}(\boldsymbol{A}) \qquad \text{ja} \qquad \prod_{i=1}^{n} \lambda_i = \det(\boldsymbol{A})\ ,$$

missä $\operatorname{tr}(\boldsymbol{A}) = a_{1,1} + a_{2,2} + \cdots + a_{n,n}$ on $\boldsymbol{A}$:n *jälki* (engl. trace).

Matriisin $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ ominaisarvojen joukkoa eli *spektriä* merkitään $\Lambda(\boldsymbol{A})$:lla. Tämä on $\mathbb{C}$:n ei-tyhjä osajoukko, jossa on korkeintaan n alkiota. $\boldsymbol{A}$:n *spektraalisäde* $\rho(\boldsymbol{A})$ on suurin $\boldsymbol{A}$:n ominaisarvojen itseisarvoista eli

$$\rho(\boldsymbol{A}) = \max_{\lambda \in \Lambda(\boldsymbol{A})} |\lambda|\ .$$

Tehtävä 4.3. Näytä: $\qquad \Lambda(\boldsymbol{A}^T) = \Lambda(\boldsymbol{A}) \qquad \text{ja} \qquad \Lambda(\boldsymbol{A}^*) = \overline{\Lambda(\boldsymbol{A})} = \left\{\overline{\lambda} \,\middle|\, \lambda \in \Lambda(\boldsymbol{A})\right\}\ .$

Jos $\boldsymbol{A}$ on ylä- tai alakolmiomatriisi, niin sitä on myös $\lambda \boldsymbol{I} - \boldsymbol{A}$, jolloin

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - a_{1,1})(\lambda - a_{2,2})\cdots(\lambda - a_{n,n})\ .$$

Saatiin

Lause 4.2. *Ylä- tai alakolmiomatriisin ominaisarvot ovat samat kuin sen lävistäjäalkiot.*

Esimerkki 4.4. Lasketaan matriisin

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{bmatrix}$$

ominaisarvot ja -vektorit. Muodostetaan karakteristinen polynomi

$$\begin{aligned}
p(\lambda) &= \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda-1 & -2 & 1 \\ -1 & \lambda & -1 \\ -4 & 4 & \lambda-5 \end{vmatrix} \\
&= (\lambda-1)\big(\lambda(\lambda-5)+4\big) + 2\big(-(\lambda-5)-4\big) + 1\big(-4+4\lambda\big) \\
&= (\lambda-1)(\lambda^2-5\lambda+4) - 2(\lambda-1) + 4(\lambda-1) \\
&= (\lambda-1)(\lambda^2-5\lambda+6) = (\lambda-1)(\lambda-2)(\lambda-3).
\end{aligned}$$

$\boldsymbol{A}$:n ominaisarvot ovat siis $1, 2$ ja 3. Ominaisarvoon $\lambda_1 = 1$ liittyvä ominaisvektori saadaan yhtälön $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{v}^1 = 0$ ratkaisuna

$$\boldsymbol{I} - \boldsymbol{A} = \begin{bmatrix} 0 & -2 & 1 \\ -1 & 1 & -1 \\ -4 & 4 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix},$$

jossa matriisille on tehty Gaussin eliminaatio porrasmuotoon asti. Siten ominaisarvoon 1 liittyviä ominaisvektoreita ovat vektorit $\alpha(1, -1, -2)$. Olkoon siis $\boldsymbol{v}^1 = (1, -1, -2)$. Vastaavasti ominaisarvoon $\lambda_2 = 2$ liittyvä ominaisvektori saadaan eliminoinnin

$$2\boldsymbol{I} - \boldsymbol{A} = \begin{bmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \\ -4 & 4 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & -4 & 1 \end{bmatrix},$$

kautta $\boldsymbol{v}^2 = (2, -1, -4)$. Ja kolmas saadaan

$$3\boldsymbol{I} - \boldsymbol{A} = \begin{bmatrix} 2 & -2 & 1 \\ -1 & 3 & -1 \\ -4 & 4 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 2 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix},$$

josta $\boldsymbol{v}^3 = (1, -1, -4)$.

Kuten alussa todettiin, yhteen ominaisarvoon liittyvät ominaisvektorit muodostavat nollan kanssa aliavaruuden. Matriisin ominaisarvoon λ liittyvä ominaisavaruus on

$$E_{\boldsymbol{A}}(\lambda) = \big\{\boldsymbol{v} \in \mathbb{C}^n \,\big|\, \boldsymbol{A}\,\boldsymbol{v} = \lambda\,\boldsymbol{v}\big\} = N(\lambda\,\boldsymbol{I} - \boldsymbol{A}).$$

Jos $\boldsymbol{A}$ ja λ ovat reaalisia, ominaisavaruus voidaan tulkita myös $\mathbb{R}^n$:n aliavaruudeksi.

Ominaisarvoihin liittyy kaksi erilaista kertaluvun käsitettä. Olkoon λ matriisin $\boldsymbol{A}$ ominaisarvo.

- Jos λ on $\boldsymbol{A}$:n karakteristisen polynomin k-kertainen juuri, niin $m_a(\lambda) = k$ on λ:n *algebrallinen kertaluku.*
- Ominaisarvon λ *geometrinen kertaluku* on $m_g(\lambda) = \dim(E_{\boldsymbol{A}}(\lambda))$.

Näillä on mm. seuraavat ominaisuudet:

- Jokaisella $\lambda \in \Lambda(\boldsymbol{A})$ selvästi pätee $m_a(\lambda) \geq 1$ ja $m_g(\lambda) \geq 1$.
- Edelleen $\sum_{\lambda \in \Lambda(\boldsymbol{A})} m_a(\lambda) = n$.

- Lauseesta 4.1 saadaan: $\sum_{\lambda\in\Lambda(\boldsymbol{A})} m_g(\lambda) \le n$.
- Myöhemmin osoitetaan: $m_a(\lambda) = \dim\Big(N\big((\lambda\boldsymbol{I}-\boldsymbol{A})^n\big)\Big)$, jolloin saadaan
$$m_g(\lambda) = \dim\big(N(\lambda\boldsymbol{I}-\boldsymbol{A})\big) \le m_a(\lambda)\ .$$

Esimerkin 4.4 kaikille ominaisarvoille pätee $m_a(\lambda) = m_g(\lambda) = 1$. Tämä on yleinen totuus (miksi?) silloin, kun $n\times n$ -matriisilla on n erisuurta ominaisarvoa.

Esimerkki 4.5. Matriisin
$$\boldsymbol{A} = \begin{bmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{bmatrix}$$
karakteristinen polynomi on $p(\lambda) = (\lambda+2)(\lambda-4)^2$. Siten ominaisarvot ovat $\lambda_1 = -2$ ja $\lambda_2 = \lambda_3 = 4$. Siis $m_a(-2) = 1$ ja $m_a(4) = 2$. Yksinkertaisen ominaisarvon geometrinen kertaluku on aina myös yksi, joten $m_g(-2) = 1$. Lasketaan tähän liittyvä ominaisvektori ratkaisemalla Gaussin eliminaatiolla $(-2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{v} = 0$:
$$-2\boldsymbol{I}-\boldsymbol{A} = \begin{bmatrix} -7 & -4 & -3 \\ 1 & -2 & 3 \\ -1 & 2 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 3 \\ 0 & -18 & 18 \\ 0 & 0 & 0 \end{bmatrix}\ ,$$
joten $\boldsymbol{v}^1 = (1,-1,-1)$ ja $E_{\boldsymbol{A}}(-2) = \mathrm{sp}(\boldsymbol{v}^1)$.

Moninkertaisen ominaisarvon geometrista kertalukua ei yleensä saa selville laskematta ominaisvektoreita (-avaruutta). Siis ratkaistaan $(4\,\boldsymbol{I}-\boldsymbol{A})\,\boldsymbol{v} = 0$:
$$4\boldsymbol{I}-\boldsymbol{A} = \begin{bmatrix} -1 & -4 & -3 \\ 1 & 4 & 3 \\ -1 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 4 & 3 \\ 0 & 0 & 0 \\ 0 & 6 & 6 \end{bmatrix}\ ,$$
joten ominaisarvoa 4 vastaavia lineaarisesti riippumattomia ominaisvektoreita löytyy yksi: $\boldsymbol{v}^2 = (1,-1,1)$ ja $E_{\boldsymbol{A}}(4) = \mathrm{sp}(\boldsymbol{v}^2)$. Siis $m_g(4) = 1 < 2 = m_a(4)$.

Näin 3×3 -matriisille löytyi vain kaksi lineaarisesti riippumatonta ominaisvektoria: $\boldsymbol{v}^1$, $\boldsymbol{v}^2$.

Monissa sovelluksissa matriisit ovat hermiittisiä ($\boldsymbol{A}^* = \boldsymbol{A}$) (reaalisessa tapauksessa symmetrisiä $\boldsymbol{A}^T = \boldsymbol{A}$). Kaikilla matriiseilla (ja lineaarikuvauksilla) erisuuriin ominaisarvoihin liittyvät ominaisvektorit ovat lineaarisesti riippumattomat. Hermiittisellä matriisilla ne ovat keskenään kohtisuorat[15].

Lause 4.3. *Hermiittisen matriisin ominaisarvot ovat reaaliset ja erisuuriin ominaisarvoihin liittyvät ominaisvektorit ovat ortogonaaliset.*

Tod. Olkoon λ hermiittisen matriisin A:n ominaisarvo ja $x \ne 0$ siihen liittyvä ominaisvektori. Tällöin yhtälöstä $\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$ saadaan sisätulon ominaisuuksilla
$$\lambda\,\langle\boldsymbol{x},\boldsymbol{x}\rangle = \langle\lambda\boldsymbol{x}\,,\,\boldsymbol{x}\rangle = \langle\boldsymbol{A}\boldsymbol{x}\,,\,\boldsymbol{x}\rangle = \langle\boldsymbol{x}\,,\,\boldsymbol{A}^*\boldsymbol{x}\rangle = \langle\boldsymbol{x}\,,\,\boldsymbol{A}\boldsymbol{x}\rangle = \langle\boldsymbol{x}\,,\,\lambda\boldsymbol{x}\rangle = \overline{\lambda}\,\langle\boldsymbol{x},\boldsymbol{x}\rangle\ \ ,$$

[15] Erisuuriin ominaisarvoihin liittyvien ominaisvektoreiden kohtisuoruus pätee kaikille matriiseille $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, joille pätee $\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A}$ eli jotka kommutoivat liittomatriisinsa kanssa. Tällaisia matriiseja kutsutaan *normaaleiksi.*

ja koska $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \neq 0$, on oltava $\lambda = \overline{\lambda}$ eli $\lambda \in \mathbb{R}$.

Jos $\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$ ja $\boldsymbol{A}\,\boldsymbol{y} = \mu\,\boldsymbol{y}$, missä $\lambda \neq \mu$, niin vastaavaan tapaan kuin edellä

$$\lambda\,\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \lambda \boldsymbol{x}\,,\, \boldsymbol{y} \rangle = \langle \boldsymbol{A}\boldsymbol{x}\,,\, \boldsymbol{y} \rangle = \langle \boldsymbol{x}\,,\, \boldsymbol{A}^* \boldsymbol{y} \rangle = \langle \boldsymbol{x}\,,\, \boldsymbol{A}\boldsymbol{y} \rangle = \langle \boldsymbol{x}\,,\, \mu \boldsymbol{y} \rangle = \overline{\mu}\,\langle \boldsymbol{x}, \boldsymbol{y} \rangle\ .$$

Nyt $\lambda = \overline{\lambda} \neq \overline{\mu}$, joten on oltava $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$. □

4.1. **Similaarisuus ja diagonalisointi.**

Neliömatriisi $\boldsymbol{A}$ on *similaarinen* matriisin $\boldsymbol{B}$ kanssa, jos on olemassa säännöllinen matriisi $\boldsymbol{S}$ siten, että $\boldsymbol{B} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{S}^{-1}$. Tällöin merkitään $\boldsymbol{A} \sim \boldsymbol{B}$. Muotoa $\boldsymbol{S}\boldsymbol{A}\boldsymbol{S}^{-1}$ olevaa matriisia kutsutaan $\boldsymbol{A}$:n *similaarimuunnokseksi.*

Tehtävä 4.4. Näytä, että $\sim$ on *ekvivalenssirelaatio* eli että sillä on ominaisuudet:

- [Refleksiivisyys:] $\boldsymbol{A} \sim \boldsymbol{A}$ kaikilla $\boldsymbol{A} \in \mathbb{K}^{n\times n}$.
- [Symmetrisyys:] $\boldsymbol{A} \sim \boldsymbol{B} \implies \boldsymbol{B} \sim \boldsymbol{A}$.
- [Transitiivisuus:] $\boldsymbol{A} \sim \boldsymbol{B}$ ja $\boldsymbol{B} \sim \boldsymbol{C} \implies \boldsymbol{A} \sim \boldsymbol{C}$.

Similaarisilla matriiseilla on monia yhteisiä ominaisuuksia.

Lause 4.4. *Keskenään similaarisilla matriiseilla on sama karakteristinen polynomi ja siten myös samat ominaisarvot samoine algebrallisine kertalukuineen. Myös ominaisarvojen geometriset kertaluvut ovat näillä samat.*

Tod. Olkoon $\boldsymbol{B} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{S}^{-1}$. Tällöin

$$\lambda \boldsymbol{I} - \boldsymbol{B} = \lambda\,\boldsymbol{S}\boldsymbol{I}\boldsymbol{S}^{-1} - \boldsymbol{S}\boldsymbol{A}\boldsymbol{S}^{-1} = \boldsymbol{S}\,(\lambda\,\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{S}^{-1}\ , \tag{4.6}$$

joten

$$\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = \det(\boldsymbol{S})\det(\lambda\,\boldsymbol{I} - \boldsymbol{A})\det(\boldsymbol{S}^{-1}) = \det(\lambda\,\boldsymbol{I} - \boldsymbol{A})\ ,$$

eli karakteristiset polynomit ovat samat. Koska $\boldsymbol{S}$ on inveroituva, (4.6):stä saadaan

$$(\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{v} = 0 \quad \iff \quad (\lambda\,\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{S}^{-1}\boldsymbol{v} = 0\ .$$

Siispä $E_{\boldsymbol{B}}(\lambda) = \left\{ \boldsymbol{v} = \boldsymbol{S}\boldsymbol{u} \,\middle|\, \boldsymbol{u} \in E_{\boldsymbol{A}}(\lambda) \right\}$. Koska invertoituva matriisi $\boldsymbol{S}$ kuvaa lineaarisesti riippumattomat vektorit lineaarisesti riippumattomiksi, saadaan $\dim(E_{\boldsymbol{B}}(\lambda)) = \dim(E_{\boldsymbol{A}}(\lambda))$. □

Kannanvaihdon muunnoskaavasta (2.3) nähdään, että lineaarioperaattorin matriisiesitykset eri kantojen suhteen ovat keskenään similaariset.

Olkoon matriisilla $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ ominaisarvot $\lambda_1, \ldots, \lambda_m$ ja näitä vastaavat ominaisvektorit $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^m$. Muodostetaan matriisi $\boldsymbol{X} = [\boldsymbol{x}^1\ \ldots\ \boldsymbol{x}^m]$. Tällöin saadaan

$$\boldsymbol{A}\,\boldsymbol{X} = [\boldsymbol{A}\,\boldsymbol{x}^1\ \ldots\ \boldsymbol{A}\,\boldsymbol{x}^m] = [\lambda_1\,\boldsymbol{x}^1\ \ldots\ \lambda_m\,\boldsymbol{x}^m] = [\boldsymbol{x}^1\ \ldots\ \boldsymbol{x}^m]\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix} = \boldsymbol{X}\,\boldsymbol{\Lambda}\ ,$$

missä merkittiin $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}$.

Jos tässä $m = n$ ja jos $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^m$ ovat lineaarisesti riippumattomat, niin $\boldsymbol{X}$ on invertoituva ja $\boldsymbol{A} = \boldsymbol{X}\,\boldsymbol{\Lambda}\,\boldsymbol{X}^{-1}$ eli $\boldsymbol{A}$ on similaarinen diagonaalimatriisin kanssa. Tällaista matriisia kutsutaan *diagonalisoituvaksi.*
Toisaalta, jos $\boldsymbol{A} = \boldsymbol{S}\,\boldsymbol{D}\,\boldsymbol{S}^{-1}$, missä $\boldsymbol{D}$ on diagonaalimatriisi, niin helposti nähdään, että $\boldsymbol{S}$:n sarakkeet ovat $\boldsymbol{A}$:n ominaisvektoreita, joten tällaiset muodostavat kannan. Saatiin:

Lause 4.5. *Matriisi on similaarinen diagonaalimatriisin kanssa täsmälleen silloin, kun sen ominaisvektoreista voidaan muodostaa kanta. Näin on erityisesti silloin, kun ominaisarvot ovat erisuuret.*

Jos $\boldsymbol{A}$:lla on moninkertaisia ominaisarvoja ja kaikille ominaisarvoille pätee $m_g(\lambda) = m_a(\lambda)$, niin voimme muodostaa ensin kullekin $E_{\boldsymbol{A}}(\lambda)$:lle kannan, joista yhteensä saadaan

$$\textstyle\sum_{\lambda\in\Lambda(\boldsymbol{A})} m_g(\lambda) = \sum_{\lambda\in\Lambda(\boldsymbol{A})} m_a(\lambda) = n$$

lineaarisesti riippumatonta ominaisvektoria, joten ne muodostavat kannan. Edellisen lauseen voi siten kirjoittaa myös muotoon:

Seuraus 4.6. *Matriisi on similaarinen diagonaalimatriisin kanssa täsmälleen silloin, kun kaikille ominaisarvoille pätee* $m_g(\lambda) = m_a(\lambda)$.

Matriisin saattaminen similaarimuunnoksin yksinkertaiseen (parhaassa tapauksessa diagonaaliseen) muotoon on kätevää monissa dynamiikan tehtävissä, kuten jatkossa tullaan näkemään.

Huomautus 4.1. Oletetaan, että matriisilla $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ on erisuuret ominaisarvot, joten se on diagonalisoituva. Jos $\boldsymbol{A}$:lla on kompleksisia ominaisarvoja, tällöin sekä $\boldsymbol{\Lambda}$ että muunnosmatriisi $\boldsymbol{X}$ ovat kompleksisia. Usein on kuitenkin mukava pysytellä reaalisissa matriiseissa. Seuraava muoto saattaa tällöin olla käyttökelpoinen.

Koska $\boldsymbol{A}$ on reaalinen, sen kompleksiset ominaisarvot ja -vektorit esiintyvät konjugaattipareina: $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ ja $\boldsymbol{A}\overline{\boldsymbol{x}} = \overline{\lambda}\,\overline{\boldsymbol{x}}$. Merkitään $\lambda = \alpha + i\beta$ ja $\boldsymbol{x} = \boldsymbol{u} + i\,\boldsymbol{v}$. Tällöin $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ ovat lineaarisesti riippumattomat (koska $\boldsymbol{x}$ ja $\overline{\boldsymbol{x}}$ ovat) ja

$$\boldsymbol{A}(\boldsymbol{u} + i\boldsymbol{v}) = (\alpha + i\beta)\,(\boldsymbol{u} + i\boldsymbol{v}) = \alpha\boldsymbol{u} - \beta\boldsymbol{v} + i\,(\beta\boldsymbol{u} + \alpha\boldsymbol{v})\ .$$

Poimimalla tästä reaali- ja imaginääriosat saadaan $\boldsymbol{A}\boldsymbol{u} = \alpha\boldsymbol{u} - \beta\boldsymbol{v}$ ja $\boldsymbol{A}\boldsymbol{v} = \beta\boldsymbol{u} + \alpha\boldsymbol{v}$ eli

$$\boldsymbol{A}\,[\boldsymbol{u}\ \boldsymbol{v}] = [\boldsymbol{u}\ \boldsymbol{v}]\left[\begin{smallmatrix} \alpha & \beta \\ -\beta & \alpha \end{smallmatrix}\right]\ .$$

Oletetaan, että matriisissa $\boldsymbol{X}$ tällaiset kompleksiset ominaisvektoriparit esiintyvät peräkkäin. Korvataan ne vastaavilla $\boldsymbol{u}, \boldsymbol{v}$-pareilla. Olkoon näin saatu matriisi $\widehat{\boldsymbol{X}} \in \mathbb{R}^{n\times n}$ (reaalisia ominaisarvoja vastaavat ominaisvektorit voidaan heti valita reaalisiksi). Tällöin $\boldsymbol{A}\widehat{\boldsymbol{X}} = \widehat{\boldsymbol{X}}\widehat{\boldsymbol{\Lambda}}$ eli $\boldsymbol{A} = \widehat{\boldsymbol{X}}\widehat{\boldsymbol{\Lambda}}\widehat{\boldsymbol{X}}^{-1}$, missä $\widehat{\boldsymbol{\Lambda}}$ on *lohkodiagonaalimatriisi* $\widehat{\boldsymbol{\Lambda}} = \left[\begin{smallmatrix} \boldsymbol{L}_1 & & \\ & \ddots & \\ & & \boldsymbol{L}_k \end{smallmatrix}\right]$, ja kukin $\boldsymbol{L}_j$ on joko reaaliluku tai muotoa $\left[\begin{smallmatrix} \alpha & \beta \\ -\beta & \alpha \end{smallmatrix}\right]$ oleva matriisi.

4.2. **Schurin hajotelma.**

Tässä todistamme, että jokainen matriisi on similaarinen yläkolmiomatriisin kanssa, vieläpä niin, että tämän similaarimuunnoksen suorittava matriisi on unitaarinen. Tämä nk. Schur-dekompositio on lähtökohta moniin muihin käyttökelpoisiin hajotelmiin.

Lause 4.7 (Schurin hajotelma). *Jokaiselle* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *on olemassa unitaarinen* $\boldsymbol{Q}$ *siten, että* $\boldsymbol{S} = \boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{Q}$ *on yläkolmiomatriisi. Tässä* $\boldsymbol{S}$ *:n diagonaalialkiot (* $\boldsymbol{A}$ *:n ominaisarvot) voidaan asettaa mielivaltaiseen järjestykseen.*

Tästä saatavaa esitystä $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{S}\boldsymbol{Q}^*$ $(= \boldsymbol{Q}\boldsymbol{S}\boldsymbol{Q}^{-1})$ kutsutaan $\boldsymbol{A}$:n Schurin hajotelmaksi (dekompositioksi).

Tod. Induktiolla n:n suhteen: tapaus $n = 1$ on selvä.

Oletetaan, että lause pätee $(n-1)\times(n-1)$ matriiseille. Olkoon λ matriisin $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ ominaisarvo ja $\boldsymbol{x}$ vastaava ominaisvektori, jolle $\|\boldsymbol{x}\|_2 = 1$. Olkoon matriisi $\boldsymbol{V} \in \mathbb{C}^{n\times(n-1)}$ siten, että $\boldsymbol{U} = [\boldsymbol{x}\ \boldsymbol{V}]$ on unitaarinen (eli täydennetään $\boldsymbol{x}$ avaruuden $\mathbb{C}^n$ ortonormaaliksi kannaksi). Tällöin $\boldsymbol{V}^*\boldsymbol{x} = 0$, joten[16]

$$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} = \begin{bmatrix} \boldsymbol{x}^* \\ \boldsymbol{V}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{A}\boldsymbol{x}\ \boldsymbol{A}\boldsymbol{V} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}^* \\ \boldsymbol{V}^* \end{bmatrix} \begin{bmatrix} \lambda\boldsymbol{x}\ \boldsymbol{A}\boldsymbol{V} \end{bmatrix} = \begin{bmatrix} \lambda & \# \\ 0 & \boldsymbol{C} \end{bmatrix},$$

missä $\boldsymbol{C} = \boldsymbol{V}^*\boldsymbol{A}\boldsymbol{V} \in \mathbb{C}^{(n-1)\times(n-1)}$. Koska lause pätee matriisille $\boldsymbol{C}$, on olemassa unitaarinen $\widetilde{\boldsymbol{U}}$ siten, että $\widetilde{\boldsymbol{U}}^*\boldsymbol{C}\widetilde{\boldsymbol{U}}$ on yläkolmiomatriisi. Asetetaan $\boldsymbol{Q} = \boldsymbol{U}\begin{bmatrix} 1 & \\ & \widetilde{\boldsymbol{U}} \end{bmatrix}$, jolloin $\boldsymbol{Q}$ on unitaarinen ja

$$\boldsymbol{S} = \boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{Q} = \begin{bmatrix} 1 & \\ & \widetilde{\boldsymbol{U}} \end{bmatrix}^* \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} \begin{bmatrix} 1 & \\ & \widetilde{\boldsymbol{U}} \end{bmatrix} = \begin{bmatrix} \lambda & \# \\ 0 & \widetilde{\boldsymbol{U}}^*\boldsymbol{C}\widetilde{\boldsymbol{U}} \end{bmatrix}$$

on yläkolmiomatriisi. Näin lause on voimassa myös $n \times n$ matriiseille. □

Huomautus 4.2. Olkoon $\boldsymbol{H} = \boldsymbol{I} - 2\,\frac{\boldsymbol{v}\boldsymbol{v}^*}{\boldsymbol{v}^*\boldsymbol{v}}$ Householderin muunnos siten, että $\boldsymbol{H}\boldsymbol{x} = \boldsymbol{e}^1$. Tämä $\boldsymbol{H}$ kelpaa edellisen todistuksen matriisiksi $\boldsymbol{U}$, sillä se on unitaarinen ja $\boldsymbol{H}\boldsymbol{e}^1 = \boldsymbol{H}^2\boldsymbol{x} = \boldsymbol{x}$, joten $\boldsymbol{x}$ on $\boldsymbol{H}$:n ensimmäinen sarake.

Seuraus 4.8. *Hermiittinen matriisi on diagonalisoituva unitaarisella similaarimuunnoksella eli sen ominaisvektoreista voidaan muodostaa ortonormaali kanta.*

Tod. Jos $\boldsymbol{A}^* = \boldsymbol{A}$ ja $\boldsymbol{S} = \boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{Q}$ on sen Schurin muoto, niin $\boldsymbol{S}^* = (\boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{Q})^* = \boldsymbol{Q}^*\boldsymbol{A}^*(\boldsymbol{Q}^*)^* = \boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{S}$. Siis $\boldsymbol{S}$ on hermiittinen yläkolmiomatriisi, joten se on (reaalinen) diagonaalimatriisi. □

Seuraus 4.9. *Matriisin determinantti on sen ominaisarvojen tulo.*

Tod. Schurin hajotelmasta $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{S}\boldsymbol{Q}^*$ saadaan: $\det(\boldsymbol{A}) = \det(\boldsymbol{Q})\det(\boldsymbol{S})\det(\boldsymbol{Q})^{-1} = \det(\boldsymbol{S})$ ja tämähän on $\boldsymbol{S}$:n diagonaalialkioiden eli $\boldsymbol{A}$:n ominaisarvojen tulo. □

[16] Merkinnällä # tarkoitamme matriisin osaa, josta emme välitä mitä se on ja siksi emme sitä tarkemmin laske.

Matriisin ominaisarvot muuntuvat luonnollisella tavalla, kun matriisia kerrotaan tai korotetaan potenssiin. Selvästi

$$\Lambda(\alpha \boldsymbol{A}) = \alpha\, \Lambda(\boldsymbol{A}) = \big\{\alpha\lambda \,\big|\, \lambda \in \Lambda(\boldsymbol{A})\big\} \ .$$

Jos $(\lambda, \boldsymbol{x})$ on $\boldsymbol{A}$:n ominaispari, niin $\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}(\lambda\boldsymbol{x}) = \lambda\, \boldsymbol{A}\boldsymbol{x} = \lambda^2\boldsymbol{x}$ ja

$$\begin{aligned}\boldsymbol{A}^k\boldsymbol{x} &= \boldsymbol{A}^{k-1}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}^{k-1}(\lambda\boldsymbol{x}) = \lambda\, \boldsymbol{A}^{k-1}\boldsymbol{x} \\ &= \lambda^2\, \boldsymbol{A}^{k-2}\boldsymbol{x} = \cdots = \lambda^{k-1}\, \boldsymbol{A}\boldsymbol{x} = \lambda^k\, \boldsymbol{x} \ .\end{aligned}$$

Edelleen, jos $\boldsymbol{A}$ on invertoituva, kerrotaan yhtälö $\lambda\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}$ matriisilla $\boldsymbol{A}^{-1}$:

$$\lambda\boldsymbol{A}^{-1}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}\ , \qquad \text{joten} \qquad \boldsymbol{A}^{-1}\boldsymbol{x} = \lambda^{-1}\boldsymbol{x}\ .$$

Siispä $\big\{1/\lambda \,\big|\, \lambda \in \Lambda(\boldsymbol{A})\big\} \subset \Lambda(\boldsymbol{A}^{-1})$. Yhdistetään tämä edelliseen tulokseen, jolloin saadaan: $\big\{\lambda^k \,\big|\, \lambda \in \Lambda(\boldsymbol{A})\big\} \subset \Lambda(\boldsymbol{A}^k)$ kaikilla $k \in \mathbb{Z}$ ($k \geq 1$, jos $\boldsymbol{A}$ on singulaarinen). Itse asiassa tässä pätee joukkojen yhtäsuuruus. Tämä jätetään harjoitustehtäväksi seuraavien tarkastelujen jälkeen.

Jos $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}$, niin $\boldsymbol{A}^2 = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}\boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{B}^2\boldsymbol{S}^{-1}$ ja

$$\boldsymbol{A}^k = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}\boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}\ldots\boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{B}^k\boldsymbol{S}^{-1}\ .$$

Annetulle polynomille $p(z) = \alpha_k z^k + \alpha_{k-1}z^{k-1} + \cdots + \alpha_1 z + \alpha_0$ määritellään

$$p(\boldsymbol{A}) = \alpha_k\boldsymbol{A}^k + \alpha_{k-1}\boldsymbol{A}^{k-1} + \cdots + \alpha_1\boldsymbol{A} + \alpha_0\boldsymbol{I}\ ,$$

jolloin tapauksessa $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}$ saadaan

$$\begin{aligned}p(\boldsymbol{A}) &= \alpha_k\boldsymbol{S}\boldsymbol{B}^k\boldsymbol{S}^{-1} + \alpha_{k-1}\boldsymbol{S}\boldsymbol{B}^{k-1}\boldsymbol{S}^{-1} + \cdots + \alpha_1\boldsymbol{S}\boldsymbol{B}^k\boldsymbol{S}^{-1} + \alpha_0\boldsymbol{S}\boldsymbol{I}\boldsymbol{S}^{-1} \\ &= \boldsymbol{S}\left(\alpha_k\boldsymbol{B}^k + \alpha_{k-1}\boldsymbol{B}^{k-1} + \cdots + \alpha_1\boldsymbol{B} + \alpha_0\boldsymbol{I}\right)\boldsymbol{S}^{-1} \\ &= \boldsymbol{S}\, p(\boldsymbol{B})\, \boldsymbol{S}^{-1}\ .\end{aligned}$$

Saatiin, että matriisit $p(\boldsymbol{A})$ ja $p(\boldsymbol{B})$ ovat myös similaariset ja muunnoksen välittää sama matriisi $\boldsymbol{S}$.

Tehtävä 4.5 (Spektraalikuvauslause)**.** Olkoon p mielivaltainen polynomi. Tällöin pätee

$$\Lambda(p(\boldsymbol{A})) = p(\Lambda(\boldsymbol{A})) = \big\{p(\lambda) \,\big|\, \lambda \in \Lambda(\boldsymbol{A})\big\}\ .$$

Vihje: näytä tämä ensin yläkolmiomatriiseille ja sovella sitten Schurin hajotelmaa ja yllä olevaa tulosta $p(\boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}) = \boldsymbol{S}\, p(\boldsymbol{B})\, \boldsymbol{S}^{-1}$.

Tehtävä 4.6. Yleistä edellisen tehtävän tulos rationaalifunktioille $r(z) = p(z)/q(z)$, kun ensin määrittelet tämän sopivasti matriisimuuttujan tapaukseen.

Matriisin potenssien ja muiden funktioiden laskeminen similaarimuunnoksen avulla on ominaisarvoteorian sovellusten eräs tärkeimpiä tekniikoita.

Esimerkki 4.6. Ihmisillä on taipumus helpommin muuttaa keskustasta esikaupunkialueelle kuin päinvastoin. Oletetaan, että K:n kaupungissa muuttamistodennäköisyydet vuoden aikana ovat:

- keskustasta esikaupunkiin: 0.1 ja

- esikaupungista keskustaan: 0.04 .

Vuoden 2000 alussa keskustassa asui 60% väestöstä ja loput esikaupunkialueella. Miten tämä prosenttiluku muuttuu? Tyhjeneekö keskusta?

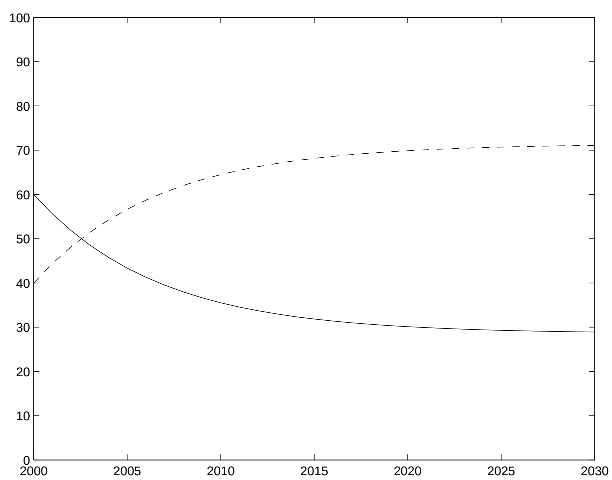
Merkitään:

$$\begin{array}{cccl} & \text{nyt} & \text{nyt} & \\ & \text{keskustassa} & \text{esikaup.} & \\ \boldsymbol{A} = & \left[\begin{array}{c} 0.9 \\ 0.1 \end{array}\right. & \left.\begin{array}{c} 0.04 \\ 0.96 \end{array}\right] & \begin{array}{l} \text{ensi vuonna keskustassa} \\ \text{ensi vuonna esikaupungissa.} \end{array} \end{array}$$

Näin, kun $\boldsymbol{x}^{2000} = \begin{bmatrix} 60 \\ 40 \end{bmatrix}$, tulevina vuosina saadaan $\boldsymbol{x}^{v+1} = \boldsymbol{A}\,\boldsymbol{x}^v$ eli (pyöristettynä)

$$\boldsymbol{x}^{2000} = \begin{bmatrix} 60 \\ 40 \end{bmatrix}, \quad \boldsymbol{x}^{2001} = \begin{bmatrix} 55.6 \\ 44.4 \end{bmatrix}, \quad \boldsymbol{x}^{2002} = \begin{bmatrix} 51.8 \\ 48.2 \end{bmatrix}, \quad \boldsymbol{x}^{2003} = \begin{bmatrix} 45.8 \\ 54.2 \end{bmatrix}, \quad \boldsymbol{x}^{2004} = \begin{bmatrix} 43.4 \\ 56.6 \end{bmatrix},$$

ja käyrinä:



Nämä näyttävät tasoittuvan noin $\begin{bmatrix} 29 \\ 71 \end{bmatrix}$ -tasoille. Onko olemassa $\lim_{v\to\infty} \boldsymbol{x}^v$? $\boldsymbol{A}$:lla on ominaisarvot

$$\begin{vmatrix} \lambda - 0.9 & -0.04 \\ -0.1 & \lambda - 0.96 \end{vmatrix} = \lambda^2 - 1.86\lambda + 0.86 = 0 \qquad \text{eli} \qquad \begin{cases} \lambda_1 = 1 \ \text{ ja} \\ \lambda_2 = 0.86 \ , \end{cases}$$

sekä ominaisvektorit $\boldsymbol{v}^1 = \left[\begin{smallmatrix} 2 \\ 5 \end{smallmatrix}\right]$, $\boldsymbol{v}^2 = \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. Siispä $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$, missä $\boldsymbol{V} = \left[\begin{smallmatrix} 2 & 1 \\ 5 & -1 \end{smallmatrix}\right]$ ja $\boldsymbol{\Lambda} = \left[\begin{smallmatrix} 1 & \\ & 0.86 \end{smallmatrix}\right]$. Näin $\boldsymbol{A}^k = \boldsymbol{V}\boldsymbol{\Lambda}^k\boldsymbol{V}^{-1}$ ja

$$\lim_{v\to\infty} \boldsymbol{x}^v = \lim_{k\to\infty} \boldsymbol{A}^k \boldsymbol{x}^{2000} = \begin{bmatrix} 2 & 1 \\ 5 & -1 \end{bmatrix} \begin{bmatrix} 1 & \\ & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 60 \\ 40 \end{bmatrix} = \tfrac{100}{7} \begin{bmatrix} 2 \\ 5 \end{bmatrix} \approx \begin{bmatrix} 28.6 \\ 71.4 \end{bmatrix} .$$

Esimerkki 4.7. *Fibonacci*–luvut f_k , $k \geq 0$, määritellään $f_0 = 1$, $f_1 = 1$ ja $f_{k+1} = f_k + f_{k-1}$ eli saadaan lukujono

$$1\,,\ 1\,,\ 2\,,\ 3\,,\ 5\,,\ 8\,,\ 13\,,\ 21\,,\ 34\,,\ 55\,,\ldots$$

Halutaan lausua f_k k:n funktiona. Olkoon

$$\boldsymbol{x}^k = \begin{bmatrix} f_k \\ f_{k-1} \end{bmatrix}, \qquad \text{jolloin} \qquad \boldsymbol{x}^{k+1} = \begin{bmatrix} f_{k+1} \\ f_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_k \\ f_{k-1} \end{bmatrix} = \boldsymbol{A}\boldsymbol{x}^k ,$$

missä $\boldsymbol{A} = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. Siispä $\boldsymbol{x}^k = \boldsymbol{A}^{k-1}\boldsymbol{x}^1$, missä $\boldsymbol{x}^1 = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.

Lasketaan $\boldsymbol{A}$:n ominaisarvot ja -ominaisvektorit:

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda - 1 & -1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 - \lambda - 1 .$$

Saadaan ominaisarvot $\lambda_1 = \frac{1}{2}(1+\sqrt{5})$, $\lambda_2 = \frac{1}{2}(1-\sqrt{5})$, ja ominaisvektorit $\boldsymbol{v}^j = \left[\begin{smallmatrix} \lambda_j \\ 1 \end{smallmatrix}\right]$. Näin, asettamalla $\boldsymbol{S} = [\boldsymbol{v}^1 \ \boldsymbol{v}^2] = \left[\begin{smallmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{smallmatrix}\right]$ ja $\boldsymbol{\Lambda} = \left[\begin{smallmatrix} \lambda_1 & \\ & \lambda_2 \end{smallmatrix}\right]$, saadaan $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$ ja $\boldsymbol{A}^k = \boldsymbol{S}\boldsymbol{\Lambda}^k\boldsymbol{S}^{-1}$. Näillä saadaan:

$$\begin{aligned} \boldsymbol{x}^k = \boldsymbol{A}^{k-1}\boldsymbol{x}^1 &= \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^{k-1} & \\ & \lambda_2^{k-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1^k(1-\lambda_2) + \lambda_2^k(\lambda_1 - 1) \\ \lambda_1^{k-1}(1-\lambda_2) + \lambda_2^{k-1}(\lambda_1 - 1) \end{bmatrix} , \end{aligned}$$

ja (käyttäen: $\lambda_1 + \lambda_2 = 1$) lopulta

$$f_k = (\boldsymbol{x}^k)_1 = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2} = \frac{1}{2^{k+1}\sqrt{5}} \left((1+\sqrt{5})^{k+1} - (1-\sqrt{5})^{k+1} \right) .$$

Tehtävä 4.7. Olkoon $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{X} & \# \\ 0 & \boldsymbol{Y} \end{bmatrix} \in \mathbb{C}^{n\times n}$, missä $\boldsymbol{X}$ ja $\boldsymbol{Y}$ ovat sopivan kokoisia neliömatriiseja. Olkoon p mielivaltainen polynomi. Näytä, että $p(\boldsymbol{A}) = \begin{bmatrix} p(\boldsymbol{X}) & \# \\ 0 & p(\boldsymbol{Y}) \end{bmatrix}$.

Tehtävä 4.8. Näytä: $m_a(\lambda) = \dim\Big(N\big((\lambda\boldsymbol{I} - \boldsymbol{A})^n\big)\Big)$. Vihje: oleta ensin, että $\boldsymbol{A}$ on yläkolmiomatriisi.

Lause 4.10 (Cayley-Hamilton)**.** *Olkoon $p_{\boldsymbol{A}}(z) = \det(z\boldsymbol{I} - \boldsymbol{A})$ matriisin $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ karakteristinen polynomi. Tällöin $p_{\boldsymbol{A}}(\boldsymbol{A}) = 0$.*

Tod. Tehdään Schurin hajotelma $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{S}\boldsymbol{Q}^*$, jolloin $p_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{Q}\, p_{\boldsymbol{A}}(\boldsymbol{S})\, \boldsymbol{Q}^*$ ja toisaalta lauseen 4.4 perusteella $p_{\boldsymbol{A}} = p_{\boldsymbol{S}}$. Riittää siis näyttää, että $p_{\boldsymbol{S}}(\boldsymbol{S}) = 0$.

Todistetaan siis, että lause pätee yläkolmiomatriiseille. Tehdään tämä induktiolla: tapaus $n = 1$: $p_s(z) = z - s$, joten $p_s(s) = 0$.

Oletetaan, että lause pätee $(n-1)\times(n-1)$ yläkolmiomatriiseille. Olkoon

$$\boldsymbol{S} = \begin{bmatrix} s & \# \\ 0 & \widetilde{\boldsymbol{S}} \end{bmatrix} \in \mathbb{C}^{n\times n}$$

yläkolmiomuotoa. Tällöin $\widetilde{\boldsymbol{S}} \in \mathbb{C}^{(n-1)\times(n-1)}$ on myös yläkolmiomatriisi, joten induktio-oletuksen mukaan $p_{\widetilde{\boldsymbol{S}}}(\widetilde{\boldsymbol{S}}) = 0$. Nyt pätee

$$p_{\boldsymbol{S}}(z) = \prod_{\lambda\in\Lambda(\boldsymbol{S})} (z-\lambda) = (z-s)\, p_{\widetilde{\boldsymbol{S}}}(z) \ .$$

Siis $p_{\boldsymbol{S}}(\boldsymbol{S}) = (\boldsymbol{S} - s\boldsymbol{I})\, p_{\widetilde{\boldsymbol{S}}}(\boldsymbol{S})$. Tehtävän 4.7 perusteella saadaan siten

$$p_{\boldsymbol{S}}(\boldsymbol{S}) = \begin{bmatrix} s-s & \# \\ 0 & \widetilde{\boldsymbol{S}} - s\boldsymbol{I} \end{bmatrix} \begin{bmatrix} p_{\widetilde{\boldsymbol{S}}}(s) & \# \\ 0 & p_{\widetilde{\boldsymbol{S}}}(\widetilde{\boldsymbol{S}}) \end{bmatrix} = \begin{bmatrix} 0 & \# \\ 0 & \widetilde{\boldsymbol{S}} - s\boldsymbol{I} \end{bmatrix} \begin{bmatrix} p_{\widetilde{\boldsymbol{S}}}(s) & \# \\ 0 & 0 \end{bmatrix} = 0 \ .$$

Näin tulos pätee myös $n \times n$ yläkolmiomatriiseille ja todistuksen alun mukaan kaikille neliömatriiseille. □

Tehtävä 4.9. Olkoon $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ invertoituva. Näytä, että on olemassa kertoimet $c_0, \ldots, c_{n-1}$ siten, että $\boldsymbol{A}^{-1} = c_0\, \boldsymbol{I} + c_1\, \boldsymbol{A} + \cdots + c_{n-1}\, \boldsymbol{A}^{n-1}$.

Tehtävä 4.10. Olkoon $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ ja $\mathbb{P}_\infty(\boldsymbol{A}) = \{ p(\boldsymbol{A}) \,|\, p \in \mathbb{P}_\infty \}$ eli kaikkien $\boldsymbol{A}$:n polynomien joukko. Näytä, että $\dim(\mathbb{P}_\infty(\boldsymbol{A})) \le n$.

Esimerkki 4.8. Tarkastellaan tähänastisen teorian ei-triviaalina sovelluksena nk. *Sylvesterin yhtälöä*

$$\boldsymbol{AX} - \boldsymbol{XB} = \boldsymbol{C} \ . \tag{4.7}$$

Tässä matriisit $\boldsymbol{A} \in \mathbb{R}^{m\times m}$, $\boldsymbol{B} \in \mathbb{R}^{n\times n}$, $\boldsymbol{C} \in \mathbb{R}^{m\times n}$ on annettu ja yhtälöstä pitäisi ratkaista matriisi $\boldsymbol{X} \in \mathbb{R}^{m\times n}$. Tällaisia yhtälöitä esiintyy mm. säätöteorian yhteydessä.

Näytetään, että yhtälöllä on olemassa yksikäsitteinen ratkaisu, jos matriiseilla $\boldsymbol{A}$ ja $\boldsymbol{B}$ ei ole yhteisiä ominaisarvoja, eli kun $\Lambda(\boldsymbol{A}) \cap \Lambda(\boldsymbol{B}) = \emptyset$.

Kuvaus $\boldsymbol{X} \mapsto T(\boldsymbol{X}) = \boldsymbol{AX} - \boldsymbol{XB}$ on lineaarikuvaus $T : \mathbb{R}^{m\times n} \mapsto \mathbb{R}^{m\times n}$. Siis aiemmin olleen harjoitustehtävän nojalla T on invertoituva, jos sen ydin on triviaali, eli $N(T) = \{0\}$. Näytetään tämä. Olkoon $\boldsymbol{X} \in \mathbb{R}^{m\times n}$ siten, että $T(\boldsymbol{X}) = 0$. Tällöin $\boldsymbol{AX} = \boldsymbol{XB}$, josta $\boldsymbol{A}^2\boldsymbol{X} = \boldsymbol{A}(\boldsymbol{AX}) = \boldsymbol{AXB} = \boldsymbol{XB}^2$ ja (induktiivisesti) $\boldsymbol{A}^k\boldsymbol{X} = \boldsymbol{XB}^k$ kaikilla $k \ge 1$. Laskemalla näitä yhteen saadaan: $p(\boldsymbol{A})\boldsymbol{X} = \boldsymbol{X}\, p(\boldsymbol{B})$ jokaisella polynomilla p . Olkoon $p_{\boldsymbol{B}}$ nyt $\boldsymbol{B}$:n karakteristinen polynomi, jolloin Cayley-Hamiltonin lauseen mukaan $p_{\boldsymbol{B}}(\boldsymbol{B}) = 0$. Siispä $p_{\boldsymbol{B}}(\boldsymbol{A})\boldsymbol{X} = 0$. Nyt spektraalikuvauslauseen perusteella $\Lambda(p_{\boldsymbol{B}}(\boldsymbol{A})) = \{ p_{\boldsymbol{B}}(\lambda) \,|\, \lambda \in \Lambda(\boldsymbol{A}) \}$. Koska $\boldsymbol{A}$:lla ja $\boldsymbol{B}$:llä ei ole yhteisiä ominaisarvoja, saadaan $p_{\boldsymbol{B}}(\lambda) \ne 0$ kaikilla $\lambda \in \Lambda(\boldsymbol{A})$. Siispä $0 \notin \Lambda(p_{\boldsymbol{B}}(\boldsymbol{A}))$, joten $p_{\boldsymbol{B}}(\boldsymbol{A})$ on invertoituva ja on oltava $\boldsymbol{X} = 0$. Näin ollen T:n ydin on triviaali, joten Sylvesterin yhtälöllä on olemassa yksikäsitteinen ratkaisu.

Kurssilla Numeerinen matriisilaskenta tullaan käsittelemään menetelmiä, joilla Sylvesterin yhtälön voi numeerisesti ratkaista. Nämä perustuvat matriisien $\boldsymbol{A}$ ja $\boldsymbol{B}$ Schurin hajotelmiin.

Tehtävä 4.11. Olkoon $\boldsymbol{M} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{C} \\ 0 & \boldsymbol{B} \end{bmatrix}$, missä $\boldsymbol{A}$ ja $\boldsymbol{B}$ ovat neliömatriiseja ja $\Lambda(\boldsymbol{A}) \cap \Lambda(\boldsymbol{B}) = \emptyset$. Näytä, että $\boldsymbol{M} \sim \begin{bmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{B} \end{bmatrix}$.

Vihje: etsi similaarimuunnosta muodossa $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{X} \\ 0 & \boldsymbol{I} \end{bmatrix}$.

4.3. **Jordanin muoto.** Kaikki matriisit eivät diagonalisoidu, mutta kaikki matriisit ovat similaarisia yläkolmiomatriisin kanssa. Kuinka yksinkertaiseen muotoon matriisin voi similaarimuuntaa? Erityisen mielenkiinnon kohteena ovat tässä matriisit, joille jonkin ominaisarvon geometrinen kertaluku on pienempi kuin algebrallinen, sillä muuthan diagonalisoituvat.

Jos matriisin $\boldsymbol{A}$ ominaisarvon λ geometrinen ja algebrallinen kertaluku ($= k$) yhtyvät, poimitaan ominaisavaruudesta $E_{\boldsymbol{A}}(\lambda) = N(\boldsymbol{A} - \lambda \boldsymbol{I})$ kanta $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k$, jolloin

$$\boldsymbol{A}[\boldsymbol{x}^1 \ \ldots \ \boldsymbol{x}^k] = [\boldsymbol{x}^1 \ \ldots \ \boldsymbol{x}^k] \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix} . \tag{4.8}$$

Jos $\lambda_1, \ldots, \lambda_q$ ovat matriisin $\boldsymbol{A}$ erisuuret ominaisarvot ja $k_1, \ldots, k_q$ näiden algebralliset kertaluvut, niin $\boldsymbol{A}$:n karakteristinen polynomi voidaan kirjoittaa: $p_{\boldsymbol{A}}(z) = \prod_{j=1}^{q} (z - \lambda_j)^{k_j}$ ja Cayley-Hamiltonin lause sanoo, että

$$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})^{k_1} (\boldsymbol{A} - \lambda_2 \boldsymbol{I})^{k_2} \ldots (\boldsymbol{A} - \lambda_q \boldsymbol{I})^{k_q} = 0 .$$

Määritellään kullekin ominaisarvolle *invariantti aliavaruus*

$$\widehat{E}_{\boldsymbol{A}}(\lambda_j) = \left\{ \boldsymbol{v} \in \mathbb{C}^n \,\middle|\, (\boldsymbol{A} - \lambda_j \boldsymbol{I})^{k_j} \boldsymbol{v} = 0 \right\} = N\big((\boldsymbol{A} - \lambda_j \boldsymbol{I})^{k_j} \big) .$$

Selvästi $E_{\boldsymbol{A}}(\lambda_j) \subset \widehat{E}_{\boldsymbol{A}}(\lambda_j)$. Jos ominaisarvon λ_j geometriset ja algebralliset kertaluvut ovat samat, niin osoittautuu, että nämä avaruudet ovat samat (ks. tehtävä 4.8).

Olkoon λ $\boldsymbol{A}$:n ominaisarvo, jolle $E_{\boldsymbol{A}}(\lambda) \neq \widehat{E}_{\boldsymbol{A}}(\lambda)$. Tällöin on olemassa $\kappa \geq 2$ ja $\boldsymbol{w}^\kappa \in \widehat{E}_{\boldsymbol{A}}(\lambda)$ siten, että $(\boldsymbol{A} - \lambda \boldsymbol{I})^\kappa \boldsymbol{w}^\kappa = 0$, mutta $(\boldsymbol{A} - \lambda \boldsymbol{I})^{\kappa - 1} \boldsymbol{w}^\kappa \neq 0$. Asetetaan

$$\boldsymbol{w}^j = (\boldsymbol{A} - \lambda \boldsymbol{I})^{\kappa - j} \boldsymbol{w}^\kappa .$$

Tällöin

$$\boldsymbol{A}\boldsymbol{w}^1 = \lambda \boldsymbol{w}^1 \qquad \text{ja} \qquad \boldsymbol{A}\boldsymbol{w}^j = \lambda \boldsymbol{w}^j + \boldsymbol{w}^{j-1}, \ j = 2, \ldots, \kappa .$$

Saadaan matriisimuodossa

$$\boldsymbol{A}[\boldsymbol{w}^1 \ \boldsymbol{w}^2 \ \ldots \ \boldsymbol{w}^\kappa] = [\boldsymbol{w}^1 \ \boldsymbol{w}^2 \ \ldots \ \boldsymbol{w}^\kappa] \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} . \tag{4.9}$$

Tällaista jonoa $(\boldsymbol{w}^1, \ldots, \boldsymbol{w}^\kappa)$ kutsutaan ominaisarvoon λ liittyväksi ja ominaisvektorista $\boldsymbol{w}^1$ lähteväksi *ketjuksi*. Tämän vektorit ovat lineaarisesti riippumattomat (harjoitustehtävä). Osoittautuu (ei tod.) että kuhunkin ominaisarvoon λ voidaan liittää $m_g(\lambda)$ kappaletta tällaisia ketjuja, joiden yhteenlaskettu pituus on $m_a(\lambda)$.

Esimerkki 4.9. Matriisin $\boldsymbol{M} = \begin{bmatrix} 3 & 1 & & & & \\ & 3 & & & & \\ & & 3 & 1 & & \\ & & & 3 & 1 & \\ & & & & 3 & \\ & & & & & 3 \end{bmatrix}$ ainoa ominaisarvo on $\lambda = 3$ (joten $m_a(3) = 6$). Tällä on kolme lineaarisesti riippumatonta ominaisvektoria $\boldsymbol{e}^1$, $\boldsymbol{e}^3$ ja $\boldsymbol{e}^6$

(siis $m_g(3) = 3$). Näistä lähtee ketjut $(\boldsymbol{e}^1, \boldsymbol{e}^2)$, $(\boldsymbol{e}^3, \boldsymbol{e}^4, \boldsymbol{e}^5)$ ja $(\boldsymbol{e}^6)$, joiden yhteenlaskettu pituus on $2 + 3 + 1 = 6$.

Muotoa

$$\boldsymbol{J}(\lambda, r) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{r \times r}$$

olevaa matriisia kutsutaan *Jordan-lohko*ksi. *Jordan-matriisi* on (lohkodiagonaalinen) yläkolmiomatriisi

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}(\lambda_1, r_1) & & \\ & \ddots & \\ & & \boldsymbol{J}(\lambda_k, r_k) \end{bmatrix} ,$$

jonka lävistäjä koostuu $r_j \times r_j$ kokoisista Jordan-lohkoista.

Esimerkin 4.9 matriisi $\boldsymbol{M}$ on Jordan-matriisi $\begin{bmatrix} \boldsymbol{J}(3,2) & & \\ & \boldsymbol{J}(3,3) & \\ & & (3,1) \end{bmatrix}$.

Edellä olevia tarkasteluja jatkaen voidaan todistaa[17]:

Lause 4.11. *Jokainen neliömatriisi* $\boldsymbol{A}$ *on similaarinen jonkun Jordan-matriisin kanssa. Tätä kutsutaan* $\boldsymbol{A}$ *:n Jordanin muodoksi. Se on lohkojen järjestystä vaille yksikäsitteinen.*

Esimerkki 4.10. Esimerkissä 4.5 laskettiin matriisin

$$\boldsymbol{A} = \begin{bmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{bmatrix}$$

ominaisarvot $\lambda_1 = -2$ ja $\lambda_2 = 4$ algebrallisine kertalukuineen $m_a(-2) = 1$ ja $m_a(4) = 2$. Ominaisvektoreita näille löytyi $\boldsymbol{v}^1 = (1, -1, -1)$ ja $\boldsymbol{v}^2 = (1, -1, 1)$. Näin molempien ominaisarvojen geometrinen kertaluku on yksi. Etsitään vektoria $\boldsymbol{w}^1 = \boldsymbol{v}^2$ vastaava ketju $(\boldsymbol{w}^1, \boldsymbol{w}^2)$ eli ratkaistaan yhtälö $(\boldsymbol{A} - 4\boldsymbol{I})\, \boldsymbol{w}^2 = \boldsymbol{w}^1$. Eliminoinnilla

$$[\boldsymbol{A} - 4\boldsymbol{I} \mid \boldsymbol{w}^1] = \left[\begin{array}{ccc|c} 1 & 4 & 3 & 1 \\ -1 & -4 & -3 & -1 \\ 1 & -2 & -3 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -6 & -6 & 0 \end{array}\right]$$

(tai suoraan näkemällä) saadaan (eräs ratkaisu) $\boldsymbol{w}^2 = (1, 0, 0)$. Näin saadaan:

$$\boldsymbol{A}\, [\boldsymbol{v}^1\ \boldsymbol{w}^1\ \boldsymbol{w}^2] = [\boldsymbol{v}^1\ \boldsymbol{w}^1\ \boldsymbol{w}^2] \begin{bmatrix} -2 & & \\ & 4 & 1 \\ & & 4 \end{bmatrix} .$$

Koska $\boldsymbol{X} = [\boldsymbol{v}^1\ \boldsymbol{w}^1\ \boldsymbol{w}^2] = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}$ on invertoituva, saadaan: $\boldsymbol{A} \sim \begin{bmatrix} J(-2,1) & \\ & J(4,2) \end{bmatrix}$.

Esimerkeissä 4.6 ja 4.7 tarkasteltiin muotoa

$$\boldsymbol{x}^{k+1} = \boldsymbol{A}\, \boldsymbol{x}^k$$

olevia diskreettejä systeemejä. Tällaisen ratkaisu saadaan: $\boldsymbol{x}^k = \boldsymbol{A}^k\, \boldsymbol{x}^0$. Diagonalisoituvassa tapauksessa $\boldsymbol{A} = \boldsymbol{V} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \boldsymbol{V}^{-1}$ saadaan potenssit laskettua yksinkertaisesti

[17] Katso esim. Horn and Johnson: *Matrix analysis*, Cambridge University Press, 1985

$\boldsymbol{A}^k = \boldsymbol{V} \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} \boldsymbol{V}^{-1}$. Yleisesti, jos $\boldsymbol{J_A}$ on $\boldsymbol{A}$:n Jordan-muoto, eli $\boldsymbol{A} = \boldsymbol{V}\, \boldsymbol{J_A} \boldsymbol{V}^{-1}$, missä $\boldsymbol{J_A} = \begin{bmatrix} \boldsymbol{J}(\lambda_1,r_1) & & \\ & \ddots & \\ & & \boldsymbol{J}(\lambda_m,r_m) \end{bmatrix}$, niin

$$\boldsymbol{A}^k = \boldsymbol{V} \begin{bmatrix} \boldsymbol{J}(\lambda_1,r_1)^k & & \\ & \ddots & \\ & & \boldsymbol{J}(\lambda_m,r_m)^k \end{bmatrix} \boldsymbol{V}^{-1} ,$$

joten tarvitaan Jordan-lohkojen potensseja. Nämä saadaan seuraavasti.

Lemma 4.12. *Kun* $\boldsymbol{J} = \boldsymbol{J}(\lambda, r) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{r\times r}$, *niin*

$$\boldsymbol{J}^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \dots & \binom{k}{r-1}\lambda^{k-r+1} \\ & \lambda^k & k\lambda^{k-1} & \ddots & \vdots \\ & & \lambda^k & \ddots & \binom{k}{2}\lambda^{k-2} \\ & & & \ddots & k\lambda^{k-1} \\ & & & & \lambda^k \end{bmatrix} ,$$

eli $(\boldsymbol{J}^k)_{i,j} = \binom{k}{j-i} \lambda^{k-j+i}$, *missä* $\binom{k}{j} = \begin{cases} \frac{k!}{j!\,(k-j)!} , & \textit{kun } j \le k , \\ 0 , & \textit{kun } j > k . \end{cases}$

Tod. Induktiolla: tapaus $k = 1$ on selvä.

Oletetaan, että $(\boldsymbol{J}^k)_{i,j} = \binom{k}{j-i} \lambda^{k-j+i}$, jolloin

$$\begin{aligned} (\boldsymbol{J}^{k+1})_{i,j} &= (\boldsymbol{J}\, \boldsymbol{J}^k)_{i,j} = \textstyle\sum_{l=1}^r \boldsymbol{J}_{i,l}\, (\boldsymbol{J}^k)_{l,j} = \boldsymbol{J}_{i,i}\, (\boldsymbol{J}^k)_{i,j} + \boldsymbol{J}_{i,i+1}\, (\boldsymbol{J}^k)_{i+1,j} \\ &= \lambda \binom{k}{j-i} \lambda^{k-j+i} + \binom{k}{j-i-1} \lambda^{k-j+i+1} = \binom{k+1}{j-i} \lambda^{k+1-j+i} , \end{aligned}$$

missä käytettiin

$$\binom{k}{l} + \binom{k}{l-1} = \frac{k!}{l!\,(k-l)!} + \frac{k!}{(l-1)!\,(k-l+1)!} = \frac{(k-l+1)\,k! + l\,k!}{l!\,(k-l+1)!} = \binom{k+1}{l} .$$

Näin induktioaskelkin on näytetty. □

Näin esimerkin 4.9 matriisille saadaan

$$\boldsymbol{M}^k = \begin{bmatrix} 3^k & k\,3^{k-1} & & & & \\ & 3^k & & & & \\ & & 3^k & k\,3^{k-1} & \frac{k(k-1)}{2}\,3^{k-2} & \\ & & & 3^k & k\,3^{k-1} & \\ & & & & 3^k & \\ & & & & & 3^k \end{bmatrix} .$$

4.4. **Ominaisarvojen ja -vektoreiden numeerinen laskeminen.**

Matriisin ominaisarvot saadaan laskettua tarkasti vain erikoistapauksissa. Usein joudutaan käyttämään approksimaatioita tai numeerisia menetelmiä. Jos matriisin diagonaalin ulkopuoliset elementit ovat pieniä, niin diagonaalielementit ovat tällöin hyviä ominaisarvojen approksimaatioita. Tämä on seuraavan lauseen sisältö.

Lause 4.13 (Gershgorinin ympyrät). *Olkoon* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *ja kullakin* $i = 1, \dots, n$ *olkoon* B_i *kompleksitason* $a_{i,i}$ *-keskinen ja* r_i *-säteinen kiekko, missä* $r_i = \sum_{j\neq i} |a_{i,j}|$. *Tällöin* $\Lambda(\boldsymbol{A})$ *on näiden kiekkojen unionissa eli*

$$\Lambda(\boldsymbol{A}) \subset \bigcup_{i=1}^{n} B_i \ .$$

Tod. Jos $\lambda \in \Lambda(\boldsymbol{A})$, niin $\lambda\boldsymbol{I}-\boldsymbol{A}$ on singulaarinen. Olkoon $\boldsymbol{x} \neq 0$ siten, että $(\lambda\boldsymbol{I}-\boldsymbol{A})\,\boldsymbol{x} = 0$ ja olkoon i siten, että $|x_i| = \max_j |x_j|$. Tällöin

$$(\lambda - a_{i,i})x_i = \sum_{j\neq i} a_{i,j}x_j \ , \qquad \text{joten} \qquad |\lambda - a_{i,i}| \leq \sum_{j\neq i} |a_{i,j}| \,\frac{|x_j|}{|x_i|} \leq \sum_{j\neq i} |a_{i,j}| = r_i \ ,$$

eli $\lambda \in B_i$. □

Ominaisarvojen ratkaiseminen karakteristisen polynomin juurina on yleensä mahdollista vain pienille matriiseille. Koska karakteristisen polynomin kertoimet saadaan kertomalla matriisialkioita keskenään ja laskemalla yhteen tuloja, on jo karakteristisen polynomin muodostaminen altis pyöristysvirheille. Siksi matriisin ominaisarvoja ei yleensä lasketa numeerisesti karakteristisen polynomin avulla[18].

Potenssimenetelmä on iteratiivinen menetelmä, jolla saadaan selville yksitellen matriisin ominaisvektoreita ja sitä kautta myös ominaisarvoja. Oletetaan, että matriisille $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ löytyy n kappaletta lineaarisesti riippumattomia ominaisvektoreita $\boldsymbol{x}^1, \dots, \boldsymbol{x}^n$ (eli A on diagonalisoituva). Silloin mielivaltainen vektori $\boldsymbol{u}^0 \in \mathbb{C}^n$ on lausuttavissa näiden lineaarikombinaationa

$$\boldsymbol{u}^0 = c_1\,\boldsymbol{x}^1 + \cdots + c_n\,\boldsymbol{x}^n \ .$$

Muodostetaan vektorijono $\boldsymbol{u}^1 = \boldsymbol{A}\boldsymbol{u}^0/\|\boldsymbol{A}\boldsymbol{u}^0\|$, $\boldsymbol{u}^2 = \boldsymbol{A}\boldsymbol{u}^1/\|\boldsymbol{A}\boldsymbol{u}^1\|$, eli iteraatio

$$\boldsymbol{u}^{k+1} = \boldsymbol{A}\boldsymbol{u}^k/\|\boldsymbol{A}\boldsymbol{u}^k\| \ , \qquad k = 1, 2, \dots$$

Tällöin induktiolla saadaan helposti $\boldsymbol{u}^k = \boldsymbol{A}^k\boldsymbol{u}^0/\|\boldsymbol{A}^k\boldsymbol{u}^0\|$. Oletetaan, että λ_1 on itseisarvoltaan suurin ominaisarvo ja että

$$|\lambda_1| > |\lambda_i| \ , \quad i = 2, \dots, n \, .$$

Tällöin, koska $\boldsymbol{A}^k\boldsymbol{x}^j = \lambda_j^k\,\boldsymbol{x}^j$, saadaan:

$$\begin{aligned}
\boldsymbol{u}^k &= \frac{c_1\,\boldsymbol{A}^k\boldsymbol{x}^1 + c_2\,\boldsymbol{A}^k\boldsymbol{x}^2 + \cdots + c_n\,\boldsymbol{A}^k\boldsymbol{x}^n}{\|c_1\,\boldsymbol{A}^k\boldsymbol{x}^1 + c_2\,\boldsymbol{A}^k\boldsymbol{x}^2 + \cdots + c_n\,\boldsymbol{A}^k\boldsymbol{x}^n\|} \\
&= \frac{c_1\,\lambda_1^k\,\boldsymbol{x}^1 + c_2\,\lambda_2^k\,\boldsymbol{x}^2 + \cdots + c_n\,\lambda_n^k\,\boldsymbol{x}^n}{\|c_1\,\lambda_1^k\,\boldsymbol{x}^1 + c_2\,\lambda_2^k\,\boldsymbol{x}^2 + \cdots + c_n\,\lambda_n^k\,\boldsymbol{x}^n\|} \\
&= \left(\tfrac{\lambda_1}{|\lambda_1|}\right)^k \frac{c_1\,\boldsymbol{x}^1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{x}^2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^n}{\|c_1\,\boldsymbol{x}^1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{x}^2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}^n\|} \ .
\end{aligned}$$

[18] Itse asiassa karakteristisen polynomin kertoimet lasketaan numeerisesti yleensä ominaisarvoista!

Koska $\lim_{k\to\infty}\left(\frac{\lambda_j}{\lambda_1}\right)^k = 0$ kaikilla $j = 2,\dots,n$, nähdään, että $\boldsymbol{u}^k$:n suunta lähestyy yhä $\boldsymbol{x}^1$:n suuntaa edellyttäen, että $c_1 \neq 0$. Tarkemmin sanottuna[19]

$$\lim_{k\to\infty}(\boldsymbol{u}^k - c\mu^k\,\boldsymbol{x}^1) = 0 \ ,$$

missä $\mu = \frac{\lambda_1}{|\lambda_1|}$ ja $c = \frac{c_1}{\|c_1\boldsymbol{x}^1\|}$. Erityisesti (näytä tämä)

$$\lim_{k\to\infty}\left\langle \boldsymbol{A}\boldsymbol{u}^k, \boldsymbol{u}^k\right\rangle = \lambda_1 \ .$$

Siten potenssimenetelmän tuottamat vektorit $\boldsymbol{u}^k$ kääntyvät ominaisvektorin $\boldsymbol{x}^1$ suuntaiseksi. Periaatteessa menetelmä ei toimi, jos alkuarvausvektorille $\boldsymbol{u}^0$ kerroin c_1 sattui olemaan tarkalleen nolla. Käytännön laskennassa näin ei kuitenkaan pyöristysvirheiden takia pääse käymään.

Iteraation suppenemisnopeus riippuu suhteen $r = \max_{i>1}|\lambda_i/\lambda_1|$ suuruudesta. Jos matriisin itseisarvoltaan suurin ominaisarvo on selvästi muita suurempi, suppeneminen on nopeaa.

Jos $\boldsymbol{A}$ on säännöllinen, sen käänteismatriisin $\boldsymbol{A}^{-1}$ ominaisarvot ovat $1/\lambda_i$, joten soveltamalla potenssimenetelmää käänteismatriisiin, saadaan selville itseisarvoltaan pienin ominaisarvo. Tässä iteraatio kannattaa laskea muodossa

$$\boldsymbol{A}\boldsymbol{w}^k = \boldsymbol{u}^{k-1} \ , \qquad \boldsymbol{u}^k = \boldsymbol{w}^k/\|\boldsymbol{w}^k\| \ ,$$

jolloin riittää ratkaista yhtälöryhmä, eikä itse käänteismatriisia tarvitse muodostaa.

Jos halutaan laskea $\boldsymbol{A}$:n ominaisvektori $\boldsymbol{x}^p$ ja ominaisarvo λ_p, jolle tunnetaan likiarvo $\widehat{\lambda}$, niin potenssimenetelmää voidaan käyttää hyväksi edellyttäen, että likiarvo toteuttaa

$$|\lambda_p - \widehat{\lambda}| < |\lambda_i - \widehat{\lambda}| \ , \qquad i = 1,\dots,n, i \neq p \, .$$

Koska matriisin $\boldsymbol{A} - \widehat{\lambda}\boldsymbol{I}$ ominaisarvot ovat $\lambda_i - \widehat{\lambda}$ ja ominaisvektorit samat kuin $\boldsymbol{A}$:lla, niin matriisin $(\boldsymbol{A} - \widehat{\lambda}\boldsymbol{I})^{-1}$ itseisarvoltaan suurin ominaisarvo on $\frac{1}{\lambda_p - \widehat{\lambda}}$ ja vastaava ominaisvektori $\boldsymbol{x}^p$. Siten potenssiinkorotusmenetelmällä saadaan laskettua $\boldsymbol{A}$:n mielivaltainen ominaisarvo.

Käytännössä approksimaatiota $\widehat{\lambda}$ korjataan joka kierroksella. Seuraava nk. *käänteinen iteraatio* on melko tehokas:

- ratkaistaan yhtälö $(\boldsymbol{A} - \widehat{\lambda}_k\boldsymbol{I})\,\boldsymbol{w}^{k+1} = \boldsymbol{u}^k$
- lasketaan: $\boldsymbol{u}^{k+1} = \boldsymbol{w}^{k+1}/\|\boldsymbol{w}^{k+1}\|$, $\qquad \widehat{\lambda}_{k+1} = \left\langle \boldsymbol{A}\boldsymbol{u}^{k+1}, \boldsymbol{u}^{k+1}\right\rangle$.

4.5. **Positiividefiniitit matriisit.**

Olkoon $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ hermiittinen (tai reaalinen ja symmetrinen). Sitä vastaava *neliömuoto* $Q_{\boldsymbol{A}}$ on kuvaus $\mathbb{C}^n \to \mathbb{R}$, $Q_{\boldsymbol{A}}(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x}\rangle$. Tämän arvot tosiaan ovat reaaliset, sillä

$$\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x}\rangle = \langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{x}\rangle = \langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}\rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{A}\boldsymbol{x}\rangle} \ .$$

[19] μ^k ei suppene, ellei $\lambda_1 > 0$.

Neliömuoto $Q_{\boldsymbol{A}}$ – ja vastaava matriisi $\boldsymbol{A}$ – on

- *Positiivisemidefiniitti (negatiivisemidefiniitti)*, jos $\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle \geq 0 \ \ (\leq 0)$ kaikille $\boldsymbol{x}$.
- *Positiividefiniitti (negatiividefiniitti)*, jos $\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle > 0 \ \ (< 0)$ kun $\boldsymbol{x} \neq 0$.
- *Indefiniitti*, jos se ei ole semidefiniitti, eli jos löytyy vektorit $\boldsymbol{x}$ ja $\boldsymbol{y}$ siten, että $\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle > 0$ ja $\langle \boldsymbol{y}, \boldsymbol{Ay}\rangle < 0$.

Esimerkki 4.11. Tridiagonaalinen $n \times n$-matriisi

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

esiintyy esim. approksimoitaessa toista derivaattaa $\frac{d^2 f}{dx^2}$ tasavälisessä hilassa. $\boldsymbol{A}$ on symmetrinen ja positiividefiniitti, mikä nähdään seuraavasti

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle &= 2\,|x_1|^2 - x_1\overline{x_2} - x_2\overline{x_1} + 2\,|x_2|^2 - x_2\overline{x_3} - x_3\overline{x_2} + 2\,|x_3|^2 - \cdots + 2\,|x_n|^2 \\
&= |x_1|^2 + (x_1 - x_2)\overline{(x_1 - x_2)} + (x_2 - x_3)\overline{(x_2 - x_3)} + \cdots + (x_{n-1} - x_n)\overline{(x_{n-1} - x_n)} + |x_n|^2 \\
&= |x_1|^2 + |x_1 - x_2|^2 + |x_2 - x_3|^2 + \cdots + |x_{n-1} - x_n|^2 + |x_n|^2 \geq 0 \ ,
\end{aligned}$$

ja nolla saadaan vain tapauksessa $x_1 = x_2 = \cdots = x_n = 0$.

Neliömuodon ja sitä vastaavan matriisin ominaisarvojen välistä yhteyttä kuvaa seuraava lause. Muistamme, että hermiittisen matriisin ominaisarvot ovat reaaliset.

Lause 4.14. *Olkoon* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *hermiittinen ja sen ominaisarvot numeroitu kasvavaan järjestykseen* $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. *Tällöin mielivaltaiselle* $x \in \mathbb{C}^n \setminus \{0\}$ *pätee*

$$\lambda_1 \leq \frac{\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle}{\langle \boldsymbol{x}, \boldsymbol{x}\rangle} \leq \lambda_n \ .$$

Alaraja (yläraja) λ_1 *(* λ_n *) saavutetaan vastaavalla ominaisvektorilla.*

Tod. Lauseen 4.8 perusteella hermiittinen matriisi voidaan kirjoittaa $\boldsymbol{A} = \boldsymbol{U\Lambda U}^*$, missä $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ ja $\boldsymbol{U}$ on unitaarinen. Olkoon $\boldsymbol{x} \neq 0$ ja $\boldsymbol{v} = \boldsymbol{U}^*\boldsymbol{x}$. Tällöin $\langle \boldsymbol{v}, \boldsymbol{v}\rangle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle$ ja

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle &= \boldsymbol{x}^T\,\overline{\boldsymbol{U\Lambda U}^*\boldsymbol{x}} = (\boldsymbol{U}^*\boldsymbol{x})^T\boldsymbol{\Lambda}\overline{\boldsymbol{U}^*\boldsymbol{x}} \\
&= \boldsymbol{v}^T\boldsymbol{\Lambda}\overline{\boldsymbol{v}} = \lambda_1\ |v_1|^2 + \lambda_2\ |v_2|^2 + \cdots + \lambda_n\ |v_n|^2 \\
&\geq \lambda_1\,(|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2) = \lambda_1\ \langle \boldsymbol{v}, \boldsymbol{v}\rangle \\
&= \lambda_1\ \langle \boldsymbol{x}, \boldsymbol{x}\rangle \ ,
\end{aligned}$$

joten alaraja on todistettu. Yläraja saadaan vastaavasti.

Se, että rajat saavutetaan vastaavilla ominaisvektoreilla on selvä. □

Lauseketta $\frac{\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle}{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$ kutsutaan Rayleigh'n osamääräksi. Sen minimointi ja maksimointi antavat hermiittisen matriisin pienimmän ja suurimman ominaisarvon.

Positiividefiniiteille matriiseille saadaan siten seuraava karakterisointi

Seuraus 4.15. *Hermiittinen matriisi on positiividefiniitti jos ja vain jos sen kaikki ominaisarvot ovat positiiviset.*

Olkoon $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ hermiittinen ja positiividefiniitti ja olkoon $\boldsymbol{A}_{1:r,1:r} = \left[\begin{smallmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{smallmatrix}\right]$. Toisin sanoen $\boldsymbol{A}$:sta otetaan $r \times r$ vasen ylänurkka, $1 \le r \le n$. Tällaista kutsutaan $\boldsymbol{A}$:n *pääneliömatriisiksi.*

Jos $\boldsymbol{v} \in \mathbb{C}^r \setminus \{0\}$, niin lisätään $\boldsymbol{v}$:hen nollia, niin että saadaan $\tilde{\boldsymbol{v}} = \left[\begin{smallmatrix} \boldsymbol{v} \\ 0 \end{smallmatrix}\right] \in \mathbb{C}^n \setminus \{0\}$, joten

$$0 < \langle \tilde{\boldsymbol{v}}, \boldsymbol{A}\tilde{\boldsymbol{v}} \rangle = \tilde{\boldsymbol{v}}^T\, \overline{\boldsymbol{A}\tilde{\boldsymbol{v}}} = \boldsymbol{v}^T\, \overline{\boldsymbol{A}_{1:r,1:r}\,\boldsymbol{v}} = \langle \boldsymbol{v}\,,\, \boldsymbol{A}_{1:r,1:r}\,\boldsymbol{v} \rangle \ .$$

Näin $\boldsymbol{A}_{1:r,1:r}$ on myös positiividefiniitti. Täten $\boldsymbol{A}$:n kaikki pääneliömatriisit ovat positiividefiniittejä. Erityisesti niiden determinantit ovat positiiviset. Yllättäen tämä myös riittää karakterisoimaan positiividefiniitin matriisin:

Lause 4.16. *Hermiittinen matriisi on positiividefiniitti jos ja vain jos sen jokaisen pääneliömatriisin determinantti on positiivinen.*

Tod. Todistus sivuutetaan. Katso esim. Horn and Johnson: *Matrix analysis* □

Esimerkki 4.12. Matriisi

$$\boldsymbol{A} = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{bmatrix}$$

on positiividefiniitti, sillä

$$2 > 0\ , \qquad \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3 > 0 \qquad \text{ja} \qquad \begin{vmatrix} 2 & -1 & -3 \\ -1 & 2 & 4 \\ -3 & 4 & 9 \end{vmatrix} = 1 > 0\ .$$

Matriisin $\|\cdot\|_2$-normin laskeminen palautuu ominaisarvotehtäväksi:

Lause 4.17. *Olkoon* $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. *Tällöin* $\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^*\boldsymbol{A})}$.

Tod. Matriisi $\boldsymbol{A}^*\boldsymbol{A}$ on hermiittinen ja positiivisesti semidefiniitti (miksi), joten lauseen 4.14 perusteella

$$\|\boldsymbol{A}\|_2^2 = \max_{\|\boldsymbol{x}\|_2=1} \langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{A}\boldsymbol{x} \rangle = \max_{\|\boldsymbol{x}\|_2=1} \langle \boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x} \rangle = \max_{\mu\in\Lambda(\boldsymbol{A}^*\boldsymbol{A})} \mu = \rho(\boldsymbol{A}^*\boldsymbol{A})\ .$$

□

Olkoon $\|\cdot\|$ jokin $\mathbb{C}^n$:n normi ja sitä vastaava matriisinormi. Olkoon $\boldsymbol{A}$ neliömatriisi, λ sen ominaisarvo ja $\boldsymbol{v}$ siihen liittyvä normeerattu ominaisvektori $\|\boldsymbol{v}\| = 1$. Tällöin

$$\|\boldsymbol{A}\boldsymbol{v}\| = \|\lambda\boldsymbol{v}\| = |\lambda|\ \|\boldsymbol{v}\| = |\lambda|\ ,$$

joten $\|\boldsymbol{A}\| \ge |\lambda|$. Tämä pätee kaikille ominaisarvoille, joten

$$\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|\ . \tag{4.10}$$

Hermiittiselle matriisille ja matriisin $\|\cdot\|_2$-normille tässä pätee yhtäsuuruus.

Lause 4.18. *Olkoon* $\boldsymbol{A}$ *hermiittinen. Tällöin* $\rho(\boldsymbol{A}) = \|\boldsymbol{A}\|_2$.

Tod. Hermiittiselle $\boldsymbol{A}$ saadaan

$$\rho(\boldsymbol{A}^*\boldsymbol{A}) = \rho(\boldsymbol{A}^2) = \max_{\mu\in\Lambda(\boldsymbol{A}^2)} |\mu| = \max_{\lambda\in\Lambda(\boldsymbol{A})} \left|\lambda^2\right| = \big(\max_{\lambda\in\Lambda(\boldsymbol{A})} |\lambda|\big)^2 = \rho(\boldsymbol{A})^2 ,$$

joten väite saadaan lauseesta 4.17. □

Lauseen 4.17 mukaan matriisin $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ kakkosnormi on matriisin $\boldsymbol{A}^*\boldsymbol{A} \in \mathbb{C}^{n\times n}$ suurimman ominaisarvon neliöjuuri. Olkoot $\boldsymbol{A}^*\boldsymbol{A}$:n ominaisarvot numeroitu suurimmasta pienimpään $\mu_1 \geq \cdots \geq \mu_n \geq 0$. Näiden neliöjuuria $\sigma_1 = \sqrt{\mu_1} \geq \cdots \geq \sigma_n = \sqrt{\mu_n}$ kutsutaan matriisin $\boldsymbol{A}$ *singulaariarvoiksi.* Toisin kuin ominaisarvot, nämä on määritelty kaikille matriiseille. Lisäksi pätee

Lause 4.19 (Singulaariarvohajotelma)**.** *Jokainen matriisi $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ voidaan esittää tulona $\boldsymbol{A} = \boldsymbol{U\Sigma V}^*$, missä $\boldsymbol{U} \in \mathbb{C}^{m\times m}$ ja $\boldsymbol{V} \in \mathbb{C}^{n\times n}$ ovat unitaarisia ja $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & & \end{bmatrix} \in \mathbb{C}^{m\times n}$. Tässä $\sigma_1, \ldots, \sigma_r$ ovat $\boldsymbol{A}$:n nollasta poikkeavat singulaariarvot.*

Tod. Emme todista tätä kokonaan, mutta lausekkeista

$$\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{V\Sigma}^T\boldsymbol{\Sigma V}^* \qquad \text{ja} \qquad \boldsymbol{AA}^* = \boldsymbol{U\Sigma\Sigma}^T\boldsymbol{U}^*$$

nähdään, että $\boldsymbol{V}$:n ja $\boldsymbol{U}$:n sarakkeet koostuvat matriisien $\boldsymbol{A}^*\boldsymbol{A}$ ja $\boldsymbol{AA}^*$ ortonormaaleista ominaisvektoreista. □

Lauseen 4.17 perusteella

$$\begin{aligned}\|\boldsymbol{A}^{-1}\|_2 &= \sqrt{\rho\big((\boldsymbol{A}^{-1})^*\boldsymbol{A}^{-1}\big)} = \sqrt{\rho\big((\boldsymbol{AA}^*)^{-1}\big)} \\ &= \max_{\mu\in\Lambda\left((\boldsymbol{AA}^*)^{-1}\right)} \sqrt{\mu} = \frac{1}{\min_{\mu\in\Lambda(\boldsymbol{AA}^*)}\sqrt{\mu}} = \frac{1}{\sigma_n} ,\end{aligned}$$

missä σ_n on $\boldsymbol{A}$:n pienin singulaariarvo[20]. Täten kakkosnormissa mitattuna matriisin häiriöalttius on

$$\kappa_2(\boldsymbol{A}) = \|\boldsymbol{A}\|_2\|\boldsymbol{A}^{-1}\|_2 = \frac{\sigma_1}{\sigma_n} .$$

Tehtävä 4.12. Näytä, että hermiittisen matriisin $\boldsymbol{A}$ häiriöalttius saadaan

$$\kappa_2(\boldsymbol{A}) = \frac{\max\left\{|\lambda| \,\middle|\, \lambda \in \Lambda(\boldsymbol{A})\right\}}{\min\left\{|\lambda| \,\middle|\, \lambda \in \Lambda(\boldsymbol{A})\right\}} = \frac{\rho(\boldsymbol{A})}{\rho(\boldsymbol{A}^{-1})} .$$

Esimerkki 4.13. Aiemmin laskettiin matriisille $\boldsymbol{A} = \begin{bmatrix} 1 & \varepsilon \\ 1 & -\varepsilon \end{bmatrix}$ häiriöalttiudeksi 1-normissa mitattuna $\kappa_1(\boldsymbol{A}) = 1 + 1/\varepsilon$. Lasketaan nyt $\kappa_2(\boldsymbol{A})$. Saadaan

$$\boldsymbol{A}^*\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ \varepsilon & -\varepsilon \end{bmatrix} \begin{bmatrix} 1 & \varepsilon \\ 1 & -\varepsilon \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2\varepsilon^2 \end{bmatrix} .$$

Siispä $\boldsymbol{A}$:n singulaariarvot ovat: $\sigma_1 = \sqrt{2}$ ja $\sigma_2 = \varepsilon\sqrt{2}$. Näin häiriöalttius saadaan: $\kappa_2(\boldsymbol{A}) = 1/\varepsilon$, joka (pienillä ε) on oleellisesti saman kokoinen kuin $\kappa_1(\boldsymbol{A})$.

[20] $\boldsymbol{A}^*\boldsymbol{A}$ ja $\boldsymbol{AA}^*$ ovat similaariset, kun $\boldsymbol{A}$ on neliömatriisi (harjoitustehtävä).

4.6. **Matriisin sarjat ja eksponenttifunktio.**

Mielivaltaisessa normiavaruudessa V jono $\{\boldsymbol{v}^k\}_{k=1}^{\infty}$ *suppenee* kohti alkiota $\boldsymbol{v} \in V$, jos pätee: $\lim_{k\to\infty} \|\boldsymbol{v}^k - \boldsymbol{v}\| = 0$.

Tässä tarkastelemme eräitä matriisijonoja ja -sarjoja sekä näiden suppenemista.

Lause 3.6 sanoo, että jos $\|\boldsymbol{A}\| < 1$, niin $\boldsymbol{I} - \boldsymbol{A}$ on invertoituva. Seuraavassa esitämme tälle toisen todistuksen ja laskemme tälle inverssille sarjan.

Lause 4.20. *Olkoon* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *siten, että* $\|\boldsymbol{A}\| < 1$. *Tällöin matriisi* $\boldsymbol{I} - \boldsymbol{A}$ *on invertoituva ja* $(\boldsymbol{I} - \boldsymbol{A})^{-1}$ *voidaan esittää suppenevana sarjana*

$$(\boldsymbol{I} - \boldsymbol{A})^{-1} = \sum_{k=0}^{\infty} \boldsymbol{A}^k = \boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2 + \dots$$

Tätä sarjaa kutsutaan $\boldsymbol{A}$:n *Neumannin sarjaksi.*

Tod. Olkoon $a_{ij}(k) = (\boldsymbol{A}^k)_{ij}$. Tällöin $|a_{ij}(k)| \leq \|\boldsymbol{A}^k\| \leq \|\boldsymbol{A}\|^k$ (ks. (3.12)). Koska $\sum_{k=0}^{\infty} \|\boldsymbol{A}\|^k$ suppenee (summa $= 1/(1-\|\boldsymbol{A}\|)$), niin kaikki sarjat $s_{ij} = \sum_{k=0}^{\infty} a_{ij}(k)$, $i, j = 1, \dots, n$, suppenevat. Olkoon $\boldsymbol{S} = \{s_{ij}\}$ ja $\boldsymbol{S}^n = \sum_{k=0}^{n} \boldsymbol{A}^k$. Tällöin

$$\|\boldsymbol{S} - \boldsymbol{S}^n\| = \|\sum_{k=n+1}^{\infty} \boldsymbol{A}^k\| \leq \sum_{k=n+1}^{\infty} \|\boldsymbol{A}\|^k = \frac{\|\boldsymbol{A}\|^{n+1}}{1 - \|\boldsymbol{A}\|} \underset{n\to\infty}{\to} 0 .$$

Edelleen

$$\begin{aligned} \|\boldsymbol{S}(\boldsymbol{I} - \boldsymbol{A}) - \boldsymbol{I}\| &= \lim_{n\to\infty} \|\sum_{k=0}^{n} \boldsymbol{A}^k(\boldsymbol{I} - \boldsymbol{A}) - \boldsymbol{I}\| = \lim_{n\to\infty} \|\sum_{k=0}^{n} (\boldsymbol{A}^k - \boldsymbol{A}^{k+1}) - \boldsymbol{I}\| \\ &= \lim_{n\to\infty} \|-\boldsymbol{A}^{n+1}\| \leq \lim_{n\to\infty} \|\boldsymbol{A}\|^{n+1} = 0 . \end{aligned}$$

Siis $\boldsymbol{S}(\boldsymbol{I} - \boldsymbol{A}) = \boldsymbol{I}$. Samoin $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{S} = \boldsymbol{I}$, joten $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{S}$. □

Tehtävä 4.13. Olkoot $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^n$ siten, että $\|\boldsymbol{A}\| < 1$ ja $a_{i,j} \geq 0$, $b_i > 0$, $i, j = 1, \dots, n$. Näytä, että yhtälön $\boldsymbol{x} - \boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ ratkaisulle pätee: $x_j > 0$, $j = 1, \dots, n$.

Sarjateoriasta tiedetään, että eksponenttifunktion potenssisarja $e^x = \sum_0^{\infty} \frac{x^k}{k!}$ suppenee $\forall x \in \mathbb{C}$. Tämän analogian perusteella voidaan määritellä neliömatriisin eksponenttifunktio kaavalla

$$e^{\boldsymbol{A}} = \sum_{k=0}^{\infty} \tfrac{1}{k!} \boldsymbol{A}^k .$$

Samaan tapaan kuin edellisen lauseen todistuksessa kullekin elementtisarjalle saadaan suppeneva majorantti $e^{\|\boldsymbol{A}\|} = \sum_{k=0}^{\infty} \frac{1}{k!} \|\boldsymbol{A}\|^k$.

Sarja on siis suppeneva, mutta näyttää hankalalta laskea $e^{\boldsymbol{A}}$ suoraan määritelmästä. Tarkastellaan muutamia erikoistapauksia. Lävistäjämatriisin $\boldsymbol{D} = \operatorname{diag}(d_1, d_2, \dots, d_n)$

potenssit ovat $\boldsymbol{D}^k = \mathrm{diag}(d_1^k, d_2^k, \ldots, d_n^k)$, joten

$$e^{\boldsymbol{D}} = \sum_{k=0}^{\infty} \tfrac{1}{k!}\, \boldsymbol{D}^k = \mathrm{diag}\big(\textstyle\sum_{k=0}^{\infty} \frac{d_1^k}{k!}, \ldots, \sum_{k=0}^{\infty} \frac{d_n^k}{k!}\big) = \mathrm{diag}(\, e^{d_1}, \ldots, e^{d_n}) \ .$$

Matriisia $\boldsymbol{N}$ sanotaan *nilpotentiksi*, jos $\boldsymbol{N}^l = 0$ jollekin $l \geq 0$. Selvästi tällöin myös kaikki korkeammat potenssit ovat nollia. Siten eksponenttifunktion $e^{\boldsymbol{N}}$ sarja katkeaa, joten

$$e^{\boldsymbol{N}} = \sum_{k=0}^{l-1} \tfrac{1}{k!}\boldsymbol{N}^k = \boldsymbol{I} + \boldsymbol{N} + \tfrac{1}{2}\boldsymbol{N}^2 + \cdots + \tfrac{1}{(l-1)!}\boldsymbol{N}^{l-1} \ .$$

Esimerkki 4.14. Matriisi

$$\boldsymbol{N} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

on nilpotentti, sillä $\boldsymbol{N}^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ja $\boldsymbol{N}^3 = 0$. Siten

$$e^{\boldsymbol{N}} = \boldsymbol{I} + \boldsymbol{N} + \tfrac{1}{2}\boldsymbol{N}^2 = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \ .$$

Seuraavan lauseen avulla $e^{\boldsymbol{A}}$ voidaan laskea yleiselle neliömatriisille $\boldsymbol{A}$.

Lause 4.21. *Olkoot $\boldsymbol{A}$, $\boldsymbol{B}$ ja $\boldsymbol{P}$ $n \times n$ -matriiseja ja $\boldsymbol{P}$ lisäksi säännöllinen.*

a) *Jos* $\boldsymbol{C} = \boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}$, *niin* $e^{\boldsymbol{C}} = \boldsymbol{P}\, e^{\boldsymbol{A}}\, \boldsymbol{P}^{-1}$.
b) *Jos* $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$, *niin* $e^{\boldsymbol{A}+\boldsymbol{B}} = e^{\boldsymbol{A}}\, e^{\boldsymbol{B}}$.
c) $e^{-\boldsymbol{A}} = (e^{\boldsymbol{A}})^{-1}$.
d) $e^{(\boldsymbol{A}^*)} = (e^{\boldsymbol{A}})^*$.

Todistus. a)

$$e^{\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}} = \sum_{0}^{\infty} \frac{1}{k!}(\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1})^k \ ,$$

jossa esiintyy potensseja

$$(\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1})^k = \boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1} \cdots \boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1} = \boldsymbol{P}\boldsymbol{A}^k\boldsymbol{P}^{-1} \ .$$

Siten

$$\lim_{m\to\infty} \sum_{k=0}^{m} \frac{1}{k!}(\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1})^k = \lim_{m\to\infty} \boldsymbol{P} \sum_{k=0}^{m} \frac{1}{k!}\boldsymbol{A}^k\boldsymbol{P}^{-1} = \boldsymbol{P}e^{\boldsymbol{A}}\boldsymbol{P}^{-1} \ .$$

b)-kohdassa voidaan käyttää binomikaavaa potenssien $(\boldsymbol{A}+\boldsymbol{B})^k$ laskemiseksi, koska $\boldsymbol{A}$ ja $\boldsymbol{B}$ kommutoivat:

$$(\boldsymbol{A}+\boldsymbol{B})^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!}\ \boldsymbol{A}^j\, \boldsymbol{B}^{k-j} \ .$$

Täten

$$\begin{aligned} e^{\boldsymbol{A}+\boldsymbol{B}} &= \sum_{k=0}^{\infty} \frac{1}{k!}(\boldsymbol{A}+\boldsymbol{B})^k = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k} \frac{\boldsymbol{A}^j}{j!}\frac{\boldsymbol{B}^{k-j}}{(k-j)!}\right) \\ &= \left(\sum_{j=0}^{\infty}\frac{\boldsymbol{A}^j}{j!}\right)\left(\sum_{k=0}^{\infty}\frac{\boldsymbol{B}^k}{k!}\right) = e^{\boldsymbol{A}}e^{\boldsymbol{B}}. \end{aligned}$$

Viimeisellä rivillä on käytetty "summien konvoluutiolle"tulosta (Cauchyn tulo), joka voidaan todistaa induktiolla. Se pätee tässä myös matriisisarjoille, koska $\boldsymbol{A}$ ja $\boldsymbol{B}$ kommutoivat.

c)-kohta seuraa b):stä: $\boldsymbol{I} = e^0 = e^{\boldsymbol{A}-\boldsymbol{A}} = e^{\boldsymbol{A}}e^{-\boldsymbol{A}}$. d) jätetään harjoitustehtäväksi. □

Jos matriisi A on diagonalisoituva, eli löytyy $\boldsymbol{P}$ siten, että $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$, jossa $\boldsymbol{\Lambda}$ on lävistäjämatriisi, niin lauseen a)-kohdan mukaan

$$e^{\boldsymbol{A}} = \boldsymbol{P}e^{\boldsymbol{\Lambda}}\boldsymbol{P}^{-1}.$$

Koska lävistäjämatriisin eksponenttifunktio osataan laskea, on $e^{\boldsymbol{A}}$:n laskeminen diagonalisoituville matriiseille ratkaistu.

Esimerkki 4.15. Matriisin $\boldsymbol{A} = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ ominaisarvot ovat $\lambda_1 = 1$ ja $\lambda_2 = -1$. Niitä vastaavat ominaisvektorit ovat esim. $\boldsymbol{v}^1 = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ ja $\boldsymbol{v}^2 = \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. Siten

$$\boldsymbol{\Lambda} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right], \quad \boldsymbol{P} = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$$

ja

$$\begin{aligned} e^{\boldsymbol{A}} &= \boldsymbol{P}e^{\boldsymbol{\Lambda}}\boldsymbol{P}^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^1 & 0 \\ 0 & e^{-1} \end{bmatrix}\begin{bmatrix} \frac12 & \frac12 \\ \frac12 & -\frac12 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} e+e^{-1} & e-e^{-1} \\ e-e^{-1} & e+e^{-1} \end{bmatrix} = \begin{bmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{bmatrix}. \end{aligned}$$

Jos $\boldsymbol{A}$ ei ole diagonalisoituva, niin se voidaan kuitenkin similaarimuuntaa Jordanin muotoon $\boldsymbol{J}_{\boldsymbol{A}}$. Nyt $e^{\boldsymbol{J}_{\boldsymbol{A}}}$ on lohkolävistäjämatriisi, jonka lohkot koostuvat muotoa $e^{\boldsymbol{J}(\lambda,r)}$ olevista matriiseista. Nämä voidaan laskea seuraavasti: Jordan-lohko voidaan kirjoittaa lävistäjämatriisin ja nilpotentin matriisin summana

$$\boldsymbol{J}(\lambda,r) = \lambda\boldsymbol{I} + \boldsymbol{N} = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} = \begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

Koska $\lambda\boldsymbol{I}$ ja $\boldsymbol{N}$ kommutoivat, saadaan

$$e^{\boldsymbol{J}(\lambda,r)} = e^{\lambda\boldsymbol{I}+\boldsymbol{N}} = e^{\lambda}\sum_{j=0}^{r-1}\tfrac{1}{j!}\boldsymbol{N}^j.$$

Esimerkki 4.16. Esimerkissä 4.10 laskettiin matriisin

$$\boldsymbol{A} = \left[\begin{smallmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{smallmatrix}\right]$$

similaarimuunnos Jordan-muotoon

$$\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{J}_{\boldsymbol{A}} = \left[\begin{smallmatrix} -2 & & \\ & 4 & 1 \\ & & 4 \end{smallmatrix}\right], \qquad \text{missä} \qquad \boldsymbol{X} = \left[\begin{smallmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ -1 & 1 & 0 \end{smallmatrix}\right].$$

Näin ollen

$$\begin{aligned} e^{\boldsymbol{A}} &= \boldsymbol{X} e^{\boldsymbol{J}_{\boldsymbol{A}}} \boldsymbol{X}^{-1} = \boldsymbol{X} \begin{bmatrix} e^{-2} & \\ & e^{\boldsymbol{J}(4,2)} \end{bmatrix} \boldsymbol{X}^{-1} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} e^{-2} & \\ & e^{4}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}^{-1} \\ &= \frac{1}{2} \begin{bmatrix} 4e^4 & -e^{-2} + 3e^4 & -e^{-2} + e^4 \\ -2e^4 & e^{-2} - e^4 & e^{-2} - e^4 \\ 2e^4 & e^{-2} + e^4 & e^{-2} + e^4 \end{bmatrix} . \end{aligned}$$

Harjoitustehtäviä

1. Osoita, että
$$\det(e^{\boldsymbol{A}}) = e^{\operatorname{tr}(\boldsymbol{A})} .$$
2. Olkoon $\boldsymbol{V} \in \mathbb{C}^{n \times n}$ vinohermiittinen: $\boldsymbol{V}^* = -\boldsymbol{V}$. Näytä, että $e^{\boldsymbol{V}}$ on unitaarinen. Vastaavasti eksponenttifunktio reaalisesta vinosymmetrisestä matriisista on ortogonaalinen.