Matriisinormeista

Sanni Carlson

Matematiikan pro gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Kevät 2015

Tiivistelmä: Sanni Carlson, *Matriisinormeista* (engl. *On matrix norms*), matematiikan pro gradu -tutkielma, 55. s., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, kevät 2015.

Tässä tutkielmassa käsitellään vektori- ja matriisinormeja, niiden ominaisuuksia ja niihin liittyviä tuloksia. Matriisinormien tarkastelemiseksi on ensin mielekästä tietää, mikä on vektorinormi ja millaisia ominaisuuksia siltä vaaditaan. Vektorinormilla voidaan esimerkiksi laskea vektorin pituus. Matriisinormi taas mittaa esimerkiksi sitä, kuinka paljon maksimissaan vektori venyy matriisilla kerrottaessa.

Vektorinormeille asetetaan kolme vaatimusta, joiden kaikkien tulee olla voimassa: positiivisuus, homogeenisuus ja kolmioepäyhtälö. Koska matriisit koostuvat vektoreista, siirtyvät vektorinormien vaatimukset suoraan matriisinormeille. Vektorinormien vaatimusten lisäksi matriisinormeille määritellään vielä yksi ehto lisää; matriisitulon submultiplikatiivisuuden tulee olla voimassa. Neljännen ehdon lisääminen matriisinormeille aiheuttaa sen, että kaikki vektorinormit eivät ole matriisinormeja.

On olemassa useita erilaisia vektori- ja matriisinormeja, joita hyödynnetään tapauskohtaisesti. Yleisin tunnettu vektorinormi on euklidinen normi. Matriisinormeista tutkielman kannalta oleellinen on spektraalinormi, jossa tarvitaan matriisin ominaisarvojen hallintaa. Joillekin matriisinormeille voidaan määrätä yhteensopiva vektorinormi, jolloin sanotaan, että matriisinormi sopeutuu vektorinormiin. Lisäksi jokainen vektorinormi indusoi matriisinormin. Indusoitu matriisinormi aina myös sopeutuu vektorinormiin, josta se on indusoitu. Esimerkiksi euklidinen normi indusoi spektraalinormin ja samalla siis spektraalinormi sopeutuu euklidiseen normiin. Kaikki matriisinormit eivät ole indusoituja matriisinormeja.

Spektraalisäde määritellään matriisin itseisarvoltaan suurimmaksi ominaisarvoksi. Spektraalisäde ei ole vektori- eikä matriisinormi. Kuitenkin voidaan todistaa, että jokin matriisinormi saadaan äärimmäisen lähelle spektraalisädettä. Eräs spektraalisäteeseen liittyvä tärkeä tulos antaa yhteyden myös matriisijonon suppenemiselle: matriisitulo suppenee, jos ja vain jos spektraalisäde on pienempi kuin yksi. Matriisijonon suppeneminen määritellään alkioittain eli matriisijono suppenee kohti jotakin tiettyä matriisia, jos kaikki matriisin alkiot suppenevat kohti jotakin tiettyä alkiota.

Matriisinormeja sovelletaan muun muassa käänteismatriisien virheiden arvioinnissa. Matriisille voidaan määrittää ehtoluku, joka mittaa esimerkiksi matriisin alkioissa mahdollisesti sattuneiden pyöristysvirheiden suuruutta. Matriisin ehtoluku riippuu valitusta matriisinormista. Mitä suurempi ehtoluku on, sitä suurempi on matriisin virhealttius.

Tämän tutkielman lopussa esitetään opitun teorian pohjalta kurssisuunnitelma ”Matriisit tutuiksi”, jonka on tarkoitus olla lukion pitkän matematiikan syventävä kurssi.

Sisältö

Johdanto

Tämän tutkielman tarkoituksena on käsitellä vektori- ja matriisinormeja. Näiden normien lisäksi tutkielmassa esitetään myös matriiseihin yleisesti liittyviä ominaisuuksia, kuten matriisien ominaisarvojen ja virheiden laskemista. Tutkielman pääpaino on kuitenkin matriisinormeissa ja niihin liittyvien tulosten tarkastelussa.

Matriisien käytöstä on merkkejä jo muinaishistoriassa, jolloin niitä hyödynnettiin mm. lineaaristen yhtälöryhmien ratkaisemisessa. Tällöin matriisit tunnettiin taulukoina. Ensimmäinen esimerkki taulukoiden käytöstä yhtälön ratkaisemisessa on kiinalaisessa kirjassa "The Nine Chapters on the Mathematical Art". Tämä matemaattinen kirja on kokoelma teksteistä useiden sukupolvien ajoilta, noin 900 vuotta ennen ajanlaskua aina ajanlaskun alkupuolelle saakka. [**16**] Myös matriisin determinantti on peräisin näistä teksteistä. James Joseph Sylvester loi nykyisen matriisi -käsitteen vuonna 1850.

Matriisien kehityksen kannalta 1800-luvun loppu oli tärkeä. Tällöin mm. Wilhelm Jordan julkaisi nykypäivänä paljon käytetyn Gaussin & Jordanin -menetelmän. Siitä lähtien matriisit ovat olleet oleellisia monilla matematiikan osa-alueilla. Erityisesti matriisit ovat merkityksellisiä lineaarialgebrassa ja numeerisessa matematiikassa. Nykyään niitä sovelletaan myös paljon tietotekniikassa ja tilastotieteessä.

Tämä tutkielma koostuu yhdeksästä luvusta, joista viimeiseen on koottu käytettyjen merkintöjen selitykset. Tutkielman pääpaino on luvuissa 4, 5 ja 6, joissa käsitellään vektori- ja matriisinormeja ja niihin liittyviä tuloksia. Näitä tuloksia varten tullaan aiemmissa luvuissa esittämään tarvittavia "työkaluja". Erityisesti luvussa kolme käsiteltävien ominaisarvojen ymmärtäminen on hyödyksi. Tutkielmassa tarkastellaan pääsääntöisesti reaali- tai kompleksikertoimisia neliömatriiseja.

Ensimmäisissä kolmessa luvussa tutustutaan matriiseiden perusoperaatioihin, kääntyvyyteen ja ominaisarvoihin. Perusoperaatioista mainittakoon hermitointi, joka vastaa kompleksikertoimisen matriisin transponointia. Reaalisen avaruuden transponoinnista erona on alkioiden konjugointi. Toisen ja kolmannen luvun kannalta keskeisessä roolissa on determinantti. Determinantin avulla saadaan selville tärkeää tietoa matriisin kääntyvyydestä ja ominaisarvoista. Näitä tietoja tarvitaan myöhemmin spektraalinormin ja spektraalisäteen yhteydessä luvussa 6.

Neljäs ja viides luku keskittyvät tämän tutkielman kannalta oleellisiin aiheisiin eli normeihin ja niiden ominaisuuksiin. Nämä luvut sisältävät esimerkkejä ja todistuksia yleisimpiin normeihin liittyen. Yleiselle vektorinormille määritellään seuraavat kolme ehtoa: positiivisuus, homogeenisuus ja kolmioepäyhtälö. Yleisin tunnettu vektorinormi on euklidinen normi, joka täyttää kaikki nämä kolme ehtoa. On olemassa myös muita vektorinormeja, joita esitellään neljännen luvun lopussa. Viidennessä luvussa laajennetaan yleinen vektorinormi yleiseksi matriisinormiksi; vektorinormin ehtojen lisäksi submultiplikatiivuus -ehdon tulee olla voimassa. Osoittautuu, että kaikki vektorinormit eivät ole matriisinormeja, koska ne eivät täytä tätä lisättyä ehtoa. Tässä luvussa paneudutaankin matriisi- ja vektorinormien välisiin yhteyksiin ja osoitetaan esimerkiksi, kuinka jokaisesta vektorinormista voidaan indusoida matriisinormi.

Kuudennessa luvussa käsitellään spektraalinormia ja spektraalisädettä. Spektraalinormi on matriisinormi, jonka euklidinen vektorinormi indusoi. Spektraalisäde puolestaan määritellään itseisarvoltaan matriisin suurimmaksi ominaisarvoksi. Sopivan matriisin avulla voidaan osoittaa, että spektraalisäde ei ole vektori- eikä matriisinormi. Yleisesti jokin matriisinormi kuitenkin saadaan aina hyvin lähelle spektraalisädettä. Spektraalisäde antaa ehtoja myös matriisijonon suppenemiselle. Matriisijonojen suppenemisiin ja muihin aiemmin opittuihin asioihin kuten matriisin kääntyvyyteen liittyen osoitetaan tärkeitä tuloksia.

Matriisinormeille on olemassa erilaisia sovelluksia. Esimerkiksi tietotekniikassa käsitellään matriisien virheiden analysointiin liittyviä algoritmeja. Luvussa seitsemän tarkastellaankin, millaisia virheitä käänteismatriisin laskemisessa voi tulla ja kuinka näitä virheitä voidaan arvioida ehtoluvun avulla. Lopussa käydään läpi esimerkki sarakesummanormilla lasketusta virheestä. Matriisin ehtoluku ja samalla matriisin virheen suuruus riippuvat valitusta matriisinormista.

Tutkielman kahdeksannessa luvussa on aiemmin opitun teorian pohjalta suunniteltu kurssisuunnitelma lukion pitkän matematiikan syventäväksi kurssiksi. Kurssin ideana on tutustua matriiseihin ja niiden ominaisuuksiin.

Tämän tutkielman lukijalta edellytetään lineaarialgebrassa ja matriisilaskennassa opittujen asioiden hallintaa. Tutkielmassa pyritään käyttämään paljon esimerkkejä, joiden avulla lukijan on helpompi ymmärtää lukemaansa.

LUKU 1

Matriisit ja vektorit

1.1. Matriisi

Määritellään aluksi, mikä on matriisi ja miten sitä merkitään. Määritelmän 1.1. matriisissa on vaakasuoria *rivejä* i, missä *rivi-indeksi* $i = 1, 2, 3 \ldots, m$ ja pystysuoria *sarakkeita* j, missä *sarakeindeksi* $j = 1, 2, 3 \ldots, n$ siten, että jokaisella rivillä ja jokaisella sarakkeella on yhtä monta *alkiota* eli reaalilukua. [**8**, Luku II]

Määritelmä 1.1. Matriisia $A \in \mathbb{R}^{m \times n}$, jossa on m-kappaletta rivejä ja n-kappaletta sarakkeita, kutsutaan tyypin $m \times n$ matriisiksi ja sitä merkitään seuraavasti:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = (a_{ij})_{m \times n} = (a_{ij}).$$

Matriisin A rivillä i ja sarakkeessa j olevaa matriisin alkiota merkitään a_{ij}.

Määritelmässä 1.1. matriisin A alkiot ovat reaalilukuja eli $a_{ij} \in \mathbb{R}$. Jos kuitenkin jokin matriisin A alkio olisi kompleksiluku eli $a_{ij} \in \mathbb{C}$, niin tällöin merkittäisiin $A \in \mathbb{C}^{m \times n}$. Tässä tutkielmassa käytetään jatkossa yleisesti merkintää $F^{m \times n}$, joka tarkoittaa joko $\mathbb{R}^{m \times n}$ tai $\mathbb{C}^{m \times n}$.

Matriisin $A \in F^{m \times n}$ lukuja m ja n kutsutaan matriisin A *kertaluvuiksi*. Kertaluvut kertovat matriisin tyypin $m \times n$, jota kutsutaan myös *dimensioksi*. Sanotaan, että matriisi A on $m \times n$ -dimensioinen.

Määritelmä 1.2. Olkoon $A \in F^{m \times n}$.
a) Jos matriisin A kertaluvut ovat samat eli $m = n$, niin A on *neliömatriisi*.
b) Jos $m > n$, niin A on *korkea matriisi*.
c) Jos $m < n$, niin A on *leveä matriisi*.

Tässä tutkielmassa tarkastellaan pääsääntöisesti neliömatriiseja ja paneudutaan niiden ominaisuuksiin. Käydään läpi esimerkki reaalisesta neliömatriisista.

Esimerkki 1.3. Olkoon B kokoa 2×2 oleva matriisi, jonka alkiot ovat $b_{11} = 3$, $b_{12} = -5$, $b_{21} = 6$ ja $b_{22} = 4$. Sijoittamalla alkiot omille paikoilleen saadaan matriisi

$$B = \begin{bmatrix} 3 & -5 \\ 6 & 4 \end{bmatrix}.$$

Koska matriisissa B on 2 riviä ja 2 saraketta eli $m = n$, niin yllä olevan määritelmän 1.2 a) -kohdan nojalla B on neliömatriisi, jonka alkiot ovat reaalisia. Tällöin merkitään $B \in \mathbb{R}^{2 \times 2}$.

MÄÄRITELMÄ 1.4. Neliömatriisin $A \in F^{n\times n}$ alkiota a_{ij} kutsutaan *lävistäjäalkioksi*, jos $i = j$ eli rivi- ja sarakeindeksit ovat samat. Lävistäjäalkiosta voidaan käyttää myös nimitystä *diagonaalialkio*. Tällöin yleensä merkitään

$$diag(A) = (a_{11}, a_{22}, \dots, a_{nn}).$$

Esimerkiksi esimerkin 1.3. matriisin B lävistäjäalkiot ovat $b_{11} = 3$ ja $b_{22} = 4$, koska tällöin $i = j$. Näin ollen $diag(B) = (3, 4)$.

1.2. Matriisityyppejä

Seuraavassa esitellään muutamia matriisityyppejä, joiden oletetaan myöhemmin olevan tunnettuja. *Identtinen matriisi* on neliömatriisi, jonka kaikki lävistäjäalkiot ovat ykkösiä ja muut alkiot ovat nollia. Tällöin, jos $A \in \mathbb{R}^{n\times n}$ on identtinen matriisi, niin $a_{ij} = 0$, kun $i \neq j$, ja $a_{ii} = 1$, kaikilla $i, j = 1, 2 \dots n$. Identtistä matriisia merkitään yleensä isolla I -kirjaimella:

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Identtinen matriisi on erikoistapaus *diagonaalimatriisista*. Sanotaan, että neliömatriisi $A \in \mathbb{R}^{n\times n}$ on diagonaalimatriisi, jos $a_{ij} = 0$ kaikilla $i \neq j$.

Neliömatriisi on *alakolmiomatriisi*, jos $a_{ij} = 0$, kun $i < j$. Tällöin alakolmiomatriisissa kaikki lävistäjäalkioiden yläpuolella olevat alkiot ovat nollia. Esimerkiksi $A \in F^{n\times n}$ on alakolmiomatriisi:

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Neliömatriisi on *yläkolmiomatriisi*, jos $a_{ij} = 0$, kun $i > j$. Tällöin yläkolmiomatriisissa kaikki lävistäjäalkioiden alapuolella olevat alkiot ovat nollia. Esimerkiksi $B \in F^{n\times n}$ on yläkolmiomatriisi:

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{nn} \end{bmatrix}.$$

Ala- ja yläkolmiomatriisien tunnistamisesta on hyötyä, koska niille pätevät seuraavat erityisominaisuudet: [**10**, Luku 2]

i) Determinantti on diagonaalialkioiden tulo.
ii) Alakolmiomatriisin käänteismatriisi on alakolmiomatriisi ja vastaavasti yläkolmiomatriisin käänteismatriisi on yläkolmiomatriisi.
iii) Kahden alakolmiomatriisin tulo on alakolmiomatriisi. Vastaavasti myös kahden yläkolmiomatriisin tulo on yläkolmiomatriisi.

Ominaisuuksissa esiintyvät determinantti ja käänteismatriisi käsitellään luvussa 2 ja matriisitulo kohdassa 1.13. Erityisesti determinantin laskeminen isoille neliömatriiseille, jotka eivät ole ala- tai yläkolmiomatriiseja, voi olla työlästä.

1.3. Vektori

Matriisit koostuvat *vektoreista*, jotka ovat joko pysty- tai vaakasuuntaisia. Tässä tutkielmassa käsiteltävät vektorit voivat olla joko reaalisia tai kompleksisia.

MÄÄRITELMÄ 1.5. Matriisia, jossa on vain yksi rivi i tai sarake j, kutsutaan vektoriksi.

ESIMERKKI 1.6. Esimerkiksi alla olevassa matriisissa a on vain yksi sarake. Tällaista yksisarakkeista matriisia kutsutaan *pysty-* tai *sarakevektoriksi*. Vektoria merkitään yleensä pienellä kirjaimella. Pystyvektori a on $m \times 1$ -dimensioinen

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}.$$

Yleensä merkitään yksinkertaisesti $a \in \mathbb{R}^m$ sen sijaan, että käytettäisiin merkintää $a \in \mathbb{R}^{m\times 1}$. Jossain mielessä voidaan ajatella, että *skalaari* eli luku $t \in \mathbb{R}$ on 1×1 -dimensioinen matriisi.

Vastaavaan tapaan kuin esimerkissä 1.6 voidaan määritellä *vaaka-* tai *rivivektori*, joka on rivin suuntainen. Vaakavektori b on $1 \times n$ -dimensioinen

$$b = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}.$$

Vektoreiden muodostamaa joukkoa kutsutaan *vektoriavaruudeksi* V. Tässä tutkielmassa käsitellään F-kertoimista vektoriavaruutta V. Vektoriavaruudessa V on määritelty yhteenlasku $(+)$ ja luvulla kertominen $(\cdot)$, joilla ovat voimassa seuraavat ehdot [**7**]:

1) Yhteenlaskulle $(V, +)$ pätee
   i) $(x + y) + z = x + (y + z)$ kaikilla $x, y, z \in V$, (assosiatiivisuus)
   ii) $x + y = y + x$ kaikilla $x, y \in V$, (kommutatiivisuus)
   iii) On olemassa nollavektori $0 \in V$ siten, että $x + 0 = x$ kaikilla $x \in V$,
   iv) Jokaiselle $x \in V$ on vastavektori $-x \in V$ siten, että $x + (-x) = 0$.
2) Luvulla kertomiselle $(V, \cdot)$ pätee
   v) $\lambda(x + y) = \lambda x + \lambda y$ kaikilla $\lambda \in F$ ja $x, y \in V$, (distributiivisuus)
   vi) $(\lambda + \mu)x = \lambda x + \mu x$ kaikilla $\lambda, \mu \in F$ ja $x \in V$, (distributiivisuus)
   vii) $(\lambda\mu)x = \lambda(\mu x)$ kaikilla $\lambda, \mu \in F$ ja $x \in V$,
   viii) $1 \cdot x = x$ kaikilla $x \in V$.

Esimerkiksi

$$F^{m\times n} = \{A = (a_{ij})_{m\times n} \mid a_{ij} \in F,\ i = 1, 2, \ldots, m,\ j = 1, 2 \ldots, n\}$$

on vektoriavaruus seuraavassa kappaleessa 1.4 määriteltävien laskutoimitusten suhteen.

1.4. Matriisien perusoperaatiot

Matriiseilla on joukko perusoperaatioita, jotka on syytä hallita ennen kuin voi paneutua syvemmälle matriiseihin ja niiden ominaisuuksiin. Tässä osiossa esitetään matriisin perusoperaatiot: transpoosi, yhteen- ja vähennyslasku, skalaarilla kertominen sekä matriisitulo. Lopussa käydään läpi matriisien laskusääntöjä.

Transpoosissa matriisin rivit ja sarakkeet vaihdetaan keskenään toisin päin. Transpoosi on matriisin perusoperaatio, jota merkitään yleensä isolla T -kirjaimella.

MÄÄRITELMÄ 1.7. Olkoon $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ matriisi. Matriisin A transpoosi on

$$A^T = (a_{ji}).$$

Jos transponoitu matriisi A^T on sama kuin alkuperäinen matriisi A eli $A^T = A$, niin matriisi A on *symmetrinen*.

Esimerkiksi esimerkin 1.6 pystyvektori a voidaan muuttaa vaakavektoriksi transpoosin avulla:

$$a^T = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix}.$$

Katsotaan seuraavaksi esimerkki matriisin transponoinnista.

ESIMERKKI 1.8. Olkoon $A \in \mathbb{R}^{2\times 3}$ matriisi, jonka dimensio on 2×3

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

Tällöin transponoinnin jälkeen matriisi muuttuu 3×2 -dimensioiseksi matriisiksi ja saadaan matriisi

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Huomaa, että jos nyt transponoitaisiin transponoitu matriisi A^T, niin saataisiin taas matriisi A. Transpoosin transpoosi on siis sama kuin alkuperäinen matriisi:

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

Kompleksinen avaruus $\mathbb{C}^n$ koostuu $n \times 1$ -dimensioisista kompleksivektoreista

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ missä kertoimet } x_i \in \mathbb{C}.$$

Kompleksisessa avaruudessa on ominaisuuksia, joita ei ole reaalisessa avaruudessa: kompleksiluvun $z = a + bi$ *konjugaatti* eli *liittoluku* on $\bar{z} = a - bi$. Tällöin myös vektorilla $x \in \mathbb{C}^n$ on konjugaatti eli *liittovektori* $\bar{x}$, joka määritellään seuraavasti:

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix}.$$

Kompleksisen matriisin transponointia, jossa samalla konjugoidaan alkiot, kutsutaan *hermitoinniksi.* Seuraavassa määritelmässä 1.9 määritellään hermitointi kompleksiselle matriisille ja esimerkissä 1.10 hermitoidaan 2×2 -dimensioinen kompleksinen matriisi.

MÄÄRITELMÄ 1.9. Olkoon $A = (a_{ij}) \in \mathbb{C}^{m \times n}$ matriisi. Tällöin hermitointia merkitään

$$A^* = \bar{A}^T = (\bar{a}_{ji}),$$

missä $\bar{A}$ tarkoittaa matriisin A jokaisen alkion konjugointia.

Huomaa, että kompleksinen vektori ja matriisi voidaan pelkästään transponoida määritelmän 1.7 mukaisesti, jolloin kyse ei ole hermitoinnista.

ESIMERKKI 1.10. Olkoon $A \in \mathbb{C}^{2 \times 2}$ kompleksinen matriisi

$$A = \begin{bmatrix} 2-i & 3i \\ -4 & 1+i \end{bmatrix}.$$

Tällöin, kun matriisi A hermitoidaan, saadaan

$$A^* = \begin{bmatrix} 2+i & -4 \\ -3i & 1-i \end{bmatrix}.$$

Jatkossa käytetään samaa * -merkintää kompleksisen vektorin hermitoinnissa. Esimerkiksi, jos $x \in \mathbb{C}^n$, merkitään $x^* = \bar{x}^T$.

Seuraavassa määritellään lyhyesti matriiseille tarpeellisia laskutoimituksia: yhteen- ja vähennyslasku, skalaarilla kertominen ja matriisien kertolasku. Näissä laskutoimituksissa matriisien tulee olla tietyn tyyppisiä, jotta laskutoimitus on hyvin määritelty.

Matriiseja voi laskea yhteen ja vähentää, jos ne ovat samantyyppisiä eli niiden dimensiot ovat keskenään samat. [**15**]

MÄÄRITELMÄ 1.11. Olkoot A ja B matriiseja, joiden dimensiot ovat $m \times n$. Matriisien yhteenlasku määritellään $(A + B) = (a_{ij} + b_{ij})$, missä $i = 1, 2 \ldots, m$ ja $j = 1, 2 \ldots, n$.

Yhteenlaskussa siis lasketaan yhteen jokainen alkio matriisien A ja B riviltä i ja sarakkeesta j. Tämän vuoksi matriisien tulee olla samantyyppiset. Esimerkiksi, jos kaksi identtistä 2×2 -dimensioista matriisia lasketaan yhteen saadaan 2×2 -dimensioinen diagonaalimatriisi:

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

MÄÄRITELMÄ 1.12. Olkoot A ja B matriiseja, joiden dimensiot ovat $m \times n$. Matriisien vähennyslasku määritellään $(A - B) = (a_{ij} - b_{ij})$, missä $i = 1, 2 \ldots, m$ ja $j = 1, 2 \ldots, n$.

Voidaan ajatella, että vähennyslasku on yhteenlasku, jossa otetaan matriisin B alkioiden vastaluvut, toisin sanoen $(A - B) = (A + (-B)) = (a_{ij} + (-b_{ij}))$.

Matriisi $A \in F^{m \times n}$ voidaan kertoa jollakin skalaarilla eli luvulla $c \in F$ seuraavasti:

$$cA = (ca_{ij}) = (a_{ij}c) = Ac, \text{ missä } i = 1, 2 \ldots, m \text{ ja } j = 1, 2 \ldots, n.$$

Skalaarilla kertomisessa jokainen matriisin A alkio siis kerrotaan luvulla c.

Matriisien kertolasku on aiempia laskutoimituksia hieman monimutkaisempi. Matriisien kertolasku on kuitenkin hyvin määritelty, jos kerrottavat matriisit ovat keskenään oikean dimensioisia. [**14**]

MÄÄRITELMÄ 1.13. Olkoon $A \in F^{m\times n}$ ja $B \in F^{p\times q}$ matriiseja ja $p = n$. Tällöin matriisien A ja B tulo on matriisi $C = AB \in F^{m\times q}$, jonka alkiot ovat:

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj},$$

missä $i = 1, 2, \ldots, m$ ja $j = 1, 2, \ldots, q$.

Tulomatriisissa jokainen alkio rivillä i ja sarakkeessa j on summa vastinalkioiden tuloista. Esimerkiksi jos $i = 1$ ja $j = 1$, niin tällöin lasketaan yhteen ensimmäisen rivin ja ensimmäisen sarakkeen alkioiden tulot. Katsotaan seuraavaksi esimerkki matriisien kertolaskusta.

ESIMERKKI 1.14. Olkoon $A \in \mathbb{R}^{4\times 2}$ ja $B \in \mathbb{R}^{2\times 3}$ matriiseja

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

Nyt voidaan laskea tulomatriisi C, koska dimensiot ovat kohdallaan $(2 = 2)$,

$$\begin{aligned} C = AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} &= \begin{bmatrix} 1\cdot 1 + 2\cdot 4 & 1\cdot 2 + 2\cdot 5 & 1\cdot 3 + 2\cdot 6 \\ 3\cdot 1 + 4\cdot 4 & 3\cdot 2 + 4\cdot 5 & 3\cdot 3 + 4\cdot 6 \\ 5\cdot 1 + 6\cdot 4 & 5\cdot 2 + 6\cdot 5 & 5\cdot 3 + 6\cdot 6 \\ 7\cdot 1 + 8\cdot 4 & 7\cdot 2 + 8\cdot 5 & 7\cdot 3 + 8\cdot 6 \end{bmatrix} \\ &= \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \\ 39 & 54 & 69 \end{bmatrix}. \end{aligned}$$

Nyt voidaan vielä tarkistaa, että tulomatriisin dimensio on "oikea": $C \in \mathbb{R}^{4\times 3}$ eli neljä riviä ja kolme saraketta, niin kuin pitääkin. Huomataan, että matriisitulolle ei päde *kommutatiivisuus* eli vaihdannaisuus. Tällöin siis yleensä $AB \neq BA$. Tässä esimerkissä matriisitulo BA ei ole edes määritelty.

Seuraavassa esitellään matriiseille voimassa olevia laskusääntöjä. [**14**] Jokaisessa säännössä matriisin tyypin tulee olla sellainen, että haluttu laskutoimitus on määritelty. Sääntöjen kohdissa d), f), h) ja j) esiintyvät skalaarit eli luvut $0, 1$ ja c. Kohdan a), d) ja e) säännöissä O on sopivan kokoinen *nollamatriisi*. Nollamatriisilla tarkoitetaan matriisia, jonka kaikki alkiot ovat nollia. Lisäksi kohdassa g) esiintyvä I on sopivan kokoinen identtinen matriisi. Säännöt ovat luokiteltu laskutoimitusten perusteella joko yhteen- tai kertolaskuihin.

Yhteenlaskulle pätee:

a) $A + O = A$

b) $A + B = B + A$ (kommutatiivisuus)

c) $(A + B) + C = A + (B + C)$ (assosiatiivisuus)

Kertolaskulle pätee:

d) $0A = O$
e) $OA = O$ ja $AO = O$
f) $1A = A$
g) $IA = AI = A$
h) $A(BC) = (AB)C$ (assosiatiivisuus)
i) $c(AB) = (cA)B = A(cB)$ (assosiatiivisuus)
j) $c(A + B) = cA + cB$ (distributiivisuus)
k) $A(B + C) = AB + AC$ (distributiivisuus)
l) $(A + B)C = AC + BC$ (distributiivisuus).

Erityisesti näistä seuraa, että aiemmin kappaleen 1.3 lopussa määritelty $F^{m\times n}$ on vektoriavaruus.

LUKU 2

Matriisin kääntyvyys

Tässä luvussa esitetään sääntöjä matriisin *kääntyvyydelle*. Sanotaan, että neliömatriisi $A \in \mathbb{R}^{n\times n}$ on kääntyvä, jos on olemassa samankokoinen matriisi $B \in \mathbb{R}^{n\times n}$ siten, että matriisitulo AB on identtinen matriisi. Matriisien kääntyvyyden tarkasteluun liittyy matriisin *determinantti*, jota käsitellään seuraavaksi. Luvun lopussa käydään myös läpi mikä on *unitaarinen matriisi* ja miten se liittyy kääntyvyyteen.

2.1. Determinantti

Matriisien determinantti on määritelty ainoastaan neliömatriiseille. Jos A on neliömatriisi, niin determinantti on luku, josta käytetään merkintää $\det A$ tai $\mid A \mid$.

Määritellään determinantti ensin 2×2 -tyyppisille matriiseille, jonka jälkeen 3×3 -tyyppiset matriisit pystytään palauttamaan astetta pienempään eli 2×2 -tyyppisiin matriiseihin. Tätä menettelyä voidaan jatkaa suuremmilla matriiseilla, palautetaan aste kerrallaan pienempään matriisiin, kunnes ollaan tapauksessa 2×2. [**15**]

Määritelmä 2.1. Olkoon $A \in \mathbb{R}^{2\times 2}$ ja alkiot $a, b, c, d \in \mathbb{R}$ lukuja. Tällöin

$$\det A = \mid A \mid = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Yleistetään 2×2 -tyyppisten matriisien tapaus $n \times n$ -tyyppisiin matriiseihin. Tällöin determinantissa muodostetaan annetusta matriisista $A \in \mathbb{R}^{n\times n}$ *alimatriisi*, jonka dimensio on $n-1\times n-1$. Tällöin muodostuneessa matriisissa $A_{n-1} \in \mathbb{R}^{n-1\times n-1}$ on yksi rivi i ja sarake j vähemmän. Näin jatketaan kunnes saadaan aikaiseksi luku $\det A$. Tätä kutsutaan matriisin *kehittämiseksi*. Kehittäminen voidaan tehdä minkä tahansa rivin tai sarakkeen suhteen. [**12**, Luku I, kappale 6]

Seuraavassa määritelmässä on kehitetty matriisia ensimmäisen rivin suhteen.

Määritelmä 2.2. Olkoon $A \in \mathbb{R}^{n\times n}$ matriisi. Tällöin determinantti on luku

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j},$$

missä $A_{1j} \in \mathbb{R}^{n-1\times n-1}$ on alimatriisi, joka saadaan poistamalla matriisista A ensimmäinen rivi ja sarake j.

Jos $\det A = 0$, sanotaan, että matriisi A on *singulaarinen*. Jos $\det A \neq 0$, niin sanotaan, että matriisi A on *ei-singulaarinen*. Tarkastellaan seuraavassa esimerkissä, kuinka matriisia kehitetään määritelmän 2.2 mukaisesti.

Esimerkki 2.3. Olkoon $B \in \mathbb{R}^{3\times 3}$ matriisi, missä $a, b, c, d, e, f, g, h, i \in \mathbb{R}$. Nyt kehittämällä determinanttia ensimmäisen vaakarivin suhteen määritelmän 2.2 mukaan, saadaan aikaiseksi 2×2 -dimensioisia *alideterminantteja*, jotka osataan laskea:

$$\det B = \mid B \mid = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \mid B_{11} \mid - b \mid B_{12} \mid + c \mid B_{13} \mid$$

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

ja nyt osataan laskea 2×2 -tyyppiset determinantit

$$a(ei - hf) - b(di - gf) + c(dh - ge) = aei - ahf - bdi + bgf + cdh - cge.$$

Käydään läpi vielä toinen esimerkki, jossa lasketaan isomman neliömatriisin determinantti. Nähdään, että mitä suurempien neliömatriisien determinantteja lasketaan sitä enemmän niitä joudutaan kehittämään.

Esimerkki 2.4. Olkoon

$$A = \begin{bmatrix} 5 & -2 & 6 & 4 \\ 1 & 0 & -7 & 8 \\ 9 & 8 & 4 & 6 \\ -1 & 4 & 3 & 2 \end{bmatrix} \in \mathbb{R}^{4\times 4}.$$

Kehitetään matriisia ensimmäisen vaakarivin suhteen:

$$\det A = 5 \begin{vmatrix} 0 & -7 & 8 \\ 8 & 4 & 6 \\ 4 & 3 & 2 \end{vmatrix} - (-2) \begin{vmatrix} 1 & -7 & 8 \\ 9 & 4 & 6 \\ -1 & 3 & 2 \end{vmatrix} + 6 \begin{vmatrix} 1 & 0 & 8 \\ 9 & 8 & 6 \\ -1 & 4 & 2 \end{vmatrix} - 4 \begin{vmatrix} 1 & 0 & -7 \\ 9 & 8 & 4 \\ -1 & 4 & 3 \end{vmatrix}.$$

Ja jatketaan determinanttien kehittämistä edelleen...

$$5 \left(0 \begin{vmatrix} 4 & 6 \\ 3 & 2 \end{vmatrix} - (-7) \begin{vmatrix} 8 & 6 \\ 4 & 2 \end{vmatrix} + 8 \begin{vmatrix} 8 & 4 \\ 4 & 3 \end{vmatrix} \right) + 2 \left(1 \begin{vmatrix} 4 & 6 \\ 3 & 2 \end{vmatrix} - (-7) \begin{vmatrix} 9 & 6 \\ -1 & 2 \end{vmatrix} + 8 \begin{vmatrix} 9 & 4 \\ -1 & 3 \end{vmatrix} \right)$$
$$+ 6 \left(1 \begin{vmatrix} 8 & 6 \\ 4 & 2 \end{vmatrix} - 0 \begin{vmatrix} 9 & 6 \\ -1 & 2 \end{vmatrix} + 8 \begin{vmatrix} 9 & 8 \\ -1 & 4 \end{vmatrix} \right) - 4 \left(1 \begin{vmatrix} 8 & 4 \\ 4 & 3 \end{vmatrix} - 0 \begin{vmatrix} 9 & 4 \\ -1 & 3 \end{vmatrix} - 7 \begin{vmatrix} 9 & 8 \\ -1 & 4 \end{vmatrix} \right)$$
$$= 5\,(0(8-18) + 7(16-24) + 8(24-16)) + 2\,(1(8-18) + 7(18+6) + 8(27+4))$$
$$+ 6\,(1(16-24) - 0(18+6) + 8(36+8)) - 4\,(1(24-16) - 0(27+4) - 7(36+8))$$
$$= 5(-56+64) + 2(-10+168+248) + 6(-8+352) - 4(8-308)$$
$$= 40 + 812 + 2064 + 1200 = 4116.$$

2.2. Kääntyvä matriisi

Tässä osiossa tarkastellaan, mitä tarkoitetaan matriisin kääntyvyydellä ja kuinka saadaan selville, onko tarkasteltava matriisi kääntyvä. Osion lopussa esitellään kääntyvän matriisin laskusääntöjä.

Määritelmä 2.5. Neliömatriisi $A \in \mathbb{R}^{n\times n}$ on kääntyvä, jos löytyy samankokoinen neliömatriisi $B \in \mathbb{R}^{n\times n}$ siten, että $AB = BA = I$. Tällöin matriisia B sanotaan matriisin A käänteismatriisiksi ja sitä merkitään $B = A^{-1}$.

Määritelmän 2.5 matriisin A käänteismatriisi $B = A^{-1}$ on yksikäsitteinen, koska jos oletetaan lisäksi, että $C \in \mathbb{R}^{n\times n}$ ja $CA = I = AC$, niin saadaan

$$C = CI = CAB = IB = B.$$

Seuraava lause 2.6 kertoo, kuinka matriisin kääntyvyyttä voidaan tarkastella matriisin determinantin avulla.

LAUSE 2.6. *Matriisi* $A \in \mathbb{R}^{n \times n}$ *on kääntyvä jos ja vain jos* $\det A \neq 0$.

TODISTUS. Katso [**8**, Lause 11.1] □

Lauseessa 2.6 sanotaan, että matriisi on kääntyvä täsmälleen silloin, kun matriisi on ei-singulaarinen. Käänteismatriisia voidaan kutsua myös matriisin *inverssiksi.*

ESIMERKKI 2.7. Tutkitaan esimerkissä 1.3 esiintyvän matriisin B kääntyvyyttä. Lasketaan ensin matriisin $B \in \mathbb{R}^{2 \times 2}$ determinantti ja saadaan

$$\det B = \begin{vmatrix} 3 & -5 \\ 6 & 4 \end{vmatrix} = (3 \cdot 4 - 6 \cdot (-5)) = 12 - (-30) = 42.$$

Nyt koska $\det B \neq 0$, niin matriisi B on ei-singulaarinen ja lauseen 2.6 nojalla B on kääntyvä. Näin ollen voidaan laskea käänteismatriisi B^{-1}. Merkitään $B^{-1} = (a_{ij})$. Määritelmän 2.5 mukaan on siis oltava $BB^{-1} = I$. Tällöin voidaan muodostaa seuraava yhtälö:

$$\begin{bmatrix} 3 & -5 \\ 6 & 4 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Nyt on siis oltava

$$\begin{aligned} &3 \cdot a_{11} + (-5) \cdot a_{21} = 1, \\ &6 \cdot a_{11} + 4 \cdot a_{21} = 0, \\ &3 \cdot a_{12} + (-5) \cdot a_{22} = 0, \\ &6 \cdot a_{12} + 4 \cdot a_{22} = 1. \end{aligned}$$

Tästä saadaan yhtälöryhmät:

$$\begin{cases} 3 \cdot a_{11} + (-5) \cdot a_{21} = 1 \\ 6 \cdot a_{11} + 4 \cdot b_{21} = 0, \end{cases} \tag{2.1}$$

$$\begin{cases} 3 \cdot a_{12} + (-5) \cdot a_{22} = 0 \\ 6 \cdot a_{12} + 4 \cdot a_{22} = 1. \end{cases} \tag{2.2}$$

Ratkaisemalla nyt ensin yhtälöryhmä (2.1) saadaan,

$$a_{11} = \frac{2}{21}\,, \quad a_{21} = \frac{-1}{7}$$

ja yhtälöryhmästä (2.2),

$$a_{12} = \frac{5}{42}\,, \quad a_{22} = \frac{1}{14}.$$

Näin ollaan saatu ratkaistua käänteismatriisi

$$B^{-1} = \begin{bmatrix} \frac{2}{21} & \frac{-1}{7} \\ \frac{5}{42} & \frac{1}{14} \end{bmatrix}.$$

Esimerkin 2.7 käänteismatriisin voi laskea myös Gaussin & Jordanin menetelmällä, jota ei tässä tutkielmassa käsitellä. Katso esimerkiksi [**8**, luku II].

Sanotaan, että matriisit A ja B ovat *similaariset*, jos on olemassa sellainen kääntyvä matriisi S, jolle $S^{-1}AS = B$. Similaarisuudesta seuraa tärkeitä ominaisuuksia, joita tullaan käsittelemään lisää lukujen 3 ja 6 yhteydessä.

Esitellään seuraavaksi kääntyville neliömatriiseille tyypillisiä ja jatkossa myös hyödyllisiä tuloksia. Olkoot A ja B kääntyviä samankokoisia neliömatriiseja. Tällöin seuraavat laskusäännöt ovat voimassa [**8**, Luku II]:

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

Laskusäännöt voi todistaa suoraan laskemalla. Esimerkiksi 1) sääntö pätee, koska jos merkitään $C = A^{-1}$, niin

$$CA = AC = I$$

ja tällöin määritelmän nojalla

$$C^{-1} = (A^{-1})^{-1} = A.$$

2.3. Unitaarinen matriisi

Kompleksikertoiminen neliömatriisi $A \in \mathbb{C}^{n\times n}$ on unitaarinen, jos

$$A^*A = AA^* = I$$

eli matriisi A kerrottuna hermitoidulla matriisilla A^* on identtinen neliömatriisi. Yleensä unitaarista matriisia merkitään kirjaimella U.

Esimerkki 2.8. Matriisi

$$U = \begin{bmatrix} -1 & 0 \\ 0 & i \end{bmatrix}$$

on unitaarinen, koska

$$U^*U = \begin{bmatrix} -1 & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Koska unitaariselle matriisille $U \in \mathbb{C}^{2\times 2}$ pätee $U^*U = UU^* = I$, niin tällöin täytyy olla $U^* = U^{-1}$ ja näin ollen U on myös kääntyvä.

Huomaa, että jos $A \in \mathbb{R}^{n\times n}$ ja

$$A^*A = AA^* = I,$$

niin tällöin hermitointi transponoi matriisin A eli $A^* = A^T$. Nyt erityisesti saadaan, että $A^TA = I$, josta edelleen saadaan, että $A^T = A^{-1}$.

LUKU 3

Matriisin ominaisarvot

Matriisin ominaisarvot ovat joukko matriisille ominaisia lukuja. Ominaisarvoista koostuu matriisin *spektri*. Luvussa 6 käsitellään spektraalinormia, jota varten esitetään ominaisarvot ja ominaisvektorit. Tässä luvussa määritelläänkin ensin ominaisarvoyhtälö ja esitetään lause, jolla saadaan laskettua kyseisen matriisin ominaisarvot. Luvun lopussa listataan ominaisarvoihin liittyviä tärkeitä tuloksia.

3.1. Ominaisarvoyhtälö

MÄÄRITELMÄ 3.1. Olkoon $A \in F^{n \times n}$ matriisi. Lukua $\lambda \in \mathbb{C}$ kutsutaan matriisin A *ominaisarvoksi*, jos on olemassa vektori $x \neq 0$ siten, että

$$Ax = \lambda x. \tag{3.1}$$

Määritelmän 3.1 vektorit $x \neq 0$, jotka toteuttavat yhtälön (3.1), ovat ominaisarvoon λ liittyviä *ominaisvektoreita*. Yhtälö (3.1) on yhtäpitävä yhtälön

$$Ax = \lambda I x$$

kanssa, koska $\lambda x = \lambda I x$. Edelleen saadaan uudelleen sijoittelemalla

$$(A - \lambda I)x = 0. \tag{3.2}$$

Yhtälö (3.2) on homogeeninen lineaarinen yhtälö, jonka kerroinmatriisi on $A - \lambda I$. Nyt koska kyseessä on neliömatriisi, niin yhtälöllä on ei-triviaaleja ratkaisuja täsmälleen silloin, kun

$$\det(A - \lambda I) = 0. \tag{3.3}$$

Laskemalla seuraavaksi determinantti saadusta lausekkeesta (3.3) saadaan matriisin A *karakteristinen polynomi*. Polynomin nollakohdat ovat matriisin A ominaisarvot. Tällöin ominaisarvot voivat olla myös kompleksisia. Lisäksi on tärkeää muistaa, että algebran peruslauseen nojalla matriisilla $A \in F^{n \times n}$ on aina kertaluvut huomioiden n kappaletta nollakohtia eli tässä tilanteessa n kappaletta ominaisarvoja.

Katsotaan seuraavaksi esimerkki matriisin ominaisarvojen ja ominaisvektoreiden laskemisesta.

ESIMERKKI 3.2. Olkoon $A \in \mathbb{R}^{2 \times 2}$ matriisi

$$A = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix} \quad \text{ja} \quad \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}.$$

Nyt voidaan laskea determinantti

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & -2 \\ -1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) - (-1(-2)) \\ &= (2 - 2\lambda - \lambda + \lambda^2) - 2 = \lambda^2 - 3\lambda, \end{aligned}$$

ja nyt asettamalla saatu polynomi yhtäsuureksi nollan kanssa saadaan

$$\lambda^2 - 3\lambda = 0 \Leftrightarrow \lambda_1 = 3 \text{ ja } \lambda_2 = 0,$$

jotka ovat matriisin A ominaisarvot.

Lasketaan nyt ominaisvektorit annetuille ominaisarvoille. Sijoitetaan saatu ratkaisu $\lambda_1 = 3$ yhtälöön 3.2 ja saadaan $(A - 3I)x = 0$. Nyt siis

$$\begin{bmatrix} 2-3 & -2 \\ -1 & 1-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0,$$

mistä saadaan yhtälöryhmä

$$\begin{cases} -x_1 - 2x_2 = 0 \\ -x_1 - 2x_2 = 0. \end{cases}$$

Tästä nähdään, että $x_1 = -2x_2$, jolloin yksi ominaisarvoon $\lambda_1 = 3$ liittyvä ominaisvektori on

$$x_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Tällöin kaikki ratkaisut ovat muotoa

$$x_1 = t \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

missä $t \in \mathbb{R}, t \neq 0$.

Samaan tapaan ratkaisemalla ominaisarvoon $\lambda_2 = 0$ liittyvät ominaisvektorit saadaan

$$\begin{bmatrix} 2-0 & -2 \\ -1 & 1-0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \quad \Leftrightarrow \quad \begin{cases} 2x_1 - 2x_2 = 0 \\ -x_1 + x_2 = 0. \end{cases}$$

Nyt nähdään, että on oltava $x_1 = x_2$, jolloin eräs ominaisarvoon $\lambda_2 = 0$ liittyvä ominaisvektori on

$$x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ja kaikki ratkaisut ovat muotoa

$$x_2 = t \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

missä $t \in \mathbb{R}, t \neq 0$.

3.2. Ominaisarvoihin liittyviä tuloksia

Luvussa 2 käsiteltiin matriisin kääntyvyyttä ja unitaarisuutta. Nyt kun olemme saaneet käsiteltyä ominaisarvot, voidaan tutustua tunnettuun lauseeseen nimeltä *Schurin lause* ja yleisesti ominaisarvoihin lliittyviin tuloksiin.

Lause 3.3. *Olkoon $A \in F^{n\times n}$, jonka ominaisarvot ovat $\lambda_1, \lambda_2 \ldots, \lambda_n$ missä tahansa järjestyksessä. Tällöin on unitaarinen matriisi $U \in F^{n\times n}$ siten, että $U^*AU = T = (t_{ij})$ on yläkolmiomatriisi, jonka diagonaalilla on ominaisarvot $t_{ii} = \lambda_i$, missä $i = 1, 2, \cdots, n$.*

Todistus. Katso [**3**, Lause 2.3.1] □

Lause 3.3 kertoo, että mille tahansa neliömatriisille A löytyy vastaava yläkolmiomatriisi, jonka diagonaalilla ovat matriisin A ominaisarvot annetussa järjestyksessä.

Käydään seuraavaksi läpi hyödyllisiä tuloksia ominaisarvoihin ja ominaisvektoreihin liittyen. Olkoon $A \in F^{n\times n}$ matriisi, jonka ominaisarvot ovat $\lambda_1, \lambda_2, \dots, \lambda_n$, tällöin seuraavat ominaisuudet ovat voimassa [**13**]:

a) $\det A = \prod_{i=1}^{n} \lambda_i$.

b) Käänteismatriisin A^{-1} ominaisarvot ovat $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}$.

c) Vakiolla $c \in \mathbb{R}$ kerrotun matriisin cA ominaisarvot ovat $c\lambda_1, c\lambda_2, \dots, c\lambda_n$ ja ominaisvektorit ovat samat kuin matriisilla A.

d) Matriisin A^T ominaisarvot ovat samat kuin matriisin A, mutta ominaisvektorit eivät välttämättä ole samat.

e) Diagonaali-, alakolmio- ja yläkolmiomatriisin ominaisarvot ovat samat kuin sen diagonaalialkiot.

Kohdan (e) ominaisuutta hyödynnetään myöhemmin luvun 6 yhteydessä. Muita edellä esitettyjä tuloksia ei tässä tutkielmassa myöhemmin tulla tarvitsemaan, joten niitä ei myöskään todisteta.

3.3. Similaarisuus ja diagonalisoituvuus

Aiemmin luvussa 2 todettiin, että matriisit A ja B ovat similaariset, jos

$$S^{-1}AS = B,$$

jollekin kääntyvälle matriisille S. Nyt jos matriisit A ja B ovat similaarisia, niin niillä on samat ominaisarvot. Tämä sen vuoksi, että niiden karakteristiset polynomit ovat samat. Tämän voi tarkistaa suoralla laskulla. [**13**]

Otetaan esimerkki similaarisista matriiseista.

Esimerkki 3.4. Olkoon $A \in \mathbb{R}^{2\times 2}$, $B \in \mathbb{R}^{2\times 2}$ ja $S \in \mathbb{R}^{2\times 2}$ seuraavat matriisit:

$$A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Tarkastellaan ovatko matriisit A ja B similaarisia. Lasketaan ensin matriisin S determinantti ja tutkitaan onko se kääntyvä.

Matriisi S on kääntyvä, koska $\det S = (1 - (-1) = 2 \neq 0$. Nyt jos

$$S^{-1}AS = B,$$

niin täytyy olla myös

$$AS = SB,$$

josta edelleen saadaan yhtäpitävästi

$$A = SBS^{-1}.$$

Lasketaan seuraavaksi onko $AS = SB$:

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

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

Tällöin siis $A = SBS^{-1}$. Nyt tiedetään, että matriisit A ja B ovat similaarisia ja niillä on samat ominaisarvot. Riittää siis laskea vain matriisin A ominaisarvot, jonka jälkeen tiedetään matriisin B ominaisarvot.

Similaarisuutta vahvempi käsite on *diagonalisoituvuus*. Sanotaan, että matriisi $A \in F^{n\times n}$ on diagonalisoituva, jos se on similaarinen diagonaalimatriisin $D \in F^{n\times n}$ kanssa. Tällöin siis $S^{-1}AS = D$ ja lävistäjäalkioina matriisissa D ovat matriisin A ominaisarvot. [**13**]

LUKU 4

Vektorinormi

Vektoreita tarkasteltaessa on mielekästä pystyä määräämään vektoreiden välinen etäisyys. Esimerkiksi tasossa vektoreiden välisen etäisyyden laskeminen onnistuu Pythagoraan lauseen avulla. Tason tapauksessa vektorien välisen etäisyyden laskemisesta käytetään tarkemmin termiä *euklidinen etäisyys*. Etäisyyksiä laskettaessa on olennaista tietää vektorin pituus eli *normi*.

Tässä luvussa määritellään *sisätulo* ja *vektorinormi* -funktioiden ominaisuudet. Vektorinormi perustuu euklidisen etäisyyden laskemiseen. Etäisyyksien laskemiseen puolestaan tarvitaan myös sisätuloa.

4.1. Sisätulo

MÄÄRITELMÄ 4.1. Olkoon x ja y reaalisia (sarake)vektoreita eli $x \in \mathbb{R}^n$ ja $y \in \mathbb{R}^n$. Vektoreiden x ja y sisätulo määritellään seuraavasti

$$\langle x, y\rangle = x \cdot y = x^T y = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Huomataan, että määritelmässä 4.1 vektori x transponoidaan. Näin täytyy tehdä, koska muuten vektoreiden välinen kertolasku ei olisi määritelty. Jatkossa käsitelläänkin määritelmän 4.1 mukaisesti sarakevektoreita ellei toisin mainita.

Nyt on syytä laajentaa sisätulon määritelmä myös kompleksiseen avaruuteen, koska matriisit voivat olla joko reaali- tai kompleksikertoimisia. Kompleksisen avaruuden sisätulossa otetaan toisesta vektorista konjugaatti.

MÄÄRITELMÄ 4.2. Olkoon x ja y kompleksisia vektoreita eli $x \in \mathbb{C}^n$ ja $y \in \mathbb{C}^n$. Vektoreiden x ja y sisätulo määritellään seuraavasti

$$\langle x, y\rangle = x \cdot y = x^T \bar{y} = \sum_{i=1}^{n} x_i \bar{y}_i = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \cdots + x_n \bar{y}_n.$$

Seuraavassa määritellään, mitä eri ominaisuuksia sisätulo -funktiolta vaaditaan yleisessä vektoriavaruudessa. Sekä euklidinen että kompleksinen sisätulo (määritelmissä 4.1 ja 4.2) toteuttaa alla olevan määritelmän 4.3 ehdot.

MÄÄRITELMÄ 4.3. Olkoon V F-kertoiminen vektoriavaruus.
Funktio $\langle \bullet, \bullet\rangle\colon V \times V \to F$ on sisätulo, jos kaikille $x, y, z \in V$ pätee [**3**, Luku 5]

(1) $\langle x, x\rangle \geq 0$,
(1a) $\langle x, x\rangle = 0$, jos ja vain jos $x = 0$,
(2) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$,
(3) $\langle cx, y\rangle = c\langle x, y\rangle$, kaikille $c \in F$,
(4) $\langle x, y\rangle = \overline{\langle y, x\rangle}$.

Määritelmän 4.3 F-kertoimisella vektoriavaruudella V tarkoitetaan epätyhjää joukkoa, jossa on määritelty yhteenlasku ja skalaarilla kertominen kohdan 1.3 lopussa olevien ehtojen mukaisesti. Tästä käytetään myös nimitystä *F-vektoriavaruus*.

Sisätulon avulla saadaan selville tärkeitä tuloksia, kuten vektoreiden välinen kulma, kohtisuoruus sekä vektorin euklidinen pituus eli normi. Erityisesti sanotaan, että vektorit $x \in \mathbb{R}^n$ ja $y \in \mathbb{R}^n$ ovat keskenään *ortogonaaliset*, jos niiden sisätulo on nolla:

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n = 0. \tag{4.1}$$

Reaaliavaruudessa ortogonaalisuus tarkoittaa, että vektorit x ja y ovat kohtisuorassa toisiaan vastaan.

Määritelmä 4.4. Olkoon $x \in \mathbb{R}^n$. Tällöin vektorin euklidinen pituus eli normi määritellään seuraavasti

$$\|x\| = \sqrt{|\, x_1\,|^2 + |\, x_2\,|^2 + \cdots + |\, x_n\,|^2} = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}} = \sqrt{x^T x} = \sqrt{\langle x, x\rangle}.$$

Euklidinen pituus saadaan ottamalla neliöjuuri vektorin itsensä kanssa lasketusta sisätulosta. Kompleksisessa avaruudessa normi lasketaan vastaavalla tavalla kuin reaalisessa avaruudessa, mutta kompleksisella sisätulolla. Olkoon $x \in \mathbb{C}^n$, tällöin

$$\|x\| = \sqrt{|\, x_1\,|^2 + |\, x_2\,|^2 + \cdots + |\, x_n\,|^2} = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}} = \sqrt{x^T \bar{x}} = \sqrt{\langle x, x\rangle}.$$

Jos yhtälön (4.1) vektoreiden $x \in \mathbb{R}^n$ ja $y \in \mathbb{R}^n$ pituudet ovat ykkösiä, niin sanotaan, että vektorit ovat *ortonormaaleja.*

4.2. Vektorinormin ominaisuudet

Määritellään seuraavaksi yleisen vektorinormi -funktion ominaisuudet:

Määritelmä 4.5. Olkoon V F-kertoiminen vektoriavaruus.
Funktio $\|\bullet\|\colon V \to \mathbb{R}$ on vektorinormi, jos kaikille $x, y \in V$ pätee

(1) $\|x\| \geq 0$,
(1a) $\|x\| = 0$ jos ja vain jos $x = 0$,
(2) $\|cx\| = |c|\,\|x\|$ kaikille $c \in F$,
(3) $\|x + y\| \leq \|x\| + \|y\|$.

Nämä neljä aksioomaa ovat tuttuja ominaisuuksia tason euklidiselle pituudelle. Määritelmän 4.5 mukaan vektorinormi on aina ei-negatiivinen ja kolmioepäyhtälö pätee. Funktiota, joka toteuttaa ehdot (1), (2) ja (3), mutta ei ehtoa (1a), kutsutaan *seminormiksi.*

Käsitellään seuraavaksi esimerkki vektorin euklidisen pituuden ja sisätulon laskemisesta.

Esimerkki 4.6. Olkoon $x \in \mathbb{R}^3$ ja $y \in \mathbb{R}^3$ vektoreita

$$x = \begin{bmatrix} 2 \\ -8 \\ 0 \end{bmatrix}, \quad y = \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}.$$

Vektoreiden x ja y euklidiset pituudet ja sisätulo voidaan nyt laskea. Pituudet ovat

$$\begin{aligned} \|x\| &= \sqrt{|\ 2\ |^2 + |\ -8\ |^2 + |\ 0\ |^2} = \sqrt{4+64+0} = \sqrt{68} = 2\sqrt{17}, \\ \|y\| &= \sqrt{|\ -4\ |^2 + |\ 1\ |^2 + |\ 3\ |^2} = \sqrt{16+1+9} = \sqrt{26}, \end{aligned}$$

ja sisätulo on

$$\begin{aligned} \langle x, y\rangle = x^T y &= \begin{bmatrix} 2 & -8 & 0 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix} \\ &= 2 \cdot (-4) + (-8) \cdot 1 + 0 \cdot 3 = -16. \end{aligned}$$

Nyt edellä määriteltyjen sisätulon ja vektorinormin avulla saadaan muutamia hyödyllisiä tuloksia. Yksi tunnetuimmista lauseista on Cauchy-Schwarzin epäyhtälö. Cauchy-Schwarzin epäyhtälöä eli lausetta 4.7 tullaan tässä tutkielmassa käyttämään myöhemmin muun muassa luvun 5 todistuksissa. Cauchy-Swarzin lauseen todistus jätetään kuitenkin lukijan itse opiskeltavaksi. Lisäksi lauseesta saadaan seuraus 4.8, johon tutustuminen jätetään myös lukijan vastuulle.

LAUSE 4.7. *Jos $\langle \bullet, \bullet \rangle$ on sisätulo F-kertoimisessa vektoriavaruudessa V, niin tällöin*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle \langle y, y\rangle \text{ kaikille } x, y \in V.$$

Yhtäsuuruus pätee jos ja vain jos x ja y ovat lineaarisesti riippumattomia eli siis $x = \alpha y$ tai $y = \alpha x$ jollekin $\alpha \in F$.

TODISTUS. Katso [**3**, Lause 5.1.4] □

SEURAUS 4.8. *Jos $\langle \bullet, \bullet \rangle$ on sisätulo vektoriavaruudessa V, niin*

$$\|x\| = (\langle x, x\rangle)^{\frac{1}{2}}$$

on vektorinormi vektoriavaruudesssa V.

TODISTUS. Katso esimerkiksi [**8**, Luku I, kohta 2.3.] □

4.3. Vektorinormien ekvivalenttisuus

Normien *ekvivalenttisuus* tarkoittaa, että on olemassa luvut $t > 0, s > 0$ siten, että kahdelle normille $\|\bullet\|_a$, $\|\bullet\|_b$ pätee

$$t\,\|x\|_a \leq \|x\|_b \leq s\,\|x\|_a\,, \text{ kaikilla } x \in V.$$

Ekvivalenteille normeille saadaan nyt samanlainen *konvergenssi* eli raja-arvon käyttäytyminen. Esimerkiksi *suppenevat jonot* ja *avoimet joukot* ovat samat kummankin normin suhteen. [**12**, Luku V] Sarjan suppenemista ja raja-arvoja tarkastellaan myöhemmin luvussa 6.

Sanotaan, että joukon A alkioiden välillä määritelty relaatio R on *ekvivalenssirelaatio*, jos se toteuttaa seuraavat kolme ehtoa kaikille $a, b, c \in A$:

(1) aRa, (refleksiivisyys)
(2) *Jos* aRb, niin myös bRa, (symmetrisyys)
(3) *Jos* aRb ja bRc, niin aRc. (transitiivisuus)

Normien ekvivalenssi on ekvivalenssirelaatio kaikkien normien joukossa. Tiedetään, että kaikki avaruuden $\mathbb{C}^n$ normit ovat ekvivalentteja. Tämä johtuu yleisemmästä tuloksesta: äärellisulotteisessa F-vektoriavaruudessa kaikki normit ovat keskenään ekvivalentteja. [**5**, Luku 10]

4.4. Erilaisia vektorinormeja

Seuraavassa esitellään hyödyllisiä vektorinormeja.

4.4.1. Euklidinen normi. Euklidinen normi eli l_2 -normi on ehkä yksi tunnetuimmista normeista:

$$\text{Olkoon } x \in \mathbb{C}^n. \text{ Tällöin } \|x\|_2 = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}} = (|x_1|^2 + \cdots + |x_n|^2)^{\frac{1}{2}}.$$

Huomaa, että aiemman määritelmän 4.4 esitys vastaa euklidista normia. Euklidisen normin avulla voidaan laskea kahden pisteen $x, y \in \mathbb{C}^n$ välinen etäisyys vektoreiden erotuksen normina $\|x-y\|_2$.

Esimerkki 4.9. Olkoon $x = (1,4)$ ja $y = (-3,-4)$ vektoreita tasossa. Lasketaan näiden pisteiden välinen etäisyys euklidisella normilla:

$$\begin{aligned}\|x-y\|_2 &= \|(1,4)-(-3,-4)\|_2 = \|(1-(-3), 4-(-4))\|_2 = \|(4,8)\|_2 \\ &= (|4|^2 + |8|^2)^{\frac{1}{2}} = \sqrt{16+64} = \sqrt{80} = 4\sqrt{5}.\end{aligned}$$

4.4.2. Summanormi. Summanormi eli l_1 -normi on nimensä mukaisesti summa kaikkien komponenttien itseisarvoista.

$$\text{Olkoon } x \in \mathbb{C}^n. \text{ Tällöin } \|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + \cdots + |x_n|.$$

4.4.3. Yleinen l_p -normi.

$$\text{Olkoon } x \in \mathbb{C}^n. \text{ Tällöin } \|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}}.$$

Tästä nähdään helposti, että edellä esitetetyt l_1- ja l_2 -normit on saatu yleisestä l_p -normista arvoilla $p = 1$ ja $p = 2$. Nyt yleisen l_p -normin raja-arvona, kun p lähestyy ääretöntä, saadaan vielä maksiminormi eli l_∞ -normi.

4.4.4. Maksiminormi. Maksiminormi eli l_∞ -normi on itseisarvoltaan suurin vektorin alkio.

$$\text{Olkoon } x \in \mathbb{C}^n. \text{ Tällöin } \|x\|_\infty = \max_{1 \le i \le n} |x_i|.$$

Nyt edellä esitellyille normeille saadaan ekvivalenttisuuteen liittyen seuraavat epäyhtälöt [**12**, Luku V]:

$$\begin{aligned}\|x\|_2 \le \|x\|_1 &\le \sqrt{n}\,\|x\|_2\,, \text{ kaikille } x \in \mathbb{C}^n, \\ \|x\|_\infty \le \|x\|_2 &\le \sqrt{n}\,\|x\|_\infty\,, \text{ kaikille } x \in \mathbb{C}^n, \\ \|x\|_\infty \le \|x\|_1 &\le n\,\|x\|_\infty\,, \text{ kaikille } x \in \mathbb{C}^n.\end{aligned}$$

LUKU 5

Matriisinormi

Tässä luvussa määritellään matriisinormi -funktion ominaisuudet ja osoitetaan vektori- ja matriisinormeihin liittyviä tärkeitä tuloksia. Lisäksi tässä luvussa tarkastellaan, mitä yhteyksiä edellisen luvun vektorinormeilla on tämän luvun matriisinormeihin. Matriisinormien avulla saadaan osviittaa matriisin ”suuruusluokasta”. Matriisin ”suuruusluokasta” puhuttaessa ei välttämättä ole kyse siitä, kuinka monta riviä tai saraketta matriisissa on. Matriisinormin voidaan ajatella kertovan, kuinka paljon vektorit enintään venyvät matriisilla kerrottaessa.

Luvussa 1 opittiin, että matriisit koostuvat rivi- ja sarakevektoreista. Tällöin voidaan ajatella, että matriisinormin laskeminen perustuu yksittäisten vektorinormien laskemiseen. Vastaavalla tavalla kuin vektorinormi -funktiolta vaadittiin tiettyjä ominaisuuksia myös matriisinormi -funktiolle tulee määrittää tiettyjä ominaisuuksia.

Tutkielmassa käytetään selvyyden vuoksi matriisinormille merkintää, jossa on kolme pystyviivaa. Tällöin merkinnästä saadaan selville onko kyseessä vektori- vai matriisinormi. Lisäksi jatkossa käytetään merkintää M_n, joka tarkoittaa $n \times n$ neliömatriisien joukkoa joko reaali- tai kompleksiavaruudessa. Joukko M_n muodostaa *renkaan* matriisien yhteenlaskun ja matriisitulon suhteen.

5.1. Matriisinormin ominaisuudet

Matriisinormi -funktion ominaisuudet ovat seuraavat:

MÄÄRITELMÄ 5.1. Funktio $|||\bullet|||\colon M_n \to \mathbb{R}$ on matriisinormi, jos seuraavat ehdot toteutuvat kaikille $A, B \in M_n$:

(1) $|||A||| \geq 0$,
(1a) $|||A||| = 0$ jos ja vain jos $A = 0$,
(2) $|||cA||| = |c|\,|||A|||$ kaikille $c \in F$,
(3) $|||A + B||| \leq |||A||| + |||B|||$,
(4) $|||AB||| \leq |||A|||\;|||B|||$.

Huomataan, että ehdot (1) - (3) ovat täysin samat kuin vektorinormille määritellyt ominaisuudet määritelmässä 4.5.

Matriisinormille määriteltyjen ominaisuuksien perusteella saadaan selville alaraja käänteismatriisin normille. Koska kaikille matriiseille $A \in M_n$ pätee

$$\left|\left|\left|A^2\right|\right|\right| = |||AA||| \leq |||A|||\;|||A||| = |||A|||^2,$$

niin silloin täytyy olla $|||A||| \geq 1$ kaikille nollasta eroaville matriiseille, joille $A^2 = A$. Erityisesti siis $|||I||| \geq 1$ kaikilla matriisinormeilla. Jos A on kääntyvä matriisi, niin silloin $I = AA^{-1}$, jolloin saadaan

$$|||I||| = \left|\left|\left|AA^{-1}\right|\right|\right| \leq |||A|||\;\left|\left|\left|A^{-1}\right|\right|\right|,$$

josta edelleen saadaan

$$|||A^{-1}||| \geq \frac{|||I|||}{|||A|||} \geq \frac{1}{|||A|||}.$$

5.2. Yhteensopiva ja indusoitu matriisinormi

Olkoon $\|\bullet\|$ F:n vektorinormi. Sanotaan, että matriisinormi $|||\bullet|||$ on *yhteensopiva* vektorinormin $\|\bullet\|$ kanssa, jos

$$\|Ax\| \leq |||A|||\,\|x\|\,, \textit{ kaikille } x \in F^n \textit{ ja kaikille } A \in M_n.$$

Yhteensopivuudesta käytetään myös termiä *sopeutuvuus*. Tällöin sanotaan, että matriisinormi sopeutuu vektorinormiin yllä olevan ehdon mukaisesti. Seuraavaksi määritellään, kuinka jokaisesta vektorinormista voidaan *indusoida* matriisinormi.

Määritelmä 5.2. Olkoon F:n vektorinormi $\|\bullet\|$ mikä tahansa. Tällöin vektorinormi indusoi matriisinormin $|||\bullet|||$ seuraavasti:

$$|||A||| = \max_{\|x\|=1} \|Ax\|\,, \text{ missä } A \in M_n.$$

Määritelmän 5.2 normi $|||A|||$ on hyvin määritelty, koska $\|Ax\|$ on jatkuva suljetun ja rajoitetun pallon kuorella ja näin ollen $\|Ax\|$ saavuttaa suurimman ja pienimmän arvonsa.

Määritelmän 5.2 esitystapa on yhtäpitävä myös seuraavien esitystapojen kanssa:

$$\begin{aligned} |||A||| &= \max_{\|x\|=1} \|Ax\| \\ &= \max_{\|x\|\leq 1} \|Ax\| \\ &= \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} \\ &= \max_{\|x\|_\alpha=1} \frac{\|Ax\|}{\|x\|}, \end{aligned}$$

missä $\|\bullet\|_\alpha$ on mikä tahansa vektorinormi. Koska $\|\lambda x\| = |\lambda|\,\|x\|$ (eli vektorinormi on homogeeninen), niin yhtäsuuruus

$$\max_{\|x\|=1} \|Ax\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|}$$

pätee.

Todistetaan seuraavaksi lause 5.3 joka todistaa, että määritelmän 5.2 funktio on todella matriisinormi eli se toteuttaa määritelmän 5.1 ehdot.

Lause 5.3. *Määritelmän 5.2 funktio $|||\bullet|||$ on matriisinormi joukossa M_n. Tällöin $|||I||| = 1$ ja $|||Ax||| \leq |||A|||\,\|x\|$ kaikilla $A \in M_n$ ja $x \in \mathbb{C}^n$.*

Todistus. Jotta funktio $|||A|||$ olisi matriisinormi tulee tarkastaa, että se toteuttaa määritelmän 5.1 ehdot. Ehto (1) pätee, koska $|||A||| \geq 0$ kaikilla $A \in M_n$. Ehto (1a) pätee, koska kaikille $x \neq 0$ saadaan

$$|||A||| = 0 \Leftrightarrow \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = 0 \Leftrightarrow \|Ax\| = 0 \Leftrightarrow Ax = 0 \Leftrightarrow A = 0.$$

Ehto (2) seuraa suorasta laskusta:

$$|||cA||| = \max_{\|x\|=1} \|cAx\| = \max_{\|x\|=1} |c| \, \|Ax\| = |c| \max_{\|x\|=1} \|Ax\| = |c| \, |||A|||.$$

Vastaavalla tavalla suoraan laskemalla nähdään, että kolmioepäyhtälö eli ehto (3) pätee:

$$\begin{aligned} |||A+B||| &= \max_{\|x\|=1} \|(A+B)x\| = \max_{\|x\|=1} \|Ax+Bx\| \le \max_{\|x\|=1} (\|Ax\| + \|Bx\|) \\ &\le \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = |||A||| + |||B|||. \end{aligned}$$

Vielä viimeinen ehto (4) saadaan myös laskemalla:

$$\begin{aligned} |||AB||| &= \max_{\|x\|=1} \|ABx\| \overset{(*)}{=} \max_{x\neq 0} \frac{\|ABx\|}{\|x\|} \overset{(**)}{=} \max_{x\neq 0, Bx\neq 0} \frac{\|ABx\|}{\|Bx\|} \frac{\|Bx\|}{\|x\|} \\ &\le \max_{y\neq 0} \frac{\|Ay\|}{\|y\|} \max_{x\neq 0} \frac{\|Bx\|}{\|x\|} \overset{(*)}{=} \max_{\|y\|=1} \|Ay\| \max_{\|x\|=1} \|Bx\| \\ &= |||A||| \, |||B|||. \end{aligned}$$

Ylläoleva yhtäsuuruus (*) pätee, koska

$$\max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x\neq 0} \left\| A \frac{x}{\|x\|} \right\| = \max_{\|x\|=1} \|Ax\|, \text{ kaikille } A \in M_n.$$

Puolestaan kohdassa (**) epäyhtälö $\geq$ seuraa inkluusiosta eli siitä, että

$$\{x \neq 0, Bx \neq 0\} \subset \{x \neq 0\}.$$

Toisaalta myös kaikille x, jolle $Bx = 0$ pätee

$$\frac{\|ABx\|}{\|x\|} = 0 \le \max_{x\neq 0, Bx\neq 0} \frac{\|ABx\|}{\|x\|},$$

jolloin epäyhtälö pätee myös toiseen suuntaan. Huomataan, että jos $x \neq 0$, niin normin suurimman arvon määritelmän avulla saadaan väite seuraavaan muotoon:

$$\frac{\|Ax\|}{\|x\|} \le |||A|||.$$

Nyt kertomalla $\|x\|$:llä saadaan

$$\|Ax\| \le |||A||| \, \|x\|,$$

joka pätee myös, kun $x = 0$. Lopuksi saadaan, että

$$|||I||| = \max_{\|x\|=1} \|Ix\| = \max_{\|x\|=1} \|x\| = 1.$$

□

Lause 5.3 sanoo, että vektorinormista indusoitu matriisinormi on yhteensopiva indusoivan vektorinormin kanssa ja erityisesti se sanoo, että mille tahansa vektorinormille $\|\bullet\|$ löytyy yhteensopiva matriisinormi $|||\bullet|||$ joukossa M_n. Toisin sanoen nyt tiedetään, että indusoitu matriisinormi on aina myös sopeutuva matriisinormi. Näin ollen yksi tapa todistaa jokin funktio matriisinormiksi on näyttää, että se on indusoitu jostakin vektorinormista.

Jokainen $n \times n$ -dimensioinen matriisi voidaan ajatella n^2 -dimensioisena vektorina. Esimerkiksi 2×2 -dimensioisen matriisin jokaisen sarakkeen voidaan ajatella vastaavan yhtä koordinaatiston akselia. Tällöin ensimmäinen sarake vastaa x-akselia ja toinen sarake y-akselia, jolloin matriisi vastaa 2-ulotteista koordinaatistoa. Sanotaankin, että matriisi virittää avaruuden. Tämän vuoksi voidaan osoittaa, että esimerkiksi luvussa 4.4 esitetyt l_1- ja l_2 -vektorinormit ovat "automaattisesti" myös matriisinormeja. Kuitenkaan kaikki vektorinormit eivät käyttäydy samalla tavalla. Esimerkiksi aiemmin kohdassa 4.4.4 esitetty l_∞ -normi ei ole matriisinormi, koska se ei täytä määritelmän 5.1 ehtoa (4).

Todistetaan seuraavaksi, että maksiminormi,

$$\|A\|_\infty = \max_{1\le i,j\le n} |a_{ij}|\,,$$

ei ole matriisinormi. Olkoon

$$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Tällöin tulomatriisiksi saadaan

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

Nyt määritelmän 5.1 ehdon (4) mukaan pitäisi olla

$$\|AB\|_\infty \le \|A\|_\infty \|B\|_\infty\,.$$

Kuitenkin laskemalla saadaan $\|AB\|_\infty = 2$ ja $\|A\|_\infty = \|B\|_\infty = 1$, jolloin $2 \le 1$, mikä on ristiriita. Näin ollen l_∞ -normi ei ole matriisinormi.

Seuraavassa osassa 5.3 esitellään erilaisia matriisinormeja ja osoittautuukin, että jos määritellään

$$|||A|||_n = n\,\|A\|_\infty\,, \quad \text{missä } A \in M_n,$$

niin saadaan aikaiseksi *kokonaisnormi*, joka täyttää määritelmän 5.1 ehdon (4).

Näytetään vielä, että kohdan 4.4.2 summanormi eli l_1 -normi on matriisinormi. Riittää osoittaa, että määritelmän 5.1 ehto (4) pätee. Olkoon $A, B \in M_n$. Nyt laskemalla saadaan

$$\begin{aligned}\|AB\|_1 &= \sum_{i,j=1}^{n} \left| \sum_{k=1}^{n} a_{ik}b_{kj} \right| \le \sum_{i,j,k=1}^{n} |a_{ik}b_{kj}| \le \sum_{i,j,k,m=1}^{n} |a_{ik}b_{mj}| \\ &\overset{(*)}{=} \left(\sum_{i,k=1}^{n} |a_{ik}| \right)\left(\sum_{j,m=1}^{n} |b_{mj}| \right) = \|A\|_1 \|B\|_1\end{aligned}$$

ja näin ollen summanormi on matriisinormi. Yhtäsuuruus (*) seuraa tulon ja summan laskusäännöistä: summan voi "halkaista" kahdeksi osasummaksi.

5.3. Erilaisia matriisinormeja

On olemassa useita eri matriisinormeja, jotka täyttävät määritelmän 5.1 ehdot. "Oikea" matriisinormi valitaankin sopivasti käyttötarkoitusten mukaan.

Seuraavaksi esitellään neljä yleisintä matriisinormia: kokonaisnormi, rivisummanormi, sarakesummanormi ja euklidinen matriisinormi. Jokaisen matriisinormin kohdalla kerrotaan, mihin vektorinormiin se sopeutuu ja mistä vektorinormista se on mahdollisesti indusoitu. Myöhemmin luvussa 6 käsitellään vielä spektraalinormia.

5.3.1. Kokonaisnormi. Kokonaisnormissa kerrotaan itseisarvoltaan matriisin suurin alkio matriisin kertaluvulla n. Kokonaisnormi on

$$|||A|||_n = n \max_{i,j} |a_{ij}| = n \left\| A \right\|_\infty .$$

Osoitetaan, että kokonaisnormi on matriisinormi. Koska kokonaisnormi $|||A|||_n$ on vektorinormi $\|A\|_\infty$ kerrottuna luvulla n, niin on selvää, että kokonaisnormi toteuttaa vektorinormille vaaditut ominaisuudet eli määritelmän 5.1 ehdot (1)-(3). Riittää siis osoittaa, että määritelmän 5.1 ehto (4) eli $\|AB\|_n \leq \|A\|_n \|B\|_n$ pätee. Laskemalla saadaan

$$\begin{aligned}
|||AB|||_n &= n \max_{1\leq i,j\leq n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right| \leq n \max_{1\leq i,j\leq n} \sum_{k=1}^{n} |a_{ik} b_{kj}| \\
&\leq n \max_{1\leq i,j\leq n} \sum_{k=1}^{n} \|A\|_\infty \|B\|_\infty = n \|A\|_\infty n \|B\|_\infty \\
&= |||A|||_n \; |||B|||_n,
\end{aligned}$$

ja näin ollen kokonaisnormi on matriisinormi. Kokonaisnormi sopeutuu kaikkiin luvussa 4 esitettyihin vektorinormeihin, joita olivat l_1-, l_2- ja l_∞ -normit. Näiden lisäksi se sopeutuu muihinkin l_p -normeihin. Tällöin siis $\|Ax\|_p \leq |||A|||_n \|x\|_p$, missä $p \in [1, \infty]$. Kokonaisnormi ei kuitenkaan ole minkään vektorinormin indusoima. Kuten jo aiemmin todettiin, niin maksiminormi eli l_∞ -normi ei ole matriisinormi, mutta pienillä muutoksilla siitä saadaan kokonaisnormi, joka on matriisinormi. Näin ollen kokonaisnormi ei voi olla minkään vektorinormin indusoima [**3**].

5.3.2. Rivisummanormi. Rivisummanormissa lasketaan jokaisen rivin alkioiden itseisarvojen summa ja valitaan niistä suurin. Rivisummanormi määritellään siis seuraavasti:

$$|||A|||_r = \max_{1\leq i\leq n} \Big(\sum_{j=1}^{n} |a_{ij}| \Big).$$

Maksiminormi eli l_∞ -normi indusoi rivisummanormin. Näin ollen rivisummanormi samalla sopeutuu l_∞ -normiin ja tällöin $\|Ax\|_\infty \leq |||A|||_r \|x\|_\infty$.

Osoitetaan seuraavaksi, että l_∞ -normi todella indusoi rivisummanormin eli

$$\max_{\|x\|=1} \|Ax\|_\infty = |||A|||_r.$$

Todistus. Todistetaan indusointi tarkastelemalla molempien puolien epäyhtälöitä. Osoitetaan ensin, että $\max\limits_{\|x\|=1} \|Ax\|_\infty \leq |||A|||_r$:

$$\begin{aligned}
\|Ax\|_\infty &= \max_{1\leq i\leq n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij} x_j| \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| \|x\|_\infty \\
&= |||A|||_r \|x\|_\infty .
\end{aligned}$$

Todistetaan seuraavaksi yhtälön toinen puoli eli, että

$$|||A|||_r \leq \max_{\|x\|=1} \|Ax\|_\infty .$$

Jos $A = 0$, niin väite on todistettu. Voidaan siis olettaa, että $A \neq 0$. Oletetaan, että matriisin A rivi k ei ole nollarivi ja valitaan k siten, että

$$\sum_{j=1}^{n} |a_{kj}| = |||A|||_r.$$

Määritellään lisäksi vektori $z \in \mathbb{C}^n$ seuraavasti:

$$z_i = \frac{\bar{a}_{ki}}{|a_{ki}|}, \qquad \text{jos } a_{ki} \neq 0,$$
$$z_i = 1, \qquad \text{jos } a_{ki} = 0.$$

Nyt tiedetään, että $\|z\|_\infty = 1$ ja $a_{kj}z_j = |a_{kj}|$ kaikilla $j = 1, 2, \ldots, n$. Näin ollen saadaan

$$\max_{\|x\|=1} \|Ax\|_\infty \geq \|Az\|_\infty = \max_{1\leq i\leq n} \left| \sum_{j=1}^{n} a_{ij} z_j \right| \geq \left| \sum_{j=1}^{n} a_{kj} z_j \right| = \sum_{j=1}^{n} |a_{kj}| = |||A|||_r.$$

Nyt siis saatiin, että

$$\max_{\|x\|=1} \|Ax\|_\infty \leq |||A|||_r \quad \text{ja} \quad \max_{\|x\|=1} \|Ax\|_\infty \geq |||A|||_r.$$

Tällöin on oltava, että

$$\max_{\|x\|=1} \|Ax\|_\infty = |||A|||_r,$$

mikä tarkoittaa, että maksiminormi indusoi rivisummanormin. □

Katsotaan seuraavaksi esimerkki rivisummanormin käytöstä.

Esimerkki 5.4. Olkoon $A \in M_3$ matriisi,

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

Lasketaan jokaisen rivin alkioiden itseisarvot yhteen:

$$|5| + |-2| + |1| = 8,$$
$$|4| + |6| + |-3| = 13,$$
$$|7| + |5| + |0| = 12,$$

ja valitsemalla näistä suurin saadaan $|||A|||_r = 13$.

5.3.3. Sarakesummanormi. Sarakesummanormi lasketaan samaan tapaan kuin rivisummanormi, mutta nyt sarakkeiden suhteen. Sarakesummanormi on siis

$$|||A|||_c = \max_{1\leq j\leq n}\Big(\sum_{i=1}^{n}|a_{ij}|\Big).$$

Sarakesummanormi on indusoitu l_1 -normista. [**3**, Luku 5] Tällöin sarakesummanormi myös sopeutuu samaan vektorinormiin, josta se on indusoitu, eli $\|Ax\|_1 \leq |||A|||_c\,\|x\|_1$. Sarakesummanormin indusointia l_1 -normista ei todisteta, koska se voidaan osoittaa lähes vastaavalla tavalla kuten edellä esitetyssä rivisummanormin tapauksessa. Katsotaankin seuraavaksi esimerkki sarakesummanormin käytöstä.

Esimerkki 5.5. Olkoon meillä käytössä sama matriisi kuin esimerkissä 5.4 Tällöin sarakesummanormi saadaan laskemalla jokaisen sarakkeen alkioiden itseisarvojen summat ja valitsemalla niistä suurin.

$$\begin{aligned}|||A|||_c &= \max(|5|+|4|+|7|\,,|-2|+|6|+|5|\,,|1|+|-3|+|0|)\\ &= \max(16,13,4)=16.\end{aligned}$$

5.3.4. Euklidinen matriisinormi. Euklidinen matriisinormi vastaa euklidista normia, jossa vektoreiden pituus määrätään Pythagoraan lauseen avulla. Euklidisessa matriisinormissa otetaan neliöjuuri jokaisen alkion neliön summasta. Euklidista matriisinormia kutsutaan myös *Frobeniuksen normiksi.* Euklidinen matriisinormi on siis

$$|||A|||_e = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}.$$

Euklidinen matriisinormi sopeutuu l_2 -normiin eli euklidiseen normiin. Näin ollen $\|Ax\|_2 \leq |||A|||_e\,\|x\|_2$. Euklidinen matriisinormi ei kuitenkaan ole minkään vektorinormin indusoima. [**9**, 1.3-18]

Euklidinen matriisinormi voidaan esittää myös l_2 -normin avulla, jolloin

$$|||A|||_e = \sqrt{\sum_{j=1}^{n}\|a_j\|_2^2}, \tag{5.1}$$

missä merkinnällä a_j tarkoitetaan matriisin $A \in M_n$ saraketta $j = 1, 2, \ldots n$. Esitystapa (5.1) on yhtäpitävä myös seuraavan esitystavan kanssa:

$$|||A|||_e = \sqrt{tr(A^*A)}, \tag{5.2}$$

missä A^* tarkoittaa matriisin A hermitointia ja $tr(A^*A)$ matriisin A^*A *jälkeä.* Hermitoinnilla tarkoitetaan matriisin A transponointia ja alkioiden a_{ij} konjugointia määritelmän 1.9 mukaisesti. Jäljellä puolestaan tarkoitetaan matriisin diagonaalialkioiden summaa, eli jos $A \in M_n$ on matriisi, niin tällöin

$$tr(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Nyt jos tarkastellaan matriisituloa $C = A^*A$, niin tällöin

$$c_{ij} = \sum_{k=1}^{n}\bar{a}_{ki}a_{kj}, \text{ kaikilla } i, j \in 1, 2, \ldots, n.$$

Nyt siis

$$tr(C) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} \bar{a}_{ki}a_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n} a_{ki}\bar{a}_{ki} = \langle A, A\rangle,$$

jolloin matriisitulon (A^*A) jäljelle saadaan siis

$$tr(A^*A) = \langle A, A\rangle.$$

Jäljelle on lisäksi voimassa seuraavat laskusäännöt:

(1) $tr(A + B) = tr(A) + tr(B)$,
(2) $tr(AB) = tr(BA)$.

Osoitetaan seuraavaksi, että euklidinen matriisinormi on matriisinormi eli se toteuttaa erityisesti määritelmän 5.1 ehdon (4):

$$\begin{aligned}
|||AB|||_e^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2 \overset{(*)}{\leq} \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|a_{ik}|^2\right)\left(\sum_{m=1}^{n}|b_{jm}|^2\right)\\
&= \left(\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2\right)\left(\sum_{j=1}^{n}\sum_{m=1}^{n}|b_{jm}|^2\right)\\
&= |||A|||_e^2|||B|||_e^2,
\end{aligned}$$

missä (*) seuraa Cauchy-Schwarzin epäyhtälöstä eli lauseesta 4.7. Näin ollen euklidinen matriisinormi on matriisinormi. Katsotaan vielä esimerkki euklidisen matriisinormin käytöstä.

Esimerkki 5.6. Käytetään edelleen esimerkin 5.4 matriisia ja saadaan

$$\begin{aligned}
|||A|||_e &= \sqrt{(5^2 + (-2)^2 + 1^2 + 4^2 + 6^2 + (-3)^2 + 7^2 + 5^2 + 0^2)}\\
&= \sqrt{(25 + 4 + 1 + 16 + 36 + 9 + 49 + 25 + 0)} = \sqrt{165}.
\end{aligned}$$

Listataan vielä esimerkin 5.4 matriisille saatujen matriisinormien tulokset:

$$\begin{aligned}
|||A|||_r &= 13,\\
|||A|||_c &= 16,\\
|||A|||_e &= \sqrt{165}.
\end{aligned}$$

Huomataan, että jokainen matriisinormi antaa eri tuloksen. Kuten jo aiemmin mainittiin, niin ”oikean” matriisinormin valinta tuleekin tehdä sen mukaan, mitä on tarkoitus tai halu mitata.

Lisäksi on tärkeää muistaa, että myös matriisinormeille pätee luvussa 4 esitetty ekvivalenttisuus: Aina on olemassa luvut $t > 0, s > 0$ siten, että kahdelle matriisinormille $|||\bullet|||_a$, $|||\bullet|||_b$ pätee

$$t|||A|||_a \leq |||A|||_b \leq s|||A|||_a, \text{ kaikilla } A \in M_n.$$

Koska $\dim F^{n\times n} = n^2 < \infty$, eli F-kertoiminen vektoriavaruus on äärellisulotteinen, niin kaikki sen matriisinormit ovat ekvivalentteja keskenään.

LUKU 6

Spektraalinormi ja spektraalisäde

Kuten jo luvussa 3 mainittiin, matriisin ominaisarvojen joukkoa kutsutaan matriisin spektriksi. Käytetään jatkossa matriisin $A \in M_n$ spektristä merkintää $\sigma(A)$. Tässä luvussa määriteltävissä spektraalinormissa ja spektraalisäteessä jo nimensä mukaisesti tarvitaan spektriä. Spektrin rajojen määrääminen onkin yksi tärkeistä käyttötarkoituksista matriisinormeille.

Tämän luvun alussa määritellään spektraalinormi ja osoitetaan, että se on euklidisen normin indusoima. Tämän jälkeen määritellään spektraalisäde ja osoitetaan, että se ei ole vektori- eikä matriisinormi. Lisäksi tähän lukuun on koottu merkittäviä tuloksia spektraalisäteeseen ja matriisinormeihin liittyen, kuten ehtoja matriisijonon suppenemiselle ja matriisin kääntyvyydelle.

6.1. Spektraalinormi

MÄÄRITELMÄ 6.1. Matriisin $A \in M_n$ spektraalinormi $|||\bullet|||_s$ on

$$|||A|||_s = \max\{\sqrt{\lambda}: \text{ missä } \lambda \text{ on tulon } A^*A \text{ ominaisarvo}\}$$

Huomataan, että normi on hyvin määritelty, koska, jos $A^*Ax = \lambda x$ ja $x \neq 0$, niin kertomalla yhtälön vasen puoli hermitoidulla vektorilla x^* saadaan

$$x^*A^*Ax = \|Ax\|_2^2 = \lambda \|x\|_2^2 .$$

Tällöin $\lambda \geq 0$ ja näin ollen $\sqrt{\lambda}$ on reaalinen ja ei-negatiivinen luku.

Seuraavassa lauseessa 6.2 osoitetaan, että euklidinen normi eli l_2 -normi indusoi spektraalinormin [**6**, Luku 10] eli

$$|||A|||_s = \max_{\|x\|=1} \|Ax\|_2 .$$

Tällöin spektraalinormi samalla sopeutuu l_2 -normiin, jolloin

$$\|Ax\|_2 \leq |||A|||_s \|x\|_2 .$$

LAUSE 6.2. *Olkoon $A \in M_n$. Tällöin*

$$|||A|||_s = \max_{\|x\|=1} \|Ax\|_2 .$$

TODISTUS. Oletetaan, että λ_A on matriisitulon $A^*A \in M_n$ suurin ominaisarvo. Tällöin tiedetään, että matriisitulolla (A^*A) on *ortonormaali ominaiskanta* $\{x_1 \cdots , x_n\}$ [**12**, Luku II, lause 2.3], koska

$$(A^*A)^* = A^*A^{**} = A^*A.$$

Koska siis tulomatriisi A^*A on symmetrinen, on sillä ortonormaali ominaiskanta. Tällöin tulomatriisin A^*A ominaisvektorit muodostavat *lineaarisesti riippumattoman joukon* ja tulomatriisilla on kertaluvut huomioiden n kpl ominaisarvoja. [**8**] Nyt siis

$x_1, x_2, \ldots, x_n$ ovat matriisitulon A^*A ortonormaaleja ominaisvektoreita ja $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ ominaisvektoreihin liittyviä ominaisarvoja. Tällöin on oltava $\lambda_A = \lambda_n$, koska λ_A oli matriisitulon A^*A suurin ominaisarvo. Olkoon nyt $x \in \mathbb{C}^n$, jolle $\|x\|_2 = 1$. Tällöin vektorille x saadaan seuraava esitys ortonormaalissa ominaiskannassa: [**11**, Lause 12.4]

$$x = \sum_{i=1}^{n} a_i x_i, \quad \text{josta saadaan} \quad A^*Ax = \sum_{i=1}^{n} a_i \lambda_i x_i.$$

Nyt euklidiselle normille saadaan arvio

$$\|Ax\|_2 = \sqrt{\langle Ax, Ax \rangle} = \sqrt{\langle x, A^*Ax \rangle} = \sqrt{\sum_{i=1}^{n} |a_i|^2 \lambda_i} \leq \sqrt{\lambda_A \sum_{i=1}^{n} |a_i|^2} = \sqrt{\lambda_A},$$

sillä $\lambda_i \leq \lambda_A$ kaikilla $i = 1, \ldots, n$. Tästä seuraa, että $|||A|||_2 \leq \sqrt{\lambda_A} = |||A|||_s$. Toisaalta vektorille x_n saadaan

$$\|Ax_n\|_2 = \sqrt{\langle x_n, A^*Ax_n \rangle} = \sqrt{\langle x_n, \lambda_A x_n \rangle} = \sqrt{\lambda_A} = |||A|||_s.$$

Näin ollen siis

$$|||A|||_s = \max_{\|x\|=1} \|Ax\|_2$$

ja väite on todistettu. □

6.2. Spektraalisäde

Määritelmä 6.3. Matriisin $A \in M_n$ spektraalisäde $\rho(A)$ on

$$\rho(A) = \max\{|\lambda| : \text{ missä } \lambda \text{ on matriisin } A \text{ ominaisarvo}\}.$$

Koska $\rho(A)$ on määritelty itseisarvoiltaan suurimmaksi matriisin A ominaisarvoksi λ, niin on selvää, että kaikille ominaisarvoille pätee $|\lambda| \leq \rho(A)$. Tällöin on myös ainakin yksi ominaisarvo λ, jolle $|\lambda| = \rho(A)$.

Lause 6.4. *Jos $|||\bullet|||$ on mikä tahansa matriisinormi ja $A \in M_n$, niin tällöin $\rho(A) \leq |||A|||$.*

Todistus. Kuten jo aiemmin todettiin, niin on selvää, että jollakin matriisin A ominaisarvolla λ on $|\lambda| = \rho(A)$. Valitaan tämä ominaisarvo λ ja olkoon $x \neq 0$ ominaisarvoa λ vastaava ominaisvektori. Ominaisarvoyhtälö on tällöin $Ax = \lambda x$. Olkoon $X \in M_n$ matriisi, jonka jokaisena sarakevektorina on ominaisvektori x eli $X = [x\ x \cdots x]$. Tällöin saadaan ominaisarvoyhtälö $AX = \lambda X$ eli

$$\begin{bmatrix} Ax & Ax & \cdots & Ax \end{bmatrix} = \begin{bmatrix} \lambda x & \lambda x & \cdots & \lambda x \end{bmatrix}.$$

Jos nyt $|||\bullet|||$ on mikä tahansa matriisinormi, niin matriisinormin ominaisuuksien perusteella

$$|\lambda|\, |||X||| = |||\lambda X||| = |||AX||| \leq |||A|||\, |||X|||,$$

jolloin $|\lambda| = \rho(A) \leq |||A|||$ ja väite on todistettu. □

Spektraalisäde -funktio ei ole vektori- eikä matriisinormi. Esimerkiksi on olemassa matriisi A, jolle $\rho(A) = 0$, mutta $A \neq 0$. Näin ollen sekä vektori- että matriisinormille määritelty ehto (1a) ei toteudu. Esimerkkinä tällaisesta matriisista voidaan tarkastella matriisia

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

Nyt luvun 3 ominaisarvoyhtälön perusteella saadaan

$$\det(A - \lambda I) = \begin{vmatrix} 0-\lambda & 0 \\ 1 & 0-\lambda \end{vmatrix} = -\lambda \cdot (-\lambda) - 1 \cdot 0 = \lambda^2.$$

Näin ollen ominaisarvot ovat $\lambda_1 = 0$ ja $\lambda_2 = 0$. Tällöin $\rho(A) = 0$, mutta $A \neq 0$.

Tarkoituksena on päästä tarkastelemaan lausetta 6.10 ja seurausta 6.13, joissa käsitellään spektraalisäteen ja matriisitulon raja-arvojen välisiä yhteyksiä. Myöskin halutaan osoittaa seuraus 6.12, joka antaa tärkeän tuloksen käänteismatriisista. Näitä tuloksia varten tarvitaan muutamia tärkeitä aputuloksia eli *lemmoja* ja määritelmiä raja-arvoihin ja suppenemiseen liittyen. Osoitetaan seuraavaksi lemma 6.5, joka sanoo, että jokin matriisinormi saadaan aina hyvin lähelle spektraalisädettä. Tämän jälkeen osoitetaan similaarisuuteen liittyvä lause 6.6, jonka jälkeen jatketaan määrittelemällä raja-arvo ja suppeneminen.

LEMMA 6.5. *Olkoon $A \in M_n$ matriisi ja $\epsilon > 0$. On olemassa matriisinormi $|||\bullet|||$ siten, että*

$$\rho(A) \leq |||A||| \leq \rho(A) + \epsilon.$$

TODISTUS. Luvun 3 lauseen 3.3 mukaan on olemassa unitaarinen matriisi U ja yläkolmiomatriisi Δ siten, että $A = U^*\Delta U$. Merkitään $\Delta = (d_{ij})$, missä $i, j = 1, 2, \ldots, n$ ja asetetaan $D_t = diag(t, t, t^2, t^3, \ldots, t^n)$. Nyt laskemalla saadaan

$$D_t \Delta D_t^{-1} = \begin{bmatrix} \lambda_1 & t^{-1}d_{12} & t^{-2}d_{13} & t^{-3}d_{14} & \cdots & t^{-n+1}d_{1n} \\ 0 & \lambda_2 & t^{-1}d_{23} & t^{-2}d_{24} & \cdots & t^{-n+2}d_{2n} \\ 0 & 0 & \lambda_3 & t^{-1}d_{34} & \cdots & t^{-n+3}d_{3n} \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 & \ddots & t^{-1}d_{n-1,n} \\ 0 & 0 & 0 & 0 & 0 & \lambda_n \end{bmatrix}.$$

Näin ollen, kun $t > 0$ on tarpeeksi iso, voidaan olla varmoja, että ei-diagonaalilla olevien alkioiden itseisarvojen summa on vähemmän kuin ϵ, sillä kun valitaan tarpeeksi iso t, niin kertoimet t^{-k}, missä $k \in \mathbb{N}$, ovat pieniä. Erityisesti tiedetään, että

$$\left|\left|\left| D_t \Delta D_t^{-1} \right|\right|\right|_c \leq \rho(A) + \epsilon,$$

kun t on tarpeeksi suuri. Määritellään nyt matriisinormi $|||\bullet|||$ seuraavasti mille tahansa $B \in M_n$

$$|||B||| = \left|\left|\left| D_t U^* B U D_t^{-1} \right|\right|\right|_c = \left|\left|\left| (U D_t^{-1})^{-1} B (U D_t^{-1}) \right|\right|\right|_c$$

ja nyt valitaan tarpeeksi iso t. Lauseessa 6.6 todistetaan, että näin todella saadaan aikaiseksi matriisinormi. Näin ollaan saatu konstruoitua matriisinormi siten, että $|||A||| \leq \rho(A) + \epsilon$. Nyt lause 6.4 antaa yhtälön toisen puolen

$$\rho(A) \leq |||A||| \leq \rho(A) + \epsilon$$

ja väite on todistettu. □

Seuraava lause 6.6 antaa hyödyllisen tuloksen, kuinka matriisinormi voidaan muuntaa toiseksi similaarisuuden avulla. Lisäksi se antaa perustelun sille, että edellä esitetyssä lemmassa 6.5 konstruoitu matriisinormi on todella matriisinormi.

LAUSE 6.6. *Olkoon* $|||\bullet|||$ *matriisinormi avaruudessa* M_n *ja* $S \in M_n$ *ei-singulaarinen matriisi. Tällöin* $|||A|||_M = |||S^{-1}AS|||$ *on matriisinormi kaikilla* $A \in M_n$.

TODISTUS. Tulee osoittaa, että $|||A|||_M$ on matriisinormi eli että se toteuttaa määritelmän 5.1 ehdot. Ehtojen (1)-(3) todistukset ovat suoraviivaisia. Osoitetaan, että määritelmän 5.1 ehto (4) pätee:

$$\begin{aligned}|||AB|||_M &= \left|\left|\left|S^{-1}ABS\right|\right|\right| = \left|\left|\left|(S^{-1}AS)((S^{-1}BS)\right|\right|\right| \leq \left|\left|\left|(S^{-1}AS)\right|\right|\right|\left|\left|\left|(S^{-1}BS)\right|\right|\right| \\ &= |||A|||_M |||B|||_M.\end{aligned}$$

□

Tarkastellaan nyt seuraavaksi *raja-arvoja* ja *suppenemista*. Yleisesti voidaan sanoa, että lukujonolla $(a_n)_{n\in\mathbb{N}}$ on raja-arvo $a \in \mathbb{R}$, jos luvun n lähestyessä ääretöntä lukujonon termit lähestyvät lukua a. Tällöin sanotaan, että lukujono $(a_n)_{n\in\mathbb{N}}$ suppenee. Tarkastellaan seuraavaksi mitä tarkoitetetaan sarjan ja matriisijonon suppenemisella.

MÄÄRITELMÄ 6.7. Olkoon $x_i \in \mathbb{R}$. Tällöin summa eli sarja

$$\sum_{i=1}^{\infty} x_i = x_1 + x_2 + \cdots + x_n + x_{n+1} + \ldots$$

suppenee, jos sen osasummien jono $(S_n)_{n\in\mathbb{N}}$ suppenee. Tällöin on olemassa $S \in \mathbb{R}$ eli *raja-arvo*, jota merkitään $\lim_{n\to\infty} S_n = S$. Tällöin S on sarjan summa ja merkitään

$$\sum_{i=1}^{\infty} x_i = x_1 + x_2 + \cdots + x_n + x_{n+1} + \cdots = S.$$

Eräs suppenemiseen liittyvä tärkeä lause sanoo, että suppenevan sarjan termien raja-arvo on nolla. Jos esimerkiksi määritelmän 6.7 sarja suppenee eli *konvergoi*, niin

$$\lim_{i\to\infty} x_i = 0.$$

Tämä ei kuitenkaan välttämättä päde toisinpäin. Jos sarja ei suppene, niin silloin sanotaan, että sarja *hajaantuu* eli *divergoi*.

MÄÄRITELMÄ 6.8. Olkoon $A^{(k)} \in M_n$ jono matriiseja. Sanotaan, että matriisijono $(A^{(k)})_k$ suppenee kohti matriisia $A \in M_n$, jos $a_{ij}^{(k)} \to a_{ij}$ kaikilla $i, j \in \mathbb{N}$. Tällöin merkitään

$$\lim_{k\to\infty} A^{(k)} = A \quad \text{tai} \quad A^{(k)} \to A.$$

Matriisin tapauksessa suppeneminen määritellään siis alkioittain eli $a_{ij}^{(k)} \to a_{ij}$.

Seuraavissa tuloksissa tullaan tarkastelemaan potenssijonon $(A^k)_k$ suppenemista. Lemmassa 6.9 osoitetaan, että jos jonkin matriisinormin arvo on pienempi kuin yksi, niin tällöin matriisijono suppenee kohti nollaa.

LEMMA 6.9. *Olkoon* $A \in M_n$ *matriisi. Jos on olemassa matriisinormi* $|||\bullet|||$ *siten, että* $|||A||| < 1$*, niin tällöin* $\lim_{k\to\infty} A^k = 0$ *eli kaikki matriisin* A^k *alkiot menevät kohti nollaa, kun indeksi* k *lähestyy ääretöntä.*

TODISTUS. Matriisinormille pätee $\left|\left|\left|A^k\right|\right|\right| \leq |||A|||^k$, joten jos $|||A||| < 1$, niin $|||A|||^k \to 0$, kun $k \to \infty$. Tällöin siis $\lim\limits_{k\to\infty} \left|\left|\left|A^k\right|\right|\right| \leq 0$ ja sen vuoksi täytyy myös olla $A^k \to 0$ matriisinormin suhteen. Lisäksi, koska kaikki vektorinormit ovat ekvivalentteja n^2 - ulotteisessa avaruudessa M_n, saadaan erityisesti, että $A^k \to 0$ myös vektorinormin $\|\bullet\|_\infty$ suhteen.

Näin ollen $\lim\limits_{k\to\infty} A^k = 0$ ja väite on todistettu. □

Suppeneville matriisijonoille on olemassa monia tärkeitä sovelluksia. Suppenemista hyödynnetetään mm. *iteratiivisten* ohjelmien analysoimisessa. Näin ollen on tärkeää pystyä määräämään ehtoja matriisijonon suppenemiselle. Seuraavassa lauseessa 6.10 osoitetaan, että matriisitulo suppenee, jos ja vain jos spektraalisäde on pienempää kuin yksi.

LAUSE 6.10. *Olkoon $A \in M_n$. Tällöin $\lim\limits_{k\to\infty} A^k = 0$, jos ja vain jos $\rho(A) < 1$.*

TODISTUS. Todistetaan ensin, että jos $\lim_{k\to\infty} A^k = 0$, niin $\rho(A) < 1$. Olkoon λ matriisin A ominaisarvo ja $x \neq 0$ sitä vastaava ominaisvektori. Tällöin $A^k x = \lambda^k x$. Koska $A^k \to 0$, niin

$$A^k x = \lambda^k x \to 0, \text{ mikä pätee vain ja ainoastaan, jos } |\lambda| < 1.$$

Epäyhtälön $|\lambda| < 1$ tulee toteutua kaikilla matriisin A ominaisarvoilla. Näin ollen saadaan, että $\rho(A) < 1$ ja väitteen toinen puoli on todistettu.

Todistetaan seuraavaksi, että jos $\rho(A) < 1$, niin $\lim_{k\to\infty} A^k = 0$. Lemma 6.5 sanoo, että on olemassa matriisinormi $|||\bullet|||$ siten, että $\rho(A) \leq |||A||| \leq \rho(A) + \epsilon$. Nyt siis halutaan, että $|||A||| < 1$ ja tiedetään, että

$$|||A||| < \rho(A) + \epsilon.$$

Tällöin tulee valita ϵ siten, että

$$\rho(A) + \epsilon < 1$$

eli $\epsilon < 1 - \rho(A)$. Nyt jos $\rho(A) < 1$, niin $|||A||| < 1$. Tällöin lemman 6.8 mukaan $A^k \to 0$, kun $k \to \infty$. Nyt yhdistämällä molempien suuntien todistukset saadaan, että $\lim\limits_{k\to\infty} A^k = 0$ jos ja vain jos $\rho(A) < 1$ ja väite on todistettu. □

Seurauksen 6.12 todistusta varten tarvitaan lemmaa 6.11, missä osoitetaan äärettömän matriisijonon suppeneminen [**1**, Lemma 3.2.3]. Sovitaan, että jatkossa käytetään merkintää $A^0 = I$.

LEMMA 6.11. *Olkoon $|||\bullet|||$ matriisinormi joukossa M_n. Jos $|||A||| = \lambda < 1$, niin tällöin matriisijono $(A_n)_n \subset M_n$ suppenee, missä $A_n = I + A + A^2 + \cdots + A^n$.*

TODISTUS. Näytetään, että $(A_n)_n$ on Cauchy-jono: kaikille $n, m \in \mathbb{N}$, joille $n > m$ saadaan, että

$$A_n - A_m = \sum_{k=m+1}^{n} A^k.$$

Näin ollen saadaan

$$\begin{aligned}
|||A_n - A_m||| &= \left|\left|\left| \sum_{k=m+1}^{n} A^k \right|\right|\right| \\
&\le \sum_{k=m+1}^{n} |||A^k||| \\
&\le \sum_{k=m+1}^{n} |||A|||^k \\
&= \sum_{k=m+1}^{n} \lambda^k = \lambda^{m+1} \sum_{j=0}^{n-m-1} \lambda^j \\
&\overset{(*)}{\le} \lambda^{m+1}(1-\lambda)^{-1} \to 0, \text{ kun } m \to \infty,
\end{aligned}$$

missä epäyhtälö (*) on voimassa, koska $0 \le \lambda < 1$. Nyt koska $(A_n)_n$ on Cauchy-jono, niin alkioiden muodostamat jonot ovat myös Cauchy-jonoja ja tällöin $(A_n)_n$ suppenee. [**4**, Luku 3] Tällöin ääretön sarja $\displaystyle\sum_{k=0}^{\infty} A_k$ suppenee kohti jotakin matriisia $C \in M_n$. $\square$

Nyt voidaan todistaa seuraus 6.12, jossa saadaan ehtoja matriisin kääntyvyydelle ja määrättyä käänteismatriisi potenssisarjan avulla. Tätä seurausta tullaan jatkossa hyödyntämään myös seuraavassa luvussa, jossa käsitellään matriisien virhearviointia.

Seuraus 6.12. *Matriisi $A \in M_n$ on kääntyvä, jos on olemassa matriisinormi $|||\bullet|||$ siten, että $|||I - A||| < 1$. Tällöin*

$$A^{-1} = \sum_{k=0}^{\infty} (I - A)^k.$$

Todistus. Merkitään käsiteltäviä sarjoja

$$S = \sum_{k=0}^{\infty} (I - A)^k \quad \text{ja} \quad S_N = \sum_{k=0}^{N} (I - A)^k.$$

Koska tiedetään, että

$$\sum_{k=1}^{\infty} \left|\left|\left|(I - A)^k\right|\right|\right| \le \sum_{k=1}^{\infty} |||I - A|||^k$$

ja $|||I - A||| < 1$, niin sarja $\sum_{k=1}^{\infty} \left|\left|\left|(I - A)^k\right|\right|\right|$ suppenee, ja siten lemman 6.11 nojalla myös sarja S suppenee jotakin matriisia $C \in M_n$ kohti. Nyt siis $S = C$. Tällöin määritelmän 6.7 nojalla sarjan osasumma S_N suppenee eli $S_N \to C$, kun $N \to \infty$. Riittää siis osoittaa, että $C = A^{-1}$.

Tarkastellaan mitä tapahtuu sarjan osasummalle S_N. Kerrotaan sarjaa S_N vasemmalta puolelta matriisilla A:

$$\begin{aligned} AS_N &= A\sum_{k=0}^{N}(I-A)^k = [I-(I-A)]\sum_{k=0}^{N}(I-A)^k \\ &= I\sum_{k=0}^{N}(I-A)^k - (I-A)\sum_{k=0}^{N}(I-A)^k \\ &= \sum_{k=0}^{N}(I-A)^k - \sum_{k=1}^{N+1}(I-A)^k \\ &= (I-A)^0 + \dots (I+A)^N - [(I-A)^1 + \dots (I-A)^N + (I-A)^{N+1})] \\ &= I - (I-A)^{N+1} \to I, \text{ kun } N \to \infty, \text{ lemman 6.9 nojalla.} \end{aligned}$$

Vastaavalla tavalla voidaan osoittaa, että $S_N A \to I$, kun sarjaa S_N kerrotaan oikealta puolelta matriisilla A. Nyt matriisin C on siis oltava A^{-1}. Tällöin summalle S saadaan

$$S = \sum_{k=0}^{\infty}(I-A)^k = C = A^{-1}$$

ja väite on todistettu. □

Edelllä esitettyjen lemmojen ja lauseen 6.10 avulla pystytään osoittamaan merkittävä seuraus 6.13, jossa spektraalisäde voidaan ilmaista myös niin kutsussa ”Gelfand'sin muodossa”. [**2**, Theorem 12]

Seuraus 6.13. *Olkoon $|||\bullet|||$ matriisinormi avaruudessa M_n. Tällöin*

$$\rho(A) = \lim_{k\to\infty} |||A^k|||^{\frac{1}{k}} \text{ kaikilla } A \in M_n.$$

Todistus. Koska $\rho(A)^k \overset{(*)}{=} \rho(A^k) \leq |||A^k|||$, niin $\rho(A)^{k\cdot\frac{1}{k}} = \rho(A) \leq |||A^k|||^{\frac{1}{k}}$. Yhtäsuuruus (*) perustuu Schurin lauseeseen 3.3 ja siihen, että yläkolmiomatriisien ominaisarvot ovat samat kuin sen diagonaalialkiot. Olkoon nyt $\epsilon > 0$ ja matriisi

$$\hat{A} = \frac{A}{\rho(A)+\epsilon}.$$

Tällöin saadaan, että

$$\rho(\hat{A}) = \frac{\rho(A)}{\rho(A)+\epsilon} < 1$$

Nyt lauseen 6.10 nojalla tiedetään, että $\lim_{k\to\infty}\hat{A}^k = 0$. Tämä taas tarkoittaa, että on olemassa luonnollinen luku $N \in \mathbb{N}$ siten, että kaikilla $k \geq N$ on

$$|||\hat{A}^k||| < 1.$$

Toisaalta tiedetään myös, että kaikilla $k \geq N$ on

$$|||A^k||| < (\rho(A)+\epsilon)^k,$$

josta edelleen

$$|||A^k|||^{\frac{1}{k}} < \rho(A)+\epsilon.$$

Lisäksi lauseen 6.4 perusteella tiedetään, että $\rho(A) \leq |||A|||$, jolloin $\rho(A) \leq \left|\left|\left|A^k\right|\right|\right|^{\frac{1}{k}}$. Tällöin kokoamalla nämä tiedot yhteen saadaan: Olkoon $\epsilon > 0$. Tällöin on olemassa luku $N \in \mathbb{N}$ siten, että kaikille $k \geq N$ pätee

$$\rho(A) - \epsilon < \left|\left|\left|A^k\right|\right|\right|^{\frac{1}{k}} < \rho(A) + \epsilon,$$

mikä tarkoittaa, että

$$\lim_{k\to\infty} \left|\left|\left|A^k\right|\right|\right|^{\frac{1}{k}} = \rho(A).$$

□

LUKU 7

Matriisien virhearvioinneista

Vektori- ja matriisinormien sovelluksena voidaan arvioida, minkälaisia ja miten suuria virheitä on tehty laskettaessa käänteismatriisia ja lineaarista yhtälöryhmää.

Olkoon $A \in M_n$ ei-singulaarinen matriisi. Aiemmin opitun perusteella voidaan ajatella, että matriisin A käänteismatriisi voidaan määrätä täsmällisesti. Näin ei kuitenkaan aina ole, koska jos käänteismatriisiin määräämiseksi vaadittavat laskut ovat esimerkiksi toteutettu tietokoneella äärellisen pitkälle, tällöin tulee väistämättä eteen pyöristysvirheitä ja ennenaikaisia laskujen keskeytyksiä.

Tässä luvussa haetaan vastauksia siihen, kuinka laskennassa ja datassa eli matriisin A tai vektorin a alkioissa esiintyvät virheet vaikuttavat etsittävän käänteismatriisiin virheeseen. Datan kertoimet on voitu saada esimerkiksi mittaustulosten perusteella.

On saatu selville, että monien tuttujen algoritmien laskemisesta saadut pyöristysvirheet voidaan mallintaa samaan tapaan kuin datassa esiintyvät virheet. Seuraavaksi johdatellaankin kohti suuretta, jota käytetään virheen laskemisessa.

7.1. Virheen arviointia

Olkoon $A \in M_n$ ei-singulaarinen matriisi. Halutaan määrätä matriisin A käänteismatriisi A^{-1}. Koska A^{-1} on ”täsmällinen” käänteismatriisi, tulee lisätä matriisiin A jotain pientä, mistä saadaan virheen suuruus selville. Tällöin täytyy määrätä matriisi $(A+E)^{-1}$, missä $E \in M_n$ on tarpeeksi pieni, jolloin $(A+E)$ on kääntyvä. Oletetaan, että virhe $E \in M_n$ on niin pieni, että $\rho(A^{-1}E) < 1$. Nyt seurauksen 6.12 nojalla tiedetään, että $(I + A^{-1}E)$ on kääntyvä, jolloin myös $(A+E)$ on kääntyvä. Tämä siksi, että $(I + A^{-1}E) = (A^{-1}A + A^{-1}E) = A^{-1}(A + E)$. [**3**, Luku 5] Näin ollen virheen suuruus voidaan laskea vähentämällä ”täsmällisestä” A^{-1} käänteismatriisista ”virheen sisältävä” käänteismatriisi $(A + E)^{-1}$ ja saadaan seuraava esitys:

$$A^{-1} - (A+E)^{-1} = A^{-1} - [A(I + A^{-1}E)]^{-1} = A^{-1} - (I + A^{-1}E)^{-1}A^{-1}.$$

Nyt lausekkeen $(I + A^{-1}E)$ tulo $A^{-1}E$ voidaan kirjoittaa seurauksen 6.12 nojalla potenssisarjojen avulla seuraavasti:

$$\begin{aligned} A^{-1} - (A+E)^{-1} &= A^{-1} - \sum_{k=0}^{\infty}[I - (I + A^{-1}E)]^k A^{-1} \\ &= A^{-1} - \sum_{k=0}^{\infty}(-1)^k (A^{-1}E)^k A^{-1} \\ &= \sum_{k=1}^{\infty}(-1)^{k+1}(A^{-1}E)^k A^{-1}, \end{aligned}$$

joka on täsmällinen kaava virheelle.

MÄÄRITELMÄ 7.1. Olkoon $A \in M_n$ ei-singulaarinen matriisi. Jos $\rho(A^{-1}E) < 1$, niin virheen kaava on

$$A^{-1} - (A+E)^{-1} = \sum_{k=1}^{\infty} (-1)^{k+1} (A^{-1}E)^k A^{-1}. \tag{7.1}$$

Määritelmän 7.1 pohjalta voidaan määrittää yläraja suhteelliselle virheelle: oletetaan, että $|||\bullet|||$ on annettu matriisinormi ja $|||A^{-1}E||| < 1$. Erityisesti $\rho(A^{-1}E) < 1$ ja yhtälö (7.1) pätee. Nyt voidaan arvioida lauseketta ylöspäin seuraavasti:

$$\begin{aligned} \left|\left|\left| A^{-1} - (A+E)^{-1} \right|\right|\right| &= \left|\left|\left| \sum_{k=1}^{\infty} (-1)^{k+1} (A^{-1}E)^k A^{-1} \right|\right|\right| \\ &\leq \sum_{k=1}^{\infty} \left|\left|\left| A^{-1}E \right|\right|\right|^k \left|\left|\left| A^{-1} \right|\right|\right| \\ &= \frac{|||A^{-1}E|||}{1 - |||A^{-1}E|||} \left|\left|\left| A^{-1} \right|\right|\right|. \end{aligned}$$

Tällöin saadaan yläraja suhteelliselle virheelle jakamalla molemmat puolet termillä $|||A^{-1}|||$

$$\frac{|||A^{-1} - (A+E)^{-1}|||}{|||A^{-1}|||} = \frac{|||A^{-1}E|||}{1 - |||A^{-1}E|||},$$

jos $|||A^{-1}E||| < 1$. Jos lisäksi oletetaan, että E on riittävän pieni

$$|||E||| < \frac{1}{|||A^{-1}|||}, \text{ niin tällöin } \rho(A^{-1}E) < \left|\left|\left| A^{-1}E \right|\right|\right| \leq \left|\left|\left| A^{-1} \right|\right|\right| \, |||E|||$$

ja saadaan lopullinen arvio seuraavasti:

$$\frac{|||A^{-1} - (A+E)^{-1}|||}{|||A^{-1}|||} \leq \frac{|||A^{-1}||| \; |||E|||}{1 - |||A^{-1}||| \; |||E|||} = \frac{|||A^{-1}||| |||A||| \frac{|||E|||}{|||A|||}}{1 - |||A^{-1}||| |||A||| \frac{|||E|||}{|||A|||}}. \tag{7.2}$$

7.2. Matriisin ehtoluku

Seuraavaa suuretta kutsutaan *matriisinormin* $|||\bullet|||$ *ehtoluvuksi*

$$\kappa(A) = \begin{cases} \left|\left|\left| A^{-1} \right|\right|\right| |||A|||, & \text{jos } A \text{ on ei-singulaarinen} \\ \infty, & \text{jos } A \text{ on singulaarinen.} \end{cases} \tag{7.3}$$

Huomataan, että $\kappa(A) \geq 1$, mille tahansa matriisinormille. Tämä nähdään suoraan laskemalla matriisinormien ominaisuuksien avulla, kuten kohdassa 5.1:

$$\kappa(A) = \left|\left|\left| A^{-1} \right|\right|\right| |||A||| \geq \left|\left|\left| A^{-1}A \right|\right|\right| = |||I||| \geq 1$$

ja näin ollen, $\kappa(A) \geq 1$.

Nyt sijoittamalla saatu ehtoluku (7.3) yhtälöön (7.2) saadaan:

$$\frac{|||A^{-1} - (A+E)^{-1}|||}{|||A^{-1}|||} \leq \frac{\kappa(A)}{1 - \kappa(A) \frac{|||E|||}{|||A|||}} \frac{|||E|||}{|||A|||}, \tag{7.4}$$

jos $|||E||| \, |||A^{-1}||| < 1$. Yhtälö (7.4) antaa rajat käänteismatriisin datasta löytyvälle suhteelliselle virheelle. Nähdään, että mitä suurempi ehtoluku on, sitä enemmän matriisilla on taipuvaisuutta virheeseen. On tärkeää huomata, että ehtoluku riippuu myös käytetystä matriisinormista. Seuraavassa esimerkissä 7.2 lasketaan käänteismatriisin suhteellinen virhe sarakesummanormin avulla.

Esimerkki 7.2. Oletetaan, että matriisissa

$$A = \begin{bmatrix} 1,00 & 0,50 & 0,10 \\ 0,50 & 1,00 & 0,25 \\ 0,10 & 0,25 & 1,00 \end{bmatrix}$$

lukujen tarkkuus on $\epsilon = 0{,}01$. Lasketaan nyt A^{-1} ja arvioidaan saadun käänteismatriisin tarkkuutta sarakesummanormin avulla. Alkuperäisen matriisin A käänteismatriisi on

$$A^{-1} = \frac{1}{281} \begin{bmatrix} 375 & -190 & 10 \\ -190 & 396 & -80 \\ 10 & -80 & 300 \end{bmatrix}.$$

Tällöin sarakesummanormeista saadaan

$$|||A|||_c = 1,75,$$
$$|||E|||_c = 0,03,$$
$$\left|\left|\left|A^{-1}\right|\right|\right|_c = \frac{666}{281},$$

ja mistä erityisesti huomataan, että $|||E||| \, |||A^{-1}||| < 1$. Nyt ehtoluku on

$$\kappa(A) = |||A|||_c \left|\left|\left|A^{-1}\right|\right|\right|_c = 1,75 \cdot \frac{666}{281} = \frac{2331}{562},$$

ja sijoittamalla ehtoluku kaavaan (7.4) saadaan matriisin A^{-1} suhteelliselle virheelle arvio

$$\frac{\frac{2331}{562}}{1 - \frac{2331}{562}\frac{0{,}03}{1{,}75}} \frac{0,03}{1,75} = \frac{999}{13051} \approx 0,0765.$$

LUKU 8

Suunnitelma lukion pitkän matematiikan syventäväksi kurssiksi - Matriisit tutuiksi

Lukion pitkän matematiikan opetussuunnitelmassa on yhtenä pakollisena kurssina vektorit (MAA5). Syventävän kurssin ”Matriisit tutuiksi” on tarkoituksena olla jatkokurssi tälle pakolliselle kurssille. Kurssin ideana on yleistää vektorit matriiseiksi ja tutustua niiden perusominaisuuksiin. Lisäksi kurssilla ratkaistaan yhtälöryhmiä matriiseja hyödyntäen. Kurssin lopussa opiskelijat saavat todistaa ryhmissä yliopistotasoisia lineaarialgebran lauseita. Kurssista ei järjestetä koetta, vaan kurssi arvostellaan opiskelijan tekemän itsearvioinnin ja tuntiaktiivisuuden pohjalta joko hyväksytty tai hylätty -arvolauseella.

Oletetaan, että tälle syventävälle kurssille on varattu viikossa 3 puolentoista tunnin tapaamista. Yksi jakso on lukiossa noin 7 viikon mittainen. Alla on tehty karkeahko viikkokohtainen kurssisuunnitelma.

8.1. Kurssin viikkoaikataulu

Viikko	Aihe
1	Kurssin esittely
1	Vektoreiden kertausta
1	Mikä on matriisi ja mihin sitä tarvitaan?
2	Erilaisten matriisien tarkastelua
2	Matriisin perusoperaatiot
2	Matriisin perusoperaatiot
3	Matriisin determinantti
3	Matriisin determinantti
3	Matriisin kääntyvyys
4	Gaussin & Jordanin menetelmä
4	Gaussin & Jordanin menetelmä
4	Sisätulo
5	Vektorinormin ja matriisinormin esittely
5	Laskemista matriisinormeilla
5	Kertausta kurssin asioista ja todistusaiheiden esittely
6	Pienryhmässä todistusaiheen valinta
6	Pienryhmätyöskentelyä
6	Pienryhmätyöskentelyä
7	Todistuksien esittelyä
7	Todistuksien esittelyä
7	Todistuksien esittelyä ja itsearviointi

Ensimmäisen viikon aikana on tarkoitus saada opiskelijoille näkemys siitä, mikä on kurssin suunnitelma ja mitä oppimistavoitteita kurssilta halutaan saavuttaa. Ensimmäisen viikon jälkeen on käyty myös läpi kertausta vektoreista ja esitelty matriisi.

Toisella viikolla paneudutaan jo tunnistamaan, minkä tyyppisiä matriisit ovat ja minkälaista dataa niihin voidaan tallettaa. Tällä viikolla käsitellään muun muassa tutkielman alussa olevissa luvuissa esiintyneet termit ylä- ja alakolmiomatriisi, identtinen matriisi, diagonaalimatriisi ja nollamatriisi. Lisäksi tämän viikon jälkeen opiskelijat osaavat matriisien perusoperaatiot: yhteen- ja vähennyslaskun, transponoinnin ja pienien matriisien kertolaskun.

Kolmas viikko täyttyy determinanttiin liittyvistä asioista. Ensin tarkastellaan 2×2 -dimensioisia matriiseja, jonka jälkeen mietitään, miten lasketaan 3×3 -dimensioisen matriisin determinantti. Tällä kurssilla ei mennä enää suurempien matriisien determinanttien tarkasteluun. Viikon viimeisellä tunnilla otetaan aiheeksi kääntyvä matriisi, jonka laskemista harjoitellaan seuraavalla viikolla Gaussin & Jordanin menetelmällä.

Neljännellä viikolla paneudutaan siihen, kuinka käänteismatriisi saadaan määrättyä. Tähän avuksi tarvitaan Gaussin & Jordanin menetelmää. Tarkoituksena on antaa opiskelijoille edellytykset ratkaista käänteismatriisit 3×3 -dimensioisille matriiseille. Loppuviikolla käsitellään sisätulo ja mietitään, mitä kaikkea sillä saadaan vektoreista selville.

Viides viikko täyttyy normeista. Ensin tutustutaan vektorinormiin perusteellisesti ja lasketaan erilaisilla normeilla vektorien pituuksia. Tämän jälkeen esitellään lyhyesti muutama matriisinormi ja lasketaan esimerkkejä. Loppuviikosta esitellään todistusaiheet, joista opiskelijat saavat ryhmissä valita mieleisensä. Opiskelijat saavat myös itse ehdottaa todistusaiheita.

Kahdella viimeisellä viikolla opiskelijat työstävät todistusta yhdessä ja etsivät siihen sopivaa lisämateriaalia. Viimeisellä viikolla opiskelijat esittelevät omat tuotoksensa muille kurssilaisille. Seuraavassa on listattuna muutamia esimerkkejä todistusaiheista. Todistusten lisäksi opiskelijoita kannustetaan esittämään esimerkkejä valitsemastaan aiheesta.

1) Osoita, että matriiseille pätevät seuraavat osittelulait:
   i) $A(B+C) = AB + AC$,
   ii) $(A+B)C = AC + BC$.
2) Osoita, että transpoosille pätee:
   i) $(A^T)^T = A$,
   ii) $(A+B)^T = A^T + B^T$,
   ii) $(\lambda A)^T = \lambda A^T$,
3) Osoita, että suunnikasyhtälö pätee:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2),$$

   kaikilla $x, y \in \mathbb{R}^n$.
4) Osoita, että jos $ad - bc \neq 0$. niin

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

   kaikilla $a, b, c, d \in \mathbb{R}$.

5) Osoita Cauchyn-Bunjakovskin-Schwarzin epäyhtälö:
$$|\langle x, y\rangle| = \|x\| \, \|y\| ,$$
kaikilla $x, y \in \mathbb{R}^n$.
6) Olkoot matriisit A ja B samankokoisia symmetrisiä neliömatriiseja. Osoita, että tällöin matriisitulo AB ei ole symmetrinen.

LUKU 9

Merkintöjä

Merkintä	*Selitys*
$\mathbb{R}$	Reaalilukujen joukko
$\mathbb{C}$	Kompleksilukujen joukko
$\mathbb{N}$	Luonnollisten lukujen joukko
F	Joko reaalilukujen tai kompleksilukujen joukko
V	Vektoriavaruus V, jossa määritelty yhteenlasku ja kertominen luvulla
M_n	$n \times n$ -dimensioiset joko reaaliset tai kompleksiset matriisit
I	Identtinen neliömatriisi
$\| \; \|$	Itseisarvo
$\\| \; \\|$	Vektorinormi F-kertoimisessa vektoriavaruudessa V
$\vert\vert\vert \; \vert\vert\vert$	Matriisinormi joukossa M_n
$\det A$	Matriisin A determinantti, josta käytetään myös merkintää $\vert A \vert$
$\langle x, y \rangle$	Vektoreiden $x, y \in V$ sisätulo
A^T	Reaalisen matriisin A transpoosi
A^*	Kompleksisen matriisin A hermitoitu matriisi
A^{-1}	Matriisin A käänteismatriisi
$\bar{z}$	Kompleksiluvun $z = a + ib$ konjugaatti
$tr(A)$	Matriisin A jälki
λ	Matriisin A ominaisarvo
$\lim_{k \to \infty} a_k$	Jonon termien $a_k \in \mathbb{R}$ raja-arvo, kun $k \in \mathbb{N}$ lähestyy ääretöntä
$\rho(A)$	Matriisin A spektraalisäde
$\sigma(A)$	Matriisin A spektri eli ominaisarvojen joukko
$\kappa(A)$	Matriisin A ehtoluku
Δ	Yläkolmiomatriisi

Lähdeluettelo

[1] Donald Allen: *Linear Algebra: Chapter 3, Eigenvalues and Eigenvectors.* http://www.math.tamu.edu/~dallen/papers.html (viitattu 15.1.2015)

[2] Simon Foucart: *Linear Algebra and Matrix Analysis: Lecture 6.* http://www.math.uga.edu/~foucart/teaching.htm#hide1 (viitattu 15.1.2015)

[3] Roger A. Horn ja Charles R. Jonhson: *Matrix Analysis.* Cambridge University Press, 1985.

[4] Tero Kilpeläinen: *Analyysi 3.* Luentomuistiinpanoja syksyltä 2005. http://www.math.jyu.fi/opiskelu/monisteet/MATA113.pdf (viitattu 15.1.2015)

[5] Lassi Kurittu: *Metriset avaruudet.* Luentomuistiinpanoja syksyltä 2012. http://users.jyu.fi/~lkurittu/metravar.pdf (viitattu 15.1.2015)

[6] Peter Lancaster ja Miron Tismenetsky: *The Theory of Matrices.* Second Edition with Applications, Academic Press, 1985

[7] Clare Parnell: *Vector and Matrix norms.* http://www-solar.mcs.st-and.ac.uk/~clare/Lectures/num-analysis/Numan_chap1.pdf (viitattu 15.1.2015)

[8] Veikko T. Purmonen: *Lineaarinen algebra ja geometria 1.* Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, Luentomoniste 58, 2008

[9] Anthony Ralston, Philip Rabinowitz: *A First Course in Numerical Analysis.* Second Edition, 2001

[10] Keijo Ruotsalainen: *Numeeriset menetelmät*, Luentomoniste keväältä 2010. http://s-mat-pcs.oulu.fi/~keba/NumMen/num_luentomoniste.pdf (viitattu 15.1.2015)

[11] Mikko Saarimäki: *Matriisilaskenta*. Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, Luentomoniste, kevät 2014.

[12] Mikko Saarimäki: *Matriisiteoria*. Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, Luentomoniste 31, 1994.

[13] Risto Silvennoinen: *Moniste ominaisarvoista ja diagonalisoinnista.* http://matriisi.ee.tut.fi/courses/7303045 (viitattu 15.1.2015)

[14] Risto Silvennoinen: *Moniste matriisilaskennan kertauksesta.* http://matriisi.ee.tut.fi/courses/7303045 (viitattu 15.1.2015)

[15] Joonas Varso: *Matriisilaskentaa*. Aalto-yliopisto. https://noppa.aalto.fi/noppa/kurssi/as-74.1106/materiaali/matriisilaskenta.pdf (viitattu 15.1.2015)

[16] Wikipedia: http://en.wikipedia.org/wiki/Matrix_%28mathematics%29 (viitattu 15.1.2015)