# MATRIISIN HESSENBERGIN MUOTO

Niko Holopainen

Matematiikan pro gradu
Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Syksy 2013

**Tiivistelmä:** Niko Holopainen, *Matriisin Hessenbergin muoto*. Matematiikan pro gradu-tutkielma, 55 sivua. Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, syksy 2013.

Tämä tutkielma käsittelee lineaarialgebraa ja erityisesti matriisiteoriaa. Erityisesti tutustutaan neliömatriiseihin ja niiden ominaisarvojen karakteristisen polynomin löytämiseen ja sitä kautta ominaisarvojen ratkaisemiseen. Mitä yksinkertaisemmassa muodossa matriisi on, sitä helpommin sen karakteristinen polynomi on ratkaistavissa. Helpompi muoto jollekin neliömatriisille $A$ tarkoittaa, että etsitään sille jokin yksinkertaisemmassa muodossa oleva matriisi $B$ niin, että ne ovat keskenään similaarisia.

Tutkielman tarkoituksena on selventää, kuinka yksinkertaisessa muodossa jokin mielivaltainen neliömatriisi on mahdollista esittää. Eräs yksinkertainen neliömatriisityyppi on Hessenbergin matriisi, johon tässä tutkielmassa tutustutaan. Hessenbergin matriisia on kahta tyyppiä: ylämuotoa ja alamuotoa. Hessenbergin ylämuodossa oleva matriisi voi poiketa yläkolmiomatriisista ainoastaan sen ensimmäisen alemman sivudiagonaalin kohdalla. Hessenbergin alamuodossa oleva matriisi voi puolestaan poiketa alakolmiomatriisista ainoastaan sen ensimmäisen ylemmän sivudiagonaalin kohdalla. Hessenbergin muodon etuna on, että kaikki neliömatriisit voidaan esittää kyseisessä muodossa. Tutkielmassa selvitetään erilaisia algoritmeja neliömatriisin muuntamiseksi sekä Hessenbergin ylämuotoon että alamuotoon similaarimuunnoksilla. Lisäksi käydään myös läpi neliömatriisin muuntaminen Hessenbergin ylämuotoon Householderin muunnoksen avulla. Tutkielmassa näytetään, kuinka redusoimattoman Hessenbergin matriisin ominaisarvot voidaan ratkaista Krylovin menetelmällä. Lisäksi näytetään myös, kuinka redusoituvan Hessenbergin matriisin ominaisarvot voidaan selvittää.

**Avainsanat:** Neliömatriisi, karakteristinen polynomi, ominaisarvo, similaarisuus, Hessenbergin matriisi, Householderin muunnos, Krylovin menetelmä.

# Sisältö

# Johdantoa

Matriisit ovat olennainen osa lineaarialgebraa. Tässä työssä tarkastelun kohteena ovat erityisesti neliömatriisit. Monissa eri sovelluksissa ollaan kiinnostuneita neliömatriisin ominaisarvoista ja karakteristisesta polynomista. Karakteristinen polynomi ja ominaisarvot ovat kytköksissä toisiinsa, koska ominaisarvot saadaan karakteristisen polynomin nollakohdista. Matriisin $A$ karakteristinen polynomi $p(\lambda)$ on

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

Mitä yksinkertaisemmassa muodossa matriisi on, sitä helpompi karakteristinen polynomi on selvittää. Jos tutkittava neliömatriisi $A$ ei alunperin ole "siistissä"muodossa, se pyritään saattamaan sellaiseen muotoon. Käytännössä on löydettävä jokin yksinkertaisemmassa muodossa oleva matriisi $B$, joka on similaarinen matriisin $A$ kanssa. Matriisit $A$ ja $B$ ovat similaarisia keskenään, jos on olemassa sellainen kääntyvä matriisi $C$, että

$$C^{-1}AC = B.$$

Jos matriisit ovat similaarisia keskenään, niillä on sama karakteristinen polynomi ja tällöin myös samat ominaisarvot.

Jos matriisi $A$ on similaarinen jonkin diagonaalimatriisin $D$ kanssa, se on diagonalisoituva. Diagonaalimatriisin ominaisarvot saadaan luettua suoraan diagonaalilta, joten tällainen muoto olisi ideaali. Kaikki matriisit eivät kuitenkaan ole diagonalisoituvia. Yksi vaihtoehto yksinkertaisemmaksi muodoksi on Hessenbergin matriisi, jota on kahta tyyppiä: ylämuotoa ja alamuotoa. Hessenbergin ylämuodossa oleva matriisi poikkeaakin yläkolmiomatriisista korkeintaan ensimmäisen alemman sivudiagonaalin (engl. subdiagonal) kohdalla ja alamuodossa oleva voi poiketa alakolmiomatriisista ainoastaan ensimmäisen ylemmän sivudiagonaalin (engl. superdiagonal) kohdalla. Tutkielmassa selviää, että itse asiassa kaikki neliömatriisit on mahdollista muuntaa sekä Hessenbergin ylämuotoon että alamuotoon.

Ensimmäisessä luvussa tutustutaan tutkielman kannalta tärkeimpiin peruskäsitteisiin. Aivan ensimmäisenä määritellään kompleksinen sisätulo ja normi. Tämän jälkeen tutustutaan lineaarikuvaukseen ja sitä kautta matriisin määritelmään. Luvussa tutustutaan muiden muassa erilaisiin neliömatriisityyppeihin, matriisituloon ja determinanttiin. Näytetään, kuinka matriisi voidaan osittaa jakamalla se pienempiin lohkoihin. Luvun lopuksi näytetään, kuinka matriiseista voidaan muodostaa polynomeja tavallisten lukujen tapaan.

Ensimmäisessä luvussa käsitellyt perusasiat löytyvät kirjallisuuslähteistä [**1**] ja [**5**]-[**9**]. Osa esiintyvistä lauseista on todistettu, mutta osa sivuutetaan. Edellä mainituista lähteistä löytyvät todistukset lauseisiin.

Toisessa luvussa keskitytään ominaisarvoteoriaan ja erityisesti matriisin karakteristiseen polynomiin ja ominaisarvoihin. Selvitetään, miksi matriisin $A$ ominaisarvoja ratkoessa päädytään tutkimaan juuri yhtälöä

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

Yhtälön vasen puoli määrittelee matriisin karakteristisen polynomin $p_A(\lambda)$. Tarkastellaan karakteristisen polynomin yhteyttä matriisipolynomeihin. Määritellään käsite similaarisuus ja todetaan, että similaarisilla matriiseilla todellakin on sama karakteristinen polynomi ja samat ominaisarvot. Toisessa luvussa käsiteltävien asioiden pohjana on käytetty pääosin lähdettä [**1**] ja myös lähteitä [**4**],[**6**],[**8**] ja [**9**].

Kolmannessa luvussa tutustutaan matriisin Hessenbergin muotoon. Selvitetään, kuinka matriisi voidaan muuntaa Hessenbergin ylämuotoon ja alamuotoon similaarimuunnoksia käyttäen. Luvussa myös todetaan, että jokainen matriisi todellakin voidaan esittää sekä Hessenbergin ylämuodossa että alamuodossa. Luvun lopuksi tutustutaan vielä Householderin muunnokseen, jonka avulla matriisi voidaan muuntaa Hessenbergin ylämuotoon. Tässä luvussa on käytetty pohjana lähdettä [**2**] laajentaen sen reaaliset tarkastelut kompleksiseen tapaukseen.

Viimeisessä luvussa on tutustuttu Hessenbergin matriisin ominaisarvojen ratkaisemiseen. Redusoitumattoman Hessenbergin matriisin tapauksessa käytetään Krylovin menetelmää. Redusoituva Hessenbergin matriisi voidaan puolestaan jakaa sopiviin lohkoihin. Tämän luvun pohjana on pääosin käytetty lähdettä [**2**] laajentaen sen reaaliset tarkastelut kompleksiseen tapaukseen ja myös lähteitä [**3**] ja [**9**].

LUKU 1

# Matriisi

## 1.1. Kompleksinen sisätulo ja normi

Kompleksilukujen kunta $\mathbb{C}$ koostuu lukupareista $(x, y) \in \mathbb{R}^2$ ja ne ovat muotoa

$$z = x + iy,$$

jossa $i$ on imaginaariyksikkö, jolle pätee $i^2 = -1$. Kompleksiluvun $z \in \mathbb{C}$ liittoluku (tai kompeksikonjugaatti) määritellään asettamalla

$$\overline{z} = x - iy.$$

Kompleksiluvun $z$ reaaliosa on

$$x =: \mathrm{Re}(z) \in \mathbb{R}.$$

ja imaginääriosa on

$$y =: \mathrm{Im}(z) \in \mathbb{R}.$$

Kompleksiluvun $z \in \mathbb{C}$ itseisarvo (tai moduli) määritetään asettamalla

$$|z| := \sqrt{z\overline{z}} = \sqrt{(x + iy)(x - iy)} = \sqrt{x^2 - i^2y^2} = \sqrt{x^2 + y^2}.$$

Selvästi $\mathrm{Re}(z) \leq |\mathrm{Re}(z)| \leq |z|$ ja $\mathrm{Im}(z) \leq |\mathrm{Im}(z)| \leq |z|$ kaikilla $(x, y) \in \mathbb{R}^2$. Tässä $|\mathrm{Re}(z)|$ ja $|\mathrm{Im}(z)|$ ovat tavallisia itseisarvoja avaruudessa $\mathbb{R}$. Kannattaa myös huomata, että

$$z + \overline{z} = (x + iy) + (x - iy) = 2x = 2\,\mathrm{Re}(z).$$

Reaalilukujen joukko $\mathbb{R}$ voidaan ajatella kompleksilukujen kunnan $\mathbb{C}$ osajoukkona, koska jokainen reaaliluku $a \in \mathbb{R}$ voidaan samaistaa kompleksiluvuksi $(a, 0) = a + i \cdot 0 = a$. Jokaiselle muotoa $(a, 0)$ olevalle kompleksiluvulle pätee tällöin selvästi $a = \overline{a}$.

Kompleksinen vektori $z \in \mathbb{C}^n$ on muotoa $z = (z_1, \ldots, z_n)$, jossa komponentit $z_1, \ldots, z_n$ ovat kompleksilukuja. Kompleksisen vektorin $z \in \mathbb{C}^n$ kompleksikonjugaatti saadaan ottamalla jokaisesta komponentista erikseen liittoluku:

$$\overline{z} = (\overline{z_1}, \ldots, \overline{z_n}).$$

Vektoreiden $z = (z_1, \ldots, z_n)$ ja $w = (w_1, \ldots, w_n)$ kompleksinen sisätulo määritetään asettamalla

$$(z \,|\, w) = \sum_{i=1}^{n} z_i\overline{w_i} = z_1\overline{w_1} + \ldots z_n\overline{w_n}.$$

Jos vektoreiden $a, b \in \mathbb{C}^n$ kaikki komponentit ovat reaalisia (se on reaalinen vektori), tällöin

$$(a \,|\, b) = \sum_{i=1}^{n} a_i \overline{b_i} = a_1 \overline{b_1} + \ldots + a_n \overline{b_n} = a_1 b_1 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i,$$

jolloin kompleksinen sisätulo vastaa tavallista Eukleideen sisätuloa reaalisille vektoreille.

Kuvaus $\| \cdot \| \colon \mathbb{C}^n \to \mathbb{R}$ on normi avaruudessa $\mathbb{C}^n$, jos se toteuttaa seuraavat ehdot:

(1) $\| x \| \geq 0$ kaikilla $x \in \mathbb{C}^n$.
(2) $\| x \| = 0$, jos ja vain jos $x = 0$.
(3) $\| x + y \| \leq \| x \| + \| y \|$ kaikilla $x, y \in \mathbb{C}^n$.
(4) $\| \alpha x \| = |\alpha| \| x \|$ kaikilla $\alpha \in \mathbb{C}$ ja $x \in \mathbb{C}^n$.

Tavallinen Eukleideen normi saadaan ottamalla neliöjuuri vektorin sisätulosta itsensä kanssa. Kompleksisessa tapauksessa normi määritetään muuten samalla tavalla, mutta tavallisen reaalisen sisätulon sijaan käytetään kompleksista sisätuloa. Kompleksisen vektorin $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$ normiksi saadaan tällöin

$$\|z\| = \sqrt{(z \,|\, z)} = \sqrt{\sum_{i=1}^{n} z_i \overline{z_i}} = \sqrt{z_1 \overline{z_1} + \ldots z_n \overline{z_n}}.$$

Vektorit $x, y \in \mathbb{C}^n$ ovat ortogonaaliset eli kohtisuorat, jos

$$(x \,|\, y) = (y \,|\, x) = 0.$$

Vektorin $y$ ortogonaaliprojektio vektorin $x$ suuntaan määritetään asettamalla

$$P_x y := \frac{(y \,|\, x)}{\|x\|^2} x.$$

Tällöin pätee

$$(y - P_x y \,|\, x) = (y \,|\, x) - (P_x y \,|\, x) = (y \,|\, x) - \frac{(y \,|\, x)}{\|x\|^2} \|x\|^2 = 0,$$

jolloin $y - P_x y$ ja $x$ ovat kohtisuorassa keskenään. Koska

$$\begin{aligned} (P_x y - y \,|\, P_x y) &= (P_x y \,|\, P_x y) - (y \,|\, P_x y) \\ &= \frac{(y \,|\, x)^2}{\|x\|^4} \|x\|^2 - \frac{(y \,|\, x)^2}{\|x\|^2} \\ &= 0, \end{aligned}$$

myös $P_x y - y$ ja $P_x y$ ovat kohtisuorassa keskenään.

Ortogonaalisille vektoreille pätee Pythagoraan lause, koska

$$\begin{aligned} \|x + y\|^2 = (x + y \,|\, x + y) &= \|x\|^2 + \underbrace{(x \,|\, y)}_{=0} + \underbrace{(y \,|\, x)}_{=0} + \|y\|^2 \\ &= \|x\|^2 + \|y\|^2. \end{aligned}$$

Seuraava tulos antaa hyödyllisen arvion kompleksisille vektoreille avaruudessa $\mathbb{C}^n$.

LAUSE 1.1 (Cauchyn-Schwarzin epäyhtälö). Kaikille $x, y \in \mathbb{C}^n$ pätee

$$|(x \,|\, y)| \leq \|x\|\|y\|.$$

TODISTUS. Tapaus on selvä, kun $x = 0$. Käsitellään tapaus $x \neq 0$. Koska $y - P_x y$ ja $P_x y$ ovat kohtisuorassa keskenään, pätee Pythagoraan lauseen perusteella

$$\begin{aligned}\|y\|^2 &= \|y - P_x y + P_x y\|^2 = \|y - P_x y\|^2 + \|P_x y\|^2 \geq \|P_x y\|^2 \\ &= \frac{|(y \,|\, x)|^2}{\|x\|^4}\|x\|^2 = \frac{|(y \,|\, x)|^2}{\|x\|^2}.\end{aligned}$$

Edellistä muokkaamalla saadaan

$$|(x \,|\, y)|^2 \leq \|x\|^2\|y\|^2,$$

joka on yhtäpitävää sen kanssa, että

$$|(x \,|\, y)| \leq \|x\|\|y\|.$$

$\square$

LAUSE 1.2 (Kolmioepäyhtälö). Kaikille $x, y \in \mathbb{C}^n$ pätee

$$\|x + y\| \leq \|x\| + \|y\|.$$

TODISTUS.

$$\begin{aligned}\|x + y\|^2 &= (x + y \,|\, x + y) = \|x\|^2 + (x \,|\, y) + (y \,|\, x) + \|y\|^2 \\ &= \|x\|^2 + (x \,|\, y) + \overline{(x \,|\, y)} + \|y\|^2 \\ &= \|x\|^2 + 2\,\mathrm{Re}((x \,|\, y)) + \|y\|^2 \\ &\leq \|x\|^2 + 2|(x \,|\, y)| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

Edellinen on yhtäpitävää sen kanssa, että

$$\|x + y\| \leq \|x\| + \|y\|.$$

$\square$

Kompleksiselle normille pätevät selvästi normin ehdot (1) ja (2). Ehto (3) saadaan lauseesta 1.2. Lisäksi kompleksiselle normille pätee

$$\|\, \alpha x \,\| = \sqrt{\alpha x \overline{\alpha x}} = \sqrt{\alpha\overline{\alpha}}\sqrt{x\overline{x}} = |\alpha|\|\, x \,\|,$$

mikä on ehto (4). Kompleksinen normi näin ollen täyttää normin ehdot. Jos vektorin $z$ kaikki komponentit ovat reaalisia, kompleksinen normi vastaa tällöin ”tavallista” Eukleideen vektorinormia reealisille vektoreille.

## 1.2. Lineaarikuvaus ja matriisi

Tarkastellaan aluksi lineaarista yhtälöryhmää

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \quad\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{1}$$

jossa $a_{ij}$ ja $b_i$ (tässä $i = 1, \ldots, m$ ja $j = 1, \ldots, n$) ovat kompleksisia vakioita ja $x_1, \ldots, x_n$ kompleksisia muuttujia. Merkitään $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$, jolloin $x$ voidaan samaistaa vektoriksi, jolla on $n$ komponenttia: $x_1, \ldots, x_n$. Yhtälöryhmän (1) yhtälöiden vasemmat puolet määrittelevät kuvauksen

$$L\colon \mathbb{C}^n \to \mathbb{C}^m\colon x \mapsto (a_{11}x_1 + \ldots + a_{1n}x_n, \ldots, a_{m1}x_1 + \ldots + a_{mn}x_n).$$

Tällaista kuvausta $L$ kutsutaan lineaarikuvaukseksi ja sillä on selvästi seuraavat ominaisuudet:

(1) $L(x + y) = L(x) + L(y)$ kaikilla $x, y \in \mathbb{C}^n$,
(2) $L(\lambda x) = \lambda L(x)$ kaikilla $x \in \mathbb{C}^n$ ja $\lambda \in \mathbb{C}$.

Tarkastellaan edelleen lineaarista yhtälöryhmä (1) ja esitetään se muodossa

$$\begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}. \tag{2}$$

Vektori $x \in \mathbb{C}^n$ on mahdollista esittää sarakevektorina

$$x = (x_1, \ldots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Nyt yhtälö (2) voidaan esittää muodossa

$$\begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix} =: \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Yhtälöryhmän kertoimien $a_{ij}$ muodostamaa kaaviota

$$A = A_{m \times n} = \begin{bmatrix} a_{ij} \end{bmatrix} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}$$

kutsutaan matriisiksi, tarkemmin $m \times n$-matriisiksi. Matriisilla $A$ on selvä yhteys lineaarikuvaukseen $L$, koska jokainen lineearikuvaus voidaan esittää aina matriisina. Matriisi $A$ onkin lineaarikuvausta $L$ vastaava matriisi.

Matriisi koostuu riveista ja sarakkeista. Matriisin $A$ sarakkeet

$$\bar{a}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

ovat vektoreina sen sarakevektoreita $\bar{a}_j = (a_{1j}, \ldots, a_{mj}) \in \mathbb{C}^m$. Matriisin $A$ rivit

$$\vec{a}_i = \begin{bmatrix} a_{i1} & \ldots & a_{in} \end{bmatrix}$$

ovat vektoreina sen rivivektoreita $\vec{a}_i = (a_{i1}, \ldots, a_{in}) \in \mathbb{C}^n$. Matriisi voidaan näin ollen esittää myös muodossa

$$A = \begin{bmatrix} \bar{a}_1 & \ldots & \bar{a}_n \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix}$$

## 1.3. Matriisilaskentaa

**1.3.1. Neliömatriisi.** Matriisia, jolla on sama määrä rivejä ja sarakkeita, kutsutaan neliömatriisiksi:

$$A = A_{n\times n} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}.$$

Neliömatriisilla on joitakin erikoistapauksia. Neliömatriisia, joiden alkioille pätee $a_{ij} = 0$ silloin, kun $i > j$, sanotaan yläkolmiomatriisiksi:

$$A = A_{n\times n} = \begin{bmatrix} a_{11} & \ldots & \ldots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & a_{nn} \end{bmatrix}.$$

Vastaavasti neliömatriisia, joiden alkioille pätee $a_{ij} = 0$ silloin, kun $i < j$, sanotaan alakolmiomatriisiksi:

$$A = A_{n\times n} = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ a_{n1} & \ldots & \ldots & a_{nn} \end{bmatrix}.$$

Ylä- ja alakolmiomatriiseille käytetään yhteisnimitystä kolmiomatriisi.

Neliömatriisia, joiden alkioille pätee $a_{ij} = 0$ silloin, kun $i \neq j$, kutsutaan lävistäjä- tai diagonaalimatriisiksi:

$$A = A_{n\times n} = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & a_{nn} \end{bmatrix}.$$

Diagonaalimatriisille käytetään myös merkintää $A = \operatorname{diag}(a_{11}, \ldots, a_{nn})$. Diagonaalimatriisin erikoistapaus on yksikkömatriisi $I$, joissa kaikki diagonaalin alkiot ovat ykkösiä. Tällöin $I = \operatorname{diag}(1, \ldots, 1)$:

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

LAUSE 1.3. Matriisille $A_{m\times n}$ pätee aina

$$AI_n = A = I_m A.$$

Jos neliömatriisille $A$ löytyy sellainen neliömatriisi $B$, että

$$AB = I,$$

sanotaan, että matriisi $A$ on kääntyvä ja silloin $B$ on sen käänteismatriisi. Tällöin myös $BA = I$. Matriisin $A$ käänteismatriisille käytetään merkintää $A^{-1}$. Jos matriisi ei ole kääntyvä, sitä kutsutaan singulaariseksi.

LAUSE 1.4. Matriisi $A_{n\times n}$ on singulaarinen, jos ja vain jos on olemassa sellainen $x \in \mathbb{C}^n$ $(x \neq 0)$, jolle pätee

$$Ax = 0.$$

Neliömatriisille $A_{n\times n}$ voidaan määrittää potensseja samaan tapaan kuin reaaliluvuille. Asetetaan ensin, että $A^0 = I$, jolloin yksikkömatriisi $I$ voidaan ikään kuin samaistaa ykköseksi matriisien tapauksessa. Matriisia voidaan tunnetusti kertoa reaaliluvulla. Yleisemmin reaaliluku $\lambda$ voidaan samaistaa matriisiksi $\lambda I$. Toisaalta voidaan ajatella, että $A^1 = A$ ja $A^2 = AA$. Sitä suuremmat potenssit voidaan määrittää rekursiivisesti $A^k = AA^{k-1}$, jolloin edellinen voidaan puolestaan samaistaa matriisien potensseiksi.

Matriisia, jonka kaikki alkiot ovat nollia, kutsutaan nollamatriisiksi ja sitä merkitään usein yksinkertaisesti vain nollalla:

$$\begin{bmatrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{bmatrix} =: 0.$$

Nollamatriisi ei tosin aina ole neliömatriisi.

**1.3.2. Matriisin adjungaatti.** Matriisin $A_{m\times n} = \big[a_{ij}\big]$ transpoosi on matriisi $B_{n\times m}$, jonka alkioille pätee $a_{ij} = b_{ji}$ kaikilla indekseillä $i$ ja $j$. Matriisin $A$ transpoosi saadaankin vaihtamalla sarakkeet riveiksi ja rivit sarakkeiksi. Matriisin $A$ transpoosille käytetään merkintää $A^T$.

Matriisin $A_{m\times n} = [a_{ij}]$ kompleksikonjugaatti on matriisi $B_{n\times m}$, jonka alkioille pätee $b_{ij} = \overline{a_{ij}}$ kaikilla indekseillä $i$ ja $j$. Matriisin $A$ kompleksikonjugaatille käytetään merkintää $\overline{A}$.

Matriisin $A$ adjungaatti määritetään asettamalla

$$A^* = \overline{A}^T.$$

Matriisin $A$ adjungaatti on sen kompleksikonjugaatin transpoosi. Matriisin $A$ adjungaatin $B$ alkioille pätee $b_{ji} = \overline{a_{ij}}$ kaikilla indekseillä $i$ ja $j$. Reaalisessa tapauksessa matriisin $A$ adjungaatti vastaa sen transpoosia.

Matriisia $A$ kutsutaan itseadjungoituvaksi, jos $A = A^*$. Reaalisessa tapauksessa tämä vastaa symmetristä matriisia.

Neliömatriisia $U_{n\times n}$ kutsutaan unitaariseksi, jos sen käänteismatriisi on sen adjungaatti, jolloin $U^{-1} = U^*$. Reaalisessa tapauksessa tämä vastaa ortogonaalista matriisia.

**1.3.3. Matriisitulo.** Matriisien $A_{l\times m}$ ja $B_{m\times n}$ tulomatriisin $C_{l\times n}$ alkioille pätee

$$c_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj} \quad \text{kaikilla indekseillä } i \text{ ja } j.$$

Matriisien $A_{l\times m}$ ja $B_{m\times n}$ tulo on tällöin muotoa

$$\begin{aligned} AB &= \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{l1} & \dots & a_{lm} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + \ldots + a_{1m}b_{m1} & \dots & a_{11}b_{1n} + \ldots + a_{1m}b_{mn} \\ \vdots & & \vdots \\ a_{l1}b_{11} + \ldots + a_{lm}b_{m1} & \dots & a_{l1}b_{1n} + \ldots + a_{lm}b_{mn} \end{bmatrix}. \end{aligned}$$

Jotta tulo olisi mahdollista laskea, matriisin $A$ sarakkeiden ja matriisin $B$ rivien lukumäärät on oltava samat.

**1.3.4. Determinantti.** Determinantin määrittämiseen on useita erilaisia tapoja. Käydään lyhyesti läpi yhtä tapaa näistä. Määritellään ensin alimatriisin käsite. Kun matriisilta $A_{n\times n}$ poistetaan $i$. rivi ja $j$. sarake, muodostuvaa $(n-1)\times(n-1)$-matriisia kutsutaan matriisin $A$ alimatriisiksi paikan $(i,j)$ suhteen, ja merkitään $A_{ij}$.

Kokoa $1\times 1$ olevan matriisin $A = [a]$ determinantti on sen ainoan alkion arvo: $\det A := a$.

Kokoa $2\times 2$ olevan matriisin

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

determinantti $\det A$ määritellään asettamalla

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

Suurempien neliömatriisien determinantit voidaan määrittää rekursiivisesti. Neliömatriisin $A_{n\times n}$ $(n \geq 3)$ determinantti voidaan hajottaa sen alimatriisien determinanttien lineaarikombinaatioksi asettamalla

$$\begin{aligned}\det A &= a_{11}\det A_{11} - a_{21}\det A_{21} + \ldots + (-1)^{n+1}a_{n1}A_{n1}\\ &= \sum_{k=1}^{n}(-1)^{k+1}a_{k1}\det A_{k1}.\end{aligned}$$

Vastaavasti jokaisen alimatriisin $A_{k1}$ determinantti voidaan hajottaa pienempien matriisien determinanttien lineaarikombinaatioksi. Samalla periaatteella edetään rekursiivisesti, kunnes päästään laskemaan kokoa $2\times 2$ olevien matriisien determinantteja. Esimerkiksi $3\times 3$-matriisin determinantti on

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

Matriisin $A_{n\times n}$ alkion $a_{ij}$ kofaktori $\operatorname{cof} a_{ij}$ määritetään asettamalla

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

Matriisia $B$, jonka alkioille pätee $b_{ij} = \operatorname{cof} a_{ij}$ kaikilla indekseillä $i$ ja $j$, sanotaan matriisin $B$ liittomatriisiksi (tai kofaktorimatriisiksi). Tällöin merkitään $B =: \tilde{A}$.

Lause 1.5. Olkoon $\tilde{A}$ matriisin $A$ liittomatriisi. Tällöin

$$A\tilde{A} = \tilde{A}A = (\det A)I.$$

Lause 1.6. Matriisi $A_{n\times n}$ on kääntyvä jos ja vain jos

$$\det A \neq 0.$$

Lause 1.7. Kolmiomatriisin $A_{n\times n} = [a_{ij}]$ determinantti on sen diagonaalialkoiden tulo, toisin sanoen

$$\det A = \prod_{j=1}^{n} a_{jj} = a_{11}\ldots a_{nn}.$$

**1.3.5. Rivioperaatiot ja alkeismatriisit.** Matriisimuotoinen lineaariyhtälöryhmä voidaan ratkaista *Gaussin-Jordanin menetelmää* käyttäen. Menetelmä perustuu ns. rivioperaatioihin, joita on kolmea eri tyyppiä. Kukin rivioperaatio saadaan tehtyä kertomalla muokattavaa matriisia vasemmalta jollakin sopivalla alkeismatriisilla.

Alkeismatriiseja muodostettaessa tarvitaan ensin kantamatriisin käsite. Kantamatriisin $E_{ij}$ alkiot ovat muotoa

$$[E_{ij}]_{kl} = \begin{cases} 1, & \text{kun } i = k \text{ ja } j = l,\\ 0, & \text{muulloin.}\end{cases}$$

Kantamatriisilla $E_{ij}$ on ykkönen paikassa $(i,j)$ ja muut alkiot ovat nollia.

Ensimmäinen alkeismatriisi määritetään asettamalla

$$M_i(\lambda) = I + (\lambda - 1)E_{ii}, \quad \text{jossa } \lambda \in \mathbb{C}, \lambda \neq 0.$$

Matriisi $M_i(\lambda)$ saadaan käytännössä vaihtamalla yksikkömatriisissa paikassa $(i,i)$ ykkönen luvuksi $\lambda$. Tämä operaatio kertoo matriisin $A$ riviä $i$ luvulla $\lambda$.

Toinen alkeismatriisi määritetään asettamalla

$$P_{ij} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}, \quad \text{jossa } i \neq j.$$

Matriisi $P_{ij}$ saadaan vaihtamalla yksikkömatriisista paikoissa $(i,i)$ ja $(j,j)$ ykköset nolliksi ja vaihtamalla paikoissa $(i,j)$ ja $(j,i)$ nollat ykkösiksi. Tämä operaatio vaihtaa matriisin $A$ rivit $i$ ja $j$ keskenään.

Kolmas ja viimeinen alkeismatriisi määritetään asettamalla

$$A_{ij}(\lambda) = I + \lambda E_{ji}, \quad \text{jossa } i \neq j.$$

Matriisi $A_{ij}(\lambda)$ saadaan vaihtamalla yksikkömatriisin paikassa $(j,i)$ nolla luvuksi $\lambda$. Tämä operaatio lisää rivin $i$ kerrottuna luvulla $\lambda$ riviin $j$.

Kaikille kolmelle operaatiolle on olemassa käänteisoperaatiot, ja ne saadaan vastaavien alkeismatriisien käänteismatriiseina. Alkeismatriisien käänteismatriisit ovat

$$\begin{aligned} M_i(\lambda)^{-1} &= M_i\left(\frac{1}{\lambda}\right), \\ P_{ij}^{-1} &= P_{ij}, \\ A_{ij}(\lambda)^{-1} &= A_{ij}(-\lambda). \end{aligned}$$

## 1.4. Matriisien lohkomuodot

Oletetaan, että matriisille $A_{m\times n}$ pätee $m_1 + m_2 = m$ joillakin $m_1, m_2 \geq 1$. Tällöin matriisi voidaan esittää lohkomuodossa

$$A = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix},$$

jossa $A_{11}$ on $m_1 \times n$-matriisi ja $A_{21}$ on $m_2 \times n$-matriisi. Tällöin matriisin $A$ rivivektorit $\vec{a}_1, \ldots, \vec{a}_{m_1}$ muodostavat matriisin $A_{21}$ rivit ja rivivektorit $\vec{a}_{m_1+1}, \ldots, \vec{a}_m$ muodostavat matriisin $A_{22}$ rivit.

Oletetaan, että matriisille $A_{m\times n}$ pätee $n \geq 2$ ja $n_1 + n_2 = n$. Tällöin matriisi voidaan esittää muodossa

$$A = \begin{bmatrix} A_{11} & A_{12} \end{bmatrix},$$

jossa $A_{11}$ $m \times n_1$-matriisi ja $A_{12}$ on $m \times n_2$-matriisi. Tällöin matriisin $A$ sarakevektorit $\bar{a}_1, \ldots, \bar{a}_{n_1}$ muodostavat matriisin $A_{11}$ sarakkeet ja sarakevektorit $\bar{a}_{n_1+1}, \ldots, \bar{a}_n$ muodostavat matriisin $A_{12}$ sarakkeet.

Jos matriisille $A_{m\times n}$ pätee $m \geq 2$ ja $m_1 + m_2 = m$, sekä $n \geq 2$ ja $n_1 + n_2 = n$, se voidaan esittää muodossa

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

Yleistyksenä matriisi voidaan vastaavalla periaatteella jakaa sekä rivien että sarakkeiden suhteen mielivaltaiseen määrään lohkoja.

Lause 1.8. Oletetaan, että matriisi $A$ voidaan esittää lohkomuodossa

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

jossa $A_{11}$ ja $A_{22}$ ovat neliömatriiseja. Tällöin

$$\det A = \det A_{11} \cdot \det A_{22}.$$

## 1.5. Matriisipolynomit

Olkoon polynomi $p$ muotoa

$$p(t) = a_k t^k + \ldots + a_1 t + a_0.$$

Tällöin matriisia

$$p(A) = a_k A^k + \ldots + a_1 A + a_0 I$$

sanotaan matriisin $A$ matriisipolynomiksi.

LUKU 2

# Ominaisarvoteoriaa

## 2.1. Ominaisarvo, ominaisvektori ja ominaisavaruus

Tutkitaan tilannetta, missä lineearikuvaukselle $L\colon \mathbb{C}^n \to \mathbb{C}^n$ halutaan löytää jokin sellainen vektori $x \in \mathbb{C}^n$, että $x$ ja $Lx$ ovat yhdensuuntaisia. Tällöin on löydettävä jokin $\lambda \in \mathbb{C}$, jolle pätee

$$Lx = \lambda x.$$

Oletetaan, että

$$A = A_{n\times n} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

on lineearikuvausta $L$ vastaava matriisi. Arvoa $\lambda \in \mathbb{C}$ kutsutaan matriisin $A$ *ominaisarvoksi*, mikäli on olemassa sellainen vektori $x \in \mathbb{C}^n$ ($x \neq 0$), että

$$Ax = \lambda x.$$

Vektoria $x$ kutsutaan puolestaan ominaisarvoa $\lambda$ vastaavaksi *ominaisvektoriksi*. Matriisin $A$ kaikkien ominaisarvojen joukkoa kutsutaan sen *spektriksi* ja merkitään $\sigma(A)$. Aliavaruudelle

$$V_\lambda = \{x \in \mathbb{C}^n \colon Ax = \lambda x\}$$

käytetään nimitystä matriisin $A$ ominaisarvoa $\lambda$ vastaava *ominaisavaruus*.

Esimerkki 2.1. Tarkastellaan $2 \times 2$ matriisia

$$A = \begin{bmatrix} 6 & -7 \\ 9 & -10 \end{bmatrix}$$

ja vektoria $x_1 = (7, 9)$. Tällöin

$$Ax_1 = \begin{bmatrix} 6 & -7 \\ 9 & -10 \end{bmatrix} \begin{bmatrix} 7 \\ 9 \end{bmatrix} = \begin{bmatrix} -21 \\ -27 \end{bmatrix} = -3 \begin{bmatrix} 7 \\ 9 \end{bmatrix}.$$

Matriisin $A$ yksi ominaisarvo on $\lambda_1 = -3$ ja eräs sitä vastaava ominaisvektori $x_1$. Toisaalta vektorilla $x_2 = (1, 1)$ saadaan

$$Ax_2 = \begin{bmatrix} 6 & -7 \\ 9 & -10 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} = -1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Matriisin $A$ toinen ominaisarvo on $\lambda_2 = -1$ ja sitä vastaava ominaisvektori $x_2$.

## 2.2. Ominaisarvojen ratkaisemisesta

Ominaisarvot voidaan ratkaista muutenkin kuin kokeilemalla. Tarkastellaan yksikkömatriisia $I$. Tällöin kaikille vektoreille $x \in \mathbb{C}^n$ pätee $Ix = x$. Tässä tapauksessa 1 on selvästi ainoa mahdollinen ominaisarvo. Matriisin $A$ ominaisarvolle $\lambda$ pätee $Ax = \lambda x = \lambda Ix$, jolloin $Ax - \lambda Ix = 0$ jollakin vektorilla $x \in \mathbb{C}^n$. Ominaisarvoja selvittäessä päädytään tutkitaan yhtälöä

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

Lause 2.1. Olkoon $A$ kokoa $n \times n$ oleva matriisi. Tällöin $\lambda$ on matriisin ominaisarvo jos, ja vain jos

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

Todistus. Osoitetaan ensin, että jos $\det(A - \lambda I) = 0$, luku $\lambda$ on ominaisarvo. Koska $\det(A - \lambda I) = 0$, lauseen 1.6 perusteella tiedetään, että matriisi $A$ on singulaarinen. On olemassa sellainen $x \in \mathbb{C}^n$ $(x \neq 0)$, että $(A - \lambda I)x = 0$ (lause 1.4). Nyt

$$(A - \lambda I)x = Ax - \lambda Ix = Ax - \lambda x = 0.$$

Tällöin $Ax = \lambda x$, jolloin $\lambda$ on ominaisarvo.

Osoitetaan vielä toisin päin: jos $\lambda$ on ominaisarvo, pätee $\det(A - \lambda I) = 0$. Tehdään antiteesi: $\lambda$ on matriisin $A$ ominaisarvo, mutta $\det(A - \lambda I) \neq 0$. Tällöin matriisi $(A - \lambda I)$ olisi kääntyvä, jolloin sille löytyisi käänteismatriisi $(A - \lambda I)^{-1}$. Kertomalla yhtälö $(A - \lambda I)x = 0$ puolittain vasemmalta käänteismatriisilla $(A - \lambda I)^{-1}$ saadaan

$$\begin{aligned}(A - \lambda I)^{-1}(A - \lambda I)x &= (A - \lambda I)^{-1} \cdot 0 \\ Ix &= 0 \\ x &= 0.\end{aligned}$$

Yhtälön $(A - \lambda I)x = 0$ ainoa mahdollinen ratkaisu on nollavektori $x = 0$, mikä on ristiriita. □

Polynomia $\det(A - \lambda I)$ kutsutaan matriisin $A$ *karakteriseksi polynomiksi* ja yhtälöä

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

matriisin $A$ *karakteristiseksi yhtälöksi.* Matriisin karakteristiselle polynomille käytetään myös merkintää

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

Laskemalla lauseke $\det(A - \lambda I)$ auki huomataan, että kyseessä todellakin on polynomi.

Esimerkki 2.2. Ratkaistaan ominaisarvot matriisille

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

Nyt

$$\begin{aligned} A-\lambda I &= \begin{bmatrix} 5 & 2 \\ -7 & -4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ -7 & -4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \\ &= \begin{bmatrix} 5-\lambda & 2 \\ -7 & -4-\lambda \end{bmatrix}. \end{aligned}$$

Karakteristiseksi polynomiksi saadaan tällöin

$$\begin{aligned} p(\lambda) &= \begin{vmatrix} 5-\lambda & 2 \\ -7 & -4-\lambda \end{vmatrix} = (5-\lambda)(-4-\lambda) - 2\cdot(-7) \\ &= \lambda^2 - \lambda + 6. \end{aligned}$$

Ratkaisemalla karakteristinen yhtälö $\lambda^2 - \lambda + 6 = 0$ saadaan matriisin $A$ ominaisarvoiksi $\lambda_1 = -2$ ja $\lambda_2 = 3$.

Lause 2.2. Kolmiomatriisin $A_{n\times n} = [a_{ij}]$ karakteristinen polynomi on

$$p(\lambda) = \prod_{j=1}^{n} (a_{jj} - \lambda) = (a_{11} - \lambda) \dots (a_{nn} - \lambda).$$

Todistus. Matriisin $A$ karakteristinen polynomi on

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

Koska $A$ on kolmiomatriisi, myös $A - \lambda I$ on kolmiomatriisi. Lauseen 1.7 perusteella matriisin $A - \lambda I$ determinantti on sen diagonaalialkioiden tulo, jolloin

$$\det(A-\lambda) = \prod_{j=1}^{n} (a_{jj} - \lambda) = (a_{11} - \lambda) \dots (a_{nn} - \lambda).$$

$\square$

## 2.3. Ominaisarvoteoriaa matriisipolynomeille

Jos matriisilla $A_{n\times n}$ on ominaisarvo $\lambda \in \mathbb{C}$ ja eräs sitä vastaava ominaisvektori $x \in \mathbb{C}^n$, määritelmän perusteella pätee

$$Ax = \lambda x.$$

Toisaalta

$$A^2 x = A(Ax) = A\lambda x = \lambda(Ax) = \lambda(\lambda x) = \lambda^2 x.$$

Vastaavasti

$$A^3 x = A(A^2 x) = A\lambda^2 x = \lambda^2(Ax) = \lambda^2(\lambda x) = \lambda^3 x.$$

Edelliset voidaankin kirjoittaa induktiivisesti yleisessä muodossa.

Lause 2.3. Olkoon $\lambda \in \mathbb{C}$ matriisin $A_{n\times n}$ ominaisarvo ja $x \in \mathbb{C}^n$ tätä ominaisarvoa vastaava ominaisvektori. Tällöin

$$A^i x = \lambda^i x \quad \text{kaikilla } i = 1, 2 \dots$$

TODISTUS. Käytetään induktiota.
Kun $i = 1$, tällöin suoraan määritelmästä saadaan

$$A^i x = Ax = \lambda x = \lambda^i x.$$

Tehdään induktio-oletus: oletetaan, että väite pätee, kun $i = k$:

$$A^k x = \lambda^k x.$$

Osoitetaan, että jos väite pätee tapauksessa $i = k$, se pätee silloin myös tapauksessa $i = k + 1$ :

$$A^{k+1}x = A(A^k x) = A\lambda^k x = \lambda^k (Ax) = \lambda^k(\lambda x) = \lambda^{k+1}x.$$

□

Käytännössä edellinen lause tarkoittaa, että jos $x \in \mathbb{C}^n$ on matriisin $A_{n\times n}$ ominaisvektori, se on myös matriisin $A^k$ ominaisvektori. Matriisin $A_k$ tapauksessa ominaisvektoria $x$ vastaa ominaisarvo $\lambda^k$.

LAUSE 2.4. Olkoot $p$ polynomi ja matriisi $A_{n\times n}$, jonka ominaisarvoa $\lambda$ vastaa ominaisvektori $x \neq 0$. Tällöin

$$p(A)x = p(\lambda)x.$$

TODISTUS. Lauseen 2.3 perusteella tiedetään, että

$$A^i x = \lambda^i x \quad \text{kaikilla } i = 1, 2 \ldots$$

Olkoon polynomi $p(t)$ muotoa

$$p(t) = a_k t^k + \ldots + a_1 t + a_0,$$

jolloin matriisipolynomiksi saadaan

$$p(A) = a_k A^k + \ldots + a_1 A + a_0 I.$$

Nyt

$$\begin{aligned}
p(A)x &= \left(a_k A^k + \ldots + a_1 A + a_0 I\right) x \\
&= a_k A^k x + \ldots + a_1 A x + a_0 x \\
&= a_k \lambda^k x + \ldots + a_1 \lambda x + a_0 x \\
&= \left(a_k \lambda^k + \ldots + a_1 \lambda + a_0\right) x \\
&= p(\lambda)x.
\end{aligned}$$

□

LAUSE 2.5 (Cayleyn-Hamiltonin lause). Olkoon $p$ matriisin $A_{n\times n}$ karakteristinen polynomi. Tällöin

$$p(A) = 0.$$

Todistus. Matriisin $A$ karakteristinen polynomi on muotoa

$$p(\lambda) = \det(A - \lambda I) = (-1)^n(\lambda^n + \ldots + a_1\lambda + a_0).$$

Olkoon $B$ matriisin $A - \lambda I$ liittomatriisi. Matriisin $B$ kaikille alkioille pätee

$$b_{ij} = (-1)^{i+j} \det(A - \lambda I)_{ij}.$$

Koska $\det(A - \lambda I)$ on muuttujan $\lambda$ suhteen $n$. asteen polynomi, $\det(A - \lambda I)_{ij}$ voi siten olla korkeintaan astetta $n - 1$, jolloin matriisin $B$ alkiot $b_{ij}$ voivat olla korkeintaan tätä astetta. Tämän perusteella tiedetään, että $B$ voidaan esittää muodossa

$$B = \lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \ldots + \lambda B_1 + B_0,$$

jossa $B_0, B_1, \ldots, B_{n-2}, B_{n-1}$ ovat matriiseja, joiden alkiot ovat komplekseja vakioita. Nyt

$$\begin{aligned} B(A - \lambda I) =&(\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \ldots + \lambda B_1 + B_0)(A - \lambda I) \\ =&\lambda^{n-1}B_{n-1}A + \lambda^{n-2}B_{n-2}A + \ldots + \lambda B_1A + B_0A \\ &- \lambda^n B_{n-1} - \lambda^{n-1}B_{n-2} - \ldots - \lambda^2 B_1 - \lambda B_0 \\ =& - \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-1}A - B_{n-2}) + \lambda(B_A1 + B_0) + B_0A. \end{aligned}$$

Edellinen hieman yleisemmin kirjoitettuna on

$$B(A - \lambda I) = B_0A + \left(\sum_{i=1}^{n-1} \lambda^i(B_iA - B_{i-1})\right) - \lambda^n B_{n-1}.$$

Toisaalta lauseen 1.5 perusteella $(A - \lambda I)B = B(A - \lambda I) = \det(A - \lambda I)I$, joten

$$(A - \lambda I)B = \det(A - \lambda I)I = p(\lambda)I = (-1)^n\lambda^n I + \ldots + a_1\lambda I + a_0 I.$$

Koska $(A - \lambda I)B = B(A - \lambda I)$, saadaan yhtälöryhmä

$$\begin{cases} B_0A = a_0I, \\ B_iA - B_{i-1} = a_iI \quad \text{kaikilla } i = 1, \ldots, n-1, \\ -B_{n-1} = (-1)^n I. \end{cases}$$

Kertomalla keskimmäiset yhtälöt puolittain matriisin potenssilla $A^i$, saadaan

$$B_iA^{i+1} - B_{i-1}A^i = a_iA^i \quad \text{kaikilla } i = 1, \ldots, n-1.$$

Kertomalla vielä viimeinen yhtälö puolittain matriisin potenssilla $A^n$, saadaan

$$-B_{n-1}A^n = (-1)^n A^n.$$

Yhtälöryhmäksi saadaan, että

$$\begin{cases} B_0A = a_0I, \\ B_iA^{i+1} - B_{i-1}A^i = a_iA^i, \quad \text{kaikilla } i = 1, \ldots, n-1, \\ -B_{n-1}A^n = (-1)^n A^n. \end{cases}$$

Laskemalla yhtälöryhmän vasemmat puolet yhteen saadaan, että

$$B_0A+B_1A^2-B_0A+B_2A^3-B_1A^2+\ldots+B_{n-1}A^n-B_{n-1}A^{n-1}-B_{n-1}A^n$$
$$= (B_0A - B_0A) + (B_1A^2 - B_1A^2) + \ldots + (B_{n-1}A^n - B_{n-1}A^n) = 0.$$

Laskemalla yhtälöryhmän oikeat puolet yhteen saadaan

$$a_0I + a_1A + \ldots + a_{n-1}A^{n-1} + (-1)^nA^n = p(A).$$

Tällöin $p(A) = 0$. □

## 2.4. Matriisien similaarisuus

Eräs oleellinen käsite neliömatriiseihin liittyen on *similaarisuus.* Matriisit $A$ ja $B$ ovat keskenään *similaarisia*, mikäli on olemassa sellainen kääntyvä matriisi $B_{n\times n}$, että

$$B = P^{-1}AP.$$

Similaarisia matriiseita vastaa sama lineaarikuvaus. Matriisi $P$ on niin sanottu *kannanvaihtomatriisi* luonnollisesta kannasta johonkin toiseen kantaan.

Lause 2.6. Similaarisilla matriiseilla $A$ ja $B$ on sama karakteristinen polynomi.

Todistus. Olkoon $\lambda \in \mathbb{C}$. Koska $B = P^{-1}AP$ jollakin kääntyvällä matriisilla $P$, voidaan kirjoittaa

$$\begin{aligned}
B - \lambda I &= P^{-1}AP - \lambda I \\
&= P^{-1}AP - \lambda P^{-1}P \\
&= P^{-1}AP - P^{-1}\lambda P \\
&= P^{-1}AP - P^{-1}\lambda IP \\
&= P^{-1}(A - \lambda I)P.
\end{aligned}$$

Determinantin ominaisuuksien avulla saadaan nyt

$$\begin{aligned}
\det(B - \lambda I) &= \det(P^{-1}(A - \lambda I)P) \\
&= \det P^{-1} \det(A - \lambda I) \det P \\
&= \det P^{-1} \det P \det(A - \lambda I) \\
&= \det(P^{-1}P) \det(A - \lambda I) \\
&= \det I \det(A - \lambda I) \\
&= \det(A - \lambda I).
\end{aligned}$$

□

Koska similaaristen matriisien karakteristiset polynomit ovat samat, niillä on tietenkin myös samat ominaisarvot.

LUKU 3

# Hessenbergin matriisi

Matriisien ominaisarvojen selvittäminen voi joskus olla laskennallisesti työlästä varsinkin isojen matriisien kohdalla. Kuten aikaisemmin todettiin, similaarisilla matriiseilla on samat ominaisarvot. Ominaisarvojen selvittämistä voidaan helpottaa etsimällä jokin similaarinen, mutta laskennallisesti yksinkertaisempi matriisi kuin alkuperäinen. Diagonaalimatriisi on yksinkertaisessa muodossa, koska ominaisarvot saadaan luettua suoraan diagonaalilta. Matriisi $A$ on *diagonalisoituva*, jos se on similaarinen jonkin diagonaalimatriisin kanssa. Käytännössä tämä tarkoittaa, että on olemassa sellainen kääntyvä matriisi $C$ ja diagonaalimatriisi $D$, että

$$C^{-1}AC = D.$$

Koska matriisit $A$ ja $D$ ovat similaarisia keskenään, niillä on samat ominaisarvot. Kun muodostettaisiin aluksi sopiva kannanvaihtomatriisi $C$, saataisiin matriisin $A$ ominaisarvot luettua suoraan matriisin $C^{-1}AC = D$ diagonaalilta. Kaikki matriisit eivät kuitenkaan diagonalisoituvia.

Miten yksinkertaisessa muodossa jokin täysin mielivaltainen matriisi voidaan esittää, jotta se säilyisi similaarisena alkuperäisen kanssa? Eräs ratkaisu tähän on *Hessenbergin matriisi*.

Hessenbergin matriisia on kahta tyyppiä. Neliömatriisi $H_{n\times n} = [h_{ij}]$ on *Hessenbergin ylämuotoa* (engl. upper Hessenberg form), jos sen alkioille pätee

$$h_{ij} = 0, \quad \text{kun } i > j+1.$$

Tällöin matriisi on muotoa

$$H = \begin{bmatrix} h_{11} & \dots & \dots & \dots & h_{1n} \\ h_{21} & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{n,n-1} & h_{nn} \end{bmatrix} \tag{3}$$

Neliömatriisi $H_{n\times n} = [h_{ij}]$ on *Hessenbergin alamuotoa* (engl. lower Hessenberg form), jos sen alkioille pätee

$$h_{ij} = 0, \quad \text{kun } j > i+1.$$

Tällöin matriisi on muotoa

$$H = \begin{bmatrix} h_{11} & h_{12} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & h_{n-1,n} \\ h_{n1} & \dots & \cdots & \cdots & h_{nn} \end{bmatrix} \tag{4}$$

Hessenbergin muodossa oleva matriisi on melkein kolmiomatriisi. Hessenbergin alamuodossa oleva matriisi voi poiketa alakolmiomatriisista ainoastaan ensimmäisen ylemmän sivudiagonaalin (engl. superdiagonal) $h_{12}, \dots, h_{n-1,n}$ kohdalla (alakolmiomatriisissa nämä alkiot olisivat kaikki nollia). Vastaavasti ylämuodossa oleva Hessenbergin matriisi voi poiketa yläkolmiomatriisista ainoastaan ensimmäisen alemman sivudiagonaalin (engl. subdiagonal) $h_{21}, \dots, h_{n,n-1}$ kohdalla (yläkolmiomatriisissa nämä alkiot olisivat kaikki nollia).

## 3.1. Muunnos Hessenbergin muotoon similaarimuunnoksilla

**3.1.1. Matriisin muuntaminen Hessenbergin ylämuotoon similaarimuunnoksilla.** Tarkastellaan johdatteluna, kuinka $4 \times 4$-matriisi

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

voidaan muuntaa Hessenbergin ylämuotoon. Oletetaan aluksi, että $a_{21} \neq 0$. Muodostetaan alkeismatriiseja (katso 1.3.5) hyödyntäen matriisi $S_1$ asettamalla

$$A_{23}\left(-\frac{a_{31}}{a_{21}}\right) A_{24}\left(-\frac{a_{41}}{a_{21}}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{a_{31}}{a_{21}} & 1 & 0 \\ 0 & -\frac{a_{41}}{a_{21}} & 0 & 1 \end{bmatrix} =: S_1.$$

Edellisten käänteisoperaatioilla muodostetaan matriisi $S_1^{-1}$ asettamalla

$$A_{23}\left(\frac{a_{31}}{a_{21}}\right) A_{24}\left(\frac{a_{41}}{a_{21}}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{a_{31}}{a_{21}} & 1 & 0 \\ 0 & \frac{a_{41}}{a_{21}} & 0 & 1 \end{bmatrix} =: S_1^{-1}.$$

Nähdään, että $S_1S_1^{-1} = I$, joten $S_1$ ja $S_1^{-1}$ ovat toistensa käänteismatriiseja. Matriisit $B := S_1AS_1^{-1}$ ja $A$ ovat selvästi similaarisia. Laskemalla matriisien kertolaskut nähdään, että matriisi $B$ on muotoa

$$B = S_1AS_1^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix}. \tag{5}$$

Matriisi $B$ ei ole vielä Hessenbergin matriisi, jos $b_{42} \neq 0$, joten sitä pitää vielä muokata. Oletetaan aluksi, että $b_{32} \neq 0$. Muodostetaan alkeismatriiseja hyödyntäen matriisi $S_2$ asettamalla

$$A_{34}\left(-\frac{b_{42}}{b_{32}}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\dfrac{b_{42}}{b_{32}} & 1 \end{bmatrix} =: S_2$$

Edellisen käänteisoperaatioilla muodostetaan matriisi $S_2^{-1}$ asettamalla

$$A_{34}\left(\frac{b_{42}}{b_{32}}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \dfrac{b_{42}}{b_{32}} & 1 \end{bmatrix} =: S_2^{-1}$$

Nyt $S_2S_2^{-1} = I$, joten $S_2$ ja $S_2^{-1}$ ovat toistensa käänteismatriiseja. Matriisit $H := S_2BS_2^{-1}$ ja $B$ ovat similaarisia. Laskemalla kertolaskut nähdään, että matriisi $H$ on muotoa

$$H = S_2BS_2^{-1} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ 0 & h_{32} & h_{33} & h_{34} \\ 0 & 0 & h_{43} & h_{44} \end{bmatrix}.$$

Matriisille $A$ on nyt löydetty Hessenbergin ylämuodossa oleva similaarinen matriisi $H$.

Jos matriisille $A$ päteekin $a_{21} = 0$, on matriisia muokattava erilaiseen muotoon. Tarkalleen ottaen halutaankin jokin matriisi $\tilde{A}$, joka on similaarinen matriisin $A$ kanssa ja jolle $\tilde{a}_{21} \neq 0$. Jos $a_{31} \neq 0$, vaihtamalla alkioiden $a_{21}$ ja $a_{31}$ paikkaa keskenään ongelma olisi ratkaistu. Hyödynnetään tässä alkeismatriiseja asettamalla

$$P_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} =: P_1$$

Nyt $P_1P_1 = I$, jolloin $P_1$ on itsensä käänteismatriisi. Erityisesti matriisi $\tilde{A} := P_1AP_1$ on similaarinen matriisin $A$ kanssa ja se on muotoa

$$\tilde{A} = \begin{bmatrix} a_{11} & a_{13} & a_{12} & a_{14} \\ a_{31} & a_{33} & a_{32} & a_{34} \\ a_{22} & a_{23} & a_{22} & a_{24} \\ a_{41} & a_{43} & a_{42} & a_{44} \end{bmatrix}.$$

Nyt $\tilde{a}_{21} = a_{31} \neq 0$, joten Hessenbergin matriisia voidaan muodostaa aiemmin esitetyllä tavalla.

Jos matriisille $A$ pätee myös $a_{31} = 0$, asetetaan

$$P_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} =: P_2$$

jolle pätee $P_2P_2 = I$. Matriisi $\tilde{A} = P_2AP_2$ on similaarinen matriisin $A$ kanssa ja se on muotoa

$$\tilde{A} = \begin{bmatrix} a_{11} & a_{14} & a_{13} & a_{12} \\ a_{41} & a_{44} & a_{43} & a_{42} \\ a_{31} & a_{34} & a_{33} & a_{32} \\ a_{21} & a_{24} & a_{23} & a_{22} \end{bmatrix}.$$

Koska $\tilde{a}_{31} = a_{31} = 0$ ja $\tilde{a}_{41} = a_{21} = 0$, matriisilla on ensimmäisessä sarakkeessa jo halutuissa alkioissa nollat Hessenbergin muodon kannalta, joten voidaan siirtyä käsittelemään toista saraketta.

Jos matriisille $A$ pätee jo alun perin $a_{21} = a_{31} = a_{41} = 0$, sillä on ensimmäisessä sarakkeessa jo sellaisenaan halutuissa alkioissa nollat, jolloin voidaan suoraan lähteä muokkaamaan toista saraketta.

Oletetaan, että matriisilla $B$ on ensimmäinen sarake halutussa muodossa Hessenbergin muodon kannalta (5), jolloin voidaan lähteä muokkaamaan toista saraketta. Jos $b_{42} = 0$, matriisi on jo valmiiksi Hessenbergin muodossa, jolloin sille ei tarvitse tehdä enää mitään. Jos $b_{32} = 0$ ja $b_{42} \neq 0$, matriisia joudutaan muokkaamaan. Asetetaan

$$P_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} =: P_3$$

jolloin pätee $P_3P_3 = I$. Erityisesti $\tilde{B} = P_3BP_3$ on similaarinen matriisin $B$ kanssa ja on muotoa

$$\tilde{B} = \begin{bmatrix} b_{11} & b_{12} & b_{14} & b_{13} \\ b_{21} & b_{22} & b_{24} & b_{23} \\ 0 & b_{42} & b_{44} & b_{43} \\ 0 & b_{32} & b_{34} & b_{33} \end{bmatrix}.$$

Nyt $\tilde{b}_{42} = b_{32} = 0$, jolloin $\tilde{B}$ on Hessenbergin muodossa.

Tarkastellaan seuraavaksi yleistä tapausta, jossa matriisille $A_{n\times n}$ etsitään sellainen Hessenbergin muoto, jossa nolla-alkiot löytyvät lävistäjän alapuolelta.

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

Olkoon $n \geq 3$. Aletaan järjestyksessä "nollata"kustakin sarakkeesta halutut alkiot. Täten on ensin löydettävä matriisin $A$ kanssa similaarinen matriisi $B_1$, jonka ensimmäisen sarakkeen $n-2$ viimeistä alkiota ovat nollia. Oletetaan, että $a_{21} \neq 0$.

Muodostetaan alkeismatriiseja hyödyntäen matriisi $S_1$ asettamalla

$$A_{23}\left(-\frac{a_{31}}{a_{21}}\right) \cdots A_{2n}\left(-\frac{a_{n1}}{a_{21}}\right) = \prod_{i=3}^{n} A_{2i}\left(-\frac{a_{i1}}{a_{21}}\right) := S_1.$$

Edellisten käänteisoperaatioilla muodostetaan matriisi $S_1^{-1}$ asettamalla

$$A_{23}\left(\frac{a_{31}}{a_{21}}\right) \cdots A_{2n}\left(\frac{a_{n1}}{a_{21}}\right) = \prod_{i=3}^{n} A_{2i}\left(\frac{a_{i1}}{a_{21}}\right) := S_1^{-1}.$$

Koska $S_1 S_1^{-1} = I$, ovat $S_1$ ja $S_1^{-1}$ toistensa käänteismatriisit. Näin ollen matriisit $B_1 := S_1 A S_1^{-1}$ ja $A$ ovat similaarisia keskenään. Laskemalla matriisien kertolaskut havaitaan, että $B_1$ on muotoa

$$B_1 = \begin{bmatrix} b_{11} & b_{12} & \dots & \dots & b_{1n} \\ b_{21} & b_{22} & & & \vdots \\ 0 & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ 0 & b_{11} & \dots & \dots & b_{11} \end{bmatrix}$$

Matriisilla $B$ on tällöin ensimmäisessä sarakkeessa halutuissa alkioissa nollat.

Seuraavaksi nollataan toisesta sarakkeesta haluttavat alkiot. Halutaan löytää matriisin $B_1$ kanssa similaarinen matriisi $B_2$, jonka toisen sarakkeen $n-3$ viimeistä alkiota ovat nollia. Oletetaan, että $b_{32} \neq 0$. Muodostetaan matriisi $S_2$ asettamalla

$$\prod_{i=4}^{n} A_{3i}\left(-\frac{b_{i2}}{b_{32}}\right) =: S_2$$

ja matriisi $S_2^{-1}$ edellisten käänteisoperaatiolla asettamalla

$$\prod_{i=4}^{n} A_{3i}\left(\frac{b_{i2}}{b_{32}}\right) =: S_2^{-1}$$

Koska $S_2 S_2^{-1} = I$, ovat $S_2$ ja $S_2^{-1}$ toistensa käänteismatriisit. Näin ollen matriisit $B_2 := S_2 B_1 S_2^{-1}$ ja $B$ ovat similaarisia keskenään. Laskemalla matriisien kertolaskut havaitaan, että $B_2$ on muotoa

$$B_2 = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \dots & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & & & \vdots \\ 0 & c_{32} & c_{33} & & & \vdots \\ \vdots & 0 & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots \\ 0 & 0 & c_{n3} & \dots & \dots & c_{nn} \end{bmatrix}$$

Nyt matriisilla $C$ on ensimmäisessä ja toisessa sarakkeessa halutuissa alkioissa nollat. Samalla periaatteella käydään läpi loput sarakkeet niin, että matriisi on saatettu Hessenbergin muotoon.

Yleisesti kirjoitettuna muokattaessa saraketta $j$ muodostetaan matriisi $S_j$ $(a_{j+1,j} \neq 0)$ asettamalla

$$\prod_{i=j+2}^{n} A_{j+1,i}\left(-\frac{a_{ij}}{a_{j+1,j}}\right) := S_j$$

ja matriisi $S_j^{-1}$ edellisten käänteisoperaatioilla:

$$\prod_{i=j+2}^{n} A_{j+1,i}\left(\frac{a_{ij}}{a_{j+1,j}}\right) := S_j^{-1}.$$

Jos $a_{j+1,j} = 0$, korjataan asia samalla periaatteella, kuin $4 \times 4$-matriisin tapauksessakin. Valitaan jokin sarakkeen $j$ lopuista alkioista $a_{j+2,j}, \dots, a_{nj}$, jolle pätee $a_{ij} \neq 0$. Jos alkio $a_{kj} \neq 0$, asetetaan $P_{j+1,k} =: P$. Koska $PP = I$, matriisi $A$ on similaarinen matriisin $\tilde{A} = PAP$ kanssa. Jos

$$a_{ij} = 0 \quad \text{kaikilla } i = j+2, \dots, n,$$

matriisilla on kyseisessä sarakkeessa jo valmiiksi halutuissa alkioissa nollat ja voidaan siirtyä muokkaamaan seuraavaa saraketta.

Matriisilla $S_j B_{j-1} S_j^{-1}$ on sarakkeissa $1, \dots, j$ nollattu tarvittavat alkiot (tässä $B_0 = A$). Toistetaan algoritmia aina $(n-2)$. sarakkeeseen asti, jonka jälkeen matriisi on Hessenbergin muodossa. Algoritmissa on oleellista, että matriisit $A, B_1, \dots, B_{n-2}$ ovat keskenään similaarisia, jolloin niillä on samat ominaisarvot.

Jokainen matriisi $B_j$ on similaarinen matriisin $A$ kanssa. Jos sarakkeella $j$ on jo valmiiksi halutuissa alkioissa nollat, asetetaan yksinkertaisesti $S_j = I$. Koska $B_j = S_j B_{j-1} S_j^{-1}$ kaikilla $j = 1, \dots, n-2$ (tässä $B_0 = A$), tällöin

$$S_{n-2} \cdots S_1 A S_1^{-1} \cdots S_{n-2}^{-1} = H,$$

jossa $H$ on Hessenbergin ylämuodossa. Merkitään $S_{n-2} \cdots S_1 =: S$, jolloin $S_1^{-1} \cdots S_{n-2}^{-1} = S^{-1}$. Nyt tiedetään, että jokaiselle matriisille

$A_{n\times n}$ löytyy jokin sellainen kääntyvä matriisi $S$, että

$$SAS^{-1} = H.$$

Erityisesti jokainen neliömatriisi on muunnettavissa Hessenbergin ylämuotoon.

Esimerkki 3.1. Etsitään similaarimuunnoksien avulla Hessenbergin ylämuoto matriisille

$$A = \begin{bmatrix} 3 & 2 & 3 & 1 \\ 1 & 5 & 6 & -5 \\ -1 & 0 & -2 & 4 \\ 2 & 5 & 12 & -10 \end{bmatrix}.$$

Aikaisemmin esitetyllä tavalla valitaan

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix} \quad \text{ja} \quad S_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}.$$

Tällöin

$$B_1 = S_1AS_1^{-1} = \begin{bmatrix} 3 & 1 & 3 & 1 \\ 1 & -11 & 6 & -5 \\ 0 & -1 & 4 & -1 \\ 0 & -5 & 0 & 0 \end{bmatrix}.$$

Matriisilla $B_1$ on ensimmäisessä sarakkeessa halutuissa alkioissa nollat Hessenbergin muodon kannalta, joten lähdetään muokkaamaan toista saraketta. Valitaan matriisit

$$S_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -5 & 1 \end{bmatrix} \quad \text{ja} \quad S_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 5 & 1 \end{bmatrix}.$$

Tällöin

$$H = S_2B_1S_2^{-1} = \begin{bmatrix} 3 & 1 & 8 & 1 \\ 1 & -11 & -19 & -5 \\ 0 & -1 & -1 & -1 \\ 0 & 0 & 5 & 5 \end{bmatrix},$$

missä $H$ on Hessenbergin matriisi.

Esimerkki 3.2. Etsitään similaarimuunnoksilla Hessenbergin ylämuoto matriisille

$$A = \begin{bmatrix} 6 & 0 & -1 & 1 \\ 0 & 3 & 1 & -1 \\ -1 & 1 & 4 & 1 \\ 1 & -1 & -1 & 8 \end{bmatrix}.$$

Ensimmäisessä sarakkeessa on nolla "väärässä" kohdassa, koska $a_{21} = 0$. Asia voidaan korjata vaihtamalla 2. ja 3. rivin ensimmäiset alkiot keskenään (säilyttämällä uusi matriisi kuitenkin similaarisena alkuperäisen kanssa.) Asetetaan

$$P_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} =: P.$$

Tällöin $PP = I$, joten $P$ on itsensä käänteismatriisi. Matriisi

$$\tilde{A} = PAP = \begin{bmatrix} 6 & -1 & 0 & 1 \\ -1 & 4 & 1 & 1 \\ 0 & 1 & 3 & -1 \\ 1 & 1 & -1 & 8 \end{bmatrix}$$

on tällöin similaarinen matriisin $A$ kanssa. Aikaisemmin esitetyllä tavalla valitaan nyt matriisit

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{ja} \quad S_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}.$$

Nyt matriisi

$$B_1 = S_1 \tilde{A} S_1^{-1} = \begin{bmatrix} 6 & -2 & 0 & 1 \\ -1 & 3 & 1 & 1 \\ 0 & 2 & 3 & -1 \\ 0 & -4 & 0 & 9 \end{bmatrix}$$

on similaarinen matriisin $\tilde{A}$ kanssa. Edelleen aikaisemmin esitetyllä tavalla saadaan matriisit

$$S_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix} \quad \text{ja} \quad S_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix}.$$

Matriisi

$$H = S_2 B_1 S_2^{-1} = \begin{bmatrix} 6 & -2 & -2 & 1 \\ -1 & 3 & -1 & 1 \\ 0 & 2 & 5 & -1 \\ 0 & 0 & -8 & 7 \end{bmatrix}$$

on Hessenbergin muodossa ja se on similaarinen alkuperäisen matriisin $A$ kanssa.

**3.1.2. Matriisin muuntaminen Hessenbergin alamuotoon similaarimuunnoksilla.** Matriisi $A_{n \times n}$ voidaan vastaavasti muokata myös Hessenbergin matriisin alamuotoon (4). Tässä tapauksessa aletaan nollata sarakkeista tarvittavat alkiot oikealta lähtien aloittaen $n$. sarakkeesta. Olkoon $n \geq 3$. Hessenbergin alamuotoa varten matriisilta

$A_{n\times n}$ halutaan nollata sarakkeesta $j$ ensimmäiset $j-2$ alkiota. Yleisesti kirjoitettuna sarakkeella $j$ muodostetaan matriisi $S_j$ $(a_{j-1,j} \neq 0)$ asettamalla

$$\prod_{i=1}^{j-2} A_{j-1,i}\left(-\frac{a_{ij}}{a_{j-1,j}}\right) = S_j$$

ja matriisiksi $S_j^{-1}$ edellisten käänteisoperaatioilla

$$\prod_{i=1}^{j-2} A_{j-1,i}\left(\frac{a_{ij}}{a_{j-1,j}}\right) = S_j^{-1}.$$

Tällöin $S_j$ ja $S_j^{-1}$ ovat toistensa käänteismatriiseja, koska $S_jS_j^{-1} = I$. Matriisilta $A_{n\times n}$ nollataan ensin halutut alkiot $n$. sarakkeesta, jolloin uusi matriisi on muotoa $S_nAS_n^{-1} =: B_n$. Seuraavaksi matriisilta $B_n$ nollataan halutut alkiot $(n-1)$. sarakkeesta. Halutut alkiot nollataan aina 3. sarakkeeseen asti. Jokainen matriisi $S_jB_{j+1}S_j^{-1} =: B_j$ kaikilla $j = 3,\ldots,n$ (tässä $B_{n+1} = A$) on similaarinen matriisin $A$ kanssa ja sillä ovat $j-2$ ensimmäistä alkiota nollia. Esimerkiksi $5\times 5$-matriisin tapauksessa 5. saraketta muokattaessa matriisiksi $S_5$ valitaan

$$A_{41}\left(\frac{a_{15}}{a_{45}}\right)A_{42}\left(\frac{a_{25}}{a_{45}}\right)A_{43}\left(\frac{a_{35}}{a_{45}}\right) = S_5,$$

ja matriisiksi $S_5^{-1}$ valitaan

$$A_{41}\left(-\frac{a_{15}}{a_{45}}\right)A_{42}\left(-\frac{a_{25}}{a_{45}}\right)A_{43}\left(-\frac{a_{35}}{a_{45}}\right) = S_5^{-1}.$$

Uusi matriisi on muotoa

$$S_5AS_5^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} & 0 \\ b_{21} & b_{22} & b_{23} & b_{24} & 0 \\ b_{31} & b_{32} & b_{33} & b_{34} & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} \end{bmatrix} = B_5.$$

Jos jotakin saraketta muokatessa onkin $a_{j-1,j} = 0$, nolla-alkion paikka kyseisessä sarakkeessa voidaan vaihtaa sopivalla alkeismatriisilla samalla periaatteella, kuin Hessenbergin matriisin ylämuodon tapauksessa.

Jokainen matriisi $B_j$ on similaarinen matriisin $A$ kanssa. Jos sarakkeella $j$ on jo valmiiksi halutuissa alkioissa nolllat, asetetaan yksinkertaisesti $S_j = I$. Koska $B_j = S_jB_{j-1}S_j^{-1}$ kaikilla $j = 3,\ldots,n$ (tässä $B_{n+1} = A$), tällöin

$$S_3\cdots S_nAS_n^{-1}\cdots S_3^{-1} = H,$$

jossa $H$ on Hessenbergin alamuodossa. Merkitään $S_3\cdots S_n =: S$, jolloin $S_n^{-1}\cdots S_3^{-1} = S^{-1}$. Nyt tiedetään, että jokaiselle matriisille $A_{n\times n}$ löytyy jokin sellainen kääntyvä matriisi $S$, että

$$SAS^{-1} = H.$$

Erityisesti jokainen neliömatriisi on muunnettavissa Hessenbergin alamuotoon.

ESIMERKKI 3.3. Muokataan matriisi

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

Hessenbergin alamuotoon similaarimuunnoksilla.

Valitaan ensin matriisit

$$S_1 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{ja} \quad S_1^{-1} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

Nyt

$$S_1 H S_1^{-1} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -3 & 1 \end{bmatrix} := H.$$

Matriisilla $H$ on viimeisessä sarakkeessa halutuissa alkioissa nollat Hessenbergin muodon kannalta ja lisäksi $h_{13} = 0$, joten se on itse asiassa jo nyt Hessenbergin muodossa.

## 3.2. Hessenbergin matriisin käyttökelpoisuus

Tässä luvussa on käyty tapoja erilaisten neliömatriisien muokkaamiseksi sekä Hessenbergin ylämuotoon että alamuotoon. On huomattu, että itse asiassa jokainen neliömatriisi voidaan muuntaa sekä Hessenbergin ylämuotoon että alamuotoon. Seuraavaksi voidaan esittää tämän tutkielman tärkein lause.

LAUSE 3.1. Jokaiselle matriisille $A_{n\times n}$ löytyy kääntyvä matriisi $S$ siten, että

$$SAS^{-1} = H,$$

jossa $H$ on Hessenbergin muodossa matriisi.

TODISTUS. Tässä luvussa on käsitelty kaikentyyppisille neliömatriiseille menetelmä Hessenbergin muotoon muuntamiseksi. □

## 3.3. Householderin muunnos

Tutustutaan toiseen tapaan matriisin Hessenbergin muodon löytämiseksi, Householderin muunnokseen. Oletetaan, että $u \in \mathbb{C}^n$ $(u \neq 0)$.
*Householderin matriisi* (tai *Householderin muunnos*) määritellään asettamalla

$$Q = I - \frac{2uu^*}{u^*u}.$$

LAUSE 3.2. Householderin matriisi on itseadjungoituva ja unitaarinen.

TODISTUS. Osoitetaan ensin itseadjungoituvuus. Määritelmän mukaan Householderin matriisi $Q_{n\times n}$ on itseadjungoituva, kun $Q = Q^*$. Merkitään lyhyemmin $b = \dfrac{2}{u^*u} \in \mathbb{R}$, jolloin

$$Q = I - buu^*.$$

Nyt

$$\begin{aligned} Q^* &= (I - buu^*)^* = I^* - (buu^*)^* \\ &= I - b\,(uu^*)^* = I - b(u^*)^*u^* \\ &= I - buu^* \\ &= Q. \end{aligned}$$

Osoitetaan vielä unitaarisuus. Määritelmän mukaan Householderin matriisi $Q_{n\times n}$ on unitaarinen, kun $Q^* = Q^{-1}$. Itseadjungoituvuuden nojalla $Q = Q^*$, joten unitaarisuus voidaan todeta osoittamalla, että $Q = Q^{-1}$. Tällöin $QQ = I$. Merkitään aluksi lyhyemmin $b = \dfrac{2}{u^*u} \in \mathbb{R}$. Nyt

$$\begin{aligned} QQ &= (I - buu^*)\,(I - buu^*) = I^2 - 2Ibuu^* + (buu^*)\,(buu^*) \\ &= I - 2buu^* + b^2\,(uu^*)\,(uu^*) \\ &= I - 2buu^* + b^2u\,\underbrace{(u^*u)}_{\in\mathbb{R}}\,u^* \\ &= I - 2buu^* + b^2(u^*u)uu^* \\ &= I - 2\frac{2}{u^*u}uu^* + \left(\frac{2}{u^*u}\right)^2(u^*u)uu^* \\ &= I - \frac{4uu^*}{u^*u} + \frac{4(u^*u)(uu^*)}{(u^*u)(u^*u)} \\ &= I - \frac{4uu^*}{u^*u} + \frac{4uu^*}{u^*u} \\ &= I. \end{aligned}$$

□

Lauseen 3.2 perusteella Householderin matriisille $Q$ pätee

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

Matriisi $Q$ onkin samaan aikaan sekä itsensä adjungaatti että käänteismatriisi.

Kun matriisille $A_{n\times n}$ etsitään Hessenbergin muotoa, lähdetään ensimmäisestä sarakkeesta lähtien nollaamaan riittävä määrä sarakevektorin komponentteja lopusta. Lähdetään tarkastelemaan, miten Householderin matriisi on muodostettava, jotta sitä voitaisiin hyödyntää. Tarkemmin halutaankin tietää, miten vektori $u \in \mathbb{C}^n$ valitaan.

Olkoon matriisilla $A$ sarakevektorit $\bar{a}_1, \ldots, \bar{a}_n$, jolloin

$$A = \begin{bmatrix} \bar{a}_1 & \ldots & \bar{a}_n \end{bmatrix}.$$

Kun matriisia $A$ kerrotaan vasemmalta puolelta matriisilla $Q$, tulomatriisin sarakevektorit ovat muotoa

$$QA = \begin{bmatrix} Q\bar{a}_1 & \cdots & Q\bar{a}_n \end{bmatrix}.$$

Oletetaan, että sarakevektorilta

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

haluttaisiin nollata $n-k$ viimeistä komponenttia kertomalla sitä Householderin matriisilla $Q$. Tutkitaan, miten päästään tulokseen $Qx = w$, jossa $w$ olisi tarkalleen muotoa

$$w = \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \\ s \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{C}^n.$$

Kun Householderin matriisilla $Q$ kerrotaan vektoria $v \in \mathbb{C}^n$, saadaan

$$Qv = \left( I - \frac{2uu^*}{u^*u} \right) v = v - \frac{2uu^*v}{u^*u}.$$

Merkitään lyhyemmin $\dfrac{2u^*v}{u^*u} = \gamma$, jolloin

$$Qv = v - \gamma u.$$

Jos haluttaisiin päätyä tulokseen $Qv = w$, olisi oltava

$$Qv = v - \gamma u = \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \\ v_k \\ \vdots \\ v_n \end{bmatrix} - \gamma \begin{bmatrix} u_1 \\ \vdots \\ u_{k-1} \\ u_k \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \\ s \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Huomataan, että edelliseen tulokseen voidaan päätyä, kun seuraavat neljä ehtoa ovat voimassa:

(1) $u_i = 0$, kun $i = 1, \ldots, k-1$,
(2) $\gamma = 1$,
(3) $v_k - u_k = s$,

(4) $u_i = v_i$, kun $i = k+1, \ldots, n$.

Householderin matriisissa esiintyvän vektorin $u \in \mathbb{C}$ on oltava muotoa

$$u = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ v_k - s \\ v_{k+1} \\ \vdots \\ v_n \end{bmatrix}.$$

Vielä on määritettävä $s$ sellaiseksi, että $\gamma = 1$. Asetetaan

$$s = \frac{v_k}{|v_k|}\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2}.$$

Koska $|v_k| \in \mathbb{R}$ ja $\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2} \in \mathbb{R}$, pätee selvästi

$$\overline{s} = \frac{\overline{v_k}}{|v_k|}\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2}.$$

Tällöin saadaan

$$\begin{aligned} |s|^2 = s\overline{s} &= \frac{v_k\overline{v_k}}{|v_k|^2}\left(|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2\right) \\ &= \frac{|v_k|^2}{|v_k|^2}\left(|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2\right) \\ &= |v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2 \in \mathbb{R}. \end{aligned}$$

Lisäksi

$$\begin{aligned} \overline{v_k}s &= \frac{\overline{v_k}v_k}{|v_k|}\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2} \\ &= \frac{|v_k|^2}{|v_k|}\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2} \\ &= |v_k|\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2}. \end{aligned}$$

Toisaalta

$$\begin{aligned} v_k\overline{s} &= \frac{v_k\overline{v_k}}{|v_k|}\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2} \\ &= \frac{|v_k|^2}{|v_k|}\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2} \\ &= |v_k|\sqrt{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2} \\ &= \overline{v_k}s. \end{aligned}$$

Edellisten perusteella saadaan

$$\begin{aligned}
u^*u &= \overline{(v_k - s)}(v_k - s) + |v_{k+1}|^2 + \ldots + |v_n|^2 \\
&= (\overline{v_k} - \overline{s})(v_k - s) + |v_{k+1}|^2 + \ldots + |v_n|^2 \\
&= |v_k|^2 - \underbrace{\overline{v_k}s}_{=v_k\overline{s}} - v_k\overline{s} + |s|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2 \\
&= \underbrace{|v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2}_{=|s|^2} + |s|^2 - v_k\overline{s} - v_k\overline{s} \\
&= 2|s|^2 - 2v_k\overline{s}.
\end{aligned}$$

Toisaalta

$$\begin{aligned}
u^*v &= \overline{(v_k - s)}v_k + |v_{k+1}|^2 + \ldots + |v_n|^2 \\
&= (\overline{v_k} - \overline{s})v_k + |v_{k+1}|^2 + \ldots + |v_n|^2 \\
&= \overline{v_k}v_k - \overline{s}v_k + |v_{k+1}|^2 + \ldots + |v_n|^2 \\
&= |v_k|^2 + |v_{k+1}|^2 + \ldots + |v_n|^2 - v_k\overline{s} \\
&= |s|^2 - v_k\overline{s}.
\end{aligned}$$

Tällöin

$$\gamma = \frac{2u^*v}{u^*u} = 1,$$

jolloin ehdossa (2) asetettu vaatimus komponentille $s \in \mathbb{C}$ pätee.

Jos Householderin matriisi $Q = [q_{ij}]$ on määritetty edellä käsitellyllä tavalla, matriisilla $QA = [b_{ij}]$ on halutuissa sarakkeissa halutuissa alkioissa nollat. Kun matriisille $A$ etsitään Hessenbergin muotoa, uuden matriisin on oltava similaarinen alkuperäisen matriisin kanssa. Koska Householderin matriisille pätee $Q = Q^{-1}$, matriisi $QAQ$ on similaarinen matriisin $A$ kanssa. Mutta onko matriisilla $QAQ$ myös halutuissa alkioissa nollat kuten matriisilla $QA$?

Hessenbergin matriisin ylämuodossa on ensimmäisessä sarakkeessa $n-2$ viimeistä alkiota nollia ja toisessa sarakkeessa $n-3$ viimeistä alkiota nollia. Yleisesti voidaan ilmaista, että kokoa $n \times n$ olevan Hessenbergin matriisin $k$. sarakkeen $n-k-1$ viimeistä alkiota ovat nollia. Muokattaessa saraketta $k$ on vektorilla $u \in \mathbb{C}^n$ nollat $k$ ensimmäisessä sarakkeessa. Havaitaan, että muodostettava Householderin matriisi lohkomuodossa ilmaistuna on tällöin

$$Q = \begin{bmatrix} I_k & 0 \\ 0 & U \end{bmatrix}, \tag{6}$$

jossa $I_k$ on kokoa $k \times k$ oleva yksikkömatriisi ja $U$ kokoa $(n-k)\times(n-k)$ oleva matriisi, jonka kaikki alkiot $u_{ij}$ voivat olla nollasta poikkeavia. Kun matriisin $A_{n\times n}$ ensimmäistä saraketta muokataan, matriisien $Q$ ja

$A$ tulo on muotoa

$$QA = \begin{bmatrix} b_{11} & b_{12} & \dots & \dots & b_{1n} \\ b_{21} & b_{22} & & & \vdots \\ 0 & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ 0 & b_{n2} & \dots & \dots & b_{nn} \end{bmatrix}.$$

Yleisesti kirjoittetuna matriisin $QAQ = [c_{ij}]$ ensimmäisen sarakkeen alkio $c_{i1}$ on matriisitulon määritelmän mukaan

$$c_{i1} = \sum_{j=1}^{n} b_{ij} q_{j1} \quad \text{kaikilla } i = 1, \dots, n.$$

Matriisille $QA$ pätee $b_{i1} = 0$ kaikilla $i = 3, \dots, n$. Toisaalta $q_{j1} = 0$ kaikilla $i = 2, \dots, n$. Täten

$$c_{i1} = 0 \quad \text{kaikilla } i = 3, \dots, n.$$

Näin ollen matriisi $QAQ$ on muotoa

$$QAQ = \begin{bmatrix} c_{11} & c_{12} & \dots & \dots & c_{1n} \\ c_{21} & c_{22} & & & c_{2n} \\ 0 & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ 0 & c_{n2} & \dots & \dots & c_{nn} \end{bmatrix},$$

jolloin silläkin on ensimmäisessä sarakkeessa halutuissa alkioissa nollat.

Tarkastellaan yleisempää tapausta, jossa matriisilla $A_{n \times n}$ on nollattu $k-1$ ensimmäisessä sarakkeessa halutut alkiot Hessenbergin muodon kannalta. Seuraavaksi matriisilta halutaan nollata sarakkeesta $k$ viimeiset $n-k-1$ alkiota. Muodostetaan aiemmin esitetyllä tavalla Householderin matriisi (6). Yleisesti kirjoitettuna matriisin $QAQ$ sarakkeen $k$ alkio $c_{i1}$ on matriisitulon määritelmän mukaan

$$c_{ik} = \sum_{j=1}^{n} b_{ij} q_{jk} \quad \text{kaikilla } i = 1, \dots, n.$$

Matriisille $QA$ pätee

$$b_{i1} = b_{i2} = \dots = b_{ik} = 0 \quad \text{kaikilla } i = k+2, \dots, n.$$

Toisaalta $q_{j1} = 0$ kaikilla $i = k+1, \dots, n$. Täten

$$c_{ik} = 0 \quad \text{kaikilla } i = k+2, \dots, n.$$

Tällöin matriisin $QAQ$ sarakkeen $k$ viimeiset $n-k-1$ alkiota ovat nollia. Havaitaan, että muokattaessa matriisia $A$ Hessenbergin muotoon matriiseilla $QA$ ja $QAQ$ on molemmilla halutuissa alkioissa nollat. Householderin muunnoksessakin lähdetään nollaamaan matriisilta sarake kerrallaan halutut alkiot. Householderin matriisilla $Q_1$ nollataan alkioita ensimmäisestä sarakkeesta, jolloin matriisi $Q_1AQ_1 =: B_1$ on similaarinen matriisin $A$ kanssa. Householderin matriisilla $Q_2$ nollataan alkioita toisesta sarakkeesta, jolloin $Q_2Q_1AQ_1Q_2 =: B_2$ on similaarinen matriisin $A$ kanssa. Nollataan tarvittavat alkiot aina $(n-2)$. sarakkeeseen asti. Yleisesti jokainen matriisi $B_j$ (jossa $i = 1, \dots, n-2$) on tällöin similaarinen matriisin $A$ kanssa. Jos sarakkeella $j$ on jo valmiiksi halutuissa alkioissa nollat, asetetaan yksinkertaisesti $Q_j = I$.

Esimerkki 3.4. Etsitään Hessenbergin ylämuoto Householderin muunnoksen avulla matriisille

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

Koska matriisi on $3 \times 3$-muotoa, se saadaan Hessenbergin ylämuotoon nollaamalla ensimmäisen sarakkeen viimeinen alkio. Merkitään nyt

$$v = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix},$$

jolloin

$$s = \sqrt{v_2^2 + v_3^2} = \sqrt{3^2 + 4^2} = 5.$$

tällöin $u_2 = v_2 - s = 3 - 5 = -2$, jolloin saadaan Householderin matriisissa esiintyväksi vektoriksi

$$u = \begin{bmatrix} 0 \\ -2 \\ 4 \end{bmatrix}.$$

Nyt $u^*u = 12$ ja

$$uu* = \begin{bmatrix} 0 \\ -2 \\ 4 \end{bmatrix} \begin{bmatrix} 0 & -2 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & -8 \\ 0 & -8 & 16 \end{bmatrix}.$$

Householderin matriisiksi saadaan tällöin

$$Q = I - \frac{2uu^*}{u^*u} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & \frac{4}{5} \\ 0 & \frac{4}{5} & -\frac{3}{5} \end{bmatrix}.$$

Nyt

$$QAQ = \begin{bmatrix} 1 & 5 & 0 \\ 5 & \frac{49}{25} & \frac{7}{25} \\ 0 & \frac{7}{25} & \frac{1}{25} \end{bmatrix},$$

mikä on Hessenbergin muodossa.

Esimerkki 3.5. Matriisilla

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

on Hessenbergin muodon kannalta ensimmäisessä sarakkeessa halutuissa sarakkeissa jo nollat. Jotta matriisi saadaan Hessenbergin muotoon, riittää nollata 2. sarakkeen viimeinen komponentti. Merkitään

$$v = \begin{bmatrix} 1 \\ 4 \\ -3 \\ 4 \end{bmatrix}.$$

Nyt

$$s = \sqrt{v_3^2 + v_4^2} = \sqrt{(-3)^2 + 4^2} = 5,$$

jolloin $u_3 = v_3 - s = -3 - 5 = -8$. Tällöin

$$u = \begin{bmatrix} 0 \\ 0 \\ -8 \\ 4 \end{bmatrix}.$$

Householderin matriisiksi saadaan nyt

$$Q = I - \frac{2uu^*}{u^*u} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{5} & \frac{4}{5} \\ 0 & 0 & \frac{4}{5} & \frac{3}{5} \end{bmatrix}.$$

Nyt

$$QAQ = \begin{bmatrix} 2 & 1 & 1 & 2 \\ 3 & 4 & \frac{4}{5} & \frac{3}{5} \\ 0 & 5 & \frac{21}{25} & \frac{47}{25} \\ 0 & 0 & \frac{22}{25} & \frac{79}{25} \end{bmatrix},$$

mikä on Hessenbergin muodossa.

LUKU 4

# Hessenbergin matriisin ominaisarvot

## 4.1. Krylovin menetelmä

Olkoon $H_{n\times n}$ Hessenbergin ylämuodossa oleva matriisi (3):

$$H = \begin{bmatrix} h_{11} & \dots & \dots & \dots & h_{1n} \\ h_{21} & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{n,n-1} & h_{nn} \end{bmatrix}$$

Hessenbergin ylämuodossa oleva matriisi $H_{n\times n} = [h_{ij}]$ on redusoitumaton, jos sen alkioille pätee

$$h_{k,k-1} \neq 0 \quad \text{kaikilla } 2, \dots, n.$$

Muulloin matriisi on redusoituva. Seuraava algoritmi ominaisarvojen selvittämiseksi toimii vain, jos matriisi on redusoitumaton. Krylovin menetelmässä määritetään matriisille $H$ yhteensä $n+1$ eri vektoria valitsemalla ensimmäiseksi jokin mielivaltainen vektori $x_0 \in \mathbb{C}^n$. Asetetaan tässä, että $x_0$ on luonnollinen kantavektori $\bar{e}_1$. Tämän jälkeen määritetään loput vektorit rekursiivisesti asettamallla

$$x_k = Hx_{k-1}, \quad \text{jossa}\, k = 1, \dots, n-1.$$

Koska

$$\begin{aligned} x_1 &= Hx_0, \\ x_2 &= Hx_1 = H(Hx_0) = H^2x_0, \\ x_3 &= Hx_2 = H(H^2x_0) = H^3x_0, \end{aligned}$$

voidaan tietenkin kirjoittaa myös yleisessä muodossa

$$x_k = H^k x_0.$$

Tällöin

$$x_0 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{ja} \quad x_1 = Hx_0 = \begin{bmatrix} h_{11} \\ h_{21} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Määrittämällä seuraavia vektoreita huomataan, että jokaisella uudella vektorilla on yksi mahdollisesti nollasta poikkeava komponentti enemmän.

LAUSE 4.1. Oletetaan, että $H_{n\times n}$ on redusoimaton Hessenbergin ylämuodossa oleva matriisi ja $x_k = H^k x_0$. Tällöin vektorilla $x_k$ ovat $n+1-k$ viimeistä komponenttia nollia ja $x_{k,k+1} \neq 0$.

TODISTUS. Käytetään induktiota.

(1) Kun $k = 1$,

$$x_1 = Hx_0 = \begin{bmatrix} h_{11} \\ h_{21} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Tässä $h_{21} \neq 0$, koska $H$ on redusoimaton.

(2) Induktio-oletuksena oletetaan, että väite pätee, kun $k = i$. Merkitään $x_i = (b_1, \ldots, b_n)$. Tällöin vektorilla $x_i$ ovat $n+1-i$ viimeistä komponenttia nollia ja $b_{i+1} \neq 0$.

(3) Osoitetaan, että väite pätee, kun $k = i+1$. Tiedetään, että matriisin $H$ rivivektorilla $\vec{h}_{i+2}$ ensimmäiset $i$ komponenttia ovat nollia ja $h_{i+2,i+1} \neq 0$ (redusoitumaton matriisi). Vektorin $x_{i+1} = Hx_i$ komponentti $x_{i+1,i+2}$ on

$$0 + \ldots + 0 + b_{i+1}h_{i+2,i+1} + 0 + \ldots + 0 = b_{i+1}h_{i+2,i+1}.$$

Induktio-oletuksen perusteella tiedetään, että erityisesti $b_{i+1}$ on nollasta poikkeava ja toisaalta redusoimattomuuden perusteella myös $a_{i+2,i+1}$ on nollasta poikkeava, jolloin komponentti $x_{i+2,i+2}$ on niin ikään nollasta poikkeava. Toisaalta matriisin $H$ rivivektoreilla $\vec{h}_{i+3}, \ldots, \vec{h}_n$ ainakin $i+1$ ensimmäistä alkiota ovat nollia, joten vektorilla $x_{i+1}$ viimeiset $n-i-1$ alkiota ovat nollia. Väite pätee myös tapauksessa $k = i+1$. □

Lauseen 4.1 perusteella tiedetäänkin nyt, että edellä määritetyt vektorit $x_0, \ldots, x_{n-1}$ ovat keskenään lineaarisesti riippumattomia.

LAUSE 4.2. Oletetaan, että $H_{n\times n}$ on redusoimaton Hessenbergin ylämuodossa oleva matriisi ja $x_k = H^k x_0$. Tällöin Krylovin menetelmällä määritetyt vektorit $x_0, \ldots, x_{n-1}$ muodostavat avaruuden $\mathbb{C}^n$ kannan.

TODISTUS. Vektoreita $x_0, \ldots, x_{n-1}$ on yhteensä $n$ kappaletta ja ne ovat kaikki lineaarisesti riippumattomia, joten ne muodostavat avaruuden $\mathbb{C}^n$ kannan. □

Cayleyn-Hamiltonin lauseen perusteella (2.5) matriisin $H_{n\times n}$ karakteristinen polynomi nollaa sen, toisin sanoen $p_H(H) = 0$. Karakteristisen polynomin $p_H(H)$ aste on $n$. Jos löydettäisiin jokin samaa astetta oleva polynomi, joka nollaisi matriisin $H$ ja se pystyttäisiin vielä osoittamaan yksikäsitteiseksi, karakteristinen polynomi olisi onnistuttu löytämään.

Koska vektorit $x_0, \ldots, x_{n-1}$ muodostavat avaruuden $\mathbb{C}^n$ kannan, jokainen vektori $y \in \mathbb{C}^n$ voidaan esittää niiden lineaarikombinaationa. Näin ollen löytyvät sellaiset yksikäsitteiset kompleksiluvut
$a_0, a_1, \ldots, a_{n-1}$, että

$$a_0 x_0 + a_1 x_1 + \ldots + a_{n-1} x_{n-1} = -x_n. \tag{7}$$

Koska

$$x_k = H^k \bar{e}_1 \quad \text{kaikilla } k = 1, 2, \ldots, n,$$

voidaan yhtälö (7) kirjoittaa myös muodossa

$$\begin{aligned} (8) \quad a_0 \bar{e}_1 + a_1 H \bar{e}_1 + \ldots a_{n-1} H^{n-1} \bar{e}_1 + H^n \bar{e}_1 \\ = \left(a_0 + a_1 H + \ldots a_{n-1} H^{n-1} + H^n\right) \bar{e}_1 = 0. \end{aligned}$$

Olkoon polynomi $\tilde{p}$ muotoa

$$\tilde{p}(t) = a_0 + a_1 t + \ldots + a_{n-1} t^{n-1} + t^n,$$

jolloin $\tilde{p}(H)\bar{e}_1 = 0$.

Vektoreita $\bar{e}_1, H\bar{e}_1, \ldots, H^n \bar{e}_1$ on yhteensä $n+1$ kappaletta, jolloin ne ovat lineaarisesti riippuvia keskenään. Yhtälöllä (8) voi näin ollen olla nollasta poikkeava ratkaisu. Selvästi

$$\begin{aligned} H^k \tilde{p}(H) &= H^k (a_0 I + a_1 H + \ldots + a_{n-1} H^{n-1} + H^n) \\ &= H^k a_0 I + H^k a_1 H + \ldots + H^k a_{n-1} H^{n-1} + H^{n+k} \\ &= a_0 H^k + a_1 H^{k+1} + \ldots + a_{n-1} H^{n+k-1} + H^{n+k} \\ &= (a_0 I + a_1 H + \ldots + a_{n-1} H^{n-1} + H^n) H_k \\ &= \tilde{p}(H) H^k. \end{aligned}$$

Koska $\tilde{p}(H)\bar{e}_1 = 0$, pätee tietenkin myös $H^k \tilde{p}(H)\bar{e}_1 = 0$, jolloin

$$\tilde{p}(H) H^k \bar{e}_1 = 0 \quad \text{kaikilla } k = 1, \ldots, n.$$

Koska $x_k = H^k \bar{e}_1$, tiedetään, että

$$\tilde{p}(H) x_k = 0 \quad \text{kaikilla } k = 1, \ldots, n.$$

Koska vektorit $x_0, \ldots, x_{n-1}$ ovat lineaarisesti riippumattomia, on välttämättä oltava $\tilde{p}(H) = 0$.

Polynomijoukko $X = \{p \in \mathcal{P} \mid p(H) = 0\}$ on polynomirenkaan $\mathcal{P}$ ideaali. Polynomirenkaan $\mathcal{P}$ kaikki ideaalit ovat yhden polynomin $p_X$ virittämiä, jolloin $X = \{q p_X \mid q \in \mathcal{P}\}$. (Katso Metsänkylä & Näätänen [**3**].) Koska $p_H \in X$, virittävän polynomin asteelle pätee $\deg p_X \leq \deg p_H = n$.

Olkoon polynomi $q \in X$ muotoa

$$q(t) = b_0 + b_1 t + \ldots + b_{k-1} t^{k-1} + b_k t^k$$

joillakin kompleksiluvuilla $b_0, \dots, b_k \in \mathbb{C}$, jossa $b_k \neq 0$. Koska $q(H) = 0$, pätee myös $q(H)\bar{e}_1 = 0$, jolloin

$$\sum_{i=0}^{k} b_i H^i \bar{e}_1 = 0.$$

Edellisen yhtälön vakioille $b_0, \dots, b_k$ voi löytyä ratkaisu vain, jos vektorit $\bar{e}_1, \dots, H^k\bar{e}_1$ ovat keskenään lineaarisesti riippuvia. Vektorit ovat lineaarisesti riippuvia keskenään, jos $n - 1 < k$, jolloin $k \geq n$. Polynomille $p_X$ on tällöin välttämättä oltava $\deg p_X \geq \deg p_H$.

Koska $\deg p_X \leq \deg p_H$ ja toisaalta $\deg p_X \geq \deg p_H$, tietenkin $\deg p_X = \deg p_H = n$. Jos virittävä polynomi $p_X$ on valittu normitetuksi, voidaan karakteristinen polynomi $p_H$ esittää muodossa

$$p_H = \alpha p_X,$$

jossa $\alpha = \pm 1$. Polynomi $\tilde{p}$ on normitettu ja se on samaa astetta virittävän polynomin kanssa, jolloin $\tilde{p} = p_X$. Näin ollen karakteristinen polynomi on muotoa

$$p_H = \alpha\tilde{p} = (-1)^n(\lambda^n + a_{n-1}\lambda^{n-1} + \dots a_1\lambda_1 + a_0).$$

Krylovin menetelmän avulla määritetty karakteristinen polynomi onkin etumerkkiä vaille yksikäsitteinen.

Edellä esitelty algoritmi tunnetaan Krylovin menetelmänä. Kerrataan vielä algoritmin vaiheet.

(1) Olkoon $x_0 = \bar{e}_1$. Lasketaan ensin vektorit $x_1, \dots, x_{n-1}$ Krylovin menetelmän mukaisesti asettamalla

$$x_k = Hx_{k-1}, \quad \text{jossa}\, k = 1, \dots, n.$$

(2) Ratkaistaan yhtälö

$$a_0x_0 + a_1x_1 + \dots a_{n-1}x_{n-1} + x_n = 0$$

vakioiden $a_0, a_1, \dots, a_{n-1}$ suhteen.

(3) Käytetään ratkaistuja vakioita karakteristisen polynomin $p(\lambda)$ kertoimina:

$$p(\lambda) = (-1)^n(\lambda^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0).$$

Tarkastellaan ominaisarvojen ratkaisua vielä Hessenbergin alamuodon tapauksessa. Tässäkin tapauksessa Krylovin menetelmä toimii vain, jos matriisi on redusoitumaton. Hessenbergin alamuodossa oleva matriisi $H_{n\times n} = [h_{ij}]$ on redusoitumaton, jos sen alkioille pätee

$$h_{k-1,k} \neq 0 \quad \text{kaikilla}\ k = 2, \dots, n.$$

Jos matriisi $H$ onkin Hessenbergin alamuotoa (4), asetetaan Krylovin menetelmän ensimmäiseksi vektoriksi viimeinen luonnollinen kantavektori $x_0 = \bar{e}_n$. Päinvastoin kuin Hessenbergin ylämuodon tapauksessa, nyt vektorilla $x_k = H^k x_0$ ensimmäiset $k$ komponenttia ovat nollia. Tässäkin tapauksessa vektorit $x_0, \ldots, x_{n-1}$ ovat lineaarisesti riippumattomia. Vektoreiden määrittämisen jälkeen algoritmia sovelletaan muuten samaan tapaan kuin ylämuodonkin tapauksessa.

Esimerkki 4.1. Hessenbergin matriisin määritelmän perusteella kaikki $2 \times 2$-matriisit ovat itse asiassa Hessenbergin matriiseja. Jokainen $2 \times 2$-matriisi voidaan tulkita olevan sekä Hessenbergin ylämuodossa että alamuodossa. Kuten aikaisemmin todettiin, Krylovin menetelmää Hessenbergin matriisin karakteristisen polynomin ratkaisemiseksi voidaan soveltaa vain, jos matriisi on redusoitumaton.

Jos matriisi $H_{2\times 2} = [h_{ij}]$ ajatellaan Hessenbergin ylämuodossa olevaksi matriisiksi, se on määritelmän mukaan redusoimaton, kun

$$h_{21} \neq 0.$$

Jos $H$ ajatellaan Hessenbergin alamuodossa olevaksi matriisiksi, se on määritelmän mukaan redusoimaton, kun

$$h_{12} \neq 0.$$

Jos matriisi on sekä ylämuodon, että alamuodon suhteen redusoitumaton, sen karakteristinen polynomi voidaan ratkaista Krylovin menetelmällä kummalla tahansa (joko ylä- tai alamuodon tapauksen) menetelmällä.

Jos $h_{21} = 0$ tai $h_{12} = 0$, kyseessä on kolmiomatriisi, jolloin lauseen 1.7 mukaan sen ominaisarvot voidaan lukea suoraan diagonaalilta.

Matriisi

$$H = \begin{bmatrix} 4 & -9 \\ 3 & -8 \end{bmatrix}$$

on redusoimaton sekä ylämuodon, että alamuodon mielessä. Ratkaistaan sen ominaisarvot Krylovin menetelmällä tulkiten se Hessenbergin ylämuodossa olevaksi. Nyt $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, joten

$$x_1 = Hx_0 = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad \text{ja} \quad x_2 = Hx_1 = \begin{bmatrix} -11 \\ -12 \end{bmatrix}.$$

Yhtälöstä

$$a_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_1 \begin{bmatrix} 4 \\ 3 \end{bmatrix} = - \begin{bmatrix} -11 \\ -12 \end{bmatrix}$$

saadaan yhtälöryhmä

$$\begin{cases} a_0 + 4a_1 = 11 \\ \quad\quad 3a_1 = 12 \end{cases},$$

josta saadaan ratkaisuiksi

$$\begin{cases} a_0 = -5 \\ a_1 = 4 \end{cases}.$$

Karakteriseksi polynomiksi saadaan

$$p(\lambda) = \lambda^2 + 4\lambda - 5,$$

jonka nollakohdista saadaan ratkaisuna ominaisarvoiksi $\lambda_1 = -5$ ja $\lambda_2 = 1$.

Jos halutaan käyttää Krylovin menetelmää alamuodon periaatteella, valitaan ensin $x_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, jolloin

$$x_1 = Hx_0 = \begin{bmatrix} -9 \\ -8 \end{bmatrix} \quad \text{ja} \quad x_2 = Hx_1 = \begin{bmatrix} 36 \\ 37 \end{bmatrix}.$$

Edellisestä saadaan yhtälö

$$a_0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + a_1 \begin{bmatrix} -9 \\ -8 \end{bmatrix} = - \begin{bmatrix} 36 \\ 37 \end{bmatrix}.$$

Ratkaisemalla loppuun asti päädytään samaan tulokseen.

Esimerkki 4.2. Ratkaistaan Hessenbergin matriisin

$$H = \begin{bmatrix} -6 & 2 & -5 \\ -4 & 2 & -3 \\ 0 & 2 & 3 \end{bmatrix}$$

ominaisarvot. Nyt

$$x_1 = Hx_0 \begin{bmatrix} -6 & 2 & -5 \\ -4 & 2 & -3 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -6 \\ -4 \\ 0 \end{bmatrix},$$

$$x_2 = Hx_1 \begin{bmatrix} -6 & 2 & -5 \\ -4 & 2 & -3 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} -6 \\ -4 \\ 0 \end{bmatrix} = \begin{bmatrix} 28 \\ 16 \\ -8 \end{bmatrix},$$

$$x_3 = Hx_2 \begin{bmatrix} -6 & 2 & -5 \\ -4 & 2 & -3 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 28 \\ 16 \\ -8 \end{bmatrix} = \begin{bmatrix} -96 \\ -56 \\ 8 \end{bmatrix}.$$

Lineaarisysteemistä

$$a_0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_1 \begin{bmatrix} -6 \\ -4 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 28 \\ 16 \\ -8 \end{bmatrix} = - \begin{bmatrix} -96 \\ -56 \\ 8 \end{bmatrix}$$

saadaan yhtälöryhmä

$$\begin{cases} a_0 - 6a_1 + 28a_2 = -96 \\ \quad -4a_1 + 16a_2 = -56 \\ \qquad\quad -8a_2 = 8 \end{cases}.$$

Yhtälöryhmän ratkaisut ovat

$$\begin{cases} a_0 = -8 \\ a_1 = 10 \\ a_2 = -1 \end{cases}.$$

Karakteriseksi polynomiksi saadaan tällöin

$$p(\lambda) = (-1)^3(\lambda^3 - 8\lambda^2 + 10\lambda - 1),$$

josta nollakohdista saadaan ratkaisuna ominaisarvoiksi $\lambda_1 = -4$, $\lambda_2 = 1$ ja $\lambda_3 = 2$.

Esimerkki 4.3. Etsitään Hessenbergin matriisin

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

ominaisarvot. Asetetaan ensin $x_0 = \bar{e}_1 \in \mathbb{C}^4$. Määritetään Krylovin menetelmän mukaisesti vektorit

$$x_{k-1} = H^{k-1}x_0 \quad \text{jossa } k = 1, \ldots, 4.$$

Tällöin

$$x_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, x_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, x_3 = \begin{bmatrix} 2 \\ 6 \\ 2 \\ 2 \end{bmatrix}, x_4 = \begin{bmatrix} 8 \\ 18 \\ 8 \\ 6 \end{bmatrix}.$$

Muodostetaan yhtälö

$$a_0 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + a_1 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} 2 \\ 6 \\ 2 \\ 2 \end{bmatrix} = - \begin{bmatrix} 8 \\ 18 \\ 8 \\ 6 \end{bmatrix},$$

josta saadaan yhtälöryhmä

$$\begin{cases} a_0 + a_2 + 2a_3 = -8 \\ a_1 + 2a_2 + 6a_3 = -18 \\ a_2 + 2a_3 = -8 \\ 2a_3 = -6 \end{cases}.$$

Yhtälöryhmän ratkaisuna saadaan vakiot

$$\begin{cases} a_0 = 0 \\ a_1 = 4 \\ a_2 = -2 \\ a_3 = -3 \end{cases},$$

joista saadaan karakteristiseksi polynomiksi

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

Karakteristisen polynomin nollakohdista saadaan ominaisarvoiksi $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 1 + \sqrt{5}$ ja $\lambda_4 = 1 - \sqrt{5}$.

Esimerkki 4.4. Matriisi

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

on Hessenbergin alamuotoa.

Ratkaistaessa sen karakteristista polynomia Krylovin menetelmällä asetetaan ensin $x_0 = \bar{e}_4 \in \mathbb{C}^4$. Määritetään vektorit

$$x_{k-1} = H^{k-1}x_0 \quad \text{jossa } k = 1, \dots, 4.$$

Tällöin

$$x_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, x_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 3 \end{bmatrix}, x_3 = \begin{bmatrix} 2 \\ 2 \\ 5 \\ 3 \end{bmatrix}, x_4 = \begin{bmatrix} 2 \\ 14 \\ 7 \\ 15 \end{bmatrix}.$$

Muodostetaan yhtälö

$$a_0 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + a_1 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 2 \\ 1 \\ 3 \end{bmatrix} + a_3 \begin{bmatrix} 2 \\ 2 \\ 5 \\ 3 \end{bmatrix} = - \begin{bmatrix} 2 \\ 14 \\ 7 \\ 15 \end{bmatrix}.$$

josta saadaan yhtälöryhmä

$$\begin{cases} 2a_3 = -2 \\ 2a_2 + 2a_3 = -14 \\ a_1 + a_2 + 5a_3 = -7 \\ a_0 + a_1 + 3a_2 + 3a_3 = -15 \end{cases}.$$

Yhtälöryhmän ratkaisuna saadaan vakiot

$$\begin{cases} a_0 = 2 \\ a_1 = 4 \\ a_2 = -6 \\ a_3 = -1 \end{cases},$$

Karakteristiseksi polynomiksi saadaan tällöin

$$p_H(\lambda) = \lambda^4 - \lambda^3 - 6\lambda^2 + 4\lambda + 2.$$

## 4.2. Redusoituvan Hessenbergin matriisin ominaisarvot

Krylovin menetelmää käytettäessä Hessenbergin matriisille $H_{n\times n}$ määritetyt vektorit $x_0, \ldots, x_{n-1}$ muodostavat avaruuden kannan vain, kun ne ovat lineaarisesti riippumattomia keskenään. Jos matriisi $H$ on redusoituva, kyseiset vektorit ovat keskenään lineaarisesti riippuvia, joten tällöin Krylovin menetelmää ei voidakaan käyttää karakteristisen polynomin löytämiseksi, joten on keksittävä jotain muuta.

Olkoon $k < n$. Redusoituva matriisi $H_{n\times n}$ voidaan kuitenkin esittää lohkomuodossa

$$H = \begin{bmatrix} H_1 & H_2 \\ 0 & H_3 \end{bmatrix},$$

jossa matriisi $H_1$ on kokoa $k \times k$, matriisi $H_2$ on kokoa $k \times (n-k)$ ja matriisi $H_3$ on kokoa $(n-k) \times (n-k)$. Redusoituvan Hessenbergin matriisin karakteristiselle polynomille päteekin seuraava tulos.

Lause 4.3. Olkoon $H_{n\times n}$ redusoituva Hessenbergin matriisi, jolle on olemassa edellä esitetty lohkomuoto. Tällöin

$$p_H(\lambda) = p_{H_1}(\lambda) p_{H_3}(\lambda).$$

Todistus. Matriisin $H_{n\times n}$ karakteristinen polynomi on $p_H(\lambda) = \det(H - \lambda I_n)$. Pienempien matriisien karakteristiset polynomit ovat $p_{H_1}(\lambda) = \det(H_1 - \lambda I_k)$ ja $p_{H_3}(\lambda) = \det(H_3 - \lambda I_{n-k})$. Koska

$$H - \lambda I_n = \begin{bmatrix} H_1 - \lambda I_k & H_2 \\ 0 & H_3 - \lambda I_{n-k} \end{bmatrix},$$

lauseen 1.8 nojalla saadaan nyt

$$\begin{aligned} p_H(\lambda) = \det(H - \lambda I_n) &= \det \begin{bmatrix} H_1 - \lambda I_k & H_2 \\ 0 & H_3 - \lambda I_{n-k} \end{bmatrix} \\ &= \det(H_1 - \lambda I_k) \det(H_1 - \lambda I_{n-k}) \\ &= p_{H_1}(\lambda) p_{H_3}(\lambda). \end{aligned}$$

□

Esimerkki 4.5. Redusoituvalle matriisille

$$H = \begin{bmatrix} 3 & 5i & i & 0 \\ i & -1 & -i & 2 \\ 0 & 0 & -i & 2 \\ 0 & 0 & 1 & i \end{bmatrix}$$

saadaan Krylovin menetelmän mukaisesti vektorit

$$x_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, x_1 = \begin{bmatrix} 3 \\ i \\ 0 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 4 \\ 2i \\ 0 \\ 0 \end{bmatrix}, x_3 = \begin{bmatrix} 2 \\ 2i \\ 0 \\ 0 \end{bmatrix}, x_4 = \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Nyt

$$4x_0 - 2x_1 + 2x_2 - x_3 + x_4 = 0,$$

jolloin vektorit $x_0, \dots, x_4$ ovat lineaarisesti riippuvia keskenään. Tällöin Krylovin menetelmää ei voida käyttää. Matriisi $H$ voidaan kuitenkin jakaa lohkoihin

$$A = \begin{bmatrix} 3 & 5i \\ i & -1 \end{bmatrix}, B = \begin{bmatrix} i & 0 \\ -i & 2 \end{bmatrix}, C = \begin{bmatrix} -i & 2 \\ 1 & i \end{bmatrix}.$$

Matriisi $H$ lohkomuodossa esitettynä on

$$H = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}.$$

Matriisin $H$ karakteristinen polynomi saadaan lohkomatriisien $A$ ja $C$ karakteristisien polynomien tulona: $p_H(\lambda) = p_A(\lambda)p_C(\lambda)$. Nyt

$$\begin{aligned} p_A(\lambda) = \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & 5i \\ i & -1-\lambda \end{vmatrix} \\ &= (3-\lambda)(-1-\lambda) - 5i^2 \\ &= \lambda^2 - 2\lambda + 2 \end{aligned}$$

ja

$$\begin{aligned} p_C(\lambda) = \det(B - \lambda I)C &= \begin{vmatrix} -i-\lambda & 2 \\ 1 & i-\lambda \end{vmatrix} \\ &= (-i-\lambda)(i-\lambda) - 2 \\ &= \lambda^2 - 1. \end{aligned}$$

Matriisin $H$ karakteristinen polynomi on

$$p_H(\lambda) = (\lambda^2 - 2\lambda + 2)(\lambda^2 - 1)$$

ja $p_H(\lambda) = 0$, kun $\lambda = 1 \pm i$ tai $\lambda = \pm 1$. Matriisin $H$ ominaisarvoiksi saadaan $\lambda_1 = 1 + i$, $\lambda_2 = 1 - i$, $\lambda_3 = 1$ ja $\lambda_4 = -1$.

Seuraava esimerkki selventää, kuinka ratkaista redusoituvan Hessenbergin matriisin karakteristinen polynomi tapauksessa, jossa alidiagonaalilla on useampia nollia.

Esimerkki 4.6. Tarkastellaan $7 \times 7$-matriisia

$$H = \begin{bmatrix} 5 & 2 & 5 & 6 & -2 & 0 & 1 \\ -7 & 4 & 1 & -5 & 0 & 2 & 3 \\ 0 & 0 & -6 & 2 & -5 & 1 & 8 \\ 0 & 0 & -4 & 2 & -3 & -7 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 4 & -9 \\ 0 & 0 & 0 & 0 & 0 & 3 & -8 \end{bmatrix}.$$

Matriisi $H$ on redusoituva, koska $h_{32} = h_{56} = 0$. Esitetään matriisi lohkomuodossa

$$H = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix},$$

jossa

$$A = \begin{bmatrix} 5 & 2 \\ -7 & 4 \end{bmatrix}, B = \begin{bmatrix} 5 & 6 & -2 & 0 & 1 \\ 1 & -5 & 0 & 2 & 3 \end{bmatrix}, C = \begin{bmatrix} -6 & 2 & -5 & 1 & 8 \\ -4 & 2 & -3 & -7 & 1 \\ 0 & 2 & 3 & 4 & 2 \\ 0 & 0 & 0 & 4 & -9 \\ 0 & 0 & 0 & 3 & -8 \end{bmatrix}.$$

Nyt matriisin $H$ karakteristiseksi polynomiksi saadaan

$$p_H(\lambda) = p_A(\lambda) p_C(\lambda).$$

Matriisin $A$ karakteristinen polynomi on helposti laskettavissa ja se on $p_A(\lambda) = \lambda^2 - \lambda + 6$ (esim. 2.2). Matriisin $C$ tapauksessa karakteristisen polynomin laskeminen on jo työläämpää. Matriisi $C$ ei myöskään ole redusoituva, koska sen alemman sivudiagonaalin alkiolle pätee $c_{34} = 0$. Jaetaan matriisi $C$ lohkomuotoon

$$C = \begin{bmatrix} D & E \\ 0 & F \end{bmatrix},$$

jossa

$$D = \begin{bmatrix} -6 & 2 & -5 \\ -4 & 2 & -3 \\ 0 & 2 & 3 \end{bmatrix}, E = \begin{bmatrix} 1 & 8 \\ -7 & 1 \\ 4 & 2 \end{bmatrix}, F = \begin{bmatrix} 4 & -9 \\ 3 & -8 \end{bmatrix}.$$

Nyt matriisin $C$ karakteristiseksi polynomiksi saadaan

$$p_C(\lambda) = p_D(\lambda) p_F(\lambda).$$

Matriisin $F$ karakteristinen polynomi on helposti laskettavissa ja se on $p_F(\lambda) = \lambda^2 + 4\lambda - 5$ (esim. 4.1). Matriisi $D$ on Hessenbergin muodossa oleva redusoimaton matriisi ja sen karakteriseksi polynomiksi saadaan Krylovin menetelmällä $p(\lambda) = (-1)^3(\lambda^3 - 8\lambda^2 + 10\lambda - 1)$ (esim. 4.2). Matriisin $H$ karakteristiseksi polynomiksi saadaan tällöin

$$\begin{aligned} p_H(\lambda) &= p_A(\lambda) p_C(\lambda) = p_A(\lambda) p_D(\lambda) p_F(\lambda) \\ &= (-1)^3(\lambda^2 - \lambda + 6)(\lambda^2 + 4\lambda - 5)(\lambda^3 - 8\lambda^2 + 10\lambda - 1). \end{aligned}$$

# Kirjallisuutta


[1] Stanley I. Grossman: *Elementary Linear Algebra*, Fifth edition, Cengage Learning, USA, 1994.

[2] Lee W. Johnson, R. Dean Riess, Jimmy T. Arnold: *Introduction to Linear Algebra*, Fourth edition, Addison-Wesley, USA, 1998.

[3] Tauno Metsänkylä, Marjatta Näätänen: *Algebra*, Limes ry, Helsinki, 2003.

[4] Peter J. Olver, Chehrzad Shakiban: *Applied Linear Algebra*, First edition, Pearson, USA, 2007.

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

[6] Veikko T. Purmonen: *Lineaarinen algebra ja geometria 2*, Luentomoniste 59, Matematiikan ja tilastotieteen laitos, Jyväskylän yliopisto, 2008.

[7] Mikko Saarimäki: *Vektorilaskentaa euklidisissa avaruuksissa*, Luentomoniste 65, Matematiikan ja tilastotieteen laitos, Jyväskylän yliopisto, 2012.

[8] Mikko Saarimäki: *Reaalisia vektoriavaruuksia ja ominaisarvoja*, Luentomoniste 66, Matematiikan ja tilastotieteen laitos, Jyväskylän yliopisto, 2012.

[9] Mikko Saarimäki: *Matriisiteoria*, Luentomoniste, Jyväskylän yliopisto, 2013.