



MATRIISIHAJOTELMIA JA NIIDEN SOVELLUKSIA

Emil Stenberg

Pro gradu -tutkielma  
Joulukuu 2020

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

Turun yliopiston laatu­järjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck-järjestelmällä

TURUN YLIOPISTO  
Matematiikan ja tilastotieteen laitos

STENBERG, EMIL: Matriisihajotelmia ja niiden sovelluksia  
Pro gradu -tutkielma, 40 s.  
Matematiikka  
Joulukuu 2020

---

Tässä Pro gradu-tutkielmassa tarkastellaan kolmea eri matriisihajotelmaa ja tutustutaan niiden erilaisiin sovelluskohteisiin. Matriisihajotelmien ideana on esittää alkuperäinen matriisi kahden tai useamman matriisin tulona. Näillä matriiseilla on usein jokin erityisominaisuus, kuten esimerkiksi ortogonaalisuus tai kolmiomuoto.

Tutkielmassa esiteltävät matriisihajotelmat ovat LU-, QR- ja singulaariarvohajotelma. Jokaisessa luvussa tutustutaan aluksi käsiteltävään matriisihajotelmaan, jonka jälkeen siirrytään sen yleisimpiin sovelluskohteisiin. Kaikista aiheista on pyritty esittämään mahdollisuuksien mukaan yksinkertainen esimerkki. Yleisimpien sovellusten lisäksi käydään läpi myös lyhyesti muita mainitsemisen arvoisia sovelluskohteita.

Tutkielma on suunnattu lineaarialgebran kurssin käyneille. Tutkielmassa ei ole erillistä lukua esitiedoille, vaan tarvittavat tiedot on pyritty esittelemään aina ennen käsiteltävää asiaa.

Asiasanat: matriisihajotelma, LU-hajotelma, QR-hajotelma, singulaariarvohajotelma.



# Sisältö

<b>1</b>	<b>Johdanto</b>	<b>1</b>
<b>2</b>	<b>LU-hajotelma</b>	<b>2</b>
2.1	Gaussin eliminaatio . . . . .	2
2.2	LU-hajotelma . . . . .	4
2.3	LU-hajotelman sovelluksia . . . . .	7
2.3.1	Lineaaristen yhtälöryhmien ratkaiseminen . . . . .	7
2.3.2	Käänteismatriisin etsiminen . . . . .	9
2.3.3	Determinantin laskeminen . . . . .	10
2.3.4	Muita sovelluksia . . . . .	10
<b>3</b>	<b>QR-hajotelma</b>	<b>11</b>
3.1	Gramin-Schmidtin ortogonalisointi . . . . .	11
3.2	QR-hajotelma . . . . .	11
3.3	QR-hajotelman sovelluksia . . . . .	14
3.3.1	Lineaaristen yhtälöryhmien ratkaiseminen . . . . .	14
3.3.2	QR-algoritmi . . . . .	18
3.3.3	Mooren-Penrosen yleistetty käänteismatriisi . . . . .	23
3.3.4	Determinantin itseisarvon laskeminen . . . . .	24
3.3.5	Muita sovelluksia . . . . .	25
<b>4</b>	<b>Singulaariarvohajotelma</b>	<b>26</b>
4.1	Symmetrinen matriisi . . . . .	26
4.2	Singulaariarvot . . . . .	27
4.3	Singulaariarvohajotelma . . . . .	28
4.4	Singulaariarvohajotelman sovelluksia . . . . .	30
4.4.1	Matriisin nolla- ja riviavaruuksien ortonormaalit kannat . . . . .	31
4.4.2	Matriisin häiriöalttius . . . . .	32
4.4.3	Matriisin approksimointi . . . . .	34
4.4.4	Mooren-Penrosen yleistetty käänteismatriisi ja pienimmän neliosumman menetelmä . . . . .	36
4.4.5	Muita sovelluksia . . . . .	39



# 1 Johdanto

Matriisit ovat hyödyllisiä pakattaessa tietoa kompaktimmaksi. Lisäksi niitä esiintyy luonnollisesti esimerkiksi lineaarisia yhtälöitä ratkaistaessa tai ilmaistessa eri asioiden välisiä lineaarisia suhteita. Monesti tällaista dataa käsittelevät matriisit ovat hyvin suuria, jolloin usein voidaan hyödyntää matriisihajotelmia laskenta-ajan tehostamiseksi. Niissä matriisi esitetään kahden tai useamman helpommin käsiteltävän matriisin tulona. Näin esimerkiksi matriisi  $A$  voitaisiin esittää muodossa  $A = BC$ , missä matriiseilla  $B$  ja  $C$  on joitain hyödyllisiä ominaisuuksia. Matriisihajotelmia on paljon erilaisia, ja niitä sovelletaan erilaisiin käyttötarpeisiin. Tässä tutkielmassa esitellään kolme matriisihajotelmaa, jotka ovat LU-, QR- ja singulaariarvohajotelma.

Näistä kolmesta ensimmäisenä tutustutaan LU-hajotelmaan. Siinä matriisi  $A$  voidaan ilmaista ala- ja yläkolmiomatriisin tulona. Gaussin eliminaatiolla muodostettava LU-hajotelma on yksi keskeisimpiä metodeja ratkaista lineaarisia yhtälöryhmiä. Lisäksi hajotelman avulla voidaan tehostaa myös esimerkiksi determinantin laskemista tai käänteismatriisin selvittämistä. LU-hajotelmaa ei voida kuitenkaan muodostaa kaikille matriiseille, joka rajoittaa hajotelman laajempaa hyödyntämistä.

Toisin kuin LU-hajotelman, kolmannessa luvussa esiteltävän QR-hajotelman voi muodostaa kaikille matriiseille. QR-hajotelmassa alkuperäisen matriisin pystyriveistä muodostetaan matriisi, jonka pystyrivit ovat ortonormaalit. Kun tämä matriisi kerrotaan yläkolmiomatriisilla, saadaan alkuperäisen matriisin QR-hajotelma. Matriisin ilmaisemista tässä muodossa voidaan hyödyntää esimerkiksi pienimmän neliösumman menetelmässä tai ominaisarvojen etsimisessä QR-algoritmilla.

Viimeisessä luvussa esitellään singulaariarvohajotelma, joka voidaan myös muodostaa kaikille matriiseille. Toisin kuin kahdessa edellä mainitussa hajotelmassa, singulaariarvohajotelmassa matriisi ilmaistaan kolmen eri matriisin tulona. Reaalissa tapauksessa hajotelmassa hyödynnetään singulaariarvoja ja ortogonaalisuutta. Matriisin esittäminen tämänkaltaisessa muodossa mahdollistaa monia erilaisia tilanteita, joissa sitä voi hyödyntää. Näitä ovat esimerkiksi matriisin approksimoiminen tai häiriöalttiuden selvittäminen.

Tutkielmassa käsitellään niin reaalisia kuin kompleksisia matriiseja. Monissa tapauksissa tulos voidaan yleistää sekä reaalille että kompleksille matriiseille, joten tutkielmassa merkitään usein kuntia  $\mathbb{R}$  ja  $\mathbb{C}$  yhteisesti kunnalla  $\mathbb{K}$ .

## 2 LU-hajotelma

Ensimmäisenä matriisihajotelmana käsitellään LU-hajotelma, jossa matriisi  $A$  voidaan ilmaista alakolmiomatriisin  $L$  ja yläkolmiomatriisin  $U$  tulona  $A = LU$ . Vaikka LU-hajotelma on olemassa myös joillekin  $m \times n$ -matriiseille [5], rajaudutaan tässä luvussa vain neliömatriiseihin. Hajotelman tärkeimpänä sovelluksena voidaan pitää lineaaristen yhtälöryhmien ratkaisemista. Sen avulla nykyiset tietokoneet pystyvät tehokkaasti ratkaisemaan lineaarisia yhtälöryhmiä  $A\mathbf{x} = \mathbf{b}$ , joissa vektori  $\mathbf{b}$  muuttuu matriisin  $A$  pysyessä samana. Lisäksi LU-hajotelmaa voidaan soveltaa muutenkin, kuten käänteismatriisin etsimisessä tai determinantin laskemisessa. LU-hajotelma muodostetaan Gaussin eliminoinnilla, johon tutustutaan seuraavaksi tarkemmin. Luvut 2.1 ja 2.2 pohjautuvat lähteeseen [1].

### 2.1 Gaussin eliminaatio

Jotta Gaussin eliminaatio LU-hajotelman muodostuksessa olisi mahdollisimman intuitiivinen, käydään ensin Gaussin eliminaatio lyhyesti läpi lineaarisille yhtälöryhmille. Tässä ideana on muuttaa lineaarista yhtälöryhmää vastaava matriisi riviopeeraatioita käyttäen kolmiomatriisi muotoon. Ratkaistaan esimerkkinä oleva lineaarinen yhtälöryhmä

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

Tämä voidaan ilmaista myös matriisimuodossa  $A\mathbf{x} = \mathbf{b}$ , jolloin

$$\begin{pmatrix} 1 & -3 & 1 \\ 3 & -4 & 1 \\ 0 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

Merkitään matriisin  $A$  alkioita  $a_{1,p}$  kirjaimella  $p = 1$ . Tätä alkioita kutsutaan matriisin  $A$  pivot-alkioksi. Nyt muokataan lineaarista yhtälöryhmää siten, että

- toisesta vaakarivistä vähennetään ensimmäinen vaakarivi kerrottuna luvulla  $\frac{3}{p}$ ,
- kolmannesta vaakarivistä vähennetään ensimmäinen vaakarivi kerrottuna luvulla  $\frac{0}{p}$ .

Näiden operaatioiden jälkeen lineaarinen yhtälöryhmä on muotoa

$$\begin{cases} x_1 - 3x_2 + x_3 = 2 \\ 5x_2 - 2x_3 = -6 \\ 2x_2 - 1x_3 = 1. \end{cases}$$

Merkitään uudeksi pivot-alkioksi  $p = 5$  (uuden matriisin (2,2)-alkio). Nyt kolmannesta vaakarivistä vähennetään toinen vaakarivi kerrottuna luvulla  $\frac{2}{p}$ , jolloin yhtä-



loryhmä saadaan muotoon

$$\begin{cases} x_1 - 3x_2 + x_3 = 2 \\ \phantom{x_1} 5x_2 - 2x_3 = -6 \\ \phantom{x_1} \phantom{x_2} - \frac{1}{5}x_3 = \frac{17}{5}. \end{cases}$$

Näin yhtäloryhmä on saatu muokattua kolmiomuotoon. Nyt alimmalta riviltä saadaan ratkaistua  $x_3 = -17$ , jonka jälkeen sijoittamalla muuttujan  $x_3$  arvo keskimäiseen riviin saadaan ratkaistua  $x_2 = -8$ . Lopuksi sijoittamalla vielä muuttujien  $x_2$  ja  $x_3$  arvot ylimpään riviin saadaan viimeiseksi muuttujaksi  $x_1 = -5$ .

Edellä käytyä ideaa voidaan soveltaa tietynlaisten neliömatriisien kohdalla, missä Gaussin eliminaation lopputulokseksi saadaan yläkolmiomatriisi  $T$ . Toisin sanoen tarkoituksena on löytää sellainen matriisi  $M$ , että  $MA = T$ . Tarkastellaan matriisin  $M$  muodostamista lähemmin alla. Menetelmässä muodostetaan matriisijono  $A^{(k)}$ , missä  $1 \leq k \leq n$ . Silloin onnistuneessa Gaussin eliminaatiossa pätee  $A^{(1)} = A$  ja  $A^{(n)} = T$ . Määritellään matriisin  $A^{(k)}$  alkioiksi  $A^{(k)} = (a_{i,j}^{(k)})_{1 \leq i,j \leq n}$ . Matriisin alkioita  $a_{k,k}^{(k)}$  kutsutaan matriisin  $A^{(k)}$  pivot-alkioksi. Pivot-alkion ollessa nolla Gaussin eliminaatiota ei voida viedä loppuun, sillä pivot-alkiolla joudutaan jakamaan eliminaatiossa. Näin ollen on tärkeää huomata, että menetelmää ei voida soveltaa kaikille neliömatriiseille. Joissakin tilanteissa voidaan kuitenkin käyttää tuentaa, jossa matriisi kerrotaan permutatiomatriisilla  $P$ . Permutatiomatriisin  $P$  jokaisella rivillä ja sarakkeella on vain yksi ykkönen muiden alkioden ollessa nollia. Tämänkaltaisella operaatiolla saadaan vaihdettua matriisin  $k$ :s vaakarivi jonkin toisen vaakarivin kanssa siten, että pivot-alkio on nolasta eroava ja eliminaatiota voidaan jatkaa. Selkeyden vuoksi rajaudutaan tässä luvussa vain tilanteisiin, jossa tuentaa ei tarvita. Lisää Gaussin eliminoinnista tuennalla voi lukea esimerkiksi lähteestä [1].

Ensimmäisessä Gaussin eliminoinnissa matriisi  $A$  kerrotaan matriisilla  $E^{(1)}$ , missä

$$E^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{a_{2,1}^{(1)}}{a_{1,1}^{(1)}} & 1 & \\ \vdots & & \ddots & 0 \\ -\frac{a_{n,1}^{(1)}}{a_{1,1}^{(1)}} & 0 & \cdots & 1 \end{pmatrix}.$$

Näin muodostamalla matriisi  $A^{(2)} = E^{(1)}A$  saadaan

$$A^{(2)} = \begin{pmatrix} a_{1,1}^{(1)} & \cdots & a_{1,n}^{(1)} \\ 0 & \left( a_{i,j}^{(2)} \right) \\ \vdots & & \\ 0 & & \end{pmatrix},$$

missä  $a_{i,j}^{(2)} = a_{i,j}^{(1)} - \frac{a_{i,1}^{(1)}}{a_{1,1}^{(1)}} \cdot a_{1,j}^{(1)}$ , kaikilla  $2 \leq i, j \leq n$ . Kun Gaussin eliminointeja on tehty  $k-1$  kertaa, ollaan saatu matriisi  $A^{(k)}$ , jonka ensimmäiset  $k-1$  pystyriviä ovat

nollia diagonaalin alapuolelta. Matriisi  $A^{(k+1)}$  muodostetaan kertomalla matriisi  $A^{(k)}$  matriisilla  $E^{(k)}$ , missä

$$E^{(k)} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -\frac{a_{k+1,k}^{(k)}}{a_{k,k}^{(k)}} & 1 & \cdots & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & -\frac{a_{n,k}^{(k)}}{a_{k,k}^{(k)}} & 0 & \cdots & 1 \end{pmatrix}.$$

Näin ollen matriisiksi  $A^{(k+1)}$  saadaan

$$A^{(k+1)} = \begin{pmatrix} a_{1,1}^{(1)} & \cdots & & \cdots & a_{1,n}^{(1)} \\ 0 & \ddots & & & \vdots \\ \vdots & & a_{k,k}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{k,n}^{(k)} \\ & & 0 & a_{k+1,k+1}^{(k+1)} & \cdots & a_{k+1,n}^{(k+1)} \\ \vdots & & \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,k+1}^{(k+1)} & \cdots & a_{n,n}^{(k+1)} \end{pmatrix},$$

missä  $a_{i,j}^{(k+1)} = a_{i,j}^{(k)} - \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}} a_{k,j}^{(k)}$ , kaikilla  $k+1 \leq i, j \leq n$ . Tehtäessä  $n-1$  Gaussin eliminointia saadaan yläkolmiomatriisi  $A^{(n)} = T$ , joka voidaan ilmaista muodossa

$$T = \underbrace{(E^{(n-1)} \cdots E^{(1)})}_{=M} A = MA.$$

Edellä kuvailtua tapaa hyödynnetään LU-hajotelman muodostamisessa.

## 2.2 LU-hajotelma

LU-hajotelma muodostamisessa käytetään Gaussin eliminaatiota. Kuten aiemmin kävi ilmi, Gaussin eliminaatiossa kaikkien pivot-alkioiden täytyy olla nolasta eroavia. Tämän vuoksi kaikille matriiseille ei voida muodostaa LU-hajotelmaa. Pohditaan seuraavaksi, millaisille matriiseille LU-hajotelma on olemassa. Esimerkiksi matriisille

$$A = \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix}$$

ei voida muodostaa LU-hajotelmaa ilman tuentaa, sillä sen pivot-alkio on nolla. Sen sijaan esimerkiksi matriisille

$$B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

saadaan muodostettua LU-hajotelma. LU-hajotelman olemassaolon selvittämisessä hyödynnetään usein matriisin  $A \in \mathcal{M}_n$  pääalimatriiseja

$$A_k = \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} \\ \vdots & \ddots & \vdots \\ a_{1,k} & \cdots & a_{k,k} \end{pmatrix},$$

missä  $1 \leq k \leq n-1$ . Kun matriisin pääalimatriiseiden tiedetään olevan säännöllisiä, matriisilla  $A$  on olemassa LU-hajotelma ilman tuentaa. Lauseessa 2.2 muodostetaan ja todistetaan LU-hajotelman olemassaolo tällaisille neliömatriiseille. Ennen lausetta 2.2 tulee kuitenkin vielä osoittaa, että kolmiomatriisin determinantti on sen diagonaalialkioiden tulo.

**Lemma 2.1.** Kolmiomatriisin  $T = (t_{i,j}) \in \mathcal{M}_n(\mathbb{K})$  determinantti on sen diagonaalialkioiden tulo  $\det(T) = t_{1,1}t_{2,2} \cdots t_{n,n}$ .

*Todistus.* Oletetaan, että matriisi  $T$  on yläkolmiomatriisi, eli  $t_{i,j} = 0$ , kun  $i > j$ . Nyt siis matriisi  $T$  on muotoa

$$T = \begin{pmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,n} \\ 0 & t_{2,2} & \cdots & t_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & t_{n,n} \end{pmatrix}.$$

Kehitetään aluksi determinantti ensimmäisen pystyrivin suhteen. Silloin  $\det(T) = t_{1,1} \det(T_{1,1})$ , missä

$$T_{1,1} = \begin{pmatrix} t_{2,2} & t_{2,3} & \cdots & t_{2,n} \\ 0 & t_{3,3} & \cdots & t_{3,n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & t_{n,n} \end{pmatrix}.$$

Toisaalta myös alideterminantti  $T_{1,1}$  yläkolmiomuotoa, jolloin voidaan kirjoittaa  $\det(T_{1,1}) = t_{2,2} \det(T_{2,2})$  ja  $\det(T) = t_{1,1}t_{2,2} \det(T_{2,2})$ . Jatkamalla samaa päättelyä saadaan lopuksi  $\det(T_{n-1,n-1}) = t_{n,n}$  ja  $\det(T) = t_{1,1}t_{2,2} \cdots t_{n,n}$ . Samalla tavalla väite voidaan todistaa myös alakolmiomatriiseille. Tällöin matriisi kehitetään joka askeleella vaakarivin suhteen pystyrivien sijaan.  $\square$

Esitellään seuraavaksi LU-hajotelman olemassaolo ilman tuentaa. Todistus mukailee lähdeä [1].

**Lause 2.2.** (LU-hajotelma) Olkoon  $A \in \mathcal{M}_n(\mathbb{K})$  matriisi, jonka pääalimatriisit  $A_k$  ovat säännöllisiä kaikilla  $1 \leq k < n$ . Silloin on olemassa sellainen alakolmiomatriisi  $L \in \mathcal{M}_n(\mathbb{K})$ , missä  $l_{i,i} = 1$ , ja yläkolmiomatriisi  $U \in \mathcal{M}_n(\mathbb{K})$ , että  $A = LU$ .

*Todistus.* Tekemällä matriisiin  $A$  Gaussin eliminointeja  $n-1$  kappaletta saadaan yläkolmiomatriisi  $U = A^{(n)}$ . Jotta tämä saataisiin suoritettua ilman tuentaa, oletetaan pivot-alkioiden olevan nollassa eroavia. Tämä oletus osoitetaan todeksi todistuksen

lopussa. Nyt suorittamalla  $n - 1$  eliminointia yhtälö voidaan kirjoittaa muodossa  $A^{(n)} = E^{(n-1)}E^{(n-2)} \dots E^{(1)}A$ , missä

$$E^{(k)} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -l_{k+1,k} & 1 & \cdots & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & -l_{n,k} & 0 & \cdots & 1 \end{pmatrix},$$

jossa  $l_{i,k} = \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}$  kaikilla  $k + 1 \leq i \leq n$ . Laskemalla matriisiin  $E^{(k)}$  käänteismatriisi saadaan

$$(E^{(k)})^{-1} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & l_{k+1,k} & 1 & \cdots & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & l_{n,k} & 0 & \cdots & 1 \end{pmatrix}.$$

Lasketaan seuraavaksi matriisitulo  $(E^{(1)})^{-1} \dots (E^{(n-2)})^{-1}(E^{(n-1)})^{-1}$  ja merkitään saatua matriisiä kirjaimella  $L$ . Matriisiksi  $L$  saadaan

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{2,1} & 1 & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ l_{n,1} & \cdots & l_{n,n-1} & 1 \end{pmatrix},$$

joka on alakolmiomatriisi. Lemman 2.1 avulla matriisin  $L$  determinantiksi saadaan  $\det(L) = l_{1,1}l_{2,2} \cdots l_{n,n} = 1$ , jolloin sen on oltava säännöllinen. Nyt matriisille  $A$  voidaan muodostaa hajotelma

$$A = \underbrace{(E^{(1)})^{-1} \dots (E^{(n-2)})^{-1}(E^{(n-1)})^{-1}}_{=L} \underbrace{A^{(n)}}_U = LU.$$

Oletetaan, että lauseessa määriteltyjen matriisin  $A \in \mathcal{M}_n(\mathbb{K})$  pääalideterminanttien on oltava nollasta eroavia. Vielä tulee osoittaa, ettei Gaussin eliminointi tarvitse tuentaa, eli ts. pivot-alkioiden on oltava nollasta eroavia. Käytetään todistuksessa induktiota. Ensimmäinen pivot-alkio  $a_{1,1} \neq 0$ , sillä  $\det(a_{1,1}) \neq 0$ . Oletetaan nyt, että ensimmäiset  $k - 1$  kappaletta Gaussin eliminointeja eivät tarvitse tuentaa. Nyt on siis osoitettava, että myöskään  $k$ :s askel ei tarvitse tuentaa. Kun on tehty  $k - 1$  kappaletta Gaussin eliminointeja, niin  $A = (E^{(1)})^{-1} \dots (E^{(k-2)})^{-1}(E^{(k-1)})^{-1}A^{(k)}$ . Tämä voidaan kirjoittaa muodossa

$$\left( \begin{array}{c|c} A_{1,1} & A_{1,2} \\ \hline A_{2,1} & A_{2,2} \end{array} \right) = \left( \begin{array}{c|c} L_{1,1}^{(k)} & 0 \\ \hline L_{2,1}^{(k)} & I \end{array} \right) \left( \begin{array}{c|c} U_{1,1}^{(k)} & A_{1,2}^{(k)} \\ \hline A_{2,1}^{(k)} & A_{2,2}^{(k)} \end{array} \right),$$

missä alimatriisit  $A_{1,1}$ ,  $L_{1,1}^{(k)}$  ja  $U_{1,1}^{(k)}$  ovat kokoa  $k \times k$ . Lohkomatriisin laskusääntöjen mukaisesti tämä voidaan kirjoittaa muodossa

$$A_{1,1} = L_{1,1}^{(k)} U_{1,1}^{(k)},$$

missä  $A_{1,1}$  on pääalimatriisi  $A_k$ ,  $U_{1,1}^{(k)}$  on yläkolmiomatriisi ja  $L_{1,1}^{(k)}$  on alakolmiomatriisi, jonka diagonaalilla on ykkösiä. Kolmiomatriisien determinantti on diagonaalialkioiden tulo, joten matriisi  $L_{1,1}^{(k)}$  on säännöllinen. Koska oletuksen mukaan myös pääalimatriisit ovat säännöllisiä, niin muodosta  $U_{1,1}^{(k)} = (L_{1,1}^{(k)})^{-1} A_{1,1}$  nähdään, että myös matriisin  $U_{1,1}^{(k)}$  on oltava säännöllinen. Silloin  $\det(U_{1,1}^{(k)}) = a_{1,1}^{(k)} \cdots a_{k,k}^{(k)} \neq 0$ , joten myös pivot-alkion  $a_{k,k}^{(k)}$  on oltava nolasta eroava. Näin ollen pivot-alkiot ovat eri suuria kuin nolla ja LU-hajotelma voidaan muodostaa.  $\square$

**Esimerkki 2.3.** Muodostetaan LU-hajotelma matriisille

$$A = \begin{pmatrix} 3 & 3 & -3 \\ 18 & 6 & 3 \\ 24 & 0 & 3 \end{pmatrix}.$$

Yläkolmiomatriisi  $U$  saadaan suorittamalla Gaussin eliminaatioita matriisille  $A$

$$\begin{aligned} & \begin{pmatrix} 3 & 3 & -3 \\ 18 & 6 & 3 \\ 24 & 0 & 3 \end{pmatrix} \xrightarrow{R2 - 6 \cdot R1} \begin{pmatrix} 3 & 3 & -3 \\ 0 & -12 & 21 \\ 24 & 0 & 3 \end{pmatrix} \\ & \xrightarrow{R3 - 8 \cdot R1} \begin{pmatrix} 3 & 3 & -3 \\ 0 & -12 & 21 \\ 0 & -24 & 27 \end{pmatrix} \xrightarrow{R3 - 2 \cdot R2} \begin{pmatrix} 3 & 3 & -3 \\ 0 & -12 & 21 \\ 0 & 0 & -15 \end{pmatrix} = U. \end{aligned}$$

Alakolmiomatriisi  $L$  saadaan muodostettua Gaussin eliminaatiossa käytettävistä kertoimista, eli  $l_{21} = 6$ ,  $l_{31} = 8$  ja  $l_{32} = 2$ . Näin ollen alakolmiomatriisiksi saadaan

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 8 & 2 & 1 \end{pmatrix}.$$

Lopuksi tarkistetaan vielä, että  $A = LU$ .  $\triangle$

## 2.3 LU-hajotelman sovelluksia

Tutustutaan seuraavaksi, minkälaisissa tilanteissa LU-hajotelmaa voidaan soveltaa. Sovelluksia käsittelevissä luvuissa on hyödynnetty lähteitä [4, 8, 12].

### 2.3.1 Lineaaristen yhtälöryhmien ratkaiseminen

Monissa sovelluksissa ollaan kiinnostuneita ratkaisemaan tehokkaasti lineaarinen yhtälöryhmä  $A\mathbf{x} = \mathbf{b}$ , jossa kerroinmatriisi  $A$  ja vektori  $\mathbf{b}$  ovat tunnettuja. Gaussin eliminaatio on yksi keskeisimpiä metodeja ratkaista lineaarisia yhtälöryhmiä. Jos

tehtävänä on kuitenkin ratkaista useampia yhtälöryhmiä, missä matriisi  $A$  pysyy samana, mutta vektori  $\mathbf{b}$  muuttuu, voidaan usein apuna käyttää LU-hajotelmaa. LU-menetelmän etuna lineaaristen yhtälöryhmien ratkaisussa on sen nopeus, sillä Gaussin eliminaatiota ei tarvitse tehdä uudestaan vektorin  $\mathbf{b}$  muuttuessa. Tämän vuoksi LU-hajotelma on tehokas tapa ratkoa esimerkiksi lineaarisia optimointiongelmia.

Aiemmin käsiteltiin jo LU-hajotelman muodostaminen. Yhtälöryhmää ratkaistaessa yhtälöryhmä saadaan kirjoitettua muotoon

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \Leftrightarrow (LU)\mathbf{x} &= \mathbf{b} \\ \Leftrightarrow L(U\mathbf{x}) &= \mathbf{b} \quad (\text{merk. } U\mathbf{x} = \mathbf{z}) \\ \Leftrightarrow L\mathbf{z} &= \mathbf{b}. \end{aligned}$$

Yhtälöryhmä ratkaistaan LU-hajotelmalla seuraavalla tavalla.

- Ratkaistaan  $\mathbf{z}$  yhtälöstä  $L\mathbf{z} = \mathbf{b}$ .
- Ratkaistaan  $\mathbf{x}$  yhtälöstä  $U\mathbf{x} = \mathbf{z}$ .

Havainnollistetaan menetelmän ideaa esimerkin avulla.

**Esimerkki 2.4.** Ratkaistaan yhtälöryhmä

$$\begin{cases} 3x_1 + 3x_2 - 3x_3 = 1 \\ 18x_1 + 6x_2 + 3x_3 = 2 \\ 24x_1 + 3x_3 = 4. \end{cases}$$

Yhtälöryhmästä saadaan sama matriisi kuin esimerkissä 2.3. Tällöin matriiseiksi  $L$  ja  $U$  saadaan

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 8 & 2 & 1 \end{pmatrix} \text{ ja } U = \begin{pmatrix} 3 & 3 & -3 \\ 0 & -12 & 21 \\ 0 & 0 & -15 \end{pmatrix}.$$

Tehtävänä on siis ratkaista yhtälöryhmä, joka voidaan kirjoittaa muotoon  $A\mathbf{x} = \mathbf{b}$ , missä  $\mathbf{b} = (1 \ 2 \ 4)^T$ . Ratkaistaan nyt yhtälöryhmä aiemmin esitetyllä tavalla. Selvitetään aluksi vektori  $\mathbf{z}$  yhtälöstä  $L\mathbf{z} = \mathbf{b}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 8 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix},$$

joka voidaan kirjoittaa yhtälöryhmämuodossa seuraavalla tavalla

$$\begin{cases} z_1 & & & = & 1 \\ 6z_1 & + & z_2 & & = & 2 \\ 8z_1 & + & 2z_2 & + & z_3 & = & 4. \end{cases}$$

Ylimmästä rivistä saadaan ratkaistua  $z_1 = 1$ . Sijoittamalla muuttujan  $z_1$  arvo keskimmäiseen riviin saadaan ratkaistua  $z_2 = -4$ . Lopuksi vielä sijoittamalla muuttujien  $z_1$  ja  $z_2$  arvot alimpaan riviin saadaan ratkaistua  $z_3 = 4$ . Vektoriksi  $\mathbf{z}$  saadaan siis  $\mathbf{z} = (1 \ -4 \ 4)^T$ . Nyt vektori  $\mathbf{x}$  saadaan ratkaistua yhtälöstä  $U\mathbf{x} = \mathbf{z}$

$$\begin{pmatrix} 3 & 3 & -3 \\ 0 & -12 & 21 \\ 0 & 0 & -15 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ 4 \end{pmatrix}.$$

Vastaavasti myös tämä voidaan kirjoittaa yhtälöryhmämuodossa

$$\begin{cases} 3x_1 + 3x_2 - 3x_3 = 1 \\ -12x_2 + 21x_3 = -4 \\ -15x_3 = 4, \end{cases}$$

jolloin saadaan ratkaistua haluttu vektori  $\mathbf{x} = \left(\frac{1}{5} \ \frac{-2}{15} \ \frac{-4}{15}\right)^T$ . △

Kuten ylläolevasta esimerkistä nähdään, LU-hajotelman etu lineaarisen yhtälöryhmän ratkaisemisessa löytyy kolmiomatriiseista, joiden avulla saadaan ratkaistua helposti muuttujat. Perinteiseen Gaussin eliminaatioon verrattuna menetelmän tehokkuus nousee kuitenkin esille vasta, kun vektori  $\mathbf{b}$  muuttuu kerroinmatriisin  $A$  pysyessä samana. Tällöin LU-hajotelmaa ei tarvitse muodostaa uudelleen, vaan voidaan suoraan edetä ratkaisemaan vektoria  $\mathbf{z}$  yhtälöstä  $L\mathbf{z} = \mathbf{b}$ . Gaussin eliminaatiolla ratkaistaessa koko yhtälöryhmän ratkaisemisen jouduttaisiin aloittamaan täysin alusta.

### 2.3.2 Käänteismatriisin etsiminen

Käänteismatriisin etsiminen saattaa olla työläs operaatio, minkä vuoksi sen laske- mista pyritään välttämään monissa tilanteissa. Esimerkiksi lineaarisen yhtälöryhmän ratkaisemisessa käänteismatriisin hyödyntäminen  $A\mathbf{x} = \mathbf{b} \implies \mathbf{x} = A^{-1}\mathbf{b}$  on yleensä tehoton ratkaisu. Jos käänteismatriisi täytyy kuitenkin selvittää, niin se voidaan etsiä ratkaisemalla yhtälöt

$$A\mathbf{x}_i = \mathbf{e}_i, \text{ missä } \mathbf{e}_1 = (1 \ 0 \ \dots \ 0 \ 0)^T, \mathbf{e}_2 = (0 \ 1 \ \dots \ 0 \ 0)^T, \mathbf{e}_n = (0 \ 0 \ \dots \ 0 \ 1)^T.$$

Ratkaistavat vektorit  $\mathbf{x}_i$  ovat käänteismatriisin  $A^{-1}$  pystysarakkeita ja ne voidaan LU-hajotelman avulla löytää samalla tavalla kuin lineaaristen yhtälöryhmien tapauksessa.

**Esimerkki 2.5.** Lasketaan esimerkin 2.3 matriisille  $A$  käänteismatriisi. Merkitään  $U\mathbf{x}_i = \mathbf{z}_i$ , jolloin  $L\mathbf{z}_i = \mathbf{e}_i$ . Vektori  $\mathbf{z}_1$  saadaan ratkaisemalla

$$\begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 8 & 2 & 1 \end{pmatrix} \begin{pmatrix} z_{11} \\ z_{12} \\ z_{13} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \implies \mathbf{z}_1 = \begin{pmatrix} 1 \\ -6 \\ 4 \end{pmatrix}.$$

Vastaavasti saadaan ratkaistua  $\mathbf{z}_2 = (0 \ 1 \ -2)^T$  ja  $\mathbf{z}_3 = (0 \ 0 \ 1)^T$ . Seuraavaksi ratkaistaan vektorit  $\mathbf{x}_i$  yhtälöstä  $U\mathbf{x}_i = \mathbf{z}_i$ , josta saadaan  $\mathbf{x}_1 = \left(\frac{1}{30} \ \frac{1}{30} \ \frac{-4}{15}\right)^T$ ,  $\mathbf{x}_2 = \left(\frac{-1}{60} \ \frac{3}{20} \ \frac{2}{15}\right)^T$  ja  $\mathbf{x}_3 = \left(\frac{1}{20} \ \frac{-7}{60} \ \frac{-1}{15}\right)^T$ . Näin ollen matriisin  $A$  käänteismatriisiksi saadaan

$$A^{-1} = \begin{pmatrix} \frac{1}{30} & \frac{-1}{60} & \frac{1}{20} \\ \frac{1}{30} & \frac{20}{2} & \frac{60}{15} \\ \frac{-4}{15} & \frac{2}{15} & \frac{-1}{15} \end{pmatrix}.$$

△

### 2.3.3 Determinantin laskeminen

Determinantti on tärkeä matemaattinen työkalu, jota sovelletaan monella eri tavalla, kuten esimerkiksi matriisin säännöllisyyden tutkimisessa. Matriisin kasvaessa determinantin laskeminen muuttuu kuitenkin jatkuvasti tehottomammaksi, jolloin LU-hajotelmalla voidaan determinantin laskemista nopeuttaa hyödyntämällä kolmio-matriisien ominaisuuksia.

**Lause 2.6.** Jos matriisilla  $A \in \mathcal{M}_n(\mathbb{K})$  on olemassa LU-hajotelma, niin silloin

$$\det(A) = \det(U) = u_{11}u_{22} \cdots u_{nn}.$$

*Todistus.* Koska lauseen 2.2 mukaan matriisin  $L$  diagonaalialkiot ovat ykkösiä, lemmän 2.1 avulla saadaan matriisin  $A$  determinantiksi laskettua  $\det(A) = \det(LU) = \underbrace{\det(L)}_{=1} \det(U) = \det(U) = u_{11}u_{22} \cdots u_{nn}$ . □

**Esimerkki 2.7.** Lasketaan esimerkissä 2.3 esiintyvän matriisin  $A$  determinantti. Samassa esimerkissä muodostettiin matriisille  $A$  LU-hajotelma. Nyt lauseen 2.6 avulla matriisin  $A$  determinantiksi saadaan  $\det(A) = \det(U) = 3 \cdot (-12) \cdot (-15) = 540$ . △

### 2.3.4 Muita sovelluksia

LU-hajotelmalla on aiemmin mainittujen lisäksi myös muita sovelluskohteita. Sitä sovelletaan esimerkiksi matriisin häiriöalttiutta selvittäessä tai nauhamatriisien kanssa [1].

Lineaarisia yhtälöryhmiä ratkaistaessa ollaan usein kiinnostuneita, kuinka altis yhtälönratkaisu on alussa esiintyneille virheille. Matriisin häiriöalttiutta määritellään  $\kappa(A) = \|A\| \|A^{-1}\|$ , missä  $\|A\|$  on matriisnormi ja yleisimmillä normeilla  $\kappa(A) \geq 1$ . Jos esimerkiksi  $\kappa = 1$ , niin alussa olleet virheet eivät ole kasvaneet ratkaisussa. Matriisin  $A$  häiriöalttiudelle saadaan LU-hajotelmalla määritettyä helposti yläraja  $\kappa(A) \leq \kappa(L)\kappa(U)$ , joten usein tarkkaa häiriöalttiuden arvoa ei tarvitse määrittää.

Nauhamatriiseja esiintyy monissa sovelluksissa, kuten esimerkiksi osittaisdifferentiaaliyhtälöiden ratkaisuisissa. Niiden erityisen rakenteen ansiosta ohjelmistoissa voidaan säästää sekä laskenta-aikaa että muistia. Neliömatriisia  $A$  sanotaan nauhamatriisiksi, jos  $a_{i,j} = 0$ , kun  $|i - j| > l$ , missä  $l \in \mathbb{N}$ . Nauhamatriisin leveys on tällöin  $2l + 1$  ja esimerkiksi tridiagonaalille matriisille  $l = 1$ . Jos nauhamatriisille voidaan muodostaa LU-hajotelma, hajotelma säilyttää nauhamatriisin rakenteen, ts. myös matriisit  $L$  ja  $U$  ovat nauhamatriiseja.



### 3 QR-hajotelma

Käsitellään seuraavaksi QR-hajotelma, jota hyödynnetään usein esimerkiksi matriisin ominaisarvojen approksimoinnissa tai lineaarisia yhtälöryhmiä ratkaistaessa pienimmän neliösumman menetelmällä. Näiden lisäksi QR-hajotelmalla on myös muita sovelluskohteita, kuten Mooren-Penrosen yleistetyn käänteismatriisin etsiminen tai determinantin itseisarvon laskeminen. QR-hajotelmassa matriisi  $A$  ilmaistaan matriisien  $Q$  ja  $R$  tulona  $A = QR$ , missä matriisin  $Q$  pystyrivit ovat ortonormaalit ja matriisi  $R$  on yläkolmiomatriisi.

QR-hajotelman etuna esimerkiksi LU-hajotelmaan verrattuna on, että sen voi muodostaa mille tahansa matriisille. Hajotelman muodostus voidaan toteuttaa kolmella eri tavalla: Gramin-Schmidtin ortogonalisoinnilla, Householderin peilauksilla tai Givensin rotaatioilla [5]. Tässä tutkielmassa esitellään QR-hajotelman muodostaminen Gramin-Schmidtin ortogonalisoinnilla, johon tutustutaan seuraavaksi tarkemmin. Luvussa 3.1 on hyödynnetty lähdettä [6] ja luvussa 3.2 lähteitä [7, 10].

#### 3.1 Gramin-Schmidtin ortogonalisointi

Gramin-Schmidtin ortogonalisointimenetelmällä voidaan muodostaa aliavaruudelle  $\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathbb{K}^n$  ortonormaalikanta. Tässä luvussa skalaarikuntana on  $\mathbb{K}$ , eli  $\mathbb{K} = \mathbb{R}$  tai  $\mathbb{K} = \mathbb{C}$ . Lisäksi kompleksiluvun  $z$  liittolukua merkitään  $\bar{z}$ . Kun  $\mathbb{K} = \mathbb{R}$ , niin sisätulo määritellään  $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n \mathbf{x}_i \mathbf{y}_i$ . Jos sen sijaan  $\mathbb{K} = \mathbb{C}$ , niin  $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n \mathbf{x}_i \bar{\mathbf{y}}_i$ . Gramin-Schmidtin ortogonalisoinnilla vektorit  $\mathbf{y}_1, \dots, \mathbf{y}_k$  määritellään rekursiivisesti:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{x}_1 \\ \mathbf{y}_j &= \mathbf{x}_j - \sum_{i=1}^{j-1} a_{ji} \mathbf{y}_i \quad (j = 2, \dots, k), \end{aligned}$$

missä

$$a_{ji} = \begin{cases} \frac{\langle \mathbf{x}_j, \mathbf{y}_i \rangle}{\|\mathbf{y}_i\|^2}, & \text{jos } \mathbf{y}_i \neq 0 \\ 0, & \text{jos } \mathbf{y}_i = 0. \end{cases}$$

Näin saadut vektorit  $\mathbf{y}_1, \dots, \mathbf{y}_k$  ovat ortogonaalisia ja  $L(\mathbf{y}_1, \dots, \mathbf{y}_j) = L(\mathbf{x}_1, \dots, \mathbf{x}_j)$  kaikilla  $j = 1, \dots, k$ . Vektoreista  $\mathbf{y}_1, \dots, \mathbf{y}_k$  saadaan vielä ortonormaalikanta poistamalla mahdolliset nollavektorit ja jakamalla jäljelle jääneet vektorit pituuksillaan.

#### 3.2 QR-hajotelma

QR-hajotelman muodostamisessa voidaan hyödyntää Gramin-Schmidtin ortogonalisointia. Todistetaan seuraavaksi QR-hajotelman olemassaolo täysiasteisille  $m \times n$ -matriiseille, missä  $m \geq n$ . Lause ja sen todistus mukailee lähdettä [10].

**Lause 3.1.** (QR-hajotelma) Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ , missä  $m \geq n$  ja aste  $r(A) = n$ . Silloin on olemassa matriisi  $Q \in \mathcal{M}_{m \times n}(\mathbb{K})$ , jonka pystyrit ovat ortonormaalit, ja säännöllinen yläkolmiomatriisi  $R \in \mathcal{M}_{n \times n}(\mathbb{K})$ , joille  $A = QR$ .

*Todistus.* Olkoon matriisin  $A \in \mathcal{M}_{m \times n}(\mathbb{K})$  pystyrit  $\mathbf{a}_1, \dots, \mathbf{a}_n$ , missä  $\mathbf{a}_i \in \mathbb{K}^m$ . Lisäksi, koska  $r(A) = n$ , niin pystyrit ovat lineaarisesti riippumattomat. Joukosta  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  saadaan Gramin-Schmidtin ortogonalisointimenetelmällä ortonormaalijoukko  $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ , missä  $\mathbf{q}_i \in \mathbb{K}^m$ . Nyt  $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$  ja  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  virittävät saman vektoriavaruuden osajoukon, tarkemmin

$$L(\mathbf{a}_1, \dots, \mathbf{a}_k) = L(\mathbf{q}_1, \dots, \mathbf{q}_k) \quad \text{kaikilla } k \in \{1, \dots, n\}.$$

Silloin on olemassa sellaiset skalaarit  $r_{ij}$ , että

$$\begin{aligned} \mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\ \mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\ &\vdots \\ \mathbf{a}_n &= r_{1n}\mathbf{q}_1 + r_{2n}\mathbf{q}_2 + \dots + r_{nn}\mathbf{q}_n. \end{aligned}$$

Tämä voidaan kirjoittaa muodossa

$$(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n) \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & r_{nn} \end{pmatrix}$$

Näin saadaan muodostettua matriisi  $R = (r_{ij}) \in \mathcal{M}_n(\mathbb{K})$ , missä  $i, j = 1, \dots, n$  ja  $r_{ij} = 0$ , kun  $i > j$ . Lisäksi määritellään matriisiksi  $Q = (\mathbf{q}_1, \dots, \mathbf{q}_n) \in \mathcal{M}_{m \times n}(\mathbb{K})$ . Matriisin  $Q$  pystyrit ovat siis ortonormaalit ja  $R$  on yläkolmiomatriisi.

Tutkitaan vielä matriisin  $R$  diagonaalialkioita. Jos  $r_{11} = 0$ , niin silloin  $\mathbf{a}_1 = 0$ , joka on ristiriidassa oletuksen  $r(A) = n$  kanssa. Näin ollen  $r_{11} \neq 0$ . Jos sen sijaan  $r_{22} = 0$ , niin silloin  $\mathbf{a}_1$  ja  $\mathbf{a}_2$  ovat lineaarisesti riippuvia keskenään, joka ei ole myöskään oletuksen mukaan mahdollista. Jatkamalla samaa päättelyä havaitaan, että  $r_{ii} \neq 0$  kaikilla  $i = 1, \dots, n$ . Näin ollen matriisi  $R$  on myös säännöllinen, sillä lemmän 2.1 avulla  $\det(R) \neq 0$ .  $\square$

Edellisessä todistuksessa matriisi  $Q$  saatiin muodostettua Gramin-Schmidtin ortogonalisoinnilla, jolloin sen pystyrit ovat ortonormaalit. Tutkitaan vielä tarkemmin tällaisen matriisin ominaisuuksia.

**Määritelmä 3.2.** Matriisin  $Q = (q_{ij})_{m \times n} \in \mathcal{M}_{m \times n}(\mathbb{C})$  adjungoiduksi matriisiksi kutsutaan matriisia  $Q^* = (\bar{q}_{ji})_{n \times m}$ .

Edellisestä määritelmästä seuraa, että  $Q^* = \bar{Q}^T$ , missä  $\bar{Q} = (\bar{q}_{ij})_{m \times n}$ . Reaaliselle matriisille pätee siis  $Q^* = Q^T$ . Koska matriisin  $Q$  pystyrit ovat ortonormaalit, niin silloin  $Q^*Q = I$ . Kuitenkaan ei ole välttämätöntä, että  $QQ^* = I$ . Tilanne muuttuu, jos lauseen 3.1 mukainen matriisi  $A$  on neliömatriisi, jolloin myös matriisi  $Q$  on neliömatriisi.

**Määritelmä 3.3.** Matriisia  $Q^* \in \mathcal{M}_n(\mathbb{C})$  sanotaan unitaariseksi, jos sen pystyivät ovat ortonormaalit. Silloin sille pätee  $Q^*Q = QQ^* = I$ .

**Huomautus 3.4.** Reaalisessa tapauksessa unitaariselle matriisille pätee  $Q^TQ = QQ^T = I$ , ja sitä kutsutaan ortogonaaliseksi matriisiksi.

Edellä olevien määritelmien avulla yläkolmiomatriisi  $R$  voidaan ratkaista helposti. Matriisi  $R$  saadaan ratkaistua kertomalla yhtälö  $A = QR$  vasemmalta matriisilla  $Q^*$ , jolloin saadaan  $R = Q^*A$ . Käydään vielä QR-hajotelman muodostamisesta läpi esimerkki.

**Esimerkki 3.5.** Muodostetaan QR-hajotelma matriisille

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

Matriisin pystyivät  $\mathbf{a}_1 = (0, 1, 1)$  ja  $\mathbf{a}_2 = (1, 0, 2)$  ovat lineaarisesti riippumattomia. Gramin-Schmidtin ortogonalisoinnilla saadaan ortogonaaliset vektorit

$$\begin{aligned} \mathbf{q}'_1 &= \mathbf{a}_1 = (0, 1, 1) \quad \text{ja} \\ \mathbf{q}'_2 &= \mathbf{a}_2 - \frac{\langle \mathbf{a}_2, \mathbf{q}'_1 \rangle}{\|\mathbf{q}'_1\|^2} \cdot \mathbf{q}'_1 = (1, 0, 2) - \frac{2}{2} \cdot (0, 1, 1) = (1, -1, 1). \end{aligned}$$

Nämä vielä normittamalla saadaan matriisin  $Q$  pystyivät eli

$$\begin{aligned} \mathbf{q}_1 &= \frac{\mathbf{q}'_1}{\|\mathbf{q}'_1\|} = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad \text{ja} \quad \mathbf{q}_2 = \frac{\mathbf{q}'_2}{\|\mathbf{q}'_2\|} = \left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \\ \Rightarrow Q &= \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \end{aligned}$$

Koska matriisin  $Q$  pystyivät ovat ortonormaalit, niin  $Q^*Q = I$ . Kertomalla  $A = QR$  vasemmalta matriisilla  $Q^*$  saadaan

$$R = Q^*A = Q^T A = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{pmatrix},$$

jolloin  $A = QR$ . △

Aiemmin lauseessa 3.1 osoitettiin QR-hajotelman olemassaolo täysiasteisille  $m \times n$ -matriiseille ( $m \geq n$ ). On kuitenkin huomattava, että Gramin-Schmidtin ortogonalisoinnilla voidaan muodostaa QR-hajotelma myös matriiseille, jotka eivät ole täysiasteisia. Kun matriisin  $A$  pystyivät ovat lineaarisesti riippuvia, Gramin-Schmidtin ortogonalisoinnissa jokin  $\mathbf{q}_j = 0$ . Tällöin pystyivi  $\mathbf{q}_j$  tulee valita siten, että se on ortogonaalinen jokaisen pystyivin  $\mathbf{q}_i$  kanssa, missä  $i < j$ . Tämän jälkeen ortogonalisointia voidaan jatkaa normaalisti. Näin menetelmä ei kuitenkaan ole enää kovin tehokas, joten Gramin-Schmidtin ortogonalisointia sovelletaan lähinnä täysiasteisille matriiseille.

**Esimerkki 3.6.** Muodostetaan QR-hajotelma matriisille

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 0 \end{pmatrix},$$

joka ei ole täysiasteinen. Gramin-Schmidtin ortogonalisoinnilla

$$\begin{aligned} \mathbf{q}'_1 &= \mathbf{a}_1 = (0, 1, 1) \text{ ja} \\ \mathbf{q}'_2 &= \mathbf{a}_2 - \frac{\langle \mathbf{a}_2, \mathbf{q}'_1 \rangle}{\|\mathbf{q}'_1\|^2} \cdot \mathbf{q}'_1 = (0, 2, 2) - \frac{4}{2} \cdot (0, 1, 1) = (0, 0, 0). \end{aligned}$$

Koska  $\mathbf{q}'_2$  on nollavektori, vektori  $\mathbf{q}'_2$  tulee valita siten, että se on ortogonaalinen pystyriivien  $\mathbf{q}'_1$  kanssa. Pystyriiviksi  $\mathbf{q}'_2$  voidaan valita esimerkiksi vektori  $(1, 0, 0)$ . Nyt jatkamalla ortogonalisointia saadaan pystyriiviksi  $\mathbf{q}'_3$

$$\begin{aligned} \mathbf{q}'_3 &= \mathbf{a}_3 - \frac{\langle \mathbf{a}_3, \mathbf{q}'_1 \rangle}{\|\mathbf{q}'_1\|^2} \cdot \mathbf{q}'_1 - \frac{\langle \mathbf{a}_3, \mathbf{q}'_2 \rangle}{\|\mathbf{q}'_2\|^2} \cdot \mathbf{q}'_2 \\ &= (1, 1, 0) - \frac{1}{2} \cdot (0, 1, 1) - 1 \cdot (1, 0, 0) \\ &= \left( 0, \frac{1}{2}, -\frac{1}{2} \right), \end{aligned}$$

joka on ortogonaalinen pystyriivien  $\mathbf{q}'_1$  ja  $\mathbf{q}'_2$  kanssa. Nämä vielä normittamalla saadaan matriisiksi

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Koska matriisin  $Q$  pystyriivit ovat ortonormaalit, niin  $Q^T Q = I$ . Kertomalla  $A = QR$  vasemmalta matriisilla  $Q^T$  saadaan

$$R = Q^T A = \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 2\sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & 0 & 1 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{pmatrix},$$

jolloin  $A = QR$ . △

### 3.3 QR-hajotelman sovelluksia

Tarkastellaan seuraavaksi, millaisissa tilanteissa QR-hajotelmaa voidaan hyödyntää. Sovelluksia käsittelevissä luvuissa on hyödynnetty lähteitä [1, 3, 4, 7, 8].

#### 3.3.1 Lineaaristen yhtälöryhmien ratkaiseminen

Myös QR-hajotelmaa voidaan hyödyntää lineaaristen yhtälöryhmien ratkaisemisessa. LU- ja QR-hajotelmia voidaankin käyttää lähes samalla tavalla lineaarisille yhtälöryhmille  $A\mathbf{x} = \mathbf{b}$ , missä matriisi  $A$  on säännöllinen. On kuitenkin huomattava,

että tämän kaltaisessa tilanteessa LU-menetelmä on nopeampi ratkaisemaan ongelman. Tilanne muuttuu kuitenkin, kun  $A$  on singulaarinen tai se ei ole neliömatriisi, jolloin yhtälö  $A\mathbf{x} = \mathbf{b}$  voi olla ristiriitainen. QR-hajotelmalla voidaan määrittää yhtälölle  $A\mathbf{x} = \mathbf{b}$  pienimmän neliösumman ratkaisu, kun matriisi  $A$  on täysiasteinen. Tässä alaluvussa skalaarikuntana on  $\mathbb{R}$ .

Matriisin  $A$  ollessa säännöllinen, lineaarinen yhtälöryhmä  $A\mathbf{x} = \mathbf{b}$  voidaan ratkaista QR-hajotelmalla samankaltaisella idealla kuin LU-hajotelmalla luvussa 2.3.1. QR-hajotelman avulla se voidaan kirjoittaa muotoon

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \Leftrightarrow (QR)\mathbf{x} &= \mathbf{b} \quad \| \cdot Q^T \text{ (vas.)} \\ \Leftrightarrow \underbrace{Q^T Q}_{=I} R\mathbf{x} &= Q^T \mathbf{b} \\ \Leftrightarrow R\mathbf{x} &= Q^T \mathbf{b}. \end{aligned}$$

Yhtälöryhmä voidaan ratkaista QR-hajotelman avulla seuraavalla tavalla.

- Ratkaistaan  $Q^T$  matriisin  $Q$  avulla.
- Lasketaan yhtälön oikea puoli:  $Q^T \mathbf{b}$ .
- Ratkaistaan  $\mathbf{x}$  yhtälöstä  $R\mathbf{x} = Q^T \mathbf{b}$ .

**Esimerkki 3.7.** Ratkaistaan sama yhtälöryhmä kuin esimerkissä 2.4 eli

$$\begin{cases} 3x_1 + 3x_2 - 3x_3 = 1 \\ 18x_1 + 6x_2 + 3x_3 = 2 \\ 24x_1 + 3x_3 = 4, \end{cases}$$

Yhtälöryhmästä saadaan muodostettua matriisi

$$A = \begin{pmatrix} 3 & 3 & -3 \\ 18 & 6 & 3 \\ 24 & 0 & 3 \end{pmatrix},$$

jonka pystyivät ovat lineaarisesti riippumattomat. Gramin-Schmidtin ortogonali-soinnilla saadaan ratkaistua matriisi  $Q$  ja yläkolmiomatriisi  $R$  muodosta  $R = Q^T A$ , jolloin

$$Q = \begin{pmatrix} \frac{1}{\sqrt{101}} & \frac{22}{\sqrt{2121}} & \frac{-4}{\sqrt{21}} \\ \frac{6}{\sqrt{101}} & \frac{31}{\sqrt{2121}} & \frac{2}{\sqrt{21}} \\ \frac{8}{\sqrt{101}} & \frac{-26}{\sqrt{2121}} & \frac{-1}{\sqrt{21}} \end{pmatrix} \text{ ja } R = \begin{pmatrix} 3\sqrt{101} & \frac{39}{\sqrt{101}} & \frac{39}{\sqrt{101}} \\ 0 & 12\sqrt{\frac{21}{101}} & -17\sqrt{\frac{3}{707}} \\ 0 & 0 & 5\sqrt{\frac{3}{7}} \end{pmatrix}$$

Lisäksi matriisin  $Q$  transpoosiksi saadaan

$$Q^T = \begin{pmatrix} \frac{1}{\sqrt{101}} & \frac{6}{\sqrt{101}} & \frac{8}{\sqrt{101}} \\ \frac{22}{\sqrt{2121}} & \frac{31}{\sqrt{2121}} & \frac{-26}{\sqrt{2121}} \\ \frac{-4}{\sqrt{21}} & \frac{2}{\sqrt{21}} & \frac{-1}{\sqrt{21}} \end{pmatrix}.$$

Ratkaistaan yhtälöryhmä  $R\mathbf{x} = Q^T\mathbf{b}$ , missä  $\mathbf{b} = (1 \ 2 \ 4)^T$ . Lasketaan  $Q^T\mathbf{b}$ , jonka jälkeen ratkaistaan vektori  $\mathbf{x}$  samalla idealla kuin esimerkissä 2.4

$$\underbrace{\begin{pmatrix} 3\sqrt{101} & \frac{39}{\sqrt{101}} & \frac{39}{\sqrt{101}} \\ 0 & 12\sqrt{\frac{21}{101}} & -17\sqrt{\frac{3}{707}} \\ 0 & 0 & 5\sqrt{\frac{3}{7}} \end{pmatrix}}_R \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} \frac{45}{\sqrt{101}} \\ \frac{-20}{\sqrt{2121}} \\ \frac{-4}{\sqrt{21}} \end{pmatrix}}_{Q^T\mathbf{b}},$$

jolloin tehtävän ratkaisuksi saadaan  $\mathbf{x} = (\frac{1}{5} \ \frac{-2}{15} \ \frac{-4}{15})^T$ . △

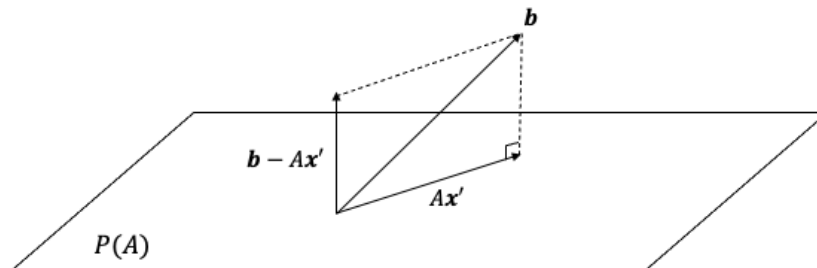
Kuten aiemmin jo mainittiin, yhtälöllä  $A\mathbf{x} = \mathbf{b}$  ei aina ole olemassa ratkaisua. Tällöin yhtälöä sanotaan ristiriitaiseksi. Pienimmän neliösumman menetelmässä (lyh. PNS) ideana on löytää ratkaisu, joka toteuttaa yhtälön mahdollisimman hyvin. Toisin sanoen menetelmällä yritetään etsiä sellainen vektori  $\mathbf{x}'$ , että  $A\mathbf{x}'$  on mahdollisimman lähellä vektoria  $\mathbf{b}$ . On kuitenkin huomattava, että jos lineaarisella yhtälöryhmällä  $A\mathbf{x} = \mathbf{b}$  on olemassa ratkaisu, niin silloin se on myös PNS-menetelmän ratkaisu. PNS-menetelmää sovelletaan paljon esimerkiksi tilastotieteessä ja sitä käytetään mm. kuvaajien suorien sovittamiseen. Ennen kuin käsitellään QR-hajotelman hyötyjä, käydään lyhyesti läpi PNS-menetelmä.

**Määritelmä 3.8.** Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  ja vektori  $\mathbf{b} \in \mathbb{R}^m$ . PNS-menetelmän ratkaisu yhtälölle  $A\mathbf{x} = \mathbf{b}$  on sellainen  $\mathbf{x}' \in \mathbb{R}^n$ , että

$$\|\mathbf{b} - A\mathbf{x}'\| = \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{b} - A\mathbf{x}\|,$$

missä  $\|\cdot\|$  on euklidinen vektorinormi eli vektorin pituus.

Tarkastellaan PNS-menetelmää geometrisesta näkökulmasta. Merkitään matriisin  $A$  pystyriviavaruutta  $P(A)$ . Kun yhtälöllä  $A\mathbf{x} = \mathbf{b}$  ei ole ratkaisua, niin silloin  $\mathbf{b} \notin P(A)$ . Tällöin pyritään etsimään sellainen  $\mathbf{x}' \in \mathbb{R}^n$ , että  $A\mathbf{x}' \in P(A)$  on mahdollisimman lähellä vektoria  $\mathbf{b}$ . Lähin ratkaisu matriisin  $A$  pystyriviavaruudessa on vektorin  $\mathbf{b}$  ortogonaaliprojektio  $A\mathbf{x}' = \text{proj}_{P(A)}\mathbf{b}$ . Alla olevassa kuvassa 1 on havainnollistettuna tilanne.



Kuva 1: PNS-menetelmällä yhtälön  $A\mathbf{x} = \mathbf{b}$  ratkaisuksi saadaan  $\mathbf{x}'$  niin, että  $A\mathbf{x}' = \text{proj}_{P(A)}\mathbf{b}$ . Kuvassa on hyödynnetty lähdeettä [1].

Nyt vektori  $\mathbf{b} - A\mathbf{x}'$  on ortogonaalinen avaruuden  $P(A)$  kanssa. Näin ollen siis jokainen matriisin  $A$  pystyriivi  $(\mathbf{a}_1, \dots, \mathbf{a}_k)$  on kohtisuorassa vektorin  $\mathbf{b} - A\mathbf{x}'$  kanssa eli  $\langle \mathbf{b} - A\mathbf{x}', \mathbf{a}_j \rangle = 0$  kaikilla  $j = 1, \dots, k$ . Vektorien sisätulo voidaan tulkita matriisitulona  $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T \mathbf{x}$ , kun  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ . Näin ollen sisätulo voidaan kirjoittaa muodossa  $\mathbf{a}_j^T (\mathbf{b} - A\mathbf{x}') = 0$ . Nyt koska vektorit  $\mathbf{a}_j^T$  ovat matriisin  $A^T$  vaakarivejä, saadaan

$$\begin{aligned} A^T (\mathbf{b} - A\mathbf{x}') &= \mathbf{0} \\ \Leftrightarrow A^T \mathbf{b} - A^T A\mathbf{x}' &= \mathbf{0} \\ \Leftrightarrow A^T A\mathbf{x}' &= A^T \mathbf{b}. \end{aligned}$$

Jokainen PNS-menetelmän ratkaisu yhtälölle  $A\mathbf{x} = \mathbf{b}$  toteuttaa siis yhtälön

$$A^T A\mathbf{x} = A^T \mathbf{b},$$

jota kutsutaan normaaliyhtälöksi.

QR-hajotelman etuna PNS-menetelmässä on kolmiomatriisi, jonka avulla yhtälöiden ratkaiseminen nopeutuu matriisikoon kasvaessa. Tarkastellaan seuraavaksi normaaliyhtälöä QR-hajotelmalla. Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  täysiasteinen. Silloin

$$\begin{aligned} A^T A\mathbf{x} &= A^T \mathbf{b} \\ \Leftrightarrow (QR)^T QR\mathbf{x} &= (QR)^T \mathbf{b} \\ \Leftrightarrow R^T \underbrace{Q^T Q}_{=I} R\mathbf{x} &= R^T Q^T \mathbf{b} \\ \Leftrightarrow R^T R\mathbf{x} &= R^T Q^T \mathbf{b}. \end{aligned}$$

Lauseen 3.1 mukaan yläkolmiomatriisi  $R$  on säännöllinen neliömatriisi. Näin ollen kerrotaan yhtälö vasemmalta matriisilla  $(R^T)^{-1}$ , jolloin

$$\begin{aligned} \underbrace{(R^T)^{-1} R^T}_{=I} R\mathbf{x} &= \underbrace{(R^T)^{-1} R^T}_{=I} Q^T \mathbf{b} \\ \Leftrightarrow R\mathbf{x} &= Q^T \mathbf{b}. \end{aligned}$$

Normaaliyhtälö sievenee QR-hajotelmalla siis täysin samaan muotoon kuin ratkaistaessa aiemmin lineaarisia yhtälöryhmiä, joilla on olemassa ratkaisu. Näin ollen siis QR-hajotelmalla täysiasteiselle matriisille  $A$  lineaarisen yhtälöryhmän ratkaisu PNS-menetelmällä on muotoa  $R\mathbf{x} = Q^T \mathbf{b}$  riippumatta siitä onko yhtälöllä  $A\mathbf{x} = \mathbf{b}$  ratkaisua vai ei.

**Lause 3.9.** Olkoon matriisilla  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  lauseen 3.1 mukainen QR-hajotelma. Tällöin yhtälöllä  $A\mathbf{x} = \mathbf{b}$  on olemassa pienimmän neliösumman ratkaisu muotoa  $\mathbf{x}' = R^{-1} Q^T \mathbf{b}$ , missä  $\mathbf{b} \in \mathbb{R}^m$ .

**Huomautus 3.10.** Koska lauseen 3.1 mukaan  $R$  on yläkolmiomatriisi, on vektori  $\mathbf{x}'$  nopeampaa ratkaista muodosta  $R\mathbf{x} = Q^T \mathbf{b}$  kuin ratkaista matriisi  $R^{-1}$  ja laskea lauseen 3.9 mukaisella yhtälöllä.

**Esimerkki 3.11.** Etsitään ristiriitaiselle yhtälölle  $A\mathbf{x} = \mathbf{b}$  ratkaisua QR-hajotelman avulla käyttämällä PNS-menetelmällä, kun

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}.$$

Esimerkistä 3.5 saadaan matriisille  $A$  QR-hajotelma, jossa

$$Q^T = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \quad R = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{pmatrix}.$$

Ratkaistaan  $\mathbf{x}'$  käyttämällä huomautuksen 3.10 mukaista kaavaa  $R\mathbf{x}' = Q^T\mathbf{b}$ . Näin ollen yhtälö on muotoa

$$\begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3\sqrt{2} \\ \frac{4}{\sqrt{3}} \end{pmatrix},$$

jolloin ratkaisuksi saadaan  $\mathbf{x}' = \left(\frac{5}{3} \quad \frac{4}{3}\right)^T$ . △

### 3.3.2 QR-algoritmi

Tässä luvussa esitellään QR-algoritmi, jonka avulla neliömatriisille voidaan löytää Schurin hajotelma. Schurin hajotelmasta saadaan määritettyä matriisin  $A$  ominaisarvot. Ominaisarvojen laskeminen on yksi numeerisen lineaarialgebran tärkeimpiä ongelmia, ja sitä sovelletaan monilla aloilla, kuten esimerkiksi tietojenkäsittelyteissä. QR-algoritmiin pohjautuvat menetelmät ovat käytetyimpiä tapoja ratkaista ominaisarvot. Ennen QR-algoritmin tutkimista esitellään Schurin lause (lause 3.14), josta saadaan Schurin hajotelma. Schurin lauseen todistus sivuutetaan ja sen voi lukea lähteestä [1].

**Määritelmä 3.12.** Neliömatriisit  $A$  ja  $B$  ovat similaarisia, jos on olemassa sellainen neliömatriisi  $P$ , että  $A = PBP^{-1}$ .

Similaarisilla matriiseilla on oltava samat ominaisarvot, sillä  $\det(A - \lambda I) = \det(PBP^{-1} - \lambda I) = \det(P)\det(B - \lambda I)\det(P^{-1}) = \det(B - \lambda I)$ . Jos sen sijaan matriisi  $P$  voidaan valita siten, että se on unitaarinen, matriiseja  $A$  ja  $B$  sanotaan unitaarisesti similaarisiksi.

**Määritelmä 3.13.** Matriisit  $A, B \in \mathcal{M}_n(\mathbb{C})$  ovat unitaarisesti similaarisia, jos on olemassa sellainen unitaarinen matriisi  $P \in \mathcal{M}_n(\mathbb{C})$ , että  $A = PBP^*$ .

Vastaavasti reaaliset neliömatriisit  $A, B \in \mathcal{M}_n(\mathbb{R})$  ovat ortogonaalisesti similaariset, jos on olemassa sellainen ortogonaalinen matriisi  $P \in \mathcal{M}_n(\mathbb{R})$ , että  $A = PBP^T$ . Myös näillä matriiseilla ominaisarvot ovat samat.

**Lause 3.14.** (Schurin lause) Jokainen neliömatriisi  $A \in \mathcal{M}_n(\mathbb{C})$  on unitaarisesti similaarinen yläkolmiomatriisin kanssa, eli  $A = UTU^*$ , missä yläkolmiomatriisin  $T$  diagonaalilla on matriisin  $A$  ominaisarvot. Yhtälöä  $A = UTU^*$  sanotaan Schurin hajotelmaksi.



**Huomautus 3.15.** Schurin lause voidaan ilmaista myös, kun skalaarikuntana on  $\mathbb{R}$ . Jos matriisin  $A \in \mathcal{M}_n(\mathbb{R})$  ominaisarvot ovat reaaliset, se on ortogonaalisesti similaarinen yläkolmiomatriisiin  $T \in \mathcal{M}_n(\mathbb{R})$  kanssa. Schurin hajotelma voidaan siis kirjoittaa muodossa  $A = UTU^T$ .

Tarkastellaan seuraavaksi, millaisissa olosuhteissa Schurin hajotelmaa voidaan hyödyntää ominaisarvojen selvittämisessä. Vaikka QR-algoritmia voidaan käyttää hyvin myös kompleksisille matriiseille, tutkitaan tästä eteenpäin vain säännöllisiä ja reaalisia neliömatriiseja  $A$ , joiden ominaisarvot ovat reaalisia sekä toteuttavat ehdon

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0. \quad (1)$$

Koska ominaisarvot ovat erisuuret, niitä vastaavat ominaisvektorit  $\mathbf{v}_i$  ovat lineaarisesti riippumattomia. Silloin matriisin  $A$  voidaan osoittaa olevan diagonalisoituva ja diagonaalimatriisin diagonaalilla ovat matriisin  $A$  ominaisarvot järjestyksessä.

**Lause 3.16.** Olkoon matriisi  $A \in \mathcal{M}_n(\mathbb{R})$ , jonka ominaisarvot  $\lambda_1, \lambda_2, \dots, \lambda_n$  toteuttavat ehdon (1). Silloin on olemassa sellainen säännöllinen matriisi  $P$ , että matriisi  $A$  on diagonalisoituva. Tämä voidaan kirjoittaa muodossa

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = D.$$

*Todistus.* Olkoon matriisin  $A$  ominaisvektorit  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ . Merkitään nyt matriisin  $P$  pystyriveiksi  $P = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ . Koska ehdon (1) toteutuessa matriisin  $A$  ominaisvektorit  $\mathbf{v}_i$  ovat lineaarisesti riippumattomia, matriisi  $P$  on säännöllinen. Tällöin saadaan

$$\begin{aligned} AP &= A(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = (A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_n) \\ &= (\lambda_1\mathbf{v}_1, \lambda_2\mathbf{v}_2, \dots, \lambda_n\mathbf{v}_n) \\ &= (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = PD. \end{aligned}$$

Kertomalla yhtälö vasemmalta matriisilla  $P^{-1}$  saadaan  $D = P^{-1}AP$ . □

Tutustutaan nyt tarkemmin QR-algoritmin ideaan. Lauseen 3.1 mukaan matriisille  $A$  on olemassa QR-hajotelma, missä matriisi  $Q$  on ortogonaalinen ja matriisi  $R$

on yläkolmiomatriisi. QR-algoritmissa muodostetaan matriisit  $A_k$  seuraavasti:

$$\begin{aligned}
 A_1 &= Q_1 R_1 & \| \cdot Q_1^T \text{ (vas.)} \cdot Q_1 \text{ (oik.)} \\
 \underbrace{Q_1^T A_1 Q_1}_{=A_2} &= R_1 Q_1 \\
 A_2 &= Q_2 R_2 & \| \cdot Q_2^T \text{ (vas.)} \cdot Q_2 \text{ (oik.)} \\
 A_3 &= R_2 Q_2 \\
 &\vdots \\
 A_k &= Q_k R_k \\
 A_{k+1} &= R_k Q_k.
 \end{aligned}$$

Näin ollen matriisi  $A_{k+1}$  muodostuu matriisin  $A_k$  QR-hajotelman tekijöistä.

**Määritelmä 3.17.** Olkoon matriisi  $A \in \mathcal{M}_n(\mathbb{R})$ , joka toteuttaa ehdon (1). Merkitään  $k \geq 1$  ja  $A = A_1$ . Silloin ominaisarvojen laskemiseen käytetty QR-algoritmi koostuu matriisijonosta  $\{A_{k+1}\} = \{R_k Q_k\}$ , missä  $Q_k$  ja  $R_k$  ovat säännöllisen matriisin  $A_k$  QR-hajotelman tekijät.

Osoitetaan seuraavaksi, että QR-algoritmissa matriisit  $A$  ja  $A_{k+1}$  ovat ortogonaalisesti similaarisia. Merkitään  $P_k = Q_1 Q_2 \cdots Q_k$ , jolloin jokainen  $P_k$  on sellainen ortogonaalinen matriisi, että

$$\begin{aligned}
 P_1^T A P_1 &= Q_1^T \underbrace{Q_1 R_1 Q_1}_{=A} = A_2 \\
 P_2^T A P_2 &= Q_2^T \underbrace{Q_1^T A Q_1}_{=A_2} Q_2 = A_3 \\
 &\vdots \\
 P_k^T A P_k &= A_{k+1}.
 \end{aligned}$$

**Lause 3.18.** Olkoon matriisi  $A \in \mathcal{M}_n(\mathbb{R})$ , joka toteuttaa ehdon (1). Silloin jokainen matriisi  $A_k$  on ortogonaalisesti similaarinen matriisin  $A$  kanssa kaikilla  $k \geq 2$ .

Nyt tiedetään, että QR-algoritmissa matriisit  $A_k$  ovat ortogonaalisesti similaarisia, eli ominaisarvot pysyvät samoina. Osoitetaan vielä seuraavassa lauseessa, että matriisijono lähestyy kohti Schurin hajotelmaa. Ennen tämän osoittamista tulee kuitenkin vielä määrittää todistusta varten QR-hajotelma matriisille  $A^k$ . Merkitään yläkolmiomatriisi  $U_k = R_k \cdots R_1$ , jolloin

$$\begin{aligned}
 P_k U_k &= (Q_1 \cdots Q_k) (R_k \cdots R_1) = (Q_1 \cdots Q_{k-1}) A_k (R_{k-1} \cdots R_1) \\
 &= (Q_1 \cdots Q_{k-1}) (Q_{k-1}^T \cdots Q_1^T A Q_1 \cdots Q_{k-1}) (R_{k-1} \cdots R_1) \\
 &= A (Q_1 \cdots Q_{k-1}) (R_{k-1} \cdots R_1) = A P_{k-1} U_{k-1} = \cdots = A^k.
 \end{aligned}$$

Seuraavan lauseen todistuksessa on hyödynnetty lähteitä [1, 3].

**Lause 3.19.** Olkoon matriisi  $A \in \mathcal{M}_n(\mathbb{R})$ , joka toteuttaa ehdon (1). Lisäksi oletetaan, että lauseen 3.16 mukaiselle matriisille  $P^{-1}$  voidaan muodostaa LU-hajotelma. Silloin matriisijono  $\{A_k\}$  lähestyy kohti yläkolmiomatriisia, jonka diagonaalilla ovat matriisin  $A$  ominaisarvot.

*Todistus.* Koska matriisi  $A$  toteuttaa ehdon (1), se on lauseen 3.16 mukaan diagonalisoituva. Olkoon  $P$  säännöllinen matriisi, jolle  $P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n) = D$ . Silloin matriisille  $A^k$  pätee  $A^k = PD^kP^{-1}$ . Oletetaan, että matriisilla  $P^{-1}$  on olemassa lauseen 2.2 mukainen LU-hajotelma  $P^{-1} = LU$ . Nyt

$$A^k = PD^kLU = P(D^kLD^{-k})D^kU, \quad (2)$$

missä alakolmiomatriisin  $D^kLD^{-k}$  alkiot ovat

$$(D^kLD^{-k})_{i,j} = \begin{cases} 0, & \text{kun } i < j, \\ 1, & \text{kun } i = j, \\ \left(\frac{\lambda_i}{\lambda_j}\right)^k l_{i,j}, & \text{kun } i > j. \end{cases}$$

Ehdon (1) mukaan  $|\lambda_j| > |\lambda_i| > 0$ , kun  $i > j$ . Tästä seuraa, että  $\lim_{k \rightarrow \infty} (D^kLD^{-k})_{i,j} = 0$ , kun  $i > j$ . Näin ollen matriisi  $D^kLD^{-k}$  lähestyy kohti identiteettimatriisiä, joka voidaan kirjoittaa  $D^kLD^{-k} = I + E_k$ , missä  $\lim_{k \rightarrow \infty} E_k = 0$ .

Aiemmin matriisi  $P$  määriteltiin säännölliseksi. Näin ollen matriisi  $P$  on myös täysiasteinen ja sillä on lauseen 3.1 mukainen QR-hajotelma. Sijoittamalla yhtälöön (2)  $P = QR$  saadaan

$$\begin{aligned} A^k &= QR(D^kLD^{-k})D^kU = QR(I + E_k)D^kU \\ &= QR(I + E_k)R^{-1}RD^kU = Q(R + RE_k)R^{-1}RD^kU \\ &= Q(I + RE_kR^{-1})RD^kU. \end{aligned}$$

Kun  $k$  on tarpeeksi suuri, matriisi  $(I + RE_kR^{-1})$  on säännöllinen ja myös sillä on olemassa lauseen 3.1 mukainen QR-hajotelma, jota merkitään  $(I + RE_kR^{-1}) = \tilde{Q}_k\tilde{R}_k$ . Sijoittamalla tämä edelliseen yhtälöön saadaan

$$\begin{aligned} A^k &= Q(\tilde{Q}_k\tilde{R}_k)RD^kU \\ &= (Q\tilde{Q}_k)(\tilde{R}_kRD^kU), \end{aligned}$$

missä  $Q\tilde{Q}_k$  on ortogonaalinen matriisi ja  $\tilde{R}_kRD^kU$  on yläkolmiomatriisi. Ennen todistusta saatiin jo, että  $A^k = P_kU_k$ , joten matriisilla  $A^k$  on kaksi eri QR-hajotelmaa. Tutkitaan tämänkaltaista tilannetta tarkemmin.

Oletetaan, että jollakin matriisilla on kaksi eri QR-hajotelmaa, eli  $Q_1R_1 = Q_2R_2$ . Silloin kertomalla yhtälö vasemmalta matriisilla  $Q_2^T$  ja oikealta matriisilla  $R_1^{-1}$  saadaan

$$Q_2^TQ_1 = R_2R_1^{-1}, \quad (3)$$

missä  $Q_2^TQ_1$  on ortogonaalinen matriisi ja  $R_2R_1^{-1}$  on yläkolmiomatriisi. Toisaalta kertomalla alkuperäinen yhtälö vasemmalta matriisilla  $Q_1^T$  ja oikealta matriisilla  $R_2^{-1}$  saadaan  $Q_1^TQ_2 = R_1R_2^{-1}$ , missä matriisi  $R_1R_2^{-1}$  on yläkolmiomatriisi. Ottamalla tästä transpoosi saadaan

$$\begin{aligned} (Q_1^TQ_2)^T &= (R_1R_2^{-1})^T \\ \Leftrightarrow Q_2^TQ_1 &= (R_1R_2^{-1})^T, \end{aligned}$$

missä  $(R_1 R_2^{-1})^T$  on alakolmiomatriisi. Silloin siis myös matriisin  $Q_2^T Q_1$  on oltava alakolmiomatriisi. Toisaalta yhtälön (3) mukaan  $Q_2^T Q_1$  on yläkolmiomatriisi. Näin sen on oltava diagonaalimatriisi, jota merkitään kirjaimella  $D$ . Nyt  $Q_2^T Q_1 = R_2 R_1^{-1} = D$ , josta saadaan ratkaistua  $Q_1 = Q_2 D$  ja  $R_1 = D^{-1} R_2$ . Lisäksi matriisin  $D$  diagonaalialkiot ovat  $|d_{i,i}| = 1$ , sillä

$$I = Q_1^T Q_1 = (Q_2 D)^T Q_2 D = D^T Q_2^T Q_2 D = D^T D = D^2.$$

Hyödynnetään edellä saatua tietoa matriisin  $A^k$  QR-hajotelmille. Äskeisen perusteella

$$P_k = Q \tilde{Q}_k \tilde{D}_k \quad \text{ja} \quad U_k = \tilde{D}_k^{-1} \tilde{R}_k R D^k U,$$

missä  $\tilde{D}_k$  on diagonaalimatriisi, jolle  $|\tilde{d}_{k,i}| = 1 \quad \forall i = 1, \dots, n$ . Lauseen 3.18 mukaan matriisit  $A$  ja  $A_{k+1}$  ovat ortogonaalisesti similaarisia, jolloin

$$\begin{aligned} A_{k+1} &= P_k^T A P_k = \left( Q \tilde{Q}_k \tilde{D}_k \right)^T A \left( Q \tilde{Q}_k \tilde{D}_k \right) \\ &= \tilde{D}_k \tilde{Q}_k^T (Q^T A Q) \tilde{Q}_k \tilde{D}_k. \end{aligned}$$

Aiemmin määriteltiin, että  $P = QR$ . Tästä voidaan ratkaista  $Q^T = RP^{-1}$  ja  $Q = PR^{-1}$ . Sijoittamalla nämä aiempaan saadaan

$$A_{k+1} = \tilde{D}_k \tilde{Q}_k^T R P^{-1} A P R^{-1} \tilde{Q}_k \tilde{D}_k.$$

Lauseen 3.16 mukaan matriisi  $A$  on diagonalisoituva, eli  $D = P^{-1} A P$ . Sijoittamalla matriisi  $D$  yhtälöön saadaan

$$A_{k+1} = \tilde{D}_k \tilde{Q}_k^T (R D R^{-1}) \tilde{Q}_k \tilde{D}_k,$$

missä matriisi  $R D R^{-1}$  on yläkolmiomatriisi, jonka diagonaalilla ovat ominaisarvot  $\lambda_1, \dots, \lambda_n$ .

Aiemmin määriteltiin, että matriisin  $(I + R E_k R^{-1})$  QR-hajotelma on muotoa  $\tilde{Q}_k \tilde{R}_k = I + R E_k R^{-1}$ . Koska  $\lim_{k \rightarrow \infty} E_k = 0$ , niin silloin matriisi  $(I + R E_k R^{-1})$  lähestyy kohti identiteettimatriisia, eli  $\lim_{k \rightarrow \infty} (I + R E_k R^{-1}) = I$ . Nyt siis  $\tilde{R}_k = \tilde{Q}_k^T (I + R E_k R^{-1}) \rightarrow \tilde{Q}_k^T$ , kun  $k \rightarrow \infty$ . Tällöin matriisi  $\tilde{Q}_k^T$  on ortogonaalinen ja yläkolmiomatriisi. Aiemmin todistuksessa saatiin, että ortogonaalisen yläkolmiomatriisin on oltava diagonaalinen, jonka diagonaalilla on  $\pm 1$ . Tällöin matriisi  $A_{k+1}$  lähestyy kohti yläkolmiomatriisia, jonka diagonaalilla on matriisin  $A$  ominaisarvot  $\lambda_1, \dots, \lambda_n$ , sillä

$$\lim_{k \rightarrow \infty} (A_{k+1})_{i,j} = \pm (R D R^{-1})_{i,j}, \quad \text{kun } i > j$$

ja

$$\lim_{k \rightarrow \infty} (A_{k+1})_{i,i} = (R D R^{-1})_{i,i} = D_{i,i} = \lambda_{n+1-i}, \quad \text{kun } i = j.$$

□

Kuten yllä olevasta todistuksesta nähdään QR-algoritmissa ominaisarvot voidaan löytää iteratiivisesti siten, että matriisijono  $\{A_k\}$  lähestyy Schurin hajotelmaa, missä yläkolmiomatriisin diagonaalilta nähdään matriisin  $A$  ominaisarvot. QR-algoritmia ei kuitenkaan yleensä käytetä sellaisenaan ominaisarvojen etsimiseen. Se on raskas laskennallisesti, sillä arvon  $k$  on oltava usein hyvin suuri, jotta ominaisarvot saadaan iteroitua sopivalla tarkkuudella. Erityisesti ominaisarvojen ollessa lähellä toisiaan se voi olla menetelmänä hyvin hidas. QR-algoritmin tehokkuutta voidaan kuitenkin parantaa merkittävästi muokkaamalla matriisi  $A$  ensin Hessenbergin matriisiksi. Matriisin  $A$  muokkaaminen Hessenbergin matriisiksi onkin monissa QR-algoritmiin pohjautuvissa algoritmeissa ensimmäinen askel. Matriisia  $H \in \mathcal{M}_n(\mathbb{R})$  kutsutaan Hessenbergin matriisiksi, jos  $h_{i,j} = 0$ , kun  $i > j + 1$ . Muokkaus tehdään niin, että Hessenbergin matriisi on ortogonaalisesti similaarinen matriisin  $A$  kanssa eli  $H = P^T A P$ . Aiheeseen voi tutustua lisää esimerkiksi lähteessä [4].

### 3.3.3 Mooren-Penrosen yleistetty käänteismatriisi

Singulaarisilla matriiseilla ei ole olemassa käänteismatriisia. QR-hajotelmalla voidaan kuitenkin löytää Mooren-Penrosen yleistetty käänteismatriisi myös singulaarisille matriiseille. Tässä luvussa tutkitaan QR-menetelmän käyttöä Mooren-Penrosen yleistetyn käänteismatriisin etsimiseen lauseen 3.1 mukaisille matriiseille. Määritellään kuitenkin ennen sitä Mooren-Penrosen yleistetty käänteismatriisi.

**Määritelmä 3.20.** Matriisin  $A \in \mathcal{M}_{m \times n}(\mathbb{K})$  Mooren-Penrosen yleistetty käänteismatriisi on  $A^+ \in \mathcal{M}_{n \times m}(\mathbb{K})$ , jos seuraavat ehdot ovat voimassa:

$$(AA^+)^* = AA^+, \quad (A^+A)^* = A^+A, \quad AA^+A = A, \quad A^+AA^+ = A^+.$$

**Lemma 3.21.** Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ , jolle  $m \geq n$  ja  $r(A) = n$ . Silloin matriisin  $A$  Mooren-Penrosen yleistetty käänteismatriisi on muotoa  $A^+ = (A^*A)^{-1}A^*$ .

*Todistus.* Matriisilla  $A$  on olemassa lauseen 3.1 mukainen QR-hajotelma. Silloin

$$A^*A = (QR)^*QR = R^*Q^*QR = R^*R.$$

Lauseen 3.1 mukaan  $R$  on säännöllinen matriisi, jolloin myös matriisin  $R^*$  on oltava säännöllinen. Tällöin myös matriisi  $A^*A$  on säännöllinen ja sille voidaan määritellä käänteismatriisi  $(A^*A)^{-1}$ . Lisäksi tiedetään, että säännölliselle matriisille  $M$  pätee  $(M^{-1})^* = (M^*)^{-1}$ , sillä

$$I = I^* = (M^{-1}M)^* = M^*(M^{-1})^* \quad \text{ja} \quad I = M^*(M^*)^{-1}.$$

Yhdistämällä nämä kaksi yhtälöä saadaan

$$\begin{aligned} M^*(M^{-1})^* &= M^*(M^*)^{-1} \\ \Leftrightarrow (M^{-1})^* &= (M^*)^{-1}. \end{aligned}$$

Käydään vielä lopuksi määritelmän 3.20 ehdot läpi.

- 1°  $(A(A^*A)^{-1}A^*)^* = A((A^*A)^{-1})^*A^* = A((A^*A)^*)^{-1}A^* = A(A^*A)^{-1}A^* = AA^+$
- 2°  $((A^*A)^{-1}A^*A)^* = A^*A((A^*A)^{-1})^* = A^*A(A^*A)^{-1} = I = (A^*A)^{-1}A^*A = A^+A$
- 3°  $A(A^*A)^{-1}A^*A = AI = A$
- 4°  $(A^*A)^{-1}A^*A(A^*A)^{-1}A^* = (A^*A)^{-1}A^* = A^+.$

□

Nyt lauseen 3.1 ja lemmän 3.21 avulla voidaan selvittää Mooren-Penrosen yleistetty käänteismatriisi  $A^+$  QR-hajotelman avulla täysiasteisille matriiseille. Yhtälöstä  $A^+ = (A^*A)^{-1}A^*$  saadaan

$$\begin{aligned} A^+ &= ((QR)^*QR)^{-1}(QR)^* \\ &= (R^* \underbrace{Q^*Q}_{=I} R)^{-1}R^*Q^* \\ &= (R^*R)^{-1}R^*Q^* \\ &= R^{-1} \underbrace{(R^*)^{-1}R^*}_{=I} Q^* \\ &= R^{-1}Q^*. \end{aligned}$$

**Lause 3.22.** Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ , jolle  $m \geq n$  ja  $r(A) = n$ . Silloin matriisin  $A$  yleistetty käänteismatriisi on muotoa  $A^+ = R^{-1}Q^*$ , missä matriisit  $Q$  ja  $R$  ovat matriisin  $A$  QR-hajotelman tekijät.

**Huomautus 3.23.** Singulaariarvohajotelmaa käsittelevässä osiossa  $m \times n$ -matriisille  $A$  paras PNS-menetelmän ratkaisu on muotoa  $\mathbf{x}' = A^+\mathbf{b}$  (lause 4.22). Tällöin lauseen 3.22 mukaisille matriiseille paras PNS-ratkaisu on muotoa  $\mathbf{x}' = R^{-1}Q^*\mathbf{b}$ , joka on yhtäpitävä lauseen 3.9 kanssa.

Käydään vielä läpi esimerkki Mooren-Penrosen yleistetyn käänteismatriisin muodostamisesta QR-hajotelman avulla.

**Esimerkki 3.24.** Lasketaan Mooren-Penrosen yleistetty käänteismatriisi esimerkin 3.5 matriisille  $A$ . Matriisin  $A$  pystyriivit ovat lineaarisesti riippumattomat, joten lauseen 3.22 mukaan  $A^+ = R^{-1}Q^*$ . Esimerkistä 3.5 saadaan

$$R = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{pmatrix} \Rightarrow R^{-1} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \\ 0 & \frac{\sqrt{3}}{3} \end{pmatrix} \quad \text{ja} \quad Q^* = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix},$$

jolloin Mooren-Penrosen yleistetyksi käänteismatriisiksi saadaan

$$A^+ = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}}{3} \\ 0 & \frac{\sqrt{3}}{3} \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{-1}{3} & \frac{5}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{-1}{3} & \frac{1}{3} \end{pmatrix}.$$

△

### 3.3.4 Determinantin itseisarvon laskeminen

Luvussa 2.3.3 esiteltiin, kuinka determinantti voidaan laskea tehokkaammin LU-hajotelmalla matriisikoon kasvaessa yhä suuremmaksi. Puolestaan QR-hajotelman avulla determinanttia voidaan arvioida laskemalla sen itseisarvo hyödyntäen unitaarisen matriisin  $Q$  ja yläkolmiomatriisin  $R$  ominaisuuksia, jotka esitellään seuraavaksi.

**Lemma 3.25.** Unitaarisen matriisin determinantin itseisarvo on aina yksi.

*Todistus.* Unitaariselle matriisille pätee  $U^* = U^{-1}$ , jolloin

$$\begin{aligned} 1 &= \det(I) = \det(UU^{-1}) = \det(UU^*) = \det(U) \det(U^*) \\ &= \det(U) \det(U)^* = (\det(U))^2 \Rightarrow |\det(U)| = 1. \end{aligned}$$

□

**Lause 3.26.** Olkoon  $A \in \mathcal{M}_n(\mathbb{K})$ . Jos unitaarinen matriisi  $Q \in \mathcal{M}_n(\mathbb{K})$  ja yläkolmiomatriisi  $R \in \mathcal{M}_n(\mathbb{K})$  ovat sellaiset, että  $A = QR$ , niin silloin

$$|\det(A)| = |\det(R)| = |r_{11}r_{22} \cdots r_{nn}|.$$

*Todistus.* Lemman 2.1 mukaan tiedetään, että  $\det(R) = r_{11}r_{22} \cdots r_{nn}$ . Lisäksi edellisen lemmän 3.25 mukaan  $|\det(U)| = 1$ . Nyt matriisin  $A$  determinantin itseisarvoksi saadaan  $|\det(A)| = |\det(QR)| = |\det(Q)| \cdot |\det(R)| = |\det(R)| = |r_{11}r_{22} \cdots r_{nn}|$ .

□

Lemman 3.25 mukaan unitaarisen matriisin determinantin itseisarvo on aina yksi. Lisäksi lauseessa 3.26 saatiin osoitettua, että matriisin  $A$  determinantti on matriisin  $R$  diagonaalialkioiden tulo itseisarvo. Kun matriisin QR-hajotelma tiedetään, näiden kahden tuloksen avulla determinantin itseisarvo voidaan laskea tehokkaasti.

**Esimerkki 3.27.** Käytetään esimerkin 3.7 QR-hajotelmaa. Lauseen 3.26 avulla matriisin  $A$  determinantin itseisarvoksi saadaan

$$|\det(A)| = |\det(R)| = \left| 3\sqrt{101} \cdot 12\sqrt{\frac{21}{101}} \cdot 5\sqrt{\frac{3}{7}} \right| = 540.$$

△

### 3.3.5 Muita sovelluksia

Edellä esiteltyjen sovellusten lisäksi QR-hajotelmalla on myös monia muita sovelluskohteita, joista esitellään kaksi vielä lyhyesti. QR-hajotelmaa voidaan hyödyntää esimerkiksi matriisin käänteismatriisia tai häiriöalttiutta selvittäessä.

Olkoon matriisi  $A$  säännöllinen. Tällöin lauseen 3.1 avulla matriisin  $A$  käänteismatriisiksi saadaan

$$A = QR \Rightarrow A^{-1} = R^{-1}Q^*.$$

Aiemmin lausessa 3.22 saatiin, että matriisin  $A$  Mooren-Penrosen yleistetty käänteismatriisi on muotoa  $A^+ = R^{-1}Q^*$ . Näin ollen säännöllisille matriiseille on voimassa  $A^+ = A^{-1}$ .

Jos matriisin QR-hajotelma tiedetään, voidaan sen avulla ratkaista matriisin häiriöalttiisuus helpommin. Tiedetään, että euklidisella matriisinnormilla unitaarille matriisille  $U$  pätee (ks. lähde [1] s. 50)

$$\|X\| = \|XU\| = \|UX\|.$$

Näin ollen matriisin  $A$  häiriöalttiudeksi saadaan

$$\kappa(A) = \|A\| \|A^{-1}\| = \|QR\| \|R^{-1}Q^{-1}\| = \|R\| \|R^{-1}\| = \kappa(R).$$

## 4 Singulaariarvohajotelma

Tutkielman viimeisenä matriisihajotelmana esitellään singulaariarvohajotelma, joka voidaan muodostaa mille tahansa matriisille. Singulaariarvohajotelmaa käsiteltäessä skalaarikuntana on  $\mathbb{R}$ . Singulaariarvohajotelma on myös mahdollista muodostaa kompleksisille matriiseille, ja siihen voi tutustua tarkemmin esimerkiksi lähteen [1] avulla.

Singulaariarvohajotelma mahdollistaa monia uusia soveltamiskohteita, kuten esimerkiksi matriisin asteen selvittämisen, approksimoimisen ja matriisin nolla- ja riviavaruuksien ortonormaalien kantojen selvittämisen. Ennen singulaariarvohajotelman käsittelemistä käydään läpi sen määrittelyyn tarvittavia tuloksia ja tutustutaan matriisin singulaariarvoihin. Symmetrisiä matriiseja käsittelevässä luvussa 4.1 on hyödynnetty lähdeä [7], kun taas singulaariarvoja ja singulaariarvohajotelmaa käsittelevissä luvuissa 4.2 ja 4.3 on hyödynnetty lähdeä [8].

### 4.1 Symmetrinen matriisi

Tutustutaan aluksi symmetriseen matriisiin ja tutkitaan sen ominaisuuksia.

**Määritelmä 4.1.** Matriisi  $A \in \mathcal{M}_n(\mathbb{R})$  on symmetrinen, jos  $A^T = A$ .

Edellisen määritelmän mukaan matriisi  $A = (a_{ij})_{n \times n}$  on siis symmetrinen jos ja vain jos  $a_{ij} = a_{ji}$  kaikilla  $i, j$ . Osoitetaan, että symmetrisen matriisin ominaisarvot ovat reaalisia.

**Lause 4.2.** Olkoon matriisi  $A \in \mathcal{M}_n(\mathbb{R})$  symmetrinen. Silloin sen ominaisarvot ovat reaalisia.

*Todistus.* Oletetaan, että ominaisarvot ja ominaisvektorit voivat olla myös kompleksisia. Olkoon  $A\mathbf{v} = \lambda\mathbf{v}$ , kun  $\mathbf{v} \neq 0$ . Ottamalla kompleksikonjugaatti molemmilta puolilta saadaan  $\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}}$ . Koska matriisi  $A$  on reaalinen, niin  $\overline{A} = A$ , eli  $A\overline{\mathbf{v}} = \overline{\lambda\mathbf{v}}$ . Sisätulo voidaan tulkita matriisitulona:  $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\mathbf{y}}^T \mathbf{x}$ , kun  $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ . Silloin

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \overline{\mathbf{v}}^T \lambda \mathbf{v} = \overline{\mathbf{v}}^T A \mathbf{v} = \overline{\mathbf{v}}^T A^T \mathbf{v} = (A\overline{\mathbf{v}})^T \mathbf{v} = (\overline{\lambda\mathbf{v}})^T \mathbf{v} = \overline{\mathbf{v}}^T \overline{\lambda\mathbf{v}} = \overline{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle,$$

joten  $\overline{\lambda} = \lambda$ , sillä  $\langle \mathbf{v}, \mathbf{v} \rangle \neq 0$ . Tällöin ominaisarvojen on oltava reaalisia  $\lambda \in \mathbb{R}$ .  $\square$

**Määritelmä 4.3.** Neliömatriisia  $A$  sanotaan ortogonaalisesti diagonalisoituvaksi, jos on olemassa sellainen ortogonaalinen matriisi  $U$ , että

$$D = U^T A U = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Osoitetaan seuraavaksi, että symmetrinen matriisi on ortogonaalisesti diagonalisoituva. Tämä voidaan osoittaa Schurin lauseen (lause 3.14) avulla, sillä symmetrisen matriisin ominaisarvojen tiedetään olevan reaalisia.

**Lause 4.4.** Symmetrinen matriisi  $A \in \mathcal{M}_n(\mathbb{R})$  on ortogonaalisesti diagonalisoituva.



*Todistus.* Olkoon  $A \in \mathcal{M}_n(\mathbb{R})$ . Koska symmetrisen matriisin  $A$  ominaisarvot ovat reaaliset, Schurin lausetta koskevan huomautuksen 3.15 mukaan se on ortogonaalisesti similaarinen yläkolmiomatriisiin  $T$  kanssa, eli  $T = U^T A U$ . Koska

$$T^T = (U^T A U)^T = U^T A^T (U^T)^T = U^T A U = T,$$

niin matriisi  $T$  on sekä yläkolmiomatriisi että symmetrinen matriisi. Näin ollen sen on oltava diagonaalinen.  $\square$

Ortogonaalisesta diagonalisoituvuudesta seuraa, että symmetrisen matriisin ominaisvektoreista voidaan muodostaa ortonormaalikanta.

**Lause 4.5.** Symmetrisen matriisin  $A \in \mathcal{M}_n(\mathbb{R})$  ominaisvektoreista  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  voidaan muodostaa ortonormaalikanta avaruudelle  $\mathbb{R}^n$ .

*Todistus.* Edellisen lauseen mukaan symmetrinen matriisi  $A$  on ortogonaalisesti diagonalisoituva, eli  $U^T A U = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ . Tämä voidaan ilmaista myös muodossa  $A U = U D$ . Ominaisarvoille ja -vektoreille pätee  $A \mathbf{v} = \lambda \mathbf{v}$ . Tutkimalla yhtälöä  $A U = U D$  tarkemmin nähdään, että matriisin  $U$  pystyriivit ovat matriisin  $A$  ominaisvektorit. Koska matriisi  $U$  on ortogonaalinen, ne muodostavat avaruuden  $\mathbb{R}^n$  ortonormaalikannan.  $\square$

## 4.2 Singulaariarvot

Tutustutaan seuraavaksi matriisin singulaariarvoihin. Olkoon  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ . Silloin matriisi  $A^T A$  on  $n \times n$ -matriisi. Lisäksi se on symmetrinen, sillä

$$(A^T A)^T = A^T (A^T)^T = A^T A.$$

Edellisten lauseiden perusteella matriisin  $A^T A$  ominaisarvot ovat reaaliset ja sen ominaisvektoreista  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  voidaan muodostaa ortonormaalikanta avaruudelle  $\mathbb{R}^n$ . Osoitetaan vielä, että ominaisarvot ovat myös epänegatiivisia. Olkoon  $A^T A \mathbf{v} = \lambda \mathbf{v}$ , kun  $\mathbf{v} \neq 0$ . Silloin

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \mathbf{v}^T \lambda \mathbf{v} = \mathbf{v}^T A^T A \mathbf{v} = (A \mathbf{v})^T A \mathbf{v} = \langle A \mathbf{v}, A \mathbf{v} \rangle = \|A \mathbf{v}\|^2 \geq 0,$$

sillä  $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ . Nyt matriisin  $A^T A$  ominaisarvot voidaan numeroida ja järjestää niin, että

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0 \text{ ja } \lambda_{k+1} = \dots = \lambda_n = 0. \quad (4)$$

Kun matriisin  $A^T A$  ominaisarvoista  $\lambda_1, \dots, \lambda_k$  otetaan neliöjuuri, saadaan matriisin  $A$  singulaariarvot  $\sigma_i = \sqrt{\lambda_i}$ ,  $\forall i = 1, \dots, k$ . Singulaariarvot ovat siis oikeastaan vektoreiden  $A \mathbf{v}_1, \dots, A \mathbf{v}_k$  pituuksia, sillä

$$\begin{aligned} \|A \mathbf{v}_i\| &= \sqrt{\langle A \mathbf{v}_i, A \mathbf{v}_i \rangle} = \sqrt{(A \mathbf{v}_i)^T A \mathbf{v}_i} = \sqrt{\mathbf{v}_i^T A^T A \mathbf{v}_i} \\ &= \sqrt{\mathbf{v}_i^T \lambda \mathbf{v}_i} = \sqrt{\lambda \langle \mathbf{v}_i, \mathbf{v}_i \rangle} = \sigma_i. \end{aligned}$$

Singulaariarvojen avulla saadaan selvitettyä uusia mielenkiintoisia ominaisuuksia matriisille  $A$ . Esimerkiksi, kun tiedetään matriisin  $A$  singulaariarvojen lukumäärä, tiedetään myös matriisin aste. Lisäksi nyt voidaan osoittaa vektoreiden  $\{A \mathbf{v}_1, \dots, A \mathbf{v}_k\}$  muodostavan matriisin  $A$  pystyriiviavaruuden kannan  $P(A) \subseteq \mathbb{R}^m$ .

**Lemma 4.6.** Oletetaan, että matriisin  $A^T A$  ominaisarvoja  $\lambda_1, \dots, \lambda_n$ , jotka on järjestetty kohdan (4) mukaisesti, vastaavat ominaisvektorit  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  muodostavat avaruuden  $\mathbb{R}^n$  ortonormaalin kannan. Lisäksi oletetaan matriisin  $A$  singulaariarvoiksi  $\sigma_1, \dots, \sigma_k$ . Silloin  $\{A\mathbf{v}_1, \dots, A\mathbf{v}_k\}$  on matriisin  $A$  pystyriviavaruuden  $P(A)$  kanta ja matriisin aste  $r(A) = k$ .

*Todistus.* Koska  $\lambda_i \mathbf{v}_i$  ja  $\mathbf{v}_j$  ovat ortogonaalisia, kun  $i \neq j$ , niin

$$\langle A\mathbf{v}_i, A\mathbf{v}_j \rangle = (A\mathbf{v}_j)^T A\mathbf{v}_i = \mathbf{v}_j^T A^T A\mathbf{v}_i = \mathbf{v}_j^T (\lambda_i \mathbf{v}_i) = \langle \lambda_i \mathbf{v}_i, \mathbf{v}_j \rangle = 0.$$

Näin ollen  $\{A\mathbf{v}_1, \dots, A\mathbf{v}_k\}$  on ortogonaalinen joukko. Aiemmin saatiin, että  $\|A\mathbf{v}_i\| = \sigma_i$ , kun  $1 \leq i \leq k$ . Koska  $\sigma_i \neq 0$  niin myös  $A\mathbf{v}_i \neq 0 \forall i = 1, \dots, k$ . Vektorit  $A\mathbf{v}_1, \dots, A\mathbf{v}_k$  ovat siis lineaarisesti riippumattomia ja ne kuuluvat matriisin  $A$  pystyriviavaruuteen. Olkoon  $\mathbf{y} \in P(A)$ , jolle pätee  $\mathbf{y} = A\mathbf{x}$ , missä  $\mathbf{x} = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n$ . Silloin vektori  $\mathbf{y}$  voidaan kirjoittaa muodossa

$$\begin{aligned} \mathbf{y} &= A\mathbf{x} = c_1 A\mathbf{v}_1 + \dots + c_k A\mathbf{v}_k + c_{k+1} A\mathbf{v}_{k+1} + \dots + c_n A\mathbf{v}_n \\ &= c_1 A\mathbf{v}_1 + \dots + c_k A\mathbf{v}_k + 0 + \dots + 0 \\ &= c_1 A\mathbf{v}_1 + \dots + c_k A\mathbf{v}_k. \end{aligned}$$

Näin ollen vektori  $\mathbf{y}$  kuuluu  $L(A\mathbf{v}_1, \dots, A\mathbf{v}_k)$ , joten  $\{A\mathbf{v}_1, \dots, A\mathbf{v}_k\}$  on pystyriviavaruuden  $P(A)$  kanta. Tällöin matriisin aste on  $r(A) = k$ .  $\square$

### 4.3 Singulaariarvohajotelma

Edellä käytyjen tulosten pohjalta voidaan muodostaa matriisille singulaariarvohajotelma. Todistus mukailee lähdeä [8].

**Lause 4.7.** (Singulaariarvohajotelma) Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  ja sen singulaariarvot  $\sigma_1 \geq \dots \geq \sigma_k > 0$ . Silloin on olemassa sellaiset ortogonaaliset matriisit  $U \in \mathcal{M}_m(\mathbb{R})$  ja  $V \in \mathcal{M}_n(\mathbb{R})$ , että  $A = U\Sigma V^T$ , missä

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_{m \times n}(\mathbb{R}), \quad D = \text{diag}(\sigma_1, \dots, \sigma_k) \quad \text{ja} \quad r(A) \leq \min\{m, n\}.$$

*Todistus.* Määritellään  $\lambda_i$  ja  $\mathbf{v}_i$  lemmän 4.6 mukaisesti, jolloin  $\{A\mathbf{v}_1, \dots, A\mathbf{v}_k\}$  on avaruuden  $P(A)$  ortogonaalinen kanta. Muodostetaan tästä kannasta ortonormaalikanta  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ , missä

$$\mathbf{u}_i = \frac{1}{\|A\mathbf{v}_i\|} A\mathbf{v}_i = \frac{1}{\sigma_i} A\mathbf{v}_i \Leftrightarrow \mathbf{u}_i \sigma_i = A\mathbf{v}_i, \quad \text{kun} \quad 1 \leq i \leq k. \quad (5)$$

Täydennetään ortonormaalikanta  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  koko avaruuden  $\mathbb{R}^m$  ortonormaalikannaksi  $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ . Merkitään  $U = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m) \in \mathcal{M}_m(\mathbb{R})$ , jolloin matriisi  $U$  on ortogonaalinen. Lisäksi merkitään matriisiksi  $V = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \in \mathcal{M}_n(\mathbb{R})$ , missä  $\mathbf{v}_i$  ovat matriisin  $A^T A$  ominaisvektorit. Näin ollen myös matriisi  $V$  on ortogonaalinen. Nyt kohdasta (5) saadaan

$$AV = (A\mathbf{v}_1, \dots, A\mathbf{v}_k, \mathbf{0}, \dots, \mathbf{0}) = (\sigma_1 \mathbf{u}_1, \dots, \sigma_k \mathbf{u}_k, \mathbf{0}, \dots, \mathbf{0}).$$

Määritellään matriisi  $\Sigma$  siten, että

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_{m \times n}(\mathbb{R}), \text{ missä } D = \text{diag}(\sigma_1, \dots, \sigma_k) \in \mathcal{M}_k(\mathbb{R}).$$

Silloin

$$U\Sigma = (\mathbf{u}_1, \dots, \mathbf{u}_m) \left( \begin{array}{ccc|c} \sigma_1 & & 0 & 0 \\ & \ddots & & 0 \\ 0 & & \sigma_k & 0 \\ \hline & & 0 & 0 \end{array} \right) = (\sigma_1 \mathbf{u}_1, \dots, \sigma_k \mathbf{u}_k, \mathbf{0}, \dots, \mathbf{0}) = AV.$$

Nyt siis  $U\Sigma = AV$ . Koska  $V$  on ortogonaalinen, voidaan yhtälö kertoa oikealta matriisilla  $V^T$ , jolloin  $A = U\Sigma V^T$ .

Lemmasta 4.6 saadaan, että  $r(A) = k$ . Matriisin  $A$  singulaariarvojen lukumäärästä voidaan siis päätellä myös matriisin aste. Näin ollen on selvää myös, että  $r(A) \leq \min\{m, n\}$ .  $\square$

**Huomautus 4.8.** Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  astetta  $k$  ja  $D = \text{diag}(\sigma_1, \dots, \sigma_k)$ . Silloin matriisilla  $\Sigma \in \mathcal{M}_{m \times n}(\mathbb{R})$  on jokin seuraavista muodoista:

- Jos  $k < \min\{m, n\}$ , niin  $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$
- Jos  $k = m < n$ , niin  $\Sigma = \begin{pmatrix} D & 0 \end{pmatrix}$
- Jos  $k = n < m$ , niin  $\Sigma = \begin{pmatrix} D \\ 0 \end{pmatrix}$
- Jos  $k = m = n$ , niin  $\Sigma = D$ .

Matriisin singulaariarvohajotelman matriisi  $V$  diagonalisoi matriisin  $A^T A$ , sillä  $A^T A = V\Sigma^T U^T U \Sigma V^T = V(\Sigma^T \Sigma)V^{-1}$ . Aiemmin saatiin, että matriisin  $A^T A$  ominaisvektoreista voidaan muodostaa kanta avaruudelle  $\mathbb{R}^n$ . Näin ollen matriisi  $V$  koostuu  $A^T A$  ominaisvektoreista. Vastaavasti matriisin  $U$  voidaan osoittaa diagonalisoivan matriisin  $AA^T$  ja matriisin  $U$  pystyriivit ovat siis matriisin  $AA^T$  ominaisvektoreita. Singulaariarvohajotelmassa matriisin  $U$  ja  $V$  pystyriivejä kutsutaankin matriisin  $A$  vasemman- ja oikeanpuoleisiksi singulaarivektoreiksi. Käydään seuraavaksi läpi esimerkki singulaariarvohajotelman muodostamisesta.

**Esimerkki 4.9.** Etsitään singulaariarvohajotelma matriisille

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

Muodostetaan ensin matriisi  $A^T A$ , jonka jälkeen ratkaistaan sen ominaisarvot. Matriisiksi saadaan

$$A^T A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

jonka ominaisarvot ovat  $\lambda_1 = 3$  ja  $\lambda_2 = 1$ . Tällöin matriisin  $A$  singulaariarvot ovat  $\sigma_1 = \sqrt{3}$  ja  $\sigma_2 = 1$ . Lemman 4.6 mukaan matriisin  $A$  aste on siis kaksi. Nyt singulaariarvojen avulla voidaan muodostaa matriisi  $\Sigma$ , joka on muotoa

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Seuraavaksi etsitään kuhunkin matriisin  $A^T A$  ominaisarvoon kuuluva ominaisvektori ja muodostetaan niistä avaruuden  $\mathbb{R}^2$  ortonormaalikanta

$$\mathbf{v}_1 = \frac{\sqrt{2}}{2}(1 \quad -1)^T \quad \text{ja} \quad \mathbf{v}_2 = \frac{\sqrt{2}}{2}(1 \quad 1)^T.$$

Muodostetaan näistä ominaisvektoreista matriisi  $V$ , eli

$$V = (\mathbf{v}_1 \quad \mathbf{v}_2) = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Matriisin  $U$  kaksi pystyriviä  $\mathbf{u}_1$  ja  $\mathbf{u}_2$  saadaan laskemalla

$$\begin{aligned} \mathbf{u}_1 &= \frac{1}{\sigma_1} A \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \frac{-\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ -\frac{\sqrt{6}}{6} \end{pmatrix} \\ \mathbf{u}_2 &= \frac{1}{\sigma_2} A \mathbf{v}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \frac{-\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix}. \end{aligned}$$

Koska vektoreista  $\mathbf{u}_i$  muodostetaan ortonormaalikanta avaruudelle  $\mathbb{R}^3$ , kolmas pystyrivi  $\mathbf{u}_3$  tulee valita siten, että se on ortonormaali vektoreiden  $\mathbf{u}_1$  ja  $\mathbf{u}_2$  kanssa. Esimerkiksi vektorille  $\mathbf{u}_3 = \frac{1}{\sqrt{3}}(1 \quad 1 \quad 1)^T$  pätee  $\langle \mathbf{u}_1, \mathbf{u}_3 \rangle = 0$  ja  $\langle \mathbf{u}_2, \mathbf{u}_3 \rangle = 0$ . Nyt matriisiksi  $U$  saadaan

$$U = \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{-\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{pmatrix}.$$

Lopuksi tarkistetaan vielä, että  $A = U \Sigma V^T$ . △

#### 4.4 Singulaariarvohajotelman sovelluksia

Tutustutaan seuraavaksi, miten singulaariarvohajotelmaa voidaan hyödyntää eri tilanteissa. Sovelluksia käsittelevissä luvuissa on hyödynnetty lähteitä [1, 8, 9, 11].

#### 4.4.1 Matriisin nolla- ja riviavaruuksien ortonormaalit kannat

Singulaariarvohajotelmaa voidaan hyödyntää muodostaessa matriisin nolla- ja riviavaruuksien ortonormaaleja kantoja. Kuten aiemmin on jo mainittu, matriisien  $U$  ja  $V$  pystyrivejä kutsutaan matriisin  $A$  singulaarivektoreiksi. Nämä singulaarivektorit muodostavat ortonormaalit kannat matriisin  $A$  pystyriviavaruudelle  $P(A)$ , vaakariviavaruudelle  $V(A)$ , nolla-avaruudelle  $N(A)$  ja nolla-avaruuden transpoosille  $N(A^T)$ . Ennen tämän osoittamista käydään kuitenkin läpi muutama todistukseen tarvittava tulos.

**Määritelmä 4.10.** Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ . Tällöin matriisin  $A$  nolla-avaruudeksi kutsutaan joukkoa

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

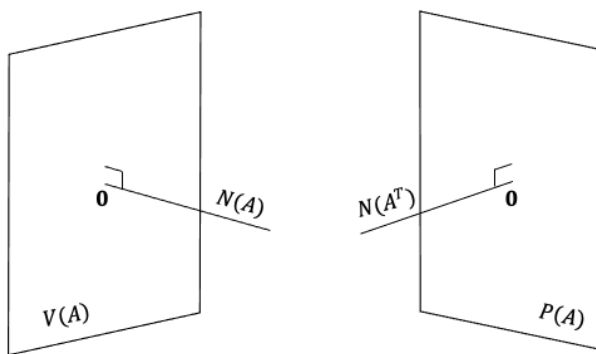
Esitellään seuraavaksi dimensiolause, jonka todistus sivuutetaan.

**Lause 4.11.** (Dimensiolause) Jos matriisilla  $A$  on  $n$  pystyriviä, niin silloin pätee  $n = r(A) + \dim N(A)$ .

**Lemma 4.12.** Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ . Silloin matriisin  $A$  vaakariviavaruuden ortogonaalikomplementti on matriisin  $A$  nolla-avaruus, ja matriisin  $A$  pystyriviavaruuden ortogonaalikomplementti on matriisin  $A^T$  nolla-avaruus, eli

$$V(A)^\perp = N(A) \quad \text{ja} \quad P(A)^\perp = N(A^T).$$

*Todistus.* Määritelmästä 4.10 saadaan, että jos vektori  $\mathbf{x} \in N(A)$ , niin  $A\mathbf{x} = \mathbf{0}$ . Toisin sanoen vektori  $\mathbf{x}$  on ortogonaalinen jokaisen matriisin  $A$  vaakarivin kanssa, eli vektori  $\mathbf{x}$  on ortogonaalinen vaakariviavaruuden  $V(A)$  kanssa. Näin ollen  $V(A)^\perp = N(A)$ . Sama päättely voidaan tehdä matriisin  $A$  transpoosille  $V(A^T)^\perp = N(A^T)$ . Koska tiedetään kuitenkin, että  $V(A^T) = P(A)$ , niin silloin  $P(A)^\perp = N(A^T)$ . Kuvassa 2 on havainnollistettu tilannetta.  $\square$



Kuva 2: Matriisin  $A$  pysty- ja vaakariviavaruus sekä niiden ortogonaalikomplementit. Kuvassa on hyödynnetty lähdettä [8].

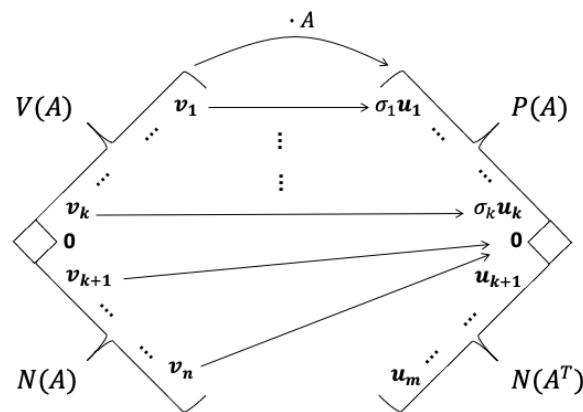
Merkitään nyt ortogonaalisten matriisien  $U$  ja  $V$  pystyrivejä  $U = (\mathbf{u}_1, \dots, \mathbf{u}_m)$  ja  $V = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ . Nyt dimensiolauseen ja lemmän 4.12 avulla voidaan osoittaa, että nämä pystyrivit muodostavat matriisin  $A$  nolla- ja riviavaruuksien ortonormaalit kannat.

**Lause 4.13.** Olkoon  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ , jonka singulaariarvot  $\sigma_1 \geq \dots \geq \sigma_k > 0$ . Lisäksi matriisin  $A$  aste on  $r(A) = k$  ja sen singulaariarvohajotelma on muotoa  $A = U\Sigma V^T = (\mathbf{u}_1, \dots, \mathbf{u}_m)\Sigma(\mathbf{v}_1, \dots, \mathbf{v}_n)^T$ . Silloin

1.  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  on avaruuden  $P(A)$  ortonormaalikanta
2.  $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_m\}$  on avaruuden  $N(A^T)$  ortonormaalikanta
3.  $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  on avaruuden  $V(A)$  ortonormaalikanta
4.  $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$  on avaruuden  $N(A)$  ortonormaalikanta.

*Todistus.* Ensimmäinen kohta on osoitettu jo lauseen 4.7 todistuksessa, jossa muodostettiin avaruuden  $P(A)$  ortonormaalikanta. Lisäksi saman lauseen todistuksessa ortonormaalikanta  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  täydennettiin koko avaruuden  $\mathbb{R}^m$  ortonormaaliksi kannaksi  $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ . Näin ollen vektorijoukko  $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_m\}$  on avaruuden  $P(A)^\perp$  ortonormaalikanta. Toisaalta lemmän 4.12 mukaan  $P(A)^\perp = N(A^T)$ , joten myös toinen kohta on osoitettu. Käydään neljäs kohta läpi ennen kolmatta. Muistetaan, että  $\|A\mathbf{v}_i\| = 0$  joss  $i > k$ . Näin ollen vektorit  $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$  virittävät nolla-avaruuden, jonka dimensio on  $n - k$ . Dimensiolauseen mukaan  $\dim N(A) = n - r(A)$ , joten joukko  $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$  on avaruuden  $N(A)$  ortonormaalikanta. Käydään vielä läpi kolmas kohta. Lemmän 4.12 mukaan  $N(A) = V(A)^\perp$ , jolloin selvästi nähdään, että  $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  on  $V(A)$  ortonormaalikanta.  $\square$

Kuvassa 3 on havainnollistettu lauseen 4.13 tulosta. Kuvasta on tärkeää huomata, että matriisin pystyriviavaruuden ortonormaalien kannan  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  tilalle on merkitty sen ortogonaalikanta  $\{\sigma\mathbf{u}_1, \dots, \sigma\mathbf{u}_k\}$  muistuttamaan, että  $\mathbf{u}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$ .



Kuva 3: Matriisin nolla- ja vaakariviavaruuksien ortonormaalit kannat. Lisäksi kuvaan on merkitty pystyriviavaruuden ortogonaalikanta. Kuvassa on hyödynnetty lähdetä [8].

#### 4.4.2 Matriisin häiriöalttius

Luvussa 2.3.4 käytiin jo lyhyesti läpi matriisin häiriöalttiutta. Silloin häiriöalttius määriteltiin yleisesti matriisille matriisinnormien avulla. Tämän luvun tiedoilla häiriöalttius voidaan ratkaista myös singulaariarvojen avulla. Määritellään ensin kuitenkin euklidinen matriisinnormi. Lause ja sen todistus pohjautuvat lähteeseen [11].

**Lause 4.14.** Merkitään  $\|\cdot\|$  euklidiseksi vektorinormiksi joukossa  $\mathbb{R}^n$ . Olkoon  $A$  nollamatriisista eroava matriisi ja sen singulaariarvot  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k$ . Silloin matriisin  $A$  euklidinen matriisnormi on sen suurin singulaariarvo eli

$$\|A\| = \max\{ \|A\mathbf{y}\| \mid \|\mathbf{y}\| = 1 \} = \sigma_1.$$

*Todistus.* Luvussa 4.3 saatiin, että matriisin  $A$  singulaariarvohajotema saadaan matriisien  $U$  ja  $V$  avulla, missä  $U = (\mathbf{u}_1, \dots, \mathbf{u}_m)$  ja  $V = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ . Lauseen 4.5 mukaan joukosta  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  voidaan muodostaa avaruuden  $\mathbb{R}^n$  ortonormaalikanta. Lisäksi lauseen 4.13 mukaan  $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  on avaruuden  $V(A)$  ortonormaalikanta,  $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$  on avaruuden  $N(A)$  ortonormaalikanta ja  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  on avaruuden  $P(A)$  ortonormaalikanta. Nyt lauseesta 4.7 saadaan

$$A\mathbf{v}_i = \begin{cases} \sigma_i \mathbf{u}_i, & \text{kun } i = 1, \dots, k \\ 0, & \text{kun } i = k+1, \dots, n. \end{cases}$$

Olkoon vektori  $\mathbf{y} \in \mathbb{R}^n$  yksikkövektori. Koska vektorit  $\mathbf{v}_1, \dots, \mathbf{v}_n$  muodostavat avaruuden  $\mathbb{R}^n$  ortonormaalikannan, vektori  $\mathbf{y}$  voidaan ilmaista muodossa

$$\mathbf{y} = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n,$$

missä  $c_i = \langle \mathbf{y}, \mathbf{v}_i \rangle$  kaikilla  $i = 1, \dots, n$ , sillä

$$\langle \mathbf{y}, \mathbf{v}_i \rangle = \left\langle \sum_{j=1}^n c_j \mathbf{v}_j, \mathbf{v}_i \right\rangle = \sum_{j=1}^n c_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle = c_i \|\mathbf{v}_i\|^2 = c_i.$$

Toisaalta

$$\|\mathbf{y}\|^2 = \left\langle \sum_{i=1}^n c_i \mathbf{v}_i, \sum_{j=1}^n c_j \mathbf{v}_j \right\rangle = \sum_{i,j=1}^n c_i c_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^n c_i^2,$$

jolloin sen euklidinen vektorinormi (vektorin pituus) on  $\|\mathbf{y}\| = \sqrt{c_1^2 + \dots + c_n^2} = 1$ . Koska myös vektorit  $\mathbf{u}_1, \dots, \mathbf{u}_k$  ovat ortonormaaaleja, vektori  $A\mathbf{y} \in \mathbb{R}^m$  voidaan ilmaista muodossa

$$A\mathbf{y} = c_1 A\mathbf{v}_1 + \dots + c_n A\mathbf{v}_n = c_1 \sigma_1 \mathbf{u}_1 + \dots + c_k \sigma_k \mathbf{u}_k,$$

ja sen euklidinen vektorinormi on  $\|A\mathbf{y}\| = \sqrt{c_1^2 \sigma_1^2 + \dots + c_k^2 \sigma_k^2}$ .

Koska singulaariarvoille pätee  $\sigma_1 \geq \dots \geq \sigma_k \geq 0$ , vektorinormia  $\|A\mathbf{y}\|$  voidaan arvioida ylöspäin, jolloin saadaan

$$\begin{aligned} \|A\mathbf{y}\| &= \sqrt{c_1^2 \sigma_1^2 + \dots + c_k^2 \sigma_k^2} \leq \sqrt{c_1^2 \sigma_1^2 + \dots + c_k^2 \sigma_1^2} \\ &= \sigma_1 \sqrt{c_1^2 + \dots + c_k^2} \leq \sigma_1 \sqrt{c_1^2 + \dots + c_n^2} \\ &= \sigma_1. \end{aligned}$$

Toisaalta, jos  $c_1 = 1, c_2 = \dots = c_n = 0$ , niin silloin  $\mathbf{y} = \mathbf{v}_1$ . Tällöin pätee  $\|A\mathbf{y}\| = \|A\mathbf{v}_1\| = \|\sigma_1 \mathbf{u}_1\| = \sigma_1$ .  $\square$

Tutkitaan matriisin häiriöalttiutta euklidisella matriisinormilla. Olkoon  $A$  säännöllinen  $n \times n$ -matriisi. Silloin sillä on olemassa käänteismatriisi  $A^{-1}$ , jonka suurin singulaariarvo on  $\sigma_n^{-1}$  (ks. luku 4.4.4 ja 4.4.5). Nyt matriisin  $A$  häiriöalttiudeksi saadaan  $\kappa(A) = \|A\| \|A^{-1}\| = \sigma_1 \cdot \sigma_n^{-1}$ . Näin ollen matriisin suurimman ja pienimmän singulaariarvon välisellä suhteella voidaan arvioida, kuinka häiriöaltis systeemi on virheille.

**Määritelmä 4.15.** Olkoon matriisi  $A \in \mathcal{M}_n$  säännöllinen ja sen singulaariarvot  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ . Silloin matriisin  $A$  häiriöalttiuus on  $\kappa(A) = \frac{\sigma_1}{\sigma_n}$ .

Tiedetään, että singulaariarvojen lukumäärä kertoo matriisin asteen. Koska singulaarisella  $n \times n$ -neliomatriisilla on vähemmän kuin  $n$  singulaariarvoa, sen häiriöalttiudeksi voidaan merkitä  $\infty$ . Mitä suurempi matriisin  $A$  häiriöalttiuus siis on, sitä lähempänä se on singulaarista matriisiä.

**Esimerkki 4.16.** Tutkitaan matriiseja

$$A = \begin{pmatrix} 5 & 5 & 5 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} \quad \text{ja} \quad A' = \begin{pmatrix} 5,00001 & 5 & 5 \\ 2 & 2,00001 & 2 \\ 3 & 3 & 3,00001 \end{pmatrix}.$$

Koska matriisilla  $A$  on vähemmän kuin kolme singulaariarvoa, sen häiriöalttiuus on  $\kappa(A) = \infty$ . Tarkastellaan seuraavaksi matriisiä  $A'$ , jonka singulaariarvot ovat  $\sigma'_1 \approx 10,67709$ ,  $\sigma'_2 \approx 1 \times 10^{-5}$  ja  $\sigma'_3 \approx 9,36586 \times 10^{-6}$ . Nyt matriisin  $A'$  häiriöalttiudeksi saadaan

$$\kappa(A') = \frac{\sigma_1}{\sigma_3} = \frac{10,67709}{9,36586 \times 10^{-6}} \approx 1,14 \cdot 10^6.$$

△

### 4.4.3 Matriisin approksimointi

Tutustutaan seuraavaksi, miksi matriisiä approksimoidaan sekä miten siinä voidaan hyödyntää singulaariarvohajotelmaa. Aiemmin saatiin, että matriisin aste on yhtä suuri singulaariarvojen lukumäärän kanssa. Matriisin aste voi kuitenkin olla herkkä hyvinkin pienille muutoksille matriisin alkioissa. Esimerkissä 4.16 esiintyvät matriisit

$$A = \begin{pmatrix} 5 & 5 & 5 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} \quad \text{ja} \quad A' = \begin{pmatrix} 5,00001 & 5 & 5 \\ 2 & 2,00001 & 2 \\ 3 & 3 & 3,00001 \end{pmatrix}.$$

Tutkimalla matriiseja  $A$  ja  $A'$  havaitaan, että matriisin  $A$  aste on yksi, jolloin se on singulaarinen. Sen sijaan matriisin  $A'$  aste on kolme, sillä sen singulaariarvot ovat  $\sigma'_1 \approx 10,67709$ ,  $\sigma'_2 \approx 1 \times 10^{-5}$  ja  $\sigma'_3 \approx 9,36586 \times 10^{-6}$ . Näin ollen se on määritelmien mukaan säännöllinen. Todellisuudessa matriisi  $A'$  on kuitenkin hyvin lähellä singulaarista matriisiä, sillä sen häiriöalttiuus  $\kappa(A') \approx 1,14 \cdot 10^6$  on suuri. Koska matriisin  $A'$  toinen ja kolmas singulaariarvo ovat hyvin pieniä, näiden voidaan ajatella vaikuttavan vain hyvin vähän matriisin ominaisuuksiin. Siksi tehokas menetelmä tämän kaltaiseen ongelmaan on määrittää singulaariarvoille jokin kynnsarvo, jonka



jälkeen ne voidaan olettaa nollassa. Tutkitaan seuraavaksi, miten tämä käytännössä tapahtuu.

Olkoon  $A \in M_{m \times n}(\mathbb{R})$  matriisi, jonka aste on  $k$ . Lisäksi sen singulaariarvot ovat  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$ . Merkitään singulaariarvohajotelmassa olevien ortogonaalisten matriisien  $U$  ja  $V$  pystyrivejä  $\mathbf{u}_i$  ja  $\mathbf{v}_i$ . Silloin matriisin  $A$  singulaariarvohajotelma voidaan esittää muodossa

$$A = U\Sigma V^T = \sum_{i=1}^k \mathbf{u}_i \sigma_i \mathbf{v}_i^T.$$

Koska singulaariarvot ovat kasvavassa järjestyksessä, matriisia  $A$  voidaan approksimoida matriisiksi  $A_t$  ottamalla vain ensimmäiset  $t$  kappaletta singulaariarvoista, jolloin

$$A_t = \sum_{i=1}^t \mathbf{u}_i \sigma_i \mathbf{v}_i^T,$$

missä  $1 \leq t \leq k$ . Tällä tavalla esimerkiksi alun matriisi  $A'$  voidaan approksimoida matriisiksi  $A$ , jonka aste on yksi. Esitellään seuraavaksi lause, jonka mukaan matriisi  $A_t$  onkin matriisin  $A$  paras approksimaatio matriiseista, joiden aste on  $t$ . Lauseen todistus sivuutetaan, mutta se on luettavissa esimerkiksi lähteestä [1].

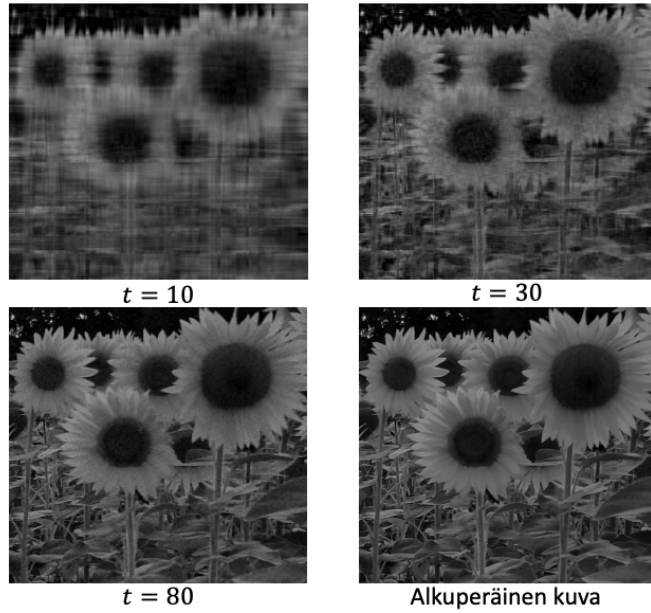
**Lause 4.17.** Olkoon  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  matriisi, jonka aste on  $k$ . Lisäksi sillä on singulaariarvohajotelma  $A = U\Sigma V^T$ , jonka singulaariarvot ovat  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$ . Matriisia  $A$  voidaan approksimoida matriisiksi  $A_t = \sum_{i=1}^t \mathbf{u}_i \sigma_i \mathbf{v}_i$ , missä  $1 \leq t < k$ . Silloin jokaiselle matriisille  $B$ , jonka aste on  $t$ , pätee  $\|A - A_t\| \leq \|B - A\|$ , missä  $\|\cdot\|$  on euklidinen matriisnormi.

Matriisin approksimointi on yksi singulaariarvohajotelman merkittävimpiä sovelluksia ja sitä voidaan soveltaa hyvin monella eri tavalla. Eräs tällainen sovellus on kuvanpakkaus, joka esitellään seuraavaksi esimerkin kautta.

**Esimerkki 4.18.** Tarkastellaan kuvassa 4 esiintyvää mustavalkokuvaa, jonka koko on  $765 \times 705$ . Se voidaan tallentaa tietokoneelle matriisina, jossa matriisin koko vastaa kuvan pikseleiden lukumäärää. Matriisin  $A$  alkiot määritellään välille  $[0, 1]$ , missä 0 kuvaa valkoista ja 1 mustaa pikseliä. Välillä  $0 < a_{ij} < 1$  olevat alkiot sen sijaan kuvaavat harmaan eri sävyjä. Joskus kuvatiedosto voi olla liian suuri, jolloin se voidaan pakata pienemmäksi singulaariarvohajotelmalla. Kuvan pakkauksessa muistin säästämiseksi alkuperäinen kuva korvataan approksimaatiolla, joka näyttää visuaalisesti samalta, mutta vie vähemmän muistia. Määritellään matriisi  $A \in \mathcal{M}_{765 \times 705}(\mathbb{R})$ , jonka aste on 705. Lisäksi sen singulaariarvot ovat  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{705} > 0$ , jolloin matriisi  $A$  voidaan esittää muodossa  $\sum_{i=1}^{705} \mathbf{u}_i \sigma_i \mathbf{v}_i^T$ . Approksimoidaan matriisia  $A$  ottamalla vain ensimmäiset  $t$  kappaletta singulaariarvoista. Tällöin matriisin  $A_t$  tallentamisessa tarvitsee tallettaa  $t$  kappaletta vektoreita  $\mathbf{u}_i \sigma_i \in \mathbb{R}^{765}$  ja  $t$  kappaletta vektoreita  $\mathbf{v}_i \in \mathbb{R}^{705}$ . Näin ollen tällaisen matriisin tallentaminen vaatii  $t(765 + 705)$  alkiota, kun taas alkuperäinen matriisin vaatii  $765 \cdot 705$  alkiota. Esimerkiksi kuvan 4 approksimoiminen niin, että  $t = 80$ , vie muistia alkuperäiseen kuvaan nähden

$$\frac{t(m+n)}{m \cdot n} = \frac{80(765+705)}{765 \cdot 705} \cdot 100\% \approx 22\%.$$

△



Kuva 4: Esimerkin 4.18 kuvanpakkaus singulaariarvohajotelman avulla. Alkuperäinen kuva sekä sen kolme eri approksimaatiota. Kuvanpakkaus on tehty lähteen [2] avulla.

#### 4.4.4 Mooren-Penrosen yleistetty käänteismatriisi ja pienimmän neliösumman menetelmä

Tässä luvussa tutustutaan, miten matriisille saadaan singulaariarvohajotelman avulla muodostettua Mooren-Penrosen yleistetty käänteismatriisi  $A^+$ . Tämän jälkeen käydään vielä läpi, miten Mooren-Penrosen yleistetyllä käänteismatriisilla voidaan löytää lineaariselle yhtälöryhmälle ratkaisu pienimmän neliösumman (lyh. PNS) menetelmällä.

**Lause 4.19.** Olkoon  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  ja sen singulaariarvohajotelma  $A = U\Sigma V^T$ . Silloin matriisin  $A$  Mooren-Penrosen yleistetty käänteismatriisi on muotoa  $A^+ = V\Sigma^+U^T$ , missä  $\Sigma^+ = \text{diag}(\sigma_1^{-1}, \dots, \sigma_k^{-1}, 0, \dots, 0) \in \mathcal{M}_{n \times m}(\mathbb{R})$  ja  $r(A) = k$ .

*Todistus.* Merkitään ensin matriisi  $\Sigma^+$  muotoon  $\Sigma^+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ , missä  $D^{-1} = \text{diag}(\sigma_1^{-1}, \dots, \sigma_k^{-1})$ . Ennen kuin saadaan itse väite osoitettua, on todistettava ensin, että matriisi  $\Sigma^+$  toteuttaa kaikki määritelmän 3.20 neljä ehtoa:

$$\begin{aligned}
 1^\circ \quad (\Sigma\Sigma^+)^T &= \left( \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} \right)^T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \Sigma\Sigma^+ \\
 2^\circ \quad (\Sigma^+\Sigma)^T &= \left( \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \right)^T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \Sigma^+\Sigma \\
 3^\circ \quad \Sigma\Sigma^+\Sigma &= \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = \Sigma \\
 4^\circ \quad \Sigma^+\Sigma\Sigma^+ &= \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} = \Sigma^+.
 \end{aligned}$$

Näin ollen siis  $\Sigma^+ = \text{diag}(\sigma_1^{-1}, \dots, \sigma_k^{-1}, 0, \dots, 0)$ . Osoitetaan vielä samalla tavalla, että  $A^+ = V\Sigma^+U^T$ :

$$\begin{aligned}
1^\circ (AA^+)^T &= \left( U \underbrace{\Sigma V^T V \Sigma^+}_{=I} U^T \right)^T = U \underbrace{(\Sigma \Sigma^+)^T}_{=\Sigma \Sigma^+} U^T = U \Sigma \Sigma^+ U^T \\
&= U \Sigma V^T V \Sigma^+ U^T = AA^+ \\
2^\circ (A^+A)^T &= \left( V \Sigma^+ \underbrace{U^T U \Sigma V^T}_{=I} \right)^T = V \underbrace{(\Sigma^+ \Sigma)^T}_{=\Sigma^+ \Sigma} V^T = V \Sigma^+ \Sigma V^T \\
&= V \Sigma^+ U^T U \Sigma V^T = A^+A \\
3^\circ AA^+A &= U \Sigma \underbrace{V^T V \Sigma^+}_{=I} \underbrace{U^T U \Sigma V^T}_{=I} = U \underbrace{\Sigma \Sigma^+ \Sigma}_{=\Sigma} V^T = U \Sigma V^T = A \\
4^\circ A^+AA^+ &= V \Sigma^+ \underbrace{U^T U \Sigma}_{=I} \underbrace{V^T V \Sigma^+}_{=I} U^T = V \underbrace{\Sigma^+ \Sigma \Sigma^+}_{=\Sigma^+} U^T = V \Sigma^+ U^T = A^+.
\end{aligned}$$

□

Edellisen lauseen mukaan matriisin  $A$  Mooren-Penrosen yleistetty käänteismatriisi on  $A^+ = V\Sigma^+U^T$ . Käydään läpi tämän muodostamisesta vielä esimerkki.

**Esimerkki 4.20.** Muodostetaan matriisille

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$$

Mooren-Penrosen yleistetty käänteismatriisi  $A^+$ . Esimerkistä 4.9 saadaan

$$U = \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{-\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ \frac{-\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{ja} \quad V = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Lauseesta 4.19 saadaan, että  $A^+ = V\Sigma^+U^T$ . Nyt laskemalla saadaan

$$\Sigma^+ = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{ja} \quad U^T = \begin{pmatrix} \frac{-\sqrt{6}}{6} & \frac{-\sqrt{6}}{3} & \frac{-\sqrt{6}}{6} \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{pmatrix}.$$

Näin ollen matriisiksi  $A^+$  saadaan  $A^+ = V\Sigma^+U^T = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$ . △

Käsitellään seuraavaksi, miten edellä esiteltyjä tietoja voidaan hyödyntää yhtälöryhmän ratkaisemisessa PNS-menetelmällä. Luvun 3.3.1 mukaan, jos yhtälöllä  $A\mathbf{x} = \mathbf{b}$  on ratkaisu, niin silloin se on myös PNS-menetelmän ratkaisu. Jos yhtälöllä  $A\mathbf{x} = \mathbf{b}$  ei sen sijaan ole ratkaisua, niin silloin  $\mathbf{b} \notin P(A)$ . Tällöin määritelmän

3.8 mukaan tulee etsiä sellainen vektori  $\mathbf{x}'$ , että  $A\mathbf{x}' \in P(A)$  on vektorin  $\mathbf{b}$  ortogonaaliprojektio, jolloin  $\|\mathbf{b} - A\mathbf{x}'\| = \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{b} - A\mathbf{x}\|$ . Lauseessa 4.22 osoitetaan, että  $\mathbf{x}' = A^+\mathbf{b}$  on paras PNS-menetelmän ratkaisu lineaariselle yhtälölle  $A\mathbf{x} = \mathbf{b}$ . Lauseen todistuksessa on hyödynnetty lähdeä [1]. Ennen sitä käydään kuitenkin siihen tarvittava tulos läpi.

**Lemma 4.21.** Olkoon  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  ja  $A^+ \in \mathcal{M}_{n \times m}(\mathbb{R})$  on sen Mooren-Penrosen yleistetty käänteismatriisi. Silloin matriisin  $A^+$  pystyriiviavuus on sama kuin matriisin  $A$  vaakariiviavuus, eli  $P(A^+) = V(A)$ .

*Todistus.* Oletetaan, että vektori  $\mathbf{x} \in N(A)$ , jolloin  $A\mathbf{x} = \mathbf{0}$ . Oletetaan lisäksi, että  $A^+\mathbf{b}_1 = \mathbf{y}$ , missä  $\mathbf{y} \in P(A^+)$  ja  $\mathbf{b}_1 \in \mathbb{R}^n$ . Hyödyntämällä määritelmää 3.20 saadaan vektoreiden  $\mathbf{x}$  ja  $\mathbf{y}$  sisätuloksi

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle &= \mathbf{y}^T \mathbf{x} = (A^+\mathbf{b}_1)^T \mathbf{x} = (A^+AA^+\mathbf{b}_1)^T \mathbf{x} = \mathbf{b}_1^T (A^+)^T A^T (A^+)^T \mathbf{x} \\ &= \mathbf{b}_1^T (A^+)^T \underbrace{(A^+A)^T}_{=A^+A} \mathbf{x} = \mathbf{b}_1^T (A^+)^T A^+ \underbrace{A\mathbf{x}}_{=\mathbf{0}} = 0. \end{aligned}$$

Koska sisätulo  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ , vektorit  $\mathbf{x}$  ja  $\mathbf{y}$  ovat ortogonaalisia, jolloin vektori  $\mathbf{y} \in N(A)^\perp$  ja  $P(A^+) \subseteq N(A)^\perp$ . Toisaalta lemmän 4.12 mukaan  $N(A)^\perp = V(A)$ , joten  $P(A^+) \subseteq V(A)$ .

Oletetaan seuraavaksi, että  $\mathbf{y} \in N(A)^\perp$ , jolloin edellisen perusteella pätee myös  $\mathbf{y} \in V(A)$ . Tällöin voidaan merkitä  $A^T\mathbf{b}_2 = \mathbf{y}$ , missä vektori  $\mathbf{b}_2 \in \mathbb{R}^m$ . Hyödyntämällä jälleen määritelmää 3.20 saadaan

$$\mathbf{y} = A^T\mathbf{b}_2 = A^T \underbrace{(A^T)^+}_{(A^+)^T} A^T\mathbf{b}_2 = A^T (A^+)^T A^T\mathbf{b}_2 = \underbrace{(A^+A)^T}_{=A^+A} \underbrace{A^T\mathbf{b}_2}_{=\mathbf{y}} = A^+A\mathbf{y},$$

joten  $\mathbf{y} \in P(A^+)$  ja silloin  $V(A) \subseteq P(A^+)$ .

Koska nyt  $P(A^+) \subseteq V(A)$  ja  $V(A) \subseteq P(A^+)$ , niin pätee  $P(A^+) = V(A)$ .  $\square$

**Lause 4.22.** Olkoon matriisi  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  sekä vektorit  $\mathbf{x} \in \mathbb{R}^n$  ja  $\mathbf{b} \in \mathbb{R}^m$ . Silloin vektori  $\mathbf{x}' = A^+\mathbf{b}$  on paras PNS-menetelmän ratkaisu lineaariselle yhtälölle  $A\mathbf{x} = \mathbf{b}$ .

*Todistus.* Mille tahansa  $\mathbf{x} \in \mathbb{R}^n$  pätee

$$A\mathbf{x} - \mathbf{b} = A\mathbf{x} - A\mathbf{x}' + A\mathbf{x}' - \mathbf{b} = A(\mathbf{x} - \mathbf{x}') - (I - AA^+)\mathbf{b}.$$

Selvästi nähdään, että  $A(\mathbf{x} - \mathbf{x}') \in P(A)$ . Seuraavaksi on osoitettava, että termi  $(I - AA^+)\mathbf{b} \in N(A^T)$ , eli ts.  $\langle (I - AA^+)\mathbf{b}, \mathbf{y} \rangle = 0$  kaikilla  $\mathbf{y} \in P(A)$ . Saadaan

$$\begin{aligned} \langle (I - AA^+)\mathbf{b}, \mathbf{y} \rangle &= \mathbf{y}^T (I - AA^+)\mathbf{b} \\ &= \mathbf{y}^T \mathbf{b} - \mathbf{y}^T AA^+\mathbf{b} \\ &= \mathbf{y}^T \mathbf{b} - \underbrace{((AA^+)^T \mathbf{y})^T}_{=AA^+} \mathbf{b} \\ &= \mathbf{y}^T \mathbf{b} - (AA^+\mathbf{y})^T \mathbf{b}. \end{aligned}$$

Koska  $\mathbf{y} \in P(A)$ , niin on olemassa sellainen vektori  $\mathbf{z} \in \mathbb{R}^n$ , että  $A\mathbf{z} = \mathbf{y}$ . Nyt sijoittamalla edelliseen saadaan

$$\mathbf{y}^T \mathbf{b} - (\underbrace{AA^+A}_{=A} \mathbf{z})^T \mathbf{b} = \mathbf{y}^T \mathbf{b} - (A\mathbf{z})^T \mathbf{b} = \mathbf{y}^T \mathbf{b} - \mathbf{y}^T \mathbf{b} = 0,$$

eli  $(I - AA^+) \mathbf{b} \in N(A^T)$ . Koska  $A(\mathbf{x} - \mathbf{x}')$  ja  $(I - AA^+) \mathbf{b}$  ovat ortogonaalisia (ks. kuva 2), Pythagoraan lauseella saadaan

$$\|A\mathbf{x} - \mathbf{b}\|^2 = \|A\mathbf{x} - A\mathbf{x}'\|^2 + \|A\mathbf{x}' - \mathbf{b}\|^2 \geq \|A\mathbf{x}' - \mathbf{b}\|^2, \quad (6)$$

eli  $\mathbf{x}'$  on PNS-menetelmän ratkaisu lineaariselle yhtälölle  $A\mathbf{x} = \mathbf{b}$ .

Osoitetaan vielä, että vektori  $\mathbf{x}'$  on lyhin PNS-menetelmän ratkaisu, eli  $\|\mathbf{x}'\| \leq \|\mathbf{x}\|$ . Jos  $\|A\mathbf{x}' - \mathbf{b}\| = \|A\mathbf{x} - \mathbf{b}\|$ , niin yhtälön (6) mukaan  $A\mathbf{x} = A\mathbf{x}'$ . Merkitään vektoriksi  $\mathbf{n} = \mathbf{x} - \mathbf{x}'$ , jolloin  $\mathbf{n} \in N(A)$ . Toisaalta lemmän 4.21 mukaan  $P(A^+) = V(A)$ , joten  $\mathbf{x}' \in V(A)$ , sillä  $\mathbf{x}' = A^+ \mathbf{b}$ . Näin ollen vektorit  $\mathbf{x}'$  ja  $\mathbf{n}$  ovat lemmän 4.12 mukaan ortogonaalisia. Hyödyntämällä jälleen Pythagoraan lausetta saadaan

$$\|\mathbf{x}\|^2 = \|\mathbf{x}' + \mathbf{n}\|^2 = \|\mathbf{x}'\|^2 + \|\mathbf{n}\|^2 \geq \|\mathbf{x}'\|^2,$$

joten  $\|\mathbf{x}'\| \leq \|\mathbf{x}\|$  kaikilla  $\mathbf{x}' \neq \mathbf{x}$ . □

Käydään vielä lopuksi läpi esimerkki, miten singulaariarvohajotelmaa voidaan hyödyntää ristiriitaisen lineaarisen yhtälöryhmän ratkaisemisessa.

**Esimerkki 4.23.** Etsitään yhtälölle  $A\mathbf{x} = \mathbf{b}$  ratkaisua PNS-menetelmällä, kun

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{ja} \quad \mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}.$$

Esimerkistä 4.20 saadaan, että matriisin  $A$  Mooren-Penrosen yleistetty käänteismatriisi on

$$A^+ = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

Nyt lauseen 4.22 mukaan PNS-menetelmän ratkaisuksi saadaan

$$\mathbf{x}' = A^+ \mathbf{b} = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{4}{3} \end{pmatrix}.$$

△

#### 4.4.5 Muita sovelluksia

Edellä mainittujen lisäksi singulaariarvohajotelmaa voidaan soveltaa esimerkiksi käänteismatriisin olemassaolon selvittämisessä tai determinantin itseisarvon laske-  
misessa.

Aiemmin ollaan nähty, miten LU- ja QR-hajotelmalla pystytään ratkaisemaan matriisin  $A$  käänteismatriisi. Myös singulaariarvohajotelmalla tämä voidaan ratkaista. Tämän lisäksi singulaariarvohajotelmalla voidaan myös helposti nähdä, milloin matriisilla  $A$  on olemassa käänteismatriisi  $A^{-1}$ . Matriisilla  $A$  on olemassa käänteismatriisi, jos ja vain jos lauseen 4.7 singulaariarvohajotelmassa  $\Sigma = D$ . Tällöin siis huomautuksen 4.8 mukaan  $k = m = n$ , eli matriisin täytyy olla säännöllinen. Kun edellä oleva ehto on voimassa, saadaan käänteismatriisi selvitettyä. Koska matriisit  $U$  ja  $V$  ovat ortogonaalisia, niin niille pätee  $U^T U = I$  ja  $V^T V = I$ . Näin ollen käänteismatriisiksi saadaan helposti  $A^{-1} = V D^{-1} U^T$ .

Myös determinantin itseisarvo voidaan ratkaista matriisin  $A \in \mathcal{M}_n$  singulaariarvohajotelmalla. Tiedetään, että matriisit  $U \in \mathcal{M}_n$  ja  $V \in \mathcal{M}_n$  ovat ortogonaalisia, joten lemmän 3.25 mukaan niiden determinantin itseisarvo on yksi. Lisäksi matriisi  $\Sigma \in \mathcal{M}_n$  on diagonaalimatriisi, jonka determinantti on sen diagonaalialkioiden tulo. Näin ollen matriisin  $A$  determinantin itseisarvo on  $|\det(A)| = |\det(\Sigma)|$ . Merkitään matriisin  $A$  asteeksi  $r(A) = k$ . Jos  $k < n$ , niin matriisin  $\Sigma$  diagonaalilla on myös nollia. Silloin  $|\det(A)| = 0$  ja matriisi on singulaarinen. Näin ollen metodilla selviää helposti myös matriisin säännöllisyys.

## Viitteet

- [1] G. Allaire, S. M. Kaber: *Numerical Linear Algebra*, Springer, New York, 2008.
- [2] T. Baumann: *Image Compression with Singular Value Decomposition* <http://timbaumann.info/svd-image-compression-demo/> (3.10.2020)
- [3] M. T. Chu: Algebraic Eigenvalue Problems, <https://mtchu.math.ncsu.edu/Teaching/Lectures/MA529/chapter5.pdf> (12.11.2020).
- [4] W. Ford: *Numerical Linear Algebra with Applications: Using MATLAB*, Elsevier Science & Technology, San Diego, 2014.
- [5] J. E. Gentle: *Matrix Algebra: Theory, Computations, and Applications in Statistics*, Springer, New York, 2007.
- [6] M. Koppinen: *Lineaarialgebra Osa 2*, luentomoniste, 2006.
- [7] M. Koppinen: *Matriisilaskenta*, luentomoniste, 2012.
- [8] D. C. Lay, S. Lay, J. McDonald: *Linear Algebra and Its Application* (5th edition), Pearson, U.S., 2014.
- [9] T. Lyche: *Numerical Linear Algebra and Matrix Factorization*, Springer, Cham, 2020.
- [10] M. T. Nair, A. Singh: *Linear Algebra*, Springer, Singapore, 2018.
- [11] P. J. Olver, C. Shakiban: *Applied Linear Algebra* (2nd edition), Springer, Cham, 2018.
- [12] K. Parvinen: *Numeerinen analyysi*, luentomoniste, 2019.