Lineaarialgebra

Mitri Kitti
Marko Korhonen
Ville Korpela

Kevät 2012–kevät 2014

Sisältö

1 Johdanto **5**

1.1 Mitä on lineaarialgebra? 5

1.2 Esimerkki markkinatasapainosta 6

2 Vektorit **8**

2.1 Vektoreilla laskeminen 9

2.2 Sisätulo ja normi 9

2.3 Hypertasot ja puoliavaruudet 10

2.4 Lineaarinen riippuvuus ja kanta 11

2.4.1 Vektorialiavaruudet 12

3 Matriisialgebraa **14**

3.1 Matriisin käsite 14

3.2 Matriiseilla laskeminen 15

3.2.1 Matriisien summa 16

3.2.2 Matriisin kertominen skalaarilla 16

3.2.3 Matriisien laskulait 17

3.3 Matriisitulo 17
3.3.1 Matriisin transponointi 19
3.4 Tärkeitä matriisityyppejä 21
3.4.1 Diagonaalimatriisi 21
3.4.2 Identiteettimatriisi 22
3.4.3 Kolmiomatriisi 22
3.4.4 Idempotentti matriisi 22

4 Lineaarikuvaukset ja yhtälöryhmät 24

4.1 Matriisit lineaarikuvauksina 24
4.2 Lineaariset yhtälöryhmät 25
4.2.1 Lineaarisen yhtälöryhmän esitys 25
4.2.2 Lineaaristen yhtälöryhmien ominaisuuksia 27
4.3 Yhtälöryhmän ratkaisu ja käänteismatriisi 28
4.4 Lineaarikuvauksiin liittyvät avaruudet 29

5 Determinantti 31

5.1 Matriisin determinantti 31
5.1.1 Determinantti 2×2-tapauksessa 31
5.1.2 Determinantin laskeminen yleisessä tapauksessa 32
5.2 Determinantin ominaisuudet 35
5.2.1 Laskusääntöjä 35
5.2.2 Determinantti tilavuuden skaalauskertoimena 37

6 Matriisin aste **39**

6.1 Aste ja lineaarinen riippumattomuus 39

6.2 Asteen ominaisuuksia 40

6.2.1 Lineaarinen implisiittifunktiolause 41

6.2.2 Esimerkki: portfolion valinta 42

6.3 Asteen laskeminen 43

7 Käänteismatriisin laskeminen **45**

7.1 Menetelmä 1: suora eliminointi 45

7.2 Menetelmä 2: Cramerin sääntö käänteismatriisille 46

7.3 Menetelmä 3: Gaussin-Jordanin eliminointimenettely 48

8 Lineaarisen yhtälöryhmän ratkaiseminen **50**

8.1 Cramerin sääntö 51

8.2 Gaussin ja Gaussin-Jordanin eliminointimenettelyt 53

8.3 Sovellus: tasapaino oligopolimalleissa 58

9 Ominaisarvot ja ominaisvektorit **60**

9.1 Ominaisarvojen laskeminen 61

9.2 Ominaisarvojen ominaisuuksia 63

9.3 Ominaisvektoreiden laskeminen 64

9.4 Ominaisarvohajotelma 66

10 Lineaariset differenssiyhtälöt **68**

10.1 Ratkaiseminen diagonalisoinnilla 69

10.2 Yleinen ratkaisu 70

10.3 Kompleksiset ominaisarvot 71

10.3.1 Kompleksiluvut 72

10.3.2 Differenssiyhtälön ratkaisu 72

10.3.3 Stabiilisuus 73

10.4 Markov-prosessit 74

11 Neliömuodot ja matriisin definiittisyys **77**

11.1 Definiittisyyden määrittäminen 79

11.2 Cholesky-hajotelma 81

12 Harjoitustehtäviä **83**

Luku 1

Johdanto

1.1 Mitä on lineaarialgebra?

Lineaarialgebra tutkii vektoreita, lineaarisia kuvauksia ja lineaarisia yhtälöryhmiä. Lineaarialgebran keskeisiä työkaluja ovat vektorit ja matriisit. Vektori- ja matriisimerkinnät helpottavat ja selkeyttävät monia asioita. Esimerkiksi monen muuttujan differentiaalilaskennan määritelmiä ja tuloksia on helppo esittää matriisien ja vektorien avulla. Lineaarialgebra tarjoaa menetelmiä yhtälöryhmien ratkaisemiseksi. Lineaariset yhtälöryhmät ovat tärkeitä useissa sovelluksissa. Myös silloin, kun mallinnettava ilmiö on epälineaarinen, voidaan päätyä lineaarisiin yhtälöihin linearisoimalla alkuperäinen malli.

Tarkastellaan paria johdattelevaa esimerkkiä.

Esimerkki 1.1.1. Mitä voit sanoa ratkaisujen määrästä ja olemassaolosta yhtälölle $ax = b$, missä $x \in \mathbb{R}$ on *endogeeninen* muuttuja ja $a, b \in \mathbb{R}$?
Jos $a \neq 0$, niin ratkaisu on $x = b/a$. Jos $a = 0$, niin ratkaisu on olemassa jos $b = 0$ ja tällöin ratkaisuja on äärettömän monta. Jos $b \neq 0$, niin yhtälöllä ei ole lainkaan ratkaisua.

Esimerkki 1.1.2. Ratkaise yhtälöpari

$$\begin{aligned} 2x_1 + x_2 &= 1 \\ x_1 + x_2 &= 3. \end{aligned}$$

Alemmasta voidaan eliminoida x_2, jolloin saadaan $x_2 = 3 - x_1$. Tämä voidaan sijoittaa ylempään, jolloin $2x_1 + 3 - x_1 = 1$, mistä saadaan $x_1 = -2$ ja edelleen $x_2 = 5$.

Esimerkki 1.1.3. Ratkaise yhtälöpari

$$\begin{aligned} 5x_1 - x_2 &= -1 \\ 10x_1 - 2x_2 &= 2. \end{aligned}$$

Eliminoidaan alemmasta x_2, eli $x_2 = 5x_1 + 1$ ja sijoitetaan ylempään. Saadaan $10x_1 - 2(5x_1 + 1) = 2$, mistä $-2 = 2$, eli yhtälöparilla ei ole ratkaisua.

Yleinen lineaarinen yhtälöryhmä on muotoa

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned}$$

Yllä olevassa yhtälöryhmässä on m yhtälöä ja n tuntematonta. Huom. tuntemattomia muuttujia ovat tässä $x_1, \ldots, x_n \in \mathbb{R}$, joita sanotaan myös endogeenisiksi muuttujiksi. Luvut $a_{ij}, b_i \in \mathbb{R}$ ovat eksogeenisia muuttujia (tai parametreja). Yhtälöryhmän toteuttavaa lukujonoa $(x_1, \ldots, x_n)$ sanotaan yhtälöryhmän ratkaisuksi. Huomaa, että tämä yhtälöryhmä on lineaarinen sen takia, että siinä eksogeenisista muuttujista esiintyy vain niiden lineaarsia funktioita. Esim. yhtälö $ax + bxy + cy = d^2$ ei ole lineaarinen x:lle ja y:lle termin xy takia.

Eräänä tavoitteena on analysoida seuraavia kysymyksiä. Milloin yhtälöryhmällä on ratkaisu? Montako ratkaisua yhtälöryhmällä? Miten ratkaisu voidaan laskea?

1.2 Esimerkki markkinatasapainosta

Analysoidaan kysyntä- ja tarjontakehikossa kahden tuotteen tasapainohintoja ja -määriä. Tuotteiden kysynnät ovat muotoa:

$$Q_i^d = K_i P_1^{\alpha_{i1}} P_2^{\alpha_{i2}} Y^{\beta_i}, \ i = 1, 2.$$

Yllä Q_i^d on tuotteen i kysytty määrä, P_i on tuotteen i hinta, Y on tulotaso, ja K_i kokoaa kaikki analyysin ulkopuolistet muuttujat (eksogeeniset muuttujat), jotka vaikuttavat kysyntään. Tulkitaan parametrit α_{ij} ja β_i mikroteorian aineopintojen mukaisesti: α_{ij} on tuotteen i kysynnän jousto tuotteen j hinnan mukaan. Siis esimerkiksi α_{11} on tuotteen hintajousto oman hinnan mukaan ja α_{12} on tuotteen 1 ristihintajousto. β_i on tuotteen i kysynnän tulojousto. Mitä etumerkkejä odotat joustojen saavan?

Tarjontafunktiot tuotteille riippuvat vain omista hinnoista (ja tämän markkinan ulkopuolisista tekijoistä): $Q_i^s = M_i P_i^{\gamma_i}$, $i = 1, 2$ Tulkinta γ_i:lle on tarjonnan hintajousto. Tasapainossa kysyntä kohtaa tarjonnan $Q_i^d = Q_i^s$, $i = 1, 2$. Toisin sanoen mallissa on kuusi yhtälöä kuudelle tuntemattomalle. Malli on melko monimutkaisen näköinen, mutta sitä voi yksinkertaistaa huomattavasti muuttujanvaihdoksella. Merkitään $q_i^d = \ln Q_i^d$, $q_i^s = \ln Q_i^s$, $p_i = \ln P_i$, $y = \ln Y$, $m_i = \ln M_i$, $k_i = \ln K_i$.

Ottamalla logaritmit puolittain mallin yhtälöistä, voidaan tasapainoehdot kirjoittaa muodossa: kaikille $i \in 1,2$

$$\begin{aligned} q_i^d &= k_i + \alpha_{ii} p_i + \alpha_{ij} p_j + \beta_i y, \\ q_i^s &= m_i + \gamma_i p_i, \\ q_i^s &= q_i^d. \end{aligned}$$

Näin mallin tasapainoehdot saatiin lineaariseksi yhtälöryhmäksi. Sijoittamalla ensin kolmannesta yhtälöstä toiseen ja sitten ensimmäisestä toiseen saadaan:

$$k_i + \alpha_{ii} p_i + \alpha_{ij} p_j + \beta_i y = m_i + \gamma_i p_i \quad i \in \{1,2\}.$$

Huom. osittaisen tasapainon malleissa, kuten tässä, y oletetaan eksogeeniseksi, jolloin ainoat tuntemattomat ovat hinnat p_1 ja p_2.

Kirjoitetaan malli edellä esitettyyn lineaarisen yhtälöryhmän "standardimuotoon":

$$\begin{array}{ccc} (\alpha_{11} - \gamma_1) p_1 & +\alpha_{12} p_2 & = & m_1 - k_1 - \beta_1 y \\ \alpha_{21} p_1 & +(\alpha_{22} - \gamma_2) p_2 & = & m_2 - k_2 - \beta_2 y \end{array}$$

Kun opimme käyttämään matriiseja, tulemme huomaamaan, että tämä yhtälö olisi kirjoitettavissa muodossa

$$\begin{pmatrix} \alpha_{11} - \gamma_1 & \alpha 12 \\ \alpha_{21} & \alpha_{11} - \gamma_1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} m_1 - k_1 - \beta_1 y \\ m_2 - k_2 - \beta_1 y \end{pmatrix}. \tag{1.1}$$

Yllä oleva malli on esimerkki lineaarisesta mallista ja sen ratkaisu saadaan helposti esimerkiksi ratkaisemalla p_1 ensimmäisestä yhtälöstä ja sijoittamalla ratkaisu alempaan yhtälöön, jolloin

$$p_1 = \frac{m_1 - k_1 - \beta_1 y - \alpha_{12} p_2}{\alpha_{11} - \gamma_1}$$

ja sijoittamalla alempaan

$$\begin{aligned} p_2 = & \frac{m_2 - k_2 - \beta_2 y - \alpha_{21} p_1}{\alpha_{22} - \gamma_2} \\ & \frac{m_2 - k_2 - \beta_2 y - \alpha_{21} \frac{m_1 - k_1 - \beta_1 y - \alpha_{12} p_2}{\alpha_{11} - \gamma_1}}{\alpha_{22} - \gamma_2} \\ & \frac{(\alpha_{11} - \gamma_1)(m_2 - k_2 - \beta_2 y) - \alpha_{21}(m_1 - k_1 - \beta_1 y - \alpha_{12} p_2)}{(\alpha_{11} - \gamma_1)(\alpha_{22} - \gamma_2) - \alpha_{12} \alpha_{21}}. \end{aligned}$$

Sijoittamalla tämä p_1:n lausekkeeseen saadaan p_1 ratkaistua.

Edellä esitetty ratkaisutapa on nimeltään suora sijoittaminen tai suora eliminointi (*direct elimination*). Huomaa kuinka monimutkaista ja virhealtista ratkaiseminen tällaisella suoralla sijoittamisella on. Jos kyseessä olisi 3,4 tai 5 hyödykkeen markkina, tehtävää olisi mahdotonta lähestyä näin. Tulemme jatkossa huomaamaan, että parempia menetelmiä yhtälöryhmien ratkaisemiseksi on tarjolla.

Luku 2

Vektorit

Kerrataan aluksi mitä tarkoitetaan joukkojen karteesisella tulolla. Joukkojen X_1 ja X_2 karteesinen tulo on joukko

$$X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1,\ x_2 \in X_2\}.$$

Vastaavasti joukoille $X_1, X_2, \ldots, X_n$;

$$X_1 \times X_2 \times \cdots X_n = \{(x_1, x_2, \ldots, x_n) : x_i \in X_1,\ i = 1, 2, \ldots, n\}.$$

Jos edellä olevassa karteesisessa tulossa $X_i = X$ kaikilla $i = 1, \ldots, n$, merkitään karteesista tuloa joukolla X^n.

Reaaliarvoiset vektorit voidaan ymmärtää $\mathbb{R}$:n karteesisen tulon alkioina. Tasovektorit kuuluvat joukkoon $\mathbb{R}^2$, missä $\mathbb{R}^2$ on joukon $\mathbb{R}$ karteesinen tulo itsensä kanssa, eli $\mathbb{R}^2 = \{(x_1, x_2) : x_1 \in \mathbb{R},\ x_2 \in \mathbb{R}\}$. Yleisemmät n-ulotteiset vektorit kuuluvat vastaavasti joukkoon $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_1 \in \mathbb{R}, \ldots, x_n \in \mathbb{R}^n\}$. Joukkoa $\mathbb{R}^n$ sanotaan n-ulotteiseksi reaaliavaruudeksi tai n-ulotteiseksi Euklidiseksi avaruudeksi.

Määäritelmä 2.0.1. Jono $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, missä $x_i \in \mathbb{R}$, $n \in \{1, 2, 3, \ldots\}$, on n-ulotteinen vektori.

Vektorit voidaan tulkita suuntavektoreina tai paikkavektoreina. Jokainen tason $\mathbb{R}^2$ vektori (x_1, x_2) määrää suuntavektorin, joka voidaan esittää piirtämällä nuoli origosta pisteeseen (x_1, x_2). Jos vektorin suuntaa tai pituutta muuttaa, muuttuu piste (x_1, x_2).

Vektoreille käytetään seuraavia merkintöjä.

1. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} = \mathbf{y} \Leftrightarrow x_i = y_i \forall i = 1, \ldots, n$
2. $\mathbf{x} \geq \mathbf{y}$ jos $x_i \geq y_i \forall i$ ja $x_i > y_i$ jollakin i
3. $\mathbf{x} \gg \mathbf{y}$ jos $x_i > y_i$ kaikilla i

Huom. toisin kuin reaaliluvuille vektoreille ei päde, että joko $\mathbf{x} \geq \mathbf{y}$ tai $\mathbf{x} \leq \mathbf{y}$

Esimerkki 2.0.1. Olkoot $\mathbf{x} = (2a + 3b + 5c, a - 3c, 5b - 3c)$ ja $\mathbf{y} = (10, -2, 2)$. Tällöin $\mathbf{x} = \mathbf{y}$ vastaa yhtälöryhmää

$$\begin{array}{cccccc} 2a & +3b & +5c & = & 10 \\ a & & -3c & = & -2 \\ & 5b & -3c & = & 2. \end{array}$$

2.1 Vektoreilla laskeminen

Vektoreille voidaan määritellä summa ja skalaarilla kertominen. Summa on määritelty komponenteittain, eli $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ tällöin $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$. Summa vektori voidaan luonnollisesti visualisoida suuntavektoreiden avulla. Skalaarilla kertominen määritellään myös komponenteittain, $\alpha \in \mathbb{R}$ (skalaari), $\mathbf{x} \in \mathbb{R}^n$: $\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$. Skalaarilla kertominen skaalaa (tai kääntää ja skaalaa) vektoria.

Vektoreilla laskemisessa pätee seuraavat laskusäännöt, jotka voidaan todistaa edellä annetuttujen summan ja skalaarilla kertomisen avulla.

Lause 2.1.1. *Olkoon* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*, ja* $\alpha, \beta \in \mathbb{R}$*. Tällöin seuraavat ominaisuudet vektoreilla laskemiselle ovat voimassa.*

1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ *(vaihdannaisuus)*
2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ *(liitännäisyys)*
3. *nollavektorille* $(0, \ldots, 0)$ *pätee* $\mathbf{x} + \mathbf{0} = \mathbf{x}$
4. $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ *(vektorilla on aina vastavektori)*
5. $\alpha\mathbf{x}$ *on* $\mathbb{R}^n$*:nnän vektori*
6. $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$
7. $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$
8. $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$
9. $1\mathbf{x} = \mathbf{x}$

Koska vektorilla on vastavektori merkitään $\mathbf{x} - \mathbf{y} = \mathbf{x} + (-\mathbf{y})$. Kohta (5) on ilmeinen, mutta se on tärkeä ominaisuus, joka vaaditaan yleisemmin vektoriavaruuksilta, eli skalaarilla kerrottu vektori kuuluu aina avaruuteen. Tämä ominaisuus puuttuu esimerkiksi jos rajoitutaan vektoreihin, joilla $\mathbf{x} \geq \mathbf{0}$; kyseiset vektorit eivät siis muodosta vektoriavaruutta.

2.2 Sisätulo ja normi

Määäritelmä 2.2.1. Olkoon $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, tällöin vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ sisätulo on $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$

Sisätulolla on seuraavat ominaisuudet.

Lause 2.2.1. *Olkoon* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *ja* $\alpha \in \mathbb{R}$*. Tällöin*

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
2. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$
3. $(\alpha\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\alpha\mathbf{y}) = \alpha(\mathbf{x} \cdot \mathbf{y})$
4. $\mathbf{x} \cdot \mathbf{x} \geq 0$ *ja* $\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

Jos $\mathbf{x} \cdot \mathbf{y} = 0$ sanotaan, että $\mathbf{x}$ ja $\mathbf{y}$ ovat ortogonaaliset.

Määäritelmä 2.2.2. Vektorin $\mathbf{x} \in \mathbb{R}^n$ normi on

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

Normi mittaa vektorin pituutta. Huom. normaalissa tasogeometriassa normi vastaa Pythagoraan lauseen mukaista kolmion hypotenuusan pituutta, kun kateettien pituudet on annettu.

Lause 2.2.2. *Olkoon* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *ja* $\alpha \in \mathbb{R}$. *Normilla on seuraavat ominaisuudet.*

1. $\|\mathbf{x}\| \geq 0$ *ja* $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ *(kolmioepäyhtälö)*

Sisätulolle ja normille pätee lisäksi nk. Schwarzin epäyhtälö $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$.

Sisätulon ja normin avulla voidaan määrittää vektoreiden välinen kulma. Vektoreiden $\mathbf{x}$ ja $\mathbf{y}$ välinen kulma $\measuredangle(\mathbf{x}, \mathbf{y})$ toteuttaa yhtälön $\cos \measuredangle(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}/(\|\mathbf{x}\|\|\mathbf{y}\|)$, kun $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$. Huomaa, että ortogonaaliset vektorit ovat toisiaan vastaan kohtisuorassa.

Normin avulla voidaan määritellä vektoreiden välinen etäisyys; vektoreiden $\mathbf{x} \in \mathbb{R}^n$ ja $\mathbf{y} \in \mathbb{R}^n$ välinen etäisyys on

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_i (x_i - y_i)^2}.$$

Tämän kaltainen etäisyyden mittaaminen on tärkeätä esimerkiksi ekonometriassa kun sovitetaan aineestoon mahdollisimman hyvää matemaattista mallia.

Etäisyyden avulla voidaan määritellä pisteille ympäristöt. Monet analyysin tuloksista ovat luonteeltaan lokaaleja, eli pätevät annetussa pisteen ympäristössä tai läheisyydessä.

Määäritelmä 2.2.3. Pisteen $\mathbf{x} \in \mathbb{R}^n$ avoin ε-ympäristö on

$$N_\varepsilon(\mathbf{x}) = \{\mathbf{y} : \|\mathbf{y} - \mathbf{x}\| < \varepsilon\}.$$

2.3 Hypertasot ja puoliavaruudet

Yhtälö $ax_1 + bx_2 = c$, $a \neq 0$ tai $b \neq 0$, kuvaa suoraa tasossa. Esim. kun $b \neq 0$ saadaan $x_2 = -ax_1/b + c/b$, eli kyseessä on (x_1, x_2) tason suora jonka kulmakerroin on $-a/b$.

Tämä antaa myös geometrisen tulkinnan yhtälöparin ratkaisulle. Yhtälöparin

$$\begin{aligned} ax_1 + bx_2 &= e \\ cx_1 + dx_2 &= f \end{aligned}$$

ratkaisu on kahden suoran leikkauspiste. Ratkaisua ei ole jos suorat ovat yhden suuntaiset (ja toisistaan poikkeavat). Ratkaisuja on äärettömästi, jos suorat ovat samat. Muuten ratkaisu on yksikäsitteinen. Tulkinta yleistyy jossain määrin myös useampiuloitteisille veltoreille.

Tason $\mathbb{R}^2$ suoran luonteva yleistys $\mathbb{R}^n$:ssä on hypertaso. Huom. $\mathbb{R}^3$:n hypertaso on tavallinen taso.

Määäritelmä 2.3.1. Vektorin $\mathbf{p} \in \mathbb{R}^n$ ja skalaarin $\alpha \in \mathbb{R}$ määräämä hypertaso on joukko $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} = \alpha\}$.

Hypertaso voidaan määritellä myös sen normaalin $\mathbf{p}$ ja yhden hypertason vektorin $\mathbf{y}$ avulla. Hypertason muodostaa ne vektorit $\mathbf{x}$, jotka toteuttavat yhtälön $\mathbf{p} \cdot (\mathbf{x} - \mathbf{y}) = 0$. Huomaa, että $\mathbf{p}$ on siis ortogonaalinen kaikkia vektoreita $\mathbf{x} - \mathbf{y}$ vastaan. Siksi $\mathbf{p}$:tä kutsutaan tason normaaliksi.

Hypertason ylä- ja alapuolisia joukkoja kutsutaan hypertason määräämiksi puoliavaruuksiksi. Hypertason $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} = \alpha\}$ määrittämät avoimet puoliavaruudet ovat $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} > \alpha\}$ ja $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} < \alpha\}$. Vastaavasti suljetut puoliavaruudet ovat $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} \geq \alpha\}$ ja $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} \leq \alpha\}$.

Esimerkki 2.3.1. Oletetaan että $\mathbf{x} \in \mathbb{R}^n$ kuvaa kuluttajan hyödykekoria ja $\mathbf{p} \in \mathbb{R}^n$ kuvaa hyödykkeiden hintoja. Kuluttajan teoriassa budjettiyhtälö $\sum_i p_i x_i = w$ määrittää hypertason. Tässä w on kuluttajan varallisuus. Hypertaso voidaan kirjoittaa myös muodossa $\mathbf{p} \cdot \mathbf{x} = w$. Kuluttajan budjettijoukko on puoliavaruuksien $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} \leq w\}$ ja $\{\mathbf{x} \in \mathbb{R}^n : x_i \geq 0\}$, $i = 1, \ldots, n$ leikkaus $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} \leq w, x_i \geq 0, i = 1, \ldots, n\}$.

2.4 Lineaarinen riippuvuus ja kanta

Määritellään aluksi vektoreiden lineaarikombinaatiot.

Määäritelmä 2.4.1. Vektoreista $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ muodostettu uusi vektori $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k$ on vektoreiden $\mathbf{x}_1, \ldots, \mathbf{x}_k$ lineaarikombinaatio.

Lineaarisella riippuvuudella tarkoitetaan sitä että joukko vektoreita voidaan esittää toistensa lineaarikombinaatioina.

Määäritelmä 2.4.2. Vektorit ovat lineaarisesti riippuvia jos $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}$ joillakin kertoimilla $\alpha_1, \ldots. \alpha_k$ joista ainakin yksi poikkeaa nollasta.

Määäritelmä 2.4.3. Jos $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}$ vain jos $\alpha_i = 0$ kaikilla i niin vektorit ovat lineaarisesti riippumattomia

Tulemme havaitsemaan, että lineaarisen yhtälöryhmän ratkaisujen määrä on yhteydessä lineaariseen riippuvuuteen. Esim. johdannon viimeisessä esimerkissä lineaarisen riippuvuuden avulla voidaan selittää miksi ratkaisuja ei ole.

Määäritelmä 2.4.4. Vektorit $\mathbf{x}_1, \ldots, \mathbf{x}_n$ muodostavat avaruuden $\mathbb{R}^n$ kannan jos ne ovat lineaarisesti riippumattomia.

Kannalla on seuraavat ominaisuudet:

1. $\mathbb{R}^n$:n kannassa on aina n vektoria
2. mikä hyvänsä $\mathbb{R}^n$:n vektori voidaan esittää kantavektoreiden lineaarikombinaationa ja tämä esitys on aina yksikäsitteinen.

Jos $\mathbf{y}$:n esitys kannassa $\mathbf{x}_1, \ldots, \mathbf{x}_n$ on $\mathbf{y} = \sum_i \alpha_i \mathbf{x}_i$ luvut $\alpha_1, \ldots, \alpha_n$ muodostavat $\mathbf{y}$:n koordinaatit kyseisessä kannassa.

Avaruuden $\mathbb{R}^n$ luonnollinen kanta on vektoreiden $\mathbf{e}_i$, $i = 1, \ldots, n$, muodostama kanta, vektorin $\mathbf{e}_i$ kaikki muut alkiot ovat nollia paitsi i:s joka on yksi.

2.4.1 Vektorialiavaruudet

Luvussa 2.1 käytiin läpi ominaisuudet, joita vektoreilla on, eli niiden yhteenlaskun ja skalaarilla kertomisen säännöt. Yleisesti joukkoa, jolla vastaavat laskusäännöt ovat voimassa, eli yhteenlasku ja skalaarilla kertominen sanotaan vektoriavaruudeksi. Huom. V on siis vektoriavaruus, jos skalaarilla kerrotut V:n elementit ja yhteenlasketut vektorit ovat yhä V:n elementtejä.

Määäritelmä 2.4.5. Joukko V on $\mathbb{R}^n$:nnän vektorialiavaruus, jos $\mathbf{x} + \mathbf{y} \in V$ kaikilla $\mathbf{x}, \mathbf{y} \in V$ ja $\alpha\mathbf{x} \in V$ kaikilla $\mathbf{x} \in V$ ja $\alpha \in \mathbb{R}$.

Esimerkki 2.4.1. Tarkastellaan joukkoja $V_0 = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}$ ja $V_1 = \{(x_1, x_2, 1) : x_1, x_2 \in \mathbb{R}\}$. Geometrisesti V_0 esittää (x_1, x_2)-tasoa avaruudessa ja V_1 toista samansuuntaista tasoa. Voimme huomata, että näistä V_0 on vektoriavaruus, koska $x_1, x_2 \in V_0$, implikoi, että $x_1 + x_2 \in V_0$ ja $\alpha x \in V_0$ kaikilla $\alpha \in \mathbb{R}$. Joukko V_1 ei ole vektorialiavaruus. Erityisesti havaitsemme, että $\mathbf{0} \notin V_1$. Huom. $\mathbf{0}$ kuuluu aina vektorialiavaruuteen.

Vektorialiavaruudet ovat siis $\mathbb{R}^n$:nnän osajoukkoja, jotka ovat vektoriavaruuksia kuten $\mathbb{R}^n$.

Esimerkki 2.4.2. Hypertaso joka kulkee origon kautta, eli vektorit $\mathbf{x}$, jotka toteuttavat yhtälö $\mathbf{p} \cdot \mathbf{x} = 0$ muodostavat vektorialiavaruuden.

Voimme määritellä vektorialiavaruuden kannan samoin kuten aikaisemmin, eli vektoreiden on oltava lineaarisesti riippumattomia ja niiden lineaarikombinaationa on pystyttävä esittämään koko vektorialiavaruus.

Määäritelmä 2.4.6. Vektorit $\mathbf{v}_1, \ldots, \mathbf{v}_k$ muodostavat V:n kannan jos ne ovat lineaarisesti riippumattomia ja $V = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_i \alpha_i \mathbf{v}_i,\ \alpha_i \in \mathbb{R},\ i = 1, \ldots, k\}$.

Yllä esitetyssä määritelmässä joukkoa $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_i \alpha_i \mathbf{v}_i,\ \alpha_i \in \mathbb{R},\ i = 1, \ldots, k\}$ sanotaan vektoreiden $\mathbf{v}_1, \ldots, \mathbf{v}_k$ virittämäksi avaruudeksi. Tälle joukolle käytetään merkintää $\mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.

Kantaa koskee seuraavat tulokset.

Lause 2.4.1. *Olkoon V avaruuden $\mathbb{R}^n$ vektorialiavaruus. Tällöin*

1. *jos V:llä on kanta, jossa on $k \leq n$ alkiota ja $\mathbf{w}^1, \ldots, \mathbf{w}^m \in V$, $m \geq k$, niin vektorit $\mathbf{w}^1, \ldots, \mathbf{w}^m$ ovat lineaarisesti riippuvia*
2. *jos V:llä on kanta, jossa on n elementtiä ja toinen kanta, jossa on m elementtiä, niin $m = n$.*

Kannan avulla voimme määritellä vektorialiavaruuden dimension.

Määäritelmä 2.4.7. Vektorialiavaruuden V dimensio $\dim(V)$ on k jos V:n kannassa on k vektoria.

Esimerkki 2.4.3. Normaalin $\mathbf{p} \neq \mathbf{0}$ määräämään origon kautta kulkevan hypertason dimensio on $n - 1$.

Luku 3

Matriisialgebraa

Matriisit ovat keskeisessä asemassa lineaarialgebrassa. Matriisilla tarkoitetaan tietynlaista järjestettyä taulukkoa lukuja. Matriisien avulla voidaan määritellä lineaarikuvauksia, jotka puolestaan ovat tärkeitä sovelluksissa, koska niiden avulla mallinnettavia ilmiöitä on runsaasti. Epälineaarisest mallit puolestaan voi linearisoida, jolloin päädytään tarkastelemaan lineaarisia malleja. Lisäksi matriisit tarjoavat näppärän tavan merkitä asioita. Matriisialgebrassa tarkastellaan matriiseilla tehtäviä laskuoperaatiota. Niiden avulla on mahdollista ratkaista lineaarisia yhtälöitä tehokkaasti ja systemaattisesti.

3.1 Matriisin käsite

Matriisilla tarkoitetaan reaalilukujen a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ muodostamaa järjestettyä taulukkoa $\mathbf{A}$, joka on muotoa

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

Lukuja a_{ij} sanotaan matriisin *alkioiksi.* Alkio a_{ij} on siis matriisin i:nnen rivin (vaakarivin) ja j:nnen sarakkeen (pystyrivin) leikkauskohdassa sijaitseva reaaliluku. Matriisi on siis taulukko jossa joka rivillä on sama määrä alkioita ja joka sarakkeessa on sama määrä alkioita. Matriisista $\mathbf{A}$ voidaan käyttää myös merkintää (a_{ij}), $i = 1, \ldots, m$, $j = 1, \ldots, n$, tai yksinkertaisesti merkintää (a_{ij}).

Matriisin, jolla on m riviä ja n saraketta, sanotaan olevan $m \times n$ -matriisi (lue m kertaa n matriisi). Tällöin voidaan käyttää tarkempaa merkintätapaa $A \in \mathbb{R}^{m \times n}$, $\mathbf{A} = \mathbf{A}_{m \times n}$ tai $(a_{ij}) = (a_{ij})_{m \times n}$. Mikäli $m = n$, matriisin sanotaan olevan *neliömatriisi.* Toisinaan on tarpeellista pystyä viittaamaan matriisin riveihin (vaakariveihin) ja sarakkeisiin (pystyriveihin). Tällöin merkinnällä $\mathbf{A}^{(j)}$ tarkoitetaan matriisin $\mathbf{A}$ j:nnettä saraketta, eli

$$\mathbf{A}^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

ja vastaavasti merkinnällä $\mathbf{A}_i$ tarkoitetaan matriisin $\mathbf{A}$ i:nnettä vaakariviä, eli

$$\mathbf{A}_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}.$$

Matriisin j:nnes sarake (pystyrivi) on siis $m \times 1$ -matriisi, jollaisia kutsutaan joskus *pystyvektoreiksi* tai *sarakevektoreiksi*, ja matriisin i:nnes rivi (vaakarivi) on $1 \times n$ -matriisi, jollaisia kutsutaan saman logiikan mukaisesti *vaakavektoreiksi* tai *rivivektoreiksi*. Näitä merkintätapoja käyttäen matriisi $\mathbf{A}$ voidaan esittää myös muodossa

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{(1)} & \mathbf{A}^{(2)} & \cdots & \mathbf{A}^{(n)} \end{pmatrix}.$$

Huomaa, että $\mathbb{R}^n$:n vektorit voidaan tulkita matriiseina: vaakavektori on $1 \times n$-matriisi ja pystyvektori on $n \times 1$-matriisi. Jatkossa, kun viitataan $\mathbb{R}^n$:n vektoriin $\mathbf{x} = (x_1, \ldots, x_n)$, tarkoitetaan pystyvektoria

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Matriisit $\mathbf{A}_{m\times n}$ ja $\mathbf{B}_{r\times s}$ ovat samat, eli $\mathbf{A}_{m\times n}$=$\mathbf{B}_{r\times s}$, silloin ja vain silloin kun seuraavat kaksi ehtoa ovat voimassa:

1. $m = r$ ja $n = s$ (ts. matriisit ovat samaa tyyppiä ts. saman kokoisia),
2. $a_{ij} = b_{ij}$, kaikilla $i,\ j$.

Esimerkki 3.1.1. Matriisi

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

on 2×3 -matriiisi, jonka rivit ovat $\begin{pmatrix} 1 & -3 & 4 \end{pmatrix}$ sekä $\begin{pmatrix} 0 & 5 & -2 \end{pmatrix}$ ja sarakkeet ovat

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ -2 \end{pmatrix}.$$

3.2 Matriiseilla laskeminen

Seuraavaksi määritellään matriisien laskutoimituksia, joiden huomataan toteuttavan samanlaisia laskusääntöjä, kuin tutut lukujen laskutoimitukset.

3.2.1 Matriisien summa

Matriisien **A** ja **B** summa **A**+**B** on määritelty ainoastaan silloin, kun kyseiset matriisit ovat samaa tyyppiä. Mikäli kummatkin ovat $m \times n$ -matriiseja, eli

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{ja} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix},$$

on summan **A**+**B** i:nnellä vaakarivillä j:nnessä sarakkeessa sijaitseva alkio vastaavassa kohdassa matriiseissa **A** ja **B** sijaitsevien alkioiden summa. Toisin sanottuna **A**+**B**=$(a_{ij})+(b_{ij}) = (a_{ij} + b_{ij})$, eli

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

3.2.2 Matriisin kertominen skalaarilla

Matriiisin **A** ja reaaliluvun k tulolla $k\mathbf{A}$ tarkoitetaan sellaista matriiisia, joka saadaan kun matriiisin **A** jokainen alkio erikseen kerrotaan luvulla k:

$$k\mathbf{A} = \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{pmatrix}.$$

Tässä kohtaa on syytä huomata, että matriisien **A** ja **B** erotus **A-B** voidaan laskea näiden määritelmien avulla kirjoittamalla **A-B**=**A**+$(-1\mathbf{B})$.

Esimerkki 3.2.1. Olkoon

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{pmatrix} \quad \text{ja} \quad \mathbf{B} = \begin{pmatrix} 3 & 0 & 2 \\ -7 & 1 & 8 \end{pmatrix}.$$

Tällöin

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

$$3\mathbf{A} = \begin{pmatrix} 3\cdot 1 & 3\cdot(-2) & 3\cdot 3 \\ 3\cdot 4 & 3\cdot 5 & 3\cdot(-6) \end{pmatrix} = \begin{pmatrix} 3 & -6 & 9 \\ 12 & 15 & -18 \end{pmatrix},$$

$$2\mathbf{A} - 3\mathbf{B} = \begin{pmatrix} 2 & -4 & 6 \\ 8 & 10 & -12 \end{pmatrix} + \begin{pmatrix} -9 & 0 & -6 \\ 21 & -3 & -24 \end{pmatrix} = \begin{pmatrix} -7 & -4 & 0 \\ 29 & 7 & -36 \end{pmatrix}.$$

3.2.3 Matriisien laskulait

Suoraan määritelmien perusteella on helppo nähdä, että matriiseilla laskeminen toteuttaa seuraavat laskulait, missä on käytetty merkintää[1]

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

(i) $(\mathbf{A}+\mathbf{B})+\mathbf{C}= \mathbf{A}+(\mathbf{B}+\mathbf{C})$,

(ii) $\mathbf{A}+\mathbf{0}=\mathbf{A}$,

(iii) $\mathbf{A}+(-\mathbf{A})=\mathbf{0}$,

(iv) $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$,

(v) $k_1(\mathbf{A}+\mathbf{B}) = k_1\mathbf{A} + k_1\mathbf{B}$,

(vi) $(k_1 + k_2)\mathbf{A} = k_1\mathbf{A} + k_2\mathbf{A}$,

(vii) $(k_1k_2)\mathbf{A} = k_1(k_2\mathbf{A})$,

(viii) $1 \cdot \mathbf{A} = \mathbf{A}$.

3.3 Matriisitulo

Matriisien tulo ei ole määritelty aivan yhtä intuitiivisesti kuin matriisien summa tai matriisin kertominen luvulla (ts. matriisien tuloa ei määritellä alkioittain). Se miksi matriisien tulo määritellään seuraavalla tavalla, selviää myöhemmin.

Matriisien $\mathbf{A}_{m\times n}$ ja $\mathbf{B}_{r\times s}$ tulo on mahdollista laskea jos ja vain jos $n = r$. Tällöin matriisissa $\mathbf{A}$ on siis yhtä monta saraketta, kuin matriisissa $\mathbf{B}$ on rivejä. Mikäli matriisien tulo on määritelty, tarkoitetaan tulolla $\mathbf{AB}$ sellaista $m \times s$ -matriisia, jonka i:nnellä rivillä j:nnessä sarakkeessa on matriisin $\mathbf{A}$ i:nnen rivin muodostaman vektorin ja matriisin $\mathbf{B}$ j:nnen sarakkeen muodostaman vektorin sisätulo. Toisin sanottuna $\mathbf{AB}=(c_{ij})_{m\times s}$, missä

$$c_{ij} = \mathbf{A}_i \cdot \mathbf{B}^{(j)} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

[1] Tällaista matriisia kutsutaan *nollamatriisiksi*.

Kirjoittamalla matriisien tulo auki eksplisiittisesti saadaan siis

$$
\begin{aligned}
\mathbf{AB} &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{ns} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{A}_1 \cdot \mathbf{B}^{(1)} & \mathbf{A}_1 \cdot \mathbf{B}^{(2)} & \dots & \mathbf{A}_1 \cdot \mathbf{B}^{(s)} \\ \mathbf{A}_2 \cdot \mathbf{B}^{(1)} & \mathbf{A}_2 \cdot \mathbf{B}^{(2)} & \dots & \mathbf{A}_2 \cdot \mathbf{B}^{(s)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_m \cdot \mathbf{B}^{(1)} & \mathbf{A}_m \cdot \mathbf{B}^{(2)} & \dots & \mathbf{A}_m \cdot \mathbf{B}^{(s)} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{k=1}^n a_{1k}b_{k1} & \sum_{k=1}^n a_{1k}b_{k2} & \dots & \sum_{k=1}^n a_{1k}b_{ks} \\ \sum_{k=1}^n a_{2k}b_{k1} & \sum_{k=1}^n a_{2k}b_{k2} & \dots & \sum_{k=1}^n a_{2k}b_{ks} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{mk}b_{k1} & \sum_{k=1}^n a_{mk}b_{k2} & \dots & \sum_{k=1}^n a_{mk}b_{ks} \end{pmatrix}.
\end{aligned}
$$

Esimerkki 3.3.1. Tarkastellaan matriiseja

$$
\mathbf{A} = \begin{pmatrix} 3 & 5 & 8 \\ 10 & 4 & 1 \end{pmatrix}_{2\times 3}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}_{2\times 2}, \quad \mathbf{C} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}_{2\times 2}.
$$

(i) $\mathbf{BC} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 0 & 1\cdot 1 + 2\cdot 2 \\ 3\cdot 1 + 4\cdot 0 & 3\cdot 1 + 4\cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 3 & 11 \end{pmatrix}.$

(ii) Matriisitulo **AC** ei ole määritelty, koska matriisissa **A** on enemmän sarakkeita (3), kuin matriisissa **B** rivejä (2). Toisin päin laskettu matriisitulo on kuitenkin määritelty,

$$
\begin{aligned}
\mathbf{CA} &= \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 3 & 5 & 8 \\ 10 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 3 + 1\cdot 10 & 1\cdot 5 + 1\cdot 4 & 1\cdot 8 + 1\cdot 1 \\ 0\cdot 3 + 2\cdot 10 & 0\cdot 5 + 2\cdot 4 & 0\cdot 8 + 2\cdot 1 \end{pmatrix} \\
&= \begin{pmatrix} 13 & 9 & 9 \\ 20 & 8 & 2 \end{pmatrix}.
\end{aligned}
$$

Matriisitulo on määritelty kummin tahansa päin ainoastaan silloin, kun toinen matriiseista on muotoa $m \times n$ ja toinen muotoa $n \times m$. Siis erikoisesti neliömatriiseilla.

Esimerkki 3.3.2. (i) Ehdosta **AB**=**0** ei seuraa, että toisen tulontekijän **A** tai **B** täytyy olla nollamatriisi **0** (toisin kuin luvuilla laskettaessa). Esimerkiksi

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

(ii) Vaikka matriisien **A** ja **B** tulo olisi määritelty kumminkin päin, silti ei tarvitse olla voimassa **AB**=**BA**. Esimerkiksi

$$\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 5 \\ 13 & 2 \end{pmatrix},$$

mutta

$$\begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 11 \\ 7 & 8 \end{pmatrix}.$$

Vaikka matriisitulo ei toteuta kaikkia totuttuja laskusääntöjä, sillä on kuitenkin useita odotettuja ominaisuuksia. Seuraavat säännöt ovat voimassa edellyttäen, että kyseiset lausekkeet ovat määriteltyjä:

1. **A**(**BC**)=(**AB**)**C**, (assosiatiivisuuslaki)
2. **A**(**B**+**C**)=**AB**+**AC**, (**A**+**B**)**C**=**AC**+**BC**, (oikea ja vasen distributiivilaki)
3. k**AB**=(k**A**)**B**=**A**(k**B**), kaikilla $k \in \mathbb{R}$. (vakion siirto)

Toisinaan matriisi halutaan osittaa useammaksi matriisiksi ja matriisitulo halutaan muodostaa ositetuille matriiseille. Olkoon esimerkiksi

$$\underset{n\times m}{\mathbf{A}} = \begin{pmatrix} \underset{n\times k}{\mathbf{B}} & \underset{n\times l}{\mathbf{C}} \end{pmatrix}$$

ja

$$\underset{m\times r}{\mathbf{D}} = \begin{pmatrix} \underset{k\times r}{\mathbf{E}} \\ \underset{l\times r}{\mathbf{F}} \end{pmatrix},$$

missä $m = k + l$. Tällöin

$$\mathbf{AD} = \mathbf{BE} + \mathbf{CF}.$$

Erityisesti kun **B** korvataan vektorilla $\mathbf{x} \in \mathbb{R}^m$ ja

$$\mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix},$$

missä $\mathbf{y} \in \mathbb{R}^k$ ja $\mathbf{z} \in \mathbb{R}^l$, saadaan

$$\mathbf{Ax} = \mathbf{By} + \mathbf{Cz}.$$

3.3.1 Matriisin transponointi

Olkoon **A** $m \times n$ -matriisi

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

Matriisin $\mathbf{A}$ *transponoidulla matriisilla* eli lyhyesti *transpoosilla*, josta käytetään merkintää $\mathbf{A}^\top$, tarkoitetaan sellaista matriisia jonka i:s rivi on matriisin $\mathbf{A}$ i:s sarake. Symbolisesti ilmaistuna

$$\mathbf{A}^\top = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}_{n\times m}.$$

Esimerkki 3.3.3. (i) Huomaa, että pystyvektorin transpoosi on vaakavektori ja päinvastoin. Siis

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}^\top = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \quad \text{ja} \quad \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}^\top = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

(ii)

$$\text{Jos } \mathbf{A} = \begin{pmatrix} 1 & -4 & 2 \\ 0 & 10 & 7 \\ 0 & 9 & -5 \end{pmatrix}, \text{ niin } \mathbf{A}^\top = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 10 & 9 \\ 2 & 7 & -5 \end{pmatrix}.$$

Eräs tärkeä luokka matriiseja ovat ne matriisit, joiden transpoosi on matriisi itse. Tämä tarkoittaa erikoisesti sitä, että ainoastaan neliömatriisi voi olla symmetrinen.

Määäritelmä 3.3.1. Matriisin $\mathbf{A}$ sanotaan olevan *symmetrinen*, mikäli $\mathbf{A} = \mathbf{A}^\top$. Tällöin siis $a_{ij} = a_{ji}$ kaikilla i, j.

Listataan vielä muutamia transpoosin ominaisuuksia. Ensinnäkin matriisin transponoinnille pätee seuraavat säännöt.

1. $(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$
2. $(\mathbf{A}^\top)^\top = \mathbf{A}$

On myös hyvä huomata että skaalaarin α transpoosille pätee $\alpha^\top = \alpha$.

Tarkastellaan seuraavaksi transpoosia kahden tai useamman matriisin tulona muodostetulle matriisille. Oleteaan, että $\mathbf{A}$ on $m \times s$ -matriisi ja $\mathbf{B}$ on $s \times n$ -matriisi. Tällöin tulo $\mathbf{AB}$ sekä tulo $\mathbf{B}^\top\mathbf{A}^\top$ ovat määriteltyjä ($\mathbf{B}^\top$ on $n \times s$ -matriisi ja $\mathbf{A}^\top$ on $s \times m$ -matriisi) ja voidaan osoittaa, että

$$(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top. \quad \text{(tulon transpoosi)}$$

Lisäksi tämä kaava voidaan induktiivisesti yleistää muotoon

$$(\mathbf{A}_1\mathbf{A}_2 \cdot \ldots \cdot \mathbf{A}_k)^\top = \mathbf{A}_k^\top \mathbf{A}_{k-1}^\top \cdot \ldots \cdot \mathbf{A}_1^\top,$$

kunhan tässä esiintyvät tulot on määritelty.

Viimeisenä huomiona transpoosista todettakoon, että matriisitulon ja transpoosin avulla voidaan lausua vektoreiden välinen sisätulo: $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\top \mathbf{y}$. Huom. tässä monisteessä noudatetaan käytäntöä, että $\mathbb{R}^n$:nnän vektorit ovat pystyvektoreita.

3.4 Tärkeitä matriisityyppejä

Tässä luvussa tarkastellaan joitakin usein vastaan tulevia tärkeitä matriisityyppejä.

3.4.1 Diagonaalimatriisi

Tarkastellaan neliömatriisia $\mathbf{A}$=$(a_{ij})_{n\times n}$. Tämän matriisin *diagonaali*, eli *päälävistäjä* (usein puhutaan vain lävistäjästä), koostuu alkioista $a_{11}, a_{22}, \ldots, a_{nn}$. Näiden päälävistäjällä sijaitsevien alkioiden summaa kutsutaan matriisin $\mathbf{A}$ *jäljeksi* ja siitä käytetään merkintää tr($\mathbf{A}$). Formaalisti ilmaistuna

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

Matriisin $\mathbf{A}$ sanotaan olevan *diagonaalimatriisi*, mikäli sen kaikki muut alkiot, paitsi mahdollisesti päälävistäjän alkiot, ovat nollia. Tällöin voidaan käyttää merkintää

$$\mathbf{A} = \mathrm{diag}(a_{11}, \ldots, a_{nn}),$$

koska diagonaalimatriisi (tai lävistäjämatriisi) pystytään spesifioimaan antamalla sen päälävistäjällä sijaitsevat alkiot.

Esimerkki 3.4.1. Olkoon

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 4 & 1 \\ 3 & 0 & 5 \end{pmatrix} \quad \text{ja} \quad \mathbf{B} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -6 \end{pmatrix}.$$

Matriisi $\mathbf{B}$ on diagonaalimatriisi, koska sen kaikki nollasta poikkeavat alkiot sijaitsevat päälävistäjällä. Voidaan siis käyttää myös merkintää $\mathbf{B}$=diag(8,0,-6). Matriisin $\mathbf{A}$ jälki on tr($\mathbf{A}$)=2+4+5=11.

3.4.2 Identiteettimatriisi

Identiteettimatriisi, eli yksikkömatriisi, josta käytetään merkintää $\mathbf{I}_{n\times n}$, on sellainen diagonaalimatriisi jonka kaikki päälävistäjän alkiot ovat ykkösiä. Edellä esitettyjä merkintöjä käyttäen

$$\underset{n\times n}{\mathbf{I}} = \mathrm{diag}(\underbrace{1, \ldots, 1}_{n\text{ kpl}}),$$

eli

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

Kun mikä tahansa matriisi $\mathbf{A}_{n\times n}$ kerrotaan identiteettimatriisilla $\mathbf{I}_{n\times n}$, saadaan tulokseksi sama matriisi, eli

$$\mathbf{IA} = \mathbf{AI} = \mathbf{A}.$$

3.4.3 Kolmiomatriisi

Neliömatriisin $\mathbf{A}=(a_{ij})$ sanotaan olevan *yläkolmiomatriisi*, mikäli sen kaikki päälävistäjän alapuolella sijaitsevat alkiot ovat nollia. Toisin sanottuna $a_{ij} = 0$ kaikilla $i > j$. Esimerkiksi matriisit

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}, \quad \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{pmatrix} \text{ ja } \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ 0 & c_{22} & c_{23} & c_{24} \\ 0 & 0 & c_{33} & c_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix}$$

ovat kaikki yläkolmiomatriiseja. *Alakolmiomatriisilla* tarkoitetaan vastaavasti sellaista matriisia, jolla kaikki päälävistäjän yläpuolella sijaitsevat alkiot ovat nollia, eli $a_{ij} = 0$ jos $i < j$.

3.4.4 Idempotentti matriisi

Jos $\mathbf{A}$ on $n \times n$ -matriisi, eli neliömatriisi, sen *potenssi* voidaan määritellä luonnollisella tavalla

$$\mathbf{A}^0 = \underset{n\times n}{\mathbf{I}} \text{ ja } \mathbf{A}^k = \underbrace{\mathbf{A}\cdot\ldots\cdot\mathbf{A}}_{k\text{ kpl}} \text{ kaikilla } k \geq 1.$$

Tällöin matriisitulon assosiatiivisuuden perusteella on voimassa

$$\mathbf{A}^k\mathbf{A}^h = \mathbf{A}^{k+h} \text{ ja } \left(\mathbf{A}^k\right)^h = \mathbf{A}^{kh} \text{ kaikilla } k, h \geq 0.$$

Huomaa kuitenkin, että

$$(\mathbf{AB})^k = \mathbf{ABAB} \cdot \ldots \cdot \mathbf{AB}$$

on yleensä eri matriisi kuin $\mathbf{A}^k\mathbf{B}^k$ (toisin kuin luvuilla laskettaessa).

Symmetrisen matriisin $\mathbf{A}$ sanotaan olevan *idempotentti*, mikäli se toteuttaa yhtälön $\mathbf{A}^2 = \mathbf{A}$. Matriisitulon assosiatiivisuuden perusteella idempotentille matriisille täytyy olla voimassa yleisemminkin $\mathbf{A}^n = \mathbf{A}$ kaikilla $n \geq 1$.

Esimerkki 3.4.2. Tarkastetaan onko matriisi

$$\mathbf{A} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix}$$

idempotentti. Tämä matriisi on symmetrinen, sillä $\mathbf{A} = \mathbf{A}^\top$. Lisäksi

$$\mathbf{A}^2 = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix} = \begin{pmatrix} \frac{1}{5}^2 + \frac{2}{5}^2 & \frac{1}{5} \cdot \frac{2}{5} + \frac{2}{5} \cdot \frac{4}{5} \\ \frac{1}{5} \cdot \frac{2}{5} + \frac{2}{5} \cdot \frac{4}{5} & \frac{2}{5}^2 + \frac{4}{5}^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix},$$

joten kyseessä on idempotentti matriisi.

Luku 4

Lineaarikuvaukset ja yhtälöryhmät

Tässä luvussa käydään läpi lineaarikuvausten ja matriisien yhteys. Luvun tarkoituksena on antaa käsitteelliset perusteet lineaarialgebran keskeisten tulosten ymmärtämiselle. Esimerkiksi luvussa hahmotellaan miten yhtälöryhmän ratkaisujen määrä ja olemassaolo liittyvät surjektion ja injektion käsitteisiin.

4.1 Matriisit lineaarikuvauksina

Funktiota $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ sanotaan lineaariseksi jos $F(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha F(\mathbf{x}) + \beta F(\mathbf{y})$ kaikilla $\alpha, \beta \in \mathbb{R}$ ja $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Esimerkki 4.1.1. Yksinkertainen esimerkki lineaarikuvauksesta on tason vektorin kääntäminen annetun kulman verran. Toinen esimerkki on vektorin pituuden skaalaaminen vakiolla, eli kuvaus $F(\mathbf{x}) = \alpha\mathbf{x}$.

Seuraava tulos kytkee matriisit ja lineaarikuvaukset toisiinsa. Erityisesti voimme tulkita kaikki matriisit lineaarikuvauksina.

Lause 4.1.1. *Jos $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ on lineaarinen funktio niin tällöin löytyy $\mathbf{A} \in \mathbb{R}^{m\times n}$ siten, että $F(\mathbf{x}) = \mathbf{A}\mathbf{x}$ kaikilla $\mathbf{x} \in \mathbb{R}^n$.*

Matriisia $\mathbf{A}$, joka määrittää lineaarikuvauksen F, sanotaan F:n esitysmatriisiksi.

Palataan vielä hetkeksi matriisituloon. Tarkastellaan kahta funktiota $f : X \mapsto Y$ ja $g : Y \mapsto Z$, tai $\mathbf{y} = f(\mathbf{x})$, $\mathbf{z} = g(\mathbf{y})$. Muodostetaan yhdistetty funktion $h(\mathbf{x}) = g(f(\mathbf{x}))$.

Oletetaan seuraavaksi, että f ja g ovat lineaarisia funktioita ja $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ ja $Z = \mathbb{R}^k$. Edellä esitetyn lauseen perusteella löytyy matriisit $\mathbf{A} \in \mathbb{R}^{m\times n}$ ja $\mathbf{B} \in \mathbb{R}^{k\times m}$ siten että $f(\mathbf{x}) = \mathbf{Ax}$ ja $g(\mathbf{y}) = \mathbf{By}$.

Olkoon nyt $\mathbf{y} = f(\mathbf{x}) = \mathbf{Ax}$, eli

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + \ldots + a_{mn}x_n, \end{aligned}$$

jolloin $z = g(f(x))$ on

$$\begin{aligned} z_1 &= b_{11}(a_{11}x_1 + \ldots + a_{1n}x_n) + \ldots + b_{1m}(a_{m1}x_1 + \ldots + a_{mn}x_n) \\ &\vdots \\ z_k &= b_{k1}(a_{11}x_1 + \ldots + a_{1n}x_n) + \ldots + b_{km}(a_{m1}x_1 + \ldots + a_{mn}x_n). \end{aligned}$$

Huomaamme, että $\mathbf{z} = \mathbf{BAx}$, eli matriisitulo vastaa yhdistetyn funktion muodostamista.

Palautetaan mieleen pari tärkeää käsitettä. Olkoon F funktio X:ltä Y:lle. Joukon $A \subseteq Y$ alkukuvaa merkitään joukolla $F^{-1}(A) \subseteq X$. Funktiolla tarkoitetaan kuvausta joka liittää jokaiseen $\mathbf{x} \in X$ täsmälleen yhden Y:n alkion jota merkitään $F(\mathbf{x})$:llä.

Määäritelmä 4.1.1. Kuvaus $F : X \mapsto Y$ on

1. surjektio, jos jokaisella $\mathbf{y} \in Y$ on alkukuva $\mathbf{x} \in X$, eli $F^{-1}(Y) = X$
2. injektio, jos jokaisella $\mathbf{y} \in Y$ on enintään yksi alkukuva $\mathbf{x} \in X$, eli $F^{-1}(\mathbf{y})$ koostuu eniintään yhdestä alkiosta $\mathbf{y} = F(\mathbf{x})$
3. bijektio, jos F on sekä injektio että surjektio.

Huom. F on bijektio jos ja vain jos F^{-1} on funktio.

4.2 Lineaariset yhtälöryhmät

4.2.1 Lineaarisen yhtälöryhmän esitys

Tarkastellaan n:n tuntemattoman $x_1, \ldots, x_n$ *lineaarista yhtälöryhmää*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

missä a_{ij} ja b_i ovat vakioita. Tällaisen yhtälöryhmän ratkaisuksi kutsutaan sellaista tuntemattomien saamien arvojen joukkoa $x_1 = k_1, x_2 = k_2, \ldots, x_n = k_n$, tai vektoria

$u = (k_1, \dots, k_n)$, joka toteuttaa kaikki yhtälöryhmän yhtälöt. Matriisitulon määritelmän perusteella lineaarinen yhtälöryhmä voidaan kirjoittaa muodossa

$$\mathbf{Ax} = \mathbf{b},$$

missä

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{ja} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Huomaa, että lineaarinen yhtälöryhmä voidaan tulkita geometrisesti niin, että yhtälöryhmän toteuttavat vektorit kuuluvat hypertasojen leikkaukseen. Jos $\mathbf{A}_i$, $i = 1, \dots m$ ovat $\mathbf{A}$:n rivivektorit niin i's yhtälö on $\mathbf{A}_i\mathbf{x} = b_i$ ja tämä määrittää hypertason.

Esimerkki 4.2.1. Tarkastellaan lineaarista yhtälöryhmää

$$x_1 + 2x_2 - 5x_3 + 4x_4 = 3$$

$$2x_1 + 3x_2 + x_3 - 2x_4 = 1.$$

Tämä yhtälöryhmä voidaan kirjoittaa muodossa

$$\begin{pmatrix} 1 & 2 & -5 & 4 \\ 2 & 3 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Arvojoukko $x_1 = -8, x_2 = 4, x_3 = 1, x_4 = 2$ ei ole yhtälöryhmän ratkaisu, sillä

$$\begin{pmatrix} 1 & 2 & -5 & 4 \\ 2 & 3 & 1 & -2 \end{pmatrix} \begin{pmatrix} -8 \\ 4 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-8) + 2 \cdot 4 + (-5) \cdot 1 + 4 \cdot 2 \\ 2 \cdot (-8) + 3 \cdot 4 + 1 \cdot 1 + (-2) \cdot 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -7 \end{pmatrix}.$$

Esimerkki 4.2.2. Ajatellaan taloutta, jossa on n hyödykettä. Kaikki tuotteet ovat sekä lopputuotteita että potentiaalisia välituotteita. Tuotantoprosessissa siis kaikkia tuotteita kulutetaan ja tuotetaan samanaikaisesti. Oletetaan tuotantoprosessi lineaariseksi siten, ett x_i yksikon tuottamiseksi tuotetta i tarvitaan $a_{ji}x_j$ yksikkoä tuotetta j. Jos taloudessa tuotetaan nettotuotantovektori $(y_1, \dots, y_n)$ kokonaistuotosta, voidaan kokonaistuotantomäärät $(x_1, \dots, x_n)$ laskea seuraavasti:

$$\begin{aligned} x_1 - a_{11}x_1 - a_{12}x_2 - \dots - a_{1n}x_n &= y_1 \\ x_2 - a_{21}x_1 - a_{22}x_2 - \dots - a_{2n}x_n &= y_2 \\ &\vdots \\ x_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{nn}x_n &= y_n. \end{aligned}$$

Vektorimuodossa tämä voidaan kirjoittaa

$$\begin{pmatrix} 1-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & 1-a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \vdots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Seuraavaksi voimme kysyä mitä nettotuotantovektoreita $\mathbf{y}$ taloudessa voidaan tuottaa? Voimmeko tuottaa minkä tahansa vektorin $\mathbf{y} \geq \mathbf{0}$? Vastaus on kyllä, mikäli $\mathbf{I} - \mathbf{A}$ on nk. diagonaalidominantti matriisi.

Määäritelmä 4.2.1. Matriisia $\mathbf{A}$ sanotaan diagonaalidominantiksi matriisiksi, jos

1. diagonaalialkiot ovat positiivisia
2. diagonaalin ulkopuoliset alkiot eivät ole positiivisia
3. jokaisen sarakkeen alkioiden summa on positiivinen.

Lause 4.2.1. *Jos panos-tuotos taulukon määräävä matriisi* $(\mathbf{I} - \mathbf{A})$ *on diagonaalidominantti, niin kaikki positiiviset lopputuosvektorit ovat mahdollisia. Toisin sanoen kaikille* $\mathbf{y} \geq \mathbf{0}$ *löytyy* $\mathbf{x} \geq \mathbf{0}$ *siten, että* $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{y}$.

4.2.2 Lineaaristen yhtälöryhmien ominaisuuksia

Lineaarisia yhtälöryhmiä koskee seuraavat perussäännöt:

1. yhtälöryhmän ratkaisut eivät muutu, jos jonkin rivin molemmat puolet kertoo samalla nollasta poikkeavalla luvulla
2. yhtälöryhmän ratkaisut eivät muutu, jos jonkin yhtälön molemmille puolille lisää saman vakion
3. yhtälöryhmän ratkaisut eivät riipu siitä missä järjestyksessä yhtälöt esittää
4. jos $\mathbf{x} = (x_1, \ldots, x_n)$ toteuttaa yhtälöt

   $$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$

   ja

   $$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2,$$

   niin silloin $\mathbf{x}$ toteuttaa myös yhtälön

   $$(a_{11} + a_{21})x_1 + (a_{12} + a_{22})x_2 + \ldots + (a_{1n} + a_{2n})x_n = b_1 + b_2.$$

Edellä olevien sääntöjen soveltamista tullaan käsittelemään myöhemmin vaakariviooperaatioita tarkasteltaessa. Huomaa, että kaksi ensimmäistä sääntöä tarkoittavat sitä, että yhtälön määrittämä hypertaso ei muutu jos hypertason yhtälön molemmat puolet kertoo samalla luvulla tai molemmille puolille lisää saman luvun. Viimeinen sääntö on seurausta kahdesta ensimmäisestä. Kolmas sääntö tarkoittaa sitä, että sillä ei ole merkitystä, että missä järjestyksessä hypertasoista ottaa leikkauksen.

Tarkastellaan edellä esitettyjen sääntöjen viimeistä kohtaa. Oletetaan nyt, että meille on annettu kolme yhtälöä

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ (a_{11} + a_{21})x_1 + (a_{12} + a_{22})x_2 + \ldots + (a_{1n} + a_{2n})x_n &= b_1 + b_2. \end{aligned}$$

Kuten edellä todettiin, kun kaksi ensimmäistä toteutuu, niin myös kolmas toteuu. Näin ollen kolmas yhtälöistä on sikäli turha, että se ei tuo uutta informaatiota kahden ensimmäisen lisäksi. Jos tässä yhtälöryhmässä olisi $n = 3$ ja olisimme kiinnostuneet sen ratkaisemisesta, voisimme todeta, että ratkaisemisen kannalta kolmas yhtälö on tarpeeton. Ratkaisu ei myöskään voisi olla yksikäsitteinen, koska se määräytyy kahdesta yhtälöstä, kun tuntemattomia on kolme. Ajatellaan vielä tapausta, jossa viimeisen yhtälön oikean puolen $b_1 + b_2$ korvataan luvulla c. Tällöin yhtälöryhmällä ei voi olla ratkaisua jos $b_1 + b_2 \neq c$. Tässä yhtälöryhmässä havaittavat ilmiöt johtuvat siitä, että kolmannen yhtälöryhmän kerroinvektori $\mathbf{A}_1 + \mathbf{A}_2$ on lineaarisesti riippuva kahden ensimmäisen kerroin vektoreista $\mathbf{A}_1$ ja $\mathbf{A}_2$. Tulemme huomaamaan, että ratkaisujen määrän kannalta yhtälöryhmän kerroinmatriisin rivien (tai sarakkeiden) lineaarinen riippumattomuus on keskeisessä asemassa.

Havaitsimme aikaisemmin, että jokaisen matriisin avulla voidaan esittää jokin lineaarinen funktio. Olkoon nyt $F(\mathbf{x}) = \mathbf{A}\mathbf{x}$, missä $\mathbf{A}$ on siis F:n esitysmatriisi. Ratkaisun olemassaolo ja lukumäärä ovat kytköksissä injektion ja surjektion käsitteisiin:

1. jos $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ on surjektio, eli jokaisella kuvajoukon pisteellä on alkukuva, niin tällöin yhtälöryhmällä $\mathbf{A}\mathbf{x} = \mathbf{b}$ on ratkaisu jokaisella $\mathbf{b}$, ratkaisuja voi olla äärettömän monta
2. jos F on injektio, niin yhtälöryhmällä $\mathbf{A}\mathbf{x} = \mathbf{b}$ on ratkaisu vain silloin kun $\mathbf{b} \in F(\mathbb{R}^n)$ (eli $\mathbf{b}$ kuuluu F:n kuva-avaruuteen) ja tällöin ratkaisu on yksikäsitteinen
3. jos F on bijektio (huom. $n = m$), niin ratkaisu on yksikäsitteinen.

4.3 Yhtälöryhmän ratkaisu ja käänteismatriisi

Tarkastellaan nyt sellaista lineaarista yhtälöryhmää, jonka kerroinmatriisi $\mathbf{A}$ on $n \times n$ -matriisi, eli neliömatriisi. Kuten yllä on todettu, yhtälöryhmällä on yksikäsitteinen ratkaisu, jos sen esitysmatriisi määrittää lineaarikuvauksen, joka on bijektio. Bijektiivisyys tarkoitaa sitä että kuvauksella on käänteiskuvaus. Osoittautuu, että tämän käänteiskuvauksen esitysmatriisi on alkuperäisen esitysmatriisin käänteismatriisi.

Neliömatriisin $\mathbf{A}$ sanotaan olevan *säännöllinen* (tai kääntyvä tai ei-singulaarinen), mikäli on olemassa jokin sellainen $n \times n$-matriisi $\mathbf{B}$, jolle on voimassa

$$\mathbf{BA} = \mathbf{AB} = \mathbf{I}_{n\times n}.$$

Matriisia $\mathbf{B}$ kutsutaan tällöin matriisin $\mathbf{A}$ *käänteismatriisiksi* ja siitä käytetään merkintää $\mathbf{B} = \mathbf{A}^{-1}$. On syytä muistaa, että matriisilla ei välttämättä ole käänteismatriisia. Jos matriisilla on käänteismatriisi, niin se on yksikäsitteinen. Alla on listattu myös muita käänteismatrisin ominaisuuksia.

Lause 4.3.1. *Oletetaan, että $n \times n$ -matriisit $\mathbf{A}$ ja $\mathbf{B}$ ovat säännöllisiä. Tällöin*

1. $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$,
2. $\left(\mathbf{A}^{-1}\right)^{\top} = \left(\mathbf{A}^{\top}\right)^{-1}$,

3. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$,

Todistetaan esimerkinomaisesti kohta 3. Täytyy siis osoittaa, että $\mathbf{B}^{-1}\mathbf{A}^{-1}$ on matriisin $\mathbf{AB}$ käänteismatriisi. Näin on, sillä[1]

$$(\mathbf{AB})\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right) = \mathbf{A}\left(\mathbf{BB}^{-1}\right)\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}_{n\times n}.$$

(toisin päin laskettu tulo käsitellään vastaavasti).

Käänteismatriisin yhteys matriisin esittämän kuvauksen käänteiskuvaukseen on ilmaistu täsmällisesti alla.

Lause 4.3.2. *Jos $F : \mathbb{R}^n \mapsto \mathbb{R}^n$ on lineaarinen bijektio ja $\mathbf{A} \in \mathbb{R}^{n\times n}$ on F:n esitysmatriisi, niin $\mathbf{A}^{-1}$ on F^{-1}:n esitysmatriisi.*

Mikäli lineaarisen yhtälöryhmän kerroinmatriisi on säännöllinen, on yhtälöryhmällä yksikäsitteinen ratkaisu, joka saadaan kertomalla yhtälöryhmä puolittain matriisilla $\mathbf{A}^{-1}$. Formaalisti esitettynä

$$\mathbf{Ax} = \mathbf{b} \;\Rightarrow\; \mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b} \;\Rightarrow\; \mathbf{Ix} = \mathbf{A}^{-1}\mathbf{b} \;\Rightarrow\; \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

Huomaa, että käänteismatriisi on olemassa täsmälleen silloin, kun matriisin esittämä lineaarikuvaus on bijektio.

Luvussa 5 tarkastellaan perusteellisemmin sitä minkälaiset matriisit ovat säännöllisiä, eli minkälaiset matriisit määrittävät yhtläryhmän, jolla on yksikäsitteinen ratkaisu. Myöhemmin palataan vielä takaisin lineaarisen yhtälöryhmän ratkaisemiseen.

4.4 Lineaarikuvauksiin liittyvät avaruudet

Tarkastellaan lineaarikuvausta $F : \mathbb{R}^n \mapsto \mathbb{R}^m$. Joukon $\mathbb{R}^n$ kuvaa sanotaan F:n kuva-avaruudeksi ja sille käytetään merkitään $F(\mathbb{R}^n)$. Voimme havaita, että kuva-avaruus on $\mathbb{R}^m$:n aliavaruus. Kuva-avaruuden lisäksi voimme määritellä nolla-avaruuden, eli ne lähtöjoukon alkiot jotka kuvautuvat nollavektoriksi: $\ker(F) = \{\mathbf{x} \in \mathbb{R}^n : F(\mathbf{x}) = \mathbf{0}\}$. Tämä joukko on $\mathbb{R}^n$:nnän aliavaruus.

Jos $\mathbf{A}$ on F:n esitysmatriisi, niin kuva-avaruus on $\mathbf{A}$:n sarakkeiden virittämä avaruus, sillä $\mathbf{Ax} = \sum_i \mathbf{A}^{(i)}x_i$. Siis havaitsemme, että $F(\mathbb{R}^n) = \mathrm{span}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(n)})$. Nolla-avaruus puolestaan käsittää ne vektorit, jotka toteuttavat yhtälöryhmän $\mathbf{Ax} = \mathbf{0}$. Tälle joukolle tulemme käyttämään merkintää $\mathcal{N}(\mathbf{A})$. Huomaa, että $\mathcal{N}(\mathbf{A}) = \ker(F)$, kun $\mathbf{A}$ on F:n esitysmatriisi.

Seuraava tulos on nimeltään lineaarialgebran peruslause ja sen avulla voidaan määritellä kuinka paljon ratkaisuja lineaarisella yhtälöryhmällä on. Palaamme lauseeseen myöhemmin.

[1] Matriisitulon assosiatiivisuuden perusteella tulo voidaan laskea missä järjestyksessä tahansa

Lause 4.4.1. *Olkoon $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ lineaarikuvaus. Tällöin $n = dim(ker(F)) + dim(F(\mathbb{R}^n))$.*

Lineaarialgebran peruslause kertoo, että jos kuvaamme n-dimensioisia vektoreita lineaarikuvauksella, niin kuva-avaruuden dimensio ei voi ylittää n:ää. Intuitiivisesti tämä on mielekästä: emme voi luoda korkeampiulotteista objektia lineaarikuvauksella kuin mikä on lähtöavaruuden dimensio. Lisäksi lause kertoo, että aina, kun kuva-avaruuden dimensio jää alhaisemmaksi kuin lähtöavaruuden, yhtälön $F(\mathbf{x}) = \mathbf{0}$ ratkaisujen joukko on vektoriavaruus, jolla on dimensioita em. avaruuksien dimensioiden erotuksen verran.

Luku 5

Determinantti

5.1 Matriisin determinantti

Matriisin determinantti on neliömatriisin liittyvä luku. Determinantti on tärkeä lineaarialgebrassa koska se kertoo milloin matriisi on säännöllinen ja eräät ratkaisumenetelmät yhtälöryhmille perustuvat determinanttien laskemiseen. Determinantin avulla voidaan myös hakea matriisin ominaisarvot.

5.1.1 Determinantti 2×2-tapauksessa

Tarkastellaan yhtälöparia

$$\begin{cases} 2x + 3y = 3 \\ 4x + ky = -1. \end{cases}$$

Tämän yhtälöryhmän ratkaisujen lukumäärä riippuu siitä, ovatko yhtälöryhmän muodostavat suorat yhdensuuntaiset. Yhtälöryhmän ratkaisujen lukumäärän suhteen on kolme eri mahdollisuutta (huomaa, että ensimmäisen suoran kulmakerroin on $-\frac{2}{3}$ ja toisen suoran kulmakerroin on $-\frac{4}{k}$): (**1**) Suorien kulmakertoimet ovat samat, eli

$$-\frac{2}{3} = -\frac{4}{k}, \quad \text{eli} \quad 2k - 3 \cdot 4 = 0.$$

Tällöin joko (**a**) suorat ovat yhdensuuntaiset mutta eivät samat, jolloin ratkaisuja ei ole yhtään, tai (**b**) suorat ovat samat, jolloin ratkaisuja on ääretön määrä. (**2**) Suorien kulmakertoimet ovat erisuuret, eli

$$2k - 3 \cdot 4 \neq 0.$$

Tällöin yhtälöryhmän muodostavat suorat leikkaavat täsmälleen yhdessä kohdassa, joten yhtälöryhmällä on yksikäsitteinen ratkaisu.

Lukua $2k - 3 \cdot 4$ kutsutaan edellä esitettyyn yhtälöryhmään liittyvän kerroinmatriisin

$$\begin{pmatrix} 2 & 3 \\ 4 & k \end{pmatrix}$$

determinantiksi ja siitä käytetään merkintöjä

$$\begin{vmatrix} 2 & 3 \\ 4 & k \end{vmatrix}$$

ja

$$\det \begin{pmatrix} 2 & 3 \\ 4 & k \end{pmatrix}.$$

Yleisemmin 2×2 -matriisin determinantti on reaaliluku

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

ja lineaarisella yhtälöryhmällä

$$\begin{cases} ax + by = n \\ cx + dy = m \end{cases}$$

on yksikäsitteinen ratkaisu täsmälleen silloin, kun yhtälöryhmän determinantti on $\neq 0$, eli $ad - bc \neq 0$.

Seuraavaksi halutaan luonnollisesti kysyä, miten tämä sääntö voidaan yleistää. Toisin sanottuna: milloin n yhtälöä ja n tuntematonta käsittävällä lineaarisella yhtälöryhmällä on yksikäsitteinen ratkaisu (tällöin kerroinmatriisi on siis $n \times n$ -matriisi). Käy ilmi, ettei edellä esitetyn säännön yleistäminen ole aivan yksinkertaista, joten se ohitetaan tässä yhteydessä. Kannattaa kuitenkin pitää mielessä, että seuraavaksi määriteltävä determinantin käsite liittyy aivan samalla tavalla lineaarisen yhtälöryhmän ratkaisun yksikäsitteisyyteen, kuin edellinen kahden yhtälön tapauskin.

5.1.2 Determinantin laskeminen yleisessä tapauksessa

Neliömatriisin $\mathbf{A}_{n\times n} = (a_{ij})$ alkioon a_{ij} liittyvä *minori* on sellainen $(n-1) \times (n-1)$ -matriisi M_{ij}, joka saadaan kun matriisista $\mathbf{A}$ pyyhitään pois i:s rivi ja j:s sarake. Alkioon a_{ij} liittyvä *alideterminantti*, on alkioon a_{ij} liittyvän minorin determinantti, eli $|\mathrm{M}_{ij}|$. ”Merkillä varustettua” alideterminanttia

$$\mathrm{A}_{ij} = (-1)^{i+j} \, |\mathrm{M}_{ij}|$$

kutsutaan alkion a_{ij} *komplementiksi.* Huomaa, että alideterminanttiin liittyvä merkki määräytyy tässä seuraavan shakkilautakuvion mukaisesti

$$\begin{pmatrix} + & - & + & - \dots \\ - & + & - & + \dots \\ + & - & + & - \dots \\ \dots & \dots & \dots & \dots \dots \end{pmatrix},$$

missä kaikki päälävistäjällä sijaitsevat merkit ovat + merkkejä.

Esimerkki 5.1.1. Tarkastellaan seuraavaa 3×3 -matriisia

$$\mathbf{A} = \begin{pmatrix} 1 & 5 & 0 \\ 8 & 2 & 2 \\ 7 & 3 & 1 \end{pmatrix}.$$

Alkioon a_{23} liittyvä minori on

$$\mathrm{M}_{23} = \begin{pmatrix} 1 & 5 \\ 7 & 3 \end{pmatrix},$$

joten alkion a_{23} komplementti on

$$\mathrm{A}_{23} = (-1)^{2+3} \begin{vmatrix} 1 & 5 \\ 7 & 3 \end{vmatrix} = -1 \cdot (1 \cdot 3 - 5 \cdot 7) = 32.$$

Matriisin **A** determinantilla tarkoitetaan nyt sellaista tulojen summaa, jossa jonkin rivin (sarakkeen) alkiot on kerrottu näitä alkioita vastaavilla komplementeilla. Toisin sanottuna

$$|\mathbf{A}| = a_{i1}\mathrm{A}_{i1} + a_{i2}\mathrm{A}_{i2} + \ldots + a_{in}\mathrm{A}_{in} = \sum_{j=1}^{n} a_{ij}\mathrm{A}_{ij}, \quad \text{kaikilla } i \in \{1, \ldots, n\}$$

ja

$$|\mathbf{A}| = a_{1j}\mathrm{A}_{1j} + a_{2j}\mathrm{A}_{2j} + \ldots + a_{nj}\mathrm{A}_{nj} = \sum_{i=1}^{n} a_{ij}\mathrm{A}_{ij}, \quad \text{kaikilla } j \in \{1, \ldots, n\}.$$

Tässä on siis annettu $2n$ erilaista tapaa laskea $n \times n$ -matriisin determinantti, joista kaikki antavat saman tuloksen. Ensimmäisessä tapauksessa sanotaan, että determinantti on *kehitetty i:nnen vaakarivin mukaan.* Toisessa tapauksessas sanotaan vastaavasti, että determinantti on *kehitetty j:nnen pystyrivin mukaan.*

Tämän menetelmän tarkoituksena on palauttaa determinantin laskeminen 2×2 -matriisien determinanttien laskemiseksi, joka on helppoa. Edellä esitetyissä kehitelmissä, nk *alide-terminanttikehitelmissä*, esiintyvien komplementtien A_{ij} laskemiseksi tarvitsee laskea ainoastaan $(n-1) \times (n-1)$ -matriisien determinantteja, vaikka matriisi **A** on $n \times n$ -matriisi. Menetelmää voidaan kuitenkin jatkaa edelleen, kunnes jäljellä on ainoastaan 2×2 -matriiseja.

Huom. 1×1 matriisi A on skalaari ja tällaisen matriisin determinantti on A.

Esimerkki 5.1.2. Lasketaan matriisin

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

determinantti **a)** kehittämällä se toisen vaakarivin mukaan ja **b)** kehittämällä se kolmannen pystyrivin mukaan. Huomataan, että tulos on sama.

a)

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

$$= 0 - 4 \cdot 14 - 2 \cdot (-5) = -46$$

b)

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

$$= -4 \cdot 4 - 2 \cdot (-5) + 5 \cdot (-8) = -46.$$

3×3 -matriisin determinantti voidaan laskea helposti käyttämällä **Sarruksen sääntönä** tunnettua muistisääntöä. Tämän muistisäännön mukaan matriisin

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

determinantti voidaan laskea kaavalla

$$\begin{array}{ccc:cc} + & + & + & & \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ - & - & - & & \end{array}$$

jolloin

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

Esimerkki 5.1.3. Lasketaan edellisen esimerkin determinantti vielä uudestaan Sarruksen sääntöä käyttämällä. Matriisia

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

vastaava "*laajennettu*" matriisi on nyt

$$\begin{pmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \\ 2 & 3 & -4 \\ 0 & -4 & 2 \end{pmatrix}.$$

Tätä käyttämällä saadaan

$$\begin{aligned} |\mathbf{A}| &= (2\cdot(-4)\cdot 5) + (0\cdot(-1)\cdot(-4)) + (1\cdot 3\cdot 2) - ((-4)\cdot(-4)\cdot 1) - (2\cdot(-1)\cdot 2) - (5\cdot 3\cdot 0) \\ &= -40 + 0 + 6 - 16 + 4 + 0 = -46, \end{aligned}$$

kuten edellisessä esimerkissä.

5.2 Determinantin ominaisuudet

5.2.1 Laskusääntöjä

Determinantilla on mm. seuraavat ominaisuudet.

1) Jos matriisin kaksi riviä (saraketta) vaihdetaan keskenään, sen determinantin merkki vaihtuu. Esimerkiksi

$$\begin{vmatrix} 2 & -6 \\ -1 & 5 \end{vmatrix} = 2\cdot 5 - (-6\cdot -1) = 4 \text{ ja } \begin{vmatrix} -1 & 5 \\ 2 & -6 \end{vmatrix} = (-1\cdot -6) - 5\cdot 2 = -4.$$

2) Jos matriisin jokin rivi (sarake) kerrotaan vakiolla $c \in \mathbb{R}$, determinantti muuttuu c kertaiseksi. Kerrotaan esimerkkinä edellisessä kohdassa esiintyvän matriisin ensimmäinen rivi luvulla 3 ja lasketaan näin saadun matriisin determinantti

$$\begin{vmatrix} 3\cdot 2 & 3\cdot(-6) \\ -1 & 5 \end{vmatrix} = \begin{vmatrix} 6 & -18 \\ -1 & 5 \end{vmatrix} = 6\cdot 5 - (-18\cdot -1) = 30 - 18 = 12 = 3\cdot 4.$$

3) Jos matriisin johonkin riviin (sarakkeeseen) lisätään jokin muu rivi (sarake) vakiolla kerrottuna, determinantin arvo ei muutu.

4) Jos matriisi $\mathbf{A} = (a_{ij})$ on ala- tai yläkolmiomatriisi, sen determinantti on päälävistäjällä olevien alkioiden tulo. Formaalisti ilmaistuna

$$|\mathbf{A}| = a_{11}a_{22}\cdot\ldots\cdot a_{nn}.$$

Tämän säännön erikoistapauksena saadaan, että diagonaalimatriisin determinantti on diagonaalilla sijaitsevien alkioiden tulo, eli

$$|\mathrm{diag}(a_{11},\ldots,a_{nn})| = a_{11}a_{22}\cdot\ldots\cdot a_{nn}.$$

Käyttämällä ominaisuuksia 1-3 mikä tahansa neliömatriisi voidaan muuntaa sellaiseksi ala- tai yläkolmiomatriisiksi, jonka determinantti on sama kuin alkuperäisen matriisin determinantti. Toisaalta ominaisuuden 4 perusteella tällaisen matriisin determinantti on erittäin helppo laskea, mikä tarjoaa meille jälleen uuden tavan laskea matriisin determinantti. Tämän kaltainen menettely on itse asiassa suurien matriisien kohdalla yleensä suositeltavampi käsin laskettaessa.

Esimerkki 5.2.1. Lasketaan matriisin

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

determinantti muuntamalla se ensin ominaisuuksia 1-3 käyttämällä yläkolmiomatriisiksi.

$$\begin{vmatrix} 1 & 2 & 3 & -1 \\ 1 & -2 & 0 & 5 \\ 0 & 1 & -2 & 2 \\ 4 & 0 & -1 & -3 \end{vmatrix} \overset{v1}{=} \begin{vmatrix} 1 & 2 & 3 & -1 \\ 0 & -4 & -3 & 6 \\ 0 & 1 & -2 & 2 \\ 0 & -8 & -13 & 1 \end{vmatrix} \overset{v2}{=} - \begin{vmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & -2 & 2 \\ 0 & -4 & -3 & 6 \\ 0 & -8 & -13 & 1 \end{vmatrix}$$

$$\overset{v3}{=} - \begin{vmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & -11 & 14 \\ 0 & 0 & -29 & 17 \end{vmatrix} \overset{v4}{=} - \begin{vmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & -11 & 14 \\ 0 & 0 & 0 & -19\frac{10}{11} \end{vmatrix} = -(1 \cdot 1 \cdot (-11) \cdot (-19\frac{10}{11}) = -219.$$

(v1) Lisätään ensimmäinen rivi -1:llä kerrottuna toiseen riviin ja tämän jälkeen ensimmäinen rivi -4:llä kerrottuna viimeiseen riviin. Ominaisuuden 3 perusteella determinantin arvo ei tällöin muutu.

(v2) Vaihdetaan toisen ja kolmannen rivin paikkaa, jolloin ominaisuuden 1 perusteella determinantin merkki vaihtuu. Operaation tarkoituksena on helpottaa seuraavaa vaihetta, jossa toisen rivin päälävistäjällä olevan alkion alapuolelle pyritään saamaan aikaiseksi nollia.

(v3) Lisätään toinen rivi 4:llä kerrottuna kolmanteen riviin ja tämän jälkeen toinen rivi 8:lla kerrottuna neljänteen riviin. Jälleen determinantin arvo säilyy ominaisuuden 3 perusteella.

(v4) Lisätään kolmas rivi luvulla $-\frac{29}{11}$ kerrottuna viimeiseen riviin.

(v5) Käytetään ominaisuutta 4.

Tarkastellaan vielä luettelonomaisesti muutamia determinantin ominaisuuksia, jotka on hyvä tietää.

5) Matriisin **A** determinantti on yhtäsuuri kuin sen transpoosin $\mathbf{A}^\top$ determinantti, eli $|\mathbf{A}| = |\mathbf{A}^\top|$.

6) Jos matriisin jokin rivi (sarake) koostuu pelkistä nollista, on sen determinantti 0 (mieti miten käy kun matriisin determinantti kehitetään tämän rivin mukaan).

7) Jos matriisissa on kaksi samaa riviä (saraketta), on sen determinantti 0. Tämä on suora seuraus ominaisuuksista 3 ja 6. Yleisemminkin nähdään, että matriisin determinantti on 0, jos siinä on kaksi samansuuntaista riviä (saraketta).

8) Olkoot $\mathbf{A}$ ja $\mathbf{B}$ kaksi $n \times n$ -matriisia. Tällöin on voimassa

$$| \mathbf{AB} |=| \mathbf{BA} |=| \mathbf{A} | \cdot | \mathbf{B} | \,.$$

Tässä kannattaa huomata, etteivät matriisit $\mathbf{AB}$ ja $\mathbf{BA}$ ole yleensä samat, mutta niiden determinantti on kuitenkin aina sama.

Esimerkki 5.2.2. Tarkastellaan esimerkissä 3.3.2 (ii) annettuja matriiseja $\mathbf{A}$ ja $\mathbf{B}$, joille pätee

$$\mathbf{AB} = \begin{pmatrix} 10 & 5 \\ 13 & 2 \end{pmatrix} \quad \text{ja} \quad \mathbf{BA} = \begin{pmatrix} 4 & 11 \\ 7 & 8 \end{pmatrix}.$$

Nämä matriisit ovat erisuuret, mutta niiden determinantit ovat kuitenkin samat

$$| \mathbf{AB} |= \begin{vmatrix} 10 & 5 \\ 13 & 2 \end{vmatrix} = 10 \cdot 2 - 5 \cdot 13 = 20 - 65 = -45$$

ja

$$| \mathbf{BA} |= \begin{vmatrix} 4 & 11 \\ 7 & 8 \end{vmatrix} = 4 \cdot 8 - 11 \cdot 7 = 32 - 77 = -45.$$

9) $|\mathbf{A}| = \frac{1}{\mathbf{A}^{-1}}$

Esimerkki 5.2.3. Todistetaan harjoituksen vuoksi tämä tulos. Determinanttia koskevan ominaisuuden 8 perusteella saadaan

$$\det(\mathbf{A}) \det\left(\mathbf{A}^{-1}\right) = \det\left(\mathbf{A}\mathbf{A}^{-1}\right) = \det\left(\mathbf{I}_{n\times n}\right) = 1$$

$$\Rightarrow \det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})}.$$

Lopuksi on syytä havaita, että determinantille ei yleisesti päde $|\mathbf{A} + \mathbf{B}| = |\mathbf{A}| + |\mathbf{B}|$.

5.2.2 Determinantti tilavuuden skaalauskertoimena

Determinantin itseisarvo kertoo miten matriisin määrittämä lineaarikuvaus skaalaa ”tilavuutta”. Ajatellaan että kuvataan lineaarikuvauksella n-ulotteinen kuutio. Kun $n = 1$ kuution sijaan kyseessä on jana, kun $n = 2$ kuvataan tason neliö, kun $n = 3$ kuvataan avaruuden kuutio, ja kun $n \geq 4$ kyseessä on kuution yleistys n-dimensioiseen avaruuteen, eli hyperkuutio. Kuten kaksi- ja kolmiuloitteisissa tapauksissa voimme laskea hyperkuutiolle tilavuuden joka on sen sivujen tulo.

Kuvataan siis lineaarikuvauksella $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ hyperkuutio ja pohditaan mitä tapahtuu hyperkuution tilavuudelle. Lineaarikuvaus kääntää ja skaalaa alkuperäistä kuutio siten, että kuva on jonkinlainen hypersuunnikas. Esimerkiksi kaksiuloitteisessa tapauksessa kuution (laatikon) kuva on suunnikas. Osoittautuu, että F:n esitysmatriisin kertoo $\mathbf{A}$ determinantti on yhteydessä siihen mikä on alkuperäisen kuution ja sen kuvan tilavuuden suhde: kuution kuvan tilavuus $= |\det(\mathbf{A})| \times$ alkuperäisen kuution tilavuus. Voit pohtia miten edellisen luvun ominaisuudet on tulkittavissa, kun determinantti kuvaa tilavuuden skaalautumista.

Determinantin perusteella voidaan siis sanoa miten matriisin lineaarikuvaus muuttaa tilavuutta. Erityisesti jos determinantti on nolla, voimme todeta, että n-ulotteinen hyperkuutio kuvautuu objektiksi, jonka tilavuus on nolla. Geometrisesti tämä tarkoittaa sitä, että hyperkuutiosta löytyy vektoreita, jotka kuvautuvat nollavektoreiksi. Huomaa, että jos $F(\mathbf{v}) = \mathbf{0}$ niin myös $F(\alpha\mathbf{v}) = \mathbf{0}$ kaikilla $\alpha \in \mathbb{R}$. Tästä seuraa, että yhtälöryhmällä $\mathbf{Ax} = \mathbf{0}$ ei ole yksikäsitteistä ratkaisua. Näin ollen $\mathbf{A}$:lla ei myöskään voi olla käänteiskuvausta ts. F ei voi olla bijektio. Tarkastelemme myöhemmin lisää determinantin ja käänteismatriisin yhteyttä toisiinsa.

Luku 6

Matriisin aste

6.1 Aste ja lineaarinen riippumattomuus

Palautetaan mieleen, mitä $\mathbb{R}^n$:n vektoreiden $\mathbf{x}_1, \ldots, \mathbf{x}_k$ lineaarisella riippumattomuudella tarkoitetaan ($k < n$). Vektoreiden $\mathbf{x}_1, \ldots, \mathbf{x}_k$ sanotaan olevan *lineaarisesti riippuvia*, jos on olemassa sellaiset reaaliluvut $r_1, \ldots, r_k$, jotka eivät ole kaikki nollia, että

$$r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k = \mathbf{0}.$$[1]

Tällöin jotkut vektorit voidaan esittää toisten lineaarisena yhdistelmänä. Jos taas ehdosta

$$r_1\mathbf{x}_1 + \ldots + r_k\mathbf{x}_k = \mathbf{0}$$

seuraa, että kaikki kertoimet ovat nollia $r_1 = r_2 = \ldots = r_k = 0$, niin tällöin vektoreiden sanotaan olevan *lineaarisesti riippumattomia*. Tällöin mitään vektoreista ei voida esittää muiden lineaarisena yhdistelmänä.

Kuten luvun 3 alussa todettiin, matriisin $\mathbf{A}_{m\times n} = (a_{ij})$ voidaan ajatella koostuvan sen vaakarivien muodostamasta vektorijoukosta $\{\mathbf{A}_1, \ldots, \mathbf{A}_m\}$ tai vaihtoehtoisesti sen pystyrivien (sarakkeiden) muodostamasta vektorijoukosta $\{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(n)}\}$. Voidaan osoittaa, että jokaisella matriisilla lineaarisesti riippumattomien vaakavektoreiden lukumäärä on lineaarisesti riippumattomien pystyvektoreiden lukumäärä.[2] Tätä lukumäärää kutsutaan matriisin $\mathbf{A}$ *asteeksi* (tai rankiksi) ja siitä käytetään merkintää $\mathrm{r}(\mathbf{A})$. Luonnollisesti täytyy olla voimassa $\mathrm{r}(\mathbf{A}) \leq \min\{m, n\}$.

Kaksi avaruuden $\mathbf{R}^n$ vektoria $\mathbf{x} = (x_1, \ldots, x_n)$ ja $\mathbf{y} = (y_1, \ldots, y_n)$ ovat lineaarisesti riippuvia ainoastaan siinä tapauksessa, että ensimmäinen saadaan toisesta kertomalla se

[1] Tässä $\mathbf{0}$ on nollavektori, eli $\mathbf{0}=(\underbrace{0, \ldots, 0}_{k \text{ kpl}})$.

[2] Tässä lukumäärällä tarkoitetaan suurinta mahdollista määrää rivejä (sarakkeita), jotka ovat lineaarisesti riippumattomia.

jollakin vakiolla. Mikäli vektorit $\mathbf{x}$ ja $\mathbf{y}$ ovat lineaarisesti riippuvia, on olemassa sellaiset reaaliluvut $a, b \in \mathbb{R}$, että

$$a\mathbf{x} + b\mathbf{y} = \mathbf{0}.$$

Näin ollen täytyy olla voimassa

$$\mathbf{y} = -\frac{a}{b}\mathbf{x},$$

eli vektorin $\mathbf{y}$ täytyy olla vektorin $\mathbf{x}$ skalaarimonikerta. Tämä on syytä jatkossa pitää mielessä.

Esimerkki 6.1.1. Määritetään matriisien

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 7 & 1 \\ 5 & 6 \end{pmatrix} \quad \text{ja} \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 7 & 14 \\ 4 & 8 \end{pmatrix}$$

asteet. Matriisin $\mathbf{A}$ ensimmäinen ja viimeinen vaakarivi ovat lineaarisesti riippumattomia, koska ensimmäisestä rivistä ei voida saada viimeistä riviä millään vakiolla kerrottaessa (tämän vakion pitäisi olla 2, jotta toisesta komponentista tulisi 6, mutta tällöin ensimmäisestä komponentista tulisi 4). Näin ollen täytyy olla voimassa r($\mathbf{A}$)$\geq$2. Koska täytyy kuitenkin olla voimassa r($\mathbf{A}$)$\leq$min$\{3, 2\}$=2, on matriisin $\mathbf{A}$ asteen oltava 2, eli r($\mathbf{A}$)=2. Matriisin $\mathbf{B}$ pystyrivit ovat lineaarisesti riippuvia, sillä toinen pystyrivi saadaan ensimmäisestä kertomalla se luvulla 2. Koska yksittäinen vektori muodostaa aina triviaalisti lineaarisesti riippumattoman joukon, saadaan r($\mathbf{B}$)=1.

6.2 Asteen ominaisuuksia

Seuraavassa on lueteltu joitakin matriisin asteen tärkeimpiä ominaisuuksia.

Lause 6.2.1. *Olkoot $\boldsymbol{A}$ ja $\boldsymbol{B}$ joitakin $n \times n$ -matriiseja, eli neliömatriiseja.*

1. *Diagonaalimatriisin aste on sen nollasta eriävien alkioiden lukumäärä. Erityisesti $r(\mathbf{I}_{n\times n}) = n$.*
2. *Matriisin ja sen transpoosin aste on sama, eli $r(\mathbf{A}) = r(\mathbf{A}^\top)$.*
3. *$r(\mathbf{AB}) \leq \min\{r(\mathbf{A}), r(\mathbf{B})\}$.*
4. *$r(\mathbf{A}) = n$ jos ja vain jos $|\mathbf{A}| \neq 0$. Tällöin matriisin $\mathbf{A}$ sanotaan olevan* täysiasteinen. *Matriisi $\mathbf{A}$ on siis täysiasteinen, kun sen pystyrivit muodostavat lineaarisesti riippumattoman joukon ja myös sen vaakarivit muodostavat lineaarisesti riippumattoman joukon.*

Matriisin aste on tärkeä käsite mm. siitä syystä, että sillä on suora yhteys matriisin säännöllisyyteen, eli käänteismatriisin olemassaoloon. Voidaan nimittäin osoittaa, että matriisi $\mathbf{A}_{n\times n}$ on säännöllinen täsmälleen silloin, kun r($\mathbf{A}$)=n. Toisin sanottuna matriisi on säännöllinen täsmälleen silloin kun se on täysiasteinen. Matriisin astetta koskevan ominaisuuden 4) perusteella tämä ehto voidaan ilmaista myös hieman toisenlaisessa muodossa.

Lause 6.2.2. $\mathbf{A}^{-1}$ *on olemassa jos ja vain jos* $|\mathbf{A}| \neq 0$.

Aiemmin esitetyn perusteella tämä tarkoittaa sitä, että n yhtälöä ja n tuntematonta käsittävällä lineaarisella yhtälöryhmällä on yksikäsitteinen ratkaisu täsmälleen silloin, kun sen kerroinmatriisin determinantti on $\neq 0$.

Koska matriisin aste kertoo kuinka monta lineaarisesti riippumatonta saraketta matriisissa on, se kertoo matriisin kuvajoukon dimension. Kuvajoukko on matriisin sarakkeiden virittämä aliavaruus. Voimme nyt muotoilla lineaarialgebran peruslauseen 4.4.1 uudelleen.

Lause 6.2.3. *Olkoon* $\mathbf{A} \in \mathbb{R}^{m\times n}$*, tällöin* $n = dim(\mathcal{N}(\mathbf{A})) + r(\mathbf{A})$.

Matriisin aste on edellisen lauseen perusteella yhteydessä siihen miten monta ratkaisua yhtälöryhmällä on. Tarkastellaan siis yhtälöryhmää $\mathbf{Ax} = \mathbf{b}$, missä $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$ ja $\mathbf{b} \in \mathbb{R}^m$. Matriisin asteen perusteella voidaan sanoa ratkaisujen lukumäärästä seuraavaa:

1. jos $\mathrm{r}(\mathbf{A}) = m$, niin yhtälöryhmällä on ratkaisu jokaisella $\mathbf{b} \in \mathbb{R}^m$ ($\mathbf{A}$:n määrittämä lineaarikuvaus on surjektio)
2. jos $\mathrm{r}(\mathbf{A}) < m$, niin yhtälöryhmällä on ratkaisu vain silloin jos $\mathbf{b}$ kuuluu $\mathbf{A}$:n kuvajoukkoon
3. jos $\mathrm{r}(\mathbf{A}) = n$, niin yhtälöryhmällä on enintään yksi ratkaisu jokaisella $\mathbf{b}$ ($\mathbf{A}$:n määrittämä lineaarikuvaus on injektio)
4. jos $\mathrm{r}(\mathbf{A}) < n$, niin ratkaisu ei ole yksikäsitteinen mikäli sellainen on olemassa.

6.2.1 Lineaarinen implisiittifunktiolause

Tarkastellaan seuraavaksi erästä sovellustusta asteelle. Oletetaan, että lineaarisesta yhtälöryhmästä osa muuttujista halutaan saada ratkaistua muiden muuttujien suhteen. Olkoon lineaarinen yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$, jossa on m yhtälöä ja n tuntematonta ja $n > m$. Oletetaan, että $\mathbf{x}$ on järjestetty siten että sen ensimmäiset m komponenttia muodostavat vektorin $\mathbf{y}$ ja loput komponenteista vektorin $\mathbf{z}$. Kirjoitetaan yhtälöryhmä muodossa $\mathbf{By} + \mathbf{Cz} = \mathbf{b}$, missä $\mathbf{B} \in \mathbb{R}^{m\times m}$ ja $\mathbf{C} \in \mathbb{R}^{m\times(n-m)}$, eli $\mathbf{A} = \begin{pmatrix}\mathbf{B} & \mathbf{C}\end{pmatrix}$. Oletetaan, että nyt halutaan ratkaista muuttujat (endogeeniset muuttujat) $\mathbf{y}$ muuttujien (eksogeenisten muuttujien) $\mathbf{z}$ avulla. Seuraavaa tulosta kutsutaan lineaariseksi implisiittifunktiolauseeksi.

Lause 6.2.4. *Jos matriisin* $\mathbf{B}$ *aste on* m *niin silloin* $\mathbf{y}$ *voidaan ratkaista* $\mathbf{z}$*:n suhteen ja tällöin* $\mathbf{y} = -\mathbf{B}^{-1}\mathbf{Cz} + \mathbf{B}^{-1}\mathbf{b}$.

Esimerkki 6.2.1. Tarkastellaan yhtälöryhmää

$$\begin{aligned} x + 2y + z - w &= 1 \\ 3x + 6y - z - 3w &= 2. \end{aligned}$$

Koska yhtälöitä on kaksi, voidaan mitkä tahansa kaksi yhtälössä esiintyvistä muuttujista valita endogeenisiksi. Valitaan x ja y ja tutkitaan voidaanko ne ratkaista muiden muuttujien suhteen. Vektorin (x, y) kerroinmatriisi on

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}.$$

Tämän matriisin determinantti on kuitenkin nolla, joten ratkaisua ei ole. Entä voidaanko (y, z) rakaista muiden funktiona? Tutkimalla determinanttia voidaan havaita, että kyllä.

6.2.2 Esimerkki: portfolion valinta

Tarkastellaan maailmaa, johon liittyy epävarmuutta ja oletetaan, että mahdolliset tulevaisuudennäkymät voidaan jaotella m erilaiseksi maailmantilaksi (state of the world). Kukin maailmantila sisältää kuvauksen kaikesta taloudelliseen päätöksentekoon vaikuttavasta datasta. Maailmantilaa ei etukäteen tiedetä, mutta seuraavalla periodilla realisoituu $s \in S = \{s_1, \ldots, s_m\}$. Konkreettisia esimerkkejä maailmantioloista:

1. palovakuutuksen hankkiminen: $S = \{s_1, s_2\}$, missä $s_1 =$ ”talo palaa”, $s_2 =$ ”talo ei pala”.
2. varutuminen tulevaan taloustilanteeseen: $S = \{s_1, s_2, s_3\}$, missä $s_1 =$ ”syvä lama”, $s_2 =$ ”pieni lama”, $s_3 =$ ”ei lamaa”.
3. sateenvarjo mukaan?: $S = \{s_1, s_2\}$, missä $s_1 =$ ”sataa”, $s_2 =$ ”ei sada”.

Rahoitusinstrumentit tarjoavat mahdollisuuden varautua tuleviin maailmantiloihin. Oletetaan, että mallissa on n rahoitusinstrumenttia ja r_{si} on instrumentin i tuotto maailmantilassa s. Toisin sanoen se on instrumentin maksama tuotto jaettuna instrumentin hinnalla investointihetkellä (ennen maailmantilan paljastumista). Kerätään kaikki rahoitusinstrumenttien tuotot matriisiksi

$$\mathbf{R} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{pmatrix}$$

Merkitään portfoliota vektorilla $\mathbf{x} = (x_1, \ldots, x_n)$. Tällöin portfolion kokonaistuotto saadaan laskettua

$$\mathbf{Rx} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Täydellisillä rahoitusmarkkinoilla instrumentteja voi ostaa ja myydä, joten x_i voi olla positiivinen tai negatiivinen.

Portfolio $\mathbf{x}$ on riskitön, jos sen tuotto on sama kaikissa maailmantiloissa, eli

$$\sum_i r_{ki} x_i = \sum_i r_{li} x_i \ \forall k, l.$$

Voidaan osoittaam että riskitön portfolio on olemassa, jos vektori $\mathbf{1} = (1, \ldots, 1)$ kuuluu matriisin $\mathbf{R}$ kuvajoukkoon, eli $\mathbf{Rx} = \mathbf{1}$. Erityisesti, kun $\mathrm{r}(\mathbf{R}) = m$ riskitön portfolio on olemassa.

Portfolio x on arbitraasiportfolio, jos sille pätee

$$\sum_i x_i = 0.$$

Koska olemme normalisoineet instrumenttien hinnat, tämä tarkoittaa yksinkertaisesti sitä, että portfolion hinta ennen maailmantilan selviämistä on nolla. Arbitraasi (tässä mallissa) tarkoittaa sitä, että on mahdollista saada varmaa tuottoa arbitraasiportfoliolla. Mallissa on siis arbitraasimahdollisuus, jos löytyy arbitraasiportfolio, jolle pätee $\mathbf{Rx} \geq \mathbf{0}$.

Mallilla on tilahinnat $\mathbf{p} = (p_1, \ldots, p_m)$ jos

$$\sum_j p_j r_{ji} = 1 \ \forall i.$$

Tällöin siis hinta p_j voidaan ajatella yhden varallisuusyksikön hintana maailmantilassa j. Tilahinnat ovat olemassa jos löytyy $\mathbf{p}$ siten, että $\mathbf{R}^\top \mathbf{p} = \mathbf{1}$. Tilahinnat ovat olemassa jos $m \geq n$ ja $\mathrm{r}(\mathbf{R}) = n$. Tilahinnat eivät ole yksikäsitteiset, jos $m > n$.

6.3 Asteen laskeminen

Aikaisemmin todettiin, että matriisille voidaan tehdä tietynlaisia rivi ja sarake operaatioita muuttamatta sen determinantin arvoa. Pyrkimyksenä oli muuntaa matriisi näitä operaatioita käyttämällä ala- tai yläkolmiomatriisiksi (tai nk. riviporrasmuotoon), koska näiden determinantti on helppo laskea. Nyt herää kysymys, voitaisiinko jotakin samankaltaista menetelmää käyttää myös matriisin asteen selvittämiseksi. Matriisin aste on nimittäin sitä helpompi selvittää, mitä enemmän siinä on nolla-alkioita. Erityisesti ala- tai yläkolmiomatriisin aste on siinä olevien portaiden lukumäärä. Toisin sanottuna, kun matriisi on saatettu yläkolmiomatriisiksi (riviporrasmuotoon), niin sen aste saadaan vähentämällä kaikkien rivien lukumäärästä niiden rivien määrä, joissa kaikki alkiot ovat nollia.

Seuraavat operaatiot säilyttävät matriisin asteen.

<u>**Matriisissa voidaan sen astetta muuttamatta:**</u>

(a) Vaihtaa kahden rivin tai kahden sarakkeen paikkaa.

(b) Kertoa mikä tahansa rivi tai sarake nollasta eroavalla vakiolla.

(c) Lisätä mihin tahansa riviin (sarakkeeseen) jokin toinen rivi

(sarake) vakiolla kerrottuna.

Nämä säännöt perustuvat yksinkertaisesti siihen, etteivät kyseiset operaatiot muuta lineaarisesti riippumattomien rivien (sarakkeiden) lukumäärää.

Esimerkki 6.3.1. Ratkaistaan matriisin

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 & 2 \\ -5 & 10 & 8 & -3 \\ -2 & 3 & 5 & -1 \\ 0 & 4 & 1 & 1 \end{pmatrix}$$

aste muuntamalla se yläkolmiomatriisiksi. Seuraavassa merkinnällä $\mathbf{A} \sim \mathbf{B}$ tarkoitetaan sitä, että matriisi $\mathbf{B}$ on saatu matriisista $\mathbf{A}$ soveltamalla siihen jotakin säännöistä a)–c). Toisin sanottuna merkinnällä tarkoitetaan sitä, että matriisin $\mathbf{B}$ aste on sama kuin matriisin **A** aste.

$$\begin{pmatrix} 1 & 0 & 3 & 2 \\ -5 & 10 & 8 & -3 \\ -2 & 3 & 5 & -1 \\ 0 & 4 & 1 & 1 \end{pmatrix} \overset{v1}{\sim} \begin{pmatrix} 1 & 0 & 3 & 2 \\ 0 & 10 & 23 & 7 \\ 0 & 3 & 11 & 3 \\ 0 & 4 & 1 & 1 \end{pmatrix} \overset{v2}{\sim} \begin{pmatrix} 1 & -2 & 3 & 2 \\ 0 & 3 & 23 & 7 \\ 0 & 0 & 11 & 3 \\ 0 & 3 & 1 & 1 \end{pmatrix}$$

$$\overset{v3}{\sim} \begin{pmatrix} 1 & -2 & 3 & 2 \\ 0 & 3 & 23 & 7 \\ 0 & 0 & 11 & 3 \\ 0 & 0 & -22 & -6 \end{pmatrix} \overset{v4}{\sim} \begin{pmatrix} 1 & -2 & 3 & 2 \\ 0 & 3 & 23 & 7 \\ 0 & 0 & 11 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(v1) Lisätään ensimmäinen rivi 5:llä kerrottuna toiseen riviin ja sen jälkeen ensimmäinen rivi 2:lla kerrotuna kolmanteen riviin.

(v2) Lisätään neljäs sarake -1:llä kerrottuna toiseen sarakkeeseen.

(v3) Lisätään toinen rivi -1:llä kerrottuna viimeiseen riviin.

(v4) Lisätään kolmas rivi 2:lla kerrottuna viimeiseen riviin.

Koska yläkolmiomatriiisin aste on siinä olevien portaiden lukumäärä, matriisin **A** aste on 3, eli r(**A**)=3.

Luku 7

Käänteismatriisin laskeminen

Käänteismatriiseja tarvitaan lineaaristen yhtälöryhmien ratkaisemiseen tai esimerkiksi silloin, kun halutaan ratkaista lineaarisesta yhtälöryhmästä osa muuttujista toisten muuttujien avulla lausuttuna. Edellä havaitsimme, että $n \times n$ matriisi on kääntyvä täsmälleen silloin, kun se on täyttä astetta. Tässä vaiheessa ei ole vielä otettu kantaa siihen miten käänteismatriisi voidaan löytää käytännössä. Seuraavaksi käydään läpi kolme käyttökelpoista tapaa. Se mikä näistä menetelmistä on suositeltavin, riippuu tilanteesta. Kaikki antavat kuitenkin luonnollisesti saman tuloksen.

7.1 Menetelmä 1: suora eliminointi

Ratkaistaan matriisin $\mathbf{A}$ käänteismatriisi $\mathbf{A}^{-1}$ yhtälöstä $\mathbf{AB}$=$\mathbf{I}$, missä matriisin $\mathbf{B}$ alkiot b_{ij} ovat tuntemattomia:

$$\mathbf{AB} = \mathbf{I} \quad \Leftrightarrow \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Tämä on lineaarinen yhtälöryhmä, jossa on n^2 yhtälöä ja n^2 tuntematonta. Kaikki menetelmät käänteismatriisin ratkaisemiseksi perustuvat oleellisesti tämän yhtälöryhmän ratkaisemiseen.

Suoralla eliminointimenetelmällä (tai suoralla sijoittamisella) tarkoitetaan sitä että yhtälöstä eliminoidaan muuttujia ja sijoitetaan ne jäljellä oleviin yhtälöihin ja jatketaan kunnes on ratkaistavana yksi yhtälö yhdelle tuntemattomalle. Tämän jälkeen ratkaistaan loput tuntemattomista yksi kerrallaan sijoittamalla aikaisemmin ratkaistut muuttujat tuntemattomien muuttujien lausekkeisiin. Alla oleva esimerkki hahmottaa mistä on kyse. Suora eliminointi toimii silloin, kun tuntemattomia on vain vähän: kaksi tai enintään kolme. Yleisemmin menettely on työläs ja virhealtis käsin laskettaessa.

Esimerkki 7.1.1. Olkoon

$$\mathbf{A} = \begin{pmatrix} 4 & 1 \\ -5 & 7 \end{pmatrix}.$$

Tämä matriisi on säännöllinen, koska $|\,\mathbf{A}\,| = 4 \cdot 7 - (1 \cdot -5) = 28 + 5 = 33 \neq 0$. Käänteismatriisi on siis olemassa ja se löydetään ratkaisemalla yhtälöryhmä

$$\begin{pmatrix} 4 & 1 \\ -5 & 7 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

eli

$$\begin{cases} 4b_{11} + b_{21} = 1 \\ 4b_{12} + b_{22} = 0 \\ -5b_{11} + 7b_{21} = 0 \\ -5b_{12} + 7b_{22} = 1 \end{cases}.$$

Kerrotaan toinen yhtälö -7:llä ja lisätään se tämän jälkeen viimeiseen yhtälöön, jolloin saadaan $-33b_{12} = 1$, eli $b_{12} = -\frac{1}{33}$. Sijoittamalla tämä toiseen yhtälöön, saadaan $b_{22} = \frac{4}{33}$. Kun tehdään aivan samalla tavalla ensimmäiselle ja kolmannelle yhtälölle, saadaan $-33b_{11} = -7$, eli $b_{11} = \frac{7}{33}$. Sijoittamalla ensimmäiseen yhtälöön saadaan ratkaistua viimeinenkin tuntematon $b_{21} = 1 - 4 \cdot \frac{7}{33} = \frac{5}{33}$. Näin ollaan siis saatu tulos

$$\mathbf{A}^{-1} = \mathbf{B} = \begin{pmatrix} \frac{7}{33} & -\frac{1}{33} \\ \frac{5}{33} & \frac{4}{33} \end{pmatrix}.$$

7.2 Menetelmä 2: Cramerin sääntö käänteismatriisille

Matriisin $\mathbf{A}_{n\times n} = (a_{ij})$ *adjungoidulla* matriisilla, josta käytetään jatkossa merkintää $\mathrm{Adj}(\mathbf{A})$, tarkoitetaan matriisia

$$\mathrm{Adj}(\mathbf{A}) = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^{\top} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},$$

missä A_{ij} on alkioon a_{ij} liittyvä komplementti, eli

$$\mathrm{A}_{ij} = (-1)^{i+j}\,|\,\mathbf{M}_{ij}\,|\,.$$

Useimmiten käänteismatriisi lasketaan kaavalla

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \cdot \mathrm{Adj}(\mathbf{A}) = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^{\top},$$

jota ei tässä yhteydessä sen paremmin perustella. Kaava tunnetaan Cramerin sääntönä käänteismatriisin laskemiseksi.

Esimerkki 7.2.1. Lasketaan aikaisemmassa esimerkissä käsitellyn matriisin

$$\mathbf{A} = \begin{pmatrix} 4 & 1 \\ -5 & 7 \end{pmatrix}$$

käänteismatriisi tämän kaavan avulla. Huomaa, että 2×2 -matriisiin liittyvät minorit ovat yksinkertaisesti lukuja. Tässä tapauksessa saadaan

$$\mathrm{A}_{11} = (-1)^{1+1} \mid \mathrm{M}_{11} \mid = 1 \cdot 7 = 7, \;\; \mathrm{A}_{12} = (-1)^{1+2} \mid \mathrm{M}_{12} \mid = -1 \cdot -5 = 5,$$
$$\mathrm{A}_{21} = (-1)^{2+1} \mid \mathrm{M}_{21} \mid = -1 \cdot 1 = -1, \;\; \mathrm{A}_{22} = (-1)^{2+2} \mid \mathrm{M}_{22} \mid = 1 \cdot 4 = 4.$$

Koska edellisen esimerkin perusteella tiedetään, että det($\mathbf{A}$) =33, saadaan käänteismatriisiksi

$$\mathbf{A}^{-1} = \frac{1}{33} \begin{pmatrix} 7 & 5 \\ -1 & 4 \end{pmatrix}^{\top} = \frac{1}{33} \begin{pmatrix} 7 & -1 \\ 5 & 4 \end{pmatrix} = \begin{pmatrix} \frac{7}{33} & -\frac{1}{33} \\ \frac{5}{33} & \frac{4}{33} \end{pmatrix}.$$

Tämä on sama matriisi kuin esimerkissä 7.1.1, kuten pitääkin.

Tarkastellaan edellä esitettyä käänteismatriisin kaavaa 2×2 -matriisien tapauksessa. Oletetaan, että

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ja det($\mathbf{A}$)=$ad - bc \neq 0$. Tällöin matriisi $\mathbf{A}$ on siis kääntyvä. Aivan samalla tavalla kuin esimerkissä 7.2.1, nyt saadaan

$$\mathrm{A}_{11} = (-1)^{1+1} \mid \mathrm{M}_{11} \mid = 1 \cdot d = d, \;\; \mathrm{A}_{12} = (-1)^{1+2} \mid \mathrm{M}_{12} \mid = -1 \cdot c = -c,$$
$$\mathrm{A}_{21} = (-1)^{2+1} \mid \mathrm{M}_{21} \mid = -1 \cdot b = -b, \;\; \mathrm{A}_{22} = (-1)^{2+2} \mid \mathrm{M}_{22} \mid = 1 \cdot a = a.$$

Matriisin $\mathbf{A}$ käänteismatriisi on siis

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{\top} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^{\top} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Tämä kaava on syytä opetella jopa ulkoa.

Esimerkki 7.2.2. Muodostetaan vielä matriisin

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

käänteismatriisi. Lasketaan aluksi tämän matriisin determinantti kehittämällä se toisen vaakarivin mukaan, jolloin saadaan

$$\det(\mathbf{A}) = 0 \cdot \mathrm{A}_{21} + 3 \cdot \mathrm{A}_{22} + 0 \cdot \mathrm{A}_{23} = 3 \cdot (-1)^{2+2} \mid \mathbf{M}_{22} \mid = 3 \begin{vmatrix} 2 & 5 \\ 1 & 1 \end{vmatrix} = 3 \cdot -3 = -9.$$

Muodostetaan seuraavaksi komplementit A_{ij} käyttämällä hyväksi sivulla 13 esitettyä merkkitaulukkoa

$$\mathrm{A}_{11} = + \begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} = 3,\ \mathrm{A}_{12} = - \begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} = 0,\ \mathrm{A}_{13} = + \begin{vmatrix} 0 & 3 \\ 1 & 0 \end{vmatrix} = -3,$$
$$\mathrm{A}_{21} = - \begin{vmatrix} 4 & 5 \\ 0 & 1 \end{vmatrix} = -4,\ \mathrm{A}_{22} = + \begin{vmatrix} 2 & 5 \\ 1 & 1 \end{vmatrix} = -3,\ \mathrm{A}_{23} = - \begin{vmatrix} 2 & 4 \\ 1 & 0 \end{vmatrix} = 4,$$
$$\mathrm{A}_{31} = + \begin{vmatrix} 4 & 5 \\ 3 & 0 \end{vmatrix} = -15,\quad \mathrm{A}_{32} = - \begin{vmatrix} 2 & 5 \\ 0 & 0 \end{vmatrix} = 0,\quad \mathrm{A}_{33} = + \begin{vmatrix} 2 & 4 \\ 0 & 3 \end{vmatrix} = 6.$$

Näitä käyttämällä saadaan

$$\mathbf{A}^{-1} = -\frac{1}{9} \begin{pmatrix} 3 & 0 & -3 \\ -4 & -3 & 4 \\ -15 & 0 & 6 \end{pmatrix}^{\top} = -\frac{1}{9} \begin{pmatrix} 3 & -4 & -15 \\ 0 & -3 & 0 \\ -3 & 4 & 6 \end{pmatrix}.$$

7.3 Menetelmä 3: Gaussin-Jordanin eliminointimenettely

Matriisin **A** käänteismatriisi voidaan ratkaista myös seuraavalla tavalla. Muodostetaan aluksi matriisi

$$(\mathbf{A} \vdots \mathbf{I}_{n\times n}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & \vdots & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & \vdots & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & \vdots & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Mikäli **A** on säännöllinen, tämä laajennettu matriisi voidaan seuraavia vaakarivioperaatioita (nk. elementaarisia riviperaatioita) käyttäen muuntaa muotooon $(\mathbf{I}_{n\times n} \vdots \mathbf{B})$:

(1) vaakarivin kertominen millä tahansa vakiolla,

(2) jonkin vaakarivin lisääminen vakiolla kerrottuna johonkin toiseen vaakariviin,

(3) vaakarivien vaihto.

Näin saatu matriisi **B** on matriisin **A** käänteismatriisi, eli $\mathbf{B} = \mathbf{A}^{-1}$. Tämä menetelmä perustuu Gaussin-Jordanin eliminointimenettelyyn, jota käsitellään vasta myöhemmin.

Lukijaa kehotetaan palaamaan tähän menetelmään uudestaan, kun Gaussin-Jordanin eliminointimenettely on käsitelty. Pohjimmiltaan tässä ratkaistaan joukko lineaarisia yhtälöryhmiä simultaanisesti.

Esimerkki 7.3.1. Ratkaistaan esimerkissä 7.2.2 tarkastellun matriisin käänteismatriisi uudestaan. Merkinnällä $\sim$ tarkoitetaan tässä sitä, että matriisit saadaan toisistaan käyttämällä operaatioita 1–3.

$$\begin{pmatrix} 2 & 4 & 5 & \vdots & 1 & 0 & 0 \\ 0 & 3 & 0 & \vdots & 0 & 1 & 0 \\ 1 & 0 & 1 & \vdots & 0 & 0 & 1 \end{pmatrix} \overset{v1}{\sim} \begin{pmatrix} 0 & 4 & 3 & \vdots & 1 & 0 & -2 \\ 0 & 3 & 0 & \vdots & 0 & 1 & 0 \\ 1 & 0 & 1 & \vdots & 0 & 0 & 1 \end{pmatrix} \overset{v2}{\sim} \begin{pmatrix} 1 & 0 & 1 & \vdots & 0 & 0 & 1 \\ 0 & 3 & 0 & \vdots & 0 & 1 & 0 \\ 0 & 4 & 3 & \vdots & 1 & 0 & -2 \end{pmatrix}$$

$$\overset{v3}{\sim} \begin{pmatrix} 1 & 0 & 1 & \vdots & 0 & 0 & 1 \\ 0 & 3 & 0 & \vdots & 0 & 1 & 0 \\ 0 & 0 & 3 & \vdots & 1 & -\frac{4}{3} & -2 \end{pmatrix} \overset{v4}{\sim} \begin{pmatrix} 1 & 0 & 1 & \vdots & 0 & 0 & 1 \\ 0 & 1 & 0 & \vdots & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & \vdots & \frac{1}{3} & -\frac{4}{9} & -\frac{2}{3} \end{pmatrix}$$

$$\overset{v5}{\sim} \begin{pmatrix} 1 & 0 & 0 & \vdots & -\frac{1}{3} & \frac{4}{9} & \frac{5}{3} \\ 0 & 1 & 0 & \vdots & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & \vdots & \frac{1}{3} & -\frac{4}{9} & -\frac{2}{3} \end{pmatrix}.$$

(v1) Lisätään viimeinen rivi -2:lla kerrottuna ensimmäiseen riviin.

(v2) Vaihdetaan viimeisen ja ensimmäisen rivin paikkaa.

(v3) Lisätään toinen rivi kolmanteen $-\frac{4}{3}$:lla kerrottuna.

(v4) Kerrotaan toinen ja kolmas rivi luvulla $\frac{1}{3}$.

(v5) Lisätään viimeinen rivi ensimmäiseen -1:llä kerrottuna.

Näin ollen

$$\mathbf{A}^{-1} = \mathbf{B} = \begin{pmatrix} -\frac{1}{3} & \frac{4}{9} & \frac{5}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & -\frac{4}{9} & -\frac{2}{3} \end{pmatrix},$$

mikä on sama matriisi kuin esimerkissä 7.2.2.

Luku 8

Lineaarisen yhtälöryhmän ratkaiseminen

Seuraavaksi esitetään kolme mahdollista tapaa ratkaista lineaarinen yhtälöryhmä. Kuten aiemmin on todettu, lineaarisella yhtälöryhmällä $\mathbf{Ax} = \mathbf{b}$ on yksikäsitteinen ratkaisu täsmälleen silloin, kun kerroinmatriisi $\mathbf{A}$ on säännöllinen. Tällöin yhtälöryhmä voidaan ratkaista kertomalla se puolittain matriisilla $\mathbf{A}^{-1}$, jolloin saadaan $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Luonnollisesti eräs tapa ratkaista yhtälöryhmä on ratkaista sen kerroinmatriisin käänteismatriisi ja sen jälkeen kertoa oikea puoli saadulla matriisilla. Käänteismatriisin laskeminen ei kuitenkaan ole aivan välttämätöntä. Erityisesti silloin, kun käänteismatriisia ei aivan välttämättä tarvita, selvitään pienemmällä määrällä laskutoimituksia, jos se jätetään ratkaisematta ja käytetään sen sijaan jotain tässä luvussa esitetyistä menetelmistä. Huomautettakoon vielä, että koska käänteismatriisin laskeminen oleellisesti edellyttää yhtälöryhmän ratkaisemista, tämän luvun menetelmän luonnollisesti soveltuvat myös käänteismatriisien laskentaan. Itse asiassa edellisessä luvussa esitetyt menetelmät käänteismatriisin laskemiseksi ovat sovelluksia menetelmistä lineaaristen yhtälöryhmien ratkaisemiseksi.

Tarkastellaan aluksi paria esimerkkiä siitä miten yhtälöryhmä ratkaistaan, kun sen kerroinmatriisin käänteismatriisi tunnetaan.

Esimerkki 8.0.2. Ratkaistaan yhtälöryhmä

$$\begin{cases} 2x + 4y + 5z = 1 \\ 3y = -3 \\ x + z = 5 \end{cases},$$

eli

$$\begin{pmatrix} 2 & 4 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix}.$$

Tämän yhtälöryhmän kerroinmatriisin käänteismatriisi laskettiin esimerkissä 7.2.2, joten ratkaisuksi saadaan

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -\frac{1}{9}\begin{pmatrix} 3 & -4 & -15 \\ 0 & -3 & 0 \\ -3 & 4 & 6 \end{pmatrix}\begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} = -\frac{1}{9}\begin{pmatrix} -60 \\ 9 \\ 15 \end{pmatrix} = \begin{pmatrix} \frac{60}{9} \\ -1 \\ -\frac{15}{9} \end{pmatrix}.$$

Vastaus kannattaa tietenkin tarkistaa sijoittamalla.

Esimerkki 8.0.3. Ratkaistaan yhtälöryhmä

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

eli

$$\begin{pmatrix} 5 & -2 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \end{pmatrix}.$$

Tähän yhtälöryhmään liittyvän kerroinmatriisin $\mathbf{A}$ determinantti $\det(\mathbf{A}) = 5 \cdot 3 - (-2 \cdot -3) = 15 - 6 = 9 \neq 0$, joten ratkaisu on yksikäsitteinen. 2×2-matriiseille johdettua kaavaa käyttämällä kerroinmatriisin käänteismatriisiksi saadaan

$$\mathbf{A}^{-1} = \frac{1}{9}\begin{pmatrix} 3 & 2 \\ 3 & 5 \end{pmatrix}$$

ja ratkaisuksi edelleen

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 3 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} 7 \\ -1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 19 \\ 16 \end{pmatrix} = \begin{pmatrix} \frac{19}{9} \\ \frac{16}{9} \end{pmatrix}.$$

8.1 Cramerin sääntö

Koska matriisin $\mathbf{A}$ determinantti kertoo onko $\mathbf{A}^{-1}$ olemassa ja näin ollen onko yhtälöryhmällä $\mathbf{Ax}=\mathbf{b}$ yksikäsitteinen ratkaisu, ei liene yllättävää, että lineaarinen yhtälöryhmä voidaan ratkaista käyttämällä pelkästään determinantteja. Tällaisen menetelmän suurena etuna on se, ettei käänteismatriisia tarvitse ratkaista. Kuten ollaan huomattu, käänteismatriisin laskeminen on huomattavasti hankalampaa kuin determinantin laskeminen.

Lause 8.1.1. ***(Cramerin sääntö.)*** *Mikäli lineaarisen yhtälöryhmän $\boldsymbol{Ax}=\boldsymbol{b}$ kerroinmatriisi $\boldsymbol{A}$ on säännöllinen, tämän yhtälöryhmän yksikäsitteinen ratkaisu $\boldsymbol{x}=(x_1, \ldots, x_n)$ saadaan ehdoista*

$$x_i = \frac{det(\mathbf{B}_i)}{det(\mathbf{A})}, \quad \textit{kaikilla } i = 1, \ldots, n,$$

missä matriisi $\mathbf{B}_i$ on saatu matriisista $\boldsymbol{A}$ korvaamalla i:s sarake vektorilla $\boldsymbol{b}$.

Cramerin sääntö voidaan todistaa Cramerin säännöstä käänteismatriisille, joka puolestaan perustuu determinantin ominaisuuksiin. Cramerin säännön etu on, että se tarjoaa analyyttisen kaavan lineaarisen yhtälöryhmän ratkaisulle.

Havainnollistetaan Cramerin sääntö kolmen yhtälön tapauksessa

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

Cramerin sääntö antaa

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

Tämä pätee aina kun kerroinmatriisi on säännöllinen, eli

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0.$$

Esimerkki 8.1.1. Ratkaistaan esimerkin 8.0.2 yhtälöryhmä

$$\begin{pmatrix} 2 & 4 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix}.$$

Cramerin sääntöä käyttämällä. Esimerkin 7.2.2 perusteella tiedetään, että

$$\begin{vmatrix} 2 & 4 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 1 \end{vmatrix} = -9.$$

Tämän lisäksi tarvitaan vielä seuraavat determinantit

$$\det(\mathbf{B}_1) = \begin{vmatrix} 1 & 4 & 5 \\ -3 & 3 & 0 \\ 5 & 0 & 1 \end{vmatrix} = 5 \cdot \begin{vmatrix} 4 & 5 \\ 3 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 4 \\ -3 & 3 \end{vmatrix} = 5 \cdot -15 + 1 \cdot 15 = -60,$$

$$\det(\mathbf{B}_2) = \begin{vmatrix} 2 & 1 & 5 \\ 0 & -3 & 0 \\ 1 & 5 & 1 \end{vmatrix} = -3 \cdot \begin{vmatrix} 2 & 5 \\ 1 & 1 \end{vmatrix} = -3 \cdot -3 = 9,$$

$$\det(\mathbf{B}_3) = \begin{vmatrix} 2 & 4 & 1 \\ 0 & 3 & -3 \\ 1 & 0 & 5 \end{vmatrix} = 1 \cdot \begin{vmatrix} 4 & 1 \\ 3 & -3 \end{vmatrix} + 5 \cdot \begin{vmatrix} 2 & 4 \\ 0 & 3 \end{vmatrix} = 1 \cdot -15 + 5 \cdot 6 = 15.$$

Ratkaisuksi saadaan

$$x_1 = \frac{60}{9}, \quad x_2 = -1, \quad x_3 = -\frac{15}{9},$$

aivan kuten aikaisemminkin.

8.2 Gaussin ja Gaussin-Jordanin eliminointimenettelyt

Gaussin ja Gaussin-Jordanin eliminointimenetelmiä voidaan soveltaa myös sellaisissa tilanteissa, joissa kerroinmatriisilla $\mathbf{A}$ ei ole käänteismatriisia. [1] Näitä menetelmiä voidaan lisäksi soveltaa myös silloin, kun yhtälöryhmän yhtälöiden ja tuntemattomien lukumäärä ei ole sama.

Tarkastellaan nyt yhtälöryhmää

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ \dots \ + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ \dots \ + a_{2n}x_n = b_2 \\ \quad \vdots \qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

missä kerroinmatriisin

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

ei välttämättä tarvitse olla neliömatriisi (mutta se voi olla). Tähän yhtälöryhmään voidaan soveltaa seuraavia muunnoksia sen ratkaisun muuttumatta:

a) yhtälöiden järjestyksen vaihtaminen,

b) yhden tai useamman yhtälön kertominen vakiolla joka $\neq 0$,

c) yhden tai useamman yhtälön lisääminen johonkin toiseen yhtälöön

vakiolla kerrottuna.

Gaussin ja Gaussin-Jordanin eliminointimenetelmissä tarkastellaan ns. *täydennettyä kerroinmatriisia*

[1] Taloustieteessä tällainen tilanne ei yleensä ole mielenkiintoinen, koska se tarkoittaa jonkinlaista determinoitumatonta tilannetta. Jos yhtälöryhmä kuvaa esimerkiksi markkinatasapainoa, tällainen tilanne tarkoittaa sitä, ettei tasapainoa ole olemassa tai sitten tasapainoja on monia. Jälkimmäisessä tapauksessa odotukset yleensä määräävät sen mihin tasapainoon lopulta päädytään.

$$(\mathbf{A} \mid \mathbf{b}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Edellä esitettyyn yhtälöryhmään kohdistuvia muunnoksia a)–c) vastaa täydennettyyn kerroinmatriisiin kohdistuvat vastaavat muunnokset. Tällöin täytyy kuitenkin pitää mielessä, etteivät sarakemuunnokset ole sallittuja, toisin kuin astetta selvitettäessä. Yhtälöryhmää ratkaistaessa sarakemuunnos vastaisi eri muuttujien lisäämistä toisiinsa, mikä ei tietenkään ole sallittua.

Gaussin eliminointimenetelmässä tarkoituksena on muuntaa täydennetty kerroinmatriisi riviporrasmuotoon, eli muotoon jossa täydennety matriisi on

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} & \vdots & c_1 \\ 0 & b_{22} & \cdots & b_{2n} & \vdots & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & b_{nn} & \vdots & c_n \end{pmatrix}.$$

Huomaa, että tällöin viimeinen rivi vastaa yhtälöä $b_{nn}x_n = c_n$, josta voidaan ratkaista x_n. Kun täydennetty kerroinmatriisi on saatu rivipporrasmuotoon voidaan kaikki tuntemattomat ratkaista takaisin sijoittamalla (*back subsitution*), jolloin ensiksi ratkaistaan x_n kuten edellä ja sijoitetaan se toiseksi viimeiseen yhtälöön, josta ratkaistaan x_{n-1} jne.

Gaussin-Jordanin menetelmä on sama kuin Gaussin eliminaatio siihen asti kun laajennettu matriisi on saatu riviporrasmuotoon. Gaussin-Jordanin menetelmässä tarkoituksena on muuntaa täydennetty kerroinmatriisi edelleen muotoon (<u>jos **A** on säännöllinen</u>)

$$\begin{pmatrix} 1 & 0 & \dots & 0 & \vdots & \\ 0 & 1 & \dots & 0 & \vdots & \mathbf{A}^{-1}\mathbf{b} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & \dots & 1 & \vdots & \end{pmatrix},$$

josta nähdään yhtälöryhmän yksikäsitteinen ratkaisu $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Tai sitten muotoon (<u>jos **A** ei ole säännöllinen tai $m \neq n$</u>)

$$\begin{pmatrix} 1 & 0 & \dots & 0 & \vdots & \\ 0 & 1 & \dots & 0 & \vdots & \mathbf{c} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & \dots & 0 & \vdots & \end{pmatrix},$$

josta nähdään onko ratkaisuja ääretön määrä vai ei yhtään. Jos ratkaisuja on ääretön määrä, ratkaisujoukko voidaan joissakin tapauksissa pystyä parametrisoimaan. Esimerkkitapaukset valaisevat asiaa parhaiten.

Esimerkki 8.2.1. Ratkaistaan jo tutuksi tullut esimerkissä 8.1.1 tarkasteltu yhtälöryhmä vielä Gaussin ja Gaussin-Jordanin eliminointimenettelyllä. Tähän yhtälöryhmään liittyvä täydennetty kerroinmatriisi on

$$\begin{pmatrix} 2 & 4 & 5 & \vdots & 1 \\ 0 & 3 & 0 & \vdots & -3 \\ 1 & 0 & 1 & \vdots & 5 \end{pmatrix}.$$

Seuraavassa merkinnällä $\sim$ tarkoitetaan sitä, että täydennettyjä kerroinmatriiseja vastaavilla yhtälöryhmillä on sama ratkaisujoukko. Gaussin eliminoinnissa matriisiin ei suoriteta enempää muunnoksia, kun se on saatettu alla olevan toisen rivin ensimmäisenä olevaan muotoon. Käydään ensin läpi miten Gaussin-Jordanin iteraatio etenee loppuun asti ja palataan sitten Gaussin eliminointiin.

$$\begin{pmatrix} 2 & 4 & 5 & \vdots & 1 \\ 0 & 3 & 0 & \vdots & -3 \\ 1 & 0 & 1 & \vdots & 5 \end{pmatrix} \overset{v1}{\sim} \begin{pmatrix} 1 & 0 & 1 & \vdots & 5 \\ 0 & 3 & 0 & \vdots & -3 \\ 2 & 4 & 5 & \vdots & 1 \end{pmatrix} \overset{v2}{\sim} \begin{pmatrix} 1 & 0 & 1 & \vdots & 5 \\ 0 & 3 & 0 & \vdots & -3 \\ 0 & 4 & 3 & \vdots & -9 \end{pmatrix}$$

$$\overset{v3}{\sim} \begin{pmatrix} 1 & 0 & 1 & \vdots & 5 \\ 0 & 3 & 0 & \vdots & -3 \\ 0 & 0 & 3 & \vdots & -5 \end{pmatrix} \overset{v4}{\sim} \begin{pmatrix} 1 & 0 & 0 & \vdots & \frac{20}{3} \\ 0 & 3 & 0 & \vdots & -3 \\ 0 & 0 & 3 & \vdots & -5 \end{pmatrix} \overset{v5}{\sim} \begin{pmatrix} 1 & 0 & 0 & \vdots & \frac{20}{3} \\ 0 & 1 & 0 & \vdots & -1 \\ 0 & 0 & 1 & \vdots & -\frac{5}{3} \end{pmatrix}.$$

(v1) Vaihdetaan ensimmäisen ja viimeisen rivin paikkaa. (Tämä ei suinkaan ole välttämätöntä)

(v2) Lisätään ensimmäinen rivi kolmanteen -2:lla kerrottuna.

(v3) Lisätään toinen rivi kolmanteen $-\frac{4}{3}$:lla kerrottuna.

(v4) Lisätään kolmas rivi ensimmäiseen $-\frac{1}{3}$:lla kerrottuna.

(v5) Kerrotaan toinen ja kolmas rivi $\frac{1}{3}$:lla.

Ratkaisu on sama kuin edellisissäkin tapauksissa. [2]

[2] Tässä vaiheessa skeptisinkin voi todeta, että ratkaisu on varmasti oikein.

Palataan nyt muotoon

$$\begin{pmatrix} 1 & 0 & 1 & \vdots & 5 \\ 0 & 3 & 0 & \vdots & -3 \\ 0 & 0 & 3 & \vdots & -5 \end{pmatrix}$$

ja jatketaan Gaussin eliminointia tästä. Viimeinen rivi vastaa yhtälöä $3x_3 = -5$, josta saadaan $x_3 = -5/3$. Tämän jälkeen ratkaistaan x_2 yhtälöstä $3x_2 + 0x_3 = -3$. Periaatteessa x_3 tulisi sijoittaa yhtälöön, mutta nyt niin ei tarvitse tehdä, koska x_3:n kerroin on nolla. Saadaan $x_2 = -1$. Lopuksi ratkaistaan x_1 yhtälöstä $x_1 + 0x_2 + x_3 = 5$ kun x_2 ja x_3 tunnetaan, jolloin saadaan $x_1 = 20/3$. Tämä vaihe Gaussin eliminoinnissa on siis takaisinsijoittaminen.

Esimerkki 8.2.2. Tarkastellaan seuraavaksi yhtälöryhmää

$$\begin{cases} x - 2y + 4z = 2 \\ 2x - 3y + 5z = 3 \\ 3x - 4y + 6z = 7 \end{cases}.$$

Ratkaistaan tämä Gaussin-Jordanin eliminointimenettelyllä (nyt kerroinmatriisi ei ole säännöllinen)

$$\begin{pmatrix} 1 & -2 & 4 & \vdots & 2 \\ 2 & -3 & 5 & \vdots & 3 \\ 3 & -4 & 6 & \vdots & 7 \end{pmatrix} \overset{v1}{\sim} \begin{pmatrix} 1 & -2 & 4 & \vdots & 2 \\ 0 & 1 & -3 & \vdots & -1 \\ 0 & 2 & -6 & \vdots & 1 \end{pmatrix} \overset{v2}{\sim} \begin{pmatrix} 1 & -2 & 4 & \vdots & 2 \\ 0 & 1 & -3 & \vdots & -1 \\ 0 & 0 & 0 & \vdots & 3 \end{pmatrix}.$$

(v1) Lisätään ensimmäinen rivi toiseen -2:lla kerrottuna ja kolmanteen -3:lla kerrottuna.

(v2) Lisätään toinen rivi -2:lla kerrottuna kolmanteen riviin.

Kolmas yhtälö vastaa nyt degeneroitunutta yhtälöä $0 = 3$, jolla ei ole ratkaisua. Koska tällä muunnetulla yhtälöryhmällä on samat ratkaisut kuin alkuperäisellä yhtälöryhmällä, ratkaisuja ei ole yhtään.

Esimerkki 8.2.3. Tarkastellaan vielä yhtälöryhmää

$$\begin{cases} 2x + 4y - 5z + 3t = 0 \\ 3x + 6y - 7z + 4t = 0 \\ 5x + 10y - 11z + 6t = 0 \end{cases},$$

jossa on nyt enemmän tuntemattomia kuin yhtälöitä. Etsitään ratkaisu Gaussin-Jordanin eliminointimenettelyä käyttämällä

$$\begin{pmatrix} 2 & 4 & -5 & 3 & \vdots & 0 \\ 3 & 6 & -7 & 4 & \vdots & 0 \\ 5 & 10 & -11 & 6 & \vdots & 0 \end{pmatrix} \overset{v1}{\sim} \begin{pmatrix} 2 & 4 & -5 & 3 & \vdots & 0 \\ 0 & 0 & 1 & -1 & \vdots & 0 \\ 5 & 10 & -11 & 6 & \vdots & 0 \end{pmatrix} \overset{v2}{\sim} \begin{pmatrix} 2 & 4 & -5 & 3 & \vdots & 0 \\ 0 & 0 & 1 & -1 & \vdots & 0 \\ 0 & 0 & 3 & -3 & \vdots & 0 \end{pmatrix}$$

$$\overset{v3}{\sim}\begin{pmatrix} 2 & 4 & -5 & 3 & \vdots & 0 \\ 0 & 0 & 1 & -1 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 \end{pmatrix}.$$

(v1) Kerrotaan toinen yhtälö 2:lla ja lisätään siihen ensimmäinen yhtälö -3:lla kerrottuna.

(v2) Kerrotaan kolmas yhtälö 2:lla ja lisätään siihen ensimmäinen yhtälö -5:llä kerrottuna.

(v3) Lisätään toinen yhtälö -3:lla kerrottuna kolmanteen yhtälöön.

Toinen rivi vastaa nyt yhtälöä $z = t$, jota käyttämällä ensimmäisestä yhtälöstä saadaan

$$2x + 4y - 5z + 3t = 2x + 4y - 2z = 0.$$

Yhtälöryhmän ratkaisuja ovat kaikki (x, y, z, t), jotka toteuttavat yhtälöt $2x+4y-5z+3t = 0$ ja $z = t$. Yhtälöryhmän ratkaisut voidaan nyt esittää myös muodossa

$$\{(x, y, x + 2y, x + 2y) \mid x, y \in \mathbb{R}\}.$$

Parametrisoidulla ratkaisulla tarkoitetaan juuri tällaisessa muodossa annettua ratkaisua. Huomaa, että ratkaisu on nyt annettu kahden parametrin, x ja y, avulla lausuttuna. Ratkaisujen joukko on siten kaksi-ulotteinen vektorialiavaruus $\mathbb{R}^4$:ssa. Voit pohtia miten tämä tulos liittyy lineaarialgebran peruslauseeseen.

Palataan vielä käänteismatriisin laskemiseen. Käänteismatriisin laskeminen edellyttää yhtälöryhmän $\mathbf{AB} = \mathbf{I}$ laskemista. Jos $\mathbf{A}$ on $n \times n$ matriisi, niin tässä yhtälöryhmässä on oikeastaan n yhtälöryhmää. Havainnollistetaan tätä esimerkin avulla.

Esimerkki 8.2.4. Olkoon

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

Ratkasitavana on nyt yhtälösysteemi

$$\begin{pmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & -b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Laskemalla tämä auki saadaan

$$\begin{pmatrix} -b_{11} + b_{21} + 2b_{31} & -b_{12} + b_{22} + 2b_{32} & \cdots \\ 3b_{11} - b_{21} + b_{31} & 3b_{12} - b_{22} + b_{32} & \cdots \\ -b_{11} + 3b_{21} + 4b_{31} & -b_{12} + 3b_{22} + 4b_{32} & \cdots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Voimme nyt havaita, että tässä on itseasiassa kolme yhtälöryhmää (sarakkeista muodostuvat yhtälöt). Niillä on kuitenkin kaikilla sama kerroinmatriisi.

Kuten edellä olevassa esimerkissä havaittiin, käänteismatriisin laskeminen edellyttää yhtälöryhmien $\mathbf{A}\mathbf{x}_i = \mathbf{e}_i$, $i = 1, \ldots, n$, ratkaisemista. Ratkaisuna saatavat vektorit $\mathbf{x}_i$ ovat käänteismatriisin sarakevektorit. Koska kerroinmatriisi kaikille näille yhtälöille on sama, voidaan ratkaisu hakea käsittelemällä yhtälöryhmiä samanaikaisesti. Juuri näin toimitaan Gaussin-Jordanin eliminoinnissa käänteismatriisille.

8.3 Sovellus: tasapaino oligopolimalleissa

Mikroteorian aineopinnoissa käydään läpi Cournot-kilpailumalli oligopolistiselle toimialalle. Toimialalla on n yritystä ja kunkin yrityksen i optimaalinen tuotannon taso riippuu sen omista (vakioksi oletetuista) rajakustannuksista c_i , ja muiden yritysten tuotannosta q_j seuraavasti:

$$q_i = \frac{a - \beta \sum_{j \neq i} q_j - c_i}{2\beta}.$$

Tasapainoehdot voidaan kirjoittaa matriisimuodossa:

$$\begin{pmatrix} 2\beta & \beta & \cdots & \beta \\ \beta & 2\beta & \cdots & \beta \\ \vdots & & \ddots & \vdots \\ \beta & \beta & \cdots & 2\beta \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} = \begin{pmatrix} a - c_1 \\ a - c_2 \\ \vdots \\ a - c_n \end{pmatrix}.$$

Tarkastellaan vaakarivioperaatioita täydennetylle matriisille

$$\begin{pmatrix} 2\beta & \beta & \cdots & \beta & \vdots & a - c_1 \\ \beta & 2\beta & \cdots & \beta & \vdots & a - c_2 \\ \vdots & & \ddots & \vdots & \vdots & \vdots \\ \beta & \beta & \cdots & 2\beta & \vdots & a - c_n \end{pmatrix}.$$

Vähennetään ensimmäinen rivi kaikista muista riveistä:

$$\begin{pmatrix} 2\beta & \beta & \cdots & \beta & \vdots & a - c_1 \\ -\beta & \beta & \cdots & 0 & \vdots & c_1 - c_2 \\ \vdots & & \ddots & \vdots & \vdots & \vdots \\ -\beta & 0 & \cdots & \beta & \vdots & c_1 - c_n \end{pmatrix}.$$

Jokaiselle $j \neq 1$ saadaan tulos

$$\beta q_j = \beta q_1 + c_1 - c_j. \tag{8.1}$$

Ensimmäinen yhtälö puolestaan kertoo

$$2\beta q_1 + \beta \sum_{j \neq i} q_j = a - c_1.$$

Ratkaistaan q_1 sijoittamalla ylempi yhtälöistä alempaan, jolloin

$$2\beta q_1 + (n-1)\beta q_1 + (n-1)c_1 - \sum_{j \neq i} c_j = a - c_1,$$

mistä voidaan ratkaista q_1:

$$q_1 = \frac{a - c_1 + \sum_{j \neq i}(c_j - c_1)}{(n+1)\beta}$$

Muut tuotannot saadaan joko yhtälöä (8.1) käyttäen tai käyttämällä j:nnettä vaakariviä matriisin muuntamisessa.

Luku 9

Ominaisarvot ja ominaisvektorit

Ominaisarvoteorian lähtökohta on se, että annettu neliömatriisi $\mathbf{A}$ halutaan muuntaa helpommin käsiteltävään muotoon. Esimerkiksi, kun tarkastelemme matriisin kertomista itsensä kanssa $\mathbf{A}^k$, $\mathbf{A}$ on siis kerrottu itsensä kanssa k kertaa, huomaamme että, jos $\mathbf{A}$ olisi diagonaalimatriisi $\mathbf{A} = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$, niin tällöin $\mathbf{A}^k = \operatorname{diag}(\alpha_1^k, \ldots, \alpha_n^k)$.

Kysymys on nyt, että voidaanko tehdä sellainen muuttujan vaihdos, että matriisin $\mathbf{A}$ sijaan voidaan uusissa muuttujissa tarkastella *helpommin käsiteltävää matriisia* kuten diagonaalimatriisia? Muuttujan vaihdoksella (tai koordinaatiston muunnoksella tai kannan vaihdolla) tarkoitetaan tässä kääntyvää kuvausta (matriisia) $\mathbf{P}$; uudet muuttujat ovat $\mathbf{P}^{-1}\mathbf{x}$. Tarkastellaan miten $\mathbf{A} \in \mathbb{R}^{n \times n}$ toimii kun muuttujat vaihdetaan: otetaan $\mathbf{y} \in \mathbb{R}^n$ (uudet muuttujat) ja haetaan sitä vastaava $\mathbf{x} \in \mathbb{R}^n$ (vanhat muuttujat), eli $\mathbf{x} = \mathbf{P}\mathbf{y}$, kuvataan $\mathbf{A}$:lla, jolloin saadaan $\mathbf{A}\mathbf{P}\mathbf{y}$ ja suoritetaan taas muuttujan vaihdos $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{y}$ ($\mathbf{y}$:n kuva uusissa muuttujissa). Siis muunnettu kuvaus on $\mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Palataan nyt alkuperäiseen kysymykseen: löytyykö $\mathbf{P}$ siten että $\mathbf{\Lambda}$ on diagonaalimatriisi?

Merkitään λ_i:llä $\mathbf{\Lambda}$:n i:nnen rivin diagonaalialkiota ja $\mathbf{v}_i$:llä $\mathbf{P}$:n i:ttä sarakevektoria. Tällöin pätee $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ (sillä $\mathbf{P}\mathbf{\Lambda} = \mathbf{A}\mathbf{P}$). Tätä yhtälöä käytetään ominaisarvojen ja niitä vastaavien ominaisvektoreiden määritelmänä. n

Määäritelmä 9.0.1. λ on $\mathbf{A}$:n ominaisarvo ja $\mathbf{v} \neq \mathbf{0}$ on vastaava ominaisvektori jos $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

Huom. oletetaan, että $\mathbf{v} \neq \mathbf{0}$, koska $\mathbf{0}$ olisi aina ominaisvektori eikä siten kovinkaan kiinnostava.

9.1 Ominaisarvojen laskeminen

Ominaisarvoja etsittäessä on ratkaistava neliömatriisia $\mathbf{A}_{n\times n}$ koskeva yhtälö

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v},$$

missä $\mathbf{v} = (v_1, \ldots, v_n)$ ja λ on tuntematon ominaisarvo, jota ollaan hakemassa. Kuten edellä määriteltiin: mikäli on olemassa jokin nollasta eroava vektori $\mathbf{v} \in \mathbb{R}^n$ ja luku λ, joille tämä yhtälö on voimassa, niin tällöin lukua λ kutsutaan matriisin $\mathbf{A}$ *ominaisarvoksi* ja vektoria $\mathbf{v} \neq \mathbf{0}$ kutsutaan ominaisarvoon λ liittyväksi *ominaisvektoriksi.* Huomaa, että nollavektori ei käy ominaisvektoriksi, mutta $\lambda = 0$ voi kuitenkin olla ominaisarvo.

Matriisien laskusääntöjen mukaan yhtälö $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ voidaan kirjoittaa myös muodossa

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{v} = \mathbf{0}.$$

Huomaa nyt, että kyseessä on lineaarinen yhtälöryhmä. Mikäli tämän yhtälöryhmän kerroinmatriisi on säännöllinen, eli $\det(\mathbf{A} - \lambda\mathbf{I}) \neq 0$, niin tällöin ainoa ratkaisu on $\mathbf{v} = (\mathbf{A} + \lambda\mathbf{I})^{-1}\,\mathbf{0} = \mathbf{0}$. Näin ollen jokaisen ominaisarvon täytyy toteuttaa yhtälö, nk. *karakteristinen yhtälö*

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix} = 0,$$

joka on astetta n oleva polynomiyhtälö (kun determinantti lasketaan auki)

$$c_{\mathbf{A}}(\lambda) = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0 = 0,$$

missä kertoimet c_i riippuvat matriisin $\mathbf{A}$ kertoimista a_{ij}. Tämän polynomiyhtälön juuret ovat matriisin $\mathbf{A}$ kaikki ominaisarvot ja sitä kutsutaan matriisin $\mathbf{A}$ *karakteristiseksi polynomiksi.* Mikäli karakteristisella polynomilla on k -kertainen juuri, sen sanotaan olevan matriisin $\mathbf{A}$ *k -kertainen ominaisarvo.*[1]

Kannattaa huomata, että $n \times n$ -matriisilla voi olla korkeintaan n eri ominaisarvoa, koska sen karakteristinen polynomi on astetta n.

Esimerkki 9.1.1. Joskus matriisin yksi tai usempia ominaisarvoja pystytään löytämään yksinkertaisesti silmäilemällä matriisia. Tarkastellaan esimerkiksi matriisia

$$\mathbf{A} = \begin{pmatrix} 5 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

[1] Jokainen ominaisarvo on karakteristisen polynomin juuri, mutta tarkalleen ottaen asiaa ei ole vielä perusteltu toisin päin: miksi jokainen karakteristisen polynomin juuri on varmasti ominaisarvo? Koska jokainen karakteristisen polynomin juuri toteuttaa ehdon $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, lineaarisella yhtälöryhmällä $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ on joko 0 tai ∞ määrä ratkaisuja. Koska $\mathbf{x} = \mathbf{0}$ on kuitenkin yhtälöryhmän ratkaisu aina, ratkaisuja täytyy olla ∞ määrä. Näin ollen yhtälö $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ toteutuu jollakin vektorilla $\mathbf{x} \neq 0$, joten kyseessä täytyy olla ominaisarvo.

Erottamalla 4 tämän matriisin jokaisesta diagonaalialkiosta, saadaan

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -2 \end{pmatrix}.$$

Tämän matriisin determinantti on 0, koska siinä on kaksi samaa saraketta (ks. kohta determinantin ominaisuudet). Luku 4 on siis tämän matriisin yksi ominaisarvo.

Esimerkki 9.1.2. Tarkastellaan diagonaalimatriisia

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Kun tämän matriisin jokaisesta diagonaalialkiosta erotetaan 2, saadaan matriisi

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

jonka determinantti on 0 (ks. kohta determinantin ominaisuudet). Nähdään, että diagonaalimatriisin ominaisarvot ovat sen päälävistäjällä olevat alkiot. Ominaisarvoja ei voi olla enempää kuin päälävistäjällä olevia alkioita, koska karakteristisen polynomin aste on sama kuin päälävistäjällä olevien alkioiden lukumäärä. Tämä tulos pätee yleisesti mille tahansa diagonaalimatriiseille. Sama tulos pätee myös ylä- ja alakolmiomatriiseille (miksi?).

Yleensä ominaisarvoja ei pystytä näkemään suoraan, vaan joudutaan etsimään karakteristisen polynomin juuret.

Esimerkki 9.1.3. Lasketaan matriisin

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$$

ominaisarvot. Tämän matriisin ominaisarvot ovat karakteristisen polynomin

$$c_{\mathbf{A}}(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 5-\lambda \end{vmatrix} = (2-\lambda)(5-\lambda) - 1 = \lambda^2 - 7\lambda + 9$$

juuret. Nämä saadaan toisen asteen yhtälön ratkaisukaavasta

$$\lambda_{1,2} = \frac{7 \mp \sqrt{(-7)^2 - 4 \cdot 1 \cdot 9}}{2} = \frac{7 \mp \sqrt{13}}{2}.$$

Esimerkki 9.1.4. Ratkaistaan matriisin

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 5 & 0 \\ 3 & 0 & 2 \end{pmatrix}$$

ominaisarvot. Karakteristinen polynomi on nyt

$$c_{\mathbf{B}}(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & 5-\lambda & 0 \\ 3 & 0 & 2-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 5-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} + 2\begin{vmatrix} 0 & 5-\lambda \\ 3 & 0 \end{vmatrix}$$

$$\begin{aligned} &= (1-\lambda)(5-\lambda)(2-\lambda) - 6(5-\lambda) = [(1-\lambda)(2-\lambda) - 6](5-\lambda) \\ &= (\lambda^2 - 3\lambda - 4)(5-\lambda) = (\lambda+1)(\lambda-4)(5-\lambda). \end{aligned}$$

Tämän matriisin ominaisarvot ovat siis

$$\lambda_1 = -1, \quad \lambda_2 = 4 \text{ ja } \lambda_3 = 5.$$

9.2 Ominaisarvojen ominaisuuksia

Ominaisarvoilla on muutamia tärkeitä ominaisuuksia, jotka tulevat usein vastaan esimerkiksi ekonometrisissa sovellutuksissa. Oletetaan, että luvut $\lambda_1, \ldots, \lambda_n$ ovat kaikki matriisin $\mathbf{A}_{n\times n}$ ominaisarvot. Nämä eivät nyt välttämättä ole kaikki erisuuria, vaan k-kertainen ominaisarvo esiintyy listassa k kertaa. Tällöin

$$\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i = \lambda_1\lambda_2 \ldots \lambda_n$$

ja

$$\operatorname{tr}(\mathbf{A}) = a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{n} \lambda_i.$$

Esimerkki 9.2.1. Tarkastellaan esimerkin 9.1.3 matriisia. Tämän matriisin determinantti on

$$\begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix} = 2 \cdot 5 - 1 \cdot 1 = 9$$

ja ominaisarvojen tulo on

$$\lambda_1\lambda_2 = \frac{7-\sqrt{13}}{2} \cdot \frac{7+\sqrt{13}}{2} = \frac{1}{4}\left(7^2 - \left(\sqrt{13}\right)^2\right) = \frac{36}{4} = 9.$$

Esimerkki 9.2.2. Tarkastellaan esimerkin 9.1.4 matriisia. Tämän matriisin jälki on

$$\operatorname{tr}\left[\begin{pmatrix} 1 & 0 & 2 \\ 0 & 5 & 0 \\ 3 & 0 & 2 \end{pmatrix}\right] = 1 + 5 + 2 = 8$$

ja ominaisarvojen summa on

$$\lambda_1 + \lambda_2 + \lambda_3 = -1 + 4 + 5 = 8.$$

Edellä annettujen ominaisarvojen ominaisuuksien perusteella voidaan todeta, että matriisi on säännöllinen täsmälleen sillon, kun 0 ei ole sen ominaisarvo. Listataan myös muita havaintoja ominaisarvoista:

1. diagonaalimatriisin diagonaalialkiot ovat ominaisarvoja
2. jos v on ominaisvektori, niin myös αv on ominaisvektori kaikilla $\alpha \neq 0$
3. ominaisvektori määrittää suunnan, jota kuvaus A ei muuta
4. jos matriisi on symmetrinen, sillä on n ominaisarvoa, jotka ovat reaalilukuja

On syytä huomata, että osa karakteristisen yhtälön ratkaisuista voi olla kompleksilukuja. Myös kompleksiset ominaisarvot antavat arvokasta tietoa matriisin määärittämästä lineaarikuvauksesta. Kompleksisiin ominaisarvoihin palataan myöhemmin. Kohdassa kolme todetaan, että ominaisvektori on määrittää suunnan, joka ei muutu sitä kuvatteassa matriisin määrittämällä lineaarikuvauksella. Ominaisvektoreiden avulla voi siis hahmotella matriisin määräämän lineaarikuvauksen luonnetta:

1. Jos A:n ominaisarvo on positiivinen, se tarkoittaa sitä, että A skaalaa vektoreita vastaavan ominaisvektorin suuntaan.
2. Jos ominaisarvo on negatiivinen, A peilaa ominaisvektorin suunnat ja skaalaa niitä.
3. Jos reaalisia ominaisarvoja ei ole, niin A ei toimi edellä kuvatulla tavalla mihinkään suuntaan, joten matriisi kääntää kaikkia suuntia (ja mahdollisesti skaalaa niitä).

Esimerkki 9.2.3. Pohdi miten matriisin

$$\mathbf{A} = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}$$

määrittämä lineaarikuvaus toimii ja mitkä ovat sen ominaisarvot. Entä alla olevan matriisin määrittämä lineaarikuvaus

$$\mathbf{A} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}?$$

9.3 Ominaisvektoreiden laskeminen

Joskus joudutaan tilanteeseen, jossa täytyy pystyä laskemaan myös eri ominaisarvoihin liittyvät ominaisvektorit. Ominaisvektoreita tarvitaan esimerkiksi joidenkin matriisien hajotelmien yhteydessä. Tähän palataan seuraavassa luvussa. Kun ominaisarvot tunnetaan, voidaan näitä vastaavat ominaisvektorit ratkaista yhtälöstä

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{v} = \mathbf{0}.$$

Kannattaa kuitenkin muistaa, ettei tämän yhtälöryhmän ratkaisu ole ikinä yksikäsitteinen. Mikäli vektorit $\mathbf{v}_1$ ja $\mathbf{v}_2$ ovat kummatkin matriisin $\mathbf{A}$ ominaisarvoon λ liittyviä ominaisvektoreita, niin tällöin myös $a\mathbf{v}_1 + b\mathbf{v}_2$ on ominaisarvoon λ liittyvä ominaisvektori kaikilla $a, b \in \mathbb{R}$, sillä

$$\begin{aligned}\mathbf{A}(a\mathbf{v}_1 + b\mathbf{v}_2) = \mathbf{A}(a\mathbf{v}_1) + \mathbf{A}(b\mathbf{v}_2) = a\mathbf{A}\mathbf{v}_1 + b\mathbf{A}\mathbf{v}_2 &= a\lambda\mathbf{v}_1 + b\lambda\mathbf{v}_2 \\ &= \lambda(a\mathbf{v}_1 + b\mathbf{v}_2).\end{aligned}$$

Tämä tulos yleistyy hyvin suoraviivaisesti myös useampien ominaisvektoreiden tapaukseen: kaikki yhtä ominaisarvoa vastaavat ominaisvektorit nollavektori mukaanlukien muodostavat vektorialiavaruuden nk. ominaisavaruuden. Koska mikä tahansa ominaisvektoreiden lineaarikombinaatio on edelleen ominaisvektori, saadaan kaikki tiettyyn ominaisarvoon liittyvät ominaisvektorit spesifioitua antamalla suurin mahdollinen lineaarisesti riippumattomien ominaisvektoreiden joukko. Kaikki ominaisvektorit saadaan näiden lineaarikombinaatioina.

Ominaisvektoreita etsittäessä on hyvä tietää, että ominaisarvoon voi liittyä korkeintaan sen kertaluvun verran lineaarisesti riippumattomia ominaisvektoreita. Mikäli ominaisarvon kertaluku on siis 1, niin tällöin sen kaikki ominaisvektorit voidaan spesifioida antamalla mikä tahansa ominaisvektori — loput ominaisvektorit ovat tämän skalaarimonikertoja.

Esimerkki 9.3.1. Ratkaistaan jokin esimerkissä 9.1.4 tarkastellun matriisin ominaisarvoon 4 liittyvä ominaisvektori. Nyt täytyy ratkaista yhtälö

$$\left[\begin{pmatrix} 1 & 0 & 2 \\ 0 & 5 & 0 \\ 3 & 0 & 2 \end{pmatrix} - 4\mathbf{I}_{3\times 3}\right]\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

joka auki kerrottuna vastaa lineaarista yhtälöryhmää

$$\begin{cases} -3v_1 \qquad + 2v_3 = 0 \\ \qquad\quad v_2 \qquad\quad = 0 \\ \;\;3v_1 \qquad - 2v_3 = 0 \end{cases}.$$

On helppo nähdä, että tämän yhtälöryhmän eräs ratkaisu on $(v_1, v_2, v_3) = (2, 0, 3)$. Vektori $\mathbf{v} = (2, 0, 3)$ on siis yksi ominaisarvoon 4 liittyvistä ominaisvektoreista. Koska tämän ominaisarvon kertaluku on 1, kaikki ominaisvektorit ovat muotoa $a(2, 0, 3)$, missä $a \in \mathbb{R}$.

Esimerkki 9.3.2. Ratkaistaan matriisin

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}$$

ominaisarvot ja näitä vastaavat ominaisvektorit. Matriisin karakteristinen polynomi on

$$c_{\mathbf{A}}(\lambda) = \begin{vmatrix} 1-\lambda & 3 \\ -1 & 5-\lambda \end{vmatrix} = (1-\lambda)(5-\lambda) + 3 = \lambda^2 - 6\lambda + 8.$$

Ominaisarvot ovat siis

$$\lambda_{1,2} = \frac{6 \mp \sqrt{(-6)^2 - 4 \cdot 1 \cdot 8}}{2} = \frac{6 \mp \sqrt{4}}{2} = \frac{6 \mp 2}{2} = 4 \text{ ja } 2.$$

Ominaisarvoa 4 vastaava ominaisvektori löydetään yhtälön $(\mathbf{A} - 4\mathbf{I})\,\mathbf{v} = \mathbf{0}$, eli lineaarisen yhtälöryhmän

$$\begin{cases} -3v_1 + 3v_2 = 0 \\ -v_1 + v_2 = 0 \end{cases}$$

ratkaisuna. Tämän eräs ratkaisu on selvästi $\mathbf{v} = (1,1)^\top$. Ominaisarvoa 2 vastaava ominaisvektori saadaan vastaavasti yhtälön $(\mathbf{A} - 2\mathbf{I})\,\mathbf{v} = \mathbf{0}$, eli lineaarisen yhtälöryhmän

$$\begin{cases} -v_1 + 3v_2 = 0 \\ -v_1 + 3v_2 = 0 \end{cases}$$

ratkaisuna. Tämän eräs ratkaisu on selvästi $\mathbf{v} = (3,1)$.

9.4 Ominaisarvohajotelma

Olkoon $\mathbf{A}$ $n \times n$-matriisi, jolla on olemassa n lineaarisesti riippumatonta ominaisvektoria. Voidaan osoittaa, että eri ominaisarvoihin liittyvät ominaisvektorit ovat aina lineaarisesti riippumattomia. Näin ollen matriisilla on n lineaarisesti riippumatonta ominaisvektoria, mikäli jokaiseen ominaisarvoon liittyy sen kertaluvun verran lineaarisesti riippumattomia ominaisvektoreita. Tällaista matriisia sanotaan diagonalisoituvaksi matriisiksi.

Olkoot $\lambda_1, \ldots, \lambda_n$ matriisin $\mathbf{A}$ ominaisvektorit, jotka eivät välttämättä ole kaikki erisuuria. Olkoot $\mathbf{v}_1, \ldots, \mathbf{v}_n$ näitä vastaavat ominaisvektorit, jotka nyt siis ovat lineaarisesti riippumattomia. Matriisi $\mathbf{A}$ voidaan tällöin esittää muodossa

$$\mathbf{A} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^{-1},$$

missä Λ on diagonaalimatriisi, jonka päälävistäjän alkiot ovat ominaisarvot, ja matriisin $\mathbf{P}$ sarakkeet ovat näitä vastaavat ominaisvektorit samassa järjestyksessä. Tällaista hajotelmaa kutsutaan matriisin *spektraaliesitykseksi* tai *ominaisarvohajotelmaksi*. Matriisia $\mathbf{P}$ kutsutaan joskus similariteettimuunnosmatriisiksi ja diagonaalimatriisia $\boldsymbol{\Lambda}$ similariteettimuunnokseksi. Kun käytetään merkintää $\mathbf{v}_i = (v_{1i}, \ldots, v_{ni})$, hajotelma voidaan kirjoittaa vielä tarkemmin muodossa

$$\mathbf{A} = \begin{pmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} & \ldots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \ldots & v_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} & \ldots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \ldots & v_{nn} \end{pmatrix}^{-1}.$$

Joskus matriisin $\mathbf{P}$ sarakkeina käytetään normalisoituja, eli omalla pituudellaan jaettuja, ominaisvektoreita. Tällöin sarakkeina ovat siis vektorit

$$\frac{\mathbf{v}_1}{\parallel \mathbf{v}_1 \parallel}, \dots, \frac{\mathbf{v}_n}{\parallel \mathbf{v}_n \parallel}.$$

Intuitiivisesti on helppo ymmärtää miksi näin voidaan tehdä. Koska spektraaliesityksessä esiintyy kummatkin matriisit $\mathbf{P}$ ja $\mathbf{P}^{-1}$, jonkinlainen normalisointi tapahtuu joka tapauksessa.

Esimerkki 9.4.1. Esimerkissä 9.3.2 ratkaistiin matriisin

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}$$

ominaisarvot ja ominaisvektorit. Ominaisarvot ovat $\lambda_1 = 4$ ja $\lambda_2 = 2$. Ominaisvektorit ovat $\mathbf{v}_1 = (1, 1)^\top$ ja $\mathbf{v}_2 = (3, 1)^\top$ vastaavasti. Lineaarisesti riippumattomia ominaisvektoreita on siis yhtä monta kuin ominaisarvoja, joten spektraaliesitys on olemassa. Tarkastetaan asia:

$$\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -\frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 6 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} -\frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} = \mathbf{A}.$$

Luku 10

Lineaariset differenssiyhtälöt

Systeemiä $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$, $k = 0, 1, \ldots$ sanotaan lineaariseksi differenssiyhtälöksi. Tässä k on aikaindeksi ja vektori $\mathbf{z}$ kuvaa jotakin ajassa etenevää ilmiötä. Siirtyminen jaksosta toiseen määräytyy matriisin $\mathbf{A}$ määrittämästä lineaarikuvauksesta rekursiivisesti. Jos halutaan tietää $\mathbf{z}$ hetkellä k on tunnettava $\mathbf{z}_0$, eli alkutila. Jos $\mathbf{z}_k$ voidaan esittää suljetussa muodossa, siten, että sen laskemiseksi ei tarvitse käydä rekursiota alusta hetkeen k asti, niin tällöin sanotaan, että tiedämme differenssiyhtälön ratkaisun. Tarkoituksena on seuraavaksi hyödyntää ominaisarvoteoriaa lineaaristen differenssiyhtälöiden ratkaisemisessa. Tarkastellaan aluksi paria esimerkkiä.

Esimerkki 10.0.2. Tarkastellaan korkoa korolle mallia: pääoma periodilla k on y_k, vuosikorko ρ, $y_{k+1} = (1+\rho)y_k$. Laskemalla periodista 0 alkaen, $y_1 = (1+\rho)y_0$, $y_2 = (1+\rho)^2 y_0$, jne. havaitaan, että $y_k = (1+\rho)^k y_0$ (differenssiyhtälön ratkaisu).

Esimerkki 10.0.3. Oletetaan, että mallissa työttömyydelle: työssäkäyvät keskimäärin x, työttömät keskimäärin y, työllistymistodennäköisyys p, työssäpysymistodennäköisyys q, jolloin

$$\begin{aligned} x_{k+1} &= qx_k + py_k \\ y_{k+1} &= (1-q)x_k + (1-p)y_k \end{aligned}$$

tai matriisimuodossa

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} q & p \\ 1-q & 1-p \end{pmatrix} \begin{pmatrix} x_k \\ y_k \end{pmatrix}$$

Kahden muuttujan mallista havaitsemme, että jos kerroinmatriisi $\mathbf{A}$ olisi diagonaalimatriisi, niin muuttujat (z:n komponentit) olisivat toisistaan riippumattomia (*uncoupled*) ja komponentit periodilla k voisi laskea erikseen kuten korkoesimerkissä, esim. työttömyysmallissa x ja y eivät näytä erikseen ratkeavan kuten korkoesimerkissä (kokeile). Ominaisarvohajotelmalla saamme muuttujan vaihdoksen, joka antaa diagonaaliseen kerroinmatriisiin, jossa lävistäjällä ovat ominaisarvot.

10.1 Ratkaiseminen diagonalisoinnilla

Oletetaan, että $\mathbf{A}$ on diagonalisoituva. Differenssiyhtälön $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$ ratkaiseminen diagonalisoinnilla voidaan ymmärtää helpoiten siten, että ratkaisuna haetaan esitys $\mathbf{A}^k$:lle. Voimme nimittäin havaita, että $\mathbf{z}_k = \mathbf{A}^k\mathbf{z}_0$. Induktiopäättelyllä saamme:

$$\mathbf{A}^k = \mathbf{P}^{-1}\mathbf{\Lambda}^k\mathbf{P}$$

ja koska $\mathbf{\Lambda}$ on diagonaalimatriisi voidaan $\mathbf{\Lambda}^k$ laskea helposti; $\mathbf{\Lambda}^k$:n diagonaalilla on $\mathbf{\Lambda}$:n diagonaalialkiot k:nnessa potenssissa. Ominaisarvohajotelma johtaa siis siihen että hankalan k-kertaisen matriisitulon sijaan saamme ratkaisun ominaisarvojen avulla varsin helposti.

Esimerkki 10.1.1. Tässä esimerkissä perustellaan ominaisarvojen käyttö muuttujan (koordinaatiston) vaihdoksella. Ratkaistaan matriisin

$$\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 1/2 & 0 \end{pmatrix}$$

määrittämä differenssiyhtälö $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$. Lasketaan ominaisarvot ja ominaisvektorit: karakteristinen yhtälö $\det(\mathbf{A}-\lambda\mathbf{I}) = 0 \Leftrightarrow (1-\lambda)(0-\lambda)-4\cdot 1/2 = 0$. Saadaan ominaisarvot $\lambda_{1,2} = 2, -1$. Ominaisvektorit, (i)

$$(\mathbf{A} - \lambda_1\mathbf{I}) = \begin{pmatrix} -1 & 4 \\ 1/2 & -2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

joten mm. $\mathbf{v}^1 = (4, 1)$ kelpaa ominaisvektoriksi.
(ii)

$$(\mathbf{A} - \lambda_2\mathbf{I}) = \begin{pmatrix} 2 & 4 \\ 1/2 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

joten mm. $\mathbf{v}^2 = (-2, 1)$ kelpaa toiseksi ominaisvektoriksi. Edelleen muunnosmatriisiksi saadaan

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

ja

$$\mathbf{P}^{-1} = \begin{pmatrix} 1/6 & 1/3 \\ -1/6 & 2/3 \end{pmatrix}.$$

Tehdään muuttujan vaihdos, merkitään $\mathbf{z} = (x, y)$, uudet muuttujat X ja Y, $\mathbf{Z} = (X, Y)$:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1/6 & 1/3 \\ -1/6 & 2/3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

ja

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$$

Johdetaan differenssiyhtälö $\mathbf{Z}$:lle: $\mathbf{Z}_{k+1} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{Z}_k$ ja $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ on diagonaalimatriisi, jossa ominaisvektorit diagonaalilla, eli

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

Tässä differenssiyhtälössä X ja Y eivät riipu toisistaan, koska $\mathbf{B}$ on diagonaalimatriisi!

Ratkaistaan $\mathbf{Z}_k$: $X_k = 2^k X_0$ ja $Y_k = (-1)^k Y_0$. Siirrytään takaisin alkuperäisiin muuttujiin

$$\begin{aligned}
\mathbf{z}_k = \mathbf{P}\mathbf{Z}_k &= \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^k X_0 \\ (-1)^k Y_0 \end{pmatrix} \\
&= \begin{pmatrix} 4 \cdot 2^k X_0 - 2 \cdot (-1)^k Y_0 \\ 2^k X_0 + (-1)^k Y_0 \end{pmatrix} = (2^k X_0) \begin{pmatrix} 4 \\ 1 \end{pmatrix} + ((-1)^k Y_0) \begin{pmatrix} -2 \\ 1 \end{pmatrix}
\end{aligned}$$

Alkuarvot X_0 ja Y_0 saadaan taas muunnoksella $\mathbf{Z}_0 = \mathbf{P}^{-1}\mathbf{z}_0$.

Esimerkki 10.1.2. Ominaisarvojen avulla voimme myös määritellä monimutkaisempia matriisiarvoisia funktioita. Esimerkiksi eksponenttifunktiolla on tunnettu sarjakehitelmä

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

Oletetaan, että $\mathbf{A}$ on diagonalisoituva matriisi. Tällöin voimme määritellä funktion $e^{\mathbf{A}}$ edellä kuvatulla sarjakehitelmällä ja käyttämällä ominaisarvohajotelmaa saamme

$$e^{\mathbf{A}} = \mathbf{P}^{-1}\mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})\mathbf{P},$$

missä $\lambda_1, \ldots, \lambda_n$ ovat matriisin ominaisarvot.

10.2 Yleinen ratkaisu

Oletetaan, että $\mathbf{A}$ diagonalisoituva ja sen ominaisarvot ovat reaalilukuja, ominaisarvot $\lambda_1, \ldots, \lambda_n$ ja ominaisvektorit $\mathbf{v}_1, \ldots, \mathbf{v}_2$. Merkitään

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix},$$

jolloin

$$\mathbf{\Lambda}^k = \begin{pmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{pmatrix}.$$

Lause 10.2.1. *Differenssiyhtälön $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$ yleinen ratkaisu on*

$$\mathbf{z}_k = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \ldots + c_n\lambda_n^k\mathbf{v}_n,$$

missä c_i, $i = 1, \ldots, n$, ovat vakioita.

Yleisellä ratkaisulla tarkoitetaan sitä että ratkaisu on aina edellä kuvattua muotoa riippumatta alkuarvosta $\mathbf{z}_0$. Vakiot c_i, $i = 1\ldots, n$, voidaan ratkaista, kun alkuarvo $\mathbf{z}_0$ on annettu; merkitään $\mathbf{c} = (c_1, \ldots, c_n)$, jolloin $\mathbf{c} = \mathbf{P}^{-1}\mathbf{z}_0$.

Lause 10.2.2. *Kun alkuarvo on* $\mathbf{z}_0$ *niin differenssiyhtälön ratkaisu on* $\mathbf{z}_k = \mathbf{P}\mathbf{\Lambda}^k\mathbf{P}^{-1}\mathbf{z}_0$.

Esimerkki 10.2.1. Ratkaistaan $\mathbf{P}\mathbf{\Lambda}^k\mathbf{P}^{-1}$ differenssiyhtälölle $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$,

$$\mathbf{A} = \begin{pmatrix} -1 & 3 \\ 2 & 0 \end{pmatrix}.$$

Karakteristinen yhtälö $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$:

$$\begin{vmatrix} -1-\lambda & 3 \\ 2 & -\lambda \end{vmatrix} = \lambda(1+\lambda) - 6 = \lambda^2 + \lambda - 6 = 0.$$

Ominaisarvot ovat $\lambda_{1,2} = (-1 \pm \sqrt{1+4\cdot 6})/2 = (-1 \pm 5)/2$, eli $\lambda_1 = -3$ ja $\lambda_2 = 2$. Ratkaistaan ominaisvektorit: (i)

$$(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{v} = \begin{pmatrix} -1+3 & 3 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

joten on oltava $2v_1 + 3v_2 = 0$, esim. $\mathbf{v} = (3, -2)$ on ominaisvektori.
(ii)

$$(\mathbf{A} - \lambda_2\mathbf{I})\mathbf{v} = \begin{pmatrix} -1-2 & 3 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

joten $-3v_1 + 3v_1 = 0$, esim. $\mathbf{v} = (1, 1)$. Saadaan muunnosmatriisiksi

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

ja sen käänteismatriisi on

$$\mathbf{P}^{-1} = \begin{pmatrix} 0.2 & -0.2 \\ 0.4 & 0.6 \end{pmatrix}.$$

Saadaan

$$\begin{aligned} \mathbf{P}\mathbf{\Lambda}^k\mathbf{P}^{-1} &= \begin{pmatrix} 3 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} (-3)^k & 0 \\ 0 & 2^k \end{pmatrix}\begin{pmatrix} 0.2 & -0.2 \\ 0.4 & 0.6 \end{pmatrix} \\ &= \begin{pmatrix} 3(-3)^k & 2^k \\ (-2)(-3)^k & 2^k \end{pmatrix}\begin{pmatrix} 0.2 & -0.2 \\ 0.4 & 0.6 \end{pmatrix} \\ &= \begin{pmatrix} 0.2\cdot(-1)^k\cdot 3^{k+1} + 0.4\cdot 2^k & 0.2\cdot(-3)^{k+1} + 0.6\cdot 2^k \\ 0.2\cdot(-1)^{k+1}\cdot 3^k + 0.4\cdot 2^k & 0.2\cdot(-1)^{k+2}\cdot 3^k + 0.6\cdot 2^k \end{pmatrix}. \end{aligned}$$

10.3 Kompleksiset ominaisarvot

Diagonalisointiin tarvitaan n ominaisarvoa, mutta mitä jos kaikkia näitä ei löydy reaalilukujen joukosta? Vastaus on, että matriisi voidaan silti diagonalisoida käyttämällä kompleksisia ominaisarvoja ja toimimalla aivan samoin kuin reaalisten ominaisarvojen tapauksessa. Huomaa, että osa ominaisarvoista on kompleksisia silloin, kun osa karakteristisen yhtälön juurista on kompleksilukuja. Kerrataan aluksi muutama asia kompleksiluvuista.

10.3.1 Kompleksiluvut

Määäritelmä 10.3.1. kompleksiluku on muotoa $a+ib$ oleva luku missä i:lle pätee $i^2 = -1$ ja a ja b ovat reaalilukuja

Lukua a sanotaan reaaliosaksi ja lukua b imaginaariosaksi, i on imaginaariyksikkö.

Määäritelmä 10.3.2. Luvun $z = a+ib$ kompleksikonjugaatiksi sanotaan lukua $\bar{z} = a-ib$.

Lause 10.3.1. *Jos kompleksiluku z on polynomin juuri niin myös $\bar{z}$ on polynomin juuri.*

Huom. n:nnen asteen polynomilla on enintään n juurta, jos $n = 2$ ja yksi juurista on kompleksiluku niin toisenkin on oltava Huom. vaikka kompleksiluvut määritellään toisen asteen yhtälön ratkaisujen kautta (i toteuttaa yhtälön $x^2 = -1$), niillä saadaan kaikkien korkeampaakin astetta olevien polynomiyhtälöiden kaikki ratkaisut.

Listataan muutamia tärkeitä kompleksilukujen ominaisuuksia:

1. summa $z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2)$ $(z_k = a_k + b_k i,\ k = 1, 2)$
2. tulo $z_1 z_2 = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 - a_2 b_1)$
3. DeMoivren kaava $z^k = r^k(\cos k\theta + i \sin k\theta)$, missä $r = \sqrt{a^2 + b^2}$ ja kulma θ määräytyy yhtälöstä $\cos\theta = a/r$.

Tarkastellaan nyt matriisia, jolla on kompleksisia ominaisarvoja. Jos jokin ominaisarvo on kompleksinen, niin luonnollisesti ominaisvektorista tulee kompleksivektori. Kompleksisille ominaisarvoille ja vektoreille pätee seuraava tulos.

Lause 10.3.2. *Jos $z = \alpha + i\beta$ on ominaisarvo $n \times n$ matriisille $\mathbf{A}$ ja $\mathbf{u} + i\mathbf{v}$, $(\mathbf{u}, \mathbf{v} \in \mathbb{R}^n)$ on tätä vastaava ominaisvektori, niin tällöin $\bar{z}$ on myös $\mathbf{A}$:n ominaisarvo ja $\mathbf{u} - i\mathbf{v}$ on tätä vastaava ominaisvektori*

Pohditaan vielä lyhyesti miten lineaarikuvaus käyttäytyy, jos sillä on kompleksisia ominaisarvoja. Määritelmän mukaan reaalinen ominaisvektori on sellainen, että kuvaukset tämän vektorin suuntaan pysyvät saman suuntaisina. Siten voimme päätellä, että jos kaikki ominaisarvot ovat kompleksisia, niin matriisi kääntää kaikkia suuntia.

10.3.2 Differenssiyhtälön ratkaisu

Olkoon $\mathbf{A}$ on 2×2 matriisi, jonka kaikki alkiot ovat reaalilukuja. Oletetaan, myös, että matriisilla on kompleksiset ominaisarvot $z = \alpha \pm i\beta$ ja niitä vastaavat ominaisvektorit $\mathbf{u} \pm i\mathbf{v}$. Tällöin differenssiyhtälön $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$ yleinen ratkaisu on

$$\mathbf{z}_k = 2r^k[(c_1 \cos k\theta - c_2 \sin k\theta)u - (c_2 \cos k\theta + c_1 \sin k\theta)v],$$

missä c_1 ja c_2 ovat reaalilukuja, jotka määräytyvät alkuarvosta $\mathbf{z}_0$. Kompleksiset ominaisarvot johtavat siihen että ratkaisu “värähtelee” sini- ja kosinitermien johdosta.

Edellä oleva tulos voidaan perustella seuraavasti.

1. Diagonalisoidaan matriisi kuten reaalisten ominaisarvojen tapauksessa ja ratkaistaan muuttujan vaihdoksen avulla differenssiyhtälö:
$$\mathbf{P} = \begin{pmatrix}\mathbf{u}+i\mathbf{v} & \mathbf{u}-i\mathbf{v}\end{pmatrix}$$
2. Kompleksilukujen laskusääntöjen (summa ja tulo) avulla havaitaan, että $\mathbf{A}$ diagonalisoituu ja
$$\mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix}\alpha+i\beta & 0\\ 0 & \alpha-i\beta\end{pmatrix}$$
3. Merkitään $\mathbf{Z} = (X, Y)$ (uudet muuttujat), differenssiyhtälö $\mathbf{Z}$:lle
$$\begin{aligned} X_{k+1} &= (\alpha+i\beta)X_k \\ Y_{k+1} &= (\alpha-i\beta)Y_k \end{aligned}$$
Yleinen ratkaisu on muotoa $X_k = k_1(\alpha+i\beta)^k$, $Y_k = k_2(\alpha-i\beta)^k$, missä k_1 ja k_2 ovat kompleksilukuja
4. Kun tehdään muunnos takaisin alkuperäisiin muuttujiin, tuleeko niistä reaalisia? Kyllä kun k_1 ja k_2 ovat toistensa kompleksikonjugaatit
5. Yleinen ratkaisu (edellä) saadaan kompleksilukujen laskusäännöillä ja DeMoivren kaavalla.

Esimerkki 10.3.1. Olkoon differenssiyhtälön kerroinmatriisi
$$\mathbf{A} = \begin{pmatrix}1 & 1\\ -9 & 1\end{pmatrix}.$$
Kirjoitetaan karakteristinen yhtälö: $\lambda^2 - 2\lambda + 10 = 0$. Tämän ratkaisut ovat
$$\lambda_{1,2} = (2 \pm \sqrt{4-4\cdot 10})/2 = (2 \pm \sqrt{-36}))/2 = (2 \pm 6\sqrt{-1})/2 = 1 \pm i3.$$
Etsitään ominaisvektorit:
$$(\mathbf{A} - \lambda_1\mathbf{I}) = \begin{pmatrix}-3i & 1\\ -9 & -3i\end{pmatrix}\begin{pmatrix}v_1\\ v_2\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix},$$
eli $-3iv_1 + v_2 = 0$, esim. $\mathbf{v} = (1, 3i)$ kelpaa. Toinen ominaisvektori on $(1, -3i)$. Saadaan, että $r = \sqrt{1^2+3^2} = \sqrt{10}$, $\theta = \arccos(1/\sqrt{10}) \approx 1.249$. Kirjoitetaan differenssiyhtälön ratkaisu. Merkitään $\mathbf{z} = (x, y)$
$$\begin{pmatrix}x_k\\ y_k\end{pmatrix} = \sqrt{10}^k\left[(c_1\cos k\theta - c_2\sin k\theta)\begin{pmatrix}1\\ 0\end{pmatrix} - (c_2\cos k\theta + c_1\sin k\theta)\begin{pmatrix}0\\ 3\end{pmatrix}\right].$$

10.3.3 Stabiilisuus

Kompleksisten ominaisarvojen eräs tärkeä sovellutus on stabiilisuusanalyysi. Määritellään ensin mitä tarkoitetaan stabiilisuudella.

Mååäritelmä 10.3.3. Lineaarisen differenssiyhtälön $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$ tasapaino $z = \mathbf{0}$ on (globaalisti asymptoottisesti) stabiili, jos differenssiyhtälön ratkaisut suppenevat kohti vektoria $\mathbf{0}$ kaikilla alkuarvoilla $\mathbf{z}_0$.

Huomaa, että origo on tasapaino, eli toteuttaa yhtälön $\mathbf{Ax} = \mathbf{x}$. Jos alkuarvoksi asetetaan origo niin siinä pysytään kaikkina ajanhetkinä. Stabiilisuus tarkoittaa sitä, että jos lähdetään origon ulkopuolelta, niin jan kuluessa päädytään mielivaltaisen lähelle origoa, eli ratkaisu suppenee kohti origoa. Tämä tarkoittaa, että $\|\mathbf{z}_k\| \to 0$ kun $k \to \infty$.

Reaalisten ominaisarvojen tapauksessa tasapaino on stabiili kun ominaisarvot ovat itseisarvoltaan pienempiä kuin 1. Yleisesti ominaisarvojen on oltava kompleksitasossa yksikköympyrän sisäpuolella.

Lause 10.3.3. *Origo on stabiili tasapaino differenssiyhtälölle $\mathbf{z}_{k+1} = \mathbf{A}\mathbf{z}_k$ jos mille hyvänsä $\mathbf{A}$:n ominaisarvolle $a + bi$ pätee $r = \sqrt{a^2 + b^2} < 1$.*

10.4 Markov-prosessit

Eräs sovellus differenssiyhtälöille on Markov-prosessit. Ne ovat ajassa sattumanvaraisesti eteneviä ilmiöitä, jotka alkavat joka hetki uudestaan. Toisin sanottuna niiden tulevaisuuteen vaikuttaa vain nykyhetken tilanne eikä se, miten siihen on tultu.

Määäritelmä 10.4.1. Oletetaan äärellinen määrä tiloja $i = 1, \ldots, n$. Stokastinen prosessi on sääntö, joka määrittää todennäköisyydet sille, että periodilla $k+1$ ollaan tilassa i. Jos tilojen todennäköisyydet riippuvat vain edellisen periodin tilasta, stokastista prosessia sanotaan Markov-prosessiksi.

Esimerkki 10.4.1. Luokitellaan perheet kotipaikan mukaan: (1) kaupunkilaisiin, (2) esikaupunkien asukkaisiin ja (3) maalaisiin. Eri luokkia voidaan pitää tiloina $1, 2, 3$, jos perheen kotipaikan tyyppi muuttuu, niin tila muuttuu.

Tilansiirtotodennäköisyydet; merkitään $x^i(k)$ todennäköisyys että jaksolla k ollaan tilassa i. Siirtymätodennäköisyydet m_{ij} = todennäköisyys että jaksolla $k+1$ tila on i kun jaksolla k oltiin tilassa j. Nämä voidaan koota tilansiirtomatriisiksi (Markov-matriisiksi):

$$\mathbf{M} = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix}.$$

Esimerkki 10.4.2. Voidaan ajatella että todennäköisyys $x^i(k)$ kuvaa osuutta populaatiosta, joka asuu ($i = 1$) kaupungissa, ($i = 2$) esikapungissa, ($i = 3$) maalla jaksolla k. Esim. m_{1j} = todennäköisyys, jolla pysytään kaupunkilaisena ($j = 1$) tai tullaan kaupunkilaiseksi ($j = 2$ tai $j = 3$), voidaan tulkita osuuksina populaatiosta, jotka pysyvät paikoillaan ja muuttavat. Oletetaan

$$\begin{pmatrix} 0.75 & 0.02 & 0.1 \\ 0.2 & 0.9 & 0.2 \\ 0.05 & 0.08 & 0.7 \end{pmatrix}.$$

Markov-prosessi voidaan esittää differenssiyhtälönä. Todennäköisyys, että jaksolla $k+1$ ollaan tilassa i: $x^i(k+1)=$ (tod. näk. että siirrytään tilasta 1 tilaan i) $\times$ (tod. näk. että lähdetään tilasta 1) $+\ldots+$ (tod. näk että siirrytään tilasta n tilaan i) $\times$ (tod. näk. että lähdetään tilasta n) $=\sum_j m_{ij}x^j(k)$. Tilansiirtomatriisin avulla $\mathbf{x}(k+1)=\mathbf{M}\mathbf{x}(k)$, eli

$$\begin{pmatrix} x^1(k+1) \\ \vdots \\ x^n(k+1) \end{pmatrix} = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix} \begin{pmatrix} x^1(k) \\ \vdots \\ x^n(k) \end{pmatrix}.$$

Esimerkki 10.4.3. Tarkastellaan seuraaksi esimerkkiä Markov-prosessista ja lasketaan Markov-matriisin ominaisarvot. Olkoon Markov matriisi

$$\mathbf{M} = \begin{pmatrix} 0.9 & 0.4 \\ 0.1 & 0.6 \end{pmatrix}.$$

Ratkaistaan ominaisarvot:

$$\det(\mathbf{M}-\lambda\mathbf{I}) = (0.9-\lambda)(0.6-\lambda) - 0.1\cdot 0.4 = \lambda^2 - (0.9+0.6)\lambda + 0.9\cdot 0.6 - 0.04 = 0.$$

Ominaisarvoiksi saadaan $\lambda_{1,2} = (1.5 \pm \sqrt{1.5^2 - 4\cdot 0.5})/2 = (1.5\pm 0.5)/2 = 1, 0.5$. Vastaavat ominaisvektorit ovat $\mathbf{v}_1 = (4,1)$ ja $\mathbf{v}_2 = (1,-1)$. Differenssiyhtälön yleinen ratkaisu on siten

$$\begin{aligned} x^1(k+1) &= c_1\cdot 4\cdot 1^k + c_2\cdot 1\cdot 0.5^k \\ x^2(k+1) &= c_1\cdot 1\cdot 1^k + c_2\cdot(-1)\cdot 0.5^k. \end{aligned}$$

Voimme tässä esimerkissä tehdä seuraavat havainnot:

1. toinen ominaisarvo on yksi ja toinen pienempi kuin yksi
2. kun $k\to\infty$, niin raja-arvona on $c_1(4,1)$ (ominaisarvoa 1 vastaavan ominaisvektorin suunta), koska raja-arvon on oltava todennäköisyysjakauma on $4c_1 + c_1 = 1$, eli $c_1 = 1/5$.

Edellä tehdyt havainnot pätevät yleisesti Markov-matriiseille. Kun $\mathbf{M}$ on Markov-matriisi, jonka kaikki komponentit ovat nollasta poikkeavia (tai $\mathbf{M}^k$:n alkiot nollasta poikkeavia jollakin k), niin yksi ominaisarvoista on yksi ja loput välillä $(-1,1)$. Tällöin prosessi suppenee kohti ominaisarvoa 1 vastaavan ominaisvektorin määrittämää rajajakaumaa. Tämä rajajakauma on differenssiyhtälön stabiili tasapaino (kaikilla todennäköisyysjakaumilla alkutiloina). Kokeile eo. esimerkissä mitä on

$$M\begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}.$$

Esimerkki 10.4.4. Tarkastellaan prosessia

$$\begin{pmatrix} x^1(k+1) \\ x^2(k+1) \\ x^3(k+1) \end{pmatrix} = \begin{pmatrix} 0.75 & 0.02 & 0.1 \\ 0.2 & 0.9 & 0.2 \\ 0.05 & 0.08 & 0.7 \end{pmatrix} \begin{pmatrix} x^1(k) \\ x^2(k) \\ x^3(k) \end{pmatrix}.$$

Tiedetään, että yksi kerroinmatriisin ominaisarvoista on 1, joten hyödyntämällä tuloksia matriisin ominaisarvojen yhteydestä matriisin jälkeen ja determinanttiin saadaan $1+\lambda_1+\lambda_2 = 0.75+0.9+0.7 = 2.35$ ja $\lambda_1\lambda_2 = 0.455$. Tästä voidaan ratkaista kaksi tuntematonta ominaisarvoa: 0.7 ja 0.65. Ominaisvektorit ovat

$$\begin{pmatrix} 2/15 \\ 10/15 \\ 3/15 \end{pmatrix}, \begin{pmatrix} 8 \\ -5 \\ -3 \end{pmatrix} \text{ ja } \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Yleiseksi ratkaisuksi saadaan

$$\begin{pmatrix} x^1(k+1) \\ x^2(k+1) \\ x^3(k+1) \end{pmatrix} = \begin{pmatrix} 2/25 \\ 10/15 \\ 3/15 \end{pmatrix} + c_2 \begin{pmatrix} 8 \\ -5 \\ -3 \end{pmatrix} (0.7)^k + c_3 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} (0.65)^k.$$

Luku 11

Neliömuodot ja matriisin definiittisyys

Jokaisen neliömatriisin $\mathbf{A} \in \mathbb{R}^{n \times n}$ avulla voidaan määritellä funktio

$$\mathrm{q} : \mathbb{R}^n \to \mathbb{R}, \quad \mathrm{q}(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,$$

jota kutsutaan matriisin $\mathbf{A}$ määrittämäksi *neliömuodoksi*. Huomaa, että neliömuodolla tarkoitetaan vektoreiden $\mathbf{x}$ ja $\mathbf{Ax}$ välistä sisätuloa. Neliömuotojen avulla voidaan esittää neliöllisiä funktioita, eli funktioita, joissa muuttujat (ja eri muuttujien tulot) ovat korkeintaan toista astetta. Yleisesti usean muuttujat toisen asteen funktiot ovat muotoa

$$f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c.$$

On helppo huomata, että aukikirjoitettuna neliöllisen funktion lausekkeesta voi tulla varsin pitkä, joten matriisimerkintä lyhentää neliömuodon merkitsemistä huomattavasti. Neliöllisistä funktioista on hyvä huomata, myös se, että niiden esitysmatriisi $\mathbf{A}$ voidaan valita symmetriseksi, sillä $x_i x_j = x_j x_i$.

Esimerkki 11.0.5. Kirjoitetaan auki miltä näyttää

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Jälkimmäinen tulo on

$$\begin{pmatrix} x_1 + 3x_2 + 2x_3 \\ 3x_1 + 2x_2 + x_3 \\ 2x_1 + x_2 + 2x_3 \end{pmatrix}.$$

Kertomalla tämän vektorin $\mathbf{x}$ kanssa saamme

$$x_1^2 + 3x_1x_2 + 2x_1x_3 + 3x_1x_2 + 2x_2^2 + x_2x_3 + 2x_1x_3 + x_2x_3 + 2x_3^2 =$$
$$x_1^2 + 2x_2^2 + 2x_3^2 + 6x_1x_2 + 4x_1x_3 + 2x_2x_3.$$

Esimerkki 11.0.6. Muodostetaan matriisiesitys funktiolle

$$f(\mathbf{x}) = 6 + 7x_1 + 3x_2 + 2x_1^2 + 5x_1x_2 + 4x_2^2.$$

Vakiotermi on $c = 6$, lineaarisen termin kerroinvektori on $\mathbf{b} = (7, 3)$ ja neliöllisen termin kerroinmatriisi on

$$\mathbf{A} = \begin{pmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 4 \end{pmatrix}.$$

Neliömuotoa $q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}$ sanotaan *positiivisesti definiitiksi*, jos $q(\mathbf{x}) > 0$ kaikilla $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq 0$. Neliömuotoa sanotaan *positiivisesti semidefiniitiksi* (joskus myös ei-negatiivisesti definiitiksi), jos

$$\mathrm{q}(\mathbf{x}) = \begin{cases} \geq 0 & \text{kaikilla } \mathbf{x} \in \mathbb{R}^n, \\ = 0 & \text{jollakin vektorilla } \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq 0. \end{cases}$$

Symmetristä matriisia $\mathbf{A}$ sanotaan nyt *positiivisesti definiitiksi* (semidefiniitiksi), mikäli sitä vastaava neliömuoto $\mathbf{x}^\top \mathbf{A}\mathbf{x}$ on positiivisesti definiitti (semidefiniitti). Vastaavalla tavalla symmetristä matriisia $\mathbf{A}$ sanotaan *negatiivisesti definiitiksi* (semidefiniitiksi), mikäli matriisi $-\mathbf{A}$ on positiivisesti definiitti (semidefiniitti). Tämä tarkoittaa sitä, että neliömuoto $-\mathbf{x}^\top \mathbf{A}\mathbf{x}$ on positiivisesti definiitti (semidefiniitti). Mikäli matriisin määrittämä neliömuoto saa positiivisia ja negatiivisia arvoja, tällöin matriisin sanotaan olevan *indefiniitti.*

Esimerkki 11.0.7. Matriisin

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 3 \\ 0 & 3 & -2 \end{pmatrix}$$

määrittämä neliömuoto on

$$\mathrm{q}(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} = (x_1, x_2, x_3) \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 3 \\ 0 & 3 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= (x_2, x_1 + 3x_3, 3x_2 - 2x_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1x_2 + x_2(x_1 + 3x_3) + x_3(3x_2 - 2x_3)$$

$$= -2x_3^2 + 2x_1x_2 + 6x_2x_3.$$

Kyseessä on indefiniitti matriisi, koska esimerkiksi q[(0,0,1)]=−2 ja q[(1,1,0)]=2.

Esimerkki 11.0.8. Matriisi

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

on positiivisesti semidefiniitti, sillä

$$q(\mathbf{x}) = \mathbf{x}^{\top}\mathbf{B}\mathbf{x} = (x_1, x_2, x_3)\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= (x_1 - x_2, -x_1 + x_2, 5x_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1(x_1 - x_2) + x_2(-x_1 + x_2) + 5x_3^2$$

$$= x_1^2 + x_2^2 - 2x_1x_2 + 5x_3^2 = (x_1 - x_2)^2 + 5x_3^2 \geq 0 \text{ kaikilla } \mathbf{x} \in \mathbb{R}^3.$$

Huomaa, että matriisi ei ole positiivisesti definiitti, sillä esimerkiksi q[(1,1,0)]=0.

11.1 Definiittisyyden määrittäminen

Matriisin definiittisyyden tarkistamiseksi on olemassa tuloksia, joiden avulla tehtävä helpottuu huomattavasti. Esimerkeissä 11.0.7 ja 11.0.8 tarkasteltu suoraviivainen tapa on yleensä liian vaikea. Tarkastellaan jälleen matriisia

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Tämän matriisin *johtavilla pääalideterminateilla* tarkoitetaan determinantteja

$$\mathbf{A}(i) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1i} \\ a_{21} & a_{22} & \dots & a_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ii} \end{vmatrix}, \quad i = 1, 2, \dots, n.$$

Johtavat pääalideterminantit saadaan ottamalla matriisin nk pääminoreista determinantit. Ensimmäinen pääminori on se matriisi josta on poistettu $n-1$ viimeistä riviä ja saraketta, eli a_{11}, toinen pääminori on matriisi, josta on poistettu $n-2$ vimeistä riviä ja saraketta.

Voidaan osoittaa, että matriisi $\mathbf{A}$ on positiividefiniitti jos sen kaikki pääalideterminantit ovat positiivisia, eli

$$\mathbf{A}(i) > 0 \text{ kaikilla } i = 1, 2, \dots, n.$$

Voidaan myös osoittaa, että matriisi $\mathbf{A}$ on negatiividefiniitti jos sen pääalideterminantin merkki vaihtelee seuraavan säännön mukaisesti

$$\mathbf{A}(1) < 0, \ \mathbf{A}(2) > 0, \ \mathbf{A}(3) < 0, \ \dots, (-1)^n\mathbf{A}(n) > 0.$$

Esimerkki 11.1.1. Osoitetaan, että matriisi

$$\mathbf{A} = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

on negatiivisesti definiitti. Tämän matriisin pääalideterminantit ovat

$$\mathbf{A}(1) = -2 < 0,$$

$$\mathbf{A}(2) = \begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix} = (-2) \cdot (-1) - 1 \cdot 1 = 2 - 1 = 1 > 0,$$

$$\mathbf{A}(3) = \begin{vmatrix} -2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -3 \end{vmatrix} = -3 \begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix} = -3 \cdot 1 = -3 < 0,$$

joten kyseessä on todellakin negatiivisesti definiitti matriisi.

Esimerkki 11.1.2. Millä vakion $u \in \mathbb{R}$ arvoilla matriisi

$$\mathbf{A} = \begin{pmatrix} 1 & u & u \\ u & 1 & u \\ u & u & 1 \end{pmatrix}$$

on positiivisesti definiitti. Kaikkien pääalideterminanttien tulee olla positiivisia, joten saadaan ehdot

$$1 > 0,$$

$$\begin{vmatrix} 1 & u \\ u & 1 \end{vmatrix} = 1 - u^2 > 0 \Leftrightarrow -1 < u < 1,$$

$$\begin{vmatrix} 1 & u & u \\ u & 1 & u \\ u & u & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & u \\ u & 1 \end{vmatrix} - u \begin{vmatrix} u & u \\ u & 1 \end{vmatrix} + u \begin{vmatrix} u & 1 \\ u & u \end{vmatrix}$$

$$= (1 - u^2) - u(u - u^2) + u(u^2 - u) = 2u^3 - 3u^2 + 1 = 2(u-1)^2(u + \frac{1}{2}) > 0$$

$$\Leftrightarrow u + \frac{1}{2} > 0 \text{ ja } u \neq 1 \Leftrightarrow u > -\frac{1}{2} \text{ ja } u \neq 1.$$

Koska kaikkien näiden ehtojen täytyy olla voimassa yhtä aikaa, on matriisi positiivisesti definiitti täsmälleen silloin kun

$$-\frac{1}{2} < u < 1.$$

Myös ominaisarvojen avulla voi tutkia matriisin definiittisyyttä.

Lause 11.1.1. *Symmetrinen matriisi* $\mathbf{A}$ *on positiivisesti semidefiniitti (neg.sem.def) jos ja vain jos* $\mathbf{A}$*:n ominaisarvot ovat suurempia tai yhtä suuria kuin nolla* (≤ 0 *neg.sem.def. matriisille). Matriisi on positiivisesti definiitti (neg.def.) jos ja vain jos sen ominaisarvot ovat positiiviset* (< 0 *neg.def). Matriisi on indefiniitti jos ja vain jos sillä on positiivinen ja negatiivinen ominaisarvo.*

Esimerkki 11.1.3. Ominaisarvojen perusteella esimerkeissä 9.1.2 ja 9.3.2 olevat matriisit ovat positiividefiniittijä. Esimerkissä 9.1.4 esitetty matriisi on indefinitiitti.

Semidefiniittisyyttä voi tutkia myös samaan tapaan kuin definiittisyyttä minorien avulla. Merkitään matriisilla $M^n_{\{i_1,2,\ldots,i_n\}}$ sitä matriisia, joka saadaan $\mathbf{A}$:sta jättämällä siitä rivit ja sarakkeet $i_1, i_2, \ldots, i_n$, eli

$$M^n_{\{i_1,2,\ldots,i_n\}} = \begin{pmatrix} a_{i_1i_1} & a_{i_1i_2} & \cdots & a_{i_1i_n} \\ \vdots & & \ddots & \vdots \\ a_{i_ni_1} & a_{i_ni_2} & \cdots & a_{i_ni_n} \end{pmatrix}.$$

Merkitään vielä

$$A^n_{\{i_1,i_2,\ldots,i_n\}} = \det(M^n_{\{i_1,i_2,\ldots,i_n\}}).$$

Matriisi $\mathbf{A}$ on positiivisemidefiniitti, jos $A^n_{\{i_1,i_2,\ldots,i_n\}} \geq 0$ kaikilla n ja $\{i_1, \ldots, i_n\}$. Vastaavasti matriisi on negatiivisemidefiniitti, jos $A^n_{\{i_1,i_2,\ldots,i_n\}} \geq 0$ parillisilla n ja kaikilla $\{i_1, \ldots, i_n\}$ ja parittomilla n pätee $A^n_{\{i_1,i_2,\ldots,i_n\}} \leq 0$ kaikilla $\{i_1, \ldots, i_n\}$. Huom. jos matriisi ei ole negatiisisemidefiniitti tai positiivisemidefiniitti, niin se on idefiniitti.

Esimerkki 11.1.4. Tarkastellaan matriisia

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

Tämä matriisi on indefiniitti, koska $A_1 = 2$ ja

$$A^2_{\{1,3\}} = \det \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = -1.$$

11.2 Cholesky-hajotelma

Olkoon $\mathbf{A}$ symmetrinen ja positiivisesti definiitti matriisi. Tällöin matriisi $\mathbf{A}$ voidaan esittää muodossa

$$\mathbf{A} = \mathbf{L}\mathbf{L}^\top,$$

missä $\mathbf{L}$ on jokin alakolmiomatriisi ($\mathbf{L}^\top$ on tällöin tietenkin yläkolmiomatriisi). Tällaista hajotelmaa kutsutaan matriisin *Cholesky hajotelmaksi.* Kirjoittamalla tämä hajotelma eksplisiittisesti muodossa

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & \dots & l_{n1} \\ 0 & l_{22} & \dots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & l_{nn} \end{pmatrix},$$

voidaan osoittaa, että matriisin $\mathbf{L}$ alkiot pystytään ratkaisemaan rekursiokaavalla

$$\begin{cases} l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right), \quad i > j, \\ \\ l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}. \end{cases}$$

Tässä rekursiokaavassa matriisin $\mathbf{L}$ alkiot ratkaistaan sarakkeittain ja ylhäältä alaspäin. Algoritmia kutsutaan *Cholesky-Banachiewicz algoritmiksi.* Esimerkki valaisee algoritmin toimintaa parhaiten.

Esimerkki 11.2.1. Esimerkin 11.1.2 perusteella tiedetään, että matriisi

$$\mathbf{A} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$$

on positiivisesti definiitti. Etsitään tämän matriisin Cholesky hajotelma. Edellä annettua rekursiokaavaa käyttämällä saadaan

$$\begin{aligned} l_{11} &= \sqrt{a_{11}} = \sqrt{1} = 1, \\ l_{21} &= \frac{a_{21}}{l_{11}} = \frac{1}{2}, \\ l_{31} &= \frac{a_{31}}{l_{11}} = \frac{1}{2}, \\ l_{22} &= \sqrt{a_{22} - l_{21}^2} = \sqrt{1 - \frac{1}{4}} = \frac{\sqrt{3}}{2}, \\ l_{32} &= \frac{1}{l_{22}} \left(a_{32} - l_{31} l_{21} \right) = \frac{2}{\sqrt{3}} \left(\frac{1}{2} - \frac{1}{4} \right) = \frac{1}{2\sqrt{3}}, \\ l_{33} &= \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = \sqrt{1 - \frac{1}{4} - \frac{1}{12}} = \frac{\sqrt{2}}{\sqrt{3}}. \end{aligned}$$

Näin ollen saadaan

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix},$$

mikä voidaan tarkastaa helposti suoraan laskemalla.

Luku 12

Harjoitustehtäviä

Vektorit

1. Muodosta sen tason yhtälö, jonka normaali on $(1, -1, 2)$ ja joka kulkee pisteen $(1, 2, 3)$ kautta.

2. Olkoon tason suora annettu parametrisoidussa muodossa $x_1(t) = 1 + 2t$, $x_2(t) = -1 + 2t$. Piirrä kuva suorasta. Muodosta suoralle esitys muodossa $ax_1 + bx_2 = c$, eli hae sopivat a, b ja c.

3. Oletetaan, että tuotannossa käytetään panoksena kolmea tuotetta; tuotteet 1, 2 ja 3. Yksikkökustannukset kullekin panokselle (eo. järjestyksessä) ovat 7, 3 ja 5. Määritä kaikkien niiden se panosvektoreiden muodostama hypertaso, jolla panosten yhteiskustannus on 100.

4. Kuluttajalla on käytössä pääomaa 100 yksikköä. Oletetaan, että kuluttaja voi allokoida tämän varallisuudeen kahteen hyödykkeeseen, määrät x ja y. Hyödykkeiden yksikköhinnat ovat 1 ja 1/2. Piirrä kuva kuluttajan budjettijoukosta: ne ei-negativiset määrät x, y, joiden arvo ei ylitä käytettävää pääomaa. Mitkä kolme hypertasoa rajaavat kyseisen joukon (anna niiden yhtälöt)?

5. Alla on esitetty neljä joukkoa vektoreita. Mitkä joukot ovat lineaarisesti riippumattomia ja mitkä eivät ole (perustele)?

   (a) $(2, 1)$ ja $(1, 2)$

   (b) $(2, 1)$ ja $(-4, -2)$

   (c) $(1, 1, 0)$ ja $(0, 1, 1)$

   (d) $(1, 1, 0)$, $(0, 1, 1)$, $(1, 0, 1)$

6. Ovatko vektorit $(1,1,1,5)$, $(1,2,1,2)$ ja $(1,3,1,-1)$ lineaarisesti riippumattomia?

7. Osoita, että vektorit $\mathbf{u}_1 = (1,1,1), \mathbf{u}_2 = (0,1,1)$ ja $\mathbf{u}_3 = (0,0,1)$ muodostavat $\mathbb{R}^3$:n kannan. Esitä vektori $\mathbf{u} = (-1,2,4)$ näiden kantavektoreiden lineaarikombinaationa.

8. Olkoot $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$. Osoita, että $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.

9. Laske avaruuden $\mathbb{R}^4$ vektoreiden $(3,-2,0,-1)$ ja $(2,6,9,3)$ välisen kulman kosini ja määritä välisen kulman suuruus.

10. Etsi suora, joka sisältää suorien $3x - 4y + 7 = 0$ ja $6x - 2y - 3 = 0$ leikkauspisteen ja täyttää seuraavat ehdot
    (a) kulkee origon kautta
    (b) on yhdensuuntainen suoran $3x - 2y + 7 = 0$ kanssa
    (c) on kohtisuorassa suoraa $3x - 2y + 7 = 0$ vastaan.

Matriisialgebraa

11. Muodosta seuraavat matriisit:
    (a) rivit muodostuvat vektoreista $(2,1)$ ja $(1,2)$
    (b) edellisen kohdan matrisii, josta on vähennetty 2×2-yksikkömatriisi
    (c) sarakkeet muodostuvat vektoreista $(1,0,3,0)$, $(2,0,0,1)$ ja $(0,0,3,2)$
    (d) edellisen kohdan matriisin transpoosi
    (e) matriisi, jossa diagonaalilla on luvut $(3,4,5,6)$ vähennettynä vastaavan kokoisella neliömatriisilla, jossa on pelkkiä ykkösiä.

12. Olkoot
$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 1 \\ 1 & -2 & 0 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 3 & 2 \end{pmatrix} \text{ ja } \mathbf{C} = \begin{pmatrix} 3 & 0 \\ 1 & 2 \\ 1 & -1 \end{pmatrix}.$$
Laske (jos laskuoperaatio on määritelty)
a) $\mathbf{A} + 2\mathbf{B}$, b) $\mathbf{A} + \mathbf{C}$, c) $\mathbf{AB}$,
d) $\mathbf{BA}$, e) $\mathbf{AC}$, f) $\mathbf{CA}$,
g) $\mathbf{ACA}$, h) $\mathbf{BAC}$ ja i) $\mathbf{A}^\top\mathbf{B}$.

13. Olkoon
$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix}.$$
a) Minkä kokoisia ovat matriisit $\mathbf{x}^\top\mathbf{x}$, $\mathbf{x}\mathbf{x}^\top$, $\mathbf{A}\mathbf{x}$, $\mathbf{x}^\top\mathbf{A}\mathbf{x}$ ja $\mathbf{x}\mathbf{x}^\top\mathbf{A}$?
b) Laske $\mathbf{x}^\top\mathbf{x}$, $\mathbf{x}\mathbf{x}^\top$.
c) Laske $\mathbf{A}\mathbf{x}$, $\mathbf{x}^\top\mathbf{A}\mathbf{x}$.

14. Minkä kokoisia matriisien $\mathbf{A}$, $\mathbf{B}$ ja $\mathbf{C}$ on oltava, jotta lausekkeet

$$(\mathbf{A}+\mathbf{B})\mathbf{C} \quad \text{ja} \quad \mathbf{AC}+\mathbf{BC}$$

ovat määriteltyjä? Osoita, että tällöin pätee $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$.

15. Muodosta matriisille

$$\mathbf{A} = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

ositus $\mathbf{A} = (\mathbf{B} \;\; \mathbf{C})$, missä $\mathbf{B}$ on 2×2 matriisi ja $\mathbf{C}$ on 2×1 matriisi.

16. Määritä esitysmatriisit seuraaville lineaarikuvauksille, eli hae kuvausta F vastaava matriisi $\mathbf{A}$ siten, että $F(\mathbf{x}) = \mathbf{Ax}$.
a) $F(x_1, x_2) = (x_1, 2x_2)$, b) $F(x_1, x_2, x_3) = (x_2, 0, 3x_1 + x_3)$.

17. Määritä matriisin $\mathbf{C} = \mathbf{AB}$ käänteismatriisi, kun $\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ ja $\mathbf{B} = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$.

Yhtälöryhmät ja matriisit

18. Mitkä seuraavista yhtälöryhmistä ovat lineaarisia, kun tuntemattomia (eli endogeenisia) muuttujia ovat x, y, z?

(a) $xz = ay$,

(b) $ax + y = bz + c^2$

(c)
$$\begin{aligned} \ln(a)y + e^b z &= x \\ \sqrt{a} + x &= z \end{aligned}$$

(d)
$$\begin{aligned} zx + ay + 10c + d &= 0 \\ cx + by - 5d &= 0 \end{aligned}$$

19. Muodosta yhtälöryhmälle

$$\begin{aligned} zx + ay + 10c + d &= 0 \\ cx + by - 5d &= 0 \end{aligned}$$

esitys matriisimuodossa, eli muodossa $\mathbf{Ax} = \mathbf{b}$, kun tuntemattomia (endogeenisia) ovat

(a) x ja y,

(b) c, z ja y,

(c) a ja d.

20. Kirjoita auki yhtälöryhmä $\mathbf{Ax} = \mathbf{b}$, kun

$$\mathbf{A} = \begin{pmatrix} 2 & 7 \\ 0 & 3 \end{pmatrix} \text{ ja } \mathbf{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}.$$

Totea, että vektori $(-11/2, 2)$ on yhtälöryhmän ratkaisu. Voiko ratkaisuja olla muita?

21. Piirrä kuva alla olevien yhtälön määräämistä suorista ja ratkaise yhtälöpari suoralla eliminointimenettelyllä.

$$\begin{aligned} -4x + 2y &= 2 \\ 3x + 9y &= 9 \end{aligned}$$

22. Oletetaan, että yrityksen tuotto on 50000 euroa ennen veroja. Yritys lahjoittaa Punaiselle ristille summan x, joka vastaa 10% sen verojen jälkeisestä tuotosta. Tämän lahjoituksen voi kuitenkin vähentää verotettavasta tuotosta. Verotus suoritetaan siten, että ensin yritys maksaa 5% verotettavasta tuotostaan valtiollista veroa (määrä y), jäljelle jäävästä summasta maksetaan 40% osavaltion veroa (määrä z). Muodosta edellä kuvatusta tilanteesta kolmen lineaarisen yhtälön ryhmä muuttujille x, y ja z.

23. Olkoon $\mathbf{A}$ kääntyvä neliömatriisi. Osoita, että $\mathbf{A}^\top$ on kääntyvä ja $(\mathbf{A}^\top)^{-1} = (\mathbf{A}^{-1})^\top$.

Vektorialiavaruudet

24. Osoita, että joukko $\{\mathbf{x} \in \mathbb{R}^3 : x_1 + 2x_2 + x_3 = 0 \text{ ja } x_3 = -3x_1 - 2x_2\}$ on $\mathbb{R}^3$:n aliavaruus.

25. Onko joukko $\{\mathbf{x} \in \mathbb{R}^2 : |x_1| \geq |x_2|\}$ tason aliavaruus?

Determinantti

26. Laske matriisin

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

determinantti kehittämällä se a) 2. rivin ja b) 3. sarakkeen suhteen.

27. Laske alla olevan matriisin determinantti:

$$\begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \end{pmatrix}.$$

28. Perustele determinanttia käyttämällä että vektorit $(\pi, 0, e)$, $(0, 1, 75)$ ja $(2012, 0, 49)$ muodostavat $\mathbb{R}^3$:n kannan.

29. Olkoon

$$\mathbf{A} = \begin{pmatrix} 13 & 8 & 6 \\ -13 & -8 & -4 \\ 8 & 5 & 5 \end{pmatrix}.$$

Laske matriisin $\det[(-5\mathbf{A}\mathbf{A}^\top)^\top]$ determinantin laskusääntöjä käyttäen.

Matriisin aste

30. Laske seuraavien matriisien asteet.

$$\text{a)} \begin{pmatrix} 2 & -4 & 2 \\ -1 & 2 & 1 \end{pmatrix} \quad \text{b)} \begin{pmatrix} 1 & 6 & -7 & 3 \\ 1 & 9 & -6 & 4 \\ 1 & 3 & -8 & 4 \end{pmatrix}.$$

31. Miten yhtälöparin

$$\begin{aligned} y - ax &= b \\ x + y &= 1 \end{aligned}$$

ratkaisujen määrä riippuu vakioista a ja b?

32. Seuraavat viisi matriisia esittävät lineaarisen yhtälöryhmän $\mathbf{Ax} = \mathbf{b}$ kerroinmatriisia. Mitä voit sanoa kussakin tapauksessa ratkaisujen määrästä yhtälöryhmälle, kun a) $\mathbf{b} = \mathbf{0}$ ja b) $\mathbf{b}$ on mielivaltainen dimensioiltaan sopiva vektori?

$$\text{i)} \begin{pmatrix} 1 & 4 \\ 2 & 1 \end{pmatrix} \qquad \text{ii)} \begin{pmatrix} 1 & 4 & 0 \\ 2 & 1 & 0 \end{pmatrix} \quad \text{iii)} \begin{pmatrix} 2 & 1 \\ 1 & 4 \\ 0 & 3 \end{pmatrix}$$

$$\text{iv)} \begin{pmatrix} 1 & 4 & 3 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{v)} \begin{pmatrix} 1 & 4 & 3 \\ 2 & 1 & 0 \\ 0 & 7 & 6 \end{pmatrix}$$

33. Tarkastellaan lineaarista IS-LM mallia:

$$\begin{aligned} sY + ar &= I^* + G \\ mY - hr &= M_s - M^*. \end{aligned}$$

(a) Määritä yhtälöparin kerroinmatriisi, kun tuntemattomia (endogeenisia) muuttujia ovat Y ja r.

(b) Muodosta yhtälö, jonka toteutuessa edellä saamasi kerroinmatriisi on täyttä astetta.

34. Voidaanko alla olevasta yhtälöparista ratkaista vektori (x_1, x_2) muuttujan x_3 suhteen, entä (x_1, x_3) muuttujan x_2 suhteen?

$$\begin{aligned} -x_1 + 3x_2 - x_3 &= 0 \\ 2x_1 - 6x_2 + 3x_3 &= 0. \end{aligned}$$

35. Muuta matriisi

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

rivi-porrasmuotoon elementaarisilla rivioperaatioilla. Mikä on matriisin aste?

36. Tarkastellaan mallia, jossa on kaksi rahoitusinstrumenttia ja kolme mahdollista maailman tilaa. Jos tila 1 realisoittuu niin instrumentin 1 tuottoon $R_{11} = 1$ ja vastaavasti kaksi muuta maailman tilaa tuottavat intrumentille tuoton $R_{21} = 2$ ja $R_{31} = 3$. Vastaavasti toiselle instrumentille tuotot ovat: $R_{12} = 3$, $R_{22} = 2$ ja $R_{32} = 1$.
a) Muodosta tuotto matriisi $\mathbf{R}$, jonka rivit vastaavat maailman tiloja ja sarakkeet instrumentteja.
Tila i on vakuutettavissa, jos löytyy portfolio $x = (x_1, x_2)$ siten, että portfolion tuotto on positiivinen jos tila i toteutuu ja nolla muissa tiloissa.
b) Muotoile yllä oleva ehto matemaattisesti.
Voidaan osoittaa, että tila i on vakuutettavissa, jos yhtälöllä $\mathbf{Rx} = \mathbf{e}_i$ on ratkaisu.
c) Olkoon $\mathbf{y}$ pystyvektori, jonka komponentit ovat a, b ja c. Muuta Gaussin eliminaatiolla, eli rivioperaatioilla laajennettu matriisi $\begin{pmatrix} \mathbf{R} & \vdots & \mathbf{y} \end{pmatrix}$ riviporrasmuotoon (yläkolmiomatriisiksi) ja perustele sen avulla, että mikään tiloista ei ole vakuutettavissa. Huom. esimerkiksi $y = (1, 0, 0)$ vastaa vektoria $\mathbf{e}_1$, jolloin Gaussin eliminaatio vastaa yllä olevan yhtälön ratkaisemista.

Käänteismatriisin laskeminen

37. Määritä matriisi **C**, kun tiedetään, että

$$\mathbf{C}^{-1} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

38. Laske seuraavien matriisien käänteismatriisit Gaussin-Jordanin algoritmilla.

a) $\begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 2 & 0 & 4 \end{pmatrix}$ c) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$

39. Määritä alideterminanttien avulla (Cramerin sääntö käänteismatriiseille) alla olevan matriisin käänteismatriisi.

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

40. Laske seuraavien matriisien käänteismatriisit Cramerin säännöllä.

a) $\begin{pmatrix} 4 & 3 \\ 1 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 1 & 0 & 8 \end{pmatrix}$ c) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Yhtälöryhmien ratkaiseminen

41. Ratkaise seuraavat yhtälöryhmät Cramerin säännöllä.

(a)

$$\begin{aligned} 5x_1 + x_2 &= 3 \\ 2x_1 - x_2 &= 4 \end{aligned}$$

(b)

$$\begin{aligned} 2x_1 - 3x_2 &= 2 \\ 4x_1 - 6x_2 + x_3 &= 7 \\ x_1 + 10x_2 &= 1 \end{aligned}$$

42. Edellä olevan tehtävän a-kohdan kerroinmatriisin käänteismatriisi ja ratkaise yhtälöryhmä sen avulla.

43. Tarkastellaan mallia:

$$\begin{pmatrix} s & a \\ m & -h \end{pmatrix} \begin{pmatrix} Y \\ r \end{pmatrix} = \begin{pmatrix} I^* + G \\ M_s - M^* \end{pmatrix}.$$

Tuntemattomia (endogeenisia) muuttujia ovat siis Y ja r. Ratkaise tuntemattomat Cramerin säännöllä olettaen, että yhtälöryhmän kerroinmatriisi on kääntyvä.

44. Ratkaise Gaussin eliminointimenetelmällä yhtälöryhmä

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 8 \\ 3x_1 + x_2 + x_3 &= 3 \\ 4x_1 + 5x_2 - x_3 &= 12. \end{aligned}$$

45. Ratkaise Gaussin-Jordanin eliminointimenetelmällä yhtälöryhmä

$$\begin{aligned} 2x_1 - x_2 + x_4 &= -1 \\ x_1 + x_2 + 2x_3 + x_4 &= 5 \\ x_1 + 2x_3 + 2x_4 &= 4. \end{aligned}$$

46. Luentomonisteessa esitetään log-lineaarinen kysyntä-tarjontamalli, jossa

$$Q_i^d = K_i P_1^{\alpha_{i1}} P_2^{\alpha_{i2}} Y^{\beta_i} \quad i = 1, 2.$$

Yllä Q_i^d on tuotteen i kysytty määrä. P_i on tuotteen i hinta, Y on tulotaso. K_i kokoaa kaikki analyysin ulkopuolistet muuttujat (eksogeeniset muuttujat), jotka vaikuttavat kysyntään. Tarjontafunktiot vat $Q_i^s = M_i P_1^{\gamma_i}$, $i = 1, 2$ Tasapainossa $Q_i^d = Q_i^s$, $i = 1, 2$. Olkoon tuote 1 vientisektorin tuote ja oletetaan, että laman torjumiseksi hallitus tukee vientiä prosentuaalisella tuella korottamalla markkinahinnan P_1 arvoon $(1 + s)P_1$. Tuella ei ole vaikutusta kysyntään.

(a) Kirjoita malli lineaariseen muotoon ottamalla logaritmit kaikista yhtälöistä ja kirjoita malli kahden yhtälön lineaarisena mallina.

(b) Ratkaise malli esimerkiksi Cramerin säännöllä ja analysoi tuen vaikutusta markkinahintoihin kummankin tuotteen osalta.

Ominaisarvot ja -vektorit

47. Totea, että vektori $(1, 0, -1)$ on ominaisvektori matriisille
$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 5 & 0 \\ 3 & 0 & 2 \end{pmatrix}.$$

48. Osoita muodostamatta polynomiyhtälöä, että matriisin
$$\begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}$$
ominaisvektorit ovat $(1, 2)$ ja $(-2, 1)$. Mitkä ovat vastaavat ominaisarvot?

49. Jos 2×2 matriisin $\mathbf{A}$ ominaisarvot ovat λ_1 ja λ_2 sekä $\mathbf{v}_1$ ja $\mathbf{v}_2$ näitä vastaavat ominaisvektorit, niin voidaan määritellä
$$e^{\mathbf{A}} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v_2} \end{pmatrix}^{-1} \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 & \mathbf{v_2} \end{pmatrix}.$$
Laske $e^{\mathbf{A}}$ edellisen tehtävän matriisille.

50. Laske ominaisarvot ja ominaisvektorit matriiseille
a) $\begin{pmatrix} -1 & 3 \\ -2 & 4 \end{pmatrix}$ b) $\begin{pmatrix} 0 & 0 & -2 \\ 0 & 7 & 0 \\ 1 & 0 & -3 \end{pmatrix}$.

51. (a) Laske matriisien
$$\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \quad \text{ja } \mathbf{C} = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$$
ominaisarvot.
(b) Laske a)-kohdan matriisille $\mathbf{B}$ ominaisvektorit, diagonalisoi $\mathbf{B}$ ja esitä diagonalisoinnissa tarvittava kuvaus/matriisi $\mathbf{P}$, jolla $\mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P}$, missä $\mathbf{\Lambda}$ on $\mathbf{B}$:n diagonaalimuunnos.
(c) Esitä yleinen ratkaisu differenssiyhtälölle $\mathbf{z}_{k+1} = \mathbf{C}\mathbf{z}_k$,

52. Yhdysvaltojen työttömyystilastoista on estimoitu Markov-prosessi suhdannevaihteluille (Hamilton 2005). Mallissa on kolme tilaa: normaali kasvu (tila 1), lievä taantuma (tila 2) ja taantuma (tila 3). Markov-matriisi on
$$M = \begin{pmatrix} 0.97 & 0.15 & 0 \\ 0.03 & 0.77 & 0.5 \\ 0 & 0.08 & 0.5 \end{pmatrix},$$
missä rivin i ja sarakkeen j luku kertoo todennäköisyyden, jolla seuraavan kuukauden tila on i, kun edellisen kuukauden tila oli j. Yksi matriisin ominaisarvoista on 1, totea tämä ja laske vastaava ominaisvektori. Kaksi muuta ominaisarvoa ovat 0.848 ja 0.392. Voit päätellä ominaisarvoa 1 vastaavasta ominaisvektorista jotain prosessin käyttäytymisestä ilman, että lasket muita ominaisvektoreita — mitä voit päätellä?

Definiittisyys

53. Määrää seuraavien matriisien definiittisyys:

a) $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ b) $\begin{pmatrix} -1 & 1 & 5 \\ 1 & -2 & 2 \\ 5 & 2 & -5 \end{pmatrix}$ c) $\begin{pmatrix} 2 & 3 \\ 3 & -1 \end{pmatrix}$

54. Kirjoita neliömuotojen

a) $16x^2 + 9y^2 + 24xy$, b) xy, c) $x^2 + y^2 + z^2 + xy + yz + xz$

esitysmatriisit, hae niiden ominaisarvot ja määrittele definiittisyys.

55. Yritys valmistaa kahta tuotetta. Olkoon x ja y näiden tuotteiden määrät. Tuotteiden yksikköhinnat ovat 4 ja 7, jolloin tuotto määristä x ja y on $4x + 7y$. Tuotantokustannus määrille x ja y on $x + 2y + 3x^2 + 4y^2 - 2xy$.

a) Muotoile yrityksen voittofunktio muodossa $\mathbf{a} \cdot \mathbf{x} - \mathbf{x}^\top \mathbf{A}\mathbf{x}$, missä $\mathbf{x} = (x, y)$ ja $\mathbf{a} \in \mathbb{R}^2$ on sopiva kerroinvektori ja $\mathbf{A}$ on sopivan kokoinen esitysmatriisi kustannusfunktiossa esiintyvälle neliömuodolle.

b) Määritä $\mathbf{A}$:n definiittisyys.