Polynomimatriisit

Antti Lindberg

Matematiikan pro gradu

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Syksy 2014

Tiivistelmä: Antti Lindberg, *Polynomimatriisit.* Matematiikan pro gradu -tutkielma, 50 sivua, Jyväskylän Yliopiston matematiikan ja tilastotieteen laitos, syksy 2014.

Tähän tulee tiivistelmä.

Avainsanat: Polynomi, matriisi, Frobeniuksen muoto

Kiitokset

Tähän tulee kiitokset.

Sisältö

Johdanto

Tähän tulee johdanto.

LUKU 1

Polynomit

1. Kuntakertoiminen polynomirengas

Polynomit ovat tässä kirjoitelmassa erittäin keskeisessä roolissa. Käydään lyhyesti läpi jatkossa käytettävät polynomeihin liittyvät merkinnät ja määritelmät. Rajoitutaan tarkastelemaan ainoastaan kuntakertoimisia polynomeja. Sovellusten kannalta tärkeimpiä kerroinkuntia lienevät reaali- ja kompleksilukujen kunnat. Kuitenkin tämän tutkielman keskeisin teoria toimii yleiselle karakteristikan nolla kunnalle aivan vastaavasti kuin reaali- ja kompleksiluvuillekin. Sen vuoksi on oikeastaan turha rajoittaa tarkastelua vain reaali- ja kompleksiluvuille. Sen sijaan jatkossa käytetään polynomien kerroinkuntana yleistä karakteristikan nolla kuntaa. Myöhemmin selvinnee myös syitä sille, miksi karakteristikan halutaan olevan nolla. Nollasta poikkevan karakteristikan kunnat vaatisivat usein erillisen tarkastelun, ja kaikki tämän tutkielman tulokset eivät niille olisi edes voimassa. Tästä syystä tällaiset kunnat jätetään tässä tutkielmassa tarkastelun ulkopuolelle. Jatkossa merkintä $\mathbb{K}$ tarkoittaa aina kuntaa, jonka karakteristika on nolla. Koska minkään muunlaisia kuntia ei tässä käsitellä, ei tätä vaatimusta enää myöhemmin erikseen toisteta.

Tarkasti määriteltynä $\mathbb{K}$-kertoiminen *polynomi* p on jono $(a_k)_{k=0}^{\infty}$, missä on vain äärellisen monta nollasta eroavaa alkiota $a_k \in \mathbb{K}$. Jos $a_{n+k} = 0$ kaikilla $k = 1, 2, \ldots$, käytetään polynomille p muodollista merkintää

$$p = \sum_{k=0}^{n} a_k x^k.$$

Jos tässä esityksessä $a_n \neq 0$, sitä kutsutaan polynomin p *johtavaksi kertoimeksi*. Tällöin $n \in \mathbb{N} \cup \{0\}$ on polynomin p *aste*, merkitään $\deg(p) := n$. Polynomin sanotaan olevan perusmuotoinen, jos sen johtava kerroin $a_n = 1$. Jos $a_k = 0$ kaikilla $k = 0, 1, 2, \ldots$, kyseessä on nollapolynomi $p = 0$. Nollapolynomin asteeksi voidaan sopia $-\infty$, jolle pätevät seuraavat:

(1) $-\infty < a$ kaikille $a \in \mathbb{N} \cup \{0\}$
(2) $a + (-\infty) = -\infty + a = -\infty$ kaikille $a \in \mathbb{N} \cup \{0\}$
(3) $-\infty + (-\infty) = -\infty$

Jos $\deg(p) \leq 0$, p on *vakiopolynomi*. Vakiopolynomit voidaan samaistaa kunnan $\mathbb{K}$ kanssa. Usein polynomeja merkitään isoilla kirjaimilla, mutta tässä tutkielmassa käytetään kuitenkin aina pieniä kirjaimia. Tällä pyritään välttämään sekaannuksia, sillä isoilla kirjaimilla merkitään jatkossa matriiseja, joiden kertoimet voivat olla polynomeja. Muuttujasymbolina käytetään pientä x-kirjainta. $\mathbb{K}$-kertoimisten polynomien joukolle käytetään merkintää $\mathbb{K}[x]$.

Polynomeille määritellään yhteen- ja kertolasku asettamalla

$$\left(\sum_{k=0}^{n} a_k x^k\right) + \left(\sum_{k=0}^{m} b_k x^k\right) = \left(\sum_{k=0}^{\max\{n,m\}} (a_k + b_k) x^k\right) \text{ ja}$$

$$\left(\sum_{k=0}^{n} a_k x^k\right) \cdot \left(\sum_{k=0}^{m} b_k x^k\right) = \sum_{k=0}^{n+m} \left(\sum_{i+j=k} a_i b_j\right) x^k.$$

Joukko $\mathbb{K}[x]$ varustettuna näillä laskutoimituksilla muodostaa kommutatiivisen renkaan, jota kutsutaan $\mathbb{K}$-kertoimiseksi polynomirenkaaksi. Polynomien tulon ja summan asteille pätevät seuraavat laskusäännöt

$$\deg(p_1 \cdots p_n) = \deg(p_1) + \cdots + \deg(p_n) \text{ ja}$$
$$\deg(p_1 + \cdots + p_n) = \max\{\deg(p_1), \ldots, \deg(p_n)\}.$$

Jokainen polynomi $p = \sum_{k=0}^{n} a_k x^k \in \mathbb{K}[x]$ määrittelee *polynomikuvauksen* $\overline{p} : \mathbb{K} \to \mathbb{K}$, jolle

$$\overline{p}(t) = \sum_{k=0}^{n} a_k t^k.$$

Polynomikuvausten joukko $\mathcal{P}(\mathbb{K})$ voidaan varustaa pisteittäisellä yhteen- ja kertolaskulla, jolloin saadaan polynomikuvausten rengas. Yleisen kunnan tapauksessa voi kuitenkin käydä niin, että eri polynomit määrittävät saman polynomikuvauksen. Esimerkiksi, jos kuntana on $\mathbb{Z}_2$, polynomit $x + 1 \in \mathbb{Z}_2[x]$ ja $x^2 + 1 \in \mathbb{Z}_2[x]$ määrittävät saman polynomikuvauksen, vaikka ne ovat eri polynomeja. Toisaalta kunnan $\mathbb{Z}_2$ karakteristika on 2, ja edellä todettiin, että tällaisia kuntia ei tässä tutkielmassa käytetä. Käytettäessä kuntia, joiden karakteristika on nolla, voidaan huoletta samaistaa polynomit ja polynomikuvaukset. Karakteristikan nolluus ei kuitenkaan ole välttämätön ehto, sillä pelkkä äärettömyys riittää.

LAUSE 1.0.1. *Olkoon K ääretön kunta. Tällöin renkaat $K[x]$ ja $\mathcal{P}(\mathbb{K})$ ovat isomorfiset.*

TODISTUS. Kuvaus $\Phi : K[x] \to \mathcal{P}(\mathbb{K})$, jolle $\Phi(p) = \overline{p}$, on helppo todeta rengashomomorfismiksi. Tällöin riittää soitaa, että $\mathrm{Ker}(\Phi) = \{0\}$. Rengashomomorfismin perusominaisuuksien mukaan pätee $\Phi(0) = 0$. Osoitetaan, että $\mathrm{Ker}(\Phi) \subset \{0\}$. Oletetaan, että polynomi $p \in K[x]$ määrittelee nollakuvauksen. Tällöin riittää osoittaa, että $p = 0$. Koska polynomin p määrittelemä kuvaus on nollakuvaus, jokainen alkio $t \in K$ on sen juuri. Jos p ei olisi nollapolynomi, sillä voisi olla enintään asteensa $\deg(p)$ verran juuria [ks. esim. Lang, Theorem 4 s. 121]. Tämä on kuitenkin mahdotonta kunnan K äärettömyyden vuoksi. Näin ollen p on nollapolynomi, mikä todistaa väitteen. □

Samaistetaan jatkossa polynomit ja vastaavat polynomikuvaukset, jolloin voidaan puhua vain polynomeista. Tällöin käytetään merkintää p tai $p(x)$. Joskus on kuitenkin oltava tarkkana siitä, puhutaanko polynomista vai polynomikuvauksen arvosta. Sekaannuksien välttämiseksi varataan muuttujasymboli x vain polynomeille. Jos halutaan merkitä polynomin arvoa jossakin kohdassa, käytetään kirjainta t tai tarvittaessa jotain muuta kirjainta. Merkintä $p(x)$ tarkoittaa siis polynomia, jonka muuttujasymbolina on x, ja merkintä $p(t)$ taas polynomin p arvoa kohdassa $t \in \mathbb{K}$.

2. Jaollisuudesta

Sanotaan, että polynomi $q \in \mathbb{K}[x]\backslash\{0\}$ jakaa polynomin $p \in \mathbb{K}[x]$, jos on olemassa sellainen $r \in \mathbb{K}[x]$, jolle $p = rq$. Tällöin käytetään myös merkintää $q|p$. Sanotaan, että polynomi p on jaoton, jos se ei ole jaollinen millään sellaisella polynomilla $r \in \mathbb{K}[x]\backslash\{0\}$, jolle $1 \leq \deg(r) < \deg(p)$. Polynomeille on voimassa ns. jakoyhtälö.

LAUSE 2.0.2 (Polynomien jakoyhtälö). *Olkoot $p, q \in \mathbb{K}[x]$ ja $q \neq 0$. Tällöin on olemassa yksikäsitteiset polynomit $r, s \in \mathbb{K}[x]$, joille*

$$p = rq + s \text{ ja } deg(s) < deg(q).$$

TODISTUS. Sivuutetaan, ks. Metsänkylä & Näätänen, s. 133. $\square$

LAUSE 2.0.3. *Olkoon $q \in \mathbb{K}[x]$. Tällöin on olemassa jaottomat perusmuotoiset polynomit $p_1, \dots, p_n$ ja $a \in \mathbb{K}$, joille*

$$q = a \prod_{j=1}^{n} p_j.$$

TODISTUS. Jos $q = 0$, väite on selvä. Oletetaan, että $q \neq 0$. Todistus tehdään induktiolla polynomin q asteen n suhteen. Kun $n = 0$, polynomi q on vakiopolynomi, jolloin $q = a \cdot 1$ jollekin $a \in \mathbb{K}$. Olkoon $n \geq 1$, ja tehdään induktio-oletus, että väite pätee kaikille enintään astetta $n - 1$ oleville polynomeille.

Jos q on jaoton, voidaan kirjoittaa $q = a(a^{-1}q)$, missä $a \in \mathbb{K}$ on polynomin q johtava kerroin. Voidaan siis olettaa, että q ei ole jaoton. Tällöin on olemassa polynomit $r, s \in \mathbb{K}[x]\backslash\{0\}$, jolle $1 \leq \deg(r) < \deg(q)$, $1 \leq \deg(s) < \deg(q)$ ja $q = rs$. Tällöin induktio-oletuksen nojalla on olemassa jaottomat perusmuotoiset polynomit $p_1, \dots, p_k$ ja $p_{k+1}, \dots, p_n$ sekä $b, c \in \mathbb{K}$, joille $r = bp_1 \cdots p_k$ ja $s = cp_{k+1} \cdots p_n$. Väite seuraa, kun valitaan $a = bc$. $\square$

LAUSE 2.0.4. *Olkoon $q \in \mathbb{K}[x]\backslash\{0\}$. Olkoot lisäksi polynomit $p_1, \dots, p_n$ ja skalaari $a \in \mathbb{K}$, kuten lauseessa 2.0.3. Tällöin, jos $\mathbb{K} = \mathbb{C}$, $deg(p_j) \leq 1$ kaikille $j = 1, \dots, n$. Jos taas $\mathbb{K} = \mathbb{R}$, $deg(p_j) \leq 2$ kaikille $j = 1, \dots, n$.*

TODISTUS. Jos yleisesti polynomilla $p \in \mathbb{K}[x]$ on juuri $c \in \mathbb{K}$, se on jaollinen polynomilla $x - c$. Algebran peruslauseen nojalla jokaisella sellaisella kompleksikertoimisilla polynomilla, joka ei ole vakiopolynomi, on juuri. Väitteen ensimmäinen osa seuraa suoraan tästä.

Tapauksessa $\mathbb{K} = \mathbb{R}$ polynomit p_j voidaan ajatella myös kompleksikertoimisiksi tulkinnalla $\mathbb{R} \subset \mathbb{C}$. Jos polynomilla p_j on kompleksisenakin polynomina vain reaalisia juuria, se voi olla enintään astetta yksi, sillä muutoin se olisi jaollinen renkaassa $\mathbb{R}[x]$. Jos $c \in \mathbb{C}$ on polynomin p_j juuri, on sitä myös sen kompleksikonjugaatti $\bar{c}$. Siten polynomin p_j aidosti kompleksiset juuret esiintyvät aina konjugaattipareina. Olkoon polynomilla p_j aidosti kompleksinen juuri $c \in \mathbb{C}\backslash\mathbb{R}$. Tällöin se on jaollinen renkaassa $\mathbb{C}[x]$ polynomilla

$$(x - c)(x - \bar{c}) = x^2 - 2\mathrm{Re}(c)x + |c|^2.$$

Toisaalta $x^2 - 2\mathrm{Re}(c)x + |c|^2 \in \mathbb{R}[x]$. Tällöin p_j on jaollinen tällä polynomilla myös renkaassa $\mathbb{R}[x]$, joten sen aste voi olla enintään kaksi. $\square$

Olkoot polynomit $p_1, \dots, p_k \in \mathbb{K}[x]$, ja oletetaan, että $p_j \neq 0$ jollakin $j \in \{1, \dots, k\}$. Sanotaan, että polynomi $0 \neq s \in \mathbb{K}[x]$ on polynomien $p_1, \dots, p_k \in \mathbb{K}[x]$ *suurin yhteinen tekijä*, merkitään $\mathrm{syt}\{p_1, \dots, p_k\} = s$, mikäli seuraavat ehdot pätevät:

(1) Polynomi s on perusmuotoinen ja jakaa jokaisen polynomin p_j, kun $j = 1 \dots, k$.
(2) Ehdoista $0 \neq r \in \mathbb{K}[x]$ ja $r|p_j$ kaikilla $j = 1, \dots, k$ seuraa, että $r|s$.

Tässä vaaditaan suurimmalta yhteiseltä tekijältä perusmuotoisuus, jotta siitä saadaan yksikäsitteinen.

LAUSE 2.0.5. *Olkoot $q_1, \dots, q_k \in \mathbb{K}[x]$, ja oletetaan, että $q_j \neq 0$ jollakin j. Tällöin on olemassa yksikäsitteinen $s = syt\{q_1, \dots, q_k\}$.*

TODISTUS. Osoitetaan ensin, että suurin yhteinen tekijä on yksikäsitteinen, jos se on olemassa. Olkoon polynomit s ja s' polynomien $\{q_1, \dots, q_k\}$ suurimpia yhteisiä tekijöitä. Tällöin suoraan määritelmästä seuraa, että $s|s'$ ja $s'|s$. Koska molemmat s ja s' ovat perusmuotoisia, tästä seuraa, että $s = s'$.

Olemassaolotodistus tehdään induktiolla polynomien lukumäärän k suhteen. Tapaus $k = 1$ on triviaali, joten oletetaan, että $k = 2$. Jos toinen polynomeista q_1 tai q_2 on nollapolynomi, väite on selvä. Jos esimerkiksi $q_1 = 0$, $\mathrm{syt}\{q_1, q_2\} = a^{-1}q_2$, missä $a \in \mathbb{K}\backslash\{0\}$ on polynomin q_2 johtava kerroin. Oletetaan siis, että $q_1 \neq 0 \neq q_2$. Voidaan olettaa, että $\deg(q_1) \geq \deg(q_2)$. Tehdään induktio polynomin q_2 asteen n_2 suhteen. Kun $n_2 = 0$, suoraan määritelmästä nähdään, että $1 = \mathrm{syt}\{q_1, q_2\}$. Oletetaan seuraavaksi, että $\mathrm{syt}\{q_1, q_2\}$ on olemassa, kun $\deg(q_2) \leq n - 1$, missä $n \geq 1$. Olkoon $\deg(q_2) = n$. Jakoyhtälön nojalla on olemassa polynomit $r, s \in \mathbb{K}[x]$, joille

$$q_1 = rq_2 + u \text{ ja } \deg(u) < \deg(q_2). \tag{1.1}$$

Induktio-oletuksen nojalla on olemassa $\mathrm{syt}\{q_2, u\} =: s$. Riittää osoittaa, että s on myös polynomien q_1 ja q_2 suurin yhteinen tekijä. Määritelmänsä mukaan s on perusmuotoinen ja $s|q_2$. Lisäksi $s|u$, joten yhtälön 1.1 nojalla $s|q_1$. Oletetaan, että $w \in \mathbb{K}[x]\backslash\{0\}$ jakaa polynomit q_1 ja q_2. Tällöin yhtälön 1.1 nojalla $w|u$. Silloin suurimman yhteisen tekijän määritelmän mukaan $w|s$, joten $s = \mathrm{syt}\{q_1, q_2\}$.

Tehdään seuraavaksi induktio-oletus, että jollakin $k \geq 3$ suurin yhteinen tekijä on olemassa, kun polynomeja on enintään $k - 1$ kappaletta. Olkoot $q_1, \dots, q_k \in \mathbb{K}[x]$ ja $q_j \neq 0$ jollakin j. Jos $q_j = 0$ kaikilla $j = 1, \dots, k-1$, nähdään suoraan määritelmästä, että $\mathrm{syt}\{q_1, \dots, q_k\} = a^{-1}q_k$, missä $a \in \mathbb{K}\backslash\{0\}$ on polynomin q_k johtava kerroin. Voidaan siis olettaa, että $q_j \neq 0$ jollakin $j \in \{1, \dots, k-1\}$. Induktio-oletuksen nojalla on olemassa $s' := \mathrm{syt}\{q_1, \dots, q_{k-1}\}$ ja $s = \mathrm{syt}\{s', q_k\}$. Riittää osoittaa, että $s = \mathrm{syt}\{q_1, \dots, q_k\}$. Määritelmänsä mukaan s on perusmuotoinen ja jakaa polynomit q_k ja s'. Toisaalta s' jakaa määritelmänsä mukaan polynomit $q_1, \dots, q_{k-1}$, jolloin myös s jakaa polynomit $q_1, \dots, q_{k-1}$. Oletetaan, että $w \in \mathbb{K}[x]\backslash\{0\}$ jakaa polynomit $q_1, \dots, q_k$. Tällöin suurimman yhteisen tekijän määritelmän mukaan $w|s'$, jolloin määritelmästä nähdään edelleen, että myös $w|s$. Tämä tarkoittaa, että $s = \mathrm{syt}\{q_1, \dots, q_k\}$. □

LUKU 2

Johdatus polynomimatriiseihin

1. Polynomimatriisi ja matriisipolynomi

Tässä tutkielmassa käytetään nimitystä *polynomimatriisi* matriisille, jonka alkiot ovat polynomirenkaasta $\mathbb{K}[x]$. Kirjallisuudessa käytetään myös nimitystä λ-matriisi, jolloin kerroinpolynomien muuttujasymbolina on yleensä λ. Rajoitutaan tarkastelamaan ainoastaan neliömatriiseja. Olkoon $n \geq 1$. Tällöin $\mathbb{K}[x]$-kertoimisten $n \times n$-polynomimatriisien joukolle käytetään merkintää $\mathrm{Mat}_n(\mathbb{K}[x])$. Jatkossa $n \times n$-matriisia voidaan kutsua myös yksinkertaisesti kokoa n olevaksi matriisiksi. Kokoa 1 olevat polynomimatriisit tulkitaan aina polynomeiksi. Vastaavasti kokoa 1 olevat kuntakertoimiset matriisit tulkitaan skalaareiksi. Useimmat myöhemmin esitettävät tulokset ovat selviä 1×1-matriiseille, joten tapaus $n = 1$ sivuutetaan yleensä erikseen mainitsematta.

Polynomimatriisien summa ja tulo määritellään aivan samoin kuin esimerkiksi reaalikertoimisille matriiseillekin. Näillä laskutoimituksilla varustettuna joukko $\mathrm{Mat}_n(\mathbb{K}[x])$ muodostaa renkaan. Polynomimatriisin $A = A(x)$ *aste* $\deg(A(x))$ määritellään asettamalla

$$\deg(A(x)) := \max\{\deg([A(x)]_{ij})\}.$$

Jos $\deg(A) \leq 0$, kyseessä on *vakiomatriisi*. Ne voidaan samaistaa $\mathbb{K}$-kertoimisten matriisien kanssa.

Merkittävä ero kuntakertoimisiin matriiseihin on se, että nollasta eroavilla kertoimilla ei välttämättä ole käänteisalkioita. Tämä rajoittaa tietysti kääntyvien matriisien joukkoa. Vaikka ainoastaan nollasta eroavilla vakiopolynomeilla on käänteisalkiot, käänteismatriisi voi kuitenkin olla myös monilla sellaisilla polynomimatriiseilla, joiden kaikki kerroinpolynomit eivät ole vakiopolynomeja. Esimerkiksi

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

Polynomimatriiseille voidaan määritellä myös determinantti aivan samoilla kehityssäännöillä, joilla reaali- ja kompleksikertoimisten matriisien determinanttikin määritellään. Tällöin determinantti on aina polynomi. Polynomimatriisin sanotaan olevan *singulaarinen*, jos sen determinantti on nollapolynomi.

Polynomimatriisi $A \in \mathrm{Mat}_n(\mathbb{K}[x])$ määrittelee luonnollisesti kuvauksen $A : \mathbb{K} \to \mathrm{Mat}_n(\mathbb{K})$, $t \mapsto A(t)$, missä $[A(t)]_{ij} = [A]_{ij}(t)$. Koska polynomit ja vastaavat polynomikuvaukset voidaan lauseen 1.0.1 nojalla samaistaa keskenään, voidaan myös yleistäen polynomimatriisi A samaistaa kuvaukseen $t \mapsto A(t)$. Tämän tiedon avulla voidaan todistaa seuraavat tulokset polynomimatriiseille yleistämällä vastaavat kuntakeroimisia matriiseja koskevat tulokset.

LAUSE 1.0.6. *Olkoot $A, B \in Mat_n(\mathbb{K}[x])$ ja oletetaan, että $AB = I$. Tällöin myös $BA = I$.*

TODISTUS. Käytetään hyväksi tietoa, että vastaava tulos pätee $\mathbb{K}$-kertoimisille matriiseille. Oletuksen mukaan polynomimatriiseja A ja B vastaaville kuvauksille pätee $A(t)B(t) = I$ kaikille $t \in \mathbb{K}$. Silloin pätee myös pisteittäin $B(t)A(t) = I$ kaikille $t \in \mathbb{K}$. Kuvaukset $t \mapsto B(t)A(t) = BA(t)$ ja $t \mapsto I$ ovat samoja ja näin ollen myös polynomimatriisit BA ja I. □

LAUSE 1.0.7. *Olkoon $A \in Mat_n(\mathbb{K}[x])$ lohkoyläkolmiomuotoa*

$$A = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_k \end{bmatrix}.$$

Tällöin

$$det(A) = \prod_{j=1}^{k} det(A_j). \tag{2.1}$$

TODISTUS. Oletetaan tunnetuksi, että tulos pätee $\mathbb{K}$-kertoimisille matriiseille [ks. Broida & Williamson, Theorem 4.14 s. 205]. Tällöin jokaiselle $t \in \mathbb{K}$ pätee

$$\det(A)(t) = \det(A(t)) = \det(A_1(t)) \cdots \det(A_k(t)) = \det(A_1)(t) \cdots \det(A_k)(t),$$

joten yhtälön (2.1) polynomien kuvapisteet ovat samat kaikille $t \in \mathbb{K}$. Tällöin myös itse polynomit ovat samat. □

LAUSE 1.0.8 (Cauchyn ja Binet'n kaava). *Olkoot $A, B \in Mat_n(\mathbb{K}[x])$. Olkoot $I, J \subset \{1, \ldots, n\}$, joille $\#I = \#J = k \in \{1, \ldots, n\}$, ja $(AB)_{I,J}$ vastaava alimatriisi. Tällöin*

$$det((AB)_{I,J}) = \sum_{\substack{K \subset \{1,\ldots,n\} \\ \#K = k}} det(A_{I,K}) det(B_{K,J}). \tag{2.2}$$

Erityisesti

$$det(AB) = det(A) det(B).$$

TODISTUS. Lauseen todistus $\mathbb{K}$-kertoimisille matriiseille sivuutetaan [ks. esim. Broida & Williamson s.212-214]. Yleistys polynomimatriiseille menee vastaavasti kuin lauseessa 1.0.7. □

Kuntakertoimisista matriiseista poiketen, determinantti ei toimi aivan samalla tavalla kääntyvyysmittarina. Esimerkiksi matriisi $A = \mathrm{diag}(x, x) \in \mathrm{Mat}_2(\mathbb{K}[x])$ ei ole kääntyvä, vaikka $\det(A) = x^2 \neq 0$. Kääntyvän polynomimatriisin determinantti ei kuitenkaan voi olla nolla. Itseasiassa se on aina nollasta eroava skalaari.

LAUSE 1.0.9. *Olkoon $A \in Mat_n(\mathbb{K}[x])$ kääntyvä polynomimatriisi. Tällöin $det(A) \in \mathbb{K} \setminus \{0\}$.*

TODISTUS. Oletuksen nojalla on olemassa $A^{-1} \in \mathrm{Mat}_n(\mathbb{K}[x])$, jolle $A^{-1}A = I$. Tällöin lauseen 1.0.8 nojalla $\det(A^{-1})\det(A) = 1$, joten polynomi $\det(A)$ jakaa polynomin 1. Tällöin $\det(A)$ on välttämättä nollasta eroava vakiopolynomi. □

Myös lauseen 1.0.9 käänteinen väite pätee. Toisin sanoen polynomimatriisi on kääntyvä, jos sen determinantti on nollasta eroava skalaari. Tämän todistamiseen palataan myöhemmin kääntyvien polynomimatriisien ryhmää käsittelevässä luvussa. Polynomimatriisille $A \in \mathrm{Mat}_n(\mathbb{K}[x])$ voidaan määritellä *ranki* $\mathrm{rank}(A)$ asettamalla

$$\mathrm{rank}(A) = \max_{t \in \mathbb{K}} \{\mathrm{rank}(A(t))\}.$$

Polynomimatriisi $A \in \mathrm{Mat}_n(\mathbb{K}[x]) \backslash \{0\}$ voidaan aina esittää myös muodossa

$$A = \sum_{k=0}^{\deg(A)} A_k x^k,$$

missä $[A_k]_{ij}$ on polynomin $[A]_{ij}$ termin x^k kerroin. Tällöin $A_k \in \mathrm{Mat}_n(\mathbb{K})$ kaikilla indekseillä k. Tapauksessa $A = 0$ vastaava esitys on triviaali ($A = A$), sillä nollamatriisi voidaan polynomimatriisinakin tulkita skalaarikertoimiseksi matriisiksi. Polynomimatriiseille pätee samankaltainen jakoyhtölö kuin polynomeillekin. Esitetään siitä hieman vaillinainen versio, jota kutsutaan myös jäännöslauseeksi [vrt. Ayres, The Remainder Theorem, s. 181].

LAUSE 1.0.10 (Polynomimatriisien jakoyhtälö, Jäännöslause). *Olkoon*

$$A = \sum_{k=0}^{deg(A)} A_k x^k \in Mat_n(\mathbb{K}[x]),$$

missä $A_k \in Mat_n(\mathbb{K})$ kaikilla k. Olkoon lisäksi $Bx + C \in Mat_n(\mathbb{K}[x])$, missä $B, C \in Mat_n(\mathbb{K})$ ja B on kääntyvä. Tällöin on olemassa polynomimatriisit $R, Q \in Mat_n(\mathbb{K}[x])$ ja matriisit $G, F \in Mat_n(\mathbb{K})$, joille

$$A = (Bx + C)R + G = Q(Bx + C) + F.$$

TODISTUS. □

Tässä polynomimatriisia ei pidä sekoittaa *matriisipolynomiin*, jolla tarkoitetaan seuraavaa. Olkoot $A \in \mathrm{Mat}_n(\mathbb{K})$ ja $p = a_k x^k + a_{k-1} x^{k-1} + \cdots + a_0 \in \mathbb{K}[x]$. Silloin näitä vastaava matriisipolynomi on

$$p(A) = a_k A^k + a_{k-1} A^{k-1} + \cdots + a_0 I.$$

Matriisipolynomille pätevät esimerkiksi seuraavat laskusäännöt.

LAUSE 1.0.11. *Olkoot $A \in Mat_n(\mathbb{K})$ ja $p = a_k x^k + a_{k-1} x^{k-1} + \cdots + a_0 \in \mathbb{K}[x]$.*

- **(a):** *Olkoon $R \in Mat_n(\mathbb{K})$ kääntyvä. Tällöin $p(RAR^{-1}) = Rp(A)R^{-1}$.*
- **(b):** *Oletetaan, että A on lohkoyläkolmiomatriisi, jonka diagonaalimatriisit ovat $A_1, \ldots, A_l$. Tällöin $p(A)$ on muotoa*

$$p(A) = \begin{bmatrix} p(A_1) & & * \\ & \ddots & \\ 0 & & p(A_l) \end{bmatrix}.$$

Vastaavasti, jos A on lohkodiagonaalimatriisi $A = diag(A_1, \ldots, A_l)$, pätee

$$p(A) = diag(p(A_1), \ldots, p(A_l)).$$

TODISTUS. Väite (a) nähdään oikeaksi laskemalla

$$p(RAR^{-1}) = \sum_{j=0}^{k} a_j(RAR^{-1})^j = \sum_{j=0}^{k} a_j RA^j R^{-1} = R\left(\sum_{j=0}^{k} a_j A^j\right) R^{-1}.$$

Olkoot $A, B \in \mathrm{Mat}_n(\mathbb{K})$ lohkoyläkolmiomatriiseja, joiden diagonaalimatriisit ovat $A_1, \ldots, A_l$ ja $B_1, \ldots, B_l$. Oletetaan lisäksi, että neliömatriisit A_j ja B_j ovat samaa kokoa kaikille $j = 1, \ldots, l$. Tällöin tarkastelemalla matriisisumman ja -tulon määritelmiä nähdään, että

$$A + B = \begin{bmatrix} A_1 + B_1 & & * \\ & \ddots & \\ 0 & & A_l + B_l \end{bmatrix} \text{ ja } AB = \begin{bmatrix} A_1 B_1 & & * \\ & \ddots & \\ 0 & & A_l B_l \end{bmatrix}.$$

Nämä laskusäännöt yleistyvät induktiolla äärellisille summille ja tuloille. Lisäksi skalaarilla kertominen voidaan tehdä lohkoittain. Väite (b) lohkoyläkolmiomatriiseille seuraa näistä laskusäännöistä. Lohkodiagonaalimatriiseille päättely menee vastaavasti. □

2. Gauss-Jordan-muunnokset ja alkeispolynomimatriisit

Selvitetään, miten tunnetut reaali- ja kompleksikertoimisten matriisien muokkaamiseen käytetyt Gauss-Jordan -muunnokset yleistyvät polynomimatriiseille. Näitä muunnoksia on kolmea tyyppiä, ja niitä vastaavia matriiseja kutsutaan alkeismatriiseiksi. Polynomimatriisien tapauksessa vastaavia matriiseja kutsutaan tässä tutkielmassa *alkeispolynomimatriiseiksi*. Ennen niiden tarkkaa määrittelemistä esitellään kuitenkin vielä kantamatriisit. *Kantamatriiseiksi* kutsutaan matriiseja $E_{ij} \in Mat_n(\mathbb{K}[x])$, joille

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

Alkeispolynomimatriisit määritellään kantamatriisien avulla seuraavasti:

(1) $M_i(\alpha) = I + (\alpha - 1)E_{ii}$ $(\alpha \in \mathbb{K}\backslash\{0\})$.
(2) $P_{ij} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}$ $(i \neq j)$.
(3) $A_{ij}(r(x)) = I + r(x)E_{ji}$ $(i \neq j)$.

Ne ovat kääntyviä ja niiden käänteismatriisit ovat

$$M_i(\alpha)^{-1} = M_i(\alpha^{-1}),\ P_{ij}^{-1} = P_{ij} \text{ ja } A_{ij}(r(x))^{-1} = A_{ij}(-r(x)).$$

Muunnoksen suorittaminen vastaa siis kyseisellä alkeispolynomimatriisilla kertomista vasemmalta. Ensimmäinen muunnos $M_i(\alpha)$ on rivin i kertominen nollasta eroavalla skalaarilla α. Erona reaali- ja kompleksikertoimisten matriisiseihin huomataan, että kertoimeksi α ei sovi mikä tahansa nollasta eroava polynomi, vaan se on pidettävä skalaarina. Muunnosta vastaavan matriisilta vaaditaan nimittäin kääntyvyys myös polynomimatriisien tapauksessa. Jos α olisi ei-vakio polynomi, matriisi $M_i(\alpha)$ ei olisi kääntyvä. Toinen muunnos on rivien i ja j vaihto, joka on siis aivan sama kuin reaali- ja kompleksikertoimisten matriisiseille.

Kolmantena muunnoksena on $A_{ij}(r(x))$ eli rivin i lisääminen riville j kerrottuna polynomilla $r(x)$. Tässä $r(x)$ voi siis olla mikä tahansa polynomi, sillä matriisi $A_{ij}(r(x))$ on aina kääntyvä polynomista $r(x)$ riippumatta.

3. Polynomimatriisista yläkolmiomatriisiksi

Tässä kappaleessa selvitetään, miten polynomimatriisit voidaan muokata yläkolmiomuotoon käyttäen pelkästään edellä esitettyjä Gauss-Jordan -muunnoksia, joita jatkossa kutsutaan rivioperaatioiksi. Tavallisessa Gauss-Jordan-algoritmissa pyritään muokkaamaan matriisi porrasmuotoon tai yksinkertaiseen porrasmuotoon, josta esimerkiksi vastaavan yhtälöryhmän ratkaisut on helppo lukea. Tässä tavoitteet ovat kuitenkin erilaiset, ja porrasmuodon sijaan tavoitteena on yläkolmiomuoto. Tämä polynomimatriisin muokkaaminen yläkolmiomuotoon tulee myöhemmin olemaan tärkeässä roolissa määriteltäessä $\mathbb{K}$-kertoimisen matriisin karakteristista polynomia.

LAUSE 3.0.12. *Olkoon $A \in Mat_n(\mathbb{K}[x])$. Tällöin on olemassa kääntyvä $R \in Mat_n(\mathbb{K}[x])$ ja yläkolmiomatriisi $U \in Mat_n(\mathbb{K}[x])$, joille $A = RU$. Lisäksi matriisi R on alkeispolynomimatriisien tulo, ja siten kääntyvä.*

TODISTUS. Todistus tehdään induktiolla matriisin koon n suhteen. Ideana on muokata matriisi A Gauss-Jordan-operaatioilla yläkolmiomatriisiksi.

Aloitetaan tarkastelemalla ensin yleisesti $n \times n$-polynomimatriiseja, kun $n \geq 2$. Olkoon tätä varten $A \in \mathrm{Mat}_n(\mathbb{K}[x])$, ja merkitään

$$A = \begin{bmatrix} p_{11}^0 & * & \cdots & * \\ p_{21}^0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ p_{n1}^0 & * & \cdots & * \end{bmatrix}.$$

Tarkoituksena on määritellä rekursiivisesti matriisit $A_k := R_k \cdots R_1 A$, kun $k \geq 1$. Valitsemalla matriisit R_k sopiviksi alkeispolynomimatriisien tuloiksi pyritään siihen, että matriisi A_k saadaan lohkoyläkolmiomuotoon

$$A_k = \begin{bmatrix} p_{11}^k & * \\ 0 & D \end{bmatrix}$$

jollakin $k \in \mathbb{N}$. Polynomimatriisi D on tällöin kokoa $n - 1$ oleva neliömatriisi tai pelkkä polynomi, jos $n = 2$. Jos $p_{j1}^0 = 0$ kaikilla $j = 2, \ldots, n$, matriisi on jo haluttua muotoa, joten voidaan olettaa, että $p_{j1}^0 \neq 0$ jollakin $j \in \{2, \ldots, n\}$

Asetetaan $A_0 = A$. Oletetaan, että jollakin parillisella $k \in \mathbb{N} \cup \{0\}$ polynomimatriisit $A_0, \ldots, A_k$ ovat jo määriteltyjä. Merkitään

$$A_k := \begin{bmatrix} p_{11}^k & * & \cdots & * \\ p_{21}^k & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ p_{n1}^k & * & \cdots & * \end{bmatrix}$$

ja lisäksi voidaan oletetaan, että $p_{j1}^k \neq 0$ jollakin $j \in \{1, \ldots, n\}$. Polynomimatriisit A_{k+1} ja A_{k+2} määritellään valitsemalla rivioperaatioita vastaavat matriisit R_{k+1} ja R_{k+2} valitaan seuraavasti:

Vaihe 1: Olkoon p_{1l}^k ensimmäisen sarakkeen asteeltaan pienin nollasta eroava polynomi. Jos $l = 1$, asetetaan $R_{k+1} = I$. Jos $l \neq 1$, asetetaan $R_{k+1} = P_{1l}$. Merkitään

$$A_{k+1} = R_{k+1}A_k := \begin{bmatrix} p_{11}^{k+1} & * & \cdots & * \\ p_{21}^{k+1} & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ p_{n1}^{k+1} & * & \cdots & * \end{bmatrix}.$$

Vaiheessa 1 tehdään siis tarvittaessa rivinvaihto, jotta ensimmäisen sarakkeen asteeltaan pienin nollasta eroava polynomi saadaan vasempaan ylänurkkaan. Erityisesti tällöin $p_{11}^{k+1} \neq 0$.

Vaihe 2: Jos $p_{j1}^{k+1} = 0$ kaikilla $j = 2, \ldots, n$, matriisi on jo haluttua muotoa, joten asetetaan $R_{k+2} = I$. Muussa tapauksessa olkoon $j \in \{2, \ldots, n\}$ sellainen, että $p_{j1}^{k+1} \neq 0$. Tällöin vaiheen 1 mukaan $\deg(p_{11}^{k+1}) \leq \deg(p_{j1}^{k+1})$. Polynomien jakoyhtälön mukaan

$$p_{j1}^{k+1} = r_j^{k+1} p_{11}^{k+1} + p_{j1}^{k+2},$$

missä $r_j^{k+1}, p_{j1}^{k+2} \in \mathbb{K}[x]$ ja

$$\deg(p_{j1}^{k+2}) < \deg(p_{11}^{k+1}). \tag{2.3}$$

Jako voi mennä myös tasan, eli voi olla $p_{j1}^{k+2} = 0$. Asetetaan kaikille $j = 2, \ldots, n$

$$R_{k+2,j} = \begin{cases} A_{1j}(-r_j^{k+1}), \text{ jos } p_{j1}^{k+1} \neq 0 \\ I, \text{ jos } p_{j1}^{k+1} = 0 =: p_{j1}^{k+2} \end{cases}$$

ja $R_{k+2} := R_{k+2,n} \cdots R_{k+2,2}$. Tällöin

$$A_{k+2} = R_{k+2}A_{k+1} = \begin{bmatrix} p_{11}^{k+2} & * & \cdots & * \\ p_{21}^{k+2} & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ p_{n1}^{k+2} & * & \cdots & * \end{bmatrix},$$

missä $p_{11}^{k+2} = p_{11}^{k+1} \neq 0$ ja $\deg(p_{j1}^{k+2}) < \deg(p_{11}^{k+1})$ kaikilla $j = 2, \ldots, n$.

Näin matriisit A_k ja R_k tulevat rekursiivisesti määritellyiksi kaikille $k \in \mathbb{N}$. Kaikilla $k \in \mathbb{N} \cup \{0\}$ ja $j = 2, \ldots n$ polynomeille p_{j1}^k pätevät seuraavat:

(a): Jos $p_{j1}^k = 0$, myös $p_{j1}^{k+1} = 0$.

(b): Jos $k + 2$ on parillinen ja $p_{j1}^{k+2} \neq 0$, pätee $\deg(p_{j1}^{k+2}) < \deg(p_{j1}^k)$.

Todistetaan väitteet (a) ja (b). Olkoon $j \in \{2, \ldots, n\}$. Väite (a) seuraa tällöin suoraan vaiheen 1 määrittelystä, jos k on parillinen. Jos taas k on pariton, väite (a) seuraa vaiheen 2 määrittelystä. Jos siis $p_{j1}^{k+2} \neq 0$, (a)-kohdan nojalla myös $p_{j1}^k \neq 0$. Silloin

$$\deg(p_{j1}^{k+2}) \overset{(i)}{<} \deg(p_{11}^{k+1}) \overset{(ii)}{\leq} \deg(p_{j1}^k).$$

Epäyhtälö (i) seuraa yhtälöstä (2.3). Epäyhtälö (ii) seuraa vaiheen 1 määrittelystä, jonka mukaan p_{11}^{k+1} on matriisin A_k ensimmäisen sarakkeen asteeltaan pienin nollasta eroava polynomi.

Olkoon $j \in \{2, \ldots, n\}$. Oletetaan, että $p_{j1}^k \neq 0$ kaikilla $k \in \mathbb{N} \cup \{0\}$. Tällöin tuloksen (b) nojalla $\deg(p_{j1}^{k+2}) < \deg(p_{j1}^k) < \deg(p_{j1}^0)$ kaikilla parillisilla $k \in \mathbb{N}$, mikä on ristiriita. Siispä on olemassa sellainen $m_j \in \mathbb{N} \cup \{0\}$, jolle $p_{j1}^{m_j} = 0$. Asetetaan

$m := \max\{m_j | j = 2, \ldots, n\}$. Tällöin tuloksen (a) nojalla $p_{j1}^m = 0$ kaikilla $j = 2, \ldots, n$. Tämä tarkoittaa sitä, että matriisi A_m on haluttua muotoa. Toisin sanoen

$$A_m = \begin{bmatrix} p_{11}^m & * \\ 0 & D \end{bmatrix},$$

missä D on kokoa $n-1$ oleva neliömatriisi tai polynomi.

Aloitetaan seuraavaksi varsinainen induktiotodistus. Osoitetaan siis ensin, että alkuperäinen väite pätee 2×2-matriiseille. Olkoon $A \in \mathrm{Mat}_2(\mathbb{K}[x])$. Edellä todistetun nojalla on olemassa polynomimatriisit $R', U \in \mathrm{Mat}_n(\mathbb{K}[x])$, joille $U = P'A$ on muotoa

$$U' = \begin{bmatrix} p_{11}^m & * \\ 0 & d \end{bmatrix},$$

eli U on yläkolmiomatriisi. Lisäksi R' on alkeispolynomimatriisien tulo. Tällöin myös $R := R'^{-1}$ on alkeispolynomimatriisien tulo ja toisaalta $A = RU$, mikä oli todistettava.

Tehdään seuraavaksi induktio-oletus. Oletetaan siis, että jollakin $n \geq 3$ väite pätee kaikille enintään kokoa $n-1$ oleville polynomimatriiseille. Induktio väitteenä on silloin, että väite pätee kokoa n oleville polynomimatriiseille. Olkoon $A \in \mathrm{Mat}_n(\mathbb{K}[x])$. Kuten edellä osoitetiin, on siis olemassa $R', U' \in \mathrm{Mat}_n(\mathbb{K}[x])$, joille $U' = R'A$ on lohkoyläkolmiomuotoa

$$U' = \begin{bmatrix} p_{11}^m & X \\ 0 & D \end{bmatrix}$$

ja R' on alkeispolynomimatriisien tulo. Erityisesti siis $A = R'^{-1}U'$. Polynomimatriisi D on kokoa $n-1$, joten induktio oletuksen nojalla on olemassa $R_0, U_0 \in \mathrm{Mat}_{n-1}(\mathbb{K}[x])$, joille $D = R_0U_0$. Lisäksi U_0 on yläkolmiomatriisi ja R_0 on alkeispolynomimatriisien tulo. Määritellään polynomimatriisit

$$U := \begin{bmatrix} p_{11}^m & X \\ 0 & U_0 \end{bmatrix} \text{ ja } R_0' := \begin{bmatrix} 1 & 0 \\ 0 & R_0 \end{bmatrix}.$$

Tällöin U on yläkolmiomatriisi ja R_0' on alkeispolynomimatriisien tulo. Lisäksi $U' = R_0'U$, joten $A = R'^{-1}R_0'U$. Asetetaan $R := R'^{-1}R_0'$, jolloin $A = RU$ ja R on alkeispolynomimatriisien tulo. □

HUOMAUTUS 3.0.13. Lauseen 3.0.12 todistuksessa käytetään vain tyypin 2 ja 3 alkeispolynomimatriiseja. Skalaarilla kertomista ei tarvita.

Lauseen 3.0.12 todistus antaa myös menetelmän kyseisen muodon laskemiseksi polynomimatriisille $A \in \mathrm{Mat}_n(\mathbb{K}[x])$. Laskeminen kannattaa suorittaa seuraavissa vaiheissa.

(1) Suoritetaan tarvittava rivien vaihto, jotta ensimmäisen sarakkeen asteeltaan pienin nollasta eroava polynomi saadaan paikalle $(1, 1)$.
(2) Kirjoitetaan ensimmäisen sarakkeen polynomit jakoyhtälön avulla muodossa $p_{i1} = r_ip_{11} + s_i$, kun $i > 1$, ja sovelletaan rivioperaatioita $A_{1i}(-r_i)$. Jos jokin polynomeista s_i on nollasta eroava, siirrytään takaisin vaiheeseen 1. Muussa tapauksessa siirrytään vaiheeseen 3.
(3) Tässä vaiheessa matriisin pitäisi olla muotoa

$$\begin{bmatrix} p_1 & * \\ 0 & D \end{bmatrix}.$$

Jos D ei ole yläkolmiomatriisi tai pelkkä polynomi, siirrytään takaisin vaiheeseen 1 ja jatketaan rivioperaatioiden soveltamista matriisiin D.

Esimerkki 3.0.14. Muunnetaan matriisi

$$\begin{bmatrix} -x+1 & 2 & 4 \\ -1 & -x & 2 \\ 3 & -1 & -x+5 \end{bmatrix} \in \mathrm{Mat}_3(\mathbb{R}[x])$$

yläkolmiomuotoon:

$$\begin{bmatrix} -x+1 & 2 & 4 \\ -1 & -x & 2 \\ 3 & -1 & -x+5 \end{bmatrix} \overset{I_{12}}{\to} \begin{bmatrix} -1 & -x & 2 \\ -x+1 & 2 & 4 \\ 3 & -1 & -x+5 \end{bmatrix} \overset{A_{12}(-x+1),\, A_{13}(3)}{\to}$$

$$\begin{bmatrix} -1 & -x & 2 \\ 0 & x^2-x+2 & -2x+6 \\ 0 & -3x-1 & -x+11 \end{bmatrix} \overset{I23}{\to}$$

$$\begin{bmatrix} -1 & -x & 2 \\ 0 & -3x-1 & -x+11 \\ 0 & \left(-\frac{1}{3}x+\frac{4}{9}\right)(-3x-1)+\frac{22}{9} & -2x+6 \end{bmatrix} \overset{A_{23}(\frac{1}{3}x-\frac{4}{9})}{\to}$$

$$\begin{bmatrix} -1 & -x & 2 \\ 0 & -3x-1 & -x+11 \\ 0 & \frac{22}{9} & (-x+11)\left(\frac{1}{3}x-\frac{4}{9}\right)-2x+6 \end{bmatrix} \overset{I23}{\to}$$

$$\begin{bmatrix} -1 & -x & 2 \\ 0 & \frac{22}{9} & \frac{1}{9}(-3x^2+19x+10) \\ 0 & -3x-1 & -x+11 \end{bmatrix} \overset{A_{23}(\frac{9}{22}(3x+1))}{\to}$$

$$\begin{bmatrix} -1 & -x & 2 \\ 0 & \frac{22}{9} & \frac{1}{9}(-3x^2+19x+10) \\ 0 & 0 & \frac{1}{9}(-3x^2+19x+10)\left(\frac{9}{22}(3x+1)\right)-x+11 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -x & 2 \\ 0 & \frac{22}{9} & \frac{1}{9}(-3x^2+19x+10) \\ 0 & 0 & -\frac{9}{22}(x^3-6x^2-3x-28) \end{bmatrix}$$

Lauseessa 3.0.12 yläkolmiomatriisin diagonaalipolynomit eivät ole yksikäsitteisiä. Tulevaa yläkolmiomuodon soveltamista varten yksikäsitteisyys olisi kuitenkin tärkeää, ja lausetta 3.0.12 voidaankin tietyin ehdoin tältä osin parantaa.

LAUSE 3.0.15. *Olkoon* $A \in Mat_n(\mathbb{K}[x])$ *ja* $A = \tilde{R}\tilde{U}$ *lauseen 3.0.12 antama esitys. Merkitään* $diag(\tilde{U}) = (\tilde{p}_1, \ldots, \tilde{p}_n)$*, ja oletetaan, että* $\tilde{p}_j \neq 0$ *kaikilla* $j = 1, \ldots, n$*. Tällöin polynomimatriisilla* A *on esitys* $A = RU$*, missä matriiseille* $R, U \in Mat_n(\mathbb{K}[x])$ *pätevät seuraavat:*

- R *on alkeispolynomimatriisien tulo.*
- U *on yläkolmiomatriisi.*
- *Kun merkitään* $diag(U) = (p_1, \ldots, p_n)$*, polynomit* $p_j \in \mathbb{K}[x]$ *ovat perusmuotoisia ja yksikäsitteisiä.*

TODISTUS. Koska $\tilde{p}_j \neq 0$ jokaisella j, voidaan kullekin j valita sellaiset kertoimet $\alpha_j \in \mathbb{K}\backslash\{0\}$, joille polynomi $p_j := \alpha_j \tilde{p}_j$ on perusmuotoinen. Kertomalla matriisin $\tilde{U}$ jokainen rivi vastaavalla kertoimella α_j saadaan diagonaalipolynomit perusmuotoisiksi. Näitä rivioperaatioita vastaavat alkeispolynomimatriisit $M_j(\alpha_j)$, joten asetetaan

$$R_0 := \prod_{j=1}^{n} M_j(\alpha_j).$$

Valitaan vielä $U := R_0\tilde{U}$ ja $R := \tilde{R}R_0^{-1}$, jolloin $A = RU$ on haluttua muotoa oleva esitys.

Osoitettavaksi jää siis polynomien p_j yksikäsitteisyys. Oletetaan, että matriisilla A on myös esitys $A = SV$. Tässä siis S on alkeispolynomimatriisien tulo ja V on yläkolmiomatriisi, jonka diagonaalipolynomit q_j ovat perusmuotoisia. Riittää osoittaa, että $p_j = q_j$ kaikilla $j = 1, \ldots, n$.

Oletuksen mukaan $RU = SV$. Merkitsemällä $W := S^{-1}R$ ja $\tilde{W} := R^{-1}S$ saadaan yhtälöt

$$WU = V, \tag{2.4}$$

ja

$$\tilde{W}V = U, \tag{2.5}$$

Kirjoittamalla matriisit auki yhtälöissä (2.4) ja (2.5) saadaan

$$\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \begin{bmatrix} p_1 & p_{12} & \cdots & p_{1n} \\ 0 & p_2 & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{bmatrix} = \begin{bmatrix} q_1 & q_{12} & \cdots & q_{1n} \\ 0 & q_2 & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_n \end{bmatrix} \tag{2.6}$$

ja

$$\begin{bmatrix} \tilde{w}_{11} & \tilde{w}_{12} & \cdots & \tilde{w}_{1n} \\ \tilde{w}_{21} & \tilde{w}_{22} & \cdots & \tilde{w}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{w}_{n1} & \tilde{w}_{n2} & \cdots & \tilde{w}_{nn} \end{bmatrix} \begin{bmatrix} q_1 & q_{12} & \cdots & q_{1n} \\ 0 & q_2 & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} p_1 & p_{12} & \cdots & p_{1n} \\ 0 & p_2 & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{bmatrix}. \tag{2.7}$$

Todistetaan seuraavaksi, että polynomimatriisit W ja $\tilde{W}$ ovat yläkolmiomatriiseja. Tehdään tämä todistus vain matriisille W, sillä matriisille $\tilde{W}$ todistus menee täysin vastaavasti. On osoitettava, että $w_{ij} = 0$, kun $i > j$. Tehdään tämä induktiolla j:n suhteen. Vertaamalla ensimmäisen sarakkeen kertoimia yhtälössä (2.6) nähdään, että

$$\begin{aligned} w_{11}p_1 &= q_1 \\ w_{21}p_1 &= 0 \\ &\vdots \\ w_{n1}p_1 &= 0. \end{aligned}$$

Koska oletuksen mukaan $p_1 \neq 0$, on välttämättä $w_{j1} = 0$ kaikilla $j = 2, \ldots, n$. Tapaus $j = 1$ on siis kunnossa. Oletetaan seuraavaksi, että jollakin $k \in \{2, \ldots, n\}$ pätee jokaiselle $j \in \{1, \ldots, k-1\}$, että $w_{ij} = 0$ kaikilla $i > j$. Induktioväitteenä on tällöin, että $w_{ik} = 0$ kaikilla $i > k$. Jos $k = n$ ei ole mitään todistettavaa, joten voidaan olettaa, että $k < n$. Olkoon $i > k$. Yhtälöstä (2.6) nähdään, että

$$\sum_{j=1}^{k} w_{ij}p_{kj} = 0 \tag{2.8}$$

Toisaalta induktio-oletuksen nojalla $w_{ij} = 0$ kaikilla $j = 1, \ldots, k-1$, joten yhtälö (2.8) sievenee muotoon

$$w_{ik}p_k = 0.$$

Tällöin $w_{ik} = 0$, koska oletuksen nojalla $p_k \neq 0$, ja indutioväite on siten todistettu.

Olkoon $j \in \{1, \ldots, n\}$. Koska $w_{ij} = 0 = \tilde{w}_{ij}$ kaikilla $i > j$, nähdään yhtälöistä (2.6) ja (2.7), että

$$w_{jj}p_j = q_j \text{ ja } \tilde{w}_{jj}q_j = p_j.$$

Siis p_j jakaa polynomin q_j ja myös q_j jakaa polynomin p_j. Koska sekä p_j että q_j ovat oletuksen mukaan perusmuotoisia, tästä seuraa, että $p_j = q_j$. $\square$

LUKU 3

Matriiseja ja lineaarisia operaattoreita

1. Matriisin ominaisarvo ja lineaariset operaattorit

Sanotaan, että $\lambda \in \mathbb{K}$ on matriisin $A \in \mathrm{Mat}_n(\mathbb{K})$ ominaisarvo, jos $Av = \lambda v$ jollekin $v \in \mathbb{K}^n \setminus \{0\}$. Tällöin siis λ on matriisin A ominaisarvo täsmälleen silloin, kun yhtälöllä

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

on epätriviaali ratkaisu. Tämä puolestaan on yhtäpitävää sen kanssa, että matriisi $A - \lambda I$ ei ole kääntyvä.

Usein matriisin karakteristinen polynomi määritellään determinantin avulla. Matriisi $A - \lambda I$ ei ole kääntyvä täsmälleen silloin, kun $\det(A - \lambda I) = 0$. Silloin matriisin A ominaisarvot ovat täsmälleen polynomin $\det(A - xI)$ juuret. Matriisin A *karakteristinen polynomi* χ_A määritellään polynomina $\det(A - xI)$. Determinantin käyttäminen karakteristisen polynomin määrittelemisessä ei kuitenkaan ole välttämätöntä. Myöhemmin luvussa 5 esitetään hieman toisenlainen tapa määrittelylle käyttäen hyväksi edellä todistettuja tuloksia polynomimatriiseille ja todistetaan tunnettu Cayleyn ja Hamiltonin lause $\mathbb{K}$-kertoimisille matriiseille tätä määritelmää hyödyntäen. Lauseen mukaan matriisin karakteristinen polynomi nollaa sen, eli toisin sanoen $\chi_A(A) = 0$. Sitä ennen on kuitenkin tarkoitus todistaa Caylen ja Hamiltonin lause reaali ja kompleksikertoimisissa tapauksissa hyödyntäen reaali- ja kompleksilukujen ominaisuuksia. Tämän yhteydessä käy ilmi myös reaali- ja kompleksikertoimisille matriiseille soveltuva karakteristisenpolynomin vaihtoehtoinen määrittely, johon ei kuitenkaan ole tarkoitus syventyä tässä sen enempää.

Renkaan $\mathrm{Mat}_n(\mathbb{K})$ matriisit liittyvät tunnetusti n-ulotteisen $\mathbb{K}$-kertoimisen lineaarisen vektoriavaruuden lineaarisiin operaattoreihin, ja näiden matriisien teoria on näin ollen myös vektoriavaruuden teoriaa ja päinvastoin. Tässä tutkielmassa tarkastellaan ensisijaisesti matriiseja, eikä niinkään olla kiinnostuneita vektoriavaruuksista ja lineaarisista operaattoreista. Jatkossa tarvitaan kuitenkin hieman myös lineaarisia vektoriavaruuksia $\mathbb{R}^n$, $\mathbb{C}^n$ ja $\mathbb{K}^n$. Päämielenkiinto pidetään kuitenkin matriiseissa, ja lineaariset operaattorit ovat lähinä apuna niiden tarkastelussa. Lisäksi avaruudessa $\mathbb{C}^n$ käytetään myös tavallista sisätuloa $(\cdot|\cdot)$, jolle

$$(x|y) = \sum_{j=1}^{n} x_j \overline{y_j},$$

missä $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$. Äärellisiulotteisten lineaaristen vektori- ja sisätuloavaruuksien teorian perusteita ei ole tarkoitus kerrata tässä tämän tarkemmin. Näitä asioita on käsitelty esimerkiksi S. Axlerin kirjassa Linear Algebra Done Right luvuissa 1,2 ja 6. Kerroinkuntina on tosin käytetty vain kuntia $\mathbb{R}$ ja $\mathbb{C}$. Yleiselle kerroinkunnalle vastaavaa teoriaa on käsitelty esimerkiksi Jonathan S. Golanin kirjassa The Linear Algebra a Beginning Graduate Student Ought to Know luvuissa 3,5 ja 15. Perusasiat eivät tosin juurikaan poikkea toisistaan oli kerroinkuntaa rajoitettu tai ei.

Matriisia $A \in \mathrm{Mat}_n(\mathbb{K})$ vastaa kiinnitetyssä avaruuden $\mathbb{K}^n$ kannassa yksikäsitteisesti lineaarinen operaattori $L_A : \mathbb{K}^n \to \mathbb{K}^n$. Sanotaan, että matriisit $A, B \in \mathrm{Mat}_n(\mathbb{K})$ ovat *yhtäläiset* tai *similaariset*, jos on olemassa sellainen kääntyvä matriisi $R \in \mathrm{Gl}_n(\mathbb{K})$, jolle $A = R^{-1}BR$. Tällöin matriisit A ja B vastaavat samaa lineaarista operaattoria mutta eri kannoissa. Aliavaruuden $V \in \mathbb{K}^n$ sanotaan olevan *L_A-invariantti*, mikäli $L_A(V) \subset V$. Lineaarisen operaattorin ominaisarvot ja ominaisvektorit ovat samat kuin sitä vastaavan matriisin. Sillä ei ole merkitystä, missä kannassa vastaavuus on. Lineaaristen operaattoreiden perusasioita ei käsitellä tässä enempää [ks. tarvittaessa esim. Axler, luvut 3 ja 5 tai Golan, luvut 6 ja 8]. Lineaariselle operaattorille ja sitä standardikannassa vastaavalle matriisille saatetaan käyttää joskus samaa merkintää, jos sekaannuksen vaaraa ei ole. Eri merkintöjä käytetään, jos sen voidaan katsoa selkeyttävän tilannetta. Lineaariselle operaattorille voidaan määritellä polynomi matriisipolynomin avulla asettamalla $p(L_A) := L_{p(A)}$.

2. Kompleksi- ja reaalikertoimisista matriiseista

Matriisin $U \in \mathrm{Mat}_n(\mathbb{C}^n)$ sanotaan olevan unitaarinen, mikäli $U^*U = UU^* = I$ tai yhtäpitävästi sen sarakevektorit muodostavat avaruuden $\mathbb{C}^n$ ortonormaalin kannan. Tässä U^* on matriisin U kompleksikonjugaatin transpoosi eli $U^* = (\overline{U})^T$. Seuraava Schurin tai Schurin ja Toeplizin lauseena tunnettu tulos on eräs kompleksikertoimisten matriisien teorian keskeisimmistä perustuloksista [ks. Horn & Johnson, SchurŽs unitary triangularization theorem, s. 79].

Lause 2.0.16 (Schurin ja Toeplizin lause). *Olkoon $A \in Mat_n(\mathbb{C})$. Tällöin on olemassa unitaarinen $U \in Mat_n(\mathbb{C})$ ja yläkolmiomatriisi $B \in Mat_n(\mathbb{C})$, joille $A = UBU^*$.*

Todistus. Todistus tehdään induktiolla matriisin koon n suhteen. Olkoon λ_1 jokin matriisin A ominaisarvo ja u_1 vastaava normitettu ominaisvektori. Kompleksikertoimisella matriisilla on aina vähintään yksi ominaisarvo, joten λ_1 on olemassa. Reaalisessa tapuksessa näin ei välttämättä ole, ja siksi tämä todistus ei toimi reaalisille matriiseille. Täydennetään vektori u_1 ortonormaaliksi kannaksi $\{u_1, \ldots, u_n\}$, ja merkitään $U = [u_1, \ldots, u_n] \in \mathrm{Mat}_n(\mathbb{C})$. Tällöin U on unitaarinen ja

$$\begin{aligned} U^*AU &= [\overline{u_1}, \ldots, \overline{u_n}]^T [\lambda_1 u_1, Au_2, \ldots, Au_n] \\ &= \begin{bmatrix} \lambda_1 \overline{(u_1|u_1)} & * \\ \lambda_1 \overline{(u_2|u_1)} & \\ \vdots & Y \\ \lambda_1 \overline{(u_n|u_1)} & \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ 0 & Y \end{bmatrix}, \end{aligned}$$

missä $Y \in \mathrm{Mat}_{n-1}(\mathbb{C})$. Tällöin väite pätee, kun $n = 2$, sillä silloin $Y \in \mathbb{C}$. Oletetaan seuraavaksi, että jollakin $n \geq 3$ väite pätee kaikille enintään kokoa $n - 1$ oleville matriiseille. Tällöin edellä todetun nojalla on olemassa unitaarinen matriisi V_1 jolle

$$V_1^* A V_1 = \begin{bmatrix} \lambda_1 & * \\ 0 & Y \end{bmatrix},$$

missä $Y \in \mathrm{Mat}_{n-1}(\mathbb{C})$. Edelleen induktio-oletuksen nojalla on olemassa unitaarinen matriisi U_1, jolle $U_1^* Y U_1 =: B_1$ on yläkolmiomatriisi. Asetetaan $V_2 := \mathrm{diag}(1, U_1)$ ja $U := V_1 V_2$, jolloin matriisit V_2 ja U on myös unitaarisia. Lisäksi tällöin

$$U^* A U = \begin{bmatrix} \lambda_1 & * \\ 0 & B_1 \end{bmatrix}$$

on yläkolmiomatriisi. □

Schurin ja Toeplizin lause ei siis päde yleisesti reaalisille matriiseille. Seuraavaksi on tarkoitus osoittaa, että reaalisillekin matriiseille pätee kuitenkin hyvin samankaltainen tulos.

LEMMA 3.1. *Olkoon $A : \mathbb{R}^n \to \mathbb{R}^n$ lineaarinen operaattori. Tällöin on olemassa A-invariantti aliavaruus $V \subset \mathbb{R}^n$, jolle $\dim V \in \{1, 2\}$.*

TODISTUS. Olkoon $0 \neq v \in \mathbb{R}^n$. Tällöin joukko

$$\{v, Av, A^2 v, \dots, A^n v\}$$

on lineaarisesti riippuva, koska se sisältää $n + 1$ vektoria. Silloin

$$\sum_{j=0}^{n} a_j A^j v = 0$$

joillekin $a_j \in \mathbb{R}$. Lemman 2.0.4 nojalla on olemassa jaottomat perusmuotoiset polynomit p_j ja $c \in \mathbb{R}$, joille

$$\sum_{j=0}^{n} a_j x^j = c \prod_{j=1}^{k} p_j(x)$$

ja $\deg(p_j) \leq 2$ kaikilla $j = 1, \dots, k$. Silloin

$$0 = \left(\sum_{j=0}^{n} a_j A^j \right) v = c \left(\prod_{j=1}^{k} p_j(A) \right) v.$$

Tällöin jollakin $j \in \{1, \dots, k\}$ kuvaus $p_j(A)$ ei ole injektio. Jollekin $0 \neq u \in \mathbb{R}^n$ pätee siis $p_j(A)u = 0$. Jos $\deg(p_j) = 1$ eli $p_j(A) = A + \alpha$ jollekin $\alpha \in \mathbb{R}$, pätee $Au = -\alpha u$. Tällöin $\langle u \rangle$ on yksiulotteinen A-invariantti aliavaruus.

Voidaan siis olettaa, että $\deg(p_j) = 2$. Tällöin $p_j(A) = A^2 + \alpha A + \beta$ joillekin α, $\beta \in \mathbb{R}$. Aliavaruus $\langle u, Au \rangle$ on selvästi vähintään yksi- ja enintään kaksiulotteinen, joten riittää osoittaa, että se on A-invariantti. Laskemalla nähdään, että

$$A(au + bAu) = aAu + bA^2 u = aAu - b(\alpha A + \beta)u = (a - \alpha)Au - (b\beta)u \in \langle u, Au \rangle$$

aina kun $au + bAu \in \langle u, Au \rangle$. □

LAUSE 2.0.17. *Olkoon $A \in Mat_n(\mathbb{R})$. Tällöin on olemassa kääntyvä matriisi $R \in Mat_n(\mathbb{R})$, jolle*

$$R^{-1}AR = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_k \end{bmatrix},$$

missä jokaiselle $j \in \{1, \dots, k\}$ joko $A_j \in \mathbb{R}$ tai A_j on 2×2-matriisi, jolla ei ole ominaisarvoja.

TODISTUS. Matriisia A vastaa standardikannassa yksikäsitteisesti lineaarinen operaattori $L : \mathbb{R}^n \to \mathbb{R}^n$. Riittää osoittaa, että on olemassa sellainen avaruuden $\mathbb{R}^n$ kanta, jonka suhteen operaattorilla L on haluttua muotoa oleva lohkodiagonaaliesitys. Todistus tehdään induktiolla matriisin koon ja vektoriavaruuden dimension n suhteen. Olkoon $n = 2$. Jos matriisilla A ei ole ominaisarvoja, se on jo haluttua muotoa. Voidaan siis olettaa, että matriisilla A on ominaisarvo $\lambda \in \mathbb{R}$, ja olkoon $v_1 \in \mathbb{R}^2$ vastaava ominaisvektori. Tällöin $\{v_1\}$ voidaan täydentää avaruuden $\mathbb{R}^2$ kannaksi $\{v_1, v_2\}$. Tämän kannan suhteen operaattorin L matriisiesitys on muotoa

$$\begin{bmatrix} \lambda & * \\ 0 & * \end{bmatrix}.$$

Oletetaan seuraavaksi, että jollakin $n \geq 3$ väite pätee kaikille enintään kokoa $n-1$ oleville matriiseille. Pitää osoittaa, että väite pätee kokoa n olevalle matriisille A. Jos vastaavalla lineaarisella operaattorilla L on ominaisarvo, voidaan valita L-invariantti aliavaruus $V \subset \mathbb{R}^n$, jolle $\dim V = 1$. Muutoin lemman 3.1 nojalla on olemassa L-invariantti aliavaruus $V \subset \mathbb{R}^n$, jolle $\dim V = 2$. Valitaan aliavaruudelle V jokin kanta $\mathcal{K}_1$ ja olkoon A_1 operaattorin $L|_V$ matriisiesitys tämän kannan suhteen. Tällöin $A_1 \in \mathbb{R}$, jos operaattorilla L on ominaisarvo. Jos operaattorilla L ei ole ominaisarvoa, A_1 on 2×2-matriisi. Lisäksi tällöin matriisilla A_1 ei voi olla ominaisarvoja, sillä muutoin myös operaattorilla $L|_V$ olisi ominaisarvo eli erityisesti myös operaattorilla L.

Kanta $\mathcal{K}_1$ voidaan täydentää koko avaruuden $\mathbb{R}^n$ kannaksi $\mathcal{K} = \{v_1, v_2, \dots, v_n\}$. Asetetaan $R_1 = [v_1, v_2, \dots, v_n] \in \mathrm{Mat}_n(\mathbb{R})$. Tällöin R_1 on kääntyvä ja

$$R_1^{-1}AR_1 = \begin{bmatrix} A_1 & * \\ 0 & B \end{bmatrix},$$

missä B on kokoa $n-1$ tai $n-2$ oleva neliömatriisi. Induktio oletuksen nojalla jollekin kääntyvälle matriisille R_2

$$R_2^{-1}BR_2 = \begin{bmatrix} A_2 & & * \\ & \ddots & \\ 0 & & A_k \end{bmatrix},$$

missä jokaiselle $j \in \{1, \dots, k\}$ joko $A_j \in \mathbb{R}$ tai A_j on 2×2-matriisi, jolla ei ole ominaisarvoja. Kun asetetaan

$$R := R_1 \begin{bmatrix} I & * \\ 0 & R_2 \end{bmatrix},$$

matriisi $R^{-1}AR$ on haluttua muotoa. □

3. Cayleyn ja Hamiltonin lause

LAUSE 3.0.18. *Olkoot $A \in Mat_n(\mathbb{K})$ ja $B \in Mat_n(\mathbb{K})$, jolle $A = UBU^{-1}$ jollekin kääntyvälle matriisille $U \in Mat_n(\mathbb{K})$. Tällöin*

$$det(A - xI) = det(B - xI).$$

TODISTUS. Cauchyn ja Binet'n kaavan (lause 1.0.8) avulla nähdään, että

$$\begin{aligned}\det(A - xI) &= \det(UBU^{-1} - xUU^{-1}) = \det(U(B - xI)U^{-1}) \\ &= \det(U)\det(B - xI)\det(U^{-1}) = \det(B - xI).\end{aligned}$$

□

SEURAUS 3.0.19. *Olkoot $A \in Mat_n(\mathbb{C})$ ja $B \in Mat_n(\mathbb{C})$ yläkolmiomatriisi, jolle $A = UBU^{-1}$ jollekin unitaariselle matriisille $U \in Mat_n(\mathbb{C})$. Olkoon lisäksi $\lambda_1, \dots, \lambda_n$ matriisin B diagonaalialkiot. Tällöin*

$$\chi_A(x) = \prod_{j=1}^{n} (\lambda_j - x),$$

TODISTUS. Määritelmän mukaan $\chi_A(x) = \det(A - xI)$, joten väite seuraa lauseesta 3.0.18. □

HUOMAUTUS 3.0.20. Karakteristinen polynomi voitaisiin siis seurauksen 3.0.19 ja lauseen 2.0.16 mukaan määritellä myös tulona $(\lambda_1 - x) \cdots (\lambda_n - x)$. Alkiot $\lambda_1, \dots, \lambda_n$ voivat olla minkä tahansa kyseisen matriisin kanssa unitaarisesti yhtäläisen yläkolmiomatriisin diagonaalialkiot. Itseasiassa unitaarisuusvaatimus ei ole oleellinen, mutta lauseen 2.0.16 nojalla se voidaan asettaa.

LAUSE 3.0.21 (Cayleyn ja Hamiltonin lause). *Olkoon $A \in Mat_n(\mathbb{C})$. Tällöin $\chi_A(A) = 0$.*

TODISTUS. Olkoon $B \in \mathrm{Mat}_n(\mathbb{C})$ jokin yläkolmiomatriisi, joka on unitaarisesti yhtäläinen matriisin A kanssa. Olkoon lisäksi $U = [u_1, \dots, u_n]$ vastaava muunnosmatriisi, jolloin $A = UBU^*$. Sarakevektorit u_j muodostavat avaruuden $\mathbb{C}^n$ kannan, joten riittää osoittaa, että $\chi_A(A)u_j = 0$ kaikille $j = 1, \dots, n$. Lauseen 3.0.19 mukaan

$$\chi_A(A) = (-1)^n (A - \lambda_1 I) \cdots (A - \lambda_n I),$$

missä $\lambda_j = [B]_{jj}$ kaikilla $j = 1, \dots, n$. Matriisit $A - \lambda_j I$ kommutoivat keskenään, kun $j \in \{1, \dots, n\}$, joten riittää osoittaa, että

$$(A - \lambda_1 I) \cdots (A - \lambda_j I)u_j = 0$$

kaikille $j = 1, \dots, n$. Kun $j = 1$, $(A - \lambda_1 I)u_1 = 0$, sillä u_1 on ominaisarvoa λ_1 vastaava ominaisvektori. Tehdään induktio-oletus, että jollakin $k \in \{2, \cdots, n\}$ väite pätee kaikilla $l = 1, \ldots, k-1$. Huomataan, että

$$
\begin{aligned}
&(A - \lambda_k I)u_k = U(B - \lambda_k I)U^* u_k = \\
&U \begin{bmatrix} \lambda_1 - \lambda_k & & & & & & * \\ & \ddots & & & & & \\ & & \lambda_{k-1} - \lambda_k & & & & \\ & & & 0 & & & \\ & & & & \lambda_{k+1} - \lambda_k & & \\ & 0 & & & & \ddots & \\ & & & & & & \lambda_n - \lambda_k \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = U \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{k-1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&= \sum_{j=1}^{k-1} \alpha_j u_j,
\end{aligned}
$$

joillekin $\alpha_j \in \mathbb{C}$. Tällöin

$$
\begin{aligned}
(A - \lambda_1 I) \cdots (A - \lambda_k I)u_k &= (A - \lambda_1 I) \cdots (A - \lambda_{k-1} I) \sum_{j=1}^{k-1} \alpha_j u_j \\
&= \sum_{j=1}^{k-1} \alpha_j (A - \lambda_1 I) \cdots (A - \lambda_k I)u_j = 0,
\end{aligned}
$$

sillä $(A - \lambda_1 I) \cdots (A - \lambda_k I)u_j = 0$ kaikilla $j = 1, \ldots, k-1$ induktio-oletuksen nojalla, koska matriisit $A - \lambda_j I$ kommutoivat keskenään. □

Cayleyn ja Hamiltonin lause pätee myös reaalisille matriiseille, toisin sanoen reaalikertoimisen matriisin karakteristinen polynomi nollaa sen. Koska reaalikertoiminen matriisi voidaan aina tulkita kompleksikertoimiseksi, on tämä jo todistettu lauseessa 3.0.21. Kuitenkin todistus voidaan tehdä myös kokonaan ilman kompleksilukuja tosin samalla periaatteella kuin kompleksinen tapaus. Schurin ja Toeplizin lauseen sijaan käytetään sen reaalista vastinetta eli lausetta 2.0.17.

LAUSE 3.0.22. *Olkoon $A \in Mat_n(\mathbb{R})$ ja*

$$
A = R \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_k \end{bmatrix} R^{-1},
$$

lauseen 2.0.17 antama esitys. Tällöin

$$
\prod_{j=1}^{k} \chi_{A_j}(x) = det(A - xI).
$$

TODISTUS. Väite seuraa lauseista 1.0.7 ja 3.0.18. □

HUOMAUTUS 3.0.23. Reaalisessa tapauksessa karakteristinen polynomi voitaisiin määritellä lauseiden 2.0.17 ja 3.0.22 mukaan tulona $\chi_{A_1}(x) \cdots \chi_{A_k}(x)$, missä matriisit $A_1, \ldots, A_k$ ovat lauseen 2.0.17 antamat diagonaalilohkot ja

$$
\chi_{A_j}(x) = \begin{cases} A_j - x, \text{ kun } A_j \in \mathbb{R} \\ ([A_j]_{11} I - x)([A_j]_{22} I - x) - [A_j]_{21} [A_j]_{12}, \text{ kun } A_j \in \mathrm{Mat}_2(\mathbb{R}). \end{cases}
$$

Lemma 3.2. *Olkoon* $A \in Mat_2(\mathbb{R})$. *Tällöin* $\chi_A(A) = 0$

Todistus. □

Lause 3.0.24. *Olkoon* $A \in Mat_n(\mathbb{R})$. *Tällöin* $\chi_A(A) = 0$.

Todistus. Olkoon $B \in \mathrm{Mat}_n(\mathbb{R})$ lauseen 2.0.17 antama yläkolmiomatriisi, joka on yhtäläinen matriisin A kanssa ja jonka diagonaalimatriisit A_j, $j = 1, \ldots, k$ ovat enintään kokoa 2. Olkoon lisäksi $R = [r_1, \ldots, r_n]$ vastaava muunnosmatriisi, jolloin $A = RBR^{-1}$. Olkoon V_j, matriisiin A_j liittyvä 1- tai 2-ulotteinen aliavaruus. Aliavaruus V_j on siis yhden tai kahden matriisin A_j liittyvän sarakevektorin r_i virittämä. Sarakevektorit r_i muodostavat avaruuden $\mathbb{R}^n$ kannan, koska matriisi R on kääntyvä. Silloin $\mathbb{R}^n = V_1 \oplus \cdots \oplus V_k$, joten riittää osoittaa, että $\chi_A(A)v = \{0\}$ kaikille $v \in V_j$ jokaisella $j = 1, \ldots, k$.

Lauseen 3.0.22 mukaan

$$\chi_A(A) = q_1(A) \cdots q_k(A),$$

missä

$$q_j(A) = \begin{cases} A_jI - A, \text{ jos } A_j \in \mathbb{R} \\ ([A_j]_{11}I - A)([A_j]_{22}I - A) - [A_j]_{21}[A_j]_{12}I, \text{ jos } A_j \in \mathrm{Mat}_2(\mathbb{R}) \end{cases}$$

kaikilla $j = 1, \ldots, k$. Matriisit $q_j(A)$ kommutoivat keskenään, kun $j \in \{1, \ldots, k\}$, joten riittää osoittaa, että

$$q_1(A) \cdots q_j(A)v = \{0\}$$

kaikille $v \in V_j$ jokaisella $j = 1, \ldots, k$. Olkoon ensin $j = 1$. Tarkastellaan tapausta $q_1(A) = A_1I - A$. Tällöin A_1 on matriisin A ominaisarvo, ja $(A - A_1I)r_1 = 0$, sillä r_1 ominaisarvoa A_1 vastaava ominaisvektori. Lisäksi $V_1 = \langle r_1 \rangle$, joten väite pätee. Toinen vaihtoehto on, että

$$q_1(A) = ([A_1]_{11}I - A)([A_1]_{22}I - A) - [A_1]_{21}[A_1]_{12}I.$$

Tällöin siis $q_1(A) = \chi_{A_1}(A)$. Olkoon $v_1 \in V_1$. Matriisista B nähdään, että V_1 on L_A-invariantti. Lineaarisena operaattorina $L_A|_{V_1} = L_{A_1}$. Tällöin lineaariselle operaattorille $\chi_{A_1}(L_A)$ pätee $\chi_{A_1}(L_A)|_{V_1} = \chi_{A_1}(L_{A_1})$. Toisaalta lemman 3.2 nojalla matriisi $\chi_{A_1}(A_1) = 0$, jolloin vastaava lineaarinen operaattori on nollaoperaattori ja erityisesti siis $\chi_{A_1}(L_A)v_1 = 0$.

Tehdään induktio-oletus, että jollakin $l \in \{2, \cdots, n\}$ väite pätee kaikilla $j = 1, \dots, l-1$. Olkoon $v_l \in V_l$. Tällöin lauseen 1.0.11 nojalla

$$
\begin{aligned}
&q_l(A)v_l = R(q_l(B))R^{-1}v_l = \\
&U \begin{bmatrix} q_l(A_1) & & & & & & \\ & \ddots & & & & * & \\ & & q_l(A_{l-1}) & & & & \\ & & & 0 & & & \\ & & & & q_l(A_{l+1}) & & \\ & 0 & & & & \ddots & \\ & & & & & & q_l(A_k) \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ Y \\ 0 \\ \vdots \\ 0 \end{bmatrix} = U \begin{bmatrix} Z_1 \\ \vdots \\ Z_{l-1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&= \sum_{j=1}^{l-1} v_j,
\end{aligned}
$$

missä $Y, Z_1, \dots, Z_{l-1}$ ovat kokoa 1 tai 2 olevia pystyvektoreita ja $v_j \in V_j$ kaikilla $j = 1, \dots, l-1$. Silloin

$$
\begin{aligned}
q_1(A) \cdots q_l(A)v_l &= q_1(A) \cdots q_{l-1}(A) \sum_{j=1}^{l-1} v_j \\
&= \sum_{j=1}^{l-1} q_1(A) \cdots q_{l-1}(A)v_j = 0
\end{aligned}
$$

sillä $q_1(A) \cdots q_{l-1}(A)v_j = 0$ kaikilla $j = 1, \dots, l-1$ induktio-oletuksen ja oletuksen $v_j \in V_j$ nojalla, koska matriisit $q_j(A_j)$ kommutoivat keskenään. □

Edellä esitetyt toditukset Cayleyn ja Hamiltonin lauseelle nojaavat oleellisesti reaali- ja kompleksilukujen erityisominaisuuksiin, ja sopivat siten huonosti yleistettäviksi kuntaan $\mathbb{K}$. Seuraavassa luvussa määritellään karakteristinen polynomi uudelleen käyttämättä determinanttia ja esitetään tätä määritelmää hyödyntäen yleisessä kunnassa $\mathbb{K}$ pätevä todistus Cayleyn ja Hamiltonin lauseelle.

LUKU 4

Karakteristinen polynomi

1. Määrittely ilman determinanttia

Seuraavaksi on tarkoitus määritellä $\mathbb{K}$-kertoimisen matriisin karakteristinen polynomi hyödyntäen aiempia polynomimatriiseihin liittyviä tuloksia. Määrittely nojaa oleellisesti seuraavaan lauseeseen.

LAUSE 1.0.25. *Olkoon $A \in Mat_n(\mathbb{K})$, jolloin $A - xI \in Mat_n(\mathbb{K}[x])$. Tällöin $A - xI = R(x)U(x)$, missä*

- *$R(x) \in Mat_n(\mathbb{K}[x])$ on alkeispolynomimatriisien tulo.*
- *$U(x) \in Mat_n(\mathbb{K}[x])$ on yläkolmiomatriisi*
- *Kun merkitään $diag(U(x)) = (p_1(x), \ldots, p_n(x))$, polynomit $p_j(x)$ ovat perusmuotoisia ja yksikäsitteisiä. Erityisesti $p_j(x) \neq 0$ kaikilla $j = 1, \ldots, n$.*

TODISTUS. Lauseen 3.0.12 nojalla on kaksi ensimmäistä ehtoa toteuttavat matriisit $\tilde{R}(x), \tilde{U}(x) \in \mathrm{Mat}_n(\mathbb{K}[x])$, joille $A(x) = \tilde{R}(x)\tilde{U}(x)$. Merkitään $\mathrm{diag}(\tilde{U}(x)) = (\tilde{p}_1(x), \ldots, \tilde{p}_n(x))$. Tällöin väite seuraa lauseesta 3.0.15, kun osoitetaan, että $\tilde{p}_j(x) \neq 0$ kaikilla $j = 1, \ldots, n$.

Tehdään antiteesi, että olisikin $\tilde{p}_j(x) = 0$ jollakin $j \in \{1, \ldots, n\}$. Olkoon $t \in \mathbb{K}$. Tällöin $\tilde{p}_j(t) = 0$, joten matriisi $\tilde{U}(t) \in \mathrm{Mat}_n(\mathbb{K})$ ei ole kääntyvä. Toisaalta matriisi $\tilde{R}(t) \in \mathrm{Mat}_n(\mathbb{K})$ on kääntyvä, joten matriisi $A - tI$ ei ole kääntyvä. Tällöin t on matriisin A ominaisarvo. On siis osoitettu, että $\mathbb{K} \subset \sigma(A)$. Kunta $\mathbb{K}$ on ääretön, sillä sen karakteristika on nolla. Matriisilla A voi olla enintään n ominaisarvoa, joten tämä on ristiriita. □

Matriisin $A \in \mathrm{Mat}_n(\mathbb{K})$ *karakteristinen polynomi* χ_A määritellään asettamalla $\chi_A(x) = p_1(x) \cdots p_n(x)$, missä polynomit $p_j(x)$ ovat lauseen 1.0.25 antamat yksikäsitteiset polynomit. Polynomien $p_j(x)$ yksikäsitteisyys takaa siis sen, että karakteristinen polynomi on hyvin määritelty. Tästä lähtien matriisin karakteristiselle polynomille käytetään vain tätä määritelmää.

LAUSE 1.0.26. *Matriisin $A \in Mat_n(\mathbb{K})$ ominaisarvot ovat täsmälleen karakteristisen polynomin juuret.*

TODISTUS. Väite nähdään oikeaksi havaitsemalla, että seuraavat ovat yhtäpitäviä:

(1) $\lambda \in \mathbb{K}$ on matriisin A ominaisarvo.
(2) Matriisi $A - \lambda I$ ei ole kääntyvä.
(3) Lauseen 1.0.25 merkinnöin matriisi $U(\lambda)$ ei ole kääntyvä.
(4) $p_j(\lambda) = 0$ jollakin $j \in \{1, \ldots, n\}$.
(5) $\chi_A(\lambda) = 0$.

□

ESIMERKKI 1.0.27. Selvitetään matriisin

$$A := \begin{bmatrix} 1 & 2 & 4 \\ -1 & 0 & 2 \\ 3 & -1 & 5 \end{bmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

karakteristinen polynomi. Lasketaan

$$A - xI = \begin{bmatrix} 1-x & 2 & 4 \\ -1 & -x & 2 \\ 3 & -1 & 5-x \end{bmatrix} \overset{\text{Esim. 3.0.14}}{\rightarrow}$$

$$\begin{bmatrix} -1 & -x & 2 \\ 0 & \dfrac{22}{9} & \dfrac{1}{9}(-3x^2+19x+10) \\ 0 & 0 & -\dfrac{9}{22}(x^3-6x^2-3x-28) \end{bmatrix} \overset{M_1(-1), M_2(\frac{9}{22}), M_3(-\frac{22}{9})}{\rightarrow}$$

$$\begin{bmatrix} 1 & x & -2 \\ 0 & 1 & \dfrac{1}{22}(-3x^2+19x+10) \\ 0 & 0 & x^3-6x^2-3x-28 \end{bmatrix},$$

mistä nähdään, että $\chi_A(x) = x^3 - 6x^2 - 3x - 28$.

2. Vertailua perinteiseen määrittelyyn

Tarkoituksena on osoittaa, että edellä esitetty määrittely tuottaa etumerkkiä vaille täsmälleen saman karakteristisen polynomin kuin perinteinen määrittely determinantin avulla. Tässä vaiheessa tiedetään, että molemmilla tavoilla määritellyillä polynomeilla on samat juuret eli ominaisarvot. Se ei kuitenkaan vielä riitä takaamaan, että näillä polynomeilla olisi välttämättä mitään tekemistä keskenään, sillä eihän juuria välttämättä ole kunnassa $\mathbb{K}$ lainkaan. Toisaalta, vaikka näillä polynomeilla olisikin kaikki juuret, kuten esimerkiksi tapauksessa $\mathbb{K} = \mathbb{C}$ aina on, voisi niillä silti olla eri kertaluvut. Seuraava lause kuitenkin takaa sen, etteivät tällaiset tilanteet oikeasti ole mahdollisia, vaan karakteristinen polynomi on oleellisesti sama kummallakin määrittelyllä.

LAUSE 2.0.28. *Olkoon $A \in Mat_n(\mathbb{K})$. Tällöin*

$$\chi_A(x) = \pm det(A - xI).$$

TODISTUS. Olkoon $A - xI = R(x)U(x)$ lauseen 1.0.25 antama esitys. Koska $U(x)$ on yläkolmiomatriisi, $\det(U(x)) = p_1(x)\cdots p_n(x)$. Tällöin $\chi_A(x) = p_1(x)\cdots p_n(x) = \det(U(x))$. Matriisi $R(x)$ on kääntyvä, joten lemman 1.0.9 nojalla $\det(R(x)) = a \in \mathbb{K}\setminus\{0\}$. Toisaalta

$$\det(A - xI) = \det(R(x)U(x)) \overset{(i)}{=} \det(R(x))\det(U(x)) = a\chi_A(x),$$

missä yhtälö (i) seuraa Cauchyn ja Binet'n kaavasta (lause 1.0.8). Polynomin $\det(A - xI)$ johtavan termin kerroin on ± 1 ja polynomi $\chi_A(x)$ on perusmuotoinen, joten $a = \pm 1$. □

Kaikki karakteristiseen polynomiin liittyvät tulokset, jotka on todistettu perinteisellä tavalla määritetylle polynomille, pätevät siis edelleen. Tällaisia ovat esimerkiksi seuraavat lauseet.

LAUSE 2.0.29. *Olkoon $A \in Mat_n(\mathbb{K})$ lohkoyläkolmiomuotoa*

$$A = \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix}.$$

Tällöin $\chi_A(x) = \chi_{A_1}(x)\chi_{A_2}(x)$.

LAUSE 2.0.30. *Olkoot $A, R \in Mat_n(\mathbb{K})$, ja olkoon lisäksi R kääntyvä. Tällöin $\chi_{RAR^{-1}}(x) = \chi_A(x)$.*

Seuraavaksi todistetaan Cayleyn ja Hamiltonin lause tätä uutta määritelmää käyttäen.

3. Cayleyn ja Hamiltonin lause

Olkoon $L : \mathbb{K}^n \to \mathbb{K}^n$ lineaarinen operaattori ja $v \in \mathbb{K}^n$. Vektorin v määrittelemä *syklinen aliavaruus* C_v on vektoreiden $v, Lv, L^2v, \dots$ virittämä alivavaruus eli toisin sanoen

$$C_v := \langle v, Lv, L^2v, \dots, L^kv, \dots \rangle.$$

Syklinen aliavaruus C_v on selvästi L-invariantti. Jos $v = 0$, $C_v = \{0\}$. Jos $v \neq 0$, on olemassa sellainen $1 \leq k \leq n$, jolle

$$L^kv \in \langle v, Lv, L^2v, \dots, L^{k-1}v \rangle,$$

sillä vektorijoukko $\{v, Lv, L^2v, \dots, L^kv\}$ on lineaarisesti riippuva kaikille $k \geq n$.

LAUSE 3.0.31. *Olkoon $L : \mathbb{K}^n \to \mathbb{K}^n$ lineaarinen operaattori ja $v \in \mathbb{K}^n \backslash \{0\}$. Tällöin on olemassa $1 \leq k \leq n$, jolle vektorit $v, Lv, L^2v, \dots, L^{k-1}v$ ovat lineaarisesti riippumattomia ja*

$$C_v = \langle v, Lv, L^2v, \dots, L^{k-1}v \rangle.$$

Lisäksi operaattorin $L|_{C_v}$ matriisiesitys kannan $\{v, Lv, L^2v, \dots, L^{k-1}v\}$ suhteen on muotoa

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & \alpha_0 \\ 1 & 0 & \cdots & 0 & \alpha_1 \\ 0 & 1 & \cdots & 0 & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & \alpha_{k-1} \end{bmatrix},$$

missä $L^kv = \alpha_0 v + \alpha_1 Lv + \cdots + \alpha_{k-1}L^{k-1}v$.

TODISTUS. Olkoon $k \in \{1, \dots, n\}$ pienin sellainen indeksi, jolle

$$L^kv \in \langle v, Lv, L^2v, \dots, L^{k-1}v \rangle.$$

Jos $k = 1$ lineaarinen riippumattomuus on selvä. Oletetaan, että $k > 1$ ja vektorit $\{v, Lv, L^2v, \dots, L^{k-1}v\}$ ovat vastoin väitettä lineaarisesti riippuvia. Tällöin

$$a_0v + a_1Lv + a_2L^2v + \cdots + a_{k-1}L^{k-1}v = 0$$

joillekin $a_0, \ldots, a_{k-1} \in \mathbb{K}$ ja $a_j \neq 0$ jollakin $j \in \{1 \ldots, k-1\}$. Olkoon $l := \max\{j \in \{1 \ldots, k-1\} \mid a_j \neq 0\}$. Tällöin $1 \leq l < k$ ja $L^l v \in \{v, Lv, \ldots, L^{l-1}v\}$, mikä on ristiriita indeksin k valinnan minimaalisuuden kanssa.

Osoitetaan seuraavaksi, että $L^m v \in \langle v, Lv, L^2v, \ldots, L^{k-1}v \rangle$ kaikilla $m = 0, 1, 2, \ldots$. Tästä seuraa, että vektorit $\{v, Lv, L^2v, \ldots, L^{k-1}v\}$ virittävät aliavaruuden C_v. Väite on selvä, kun $m \leq k$. Oletetaan, että jollakin $m \geq k$ $L^m v \in \langle v, Lv, L^2v, \ldots, L^{k-1}v \rangle$. Tällöin

$$
\begin{aligned}
L^{m+1}v &= LL^m v \\
&= L(\gamma_0 v + \gamma_1 Lv + \cdots + \gamma_{k-1}L^{k-1}v) \\
&= \gamma_0 Lv + \gamma_1 L^2 v + \cdots + \gamma_{k-1}L^k v \\
&= \gamma_0 Lv + \gamma_1 L^2 v + \cdots + \gamma_{k-1}(\alpha_0 v + \alpha_1 Lv + \cdots + \alpha_{k-1}L^{k-1}v),
\end{aligned}
$$

missä $L^m v = \gamma_0 v + \gamma_1 Lv + \cdots + \gamma_{k-1}L^{k-1}v$ ja $L^k v = \alpha_0 v + \alpha_1 Lv + \cdots + \alpha_{k-1}L^{k-1}v$. Siten myös $L^{m+1}v \in \langle v, Lv, L^2v, \ldots, L^{k-1}v \rangle$, ja induktioaskel on otettu.

Operaattorin $L|_{C_v}$ matriisiesitys kannan $\{v, Lv, \ldots, L^{k-1}v\}$ suhteen on haluttua muotoa, sillä

$$
\begin{aligned}
L(v) &= Lv \\
L(Lv) &= L^2 v \\
&\vdots \\
L(L^{k-1}v) &= L^k v = \alpha_0 v + \alpha_1 Lv + \cdots + \alpha_{k-1}L^{k-1}v.
\end{aligned}
$$

□

Matriisia

$$
C_p = \begin{bmatrix} 0 & 0 & \cdots & 0 & \alpha_0 \\ 1 & 0 & \cdots & 0 & \alpha_1 \\ 0 & 1 & \cdots & 0 & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & \alpha_{k-1} \end{bmatrix},
$$

missä $k \geq 1$, kutsutaan perusmuotoisen polynomin $p = x^k - \alpha_{k-1}x^{k-1} - \cdots - \alpha_1 x - \alpha_0 \in \mathbb{K}[x]$ *kumppanimatriisiksi.* Astetta 1 olevan polynomin $x - \alpha_0$ kumppanimatriisi on 1×1-matriisi eli skalaari α_0.

LAUSE 3.0.32. *Olkoon $C_p \in Mat_n(\mathbb{K}[x])$, polynomin $p(x) = x^k - \alpha_{k-1}x^{k-1} - \cdots - \alpha_1 x - \alpha_0 \in \mathbb{K}[x]$ kumppanimatriisi. Tällöin*

$$
\chi_{C_p}(x) = p(x).
$$

TODISTUS. Tapaus $n = 1$ on selvä ja, kun $n = 2$, väite nähdään oikeaksi seuraavalla laskulla

$$
C_p - xI = \begin{bmatrix} -x & \alpha_0 \\ 1 & -x + \alpha_1 \end{bmatrix} \overset{\substack{P_{1\,2} \\ A_{1\,2}(x)}}{\longrightarrow} \begin{bmatrix} 1 & \alpha_0 \\ 0 & -x + \alpha_1 x + \alpha_0 \end{bmatrix}.
$$

Oletetaan, että $n > 2$. Asetetaan $R_0 := I$ ja $R_j := A_{j\,j+1}(x^j)P_{j\,j+1}$ kaikille $j = 1,\dots,n-2$. Asetetaan

$$B_j := \begin{bmatrix} -x^{j+1} & 0 & \cdots & 0 & \alpha_0+\alpha_1x+\cdots+\alpha_jx^j \\ 1 & -x & \ddots & \vdots & \alpha_{j+1} \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & -x & \alpha_{n-2} \\ 0 & \cdots & 0 & 1 & -x+\alpha_{n-1} \end{bmatrix} \in \mathrm{Mat}_{n-j}(\mathbb{K}[x])$$

kaikille $j = 0,\dots,n-2$. Osoitetaan, että kaikille $j = 1,\dots,n-2$

$$R_j\cdots R_0(C_p - xI) = \left[\begin{array}{cccc|cc} 1 & -x & & 0 & 0 & \alpha_1 \\ & \ddots & \ddots & & \vdots & \vdots \\ 0 & & 1 & -x & 0 & \alpha_j \\ \hline \multicolumn{3}{c|}{0} & \multicolumn{3}{c}{B_j} \end{array}\right],$$

ja $R_0(C_p - xI) = B_0$, kun $j = 0$. Väite pätee selvästi tapauksessa $j = 0$. Oletetaan, että väite pätee kaikilla $0,\dots,j-1$, kun $j \geq 1$. Tällöin matriisia R_j vastaavien rivioperaatioiden suorittaminen matriisille $R_{j-1}\cdots R_0(C_p - xI)$ vaikuttaa matriisiin B_{j-1} seuraavasti:

$$B_{j-1} = \begin{bmatrix} -x^{j} & 0 & \cdots & 0 & \alpha_0+\alpha_1x+\cdots+\alpha_{j-1}x^{j-1} \\ 1 & -x & \ddots & \vdots & \alpha_{j} \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & -x & \alpha_{n-2} \\ 0 & \cdots & 0 & 1 & -x+\alpha_{n-1} \end{bmatrix}$$

$$\xrightarrow{P_{1\,2},\,A_{1\,2}(x^j)} \begin{bmatrix} 1 & -x & \cdots & 0 & \alpha_j \\ 0 & -x^{j+1} & \ddots & \vdots & \alpha_0+\alpha_1x+\cdots+\alpha_{j-1}x^{j-1}+\alpha_jx^j \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & -x & \alpha_{n-2} \\ 0 & \cdots & 0 & 1 & -x+\alpha_{n-1} \end{bmatrix}$$

Tästä ja induktio-oletuksesta nähdään, että polynomimatriisi $R_j \cdots R_0(C_p - xI)$ on halutta muotoa, ja induktioaskel on siis otettu. Tällöin

$$
\begin{aligned}
&R_{n-2} \cdots R_0(C_p - xI) = \\
&\begin{bmatrix}
1 & -x & & 0 & \alpha_1 \\
 & \ddots & \ddots & & \vdots \\
0 & & 1 & -x & \alpha_{n-2} \\
0 & \cdots & 0 & -x^{n-1} & \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-2} x^{n-2} \\
0 & \cdots & 0 & 1 & -x + \alpha_{n-1}
\end{bmatrix}
\xrightarrow{\substack{P_{n-1\,n} \\ A_{n-1\,n}(x^{n-1})}} \\
&\begin{bmatrix}
1 & -x & & 0 & \alpha_1 \\
 & \ddots & \ddots & & \vdots \\
0 & & 1 & -x & \alpha_{n-2} \\
0 & \cdots & 0 & 1 & -x + \alpha_{n-1} \\
0 & \cdots & 0 & 0 & \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-2} x^{n-2} + \alpha_{n-1} x^{n-1} - x^n
\end{bmatrix},
\end{aligned}
$$

mistä väite seuraa. □

Lause 3.0.33 (Cayleyn ja Hamiltonin lause). *Olkoon* $A \in Mat_n(\mathbb{K})$. *Tällöin* $\chi_A(A) = 0$.

Todistus. Olkoon L matriisia A standardikannassa vastaava lineaarinen operaattori. Olkoot $x \in \mathbb{K}^n \setminus \{0\}$ ja $C_x = \langle v, Lv, L^2 v, \ldots, L^{k-1} v \rangle$ vastaava syklinen aliavaruus. Oletetaan, että $k < n$. Tällöin joukko $\{v, Lv, L^2 v, \ldots, L^{k-1} v\}$ voidaan täydentää koko avaruuden $\mathbb{K}^n$ kannaksi $\mathcal{K} := \{v, Lv, L^2 v, \ldots, L^{k-1} v, u_1, \ldots, u_l\}$. Merkitään $M := \langle u_1, \ldots, u_l \rangle$. Tällöin avaruudelle $\mathbb{K}^n$ saadaan hajotelma $\mathbb{K}^n = C_x \oplus M$.

Operaattorin L matriisiesitys kannan $\mathcal{K}$ suhteen on muotoa

$$
\begin{bmatrix} C_p & B \\ 0 & D \end{bmatrix},
$$

missä $p = x^k - \alpha_{k-1} x^{k-1} - \cdots - \alpha_0 \in \mathbb{K}[x]$, ja C_p on polynomin p kumppanimatriisi. Koska vektorin v valintaa ei ole mitenkään rajoitettu, riittää osoittaa, että $\chi_A(A)v = 0$. Tämä puolestaan seuraa osoittamalla, että $\chi_A(L|_{C_x})v = 0$, mikä taas seuraa osoittamalla, että $\chi_A(C_p) = 0$. Lauseiden 2.0.29 ja 3.0.32 nojalla

$$
\chi_A(x) = \chi_{C_p}(x) \chi_D(x) = p(x) \chi_D(x),
$$

joten lopulta riittää osoittaa, että $p(C_p) = 0$. Myös tapaus $k = n$ seuraa suoraan tästä, sillä silloin joukko $\{v, Lv, L^2 v, \ldots, L^{k-1} v\}$ muodostaa koko avaruuden $\mathbb{K}^n$ kannan, jonka suhteen operaattorin L matriisiesitys on C_p.

Jokaiselle $j = 1, \ldots, n$ pätee

$$
p(C_p) e_j = p(C_p) C_p^{j-1} e_1 = C_p^{j-1} p(C_p) e_1.
$$

Toisaalta

$$
\begin{aligned}
p(C_p) e_1 &= \left(C_p^k + \alpha_{k-1} C_p^{k-1} + \cdots + \alpha_1 C_p + \alpha_0 I \right) e_1 \\
&= C_p^k e_1 + \alpha_{k-1} C_p^{k-1} e_1 + \cdots + \alpha_1 C_p e_1 + \alpha_0 e_1 \\
&= C_p e_n + \alpha_{n-1} e_n + \cdots + \alpha_1 e_2 + \alpha_0 e_1 \\
&= -(\alpha_{n-1} e_n + \cdots + \alpha_1 e_2 + \alpha_0 e_1) + \alpha_{n-1} e_n + \cdots + \alpha_1 e_2 + \alpha_0 e_1 \\
&= 0.
\end{aligned}
$$

□

Cayleyn ja Hamiltonin lause voidaan toki todistaa monilla eri tavoilla. Edellä on esitetty kaksi tapaa, joista jälkimmäinen pätee yleisessä kunnassa $\mathbb{K}$. Esimerkiksi Joel N. Franklinin kirjassa Matrix Theory on esitetty kaksi täysin erilaista todistusta, jotka myös poikkeavat tässä tutkielmassa esitetyistä todistuksista. Toinen näistä on puhtaasti algebrallinen, kuten lauseen 3.0.33 todistuskin. Tämä kuntaan $\mathbb{K}$ yleistyvä todistus hyödyntää matriisin kofaktoreita [ks. Fanklin, lemma 1 s. 127]. Toinen todistus pätee vain kunnassa $\mathbb{C}$, kuten lauseen 3.0.21 todistus. Se tehdään osittain analyysin keinoin raja-arvoja ja Schurin ja Toeplizin lausetta hyödyntäen [ks. Franklin s. 111-114]. Esimerkiksi näihin todistuksiin verrattuna lauseen 3.0.33 todistuksen vahvuutena voitaneen sen yleispätevyyden lisäksi pitää päättelyn suoraviivaisuutta, vaikka todistus aputuloksineen voi ehkä ollakin hieman teknisen näköinen.

LUKU 5

Lisää polynomimatriiseja

1. Kääntyvien polynomimatriisien ryhmä

Kääntyvät polynomimatriisit muodostavat ryhmän kertolaskun suhteen. Käytetään tälle ryhmälle merkintää $\mathrm{Gl}_n(\mathbb{K}[x])$. Alkeispolynomimatriisit todettiin jo aiemmin kääntyviksi, ja seuraava lause kertoo, että muutkin kääntyvät polynomimatriisit ovat vain niiden tuloja.

LAUSE 1.0.34. *Alkeispolynomimatriisit virittävät ryhmän $Gl_n(\mathbb{K}[x])$.*

TODISTUS. Olkoon $A \in \mathrm{Gl}_n(\mathbb{K}[x])$, jolloin luonnollisesti myös $A^{-1} \in \mathrm{Gl}_n(\mathbb{K}[x])$. Todistuksen ideana on muokata A^{-1} yksikkömatriisiksi rivioperaatioiden avulla. Lauseen 3.0.12 nojalla on olemassa alkeispolynomimatriisien tulo $R \in \mathrm{Gl}_n(\mathbb{K}[x])$ ja yläkolmiomatriisi $U \in \mathrm{Mat}_n(\mathbb{K}[x])$, joille $RA^{-1} = U$. Tällöin yläkolmiomatriisi U on kääntyvä, joten lauseen 1.0.9 mukaan $\det(U) \in \mathbb{K}\backslash\{0\}$. Siten matriisin U diagonaalipolynomien u_{jj} täytyy olla nollasta eroavia vakiopolynomeja. Asetetaan

$$M := \prod_{j=1}^{n} M_j(u_{jj}^{-1}),$$

jolloin matriisi $V := MU = MRA^{-1}$ on yläkolmiomatriisi, jonka diagonaalialkiot ovat ykkösiä, $v_{jj} = 1 \in \mathbb{K}$ kaikilla $j = 1, \ldots, n$. Asetetaan vielä kaikille $i \in \{1, \ldots, n-1\}$

$$Q_i := \prod_{j=i+1}^{n} A_{ji}(-v_{ij}),$$

ja $Q := Q_1 \cdots Q_{n-1}$, jolloin $QV = I$. Toisin sanoen $QMRA^{-1} = I$ eli $A = QMR$. $\square$

HUOMAUTUS 1.0.35. Lauseen 1.0.34 todistus antaa menetelmän polynomimatriisin käänteismatriisin laskemiseksi. Suoritetaan ensin tarvittavat rivioperaatiot, jotta annettu polynomimatriisi saadaan yläkolmiomuotoon. Jos diagonaalialkiot ovat kaikki nollasta eroavia skalaareja alkuperäinen matriisi on kääntyvä ja voidaan jatkaa. Suoritetaan tarvittavat rivioperaatiot, jotta yläkolmiomatriisista saadaan yksikkömatriisi. Tällöin alkuperäisen polynomimatriisin käänteismatriisi on kaikkien suoritettuja rivioperaatioita vastaavien alkeispolynomimatriisien tulo käänteisessä järjestyksessä.

SEURAUS 1.0.36. *Olkoon $U \in Mat_n(\mathbb{K}[x])$ yläkolmiomatriisi ja $det(U) \in \mathbb{K}\backslash\{0\}$. Tällöin U on kääntyvä.*

TODISTUS. Oletuksen $\det(U) \in \mathbb{K}\backslash\{0\}$ nojalla $u_{jj} = [U]_{jj} \in \mathbb{K}\backslash\{0\}$ kaikilla $j = 1, \ldots, n$. Silloin lauseen 1.0.34 todistuksesta nähdään, että on olemassa $R \in \mathrm{Gl}_n(\mathbb{K}[x])$, jolle $RU = I$. Väite seuraa tällöin lauseesta 1.0.6. $\square$

Determinantin avulla polynomimatriiseille saadaan seuraavanlainen kääntyvyysehto.

LAUSE 1.0.37. *Olkoon $A \in Mat_n(\mathbb{K}[x])$. Tällöin A on kääntyvä, jos ja vain jos $det(A) \in \mathbb{K}\backslash\{0\}$*

TODISTUS. Ehdon välttämättömyys on jo todettu lauseessa 1.0.9. Oletetaan siis, että $\det(A) \in \mathbb{K}\backslash\{0\}$. Lauseen 3.0.12 nojalla on olemassa kääntyvä R, jolle $U := RA$ on yläkolmiomatriisi. Riittää osoittaa, että U on kääntyvä. Lauseen 1.0.9 nojalla $\det(R) \in \mathbb{K}\backslash\{0\}$. Tällöin oletuksen ja lauseen 1.0.8 nojalla $\det(U) \in \mathbb{K}\backslash\{0\}$, jolloin väite seuraa seurauksesta 1.0.36. □

2. Smithin normaalimuoto

Kuten aiemmin on todettu, alkeispolynomimatriisilla vasemmalta kertominen vastaa rivioperaatioita taulukossa 1 esitetyllä tavalla. Jos alkeispolynomimatriiseilla kerrotaankin oikealta puolelta, saadaan vastaavat operaatiot sarakkeille. Näitä operaatioita kutsutaan sarakeoperaatioiksi ja ne on esitetty taulukossa 2.

TAULUKKO 1. Rivioperaatiot

Alkeispolynomimatriisi	Vaikutus kerrottaessa vasemmalta
$M_i(\alpha)$	Kerrotaan i:s rivi skalaarilla $\alpha \in \mathbb{K}\backslash\{0\}$.
P_{ij}	Vaihdetaan rivit i ja $j \neq i$.
$A_{ij}(r(x))$	Lisätään rivi i riviin $j \neq i$ kerrottuna polynomilla $r(x)$.

TAULUKKO 2. Sarakeoperaatiot

Alkeispolynomimatriisi	Vaikutus kerrottaessa oikealta
$M_i(\alpha)$	Kerrotaan i:s sarake skalaarilla $\alpha \in \mathbb{K}\backslash\{0\}$.
P_{ij}	Vaihdetaan sarakkeet i ja $j \neq i$.
$A_{ij}(r(x))$	Lisätään sarake j sarakkeeseen $i \neq j$ kerrottuna polynomilla $r(x)$.

Kun rivioperaatioiden lisäksi käytetään myös sarakeoperaatioita, voidaan polynomimatriisin muokkaamista viedä entistä pidemmälle. Aiemmin muokattiin polynomimatriiseja rivioperaatioiden avulla yläkolmiomuotoon, mutta nyt tavoitteena on diagonaalimatriisi rivi- ja sarakeoperaatioita käyttäen. Yleisesti matriisien A ja B sanotaan olevan *ekvivalentit*, jos on olemassa kääntyvät matriisit R ja Q, joille $A = RBQ$. Reaali- ja kompleksikertoimisessa tapauksessa jokainen matriisi on ekvivalentti muotoa $\mathrm{diag}(I_r, 0)$ olevan diagonaalimatriisin kanssa, missä r on kyseisen matriisin ranki [ks. esim. Ayres, s. 43]. Seuraavaksi on siis tavoitteena todistaa polynomimatriiseille hieman vastaavanlainen tulos, jonka mukaan jokainen polynomimatriisi on ekvivalentti diagonaalimatriisin kanssa. Lisäksi diagonaalipolynomeille saadaan eräitä lisäehtoja.

LEMMA 5.1. *Olkoot $n \geq 2$ ja $C \in Mat_n(\mathbb{K}[x])$. Oletetaan lisäksi, että $c_{11} := [C]_{11} \neq 0$. Tällöin on olemassa kääntyvät polynomimatriisit $Q, R \in Gl_n(\mathbb{K}[x])$, joille*

$$RCQ = \begin{bmatrix} p & 0 \\ 0 & D \end{bmatrix},$$

missä D on kokoa $n-1$ oleva polynomimatriisi tai pelkkä polynomi ja $0 \neq p \in \mathbb{K}[x]$. Lisäksi $deg(p) \leq deg(d_{ij})$ kaikilla $d_{ij} := [D]_{ij} \neq 0$.

TODISTUS. Tarkoituksena on muokata polynomimatriisi C haluttuun muotoon käyttäen rivi- ja sarakeoperaatioita. Todistus noudattelee pitkälti samaa ideaa kuin lauseen 3.0.12 todistus. Asetetaan ensin

$$A_0 := C := \begin{bmatrix} p_{11}^0 & p_{12}^0 & \cdots & p_{1n}^0 \\ p_{21}^0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ p_{n1}^0 & * & \cdots & * \end{bmatrix}.$$

Tarkoituksena on määritellä rekursiivisesti matriisit $A_k = R_k A_{k-1} Q_k$, kun $k = 1, 2, \ldots$, missä matriisit R_k vastaavat rivioperaatioita ja matiriisit Q_k sarakeoperaatioita. Rekursioaskel jaetaan selvyyden vuoksi kolmeen eri vaiheeseen. Oletetaan, että jollakin parillisella k polynomimatriisit $A_0, \ldots, A_k$ ovat jo määritelty. Merkitään

$$A_k = \begin{bmatrix} p_{11}^k & p_{12}^k & \cdots & p_{1n}^k \\ p_{21}^k & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ p_{n1}^k & * & \cdots & * \end{bmatrix}.$$

ja oletetaan, että $p_{11}^k \neq 0$. Tällöin polynomimatriisit A_{k+1} ja A_{k+2} määritellään valitsemalla rivi- ja sarakeoperaatioita vastaavat matriisit R_{k+1}, R_{k+2}, Q_{k+1} ja Q_{k+2} seuraavasti:

Vaihe 1: Olkoon p_{ij}^k matriisin A_k asteeltaan pienin nollasta eroava polynomi. Polynomi p_{ij}^k ei ole välttämättä yksikäsitteinen, mutta oletuksen $p_{11}^k \neq 0$ nojalla sellainen on kuitenkin olemassa. Jos polynomi p_{11}^k on asteeltaan pienin nollasta eroava polynomi, niin valitaan $i = j = 1$. Tätä vaatimusta ei välttämättä tarvittaisi tässä todistuksessa, mutta siitä on hyötyä myöhemmin. Tällä varmistutaan siitä, että matriisi A_k ei enää muutu sen jälkeen kun se on kerran saatu haluttuun muotoon. Tehdään seuraavaksi tarvittavat rivin- ja sarakkeenvaihdot, jotta polynomi p_{ij}^k saadaan vasempaan ylänurkkaan. Asetetaan siis

$$R_{k+1} := \begin{cases} I, \text{ kun } i = 1 \\ P_{1i}, \text{ kun } i \neq 1, \end{cases}$$

ja vastaavasti

$$Q_{k+1} := \begin{cases} I, \text{ kun } j = 1 \\ P_{1j}, \text{ kun } j \neq 1. \end{cases}$$

Vaihe 2: Jos $p_{i1}^{k+1} = 0$ kaikilla $i = 2, \ldots, n$, asetetaan $R_{k+2} = I$. Muussa tapauksessa olkoon $i \in \{2, \ldots, n\}$ sellainen, että $p_{i1}^{k+1} \neq 0$. Tällöin vaiheen 1 mukaan $\deg(p_{11}^{k+1}) \leq \deg(p_{i1}^{k+1})$. Polynomien jakoyhtälön mukaan

$$p_{i1}^{k+1} = r_i^{k+1} p_{11}^{k+1} + p_{i1}^{k+2},$$

missä $r_i^{k+1}, p_{i1}^{k+2} \in \mathbb{K}[x]$ ja

$$\deg(p_{i1}^{k+2}) < \deg(p_{11}^{k+1}). \tag{5.1}$$

Asetetaan kaikille $i = 2, \ldots, n$

$$R_{k+2,i} = \begin{cases} A_{1i}(-r_i^{k+1}), \text{ jos } p_{i1}^{k+1} \neq 0 \\ I, \text{ jos } p_{i1}^{k+1} = 0 =: p_{i1}^{k+2} \end{cases}$$

ja $R_{k+2} := R_{k+2,n} \cdots R_{k+2,2}$.

Vaihe 3: Tässä vaiheessa tehdään olleellisesti sarakkeille samat operaatiot, jotka vaiheessa 2 tehtiin riveille. Jos $p_{1j}^{k+1} = 0$ kaikilla $j = 2, \ldots, n$, asetetaan $Q_{k+2} = I$. Muussa tapauksessa olkoon $j \in \{2, \ldots, n\}$ sellainen, että $p_{1j}^{k+1} \neq 0$. Tällöin vaiheen 1 mukaan $\deg(p_{11}^{k+1}) \leq \deg(p_{1j}^{k+1})$. Edelleen polynomien jakoyhtälön avulla voidaan kirjoittaa

$$p_{1j}^{k+1} = r_j^{k+1} p_{11}^{k+1} + p_{1j}^{k+2},$$

missä $r_j^{k+1}, p_{1j}^{k+2} \in \mathbb{K}[x]$ ja

$$\deg(p_{1j}^{k+2}) < \deg(p_{11}^{k+1}). \tag{5.2}$$

Asetetaan kaikille $j = 2, \ldots, n$

$$Q_{k+2,j} = \begin{cases} A_{j1}(-r_j^{k+1}), \text{ jos } p_{1j}^{k+1} \neq 0 \\ I, \text{ jos } p_{1j}^{k+1} = 0 =: p_{1j}^{k+2} \end{cases}$$

ja $Q_{k+2} := Q_{k+2,2} \cdots Q_{k+2,n}$. Tällöin

$$A_{k+2} = R_{k+2} A_{k+1} Q_{k+2} = \begin{bmatrix} p_{11}^{k+2} & p_{12}^{k+2} & \cdots & p_{1n}^{k+2} \\ p_{21}^{k+2} & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ p_{n1}^{k+2} & * & \cdots & * \end{bmatrix}.$$

missä $p_{11}^{k+2} = p_{11}^{k+1} \neq 0$, $\deg(p_{j1}^{k+2}) < \deg(p_{11}^{k+1})$ ja $\deg(p_{1j}^{k+2}) < \deg(p_{11}^{k+1})$ kaikilla $j = 2, \ldots, n$.

Näin polynomimatriisit A_k tulevat rekursiiviseti määritellyiksi kaikille $k = 0, 1, 2, \ldots$. Tehdään seuraavaksi vastaavat huomiot kuin lauseen 3.0.12 todistuksessa. Kaikilla $k \in \mathbb{N} \cup \{0\}$ ja $j = 2, \ldots n$ polynomeille p_{j1}^k ja p_{1j}^k pätevät seuraavat:

(a): Jos $p_{j1}^k = 0$, myös $p_{j1}^{k+1} = 0$. Vastaavsti, jos $p_{1j}^k = 0$, myös $p_{1j}^{k+1} = 0$.

(b): Jos k on parillinen ja $p_{j1}^{k+2} \neq 0$, pätee $\deg(p_{j1}^{k+2}) < \deg(p_{j1}^k)$. Samoin, jos k on parillinen ja $p_{1j}^{k+2} \neq 0$, pätee $\deg(p_{1j}^{k+2}) < \deg(p_{1j}^k)$.

Väitteiden (a) ja (b) todistaminen tapahtuu samoin kuin lauseen 3.0.12 tapauksessakin. Ainoa ero on, että nyt joudutaan ottamaan huomioon myös sarakeoperaatiot. Olkoon $j \in \{2, \ldots, n\}$. Väite (a) seuraa tällöin suoraan vaiheen 1 määrittelystä, jos k on parillinen tai nolla. Jos taas k on pariton, väite (a) seuraa vaiheen 2 määrittelystä rivioperaatioiden tapauksessa ja vaiheen 3 määrittelystä sarakeoperaatioiden tapauksessa.

Oletetaan (b)-kohdan todistamista varten, että $p_{j1}^{k+2} \neq 0 \neq p_{1j}^{k+2}$. Silloin (a)-kohdan nojalla myös $p_{j1}^k \neq 0 \neq p_{1j}^k$. Tällöin rivioperaatioiden tapauksessa saadaan, että

$$\deg(p_{j1}^{k+2}) \overset{(i)}{<} \deg(p_{11}^{k+1}) \overset{(ii)}{\leq} \deg(p_{j1}^k).$$

Epäyhtälö (i) seuraa yhtälöstä (5.1). Epäyhtälö (ii) seuraa vaiheen 1 määrittelystä, jonka mukaan p_{11}^{k+1} on matriisin A_k ensimmäisen sarakkeen asteeltaan pienin nollasta eroava polynomi. Vastaava tulos sarakeoperaatioille on

$$\deg(p_{1j}^{k+2}) \overset{(i)}{<} \deg(p_{11}^{k+1}) \overset{(ii)}{\leq} \deg(p_{1j}^{k}).$$

Tässä epäyhtälö (i) seuraa yhtälöstä ja (5.2). Epäyhtälö (ii) seuraa taas vaiheen 1 määrittelystä, jonka mukaan p_{11}^{k+1} on matriisin A_k ensimmäisen rivin asteeltaan pienin nollasta eroava polynomi.

Olkoon $j \in \{2, \dots, n\}$. Oletetaan, että $p_{j1}^{k} \neq 0$ kaikilla $k \in \mathbb{N} \cup \{0\}$ tai $p_{1j}^{k} \neq 0$ kaikilla $k \in \mathbb{N} \cup \{0\}$. Tällöin tuloksen (b) nojalla $\deg(p_{j1}^{k+2}) < \deg(p_{j1}^{k}) < \deg(p_{j1}^{0})$ kaikilla parillisilla $k \in \mathbb{N}$ tai $\deg(p_{1j}^{k+2}) < \deg(p_{1j}^{k}) < \deg(p_{1j}^{0})$ kaikilla parillisilla $k \in \mathbb{N}$. Kumpikaan vaihtoehto ei ole mahdollinen, joten päädytään ristiriitaan. Siispä on olemassa sellainet $m_i \in \mathbb{N} \cup \{0\}$ ja $m_j \in \mathbb{N} \cup \{0\}$, joille $p_{i1}^{m_i} = 0 = p_{1j}^{m_j}$. Asetetaan $m := \max\{m_i, m_j | i, j = 2, \dots, n\}$. Tällöin tuloksen (a) nojalla $p_{i1}^{m} = 0 = p_{1j}^{m}$ kaikilla $i, j = 2, \dots, n$. Matriisi A_m on siis haluttua muotoa eli

$$A_m = \begin{bmatrix} p_{11}^{m} & 0 \\ 0 & D \end{bmatrix},$$

missä D on kokoa $n-1$ oleva neliömatriisi tai polynomi.

Vaiheissa 2 ja 3 suoritettavissa operaatioissa polynomien p_{ij}^{k}, missä $i, j > 1$, asteet eivät pienene. Siksi vaiheen 1 määrittelystä nähdään suoraan, että $\deg(p_{11}^{m}) \leq \deg(d_{ij})$ kaikilla $d_{ij} := [D]_{ij} \neq 0$.

□

LEMMA 5.2. *Olkoon p_{11}^{1} kuten lemman 5.1 todistuksessa. Oletetaan seuraavat:*

- *Jollakin $i > 1$, polynomille p_{i1}^{1} on $deg(p_{i1}^{1}) \geq deg(p_{11}^{1})$.*
- *Polynomi p_{i1}^{1} ei ole jaollinen polynomilla p_{11}^{1}.*

Tällöin lemman 5.1 polynomimatriisit R ja Q voidaan valita niin, että

$$deg(p) < deg(c_{11}).$$

TODISTUS. Väite voidaan perustella tarkastelemalla lemman 5.1 todistusta. Vaiheiden 1,2 ja 3 määrittelyistä huomataan, että $\deg(p_{11}^{k}) \geq \deg(p_{11}^{k+1})$ kaikilla $k = 1, 2, \dots$. Koska p_{i1}^{1} ei ole jaollinen polynomilla p_{11}^{1}, vaiheen 2 määrittelyn mukaan $p_{i1}^{2} \neq 0$ ja $\deg(p_{i1}^{2}) < \deg(p_{11}^{1})$. Toisaalta vaiheen 1 määrittelyn mukaan $\deg(p_{11}^{3}) \leq \deg(p_{i1}^{2})$. Näin ollen

$$\deg(p_{11}^{m+3}) \leq \deg(p_{11}^{3}) \leq \deg(p_{i1}^{2}) < \deg(p_{11}^{1}) \leq \deg(c_{11}).$$

Vaiheista 1,2 ja 3 nähdään, että $p_{11}^{m+k} = p_{11}^{m}$ kaikille $k = 0, 1, 2 \dots$, joten erityisesti $p_{11}^{m+3} = p_{11}^{m} = p$. □

LEMMA 5.3. *Olkoot $n \geq 2$ ja $C \in Mat_n(\mathbb{K}[x])$. Oletetaan lisäksi, että $c_{11} := [C]_{11} \neq 0$. Tällöin on olemassa kääntyvät polynomimatriisit $Q, R \in Gl_n(\mathbb{K}[x])$, joille*

$$RCQ = \begin{bmatrix} p & 0 \\ 0 & D \end{bmatrix},$$

missä $0 \neq p \in \mathbb{K}[x]$ jakaa jokaisen polynomin $d_{ij} := [D]_{ij}$, kun $i, j = 1, \dots, n-1$.

TODISTUS. Asetetaan

$$C := A_0 := \begin{bmatrix} p_{11}^0 & p_{12}^0 & \cdots & p_{1n}^0 \\ p_{21}^0 & p_{22}^0 & \cdots & p_{2n}^0 \\ \vdots & & \ddots & \vdots \\ p_{n1}^0 & p_{n2}^0 & \cdots & p_{nn}^0 \end{bmatrix}$$

Tarkoituksena on määritellä rekursiivisesti polynomimatriisit $A_k = R_k A_{k-1} Q_k$, kun $k = 1, 2, \ldots$, missä polynomimatriisit R_k vastaavat jälleen rivioperaatioita ja polynomimatriisit Q_k sarakeoperaatioita. Tehdään oletus, että jollakin parillisella $k \in \mathbb{N} \cup \{0\}$ polynomimatriisit $A_0, \ldots, A_k$ ovat jo määriteltyjä. Tällöin polynomimatriisit A_{k+1} ja A_{k+2} määritellään valitsemalla rivi- ja sarakeoperaatioita vastaavat matriisit R_{k+1}, R_{k+2}, Q_{k+1} ja Q_{k+2} seuraavasti:

Vaihe 1: Lemman 5.1 nojalla on olemassa matriisit $R_{k+1}, Q_{k+1} \in \mathrm{Gl}_n(\mathbb{K}[x])$, joille

$$A_{k+1} := R_k A_k Q_k = \begin{bmatrix} p_{11}^{k+1} & p_{12}^{k+1} & \cdots & p_{1n}^{k+1} \\ p_{21}^{k+1} & p_{22}^{k+1} & \cdots & p_{2n}^{k+1} \\ \vdots & & \ddots & \vdots \\ p_{n1}^{k+1} & p_{n2}^{k+1} & \cdots & p_{nn}^{k+1} \end{bmatrix},$$

missä $p_{j1}^{k+1} = p_{1j}^{k+1} = 0$ kaikilla $j = 2, \ldots, n$. Lisäksi $\deg(p_{11}^{k+1}) \leq \deg(p_{ij}^{k+1})$ kaikilla $p_{ij}^{k+1} \neq 0$.

Vaihe 2: Asetetaan $R_{k+2} = I$. Jos jokainen polynomi p_{ij}^{k+1} on jaollinen polynomilla p_{11}^{k+1}, asetetaan myös $Q_{k+2} = I$. Muutoin olkoon $p_{ij}^{k+1} \neq 0$ se polynomi, joka ei ole jaollinen polynomilla p_{11}^{k+1}. Tällöin asetetaan

$$Q_{k+2} = A_{j1}(1) \text{ ja } A_{k+2} = R_{k+2} A_{k+1} Q_{k+2}.$$

Seuraavaksi pitää vielä varmistua siitä, että haluttu lopputulos saavutetaan äärellisellä määrällä edellä määriteltyjä matriiseja R_k ja Q_k. Suoraan vaiheiden 1 ja 2 määrittelyistä ja lemmasta 5.1 nähdään, että $p_{11}^k \neq 0$ kaikilla $k = 0, 1, 2, \ldots$. Jos jollakin parillisella $m \geq 2$ vaiheessa 2 jokainen polynomi p_{ij}^{m-1} on jaollinen polynomilla p_{11}^{m-1}, matriisi A_m toteuttaa väitteen ehdot. Riittää siis osoittaa, että jollakin parillisella $m \geq 2$ jokainen polynomi p_{ij}^{m-1} on jaollinen polynomilla p_{11}^{m-1}.

Tehdään antiteesi, eli oletetaan, että jokaisella parillisella $k \geq 2$ on olemassa polynomi p_{ij}^{k-1}, joka ei ole jaollinen polynomilla p_{11}^{k-1}. Tällöin jollakin $i \geq 2$ pätee $\deg(p_{i1}^k) \geq \deg(p_{11}^k)$ ja polynomi p_{i1}^k ei ole jaollinen polynomilla p_{11}^k. Silloin lemman 5.2 nojalla $\deg(p_{11}^{k+1}) < \deg(p_{11}^k)$. Kaikille parillisille k pätee $p_{11}^{k+2} = p_{11}^{k+1}$, joten edellä todetun nojalla

$$\deg(p_{11}^{k+2}) < \deg(p_{11}^k)$$

kaikilla parillisilla $k \geq 2$, mikä on ristiriita, sillä $\deg(p_{11}^k) \geq 0$ kaikilla $k = 0, 1, 2, \ldots$.

□

LEMMA 5.4. *Olkoon $D \in Mat_n(\mathbb{K}[x])$, $0 \neq p \in \mathbb{K}[x]$ ja $R, Q \in Gl_n(\mathbb{K}[x])$. Oletetaan, että polynomi p jakaa matriisin D jokaisen kerroinpolynomin. Tällöin p jakaa myös matriisin RDQ jokaisen kerroinpolynomin.*

TODISTUS. Lauseen 1.0.34 nojalla ryhmä $\mathrm{Gl}_n(\mathbb{K}[x])$ on alkeispolynomimatriisien virittämä, eli erityisesti matriisit R ja Q ovat silloin alkeispolynomimatriisien tuloja.

Siten riittää osoittaa, että jaollisuus säilyy kerrottaessa alkeispolynomimatriisilla vasemmalta ja oikealta eli toisin sanoen, että jaollisuus säilyy jokaisessa rivi- ja sarakeoperaatiossa. On selvää, että jaollisuus säilyy kerrottaessa riviä tai saraketta skalaarilla sekä rivien ja sarakkeiden vaihdossa.

Olkoon $r \in \mathbb{K}[x]$ ja $k, l \in \{1, \dots, n\}$. Suoritetaan rivioperaatio $A_{kl}(r)$. Tällöin

$$[A_{kl}(r)D]_{ij} = \begin{cases} d_{ij}, \text{ kun } i \neq k \\ d_{kj} + rd_{lj}, \text{ kun } i = k. \end{cases}$$

Oletuksen nojalla polynomit d_{ij} ovat jaollisia polynomilla p kaikilla $i, j = 1, \dots, n$, joten myös polynomi $d_{kj} + rd_{lj}$ on jaollinen polynomilla p. Jaollisuus säilyy siis rivioperaatiossa $A_{kl}(r)$.

Vastaavasti voidaan tarkastella sarakeoperaatiota $A_{lk}(r)$. Tällöin

$$[DA_{lk}(r)]_{ij} = \begin{cases} d_{ij}, \text{ kun } j \neq k \\ d_{ik} + rd_{il}, \text{ kun } j = k. \end{cases}$$

Polynomi $d_{ik} + rd_{il}$ on myös jaollinen polynomilla p samoin perustein kuin edellä, eli jaollisuuden säilyminen on varmistettu myös sarakeoperaatiossa $A_{lk}(r)$. □

Otetaan seuraavaa lemmaa varten käyttöön uusi merkintä. Matriisille $A \in \mathrm{Mat}_n(\mathbb{K}[x])$ merkitään

$$\Delta_{A,k} := \begin{cases} 0, \text{ jos } \det(A_{I,J}) = 0 \text{ kaikilla } I, J \subset \{1, \dots, k\}, \#I = \#J = k. \\ \mathrm{syt}\{\det(A_{I,J}) \mid I, J \subset \{1, \dots, k\} \text{ ja } \#I = \#J = k\} \text{ muulloin.} \end{cases}$$

Lemma 5.5. *Olkoon $A \in Mat_n(\mathbb{K}[x])$ ja $B = RAQ$, missä $R, Q \in Gl_n(\mathbb{K}[x])$. Tällöin*

$$\Delta_{B,k} = \Delta_{A,k}$$

kaikilla $k = 1, \dots, n$.

Todistus. Olkoot $k \in \{1, \dots, n\}$ ja $B_{I,J}$, jokin matriisin B alimatriisi, missä $I, J \subset \{1, \dots, n\}$ ja $\#I = \#J = k$. Lauseen 1.0.8 nojalla

$$\begin{aligned} \det(B_{I,J}) = \det((RAQ)_{I,J}) &= \sum_{\substack{K \subset \{1,\dots,n\} \\ \#K = k}} \det((RA)_{I,K}) \det(Q_{K,J}) \\ &= \sum_{\substack{K \subset \{1,\dots,n\} \\ \#K = k}} \left(\sum_{\substack{L \subset \{1,\dots,n\} \\ \#L = k}} \det(R_{I,L}) \det(A_{L,K}) \right) \det(Q_{K,J}) \qquad (5.3) \\ &= \sum_{\substack{K \subset \{1,\dots,n\} \\ \#K = k}} \sum_{\substack{L \subset \{1,\dots,n\} \\ \#L = k}} \det(R_{I,L}) \det(Q_{K,J}) \det(A_{L,K}). \end{aligned}$$

Tarkastellaan ensin tapausta $\Delta_{A,k} = 0$. Koska voidaan kirjoittaa myös $A = Q^{-1}BR^{-1}$, matriisien A ja B rooleissa ei ole eroa. Siksi riittää osoittaa, että myös $\Delta_{B,k} = 0$. Koska $\Delta_{A,k} = 0$, määritelmän mukaan $\det(A_{L,K}) = 0$ kaikille $K, L \in \{1, \dots, k\}$, missä $\#K = \#L = k$. Tällöin yhtälön (5.3) nojalla $\det(B_{I,J}) = 0$ kaikille $I, J \in \{1, \dots, k\}$, missä $\#I = \#J = k$, eli $\Delta_{B,k} = 0$. Siis $\Delta_{A,k} = 0$, jos ja vain jos $\Delta_{B,k} = 0$

Voidaan siis olettaa, että $\Delta_{A,k} \neq 0$, jolloin myös $\Delta_{B,k} \neq 0$. Symmetrisyyden vuoksi riittää osoittaa, että

$$\Delta_{A,k} | \det(B_{I,J}).$$

Määritelmänsä mukaan $\Delta_{A,k}$ jakaa jokaisen polynomin $\det(A_{L,K})$. Erityisesti se jakaa silloin yhtälön (5.3) oikean puolen summan jokaisen termin ja siten myös vasemman puolen eli polynomin $\det(B_{I,J})$. □

LAUSE 2.0.38 (Smithin normaalimuoto). *Olkoon $C \in Mat_n(\mathbb{K}[x])$. Tällöin on olemassa kääntyvät polynomimatriisit $Q, R \in Gl_n(\mathbb{K}[x])$, joille*

$$RCQ = diag(p_1, \dots, p_n).$$

ja seuraavat ovat voimassa:

(1) Kaikki nollasta eroavat polynomit $p_j \in \mathbb{K}[x]$ ovat perusmuotoisia.
(2) Jokainen nollasta eroava polynomi p_j jakaa polynomin p_{j+1}, kun $j < n$.
(3) Jos $p_j = 0$ jollakin $j \in \{1, \dots, n-1\}$, niin myös $p_{j+1} = 0$.
(4) Polynomit p_j ovat yksikäsitteisiä.

TODISTUS. Todistetaan ensin olemassaolo, ja jätetään yksikäsitteisyyden todistaminen viimeiseksi. Todistus voidaan suorittaa induktiolla matriisin koon n suhteen. Polynomien p_j perusmuotoisuudesta ei tarvitse erikseen huolehtia. Jos kaikki nollasta eroavat polynomit p_j eivät ole perusmuotoisia, voidaan kyseiset rivit vielä lopuksi kertoa sopivilla skalaareilla samaan tapaan kuin lauseen 3.0.15 todistuksessa on tehty, ja sisällyttää näitä operaatioita vastaavat alkeispolynomimatriisit matriisiin R.

Tarkastellaan ensin tapausta $n = 2$. Jos $C = 0$, väite on selvä. Tällöin voidaan valita $R = Q = I$. Oletetaan, että $C \neq 0$. Voidaan myös olettaa suoraan, että $c_{11} \neq 0$, sillä mikä tahansa kerroinpolynomi voidaan aina siirtää tarvittaessa vasempaan ylänurkkaan rivin- ja sarakkeen vaihtojen avulla lemman 5.1 todistuksessa esitetyllä tavalla. Tällöin väite seuraa suoraan lemmasta 5.3. Erityisesti myös listan kolmas väite seuraa suoraan tästä lemmasta, sillä lemman ja oletuksen $c_{11} \neq 0$ mukaan $p_1 \neq 0$. Tehdään siis seuraavaksi induktio-oletus. Oletetaan, että jollakin $n > 2$ väite pätee kaikille enintään kokoa $n-1$ oleville polynomimatriiseille. Olkoon C kokoa n oleva polynomimatriisi. Voidaan jälleen olettaa, että $C \neq 0$, sillä muutoin väite on selvä. Lisäksi voidaan edelleen olettaa, että $c_{11} \neq 0$. Lemman 5.3 nojalla on olemassa sellaiset $R_1, Q_1 \in \mathrm{Gl}_n(\mathbb{K}[x])$, joille

$$R_1 C Q_1 = \begin{bmatrix} p_1 & 0 \\ 0 & D \end{bmatrix},$$

missä D on kokoa $n-1$ oleva polynomimatriisi. Lisäksi polynomi $p_1 \neq 0$ jakaa jokaisen polynomin $d_{ij} := [D]_{ij}$. Induktio-oletuksen nojalla on olemassa sellaiset $R', Q' \in \mathrm{Gl}_{n-1}(\mathbb{K}[x])$, joille

$$R' D Q' = \mathrm{diag}(p_2, \dots, p_n),$$

missä polynomit p_j toteuttavat väitteen ehdot. Asetetaan

$$R_2 := \begin{bmatrix} 1 & 0 \\ 0 & R' \end{bmatrix}, \ Q_2 := \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix},$$

$R := R_2R_1$ ja $Q := Q_1Q_2$. Tällöin $R, Q \in \mathrm{Gl}_n(\mathbb{K}[x])$ ja

$$RCQ = \begin{bmatrix} p_1 & 0 \\ 0 & R'DQ' \end{bmatrix} = \begin{bmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & p_n \end{bmatrix}.$$

Koska p_1 jakaa jokaisen polynomin d_{ij}, missä $i, j = 1, \ldots, n-1$, lemman 5.4 nojalla se jakaa myös jokaisen polynomin p_j, missä $j = 2, \ldots, n$.

Vielä on todistettava polynomien p_j yksikäsitteisyys. Oletetaan, että on polynomimatriisit

$$D := RCQ = \mathrm{diag}(p_1, \ldots, p_n) \text{ ja } D' := R'CQ' = \mathrm{diag}(p_1', \ldots, p_n)',$$

missä polynomit p_j ja p_j' toteuttavat väitteessä esitetyt vaatimukset kaikille $j = 1, \ldots, n$. Koska polynomimatriisit D,D' ja C ovat ekvivalentit, niillä on oltava sama ranki. Tämä seuraa suoraan polynomimatriisien rankin määritelmästä ja vastaavasta tuloksesta kuntakertoimisille matriiseille. Silloin erityisesti $D = D' = 0$, jos $C = 0$. Voidaan siis olettaa, että $C \neq 0$. Tällöin jollakin $k \in \{1, \ldots, n\}$ $D = \mathrm{diag}(p_1, \ldots, p_k, 0, \ldots, 0)$ ja $D' = \mathrm{diag}(p_1', \ldots, p_k', 0, \ldots, 0)$, missä $p_j \neq 0 \neq p_j'$ kaikilla $j = 1, \ldots, k$. Polynomimatriisien D ja D' diagonaaleilla on sama määrä nollasta eroavia polynomeja, sillä muutoin niillä olisi eri rankit.

Smithin normaalimuodon ehdoista huomataan, että

$$\Delta_{D,j} = p_1 \cdots p_j \text{ ja } \Delta_{D',j} = p_1' \cdots p_j' \tag{5.4}$$

kaikille $j = 1, \ldots, k$. Koska polynomimatriisit D ja D' ovat ekvivalentit, lemman 5.5 nojalla

$$\Delta_{D,j} = \Delta_{D',j}$$

kaikilla $j = 1, \ldots, k$. Tällöin yhtälöiden (5.4) nojalla

$$\begin{aligned} p_1 &= p_1' \\ p_1p_2 &= p_1'p_2', \\ &\vdots \\ p_1 \cdots p_k &= p_1' \cdots p_k'. \end{aligned}$$

Tästä seuraa induktiivisesti, että $p_j = p_j'$ kaikilla $j = 1, \ldots, k$. □

Lauseessa 2.0.38 esiintyvää diagonaalimatriisia RCQ kutsutaan matriisin C *Smithin normaalimuodoksi*. Polynomimatriisin sanotaan olevan Smithin normaalimuodossa, jos se on diagonaalinen ja sen diagonaalipolynomit toteuttavat ehdot 1-3. Smithin normaalimuoto on vastaavasti olemassa myös kokonaislukukertoimisille matriiseille. Siis jokainen kokonaislukukertoiminen matriisi on ekvivalentti sellaisen diagonaalimatriisin kanssa, jonka lävistäjällä olevat kokonaisluvut toteuttavat jaollisuusehdot. Toisaalta kokonaislukujen rengas $\mathbb{Z}$ ja kuntakertoiminen polynomirengas $\mathbb{K}[x]$ ovat molemmat pääideaalialueita, ja itseasiassa Smithin normaalimuoto voidaankin yleistää myös matriiseille, joiden kertoimet ovat mielivaltaisesta pääideaalialueesta [ks. Storjohan, luku 2.2].

HUOMAUTUS 2.0.39. Ekvivalenteilla polynomimatriiseilla on sama Smithin normaalimuoto. Tämä seuraa normaalimuodon yksikäsitteisyydestä.

ESIMERKKI 2.0.40. Kääntyvän polynomimatriisin Smithin normaalimuoto on yksikkömatriisi I. Olkoon $A \in \mathrm{Gl}_n(\mathbb{K}[x])$ ja $S = RAQ$ sen Smithin normaalimuoto. Muunnosmatriisit R ja Q ovat aina kääntyviä, joten S on kääntyvien polynomimatriisien tulona kääntyvä. Tällöin lauseen 1.0.9 nojalla $\det(S) \in \mathbb{K}\backslash\{0\}$, joten diagonaalipolynomit ovat välttämättä nollasta eroavia skalaareita. Toisaalta ne ovat myös perusmuotoisia, eli ainoa mahdollisuus on, että ne ovat ykkösiä.

LAUSE 2.0.41. *Olkoot $n \geq 2$ ja $A = diag(p_1, \ldots, p_n) \in Mat_n(\mathbb{K}[x])$. Oletetaan lisäksi, että polynomit $p_1, \ldots, p_n$ ovat nollasta eroavia perusmuotoisia polynomeja, joille $syt\{p_1 \cdots p_j, p_{j+1}\} = 1$ kaikilla $j = 1, \ldots n-1$. Tällöin polynomimatriisin A Smithin normaalimuoto on $diag(1, \ldots, 1, p_1 \cdots p_n)$.*

TODISTUS. Olkoot ensin $n = 2$ ja $S = \tilde{R}A\tilde{Q} = \mathrm{diag}(q_1, q_2)$ matriisin A Smithin normaalimuoto, missä $\tilde{R}, \tilde{Q} \in \mathrm{Gl}_2(\mathbb{K}[x])$. Lemman 5.4 nojalla $q_1 | p_j$, kun $j = 1, 2$. Tällöin erityisesti $q_1 | 1 = \mathrm{syt}\{p_1, p_2\}$, joten $q_1 = 1$. Toisaalta lauseiden 1.0.8 ja 1.0.9 nojalla

$$p_1 p_2 = \det(A) = \det(\tilde{R})\det(S)\det(\tilde{Q}) = a q_1 q_2 = a q_2,$$

jollekin $a \in \mathbb{K}\backslash\{0\}$. Molemmat polynomit $p_1 p_2$ ja q_2 ovat perusmuotoisia, joten $a = 1$.

Oletetaan seuraavaksi, että jollain $n \geq 3$ väite pätee kaikille enintään kokoa $n-1$ oleville polynomimatriiseille. Olkoon

$$S_1 = R_1 \mathrm{diag}(p_1, \ldots, p_{n-1}) Q_1 = \mathrm{diag}(q_1, \ldots, q_{n-1})$$

polynomimatriisin $\mathrm{diag}(p_1, \ldots, p_{n-1})$ Smithin normaalimuoto, missä $R_1, Q_1 \in \mathrm{Gl}_{n-1}(\mathbb{K}[x])$. Tällöin induktio-oletuksen nojalla $q_j = 1$ kaikilla $j = 1, \ldots, n-2$ ja $q_{n-1} = p_1 \cdots p_{n-1}$. Oletuksen nojalla $\mathrm{syt}\{q_{n-1}, p_n\} = 1$, joten induktio-oletuksen mukaan polynomimatriisin $\mathrm{diag}(q_{n-2}, p_n)$ Smithin normaalimuoto on

$$S_2 = R_2 \mathrm{diag}(q_{n-2}, p_n) Q_2 = \mathrm{diag}(1, p_1 \cdots p_n),$$

missä $R_2, Q_2 \in \mathrm{Gl}_2(\mathbb{K}[x])$. Asetetaan

$$R := \begin{bmatrix} I_{n-2} & 0 \\ 0 & R_2 \end{bmatrix} \begin{bmatrix} R_1 & 0 \\ 0 & 1 \end{bmatrix} \text{ ja } Q := \begin{bmatrix} Q_1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I_{n-2} & 0 \\ 0 & Q_2 \end{bmatrix},$$

jolloin

$$RAQ = \begin{bmatrix} I_{n-1} & 0 \\ 0 & p_1 \cdots p_n \end{bmatrix}.$$

Induktioväite seuraa tällöin Smithin normaalimuodon yksikäsitteisyysdestä. □

3. Smithin normaalimuodon laskeminen

Jokaiselle polynomimatriisille $C \in \mathrm{Mat}_n(\mathbb{K}[x])$ löytyy siis yksikäsitteinen Smithin normaalimuodossa oleva polynomimatriisi, joka on ekvivalentti tämän matriisin kanssa. Edellä esitetystä Smithin normaalimuodon olemassaolotodistuksesta näkyy, miten kyseinen muoto voidaan laskea. Todistus on kuitenkin tehty monessa osassa, joten esitetään vielä selvyyden vuoksi laskemiseen liittyvät vaiheet tiivistetysti algoritmina:

(1) Tehdään tarvittavat rivin- ja sarakkeenvaihdot, jotta matriisin asteeltaan pienin nollasta eroava kerroinpolynomi saadaan paikalle $(1, 1)$.

(2) Kirjoitetaan ensimmäisen sarakkeen polynomit jakoyhtälön avulla muodossa $p_{i1} = r_i p_{11} + s_i$, kun $i > 1$, ja sovelletaan rivioperaatioita $A_{1i}(-r_i)$.
(3) Kirjoitetaan ensimmäisen rivin polynomit jakoyhtälön avulla muodossa $p_{1j} = r_j p_{11} + s_j$, kun $j > 1$, ja sovelletaan sarakeoperaatioita $A_{j1}(-r_j)$.
(4) Jos tässä vaiheessa jollakin nollasta eroavalla kerroinpolynomilla on aidosti pienempi aste kuin kerroinpolynomilla p_{11}, palataan takaisin vaiheeseen 1. Muutoin siirrytään vaiheeseen 5.
(5) Tässä vaiheessa polynomi p_{11} on ainoa ensimäisen rivin ja sarakkeen nollasta eroava polynomi ja se on lisäksi nollasta eroavista polynomeista asteeltaan pienin. Jos p_{11} jakaa kaikki kerroinpolynomit, voidaan siirtyä vaiheeseen 6. Muutoin olkoon p_{ij} kerroinpolynomi, joka ei ole jaollinen polynomilla p_{11}. Suoritetaan sarakeoperaatio $A_{1j}(1)$ ja palataan vaiheeseen 2.
(6) Matriisin pitäisi tässä vaiheessa olla muotoa

$$\begin{bmatrix} p & 0 \\ 0 & D \end{bmatrix},$$

missä polynomi p jakaa kaikki matriisin D kerroinpolynomit. Jos D ei ole vielä haluttua muotoa, siirrytään takaisin vaiheeseen 1, ja jatketaan rivi- ja sarakeoperaatioiden soveltamista matriisiin D.

Esimerkki 3.0.42. Lasketaan polynomimatriisin

$$\begin{bmatrix} 1 & 2x & 2x^2+2x \\ 1 & 6x & 6x^2+6x \\ 1 & 3 & x \end{bmatrix} \in \mathrm{Mat}_3(\mathbb{R}[x])$$

Smithin normaalimuoto. Nuolen yläpuolella olevat merkinnät viittaavat rivioperaatioihin ja alapuolella olevat sarakeoperaatioihin.

$$\begin{bmatrix} 1 & 2x & 2x^2+2x \\ 1 & 6x & 6x^2+6x \\ 1 & 3 & x \end{bmatrix} \underset{A_{21}(-2x),\, A_{31}(-2x^2-2x)}{\overset{A_{12}(-1),\, A_{13}(-1)}{\longrightarrow}}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4x & 4x^2+4x \\ 0 & -2x+3 & -2x^2-x \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4x & 4x(x+1) \\ 0 & -\frac{1}{2}(4x)+3 & -2x^2+-x \end{bmatrix} \underset{A_{32}(-x-1)}{\overset{A_{23}(\frac{1}{2})}{\longrightarrow}}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4x & 0 \\ 0 & 3 & -2x^2-x+2x(x+1)+3(-x-1) \end{bmatrix} \overset{P_{23}}{\rightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -2x-3 \\ 0 & 4x & 0 \end{bmatrix}$$

$$\underset{A_{32}(\frac{1}{3}(2x+3))}{\overset{A_{23}(-\frac{4}{3}x)}{\longrightarrow}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & (-\frac{4}{3}x)(-2x-3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \frac{8}{3}x^2+4x \end{bmatrix} \overset{M_2(\frac{1}{3}),\, M_3(\frac{3}{8})}{\longrightarrow}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & x^2+\frac{3}{2}x \end{bmatrix}$$

Muunnosmatriisit ovat tällöin

$$
\begin{aligned}
R &= M_3\left(\frac{3}{8}\right) M_2\left(\frac{1}{3}\right) A_{23}\left(-\frac{4}{3}x\right) P_{23}A_{23}\left(\frac{1}{2}\right) A_{13}\left(-1\right) A_{12}\left(-1\right) \\
&= \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ \frac{3}{4}x - \frac{3}{8} & -\frac{1}{4}x + \frac{3}{8} & -\frac{1}{2}x \end{bmatrix}
\end{aligned}
$$

ja

$$
\begin{aligned}
Q &= A_{21}(-2x)A_{31}(-2x^2 - 2x)A_{32}(-x-1)A_{32}\left(\frac{1}{3}(2x+3)\right) \\
&= \begin{bmatrix} 1 & -2x & -\frac{4}{3}x^2 - 2x \\ 0 & 1 & -\frac{1}{3}x \\ 0 & 0 & 1 \end{bmatrix},
\end{aligned}
$$

joten

$$
\begin{aligned}
&\begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ \frac{3}{4}x - \frac{3}{8} & -\frac{1}{4}x + \frac{3}{8} & -\frac{1}{2}x \end{bmatrix}
\begin{bmatrix} 1 & 2x & 2x^2 + 2x \\ 1 & 6x & 6x^2 + 6x \\ 1 & 3 & x \end{bmatrix}
\begin{bmatrix} 1 & -2x & -\frac{4}{3}x^2 - 2x \\ 0 & 1 & -\frac{1}{3}x \\ 0 & 0 & 1 \end{bmatrix}. \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & x^2 + \frac{3}{2}x \end{bmatrix}
\end{aligned}
$$

Lause 3.0.43. *Olkoot $\lambda \in \mathbb{K}$ ja $n \geq 2$. Tällöin polynomimatriisin*

$$
\begin{bmatrix} \lambda - x & 1 & & 0 \\ & \lambda - x & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda - x \end{bmatrix} \in Mat_n(\mathbb{K}[x]).
$$

Smithin normaalimuoto on $diag(1, \ldots, 1, (-1)^n(\lambda - x)^n)$.

Todistus. Väite voidaan todistaa induktiolla hieman yleisemmässä muodossa. Olkoon $k \in \mathbb{N}$. Osoitetaan, että polynomimatriisin

$$
A := \begin{bmatrix} \pm(\lambda - x)^k & 1 & & 0 \\ & \lambda - x & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda - x \end{bmatrix} \in \mathrm{Mat}_n(\mathbb{K}[x])
$$

smithin normaalimuoto on $\operatorname{diag}(1,\dots,1,(-1)^{n+k-1}(\lambda-x)^{n+k-1})$. Kun $n=2$,

$$\begin{bmatrix}\pm(\lambda-x)^k & 1\\ 0 & \lambda-x\end{bmatrix} \underset{P_{12}}{\rightarrow} \begin{bmatrix}1 & \pm(\lambda-x)^k\\ \lambda-x & 0\end{bmatrix}$$
$$\xrightarrow[A_{21}(\mp(\lambda-x)^k)]{A_{12}(-(\lambda-x))} \begin{bmatrix}1 & 0\\ 0 & \mp(\lambda-x)^{2+k-1}\end{bmatrix},$$

joten väite pätee. Oletetaan, että jollakin $n\geq 3$ väite pätee kaikille enintään kokoa $n-1$ oleville polynomimatriiseille. Tällöin vastaavalle kokoa n olevalle polynomimatriisille pätee

$$\begin{bmatrix}\pm(\lambda-x)^k & 1 & & 0\\ & \lambda-x & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda-x\end{bmatrix} \underset{P_{12}}{\rightarrow} \begin{bmatrix}1 & \pm(\lambda-x)^k & & 0\\ \lambda-x & 0 & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda-x\end{bmatrix}$$
$$\xrightarrow[A_{21}(\mp(\lambda-x)^k)]{A_{12}(-(\lambda-x))} \begin{bmatrix}1 & 0\\ 0 & D_0\end{bmatrix},$$

missä

$$D_0 = \begin{bmatrix}\mp(\lambda-x)^{k+1} & 1 & & 0\\ & \lambda-x & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda-x\end{bmatrix}.$$

Induktio oletuksen nojalla matriisin D_0 Smithin normaalimuoto on

$$R_0D_0Q_0 = \operatorname{diag}(1,\dots,1,(-1)^{n-1+k}(\lambda-x)^{n-1+k}),$$

missä $R_0,Q_0\in \mathrm{Gl}_{n-1}(\mathbb{K}[x])$. Asetetaan $R:=\operatorname{diag}(1,R_0)$ ja $Q=\operatorname{diag}(1,Q_0)$, jolloin

$$R\begin{bmatrix}1 & 0\\ 0 & D_0\end{bmatrix}Q = \begin{bmatrix}I_{n-1} & 0\\ 0 & (-1)^{n+k-1}(\lambda-x)^{n+k-1})\end{bmatrix},$$

mistä induktioväite seuraa. □

LUKU 6

Invariantit tekijät ja similaarisuusinvariantit

1. Similaarisuus ja Smithin normaalimuoto

Smithin normaalimuoto liittyy polynomimatriiseihin, mutta sitä voidaan hyödyntää myös tarkasteltaessa $\mathbb{K}$-kertoimisia matriiseja. Se antaa esimerkiksi vastauksen kysymykseen, milloin kaksi eri $\mathbb{K}$-kertoimista matriisia ovat similaariset. Similaarisilla matriiseilla on paljon samoja ominaisuuksia, kuten esimerkiksi sama karakteristinen polynomi ja determinantti. Toisaalta tällaisia yhteisiä ominaisuuksia voi olla myös ei-similaarisilla matriiseilla, joten niistä ei voida päätellä similaarisuutta. Ilman oikeita apuneuvoja kahden annetun matriisin similaarisuuden päätteleminen voi siis olla hyvinkin hankala pulma. Smithin normaalimuotoa voidaan lisäksi soveltaa moniin muihinkin $\mathbb{K}$-kertoimisten matriisien teoriaan liittyviin aiheisiin. Seuraavaksi onkin tarkoitus selvittää paitsi Smithin normaalimuodon yhteys similaarisuuteen myös, mitä sen avulla voidaan sanoa esimerkiksi $\mathbb{K}$-kertoimisen matriisin karakterisisesta polynomista ja minimipolynomista sekä kanonisista muodoista.

Olkoon $A \in \mathrm{Mat}_n(\mathbb{K})$ ja $S = \mathrm{diag}(p_1, \ldots, p_n)$ sen Smithin normaalimuoto. Tällöin polynomeja $p_1, \ldots, p_n$ kutsutaan polynomimatriisin A *invarianteiksi tekijöiksi*. Smithin normaalimuodon yksikäsitteisyyden nojalla invariantit tekijät ovat hyvin määriteltyjä.

Lause 1.0.44. *Olkoon $A \in Mat_n(\mathbb{K})$. Tällöin $\chi(A)(x) = p_1(x) \cdots p_n(x)$, missä polynomit $p_1, \ldots, p_n \in \mathbb{K}[x]$ ovat polynomimatriisin $A - xI$ invariantit tekijät.*

Todistus. Olkoon $S = R(a - xI)Q$, missä $R, Q \in \mathrm{Gl}_n(\mathbb{K}[x])$, polynomimatriisin $A - xI$ Smithin normaalimuoto. Tällöin

$$p_1(x) \cdots p_n(x) = \det(S) = \det(R(A - xI)Q) = \det(R)\det(A - xI)\det(Q).$$

Lauseen 1.0.9 nojalla $\det(R)$, $\det(Q) \in \mathbb{K}\backslash\{0\}$, ja lauseen 2.0.28 nojalla $\det(A-xI) = \pm\chi_A(x)$. Siten

$$p_1(x) \cdots p_n(x) = a\chi_A(x)$$

jollekin $a \in \mathbb{K}\backslash\{0\}$. Toisaalta polynomit $p_1(x) \cdots p_n(x)$ ja $\chi_A(x)$ ovat molemmat perusmuotoisia, joten $a = 1$. □

Olkoon $A \in \mathrm{Mat}_n(\mathbb{K})$. Lauseen 1.0.44 nojalla polynomimatriisin $A-xI$ invariantit tekijät $p_1, \ldots, p_n$ ovat nollasta eroavia polynomeja, sillä karakteristinen polynomi ei voi koskaan olla nollapolynomi. Lisäksi karakteristisen polynomin aste on aina n, joten $p_j \neq 1$ jollakin $j \in \{0, \ldots, n\}$. Jos $p_1 \neq 1$, kaikki invariantit tekijät p_j ovat vähintään astetta 1. Asetetaan tällöin $k := n$. Oletetaan, että $p_1 = 1$. Silloin on olemassa sellainen $k \in \{1, \ldots, n-1\}$, jolle polynomit $p_n, \ldots, p_{n-k+1}$ ovat vähintään astetta 1 ja $p_1 = \cdots = p_{n-k} = 1$. Asetetaan kaikille $j = 1, \ldots, k$, $q_j := p_{n-j+1}$. Polynomeja $q_1, \ldots, q_k$ kutsutaan matriisin A *similaarisuusinvarianteiksi.* Ne ovat siis perusmuotoisia ja ei-vakioita polynomeja, joille pätee $q_j | q_{j-1}$ kaikilla $j = 2, \ldots, k$.

SEURAUS 1.0.45. *Olkoon $A \in Mat_n(\mathbb{K})$. Tällöin $\chi(A)(x) = q_1(x) \cdots q_k(x)$, missä polynomit $q_1, \ldots, q_k \in \mathbb{K}[x]$ ovat matriisin A similaarisuusinvariantit.*

HUOMAUTUS 1.0.46. Matriisin $A \in \mathrm{Mat}_n(\mathbb{K})$ karakteristinen polynomi voitaisiin määritellä myös sen similaarisuusinvaranttien tulona tai polynomimatriisin $A - xI$ invarianttien tekijöiden tulona.

Seuraava lause on erittäin tärkeä jatkon kannalta, sillä se yhdistää $\mathbb{K}$-kertoimisten matriisien similaarisuuden polynomimatriisien ekvivalenttiuteen eli olennaisesti Smithin normaalimuotoon. Lause on erikoistapaus D. Serren kirjassa Matrices: Theory and Applications todistetusta yleisemmästä versiosta [Theorem 6.3.2, s.104], ja sen todistus noudatteleekin pitkälti tätä Serren kirjassaan esittämää todistusta.

LAUSE 1.0.47. *Olkoot matriisit $A, B \in Mat_n(\mathbb{K})$. Tällöin A ja B ovat similaariset jos ja vain jos polynomimatriisit $A - xI$ ja $B - xI$ ovat ekvivalentit.*

TODISTUS. Jos matriisit A ja B ovat similaariset, on olemassa $R \in \mathrm{Gl}_n(\mathbb{K})$, jolle $A = RBR^{-1}$. Tällöin pätee myös $R \in \mathrm{Gl}_n(\mathbb{K}[x])$ ja $A - xI = RBR^{-1} - xRR^{-1} = R(B - xI)R^{-1}$, joten väitteen tämä suunta on selvä.

Oletetaan siis, että polynomimatriisit $A - xI$ ja $B - xI$ ovat ekvivalentit. Tällöin on olemassa sellaiset $R, Q \in \mathrm{Gl}_n(\mathbb{K}[x])$, joille

$$R(-xI + A) = (-xI + B)Q. \tag{6.1}$$

Polynomimatriisien jakoyhtälön (lause 1.0.10) nojalla

$$\begin{cases} R = (-xI + B)R_1 + G \\ Q = Q_1(-xI + A) + H, \end{cases}$$

missä R_1, $Q_1 \in \mathrm{Mat}_n(\mathbb{K}[x])$ ja $G, H \in \mathrm{Mat}_n(\mathbb{K})$. Sijoittamalla nämä yhtälöön (6.1) saadaan yhtälö

$$((-xI + B)R_1 + G)(-xI + A) = (-xI + B)(Q_1(-xI + A) + H), \tag{6.2}$$

joka voidaan kirjoittaa muodossa

$$(-xI + B)(R_1 - Q_1)(-xI + A) = (-xI + B)H - G(-xI + A). \tag{6.3}$$

Jos $(R_1 - Q_1) \neq 0$, on polynomimatriisi $(-xI + B)(R_1 - Q_1)(-xI + A)$ vähintään astetta 2. Toisaalta polynomimatriisi $(-xI + B)H - G(-xI + A)$ on enintään astetta 1, joten täytyy olla $R_1 = Q_1$. Silloin yhtälö (6.3) saadaan muotoon

$$-Gx + GA = -Hx + BH. \tag{6.4}$$

Koska $A, B, G, H \in \mathrm{Mat}_n(\mathbb{K})$, saadaan kerroinmatriisien yksikäsitteisyyden nojalla yhtälöt

$$\begin{cases} G = H \\ GA = BH. \end{cases}$$

Väitteen todistamiseksi on enää osoitettava, että matriisi G on kääntyvä.

Polynomimatriisien jakoyhtälön nojalla voidaa kirjoittaa

$$R^{-1} = (-xI + A)R_2 + K,$$

missä $R_2 \in \mathrm{Mat}_n(\mathbb{K}[x])$ ja $k \in \mathrm{Mat}_n(\mathbb{K})$. Tällöin

$$\begin{aligned}
I - GK &= RR^{-1} - GK \\
&= ((-xI + B)R_1 + G)((-xI + A)R_2 + K) - GK \\
&= (-xI + B)R_1(-xI + A)R_2 + (-xI + B)R_1K + G(-xI + A)R_2 \\
&= ((-xI + B)R_1(-xI + A) + G(-xI + A))R_2 + (-xI + B)R_1K \\
&= ((-xI + B)R_1(-xI + A) + (-xI + B)H)R_2 + (-xI + B)R_1K \\
&= (-xI + B)(Q_1(-xI + A) + H)R_2 + (-xI + B)R_1K \\
&= (-xI + B)(QR_2 + R_1K).
\end{aligned}$$

Tämän yhtälön vasen puoli on vakipolynomimatriisi eli sen aste on enintään 0. Jos $QR_2 + R_1K \neq 0$, oikean puolen aste on vähintään 1, mikä on siis mahdotonta. Näin ollen on oltava $QR_2 + R_1K = 0$, jolloin $GK = I$. □

SEURAUS 1.0.48. *Olkoot matriisit $A, B \in Mat_n(\mathbb{K})$. Tällöin A ja B ovat similaariset jos ja vain jos polynomimatriisien $A - xI$ ja $B - xI$ Smithin normaalimuodot ovat samat.*

SEURAUS 1.0.49. *Olkoot matriisit $A, B \in Mat_n(\mathbb{K})$. Tällöin A ja B ovat similaariset jos ja vain jos niillä on samat similaarisuusinvariantit.*

LAUSE 1.0.50. *Olkoon C_p polynomin $p = x^n + \alpha_{n-1}x^{k-1} + \cdots + \alpha_1 x + \alpha_0 \in \mathbb{K}[x]$ kumppanimatriisi. Tällöin polynomimatriisin $C_p - xI$ invariantit tekijät ovat $p_1 = \cdots = p_{n-1} = 1$ ja $p_n = p$.*

TODISTUS. Polynomimatriisi $C_p - xI$ on muotoa

$$C_p - xI = \begin{bmatrix} -x & & \cdots & 0 & -\alpha_0 \\ 1 & -x & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{k-1} - x \end{bmatrix}.$$

Lauseen 3.0.32 todistuksen nojalla on olemassa $R \in \mathrm{Gl}_n(\mathbb{K}[x])$, jolle

$$R(C_p - xI) = \begin{bmatrix} 1 & -x & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & 1 & -x - \alpha_{n-1} \\ 0 & 0 & \cdots & 0 & p(x) \end{bmatrix}.$$

Asetetaan $Q_j = A_{j+1\,j}(x)$ ja $Q_{\alpha\,j} = A_{nj}(\alpha_j)$ kaikille $j = 1, \ldots, n-1$. Tällöin

$$R(C_p - xI)(Q_1 Q_{\alpha\,1}) \cdots (Q_{n-1} Q_{\alpha\,n-1}) = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & p(x) \end{bmatrix}.$$

Yksikäsitteisyyden nojalla tämä on polynomimatriisin $C_p - xI$ Smithin normaalimuoto, mistä väite seuraa. □

SEURAUS 1.0.51. *Perusmuotoisen astetta $n \geq 1$ olevan polynomin $p \in \mathbb{K}[x]\backslash\{0\}$ kumppanimatriisin C_p ainoa similaarisuusinvariantti on p.*

2. Frobeniuksen muoto

Jokainen matriisi $A \in \mathrm{Mat}_n(\mathbb{K})$ on similaarinen sellaisen lohkodiagonaalimatriisin C kanssa, jonka diagonaalilohkot ovat matriisin A similaarisuusinvarianttien kumppanimatriisit. Tällaista matriisia C kutsutaan matriisin A *Frobeniuksen muodoksi* tai *rationaaliseksi kanoniseksi muodoksi.*

LAUSE 2.0.52 (Frobeniuksen muoto). *Olkoon $A \in Mat_n(\mathbb{K})$. Tällöin on olemassa kääntyvä matriisi $R \in Mat_n(\mathbb{K})$, jolle*

$$C := RAR^{-1} = \begin{bmatrix} C_{q_1} & 0 & \cdots & 0 \\ 0 & C_{q_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_{q_k} \end{bmatrix},$$

missä polynomit $q_1, \ldots, q_k$ ovat matriisin A similaarisuusinvariantit ja matriisit $C_{q_1}, \ldots, C_{q_k}$ vastaavat kumppanimatriisit.

TODISTUS. Lauseen 1.0.47 nojalla riittää osoittaa, että polynomimatriisit $A - xI$ ja $C - xI$ ovat ekvivalentit, kun $C = \mathrm{diag}(C_{q_1}, \ldots, C_{q_k})$. Olkoot polynomit $q_1, \ldots, q_k$ matriisin A similaarisuusinvariantit, ja S polynomimatriisin $A-xI$ Smithin normaalimuoto. Tällöin $S = \mathrm{diag}(1, \ldots, 1, q_k, \ldots, q_1)$. Polynomimatriisi $A-xI$ on ekvivalentti Smithin normaalimuotonsa S kanssa, joten riittää osoittaa, että polynomimatriisit S ja $C - xI$ ovat ekvivalentit.

Seurauksen 1.0.45 nojalla $\chi_A = q_1 \cdots q_k$, jolloin erityisesti

$$\deg(q_1) + \cdots + \deg(q_k) = n.$$

Merkitään $D_{q_j} := \mathrm{diag}(1, \ldots, 1, q_j) \in \mathrm{Mat}_{\deg(q_j)}(\mathbb{K}[x])$ kaikille $j = 1, \ldots, k$. Alkio 1 esiintyy $\deg(q_j)-1$ kertaa jokaisessa matriisissa D_{q_j}. Toisaalta 1 esiintyy matriisissa S $n-k = \deg(q_1)-1+\cdots+\deg(q_k)-1$ kertaa eli siis yhtä monta kertaa kuin matriiseissa D_{q_j} yhteensä, kun $j = 1, \ldots, k$. Tällöin diagonaalimatriisin S diagonaalialkiot voidaan järjestää rivin- ja sarakkeenvaihto -operaatioiden avulla uudelleen niin, että

$$R_0 S Q_0 = \begin{bmatrix} D_{q_1} & & 0 \\ & \ddots & \\ 0 & & D_{q_k} \end{bmatrix},$$

missä matriisit $R_0, Q_0 \in Gl_n(\mathbb{K}[x])$ vastaavat rivin- ja sarakkeenvaihto -operaatioita.

Lauseen 1.0.50 nojalla jokaiselle $j = 1, \ldots, k$ on olemassa kääntyvät polynomimatriisit $\tilde{R}_j$, $\tilde{Q}_j \in \mathrm{Gl}_{\deg(q_j)}(\mathbb{K}[x])$, joille

$$\tilde{R}_j D_{q_j} \tilde{Q}_j = C_{q_j} - xI_{\deg(q_j)}.$$

Asetetaan $R_1 := \mathrm{diag}(\tilde{R}_1, \dots, \tilde{R}_k)$ ja $Q_1 := \mathrm{diag}(\tilde{Q}_1, \dots, \tilde{Q}_k)$, jolloin

$$\begin{aligned}(R_1R_0)S(Q_0Q_1) &= \begin{bmatrix} C_{q_1} - xI_{\deg(q_1)} & & 0 \\ & \ddots & \\ 0 & & C_{q_k} - xI_{deg(q_k)} \end{bmatrix} \\ &= \begin{bmatrix} C_{q_1} & & 0 \\ & \ddots & \\ 0 & & C_{q_k} \end{bmatrix} - xI_n.\end{aligned}$$

□

Olkoon $A \in \mathrm{Mat}_n(\mathbb{K})$ ja C sen Frobeniuksen muoto. Lause 2.0.52 takaa siis, että jollekin kääntyvälle $H \in \mathrm{Mat}_n(\mathbb{K})$ on $C = HAH^{-1}$. Jossain tapauksissa voi olla tarpeellista tietää muunnosmatriisi H, ja se voidaankin selvittää edellisten lauseiden todistuksien avulla.

Olkoon S polynomimatriisin $A - xI$ Smithin normaalimuoto, jolloin $S = R_S(a - xI)Q_S$ joillekin R_S, $Q_S \in \mathrm{Gl}_n(\mathbb{K}[x])$. Polynomimatriisit R_S ja Q_S voidaan laskea Smithin normaalimuodon laskemiseen käytettyjen rivi- ja sarakeoperaatioiden perusteella. Käytetään lauseen 2.0.52 merkintöjä, ja asetetaan $R := R_1R_0R_S$. Polynomiatriisi R_0 voidaan laskea Smithin normaalimuodon S diagonaalialkioiden uudelleenjärjestämiseen käytettävien rivioperaatioiden perusteella. Polynomimatriisi R_1 on lohkodiagonaalimuotoinen, ja sen diagonaalilohkot $\tilde{R}_j$ voidaan selvittää polynomimatriisien $C_{q_j} - xI$ Smithin normaalimuotojen laskemiseen käytettyjen rivioperaatioiden perusteella.

Lauseen 1.0.47 merkinnöin polynomimatriisien jakoyhtälöstä saadaan

$$R = (B - xI)R_1 + G,$$

missä $R_1 \in \mathrm{Mat}_n(\mathbb{K}[x])$ ja $G \in \mathrm{Mat}_n(\mathbb{K})$. Lisäksi tällöin lauseen 1.0.47 todistuksen mukaan G on kääntyvä ja $C = GAG^{-1}$.

Olkoon $A \in \mathrm{Mat}_n(\mathbb{K})$ ja polynomit $q_1, \dots, q_k$ sen similaarisuusinvariantit. Ensimmäisellä similaarisuusinvariantilla q_1 on oma erityisroolinsa. Sitä kutsutaan matriisin A *minimipolynomiksi*, ja merkitään m_A. Se on asteeltaan pienin perusmuotoinen matriisin A nollaava polynomi, kuten seuraava lause osoittaa.

LAUSE 2.0.53. *Olkoon $A \in Mat_n(\mathbb{K})$ ja m_A sen minimipolynomi. Tällöin $m_A(A) = 0$ ja $deg(A) \leq deg(p)$ kaikille sellaisille polynomeille $p \in \mathbb{K}[x]$, joille $p(A) = 0$.*

TODISTUS. Osoitetaan ensin, että väite pätee kumppanimatriiseille. Olkoon $C_q \in \mathrm{Mat}_n(\mathbb{K})$, jonkin perusmuotoisen astetta $n \geq 1$ olevan polynomin q kumppanimatriisi. Tällöin seurauksen 1.0.51 nojalla $m_{C_q} = q$. Toisaalta lauseen 3.0.33 todistuksen nojalla $q(C_q) = 0$. Olkoon $p \in \mathbb{K}[x]$, sellainen polynomi, jolle $p(A) = 0$. Riittää osoittaa, että $n = \deg(q) \leq \deg(p)$. Tehdään antiteesi, että olisikin $n = \deg(q) > \deg(p) =: k$. Tällöin

$$\begin{aligned}0 = p(C_q)e_1 &= \gamma_0 e_1 + \gamma_1 C_q e_1 + \dots + \gamma_k C_q^k e_1 \\ &= \gamma_0 e_1 + \gamma_1 e_2 + \dots + \gamma_k e_{k+1}\end{aligned},$$

missä $p(x) = \gamma_k x^k + \dots + \gamma_1 x + \gamma_0$ ja $\gamma_k \neq 0$. Tästä seuraa, että vektorijoukko $\{e_1, \dots, e_{k+1}\}$ on lineaarisesti riippuva, mikä on ristiriita.

Osoitetaan seuraavaksi yleinen tapaus. Olkoon $A \in \mathrm{Mat}_n(\mathbb{K})$, polynomit $q_1, \ldots, q_k$ sen similaarisuusinvariantit ja $C = \mathrm{diag}(C_{q_1}, \ldots, C_{q_k})$ sen Frobeniuksen muoto. Tällöin $A = R^{-1}CR$ jollekin $R \in \mathrm{Gl}_n(\mathbb{K})$, joten lauseen 1.0.11 nojalla

$$m_A(A) = q_1(A) = q_1(R^{-1}CR) = R^{-1}q_1(C)R.$$

Edelleen lauseen 1.0.11 mukaan $q_1(C) = \mathrm{diag}(q_1(C_{q_1}), \ldots, q_1(C_{q_k}))$. Alussa todistetun tuloksen nojalla $q_j(C_{q_j}) = 0$ kaikille $j = 1, \ldots, k$. Toisaalta $q_j | q_1$ jokaisella $j = 1, \ldots, k$, joten myös $q_1(C_{q_j}) = 0$ kaikille $j = 1, \ldots, k$, eli $q_1(C) = 0$. □

3. Similaarisuusinvarianttien yhteys Jordanin muotoon

Seuraavaksi tarkastellaan ainoastaan kompleksikertoimisia matriiseja. Matriisia

$$J_n(\lambda) := \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \in \mathrm{Mat}_n(\mathbb{C})$$

kutsutaan kokoa n olevaksi *Jordanin soluksi.* Jokainen kompleksikertoiminen matriisi $A \in \mathrm{Mat}_n(\mathbb{C})$ on similaarinen sellaisen lohkodiagonaalimatriisin kanssa, jonka diagonaalilohkot ovat Jordanin soluja. Tällaista matriisia kutsutaan matriisin *A Jordanin muodoksi.* Tässä tutkielmassa ei ole tarkoitus syventyä itse Jordanin muodon teoriaan tai sovelluksiin. Tarkempaa tietoa Jordanin muodosta löytyy esimerkiksi teoksesta [Horn & Johnson, luku 3]. Siinä Jordanin muotoa käsitellään ilman Smithin normaalimuotoa. Sen sijaan esimerkiksi teoksessa [Broida & Williamson, luku 8.6] Jordanin muotoa tarkastellaan Smithin normaalimuotoa hyödyntäen, kuten tässäkin tutkielmassa tehdään.

Matriisin Jordanin muoto liittyy läheisesti sen similaarisuusinvariantteihin. Itse asiassa, jos matriisin similaarisuusinvariantit tiedetään, tiedetään myös sen Jordanin muoto. Seuraavaksi on tarkoitus todistaa Jordanin muodon olemassaolo hyödyntäen edellä käsiteltyjä Smithin normaalimuotoa ja Frobeniuksen muotoa sekä niihin liittyvää teoriaa. Tästä olemassaolotodistuksesta selviää myös, miten Jordanin muoto voidaan päätellä similaarisuusinvarianteista.

Olkoot $A \in \mathrm{Mat}_n(\mathbb{C})$ ja polynomit $q_1(x), \ldots, q_k(x)$ sen similaarisuusinvariantit. Renkaassa $\mathbb{C}[x]$ jokainen polynomi hajoaa ensimmäisen asteen polynomien tuloksi. Tällöin on olemassa kuvaukset $s : \{1, \ldots, k\} \to \mathbb{N}$ ja $r : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, joille

$$q_l(x) = (x - \lambda_{l1})^{r(l,1)} \cdots (x - \lambda_{l\,s(l)})^{r(l,s(l))},$$

kaikille $l = 1, \ldots, k$. Lisäksi jokaiselle $l \in \{1, \ldots, k\}$ pätee $\lambda_{li} \neq \lambda_{lj}$, kun $i \neq j$ ja $i, j \in \{1, \ldots, s(l)\}$. Polynomeja $(x - \lambda_{lj})^{r(l,j)}$, missä $l = 1, \ldots, k$ ja $j = 1, \ldots, s(l)$, kutsutaan matriisin A *alkeistekijöiksi.*

Lause 3.0.54. *Olkoot $n \geq 1$ ja $C_p \in Mat_n(\mathbb{C})$ perusmuotoisen astetta n olevan polynomin $p \in \mathbb{C}[x]$ kumppanimatriisi. Olkoon $p(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_l)^{k_l}$, missä $\lambda_i \neq \lambda_j$, kun $i \neq j$. Tällöin C_p on similaarinen matriisin*

$$B := \begin{bmatrix} C_{(x-\lambda_1)^{k_1}} & & 0 \\ & \ddots & \\ 0 & & C_{(x-\lambda_l)^{k_l}} \end{bmatrix}$$

kanssa.

TODISTUS. Lauseen 1.0.47 nojalla riittää osoittaa, että polynomimatriisit $C_p - xI$ ja $B - xI$ ovat ekvivalentit. Lauseen 1.0.50 nojalla polynomimatriisin $C_p - xI$ Smithin normaalimuoto on $D_p = \operatorname{diag}(1, \dots, 1, p) \in \operatorname{Mat}_n(\mathbb{C})$. Saman lauseen mukaan jokaisella $j = 1, \dots, l$ polynomimatriisin $C_{(x-\lambda_j)^{k_j}}$ Smithin normaalimuoto on

$$\tilde{R}_j C_{(x-\lambda_j)^{k_j}} \tilde{Q}_j = D_{(x-\lambda_j)^{k_j}} := \operatorname{diag}(1, \dots, 1, (x - \lambda_j)^{k_j}) \in \operatorname{Mat}_{k_j}(\mathbb{C})$$

Asetetaan $R_0 := \operatorname{diag}(\tilde{R}_1, \dots, \tilde{R}_l)$ ja $Q_0 := \operatorname{diag}(\tilde{Q}_1, \dots, \tilde{Q}_l)$, jolloin

$$R_0(B - xI)Q_0 = \operatorname{diag}(D_{(x-\lambda_1)^{k_1}}, \dots, D_{(x-\lambda_l)^{k_l}}) =: D.$$

Polynomimatriisit $B - xI$ ja D ovat siis ekvivalentit, jolloin väitteen todistamiseksi riittää osoittaa, että polynomimatriisit D_p ja D ovat ekvivalentit. Tämä seuraa kuitenkin lauseesta 2.0.41, jonka mukaan D_p on matriisin D Smithin normaalimuoto. □

LAUSE 3.0.55. *Olkoot $n \geq 1$ ja $C_p \in Mat_n(\mathbb{C})$ perusmuotoisen polynomin $(x - \lambda)^n \in \mathbb{C}[x]$ kumppanimatriisi. Tällöin C_p on similaarinen Jordanin solun $J_n(\lambda)$ kanssa.*

TODISTUS. Lauseen 1.0.47 nojalla riittää osoittaa, että polynomimatriisit $J_n(\lambda) - xI$ ja $C_p - xI$ ovat ekvivalentit. Tämä seuraa kuitenkin suoraan lauseista 3.0.43 ja 1.0.50, joiden mukaan niillä on sama Smithin normaalimuoto. □

LAUSE 3.0.56 (Jordanin muoto). *Olkoon $A \in Mat_n(\mathbb{C})$. Olkoot Polynomit $(x - \lambda_{lj})^{r(l,j)}$, missä $l = 1, \dots, k$ ja $j = 1, \dots, s(l)$, matriisin A alkeistekijät. Tällöin A on similaarinen lohkodiagonaalimatriisin*

$$J := \begin{bmatrix} J_{r(1,1)}(\lambda_{11}) & & & & & & \\ & \ddots & & & & 0 & \\ & & J_{r(1,s(1))}(\lambda_{1\,s(1)}) & & & & \\ & & & \ddots & & & \\ & & & & J_{r(k,1)}(\lambda_{k1}) & & \\ & 0 & & & & \ddots & \\ & & & & & & J_{r(k,s(k))}(\lambda_{k\,s(k)}) \end{bmatrix}$$

kanssa.

TODISTUS. Olkoon $F = \operatorname{diag}(C_{q_1}, \dots, C_{q_k})$ matriisin A Frobeniuksen muoto, missä polynomit $q_1, \dots, q_k$ ovat matriisin A similaarisuusinvariantit. Riittää osoittaa, että matriisit F ja J ovat similaariset. Lauseen 3.0.54 nojalla jokaisella $l = 1, \dots, k$ matriisi C_{q_l} on similaarinen lohkodiagonaalimatriisin

$$\operatorname{diag}(C_{(x-\lambda_{l1})^{r(l,1)}}, \dots, C_{(x-\lambda_{l\,s(l)})^{r(l,s(l))}})$$

kanssa. Olkoon $l \in \{1, \dots, k\}$ kiinnitetty. Silloin lauseen 3.0.55 nojalla jokaisella $j = 1, \dots, s(l)$ matriisi $C_{(x-\lambda_{lj})^{r(l,j)}}$ on similaarinen Jordanin solun $J_{r(l,j)}(\lambda_{lj})$ kanssa. □

Kirjallisuutta

[1] Sheldon Axler: *Linear Algebra Done Right.* Springer-Verlag, New York, 1997.

[2] Frank Ayres: *Theory and Problems of Matrices.* Schaum Publishing Co, New York, 1962.

[3] Joel G. Broida & S. Gill Williamson: *Comprehensive Introduction to Linear Algebra.* Addison-Wesley, 1986.

[4] Joel N. Franklin: *Matrix Theory.* Prentice-Hall, 1968.

[5] Jonathan S. Golan: *The Linear Algebra a Beginning Graduate Student Ought to Know.* Third Edition, Springer, 2010.

[6] Roger A. Horn & Charles R. Johnson: *Matrix Analysis* Cambridge university press, 1985.

[7] Serge Lang: *Algebra* Addison-Wesley, eight printing 1978.

[8] Tauno Metsänkylä & Marjatta Näätänen: *Algebra.* 1. painos, Limes ry, 2003. Verkossa http://solmu.math.helsinki.fi/2010/algebra.pdf (luettu 16.7.2013).

[9] Peter Petersen: *Linear Algebra.* Springer, New York 2012.

[10] Denis Serre: *Matrices: Theory and Applicatons.* Springer, Graduate Texts in Mathematics 216, 2002.

[11] Arne Storjohan: *Computation of Hermite and Smith Normal Forms of Matrices.* U. of W, master thesis, 1994. Verkossa
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.127.8231&rep=rep1&type=pdf (luettu 28.8.2013).