



LINEÁRIS ALGEBRA

Verzió információk

References	GitHub: May 16, 2022
Branch	GitHub: May 16, 2022
Dirty	
Hash	GitHub: May 16, 2022
Author Iso Date	GitHub: May 16, 2022
First Tag Describe	GitHub: May 16, 2022
Reln	GitHub: May 16, 2022
Roff	
Tags	
Describe	(None)

Előszó

A LEGFONTOSABB FORRÁS (Dancs és Puskás 2001).

...

Igyekszem strukturáltan írni. Kicsi, atomszerű építőkövek egymás utáni megértése visz az anyagban előre, ezek az egymástól feltűnő módon szeparált „állítások” és azok érvekkel való alátámasztása, amit „bizonyításnak” is szokás mondani. Az írásmód oka, hogy evvel is hangsúlyozzam, hogy az olvasónak igyekeznie kell strukturáltan gondolkodni. A hátulütője, hogy hibásan azt a helytelen képzetet keltheti, mintha az egyes állítások mintegy puzzle-ként állnának össze. Nem, nem erről van szó. A puzzle-ban minden elem egyenrangú, az egyik elem hiánya éppen annyira fájdalmas mint a másiké. Ez egyetlen matematikai diszciplína esetében sem igaz! Az olvasónak igyekeznie kell, hogy meglássa mi a legfontosabb gondolat a sok-sok állításnak, mint építménynek egy-egy „nyilvánvaló következményében”.

Hogy e kis lépések egymástól még határozottabban váljanak el azt az írás tipográfiája is erősíti azzal, az állítás-szerű környezeteket a „, és a bizonyítás környezetet a □ karakterekkel zárom le.

Stb.

Magyarkuti Gyula

Budapest, 2022. május 16-án

Tartalomjegyzék

I	Ősz	9
1	ELŐZMÉNYEK	11
1.1.	Algebrai struktúrák	11
1.2.	Polinomgyűrűk	13
1.3.	Polinomok oszthatósága és a maradékos osztás	15
1.4.	Az Euklideszi-algoritmus	18
1.5.	Polinom faktorizáció	20
1.6.	Mátrixok	21
1.7.	A komplex számok mint mátrixok	25
1.8.	A komplex számok abszolútértéke	27
1.9.	A komplex számok trigonometrikus alakja	29
1.10.	Polinom faktorizáció a komplex- és a valós számtest feletti	30
2	A VEKTORTÉR FOGALMA	33
2.1.	Vektortér altíerei	34
2.2.	Elimináció	36
2.3.	Lineárisan független rendszerek	45
3	A STEINITZ-LEMMA	51
3.1.	Rang-tétel	52
3.2.	Dimenzió	53
4	KOORDINÁTÁZÁS	57
4.1.	Lineáris operátor fogalma	57
5	ALTEREK MINKOWSKI-ÖSSZEGE ÉS DIREKT ÖSSZEGE	61
5.1.	Minkowski-összeg	61
5.2.	Direkt összeg	63
5.3.	Direkt kiegészítő	65
6	VEKTORTÉR FAKTORTERE	67
6.1.	Izomorfia tételek	69
II	TAVASZ	71
7	MÁTRIXOK ÉS LINEÁRIS OPERÁCIÓK	73
7.1.	Rang-defektus-tétel következménye	73
7.2.	Mátrixok tere mint koordináta-ter	74
7.3.	Lineáris operátorok szorzata	77
8	ÁLTALÁNOS BÁZISTRANSZFORMÁCIÓ	81
8.1.	Vektor koordinátái az új bázisban	81
8.2.	Lineáris operátorok mátrixa új bázispárban	82
8.3.	Lineáris transzformáció mátrixa az új bázisban	82

9 INVARIÁNS ALTEREK	83
9.1. Transzformációk sajátértéke	85
10 TRANSZFORMÁCIÓK POLINOMJAI	87
10.1. Kis minimálpolinom	88
10.2. Minimálpolinom	91
10.3. Sajátvektorok és diagonalizálhatóság	93
11 TRANSZFORMÁCIÓK REDUKÁLÁSA	97
11.1. Minimálpolinom és diagonalizálhatóság	98
11.2. Redukálás: az általános eset	99
12 REDUKÁLÁS IRREDUCIBILIS MINIMÁLPOLINOM ESETÉN	101
12.1. Irreducibilis polinommal képzett magtér redukálása	102
13 A MINIMÁLPOLINOM FOKSZÁMÁRÓL	105
14 NILPOTENS TRANSZFORMÁCIÓK	107
14.1. Hatvány függvény alakú minimálpolinom	107
14.2. Nilpotens operátorok redukálása	107
14.3. Egyértelműség	109
14.4. Illusztrációk	111
14.5. A nilpotens felbontási tételek nélkül?	113
15 A JORDAN-NORMÁLALAK	115
15.1. Egy gyöktényezős minimálpolinom	115
15.2. Jordan-normálalak: az általános eset	116
16 DETERMINÁNS	119
16.1. Permutációk	119
16.2. Mértékek	122
16.3. A mértékek jellemzése	127
16.4. A determináns kifejtése	130
16.5. A karakterisztikus polinom	131
17 SKALÁRISZSORZATOS TEREK GEOMETRIÁJA	135
17.1. Definíciók	135
17.2. Egyenlőtlenségek	137
17.3. Pont és altér távolsága	138
17.4. Ortogonalizáció	139
17.5. Projekciós tételek	141
17.6. Ortonormált rendszer teljessége	142
18 AZ ADJUNGÁLT OPERÁTOR BEVEZETÉSE	145
18.1. Riesz-reprezentáció véges dimenziós skaláriszszorzatos-térben	145
18.2. Az adjungált operáció	145
18.3. Önadjungált transzformációk	148
18.4. Unitér transzformációk	148
19 NORMÁLIS TRANSZFORMÁCIÓK DIAGONALIZÁLHATÓSÁGA	151
20 ORTOGONÁLIS PROJEKCIÓK	155
20.1. Ortogonális projekciók lineáris kombinációja	157
21 SPEKTRÁLIS FELBONTÁSOK	161
21.1. Komplex eset	161
21.2. Valós eset	164
21.3. Operátornorma	165
FÜGGELÉKEK	169

A A KOMPLEX SZÁMOKRÓL	171
A.1. Lineáris algebrai megközelítés	171
A.2. Analízis megközelítés	173
A.3. A komplex számok egyértelműsége	175
B A FROBENIUS-NORMÁLALAK	179
C KVADRATIKUS ALAKOK DEFINITSÉGÉNEK EGY NAIV TARGYALÁSA	187
C.1. Szimmetrikus mátrix diád felbontása	188
C.2. Definitiség osztályozása a teljes négyzetek segítségével	192
IRODALOM	193
TÁRGY MUTATÓ	195

I. rész

Ősz

1. fejezet

Előzmények

A LINEÁRIS ALGEBRA tárgyalásához elengedhetetlenül szükséges általános algebrai ismereteket foglaljuk össze.

1.1. Algebrai struktúrák

1.1. definíció (n -változós művelet). Legyen H egy halmaz. Egy

$$\varphi: H^n \rightarrow H$$

függvényt n -változós műveletnek nevezünk. Egy halmazt és rajta véges sok műveletet együtt *algebrai struktúrának* mondunk. Jelölés:

$$(H, \varphi_1, \dots, \varphi_n),$$

ahol H a halmaz és $\varphi_1, \dots, \varphi_n$ a H halmazon értelmezett műveletek.

1.2. definíció (félcsoport). Egy $(S, *)$ algebrai struktúrát *félcsoportnak* mondjuk, ha $*$ egy kétváltozós asszociatív művelete az S halmaznak, azaz minden $a, b, c \in S$ mellett

$$a * (b * c) = (a * b) * c.$$

Lefordítva ez azt jelenti, hogy

1. minden $a, b \in S$ mellett $a * b \in S$, és
2. minden $a, b, c \in S$ esetén $(a * b) * c = a * (b * c)$

1.3. definíció (neutrális elem). Az $(S, *)$ félcsoportban az $s \in S$ elem *balról (jobbról) neutrális*, ha $s * t = t$ ($t * s = t$) minden $t \in S$ mellett. Ha $s \in S$ balról is és jobbról is neutrális, akkor s -et egy *neutrális elemnek* mondjuk. A félcsoportot *neutrális elemes félcsoportnak* nevezzük, ha van benne neutrális elem.

1.4. állítás. Ha egy félcsoportban, van egy balról neutrális elem és egy jobbról neutrális elem, akkor ezek megegyeznek. Emiatt egy neutrális elemes félcsoportban neutrális elem csak egy van.

Bizonyítás: Legyen s_1 balról- és s_2 jobbról neutrális elem. Ekkor $s_2 = s_1 * s_2 = s_1$. □

A félcsoport additív írásmódja esetén természetes a neutrális elemet *zérusnak*, míg multiplikatív írásmód esetén *egységnak* nevezni.

1.5. definíció (csoport). Egy $(G, *)$ algebrai struktúrát *csoportnak* nevezünk, ha neutrális elemes félcsoport, amelyben minden $g \in G$ -hez létezik $g' \in G$, hogy

$$g * g' = e = g' * g. \tag{†}$$

Itt $e \in G$ jelöli a G csoport neutrális elemét.

1.6. definíció-állítás (inverz elem). Legyen $(G, *)$ egy csoport. Ekkor minden $g \in G$ -hez, csak egyetlen $g' \in G$ létezik, amelyre a fenti (\dagger) azonosság fennáll. Adott g -hez ezt ez egyetlen $g' \in G$ elemet, amelyre (\dagger) teljesül a g elem *inverzének* mondjuk.

Bizonyítás: Jelölje e a csoport neutrális elemét, és tegyük fel, hogy g', g'' inverz elemei g -nek. Azt mutatjuk meg, hogy $g' * g = e$ és $g * g'' = e$ esetén a két inverz megegyezik.

$$g' = g' * e = g' * (g * g'') = (g' * g) * g'' = e * g'' = g''. \quad \square$$

Példaként gondoljuk meg, hogy az összes $H \rightarrow H$ függvények halmaza a kompozíció műveettel neutrális elemes félcsoportot, és az összes $H \rightarrow H$ kölcsönösen egyértelmű függvények halmaza a kompozíció műveettel csoportot alkotnak. Ez utóbbi csoportot mondjuk *permutáció csoportnak*.

1.7. állítás (egyszerűsítési szabály). *Csoportban igaz az egyszerűsítési szabály, azaz*

$$a * c = b * c \implies a = b.$$

Bizonyítás: $a = a * e = a * (c * c') = (a * c) * c' = (b * c) * c' = b * (c * c') = b * e = b.$ \square

1.8. definíció (Abel-csoport). Egy $(G, *)$ csoportot *Abel-csoportnak* nevezünk, ha a művelete *kommutatív* is, azaz minden $s, t \in G$ mellett $s * t = t * s$.

1.9. definíció (gyűrű). A kétműveletes $(R, +, \cdot)$ algebrai struktúrát *gyűrűnek* nevezzük, ha

1. $(R, +)$ Abel-csoport;
2. (R, \cdot) félcsoport;
3. és a két műveletet összeköti a következő két disztributivitás:

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Ha (R, \cdot) neutrális elemes félcsoport, akkor azt mondjuk, hogy R egy *egységelemes gyűrű*, és ha (R, \cdot) kommutatív félcsoport, akkor azt mondjuk, hogy R egy *kommutatív gyűrű*.

1.10. definíció (test). Egy $(\mathbb{F}, +, \cdot)$ kétműveletes algebrai struktúrát *testnek* nevezünk, ha olyan kommutatív egységelemes gyűrű, amelyben minden nemzérus¹ elemnek van inverze², és $0 \neq 1^3$.

A test az algebrai struktúra, ahol az összeadás és szorzás műveletekkel úgy számolhatunk, mint amit a valós számok során megszoktuk. Példaként néhány tulajdonság.

1.11. állítás. Az $(R, +, \cdot)$ egységelemes gyűrűben minden $a \in R$ mellett

$$a \cdot 0 = 0 \text{ és } (-1) \cdot a = -a.$$

Bizonyítás: $0 + a \cdot 0 = a \cdot 0 = a (0 + 0) = a \cdot 0 + a \cdot 0$. A jobboldali $a \cdot 0$ -val való egyszerűsítés után kapjuk, hogy $0 = a \cdot 0$. A második azonosságot az első felhasználásával kapjuk: $0 = 0a = (1 + (-1)) \cdot a = 1 \cdot a + (-1) \cdot a = a + (-1) \cdot a$. Az additív inverz definíciója és egyértelműsége szerint ez éppen azt jelenti, hogy $-a = (-1) \cdot a$. \square

Ami nagyon fontos, hogy egy gyűrűben nem feltétlen teljesül, hogy elemek szorzata csak úgy lehet zérus, ha legalább az egyik elem zérus. Számunkra a legfontosabb példa a mátrixok gyűrűje⁴, ahol pont ennek a hiánya jelenti nehézséget.

Egy testben ilyen nem fordulhat elő.

1.12. definíció (nullosztómentes gyűrű). Egy gyűrűt *nullosztómentesnek* nevezzük, ha két elem szorzata csak úgy lehet nulla, ha legalább az egyik elem nulla.

1.13. állítás. Egy test egyben nullosztómentes gyűrű, azaz ha \mathbb{F} egy test, és $a, b \in \mathbb{F}$. Akkor

$$ab = 0 \implies a = 0 \text{ vagy } b = 0.$$

¹Értsd: minden elemnek, amely a + műveletre nézve neutrális elemtől különbözik.

²Értsd: a · szorzás neutrális elemére mint egységelemre nézve.

³Értsd: az összeadásra nézve és a szorzásra nézve képzett neutrális elemek nem azonosak.

⁴Lásd kicsit később.

Bizonyítás: Tegyük fel, hogy $ab = 0$. Ha $b \neq 0$, akkor létezik $b' \in \mathbb{F}$, hogy $bb' = 1$. Így

$$0 = 0b' = (ab)b' = a(bb') = a1 = a.$$

□

1.14. állítás. Nullosztómentes gyűrűben nem zérus elemmel való szorzatot egyszerűsíteni lehet azaz, ha $a, b, c \in R, b \neq 0$ esetén

$$ab = cb \implies a = c.$$

Bizonyítás: $(a - c)b = ab - cb = 0 \implies a - c = 0$.

□

1.15. definíció (ideál). Egy $(R, +, \cdot)$ kommutatív gyűrű egy $J \subseteq R$ nem üres részhalmazát *ideálnak* nevezünk, ha

1. minden $a, b \in J$ mellett $a + b \in J$;
2. minden $c \in R$ és minden $a \in J$ mellett $ca \in J$.

Ha egy $d \in R$ adott, akkor a

$$\{da : a \in R\}$$

halmaz egy ideálja R -nek. Ez a d elem többszöröseiből álló ideál, amelyet *főideálnak* is nevezünk. Ha egy gyűrűben minden ideál egy főideál, akkor a gyűrűt *főideál-gyűrűnek* mondjuk.

A generált ideál fogalma nagyon fontos.

1.16. definíció-állítás (generált ideál). Legyen adott a kommutatív, egységelemes $(R, +, \cdot)$ gyűrűben véges sok a_1, \dots, a_r elem. Az e véges sok elemet tartalmazó ideálok közös része maga is ideál, és e metszet az eredeti véges halmazt tartalmazó *legszűkebb ideál*. Jelöljük ezt $J(a_1, a_2, \dots, a_r)$ módon.

Tekintsük a $\left\{ \sum_{j=1}^r a_j b_j : b_1, \dots, b_r \in R \right\}$ halmazt. Világos, hogy ez egy ideál az R gyűrűben. A gyűrű egységelemes, ezért ennek $J(a_1, \dots, a_r)$ egy részhalmaza. Másrészt minden az $\{a_1, \dots, a_r\}$ elemeket tartalmazó ideál, egyben tartalmazza a $\left\{ \sum_{j=1}^r a_j b_j : b_1, \dots, b_r \in R \right\}$ halmazt is, ami azt jelenti, hogy

$$J(a_1, \dots, a_r) = \left\{ \sum_{j=1}^r a_j b_j : b_1, \dots, b_r \in R \right\}$$

az a_1, \dots, a_r elemeket tartalmazó legszűkebb ideál. Nevezük ezt az ideált az a_1, \dots, a_r elemek *generálta ideálnak* is.

Világos, hogy $\{0\}$ és maga az egész R ideálok.

A legfontosabb struktúrák számunkra a következők:

- Egységelemes gyűrű, amelyben a nullosztómentesség nem teljesül: mátrixok.
- Kommutatív egységelemes gyűrű, amely nullosztómentes de mégsem test: polinomok.
- Test: a valós vagy a komplex számok.

1.2. Polinomgyűrűk

1.17. definíció (polinom). Legyen \mathbb{F} egy test. E test feletti polinomokon az összes

$$p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_n t^n$$

alakú formális algebrai kifejezést értjük. Itt n tetszőleges nem negatív egész és $\alpha_0, \dots, \alpha_n$ tetszőleges, az \mathbb{F} testbeli elemek. Az \mathbb{F} test feletti összes polinomok halmazát $\mathbb{F}[t]$ módon jelöljük.

A fenti definícióban az *algebrai kifejezés* szó arra utal, hogy az $\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_n t^n$ műveletek minden $t \in \mathbb{F}$ mellett értelmesek, és eredményük egy újabb \mathbb{F} testbeli elem. Ha $t \in \mathbb{F}$ konkrétan meg van adva, akkor a behelyettesítés után kapott elemet mondjuk a p polinom helyettesítési értékének.

A *formális algebrai kifejezés* arra utal, hogy egy polinomot az együtthatói határozzák meg, azaz két polinom akkor és csak akkor azonos, ha az megfelelő együtthatói azonosak. Ez szemben áll avval, hogy ha a polinomokra mint függvényekre tekintenénk, akkor a helyettesítési értékek egyenlősége jelentné a két polinom azonos voltát. A formális szó tehát azt jelenti, hogy nem mint függvényre gondolunk, hanem egyszerűen az adott $\alpha_0, \dots, \alpha_n$ rögzített elemek – ezeket mondjuk együtthatóknak –, által előírt műveletekre. Az az előírás ugyanis, hogy tetszőleges $t \in \mathbb{F}$ mellett hajtsuk végre az

$$\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_n t^n$$

műveletsort. A műveletsorról és nem annak eredményéről van szó.

Két műveletsor akkor azonos, ha ugyanazok a műveletsor meghatározó $(\alpha_0, \alpha_1, \dots, \alpha_n)$ ⁵ együtthatók.⁶ A jelölések megértése is fontos. $p(t) \in \mathbb{F}[t]$ semmi más nem jelent, minthogy $p(t)$ egy polinom. Persze a polinom nem keverendő össze a helyettesítési értékkel, hiszen az egyik egy algebrai kifejezés-együttet, a másik egy az adott testbeli elem. Szokásos viszont, hogy ha nincs konkrét t a szövegkörnyezetben, akkor is $p(t)$ jelöli a polinomot. Néha egyszerűbben csak p -vel jelöljük, főleg akkor ha nincs szó behelyettesítésről, emiatt érdektelen a változó jele. Ritkábban, de előfordul, hogy egy konkrét értékre, mondjuk $s \in \mathbb{F}$ -re kell kiértékelni a polinomot ilyenkor $p(s)$ jelöli azt a testbeli elemet, amelyet t helyett s -et téve az előírt műveletek kiértékelése után kapunk. A szövegkörnyezetben mindenkor a polinom normálisan kell lennie, hogy $p(t)$ a polinomot jelenti, vagy egy konkrét t -re kiértékelte testbeli elemet.

1.18. definíció (polinom foka). Legyen $p(t) \in \mathbb{F}[t]$ egy polinom. Azt mondjuk, hogy az n nem negatív egész szám a *polinom fokszáma*, ha n a legnagyobb indexű nem zérus együttható. A legnagyobb indexű nem zérus együtthatót *főegyütthatónak* nevezzük. Azt mondjuk, hogy egy nemzérus polinom *normált*, ha 1 a főegyütthatója.

A $p(t) = 0$ konstans zérus polinom foka megállapodás szerint legyen $-\infty$. A p polinom fokszámát $\deg p$ módon jelöljük.

Látni fogjuk, hogy a konstans zérus polinomra $\deg p = -\infty$ csak egy kényelmes jelölés. Időnként a polinom fokszámával műveleteket is végezünk. Megegyezés szerint ilyenkor $-\infty + a = -\infty$ minden a nem negatív egész számra, és $(-\infty) + (-\infty) = -\infty$. A $-\infty$ szimbólumot minden egész számnál határozottan kisebbnek gondoljuk.

Két polinom összegét és szorzatát a szokásos módon definiáljuk:

1.19. definíció. Legyen $p, q \in \mathbb{F}[t]$, két polinom.

$$p(t) = \sum_{j=0}^n \alpha_j t^j \text{ és } q(t) = \sum_{j=0}^m \beta_j t^j, \quad \alpha_j, \beta_j \in \mathbb{F}, 0 \leq n, m \in \mathbb{Z}.$$

Ekkor a p és q összegének definíciója:

$$(p+q)(t) = \sum_{j=0}^{\max\{m,n\}} (\alpha_j + \beta_j) t^j;$$

míg a két polinom szorzatának definíciója:

$$(pq)(t) = \sum_{j=0}^{n+m} c_j t^j \text{ ahol } c_j = \sum_{k=0}^j \alpha_k \beta_{j-k}.$$

Mind az összeadás, mind a szorzás definíciójában úgy kell a formulát érteni, hogy amikor a hivatkozott együtthatók nem léteznek, akkor értékük legyen zérus. Például a szorzat legmagasabb indexű együtthatójára $c_{n+m} = \alpha_n \beta_m$, hiszen az α_k, β_{n+m-k} számok csak egyetlen k mellett értelmezettek közösen, mikor $k = n$.⁷

⁵Az előbbi zárójellel azt hangsúlyozzuk, hogy az együtthatók sorrendje is számít.

⁶Persze felmerül a kérdés, hogy ha két polinom minden helyettesítési értéke azonos, akkor igaz-e, hogy mint formális polinomok is azonosak, tehát a két polinom együtthatói is rendre azonosak-e? A pozitív választ később látjuk nem véges számtest, például a valós vagy a komplex test, feletti polinomok esetén. Lásd az 1.32. állítás utáni megjegyzést a 18. oldalon.

⁷Ha már követjük a konvenciót, amely szerint a nem definiált együtthatókat zérusnak tekintjük, akkor persze $(p+q)(t) = \sum_{j=0}^{\infty} (\alpha_j + \beta_j) t^j$; és $(pq)(t) = \sum_{j=0}^{\infty} (\sum_{k=0}^{\infty} \alpha_k \beta_{j-k}) t^j$ is írható. Itt a fenti összegek nem végtelen összegek, hiszen a két polinom együtthatói csak véges sok esetben különböznek zérustól. Így ebben az esetben a formálisan végtelen összeg definíciója egyszerűen a véges sok nem zérus elem összege.

1.20. állítás. Legyenek $p, q \in \mathbb{F}[t]$ polinomok az \mathbb{F} test felett. Ekkor

1. $\deg(pq) = \deg p + \deg q$;
2. $\deg(p+q) \leq \max\{\deg p, \deg q\}$.

Bizonyítás: Figyeljünk arra, hogy a konstans zérus polinom esetében is működik a téTEL, és vegyük észre, hogy az szorzat polinomra vonatkozó állítás azért igaz, mert a test nullosztómentes. \square

A következő állítás igazolását az olvasóra bízom. Aprólékosan igazoljuk a test axiómák gyakorlásaként. Vigyázat: a szorzás asszociativitása nem is olyan egyszerű.

1.21. állítás. Egy \mathbb{F} test feletti $\mathbb{F}[t]$ formális polinomok a fent bevezetett összeadás és szorzás műveletekkel, nullosztómentes, kommutatív, egységelemes gyűrűt alkotnak.

1.3. Polinomok oszthatósága és a maradékos osztás

1.22. definíció (oszthatóság). Azt mondjuk, hogy a $p \in \mathbb{F}[t]$ osztója az $f \in \mathbb{F}[t]$ nem zérus polinomnak, ha létezik $h \in \mathbb{F}[t]$, hogy $f(t) = p(t)h(t)$. Ilyenkor f -et a p egy többszörösének is mondjuk. Jelölés: $p \mid f$.

Világos, hogy egy p polinom összes többszörösei – tehát azok, amelyeknek p osztója – ideál alkotnak. Ez a p generálta legszűkebb ideál, azaz a $J(p) = \{fp : f \in \mathbb{F}[t]\}$ főideál. Ha $q \in J(p)$, akkor $J(q) \subseteq J(p)$, azaz ha q egy többszöröse p -nek, akkor q minden többszöröse p -nek is többszöröse. Ha p, q polinomok, amelyekre $p \mid q$ és $q \mid p$ akkor a két polinom csak konstans szorzóban különbözik egymástól. Ha például a két polinom még normált is, akkor $p \mid q$ és $q \mid p$ csak $p = q$ esetben lehetséges. A polinomgyűrű ideáljaira fókuszálva, azt gondoltuk éppen meg, hogy $J(p) = J(q)$ normált p, q polinomokra csak úgy teljesülhet, ha $p = q$, azaz a polinomok gyűrűjében minden főideálnak csak egy generáló eleme van a normált polinomok körében.

A következő állítás szerint a polinomok között is működik a maradékos osztás, ahogyan azt az egész számok között megszoktuk.

1.23. állítás (maradékos osztás). Legyenek $p, q \in \mathbb{F}[t]$ polinomok, $q \neq 0$. Ekkor létezik egyetlen $h, r \in \mathbb{F}[t]$ polinom, amelyre

$$p = hq + r; \quad \deg r < \deg q.$$

Bizonyítás: Először is azt vegyük észre, hogy $\deg p < \deg q$ esetben $r = p, h = 0$ szereposztással készen is vagyunk.

Tegyük fel tehát, hogy $n = \deg p \geq \deg q = m$, és lássuk be az állítást n szerinti indukcióval. Ha $n = 0$, akkor $p(t) = \alpha_0$ és $q(t) = \beta_0 \neq 0$. Ekkor persze

$$\alpha_0 = \frac{\alpha_0}{\beta_0} \beta_0 + 0,$$

ami azt jelenti, hogy $h(t) = \frac{\alpha_0}{\beta_0}$ és $r(t) = 0$ szereposztás megfelelő.

Most tegyük fel, hogy igaz az állítás $n + 1$ -nél kisebb fokú p polinomokra ($n \geq 0$), és lássuk be egy pontosan $n + 1$ -ed fokú polinomra. Legyen tehát

$$p(t) = \alpha_{n+1}t^{n+1} + \dots + \alpha_0 \quad \text{és} \quad q(t) = \beta_m t^m + \dots + \beta_0,$$

ahol $m \leq n + 1$. Tekintsük a következő polinomot:

$$\frac{\alpha_{n+1}}{\beta_m} t^{n+1-m} q(t).$$

Világos, hogy ennek főegyütthatója éppen α_{n+1} és foka éppen $n + 1 = \deg p$. Így a

$$p_1(t) = p(t) - \frac{\alpha_{n+1}}{\beta_m} t^{n+1-m} q(t).$$

polinomra $\deg p_1 < \deg p$. Na most, ha $\deg p_1 < \deg q$, akkor a bizonyítás első mondatában említett helyzetben vagyunk, tehát nyilvánvaló szereposztással az állítás igaz p_1 -re és q -ra. Ha viszont $\deg p_1 \geq \deg q$

még mindig igaz, akkor az indukciós feltétel szerint található $h, r \in \mathbb{F}[t]$ polinom, amelyre igaz az állítás. Mindkét esetben találtunk tehát h, r polinomokat, amelyre

$$p(t) - \frac{\alpha_{n+1}}{\beta_m} t^{n+1-m} q(t) = p_1(t) = h(t)q(t) + r(t); \quad \deg r < \deg q$$

teljesül. Ekkor persze

$$p(t) = \left(h(t) + \frac{\alpha_{n+1}}{\beta_m} t^{n+1-m} \right) q(t) + r(t); \quad \deg r < \deg q$$

is fennáll. Ezt kellett belátni az állítás egzisztencia részéhez.

Az unicitaás részhez tegyük fel, hogy valamely h, h_1, r, r_1 polinomokra

$$h(t)q(t) + r(t) = p(t) = h_1(t)q(t) + r_1(t)$$

teljesül, ahol $\deg r < \deg q$ és $\deg r_1 < \deg q$. Persze átrendezve ekkor

$$(h(t) - h_1(t))q(t) = r_1(t) - r(t)$$

is fennáll. Ekkor a fokszámokra figyelembe véve

$$\deg(h - h_1) + \deg q = \deg(r_1 - r) \leq \max\{\deg r_1, \deg(-r)\} < \deg q.$$

Ez csak akkor lehetséges, ha $\deg(h - h_1) = -\infty$, ami azt jelenti, hogy $h = h_1$, amiből persze $r_1 = r$ már látszik is. \square

1.24. állítás (a polinomgyűrű egy főideál-gyűrű). *A polinomok $\mathbb{F}[t]$ kommutatív, egységelemes, nullosztómentes gyűrűjében minden a $\{0\}$ -tól különböző ideált generál az ideálban lévő egyetlen normált minimális fokszámú polinom. Igy $\mathbb{F}[t]$ egy főideál-gyűrű.*

Bizonyítás: Legyen a J egy ideálja $\mathbb{F}[t]$ -nek, amely nem csak a zérus elemből áll. Vegyünk egy minimális fokszámú de nem zérus polinomot J -ben, tehát olyat, amely maga sem zérus és nála kisebb fokszámú polinom már nincs J -ben a 0 elemen kívül. Legyen ez d . Most megmutatjuk, hogy minden $p \in J$ -re $d|p$. A maradékos osztás szerint valamely h, r polinomokra

$$p(t) = h(t)d(t) + r(t); \text{ ahol } \deg r < \deg d.$$

Mivel $p, d \in J$, és J egy ideál, ezért $r \in J$. No de, d konstrukciója szerint ilyen csak a zérus polinom van, ezért valóban $d|p$. Ez éppen azt jelenti, hogy $J = \{dh : h \in \mathbb{F}[t]\}$ azaz d generálja a J ideált. Azt viszont már korábban is meggyondoltuk, hogy egy főideált csak egyetlen normált polinom generál.

Megmutattuk tehát, hogy egyetlen normált, minimális fokszámú polinom van J -ben, és minden J -beli polinom ennek többszöröse. \square

Érdemes eltenni magunknak, hogy az ideál generáló eleme, tehát az ideálbeli elemek közös osztója éppen az ideál minimális fokszámú nem zérus polinomja. Ilyenből a normált polinomok közül csak egy van.

1.25. definíció. Legyenek most p_1, \dots, p_k polinomok.

1. A d polinom a legnagyobb közös osztója az adott polinomoknak, ha

- a) $d|p_j$ minden $j = 1, \dots, k$ -ra,
- b) ha $d_1|p_j$ minden $j = 1, \dots, k$ mellett akkor $d_1|d$ is fennáll,
- c) d normált.

A p_1, \dots, p_k polinomokat relatív prímeknek nevezzük, ha közös osztójuk csak a konstans polinomok, azaz a $d(t) = 1$ a legnagyobb közös osztó.

2. A d polinom a legkisebb közös többszöröse az adott polinomoknak, ha

- a) $p_j|d$ minden $j = 1, \dots, k$ -ra,
- b) ha $p_j|d_1$ minden $j = 1, \dots, k$ mellett akkor $d|d_1$ is fennáll.

c) d normált.

Persze az első kérdés, hogy van-e a polinomoknak legnagyobb közös osztója vagy legkisebb közös többszöröse, és hány ilyen van?

1.26. állítás. Bármely $p_1, \dots, p_r \in \mathbb{F}[t]$ nem zérus polinomoknak létezik egyetlen legkisebb közös többszöröse. Jelesül, a legkisebb közös többszörös, a $\cap_{j=1}^r J(p_j)$ ideálnak, mint főideálnak az egyetlen normált generátora.

Bizonyítás: Világos, hogy ideálok metszete is ideál, emiatt $\cap_{j=1}^r J(p_j)$ is ideál $\mathbb{F}[t]$ gyűrűben. De itt minden ideál főideál, létezik tehát $d \in \mathbb{F}[t]$ normált polinom, amelyre

$$J(d) = \cap_{j=1}^r J(p_j)$$

Világos, hogy $d \in J(p_j)$ minden j -re, ergo d többszöröse minden p_j -nek. Ha $p_j|d_1$ fennáll, minden j -re az azt jelenti, hogy $d_1 \in J(p_j)$, minden j -re, azaz $d_1 \in \cap_{j=1}^r J(p_j) = J(d)$, tehát $d|d_1$ valóban fennáll.

Ha d mellett g is legkisebb közös többszörös, akkor $d|g$ és $g|d$ szerint g és d foka azonos, így csak egymás konstans szorosai lehetnek, de mivel mindenketen normáltak, ezért e konstans csak 1 lehet. \square

1.27. állítás. Bármely $p_1, \dots, p_r \in \mathbb{F}[t]$ nem zérus polinomoknak létezik egyetlen legnagyobb közös osztója. Jelesül, a legnagyobb közös osztó a $J(p_1, \dots, p_r)$ ideálnak mint főideálnak az egyetlen generáló eleme. Ezért a legnagyobb közös osztó kifejezhető

$$d(t) = f_1(t)p_1(t) + \dots + f_r(t)p_r(t)$$

alakban, valamely $f_1, \dots, f_r \in \mathbb{F}[t]$ polinomok segítségével.

Bizonyítás: Láttuk, hogy létezik egyetlen normált d polinom, amelyre $J(d) = J(p_1, \dots, p_r)$. Világos, hogy $d|p_j$ minden $j = 1, \dots, r$ és $d \in J(p_1, \dots, p_r)$, azaz

$$d = f_1 p_1 + \dots + f_r p_r$$

valamely f_1, \dots, f_r polinomokra. Ha valamely d_1 polinomra $d_1|p_j$ minden $j = 1, \dots, r$ mellett, akkor a fenti azonosság szerint $d_1|d$ is fennáll.

Az egyértelműség mint a legkisebb közös többszörösnél. \square

Az előző állítás kiemelt azonosságát *Bezout-azonosságnak* mondjuk.

1.28. állítás. Legyenek a $p_1, \dots, p_r \in \mathbb{F}[t]$ tetszőleges \mathbb{F} test feletti polinomok. Ezek pontosan akkor relatív prímek, ha léteznek $f_1, \dots, f_r \in \mathbb{F}[t]$ polinomok, hogy

$$f_1(t)p_1(t) + f_2(t)p_2(t) + \dots + f_r(t)p_r(t) = 1$$

A szakaszt a maradékos osztás módszerének másik fontos következményeivel zárjuk. Azt gondoljuk meg, hogy a gyöktényező a polinomból mindenkielhető, emiatt egy akármilyen test feletti n -ed fokú polinom gyökeinek száma n -nél nagyobb nem lehet.

1.29. állítás. Legyen $p \in \mathbb{F}[t]$ egy nem zérus polinom, és t_0 egy gyöke, azaz $p(t_0) = 0$. Ekkor létezik $h \in \mathbb{F}[t]$ nem zérus polinom, amelyre

$$p(t) = (t - t_0)h(t).$$

Bizonyítás: Maradékos osztással p -re és az elsőfokú $t - t_0$ polinomra

$$p(t) = h(t)(t - t_0) + r(t), \text{ ahol } \deg r < 1.$$

No de, t_0 egy gyök, tehát $0 = p(t_0) = r(t_0)$. Ez azt jelenti, hogy $\deg r = -\infty$, ami éppen az állítás. \square

1.30. definíció (gyök multiplicitása). Legyen t_0 gyöke a $p(t)$ polinomnak. Azt mondjuk, hogy a k pozitív egész t_0 gyök multiplicitása, ha van olyan $h(t)$ polinom, hogy $p(t) = (t - t_0)^k h(t)$, de $h(t_0) \neq 0$. Néha azt is mondjuk, hogy t_0 egy k -szoros gyöke p -nek.

Teljesen világos, hogy a gyöktényező kiemelhetősége miatt minden gyök legalább egyszeres multiplicitású. A következő állítás szerint a gyökök száma még a multiplicitásukkal együtt számolva sem lehet több mint a polinom foka.

1.31. állítás. *Legyen(ek) a $p \in \mathbb{F}[t]$ nem zérus polinom különböző gyökei t_1, \dots, t_k , és ezen gyökök multiplicitásai rendre m_1, \dots, m_k . Ekkor $m_1 + \dots + m_k \leq \deg p$.*

Bizonyítás: A test nulosztó mentessége és a gyöktényező kiemelhetősége miatt

$$p(t) = (t - t_1)^{m_1} \cdot (t - t_2)^{m_2} \cdots (t - t_k)^{m_k} \cdot h(t),$$

ahol h olyan nem zérus polinom, amelynek már nincsen gyöke. A fokszámok összehasonlításából kapjuk, hogy $m_1 + \dots + m_k \leq m_1 + \dots + m_k + \deg h = \deg p$. \square

A fenti gondolat szerint, ha egy legfeljebb n -ed fokú polinomnak $n + 1$ különböző gyöke van, akkor csak úgy lehetséges, ha a polinom minden együtthatója nulla. Ezt úgy is szoktuk fogalmazni, hogy egy legfeljebb n -ed fokú polinomot $n + 1$ helyettesítési értéke már egyértelműen meghatározza:

1.32. állítás. *Tegyük fel, hogy a $p, q \in \mathbb{F}[t]$ polinomok legfeljebb n -ed fokúak, ahol n egy nemnegatív egész, és tegyük fel, hogy létezik $n + 1$ különböző t_0, t_1, \dots, t_n pont a testben, amelyekre $p(t_j) = q(t_j)$ minden $j = 0, \dots, n$. Ekkor $p(t) = q(t)$, azaz p és q együtthatói azonosak.*

Bizonyítás: Legyen $h = p - q$. Világos, hogy $\deg h \leq n$ és h -nak van $n + 1$ különböző gyöke. Az előző állítás szerint ez csak a $h = 0$ polinomra igaz, ami azt jelenti, hogy p és q együtthatói azonosak. \square

A maradékos osztás tételenek szép következménye tehát, hogy ha \mathbb{F} egy nem véges test, és a $p, q \in \mathbb{F}[t]$ polinomok, akkor p -nek és q -nak pontosan akkor azonosak az együtthatói, ha $p(t) = q(t)$ fennáll minden $t \in \mathbb{F}$ mellett.

Itt fontos, hogy \mathbb{F} nem véges test, hiszen például ha $\mathbb{F} = \{0, 1\}$ a két elemű test, akkor a $p(t) = t^2 + t$ polinomra minden $t \in \{0, 1\}$ mellett $p(t) = 0$, de a polinom együtthatói rendre a $\{0, 1, 1\}$ számok a testből, tehát ez nem a összeadásra nézve neutrális eleme az $\mathbb{F}[t]$ polinomgyűrűnek.

Konklúzióképpen: megnyugodhatunk, hogy az iménti szörnyűség nem véges testek esetén nem fordulhat elő, tehát mondjuk a valós vagy a komplex számtest felett mindegy, hogy a polinomokat függvényeknek, vagy formális algebrai kifejezéseknek gondoljuk. A lényeg hogy egy n -ed fokú polinomot az $n + 1$ együtthatója, definíció szerint, de az $n + 1$ különböző helyen felvett helyettesítési értéke is egyértelműen meghatározza.

1.4. Az Euklideszi-algoritmus

Algoritmust keresünk polinomok legnagyobb közös osztójának és legkisebb közös többszörösének meghatározására. Ha egy pillanatra (p_1, \dots, p_n) jelöli az adott p_1, \dots, p_n polinomok legnagyobb közös osztóját, akkor nem nehéz meggondolni, hogy

$$((p_1, \dots, p_{n-1}), p_n) = (p_1, \dots, p_n).$$

Ezt $n = 3, 4, \dots$ számokra alkalmazva azt kapjuk, hogy ha módszerünk van két polinom legnagyobb közös osztójának meghatározására, akkor evvel már akárhány – persze véges sok – polinom legnagyobb közös osztója is meghatározható. Analóg módon ugyanez igaz a legkisebb közös többszörösre is. Azt gondoltuk meg tehát, hogy ha meg tudnánk határozni két polinom legnagyobb közös osztóját és legkisebb közös többszörösét, akkor ugyanezt már meg tudnánk tenni több polinom esetén is.

Az Euklideszi-algoritmus két polinom legnagyobb közös osztójának meghatározására szolgál. Az eddigi ismereteink szerint a p, q legnagyobb közös osztója a $J(p, q)$ ideál legalacsonyabb fokú, normált tagja. Az Euklideszi-algoritmus egy véges sok lépésben végrehajtható algoritmus, egy a definíciót sokkal hatékonyabb eljárás, két polinom generálta főideál generáló elemének, azaz a legnagyobb közös osztónak a meghatározására.

Az algoritmus megértése előtt emlékeznünk kell arra, hogy ha $p, q \in \mathbb{F}[t]$ valamely polinomok, akkor $J(p, q) = \{fp + gq : f, g \in \mathbb{F}[t]\}$ jelöli a p és a q polinomokat tartalmazó legszűkebb ideált. Világos, hogy ha $r_1, r_2 \in J(p, q)$, akkor $J(r_1, r_2) \subseteq J(p, q)$.

1.33. állítás (Euklidesz). Legyen $p, q \in \mathbb{F}[t]$ nem zérus polinomok. Definiálja $p_{-1} = p, p_0 = q$. Folytatva, ha valamely $i \geq 0$ számra p_{i-1} és p_i már definiált és $p_i \neq 0$, akkor a maradékos osztás szerint létezik egyetlen $h_{i+1}, p_{i+1} \in \mathbb{F}[t]$ polinom, amelyre

$$p_{i-1} = h_{i+1}p_i + p_{i+1}; \text{ ahol } \deg p_{i+1} < \deg p_i. \quad (\dagger)$$

Mivel minden egyes lépésben csökken a fokszám, ezért van olyan $s \geq 0$, hogy $p_s \neq 0$, de $p_{s+1} = 0$. Erre a p_s polinomra $J(p_s) = J(p, q)$, ezért p_s normáltja a p és a q polinomok legnagyobb közös osztója.

Bizonyítás: A fenti algoritmussal olyan $p_{-1}, p_0, p_1, \dots, p_s, p_{s+1}$ polinomokat kapunk, amelyekre minden $i = 0, \dots, s$ mellett a (\dagger) azonosság fennáll, és $\deg p_{s+1} = -\infty$, azaz $p_{s+1} = 0$.

A (\dagger) azonosság szerint, minden $i = 0, 1, \dots, s$ mellett $p_{i-1} \in J(p_i, p_{i+1})$, amiből a $J(p_{i-1}, p_i) \subseteq J(p_i, p_{i+1})$ tartalmazás adódik. Másrészt a (\dagger) azonosságot értelmezhetjük úgy is, hogy $p_{i+1} \in J(p_{i-1}, p_i)$, amiből persze $J(p_i, p_{i+1}) \subseteq J(p_{i-1}, p_i)$ következik. Így minden $i = 0, \dots, s$ -re végül is

$$J(p_{i-1}, p_i) = J(p_i, p_{i+1}).$$

Alkalmazva ezt minden $i = 0, \dots, s$ mellett

$$J(p, q) = J(p_{-1}, p_0) = J(p_0, p_1) = J(p_1, p_2) = \dots = J(p_s, p_{s+1}) = J(p_s). \quad \square$$

Most a legkisebb közös többszörös algoritmikus meghatározására törekszünk.

1.34. állítás. Legyenek a $p, q \in \mathbb{F}[t]$ polinomok relatív prímek, és tegyük fel, hogy $p \nmid qr$. Ekkor $p \mid r$.

Bizonyítás: Mivel a p és a q polinomok relatív prímek, ezért a Bezout-azonosság szerint van f és g polinom, amelyekre $fpr + gqr = r$. A feltétel szerint qr a p többszöröse, így az iménti azonosság baloldala a p többszöröse, ami éppen azt jelenti, hogy $p \mid r$. \square

1.35. állítás. Tekintsük a $p, q \in \mathbb{F}[t]$ normált polinomokat. Jelölje d a legnagyobb közös osztót, és m a legkisebb közös többszöröst. Ekkor

$$d(t)m(t) = p(t)q(t).$$

Bizonyítás: Legyen $p = dr_1$ és $q = dr_2$. Először megmutatjuk, hogy r_1 és r_2 relatív prímek. A Bezout-azonosság szerint valamely f, g polinomra $d = fdr_1 + gdr_2$. Kihasználva, hogy nullosztómentes gyűrűben nem zérus elemmel egyszerűsíteni lehet, kapjuk az $1 = fr_1 + gr_2$ azonosságot. Ez azt jelenti, hogy r_1 és r_2 relatív prímek.

Most megmutatjuk, hogy

$$dr_1r_2 \quad (\dagger)$$

a legkisebb közös többszörös. Világos, hogy ez egy közös többszöröse a p, q polinomoknak. Tegyük fel, hogy s egy másik közös többszörös, azaz $s = ps_1$ és $s = qs_2$. Ekkor $dr_1s_1 = ps_1 = s = qs_2 = dr_2s_2$, amiből újra a nullosztómentesség szerint

$$r_1s_1 = r_2s_2.$$

Na most, a fent kiemelt azonosság szerint szerint $r_1 \mid r_2s_2$, ahol r_1 és r_2 relatív prímek. Ebből azonnal kapjuk, hogy $r_1 \mid s_2$. No de, $s = qs_2 = dr_2s_2$, amiből már látszik, hogy s egy többszöröse a dr_1r_2 polinomnak. Az is világos, hogy (\dagger) normált polinom, ez az m legkisebb közös többszörös. Innen $d \cdot m = d(dr_1r_2) = (dr_1)(dr_2) = p \cdot q$ már nyilvánvaló. \square

A fenti állítás csak két polinomra igaz, többre nem, de nekünk csak két polinomra kell. Úgy interpretáljuk, hogy ha a szorzatot osztom maradékosan a legnagyobb közös osztóval, akkor a maradék minden zérus, és a hánnyados éppen a legkisebb közös többszörös. Azt gondoltuk meg tehát, hogy az Euklideszi-algoritmus módszert ad két polinom legkisebb közös többszörösének algoritmikus meghatározására is.

Visszatérve a szakasz elején felvett gondolatra, ilyen módon véges sok lépésben végrehajtható algoritmust kapunk véges sok polinom legkisebb közös többszörösének meghatározására. Például öt polinom legkisebb közös többszöröséhez, egy Euklideszi-algoritmussal meghatározzuk az első kettő polinom legnagyobb közös osztóját, majd egy újabb maradékos osztással az első kettő legkisebb közös többszörösét. Ugyanezt teszem az így kapott és a harmadik polinommal, az eredmény az első három polinom legkisebb közös többszöröse. Az így kapott polinommal és a negyedik polinommal egy újabb Euklideszi-algoritmus és egy újabb maradékos osztás után kapjuk az első négy polinom legkisebb közös többszörösét, majd ennek eredményével és az ötödik polinommal mint két polinomnak a legkisebb közös többszörösével kapjuk az eredeti öt polinom legkisebb közös többszörösét.

1.5. Polinom faktorizáció

Kicsi korunk óta sulykolják belénk, hogy minden egész szám előáll, még ha csak egyfél leképpen prímek szorzataként. Ha ismerjük két szám prímtényezős előállítását, akkor nagyon könnyű megmondani a két szám legkisebb közös többszörösét, vagy a legnagyobb közös osztóját. Evvel a probléma csak annyi, hogy nagyon nehéz megmondani két, esetleg jó nagy, szám prímtényezős előállítását, így még az egész számok gyűrűjében is az Euklideszi-algoritmus a megfelelő, véges sok lépésben, végrehajtható módszer a legnagyobb közös osztó és a legkisebb közös többszörös konkrét felírására.

Ebben a szakaszban azt mutatjuk meg, hogy prímtényezős előállításról szóló tételek a polinomok gyűrűjében is igaz marad. Természetesen konkrét algoritmust nem adunk, hiszen még számokra sem igen van.⁸

1.36. definíció (reducibilis polinom). Egy polinomot *reducibilisnek* mondunk, ha előáll mint két legalább elsőfokú polinom szorzata. Egy nem reducibilis polinomot *irreducibilisnek* nevezünk.

Világos, hogy minden legfeljebb elsőfokú polinom tetszőleges test felett irreducibilis. A magasabb fokú polinomok esetében a probléma nagyban függ a testtől is, ahonnan a polinom együtthatói származnak. A következő állítás viszont minden test mellett igaz.

1.37. állítás (polinom faktorizáció). *Tetszőleges test feletti polinomgyűrűben, minden (normált) polinom előáll mint (normált) irreducibilis polinomok szorzata.*

Bizonyítás: Az előállítandó polinom foka szerinti teljes indukció. Elsőfokú polinom maga irreducibilis.

Most tegyük fel, hogy az állítás igaz n -nél alacsonyabb fokú polinomokra és lássuk be $\deg p = n$ mellett, ahol $n > 1$. Ha p irreducibilis, akkor megint készen vagyunk. Ha $p = fg$ valamely $\deg f \geq 1$ és $\deg g \geq 1$ mellett, akkor $\deg f < n$ és $\deg g < n$. Az indukciós feltevés szerint f és g , emiatt $p = fg$ is előáll irreducibilis polinomok szorzataként. \square

Érdemes látni, hogy ha a d polinom egy irreducibilis p polinom osztója, akkor csak $\deg d = 0$, vagy $\deg d = \deg p$ lehetséges. Ezért egy tetszőleges f polinomra igaz, hogy vagy $p \mid f$ vagy p és f relatív prímek. Ugyanis, ha az f és p polinomok d legnagyobb közös osztója nem 1, akkor csak $d = p$ lehetséges, ergo $p \mid f$.

1.38. állítás (prím tulajdonság). *Legyen $p \in \mathbb{F}[t]$ irreducibilis polinom, amelyre $p \mid (f_1 \cdots f_n)$, valamely $f_j \in \mathbb{F}[t]$ polinomokra, ahol $j = 1, \dots, n \geq 1$. Ekkor létezik $1 \leq j \leq n$, amelyre $p \mid f_j$.*

Bizonyítás: A polinomok n száma szerinti indukció. Az $n = 1$ eset semmitmondó módon teljesül. Tegyük fel, hogy igaz az állítás n -nél kevesebb polinomra, és lássuk be n -re. Itt $n \geq 2$. Indulunk ki tehát abból, hogy

$$p \mid (f_1 \cdots f_{n-1}) \cdot f_n$$

Ha p osztója lenne az $f_1 \cdots f_{n-1}$ szorzatnak, akkor az indukciós feltevés szerint készen is lennének. Ha p nem osztója a szorzatnak, akkor p irreducibilis volta miatt relatív prímek. Ekkor az 1.34. állítás szerint $p \mid f_n$. \square

Az is igaz, hogy ha egy polinom reducibilis, akkor az nem prím, de ezt a gyakorlatokra hagyjuk. Összegében látjuk tehát, hogy a polinomok irreducibilitása és prím tulajdonsága azonos fogalmak.

Az irreducibilis polinomok prím tulajdonsága segítségével a polinomok faktorizáció egyértelműségét is igazolhatjuk.

1.39. állítás. *Legyen $p \in \mathbb{F}[t]$ egy legalább elsőfokú normált polinom. Ekkor ezek sorrendjétől eltekintve egyértelműen léteznek legalább elsőfokú, normált, irreducibilis $q_1, \dots, q_s \in \mathbb{F}[t]$ polinomok, hogy $p = q_1 \cdots q_s$.*

A „sorrendtől eltekintés” alatt azt értjük, hogy ha p előáll

$$p_1 \cdots p_s = p = q_1 \cdots q_r$$

legalább elsőfokú, normált, irreducibilis polinomok szorzataként, akkor $s = r$ és a p_1, \dots, p_s polinomok alkalmás átindexelése után $p_j = q_j$, minden $j = 1, \dots, s$ mellett.

⁸Természetesen számokra működik az az algoritmus, amit általános iskolában tanultunk, csak annyira lassú, hogy egy nagy szám prímtényezős előállítását esetleg többszáz évig számolná a földkerekesség valamennyi számítógépe. Ez egy szép példája annak, mennyire hasznos lehet az a tudás, ami valaminek a nem tudásáról szól, hiszen emiatt tudunk biztonságos módon két számla között pénzt mozgatni. Lásd például: [https://en.wikipedia.org/wiki/RSA_\(cryptosystem\)](https://en.wikipedia.org/wiki/RSA_(cryptosystem))

Bizonyítás: Az 1.37. állítás szerint p előáll, mint normált, irreducibilis polinomok szorzata. Itt legalább egy polinom legalább elsőfokú, hiszen p legalább elsőfokú. Ha van a szorzat tagjai között nullafokú, akkor az csak 1 lehet a normáltság szerint, ezért egyszerűen elhagyható. Meggondoltuk tehát, hogy p legalább elsőfokú, normált, irreducibilis polinomok szorzata.

Az unicitást, a felbontandó polinom fokszáma szerinti indukcióval igazoljuk: Ha a polinom elsőfokú, akkor a felbontás egyértelműsége is nyilvánvaló. Tegyük fel, hogy igaz az egyértelműség n -nél alacsonyabb fokszámú polinomokra, és tegyük fel, hogy $\deg p = n > 1$. Nézzünk két lehetséges előállítást

$$p_1 \cdots p_s = p = q_1 \cdots q_r,$$

ahol $p_1, \dots, p_s, q_1, \dots, q_r$ legalább elsőfokú, normált, irreducibilis polinomok. A q_1 polinom osztója a jobboldalnak, ezért a baloldalnak is. A prím tulajdonság miatt, 1.34. állítás, q_1 osztója az egyik baloldali polinomnak. Alkalmas átindexelés után feltehető, hogy $q_1 | p_1$. No de, p_1 is irreducibilis, és $\deg q_1 \geq 1$ miatt csak $\deg q_1 = \deg p_1$ lehetséges, tehát a normáltság szerint $q_1 = p_1$. A polinomgyűrű nullosztó mentessége szerint az első polinomokkal egyszerűsíthetünk, ergo

$$p_2 \cdots p_s = q_2 \cdots q_r,$$

is fennáll. A fenti polinom már n -nél alacsonyabb fokú, így az indukciós feltevés szerint $s - 1 = r - 1$, és alkalmas átindexelés után minden $j = 2, \dots, s$ esetén is teljesül a $p_j = q_j$ egyenlőség. \square

Az egész szakaszt összefoglalhatjuk így is:

1.40. állítás. *Minden legalább elsőfokú normált polinomhoz léteznek – még hozzá sorrendjükű eltekintve egyértelműen – olyan q_1, \dots, q_s egymástól különböző, normált, irreducibilis polinomok; és olyan n_1, \dots, n_s pozitív egészek; amelyekre*

$$p(t) = q_1^{n_1}(t) \cdots q_s^{n_s}(t).$$

1.6. Mátrixok

1.41. definíció (mátrix). Egy tetszőleges test feletti mátrixnak nevezünk, a test elemeiből képzett táblázatot. Ha $m, n \in \mathbb{N}$ előre rögzített pozitív egészek és az A táblázatnak m sora és n oszlopa van, akkor azt mondjuk, hogy A egy $m \times n$ méretű mátrix. Az \mathbb{F} test feletti $m \times n$ -es mátrixok halmazát $\mathbb{F}^{m \times n}$ módon jelöljük.

Ha $A \in \mathbb{F}^{m \times n}$ egy mátrix, akkor A_i jelöli az i -edik sort, ami persze egy $1 \times n$ -es mátrix; A^j jelöli a j -edik oszlopot, ami persze egy $m \times 1$ -s mátrix; $A_{i,j}^j$ jelöli az i -edik sor j -edik elemét. Sokszor használjuk az $A_{i,j} = A_{i,j}^j$ jelölést is.

Időnként, azt hangsúlyozandó hogy mátrixokról van szó a mátrixot jelölő betűt kapcsos zárójelbe teszem. Pl. $[A] \in \mathbb{F}^{m \times n}$.

A mátrixot a mérete és az elemei határozzák meg. Emiatt két mátrix akkor azonos, ha azonos méretűek, és a megfelelő elemeik is azonosak.

Az azonos típusú mátrixok között műveleteket definiálunk:

1.42. definíció (mátrixok összege). Rögzített $m, n \in \mathbb{N}$ mellett, ha $A, B \in \mathbb{F}^{m \times n}$, akkor ezek összege az a $C \in \mathbb{F}^{m \times n}$ mátrix, amelyre

$$C_{i,j} = A_{i,j} + B_{i,j}$$

minden $i = 1, \dots, m$ és $j = 1, \dots, n$. Jelölés: $C = A + B$.

1.43. állítás. Az $m \times n$ méretű mátrixok az fent definiált összeadás művelettel Abel-csoportot alkotnak. A $[0]$ -val jelölt neutrális elem az az $m \times n$ -s mátrix, amelynek minden eleme a test zérus eleme:

$$[0]_{i,j} = 0;$$

az $[A]$ mátrix additív inverze az az $[-A]$ -val jelölt $m \times n$ méretű mátrix, amelyre

$$[-A]_{i,j} = -([A]_{i,j}).$$

Most definiáljuk egy számnak és egy mátrixnak a szorzatát.

1.44. definíció (szám és mátrix szorzata). Ha $\alpha \in \mathbb{F}$ egy szám és $A \in \mathbb{F}^{m \times n}$ egy mátrix, akkor ezek szorzata az $\alpha A \in \mathbb{F}^{m \times n}$ módon jelölt mátrix, melynek elemeire

$$[\alpha A]_{i,j} = \alpha [A]_{i,j}.$$

Könnyen ellenőrizhetők a következő számolási szabályok:

1.45. állítás. Legyenek $A, B \in \mathbb{F}^{m \times n}$ mátrixok, és $\alpha, \beta \in \mathbb{F}$ tetszőleges számok. Ekkor

1. $\alpha (A + B) = \alpha A + \alpha B$;
2. $(\alpha + \beta) A = \alpha A + \beta A$;
3. $(\alpha\beta) A = \alpha (\beta A)$;
4. $1A = A$.

Az utolsó két állítást, 1.43. és 1.45., együtt később úgy fogjuk fogalmazni, hogy adott test feletti tetszőleges méretű mátrixok *vektorteret* alkotnak. Ha az A_1, \dots, A_k azonos méretű mátrixok adottak, akkor tetszőleges $\alpha_1, \dots, \alpha_k \in \mathbb{F}$ számok mellett értelmes az A_1, \dots, A_k mátrixoknak az $\alpha_1, \dots, \alpha_k$ számokkal mint együtthatókkal képzett *lineáris kombinációja*:

$$\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_k A_k.$$

Minden eddig megértett dolog persze az $m = 1$ speciális esetben, és az $n = 1$ speciális esetben is igaz. Ha $m = 1$, akkor a mátrixot *sorvektornak*; ha pedig $n = 1$, akkor a mátrixot *oszlopvektornak* mondjuk.

Most a mátrixok között a szorzás műveletet definiáljuk.

1.46. definíció (mátrixok szorzata). Legyen $A \in \mathbb{F}^{m \times k}$ és $B \in \mathbb{F}^{k \times n}$ mátrix. Fontos, hogy A oszlopainak száma azonos B sorainak számával. Ezek $C = AB$ szorzata egy $C \in \mathbb{F}^{m \times n}$ mátrix, melynek elemeit az alábbi egyenlőség definiálja

$$[C]_{i,j} = \sum_{s=1}^k [A]_{i,s} [B]_{s,j}.$$

Itt persze $i = 1, \dots, m$ és $j = 1, \dots, n$.

Figyeljünk arra, hogy a fenti definícióban az A és B mátrixok sorrendje is fontos. Például, ha $n \neq m$, akkor BA nem is értelmes. Még $m = n$ esetén is BA egy $k \times k$ méretű mátrix, míg AB egy $n \times n$ méretű mátrix, azaz az AB szorzatnak általában véve semmi köze a BA szorzathoz.

A bevezetett összeadás és szorzás műveletekre teljesül a disztributivitás is:

1.47. állítás. Legyen $A \in \mathbb{F}^{m \times k}$, valamint $a, B, C \in \mathbb{F}^{k \times n}$ mátrixok. Ekkor

$$A(B + C) = AB + AC.$$

Hasonlóan, ha $A, B \in \mathbb{F}^{m \times k}$, valamint $a C \in \mathbb{F}^{k \times n}$ mátrixok, akkor

$$(A + B)C = AC + BC.$$

Bizonyítás: Az első disztributivitást mutatjuk meg, a második evvel analóg. Világos, hogy minden két oldalon azonos méretű mátrixok vannak. minden szóba jövő i, j index mellett

$$\begin{aligned} [A(B + C)]_{i,j} &= \sum_{s=1}^k [A]_{i,s} [B + C]_{s,j} = \sum_{s=1}^k [A]_{i,s} ([B]_{s,j} + [C]_{s,j}) = \sum_{s=1}^k ([A]_{i,s} [B]_{s,j}) + ([A]_{i,s} [C]_{s,j}) = \\ &= \sum_{s=1}^k [A]_{i,s} [B]_{s,j} + \sum_{s=1}^k [A]_{i,s} [C]_{s,j} = [AB]_{i,j} + [AC]_{i,j} = [AB + AC]_{i,j}. \end{aligned}$$

Ez éppen az $A(B + C) = AB + AC$ mátrixazonosságot jelenti. \square

Szinte banalitás, de egy mátrix az elemei összesége, de a sorai összesége, sőt az oszlopai összesége is. A következő ultrafontos állítás arról szól, hogy a szorzás operációt hogyan kell látnunk attól függően, hogy a szorzat mátrixra, mint elemei összeségére, mint oszlopai összeségére, vagy mint sorainak összeségére gondolunk.

1.48. állítás. Legyen $A \in \mathbb{F}^{m \times k}$ és $B \in \mathbb{F}^{k \times n}$ mátrix. Jelölje $C = AB$ ezek szorzatát ebben a sorrendben. Ekkor

1. A szorzat mátrix i -edik sorának j -edik eleme, az A mátrix i -edik sorának és a B mátrix j -edik oszlopának, mint speciális mátrixoknak a szorzata. Magyarul: minden $1 \leq i \leq m$ és $1 \leq j \leq n$ mellett

$$[C]_{i,j} = [A]_i \cdot [B]^j.$$

2. A szorzat mátrix minden oszlopa az A mátrix oszlopainak a B mátrix megfelelő oszlopából vett elemekkel képzett lineáris kombinációja. Magyarul: minden $1 \leq j \leq n$ mellett

$$[C]^j = \sum_{s=1}^k [B]^j_s [A]^s.$$

3. A szorzat mátrix minden sora a B mátrix sorainak az A mátrix megfelelő sorából vett elemekkel képzett lineáris kombinációja. Magyarul: minden $1 \leq i \leq m$ mellett

$$[C]_i = \sum_{s=1}^k [A]^s_i [B]_s.$$

4. A szorzat mátrix az A oszlopaiból, és a B soraiból alkotott diádok összege. Magyarul:

$$[C] = \sum_{s=1}^k [A]^s [B]_s.$$

Diádnak nevezzük egy oszlop- és egy sorvektor szorzatát. Ha az oszlopnak és a sornak rendre azonosak az elemei, akkor szimmetrikus diádról beszélünk.

1. bizonyítása: $[A]_i \cdot [B]^j = \sum_{s=1}^k [A]_{i,s} [B]_{s,j} = [C]_{i,j}$. □

2. bizonyítása: $\left[\sum_{s=1}^k [B]^j_s [A]^s \right]_i = \sum_{s=1}^k [B]^j_s [A]^s_i = \sum_{s=1}^k [A]^s_i [B]^j_s = [C]_{i,j} = [C]^j_i$ minden i -re. □

3. bizonyítása: $\left[\sum_{s=1}^k [A]^s_i [B]_s \right]^j = \sum_{s=1}^k [A]^s_i [B]_s^j = [C]_{i,j} = [C]_i^j$ minden j -re. □

4. bizonyítása: $\left[\sum_{s=1}^k [A]^s [B]_s \right]_{i,j} = \sum_{s=1}^k [[A]^s [B]_s]_{i,j} = \sum_{s=1}^k [A]^s_i [B]_s^j = C_{i,j}$ minden i -re j -re. □

1.49. definíció (Kronecker-delta, identitás mátrix). Kroencker-deltának nevezzük az alábbi egyszerű szimbólumot:

$$\delta_{i,j} = \begin{cases} 1, & \text{ha } i = j; \\ 0, & \text{egyébként.} \end{cases}$$

Adott $n \geq 1$ természetes számra az $n \times n$ méretű identitás mátrix azaz $I \in \mathbb{F}^{n \times n}$ mátrix, amelyre

$$[I]_{i,j} = \delta_{i,j}.$$

Nyilvánvaló, hogy ha $A \in \mathbb{F}^{m \times n}$ mátrix és $I \in \mathbb{F}^{n \times n}$ méretű identitás mátrix, akkor $A \cdot I = A$. Hasonlóan, ha most I az $m \times m$ identitás mátrixot jelöli, akkor pedig $I \cdot A = A$ azonosság teljesül. Persze, ha $A \in \mathbb{F}^{n \times n}$ négyzetes mátrix és $I \in \mathbb{F}^{n \times n}$ az ugyanilyen méretű identitás mátrix, akkor

$$IA = AI = A$$

is fennáll.

A mátrixok szorzásának legérdekesebb tulajdonsága a szorzás asszociativitása.

1.50. állítás. Legyen az A, B, C mátrixok úgy megadva, hogy AB is értelmes és BC is értelmes legyen, azaz $A \in \mathbb{F}^{m \times k}, B \in \mathbb{F}^{k \times l}, C \in \mathbb{F}^{l \times n}$. Ekkor

$$A(BC) = (AB)C.$$

Bizonyítás: Először is azt vegyük észre, hogy ha az A és a C mátrixok egyike egy szám, és a másik két mátrix összeszorozható, akkor a mátrix szorzás definíciója szerint az állítás nyilvánvaló.

Másodszor azt vegyük észre, hogy minden két oldalon azonos méretű konkrétan $m \times n$ méretű mátrixok szerepelnek.

Azt kell tehát még meggondolnunk, hogy az i -edik sor j -edik eleme minden két oldalon ugyanaz. A jobboldalon ez

$$\begin{aligned} [AB]_i \cdot [C]^j &= \left(\sum_{s=1}^k [A]_i^s [B]_s \right) [C]^j = \sum_{s=1}^k ([A]_i^s [B]_s) [C]^j = \sum_{s=1}^k [A]_i^s ([B]_s [C]^j) \\ &= \sum_{s=1}^k [A]_i^s \left(\sum_{r=1}^l [B]_s^r [C]_r^j \right) = \sum_{s=1}^k \sum_{r=1}^l [A]_i^s ([B]_s^r [C]_r^j). \end{aligned}$$

A baloldalon az i -edik sor j -edik eleme hasonló számolgatással:

$$\begin{aligned} [A]_i [BC]^j &= [A]_i \left(\sum_{r=1}^l [B]^r [C]_r^j \right) = \sum_{r=1}^l [A]_i ([B]^r [C]_r^j) = \sum_{r=1}^l ([A]_i [B]^r) [C]_r^j \\ &\quad \sum_{r=1}^l \left(\sum_{s=1}^k [A]_i^s [B]_s^r \right) [C]_r^j = \sum_{r=1}^l \sum_{s=1}^k ([A]_i^s [B]_s^r) [C]_r^j. \end{aligned}$$

A testben fennálló asszociativitás és kommutativitás miatt a bal- és a jobboldali kifejezés azonos. \square

Ha tehát adottak az A_1, A_2, \dots, A_p mátrixok, úgy hogy a sorban egymás után következőknek értelmes a szorzata, akkor ezek bármilyen zárójelzésével képzett szorzata is értelmes, és ugyanazt a mátrixot eredményezi. Mint azt a számok esetén megszoktuk, használhatjuk az $A_1 A_2 \dots A_p$ jelölést ezen mátrixok bármelyik zárójelzésével képzett szorzatára. Érdemes a fenti bizonyításból megjegyezni a formulát, amelyet három mátrix szorzatára kapunk:

$$[ABC]_{i,j} = \sum_{s=1}^k \sum_{r=1}^l [A]_{i,s} [B]_{s,r} [C]_{r,j}.$$

Foglaljuk össze az eddigieket négyzetes mátrixok mellett:

1.51. állítás. Legyen $n \in \mathbb{N}$ természetes szám, és tekintsük az $n \times n$ méretű mátrixok halmazát, ellátva ezt a halmazt a mátrix összeadással és a mátrixszorzással. Az $(\mathbb{F}^{n \times n}, +, \cdot)$ algebrai struktúra egy egységelemes gyűrű.

Ez a gyűrű, az $n > 1$ esetben biztosan nem kommutatív. Például $n = 2$ mellett

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \text{ amíg } \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Az sem igaz, hogy ez a gyűrű nullosztómentes lenne, hiszen például $n = 2$ mellett az

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

mátrix nyilván nem a zérus mátrix (az összeadásra nézve neutrális elem), de $A \cdot A = 0$. Ebből persze már az is következik, hogy a fenti A mátrixnak nincs a szorzásra nézve inverze, de ez a nélkül is nagyon egyszerűen látszik.⁹ Gyakorlatként próbálunk magasabb n számok mellett is a kommutativitás és a nullosztómentesség hiányára példát találni.

Nagyon fontos látni, hogy a kommutativitás hiánya, az eddigiek től eltérő számolási gyakorlatot eredményez. A számolás közben a mátrixok sorrendjén nem változtathatunk. Persze előfordul, hogy két mátrix szorzata nem függ a sorrendtől. Ilyenkor a két mátrixot egymással felcserélhetőnek, vagy kommutálónak mondjuk. Például, az identitás mátrixszal minden más mátrix kommutál. A mátrix azon elemeit, ahol a sor- és az oszlopindexek azonosak fődiagonálisbeli elemeknek mondjuk. Egy mátrixot *diagonális alakúnak* mondunk,

⁹Ha $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ jobboldali inverze lenne akkor a jobb alsó sarokra figyelve $0 \cdot b + 0 \cdot d = 1$ lenne, de egy testben $1 \neq 0$.

ha minden nem zérus eleme a fődiagonálisában van. Az is világos, hogy a diagonális mátrixok egymással kommutálnak. Fontos része az első féléves anyagnak, hogy ha két négyzetes mátrix szorzata az identitás mátrix, akkor e két mátrix egymással kommutál. Ez az eredmény távolról sem nyilvánvaló, és most nem is tudjuk belátni, mert ehhez már szükség van a Gauss–Jordan eliminációs algoritmusra¹⁰, vagy a lineáris függetlenség fogalmára, amiket majd később vezetünk be.

Nagyon is triviális mégis érdemes észrevenni, hogy a fentiek $n = 1$ esetben nem jelentenek problémát. Ilyenkor az 1×1 -es mátrixok tere voltaképpen azonos az \mathbb{F} -testtel, hiszen csak az a különbség, hogy egy testbeli a elemet $[a]$ módon írjuk. A szorzás és az összeadás definíciója ugyanazt adja, ha mint a testbeli elemre, vagy az ebből képzett 1×1 -es mátrixra gondolunk.

Az $n \times n$ -es négyzetes mátrixok másik érdemleges részstruktúrája az

$$\mathcal{F} = \{c \cdot I : c \in \mathbb{F}\}.$$

Itt I az $n \times n$ méretű identitás mátrix, tehát \mathcal{F} elemei azon diagonális alakú mátrixok, ahol minden elem a diagonálisban azonos. Világos, hogy két ilyen mátrix összege és szorzata is ilyen marad:

$$aI + bI = (a + b)I \text{ és } aI \cdot bI = (ab)I.$$

Ez azt jelenti, hogy az $(\mathcal{F}, +, \cdot)$ struktúra egy egységelemes gyűrű. Világos, hogy itt bármely két elem kommutál, ergo egy kommutatív egységelemes gyűrűvel állunk szemben és az is teljesen nyilván való, hogy minden nem zérus elemnek van a szorzásra nézve inverze. Azt kaptuk tehát, hogy a fenti \mathcal{F} minden n mellett egy test.

1.7. A komplex számok mint mátrixok

1.52. definíció (Izomorf testek). Legyenek \mathbb{F} és \mathbb{G} testek. Azt mondjuk, hogy a két test *izomorf* egymással, ha létezik közük *művelettartó bijekció*, azaz létezik

$$\varphi : \mathbb{F} \rightarrow \mathbb{G}$$

bijekció, amely tartja a műveleteket is, azaz

$$\varphi(a + b) = \varphi(a) + \varphi(b) \text{ és } \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b).$$

A művelettartó bijekciót *izomorfizmusnak* nevezzük.

Az izomorf testek között nem teszünk különbséget. Úgy tekintjük őket, hogy csak jelölésükben különböznek. Például, ha az \mathbb{R} valós számokra gondolunk, akkor a 2×2 -es diagonális alakú valós mátrixok közül azok, ahol a diagonális minden két eleme azonos, a valós testtel izomorf testet alkot.

$$\mathcal{R} = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} : a \in \mathbb{R} \right\} \text{ és } \varphi : \mathbb{R} \rightarrow \mathcal{R}, \text{ ahol } \varphi(a) = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

A 2×2 -es valós mátrixok egységelemes gyűrűjében tehát \mathcal{R} egy olyan részgyűrű, ami még test is, és izomorf az \mathbb{R} valós számtesttel. Voltaképpen azt csináltuk, hogy a valós számtestet beágyaztuk a 2×2 -es mátrixok közé, azaz egy a valós számot az $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ mátrixszal reprezentálunk (írunk le).

A 2×2 méretű valós mátrixok még sok-sok más testet is tartalmaznak.¹¹ Ezek közül számunkra a legfontosabb a következő részhalmaz.

1.53. definíció-állítás (komplex számtest). Jelölje két tetszőleges $a, b \in \mathbb{R}$ valós szám mellett $M_{a,b}$ az (a, b) valós számpárhoz tartozó $M_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ mátrixot. Tekintsük az ilyen típusú mátrixok \mathcal{C} -vel jelölt halmazát:

$$\mathcal{C} = \{M_{a,b} : a, b \in \mathbb{R}\}$$

E részhalmaz

¹⁰Ha viszont tudjuk mi az a Gauss–Jordan-elimináció, akkor márás megérthetjük a 2.19. állítást. Ehhez ugorjunk a 42. oldalra

¹¹Karakterizációjukat lásd: (Ebbinghaus és tsai. 1991)-ben.

1. zárt a mátrix összeadásra és a mátrix szorzásra, így a 2×2 -es valós mátrixok egy speciális egységelemes részgyűrűje.

2. E részgyűrűben a mátrix szorzás kommutatív művelet, és

3. és e részgyűrűben minden nem zérus mátrixnak van inverze is a mátrixszorzás műveletre nézve.

Eszerint a $(\mathcal{C}, +, \cdot)$ algebrai struktúra egy test. Ezt a testet nevezzük a *komplex számtestnek*, vagy a *komplex számtest mátrix reprezentációjának*.

Bizonyítás: Az $a, b, c, d \in \mathbb{R}$ valós számok mellett

$$M_{a,b} + M_{c,d} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} + \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} a+c & -b-d \\ b+d & a+c \end{pmatrix} = M_{a+c,b+d}$$

és hasonlóan a mátrix szorzás definíciója szerint

$$M_{a,b} \cdot M_{c,d} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \cdot \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac - bd & -ad - bc \\ bc + ad & -bd + ac \end{pmatrix} = M_{ac-bd,ad+bc}.$$

Mivel $M_{1,0}$ a 2×2 -es identitás mátrix, ezért \mathcal{C} a mátrix összeadásra és a mátrixszorzásra nézve egységelemes gyűrűt alkot.

A szorzás kommutativitánya is látszik a fenti számolásból, hiszen

$$M_{c,d} \cdot M_{a,b} = M_{ca-db,cb+da} = M_{ac-bd,ad+bc} = M_{a,b} \cdot M_{c,d}$$

a valós számok összeadásának és szorzásának kommutativitása miatt.

Legyen most $M_{a,b}$ egy nem zérus mátrix, így $a^2 + b^2 \neq 0$. Világos, hogy

$$M_{a,b} \cdot M_{a,-b} = M_{a^2+b^2,0} = (a^2 + b^2) M_{1,0} = (a^2 + b^2) I.$$

Ebből már látszik is, hogy $M_{a,b} \cdot M_{\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}} = I$. Ez a már igazolt kommutativitással éppen azt jelenti, hogy minden nem zérus elemnek van multiplikatív inverze, ergo \mathcal{C} valóban test. \square

Ebben a \mathcal{C} testben az $M_{0,1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ elem olyan, hogy a négyzete a szorzásra nézve reprodukáló elemek az összeadásra nézve képzett inverze:

$$M_{0,1} \cdot M_{0,1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = M_{-1,0} = -I.$$

Ez a tulajdonság azért figyelemre méltó, mert ha a szokásoknak megfelelően a test multiplikatív neutrális elemét az 1 szimbólummal jelöljük, akkor olyan elemet találtunk a komplex számtestben, amelynek négyzete éppen -1 . Tudjuk, hogy a valós számtest esetében ez nem lenne lehetséges.

Tekintsük most valamely $a, b \in \mathbb{R}$ mellett az

$$M_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}$$

felbontást. Ha bevezetjük az $i = M_{0,1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ jelölést, akkor minden komplex szám

$$M_{a,b} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + i \cdot \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}$$

alakban írható. Emlékezzünk arra, hogy a valós számtest is részhalmaza a komplex számtestnek abban az értelemben, ha minden a valós számot az $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ mátrixszal reprezentálunk. ($\mathcal{R} \subseteq \mathcal{C}$). Ha tehát megegyezünk abban, hogy az a valós számra nézve mi az $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ mátrixra gondolunk,¹² akkor azt kapjuk, hogy minden komplex szám

$$M_{a,b} = a + ib$$

¹²Kicsit pontosabban: a valós számok 2×2 -es mátrix reprezentációját használjuk.

alakú, ahol i egy olyan komplex szám, amelyre $i^2 = -1$, $a, b \in \mathbb{R}$. Ezt nevezzük a komplex szám *normálalakjának*.

Ne felejtsük a műveleteket: Láttuk, hogy $M_{a,b} + M_{c,d} = M_{a+c,b+d}$. Ez a normálalak reprezentáció mellett azt jelenti, hogy az összeadás definíciója csak

$$(a + ib) + (c + id) = (a + c) + i(b + d) \quad (\dagger)$$

lehet. Hasonlóan emlékszünk, hogy $M_{a,b} \cdot M_{c,d} = M_{ac-bd,ad+bc}$, ami normálalak reprezentáció mellett az

$$(a + ib) \cdot (c + id) = (ac - bd) + i(ad + bc) \quad (\ddagger)$$

definíciót eredményezi.

A következőket gondoltuk meg:

1.54. definíció-állítás (komplex számtest a normálakkal). Definiálja

$$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$$

a komplex számok normálalakját. Az összeadás műveletet definiálja (\dagger) , és a szorzás műveletet definiálja (\ddagger) . Az így kapott algebrai struktúra test, amely izomorf a komplex számtest mátrix reprezentációjával. Az izomorfizmust a

$$\varphi : \mathbb{C} \rightarrow \mathbb{C}, \quad \varphi \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a + ib$$

művelettartó bijekció hozza létre.

Ha már megértettük, hogy a komplex számok normálalak reprezentációja testet alkot, akkor a (\dagger) és (\ddagger) definíciók megjegyzése nagyon könnyű. Más nem is lehet: Az összeadáshoz (\dagger) csak el kell végezni a műveletet majd kiemelni i -t, a szorzás definíciójához (\ddagger) az i kiemelése után jutunk.

1.8. A komplex számok abszolútértéke

A valós számokra jól ismert abszolútérték függvényt terjesztjük ki komplex számtest elemeire.

1.55. definíció (valós rész, képzetes rész). Legyen $z \in \mathbb{C}$, $z = a + ib$. Ekkor $a \in \mathbb{R}$ a z komplex szám *valós része*, és $b \in \mathbb{R}$ a z komplex szám *képzetes része*. $\Re z$ jelöli a valós részt, és $\Im z$ a képzetes részt.

1.56. definíció (konjugált). Legyen $z \in \mathbb{C}$, $z = a + ib$. Ekkor z *konjugáltja* $\bar{z} = a - ib$.

A konjugált definíciója alapján, egyszerű számolatással kapjuk az alábbi azonosságokat.

1.57. állítás. Minden $z \in \mathbb{C}$ komplex szám mellett fennállnak a konjugálás következő szabályai:

$$\overline{(\bar{z})} = z, \quad z + \bar{z} = 2\Re z, \quad z - \bar{z} = 2i\Im z, \quad z \in \mathbb{R} \text{ akkor és csak akkor, ha } z = \bar{z}, \quad z\bar{z} = (\Re z)^2 + (\Im z)^2 \geq 0.$$

Az összeadás és szorzás művelet is felcserélhető a konjugálással, azaz bármely két $z, w \in \mathbb{C}$ komplex szám esetén igaz, hogy

$$\overline{z + w} = \bar{z} + \bar{w}, \quad \overline{z \cdot w} = \bar{z} \cdot \bar{w}, \quad \overline{z - w} = \bar{z} - \bar{w}, \quad \overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}} \quad \text{feltéve, hogy } w \neq 0.$$

Most használjuk először a valós számok rendezését. Az \mathbb{R} testen a \geq reflexív, antiszimmetrikus, tranzitív relációt az algebrai műveletekkel a következő két axióma kapcsolja össze: minden $a, b, c \in \mathbb{R}$ mellett

$$a \geq b \text{ esetén } a + c \geq b + c, \quad a, b \geq 0 \text{ esetén } ab \geq 0.$$

Az $a^2 - b^2 = (a + b)(a - b)$ azonosság szerint, az a, b nem negatív valós számok mellett $a \geq b$ akkor és csak akkor, ha $a^2 \geq b^2$. Speciálisan $a^2 = b^2$ akkor és csak akkor, ha $a = b$.

1.58. definíció (komplex szám abszolútértéke). Legyen $z \in \mathbb{C}$ egy komplex szám. Láttuk, hogy $z\bar{z} \in \mathbb{R}$ és $z\bar{z} \geq 0$. E szám négyzetgyökét nevezzük a z komplex szám *abszolútértékének*. Jelölés: $|z| = \sqrt{z\bar{z}}$.

Ha speciálisan $\Im z = 0$, tehát ha z egy valós szám, akkor e definíció szerint

$$|z| = \sqrt{z^2} = \begin{cases} z & , \text{ha } z \geq 0 \\ -z & , \text{ha } z < 0, \end{cases}$$

ami egybeesik a valós számok abszolútértékének definíciójával. A fentiből azonnal látszik, hogy minden $a \in \mathbb{R}$ valós szám mellett $a \leq |a|$.

Mivel $z\bar{z} = \bar{z}z$, ezért $|z| = |\bar{z}|$, azaz komplex számnak és konjugáltjának azonos az abszolútértéke. Szokásos technika, hogy a komplex szám abszolútértékére vonatkozó állításokat, az abszolútértékek négyzetére fogalmazzuk át. Ez megtehető, hiszen nem negatív valós számok egyenlősége és azok négyzetének egyenlősége azonos fogalmak. Például $|z| = 0$, akkor és csak akkor, ha $z\bar{z} = 0$, ami akkor és csak akkor teljesül ha $z = 0$. Hasonlóan, ha z, w komplex számok, akkor $|zw|^2 = zw(\bar{z}\bar{w}) = zw\bar{z}\bar{w} = z\bar{z}w\bar{w} = |z|^2|w|^2 = (|z||w|)^2$, ergo $|zw| = |z||w|$.

Világos, hogy $|\Re z|^2 \leq (\Re z)^2 + (\Im z)^2 = z\bar{z} = |z|^2$, emiatt $|\Re z| \leq |z|$. Ezt alkalmazva tetszőleges z, w komplex számok mellett

$$|\Re(z\bar{w})| \leq |z\bar{w}| = |z||w|,$$

amit *Cauchy–Schwartz-egyenlőtlenségnak* mondunk. Az abszolútérték legfontosabb tulajdonságai:

1.59. állítás. Legyen $z, w \in \mathbb{C}$ komplex szám. Ekkor

1. $|z| = 0$ akkor és csak akkor, ha $z = 0$,
2. $|zw| = |z||w|$,
3. $|z + w| \leq |z| + |w|$.

Az utolsó egyenlőtlenséget *háromszög-egyenlőtlenségnak* nevezik. Egy kevésbé népszerű, de ekvivalens alakja

$$|(|z| - |w|)| \leq |z - w|.$$

Bizonyítás: Csak a háromszög-egyenlőtlenséget nem gondoltuk meg eddig:

$$\begin{aligned} |z + w|^2 &= (z + w)(\bar{z} + \bar{w}) = (z + w)(\bar{z} + \bar{w}) = z\bar{z} + w\bar{w} + z\bar{w} + \bar{z}\bar{w} = \\ &|z|^2 + |w|^2 + 2\Re(z\bar{w}) \leq |z|^2 + |w|^2 + 2|\Re(z\bar{w})| \leq |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2. \end{aligned}$$

Az első egyenlőtlenség, hogy egy valós szám nem nagyobb az abszolútértékénél, a második egyenlőtlenség pedig a Cauchy–Schwartz-egyenlőtlenség.

A kevésbé populáris alakhoz használjuk a már igazolt egyenlőtlenséget: $|z| = |(z - w) + w| \leq |z - w| + |w|$, majd ugyanezt a z és a w számok felcserélése után mégegyeszer felírva

$$|z| - |w| \leq |z - w| \quad \text{és} \quad |w| - |z| \leq |w - z|.$$

Na most, azt ugyan nem tudjuk, hogy a $|z| - |w|$ szám a baloldali számok közül melyikkel azonos, de az egyikkel biztosan egybeesik. No de, mindenki is melyikkel, hiszen a jobboldal mindenkorban a kívánt $|z - w|$.

□

A definíció szerint $|z|^2 = \bar{z}z$. Véve mindenkorban abszolútértékét, azt kapjuk, hogy¹³,

$$|\bar{z}z| = |z|^2.$$

Persze valós számok esetén az iménti állítás 1., 2., 3. pontját korábban is sokszor használtuk már. Vegyük észre, hogy fent a valós esetet is újra igazoltuk.

¹³Ezt az azonosságot a komplex számok C*-azonosságának mondjuk

1.9. A komplex számok trigonometrikus alakja

Láttuk, hogy $z, w \in \mathbb{C}$ komplex szám mellett

$$\Re(z+w) = \Re z + \Re w \quad \text{és} \quad \Im(z+w) = \Im z + \Im w.$$

Ez azt jelenti, hogy ha az $a+ib$ komplex számot azonosítjuk az \mathbb{R}^2 sík (a, b) pontjával, akkor egyszerűen koordinátánként kell összeadni a komplex számokat, minthá a z komplex szám az origóból az (a, b) pontra mutató vektor lenne.

A kérdés, hogy ha így képzeljük a komplex számokat, akkor a komplex számok szorzása mit jelent a komplex számoknak megfeleltetett vektorok körében?

1.60. definíció-állítás (komplex szám trigonometrikus alakja). Legyen $z \in \mathbb{C}$ egy nem zérus komplex szám. Ekkor létezik $\varphi \in \mathbb{R}$ valós szám, hogy

$$z = |z|(\cos \varphi + i \sin \varphi) \quad (\dagger)$$

Egy ilyen φ számot nevezzük a z komplex szám egy *argumentumának*, és néha $\arg z$ módon jelöljük. A (\dagger) alak a komplex szám *trigonometrikus alakja*.

Két nem zérus komplex szám egyenlősége azt jelenti, hogy az abszolútértékük azonos, és az argumentumaiak különbsége 2π többszöröse.

Bizonyítás: Világos, hogy ha $z = a+ib \neq 0$, akkor $a^2 + b^2 > 0$, így

$$z = a+ib = \sqrt{a^2+b^2} \left(\frac{a}{\sqrt{a^2+b^2}} + i \frac{b}{\sqrt{a^2+b^2}} \right).$$

Ha x jelöli a fenti zárójelben a valós részt és y a képzetes részt, akkor $x^2 + y^2 = 1$. Így az (x, y) pár a sík egységgörének egy pontja. A trigonometrikus függvények középiskolai definíciója szerint, ha φ jelöli az $(1, 0)$ pontot az (x, y) ponttal összekötő ív hosszát, – az óramutató járásával ellentétes irányban mérve a körcikk peremén – akkor $x = \cos \varphi$ és $y = \sin \varphi$. \square

1.61. állítás. Legyenek $z, w \in \mathbb{C}$ nem nulla komplex számok a trigonometrikus alakjukban felírva, azaz

$$z = |z|(\cos \varphi + i \sin \varphi), \quad w = |w|(\cos \psi + i \sin \psi).$$

Ekkor a két szám szorzatának abszolútértéke az abszolútértékek szorzata és a szorzat egy argumentumát is megkapjuk mint az argumentumok összege. Magyarául:

$$zw = |z||w|(\cos(\varphi + \psi) + i \sin(\varphi + \psi)).$$

Speciálisan minden $n \in \mathbb{Z}$ egész számra

$$z^n = |z|^n(\cos n\varphi + i \sin n\varphi).$$

A szakasz bevetőjében feltett kérdésre tehát a válasz, hogy a z komplex számmal való szorzást a sík olyan geometriai transzformációjának képzelhetjük, amely egy $\arg z$ szöggel való forgatásból és egy $|z|$ -szeres origó középpontú nyújtásból áll.

1.62. definíció (egységgöök). Legyen az $n \in \mathbb{N}$ természetes szám rögzítve. Tetszőleges $k = 0, 1, \dots, n-1$ mellett jelölje

$$\epsilon_k^{(n)} = \cos k \frac{2\pi}{n} + i \sin k \frac{2\pi}{n}.$$

az úgynevezett n -edik komplex egységgöököt.

Az n -edik komplex egységgöök a komplex számsík pontosan n különböző pontját alkotják, és az n -edik hatványuk 1, azaz $(\epsilon_k^{(n)})^n = 1$. Emiatt minden $w \neq 0$ komplex szám mellett pontosan n komplex gyöke van a $t^n - w$ polinomnak. Ha ugyanis w trigonometrikus alakja $w = |w|(\cos \psi + i \sin \psi)$, akkor legyen például $z_0 = \sqrt[n]{|w|}(\cos(\psi/n) + i \sin(\psi/n))$, és így $(z_0 \epsilon_k^{(n)})^n = w$ minden $k = 0, \dots, n-1$ szám mellett.

1.10. Polinom faktorizáció a komplex- és a valós számtest felett

A komplex számtest felett minden nem konstans polinomnak van gyöke. Formálisabban:

Az algebra alaptétele. *Minden legalább elsőfokú komplex együtthatós polinomnak van komplex gyöke. Azaz ha $p(t) \in \mathbb{C}[t]$ egy nem konstans polinom, akkor létezik $z \in \mathbb{C}$ komplex szám, amelyre $p(z) = 0$.*

Az állítást itt nem tudjuk igazolni és egyelőre érdemes bizonyítás nélkül elfogadni. Illusztrációként megértettük, hogy miért igaz $p(t) = \alpha_0 + t^n$ alakú polinomra. A felépítés jelen pontján tekintsük a komplex számtest egy még nem igazolt tulajdonságának.

Most összefoglaljuk, hogy az algebra alaptétele mit jelent a komplex test feletti, majd a valós test feletti polinomok faktorizációjára nézve. Láttuk, hogy minden normált polinom előáll mint normált irreducibilis polinomok szorzata, emiatt azt kell meggondolnunk, hogy mik az irreducibilis polinomok.

1.63. állítás. *A $\mathbb{C}[t]$ polinomgyűrűben minden legalább másodfokú polinom reducibilis.*

Bizonyítás: Legyen p egy legalább másodfokú komplex polinom. Az algebra alaptétele miatt van gyöke, és tudjuk hogy a gyöktényező mindenkielhető. Így azt kapjuk, hogy $p(t) = (t - z)h(t)$ alakú, emiatt $2 \leq \deg p = 1 + \deg h$, ergo p valóban előállt mint két legalább elsőfokú polinom szorzata. \square

A polinomok szorzattá bontásáról szóló tételek a komplex számtest feletti speciális esete tehát:

1.64. állítás. *A $\mathbb{C}[t]$ polinomgyűrű minden legalább elsőfokú normált polinomjához léteznek a sorrendjükötől eltekintve egyetlen $z_1, \dots, z_s \in \mathbb{C}$ egymástól különböző komplex számok, és léteznek n_1, \dots, n_s pozitív egészek, amelyekre*

$$p(t) = (t - z_1)^{n_1} \cdots (t - z_s)^{n_s}.$$

Persze ez azt jelenti, hogy egy n -ed fokú komplex polinomnak minden pontosan n komplex gyöke van, ha a gyökök számát multiplicitással számoljuk. Az állítást egy kicsit egyszerűbben úgy is fogalmazhatjuk, hogy minden nem konstans komplex polinom előáll mint elsőfokú komplex polinomok szorzata. Az iménti mondat nyilván nem igaz a „komplex” szót a „valós” szóra cserélve, ugyanis minden negatív diszkriminánsú másodfokú valós polinom jó is ellenpéldának. Ez a jelenség az oka annak, hogy a komplex számtest felett sokkal kényelmesebb dolgoznunk mint a valós számok felett.

Most nézzük, hogy mit jelent az algebra alaptétele a valós test feletti polinomgyűrűre nézve.

1.65. állítás. *Legyen $p(t) \in \mathbb{C}[t]$ a komplex polinomgyűrű egy olyan eleme, amelynek minden együtthatója valós. Ekkor a $z \in \mathbb{C}$ komplex szám pontosan akkor gyöke p -nek, ha \bar{z} is gyöke p -nek. Sőt, ha z egy k -szoros multiplicitású gyök, akkor \bar{z} is pontosan k -szoros multiplicitású gyök.*

Bizonyítás: A p polinomra tehát $p(t) = \sum_{j=0}^n \alpha_j t^j$, ahol $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{R}$. A konjugálás tulajdonságai szerint egy z komplex számra

$$p(z) = \overline{\sum_{j=0}^n \alpha_j z^j} = \sum_{j=0}^n \overline{\alpha_j z^j} = \sum_{j=0}^n \bar{\alpha}_j \bar{z}^j = \sum_{j=0}^n \alpha_j \bar{z}^j = p(\bar{z}).$$

Emiatt valós együtthatós p polinomra és z komplex számra $p(z) = 0$ akkor és csak akkor, ha $p(\bar{z}) = 0$.

Ha a gyök speciálisan valós szám, akkor a multiplicitásra vonatkozó állítás semmit mondó. Ha most z egy nem valós komplex gyök, akkor $p(t) = (t - z)(\bar{t} - \bar{z})h(t)$, ahol $h(z) = 0$ pontosan akkor, ha $h(\bar{z}) = 0$. Így, ha z egy k szoros gyök, akkor

$$p(t) = (t - z)^k (t - \bar{z})^k \cdot h(t)$$

alakú, ahol h -nak már sem z sem \bar{z} nem gyöke. \square

Minden valós együtthatós polinom előáll mint első vagy másodfokú polinomok szorzata:

1.66. állítás. *Az $\mathbb{R}[t]$ polinomgyűrű minden legalább elsőfokú normált polinomjához léteznek $x_1, \dots, x_r \in \mathbb{R}$ valós számok, léteznek $\alpha_1, \beta_1, \dots, \alpha_s, \beta_s$ valós együttható párok, és léteznek $n_1, \dots, n_r, m_1, \dots, m_s$ pozitív egészek úgy, hogy*

$$p(t) = (t - x_1)^{n_1} \cdots (t - x_r)^{n_r} \cdot (\alpha_1 + \beta_1 t + t^2)^{m_1} \cdots (\alpha_s + \beta_s t + t^2)^{m_s}.$$

Itt $r, s \geq 0$, de $r + s > 0$, és $n_1 + \dots + n_r + 2(m_1 + \dots + m_s) = \deg p$, továbbá a jobboldalon szereplő másodfokú polinomok irreducibilisek.

Bizonyítás: Tekintsük a p polinomot mint a komplex számtest feletti polinomgyűrű egy elemét, és alkalmazzuk a komplex polinom faktorizációról szóló tételeit. A p így előáll mint elsőfokú, esetleg komplex polinomok szorzata. A gyököket osszuk két része. Legyenek x_1, \dots, x_r a különböző valós gyökök, amelyek multiplicitása rendre n_1, \dots, n_r . Legyenek $z_1, \bar{z}_1, \dots, z_s, \bar{z}_s$ a különböző nem valós de komplex gyökök, ahol m_1, \dots, m_s rendre a konjugált gyökpárok multiplicitása. Így

$$p(t) = (t - x_1)^{n_1} \cdot \dots \cdot (t - x_r)^{n_r} \cdot ((t - z_1) \cdot (t - \bar{z}_1))^{m_1} \cdot \dots \cdot ((t - z_s) \cdot (t - \bar{z}_s))^{m_s}.$$

Világos, hogy $r, s \geq 0$, de $r + s > 0$, és $n_1 + \dots + n_r + 2(m_1 + \dots + m_s) = \deg p$. A komplex faktorokra végezzük el a szorzást, így $(t - z_k)(t - \bar{z}_k) = t^2 - 2\Re z_k t + |z_k|^2$. Így $\alpha_k = |z_k|^2$ és $\beta_k = -2\Re z_k$, választással készen is vagyunk. \square

A fenti állításból két dolog azonnal látszik. Az első, hogy *valós számtest felett minden legalább harmadfokú polinom reducibilis*, a második pedig, hogy ha egy valós együtthatós polinomnak nincs valós gyöke, akkor csak páros fokú lehet, vagy ami ugyanaz: *minden páratlan fokú valós együtthatós polinomnak van valós gyöke*.

2. fejezet

A vektortér fogalma

A LINEÁRIS ALGEBRA kezdő fejezetéhez érkeztünk, miután áttekintettük azokat az általános algebrai ismereteket, amelyek nélkül nem tárgyalhatók a lineáris algebrához szükséges gondolatok.

2.1. definíció (Vektortér). Legyen adva egy \mathbb{F} test, és egy V halmaz. Tegyük fel, hogy adott egy $+ : V \times V \rightarrow V$ két változós művelet (ezt összeadásnak nevezzük) és adott egy $\cdot : \mathbb{F} \times V \rightarrow V$ szintén kétváltozós művelet (ezt skalárral való szorzásnak, vagy számmal való szorzásnak mondjuk). A V -t az \mathbb{F} test feletti *vektortérnek* nevezzük, ha $(V, +)$ egy Abel-csoport, és a számmal valós szorzás műveletre teljesülnek az alábbi axiómák:

1. minden $\alpha \in \mathbb{F}$ és minden $u, v \in V$ mellett $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$,
2. minden $\alpha, \beta \in \mathbb{F}$ és minden $u \in V$ mellett $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$,
3. minden $\alpha, \beta \in \mathbb{F}$ és minden $u \in V$ mellett $\alpha \cdot (\beta \cdot u) = (\alpha\beta) \cdot u$,
4. minden $u \in V$ esetén $1 \cdot u = u$.

Nagyon hasonlóan ahhoz, ahogyan testben is meggondoltuk igazak a következő számolási szabályok. minden $\alpha \in \mathbb{F}$ mellett

- i.) $\alpha \cdot 0 = 0$,
- ii.) $0 \cdot v = 0$,
- iii.) $(-1) \cdot v = -v$,
- iv.) $\alpha \cdot v = 0$ esetén $\alpha = 0$ vagy $v = 0$.

A számmal valós szorzás $\cdot : \mathbb{F} \times V \rightarrow V$ művelet eredményének szokásos rövidítése, hogy a kissé körülményes $\alpha \cdot v$ helyett csak αv -t írunk. Az is előfordul, különösen amikor egy konkrét vektortér konkrét műveletéről van szó, hogy az αv és a $v\alpha$ jelölést is ugyanarra az $\alpha \cdot v \in \mathbb{F}$ elemre használjuk.

Alapvető példa vektortérre a mátrixok tere. Az $m \times n$ méretű mátrixok $\mathbb{F}^{m \times n}$ halmaza a mátrix összeadással és a számmal valós szorzással vektorteret alkot az \mathbb{F} test feletti. Ha $n = 1$, akkor kapjuk az oszlopvektorok \mathbb{F}^m terét, ami így szintén egy \mathbb{F} feletti vektortér. Speciális esetként \mathbb{R}^m egy \mathbb{R} feletti vektortér, \mathbb{C}^m egy \mathbb{C} feletti vektortér.

Fontos példa még, egy adott X halmazból az \mathbb{F} testbe képező összes függvények halmaza a függvények között szokásos összeadás műveettel, és számmal való szorzással. Ezt a teret \mathbb{F}^X módon szokás jelölni, és minden nehézség nélkül ellenőrizhető, hogy \mathbb{F}^X egy \mathbb{F} feletti vektortér. Hasonlóan látszik, hogy például az összes valós-valós folytonos függvények is egy \mathbb{R} feletti vektorteret alkotnak, vagy ha ezek közül csak a differenciálható függvényekre szorítkozunk, akkor ezen függvénytér is vektortér az \mathbb{R} test feletti.

Rögzítéink kell magunkban, hogy a vektortér definíciójában a test is fontos szerepet játszik. Más test felett ugyanaz az Abel-csoport már egy másik vektortérhez tartozik. Világos például, hogy \mathbb{R} az \mathbb{R} test felett vektortér a szokásos műveletekkel, de látjuk majd, hogy egészen más tulajdonságai vannak annak a vektortérnek, amit akkor kapunk, ha az \mathbb{R} valós számok Abel-csoportját, mint a \mathbb{Q} racionális test feletti vektortérnek tekintjük.

Játékos példaként gondoljuk meg, hogy a pozitív valós számok egy az \mathbb{R} feletti vektorteret alkot a következő fura műveletekkel: Tetszőleges a, b pozitív valós szám mellett $a \# b = a \cdot b$, majd tetszőleges α valós szám és a pozitív valós számra legyen $\alpha * a = a^\alpha$.¹

¹Itt $a * \alpha$ mit jelent?

A vektorteret sokszor csak az additív művelet alaphalmazával jelöljük. Kicsit pontosabb ha az alaphalmaz mellett a testet is konkrétan specifikáljuk, de sokszor feltesszük, hogy a szövegkörnyezetből nyilvánvaló, hogy mely testre gondolunk. Hasonlóan pontosabb lenne a két műveletet is minden kijelölni, mikor egy vektortérre hivatkozunk, de ha világos, hogy mi a vektortérbeli elemek között az additív művelet, és hogy mit jelent egy test elemeivel szorozni, akkor elhagyjuk a műveletek kijelölését. A legfontosabb, – de persze a legkörülmenyezésből – jelölés az lenne, hogy pl. „tekintsünk egy $(V, +, \cdot)$ vektorteret az \mathbb{F} test fölött.” Ehelyett sokszor csak azt mondjuk, hogy „legyen V egy vektortér”. Ilyenkor a szövegkörnyezetből világosnak kell lennie, hogy mi a test, mit jelent a számmal való szorzás, és mi az összeadás a V halmazon.

2.1. Vektortér altérei

2.2. definíció (altér). Legyen $(V, +, \cdot)$ egy vektortér az \mathbb{F} test felett. Egy $M \subseteq V$ részhalmaz a V vektortér *altere*, ha M maga is vektorteret alkot a V -ben definiált additív műveettel, és a V -ben definiált számmal való szorzással.

2.3. állítás. Legyen V egy vektortér, és $M \subseteq V$ egy részhalmaza. Az M pontosan akkor altér, ha

1. $0 \in M$,
2. $u, v \in M$ esetén $u + v \in M$,
3. $u \in M$ és $\alpha \in \mathbb{F}$ esetén $\alpha u \in M$.

Tetszőleges V vektortérre a $\{0\}$ és maga V minden alterek, ezeket *triviális altereknek* is szokás mondani.

Most két fontos fogalmat vezetünk be. Egy halmazt tartalmazó legszűkebb altér fogalmát, és a halmaz lineáris burkának fogalmát. Ki fog derülni, hogy a lineáris burok minden egybeesik a legszűkebb altérrel.

2.4. definíció-állítás (generált altér). Egy vektortérben, akárhány altér közös része altér. Emiatt értelmes a következő definíció. Ha $H \subseteq V$ egy részhalmaza a V vektortérnek, akkor jelölje $\text{gen } H$ a H halmazt tartalmazó összes alterek metszetét. Ezt az alteret nevezzük a H halmaz által *generált altérnek*.

2.5. állítás. Legyen V egy vektortér. Ekkor

1. minden $H \subseteq V$ halmazra $H \subseteq \text{gen } H$,
2. ha $H \subseteq K \subseteq V$, akkor $\text{gen } H \subseteq \text{gen } K$,
3. minden $H \subseteq V$ mellett $\text{gen}(\text{gen } H) = \text{gen } H$.

A $\text{gen } H$ a H halmazt tartalmazó alterek között a legszűkebb. Így $H \subseteq V$ pontosan akkor altér, ha $\text{gen } H = H$.

2.6. definíció (lineáris kombináció, lineáris burok). Ha adott a V vektortérben véges sok v_1, \dots, v_r vektor, akkor a vektortér minden

$$\alpha_1 v_1 + \dots + \alpha_r v_r$$

alakú vektorát a v_1, \dots, v_r vektorok egy *lineáris kombinációjának* mondjuk.

Legyen $H \subseteq V$ egy tetszőleges halmaz. Ekkor $\text{lin } H$ jelöli H összes véges részhalmazának összes lineáris kombinációjának halmazát, azaz

$$\text{lin } H = \left\{ \sum_{j=1}^n \alpha_j v_j : n \in \mathbb{N}, v_1, \dots, v_n \in H, \alpha_1, \dots, \alpha_n \in \mathbb{F} \right\}$$

Definíció szerint nulla darab vektor lineáris kombinációja a vektortér zérus eleme, tehát $\text{lin } \emptyset = \{0\}$. A $\text{lin } H$ halmazt nevezzük a H halmaz *lineáris burkának*.

2.7. állítás. Egy V vektortér minden $H \subseteq V$ részhalmazának lineáris burka, a H halmazt tartalmazó legszűkebb altér, azaz

$$\text{lin } H = \text{gen } H.$$

Bizonyítás: Világos, hogy $H \subseteq \text{lin } H$, világos hogy $\text{lin } H$ egy altér, és az is nyilvánvaló, hogy ha $H \subseteq M$ egy tetszőleges altér, akkor $\text{lin } H \subseteq M$.

A $\text{lin } H$ egy a H -t tartalmazó altér, és $\text{gen } H$ az összes ilyen alterek metszete, ezért $\text{gen } H \subseteq \text{lin } H$. Másoldalról, a $\text{lin } H$ altér részhalmaza minden H -t tartalmazó altérnek, így azok közös részének is, ergo $\text{lin } H \subseteq \text{gen } H$. \square

Ha H egy véges halmaz, akkor $\text{lin } H$ kicsit egyszerűbben írható. Mint arról már a definícióban is szó volt $\text{lin } \emptyset = \{0\}$. Ha H egy elemű, akkor $\text{lin}(\{v_1\}) = \{\alpha v_1 : \alpha \in \mathbb{F}\}$. Ha $H = \{v_1, v_2\}$ két elemű, akkor $\text{lin}(v_1, v_2) = \{\alpha_1 v_1 + \alpha_2 v_2 : \alpha_1, \alpha_2 \in \mathbb{F}\}$. Hasonlóan, ha $H = \{v_1, \dots, v_r\}$ halmaz r elemből áll akkor elegendő csak az r elemből álló lineáris kombinációkat képezni, azaz

$$\text{lin}(\{v_1, \dots, v_r\}) = \left\{ \sum_{j=1}^r \alpha_j v_j : \alpha_1, \dots, \alpha_r \in \mathbb{F} \right\}.$$

2.8. definíció (generátorrendszer, végesen generált vektortér). Egy vektortér egy H részhalmazáról azt mondjuk, hogy generálja a vektorteret vagy, hogy H egy generátorrendszer V -nek, ha

$$\text{lin } H = V.$$

A V vektorteret végesen generáltnak mondunk, ha létezik véges generátorrendszer.

A generátorrendszer cseréről szóló 2.9. lemmának kiemelten fontos szerepe van felépítésünkben. Egyszerűen használjuk majd a Steinitz-lemma (3) igazolásában, másrészt ennek segítségével tisztázzuk majd azt a kérdést, hogy hogyan alakulnak egy vektor „koordinátái”, ha a vonatkoztatási rendszert változtatjuk.

2.9. lemma (generátorrendszer csere). Legyen $\{x_1, \dots, x_m\}$ egy generátorrendszer valamely vektortérnek, és tegyük fel, hogy valamely y vektorra

$$y = \sum_{j=1}^m \eta_j x_j,$$

ahol $\eta_k \neq 0$ valamely $1 \leq k \leq m$ mellett. Ekkor y becserélhető a k -adik helyen a generátorrendszerbe, úgy hogy az generátorrendszer maradjon, azaz a

$$\{x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_m\}$$

vektorrendszer is generátorrendszer.

Bizonyítás: Fejezzük ki x_k -t az y -ra felírt formulából: $x_k = \frac{1}{\eta_k} y + \sum_{\substack{j=1 \\ j \neq k}}^m \frac{-1}{\eta_k} \eta_j x_j$. Ha a eredetileg $a = \sum_{\substack{j=1 \\ j \neq k}}^m \alpha_j x_j$ alakú, akkor x_k helyére betéve, a fent kifejezett formulát és bevezetve a $\delta = \frac{\alpha_k}{\eta_k}$ jelölést, azt kapjuk hogy:

$$\begin{aligned} a &= \alpha_k x_k + \sum_{\substack{j=1 \\ j \neq k}}^m \alpha_j x_j = \\ &= \alpha_k \left(\frac{1}{\eta_k} y + \sum_{\substack{j=1 \\ j \neq k}}^m \frac{-1}{\eta_k} \eta_j x_j \right) + \sum_{\substack{j=1 \\ j \neq k}}^m \alpha_j x_j = \frac{\alpha_k}{\eta_k} y + \sum_{\substack{j=1 \\ j \neq k}}^m \left(\alpha_j - \frac{\alpha_k}{\eta_k} \eta_j \right) x_j = \\ &= \delta y + \sum_{\substack{j=1 \\ j \neq k}}^m (\alpha_j - \delta \eta_j) x_j. \end{aligned}$$

Azt kaptuk tehát, hogy ha egy vektor kifejezhető az eredeti vektorrendszerből az

$$(\alpha_1, \dots, \alpha_m)$$

együttetőkkel, akkor ugyanez a vektor a módosított vektorrendszerből is kifejezhető, még hozzá az

$$\left(\underbrace{\alpha_1 - \delta \eta_1}_1, \underbrace{\alpha_2 - \delta \eta_2}_2, \dots, \underbrace{\alpha_{k-1} - \delta \eta_{k-1}}_{k-1}, \underbrace{\delta}_k, \underbrace{\alpha_{k+1} - \delta \eta_{k+1}}_{k+1}, \dots, \underbrace{\alpha_m - \delta \eta_m}_m \right)$$

együttetőkkel. □

2.2. Elimináció

Szokásos jelölés és a kívánatos szemlélet kialakításában is fontos szerepet játszik a következő táblázat. Ha $\{x_1, \dots, x_m\}$ egy generátorrendszer az azt jelenti, hogy minden $a \in V$ vektor előáll mint az x_1, \dots, x_m vektorok valamelyen együtthatókkal vett lineáris kombinációja. Ha tehát $a = \alpha_1 x_1 + \dots + \alpha_m x_m$, akkor azt a következőképpen fejezzük ki.

	a
x_1	α_1
\vdots	\vdots
x_k	α_k
\vdots	\vdots
x_m	α_m

Tekinthető ez egy $m \times 1$ típusú bekeretezett mátrixnak, ahol a sorok címkéi a generátorrendszer elemei, az egyetlen oszlop címkéje pedig az a vektor, amelynek az előállításáról van szó.

A generátorrendszer csere lemma arról is szólt, hogy ha y -t a k -adik helyen cseréljük a generátorrendszerbe, akkor a fenti táblázat hogyan változik. Azt kaptuk, hogy a

$$\begin{array}{c|cc}
 \begin{array}{c|cc}
 & y & a \\
 \hline
 x_1 & \eta_1 & \alpha_1 \\
 \vdots & \vdots & \vdots \\
 x_k & \boxed{\eta_k} & \alpha_k \\
 \vdots & \vdots & \vdots \\
 x_m & \eta_m & \alpha_m \\
 \hline
 & \delta & \frac{\alpha_k}{\eta_k}
 \end{array} & \Rightarrow & \begin{array}{c|cc}
 \begin{array}{c|cc}
 & y & a \\
 \hline
 x_1 & 0 & \alpha_1 - \eta_1 \delta \\
 \vdots & \vdots & \vdots \\
 y & 1 & \delta \\
 \vdots & \vdots & \vdots \\
 x_m & 0 & \alpha_m - \eta_m \delta \\
 \hline
 & &
 \end{array} & (2.1)
 \end{array}
 \end{array}$$

transzformációt kell végrehajtani. Persze ugyanezt azt egyetlen a vektor helyett kiszámolhatjuk például az a, b, c vektorokra is. Továbbra is az y -t cseréljük a generátorrendszerbe, így a

$$\begin{array}{c|cccc}
 \begin{array}{c|cccc}
 & y & a & b & c \\
 \hline
 x_1 & \eta_1 & \alpha_1 & \beta_1 & \gamma_1 \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 x_k & \boxed{\eta_k} & \alpha_k & \beta_k & \gamma_k \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 x_m & \eta_m & \alpha_m & \beta_m & \gamma_m \\
 \hline
 & \delta & \frac{\alpha_k}{\eta_k} & \frac{\beta_k}{\eta_k} & \frac{\gamma_k}{\eta_k}
 \end{array} & \Rightarrow & \begin{array}{c|cccc}
 \begin{array}{c|cccc}
 & y & a & b & c \\
 \hline
 x_1 & 0 & \alpha_1 - \eta_1 \delta_a & \beta_1 - \eta_1 \delta_b & \gamma_1 - \eta_1 \delta_c \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 y & 1 & \delta_a & \delta_b & \delta_c \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 x_m & 0 & \alpha_m - \eta_m \delta_a & \beta_m - \eta_m \delta_b & \gamma_m - \eta_m \delta_c \\
 \hline
 & & & &
 \end{array} & (2.1)
 \end{array}
 \end{array}$$

a transzformációt hajtjuk végre. Most arra figyeljünk, hogy végül is a keretben lévő $m \times 4$ -es mátrixszal sorműveleteket hajtottunk végre:

- Eldöntöttük, hogy az y, a, b, c vektorok közül az y -t cseréljük be a generátorrendszerbe, mégpedig a k -adik helyen. Ezt jelezük az y oszlopa k -adik helyen lévő elemének bekeretezésével. A keretben csak nem zéró szám lehet.
- Első lépésként a keretezett számmal osztottuk a k -adik sort. Ez az új táblázat k -adik sora. Pusztán segítségképpen ugyanezt a sort mint egy számolási segédsort az első táblázat alá másoltuk.
- Az új táblázat első sora az eredeti első sornak és a segédsor η_1 -szeresének különbsége. A második sor az eredeti második sornak és a segédsor η_2 -szeresének különbsége. Hasonlóan, $m \neq k$ -ra az m -edik sor az eredeti m -edik sornak és a segédsor η_m -szeresének különbsége.

Látjuk tehát, hogy összesen két fajta sorművelettel alakítottuk a kiindulási mátrixot:

1. Egy sort szoroztunk egy nem zérus számmal,
2. Egy sorhoz hozzáadtuk egy másik sor számszorosát.

Világos, hogy ez ugyanaz, mintha a feladat az lett volna, hogy a fenti két sorművelet használva az y oszlopában minden számot el kell tüntetni, a k -adikat pedig 1-re kell beállítani. Ez a *Gauss – Jordan-elimináció*.

Homogén lineáris egyenletrendszer

Az alábbi feladatot lineáris egyenletrendszernek nevezzük.

$$\begin{array}{lclcl} a_{1,1}x_1 + \cdots + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n & = & b_2 \\ \vdots & & \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n & = & b_m \end{array}$$

Itt az $n, m \in \mathbb{N}$, az $a_{i,j} \in \mathbb{F}$ testbeli számok előre adottak minden $i = 1, \dots, m$ és minden $j = 1, \dots, n$ mellett. Adottak még az egyenletek jobboldalát képező $b_i \in \mathbb{F}$ számok. A feladat megoldása annyit tesz, hogy keresünk az x_1, \dots, x_n ismeretlenek összes olyan értékét az \mathbb{F} testből, amelyre fenti egyenletek minden teljesülnek. Ha a jobboldali számokra $b_i = 0$ minden $i = 1, \dots, n$ mellett, akkor *homogén lineáris egyenletrendszer*ről beszélünk, egyébként a rendszert *inhomogénnek* mondjuk. A fenti rendszer együtthatóiból álló

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

mátrixot *együttható-mátrixnak* mondjuk.

Ha az egyenletek egyikét egy nemzérus számmal szorozzuk, és az egyenletek egyikéhez hozzáadjuk egy másik egyenlet számszorosát, akkor a megoldások nem változnak, azaz ekvivalens átalakítást hajtunk végre.

Egy Gauss–Jordan-eliminációs lépésre tekinthetünk úgy is, mint rögzített i, j mellett a j -edik változó kiküszöbölésére – azaz eliminációjára – valamennyi nem az i -edik sorból, és az i -edik sorban e j -edik változó együtthatójának 1-re állítására. Valóban, ha az i -edik sort osztjuk az $a_{i,j} \neq 0$ számmal, akkor e sor j -edik eleme 1-re változik. Ha a k -adik ($k \neq i$) sorból kivonjuk az imént normált sor $a_{k,j}$ -szeresét, akkor az eredmény sor j -edik helyén zérust kapunk. Ha ezt minden $k = 1, \dots, m$ mellett kiszámoljuk, akkor éppen egy Gauss–Jordan-eliminációt hajtunk végre az $a_{i,j}$ pivot elem² választásával. Az eliminációs lépés hatására az j -edik változó az i -edik kivételével minden más sorból eltűnt, és az i -edik sorban pontosan 1 együtthatóval szerepel.

Ugyan ez egy táblázatban megfogalmazva. Itt a baloldali mátrix az együttható mátrix, amelynek a j -edik oszlopa van külön kiemelve. L_1, \dots, L_m jelöli az együttható-mátrix sorait A jobboldali mátrix az elimináció eredménye, a megfelelő sorműveletekkel az egyes sorok címkéiben:

$$\begin{array}{ccc|ccc|c} \cdots & a_{1,j} & \cdots & L_1 & \cdots & 0 & \cdots & L_1 - a_{1,j} \frac{1}{a_{i,j}} L_i \\ \vdots & \boxed{a_{i,j}} & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \cdots & a_{m,j} & \cdots & L_i & \cdots & 1 & \cdots & \frac{1}{a_{i,j}} L_i \\ \vdots & & & \vdots & & \vdots & & \vdots \\ & & & L_m & \cdots & 0 & \cdots & L_m - a_{m,j} \frac{1}{a_{i,j}} L_i \end{array}$$

Azt kell látnunk, hogy egy eliminációs lépés az $a_{i,j}$ pivot elem választásával pontosan ugyanazt eredményezi, mint az együttható-mátrix j -edik $[A]^j$ oszlopának becserélése a generátorrendszer i -edik helyére.

Az algoritmus célja, hogy a generátorrendszerbe annyi oszlopot cseréljünk be amennyit csak tudunk, de persze egy már korábban becserélt vektort nem cserélünk ki egy újabbra. Az algoritmus akkor áll meg, ha nem tudunk új oszlopot becserélni. Ez pontosan akkor fordul elő, ha a táblázat minden nem zérus sora egy már becserélt vektorhoz tartozik.

A lineáris egyenletrendszerek nyelvére átültetve ez azt jelenti, hogy az algoritmus célja, a változók eliminálása, lehetőség szerint minnél többet. Egy sorban (és egy oszlopban³) csak egy eliminált változó szerepelhet. Az algoritmus akkor áll meg, ha a táblázatnak minden nem zérus sorában már van eliminált változó.

Az algoritmus utolsó táblájából a homogén lineáris egyenletrendszer általános megoldása könnyen leolvasható: Dobjuk el a zérus sorokat, és foglalkozzunk a maradék táblázattal. Nevezzük az eliminált

²Értsd: sarok elem.

³Ez magától teljesül, erre nem kell figyelnünk.

változókat *kötöttnek* a többi változót *szabadnak*. minden sor egy és csak egy kötött, azaz eliminált, változót tartalmaz. A szabad változók, az az a többiek már ha vannak ilyenek egyáltalán, egy nem eliminált változóhoz tartoznak. Adjunk a szabad változóknak tetszőleges értéket, majd minden sorban ebből már egyértelműen kifejezhető a kötött változók értéke.

Összefoglalva Az eliminált változókat *kötött változónak* nevezzük, és a többi változót *szabad változónak* mondjuk. Persze az algoritmusban a kötött változók száma és a szabad változók száma minden az együttható mátrix oszlopainak számával azonos.

A következő tételek rendkívül fontos, még akkor is ha teljesen nyilvánvaló a most tárgyalt algoritmus alapján.

2.10. állítás. *Ha egy homogén lineáris egyenletrendszerben több ismeretlen van mint egyenlet, akkor a rendszernek van nem zéró megoldása.*

Bizonyítás: Gauss–Jordan-algoritmus legutolsó táblázatában a kötött változó száma legfeljebb a sorok száma az pedig szigorúan kisebb mint az összes ismeretlenek száma. Van tehát szabad változó, amelynek értéke tetszőleges lehet. \square

Mivel a változók vagy kötöttek vagy szabadok, ezért a kötött változók számának és a szabad változók számának az összege a változók száma. Ebben a pillanatban még nem látszik, de később ki fog derülni, hogy a kötött így a szabad változók száma is független az algoritmus során választott pivot elemektől. Érdemes a rendszer egy megoldásra úgy gondolni, mint egy $x \in \mathbb{F}^n$ oszlopvektorra, amelynek i -edik koordinátája az x_i változó aktuális értéke. Ilyen módon az x_1, \dots, x_n pontosan akkor elégít ki a lineáris egyenletrendszer, ha

$$Ax = b$$

mátrixegyenlet teljesül, ahol a $b \in \mathbb{F}^n$ az a vektor melynek i -edik koordinátája b_i .

Érdemes itt konkrét példákkal szemléltetni a homogén lineáris egyenletrendszer általános megoldásának felírását:

1. Oldjuk meg a

$$\begin{cases} x + 4y + 7z = 0 \\ 2x + 5y + 8z = 0 \\ 3x + 6y + 8z = 0 \end{cases}$$

homogén lineáris egyenletrendszer. A Gauss–Jordan-algoritmus lehet például a következő:

$$\begin{array}{c|ccc} a & b & c \\ \hline 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 8 \\ \hline \delta & 4 & 7 \end{array} \Rightarrow \begin{array}{c|ccc} a & b & c \\ \hline a & 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & -6 & -13 \\ \hline 0 & \delta & 2 \end{array} \Rightarrow \begin{array}{c|ccc} a & b & c \\ \hline a & 1 & 0 & -1 \\ b & 0 & 1 & 2 \\ 0 & 0 & -1 & \delta \\ \hline 0 & 0 & 0 & 0 \end{array} \Rightarrow \begin{array}{c|ccc} a & b & c \\ \hline a & 1 & 0 & 0 \\ b & 0 & 1 & 0 \\ c & 0 & 0 & 1 \end{array}$$

Ez azt jelenti, hogy az eredeti feladat ekvivalens az

$$\begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases}$$

feladattal, amelynek nyilvánvalóan csak a zéró vektor a megoldása.

2. Tekintsük a következő feladatot:

$$\begin{array}{rl} x_1 + 3x_2 + 4x_3 + 5x_4 - x_5 &= 0 \\ -2x_1 + x_2 - x_3 + 4x_4 - 5x_5 &= 0 \\ 2x_1 + x_2 + 3x_3 + 3x_5 &= 0 \\ 3x_1 + x_2 + 4x_3 - x_4 + 5x_5 &= 0 \end{array}$$

Az elimináció lehet a következő:

$$\begin{array}{c|ccccc}
 a_1 & a_2 & a_3 & a_4 & a_5 \\
 \hline
 1 & 3 & 4 & 5 & -1 \\
 -2 & 1 & -1 & 4 & -5 \\
 2 & 1 & 3 & 0 & 3 \\
 3 & 1 & 4 & -1 & 5 \\
 \hline
 \delta & 3 & 4 & 5 & -1
 \end{array} \Rightarrow
 \begin{array}{c|ccccc}
 a_1 & a_2 & a_3 & a_4 & a_5 \\
 \hline
 1 & 3 & 4 & 5 & -1 \\
 0 & 7 & 7 & 14 & -7 \\
 0 & -5 & -5 & -10 & 5 \\
 0 & -8 & -8 & -16 & 8 \\
 \hline
 0 & \delta & 1 & 2 & -1
 \end{array} \Rightarrow
 \begin{array}{c|ccccc}
 a_1 & a_2 & a_3 & a_4 & a_5 \\
 \hline
 1 & 0 & 1 & -1 & 2 \\
 a_2 & 0 & 1 & 2 & -1 \\
 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0
 \end{array}$$

Az eredeti egyenletrendszer tehát ekvivalens a következő egyenletrendszerrel:

$$\begin{aligned}
 x_1 + x_3 - x_4 + 2x_5 &= 0 \\
 x_2 + x_3 + 2x_4 - x_5 &= 0
 \end{aligned}$$

Itt x_1, x_2 a kötött változók és x_3, x_4, x_5 a szabad változók. Ezek értéke tetszőleges lehet, mondjuk $x_3 = s$, $x_4 = r$, $x_5 = t$ és ekkor $x_1 = -s + r - 2t$ valamint $x_2 = -s - 2r + t$. Az általános megoldás ezért

$$\left\{ \begin{pmatrix} -s + r - 2t \\ -s - 2r + t \\ s \\ r \\ t \end{pmatrix} : s, r, t \in \mathbb{R} \right\} = \left\{ s \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + r \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} : s, r, t \in \mathbb{R} \right\}.$$

A szabad változók száma tehát 3, és a rendszer megoldáshalmaza a

$$\begin{pmatrix} -1 & 1 & -2 \\ -1 & -2 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

mátrix oszlopai generált altér.

Egy homogén lineáris egyenletrendszernek a zérus vektor minden megoldása, ezt nevezzük *triviális megoldásnak*. A Gauss–Jordan-eliminációs algoritmusból világos a következő gondolat. A rendszernek pontosan akkor nincs nem triviális megoldása, ha nincs szabad változó, azaz minden változó kötött:

2.11. állítás. *Egy $Ax = 0$ homogén lineáris egyenletrendszernek pontosan akkor a triviális megoldás az egyetlen megoldása, ha az eliminációs algoritmusban minden oszlop a generátorrendszerbe cserélhető.*

Mátrixok bázisfaktorizációja 1.

Érdemes a Gauss–Jordan-eliminációt az egyenletrendszer megoldásától függetlenül is szemlélni. Ide másolom az előző feladat megoldásának táblázatát:

$$\begin{array}{c|ccccc}
 & a_1 & a_2 & a_3 & a_4 & a_5 \\
 \hline
 a_1 & 1 & 0 & 1 & -1 & 2 \\
 a_2 & 0 & 1 & 1 & 2 & -1 \\
 & 0 & 0 & 0 & 0 & 0 \\
 & 0 & 0 & 0 & 0 & 0
 \end{array}$$

Ez utolsó táblázat felfedi az eredeti a_1, a_2, a_3, a_4, a_5 vektorok közti kapcsolatot. A táblázat értelmezése szerint

$$\begin{aligned}
 a_3 &= a_1 + a_2 \\
 a_4 &= -a_1 + 2a_2 \\
 a_5 &= 2a_1 - a_2
 \end{aligned}$$

Ezt mátrixszorzásként interpretálva azt kapjuk, hogy

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

hiszen a baloldali mátrix minden oszlopa az a_1, a_2 oszlopok lineáris kombinációja.

2.12. definíció. Legyen $A \in \mathbb{F}^{m \times n}$ mátrix. Az A mátrix oszlop- (sor-) vektorterének nevezzük az A oszlopai (sorai) generálta alterét az \mathbb{F}^m (\mathbb{F}^n) vektortérnek.

Az előző példa szerint az ottani 4×5 méretű mátrix oszlopvektorterének az $\{a_1, a_2\}$ két elemű rendszer egy generátorrendszerére, hiszen ha W jelöli az oszlopok generálta alteret, azaz $W = \text{lin} \{a_1, a_2, a_3, a_4, a_5\}$, akkor

$$\{a_1, a_2, a_3, a_4, a_5\} \subseteq \text{lin} \{a_1, a_2\},$$

amiből a lineáris burkolatot képezve

$$W \subseteq \text{lin} \{a_1, a_2\} \subseteq W$$

következik, ergo $W = \text{lin} \{a_1, a_2\}$ valóban fennáll.

Nézzük újra a 40. oldalon a (+) matrix faktorizációt. Látható, hogy nem csak az oszlop vektortérnek, de a sorvektortérnek is találtunk egy két elemű generátorrendszerét: Jelölje most a_1, a_2, a_3, a_4 a (+) baloldalán lévő mátrix sorait. E mátrix legyen A . A B mátrix legyen a (+) jobboldalán az első mátrix és C a második. Persze $A = B \cdot C$, ami szerint A minden sora C sorainak lineáris kombinációja. Itt konkrétan

$$\begin{aligned} a_1 &= 1c_1 + 3c_2 \\ a_2 &= (-2)c_1 + 1c_2 \\ a_3 &= 2c_1 + 1c_2 \\ a_4 &= 3c_1 + (-1)c_2, \end{aligned}$$

ahol c_1, c_2 jelöli a C mátrix sorait. Tehát, ha most V jelöli az A mátrix sorvektorterétet, azaz $V = \text{lin} \{a_1, a_2, a_3, a_4\}$, akkor

$$\{a_1, a_2, a_3, a_4\} \subseteq \text{lin} \{c_1, c_2\},$$

amiből a lineáris burkolatot képezve

$$V \subseteq \text{lin} \{c_1, c_2\} \subseteq V$$

fennáll. Az utolsó tartalmazás most azért igaz, mert C mátrix az A utolsó eliminációs táblázatának nem zérus soraiból áll, de az elimináció minden egyes lépésében a sorok egy lináris kombinációját képezzük, ergo a táblázat minden egyes sora az eredeti sorok alkotta vektortérben – ami itt V – marad.

Vegyük észre, hogy a fenti eljárásnak azt gondoltuk meg, hogy tetszőleges $m \times n$ méretű mátrix sorvektorterének és oszlopvektorterének van azonos r elemszámú generátorrendszerére, ahol $1 \leq r \leq \min \{m, n\}$.

Az alábbi állítás a mátrix szorzás definíciójának következménye.

2.13. állítás. Legyen $A \in \mathbb{F}^{m \times n}$ egy nem zérus mátrix.

1. Tegyük fel, hogy A oszlopvektorterének adott egy r -elemű generátorrendszer. Ekkor a generátorrendszer r db. vektorát a B mátrix oszlopaiba rendezve egy $m \times r$ mátrixot kapunk. Ehhez a B mátrixhoz létezik C mátrix, amely $r \times n$ méretű és $A = B \cdot C$.
2. Most azt tegyük fel, hogy A sorvektorterében adott egy r -elemű generátorrendszer. Ekkor a generátorrendszer r db. sorát a C mátrix soraiba rendezve egy $r \times n$ mátrixot kapunk. Ehhez a C mátrixhoz létezik B mátrix, amely $m \times r$ méretű és $A = B \cdot C$.

Inhomogén lineáris egyenletrendszer

Láttuk, hogy egy m egyenletet és n ismeretlenetet tartalmazó lineáris egyenletrendszer ekvivalens az

$$Ax = b$$

feladattal, ahol $A \in \mathbb{F}^{m \times n}$ az együttható-mátrix és $b \in \mathbb{F}^m$ a jobboldalakból alkotott vektor. A következő egyszerű észrevétel szerint ha az inhomogén rendszernek találunk valahogyan egyetlen megoldását és ismerjük a homogén rendszer általános megoldását, akkor már az inhomogén rendszer általános megoldása is egyszerűen felírható.

2.14. állítás. A fenti jelölések megtartása mellett legyen $x_0 \in \mathbb{F}^m$ egy tetszőlegesen rögzített megoldása az $Ax = b$ inhomogén lineáris egyenletrendszernek. Ekkor

$$\{x : Ax = b\} = \{x_0 + z : Az = 0\},$$

azaz az inhomogén rendszer megoldáshalmaza, azonos a homogén rendszer megoldás halmazának egy partikuláris megoldással való eltoltjával.

Bizonyítás: Legyen először x egy megoldás, azaz $Ax = b$. Mivel x_0 is egy megoldás, ezért $Ax_0 = b$ is teljesül. Persze $x = x_0 + (x - x_0) = x_0 + z$, ahol $z = x - x_0$ egy olyan vektor, amelyre $Az = Ax - Ax_0 = b - b = 0$.

Másodsor tekintsünk egy $x = x_0 + z$ alakú vektort, ahol $Az = 0$. Ekkor persze $Ax = Ax_0 + Az = b + 0 = b$. \square

Egy nem nyilvánvaló következmény, hogy akármelyik partikuláris megoldással toljuk is el az inhomogén rendszer megoldását minden ugyanazt a halmazt kapjuk, emiatt mindegy melyik partikuláris megoldást rögzítettük.

Persze a kérdés, hogy hogyan találunk egyáltalán partikuláris megoldást. Máshogyan fogalmazva: mikor van egyáltalán megoldása egy inhomogén rendszernek?

2.15. állítás. Az $Ax = b$ inhomogén lineáris egyenletrendszernek pontosan akkor van megoldása, ha b előáll mint A oszlopainak valamelyen lineáris kombinációja, azaz a b vektor az A mátrix oszlopvektorteréhez tartozik.

Konkrétan adott A és b mellett Gauss–Jordan-algoritmussal el tudjuk dönteni, hogy b vektor eleme-e az A mátrix oszlopvektorterének. Egészítük ki az A mátrixot az utolsó oszlopában a b vektorral. Most hajtsuk végre a Gauss–Jordan-eliminációt, de úgy hogy pivot-elemet, csak az első n oszlopból válasszunk, ugyanúgy mintha csak a homogén rendszert oldanánk meg. Persze minden egyes eliminációs lépésben számoljuk ki az utolsó oszlopvektor elemeit is. Az algoritmus úgy ér véget, hogy az A mátrixnak megfelelő részben minden nem zérus sor tartalmaz eliminált változót.

Ha van olyan sor ahol az első n elem zérus, de az utolsó elem nem zérus, akkor nyilván nincs megoldás. minden egyéb esetben van megoldás, hiszen a csupa zérő sorokhoz tartozó egyenletek az ismeretlenek minden értéke mellett fennállnak, de maradék a nem zérus sorok esetén a kötött változót kifejezhetjük a szabad változók tetszőleges – mondjuk zérus – értéke mellett.

Az inhomogén rendszernek tehát pontosan akkor van megoldása, ha az utolsó oszlop minden nem zérus eleme olyan egyenlethez tartozik, ahol szerepel eliminált változó.

Ha a fenti eliminációra, mint a generátorrendszer csere lemma alkalmazására gondolunk, akkor azt kapjuk, hogy a megoldhatóság szükséges és elegendő feltétele, hogy a végső táblában az utolsó, a b -hez tartozó oszlopnak csak olyan helyeken lehetnek nem zérus tagjai, amely helyek korábban már a generátorrendszerbe becserélt oszlopvektorokhoz tartoznak, ami persze nem jelent többet, mint hogy b eleme az A oszlopvektorterének. Az algoritmus annyiban több, mint a korábban megértett feltétel, hogy a megoldható esetben rögtön kapunk egy partikuláris megoldást is.

Nézzünk egy példát.

$$\begin{array}{rcl} x_1 + 3x_2 + 4x_3 + 5x_4 - x_5 & = & 6 \\ -2x_1 + x_2 - x_3 + 4x_4 - 5x_5 & = & -5 \\ 2x_1 + x_2 + 3x_3 + 3x_5 & = & 7 \\ 3x_1 + x_2 + 4x_3 - x_4 + 5x_5 & = & 10 \end{array}$$

A Gauss–Jordan-algoritmus:

$$\begin{array}{c|ccccc|c} a_1 & a_2 & a_3 & a_4 & a_5 & b \\ \hline 1 & 3 & 4 & 5 & -1 & 6 \\ -2 & 1 & -1 & 4 & -5 & -5 \\ 2 & 1 & 3 & 0 & 3 & 7 \\ 3 & 1 & 4 & -1 & 5 & 10 \\ \hline \delta & 3 & 4 & 5 & -1 & 6 \end{array} \Rightarrow \begin{array}{c|ccccc|c} a_1 & a_2 & a_3 & a_4 & a_5 & b \\ \hline 1 & 3 & 4 & 5 & -1 & 6 \\ 0 & 7 & 7 & 14 & -7 & 7 \\ 0 & -5 & -5 & -10 & 5 & -5 \\ 0 & -8 & -8 & -16 & 8 & -8 \\ \hline 0 & -1 & -1 & -2 & \delta & -1 \end{array} \Rightarrow \begin{array}{c|ccccc|c} a_1 & a_2 & a_3 & a_4 & a_5 & b \\ \hline 1 & 2 & 3 & 3 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

A rendszernek tehát van megoldása. A legegyszerűbb ha a szabad változókat zérusra állítjuk, ergo $x_1 = 5$, $x_5 = -1$, $x_2 = 0$, $x_3 = 0$, $x_4 = 0$ egy partikuláris megoldás. Az általános megoldás tehát

$$\left\{ \begin{pmatrix} 5 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} + s \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + r \begin{pmatrix} -3 \\ 0 \\ 0 \\ 1 \\ 2 \end{pmatrix} : s, t, r \in \mathbb{R} \right\}.$$

2.16. állítás. Legyen $A \in \mathbb{F}^{m \times n}$ mátrix és $b \in \mathbb{F}^n$ egy előre megadott vektor. Az $Ax = b$ inhomogén lineáris egyenletrendszernek pontosan akkor van egy és csak egy megoldása, ha b az A képteréhez tartozik és A minden oszlopa a Gauss–Jordan-eliminációs algoritmussal a generátorrendszerbe cserélhető.

Ebben az esetben persze az algoritmusból adódó partikuláris megoldás az egyetlen megoldás.

Ha például a fenti harmadik eliminációs táblázatban a b oszloban a 2. elem nem zérus szám lenne, akkor a 0, 1, 0, 0 számokkal mint oszlopvektorral jelölt b jobboldal mellett az $Ax = b$ rendszernek nem lenne megoldása. Ez azért van, mert az utolsó táblázathoz tartozó homogén lineáris egyenletrendszer ekvivalens

az első táblázat rendszerével, és a $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ oszlopot helyettesítve az első táblázat b oszlopába, az elimináció

során a b oszlop nem változik, hiszen a 2. sorban nem választottunk pivot elemet, ergo a segédsorban b alatt minden lépésben zérus állna.

Ezt végig gondolhatjuk tetszőleges méretű mátrixra is, így kapjuk a következő állítást.

2.17. állítás. Legyen $A \in \mathbb{F}^{m \times n}$ mátrix. Hajtsunk végre Gauss–Jordan-eliminációt a mátrixon. Az utolsó táblázatnak pontosan akkor nincs csupa zérus elemből álló sora, ha minden $b \in \mathbb{F}^m$ vektorra az $Ax = b$ inhomogén lineáris egyenletrendszernek van megoldása.

Bizonyítás: Ha az elimináció utolsó táblázatában nincs zérus sor, akkor minden sorban van egyetlen eliminált változó. Akármi is tehát a jobboldal, ha a szabad változókat zérusra állítjuk akkor kapunk egy megoldást. Megfordítva, ha mondjuk a k -adik sor a csupa zérus elemből álló sor az utolsó táblázatban, akkor az a jobboldal, ahol egyedül a k -adik koordináta nem zérus, egy nem megoldható $Ax = b$ rendszert definiál. \square

Tudjuk, hogy az A mátrixra alkalmazott Gauss–Jordan-algoritmus akkor ér véget, amikor minden sorra teljesül, hogy az vagy a csupa zérus sor, vagy az pontosan egy kötött ismeretlen tartalmaz. Eszerint akkor és csak akkor nincs zérus sor, ha a kötött ismeretlenek száma azonos a sorok számával.

Ebből már adódik, hogy ha kevesebb ismeretlen mint sora van egy rendszernek, akkor minden megválasztható olyan jobboldali vektor, amellyel az inhomogén rendszer nem megoldható.

2.18. állítás. Legyen $A \in \mathbb{F}^{m \times r}$ olyan mátrix, ahol $r < m$. Ekkor létezik $b \in \mathbb{F}^m$ vektor, amellyel felírt $Ax = b$ inhomogén lineáris egyenletrendszernek már nincs megoldása.

Bizonyítás: Tegyük fel tehát, hogy kevesebb ismeretlen van mint sor. Ekkor persze a kötött változók száma is kisebb mint a sorok száma, ergo van zérus sor az utolsó eliminációs táblázatban. Ekkor a 2.17. állítás szerint, létezik olyan jobboldal, amivel képzett inhomogén rendszer már nem megoldható. \square

A Gauss–Jordan-elimináció utolsó táblázatára vonatkozó a 2.17. állítás, gyönyörű következménye a négyzetes mátrixok kétoldalisági tulajdonsága.

2.19. állítás. Tegyük fel, hogy $A, B \in \mathbb{F}^{m \times m}$ négyzetes mátrixok szorzata az identitás mátrix, azaz $AB = I$. Ekkor $BA = I$ is fennáll.

Bizonyítás: Ha $AB = I$, akkor nyilvánvalóan minden $b \in \mathbb{F}^m$ jobboldal mellett, az $Ax = b$ inhomogén lineáris egyenletrendszer van megoldása. Például $x = Bb$ egy megoldás. A 2.17. állítás szerint ez csak úgy lehet, ha az utolsó eliminációs táblázatban nincs csupa zérus sor. Ezek szerint az utolsó táblázatban a sorok száma megegyezik az eliminált változók számával, de négyzetes mátrixról beszélünk, tehát a oszlopok száma, azaz az $Ax = 0$ homogén rendszerben az összes változók száma is azonos a kötött változók számával. Ekkor persze nincs szabad változó, tehát az $Ax = 0$ homogén rendszernek csak a triviális megoldás a megoldása.

A mátrix szorzás definíciója szerint ez úgy is fogalmazható, hogy tetszőleges $C \in \mathbb{F}^{m \times n}$ mátrixra, az $AC = 0$ feltételből $C = 0$ következik. Na most

$$A(BA - I) = A(BA) - AI = (AB)A - AI = IA - AI = A - A = 0.$$

Meggondoltuk tehát, hogy $BA - I = 0$, ergo $BA = I$ is fennáll. \square

Inverz mátrix

Gondolunk arra, hogy milyen módszerrel automatizálhatnánk a Gauss–Jordan-algoritmust. Mivel sor-műveletekről van szó, természetes gondolat, hogy az $A \in \mathbb{F}^{m \times n}$ mátrix transzformálásához $m \times m$ méretű mátrixokat használunk, amelyekkel balról szorozzuk A -t. Valóban, ha $m \times m$ méretű identitás mátrix

1. i -edik sorát szorozzuk egy δ számmal, akkor egy olyan mátrixot kapunk, amellyel való balszorzása A -nak éppen az A mátrix i -edik sorát szorozza δ -val.
2. k -adik sorából kivonjuk az i -edik sor δ -szorosát, akkor egy olyan mátrixot kapunk, amellyel való balszorzása A -nak éppen az A mátrix k -adik sorából vonja ki az i -edik sor δ -szorosát.
3. k -adik és j -edik sorát felcseréljük, akkor egy olyan mátrixot kapunk, amellyel való balszorzása A -nak éppen az A mátrix k -adik és j -edik sorát cseréli fel.

Tekintsünk most, egy $m \times m$ méretű négyzetes mátrixot. Tegyük fel, hogy a mátrix minden oszlopát a generátorrendszerbe cserélhetjük a szokásos Gauss–Jordan-eliminációs algoritmus során. Láttuk korábban, hogy ez azt jelenti, hogy az $Ax = b$ inhomogén lineáris egyenletrendszer minden b vektor mellett egyértelműen megoldható.

Legyen $\{e_1, \dots, e_m\}$ az \mathbb{F}^m szokásos generátorrendszeré, azaz e_k -nak éppen a k -adik koordinátája 1, a többi zérus. Ha $B \in \mathbb{F}^m$ az a mátrix, amelynek k -edik oszlopa az $Ax = e_k$ inhomogén lineáris egyenletrendszer egyetlen $x_k \in \mathbb{F}^m$ megoldása, akkor a mátrix szorzás definíciója szerint

$$AB = I$$

teljesül. A B mátrixot tehát m darab m lépéses Gauss–Jordan-eliminációval meg tudnánk határozni.

Egyszerűbb, ha egyetlen eliminációs algoritmus használunk, de arra az $m \times 2m$ méretű mátrixra, amelyet úgy kapunk, hogy az a A mátrix jobboldalához illesztjük az $I \in \mathbb{F}^{m \times m}$ identitás mátrixot. Az elimináció során, A minden sorát a generátorrendszerbe cseréljük, de persze a jobbra illesztett identitás mátrix sorait is transzformáljuk. Feltevésünk szerint m lépés után áll le az algoritmus. Ekkor egy olyan táblázat van előttünk, amelynek első m oszlopa egy olyan négyzetes mátrixot alkot, amelynek minden oszlopa és minden sora egyetlen 1-est tartalmaz a többi elem zérus. Vegyük észre, hogy egy ilyen mátrix a sorok felcserélésével az identitás mátrixszá transzformálható.

Az $Ax = e_k$ egyenlet megoldása a jobboldali $m \times m$ mátrix k -adik oszlopából olvasható le. Ha például az első sor j -edik elemén van 1-es, akkor az azt jelenti, hogy az első sorba elimináltuk az j -edik ismeretlenetet, tehát a megoldás j -edik koordinátáját tartalmazza az $m + k$ -adik oszlop első eleme.

Ha tehát úgy cseréljük fel a táblázat sorait, hogy a baloldali $m \times m$ méretű mátrix váljon az identitás mátrixszá, akkor az $Ax = e_k$ egyenlet egyetlen partikuláris megoldása jelentkezik az egész táblázat $m + k$ adik oszlopában. Mivel a keresett B mátrixnak éppen ez a k -adik oszlopa, ezért maga a B mátrix áll a táblázat jobboldali $m \times m$ méretű részén.

Most meggondoljuk, hogy $BA = I$ is fennáll. Az iménti algoritmussal az $[A|I]$ mátrixot transzformáltuk a $[I|B]$ mátrixszá. A transzformáció során a Gauss–Jordan-eliminációt használtuk, amely a szakasz elején említett 1) és 2) típusú mátrixokkal mint balszorzásokkal hajtható végre. Az algoritmus végén még sorokat is felcserélünk, ami egy 3) típusú balszorzást jelent. Ha az 1) vagy 2) vagy 3) típusú mátrixot *elemi mátrixnak* nevezünk, akkor úgy fogalmazhatunk, hogy van véges sok $P_1, \dots, P_r \in \mathbb{F}^{m \times m}$ elemi mátrix, amelyre $(P_r \cdots P_1)[A|I] = [I|B]$. Ekkor persze $(P_r \cdots P_1)A = I$ és $P_r \cdots P_1 = B$, amiből már következik, hogy

$$BA = I.$$

2.20. definíció (nem szinguláris mátrix). Egy $A \in \mathbb{F}^{m \times m}$ négyzetes mátrixot *invertálhatónak* vagy *regulárisnak* vagy *nem szingulárisnak* nevezünk, ha létezik $B \in \mathbb{F}^{m \times m}$ négyzetes mátrix, amelyre

$$AB = BA = I$$

Mivel egy egységelemes gyűrűben ha van inverz, akkor csak egyetlen egy van, ezért adott A invertálható mátrixhoz csak egy B mátrix van, amely kielégíti a fenti definíciót. Ezt a B mátrixot nevezzük az A inverzének és $A^{-1} = B$ módon jelöljük.

2.21. állítás. Legyen $A \in \mathbb{F}^{m \times m}$ mátrix. Az alábbi feltevések ekvivalensek:

1. A invertálható.
2. Létezik $B \in \mathbb{F}^{m \times m}$ mátrix, amelyre $AB = I$.
3. Gauss–Jordan-eliminációval az A minden oszlopa a generátorrendszerbe cserélhető.

Ha a fenti feltevések egyike (ergo mindegyike) fennáll, akkor az 2)-ben szereplő B mátrixból csak egy van, mégpedig az A^{-1} inverz mátrix.

Bizonyítás: Körben igazolunk:

1. \Rightarrow 2. Nyilvánvaló.

2. \Rightarrow 3. Tegyük fel – indirekt –, hogy már r transzformáció után a Gauss–Jordan-algoritmus megáll, ahol $r < m$. Ez azt jelenti, hogy az A oszlop-vektorterének van r elemű generátorrendszer. E generátorrendszer vektorait mint oszlopokat egy A_1 mátrixba téve egy $m \times r$ mátrixot kapunk. Mivel az oszlopok lineáris burkában az A valamennyi oszlopa szerepel, van olyan $r \times m$ méretű A_2 mátrix, amelyre $A_1 A_2 = A$. Azt kaptuk tehát, hogy

$$A_1 (A_2 B) = (A_1 A_2) B = AB = I.$$

Ebből az következik, hogy tetszőleges $b \in \mathbb{F}^m$ vektor mellett az $x = A_2 B b \in \mathbb{F}^r$ vektor megoldása az $A_1 x = b$ inhomogén lineáris egyenletrendszernek, ami ellentmondás, hiszen A_1 mátrixnak kevesebb oszlopa van mint sora.

3. \Rightarrow 1. Éppen ezt igazoltuk a szakasz elején.

Ha a feltevések fennállnak, akkor

$$A^{-1} = A^{-1}I = A^{-1}(AB) = (A^{-1}A)B = IB = B. \quad \square$$

Ha tehát A és B négyzetes mátrixok, amelyekre $AB = I$, akkor 2) szerint A invertálható, és $A^{-1} = B$, ezért $BA = A^{-1}A = I$ is fennáll. Ahogyan azt a mátrixok bevezetésekor a 25. oldalon megígértük, most megmutattuk azt, hogy a négyzetes mátrixok gyűrűjében ha két mátrix szorzata a gyűrű egységeleme, akkor ezek a mátrixok kommutálnak.

Egy konkrét példa megoldásával ismételjük át a szakasz elején megértett algoritmust. A feladat, hogy

keressük meg az $A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 1 & 2 \\ 2 & 3 & 1 & 0 \\ 1 & 0 & 2 & 1 \end{pmatrix}$ mátrix inverzét! Az inverz létezésének szükséges és elegendő feltétele,

hogy 4 lépést tudunk végrehajtani az eliminációs algoritmusban. Egy lehetséges megoldás a következő:

$$\begin{array}{c|cccc|cccc|cccc}
 \boxed{1} & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
 0 & 3 & 1 & 2 & 0 & 1 & 0 & 0 & 0 & 3 & 1 & 2 & 0 & 1 & 0 & 0 \\
 2 & 3 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -1 & 0 & -2 & 0 & 1 & 0 \\
 1 & 0 & 2 & 1 & 0 & 0 & 0 & 1 & 0 & -1 & 1 & \boxed{1} & -1 & 0 & 0 & 1 \\
 \hline
 \delta & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & \delta & -1 & 0 & 0 & 1
 \end{array}
 \Rightarrow
 \begin{array}{c|cccc|cccc|cccc}
 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 6 & 0 & 0 & 3 & 1 & 0 & -2 \\
 0 & 5 & \boxed{-1} & 0 & 2 & 1 & 0 & -2 & 0 & -5 & 1 & 0 & -2 & -1 & 0 & 2 \\
 0 & 1 & -1 & 0 & -2 & 0 & 1 & 0 & 0 & -4 & 0 & 0 & -4 & -1 & 1 & 2 \\
 0 & -1 & 1 & 1 & -1 & 0 & 0 & 1 & 0 & 4 & 0 & 1 & 1 & 1 & 0 & -1 \\
 \hline
 0 & -5 & \delta & 0 & -2 & -1 & 0 & 2 & 0 & \delta & 0 & 0 & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2}
 \end{array}
 \Rightarrow
 \begin{array}{c|cccc|cccc|cccc}
 1 & 0 & 0 & 0 & -3 & -\frac{1}{2} & \frac{3}{2} & 1 & 1 & 0 & 0 & 0 & -3 & -\frac{1}{2} & \frac{3}{2} & 1 \\
 0 & 0 & 1 & 0 & 3 & \frac{1}{4} & -\frac{5}{4} & -\frac{1}{2} & 0 & 1 & 0 & 0 & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} \\
 0 & 1 & 0 & 0 & 1 & \frac{1}{4} & -\frac{4}{4} & -\frac{1}{2} & 0 & 0 & 1 & 0 & 3 & \frac{1}{4} & -\frac{4}{4} & -\frac{1}{2} \\
 0 & 0 & 0 & 1 & -3 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & -3 & 0 & 1 & 1
 \end{array}$$

Kaptuk hát, hogy az A mátrix inverze az $A^{-1} = \frac{1}{4} \begin{pmatrix} -12 & -2 & 6 & 4 \\ 4 & 1 & -1 & -2 \\ 12 & 1 & -5 & -2 \\ -12 & 0 & 4 & 4 \end{pmatrix}$ mátrix.

Picit kevesebb helyet foglal az algoritmus, ha lusták vagyunk a generátorrendszerbe már becserélt oszlopok kiírására. Ekkor érdemes kiírni a sorok és oszlopok címkéjét, nehogy eltévedjünk. Az előző feladat így alakul:

$$\begin{array}{c}
 \begin{array}{c|ccccc}
 & a_1 & a_2 & a_3 & a_4 & e_1 & e_2 & e_3 & e_4 \\
 \begin{array}{c} e_1 \\ e_2 \\ e_3 \\ e_4 \end{array} & \boxed{1} & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
 & 0 & 3 & 1 & 2 & 0 & 1 & 0 & 0 \\
 & 2 & 3 & 1 & 0 & 0 & 0 & 1 & 0 \\
 & 1 & 0 & 2 & 1 & 0 & 0 & 0 & 1 \end{array} & \delta & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
 \hline
 & \delta & 1 & 1 & 0 & 1 & 0 & 0 & 0
 \end{array} \Rightarrow \begin{array}{c|ccccc}
 & a_2 & a_3 & a_4 & e_1 & e_2 & e_3 & e_4 \\
 \begin{array}{c} a_1 \\ e_2 \\ e_3 \\ e_4 \end{array} & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
 & 3 & 1 & 2 & 0 & 1 & 0 & 0 \\
 & 1 & -1 & 0 & -2 & 0 & 1 & 0 \\
 & -1 & 1 & \boxed{1} & -1 & 0 & 0 & 1 \end{array} & -1 & 1 & \delta & -1 & 0 & 0 & 1 \\
 \hline
 & -1 & 1 & \delta & -1 & 0 & 0 & 1
 \end{array} \\
 \begin{array}{c}
 \begin{array}{c|ccccc}
 & a_2 & a_3 & e_1 & e_2 & e_3 & e_4 \\
 \begin{array}{c} a_1 \\ e_2 \\ e_3 \\ e_4 \end{array} & 1 & 1 & 1 & 0 & 0 & 0 \\
 & 5 & \boxed{-1} & 2 & 1 & 0 & -2 \\
 & 1 & -1 & -2 & 0 & 1 & 0 \\
 & -1 & 1 & -1 & 0 & 0 & 1 \end{array} & -5 & \delta & -2 & 0 & 2 & -1 \\
 \hline
 & -5 & \delta & -2 & -1 & 0 & 2
 \end{array} \Rightarrow \begin{array}{c|ccccc}
 & a_2 & e_1 & e_2 & e_3 & e_4 \\
 \begin{array}{c} a_1 \\ a_3 \\ e_3 \\ a_4 \end{array} & 6 & 3 & 1 & 0 & -2 \\
 & -5 & -2 & -1 & 0 & 2 \\
 & -4 & -4 & -1 & 1 & 2 \\
 & 4 & 1 & 1 & 0 & -1 \end{array} & \delta & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} \\
 \hline
 & \delta & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2}
 \end{array} \\
 \begin{array}{c}
 \begin{array}{c|ccccc}
 & e_1 & e_2 & e_3 & e_4 \\
 \begin{array}{c} a_1 \\ a_3 \\ a_2 \\ a_4 \end{array} & -3 & -\frac{1}{2} & \frac{3}{2} & 1 & 1 \\
 & 3 & \frac{1}{4} & -\frac{5}{4} & -\frac{1}{2} & \frac{1}{4} \\
 & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} & -\frac{1}{4} \\
 & -3 & 0 & 1 & 1 & 0 \end{array} & -3 & -\frac{1}{2} & \frac{3}{2} & 1 \\
 \hline
 & -3 & 0 & 1 & 1 & 0
 \end{array} \Rightarrow \begin{array}{c|ccccc}
 & e_1 & e_2 & e_3 & e_4 \\
 \begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \end{array} & -3 & -\frac{1}{2} & \frac{3}{2} & 1 & -3 \\
 & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \\
 & 3 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} & -\frac{1}{4} \\
 & -3 & 0 & 1 & 1 & 0 \end{array} & -3 & 0 & 1 & 1
 \end{array}$$

2.3. Lineárisan független rendszerek

2.22. definíció (lineárisan összefüggő vektorrendszer). Egy véges $\{y_1, \dots, y_n\}$ vektorrendszeret *lineárisan összefüggőnek* mondunk, ha van olyan vektor, amely kifejezhető a többi vektor lineáris kombinációjaként.

Úgy is fogalmazhatnánk, hogy az $\{y_1, \dots, y_n\}$ rendszer pontosan akkor lineárisan összefüggő, ha létezik $1 \leq k \leq n$ index, amelyre

$$y_k \in \text{lin} \{y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_n\}.$$

Nyilvánvaló példa, hogy ha a vektorrendszer a 0 vektort tartalmazza, akkor lineárisan összefüggő.

Hasonlóan, ha ugyanaz a vektor többször is szerepel a vektorrendszerben, akkor a rendszer szintén lineárisan összefüggő.

Fontos emlékeznünk, hogy 0 db. vektor lineáris kombinációja megegyezés szerint a vektortér zérus eleme. Így ha a vektorrendszer egyetlen elemű, például $\{v\}$, akkor a v -n kívüli vektorok rendszere az üres rendszer, így a v -n kívüli elemek lineáris kombinációja egyedül a zérus vektor. Ezek szerint $\{v\}$ pontosan akkor lineárisan összefüggő, ha $v = 0$. Olyan vektorrendszerre tehát, amelyre nem igaz a lineárisan összefüggőség definíciója egy egyedüli nem zérus vektor, de akár az $\{\}$ üres rendszer is.

2.23. állítás. Legyen $\{y_1, \dots, y_n\}$ vektorrendszer rögzítve a V vektortérben. A vektorrendszerre tett alábbi feltevések egymással ekvivalensek.

1. Lineárisan összefüggő;
2. Van olyan elem a vektortérben, amely nem csak egyféleképpen áll elő mint az y_1, \dots, y_n vektorok lineáris kombinációja,
azaz: létezik $z \in V$, amelyre $z = \sum_{j=1}^n \xi_j y_j$ és $z = \sum_{j=1}^n \eta_j y_j$ és létezik $1 \leq k \leq n$, amelyre $\xi_k \neq \eta_k$.
3. Vannak olyan nem mind zérus $\alpha_1, \dots, \alpha_n$ skalárok, amelyekkel

$$\sum_{j=1}^n \alpha_j y_j = 0,$$

azaz a vektorrendszernek van nem triviális lineáris kombinációja, amely a zérus vektort eredményezi.

Bizonyítás: Körben bizonyítunk.

1. \Rightarrow 2. Tegyük fel, hogy $y_k = \sum_{j=1}^{k-1} \eta_j y_j + \sum_{j=k+1}^n \eta_j y_j$. Ekkor az alábbi együttható rendszerek

$$(\eta_1, \dots, \eta_{k-1}, 0, \eta_{k+1}, \dots, \eta_n) \quad (0, \dots, 0, 1, 0, \dots, 0)$$

a k -adik helyen biztosan különböznek, hiszen $0 \neq 1$, és mind a két együttható rendszerrel képzett lineáris kombináció ugyanazt az y_k vektort eredményezi.

2. \Rightarrow 3. Világos, hogy

$$0 = z - z = \sum_{j=1}^n (\xi_j - \eta_j) y_j$$

és a k -adik skalár nem zérus.

3. \Rightarrow 1. Tegyük fel most, hogy $\sum_{j=1}^n \alpha_j y_j = 0$, és, hogy $\alpha_k \neq 0$. Ekkor

$$y_k = \sum_{\substack{j=1 \\ j \neq k}}^n -\frac{1}{\alpha_k} \alpha_j y_j$$

azaz a k -adik vektor tekinthető mint a többi vektor valamely lineáris kombinációja.

Ezt kellett belátni. □

Fontos észrevétel a következő.

2.24. állítás. *Minden lineárisan összefüggő vektorrendszeret tartalmazó vektorrendszer maga is lineárisan összefüggő.*

2.25. definíció. Egy nem véges vektorrendszeret *lineárisan összefüggőnek* nevezünk, ha van véges részrendszer, amely lineárisan összefüggő.

Persze továbbra is igaz, hogy egy összefüggő rendszert tartalmazó rendszer is összefüggő marad. A lineárisan összefüggőség a lineáris burok operációval is megfogható:

2.26. állítás. *Egy vektortér H vektorrendszer pontosan akkor lineárisan összefüggő, ha létezik $y \in H$, amelyre $y \in \text{lin}(H \setminus \{y\})$.*

Bizonyítás: Ha H lineárisan összefüggő, akkor van $\{x_1, \dots, x_k\}$ véges összefüggő vektorrendszer. Itt az egyik vektor előáll mint a többi lineáris kombinációja. Ha ezt a vektort y jelöli, akkor $y \in \text{lin}(H \setminus \{y\})$. Megfordítva, ha valamely $y \in H$ mellett van $\{x_1, \dots, x_k\} \subseteq H \setminus \{y\}$ vektorrendszer, amelynek egy lineáris kombinációja éppen y , akkor az $\{x_1, \dots, x_k, y\}$ véges rendszer egy lineárisan összefüggő rendszere H -nak, ergo H valóban lineárisan összefüggő. □

2.27. definíció (lineárisan független vektorrendszer). Egy vektorrendszer *lineárisan független*, ha nem lineárisan összefüggő.

Az előző állítás tagadásával azonnal meg is kapjuk a lineárisan független rendszerek jellemzését a lineáris burok operáció segítségével.

2.28. állítás. *Egy vektortér H vektorrendszer pontosan akkor lineárisan független, ha minden $y \in H$ mellett $y \notin \text{lin}(H \setminus \{y\})$ is teljesül.*

Fontos észrevétel a következő:

2.29. állítás. *Egy lineárisan független rendszer minden részrendszerre is lineáris független marad.*

A definíció szerint egy nem véges vektorrendszer akkor lineárisan független, ha minden véges részrendszer is az. Egy véges vektorrendszer lineárisan függetlenségét, pedig a következő egymással ekvivalens állítások karakterizálják.

2.30. állítás. *Legyen $\{y_1, \dots, y_n\}$ vektorrendszer rögzítve a V vektortérben. A vektorrendszerre tett alábbi feltevések egymással ekvivalensek.*

1. Lineárisan független;
2. A vektorrendszer lineáris burkában minden elem egyetlen egyféléképpen áll elő, mint az y_1, \dots, y_n vektorok lineáris kombinációja.
3. Az y_1, \dots, y_n vektoroknak csak a triviális lineáris kombinációja zérus, azaz

$$\sum_{j=1}^n \alpha_j y_j = 0 \text{ esetén } \alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

Bizonyítás: Nyilvánvaló a lineáris összefüggés karakterizációjából. \square

Egy a zéró vektort tartalmazó vektorrendszer persze lineárisan összefüggő, és egy nem zérus vektorból álló egyelemű vektorrendszer lineárisan független. A következő állítás sokszor teszi kényelmessé a gondolatmenetünket.

2.31. állítás. Legyen $\{y_1, \dots, y_n\}$ egy olyan legalább két elemű vektorrendszer, amelynek első eleme nem a zérus vektor, tehát $y_1 \neq 0$. A vektorrendszer pontosan akkor lineárisan összefüggő, ha létezik olyan eleme, amely pusztán az előző elemek lineárisan kombinációja.

Formálisabban: akkor és csak akkor, ha $\exists k \quad 2 \leq k \leq n : y_k \in \text{lin} \{y_1, \dots, y_{k-1}\}$

Bizonyítás: Tegyük fel, hogy a vektorrendszer lineárisan összefüggő. Ekkor van olyan a zérus vektort eredményező lineáris kombinációja $\alpha_1 y_1 + \dots + \alpha_n y_n = 0$, ahol nem az összes együttható nulla. Legyen k a lineáris kombinációban a legnagyobb nem nulla együttható indexe. Világos, hogy $k \neq 1$, hiszen $y_1 \neq 0$. Persze a k feletti együtthatók mind nullák, emiatt

$$\alpha_1 y_1 + \dots + \alpha_k y_k = 0.$$

Itt már $\alpha_k \neq 0$, tehát y_k kifejezhető az előző vektorok segítségével:

$$y_k = \frac{-1}{\alpha_k} \alpha_1 y_1 + \dots + \frac{-1}{\alpha_k} \alpha_{k-1} y_{k-1}.$$

\square

A lehető legsűkebb lineárisan független rendszer az $\{\}$ üres vektorrendszer, amelynek egyetlen eleme sincs. A lehető legbővebb generátorrendszer az egész vektortér. Ennél sokkal érdekesebb és fontosabb a lineárisan független rendszerek közül a lehető legbővebbet keresni, és generátorrendszerök közül a lehető legsűkebbet keresni.

2.32. definíció (maximális lineárisan független– és minimális generátorrendszer). Egy lineárisan független rendszert *maximális lineárisan független rendszernek* nevezünk, ha nem lehet bővíteni úgy, hogy lineárisan független maradjon.

Egy generátorrendszer *minimális generátorrendszernek* mondunk, ha nem lehet szűkíteni úgy, hogy generátorrendszer maradjon.

2.33. állítás. Legyen H egy (nem feltétlen véges) vektorrendszer a V vektortérben.

1. Tegyük fel, hogy $y \notin H$ és H lineárisan független.
A $H \cup \{y\}$ pontosan akkor lineárisan összefüggő, ha $y \in \text{lin } H$.
2. Tegyük fel, hogy $y \in H$ és H generátorrendszer.
A $H \setminus \{y\}$ pontosan akkor generátorrendszer, ha $y \in \text{lin } (H \setminus \{y\})$.

Bizonyítás: Voltaképpen négy különböző állítást kell igazolnunk:

1. \Leftarrow Ha $y \in \text{lin } H$, akkor H -nak van $\{x_1, \dots, x_k\}$ véges részrendszerére, amelyre az $\{x_1, \dots, x_k, y\}$ vektorrendszer y elemére igaz, hogy a többi elem lineáris kombinációja, ergo $H \cup \{y\}$ lineárisan összefüggő.
1. \Rightarrow Ha $H \cup \{y\}$ lineárisan összefüggő, akkor van $H \cup \{y\}$ -nak véges részrendszerére, amely lineárisan összefüggő. Ebben a rendszerben y -nak szerepelnie kell hiszen H lineárisan független. Azt kapjuk tehát, hogy bizonyos $x_1, \dots, x_k \in H$ vektorok mellett az $\{x_1, \dots, x_k, y\}$ vektorrendszer lineáris összefüggő. No de, H lineárisan független, ezért csak az utolsó vektor, az y vektor lehet az előző vektorok lineáris burkában.

2. \Rightarrow Ha $H \setminus \{y\}$ generátorrendszer, akkor minden vektor így az y is eleme a $\text{lin}(H \setminus \{y\})$ altérnek.
 2. \Leftarrow Ha $y \in \text{lin}(H \setminus \{y\})$, akkor $H \subseteq \text{lin}(H \setminus \{y\})$. Emiatt $\text{lin}H \subseteq \text{lin}(\text{lin}(H \setminus \{y\})) \subseteq \text{lin}(H \setminus \{y\})$. Ebből persze $\text{lin}(H \setminus \{y\}) = \text{lin}H$ következik. No de, H generátorrendszer, ergo $H \setminus \{y\}$ is az. \square

A előző állítás 1. pontját úgy érdemes értelmezni, hogy a V vektortér egy $H \subseteq V$ lineáris független rendszere akkor és csak akkor nem bővíthető a lineárisan függetlenség megtartásával, ha $\text{lin}H = V$, azaz H egy generátorrendszer.

A 2. pont is hasonló: A H generátorrendszer egyetlen eleme sem elhagyható a generátorrendszer tulajdonság megtartásával akkor és csak akkor, ha minden $y \in H$ mellett $y \notin \text{lin}(H \setminus \{y\})$, azaz H egy lineárisan független rendszer.

Bebizonyítottuk tehát az alábbi fontos állítást.

2.34. állítás. *Egy vektortér egy vektorrendszerére az alábbi feltevések ekvivalensek.*

1. *A vektorrendszer maximális lineárisan független rendszer.*
2. *A vektorrendszer egyszerre lineárisan független és generátorrendszer.*
3. *A vektorrendszer minimális generátorrendszer.*

2.35. lemma (független rendszer csere). *Legyen $\{y_1, \dots, y_n\}$ egy lineárisan független rendszere valamely vektortérnek, és tegyük fel, hogy a tér valamely x vektorára*

$$x = \sum_{j=1}^n \xi_j y_j,$$

ahol $\xi_k \neq 0$ az egyik $1 \leq k \leq n$ mellett. Ekkor x becserélhető a k -adik helyen a független rendszerbe úgy, hogy az független maradjon, azaz az

$$\{y_1, \dots, y_{k-1}, x, y_{k+1}, \dots, y_n\}$$

vektorrendszer is lineárisan független.

Bizonyítás: Megmutatjuk, hogy a cserélt rendszernek egyedül a triviális lineáris kombinációja zérus. Legyen tehát

$$0 = \sum_{\substack{j=1 \\ j \neq k}}^n \eta_j y_j + \eta_k x = \sum_{\substack{j=1 \\ j \neq k}}^n \eta_j y_j + \sum_{j=1}^n \eta_k \xi_j y_j = \sum_{\substack{j=1 \\ j \neq k}}^n (\eta_j + \eta_k \xi_j) y_j + \eta_k \xi_k y_k.$$

Mivel az eredeti rendszer lineárisan független, ezért az utóbbi lineáris kombináció minden együtthatójá a test null eleme. A k -adikkal kezdve $\eta_k \xi_k = 0$, $\xi_k \neq 0$, így $\eta_k = 0$. Ezt a nem k -adik együtthatókba visszahelyettesítve már azt kapjuk, hogy $\eta_j = 0$ minden $j \neq k$ mellett is. \square

Mátrixok bázisfaktorizációja 2.

A szakaszban bevezetett fogalmak illusztrációjaként érdemes visszatérni a 40. oldalra, ahol láttuk hogyan találunk generátorrendszeret egy mátrix oszlopvektorterében és sorvektorterében. A kényelmes olvasás okán, ide másolom az eliminációt, annyi változtatással, hogy most kiírjuk az eredeti generátorrendszer e_1, e_2, e_3, e_4 elemeit is. Persze itt $e_j \in \mathbb{R}^4$, melynek egyedül a j -edik koordinátája 1, a többi 0.

$$\begin{array}{c|ccccc}
 & a_1 & a_2 & a_3 & a_4 & a_5 \\
 \hline
 e_1 & 1 & 3 & 4 & 5 & -1 \\
 e_2 & -2 & 1 & -1 & 4 & -5 \\
 e_3 & 2 & 1 & 3 & 0 & 3 \\
 e_4 & 3 & 1 & 4 & -1 & 5 \\
 \hline
 & \delta & 3 & 4 & 5 & -1
 \end{array} \Rightarrow \begin{array}{c|ccccc}
 & a_1 & a_2 & a_3 & a_4 & a_5 \\
 \hline
 a_1 & 1 & 3 & 4 & 5 & -1 \\
 e_2 & 0 & 7 & 7 & 14 & -7 \\
 e_3 & 0 & -5 & -5 & -10 & 5 \\
 e_4 & 0 & -8 & -8 & -16 & 8 \\
 \hline
 & 0 & \delta & 1 & 2 & -1
 \end{array} \Rightarrow \begin{array}{c|ccccc}
 & a_1 & a_2 & a_3 & a_4 & a_5 \\
 \hline
 a_1 & 1 & 0 & 1 & -1 & 2 \\
 a_2 & 0 & 1 & 1 & 2 & -1 \\
 e_3 & 0 & 0 & 0 & 0 & 0 \\
 e_4 & 0 & 0 & 0 & 0 & 0
 \end{array}$$

Világos, hogy a kezdtő $\{e_1, e_2, e_3, e_4\}$ generátorrendszer egyben lineárisan független rendszer is. Az első lépésben a_1 állt be e_1 helyére, majd a második lépésben az a_2 vektort cseréltük be az e_2 helyére. A független rendszer cseréről szóló 2.35. lemmát alkalmazva azt kapjuk, hogy az utolsó táblázat $\{a_1, a_2, e_3, e_4\}$ rendszere is lineárisan független. No de, lineárisan független rendszer részrendszere is az, így az $\{a_1, a_2\}$ oszlopok az oszlopvektortér egy független rendszerét adják.

Ahogy azt korábban is láttuk – lásd a 40. oldalt – $\{a_1, a_2\}$ egy generátorrendszerére is az oszlopvektortérnek, így már láttuk is, hogy $\{a_1, a_2\}$ az oszlopvektortér egy két elemű lineárisan független generátorrendszerét alkotja.

Emlékezzünk, hogy az utolsó táblázat nem zérus sorai a sorvektortér egy generátorrendszerét szolgáltatja. Ha az a_j oszlopot az i -edik sorban cseréljük be, akkor a végső táblázat j -edik oszlopának i -edik sorában van 1-es egyébként a j -edik oszlop többi eleme zérus. Ebből látszik, hogy a végső táblázat nem zérus sorai a sorvektortér egy lineárisan független generátorrendszerét alkotják.

Összefoglalásként az eliminációból leolvasható az

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

bázisfaktORIZÁció, ahol a jobboldali első mátrix oszlopai a baloldali A mátrix oszlopvektorterének lineárisan független generátorrendszerét alkotja, és a jobboldali második mátrix sorai az A mátrix sorvektorterének lineárisan független generátorrendszerét alkotja.

A fenti bázisfaktORIZÁció tetszőleges mátrix mellett is végrehajtható, így a Gauss–Jordan-eliminációnak és a két cserélési lemmának – 2.35 és 2.9 – az alábbi szép következményét kapjuk:

2.36. állítás. *Tetszőleges $m \times n$ méretű nem zérus mátrix sorvektorterének és oszlopvektorterének is van azonos elemszámú lineárisan független, generátorrendszer. Ha r e két vektorrendszer közös elemszáma, akkor $1 \leq r \leq \min\{m, n\}$.*

Ha az itt elemzett mátrixra gondolunk, akkor felvethető, hogy mi történik, ha az eliminációs algoritmus során a fentől eltérő pivot elemeket választunk? Világos, hogy a pivot elemek más-más választása az eredeti mátrix más-más bázisfaktORIZÁCIÓJÁHOZ vezet. A bázisfaktORIZÁCIÓ tehát messze nem egyértelmű. A probléma, hogy ha adott néhány vektor, akkor mi a kapcsolat a belőlük kiválasztható maximális lineárisan független rendszerek között. Igaz-e, hogy a vektorokból kiválasztható maximális lineárisan független rendszerek – tehát a vektorok lineáris burkában mint altérben maximális lineárisan független rendszereknek – az elemszáma minden azonos? Vagy előfordulhat például, hogy egy vektortérnek van egy két elemű és egy három elemű maximális lineárisan független rendszere? Ami evvel azonos: előfordulhat-e, hogy egy vektortérben van egy 2 elemű és egy 3 elemű lineárisan független, generátorrendszer?

3. fejezet

A Steinitz-lemma

FÜGGETLEN- ÉS GENERÁTORRENDSZEREK elemszáma közti kapcsolat vezet a bázis és a dimenzió fogalmához. Az derül ki, hogy a lineárisan független, generátorrendszerök elemszáma egy az adott vektortérre jellemző szám, azaz bármely lineárisan független, generátorrendszernek ugyanannyi eleme van.

A félév legfajsúlyosabb állítása következik, az egyszerű bizonyítása ellenére. A pontos megfogalmazás előtt emlékezzünk a generátorrendszer cseréjére vonatkozó a 2.9. lemmára.

Steinitz-lemma. *Legyenek n és m nem negatív egészek. Tegyük fel, hogy az $\{y_1, \dots, y_n\}$ egy lineárisan független rendszer, és az $\{x_1, \dots, x_m\}$ egy generátorrendszer. Ekkor*

1. $n \leq m$ és;
2. az x_1, \dots, x_m vektorok alkalmaz átindexelésével kapott $\{y_1, \dots, y_n, x_{n+1}, \dots, x_m\}$ vektorrendszer is generátorrendszer.¹

Bizonyítás: Legyen k a legnagyobb a $\{0, \dots, n\}$ egészek közül, amelyre

1. $k \leq m$, és
2. az x_1, \dots, x_m vektorok alkalmaz átindexelésével az

$$\{y_1, \dots, y_k, x_{k+1}, \dots, x_m\} \tag{†}$$

vektorrendszer is generátorrendszer.

Ilyen k biztosan van, hiszen $k = 0$ triviálisan jó. Összesen azt kell meggondolnunk, hogy $k = n$. Ha $k < n$ lenne,

- akkor létezne y_{k+1} vektor. No de, ez az y_{k+1} nem szerepel az $\{y_1, \dots, y_k\}$ lineáris burkában, ami (†) generátorrendszer volta miatt csak úgy lehetséges, hogy $k \neq m$, azaz $k < m$, ergo $k + 1 \leq m$.
- A 2.9. lemma szerint a (†) vektorrendszerben az y_{k+1} vektor avval az x -el — a generátorrendszer tulajdonság megtartásával is — kicsérélhető, amely x szerepel az y_{k+1} vektornak a (†)-beli vektorokkal képzett lineáris kombinációjában.

Ez ellentmondás, hiszen k a legnagyobb olyan szám, amelyre a bizonyítás elején szereplő 1. és 2. feltételek egyszerre állnak fenn. \square

¹Úgy kell a jelöléseket érteni, hogy az $n = 0$, de az $n = m$ eset is lehetséges. Az $n = 0$ esetben az y -okkal jelölt vektorok egyike sem, míg az $n = m$ esetben az x -el jelölt vektorok egyike sem szerepel az $\{y_1, \dots, y_n, x_{n+1}, \dots, x_m\}$ vektorrendszer elemei között.

3.1. Rang-tétel

3.1. definíció (mátrix feszítőrangja). Legyen $A \in \mathbb{F}^{n \times m}$ egy tetszőleges nem zérus mátrix. Azt mondjuk, hogy *feszítőrangja* r , ha r a legkisebb olyan pozitív egész, amelyre A előáll

$$A = BC$$

alakban, ahol $B \in \mathbb{F}^{n \times r}$ és $C \in \mathbb{F}^{r \times m}$.

Világos, hogy tetszőleges nemzérus négyzetes mátrixra ez jól definiált és $1 \leq r \leq \min\{n, m\}$. A Rang-tételnek a szokásosnál egy kicsit erősebb formája következik.

3.2. állítás (Rang-tétel). *Minden nemzérus mátrixban a maximális lineárisan független oszloprendszerek és a maximális lineárisan független sorrendserek azonos elemszámúak, és ez a szám azonos a mátrix feszítőrangjával.*

Bizonyítás: Jelölje r az $A \in \mathbb{F}^{n \times m}$ mátrix feszítőrangját. Legyen r_c a mátrix egyik rögzített maximális lineárisan független oszloprendszerének elemszáma.

- Ezen oszlopokat egy $B \in \mathbb{F}^{n \times r_c}$ mátrixba téve – a maximalitás miatt – létezik olyan $C \in \mathbb{F}^{r_c \times m}$ mátrix, amelyre $A = BC$, azaz $r \leq r_c$.
- Most tekintsünk egy tetszőleges olyan $A = BC$ felbontást, ahol $B \in \mathbb{F}^{n \times r}$ és $C \in \mathbb{F}^{r \times m}$. Jelölje W a B mátrix oszlopai lineáris burkát. Az A mátrix fent rögzített maximális lineárisan független oszloprendszere egy lineárisan független rendszer a W vektortérben, és B oszlopai pedig egy generátorrendszer ugyanebben a vektortérben, így a Steinitz-lemma szerint $r_c \leq r$.

Evvel megmutattuk, hogy bármely két maximális lineárisan független oszloprendszer azonos elemszámú, és számuk megegyezik a mátrix feszítőrangjával.

Legyen r_w az A mátrix egyik rögzített maximális lineárisan független sorrendszerének elemszáma.

- Ezen sorokat egy $C \in \mathbb{F}^{r_w \times m}$ mátrixba téve – a maximalitás miatt – létezik olyan $B \in \mathbb{F}^{n \times r_w}$ mátrix, amelyre $A = BC$, azaz $r \leq r_w$.
- Most tekintsünk egy tetszőleges olyan $A = BC$ felbontást, ahol $B \in \mathbb{F}^{n \times r}$ és $C \in \mathbb{F}^{r \times m}$. Jelölje most V a C mátrix sorai lineáris burkát. Az A mátrix fent rögzített maximális lineárisan független sorrendszerének egy lineárisan független rendszer a V vektortérben, és C sorai pedig egy generátorrendszer alkotnak ugyanebben a V vektortérben, így a Steinitz-lemma szerint $r_w \leq r$.

Evvel azt is megmutattuk, hogy bármely két maximális lineárisan független sorrendszer azonos elemszámú, és számuk megegyezik a mátrix feszítőrangjával. \square

Ha $A \in \mathbb{F}^{n \times m}$ mátrix, amelynek feszítőrangja r , akkor A bármelyik maximális lineárisan független oszloprendszerének r az elemszáma. Ezeket az oszlopokat egy $B \in \mathbb{F}^{n \times r}$ mátrixba rendezve az A oszlopvektortének egy generátorrendszerét kapjuk, így létezik $C \in \mathbb{F}^{r \times m}$ mátrix, amelyre $A = BC$ teljesül. Mivel két mátrix szorzata diádok összegeként is felírható, úgy azt kaptuk, hogy egy r feszítőrangú mátrix felírható mint r darab diád, tehát 1 rangú mátrix, összegeként. Itt a diádokat alkotó oszlop rendszer és sorrendszer is lineárisan független. A következő gondolat szerint több is igaz.

3.3. állítás. *Tegyük fel, hogy az $A \in \mathbb{F}^{n \times m}$ nem zérus mátrix, amelynek feszítőrangja r előáll*

$$A = \sum_{j=1}^r b_j \cdot c_j$$

r darab diád összegének alakjában, ahol $b_j \in \mathbb{F}^{n \times 1}$ oszlopvektorok és $c_j \in \mathbb{F}^{1 \times m}$ sorvektorok. Ekkor a diádok oszlopaiból illetve soraiból alkotott

$$\{b_j : j = 1, \dots, r\} \text{ és } \{c_j : j = 1, \dots, r\}$$

rendszer lineárisan független rendszereket alkotnak.

Bizonyítás: Legyen a B mátrix j -edik oszlopa b_j és a C mátrix j -edik sora c_j . Világos, hogy $B \in \mathbb{F}^{n \times r}$, $C \in \mathbb{F}^{r \times m}$ és $A = BC$. Ha B oszlopai nem alkotnának lineárisan független rendszert, akkor lenne $B = B_1 B_2$ előállítás, ahol $B_1 \in \mathbb{F}^{n \times s}$, $B_2 \in \mathbb{F}^{s \times r}$, ahol $s < r$. Ekkor persze

$$A = BC = (B_1 B_2) C = B_1 (B_2 C)$$

Itt B_1 mátrixnak s oszlopa van, a $B_2 C$ mátrixnak s sora van, ami ellentmond annak, hogy A feszítőrangja r . A sorrendszer lineárisan függetlensége is hasonlóan adódik. \square

3.2. Dimenzió

A Steinitz-lemma következményeképpen:

3.4. következmény. Egy vektortérben bármely két véges, lineárisan független, generátorrendszer elemszáma azonos. Konkrétabban, ha

$$\{x_1, \dots, x_m\} \text{ és } \{y_1, \dots, y_n\}$$

lineárisan független generátorrendszerek, akkor $n = m$.

3.5. definíció (végesen generált vektortér). Egy vektorteret végesen generáltnak nevezünk, ha létezik véges elemszámú generátorrendszer.

Teljesen világos, hogy ha van egy vektortérben véges generátorrendszer, akkor van minimális generátorrendszer is, azaz van a térben lineárisan független generátorrendszer. Ezt rögzítjük a következőkben.

3.6. állítás. Minden végesen generált vektortérnek van olyan vektorrendszer, amely egyszerre lineárisan független és generátorrendszer.

Bizonyítás: Tekintsünk egy véges generátorrendszert. Világos, hogy ennek van minimális generátorrendszere, hiszen ha egyik eleme sem hagyható el, úgy hogy generátorrendszer maradjon, akkor minimális generátorrendszerrel állunk szemben. Véges sok elemből indulunk ki, így véges sok elem esetleges eldobása után egy minimális generátorrendszert kapunk.

Láttuk korábban (3.4. állítás), hogy egy generátorrendszer minimalitása éppen a rendszer lineárisan függetlenségét jelenti. \square

A lineárisan független generátorrendszerek olyan sűrűn fordulnak elő a tárgyalásban, hogy rövidebb külön nevet adni nekik.

3.7. definíció (bázis). Egy vektorrendszert bázisnak nevezünk, ha ez egyszerre lineárisan független és generátorrendszer.

A 3.6. állítást tehát úgy fogalmazhatjuk, hogy végesen generált vektortérnek van bázisa, és hasonlóan a 3.4. következmény pedig azt jelenti, hogy egy vektortérben bármely két bázis azonos elemszámú. Ez utóbbi tény ad értelmet a következő definíciónak:

3.8. állítás. Egy végesen generált vektortérrel azt mondjuk, hogy n dimenziós, vagy n a dimenzió száma, ha a vektortérben van n elemű bázis.

A tényt, hogy a V vektortér n dimenziós $\dim(V) = n$ módon jelöljük.

Fontos látni, hogy éppen azt gondoltuk meg, hogy minden végesen generált vektortérben van bázis,² és bármely két bázis pontosan annyi vektorból áll mint a tér dimenziója. A végesen generált vektortereket sokszor szinonimaként véges dimenziósnak is mondjuk.

A következő állítás az eddigiek összefoglalása:

3.9. állítás. Tekintsünk egy m -dimenziós vektorteret, és abban egy m -elemű $\{x_1, \dots, x_m\}$ vektorrendszert. E vektorrendszerre tett alábbi feltevések ekvivalensek.

1. Lineárisan független;
2. Maximális lineárisan független rendszer;
3. Generátorrendszer;
4. Minimális generátorrendszer;
5. Bázis.

Bizonyítás: Az első négy feltétel ekvivalenciájával kezdünk.

1. \Rightarrow 2. Mivel a tér m -dimenziós, ezért van m -elemű generátorrendszer, így a Steinitz-lemma szerint nincs m -nél több elemet tartalmazó lineárisan független rendszere, ergo bármely m elemet tartalmazó lineárisan független rendszer maximális is.

²Ez nem végesen generált vektorterekre is igaz, de itt nem igazoljuk, viszont később sem használjuk.

2.⇒3. Láttuk korábban.

3.⇒4. Mivel a tér m -dimenziós, ezért van m -elemű lineárisan független rendszere, így a Steinitz-lemma szerint nincs m -nél kevesebb elemet tartalmazó generátorrendszer, ergo bármely m elemet tartalmazó generátorrendszer rendszer minimális is.

4.⇒1. Láttuk korábban.

Az első négy feltétel tehát ugyanazt jelenti. Így ha 1.-et feltesszük, akkor 3. is fennáll, ami azt jelenti, hogy 1. feltétel és 5. feltétel is ekvivalensek. \square

A Steinitz-lemma kulcs szerepet játszott dimenzió fogalmának megértésében, hiszen a bázis elemszáma nem lehetne a tér dimenziója, anélkül hogy tudnánk a tényt: bármely két bázis azonos elemszámú! Márpedig egy vektortérben nagyon sok bázis van. A Steinitz-lemma 2. pontja segít ennek megértéséhez.

3.10. állítás. *Egy végesen generált vektortér bármely lineárisan független rendszere kiegészíthető bázissá.*

Bizonyítás: Tegyük fel, hogy a tér m dimenziós, ami azt jelenti, hogy van

$$\{x_1, \dots, x_m\}$$

m elemű lineárisan független generátorrendszer. Legyen $\{y_1, \dots, y_n\}$ egy lineárisan független rendszer. A Steinitz-lemma szerint ez a rendszer kiterjeszhető egy

$$\{y_1, \dots, y_n, x_{n+1}, \dots, x_m\}$$

generátorrendszerre, ami persze bázis is. Ezt kellett belátni. \square

Meggondoltuk tehát, hogy bármely véges generátorrendszerből elhagyható néhány elem úgy, hogy a rendszer lineárisan független generátorrendszerre váljon, és hasonlóan bármely lineárisan független rendszerhez, hozzáehető néhány elem úgy, hogy a rendszer lineárisan független generátorrendszerre váljon.

A független rendszerek cseréjéről (2.35.) és a generátor rendszerek cseréjéről (2.9.) szóló lemmák együttes alkalmazásaként kapjuk a következő lemmát.

3.11. lemma (bázis csere). *Legyen $\{x_1, \dots, x_m\}$ egy bázisa valamely vektortérnek, és tegyük fel, hogy egy y vektorra*

$$y = \sum_{j=1}^m \eta_j x_j,$$

ahol $\eta_k \neq 0$ valamely $1 \leq k \leq m$ mellett. Ekkor y becserélhető a k -adik helyen a bázisba, úgy hogy az is bázis maradjon, azaz az

$$\{x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_m\}$$

vektorrendszer is egy bázis.

3.12. állítás. *Egy végesen generált V vektortér minden $M \subseteq V$ altérre is végesen generált. Ekkor $\dim(M) \leq \dim(V)$ és egyenlőség csak $V = M$ esetben lehetséges.*

Bizonyítás: Mivel V -ben nincs tetszőleges nagy lineárisan független rendszer, ezért M -ben van véges elemszámú maximális lineárisan független rendszer. Erről tudjuk, hogy egyben generátorrendszer is.

Van tehát V -ben is M -ben is bázis. Mivel az M -beli bázis egyben lineárisan független rendszer, ezért a Steinitz-lemma szerint elemszáma nem nagyobb mint a rögzített bázis mint generátorrendszer elemszáma, azaz $\dim(M) \leq \dim(V)$.

Ha $\dim(M) = \dim(V) = n$, akkor M az altér egy $\{x_1, \dots, x_n\}$ bázisa olyan lineárisan független rendszer V -ben, amelynek annyi eleme van, mint a V dimenziója. Így ez egy V -beli generátorrendszer is, ezért $M = \text{lin}(\{x_1, \dots, x_n\}) = V$. \square

Rang és feszítőrang

A 3.2. állítást, azaz a Rang-tételt a dimenzió fogalmának ismeretében úgy is fogalmazhatjuk, hogy *tetszőleges nem zérus mátrixra az oszlopvektortér dimenziója egybeesik a feszítőranggal, és a sorvektortér dimenziója is azonos a feszítőranggal.*

3.13. definíció (mátrix rangja). Egy mátrix *rangja* az oszlopai által generált vektortér dimenziója.

A Rang-tétel tehát azt állítja, hogy nem zérus mátrixra a feszítőrang, a rang, azaz az oszlopvektortér dimenziószáma, és a sorvektortér dimenziószáma azonosak.

A zérus mátrixra a feszítőrang fogalmát nem definiáltuk, de az nyilvánvaló, hogy ebben az esetben az oszlopvektortér és a sorvektortér is 0 dimenziós.

Most felejtük el egy pillanatra, hogy igazoltuk már a Rang-tételt. A feszítőrang fogalma nélkül, de a dimenzió fogalma ismeretében gondoljuk át a Rang-tétel klasszikus alakját.

3.14. állítás. *Tetszőleges mátrixra az oszlopvektortér és a sorvektortér azonos dimenziós alterek.*

Bizonyítás: A Gauss–Jordan-elimináció algoritmusának segítségével $A = B \cdot C$ a mátrix bázisfaktorizációja, ahol B oszlopainak rendszere a mátrix oszlopvektorterének egy bázisa, és C sorainak rendszere a sorvektortér egy bázisa. Mivel itt B oszlopainak száma azonos C sorainak számával, ezért az oszlopvektortér és a sorvektortér azonos dimenziós vektorterek.³ □

Persze ebből az alakból is látszik, hogy az oszlopok közül bárhogyan is választok ki egy maximális lineárisan független rendszert e vektorrendszerek elemszáma azonos lesz. Más szavakkal: A Gauss–Jordan-elimináció során, akárhogyan is választok pivot elemeket, az utolsó táblázatban a nem csupa zérus sorok száma mindenkor azonos, mégpedig egyenlő a sorvektortér dimenziójának számával, azaz a mátrix rangjával.

³Érdemes felfigyelni rá, hogy milyen döbbenetesen szimpla gondolat, de persze a mélyén a dimenzió fogalma, azaz a Steinitz-lemma rejlik.

4. fejezet

Koordinátázás

A KOORDINÁTA-TÉR FOGALMÁT vezetjük be. Látni fogjuk, hogy véges dimenziós vektortérre „lényegében” az egyetlen példa, az \mathbb{F}^n tér.

4.1. Lineáris operátor fogalma

4.1. definíció (lineáris operátor). Legyenek V, W ugyanazon \mathbb{F} test feletti vektorterek. Az $A : V \rightarrow W$ függvényt *lineáris operációt* mondjuk, ha az

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$$

azonosság teljesül minden $x, y \in V$ és minden $\alpha, \beta \in \mathbb{F}$ mellett.

Ha $V = W$, akkor a *lineáris transzformáció* kifejezést is használjuk, ha pedig $W = \mathbb{F}$, akkor szokásos még a *lineáris funkcionál* szó összetétel is.

A $V \rightarrow W$ összes lineáris operációk halmazát $L(V, W)$ módon jelöljük. A $V = W$ esetben a rövidség kedvéért csak $L(V)$ -t írunk.

Láttuk korábban, hogy ha A egy $m \times n$ méretű mátrix, akkor $x \in \mathbb{F}^n$ oszlopvektor mellett az

$$x \mapsto A \cdot x,$$

mátrix szorzás egy $\mathbb{F}^n \rightarrow \mathbb{F}^m$ lineáris operációt definiál. Hasonlóan, ha most $x \in \mathbb{F}^m$ sorvektor jelöl akkor az

$$x \mapsto x \cdot A$$

szorzás egy $\mathbb{F}^m \rightarrow \mathbb{F}^n$ lineáris operációt jelent.

4.2. definíció-állítás. Legyen $A \in L(V, W)$ egy lineáris operáció. Ennek értékkészlete egy altér, amelyet $\text{Im } A$ módon jelölünk. Azon pontok halmaza, amelyeket A operátor a W tér zérus elemébe képez egy alteret alkotnak, amelyet $\ker A$ módon jelölünk.

Az $L(V, W)$ lineáris operációk halmaza a szokásos függvény műveletekkel vektorteret alkot.

A jelölések tehát:

$$\text{Im } A = \{y \in W : \text{létézik } x \in V, A(x) = y\}, \quad \ker A = \{x \in V : A(x) = 0\}.$$

4.3. állítás. Egy lineáris leképezés pontosan akkor injektív, ha $\ker A = \{0\}$. Egy injektív lineáris leképezés inverz függvénye egy $\text{Im } A \rightarrow V$ lineáris operátor. Az $A \in L(V, W)$ injektív lineáris leképezés inverzét A^{-1} módon jelöljük.

Lineáris leképezések kompozíciója is lineáris.

Tekinthetünk az A mátrixra mint az $x \mapsto A \cdot x$ lineáris operátorra. Ennek inverze az $x \mapsto A^{-1} \cdot x$ operátor, ahol A^{-1} az inverz mátrixot jelöli, hiszen $Ax = y$ pontosan akkor teljesül, ha $x = A^{-1}y$.

Itt jegyezem meg, hogy a tradíciókat követve egy A lineáris operáció esetén az x vektor $A(x)$ képét a zárójeleket elhagyva Ax módon írjuk. Hasonlóan az $A \circ B$ kompozíciót is AB módon jelöljük.

4.4. állítás. Legyen $A \in L(V, W)$ egy lineáris operáció. Ekkor

1. ha A injektív és $\{y_1, \dots, y_m\} \subseteq V$ lineárisan független, akkor $\{Ay_1, \dots, Ay_m\}$ is lineárisan független.
2. ha A szürjektív és $\{x_1, \dots, x_n\} \subseteq V$ generátorrendszer, akkor $\{Ax_1, \dots, Ax_n\}$ is generátorrendszer W -nek.
3. ha A bijekció és $\{e_1, \dots, e_n\} \subseteq V$ bázis, akkor $\{Ae_1, \dots, Ae_n\}$ is bázis W -ben.

4.5. definíció (izomorf vektorterek). Egy $A : V \rightarrow W$ függvényt *izomorfizmusnak* mondunk, ha A lineáris bijekció. Az ugyanazon test feletti V, W vektortereket egymással *izomorf vektortérnek* mondjuk, ha létezik $A : V \rightarrow W$ izomorfizmus.

Gondoljuk meg, hogy az azonos test feletti vektorterek izomorfizmusa egy ekvivalencia reláció, azaz (1) minden vektortér izomorf saját magával; (2) ha V izomorf W -vel, akkor W is V -vel; (3) ha V_1 és V_2 izomorfak, továbbá ha V_2 és V_3 izomorfak, akkor V_1 és V_3 is izomorfak.

4.6. állítás. minden az \mathbb{F} test feletti véges dimenziós V vektortér izomorf az $\mathbb{F}^{\dim(V)}$ koordináta-térrel.

Bizonyítás: Rögzítsünk egy $\{e_1, \dots, e_n\}$ bázist V -ben. Definiálja $\Psi : \mathbb{F}^n \rightarrow V$ a következő függvényt:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto \sum_{j=1}^n \alpha_j e_j$$

Könnyű számolással ellenőrizhető, hogy Ψ egy lineáris operáció, a bázis lineárisan függetlenségét használva adódik A injektivitása, a szürjektivitás pedig a bázis generátorrendszer tulajdonságát használva igazolható. \square

4.7. definíció (koordináta). Legyen $\{e_1, \dots, e_n\}$ egy bázisa a V vektortérnek. Tudjuk, hogy minden $v \in V$ vektor egyetlen egyfeleképpen, de előáll mint a bázisvektorok egy lineáris kombinációja. Ha ez $v = \sum_{j=1}^n \alpha_j e_j$,

akkor az $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{F}^n$ vektort az x vektor $\{e_1, \dots, e_n\}$ bázisban felírt koordináta-vektorának nevezzük.

Pont azt mutattuk meg, hogy egy rögzített bázis mellett az a leképezés, amely egy vektorhoz hozzárendeli a bázisban felírt koordinátáit, egy izomorfizmus a vektortér és a koordináta-tere közt. Az eddigiek összefoglalásaként kapjuk az alábbi állítást.

4.8. állítás. Legyenek V és W ugyanazon test feletti végesen generált vektorterek. E két vektortér pontosan akkor izomorf, ha azonos dimenziósak.

Bizonyítás: Ha V izomorf W -vel, akkor létezik köztük izomorfizmus. No de, izomorfizmus bázist bázisra visz, ami azt jelenti, hogy azonos elemszámú bázisa van minden téren.

Ha $\dim(V) = \dim(W) = n$, akkor minden vektortér izomorf \mathbb{F}^n vektortérrel, ergo egymással is izomorfak. \square

Fontos látni, hogy a vektor koordináta-vektora függ a bázis megválasztásától. Akár ha csak a bázisban az elemek sorrendjét megváltoztatjuk, már akkor is változik a vektor koordináta-vektora. A 2.9., a 2.35. és a 3.11. lemmák összefoglalásaként kapjuk, hogy milyen módon változnak egy rögzített vektor koordinátái, ha a bázisban egy elemet megváltoztatunk.

4.9. lemma (bázistranszformáció). Legyen $\{x_1, \dots, x_m\}$ egy bázisa valamely vektortérnek, és tegyük fel, hogy egy y olyan vektor $y = \sum_{j=1}^m \eta_j x_j$, ahol $\eta_k \neq 0$ valamely $1 \leq k \leq m$ mellett. Láttuk y becserélhető a k -adik helyen a bázisba, tehát az $\{x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_m\}$ vektortrendszer is bázis marad. Tegyük fel továbbá, hogy $a \in V$ vektor koordinátái ismertek az eredeti bázisban, $a = \sum_{j=1}^m \alpha_j x_j$. Ekkor ugyanez az a vektor az új bázisban kifejezve $a = \sum_{j=1, j \neq k}^m (\alpha_j - \eta_j \delta) x_j + \delta y_k$, ahol $\delta = \frac{\alpha_k}{\eta_k}$.

Egy táblázatban összefoglalva:

$$\begin{array}{c|cc}
 & y & a \\
 \hline
 x_1 & \eta_1 & \alpha_1 \\
 \vdots & \vdots & \vdots \\
 x_k & \boxed{\eta_k} & \alpha_k \\
 \vdots & \vdots & \vdots \\
 x_m & \eta_m & \alpha_m \\
 \hline
 & \delta & \frac{\alpha_k}{\eta_k}
 \end{array}
 \implies
 \begin{array}{c|c}
 & a \\
 \hline
 x_1 & \alpha_1 - \eta_1 \delta \\
 \vdots & \vdots \\
 y & \delta \\
 \vdots & \vdots \\
 x_m & \alpha_m - \eta_m \delta
 \end{array}
 \quad (4.1)$$

5. fejezet

Alerek Minkowski-összege és direkt összege

AZ ALTEREK STRUKTÚRÁJÁT vizsgáljuk a fejezetben.

5.1. Minkowski-összeg

5.1. definíció. Legyen V egy \mathbb{F} test feletti vektortér és $H_1, H_2 \subseteq V$ részhalmazok. E két halmaz összegén vagy Minkowski-összegén a

$$H_1 + H_2 = \{a + b : a \in H_1, b \in H_2\}$$

halmazt értjük. Hasonlóan, $\alpha \in \mathbb{F}$ szám mellett jelölje

$$\alpha H_1 = \{\alpha a : a \in H_1\}$$

az α szám és a H_1 halmaz szorzatát.

Világos, hogy ha például H_1 üres, akkor $H_1 + H_2$ is, és az αH_1 halmazok is üresek.

Az alábbi tulajdonságok, a vektortér axiómák közvetlen következményei. Azokat a tulajdonságokat foglaljuk össze, amelyeket a vektortér axiómákból meg tudunk menteni a Minkowski-összegre és a számmal való szorzásra.

5.2. állítás. Legyenek $A, B \subseteq V$ a V vektortér részhalmazai, továbbá $\alpha, \beta \in \mathbb{F}$ számok. Ekkor

1. $A + B = B + A$;
2. $(A + B) + C = A + (B + C)$;
3. $A + \{0\} = A$;
4. $\alpha(A + B) = \alpha A + \alpha B$;
5. $(\alpha + \beta)A \subseteq \alpha A + \beta A$ ¹;
6. $\alpha(\beta A) = (\alpha\beta)A$;
7. $1 \cdot A = A$;

Igaz továbbá, hogy ha $A \neq \emptyset$, akkor $A + V = V$ és $0 \cdot A = \{0\}$.

Ha $A \subseteq V$ legalább két elemű, akkor $A + (-A) = \{0\}$, ami azt mutatja hogy a fenti 5. tartalmazás általában nem teljesülhet egyenlőségre.

Az alábbi állítás csak átfogalmazása az altér definíciójának.

5.3. állítás. A V vektortér $N \subseteq V$ részhalmaza pontosan akkor altér V a vektortérben, ha (a) $N \neq \emptyset$ és az (b) $\alpha N + \beta N \subseteq N$ tartalmazás minden $\alpha, \beta \in \mathbb{F}$ mellett fennáll.

¹Egy $A \subseteq V$ halmazt *affin halmaznak* szokás mondani, ha zárt az 1 összegű lineáris kombinációra (*affin kombinációra*) nézve, tehát ha $\lambda A + (1 - \lambda)A \subseteq A$ tartalmazás teljesül minden $\lambda \in \mathbb{F}$ mellett. Ha A affin halmaz, akkor a fordított egyenlőség is igaz feltéve, hogy $\alpha + \beta \neq 0$. Ugyanis $\alpha A + \beta A = (\alpha + \beta) \left(\frac{1}{\alpha + \beta} (\alpha A + \beta A) \right) = (\alpha + \beta) \left(\frac{\alpha}{\alpha + \beta} A + \frac{\beta}{\alpha + \beta} A \right) \subseteq (\alpha + \beta)A$.

Látható, hogy egy affin halmaz eltoltja affin marad, hiszen ha N egy affin halmaz és $x \in V$ egy vektor, akkor $\lambda(x + N) + (1 - \lambda)(x + N) = x + \lambda N + (1 - \lambda)N \subseteq x + N$. Mivel egy altér nyilvánvalóan affin halmaz is, ezért egy altér eltoltja is affin halmaz. Még speciálisabban egy lineáris egyenletrendszer megoldáshalmaza is fontos példa affin halmazra.

Láttuk korábban, hogy alterek közös része is altér. Ezt úgy interpretálhatjuk, hogy a legbővebb olyan altér, amelyet két előre megadott altér tartalmaz ezen alterek halmazmájához közös része. Felmerül, hogy mi a legszűkebb olyan altér, amely minden két előre megadott altér tartalmazza. A válasz kézenfekvő, hiszen éppen ez a generált altér fogalma. Ha $H_1, H_2 \subseteq V$ tetszőleges halmazok, akkor $\text{lin}\{H_1 \cup H_2\}$ a H_1 és H_2 halmazokat tartalmazó legszűkebb altér. Amennyiben altereket tartalmazó legszűkebb altér keresünk, akkor többet is állíthatunk:

5.4. állítás. *Legyenek $M, N \subseteq V$ alterek. Ekkor*

$$\text{lin}(M \cup N) = M + N,$$

tehát alterek Minkowski-összege is altér, sőt ez az M és az N altereket tartalmazó legszűkebb altér.

Bizonyítás: Először is gondoljuk meg, hogy $M + N$ egy altér. Világos, hogy $M \cup N \subseteq M + N$, ezért $\text{lin}(M \cup N) \subseteq \text{lin}(M + N) = M + N$. Másrészt a $\text{lin}(M \cup N)$ halmazra, mint az $M \cup N$ halmazból képzett összes lineáris kombináció halmazára látszik, hogy $M + N \subseteq \text{lin}(M \cup N)$. \square

A következő állítás szerint, ha előre adott két altér, akkor ezek dimenzióinak összege minden azonos a két altér tartalmazó legszűkebb altér dimenziójának, és a két altér által tartalmazott legbővebb altér dimenziójának összegével.

5.5. állítás. *Legyenek M és N a V vektortér végesen generált alterei. Ekkor az $M + N$ altér is végesen generált, továbbá*

$$\dim(M + N) = \dim(M) + \dim(N) - \dim(M \cap N).$$

Bizonyítás: Világos, hogy M egy generátorrendszerének és N egy generátorrendszerének egyesítése generátorrendszer az $M + N$ altérnek is.

Legyen most $\{p_1, \dots, p_r\}$ bázis $M \cap N$ -ben. Mivel egy lineárisan független rendszer kiegészíthető egy bázissá, ezért legyen $\{m_1, \dots, m_s, p_1, \dots, p_r\}$ bázis M -ben, és $\{p_1, \dots, p_r, n_1, \dots, n_t\}$ bázis N -ben, ahol r, s, t nem negatív egészek.² Nyilvánvaló, hogy az

$$\{m_1, \dots, m_s, p_1, \dots, p_r, n_1, \dots, n_t\} \quad (\dagger)$$

rendszer egy véges generátorrendszer az $M + N$ altérnek. Most megmutatjuk, hogy a (\dagger) rendszer lineárisan független is. Legyen ehhez

$$\sum_{j=1}^s \alpha_j m_j + \sum_{j=1}^r \gamma_j p_j + \sum_{j=1}^t \beta_j n_j = 0. \quad (\ddagger)$$

Ekkor persze $\sum_{j=1}^s \alpha_j m_j = -\sum_{j=1}^r \gamma_j p_j - \sum_{j=1}^t \beta_j n_j \in M \cap N$. Léteznek tehát $\delta_1, \dots, \delta_r$ számok, amelyekre $\sum_{j=1}^s \alpha_j m_j = \sum_{j=1}^r \delta_j p_j$, emiatt a (\ddagger) egyenlőség a következő módon alakul:

$$\sum_{j=1}^r (\delta_j + \gamma_j) p_j + \sum_{j=1}^t \beta_j n_j = 0.$$

No de, $\{p_1, \dots, p_r, n_1, \dots, n_t\}$ lineárisan független, ezért itt minden együttható zérus. Speciálisan $\beta_j = 0$ minden $j = 1, \dots, t$ mellett. Ezt (\ddagger) -be visszaírva kapjuk, hogy

$$\sum_{j=1}^s \alpha_j m_j + \sum_{j=1}^r \gamma_j p_j = 0,$$

amiből az $\{m_1, \dots, m_s, p_1, \dots, p_r\}$ rendszer lineárisan függetlenségét használva kapjuk, hogy $\alpha_1 = \dots = \alpha_s = \gamma_1 = \dots = \gamma_r = 0$.

Megmutattuk tehát, hogy (\dagger) vektorrendszer bázisa $M + N$ -nek. Mivel ennek elemszáma $s + r + t = (s + r) + (r + t) - r$, ezért készen is vagyunk. \square

²Figyeljünk arra, hogy $r = 0$ is lehet, ha M és N diszjunktak. Ekkor $\{\}$ a bázis a $\{0\}$ altérben, azaz nincs egyetlen p -sem. Az összes r elemet tartalmazó szumma ilyenkor üres, tehát értéke zérus.

5.2. Direkt összeg

Szokásos szóhasználat alterek esetén, hogy két alteret *diszjunktnak* mondunk, ha metszetük csak a vektortér zérus elemét tartalmazza. A következő gondolat nem csak kettő hanem véges sok altéről szól. Gondoljuk meg, hogy akárhány de véges sok alter összege is alter.

5.6. állítás. Legyenek az $M_1, \dots, M_s \subseteq V$ alterek a V vektortér alterei ($s \geq 2$). Az alábbi feltételek ekvivalensek:

1. $\left(\sum_{j=1}^{k-1} M_j \right) \cap M_k = \{0\}$ minden $k = 2, \dots, s$ mellett.
2. A nullvektor csak egyetlen egyféléképpen áll elő, mint M_j -beli vektorok összege, azaz $\sum_{j=1}^s u_j = 0, u_j \in M_j$ esetén minden $j = 1, \dots, s$ mellett $u_j = 0$ is teljesül.
3. $\left(\sum_{\substack{j=1 \\ j \neq k}}^s M_j \right) \cap M_k = \{0\}$ minden $k = 1, \dots, s$ mellett.

Szokásos szóhasználat, hogy azt mondjuk, hogy az M_1, \dots, M_s altereknek értelmezhető a direkt összege, vagy a direkt összege értelmes, ha a fenti feltételek egyike (ezért mindegyike) fennáll.

Az 1. \Rightarrow 2. igazolása: Ha 2. nem áll fenn, akkor valamely $u_j \in M_j$ vektorokra $\sum_{j=1}^s u_j = 0$ úgy áll fenn, hogy nem minden $u_j = 0$. Legyen k a legnagyobb olyan j index, amihez nem zérus vektor tartozik. Világos, hogy van ilyen index; világos, hogy $u_k \neq 0$, de

$$u_1 + \dots + u_k = 0;$$

így az is világos, hogy $2 \leq k \leq s$. Ekkor persze u_k egy nem zérus vektor az $\left(\sum_{j=1}^{k-1} M_j \right) \cap M_k$ altérben. Megmutattuk tehát, hogy ha 2. nem áll fenn, akkor valamely k mellett 1. sem teljesül. \square

A 2. \Rightarrow 3. igazolása: Legyen $v_k \in M_k$ olyan, hogy valamely $v_j \in M_j$ vektorokra $v_k = \sum_{\substack{j=1 \\ j \neq k}}^s v_j$ teljesül. Ekkor persze $v_k + \sum_{\substack{j=1 \\ j \neq k}}^s -v_j = 0$. No de, a 2. feltétel szerint ez csak úgy lehet, ha minden vektor zérus, speciálisan $v_k = 0$. Az mutattuk meg tehát, hogy minden k mellett a k -adik alter diszjunkt az összes többi összegétől. \square

Végül a 3. \Rightarrow 1.: Nyilvánvalóan következik a $\sum_{j=1}^{k-1} M_j \subseteq \sum_{\substack{j=1 \\ j \neq k}}^s M_j$ tartalmazásból. \square

Ezek szerint a direkt összeg akkor értelmes, ha az alterek mindegyike diszjunkt az összes többi alter direkt összegétől. Azt gondoltuk meg éppen, hogy ezzel ekvivalens feltevés, hogy minden alter diszjunkt csak a sorban előtte álló alterek direkt összegétől. E második tulajdonság úgy tűnhet mintha függne az alterek sorrendjétől, de mivel ekvivalens az első tulajdonsággal, ezért persze nem függ az alterek sorrendjétől. Mind a kettő ugyanazt fejezi ki: A zérus vektor csak egyféléképpen áll elő, mint az egyes alterekből vett vektorok összege.

A következő definíció szerint, amennyiben az alterek direkt összege értelmes, úgy a direkt összeg egyszerűen a szóban forgó alterek Minkowski-összege.

5.7. definíció (direkt összeg). Legyenek az $M_1, \dots, M_s \subseteq V$ alterek a V vektortér alterei ($s \geq 2$). Azt mondjuk, hogy az $N \subseteq V$ alter az M_1, \dots, M_s alterek direkt összege, ha

1. $\sum_{j=1}^s M_j = N$,
2. $\left(\sum_{\substack{j=1 \\ j \neq k}}^s M_j \right) \cap M_k = \{0\}$ minden $k = 1, \dots, s$ mellett.

Ekkor a N -et $N = M_1 \oplus \dots \oplus M_s$ módon jelöljük.

Külön érdemes figyelni az $s = 2$, tehát csak két altér direkt összegének esetére. Ilyenkor a fenti definíció azt jelenti, hogy az M_1 és M_2 diszjunkt alterekre $N = M_1 + M_2$.

Az 5.6. állításban 3. feltétel előnye, hogy ennek alapján teljesen világos, hogy az $M_1 \oplus \dots \oplus M_s$ direkt összeg nem függ az alterek sorrendjétől. Az evvel ekivalens 1. feltevés előnye pedig, hogy ha képeznünk kell az M_1, \dots, M_s alterek direkt összegét, akkor rekurzívan járhatunk el: ha a korábbi lépésekben ellenőriztük, hogy az $M_1 \oplus \dots \oplus M_{n-1}$ értelmes, akkor a következő lépésekben csak azt kell ellenőriznünk, hogy M_n diszjunkt az első $n - 1$ altér direkt összegétől.

5.8. állítás. *Legyen az $M_1, \dots, M_s, N \subseteq V$ a V vektortér altere. Az $N = M_1 \oplus \dots \oplus M_s$ pontosan akkor teljesül, ha*

1. minden $k = 1, \dots, s$ mellett $M_k \subseteq N$, és
2. minden $v \in N$ vektorhoz létezik egyetlen $\{v_1, \dots, v_s\} \subseteq V$ vektorrendszer, amelyre $v_j \in M_j, j = 1, \dots, s$ és

$$v = \sum_{j=1}^s v_j.$$

Bizonyítás: Tegyük fel, hogy N direkt összege az M_1, \dots, M_s altereknek. minden szóba jövő k -ra $M_k \subseteq \sum_{j=1}^s M_j = N$. Mivel $N = M_1 + \dots + M_s$, ezért minden $v \in N$ -hez létezik $v_j \in M_j$, hogy $v = v_1 + \dots + v_s$. Most tegyük fel, hogy v előáll $v = u_1 + \dots + u_s$ alakban is, ahol $u_j \in M_j$ minden $j = 1, \dots, s$ mellett. Ekkor

$$0 = \sum_{j=1}^s (v_j - u_j), \text{ ahol } v_j - u_j \in M_j.$$

No de, a direkt összeg értelmes, ezért ez csak úgy lehet, hogy $v_j = u_j$ áll fenn minden $j = 1, \dots, s$ mellett, azaz az előállítás valóban egyértelmű is.

Megfordítva, most azt tegyük fel, hogy a két feltétel fennáll. Ekkor 1. szerint $\sum_{j=1}^s M_j \subseteq N + \dots + N = N$, és 2. szerint $N \subseteq M_1 + \dots + M_s$, amiből már $\sum_{j=1}^s M_j = N$ következik is.

Persze 2. speciálisan az N altér nullvektorára is igaz, ami azt jelenti, hogy az alterek direkt összegé értelmes. \square

Érdemes észrevenni, hogy a fenti állítás 2. feltétele nem más mint az alábbi két feltétel együttesének tömör megfogalmazása:

2a. $N \subseteq M_1 + \dots + M_s$,

2b. $\sum_{j=1}^s v_j = 0, v_j \in M_j$ minden $j = 1, \dots, s$ mellett $v_j = 0$.

Mégint csak a két vektortér speciális esetére figyelve azt igazoltuk, hogy N pontosan akkor az M_1 és M_2 direkt összege, ha M_1 és M_2 olyan alterek N -nek, hogy N minden eleme előáll, de csak egyféléképpen egy M_1 és egy M_2 -beli vektor összegeként.

5.9. állítás. *Legyenek az $M_1, \dots, M_s \subseteq V, (s \geq 2)$ alterek a V vektortér olyan végesen generált alterei, amelyekre $N = M_1 \oplus \dots \oplus M_s$. Ekkor $\dim(N) = \dim(M_1) + \dots + \dim(M_s)$.*

Bizonyítás: Az $s = 2$ eset éppen az 5.5. állítás, hiszen a $\{0\}$ triviális altér nulla dimenziós. Most tegyük fel, hogy s -nél kisebb számokra igaz az állítás, és lássuk be s -re. $s > 2$. Legyen $M = \sum_{j=1}^{s-1} M_j$. Világos, hogy $M = M_1 \oplus \dots \oplus M_{s-1}$ és $M \oplus M_s = N$. A már igazolt $s = 2$ eset és az indukciós feltevés szerint

$$\dim(N) = \dim(M) + \dim(M_s) = \sum_{j=1}^{s-1} \dim(M_j) + \dim(M_s) = \sum_{j=1}^s \dim(M_j). \quad \square$$

Az állításból azonnal adódik, hogy ha az M_1, M_2, \dots, M_s alterek bázisait, egy közös vektorrendszerbe tesszük, akkor az így kapott vektorrendszer az N tér egy bázisává válik. Világos ugyanis, hogy az egyesített rendszer N -nek generátorrendszeré, és az elemeinek száma az 5.9. állítás szerint éppen $\dim(N)$. Azt kaptuk tehát, hogy ez egy minimális generátorrendszer N -nek, ergo egy bázis.

5.3. Direkt kiegészítő

5.10. állítás. Legyen V egy véges dimenziós vektortér, és $M \subseteq V$ egy altér. Ekkor létezik $N \subseteq V$ altér, amelyre $M \oplus N = V$. Egy ilyen alteret az M altér direkt kiegészítőjének nevezünk. Az N altér valamennyi direkt kiegészítője egymással izomorf.

Bizonyítás: Vegyünk fel M -ben egy $\{m_1, \dots, m_r\}$ bázist. Ez persze V -ben is lineárisan független, ezért kiegészíthető a V egy $\{m_1, \dots, m_r, n_1, \dots, n_s\}$ bázisává. Definiálja $N = \text{lin} \{n_1, \dots, n_s\}$.

Ha $x \in N \cap M$, akkor x egyszer $x = \sum_{j=1}^r \alpha_j m_j$, másrészről $x = \sum_{j=1}^s \beta_j n_j$ alakú. Ebből azt kapjuk, hogy

$$\sum_{j=1}^r \alpha_j m_j + \sum_{j=1}^s -\beta_j n_j = 0.$$

No de, V fenti bázisa lineárisan független is, ezért a kiemelt lineáris kombináció minden együtthatója zérus, ami azt jelenti, hogy $x = 0$, azaz $M \cap N = \{0\}$. Az M és N alterek direkt összege tehát értelmes. Az is világos, hogy ha $x \in V$ egy tetszőleges vektor, akkor valamely $\alpha_1, \dots, \alpha_r$ és valamely β_1, \dots, β_s együtthatók mellett

$$\sum_{j=1}^r \alpha_j m_j + \sum_{j=1}^s \beta_j n_j = x,$$

hiszen a V fenti bázisa egyben V generátorrendszere is. Ha itt az első szumma értékét u -val jelöljük, a második szumma értékét pedig v -vel, akkor az $u \in M$ és a $v \in N$ vektorokra $u + v = x$, ergo $M + N = V$. Ez azt jelenti, hogy konstruáltunk $N \subseteq V$ alteret, amelyre $M \oplus N = V$ teljesül.

Ha $L \subseteq V$ egy másik olyan altér V -nek, amire $M \oplus L = V$ fennáll, akkor felhasználva a dimenziók összeadásáról szóló állítást azt kapjuk, hogy

$$\dim(L) = \dim(V) - \dim(M) = \dim(N).$$

Emlékezzünk arra, hogy végesen generált vektorterek izomorf voltának szükséges és elegendő feltétele, hogy a két altér dimenzió száma azonos legyen. Megmutattuk tehát, hogy az M altér bármely direkt kiegészítője egymással izomorf. \square

Egy altérnek sok-sok direkt kiegészítője lehet. Például tekintsük a folytonos függvények $C[0, 1]$ vektortérét \mathbb{R} felett, és legyen N az az altér, amely azon $f \in C[0, 1]$ függvényeket tartalmazza, amelyekre $f(0) = 0$. Látható, hogy tetszőleges $g \in C[0, 1]$ függvényre, amelyre $g(0) \neq 0$, a g által generált egydimenziós altér direkt kiegészítője az N altérnek.³

³Ha véges dimenziós példára vágyunk, helyettesítsük $C[0, 1]$ -et a legfeljebb n -ed fokú \mathbb{R} feletti polinomok $n + 1$ dimenziós vektortérével.

6. fejezet

Vektortér faktortere

A VEKTORTÉR KONSTRUKCIÓK végéhez érkeztünk. Láttuk, hogy alerek közörsésze, összege is vektorteret alkot. A faktortér az utolsó vektortér konstrukciós eljárásunk.

6.1. definíció. Legyen V egy vektortér és $M \subseteq V$ egy adott altér. Definiáljuk a \sim relációt a vektortér elemei felett: $x \sim y$ pontosan akkor, ha $x - y \in M$.

Látható, hogy \sim ekvivalencia reláció, hiszen

1. reflexív, ugyanis $x - x = 0 \in M$ minden $x \in V$ -re,
2. szimmetrikus, ugyanis $x - y \in M$ mellett $y - x = -(x - y) \in -M = M$,
3. tranzitív, ugyanis $x - y \in M$ és $y - z \in M$ esetén $x - z = (x - y) + (y - z) \in M + M = M$

Legyen adott $x \in V$ mellett az x elemet tartalmazó ekvivalencia osztály M_x , azaz

$$M_x = \{u \in V : u \sim x\}.$$

Tudjuk, hogy az összes ekvivalencia osztályok $\{M_x : x \in V\}$ halmazrendszere a V egy partícióját alkotja, ami azt jelenti, hogy $V = \bigcup_{x \in V} M_x$; ha valamely $x, y \in V$ mellett $M_x \cap M_y \neq \emptyset$, akkor $M_x = M_y$; és minden $x \in V$ mellett $M_x \neq \emptyset$.

Most azt gondoljuk meg, hogy minden egyes ekvivalencia osztály tekinthető úgy is mint, bármelyik elemével való eltoltja az M vektortérnek. Speciálisan az is adódik, hogy az ekvivalencia osztályok affin halmazok.

6.2. állítás. A fenti jelölések mellett $M_x = u + M$ minden $u \in M_x$ mellett. Speciálisan, minden $x \in V$ mellett $M_x = x + M$.

Bizonyítás: Megmutatjuk, hogy $M_x \subseteq u + M$. Legyen $v \in M_x$ tetszőleges. Ekkor $v \sim x$, $u \sim x$, ezért $v \sim u$. Ez azt jelenti, hogy $v - u \in M$, amiből már adódik, hogy $v \in u + M$.

Megfordítva, most igazoljuk, hogy $u + M \subseteq M_x$. Legyen tehát $v \in u + M$. Ekkor $v - u \in M$, ergo $v \sim u$. No de, $u \sim x$ is fel van téve, emiatt $v \sim x$, azaz $v \in M_x$. \square

Most definiálni szeretnénk az ekvivalencia osztályok halmazán összeadás és egy testbeli elemmel való szorzás műveletet. Az összeadás művelet legyen a korábban definiált Minkowski-összeg. Ez valóban művelet az ekvivalencia osztályokra megszorítva, hiszen ha M_{x_1} és M_{x_2} két ekvivalencia osztály, akkor

$$M_{x_1} + M_{x_2} = (x_1 + M) + (x_2 + M) = x_1 + x_2 + M + M = (x_1 + x_2) + M = M_{x_1 + x_2}, \quad (\dagger)$$

azaz két ekvivalencia osztály Minkowski-összege is egy ekvivalencia osztály. Sőt, az is adódik, hogy *egy-egy vektorhoz tartozó ekvivalencia osztályok Minkowski-összege azonos e két vektor összegéhez tartozó ekvivalencia osztályjal*.

A skalárral való szorzás már nem ilyen egyszerű, hiszen $0 \cdot M_x = \{0\}$, de ez utóbbi halmaz általában nem egy ekvivalencia osztály, ezért az ekvivalencia osztályok halmazán a skalárral való szorzás nem egy művelet.

6.3. definíció (ekvivalencia osztály számszorosa). Legyen $\alpha \in \mathbb{F}$ egy szám és M_x egy ekvivalencia osztály. Definiálja

$$\alpha * M_x = \begin{cases} \alpha M_x & , \text{ha } \alpha \neq 0 \\ M & , \text{ha } \alpha = 0. \end{cases}$$

Egyrészt vegyük észre, hogy $\alpha \neq 0$ mellett $\alpha M_x = \alpha(x + M) = \alpha x + \alpha M = \alpha \cdot x + M = M_{\alpha \cdot x}$, másrészt azt lássuk, hogy az $\alpha = 0$ esetben $M = 0 + M = M_0 = M_{0 \cdot x}$, ami azt jelenti, hogy a fenti definícióra az

$$\alpha * M_x = M_{\alpha \cdot x} \quad (\ddagger)$$

azonosság is teljesül minden $x \in V$ vektorra és minden $\alpha \in \mathbb{F}$ számra. Kiderült tehát, hogy a most bevezetett $*$ skalárral való szorzást alkalmazva egy ekvivalencia osztálynak egy számmal való szorzata egy ekvivalencia osztárral lesz, sőt egy vektorhoz tartozó ekvivalencia osztály α szorosa azonos a vektor α -szorosához tartozó ekvivalencia osztályval.

6.4. definíció-állítás (faktortér). Legyen V egy az \mathbb{F} test feletti vektortér, és $M \subseteq V$ egy rögzített altér. Jelölje \sim azt V vektortér elemein értelmezett relációt, amelyre $x \sim y$, akkor és csak akkor, ha $x - y \in M$. Láttuk, hogy ez ekvivalencia reláció. Jelölje

$$V/M = \{M_x : x \in V\}$$

az ekvivalencia osztályok halmazát. Definiáljuk a V/M halmazrendszer elemei mint halmazok közt az összeadás nevű műveletet mint a halmazok Minkowski-összegét; és a skalárral való szorzást műveletet mint a $*$ fent bevezetett szorzást. (Láttuk, hogy az

$$M_{x_1} + M_{x_2} = M_{x_1 + x_2}, \quad \text{és} \quad \alpha * M_x = M_{\alpha \cdot x}$$

azonosságok minden $x_1, x_2, x \in V$ és minden $\alpha \in \mathbb{F}$ mellett fennállnak.) Ekkor az itt bevezetett $(V/M, +, *)$ struktúra az \mathbb{F} test feletti vektorteret alkot. Ezt a vektorteret nevezzük a V vektortér M altérre szerinti faktorterének.

Bizonyítás: Meggondoltuk már, hogy a Minkowski-összeg és a $*$ szorzás eredménye egy ekvivalencia osztály. Most vegyük sorra a vektortér axiómákat: Az

1. $M_{x_1} + M_{x_2} = M_{x_2} + M_{x_1}$;
2. $(M_{x_1} + M_{x_2}) + M_{x_3} = M_{x_1} + (M_{x_2} + M_{x_3})$;

axiómák tetszőleges halmazokra, nem csak ekvivalencia osztályokra is fennállnak.

Az ekvivalencia osztályok között $M_0 = M$ neutrális elem, hiszen minden $x \in V$ mellett $M_x + M = (x + M) + M = x + (M + M) = x + M = M_x$.

3. minden M_x ekvivalencia osztályra $M_x + M = M_x$;
4. minden M_x ekvivalencia osztályra $M_x + M_{-x} = M$, hiszen $M_x + M_{-x} = M_{x+(-x)} = M_0 = M$, felhasználva a már igazolt (\ddagger) azonosságot.

A $*$ szorzás és az összeadás kapcsolatát leíró aximák ellenőrzéséhez használjuk (\dagger) és (\ddagger) azonosságokat:

5. $(\alpha + \beta) * M_x = \alpha * M_x + \beta * M_x$, hiszen $(\alpha + \beta) * M_x = M_{(\alpha+\beta)x} = M_{\alpha x+\beta x} = M_{\alpha x} + M_{\beta x} = \alpha * M_x + \beta * M_x$;
6. $\alpha * (M_{x_1} + M_{x_2}) = \alpha * M_{x_1} + \alpha * M_{x_2}$, hiszen $\alpha * (M_{x_1} + M_{x_2}) = \alpha * (M_{x_1+x_2}) = M_{\alpha(x_1+x_2)} = M_{\alpha x_1+\alpha x_2} = M_{\alpha x_1} + M_{\alpha x_2} = \alpha * M_{x_1} + \alpha * M_{x_2}$;
7. $(\alpha\beta) * M_x = \alpha * (\beta * M_x)$, hiszen $(\alpha\beta) * M_x = M_{(\alpha\beta)x} = M_{\alpha(\beta x)} = \alpha * M_{\beta x} = \alpha * (\beta * M_x)$;
8. $1 * M_x = M_x$, hiszen $1 * M_x = M_{1 \cdot x} = M_x$.

Ezt kellett belátni. □

6.1. Izomorfia tételek

6.5. állítás. Legyen M a V vektortér egy altere. Definiálja $\varphi : V \rightarrow V/M$ az $x \mapsto M_x$ függvényt.

1. Ekkor $\varphi : V \rightarrow V/M$ egy szürjektív lineáris operáció.
2. Legyen most N az M egy direkt kiegészítője, azaz $M \oplus N = V$. Definiálja Φ a φ megszorítását az N direkt kiegészítőre. Ekkor $\Phi : N \rightarrow V/M$ egy izomorfizmus.

Bizonyítás: 1. A (\dagger) és (\ddagger) szerint $\varphi(\alpha x_1 + \beta x_2) = M_{\alpha x_1 + \beta x_2} = M_{\alpha x_1} + M_{\beta x_2} = \alpha * M_{x_1} + \beta * M_{x_2} = \alpha * \varphi(x_1) + \beta * \varphi(x_2)$, ami éppen azt jelenti, hogy φ egy lineáris operáció. Ha $A \in V/M$ egy ekvivalencia osztály, és $x \in A$ tetszőleges eleme, akkor $\varphi(x) = M_x = A$, ergo φ szürjekció.

2. Világos, hogy Φ lineáris leképezés leszűkítéseként maga is lineáris.

Most nézzük, hogy Φ is szürjekció marad. Ha $A \in V/M$ egy ekvivalencia osztály, akkor létezik $x \in V$, amelyre $\varphi(x) = A$. No de, $x = u + v$ alakú, ahol $u \in M$ és $v \in N$, és $M = M_0$ az V/M vektortér neutrális eleme. Így $A = \varphi(x) = \varphi(u + v) = \varphi(u) + \varphi(v) = M + \varphi(v) = \varphi(v)$.

Az injektív tulajdonsághoz legyen $v \in \ker \Phi$, azaz $v \in N$, amelyre $\varphi(v) = M$. Ez a φ definíciója miatt azt jelenti, hogy $M_v = M$, azaz $v \in M$. Azt kaptuk tehát, hogy $v \in N \cap M = \{0\}$, ami csak úgy lehetséges, hogy $v = 0$. Megmutattuk tehát, hogy $\ker \Phi = \{0\}$, ami Φ injektivitását jelenti. \square

Az állítás legelső következménye, hogy az M altér bármely direkt kiegészítője izomorf a V/M faktortérrel, emiatt bármely direkt kiegészítő izomorf bármely direkt kiegészítővel is. Feltéve, hogy létezik direkt kiegészítő, izomorfiatól eltekintve csak egyetlen egy létezik. De van-e, minden direkt kiegészítő? Véges dimenziós esetben már láttuk az igenlő választ. Nem véges dimenziós vektorterekre is igaz az állítás, de itt nem igazoljuk.

6.6. definíció (co-dimenzió). Azt mondjuk, hogy a V vektortér M altere k co-dimenziós, ha létezik k -dimenziós direkt kiegészítője M -nek, azaz létezik olyan N altere V -nek, amelyre $M \oplus N = V$ és $\dim(N) = k$. A tényt, hogy M altérnek létezik k -dimenziós direkt kiegészítője codim $M = k$ módon jelöljük.

6.7. állítás. Legyenek V, W ugyanazon test feletti vektorterek, és $A \in L(V, W)$ egy lineáris operáció. Ekkor a $V/\ker A$ faktortér izomorf az $\text{Im } A$ vektortérrel.

Bizonyítás: Jelölje $M = \ker A$ a V vektortér alterét. Definiálni szeretnénk egy $\Omega : \text{Im } A \rightarrow V/M$ izomorfizmust. Ha $Ax_1 = y = Ax_2$, akkor $x_1 - x_2 \in M$, azaz $x_1 \sim x_2$, ergo $M_{x_1} = M_{x_2}$. Emiatt $y \in \text{Im } A$ mellett

$$y \mapsto M_x, \text{ ahol } Ax = y$$

jól definiálja az $\Omega : \text{Im } A \rightarrow V/M$ függvényt.

1. Ω lineáris operáció: Legyen $Ax_1 = y_1$ és $Ax_2 = y_2$. Ekkor $A(\alpha x_1 + \beta x_2) = \alpha y_1 + \beta y_2$, így

$$\Omega(\alpha y_1 + \beta y_2) = M_{\alpha x_1 + \beta x_2} = \alpha * M_{x_1} + \beta * M_{x_2} = \alpha * \Omega(y_1) + \beta * \Omega(y_2).$$

2. Ω szürjekció: Ha $H \in V/M$ egy ekvivalencia osztály, akkor tetszőleges $x \in H$ mellett legyen $y = Ax$. Ekkor

$$\Omega(y) = M_x = H,$$

amiivel azt igazoltuk, hogy minden ekvivalencia osztály egy $\text{Im } A$ -beli vektor Ω képe.

3. Ω injekció: Legyen $y \in \ker \Omega$, azaz $y \in W$, amelyre $\Omega(y)$ a V/M faktortér neutrális eleme, azaz $\Omega(y) = M$. Ez Ω definíciója szerint csak úgy lehet, ha $M = M_x$, ahol $Ax = y$. No de, $M = M_0$, így $y = A0 = 0$. Megmutattuk tehát, hogy $\ker \Omega = \{0\}$, ami éppen Ω injektivitása. \square

6.8. állítás. Tegyük fel, hogy M a V vektortér olyan altere, amelynek van véges dimenziós direkt kiegészítője. Ekkor a V/M faktortér is végesen generált, és

$$\text{codim } M = \dim(V/M).$$

Bizonyítás: Válasszunk N véges dimenziós alterét V -nek, amelyre $V = M \oplus N$. Láttuk, hogy V/M és N izomorf vektorterek, emiatt V/M is végesen generált, és $\dim(V/M) = \dim(N) = \text{codim}(M)$. \square

6.9. állítás. Legyenek V, W vektorterek, továbbá $A \in L(V, W)$ egy olyan lineáris operáció, amelyre az $\text{Im } A \subseteq W$ végesen generált. Ekkor a $\ker A$ altérnek van $\text{Im } A$ -val izomorf direkt kiegészítője, emiatt

$$\text{codim}(\ker A) = \dim(\text{Im } A).$$

Bizonyítás: Legyen $\text{Im } A$ egy bázisa $\{Ax_1, \dots, Ax_r\}$. Ha például x_i előállna, mint a többi x_j vektor lineáris kombinációja, akkor az A linearitása miatt Ax_i is előállna ugyanazon együtthatókkal mint a többi Ax_j lineáris kombinációja. Ez azt jelenti, hogy $\{x_1, \dots, x_r\}$ is lineárisan független vektorrendszer. Definiálja $N = \text{lin}\{x_1, \dots, x_r\}$. Persze $\{x_1, \dots, x_r\}$ egy bázisa N -nek, ergo $\dim(N) = r$.

Ha $x \in \ker A \cap N$, akkor valamely α_j skalárokkal $x = \sum_{j=1}^r \alpha_j x_j$ és $Ax = 0$. Így $0 = \sum_{j=1}^r \alpha_j Ax_j$, ami csak $\alpha_1 = \dots = \alpha_r = 0$ esetben lehetséges az $\{Ax_1, \dots, Ax_r\}$ rendszer lineárisan függetlensége szerint. Megmutattuk tehát, hogy $\ker A \cap N = \{0\}$, ergo $\ker A$ -nak és N -nek értelmes a direkt összege.

Ha $x \in V$ tetszőleges vektor, akkor $Ax \in \text{Im } A$, emiatt Ax valamely skalárokkal $Ax = \sum_{j=1}^r \alpha_j Ax_j$ alakú. Világos, hogy így $x - \sum_{j=1}^r \alpha_j x_j \in \ker A$, ami igazolja, hogy $V = \ker A + N$.

Láttuk tehát, hogy $V = \ker A \oplus N$, azaz $\ker A$ -nak valóban találtunk véges dimenziós direkt kiegészítőjét. Az is világos, hogy codim($\ker A$) = $\dim(N) = r = \dim(\text{Im } A)$. \square

Magától értetődik a következő két észrevétel.

6.10. állítás. *Ha $A \in L(V, \mathbb{F})$ egy nem zérus lineáris funkcionál a V vektortéren, akkor $\ker A$ egy 1 co-dimenziós altere V -nek.*

6.11. állítás (Rang–defektus-tétel). *Legyen V egy véges dimenziós vektortér, és W egy tetszőleges vektortér, továbbá $A \in L(V, W)$ egy lineáris operáció. Ekkor $\text{Im } A$ is véges dimenziós, és*

$$\dim(\ker A) + \dim(\text{Im } A) = \dim(V).$$

II. rész

Tavasz

7. fejezet

Mátrixok és lineáris operációk

MÁTRIXOK ÉS LINEÁRIS operációk kapcsolatát vizsgáljuk. Azt már a mátrixok bevezetésekor is láttuk, hogy egy $n \times m$ méretű mátrix egyben tekinthető valamely $\mathbb{F}^m \rightarrow \mathbb{F}^n$ lineáris operációnak olyan módon, hogy az egy $x \in \mathbb{F}^m$ oszlopvektorhoz a mátrix szorzás definíciójának megfelelően az $A \cdot x \in \mathbb{F}^n$ oszlopvektort rendeli. Ebben a fejezetben a mátrixok ezen interpretációját erősítjük tovább.

7.1. Rang–defektus–tételek következménye

Láttuk, hogy tetszőleges $A \in L(V, W)$ lineáris operáció mellett

$$\nu(A) + \rho(A) = \dim(V),$$

ahol $\nu(A) = \dim(\ker A)$ az A defektusa, $\rho(A) = \dim(\text{Im } A)$ az A rangja. Ebből adódóan, ha $A \in L(V)$ egy lineáris transzformáció, akkor $\nu(A) = 0$ és $\rho(A) = \dim(V)$ egymással ekvivalens feltevések.

Itt az első feltétel pontosan A injektivitását, míg a második feltétel pontosan A szürjektivitását jelenti. Azt látjuk tehát, hogy lineáris transzformációk esetén a transzformáció injektivitása és szürjektivitása egyszerre teljesül, vagy egyszerre nem teljesül.

7.1. állítás. Legyen V egy véges dimenziós vektortér, és $A \in L(V)$ egy lineáris transzformáció.

1. Ekkor az alábbi feltételek egymással ekvivalensek.

- a) A injektív;
- b) A szürjektív;
- c) A vektortér-izomorfizmus.
- d) Létezik $B \in L(V)$ lineáris transzformáció, amelyre $A \circ B = I$.

2. Ha a fenti d) fennáll, akkor $B \circ A = I$ is teljesül.

Bizonyítás: Az első három pont ekvivalens voltát már meggondoltuk.

Ha $A \circ B = I$, akkor A szürjektív, hiszen az $y \in V$ vektor előáll, mint a By vektor A képe.

Megfordítva, ha A szürjektív, akkor a) szerint injektív is, van tehát $V \rightarrow V$ inverze. De lineáris függvény inverze is lineáris, így a $B = A^{-1}$ jelölést bevezetve találtunk $B \in L(V)$ transzformációt, amelyre $A \circ B = I$.

Most tegyük fel, hogy valamely $B \in L(V)$ -re $A \circ B = I$. Ekkor a) szerint A injektív is, ergo $\ker A = \{0\}$. No de, minden $x \in V$ mellett

$$A((B \circ A)x - x) = A(B(Ax)) - Ax = (A \circ B)(Ax) - Ax = I(Ax) - Ax = Ax - Ax = 0,$$

ezért $(B \circ A)x - x \in \ker A = \{0\}$. $(B \circ A)x = x$ minden $x \in V$ mellett, ami éppen azt jelenti, hogy $B \circ A = I$ is fennáll. \square

Kiderült tehát, hogy ugyanúgy mint amikor mátrixok regularitásáról volt szó, az $A \in L(V)$ lineáris transzformáció pontosan akkor injektív, ha van olyan $B \in L(V)$ lineáris transzformáció, amelyre $AB = I$ (vagy $BA = I$) teljesül. Ekkor $A^{-1} = B$.

7.2. Mátrixok tere mint koordináta-tér

Világos, hogy adott \mathbb{F} test feletti $n \times m$ méretű mátrixok az \mathbb{F} feletti vektorteret alkotnak. Az is világos, hogy ha $A_{i,j}$ jelöli azt az $n \times m$ méretű mátrixot, amelynek minden helyén 0 van az i -edik sor j -edik helyének kivételével, ahol 1 szerepel, akkor az

$$\{A_{i,j} : i = 1, \dots, n; j = 1, \dots, m\}$$

mátrixok rendszere egy bázis az $\mathbb{F}^{n \times m}$ térben. Így persze $\dim \mathbb{F}^{n \times m} = n \cdot m$.

Most átérünk az $L(V, W)$ lineáris operációk vektorterének vizsgálatára. Először azt gondoljuk meg, hogy lineáris transzformáció egy bázison tetszőlegesen és egyértelműen előírható.

7.2. állítás. Legyenek V és W ugyanazon test feletti vektorterek. Legyen V -ben a $\{v_1, \dots, v_m\}$ egy bázis, és rögzítsünk W -ben egy tetszőleges m elemű $\{w_1, \dots, w_m\}$ vektorrendszert.

Ekkor létezik egyetlen egy $A \in L(V, W)$ lineáris operáció, amelyre $Av_j = w_j$ minden $j = 1, \dots, m$ esetén.

Bizonyítás: Definiáljuk az $A : V \rightarrow W$ függvényt a következőképpen. minden $v \in V$ egyértelműen áll elő mint $v = \sum_{j=1}^m \alpha_j v_j$. Egy ilyen v mellett legyen

$$A(v) = \sum_{j=1}^m \alpha_j w_j.$$

Világos, hogy $A : V \rightarrow W$ függvény jól definiált, hiszen a rögzített bázisban minden elem előáll és egyetlen egyféleképpen áll elő mint a bázis elemek egy lineáris kombinációja. Megmutatjuk, hogy az így definiált A függvény egy lineáris operáció. Legyen $x_1 = \sum_{j=1}^m \alpha_j v_j$ és $x_2 = \sum_{j=1}^m \beta_j v_j$. Tetszőleges $\alpha, \beta \in \mathbb{F}$ mellett

$$\alpha x_1 + \beta x_2 = \sum_{j=1}^m (\alpha \alpha_j + \beta \beta_j) v_j,$$

tehát A definíciója szerint

$$A(\alpha x_1 + \beta x_2) = \sum_{j=1}^m (\alpha \alpha_j + \beta \beta_j) w_j = \alpha \sum_{j=1}^m \alpha_j w_j + \beta \sum_{j=1}^m \beta_j w_j = \alpha A(x_1) + \beta A(x_2).$$

Ez éppen A függvény linearitását jelenti. Az is világos, hogy $v_j = 0v_1 + \dots + 1v_j + \dots + 0v_m$, tehát az A függvény definíciója értelmében

$$Av_j = w_j$$

valóban fennáll minden $j = 1, \dots, m$ mellett.

Most tegyük fel, hogy $B \in L(V, W)$ szintén teljesíti a $Bv_j = w_j$ feltételeket. Ekkor tetszőleges $v \in V$ mellett, ha $v = \sum_{j=1}^m \alpha_j v_j$ alakú, akkor az A definíciója és B linearitása miatt

$$Av = \sum_{j=1}^m \alpha_j w_j = \sum_{j=1}^m \alpha_j Bv_j = B\left(\sum_{j=1}^m \alpha_j v_j\right) = Bv,$$

ami éppen azt jelenti, hogy $A = B$. □

7.3. állítás. Legyenek az $\{e_1, \dots, e_m\} \subseteq V$ és az $\{f_1, \dots, f_n\} \subseteq W$ bázisok rögzítve. Definiáljuk valamely rögzített $i \in \{1, \dots, n\}$ és valamely rögzített $j \in \{1, \dots, m\}$ mellett az $A_{i,j} \in L(V, W)$ lineáris operációt az

$$A_{i,j}(e_k) = \delta_{j,k} f_i \tag{7.1}$$

azonosságokkal, ahol $k = 1, \dots, m$. Ekkor az $\{A_{i,j} : i = 1, \dots, n; j = 1, \dots, m\}$ lineáris operációk rendszere egy bázisa az $L(V, W)$ vektortérnek, ezért

$$\dim(L(V, W)) = n \cdot m.$$

Bizonyítás: Azt kell először is látni, hogy (7.1.) azonosságok összesen az előző 7.2. állítás alkalmazását írják elő rögzített i, j mellett az $\{e_1, \dots, e_j, \dots, e_m\}$ és az $\{0, \dots, 0, f_i, 0, \dots, 0\}$ két pontosan m elemű vektorrendszerre. Ezt úgy is fogalmazhatjuk, hogy az $A_{i,j}$ az a lineáris operáció, amely a bázis minden nem j -edik elemét zérusra viszi, és a j -edik bázis elemet pedig f_i -re. A 7.2. állítás szerint vannak ilyen $A_{i,j}$ lineáris transzformációk, és minden i, j pár mellett csak egyetlen egy van.

Most megmutatjuk, hogy az $\{A_{i,j} : i = 1, \dots, n; j = 1, \dots, m\}$ lineárisan független rendszer. Legyenek az $\alpha_{i,j} \in \mathbb{F}$ számok olyanok, amelyekre $\sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j} = 0$. Ekkor tetszőleges $k \in \{1, \dots, m\}$ esetén

$$0 = \left(\sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j} \right) e_k = \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j} e_k = \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} \delta_{j,k} f_i = \sum_{i=1}^n \alpha_{i,k} f_i.$$

No de, az $\{f_1, \dots, f_n\}$ egy lineárisan független rendszer, ergo csak a triviális lineáris kombinációja zérus, ergo $\alpha_{i,k} = 0$ minden $i = 1, \dots, n$ mellett. Persze ez minden k mellett megismételhető, tehát azt kaptuk, hogy valamennyi $\alpha_{i,j}$ együttható zérus.

Most azt mutatjuk meg, hogy az $\{A_{i,j} : i = 1, \dots, n; j = 1, \dots, m\}$ egy generátorrendszer az $L(V, W)$ vektortérnek. Legyen $A \in L(V, W)$ egy rögzített lineáris operáció. Definiáljuk az $\alpha_{i,j}$ számokat, mint az $Ae_j \in W$ vektor $\{y_1, \dots, y_n\}$ bázisban felírt koordináta-vektorának i -edik elemét. Magyarul

$$Ae_j = \sum_{i=1}^n \alpha_{i,j} f_i.$$

Ekkor minden $k \in \{1, \dots, m\}$ mellett

$$\left(\sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j} \right) e_k = \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j} e_k = \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} \delta_{j,k} f_i = \sum_{i=1}^n \alpha_{i,k} f_i = Ae_k.$$

Ez azt jelenti, hogy az A és az $\sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j}$ lineáris operátorok az $\{e_1, \dots, e_m\}$ bázis minden elemén megegyeznek. Láttuk, hogy bázison felvett értékek már egyértelműen meghatározzák a lineáris operációt, ezért e két lineáris operáció is azonos. Találtunk tehát $\alpha_{i,j}$ számokat, amelyekre $A = \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j}$, ami azt jelenti, hogy A előáll mint az $A_{i,j}$ függvények valamely lineáris kombinációja. \square

Érdemes későbbre is eltennünk magunkban, hogy hogyan találtunk az adott A operátorhoz azon $\alpha_{i,j}$ számokat, amelyekre az

$$A = \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} A_{i,j} = \sum_{j=1}^m \sum_{i=1}^n \alpha_{i,j} A_{i,j}$$

egyenlőség teljesül: $\alpha_{i,j}$ a j -edik bázis elem A képének i -edik koordinátája. Ugyanez a koordináta-vektor fogalmával: Az $Ae_j \in W$ vektornak az $\{y_1, \dots, y_n\}$ bázisban felírt koordináta-vektorának az elemei adják az $\alpha_{1,j}, \alpha_{2,j}, \dots, \alpha_{n,j}$ számokat, formálisabban:

$$[Ae_j]_{\{y_1, \dots, y_n\}} = \begin{pmatrix} \alpha_{1,j} \\ \alpha_{2,j} \\ \vdots \\ \alpha_{n,j} \end{pmatrix}.$$

A következő gondolat előtt arra kell emlékeznünk, hogy minden véges dimenziós vektortér izomorf a koordináta-terével. Láttuk, hogy $\dim(L(V, W)) = \dim(V) \cdot \dim(W)$. De mi lesz $L(V, W)$ koordináta-tere? A válaszhoz rögzítenünk kell a bázis elemek egy sorrendjét.

7.4. definíció (lineáris operátor mátrixa). (A 7.1.) azonosságokkal definiált $A_{i,j}$ lineáris operátorokat állítsuk a következő sorrendbe:

$$\underbrace{\{A_{1,1}, A_{2,1}, \dots, A_{n,1}\}}_{j=1}, \underbrace{\{A_{1,2}, A_{2,2}, \dots, A_{n,2}\}}_{j=2}, \underbrace{\{A_{1,3}, A_{2,3}, \dots, A_{n,3}\}}_{j=3}, \dots, \dots, \dots, \dots, \dots, \underbrace{\{A_{1,m}, A_{2,m}, \dots, A_{n,m}\}}_{j=m}$$

Láttuk, hogy a fenti konstruált $\alpha_{i,j}$ számokkal

$$A = \sum_{j=1}^m \sum_{i=1}^n \alpha_{i,j} A_{i,j}.$$

Ez azt jelenti, hogy az $A \in L(V, W)$ függvénynek a fenti bázisban felírt koordináta-vektora az $\alpha_{i,j}$ elemekből a fenti sorrendben képzett oszlopvektor, azaz

$$[A] = \begin{pmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \vdots \\ \alpha_{n,1} \\ \alpha_{1,2} \\ \alpha_{2,2} \\ \vdots \\ \alpha_{n,2} \\ \alpha_{1,3} \\ \alpha_{2,3} \\ \vdots \\ \alpha_{n,3} \\ \vdots \\ \alpha_{1,m} \\ \alpha_{2,m} \\ \vdots \\ \alpha_{n,m} \end{pmatrix} \quad (\dagger)$$

Mint minden vektortér így az $L(V, W)$ is izomorf a koordináta-terével, ergo $L(V, W)$ izomorf az $\mathbb{F}^{n \times m}$ vektortérrel.

Érdemes a koordinátatér elemeit, azaz az $n \times m$ elemből álló oszlopvektorokat n koordinátánként megtörni, és ezzel egy $n \times m$ -es mátrixba rendezni. Ezt a jelölést alkalmazva az A lineáris operáció a fenti bázisban felírt koordináta vektora az az $n \times m$ -es mátrix, amelyre

$$[A] = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} & \dots & \alpha_{1,m} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} & \dots & \alpha_{2,m} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} & \dots & \alpha_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \alpha_{n,3} & \dots & \alpha_{n,m} \end{pmatrix}.$$

Ezt a mátrixot nevezzük az $A \in L(V, W)$ lineáris operáció mátrixának az előre rögzített $\{e_1, \dots, e_m\} \subseteq V$ és $\{f_1, \dots, f_n\}$ bázisok mellett.

Meggondoltuk tehát, hogy a kapcsolat lineáris operáció és mátrixa között azonos avval a kapcsolattal, ami általában egy vektor és koordinátái között van.

Nagyon fontos, hogy bármilyen lineáris operáció mátrixát fel tudjuk írni, emiatt foglaljuk össze az eddigieket.

7.5. állítás. *Tegyük fel, hogy $A \in L(V, W)$ egy lineáris operáció. Rögzítük a V és a W vektortér egy-egy bázisát. Legyen tehát $\{e_1, \dots, e_m\} \subseteq V$ egy bázis és $\{f_1, \dots, f_n\} \subseteq W$ egy bázis. Az A operációnak a fenti rögzített bázisokban felírt mátrixa az az $n \times m$ méretű $[A]_{\{e_1, \dots, e_m\} \{f_1, \dots, f_n\}}$ mátrix, ¹ amelynek j -edik oszlopa az Ae_j vektor $\{y_1, \dots, y_n\}$ bázisban felírt koordinátája. Formálisan:*

$$[A]_{\{e_1, \dots, e_m\} \{f_1, \dots, f_n\}}^j = [Ae_j]_{\{y_1, \dots, y_n\}}.$$

¹Ha világos, hogy mely bázisokat rögzítjük akkor a nehézkes $[A]_{\{e_1, \dots, e_m\} \{f_1, \dots, f_n\}}$ jelölés helyett csak $[A]$ -t írunk. Persze minden tudnunk kell, hogy az A lineáris operátor mely bázisokban felírt mátrixáról van szó.

Speciálisan, ha $V = W$, akkor $A : V \rightarrow V$ lineáris transzformációról beszélünk. Amikor egy lineáris transzformáció mátrixáról van szó, akkor az minden úgy értendő, hogy a V vektortérnek mint az értelmezési tartománynak és a V vektortérnek mint értékkészletnek is ugyanazt a bázisát választjuk. Ekkor tehát A lineáris transzformáció $[A]_{\{e_1, \dots, e_m\}}_{\{e_1, \dots, e_m\}}$ mátrixára:²

$$[A]_{\{e_1, \dots, e_m\}}^j = [Ae_j]_{\{e_1, \dots, e_m\}}.$$

Meggondoltuk tehát, hogy az $L(V, W)$ lineáris operátorok vektortere izomorf az $\mathbb{F}^{\dim(W) \times \dim(V)}$ mátrixok vektorterével. Az izomorfizmus tehát az a leképezés, amely egy lineáris transzformációhoz hozzárendeli annak – valamely előre megadott bázisokban felírt – mátrixát. Emiatt persze $A, B \in L(V, W)$ és $\alpha, \beta \in \mathbb{F}$ mellett

$$[\alpha A + \beta B] = \alpha [A] + \beta [B]. \quad (7.2)$$

Izomorf struktúrák között nem érdemes különbséget tenni, viszont vigyáznunk kell arra, hogy olyan fogalma-kat definiálunk, amelyek invariánsak az izomorfiára. Például lineáris operáció rangja definíció szerint a képtere dimenziója, mátrix rangja definíció szerint a feszítőrang, azaz a legkisebb r szám amelyre a mátrix felírható $n \times r$ és egy $r \times m$ méretű mátrix szorzataként. Látni fogjuk, hogy lineáris operátorok és mátrixának rangja azonos.

7.6. állítás. Legyen $A \in L(V, W)$ lineáris operátor és $x \in V$ egy vektor. Rögzítsük az $\{e_1, \dots, e_m\} \subseteq V$ és az $\{f_1, \dots, f_n\} \subseteq W$ bázisokat. Ekkor

$$[Ax]_{\{f_1, \dots, f_n\}} = [A]_{\{e_1, \dots, e_m\}} \cdot [x]_{\{e_1, \dots, e_m\}}.$$

Emiatt $\text{Im } A$ és $\text{Im } [A]$ egymással izomorf alterek, hasonlóan $\ker A$ és $\ker [A]$ egymással izomorf vektorterek, így dimenziójuk is azonos. Konkrétan A lineáris operátorának és az $[A]$ mátrixának a rangja is defektusa is azonos.

Bizonyítás: Legyen $[x]_{\{e_1, \dots, e_m\}} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix}$. Ez azt jelenti, hogy $x = \sum_{j=1}^m \xi_j e_j$, így $Ax = \sum_{j=1}^m \xi_j Ae_j$. Véve az Ax vektor $\{f_1, \dots, f_n\}$ bázisban felírt koordináta-vektorát

$$[Ax]_{\{f_1, \dots, f_n\}} = \left[\sum_{j=1}^m \xi_j Ae_j \right]_{\{f_1, \dots, f_n\}} = \sum_{j=1}^m \xi_j [Ae_j]_{\{f_1, \dots, f_n\}} = \sum_{j=1}^m \xi_j [A]_{\{e_1, \dots, e_m\}}^j = [A]_{\{e_1, \dots, e_m\}} \cdot [x]_{\{e_1, \dots, e_m\}}.$$

Ebből persze már könnyen adódik, ahogy $Ax = 0$ ekvivalens $[A][x] = 0$ feltétellel, emiatt $\ker A$ és $\ker [A]$ izomorfak, így $\nu(A) = \nu([A])$.

Hasonlóan $\text{Im } A$ és $\text{Im } [A]$ is izomorfak. A rangtétel szerint $[A]$ mátrix rangja megegyezik az oszlopvektorai generálta altérben lévő maximális lineárisan független rendszer elemszámával. Mivel egy vektor pontosan akkor van $[A]$ képterében, ha előáll mint az oszlopai lineáris kombinációja, ezért $[A]$ rangja azonos $[A]$ képterének dimenziójával, így $\rho(A) = \rho([A])$. \square

7.3. Lineáris operátorok szorzata

Emlékezzünk arra, hogy a lineáris operáció mátrixát úgy kaptuk, hogy a koordinátáit n elemenként megtörve az oszlopvektort egy mátrixszá rendeztük át. Ennek az átrendezésnek az igazi értelme, hogy ilyen módon a mátrix szorzás művelet a lineáris operátorok kompozíójával kapcsolódik össze.

²A nehézkes $[A]_{\{e_1, \dots, e_m\}}$ helyett egyszerűbben $[A]_{\{e_1, \dots, e_m\}}$, vagy még inkább ha a szövegkörnyezetből nyilvánvaló, hogy mely bázisra gondolunk, akkor csak a $[A]$ jelölést használjuk.

7.7. állítás. Legyenek V, Z, W ugyanazon test feletti véges dimenziós vektorterek, $B \in L(V, Z)$ és $A \in L(Z, W)$. Rögzítsük az

$$\{e_1, \dots, e_m\} \subseteq V, \quad \{z_1, \dots, z_r\} \subseteq Z, \quad \{f_1, \dots, f_n\} \subseteq W$$

bázisokat. Jelölje $C = A \circ B$ a kompozíció lineáris operátort.

Ekkor C mátrixa az A és B mátrixszának szorzata. Formálisabban:

$$[C]_{\{e_1, \dots, e_m\}} = [A]_{\{z_1, \dots, z_r\}} \cdot [B]_{\{f_1, \dots, f_n\}}.$$

Bizonyítás: Először is ellenőrizzük, hogy értelmes-e az állításban felírt formula. $[A]$ mérete $n \times r$, $[B]$ mérete $r \times m$. Így az $[A] \cdot [B]$ szorzat értelmes és a szorzás eredménye egy $n \times m$ mátrix. A $[C]$ szintén egy $n \times m$ méretű mátrix, így a bal és a jobboldal összehasonlítása is értelmes.

Már csak azt kell meggondolni, hogy a baloldali mátrix minden eleme azonos a jobboldali szorzatmátrix megfelelő elemével. Az α, β, γ szimbólumokkal jelöljük az $[A], [B], [C]$ mátrixok megfelelő elemeit. Ez azt jelenti, hogy

$$Be_j = \sum_{k=1}^r \beta_{k,j} z_k, \quad Az_k = \sum_{i=1}^n \alpha_{i,k} f_i, \quad Ce_j = \sum_{i=1}^n \gamma_{i,j} f_i.$$

Így azt kapjuk, hogy minden $j = 1, \dots, m$ mellett

$$\begin{aligned} \sum_{i=1}^n \gamma_{i,j} f_i &= Ce_j = A(Be_j) = \sum_{k=1}^r \beta_{k,j} Az_k = \sum_{k=1}^r \beta_{k,j} \left(\sum_{i=1}^n \alpha_{i,k} f_i \right) = \\ &= \sum_{k=1}^r \sum_{i=1}^n \beta_{k,j} \alpha_{i,k} f_i = \sum_{i=1}^n \sum_{k=1}^r \alpha_{i,k} \beta_{k,j} f_i = \sum_{i=1}^n \left(\sum_{k=1}^r \alpha_{i,k} \beta_{k,j} \right) f_i. \end{aligned}$$

No de, az $\{f_1, \dots, f_n\}$ rendszer lineárisan független, tehát a lineáris burkában minden elem előállítása egyértelmű, ami éppen azt jelenti, hogy minden $i = 1, \dots, n$ és minden $j = 1, \dots, m$ mellett

$$[C]_{i,j} = \gamma_{i,j} = \sum_{k=1}^r \alpha_{i,k} \beta_{k,j} = ([A] \cdot [B])_{i,j}. \quad \square$$

Azért, hogy teljes legyen az analógia a lineáris operátorok kompozíciója és a mátrixok szorzása műveletek közt, a lineáris operátorok $A \circ B$ kompozíóját is $A \cdot B$ módon, vagy még egyszerűbben AB módon jelöljük. A jelölést a szóhasználat is követi:

7.8. definíció (lineáris transzformációk szorzata). A lineáris operátorok kompozíció műveletét *szorzatnak* is mondjuk.

Ilyen módon ha A és B két olyan lineáris operátor, amelynek kompozíciója – azaz szorzata – értelmes, akkor a mátrixaik szorzata is értelmes, továbbá

$$[AB] = [A][B]. \quad (7.3)$$

7.9. definíció (lineáris operátor hatványai). A lineáris transzformációk a bevezetett szorzás művelettel egységelemes gyűrűt alkotnak, ahol az egységelem az $I : V \rightarrow V$ az identitás operáció.

Legyen $A \in L(V)$ egy lineáris transzformáció. Ennek 0-dik hatványát definiálja $A^0 = I$. Ha valamely n nem negatív egész mellett A^n már definiált, akkor legyen $A^{n+1} = AA^n$.

Világos, hogy ha n, m nem negatív egészek, akkor $A^{n+m} = A^n \cdot A^m$. Az is nyilvánvaló, hogy az I identitás operációinak akármelyik bázisban felírt mátrixa ugyanaz a mátrix, mégpedig az identitás mátrix (ahol minden elem zérus, kivétel a diagonális elemek, amelyek értéke 1.)

7.10. állítás. Legyen $A \in L(V)$ egy lineáris transzformáció, és legyen rögzítve a tér egy bázisa, amelyben felírjuk A transzformáció $[A]$ mátrixát. Ekkor $[A^n] = [A]^n$, továbbá az A transzformáció pontosan akkor izomorfizmus, ha az $[A]$ mátrixa reguláris és $[A^{-1}] = [A]^{-1}$.

Bizonyítás: A (7.3) azonosságot alkalmazva $B = A$ mellett nyilvánvaló indukcióval kapjuk az $[A^n] = [A]^n$ azonosságot.

Világos, hogy A pontosan akkor izomorfizmus, ha létezik $B \in L(V)$, amelyre $AB = I$. No de, ez ekvivalens avval, hogy $[A][B] = [I]$, ami pedig a szükséges és elegendő feltétele $[A]$ mátrix invertálhatóságának. Ekkor $[A]^{-1} = [B] = [A^{-1}]$. \square

A (7.2) és a (7.3) azonosságok szerint egy lineáris operátorhoz hozzárendelni annak valamely bázisban felírt mátrixát egy olyan bijekció, ami tartja a gyűrű műveleteket. Az $L(V)$ és a $\mathbb{F}^{\dim(V) \times \dim(V)}$ tehát olyan egységelemes (nem kommutatív) gyűrűk, amelyek között van a gyűrű műveleteket tartó bijekció (gyűrű-izomorfizmus).

8. fejezet

Általános bázistranszformáció

EGY VETKOR KOORDINÁTÁI függnek a bázis megválasztásától. Az elemi bázistranszformáció arra szolgál hogy felírjuk az új bázisban a vektor koordinátáit, amikor az új bázis és a régi bázis csak egyetlen vektorban különbözik. A fejezetben általánosabban oldjuk meg a problémát, mikor az új bázis és a régi bázis elemei tetszőlegesen különbözhetnek.

Vegyük fel a V vektortér egy-egy bázisát. Nevezzük az $\{e_1, \dots, e_n\}$ bázist régi bázisnak, és nevezzük az $\{f_1, \dots, f_n\}$ bázis elemeit új bázisnak. A kérdések a következők:

1. Ha ismerjük egy $x \in V$ vektor régi bázisra vonatkozó koordinátáit, akkor hogyan számítható ugyanennek a vektornak az új bázisban felírt koordináta-vektora?
2. Ha ismerjük egy $A \in L(V)$ lineáris transzformációkat a régi bázisban felírt mátrixát, akkor hogyan számolható ki az A -nak valamely új bázisban felírt mátrixa?
3. Ha ismerjük egy $A \in L(V, W)$ lineáris operációt a régi bázispáron felírt mátrixát, akkor hogyan számítható A -nak az valamely új bázispáron felírt mátrixa?

8.1. definíció. Legyenek az $\{e_1, \dots, e_n\}$ és $\{f_1, \dots, f_n\}$ bázisok rögzítve. Tekintsük azt a B lineáris transzformációt, amelyre $Be_j = f_j$ teljesül minden $j = 1, \dots, n$ mellett. Ezt a lineáris transzformációt nevezzük az $\{e_1, \dots, e_n\}$ bázisról az $\{f_1, \dots, f_n\}$ bázisra való áttérés transzformációjának.

Világos, hogy ha B az áttérés transzformáció, akkor ennek a régi bázisban felírt $[B]$ mátrixa egy olyan $n \times n$ méretű mátrix, amelynek j -edik oszlopja a $Be_j = f_j$ vektornak a régi bázisban felírt koordinátája. Mivel B szürjektív, ergo injektív is, tehát B egy izomorfizmus, ennek megfelelően a $[B]$ mátrix reguláris mátrix, azaz létezik $[B]^{-1}$ inverze.¹

8.1. Vektor koordinátái az új bázisban

8.2. állítás. Legyen az $\{e_1, \dots, e_n\}$ a régi bázis, és $\{f_1, \dots, f_n\}$ az új bázis. Jelölje B a régi bázisról az új bázisra való áttérést. Ekkor tetszőleges $x \in V$ vektor mellett

$$[x]_{\text{új}} = [B]_{\text{régi}}^{-1} \cdot [x]_{\text{régi}}.$$

Bizonyítás: Legyen $[x]_{\text{új}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$. Ez azt jelenti, hogy $x = \sum_{j=1}^n \alpha_j f_j$. Így felírva az x vektornak a régi bázisra vonatkozó koordináta-vektorát

$$[x]_{\text{régi}} = \left[\sum_{j=1}^n \alpha_j f_j \right]_{\text{régi}} = \sum_{j=1}^n \alpha_j [f_j]_{\text{régi}} = \sum_{j=1}^n \alpha_j [B]_{\text{régi}}^j = [B]_{\text{régi}} \cdot [x]_{\text{új}}.$$

A kívánt azonosságot kapjuk, ha balról szorozzuk minden két oldalt $[B]_{\text{régi}}^{-1}$ inverz mátrixszal. \square

¹Könnyen látható, hogy B -nek a régi és az új bázisban felírt mátrixa azonos, de ez később egyszerű következményként is adódik.

8.2. Lineáris operátorok mátrixa új bázispárban

8.3. állítás. Legyen $A \in L(V, W)$ lineáris operáció és tegyük fel, hogy ismerjük A mátrixát az $\{e_1, \dots, e_m\} \subseteq V$ és az $\{f_1, \dots, f_n\} \subseteq W$ régi bázisokban. Legyenek az $\{v_1, \dots, v_m\} \subseteq V$ és a $\{w_1, \dots, w_n\} \subseteq W$ az új bázisok. Definiálja $B \in L(V)$ a V térről áttérés lineáris transzformációját és $D \in L(W)$ a W térről áttérés transzformációját. Ekkor

$$[A]_{ij} = [D]_{\text{régi}}^{-1} [A]_{\text{régi}} [B]_{\text{régi}}.$$

Bizonyítás: Azt mutatjuk meg, hogy a baloldali mátrix és a jobboldali szorzat mátrix megfelelő oszlopai megegyeznek. A j -edik oszlopra:

$$\begin{aligned} [A]_{ij}^j &= [A]_{\{v_1, \dots, v_m\}}^j = [Av_j]_{\{w_1, \dots, w_n\}} \\ &= [D]_{\{f_1, \dots, f_n\}}^{-1} \cdot [Av_j]_{\{f_1, \dots, f_n\}} = [D]_{\{f_1, \dots, f_n\}}^{-1} \cdot [A]_{\{e_1, \dots, e_m\}} \cdot [v_j]_{\{e_1, \dots, e_m\}} \\ &= [D]_{\{f_1, \dots, f_n\}}^{-1} \cdot [A]_{\{e_1, \dots, e_m\}} \cdot [B]_{\{e_1, \dots, e_m\}}^j \\ &= [D]_{\text{régi}}^{-1} [A]_{\text{régi}} [B]_{\text{régi}}^j = [[D]_{\text{régi}}^{-1} [A]_{\text{régi}} [B]_{\text{régi}}]^j. \end{aligned}$$

Közben használtuk a mátrix szorzás művelet asszociativitását, és azt a tényt, hogy tetszőleges $[E], [F]$ mátrixok mellett $[[E][F]]^j = [E]([F]^j)$, azaz az $[[E][F]]$ szorzat j -edik oszlopa azonos az $[E]$ mátrixnak az $[F]$ mátrix j -edik oszlopával való szorzatával. \square

8.3. Lineáris transzformáció mátrixa az új bázisban

8.4. állítás. Legyen $A \in L(V)$ egy lineáris transzformáció. Ekkor

$$[A]_{ij} = [B]_{\text{régi}}^{-1} [A]_{\text{régi}} [B]_{\text{régi}}.$$

Bizonyítás: A lineáris operátorokra vonatkozó állítás speciális esete, mikor $W = V$ és $D = B$. \square

A fenti állítás minden $A \in L(V)$ lineáris transzformáció mellett igaz. Speciálisan, ha alkalmazzuk az $A = B$ esetre, akkor látjuk, hogy az áttérés mátrixa azonos, ha az új bázisban, vagy a régi bázisban írjuk fel.

8.5. állítás. Legyen B az áttérés lineáris transzformáció. Ekkor

$$[B]_{\text{régi}} = [B]_{ij} \quad \text{és persze} \quad [B^{-1}]_{\text{régi}} = [B^{-1}]_{ij}.$$

9. fejezet

Invariáns alterek

9.1. definíció. Legyen $A \in L(V)$ egy lineáris transzformációja a V vektortérnek, és $M \subseteq V$ a térfelület egy altere. Az mondjuk, hogy M egy *invariáns altere* V -nek, ha minden $v \in M$ mellett $Av \in M$ is teljesül.

Ha nem világos, hogy mely lineáris transzformációról van szó, akkor azt is mondjuk, hogy az M alter A -invariáns, vagy azt, hogy az M alter invariáns az A transzformációra nézve.

Úgy is fogalmazhatnánk, hogy az M alter akkor invariáns alter, ha az A transzformációnak az M -re való $A|_M$ megszorítása az M alter egy lineáris transzformációja, azaz

$$A|_M \in L(M).$$

Nyilvánvaló példa invariáns alterekre a teljes V vektortér és a $\{0\}$ alter. Tetszőleges A lineáris transzformáció mellett ker A és $\text{Im } A$ minden invariáns alterek. Ugyanis, ha $u \in \ker A$, akkor $A(Au) = A0 = 0$, tehát $Au \in \ker A$. Az $\text{Im } A$ alter esete még egyszerűbb: A térfelület minden elemének képe $\text{Im } A$ -ban van, speciálisan persze ez $\text{Im } A$ elemeire is igaz.

A célunk, hogy a teret előállítsuk lehetőleg minél alacsonyabb dimenziós terek direkt összegeként.

9.2. definíció (legszűkebb invariáns alter). Legyen $A \in L(V)$ egy lineáris transzformáció. Világos, hogy akárhány A -ra nézve invariáns alter metszete is A -invariáns alter, így egy $H \subseteq V$ halmazt tartalmazó legszűkebb A -invariáns alter a H halmazt tartalmazó A -invariáns alterek közös része. Formálisan

$$\text{lin}(H; A) = \bigcap_{\substack{H \subseteq N \\ N \text{ alter} \\ N \text{ invariáns}}} N.$$

A legérdekesebb eset, amikor H egy elemű halmaz. A $H = \{v\}$ esetben a kissé nehézkes $\text{lin}(\{v\}; A)$ helyett egyszerűbben $\text{lin}(v; A)$ -t írunk.

9.3. állítás. Egy $A \in L(V)$ lineáris transzformációra és egy $v \in V$ vektorra

$$\text{lin}(v; A) = \text{lin}\{v, Av, A^2v, \dots\}.$$

Bizonyítás: Mivel $\text{lin}(v; A)$ egy v -t tartalmazó A -invariáns alter, ezért $\{v, Av, \dots, A^k v, \dots\} \subseteq \text{lin}(v; A)$, amiatt

$$\text{lin}\{v, Av, A^2v, \dots\} \subseteq \text{lin}(v; A).$$

Most vegyük észre, hogy $\text{lin}\{v, Av, A^2v, \dots\}$ is egy v -t tartalmazó A -invariáns alter és $\text{lin}(v; A)$ ilyenek között a legszűkebb, ezért

$$\text{lin}(v; A) \subseteq \text{lin}\{v, Av, A^2v, \dots\}.$$

□

9.4. lemma. Tegyük fel, hogy egy $A \in L(V)$ lineáris transzformációra és egy $v \in V$ vektorra a

$$\{v, Av, \dots, A^k v\}$$

lineárisan összefüggő ($k \geq 1$). Ekkor minden $n \geq k$ mellett

$$A^n v \in \text{lin}\{v, Av, \dots, A^{k-1} v\}.$$

Bizonyítás: Ha $v = 0$ akkor az állítás nyilvánvaló. Ha $v \neq 0$, akkor a $\{v\}$ rendszer lineárisan független. Legyen t a legkisebb olyan szám, hogy a rendszerhez $A^t v$ -t hozzávéve az lineárisan összefüggővé válik. Ilyen módon tehát

$$\{v, Av, \dots, A^{t-1}v\} \text{ lineárisan független} \quad (\dagger)$$

de

$$\{v, Av, \dots, A^{t-1}v, A^t v\} \text{ lineárisan összefüggő.} \quad (\ddagger)$$

Világos, hogy $1 \leq t \leq k$. Jelölje $N = \text{lin} \{v, Av, \dots, A^{t-1}v\}$.

Most indukcióval megmutatjuk, hogy minden $m \geq 0$ számra

$$A^{t+m}v \in N.$$

Ha $m = 0$, akkor $A^t v \in N$, hiszen a fenti (\ddagger) lineárisan összefüggő rendszerre, az egyik elem előáll mint az előző elemek lineáris kombinációja. No de, ez elem csak az utolsó lehet, hiszen az utolsó elem nélküli (\dagger) rendszer még lineárisan független.

Most tegyük fel, hogy $A^{t+m}v \in N$ és lássuk be, hogy $A^{t+m+1}v \in N$ is teljesül. Ezek szerint $A^{t+m}v = \sum_{j=0}^{t-1} \alpha_j A^j v$ alakú. Erre A -t alkalmazva

$$A^{t+m+1}v = \sum_{j=0}^{t-1} \alpha_j A^{j+1}v = \sum_{j=1}^t \alpha_{j-1} A^j v = \left(\sum_{j=1}^{t-1} \alpha_{j-1} A^j v \right) + \alpha_{t-1} A^t v \in N + N = N. \quad \square$$

9.5. állítás. Legyen V egy véges dimenziós vektortér. $A \in L(V)$ lineáris transzformáció, és $v \in V$ egy $v \neq 0$ vektor. Ekkor létezik egyetlen $1 \leq k \leq \dim(V)$ szám, amelyre

$$\{v, Av, \dots, A^{k-1}v\} \text{ lineárisan független} \quad (\dagger)$$

de

$$\{v, Av, \dots, A^{k-1}v, A^k v\} \text{ lineárisan összefüggő.} \quad (\ddagger)$$

A fenti (\dagger) rendszer bázisa $\text{lin}(v; A)$ -nak.

Bizonyítás: Először a k szám konstrukciója: Mivel $v \neq 0$, ezért $\{v\}$ lineárisan független. Ha $\{v, Av\}$ lineárisan összefüggő, akkor $k = 1$ és készen vagyunk. Egyébként tekintsük a $\{v, Av, A^2v\}$ rendszert. Ha ez lineárisan összefüggő, akkor $k = 2$ -vel készen vagyunk. Ha ez lineárisan független, akkor tekintsük a $\{v, Av, A^2v, A^3v\}$ vektorrendszeret. Ha lineárisan összefüggő, akkor $k = 3$ és készen vagyunk, ha lineárisan független, akkor folytatjuk egy újabb elem hozzá vételével. Az eljárás előbb-utóbb megáll a vektorrendszer összefüggővé válásával, hiszen a Steinitz-lemma szerint egy véges dimenziós vektortérben legfeljebb $\dim(V)$ elemszámú lineárisan független vektorrendszer van.

Mivel $\text{lin}(v; A)$ egy v -t tartalmazó A -invariáns altér, ezért

$$\{v, Av, \dots, A^{k-1}v\} \subseteq \text{lin}(v; A).$$

Az előző lemma szerint

$$\{v, Av, A^2v, \dots\} \subseteq \text{lin}\{v, Av, \dots, A^{k-1}v\}$$

ezért

$$\text{lin}(v; A) = \text{lin}\{v, Av, A^2v, \dots\} \subseteq \text{lin}\{v, Av, \dots, A^{k-1}v\} \subseteq \text{lin}(v; A).$$

Ez azt jelenti, hogy a (\dagger) vektorrendszer egy lineárisan független generátorrendszer a $\text{lin}(v; A)$ térféknél, tehát valóban bázisa is. \square

A $\text{lin}(v; A)$ tehát a fenti k -ra egy k -dimenziós altér, amelynek bázisa $\{A^{k-1}v, \dots, Av, v\}$. A játék kedvéért írjuk fel az $A|_{\text{lin}(v; A)}$ transzformáció mátrixát ebben a bázisban:

$$\begin{array}{c|ccccc} & A^k v & A^{k-1} v & A^{k-2} v & \dots & Av \\ \hline A^{k-1} v & \alpha_{k-1} & 1 & 0 & \dots & 0 \\ A^{k-2} v & \alpha_{k-2} & 0 & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ Av & \alpha_1 & 0 & 0 & \dots & 1 \\ v & \alpha_0 & 0 & 0 & \dots & 0 \end{array}, \text{ ahol } A^k v = \sum_{j=0}^{k-1} \alpha_j A^j v.$$

9.1. Transzformációk sajátértéke

A célunk, hogy a lehető legalacsonyabb dimenziós invariáns altereket találjunk. Így persze az egydimenziós invariáns alterek a legérdekesebbek. Ez vezet a sajátérték fogalmához.

9.6. definíció (sajátérték, sajátvektor, spektrum). Legyen $A \in L(V)$. Azt mondjuk, hogy a $\lambda \in \mathbb{F}$ szám az A lineáris transzformáció *sajátértéke*, ha létezik $v \in V, v \neq 0$ vektor, amelyre

$$Av = \lambda v$$

teljesül.

Ha λ egy sajátértéke A -nak, akkor az összes olyan nem zérus v vektort, amelyre $Av = \lambda v$ fennáll az A transzformáció λ sajátértékéhez tartozó *sajátvektorainak* nevezzük.

Egy A lineáris transzformáció összes sajátértékeinek halmazát az A *spektrumának* nevezzük, és $\sigma(A)$ -val jelöljük.

A spektrum tehát az \mathbb{F} test azon részhalmaza, amelyre

$$\sigma(A) = \{\lambda \in \mathbb{F} : \exists v \in V, v \neq 0, Av = \lambda v\}$$

teljesül. Világos, hogy ha λ egy sajátértéke A -nak, akkor a λ -hoz tartozó sajátvektorok halmaza éppen a

$$\ker(A - \lambda I)$$

altér nem zérus elemei. Emiatt a fenti alteret a λ sajátértékhez tartozó *sajátaltérnek* mondjuk.

Vegyük észre, hogy $v \in V$ vektor pontosan akkor sajátvektora A -nak, ha $\dim(\text{lin}(v; A)) = 1$. Hasonlóan az is könnyű, hogy λ pontosan akkor sajátértéke A -nak, ha az $A - \lambda I$ transzformáció szinguláris.

10. fejezet

Transzformációk polinomjai

10.1. definíció (transzformáció polinomja). Legyen V az \mathbb{F} test feletti vektortér, $A \in L(V)$ egy lineáris transzformáció, és legyen $p \in \mathbb{F}[t]$ egy az \mathbb{F} test feletti polinom, amely $p(t) = \sum_{j=0}^n \alpha_j t^j$ alakú. Definiálja $p(A) \in L(V)$ az A transzformáció p polinomját

$$p(A) = \sum_{j=0}^n \alpha_j A^j.$$

10.2. állítás (Számolási szabályok). Legyen $p, q \in \mathbb{F}[t]$ polinomok, $A \in L(V)$ lineáris transzformáció.

1. Ha $r = p + q$, akkor $r(A) = p(A) + q(A)$.
2. Ha $r = pq$, akkor $r(A) = p(A)q(A)$.

Bizonyítás: Az első állítás azért teljesül, mert $(\alpha A^k + \beta A^l) = (\alpha + \beta) A^k$.

A második állításhoz azt vegyük észre, hogy $A^k A^l = A^{k+l}$. Így, amikor összegyűjtjük, hogy a $q(A)p(A)$ kompozícióban, mi lesz A^j együtthatója, akkor azt kapjuk, hogy

$$\left(\sum_{\substack{k,l \\ k+l=j}} \alpha_k \beta_l \right) A^j = \left(\sum_{k=0}^j \alpha_k \beta_{j-k} \right) A^j.$$

Vegyük észre, hogy az 1.19. definíció szerint az r szorzat polinomban is a fenti zárójelben lévő szám a t^j tag együtthatója. \square

Tudjuk, hogy lineáris transzformációk szorzata függ azok sorrendjétől. Ugyanúgy mint mátrixokra, két lineáris transzformációt *kommutálónak* mondunk, ha szorzatuk a szorzás sorrendjétől független. Például az I identitás minden transzformációval kommutál. Azt is láttuk, hogy ha $AB = I$, akkor $BA = I$ is fennáll, azaz A és B kommutálnak. Nagyon fontos, de nyilvánvaló következménye a fenti számolási szabálynak, hogy ha p, q tetszőleges polinomok, akkor a $p(A)$ és $q(A)$ egymással kommutáló lineáris transzformációk lesznek, hiszen az $\mathbb{F}[t]$ egy kommutatív gyűrű, azért ha $r = pq$, akkor $qp = r$, ergo

$$p(A)q(A) = r(A) = q(A)p(A).$$

Tehát meggondoltuk, hogy

10.3. állítás. Lineáris transzformáció polinomjai egymással kommutálnak.

Polinomok segítségével sok-sok új invariáns alteret kapunk.

10.4. állítás. Tetszőleges p polinomra és $A \in L(V)$ lineáris transzformáció mellett ker $p(A)$ és $\text{Im } p(A)$ is invariáns alterek.

Bizonyítás: Legyen először $v \in \ker p(A)$. Ekkor, mivel A és $p(A)$ kommutálnak

$$p(A)Av = Ap(A)v = A0 = 0,$$

ami pont azt jelenti, hogy $Av \in \ker p(A)$.

Legyen most $v \in \text{Im } p(A)$, azaz $v = p(A)x$ valamely $x \in V$ mellett. Ekkor, mivel A és $p(A)$ kommutálnak

$$Av = Ap(A)x = p(A)(Ax),$$

amiből már látszik, hogy $Av \in \text{Im } p(A)$. \square

Speciálisan ez igaz az $t - \lambda$ polinomra is, amikor λ egy sajátértéke A -nak. Tehát a λ sajátértékhez tartozó $\ker(A - \lambda I)$ sajátáltér egy legalább egydimenziós invariáns altere A -nak.

10.5. definíció. Azt mondjuk, hogy az A lineáris transzformáció a p polinom gyöke, ha $p(A) = 0$.

Mielőtt tovább lépünk érdemes visszagondolnunk arra, hogy egy test feletti polinomgyűrű egy főideálgyűrű. Láttuk ugyanis – 1.24. állítás –, hogy minden nem csupán a $\{0\}$ elemet tartalmazó ideálnak van egyetlen legkisebb fokú és normált eleme, ami egyben az ideál generáló eleme is.

10.1. Kis minimálpolinom

10.6. állítás. Legyen $A \in L(V)$ a V véges dimenziós vektortér egy lineáris transzformációja, és legyen $v \in V$ egy rögzített vektor. Tekintsük az $\mathbb{F}[t]$ polinomgyűrű következő részhalmazát:

$$J_{A,v} = \{p \in \mathbb{F}[t] : p(A)v = 0\}.$$

Ez a halmaz egy ideálja $\mathbb{F}[t]$ -nek, amelynek van legfeljebb $\dim(V)$ -ed fokú, de nem konstans zérus polinomja.

Bizonyítás: Ha $p, q \in J_{A,v}$, akkor $(p + q)(A)v = p(A)v + q(A)v = 0 + 0 = 0$, azaz $p + q \in J_{A,v}$. Ha most $p \in J_{A,v}$ és h egy tetszőleges polinom, akkor $(hp)(A)v = h(A)p(A)v = h(A)0 = 0$, azaz $hp \in J_{A,v}$. Megmutattuk tehát, hogy $J_{A,v}$ egy ideálja a polinomgyűrűnek.

Jelölje $n = \dim(V)$ és tekintsük az $n + 1$ elemű $\{v, Av, \dots, A^n v\}$ vektorrendszeret. Mivel a Steinitz-lemma szerint $n + 1$ vektor egy n -dimenziós vektortérben lineárisan összefüggő, ezért van $\alpha_0, \dots, \alpha_n \in \mathbb{F}$ nem minden zérus szám, hogy $\sum_{j=0}^n \alpha_j A^j v = 0$. Ha tehát p jelöli a $p(t) = \sum_{j=0}^n \alpha_j t^j$ polinomot, akkor $p(A)v = 0$, azaz $p \in J_{A,v}$ és $-\infty < \deg p \leq n$. \square

Ha például $v = 0$, akkor $J_{A,0} = \mathbb{F}[t]$, azaz minden polinom az ideálhoz tartozik. Ha $v \neq 0$, akkor a $J_{A,v}$ ideálnak nincs nulladfokú polinomja, hiszen $p(t) = c$, ($c \neq 0$) mellett $p(A)v = (cI)v = cv \neq 0$.

10.7. definíció. Legyen $A \in L(V)$ a V véges dimenziós vektortér egy lineáris transzformációja, és legyen $v \in V$ egy rögzített vektor. Láttuk, hogy

$$J_{A,v} = \{p \in \mathbb{F}[t] : p(A)v = 0\}$$

az $\mathbb{F}[t]$ gyűrű egy ideálja, amely nem csak a konstans zéruspolinomot tartalmazza. Tudjuk, hogy az $\mathbb{F}[t]$ polinomgyűrű egy főideál-gyűrű, van tehát egyetlen normált polinomja $J_{A,v}$ -nek, amely generálja $J_{A,v}$ -t. Ez a $J_{A,v}$ legkisebb fokú, normált polinomja. Ezt a polinomot nevezzük az A transzformáció, *kis minimálpolinomjának*. a v pontban.

10.8. állítás. Az n polinom pontosan akkor az A transzformáció vektorhoz tartozó kis minimálpolinomja, ha

1. n normált polinom,
2. $n(A)v = 0$,
3. ha $p \in \mathbb{F}[t]$, $p \neq 0$, amelyre $p(A)v = 0$, akkor $n|p$.

A kis minimálpolinom foka legfeljebb a tér dimenziója.

Bizonyítás: Definíció szerint n azon p polinomok közül, amelyek normáltak, és $p(A)v = 0$, a legalacsonyabb fokú. Láttuk hogy ilyen polinom csak egy van, és erre a polinomra

$$J_{A,v} = J(n) = \{hn : h \in \mathbb{F}[t]\}.$$

Ezt kellett belátni. \square

10.9. állítás. Legyen $v \neq 0$ és tegyük fel, hogy $p(A)v = 0$ valamely normált, irreducibilis p polinomra. Ekkor p a v vektorhoz tartozó kis minimálpolinom.

Bizonyítás: Ha maga $\deg p = 1$, akkor készen vagyunk, hiszen nem zérus vektornak kis minimálpolinoma legalább elsőfokú. Ha $\deg p > 1$ és p nem egyezne az n kis minimálpolinommal, akkor $\deg n < \deg p$ lenne, és mivel n generálja a $J_{A,v}$ ideált, ezért $p = nh$ alakú lenne, ahol h is legalább elsőfokú. Így p két legalább elsőfokú polinom szorzata, azaz reducibilis lenne. \square

A következő állítás módszert ad a kis minimálpolinom meghatározására, amely minden használható.

10.10. állítás. Legyen $A \in L(V)$ és $v \neq 0$. Feltéve, hogy V egy véges dimenziós vektortér, létezik $1 \leq k \leq \dim(V)$, hogy $\{v, Av, \dots, A^{k-1}v\}$ lineáris független, de $\{v, Av, \dots, A^{k-1}v, A^k v\}$ lineárisan összefüggő. Léteznek tehát $\alpha_0, \dots, \alpha_{k-1} \in \mathbb{F}$ számok, amelyekre

$$A^k v + \sum_{j=0}^{k-1} \alpha_j A^j v = 0.$$

Ekkor

1. Az A operátor ezen v vektorhoz tartozó kis minimálpolinomja

$$n(t) = t^k + \alpha_{k-1}t^{k-1} + \dots + \alpha_1 t + \alpha_0.$$

2. A v -hez tartozó kis minimálpolinom foka megegyezik a v -t tartalmazó legszűkebb invariáns altér dimenziójával, azaz

$$\deg n = \dim(\text{lin}(v; A)).$$

3. Sőt, minden $x \in \text{lin}(v; A)$ mellett $n(A)x = 0$ is fennáll, azaz

$$n(A)|_{\text{lin}(v; A)} = 0|_{\text{lin}(v; A)}.$$

Bizonyítás: Láttuk, hogy $\{v, Av, \dots, A^{k-1}v\}$ bázisa $\text{lin}(v; A)$ altérnek. Világos, hogy $A^k v \in \text{lin}(v; A)$, emiatt a kívánt előállítás valóban létezik. Ezt átrendezve kapjuk, hogy n valóban olyan normált polinom, amelyre $n(A)v = 0$ fennáll. No de, k -nál alacsonyabb fokú ilyen polinom csak a konstans zérus polinom lehet, hiszen $\{v, Av, \dots, A^{k-1}v\}$ lineárisan független. Azt láttuk tehát, hogy n a legalacsonyabb fokú nem zérus eleme $J_{A,v}$ -nek, tehát n generálja a főideált.

A fentiek szerint $\dim(\text{lin}(v; A)) = k = \deg n$.

Azt már láttuk, hogy $n(A)v = 0$. A $\text{lin}(v; A)$ egy bázisa $\{v, Av, \dots, A^{k-1}v\}$ és minden bázis elemre

$$n(A)(A^j v) = A^j(n(A)v) = A^j(0) = 0,$$

hiszen $n(A)$ és A^j kommutáló transzformációk. Azt kaptuk tehát, hogy az $n(A)$ transzformáció a $\text{lin}(v; A)$ egy bázisát zérusra viszi, ezért maga $n(A)$ a $\text{lin}(v; A)$ altér zérus transzformációja. \square

Illusztráció

Legyen $\{u_1, u_2, u_3\}$ bázisa a \mathbb{C} komplex számtest feletti V vektortérnek. Az $A \in L(V)$ lineáris transzformációra

$$A(\alpha u_1 + \beta u_2 + \gamma u_3) = (-3\beta - 2\gamma)u_1 + (\alpha - \beta + \gamma)u_2 + (-\alpha + 3\beta + \gamma)u_3.$$

Írjuk fel az u_3 vektorhoz tartozó kis minimálpolinomot.

Világos, hogy A mátrixa a megadott bázisban

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

Az előző állításban megadott algoritmust használjuk:

$$\begin{array}{c|ccc} & Au_3 & A^2u_3 & A^3u_3 \\ \hline u_3 & -2 & -5 & -6 \\ & 1 & -2 & 3 \\ \hline & 1 & 6 & 5 \\ \hline & \delta & -2 & 3 \end{array} \quad \begin{array}{c|cc} & A^2u_3 & A^3u_3 \\ \hline Au_3 & -9 & 0 \\ u_3 & -2 & 3 \\ \hline & 8 & 2 \\ \hline & \delta & 0 \end{array} \quad \begin{array}{c|c} & A^3u_3 \\ \hline A^2u_3 & 0 \\ Au_3 & 3 \\ u_3 & 2 \\ \hline & \end{array}$$

Azt kapjuk tehát, hogy $\{u_3, Au_3, A^2u_3\}$ még lineárisan független, de $A^3u_3 - 3Au_3 - 2u_3 = 0$. Ezek szerint az u_3 vektorhoz tartozó kis minimálpolinom: $n(t) = t^3 - 3t - 2$, amivel a feladatot meg is oldottuk.

Érdemes még látni, hogy mit mond az előző állítás a fenti transzformációról. Mivel $\text{lin}(u_3; A)$ egy 3 dimenziós altér a V három dimenziós vektortérnek, ezért $\text{lin}(u_3; A) = V$, így $A^3 - 3A - 2I = 0$ nem csak u_3 vektort, de V minden vektorát zérusra viszi, azaz $n(A) = A^3 - 3A - 2I = 0$. Alkalmazhatjuk a 84. lapon felírt mátrixot erre a transzformációra. Így azt kapjuk, hogy A -nak az $\{A^2u_3, Au_3, u_3\}$ bázisban felírt mátrixa a következő.

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

Az érdemes látni, hogy ebben a mátrixban az első oszlopban találjuk a v_3 vektorhoz tartozó kis minimálpolinom együtthatóinak ellentétjét a $k-1$ -ediktől kezdve sorban lefelé (a k adik a legmagasabb fokú együttható mindig 1.) A mátrix szuper diagonálisában csupa 1-es szerepel, és az összes többi tag zérus. Az ilyen mátrixoknak központi szerepe van, amint azt később látni fogjuk.

Első alkalmazásként megmutatjuk, hogy minden normált polinom lehet kis minimálpolinom.

10.11. állítás. Legyen $p(t) \in \mathbb{F}[t]$ egy tetszőleges normált k -ad fokú polinom, ahol $k \geq 1$. Ekkor tetszőleges V éppen k dimenziós vektortér tetszőleges $v \in V$, $v \neq 0$ vektorához van olyan $A \in L(V)$ lineáris transzformáció, amelyre a v -hez tartozó kis minimálpolinom éppen p .

Bizonyítás: Jelölje $p(t) = t^k + \alpha_{k-1}t^{k-1} + \dots + \alpha_1t + \alpha_0$, $v_1 = v$. A $v_1 \neq 0$ vektor mint egy elemű vektorrendszer lineárisan független, ezért kiegészíthető a tér $\{v_k, \dots, v_2, v_1\}$ bázisává. Legyen $A \in L(V)$ az a lineáris transzformáció, amelynek ebben a bázisban felírt mátrixa¹

$$\begin{pmatrix} -\alpha_{k-1} & 1 & 0 & \dots & 0 \\ -\alpha_{k-2} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\alpha_1 & 0 & 0 & \ddots & 1 \\ -\alpha_0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Ez persze azonos avval, mintha A -t a $\{v_k, \dots, v_1\}$ bázison definiálnánk az alábbi módon:

$$A(v_j) = v_{j+1}, j = 1, \dots, k-1 \text{ esetén, és } A(v_k) = \sum_{j=0}^{k-1} -\alpha_j v_{j+1}.$$

Ekkor minden $j = 0, \dots, k-1$ mellett $v_{j+1} = A^j v_1$, ezért $\{v_1, Av_1, \dots, A^{k-1}v_1\} = \{v_1, v_2, \dots, v_k\}$ lineárisan független rendszer. A mátrix első oszlopa szerint

$$A^k v_1 = A(A^{k-1} v_1) = A(v_k) = \sum_{j=0}^{k-1} -\alpha_j v_{j+1} = \sum_{j=0}^{k-1} -\alpha_j A^j v_1,$$

ami pont azt jelenti, hogy $p(A)v_1 = 0$, ezért valóban $p(t)$ a v_1 -hez tartozó kis minimálpolinom.

□

¹Az első oszlopban sorban lefelé $p(t)$ együtthatóinak ellentetje, a szuper diagonálisban 1, mindenütt másutt zérus.

Vegyük észre, hogy a fenti bizonyításban V éppen k dimenziós, így csak $\text{lin}(v_1; A) = V$ lehetséges. Ez azt jelenti, hogy $p(A)$ nem csak v_1 -et, de V valamennyi más vektorát is zérusra viszi, ergo $p(A) = 0$, sőt p a legalacsonyabb fokú polinom, amire ez teljesül, hiszen $\{v_1, Av_1, \dots, A^{k-1}v_1\}$ vektorrendszer még lineárisan független.

Ez vezet a minimálpolinom fogalmához.

10.2. Minimálpolinom

10.12. állítás. Legyen $A \in L(V)$ a V véges dimenziós vektortér egy lineáris transzformációja. Tekintsük az $\mathbb{F}[t]$ polinomgyűrű következő részhalmazát.

$$J_A = \{p \in \mathbb{F}[t] : p(A) = 0\}$$

Ez a halmaz egy ideálja $\mathbb{F}[t]$ -nek, amelynek van legfeljebb $(\dim(V))^2$ -ed fokú, de nem konstans zérus polinomja.

Bizonyítás: Ha $p, q \in J_A$, akkor $(p + q)(A) = p(A) + q(A) = 0 + 0 = 0$, azaz $p + q \in J_A$. Ha most $p \in J_A$ és h egy tetszőleges polinom, akkor $(hp)(A) = h(A)p(A) = h(A)0 = 0$, azaz $hp \in J_A$. Megmutattuk tehát, hogy J_A egy ideálja a polinomgyűrűnek.

Jelölje $n = \dim(V)$ és tekintsük az $n^2 + 1$ elemű $\{I, A, \dots, A^{n^2}\}$ rendszerét az $L(V)$ vektortérnek. Mivel a Steinitz-lemma szerint $n^2 + 1$ vektor egy n^2 -dimenziós vektortérben lineárisan összefüggő, ezért van $\alpha_0, \dots, \alpha_{n^2} \in \mathbb{F}$ nem mind zérus szám, hogy $\sum_{j=0}^{n^2} \alpha_j A^j = 0$. Ha tehát p jelöli a $p(t) = \sum_{j=0}^{n^2} \alpha_j t^j$ polinomot, akkor $p(A) = 0$, azaz $p \in J_A$ és $-\infty < \deg p \leq n^2$. \square

10.13. definíció (minimálpolinom). Legyen $A \in L(V)$ a V véges dimenziós vektortér egy lineáris transzformációja. Láttuk, hogy

$$J_A = \{p \in \mathbb{F}[t] : p(A) = 0\}$$

az $\mathbb{F}[t]$ gyűrű egy ideálja, amely nem csak a konstans zéruspolinomot tartalmazza. Tudjuk, hogy az $\mathbb{F}[t]$ polinomgyűrű egy föideál-gyűrű, van tehát egyetlen normált polinomja J_A -nek, amely generálja J_A . Ez a J_A legkisebb fokú, normált polinomja, amit A transzformáció *minimálpolinomjának* nevezünk.

Ha $V \neq \{0\}$, ergo $\dim(V) \geq 1$, akkor a J_A ideálnak nincs nullafokú polinomja, hiszen $p(t) = c$, ($c \neq 0$) mellett $p(A) = cI \neq 0$, tehát legalább egydimenziós tér egy lineáris transzformációjának a minimálpolinomja legalább elsőfokú.

10.14. állítás. Az $m \in \mathbb{F}[t]$ polinom pontosan akkor az $A \in L(V)$ transzformáció minimálpolinomja, ha

1. m normált polinom,
2. $m(A) = 0$,
3. ha $p \in \mathbb{F}[t]$, $p \neq 0$, amelyre $p(A) = 0$, akkor $m|p$.

A minimálpolinom fokszámára: $\deg m \leq \dim(V)^2$.

Bizonyítás: Definíció szerint m azon p polinomok közül, amelyek normáltak, és $p(A) = 0$, a legalacsonyabb fokú. Láttuk hogy ilyen polinom csak egy van, és erre a polinomra

$$J_A = J(m) = \{hm : h \in \mathbb{F}[t]\}.$$

Ezt kellett beláttni. \square

10.15. állítás. Legyen V legalább egydimenziós, és tegyük fel, hogy $p(A) = 0$ valamely normált, irreducibilis p polinomra. Ekkor p az A minimálpolinomja.

Bizonyítás: Ha maga $\deg p = 1$, akkor készen vagyunk, hiszen a minimálpolinom legalább elsőfokú. Ha $\deg p > 1$ és p nem egyezne az m minimálpolinommal, akkor $\deg m < \deg p$ lenne, és mivel m generálja a J_A ideált, ezért $p = mh$ alakú lenne, ahol h is legalább elsőfokú. Így p két legalább elsőfokú polinom szorzata, azaz reducibilis lenne. \square

Ha összevetjük a fenti állítást az irreducibilis kis minimálpolinomra vonatkozó hasonló állítással, akkor azt kapjuk, hogy olyan normált, irreducibilis p polinomra, amelyre $p(A) = 0$, tetszőleges $v \neq 0$ vektorhoz tartozó p_v kis minimál polinomok azonosak és ez a polinom egyben a transzformáció minimálpolinomja is.

Ezek szerint, ha a minimálpolinom irreducibilis, akkor fokszáma legfeljebb a tér dimenziója. Látni fogjuk később, hogy ez minden minimálpolinomra is igaz. A következő állítás módszert ad a minimálpolinom meghatározására.

10.16. állítás. Legyen $A \in L(V)$, ahol $\dim(V) \geq 1$. Mivel $\dim L(V) = \dim(V)^2$ egy véges dimenziós vektortér, ezért létezik $1 \leq k \leq \dim(V)^2$, hogy $\{I, A, \dots, A^{k-1}\}$ lineárisan független, de $\{I, A, \dots, A^{k-1}, A^k\}$ lineárisan összefüggő. Ekkor léteznek $\alpha_0, \dots, \alpha_{k-1} \in \mathbb{F}$ számok, amelyekre $A^k = \sum_{j=0}^{k-1} \alpha_j A^j$, így az A transzformáció minimálpolinomja a következő alakú:

$$m(t) = t^k - \alpha_{k-1}t^{k-1} - \dots - \alpha_1t - \alpha_0.$$

Bizonyítás: Mivel a $k+1$ elemű rendszer lineárisan összefüggő és a k elemű rendszer lineárisan független, ezért a kívánt előállítás valóban létezik. Ezt átrendezve kapjuk, hogy m valóban olyan normált polinom, amelyre $m(A) = 0$ fennáll. No de, k -nál alacsonyabb fokú ilyen polinom csak a konstans zérus polinom lehet, hiszen $\{I, A, \dots, A^{k-1}\}$ lineárisan független. Azt láttuk tehát, hogy m a legalcsonyabb fokú nem zérus eleme J_A -nek, tehát az m polinom generálja a főideált. \square

A minimálpolinom algoritmikus meghatározásához a kis minimálpolinomok is használhatók.

10.17. állítás. Legyen az $A \in L(V)$ lineáris transzformáció minimálpolinomja m . Tegyük fel, hogy az $\{e_1, \dots, e_n\}$ egy bázisa V -nek, és p_1, \dots, p_n rendre a bázis elemekhez tartozó kis minimálpolinomok. Ekkor a p_1, \dots, p_n polinomok legkisebb közös többszöröse az m minimálpolinom.

Bizonyítás: Tetszőleges $v \in V$ mellett, ha p_v a kis minimálpolinom, akkor $p_v|m$, hiszen $m(A)$ a konstans zérus transzformáció. Emiatt m egy közös többszöröse a p_j kis minimálpolinomoknak. Most tegyük fel, hogy egy p polinom többszöröse a p_j kis minimálpolinomoknak. Világos, hogy minden j mellett

$$p(A)e_j = h_j(A)(p_j(A)e_j) = h_j(A)0 = 0.$$

Mivel egy lineáris transzformáció egy bázison egyértelműen meghatározott, ezért $p(A) = 0$. Ekkor persze $m|p$, azaz m valóban a kis minimálpolinomok legkisebb közös többszöröse. \square

Egy másik bizonyítás: Mivel a $p(A)$ lineáris transzformáció a bázison egyértelműen meghatározott, ezért $p \in J_A$ pontosan akkor teljesül, ha $p \in J_{A,e_i}$ a bázis minden e_i elemére. Így

$$\cap_{i=1}^n J(p_i) = \cap_{i=1}^n J_{A,e_i} = J_A = J(m).$$

No de, a baloldali ideál is főideál, aminek a normált generáló eleme – az 1.26. állítás szerint – a p_1, \dots, p_n polinomok legkisebb közös többszöröse. Mivel egy főidálnak csak egy normált generáló eleme van, ezért a legkisebb közös többszörös azonos az m minimálpolinommal. \square

Illusztráció

Újra vegyük elő a 89. oldalon vett mátrixot.

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

Láttuk, hogy az u_3 bázisvektorhoz tartozó kis minimálpolinom $t^3 - 3t - 2$. Az is világos, hogy az u_3 vektort tartalmazó legszűkebb invariáns altér dimenziója a kis minimálpolinom foka, ergo 3. Így $\text{lin}(u_3; A) = V$, tehát $A^3 - 3A - 2I = 0$. Azt is megmondottuk, hogy 3-nál alacsonyabb fokú normált polinom nincs, ami u_3 vektort zérusra viszi, azaz $m(t) = t^3 - 3t - 2$ az A minimálpolinomja.

Persze itt a szerencsénk, hogy u_3 tartalmazó legszűkebb invariáns altér az egész vektortér. Ha ezt nem használjuk, akkor a 10.17. állítás eliminációs algoritmusát alkalmazhatjuk a bázis elemekhez tartozó

kis minimálpolinomok felírására, majd ezek legkisebb közös többszöröseként kapjuk a transzformáció minimálpolinomját.

$$\begin{array}{c|cc} & Au_1 & A^2u_1 \\ \hline u_1 & 0 & -1 \\ & \boxed{1} & -2 \\ & -1 & 2 \\ \hline & \delta & -2 \end{array} \quad \begin{array}{c|cc} & A^2u_1 \\ \hline u_1 & -1 \\ Au_1 & -2 \\ & 0 \end{array}$$

$\{u_1, Au_1\}$ lineárisan független, de
 $A^2u_1 + 2Au_1 + u_1 = 0$,
ezért $p_1(t) = t^2 + 2t + 1 = (t + 1)^2$.

Tehát itt nem volt olyan szerencsénk mint u_3 esetében. Most $\text{lin}(u_1; A)$ egy 2 dimenziós altér V -ben. Csak $x \in \text{lin}(u_1; A)$ vektoraira tudjuk, hogy $A^2x + 2Ax + x = 0$. Folytatnunk kell a második bázisvektorra:

$$\begin{array}{c|cc} & Au_2 & A^2u_2 \\ \hline u_2 & -3 & -3 \\ & -1 & 1 \\ & \boxed{3} & 3 \\ \hline & \delta & 1 \end{array} \quad \begin{array}{c|cc} & A^2u_2 \\ \hline u_2 & 0 \\ Au_2 & 2 \\ & 1 \end{array}$$

$\{u_2, Au_2\}$ lineárisan független, de
 $A^2u_2 - Au_2 - 2u_2 = 0$,
ezért $p_2(t) = t^2 - t - 2 = (t + 1)(t - 2)$.

Ugyanúgy mint az imént az u_2 vektort tartalmazó legszűkebb invariáns altér csak két dimenziós. Látni fogjuk, hogy a minimálpolinom foka legfeljebb a tér dimenziója. Ha ezt most felhasználjuk, akkor nem is kell tovább lépnünk, hiszen egy olyan legfeljebb 3-ad fokú normált polinomot keresünk, amely minden $p_1(t) = (t + 1)^2$ polinomnak, minden $p_2(t) = (t + 1)(t - 2)$ polinomnak többszöröse. Persze ilyen polinom csak egy van az

$$m(t) = (t + 1)^2(t - 2).$$

Egy kicsi szerencsénk még volt a fenti eljárásban, hiszen már a második bázisvektor után kiderült a minimálpolinom. Általában csak azt tudjuk, hogy a kis minimálpolinomok legkisebb közös többszöröse a minimálpolinom. Ez a mi konkrét esetünkben azt jelenti, hogy az $m(t)$ minimálpolinom a $p_1(t) = (t + 1)^2$, a $p_2(t) = (t + 1)(t - 2)$ és a $p_3(t) = t^3 - 3t - 2 = (t + 1)^2(t - 2)$ polinomok legkisebb közös többszöröse, ami most nyilvánvaló módon teljesül.

Hasonlóan mint a kis minimálpolinom esetében, most is igaz, hogy minden normált polinom egy minimálpolinom. Ez a 10.11. állítás kiterjesztése. Csak azt kell még észrevennünk $A|_{\text{lin}(v; A)}$ transzformáció minimálpolinoma éppen a v -hez tartozó kis minimálpolinom.

10.18. állítás. Legyen $m(t) \in \mathbb{F}[t]$ egy tetszőleges normált k -ad fokú polinom, ahol $k \geq 1$. Ekkor tetszőleges V éppen k dimenziós vektortér tetszőleges $v \in V$, $v \neq 0$ vektorához van olyan $A \in L(V)$ lineáris transzformáció, amelyre a v -hez tartozó kis minimálpolinom éppen m . Így $\text{lin}(v; A)$ is k -dimenziós, ezért m a minimálpolinomja is az A -nak.

10.3. Sajátvektorok és diagonalizálhatóság

10.19. állítás. Tegyük fel, hogy p legalább elsőfokú osztója az A transzformáció minimálpolinomjának. Ekkor $p(A)$ szinguláris.

Bizonyítás: Jelölje m a minimálpolinomot, és $m = pq$. Így $\deg q < \deg m$. Persze $m(A) = p(A)q(A)$, így ha $p(A)$ reguláris lenne, akkor

$$q(A) = p(A)^{-1}m(A) = 0$$

Ebből persze $m|q$, így $\deg m \leq \deg q$, ami ellentmondás. \square

10.20. állítás (sajátérték és minimálpolinom). Legyen $A \in L(V)$ lineáris transzformációja a V véges dimenziós vektortérnek, és $\lambda \in \mathbb{F}$ egy szám, valamint m az A minimálpolinomja. Az alábbi feltevések ekvivalensek:

1. λ sajátértéke A -nak,
2. $\ker(A - \lambda I) \neq \{0\}$,
3. $A - \lambda I$ szinguláris,
4. létezik $v \in V$, amelyre a v -hez tartozó kis minimálpolinomja A -nak $p_v(t) = t - \lambda$.

$$5. t - \lambda | m(t),$$

$$6. m(\lambda) = 0.$$

Bizonyítás: Az első három pont ekvivalenciája nyilvánvaló, majd a $3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 6 \Rightarrow 5 \Rightarrow 3$ utat érdemes követni. Az utolsó lépéshez mutattuk meg az előző állítást. \square

Az egyik legfontosabb definícióhoz érkeztünk.

10.21. definíció (diagonalizálható transzformáció). Az $A \in L(V)$ lineáris transzformációt *diagonalizálhatónak* mondjuk, ha van a térek olyan bázisa, amelyben a transzformáció mátrixa diagonális alakú, azaz a fődiagonálisán kívül minden elem zérus.

Az $[A]$ négyzetes mátrix tehát pontosan akkor diagonális alakú, ha minden $i \neq j$ mellett $[A]_{i,j} = 0$. Ez azt jelenti hogy a j -edik bázis elem képének a nem j -edik koordinátája zérus, azaz az e_j bázisvektorra $Ae_j = \lambda e_j$ áll fenn, valamely $\lambda \in \mathbb{F}$ számmal. Ez éppen azt jelenti, hogy az e_j bázisvektor egy sajátvektor. Nyilvánvaló tehát, hogy diagonalizálhatóság szükséges és elegendő módon megragadható a sajátvektor fogalmának segítségével.

10.22. állítás (diagonalizálhatóság). Az $A \in L(V)$ egy lineáris transzformáció pontosan akkor diagonalizálható, ha van térek csupa sajátvektorokból álló bázisa.

Ebben az esetben a transzformáció $\{v_1, \dots, v_n\}$ sajátvektorokban, mint bázisban, felírt mátrixának a j -edik diagonális eleme, éppen az a λ_j sajátértéke A -nak, amelyre $Av_j = \lambda_j v_j$.

Mivel a $\begin{pmatrix} -t & 1 \\ -1 & -t \end{pmatrix}$ valós test feletti mátrix minden valós t mellett reguláris, ezért a $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ mátrix egy nagyon egyszerű példa olyan mátrixra, amelynek spektruma üres, így persze nem diagonalizálható.

Csak a játék kedvéért, ha az $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ mátrixot tekintjük \mathbb{R} felett, azt kapjuk, hogy ez sem diagonalizálható. Van ugyan egyetlen sajátértéke $\lambda = 0$, de az ehhez a sajátértékhez tartozó ker A sajátaltér egydimenziós, emiatt nincs a térben két lineárisan független sajátvektor.

A pozitív példa kedvéért nézzük az $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ mátrixot. Mind az \mathbb{R} , mind a \mathbb{C} test felett diagonalizálható, hiszen $\sigma(A) = \{0, 2\}$, továbbá a $\lambda = 0$ -hoz tartozó sajátaltérre $\ker A = \text{lin} \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$, míg a $\lambda = 2$ sajátértékhez tartozó sajátaltérre $\ker(A - 2I) = \text{lin} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$. Világos, hogy a $B = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ vektorrendszer egy sajátvektorokból álló bázisa a két dimenziós vektortérnek. A fenti bázisban a transzformáció mátrixa $\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$. Mivel az erre a bázisra való áttérés mátrixa $B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, ezért az új bázisra való áttérést formuláját használva

$$[B]^{-1}[A]_{\text{régi}}[B] = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} = [A]_{\text{új}}.$$

Ezt úgy fejezzük ki, hogy az $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ izomorfizmus diagonalizálja az $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ mátrixot.

10.23. állítás. Különböző sajátértékekhez tartozó sajátvektorok rendszere lineárisan független.

Formálisabban: Legyen $\{\lambda_1, \dots, \lambda_s\} \subseteq \sigma(A)$ a sajátérékek páronként különböző elemekből álló rendszere, és legyen $\{v_1, \dots, v_s\} \subseteq V$ sajátvektorok olyan rendszere, amelyre $v_j \in \ker(A - \lambda_j I)$ minden $j = 1, \dots, s$ mellett. Ekkor a sajátvektorok rendszere lineárisan független.

Bizonyítás: A sajátvektorok száma, azaz s szerinti indukció. Ha $s = 1$, akkor készen is vagyunk, hiszen egy sajátvektor egy nem zérus vektor.

Most tegyük fel, hogy igaz az állítás sajátvektorok s -nél kevesebb elemből álló rendszerére, és lássuk be sajátvektorok olyan s elemű rendszerére, amelyek különböző sajátértékhez tartoznak. Az indukciós

feltevés szerint tehát $\{v_1, \dots, v_{s-1}\}$ lineárisan független. Ha $\{v_1, \dots, v_{s-1}, v_s\}$ lineárisan összefüggő lenne, akkor valamely $\alpha_1, \dots, \alpha_{s-1}$ számokkal

$$v_s = \sum_{j=1}^{s-1} \alpha_j v_j$$

lenne. No de,

$$\sum_{j=1}^{s-1} \lambda_s \alpha_j v_j = \lambda_s v_s = Av_s = \sum_{j=1}^{s-1} \alpha_j Av_j = \sum_{j=1}^{s-1} \alpha_j \lambda_j v_j,$$

ami az első $s-1$ elem lineárisan függetlensége szerint csak úgy lehetséges, ha minden $j = 1, \dots, s-1$ mellett $\lambda_s \alpha_j = \alpha_j \lambda_j$, ergo $\alpha_j (\lambda_s - \lambda_j) = 0$. Mivel itt különböző sajátértékekkel van szó, ezért minden szóba jövő j mellett $\alpha_j = 0$. Ebből $v_s = 0$ következik, ami ellentmond annak, hogy v_s egy sajátvektor. \square

Egy n -dimenziós térben n -elemű lineárisan független rendszer generátorrendszer is, így azonnali következmény a diagonalizálhatóság egy elegendő feltétele:

10.24. állítás (diagonalizálhatóság elegendő feltétele). *Tegyük fel, hogy az $A \in L(V)$ lineáris transzformációnak annyi különböző sajátértéke van, mint a V vektortér dimenziója. Ekkor A diagonalizálható.*

Az identitás mátrix példája mutatja, hogy a feltétel elegendő, de nem szükséges. Mivel n -dimenziós térben legfeljebb n elemű linárisan független rendszer van, ezért kapjuk, hogy a spektrumnak több eleme nem lehet, mint a tér dimenziója:

10.25. állítás. *Legyen $A \in L(V)$ lineáris transzformáció. Ekkor A -nak legfeljebb $\dim(V)$ darab különböző sajátértéke lehet.*

10.26. definíció (geometriai multiplicitás). Ha $A \in L(V)$ egy lineáris transzformáció, és $\lambda \in \sigma(A)$ annak egy sajátértéke, akkor az $A - \lambda I$ sajátaltér defektusát, tehát a $\ker(A - \lambda I)$ altér dimenzióját, a λ sajátérték geometriai multiplicitásának mondjuk.

10.27. állítás. *Legyen $A \in L(V)$ transzformáció, és $\{\lambda_1, \dots, \lambda_s\} \subseteq \sigma(A)$, a spektrum különböző elemei. Jelölje $M_j = \ker(A - \lambda_j I)$. Ekkor értelmes az $M_1 \oplus \dots \oplus M_s$ direkt összeg.*

Bizonyítás: Megmutatjuk, hogy a $\sum_{j=1}^s M_j$ összegben minden elem előállítása egyértelmű. Ehhez elég azt belátni, hogy $\sum_{j=1}^s v_j = 0$, $v_j \in M_j$ csak úgy lehetséges, ha minden $j = 1, \dots, s$ mellett $v_j = 0$. Tegyük fel tehát, hogy valamely $v_j \in M_j$ vektorokra

$$\sum_{j=1}^s v_j = 0.$$

Ez egy olyan lineáris kombináció, amelyben minden vektor együtthatója 1. Emiatt a $\{v_1, \dots, v_s\}$ vektorrendszer nem zérus vektorai is összefüggő rendszert alkotnak, feltéve hogy vannak ilyenek. No de, egy $v_j \in M_j$ vektor ha nem zérus, akkor egy sajátvektor. Tehát ha a vektorrendszerben lenne nem zérus elem, akkor talánk különböző sajátértékekhez tartozó sajátvektorok egy lineárisan összefüggő rendszerét, ami a 10.23. állítás szerint nem lehetséges. \square

Meggondoltuk tehát, hogy ha páronként különböző $\{\lambda_1, \dots, \lambda_s\} \subseteq \sigma(A)$ sajátértékekből indulunk ki, és egyesítjük a $\ker(A - \lambda_1 I)$ sajátalerek egy-egy bázisait, akkor az így összetett vektorrendszer az

$$\ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_s I)$$

altér egy – sajátvektorokból álló – bázisa. Rögzítsük is ezt a fontos gondolatot, amelyet a feladatok megoldása során sokszor használjuk majd.

10.28. állítás. *A különböző sajátértékekhez tartozó sajátalerek bázisainak egyesítése a sajátalerek direkt összegének egy bázisa.*

10.29. állítás (diagonalizálhatóság). *Legyen $A \in L(V)$ lineáris transzformáció. Jelölje $\{\lambda_1, \dots, \lambda_s\} = \sigma(A)$ az A spektrumát, azaz valamennyi különböző sajátértékét. Az alábbi feltevések ekvivalensek.*

1. *Az A sajátértékei geometriai multiplicitásának összege $\dim(V)$,*

2. $\ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_s I) = V$,
3. *Minden vektor előáll mint sajátvektorok összege,*
4. *Az A diagonalizálható lineáris transzformáció.*

Bizonyítás: Mivel a $\ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_s I) \subseteq V$ tartalmazás minden fennáll, ezért a 2. feltétel ekvivalens avval, hogy

$$\sum_{j=1}^s \dim(\ker(A - \lambda_j I)) = \dim(\ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_s I)) = \dim(V),$$

ami éppen a geometriai multiplicitásra vonatkozó feltétel. Így az első két feltevés ekvivalenciáját megértettük.

Mivel a $\ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_s I) \subseteq V$ tartalmazás minden fennáll, ezért a 2. feltétel ekvivalens avval, hogy

$$\sum_{j=1}^s \ker(A - \lambda_j I) = V,$$

ami éppen a 3. feltétel. Így a 2. és a 3. feltételek ekvivalenciáját is megértettük.

Ha 3., ezért 2. fennáll, akkor az egyes $\ker(A - \lambda_j I)$ sajátalerek bázisait egyesítve a V tér egy sajátvektorokból álló bázisát kapjuk, ergo az A diagonalizálható transzformáció. Megfordítva, ha A diagonalizálható, akkor van sajátvektorokból álló bázisa, így minden vektor előáll mint sajátvektorok összege. Evvel a 3. és a 4. feltételek ekvivalenciáját is igazoltuk. \square

11. fejezet

Transzformációk redukálása

11.1. állítás. Tekintsünk egy $A \in L(V)$ lineáris transzformációt, amelynek minimálpolinomja $m \in \mathbb{F}[t]$. Tegyük fel, hogy m előáll mint a p, q egymással relatív prím, normált polinomok

$$m = pq$$

szorzata. Ekkor a tér szétesik a $\ker p(A)$ és a $\ker q(A)$ invariáns alterei direkt összegére, azaz

$$\ker p(A) \oplus \ker q(A) = V.$$

Ha $A_1 = A|_{\ker p(A)}$ és $A_2 = A|_{\ker q(A)}$, akkor A_1 minimálpolinomja p és A_2 minimálpolinomja q .

Bizonyítás: Mivel p és q relatív prímek, ezért a Bezout-azonosság szerint van $f, g \in \mathbb{F}[t]$ polinom, amelyekre

$$fp + gq = 1.$$

Emiatt persze minden $x \in V$ mellett

$$f(A)p(A)x + g(A)q(A)x = Ix = x. \quad (\dagger)$$

Ha $x \in \ker p(A) \cap \ker q(A)$, akkor $p(A)x = 0 = q(A)x$, tehát $x = 0$, ami azt jelenti, hogy $\ker p(A)$ és $\ker q(A)$ diszjunkt alterek.

Vegyük észre, hogy a $f(A)p(A)x \in \ker q(A)$, és hasonlóan $g(A)q(A)x \in \ker p(A)$. Ez azt jelenti, hogy $\ker q(A) + \ker p(A) = V$ azonosság is fennáll. Megmutattuk tehát, hogy V előáll mint a $\ker p(A)$ és a $\ker q(A)$ alterek direkt összege.

Világos, hogy minden $u \in \ker p(A)$ mellett $p(A_1)u = p(A)u = 0$.

Ha h egy másik olyan polinom, amelyre $h(A_1) = 0 \in L(\ker p(A))$, akkor mivel a direkt összegre vonatkozó állítást már igazoltuk

$$(hq)(A)x = h(A)q(A)(x_1 + x_2) = q(A)h(A)x_1 + h(A)q(A)x_2 = 0 + 0 = 0,$$

ahol $x = x_1 + x_2$, $x_1 \in \ker p(A)$ és $x_2 \in \ker q(A)$. Látjuk tehát, hogy az A transzformáció a hq polinomnak is gyöke, emiatt $p|m$, ezért

$$p|hq$$

is fennáll. Most újra használjuk, hogy a p és a q polinomok relatív prímek, így azt kapjuk, hogy $p|h$. Megmutattuk, hogy a J_{A_1} ideált a p normált polinom generálja, ami éppen azt jelenti, hogy p az A_1 transzformáció minimálpolinomja. Az A_2 minimálpolinomja q állítás igazolása a fentivel analóg. \square

11.2. állítás. Tekintsünk egy $A \in L(V)$ lineáris transzformációt, amelynek minimálpolinomja $m \in \mathbb{F}[t]$. Tegyük fel, hogy m előáll mint a páronként relatív prím normált polinomok

$$m = p_1 p_2 \cdots p_n$$

szorzata. Jelölje minden $i = 1, \dots, n$ mellett $V_i = \ker p_i(A)$ invariáns alteret, és $A_i = A|_{V_i}$ megszorítást. Világos, hogy $A_i \in L(V_i)$. Ekkor

1. $V = V_1 \oplus \cdots \oplus V_n$;
2. p_i az A_i transzformáció minimálpolinomja minden szóba jövő $i = 1, \dots, n$ mellett.

Bizonyítás: A polinomok n száma szerinti teljes indukcióval igazolunk. Az $n = 1$ eset triviális, de $n = 2$ éppen az előző állítás.

Most tegyük fel, hogy az állítás n -nél kevesebb polinom szorzatára igaz, és lássuk be n -re. Feltehető tehát, hogy $n \geq 3$. Legyen $p = p_1 \cdots p_{n-1}$. Világos, hogy p és p_n relatív prímek, hiszen ha d irreducibilis osztója p -nek és p_n -nek, akkor d prím tulajdonsága szerint $d \mid p_i$ valamely $i < n$ -re, tehát $d \mid p_i$ és $d \mid p_n$. A feltevés szerint ilyen csak a konstans polinom lehetséges, ami valóban igazolja, hogy p és p_n relatív prímek. Persze

$$m = pp_n.$$

Alkalmazhatjuk tehát az előző állítást, azaz

$$V = \ker p(A) \oplus V_n,$$

továbbá p a minimálpolinomja az $A \mid \ker p(A)$ -nak és persze p_n minimálpolinomja A_n -nek.

Alkalmazzuk most az indukciós feltevést a $\ker p(A)$ vektortérre. Ott az $A_{\mid \ker p(A)}$ lineáris transzformáció p minimálpolinomja előáll mint $n - 1$ páronként relatív prím polinom szorzata:

$$p = p_1 \cdots p_{n-1}.$$

Világos tehát, hogy

1. $\ker p(A) = V_1 \oplus \cdots \oplus V_{n-1}$ és
2. p_i az A_i minimálpolinomja minden $i = 1, \dots, n - 1$ mellett.

Teljesül tehát a V_1, \dots, V_{n-1}, V_n alterekre, hogy minden diszjunkt – az itt adott sorrendben – az előzőek összegétől, és a Minkowski-összegük az egész V vektortér. Az 5.6. állítás szerint tehát $V = V_1 \oplus \cdots \oplus V_n$. \square

11.1. Minimálpolinom és diagonalizálhatóság

Alkalom nyílik, hogy a lineáris transzformáció diagonalizálhatóságát karakterizáló a 10.29. állítást tovább bővítsük, a minimálpolinom szerepének hangsúlyozásával.

11.3. állítás (diagonalizálhatóság). Legyen $A \in L(V)$ lineáris transzformáció. Jelölje $\{\lambda_1, \dots, \lambda_s\} = \sigma(A)$ az A spektrumát, azaz valamennyi különböző sajátértékét. Az alábbi feltevések ekvivalensek.

1. Az A sajátértékei geometriai multiplicitásának összege $\dim(V)$;
2. $\ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_s I) = V$;
3. minden vektor előáll mint sajátvektorok összege;
4. Az A diagonalizálható lineáris transzformáció;
5. Az A transzformáció minimálpolinomja

$$m(t) = \prod_{j=1}^s (t - \lambda_j).$$

Bizonyítás: Ha az A transzformáció diagonalizálható, akkor a diagonálisában a sajátértékei vannak. Írjuk fel tehát a $\prod_{j=1}^s (A - \lambda_j I)$ transzformáció mátrixát abban a bázisban, amelyben A mátrixa is diagonális. Az eredmény egy diagonális mátrix, és ha felírjuk ezt mint az $[A - \lambda_j I]$ mátrixok szorzatát, akkor minden diagonális pozíció az egyik szorzó mátrixban zérus, ergo a szorzat mátrix is a zéró mátrix. A fenti polinomnak tehát az A transzformáció gyöke. Mivel minden sajátérték a minimálpolinom gyöke, ezért a fenti polinom a legalacsonyabb fokú normált polinom, amelynek gyöke A .¹

¹Egy másik érv: $t - \lambda_j$ a λ_j sajátértékhez tartozó sajátvektor kis minimálpolinomja. Mivel a sajátvektorok egy bázist alkotnak, ezért ezen polinomok legkisebb közös többszöröse a minimálpolinom. Persze $\dim(V)$ darab elsőfokú normált polinom legkisebb közös többszöröse, ezek közül a különbözők szorzata.

Megfordítva, legyen $p_j(t) = t - \lambda_j$ minden $j = 1, \dots, s$. Ekkor $m = p_1 \cdots p_s$ páronként relatív prím, normált polinomok szorzata, ezért az éppen igazolt 11.2. állítás szerint

$$V = \ker p_1(A) \oplus \cdots \oplus \ker p_s(A) = \ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_s I)$$

Ezt kellett belátni. \square

11.2. Redukálás: az általános eset

Mivel minden polinom előáll mint néhány irreducibilis polinom szorzata, ezért 11.2. állítás így is fogalmazható.

11.4. állítás. *Legyen $A \in L(V)$ lineáris transzformáció minimálpolinomja m . Tudjuk, hogy m egyértelműen áll elő*

$$m = p_1^{m_1} p_2^{m_2} \cdots p_r^{m_r},$$

alakban, ahol p_1, \dots, p_r egymástól páronként különböző normált, irreducibilis polinomok. Jelölje $V_i = \ker p_i^{m_i}(A)$ jelölje $A_i = A|_{V_i}$ minden $i = 1, \dots, r$ mellett. Ekkor

1. $V = V_1 \oplus \cdots \oplus V_r$ és,
2. A_i minimálpolinomja $p_i^{m_i}$ minden $i = 1, \dots, r$ mellett.

A konklúzió tehát az, hogy elég olyan transzformációkkal foglalkoznunk, amelyek minimálpolinomja egy irreducibilis polinom valamely egész kitevős hatványa. Az algebra alaptétele szerint a komplex számtest feletti irreducibilis polinom csak elsőfokú polinom lehet. Abban a speciális esetben tehát, amikor a \mathbb{C} komplex számtest feletti vektortereket vizsgálunk, ez azt jelenti, hogy elég ha olyan $A \in L(V)$ transzformációval foglalkozunk, amelynek m minimálpolinomjára

$$m(t) = (t - \alpha)^n$$

teljesül valamely $\alpha \in \mathbb{C}$ komplex szám és $n \in \mathbb{N}$ egész mellett. Ilyen A transzformációra, ha B jelöli a $B = A - \alpha I$ lineáris transzformációt, akkor a B olyan, hogy valamely egész kitevős hatványa a konstans zérus transzformáció. Ez vezet majd a *nilpotens* fogalmához. Ha felírjuk valamely bázisban egy ilyen B nilpotens transzformáció mátrixát, akkor A mátrixa is könnyen adódik B mátrixából. Ehhez csak az $\alpha[I]$ mátrixot kell $[B]$ -hez adni, ami praktikusan nem jelent többet, mint hogy a $[B]$ diagonális elemeket kell az α komplex számmal megemelni. A fenti gondolatot alapul a *Jordan-normálak* fogalma.

12. fejezet

Redukálás irreducibilis minimálpolinom esetén

A LEGEGBSZERŰBB ESET, mikor a minimálpolinom elsőfokú irreducibilis polinomok szorzata.

12.1. állítás. Legyen $A \in L(V)$ egy lineáris transzformációja a V véges dimenziós vektortérnek, az $m \in \mathbb{F}[t]$ egy k -adfokú, irreducibilis polinom, amelyre $m(A) = 0$. Ekkor

1. minden $v \neq 0$ mellett $\dim \text{lin}(v; A) = k$;
2. minden $v \in V$ vektor és minden $K \subseteq V$ invariáns altér mellett $\text{lin}(v; A) \cap K = \{0\}$ vagy $\text{lin}(v; A) \subseteq K$;
3. a V altér nulla dimenziós, vagy k dimenziós, vagy k dimenziós A invariáns alterek direkt összege. Pontosabban, ha $\dim(V) > 0$, akkor létezik $r \geq 1$ szám, és léteznek v_1, \dots, v_r vektorok, amelyekre

$$\text{lin}(v_1; A) \oplus \dots \oplus \text{lin}(v_r; A) = V.$$

Bizonyítás (1.) Legyen adott $v \neq 0$ vektor mellett p_v a v -hez tartozó kis minimálpolinom. Mivel $m(A)v = 0$, ezért $p_v|m$. No de, m irreducibilis, ezért $m = p_v$. Ekkor viszont

$$k = \deg m = \deg p_v = \dim \text{lin}(v; A).$$

Bizonyítás (2.) Ha $v = 0$, akkor az állítás nyilvánvaló. A továbbiakban emiatt $v \neq 0$. Most tegyük fel, hogy $x \in \text{lin}(v; A) \cap K$ és $x \neq 0$. Ekkor

$$\text{lin}(x; A) \subseteq \text{lin}(v; A) \cap K \subseteq \text{lin}(v; A)$$

No de, a bal- és a jobboldali altér azonos dimenziós alterek, emiatt fent mindenütt egyenlőség van. Speciálisan $\text{lin}(v; A) = \text{lin}(v; A) \cap K$, ami éppen azt jelenti, hogy $\text{lin}(v; A) \subseteq K$. \square

Bizonyítás (3.) Először is gondoljuk meg, hogy invariáns alterek Minkowski–összege is invariáns altér marad. Ha V nem tartalmaz nem zérus vektort, akkor $\dim(V) = 0$.

Ha $v_1 \in V$ egy nem zérus vektor, akkor jelölje $V_1 = \text{lin}(v_1; A)$. Ha $V = V_1$, akkor V egy k -dimenziós vektortér.

Ha $V \neq V_1$, akkor van $v_2 \in V \setminus V_1$, $v_2 \neq 0$ vektor. Mivel V_1 egy invariáns altér, ezért a már igazolt állítás szerint bevezetve a $V_2 = \text{lin}(v_2; A)$ jelölést $V_2 \cap V_1 = \{0\}$. Értelmes tehát venni e két invariáns altér direkt összegét. Ha $V = V_1 \oplus V_2$, akkor V előállt két k dimenziós invariáns alterének direkt összegeként.

Ha $V \neq V_1 \oplus V_2$, akkor van $v_3 \in V \setminus (V_1 \oplus V_2)$, $v_3 \neq 0$ vektor. Mivel $V_1 \oplus V_2$ egy invariáns altér, ezért a már igazolt állítás szerint bevezetve a $V_3 = \text{lin}(v_3; A)$ jelölést $V_3 \cap (V_1 \oplus V_2) = \{0\}$. Értelmes tehát venni a három invariáns altér direkt összegét, hiszen a V_1, V_2, V_3 alterek ebben a sorrendben véve olyanok, hogy minden diszjunkt az előzőek összegétől. Ha $V = V_1 \oplus V_2 \oplus V_3$, akkor V előállt három k dimenziós invariáns alterének direkt összegeként.

Ha $V \neq V_1 \oplus \dots \oplus V_t$, valamely $t \geq 2$ mellett, akkor van $v_{t+1} \in V \setminus (V_1 \oplus \dots \oplus V_t)$, $v_{t+1} \neq 0$ vektor. Mivel $V_1 \oplus \dots \oplus V_t$ egy invariáns altér, ezért a már igazolt állítás szerint bevezetve a $V_{t+1} = \text{lin}(v_{t+1}; A)$ jelölést $V_{t+1} \cap (V_1 \oplus \dots \oplus V_t) = \{0\}$. Értelmes tehát venni ezen $t+1$ invariáns altér direkt összegét, hiszen a $V_1, V_2, V_3, \dots, V_t, V_{t+1}$ alterek ebben a sorrendben véve olyanok, hogy minden diszjunkt az előzőek összegétől. Ha $V = V_1 \oplus \dots \oplus V_{t+1}$, akkor V előállt $t+1$ darab k dimenziós invariáns alterének direkt összegeként.

Az eljárás előbb utóbb a V vektortér véges dimenziós volta miatt megáll. \square

Következmény

Világos, hogy $\{A^{k-1}v_1, A^{k-2}v_1, \dots, Av_1, v_1\}$ bázisa a $\text{lin}(v_1; A)$ invariáns altérnek. Ha ebben a bázisban felírjuk a transzformáció mátrixát, akkor az első oszlopban vannak a minimálpolinom együtthatóinak ellentettjei, a diagonális felett 1-ek vannak és minden más elem zérus:

$$\begin{pmatrix} -\alpha_{k-1} & 1 & 0 & \dots & 0 \\ -\alpha_{k-2} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\alpha_1 & 0 & 0 & \dots & 1 \\ -\alpha_0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Meggondoltuk tehát, hogy ha a transzformáció m minimálpolinomja irreducibilis és

$$m(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_{k-1} t^{k-1} + t^k$$

alakú, akkor a térnek van olyan bázisa, amelyben a transzformáció mátrixa a fenti mátrix diagonális elrendezésű r darab másolatából áll, ahol $rk = \dim(V)$.

Korábban – lásd a 10.18. állítást – láttuk, hogy ha $m(t)$ egy előre adott k -adfokú normált polinom, akkor például a fenti mátrix által reprezentált, k -dimenziós téren értelmezett lineáris transzformációnak $m(t)$ a minimálpolinomja. Most megmutattuk azt, hogy amennyiben $m(t)$ még irreducibilis is, akkor egyedül a fenti transzformációknak lehet $m(t)$ a minimálpolinomja, amennyiben a k -dimenziós vektortereken értelmezett lineáris transzformációra hagyatkozunk. Ha feloldjuk ezt a dimenzió korlátot, akkor még az rk dimenziós téren értelmezett transzformációk is szóba jönnek fent megértett módon, tehát pl az $5k$ dimenziós téren annak és csak annak a transzformációknak a minimálpolinomja m , amelynek alkalmas bázisban felírt mátrixa 5 darab fenti típusú mátrix diagonális elrendezésével keletkezik.

12.1. Irreducibilis polinommal képzett magtér redukálása

Legyen most p egy tetszőleges k -ad fokú irreducibilis polinom, és $A \in L(V)$ egy lineáris transzformáció. Tekintsük a $V_1 = \ker p(A)$ invariáns alteret, amelyre szorítsuk meg az A transzformációt, azaz $A_1 = A|V_1$. Világos, hogy $p(A_1) = 0 \in L(V_1)$, ezért alkalmazhatjuk a fent igazolt 12.1. állítást a V_1 altérre és az A_1 transzformációra.

12.2. állítás. Legyen V egy véges dimenziós vektortér, $A \in L(V)$ egy lineáris transzformáció, és $p \in \mathbb{F}[t]$ egy k -adfokú irreducibilis polinom. Ekkor:

1. minden $v \neq 0, v \in \ker p(A)$ mellett $\dim(\text{lin}(v; A)) = \deg p$;
2. $A \ker p(A)$ altér nulla dimenziós, vagy k dimenziós, vagy k dimenziós invariáns alterek direkt összege. Ponto-sabban fogalmazva ha $v(p(A)) > 0$, akkor létezik $r \geq 1$ szám, és léteznek v_1, \dots, v_r vektorok, amelyekre

$$\text{lin}(v_1; A) \oplus \dots \oplus \text{lin}(v_r; A) = \ker p(A).$$

12.3. állítás. Legyen $A \in L(V)$ lineáris transzformáció, és $p(t) \in \mathbb{F}[t]$ egy k -ad fokú irreducibilis polinom, amelyre $p(A)^m = 0$. Ekkor az $\text{Im } p(A)^{m-1}$ altér nulla dimenziós, vagy k dimenziós, vagy k dimenziós alterek direktösszege.

Bizonyítás: Az egyszerűbb jelölések kedvéért legyen $B = \ker p(A)$. Mivel $B^m = 0$, ezért $\text{Im } B^{m-1} \subseteq \ker B$. Persze B^{m-1} is A -nak polinomja, ergo B^{m-1} képtere is A -ra nézve invariáns altere $\ker B = \ker p(A)$ -nak. \square

Összefoglalhatjuk azt az esetet, mikor a minimálpolinom különböző irreducibilis polinomok szorzata. Ez éppen a 11.2. állítás esete.

12.4. állítás. Tekintsünk egy $A \in L(V)$ lineáris transzformációt, amelynek minimálpolinomja $m \in \mathbb{F}[t]$. Tegyük fel, hogy m előáll mint a különböző normált, irreducibilis polinomok elsőfokú hatványainak

$$m = p_1 p_2 \dots p_s$$

szorzata. Ekkor V előáll mint néhány – de legalább egy-egy darab – $\deg p_1, \deg p_2, \dots, \deg p_s$ dimenziós minimális invariáns alterének direkt összege.

Bizonyítás: A 11.4. állításban láttuk, hogy $V = \ker p_1(A) \oplus \cdots \oplus \ker p_s(A)$ alakú. minden egyes p_j a minimálpolinom legalább elsőfokú osztója, így $\ker p_j(A) \neq \{0\}$. Az előző állítás szerint minden j mellett $\ker p_j(A)$ egy $\deg p_j$ dimenziós invariáns altér, vagy néhány ilyen direkt összege. Mivel ez minden j mellett igaz, ezért $\ker p_j(A)$ felbontását a V felbontásába helyettesítve készen is vagyunk. \square

A komplex és a valós eset

Nézzük meg mit jelent a 12.4. állítás abban a speciális esetben, mikor az \mathbb{F} test a komplex számok, vagy a valós számok teste.

Elsőként tegyük fel, hogy a V vektortér a \mathbb{C} komplex számtípus feletti vektortér. Ekkor az $A \in L(V)$ lineáris transzformáció minimálpolinomja egy komplex együtthatós polinom. Az algebra alaptétele szerint irreducibilis normált polinom csak elsőfokú, azaz $t - \lambda$ alakú lehet, tehát az előző, a 12.4. állítás feltétele most abba megj át, hogy A minimálpolinomja

$$m(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_s)$$

alakú, ahol $\lambda_1, \dots, \lambda_s$ az m különböző gyökei. Az állítás implikációja, hogy ekkor a vektortér előáll mint egydimenziós A -invariáns alttereinek direkt összege, ergo az A transzformáció diagonalizálható. Kiderült tehát, hogy nincs új a nap alatt, hiszen azt már korábban is tudtuk, – lásd: 11.3 – hogy A pontosan akkor diagonalizálható, ha a minimálpolinomja a fenti kiemelt alakú.¹

Ha most a valós vektortér esetére szorítkozunk, akkor arra kell emlékeznünk, hogy egy a valós test feletti irreducibilis polinomom első- vagy másodfokú lehet csak. Ezért állításunk feltétele most azt jelenti, hogy a minimálpolinom előáll mint páronként relatív prím, első vagy másodfokú polinomok szorzata. A konklúzió szerint ekkor a V vektortér is előáll mint egy vagy két dimenziós invariáns alttereinek direkt összege. Bebizonyítottuk tehát a következő állítást.

12.5. állítás. Tegyük fel, hogy V egy \mathbb{R} feletti vektortér és $A \in L(V)$, az A minimálpolinomjának irreducibilis polinomokra felbontásában minden polinom annak elsőfokú hatványával szerepel. Ekkor a térfelület előáll mint egy vagy két dimenziós invariáns alttereinek direkt összege, ergo van a térfelület olyan bázisa, amelyben felírt mátrix diagonálisában csak 1-szer 1-es vagy 2-szer 2-es blokkok szerepelnek, mindenütt másutt a mátrixban zérus szerepel.

Illusztráció

Bontsuk lehető legalacsonyabb invariáns alterek direkt összegére az alábbi \mathbb{R} feletti vektortéren értelmezett lineáris transzformáció értelmezési tartományát, és írjuk fel A mátrixát a lehető legegyszerűbb módon. A transzformáció definíciója egy $\{u_1, u_2, u_3, u_4\}$ bázis felett a következő:

$$A(\alpha u_1 + \beta u_2 + \gamma u_3 + \delta u_4) = (-2\alpha + 3\gamma)u_1 + (-2\alpha - \beta + 3\gamma + \delta)u_2 + (-\alpha + \gamma)u_3 + (-\alpha - \beta + 3\gamma)u_4.$$

Megoldás: Írjuk fel az operátor mátrixát: $A = \begin{pmatrix} -2 & 0 & 3 & 0 \\ -2 & -1 & 3 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & -1 & 3 & 0 \end{pmatrix}$. Ha a sajátvektorokat keressük láttuk, hogy

nincs valós sajátérték. Keressük tehát a minimálpolinomot a bázis egyes elemeihez tartozó kis minimálpolinomok meghatározásával.

	Au_1	A^2u_1		Au_1	A^2u_1	
u_1	-2 -2 -1 -1	1 2 1 1		u_1	-1 -1 0 0	
				Au_1		
	δ	-1				

$\{u_1, Au_1\}$ lineárisan független, de
 $A^2u_1 + Au_1 + u_1 = 0$,
ezért $p_1(t) = t^2 + t + 1$.

	Au_2	A^2u_2		Au_2	A^2u_2	
u_2	0 -1 0 -1	0 0 0 1		u_2	0 -1 0 -1	
				Au_2		
	δ	-1				

$\{u_2, Au_2\}$ lineárisan független, de
 $A^2u_2 + Au_2 + u_2 = 0$,
ezért $p_2(t) = t^2 + t + 1$.

¹ Mi több, ez utóbbi tetszőleges test feletti vektortéren is igaz nem csak \mathbb{C} felett.

	Au_3	A^2u_3		Au_3	A^2u_3
u_3	3	-3		-1	
	3	-3		0	
	1	-2		-1	
	3	-3		0	
	δ	-1			

$\{u_3, Au_3\}$ lineárisan független, de
 $A^2u_3 + Au_2 + u_3 = 0$,
ezért $p_3(t) = t^2 + t + 1$.

	Au_4	A^2u_4		Au_4	A^2u_4
u_4	0	0		0	
	1	-1		-1	
	0	0		0	
	0	-1		-1	
	δ	-1			

$\{u_4, Au_4\}$ lineárisan független, de
 $A^2u_4 + Au_4 + u_4 = 0$,
ezért $p_4(t) = t^2 + t + 1$.

Ez azt jelenti, hogy a minimálpolinom az $m(t) = t^2 + t + 1$ másodfokú, az \mathbb{R} test felett irreducibilis polinom. A mátrix, tehát két darab $\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$ mátrix diagonális elrendezésű partíciója. Mivel $Au_4 = u_2$ az első invariáns altér bázisa lehet például $\{u_2, u_4\}$. minden olyan nem zérus vektorra, amely nincs e két vektor lineáris burkában, az $\{Av, v\}$ rendszer egy invariáns direktkiegészítőt definiál. Pont erről szól a 12.1. állítás. Ilyen módon például az $\{u_2, u_4, Au_1, u_1\}$ bázisban a transzformáció mátrixa $\begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$ alakú. Emlékezve az általános bázistranszformációra azt kaptuk, hogy

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 & 1 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} -2 & 0 & 3 & 0 \\ -2 & -1 & 3 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & -1 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & -2 & 1 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}. \quad \square$$

13. fejezet

A minimálpolinom fokszámáról

A MINIMÁLPOLINOM DEFINIÁLÁSAKOR csak annyit láttunk, hogy legfeljebb n^2 fokú polinom minden konstruálható, amelynek a transzformáció gyöke, ahol n a tér dimenziója. Ebben a fejezetben látni fogjuk, hogy a fenti gondolat nagyon lényegesen erősíthető. Azt mutatjuk meg, hogy a minimálpolinom fokszáma nem lehet a tér dimenziójánál magasabb.

13.1. lemma. *Legyen $B \in L(V)$ lineáris transzformáció, tegyük fel, hogy a v_1, \dots, v_r vektorok mindenhez kiegészítve $B^m v_j = 0$, de a*

$$\{B^{m-1}v_1, B^{m-1}v_2, B^{m-1}v_3, \dots, B^{m-1}v_r\}$$

vektorrendszer lineárisan független. Ekkor a

$$\left\{ \begin{array}{cccccc} v_1 & v_2 & v_3 & \dots & v_r \\ Bv_1 & Bv_2 & Bv_3 & \dots & Bv_r \\ B^2v_1 & B^2v_2 & B^2v_3 & \dots & B^2v_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^t v_1 & B^t v_2 & B^t v_3 & \dots & B^t v_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^{m-1}v_1 & B^{m-1}v_2 & B^{m-1}v_3 & \dots & B^{m-1}v_r \end{array} \right\}$$

vektorrendszer is lineárisan független.

Bizonyítás: Tekintsük ezen vektorok egy

$$\sum_{j=0}^{m-1} \sum_{k=1}^r \alpha_{j,k} B^j v_k = 0$$

lineáris kombinációját. Meg kell mutatnunk, hogy az összes $\alpha_{j,k} = 0$. Ha $m-1 \geq t \geq 0$, az első olyan index, amelyre van $\alpha_{t,k} \neq 0$ együttható, akkor a t -edik index előtt, minden együttható zérus, ergo

$$\sum_{j=t}^{m-1} \sum_{k=1}^r \alpha_{j,k} B^j v_k = 0$$

Erre alkalmazva a B^{m-t-1} transzformációt azt kapjuk, hogy

$$0 = B^{m-t-1} 0 = \sum_{j=t}^{m-1} \sum_{k=1}^r \alpha_{j,k} B^{j+m-t-1} v_k = \sum_{k=1}^r \alpha_{t,k} B^{t+m-t-1} v_k = \sum_{k=1}^r \alpha_{t,k} B^{m-1} v_k.$$

No de, az $\{B^{m-1}v_1, \dots, B^{m-1}v_r\}$ egy lineárisan független rendszer, amiből már következik, hogy minden $\alpha_{t,k} = 0$, ami ellentmondásban van a t index definíciójával. \square

13.2. állítás. *Legyen $B \in L(V)$ lineáris transzformációra $B^m = 0$. Ekkor $\dim(V) \geq m\rho(B^{m-1})$.*

Bizonyítás: Létezik tehát olyan $\{B^{m-1}v_1, B^{m-1}v_2, \dots, B^{m-1}v_r\}$ lineárisan független vektorrendszer, ahol $r = \rho(B^{m-1})$ a B^{m-1} transzformáció képterének dimenziója. A 13.1. lemma szerint a térben van egy $m \cdot r$ elemű lineárisan független vektorrendszer, ergo a tér legalább $mr = m\rho(B^{m-1})$ dimenziós. \square

13.3. állítás. Legyen $A \in L(V)$ minimálpolinomja $p^m(t)$ alakú, ahol p egy irreducibilis polinom, melynek foka k . Ekkor a tér legalább $m \cdot k$ dimenziós (*azaz a minimálpolinom foka legfeljebb a tér dimenziója*).

Bizonyítás: Először nézzük a trivialitásokat. Ha p a konstans 1 polinom, akkor a tér csak a 0 vektort tartalmazza, tehát minden a minimálpolinom fokszáma, minden a tér dimenziója zérus. Ha $\deg p \geq 1$, és $m = 1$, akkor a minimálpolinom irreducibilis, de van $v \in V$ nem zérus vektor. Láttuk, hogy ilyenkor

$$\dim(\text{lin}(v; A)) = \deg p = k,$$

amiből persze következik, hogy a tér legalább k dimenziós.¹

Most nézzük az érdekes esetet, mikor $\deg p \geq 1$ és $m \geq 2$. Mivel a $p^{m-1}(t)$ a minimálpolinomnál alacsonyab fokú de nem zérus polinom, ezért létezik $v \in V$, melyre $B = p(A)$ jelöléssel $B^{m-1}v \neq 0$. Persze e vektor a $\ker B = \ker p(A)$ egy eleme, és $\ker p(A)$ -ban minden nem zérus elem generálta invariáns altér éppen k -dimenziós, ezért

$$\text{lin}(B^{m-1}v; A) = \text{lin}\{B^{m-1}v, AB^{m-1}v, A^2B^{m-1}v, \dots, A^{k-1}B^{m-1}v\}$$

pontosan k dimenziós. Persze $\{B^{m-1}v, B^{m-1}Av, B^{m-1}A^2v, \dots, B^{m-1}A^{k-1}v\} \subseteq \text{Im } B^{m-1}$, ergo

$$\rho(B^{m-1}) \geq k.$$

Alkalmazva B -re az imént igazolt 13.2 tért kapjuk a kívánt $\dim(V) \geq m\rho(B^{m-1}) \geq m \cdot k$ becslést. \square

13.4. állítás. Tetszőleges lineáris transzformáció minimálpolinomjának foka legfeljebb a tér dimenziója.

Bizonyítás: Legyen $m(t) = p_1^{m_1}(t) \cdot p_2^{m_2}(t) \cdots p_r^{m_r}(t)$ a minimálpolinom relativ prím, irreducibilis polinomok hatványaiként való faktorizációja. Tudjuk, hogy ekkor

$$V = \ker p_1^{m_1}(A) \oplus \ker p_2^{m_2}(A) \oplus \cdots \oplus \ker p_r^{m_r}(A).$$

Az $A|_{\ker p_i^{m_i}(A)}$ minimálpolinomja $p_i^{m_i}(t)$, így az előző tért szerint minden egyes i index mellett $m_i \cdot k_i \leq \dim(\ker p_i^{m_i}(A))$. Világos, hogy

$$\deg m = \sum_{i=1}^r k_i \cdot m_i \leq \sum_{i=1}^r \dim(\ker p_i^{m_i}(A)) = \dim(V). \quad \square$$

Később ki fog derülni, hogy a transzformáció karakterisztikus polinomjának, amely pontosan $\dim(V)$ -ed fokú, is minden gyöke a transzformáció. Ez az úgynevezett Cayley–Hamilton-tétel.

¹Sőt még azt is láttuk, hogy néhány – lehet, hogy csak egy – k dimenziós invariáns altér direkt összege.

Nilpotens transzformációk

14.1. Hatvány függvény alakú minimálpolinom

14.1. definíció (nilpotens transzformáció). Egy $A \in L(V)$ lineáris transzformációt *nilpotensnek* mondjuk, ha létezik $k \in \mathbb{N}$, melyre $A^k = 0$. Ha A egy nilpotens transzformáció, akkor azt a legkisebb m számot, melyre $A^m = 0$ a *nilpotencia rendjének* nevezzük.

Például egy 5 dimenziós téren könnyen definiálhatunk első-, másod-, harmad-, negyed-, és ötöd-rendű nilpotens transzformációkat. De van-e mondjuk hatod-rendű nilpotens transzformáció ezen öt dimenziós vektortéren? Az első észrevétel adja a negatív választ.

14.2. állítás. Legyen B egy m -ed rendben nilpotens operátor. Ekkor létezik $v \in V$ vektor, melyre $B^{m-1}v \neq 0$. minden ilyen v vektorra a

$$\{v, Bv, B^2v, \dots, B^{m-1}v\}$$

m elemű vektorrendszer lineárisan független.

Emiatt ha a V vektortérnek van m -edrendben nilpotens lineáris transzformációja, akkor $\dim(V) \geq m$, azaz a nilpotencia rendje legfeljebb a tér dimenziója.

Bizonyítás: Mivel $B^{m-1} \neq 0$, ezért valóban létezik $v \in V$ vektor, melyre $B^{m-1}v \neq 0$. Persze ekkor a $\{B^{m-1}v\}$ egy elemet tartalmazó rendszer lineárisan független, ezért a 13.1. lemma szerint a tételebéli rendszer is lineárisan független. \square

14.3. állítás. Egy lineáris transzformáció pontosan akkor nilpotens, ha valamely $m \leq \dim(V)$ mellett a minimálpolinomja $m(t) = t^m$ alakú.

Bizonyítás: Tegyük fel először, hogy B lineáris transzformáció m -ed rendben nilpotens. Legyen $p(t) = t^m$. Ekkor $p \in J_A$, így a minimálpolinom p osztója. Másrészt a 14.2. állítás szerint van a térfelvétel olyan vektorá, amelyhez tartalmazó kis minimálpolinom is p . Így p osztója a minimálpolinomnak, ergo azonos vele.

Megfordítva, ha $m(t) = t^m$ a minimálpolinomja B -nek, akkor $B^m = 0$ és ez m -nél kisebb kitevőre nem teljesülhet. Ez éppen azt jelenti, hogy B transzformáció m -ed rendben nilpotens. \square

14.2. Nilpotens operátorok redukálása

Az alábbi lemmának nincs köze a transzformációk redukálásához. Arra kell emlékeznünk, hogy egy véges dimenziós vektortérben minden altérnek van direkt kiegészítője.

14.4. lemma. Legyenek V_1 és V_2 diszjunkt alterei a véges dimenziós W vektortérnek. Ekkor V_1 -nek létezik V_2 altérét tartalmazó direktiegészítője, azaz létezik $K \subset W$ altér, amelyre $V_2 \subseteq K$ és $V_1 \oplus K = W$.

Bizonyítás: Jelölje $V = V_1 + V_2$. Világos, hogy V altér W -ben. Legyen L a direktiegészítője, azaz $V \oplus L = W$. Ha $K = V_2 + L$, akkor K olyan altér W -ben, amelyre $V_2 \subseteq K$, $K \cap V_1 = \{0\}$, valamint $V_1 + K = W$. \square

A következő lemmának nagyon fontos szerepe lesz a fejezet leglényegesebb állításában a nilpotens operátorok redukcióját biztosító állításban.

14.5. lemma. Legyen a W véges dimenziós vektortérnek H, K_0, \bar{K} altére. Tegyük fel, hogy

1. $H \cap K_0 = \{0\}$;
2. $H + \bar{K} = W$;
3. $K_0 \subseteq \bar{K}$.

Ekkor létezik K altér W -nek, melyre

1. $K_0 \subseteq K \subseteq \bar{K}$ és
2. K direkt kiegészítője H -nak, azaz $H \oplus K = W$.

Bizonyítás: Először is a K altér konstrukciója következik. Világos, hogy $H \cap \bar{K} \subseteq \bar{K}$ és $K_0 \subseteq \bar{K}$ diszjunkt alterek \bar{K} -ban, hiszen $(H \cap \bar{K}) \cap K_0 \subseteq H \cap K_0 = \{0\}$. Alkalmazzuk az előző 14.4. lemmát a \bar{K} altérben. Létezik tehát $K_0 \subseteq K \subseteq \bar{K}$ altér \bar{K} -ban, amelyre $(H \cap \bar{K}) \oplus K = \bar{K}$.
 A H és a K diszjunkt alterek: $H \cap K = H \cap (K \cap \bar{K}) = (H \cap \bar{K}) \cap K = \{0\}$.
 A H és a K Minkowski-összege az egész tér: $W = H + \bar{K} = H + ((H \cap \bar{K}) + K) \subseteq H + (H + K) = H + K$. Ez éppen azt jelenti, hogy K direkt kiegészítője H -nak. \square

14.6. állítás (nilpotens operátorok felbontása). Legyen $B \in L(W)$ egy m -ed rendben nilpotens lineáris transzformáció. Ekkor minden olyan $v \in V$ vektorhoz, amelyre $B^{m-1}v \neq 0$, a $\text{lin}(v; B)$ invariáns altérnek van invariáns altér direkt kiegészítője.

Formálisabban: létezik $K \subseteq W$ invariáns altér, amelyre $\text{lin}\{v, Bv, \dots, B^{m-1}v\} \oplus K = W$.

Bizonyítás: A nilpotens transzformáció rendje szerinti teljes indukció. Ha $m = 1$, akkor $B = 0$, de a konstans zérus operátorra nézve minden altér invariáns, így a téTEL összesen annyit állít, hogy egy nem zérus v vektor generálta egydimenziós altérnek van direkt kiegészítője.

Tegyük fel, hogy igaz az állítás minden vektortér legfeljebb $m-1$ -ed rendben nilpotens transzformációjára. Legyen tehát $m > 1$ és B egy a W vektortéren értelmezett m -ed rendben nilpotens lineáris transzformáció. Rögzítsünk egy $v \in W$ elemet, amelyre $B^{m-1}v \neq 0$. Tekintsük az $\text{Im } B$ invariáns alteret. Világos, hogy $B|_{\text{Im } B}$ egy lineáris transzformáció az $\text{Im } B$ vektortéren. Az is világos, hogy $B|_{\text{Im } B}$ egy $m-1$ -rendben nilpotens lineáris transzformáció, hiszen minden $u \in \text{Im } B$ mellett $B^{m-1}u = 0$. Azt is vegyük észre, hogy ezek szerint $v \notin \text{Im } B$.

Alkalmazhatjuk tehát az indukciós feltevést $B|_{\text{Im } B} \in L(\text{Im } B)$ mellett a Bv vektorra. Persze $B^{m-2}Bv = B^{m-1}v \neq 0$. Létezik tehát $K_0 \subseteq \text{Im } B$ a B -re is invariáns altér, amelyre

$$\text{lin}\{Bv, B^2v, \dots, B^{m-1}v\} \oplus K_0 = \text{Im } B.$$

Most megmutatjuk, hogy $\text{lin}(v; B) \cap K_0 = \{0\}$. Ugyanis, ha $x = \sum_{k=0}^{m-1} \alpha_k B^k v \in K_0 \subseteq \text{Im } B$, akkor $\alpha_0 v \in \text{Im } B$, ami csak úgy lehetséges, hogy $\alpha_0 = 0$. Ezek szerint $x \in \text{lin}(Bv; B)$ altérnek melynek direkt kiegészítője K_0 . Ez persze csak úgy lehetséges, hogy $x = 0$.

Definiálja

$$\bar{K} = \{x \in W : Bx \in K_0\}.$$

Mivel K_0 egy altér, ezért \bar{K} is az. Mivel a K_0 altér B -invariáns, azért teljesül a $K_0 \subseteq \bar{K}$ tartalmazás.

Most megmutatjuk, hogy $\text{lin}(v; B) + \bar{K} = W$. Válasszunk egy $u \in W$ vektort. Persze $Bu \in \text{Im } B$, ezért előáll

$$Bu = \sum_{k=1}^{m-1} \alpha_k B^k v + k_0 = B \left(\sum_{k=0}^{m-2} \alpha_{k+1} B^k v \right) + k_0$$

alakban, ahol $k_0 \in K_0$. Ebből azt látjuk, hogy $B(u - \sum_{k=0}^{m-2} \alpha_{k+1} B^k v) \in K_0$, ami persze \bar{K} definícióját figyelembe véve azt jelenti, hogy $u - \sum_{k=0}^{m-2} \alpha_{k+1} B^k v \in \bar{K}$. Előállítottuk tehát az u vektort egy $\text{lin}(v; B)$ -beli és egy \bar{K} -beli vektor összegeként.

Alkalmazhatjuk tehát a 14.5. lemmát. Így létezik $K_0 \subseteq K \subseteq \bar{K}$ altér, melyre $\text{lin}(v; B) \oplus K = W$. Persze ha $u \in K \subseteq \bar{K}$, akkor $Bu \in K_0 \subseteq K$, ergo K egy B -invariáns direkt kiegészítője a v -t tartalmazó legszűkebb B invariáns altérnek. Ezt kellett belátni. \square

Ahhoz, hogy megkapjuk az egész vektorteret v -invariáns altérként vagy ilyenek direkt összegeként, az előző tételt kell rekurzívan alkalmaznunk. A W vektortér véges dimenziós volta garantálja, hogy a rekurzió véget ér.

Persze $B|_K$ a K altér lineáris transzformációja, ami $m \geq n_2 \geq 1$ rendben nilpotens. Ha alkalmazzuk a fenti tételt, akkor kapjuk, hogy létezik $v_2 \in K, v_2 \neq 0$ elem és létezik $K_2 \subseteq K$ a B -re nézve invariáns altér, amelyre

$$\text{lin}\{v_2, Bv_2, \dots, B^{n_2-1}v_2\} \oplus K_2 = K.$$

Itt persze $\dim(K_2) < \dim(K)$, hiszen a baloldali első invariáns altér legalább egydimenziós. Az első két lépést összefoglalva:

$$\text{lin}(v_1; B) \oplus \text{lin}(v_2; B) \oplus K_2 = W,$$

Az eljárást folytatva minden lépésben legalább egyelőre csökken a kiegészítő invariáns altér dimenziója. Végül a vektortér előáll néhány, mondjuk r darab B -re invariáns altér direkt összegeként:

$$\text{lin}(v_1; B) \oplus \text{lin}(v_2; B) \oplus \dots \oplus \text{lin}(v_r; B) = W.$$

14.7. állítás (nilpotens transzformáció redukálása). *Legyen $B \in L(W)$ egy m -ed rendben nilpotens transzformáció. Ekkor léteznek olyan $v_1, v_2, \dots, v_r \in W$ vektorok és léteznek olyan $m = n_1 \geq n_2 \geq \dots \geq n_r \geq 1$ pozitív egészek, amelyekre*

$$\text{lin}\{v_1, Bv_1, \dots, B^{n_1-1}v_1\} \oplus \text{lin}\{v_2, Bv_2, \dots, B^{n_2-1}v_2\} \oplus \dots \oplus \text{lin}\{v_r, Bv_r, \dots, B^{n_r-1}v_r\} = W.$$

Emiatt az egyes invariáns alterek bázisainak egyesítésével kapott vektorrendszer bázisa W -nek:

$$\{B^{n_1-1}v_1, \dots, Bv_1, v_1\} \cup \{B^{n_2-1}v_2, \dots, Bv_2, v_2\} \cup \dots \cup \{B^{n_r-1}v_r, \dots, Bv_r, v_r\}$$

14.3. Egyértelműség

Azt mutatjuk meg, hogy a normálalakban $r = \nu(B)$, és az n_1, n_2, \dots, n_r számok is a B nilpotens transzformáció által egyértelműen meghatározottak. Ez azt jelenti, hogy minden nilpotens transzformációnak csak egyetlen normálalakja van.

14.8. lemma. *Legyen $A \in L(V)$ lineáris transzformáció, és $v \in V$ olyan vektor, amelyre $A^m v = 0$, de $A^{m-1}v \neq 0$. Ekkor*

1. $\{v, Av, \dots, A^{m-1}v\}$ lineárisan független, így $\text{lin}(v; A) = \text{lin}\{v, Av, \dots, A^{m-1}v\}$;
2. minden $0 \leq l \leq m$ mellett $\nu(A^l | \text{lin}(v; A)) = l$ és $\rho(A^l | \text{lin}(v; A)) = m - l$.

Bizonyítás: Ha $l = m$, akkor minden két állítás triviálisan teljesül. Legyen tehát $0 \leq l < m$. Az A^l transzformáció az $\{A^{m-l}v, \dots, A^{m-1}v\}$ lineárisan független vektorrendszer nullára viszi, így $l \leq \nu(A^l | \text{lin}(v; A))$. A maradékot, a $\{v, \dots, A^{m-l-1}v\}$ vektorrendszerét pedig az $\{A^l v, \dots, A^{m-1}v\}$ lineárisan független rendszerre képezi. Így $m - l \leq \rho(A^l | \text{lin}(v; A))$, amiből

$$l \leq \nu(A^l | \text{lin}(v; A)) = m - \rho(A^l | \text{lin}(v; A)) \leq l.$$

□

14.9. lemma. *Legyen $A \in L(V)$ és tegyük fel, hogy a V vektortér előáll a K_1, K_2 invariáns alterei direkt összegeként, azaz $V = K_1 \oplus K_2$. Ekkor $\rho(A) = \rho(A|K_1) + \rho(A|K_2)$ és $\nu(A) = \nu(A|K_1) + \nu(A|K_2)$.*

Bizonyítás: Világos, hogy $A(V) = A(K_1) + A(K_2)$, és az invariancia szerint $A(K_1) \cap A(K_2) \subseteq K_1 \cap K_2 = \{0\}$. Így persze $A(V) = A(K_1) \oplus A(K_2)$, és

$$\rho(A) = \dim(A(V)) = \dim(A(K_1)) + \dim(A(K_2)) = \rho(A|K_1) + \rho(A|K_2).$$

Ebből már

$$\nu(A) = \dim(V) - \rho(A) = \dim(K_1) + \dim(K_2) - \rho(A|K_1) - \rho(A|K_2) = \nu(A|K_1) + \nu(A|K_2)$$

könnyen adódik. □

14.10. állítás (A nilpotens felbontás egyértelműsége). *Legyen $A \in L(V)$ egy lineáris transzformáció. Tegyük fel, hogy valamely $\{v_1, \dots, v_r\}$ vektorrendszerre és valamelyen $m_1 \geq m_2 \geq \dots \geq m_r$ pozitív számokra $A^{m_k-1}v_k \neq 0$, de $A^{m_k}v_k = 0$ fennáll minden $k = 1, \dots, r$, továbbá*

$$V = \text{lin}(v_1; A) \oplus \dots \oplus \text{lin}(v_r; A).$$

Tegyük fel még azt is, hogy valamely másik $\{w_1, \dots, w_s\}$ vektorrendszerre és valamely más $n_1 \geq n_2 \geq \dots \geq n_s$ pozitív számokra $A^{n_k-1}w_k \neq 0$, de $A^{n_k}w_k = 0$ fennáll minden $k = 1, \dots, s$ mellett, és

$$V = \text{lin}(w_1; A) \oplus \dots \oplus \text{lin}(w_s; A).$$

Ekkor

1. A nilpotens lineáris transzformációja V -nek;
2. Ha m jelöli a nilpotencia rendjét, akkor $m_1 = m = n_1$;
3. A transzformáció $\nu(A)$ defektusára $r = \nu(A) = s$;
4. Valamennyi $k = 1, \dots, r$ esetén $m_k = n_k$.

Bizonyítás: Az első két állítás nyilvánvaló, mivel a rendezettség szerint $A^{m_1}|_{\text{lin}(v_k; A)} = 0$, minden $k = 1, \dots, r$ mellett

A harmadik állításhoz:

$$\nu(A) = \sum_{k=1}^r \nu(A| \text{lin}(v_k; A)) = \sum_{k=1}^r 1 = r.$$

Ugyanígy a másik direkt összeg felbontásból kapjuk, hogy $\nu(A) = s$.

A negyedik állítás. Tegyük fel – indirekt –, hogy $m_k = n_k$ nem teljesül minden $k = 1, \dots, r$ számra. Legyen $1 < t \leq r$ az a legkisebb szám, amelyre $m_t \neq n_t$. Ezek szerint $k = 1, \dots, t-1$ mellett $m_k = n_k$, de $m_t \neq n_t$. Feltehető, hogy $m_t > n_t$. Ekkor tehát

$$m_1 = n_1 \geq m_2 = n_2 \geq \dots \geq m_{t-1} = n_{t-1} \geq m_t > n_t.$$

Ekkor az első direkt összeg felbontásban a t -nél magasabb indexű tagokat elhagyva

$$\rho(A^{n_t}) \geq \left(\sum_{k=1}^{t-1} \rho(A^{n_t} | \text{lin}(v_k; A)) \right) + \rho(A^{n_t} | \text{lin}(v_t; A)) = \left(\sum_{k=1}^{t-1} (m_k - n_t) \right) + m_t - n_t.$$

Hasonlóan, a második direkt összeg felbontásban a t -edik, és a t -nél magasabb indexű rangok zérók, így

$$\rho(A^{n_t}) = \sum_{k=1}^{t-1} \rho(A^{n_t} | \text{lin}(w_k; A)) = \sum_{k=1}^{t-1} (n_k - n_t) = \sum_{k=1}^{t-1} (m_k - n_t),$$

ami ellentmond $m_t > n_t$ feltételnek. □

Az alábbiakban a nilpotens felbontási tételek annak egyértelműségével együtt foglaljuk össze.

14.11. állítás (nilpotens transzformáció normálalakja). *Legyen $B \in L(W)$ egy m -ed rendben nilpotens transzformáció. Jelölje $r = \nu(B)$ a B defektusát. Ekkor léteznek olyan $v_1, v_2, \dots, v_r \in W$ vektorok és létezik pozitív egészek egyetlen olyan $m = n_1 \geq n_2 \geq \dots \geq n_r \geq 1$ véges sorozata, amelyekre*

$$\text{lin}\{v_1, Bv_1, \dots, B^{n_1-1}v_1\} \oplus \text{lin}\{v_2, Bv_2, \dots, B^{n_2-1}v_2\} \oplus \dots \oplus \text{lin}\{v_r, Bv_r, \dots, B^{n_r-1}v_r\} = W.$$

Emiatt a

$$\{B^{n_1-1}v_1, \dots, Bv_1, v_1\} \cup \{B^{n_2-1}v_2, \dots, Bv_2, v_2\} \cup \dots \cup \{B^{n_r-1}v_r, \dots, Bv_r, v_r\} \quad (†)$$

vektorrendszer bázisa W -nek.

Ha ebben a bázisban felírjuk B mátrixát, akkor r darab diagonálisan elhelyezkedő részmátrixból álló mátrixot kapunk. Az első $n_1 \times n_1$ méretű, a második $n_2 \times n_2$ méretű, ..., az utolsó $n_r \times n_r$ méretű. minden ilyen blokkban csak a (felső) mellék diagonális elemei nem nullák. A mellék diagonális elemei 1-esek. minden más elem zérus. Ezt a mátrixot nevezzük a B nilpotens transzformáció normálalakjának.

Adott $1 \leq j \leq r$ mellett tehát, $\mathbf{B}_j \in \mathbb{F}^{n_j \times n_j}$ a fenti j -edik invariáns altérre leszorított B leképezésnek a $\{B^{n_j-1}v_j, \dots, Bv_j, v_j\}$ bázison felírt mátrixa:

$$\mathbf{B}_j = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Így az egész V vektortéren értelmezett B lineáris transzformáció normálalakja – azaz a B -nek (\dagger) bázisban felírt mátrixa – a fenti tipusú mátrixok diagonális alakú elrendezésével adódik:

$$[B] = \left(\begin{array}{cccccc|cccccc|cccccc|cccccc} 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \end{array} \right).$$

Hangsúlyozni szeretném, hogy az egyértelműség szerint a \mathbf{B}_j blokkok száma, és azok mérete is csak a nilpotens transzformációtól függ. Több bázis is lehetséges, amelyben a nilpotens transzformáció normálalakú, de nem csak hogy minden normálalakban azonos számú blokk van (B defektusa), de a blokkok mérete is azonos. A nilpotens felbontásban szereplő direkt összeadandó alerek páronként izomorfak egymással.

14.4. Illusztrációk

Egyetlen invariáns altér

Írjuk fel a W vektortéren értelmezett lineáris transzformáció normál alakját, ahol az $\{u_1, u_2, u_3, u_4\}$ bázisban

$$B(\alpha u_1 + \beta u_2 + \gamma u_3 + \delta u_4) = -(\gamma + \delta)u_1 + \gamma u_2 - (\alpha + \beta + \gamma)u_3 + (\alpha + \beta + \gamma + \delta)u_4.$$

A B mátrixa a fent rögzített bázisban:

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

Mivel két azonos oszlop is van, ezért a defektus legalább egy. Számoljuk ki a minimálpolinomot! Az u_1 bázis elemhez

	u_1	Bu_1	B^2u_1	B^3u_1	B^4u_1
u_1	1	0	0	-1	0
u_2	0	0	-1	1	0
u_3	0	-1	1	0	0
u_4	0	1	0	0	0

Mivel az első négy oszlop lineárisan független, ezért az u_1 -hez tartozó kis minimálpolinom $p_1 = t^4$. Mivel tudjuk, hogy a minimálpolinom legfeljebb 4-ed fokú, és p_1 osztja, ezért csak $m(t) = t^4$ lehetséges, ezért B transzformáció 4-ed rendben nilpotens. Éppen most számoltuk ki, hogy $B^3u_1 \neq 0$, ezért az u_1 -et tartalmazó legszűkebb B invariáns altér az egész W , tehát

$$\text{lin} \{u_1, Bu_1, B^2u_1, B^3u_1\} = W$$

és a $\{B^3u_1, B^2u_1, Bu_1, u_1\}$ sorrendű bázisban B mátrixa

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

alakú. Emlékezzünk, hogy az új bázisra való áttérés mátrixa egyszerűen az új bázis elemeiből mint oszlopokból alkotott mátrix, ami azt jelenti, hogy

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

Két invariáns altér (3- és 1-dimenziós)

Most a játék kedvéért, keressük meg a fenti felbontást és bázist, egy konkrét esetben. Legyen W egy négy dimenziós vektortér az $\{u_1, u_2, u_3, u_4\}$ rögzült bázissal. A $B \in L(W)$ lineáris transzformáció definíciója a báziselemek segítségével:

$$B(\alpha u_1 + \beta u_2 + \gamma u_3 + \delta u_4) = (\gamma - \delta)u_1 + \gamma u_2 + \delta u_3.$$

Első lépésként keressük meg a minimálpolinomot. A transzformáció mátrixa

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

Az u_1 -minimálpolinom $p_1(t) = t$, az u_2 -minimálpolinom $p_2(t) = t$. A harmadik kis minimálpolinomhoz írjuk fel u_3 hatványait:

	u_3	Bu_3	B^2u_3
u_1	0	1	0
u_2	0	1	0
u_3	1	0	0
u_4	0	0	0

Persze az első két oszlop lineárisan független, így $p_3(t) = t^2$. Hasonlóan az u_4 vektor B hatványait felírva:

	u_4	Bu_4	B^2u_4	B^3u_4
u_1	0	-1	1	0
u_2	0	0	1	0
u_3	0	1	0	0
u_4	1	0	0	0

Az első három oszlop ránézésre független, emiatt a minimálpolinom $p_4(t) = t^3$.

Emlékszünk, hogy a minimálpolinom a p_1, p_2, p_3, p_4 legkisebb közös többszöröse, ergo $m(t) = t^3$, és B egy harmadrendben nilpotens transzformáció.

Írjuk fel a fenti tételekben szereplő invariáns direktfelbontást. Mivel a B rangja ránézésre 2, így a magtere 2 dimenziós, ergo két invariáns altér direkt összege W , amiből az egyik 3 dimenziós, így a másik csak egydimenziós lehet, ezért azt csak a magtér egyik eleme generálhatja! A három dimenziós altér lehet például a $\text{lin}(u_4; B)$, hiszen éppen az imént láttuk, hogy $B^2 u_4 \neq 0$. A fenti 14.6. állítás éppen azt mondja, hogy ennek az altérnek van olyan K invariáns altér direkt kiegészítője, aminek van olyan bázisa, amely annyi magtérbeli elemet tartalmaz, ami a $B|_K$ nilpotens operátor rendje, tehát legalább 1. Így most nyilvánvaló, hogy az egydimenziós invariáns alteret generálhatja a ker A magtér bármely olyan eleme, amely lineárisan független $B^2 u_4$ -től. Például u_1 .¹ Ilyen módon

$$\text{lin}\{u_4, Bu_4, B^2 u_4\} \oplus \text{lin}\{u_1\} = W,$$

és az $\{B^2 u_4, Bu_4, u_4, u_1\}$ vektorok olyan bázisát adják a térnak, melyben B mátrixa az alábbi alakú

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

14.5. A nilpotens felbontási tételek nélkül?

Legyen $B \in L(V)$ egy m -edrendben nilpotens operátor. Világos, hogy $\nu(B) \geq 1$, hiszen egyébként B valamennyi hatványa is reguláris maradna.

Válasszuk ki a ker B egy $\{e_1, \dots, e_r\}$ bázisát, ahol a rövidség kedvéért $r = \nu(B)$. Most minden $i = 1, \dots, r$ mellett legyen m_i a legnagyobb olyan k egész, amelyre a $B^{k-1}x = e_i$ egyenletnek még van megoldása. Világos, hogy $1 \leq m_i \leq m$. Az általánosság elvesztése nélkül feltehető, hogy $m \geq m_1 \geq m_2 \geq \dots \geq m_r$, hiszen a bázis elemeket az m_i számok csökkenő sorrendjében átindexelhetjük. Jelölje $v_i \in V$ egy tetszőleges megoldását $B^{m_i-1}x = e_i$ -nek. Világos tehát, hogy minden $i = 1, \dots, r$ mellett

$$B^{m_i-1}v_i = e_i \quad \text{és} \quad B^{m_i}v_i = 0$$

Írjuk egy táblázat legalsó sorába a ker B kiválasztott bázis elemeit, majd föléjük a megfelelő csökkenő B hatványokat.²

$$\begin{array}{cccccc} v_1 & \vdots & \vdots & \dots & \vdots \\ Bv_1 & v_2 & \vdots & \dots & \vdots \\ B^2v_1 & Bv_2 & v_3 & \dots & v_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^{m_1-k}v_1 & B^{m_2-k}v_2 & B^{m_3-k}v_3 & \dots & B^{m_r-k}v_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^{m_1-1}v_1 & B^{m_2-1}v_2 & B^{m_3-1}v_3 & \dots & B^{m_r-1}v_r \end{array}$$

Felmerül, hogy a fenti vektorrendszer bázis, evvel a nilpotens felbontási tételel megkerülve kapnánk a nilpotens normálalak igazolását.

Sajnos nem feltétlen bázis a fenti vektorrendszer. A lineárisan függetlenség a korábbi technikával igazolható – tegyük ezt meg! –, de a rendszer nem minden generátorrendszer.

Tekintsük például a

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

¹Vagy bármilyen, ami $\alpha u_1 + \beta u_2$ alakú, ahol $\alpha \neq \beta$.

²Tipográfiai probléma e táblázat áttekinthető leírása. A lényeg, hogy a legalsó sorban lévő vektorok vannak csak biztosan egy sorban. A 2. oszlopnak $m_2 \leq m_1$ eleme van, tehát a v_1 akkor és csak akkor esik egy sorba v_2 vel, ha $m_1 = m_2$. Egyébként v_2 lejebb van mint v_1 . Képzeljük úgy a táblázatot, mint monoton fogyó elemszámú oszlopok, alulra zárt összeségét, úgy hogy a legnagyobbal kezdem, stb.

mátrixát. Jelölje most $\{e_1, e_2, e_3, e_4\}$ a térnak azt a bázisát, amelyben a fenti mátrixot felírtuk. Világos, hogy $\{e_4, e_3 + e_4, e_1 + e_3 + e_4\}$ vektorrendszer a ker B egy olyan bázisa, amelynek egyik eleme sem esik B képterébe, hiszen egyik elem sem az e_1 skalárszorosa. Ilyen módon a fent konstruált vektorrendszer csak három elemű lesz, ergo nem bázisa a négy dimenziós térnak.

A nilpotens felbontási tétele éppen azt mondja, hogy a ker B -nek van olyan alkalmasan megválasztott bázisa, amelyre a fenti konstrukcióban kapott vektorrendszer a térnak generátorrendszerére, ergo bázisa. Persze amikor konkrétan a normálalakot konstruáljuk, akkor elegendő olyan bázisát keresni ker B -nek, amelyből kiindulva a fenti vektorrendszer elemeinek száma a tér dimenziójával egyezik meg. Az így kapott rendszer persze bázis lesz.

15. fejezet

A Jordan-normálalak

A LEGFONTOSABB GONDOLATHOZ ÉRKEZTÜNK, DE IGÁZÁN NINCS ÚJ A FEJEZETBEN. AZ ELŐZŐ FEJEZETEKET FOGLALJUK ÖSSZE. ELŐSZÖR AZT AZ ESETET VIZSGÁLJUK, MIKOR EGY TRANSZFORMÁCIÓ MINIMÁLPOLINOMJA EGYETLEN ELSŐFOKÚ POLINOM VALAMELYEN HATVÁNYA, MAJD EZT FELHASZNÁLVA KAPJUK AZ ÚGYNEVEZETT KANONIKUS ALAKOT ABBAZ AZ ESETBEN IS, MIKOR A MINIMÁLPOLINOM KÜLÖNÖBÖZŐ ELSŐFOKÚ POLINOMOK HATVÁNYAINAK SZORZATAIKÉNT ÁLL ELŐ. Mivel egy komplex együtthatós polinom minden ilyen alakú, ezért az eredményeinket komplex test feletti vektorterek lineáris transzformációira alkalmazhatjuk.

15.1. Egy gyöktényezős minimálpolinom

15.1. definíció (Jordan-normálalak). Tegyük fel, hogy az $A \in L(V)$ lineáris transzformáció minimálpolinoma

$$(t - \lambda)^m$$

alakú, valamely $\lambda \in \mathbb{F}$ és m pozitív egész mellett. Világos, hogy $B = A - \lambda I$ éppen m -edrendben nilpotens. Alkalmazható tehát B -re a 14.11. állítás. Jelölje tehát $r = \nu(B)$ a B defektusát. Így léteznek olyan $v_1, v_2, \dots, v_r \in V$ vektorok és létezik pozitív egészek egyetlen olyan $m = n_1 \geq n_2 \geq \dots \geq n_r \geq 1$ véges sorozata, amelyekre a

$$\{B^{n_1-1}v_1, \dots, Bv_1, v_1\} \cup \{B^{n_2-1}v_2, \dots, Bv_2, v_2\} \cup \dots \cup \{B^{n_r-1}v_r, \dots, Bv_r, v_r\} \quad (\dagger)$$

vektorrendszer bázisa V -nek.

A B transzformációnak ebben a bázisban felírt mátrixa olyan, hogy a szuper diagonálisán kívül minden elem zérus, a szuper diagonálisban $n_1 - 1$ db 1-es, aztán egy zérus, majd $n_2 - 1$ db 1-es aztán egy zérus, stb. Adjuk a $B = A - \lambda I$ mátrixához a λI diagonális mátrixot! Ekkor kapjuk az A -nak a (\dagger) bázisban felírt mátrixát. Ezt a mátrixot nevezzük az A transzformáció *Jordan-normálalakjának*.

Adott $1 \leq j \leq r$ mellett, a j -edik invariáns altérre leszorított A leképezésnek a $\{B^{n_j-1}v_j, \dots, Bv_j, v_j\}$ bázison felírt mátrixát *Jordan-blokknak* mondjuk.

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & \dots & 0 & \lambda & 1 \\ 0 & 0 & \dots & \dots & 0 & 0 & \lambda \end{pmatrix}.$$

Így az egész V vektortéren értelmezett A lineáris transzformáció normálalakja – azaz a A -nak (\dagger) bázisban

felírt mátrixa – a fenti tipusú Jordan-blokkok, diagonális alakú elrendezésével adódik:

$$[A] = \left(\begin{array}{ccccccc} \lambda & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & \lambda & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & \lambda \\ \\ & \lambda & 1 & 0 & \dots & 0 & 0 & 0 \\ & 0 & \lambda & 1 & \dots & 0 & 0 & 0 \\ & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ & 0 & 0 & \dots & \lambda & 1 & 0 & 0 \\ & 0 & 0 & \dots & 0 & \lambda & 1 & 0 \\ & 0 & 0 & \dots & \dots & 0 & \lambda & 1 \\ & 0 & 0 & \dots & \dots & 0 & 0 & \lambda \\ \\ & & & & \ddots & & & \\ & & & & & \lambda & 1 & 0 & \dots & 0 & 0 & 0 \\ & & & & & 0 & \lambda & 1 & \dots & 0 & 0 & 0 \\ & & & & & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ & & & & & 0 & 0 & \dots & \lambda & 1 & 0 & 0 \\ & & & & & 0 & 0 & \dots & 0 & \lambda & 1 & 0 \\ & & & & & 0 & 0 & \dots & \dots & 0 & \lambda & 1 \\ & & & & & 0 & 0 & \dots & \dots & 0 & 0 & \lambda \end{array} \right)$$

Hangsúlyozni szeretném, hogy az egyértelműség szerint a Jordan blokkok száma, és azok mérete is csak az A transzformációtól függ. Több bázis is lehetséges, amelyben a transzformáció Jordan-normálalakú, de nem csak hogy minden normálalakban azonos számú Jordan-blokk van,¹ de a blokkok mérete is azonos. A felbontásban szereplő direkt összeadandó alterek páronként izomorfak egymással.

15.2. Jordan-normálalak: az általános eset

15.2. definíció. Legyen $A \in L(V)$ lineáris transzformáció valamely \mathbb{F} test feletti vektortéren. Tegyük fel, hogy A minimálpolinomja

$$(t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_s)^{m_s}$$

alakú, ahol $\lambda_1, \dots, \lambda_s$ az A különböző sajátértékei, és m_1, \dots, m_s pozitív egészek. Jelölje $V_j = \ker(A - \lambda_j I)^{m_j}$, és $A_j = A|V_j$ minden $j = 1, \dots, s$ mellett. A 11.4. állításban rögzítettük, hogy

$$V = V_1 \oplus \cdots \oplus V_s$$

továbbá a V_j vektortér A_j lineáris transzformációjának minimálpolinomja

$$(t - \lambda_j)^{m_j}.$$

Minden egyes j mellett írjuk fel az $A_j \in L(V_j)$ transzformáció Jordan-normálakját, majd egyesítsük a V_j alterek Jordan-bázisait a V tér bázisává. Jelölje $B_j = A - \lambda_j I$, így kapjuk minden $j = 1, \dots, s$ mellett az

$m_j = n_1^{(j)} \geq n_2^{(j)} \geq \dots \geq n_{r_j}^{(j)} \geq 1$ pozitív egészket és a $v_1^{(j)}, v_2^{(j)}, \dots, v_{r_j}^{(j)} \in V$

vektorokat, amelyekre a

$$\left\{ \begin{array}{lll} B_1^{n_1^{(1)}-1}v_1^{(1)}, \dots, B_1v_1^{(1)}, v_1^{(1)}; & B_1^{n_2^{(1)}-1}v_2^{(1)}, \dots, B_1v_2^{(1)}, v_2^{(1)}; & \dots; & B_1^{n_r^{(1)}-1}v_{r_1}^{(1)}, \dots, B_1v_{r_1}^{(1)}, v_{r_1}^{(1)}; \\ B_2^{n_1^{(2)}-1}v_1^{(2)}, \dots, B_2v_1^{(2)}, v_1^{(2)}; & B_2^{n_2^{(2)}-1}v_2^{(2)}, \dots, B_2v_2^{(2)}, v_2^{(2)}; & \dots; & B_2^{n_r^{(2)}-1}v_{r_2}^{(2)}, \dots, B_2v_{r_2}^{(2)}, v_{r_2}^{(2)}; \\ \vdots & \vdots & \ddots & \vdots \\ B_s^{n_1^{(s)}-1}v_1^{(s)}, \dots, B_sv_1^{(s)}, v_1^{(s)}; & B_s^{n_2^{(s)}-1}v_2^{(s)}, \dots, B_sv_2^{(s)}, v_2^{(s)}; & \dots; & B_s^{n_r^{(s)}-1}v_{r_s}^{(s)}, \dots, B_sv_{r_s}^{(s)}, v_{r_s}^{(s)}. \end{array} \right\}$$

¹A Jordan-blokkok száma a $B = A - \lambda I$ defektusa, azaz a λ sajátérték geometriai dimenziója.

vektorrendszer bázisa V -nek, amely bázist az A transzformáció *Jordan-bázisának* nevezünk. Az A transzformációjának ebben a bázisban felírt mátrixát nevezzük az *A Jordan-normálalakjának*.

Az A lineáris transzformáció Jordan-normálalakját tehát az egyes λ_j sajátértékekhez tartozó $A_j \in L(V_j)$ transzformációk Jordan-normálalakú mátrixainak diagonális alakú elrendezésével kapjuk. Tehát

$$[A] = \begin{pmatrix} [A_1] & & & \\ & [A_2] & & \\ & & \ddots & \\ & & & [A_{s-1}] \\ & & & & [A_s] \end{pmatrix}.$$

Itt minden egyes $[A_j]$ mátrix annyi Jordan-blokkból áll, amennyi a λ_j sajátérték geometriai dimenziója. minden egyes A_j mátrix első Jordan-blokkja egy $m_j \times m_j$ méretű részmátrix, majd rendre kisebb vagy egyenlő méretű blokkok következnek.

15.3. állítás. Legyen V egy \mathbb{C} komplex számtest feletti vektortér, és $A \in L(V)$ egy lineáris transzformáció. Ekkor A -hoz létezik Jordan-bázis, azaz A felírható Jordan-normálalakban. A Jordan-normálalak egyértelmű abban az értelemben, hogy minden λ_j sajátértékhez annyi Jordan-blokk tartozik, mint a sajátérték geometriai multiplicitása, és a Jordan-blokkok mérete csak az A transzformációtól függ.

Bizonyítás: Mivel a komplex számtest feletti igaz az algebra alaptétele, ezért minden polinom, így A minimálpolinomja is előáll $(t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_s)^{m_s}$ alakban. Alkalmazhatjuk tehát a fent leírt eljárást A Jordan-bázisának konstrukciójához. \square

Következmények

Érdemes látni, hogy a fenti Jordan normálalakról szóló téTEL speciális esetként tartalmaz néhány korábbi, de későbbi fontos eredményünket feltéve, hogy az alaptest a komplex számtest.

Az egyes sajátértékekhez tartozó első Jordan blokk mérete m_j . Ezért a minimálpolinom fokszámánál, a $\sum_{j=1}^s m_j$ számnál a téR dimenziója nem lehet kisebb. Meggondoltuk tehát, hogy a *minimálpolinom foka legfeljebb a téR dimenziója*.

Az is nyilvánvaló, hogy a Jordan-normálalak pontosan akkor diagonális, ha minden Jordan-blokk diagonális. Mivel a legnagyobb méretű Jordan-blokkok az A_j transzformációk első blokkja, aminek mérete m_j , ezért a *Jordan-normálalak pontosan akkor diagonális mátrix, ha minden j mellett $m_j = 1$* , ami éppen azt jelenti, hogy a *minimálpolinom gyökei egyszeresek*. Persze ebben az esetben a Jordan-bázis minden eleme az A egy sajátvektora.

Később teljesen nyilvánvaló lesz, hogy egy Jordan-normálalakú mátrix úgynevezett *karakterisztikus polinomja*

$$(t - \lambda_1)^{k_1} (t - \lambda_2)^{k_2} \dots (t - \lambda_s)^{k_s}$$

alakú, ahol a λ_j sajátérték pontosan k_j -szer szerepel a Jordan-normálalak diagonálisában. Mivel minden λ_j -re a legnagyobb Jordan-blokk éppen $m_j \times m_j$ méretű, ezért $m_j \leq k_j$ minden j -mellett. Ebből persze következik, hogy a minimálpolinom osztója a karakterisztikus polinomnak, ami avval ekvivalens, hogy az A transzformáció karakterisztikus polinomjának is gyöke maga az A transzformáció. Ezt az eredményt gondoltuk meg, mikor a vektortér olyan test felett van értelmezve, ahol minden polinomnak van gyöke. A gondolat általánosabban is igaz, tetszőleges test feletti vektorterek esetére. Ez a *Cayley–Hamilton-tétel*.

Mivel a karakterisztikus polinom minden pontosan egy $\dim(V)$ -ed fokú normált polinom, ezért a Cayley–Hamilton-tételből is nevetve következik, hogy a minimálpolinom minden legfeljebb $\dim(V)$ -ed fokú.

Szokás azt mondani, hogy a $\lambda \in \sigma(A)$ sajátérték *algebrai multiplicitása* k , ha a Jordan-normálalakban λ_j éppen k -szor szerepel. Világos, hogy ez azonos avval, hogy a transzformáció karakterisztikus polinomjának λ egy k -szoros multiplicitású gyöke. Mivel minden egyes Jordan-blokk legalább 1×1 méretű, ezért a λ_j geometriai multiplicitása, azaz a λ_j -hez tartozó Jordan-blokkok száma legfeljebb ezen blokkokban lévő diagonális elemek száma, ergo a λ_j algebrai multiplicitása. Az is világos, hogy e két multiplicitás pontosan akkor egyezik meg, ha minden Jordan-blokk 1×1 méretű, tehát ha diagonalizálható. Arra jutottunk tehát, hogy egy komplex számtest feletti vektortér esetén az alábbi feltételek ekvivalensek.

- A diagonalizálható;

- *A minimálpolinomjának gyökei egyszeresek;*
- *A minden sajátértékének a geometriai és algebrai multiplicitása azonos.*

16. fejezet

Determináns

SÚLYOS ÖNBIZALOMHIÁNYHOZ VEZET tapasztalataim szerint a determináns fogalmának mellőzése a lineáris algebra körében. E jelenség okait inkább a megsokások, a régi beidegződések közt kell keresnünk, mintsem a logikai szükségszerűségek közt. Tény ugyanakkor, hogy egyszerű egy mátrix szingularitására következtetni, ha egy konkrét függvény argumentumába helyettesítve a mátrixunkat zérust kapunk. Hasonlóan egyszerű, ha egy explicit formulánk van, egy képlet amibe csak be kell helyettesíteni, egy reguláris mátrixú inhomogén lineáris egyenletrendszer megoldására, vagy egy mátrix inverzének felírására. A determináns fogalmát valóban nagyon sokszor csak arra használják, hogy a lineáris algebra alap algoritmusát, a Gauss–Jordan-eliminációt, egy explicit formulára cseréljék. Azt szeretném itt hangsúlyozni, hogy pusztán emiatt nem érdemes előnyben részesíteni a determináns használatát. Az algoritmus éppen olyan jól, sőt látni fogjuk: sokkal de sokkal hatékonyabban működik, a mint a képlet.

Amiért mégis fontos ez a fejezet az kisebb részt az, hogy a lineáris algebra tanulásának az is a célja, hogy a későbbi standard felhasználásokat megértsük. Sok-sok közgazdasági és műszaki szakkönyv előszeretettel sűríti a mátrixokkal kapcsolatos tudnivalókat a determináns fogalmának segítségével. Az igazi ok viszont nem ez. Egyetlen helyen van megkerülhetetlen szerepe a transzformáció determinánsának. Ez az, amikor az válik valamiért fontossá, hogy egy lineáris transzformáció, a tér egység kockáját, amelynek a „térfogata” nyilván 1, mekkora térfogatú paralelepipedonba transzformálja.¹ Ez a kérdés a valós analízis fontos és érdemi pontjain merül fel. Például, amikor egy többváltozós függvény integrálját számoljuk valamelyen helyettesítéssel. Ezen a ponton válik majd a determináns fogalma igazából élővé. A fejezet tehát az első lépés ebben az irányban.²

16.1. Permutációk

Emlékezzünk arra, hogy egy n elemű halmaz permutációi $n!$ elemszámú csoportot alkotnak a kompozíció műveletével. Szokásos elnevezés, hogy ezt a kompozíció műveletet szorzásnak nevezzük. Ezt a csoportot a továbbiakban S_n -el jelöljük. A determináns definíciójában majd olyan függvény szerepel, amely egy permutációhoz egy számot rendel, tehát egy $\varphi : S_n \rightarrow \mathbb{F}$ függvényről van szó. Mivel a $\pi \mapsto \pi^{-1}$ hozzárendelés egy $S_n \rightarrow S_n$ bijekció, ezért tetszőleges ilyen függvény mellett

$$\sum_{\pi \in S_n} \varphi(\pi) = \sum_{\pi \in S_n} \varphi(\pi^{-1}),$$

hiszen a két $n!$ összeandóból álló összeg csak az elemek sorrendjében különbözik. Így minden oldal a φ értékkészlete elemeinek összegét jelöli.

A permutáció csoport elemeit, tehát az egyes permutációkat, érdemes a sakktábla bástyafelrakásával azonosítani. Képzeljünk el egy $n \times n$ méretű sakktáblát. A feladat, hogy helyezzünk el a táblára n db bástyát olyan módon, hogy egy bánya se legyen ütésben. Nyilván ez úgy és csak úgy lehetséges, ha minden sorba és minden oszlopba egy bánya raktunk le. A bástyák egy ilyen elhelyezését nevezzük bástyafelrakásnak.

Vegyük észre, hogy n darab bánya minden olyan elrendezése, amelynél egy sorban egy bánya van, egy $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ függénnel azonosítható. Gondolunk a sorok számozására mint az értelmezési

¹A válasz érzelmi része az lesz, hogy a transzformáció determinánsa ennek a „mértéke”.

²Persze fontos az olvasó szociális érzelmi állapota, ami pozitív mellékterméke e fejezet pontos megértésének. ☺

tartomány elemeire. Mivel így minden sorban van bábú, és minden sorban csak egy bábú van, ezért értelmes az a definíció, amely szerint $\pi(j)$ értéke legyen a j -edik sorban lévő bábú oszlopszáma.

Ha a bástyafelrakásra gondolunk, akkor további feltétel, hogy egyetlen oszlopban se legyen két bánya, ami éppen a π függvény injektivitását jelenti. No de, egy véges halmaznak önmagára képező injekciója egyben szürjekció is, emiatt a bástyafelrakások a bijekciókkal azonosíthatók.

A determináns fogalmának megértését nagyban segíti, ha n elem permutációjára mint a n darab bástyafelrakására gondolunk.³ Az id identikus permutációt például a diagonális bástyafelrakás jelképezi. Ha adott egy π permutációhoz tartozó bástyafelrakás, akkor a π^{-1} inverzhez tartozó bástyafelrakást a diagonálisra való tükrözéssel kapjuk.

16.1. definíció (transzpozíció). Egy permutációt *transzpozíciónak* nevezünk, ha az csak két elemet cserél fel.

Egy bástyafelrakásra gondolva, az akkor transzpozíció, ha az identitásnak megfelelő diagonális felrakás két sorának felcserélésével keletkezik. Emiatt transzpozíciók szorzata egy olyan permutáció, amely a diagonális felrakásból néhány sor felcserélésével keletkezik.

Na most, gondoljunk bele, hogy minden bástyafelrakás véges sok sor felcserélésével az identitás permutációt reprezentáló diagonális felrakásba vihető át. Az algoritmus a következő. Keressük meg azt a sort, ahol a bánya az első oszlopban áll. Cseréljük fel ezt a sort az első sorral. Most keressük meg azt a sort, ahol a bánya a második oszlopot foglalja. Cseréljük fel ezt a sort a második sorral. Stb. Az $n - 1$ -edik diagonális elem helyre kerülése után, az utolsó sorban, a bánya az utolsó pozíciót foglalja el, hiszen a sorok cserélgetésével megmarad a táblázatnak az a tulajdonsága, hogy minden sorban és minden oszlopban pontosan egy bánya van.

Adott π permutációhoz találtunk tehát $\sigma_1, \dots, \sigma_{n-1}$ transzpozíciókat, amelyekre $\pi \cdot \sigma_1 \cdots \sigma_{n-1} = \text{id}$, amiből persze $\pi = \sigma_{n-1}^{-1} \cdots \sigma_1^{-1}$ következik. Egy transzpozíció inverze is transzpozíció, tehát igazoltuk a könyvtárak klasszikus felíratáról szóló állítást.

16.2. állítás. *Minden permutáció transzpozíciók szorzata.*

Persze szó nincs arról, hogy a fenti előállítás egyértelmű lenne. Még nagyon sok más algoritmust is kitalálhatunk a sorba rendezésre, olyanokat is, amelyek minden lépésükben csak két elemet cserélnek fel.⁴ Annyi viszont igaz, hogy amennyiben egy permutáció páros (páratlan) sok transzpozíció szorzata, akkor minden más előállításában is páros (páratlan) sok transzpozíció szerepel. Kisvártatva látjuk majd, hogy ez miért van így.

A permutáció paritása

Egy kényelmes jelölést vezetünk be egy permutációnak és egy előre adott transzpozíciónak a szorzatára.

16.3. definíció. Legyenek $1 \leq i, j \leq n$ különböző egészek, és $\pi \in S_n$ egy permutáció. Definiálja $\pi^* \in S_n$ a π permutációinak és az (i, j) párt felcserélő transzpozícióinak a szorzatát, azaz

$$\pi^*(k) \doteq \begin{cases} \pi(k) & , \text{ha } k \neq i \text{ és } k \neq j; \\ \pi(j) & , \text{ha } k = i; \\ \pi(i) & , \text{ha } k = j. \end{cases}$$

Látható, hogy a π -t reprezentáló bástyafelrakásból az i és j sorok felcserélésével kapjuk a π^* bástyafelrakását.

Az is világos, hogy rögzített (i, j) transzpozíció mellett a $\pi \mapsto \pi^*$ egy $S_n \rightarrow S_n$ bijekció, emiatt újra elmondható, a fejezet eleji megjegyzés, tehát ha $\varphi : S_n \rightarrow \mathbb{F}$ egy tetszőleges függvény, akkor

$$\sum_{\pi \in S_n} \varphi(\pi) = \sum_{\pi \in S_n} \varphi(\pi^*).$$

³Persze ez ugyanaz mintha olyan $n \times n$ mátrixokról beszélünk, ahol csak 0 és 1-es szerepel, mégpedig úgy, hogy minden sorban és minden oszlopban pontosan egy db 1-es van. Az 1-esek játszák a bányaok szerepét. Az ilyen mátrixokat szoktuk *permutáció mátrixnak* mondani.

⁴Keressük például a 'bubble sort' kifejezésre.

16.4. definíció (Inverzió). Legyenek újra $1 \leq i, j \leq n$ különböző egészek, és $\pi \in S_n$ egy permutáció. Azt mondjuk, hogy az i -edik és a j -edik sorban álló bástyák *inverzióban vannak*, ha

$$(\pi(j) - \pi(i))(j - i) < 0.$$

A fenti jelölések mellett az i és j -edik sor bástyái pontosan akkor vannak inverzióban, ha $j > i$ esetén $\pi(j) < \pi(i)$. Gondoljuk meg, hogy az identitás permutáció az egyetlen olyan a permutációk között, melynél nincs inverzióban álló bástyapár.

16.5. állítás. minden $\pi \in S_n$ permutációra, a π -nél és π^{-1} -nél inverzióban álló bástyapárok száma azonos.

Bizonyítás: Világos, hogy

$$(\pi(j) - \pi(i))(j - i) = (j - i)(\pi(j) - \pi(i)) = (\pi^{-1}(\pi(j)) - \pi^{-1}(\pi(i)))(\pi(j) - \pi(i)).$$

Ez azt jelenti, hogy az i és j sorokat foglaló bástyapár pontosan akkor van inverzióban π -szerint, ha a $\pi(i)$ és a $\pi(j)$ sorokat foglaló bástyapár inverzióban van π^{-1} szerint. Mivel az $(i, j) \mapsto (\pi(i), \pi(j))$ leképezés az $\{(i, j) : i \neq j : 1 \leq i, j \leq n\}$ halmaznak önmagára képező bijekciója, ezért kölcsönösen egyértelmű leképezést találtunk a π és a π^* inverzióban álló bástyapárjai között. \square

16.6. definíció (Páros, páratlan permutáció). Egy $\pi \in S_n$ permutációt párosnak (páratlanak) nevezünk, ha π -nél inverzióban lévő párok száma páros (páratlan). A permutáció előjele:

$$\operatorname{sgn} \pi \doteq \begin{cases} +1 & , \text{ha } \pi \text{ páros;} \\ -1 & , \text{ha } \pi \text{ páratlan.} \end{cases}$$

Egy permutáció előjelét meghatározó algoritmus a definíció szerint a következő: Tekintsük a π permutáció bástyafelrakását. Hasonlítsuk össze az összes különböző – tehát $\binom{n}{2}$ darab – bástyapárnak az oszlopszámait a sorszámaival. Ha a nagyobb sorindexű bánya oszlopszáma a kisebb, akkor ők inverzióban vannak. Ha k jelöli az inverzióban álló bástyapárok számát, akkor $\operatorname{sgn}(\pi) = (-1)^k$.

Az identitás permutációban zérus az inverziók száma, tehát $\operatorname{sgn}(\text{id}) = +1$. Mivel egy permutációt és az inverzének azonos számú inverziója van, ezért az előjük is azonosak:

$$\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1}).$$

A következő állítás szerint, egy bástyafelrakás előjele két sorának felcserélésével az ellenkezőjére vált.

16.7. állítás. Jelölje π^* a π permutációjának és egy adott transzpozíciójának a szorzatát. Ekkor

$$\operatorname{sgn} \pi = -\operatorname{sgn} \pi^*.$$

Bizonyítás: Tegyük fel, hogy π^* a π -ből az i és j -edik sorok felcserélésével keletkezik, ahol $i < j$. Leszámoljuk, hogy hány az u, v sorokat foglaló bástyapár van inverzióban π és π^* mellett. Az esetek szétválasztása folyamán végig páros számban tér el egymástól a π^* és π -melletti inverziók száma, míg az utolsó esetben pontosan egy a különbség. Ily módon a π^* és π -melletti inverziók összességében páratlan számban különböznak, ami igazolja az állítást. Az eset szétválasztás aszerint történik, hogy az egyes esetekben számlolandó u és v sorbeli bástyák milyen kapcsolatban vannak, a felcserélendő i és j sorokkal. Lássuk tehát a különböző eseteket:

1. Az u és v egyike sem egyezik i és j egyikével sem.

Világos, hogy ebben az esetben a π melletti és a π^* melletti inverziók száma azonos, hiszen egyik itt vizsgált bástyapár sem mozdul a π -ről π^* -ra való áttérés során.

2. Az u megegyezik az i és j közül az egyikkel, de $v > j$.

Tekintsük egy u, v sorbeli bástyapárt. Az i -edik és a j -edik sor felcseréléséve az u -adik sorban álló bánya megmozdul, de az oszlop koordinátáját megtartja, és sor koordinátája is v -nél kisebb marad. Ez azt jelenti, hogy ugyanazok bástyapárok lesznek inverzióban a π -nél, mint π^* -nál, ergo a π -nél és π^* -nál inverzióban álló párok száma azonos.

3. Az u megegyezik az i és j közül az egyikkel, de $i < v < j$.

Különböztessünk meg ezen belül is három esetet.

- i. Ha $\pi(v) > \max(\pi(j), \pi(i))$.
Ekkor a v -edik sor bástyája az i -sorbeli bástyával nincs inverzióban, de a j -sorbeli bástyával inverzióban van. A π -nél tehát az inverziók száma 2. Ugyanez a helyzet π^* -nál is, ergo a π melletti és a π^* melletti inverziók száma azonos.
 - ii. Ha $\pi(v) < \min(\pi(j), \pi(i))$.
Analóg [i]-vel.
 - iii. Ha $\min(\pi(j), \pi(i)) < \pi(v) < \max(\pi(j), \pi(i))$.
Ha $\pi(j) < \pi(i)$, akkor az inverziók száma π -nél 2, π^* -nál zérus. Ha viszont $\pi(j) > \pi(i)$, akkor az inverziók száma π -nél 0, de π^* -nál 2. Mindkét esetben tehát 2 a különbség, ezért bárhogy is van, de a π melletti és a π^* melletti inverziók száma páros számban különbözik.
4. Az u megegyezik az i és j közül az egyikkel, de $v < i$.
Analóg a második esettel.
5. Az u is és a v is megegyezik az i és j egyikével.
Ekkor ha az u és v sorbeli bástyák inverzióban vannak π -nél, akkor nem lesznek inverzióban π^* -nál, és hasonlóan, ha nincsenek inverzióban π -nél, akkor inverzióban lesznek π^* mellett. Bárhogyan is van, de a π -nél és π^* -nál inverzióban lévő elemek száma 1-el változik.
- Világos, hogy minden u, v pár a fenti esetek közül pontosan az egyikbe tartozik. Összességében a π -ről π^* -ra való áttéréssel a transzpozíciók száma páratlan számmal változik, tehát párosból páratlan válik, vagy páratlanból párossá. \square
- Nyilvánvaló tehát, hogy az S_n csoporthnak $n!/2$ számú páros és ugyanennyi páratlan eleme van.
- Mivel egy transzpozícióval való szorzás a permutáció előjelét ellenérettéje állítja, ezért ha egy permutáció k darab transzpozíció szorzataként áll elő az identikus permutációból, akkor annak előjele $(-1)^k$. Ebből azonnal következik, hogy egy páros (páratlan) permutáció nem áll elő mint páratlan (páros) sok transzpozíció szorzata.
- Meggondoltuk tehát, hogy egy permutációval a transzpozíciók szorzataként előállítása ugyan nem egyértelmű, de az előállításban szereplő transzpozíciók számának paritása minden azonos. Ez egyben egy alternatív algoritmust eredményez egy permutáció előjelének meghatározására: *Tudjuk, hogy minden bástyafelrakás átrendezhető a diagonális felrakásba pusztán a sorok felcserélésével. Ha a π permutáció esetén k darab sorcserére van elhely szükség, akkor $\text{sgn } \pi = (-1)^k$.*

16.2. Mértékek

A determinánsról szóló minden további gondolatban feltesszük, hogy az \mathbb{F} testben $1 + 1 = 0$ nem teljesül. Ezt úgy fejezzük ki, hogy az \mathbb{F} test nem 2 karakaterisztikájú.

16.8. definíció (mérték). Legyenek V_1, V_2 és V ugyanazon test feletti vektorterek. Az $A : V_1 \times V_2 \rightarrow V$ leképezést *bilineáris operátornak* nevezzük, ha:

1. minden rögzített $x \in V_1$ esetén az $A(x, \cdot) : V_2 \rightarrow V$ lineáris operátor.
2. minden rögzített $x \in V_2$ esetén az $A(\cdot, x) : V_1 \rightarrow V$ lineáris operátor.

Ha valamely $A : V \times V \rightarrow V$ bilineáris operátorra minden $u, v \in V$ esetén $A(u, v) = -A(v, u)$, akkor A -t *anti-szimmetrikus bilineáris operátornak* nevezzük.

Ha valamely $A : V \times V \rightarrow V$ bilineáris operátorra minden $u \in V$ esetén $A(u, u) = 0$, akkor A -t *alternáló bilineáris operátornak* nevezzük.

Legyenek most V_1, V_2, \dots, V_k és V ugyanazon test feletti vektorterek. Az $A : V_1 \times \dots \times V_k \rightarrow V$ leképezést *multi-lineáris operátornak* vagy *n-lineáris operátornak* nevezzük, ha bármely rögzített

$$x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{j-1}, x_{j+1}, \dots, x_k$$

vektor mellett az

$$A(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_{j-1}, \cdot, x_{j+1}, \dots, x_k) : V_i \times V_j \rightarrow V$$

függvény bilineáris operátor. Ha a fenti függvény anti-szimmetrikus (alternáló) operátor minden i -re j -re és tetszőlegesen rögzített vektorokra, akkor az $A : V_1 \times \dots \times V_k \rightarrow V$ leképezést *anti-szimmetrikus (alternáló) multi-lineáris operátornak* nevezzük.

Ha V egy nem 2 karakterisztikájú test feletti n -dimenziós vektortér, akkor egy $V \times \dots \times V \rightarrow V$ n -lineáris anti-szimmetrikus függvényt a V vektortér feletti mértéknak nevezzünk.

16.9. állítás. Legyen V olyan \mathbb{F} test feletti vektortér, ahol $a \cdot 1 + 1 \neq 0$. Tegyük fel, hogy $A : V \times V \rightarrow V$ bilineáris függvény. Ekkor az alábbi három feltevés egymással ekvivalens.

1. Az A anti-szimmetrikus;
2. Az A alternáló;
3. minden $u, v \in V$ -re valamint tetszőleges $\lambda, \mu \in \mathbb{F}$ skalárokra

$$A(u, v + \lambda u) = A(u, v) = A(u + \mu v, v).$$

Bizonyítás: 1. \Rightarrow 2. és 2. \Rightarrow 3. nyilvánvaló. 3. \Rightarrow 1.: $A(u, v) = A(u, v - u) = A(u + v - u, v - u) = A(v, v - u) = A(v, -u) = -A(v, u)$. \square

Világos, hogy tetszőleges k pozitív egész mellett, k -lineáris függvények vektorteret alkotnak. E vektor tér egy altere a k -lineáris anti-szimmetrikus függvények vektortere. Ez persze fennáll akkor is, mikor n -dimenziós vektortérnek n -szeres szorzatán értelmezett n -lineáris függvényekről van szó, ergo a mértékek egy vektorteret alkotnak. A konstans zérus mérték így nyilvánvaló példa mértékre. A kérdés, hogy van-e más mérték?

Példa mértékre

Egy n dimenziós tér n -szeres Descartes-szorzatán értelmezett konstans zérus leképezés egy triviális példa mértékre. Az alábbiakban egy olyan mértéket készítünk, ami egészen biztosan nem a konstans zérus függvény.

16.10. definíció. Legyen V egy vektortér, amelyben rögzített az $\{e_1, \dots, e_n\}$ bázis. Mivel lineáris funkcionál egy bázison tetszőlegesen és egyértelműen előírható, ezért tetszőleges $1 \leq i \leq n$ -hez létezik egyetlen $e_i^* : V \rightarrow \mathbb{F}$ lineáris funkcionál, amelyre

$$e_i^*(e_j) = \delta_{i,j}$$

fennáll. Ezt az e_i^* lineáris funkcionált nevezzük az i -edik duális báziselemeinek.⁵

16.11. állítás. Rögzítsük az $\{e_1, \dots, e_n\} \subseteq V$ bázist. Ekkor minden $v \in V$ mellett $v = \sum_{j=1}^n e_j^*(v) e_j$, azaz v -nek az $\{e_1, \dots, e_n\}$ bázisban felírt j -edik koordinátája $e_j^* v$.

Bizonyítás: Legyen $v = \sum_{i=1}^n \alpha_i e_i$. Ekkor tetszőlegesen rögzített j index mellett

$$e_j^*(v) = e_j^* \left(\sum_{i=1}^n \alpha_i e_i \right) = \sum_{i=1}^n \alpha_i e_j^*(e_i) = \sum_{i=1}^n \alpha_i \delta_{j,i} = \alpha_j.$$

\square

Úgy interpretálhatjuk tehát az e_j^* lineáris funkcionált, hogy az minden v vektorhoz hozzárendeli a vektor j -edik koordinátáját.

16.12. állítás. Rögzítsünk az n -dimenziós V vektortérben egy $\{e_1, \dots, e_n\}$ bázist. Ekkor minden $v_1, \dots, v_n \in V$ vektor esetén

$$\sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_i^* v_{\pi(i)} = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_{\pi(i)}^* v_i.$$

Képzeljük el, hogy egy $n \times n$ méretű sakktábla j -edik oszlopába írjuk a v_j vektor koordinátáit, minden $j = 1, \dots, n$ mellett. Tegyük fel erre a sakktáblára egy π permutációt, azaz a π bástyafelrakását. minden i -re az i -edik sor $\pi(i)$ -edik elemén áll tehát egy bánya. No de, a $\pi(i)$ edik oszlop i -edik helyére $v_{\pi(i)}$ vektor i -edik koordinátája, azaz $e_i^* v_{\pi(i)}$ van írva. Így az $\prod_{i=1}^n e_i^* v_{\pi(i)}$ szám egyszerűen a bástyafelrakásban a bányaok taposta számok szorzata. A baloldali számhoz ezt kell megszorozni még a szóban forgó permutáció előjelével, majd ugyanezt az összes permutációra megismételni és a kapott számok összegét képezni.

⁵Az elnevezés oka, hogy $\{e_1^*, \dots, e_n^*\}$ bázisa az összes $V \rightarrow \mathbb{F}$ lineáris funkcionálok vektorterének.

A jobboldali szám interpretációját kapjuk, ha most a v_1, \dots, v_n vektorok koordinátáit rendre a sorokba írjuk. Újra gondoljunk egy π permutáció bástyafelrakására. minden i -re az i -edik sor $\pi(i)$ -edik elemén áll tehát egy bánya, ami most az i -edik vektornak tehát v_i -nek, a $\pi(i)$ -edik koordinátáját azaz $e_{\pi(i)}^* v_i$ -t tapossa. Így az $\prod_{i=1}^n e_{\pi(i)}^* v_i$ szám egyszerűen a bástyafelrakásban a bányaok taposta számok szorzata. A jobboldali számhoz ezt kell megszorozni még a szóban forgó permutáció előjelével, majd ugyanezt az összes permutációra megismételni és a kapott számok összegét képezni.

Az állítás szerint tehát, mindegy hogy a sorokra vagy az oszlopokra másoljuk a v_1, \dots, v_n vektorok koordinátáit, a fenti eljárás mindenkor ugyanazt a számot eredményezi.⁶

Bizonyítás: A számolásban kihasználjuk, hogy minden $\pi \in S_n$ mellett $\operatorname{sgn} \pi = \operatorname{sgn} \pi^{-1}$, és azt, hogy a $\pi \mapsto \pi^{-1}$ hozzárendelés egy bijekció az S_n halmazon.

$$\sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_i^* v_{\pi(i)} = \sum_{\pi \in S_n} \operatorname{sgn} \pi^{-1} \prod_{k=1}^n e_{\pi^{-1}(k)}^* v_k = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{k=1}^n e_{\pi(k)}^* v_k. \quad \square$$

16.13. állítás („Példa” mértékre). Rögzítsünk az n -dimenziós V vektortérben egy $\{e_1, \dots, e_n\}$ bázist. Ekkor

$$d_{\{e_1, \dots, e_n\}}(v_1, \dots, v_n) = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_i^* v_{\pi(i)} = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_{\pi(i)}^* v_i$$

Ekkor $d_{\{e_1, \dots, e_n\}} : V \times \dots \times V \rightarrow \mathbb{F}$ egy olyan mérték, amelyre $d_{\{e_1, \dots, e_n\}}(e_1, \dots, e_n) = 1$.⁷

Bizonyítás: Világos, hogy $d(e_1, \dots, e_n) = 1$, hiszen ha π nem az identikus permutáció, akkor van olyan i melyre $\pi(i) \neq i$, az $e_i^*(e_{\pi(i)}) = 0$. Így a $d(e_1, \dots, e_n)$ definíciójában az $n!$ darab permutáció minden elemére 0-t kapunk, kivétel az identikus permutáció, amelynek előjele +1 és $\prod_{i=1}^n e_i^* e_{\operatorname{id}(i)} = \prod_{i=1}^n \delta_{i,i} = 1$.

A linearitás könnyen következik az e_i^* duális bázis elemek linearitásából. Mondjuk ha a hátsó $n-1$ változót rögzítjük, akkor az első változót szabadon hagyva

$$d(\cdot, v_2, \dots, v_n) = \sum_{\pi \in S_n} \operatorname{sgn} \pi \cdot \left(\prod_{k=2}^n e_{\pi(k)}^* v_k \right) \cdot e_{\pi(1)}^*.$$

Mivel a lineáris funkcionálok vektorteret alkotnak, ezért a $d(\cdot, v_2, \dots, v_n)$ függvény is egy lineáris funkcionál. Analóg módon a többi változót szabadon hagyva kapjuk, hogy d egy n -lineáris függvény.

Az anti-szimmetria igazolásához szükséges, a $\operatorname{sgn} \pi = -\operatorname{sgn} \pi^*$ összefüggés. Például az első két változóra:

$$\begin{aligned} d(x, y, v_3, \dots, v_n) &= \\ \sum_{\pi \in S_n} \operatorname{sgn} \pi \left(\prod_{k=3}^n e_{\pi(k)}^* v_k \right) e_{\pi(1)}^* x \cdot e_{\pi(2)}^* y &= \sum_{\pi \in S_n} (-\operatorname{sgn} \pi^*) \left(\prod_{k=3}^n e_{\pi^*(k)}^* v_k \right) e_{\pi^*(1)}^* y \cdot e_{\pi^*(2)}^* x = \\ - \sum_{\pi \in S_n} \operatorname{sgn} \pi \left(\prod_{k=3}^n e_{\pi(k)}^* v_k \right) e_{\pi(1)}^* y \cdot e_{\pi(2)}^* x &= -d(y, x, v_3, \dots, v_n). \end{aligned}$$

Itt π^* a π permutációinak és az $(1, 2)$ elemeket felcserélő transzpozícióinak a szorzata. Az utolsó egyenlőségenben azt használtuk ki, hogy ha π befütje az S_n csoport összes elemét, akkor az evvel párhuzamosan számolt π^* permutációk is végig mennek a permutáció csoport minden elemén.

Analóg módon kapjuk, hogy bármely két változót felcserélve d értéke az ellentettjére változik. \square

Meggondoltuk tehát, hogy egy n -dimenziós vektortér feletti mértékek vektortere legalább egydimenziós, hiszen van e vektortérnek a zérustól különböző eleme. Azt mutattuk meg, hogy adott $\{e_1, \dots, e_n\}$ bázis rögzítése mellett, az ehhez a bázishoz tartozó $d_{\{e_1, \dots, e_n\}}$ függvény egy nem triviális mérték V felett. Persze sok-sok bázist tudunk rögzíteni, így sok-sok mértéket is tudunk konstruálni. Leginkább pedig az általunk konstruált mértékeken felül is lehet még mérték a V felett. Ehhez képest nem sokára látni fogjuk, hogy a V feletti mértékek vektortere pontosan egydimenziós vektortér. A legfontosabb gondolat, hogy a fenti d -nek csak konstans szorosai lehetnek a V feletti mértékek, de megszakítjuk itt a tárgyalást, azért hogy a fent definiált mérték fontosságát hangsúlyozzuk.

⁶E hosszú, de ultra fontos interpretáció után nézzük az egysoros bizonyítást.

⁷Ha a szövegkörnyezetből világos, hogy melyik a rögzített bázis, akkor a kicsit nehézkes $d_{\{e_1, \dots, e_n\}}$ jelölés helyett csak d -t használunk.

Mátrix determinánса

Gondolunk erre a szakaszra mint egy érdekes kitérőre, ha már van a kezünkben egy igazi – értsd nem triviális – mérték.

16.14. definíció (Mátrix determinánса). Legyen $[A]$ egy mátrix. E mátrix $|[A]|$ *determinánса* a következő szám:

$$|[A]| = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n [A]_{i,\pi(i)}.$$

Az előbb bevezetett d függvény és a mátrix determinánсанak kapcsolata.

16.15. állítás. Legyen V egy vektortér, $\{e_1, \dots, e_n\}$ egy bázis, és $a_1, \dots, a_n \in V$ vektorok. Az A mátrix oszlopai legyenek rendre az a_1, \dots, a_n vektorok fenti bázisban felírt koordinátavektorai. Ekkor

$$|[A]| = d_{\{e_1, \dots, e_n\}}(a_1, \dots, a_n) = |[A]|^T.$$

Bizonyítás: Elég azt látni, hogy $[A]_{i,\pi(i)}$ a mátrix $\pi(i)$ -edik oszlopának, tehát $a_{\pi(i)}$ -nek i -edik koordinátája, azaz a $e_i^*(a_{\pi(i)})$. Így

$$\begin{aligned} |[A]| &= \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n [A]_{i,\pi(i)} \\ &= \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_i^*(a_{\pi(i)}) = d(a_1, \dots, a_n) = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_{\pi(i)}^*(a_i) \\ &= \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n [A]_{\pi(i),i} = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n [A]_{i,\pi(i)}^T = |[A]_{\{e_1, \dots, e_n\}}^T|. \end{aligned}$$

Ezt kellett belátni. □

Mátrix determinánсанak legfontosabb tulajdonságai

Összefoglaljuk, amit eddig a pontig megértettünk a mátrix determinánсанak fogalmáról.⁸ Legyen tehát $[A]$ egy mátrix. Ekkor

1. $|[A]| = |[A]|^T$;
2. $|[A]|$ az oszlopok (sorok) lineáris függvénye;
3. $|[A]|$ értéke ellentetjére változik két oszlopának (sorának) felcserélése után;
4. $|[A]|$ értéke zérus, ha két azonos oszlopa (sora) van;
5. $|[A]|$ értéke nem változik, ha egyik oszlopához (sorához) egy másik oszlop (sor) skalárszorosát adjuk.

Egy permutáció mátrix determinánса, a permutáció előjele. Ugyanis, ha π permutációhoz tartozik a permutáció mátrix, akkor minden $\sigma \neq \pi$ permutáció mellett a σ permutáció egyik bástyája (sőt legalább két bástyája) zérus elemen áll. Így a determinánса definíciója szerint az $n!$ elemszámú összegben a π kivételével minden összeadandó zérus. Speciálisan, az identitás mátrix determinánса 1.

Egy felső- (alsó-)háromszög alakú mátrix determinánса a diagonális elemek szorzata. Ugyanis minden $\pi \neq \text{id}$ permutáció esetén van egy bánya a diagonális alatt (és egy a diagonális felett). Ez azt jelenti, hogy a definícióban csak az identitás permutációhoz tartozó tag nem feltétlen zérus. Az identitáshoz tartozó tag pedig a diagonális elemek szorzata. Speciálisan, az identitás mátrix determinánса 1.

⁸Mielőtt tovább megyünk adjunk képletet 1×1 -es, 2×2 -es és 3×3 -as mátrixok determinánсанak kiszámolására, a definíció alapján. Érdemes figyelni arra, mennyire nő a definíció komplexitása a méret emelésével.

Egy mátrix determinánsa Gauss–Jordan-eliminációval. Képzeljük el, hogy adott egy reguláris mátrix, tehát olyan, amelynek oszlopai lineárisan függetlenek. Az eliminációs algoritmusnak arra az alakjára gondoljunk először, amikor a bázisba bevont oszlopvektornak is kiírjuk az új bázisban a koordinátáit. Ilyen módon végig $n \times n$ méretű táblázatokkal dolgozunk. Tegyük fel, hogy egy táblázat A_1 az ebből következő táblázat A_2 . A_2 -t úgy kaptuk A_1 -ből, hogy a pivot elemmel osztottuk a pivot elem sorát, majd minden más sorhoz hozzáadtuk a pivot elem sorának skalárszorosát. Ha r_1 jelöli az A_1 -ben választott pivot-elemet, akkor

$$|A_2| = \frac{1}{r_1} |A_1|.$$

Mivel A_1 oszlopai lineárisan függetlenek, ezért minden oszlop a bázisba bevonható, ami azt jelenti, hogy az utolsó táblázat – az A_{n+1} – egy permutáció mátrix. Tehát az előbbi összefüggést lépésenként alkalmazva

$$|A_1| = r_1 \cdot |A_2| = r_1 \cdot r_2 \cdot |A_3| = \dots = r_1 \cdot r_2 \dots r_n \cdot |A_{n+1}| = \left(\prod_{j=1}^n r_j \right) \cdot \text{sgn}(A_{n+1}).$$

Ezek szerint az algoritmus a következő: *A determinánsának kiszámításához az összes oszlopot Gauss–Jordan-eliminációval bevisszük a bázisba. A determináns a pivot elemek szorzata szorozva +1, vagy -1-el. Az utolsó táblázatban, az új bázis elemeit az eredeti oszlopvektorok alkotják. Számoljuk meg, hogy hány sorcserét kell kapnunk ahoz, hogy az eredeti oszlopvektorok természetes sorrendjébe cseréljük az elimináció végén kapott új bázis elemeket. Ha ehhez páratlan sok csere kellett, akkor -1-el kell még szoroznunk a pivot elemek szorzatát. Egyébként a determináns a pivot elemek szorzata.*⁹

Az oszlopok (sorok) lineáris függetlensége elődönthető a determináns segítségével. Láttuk ugyanis, hogy a mátrix determinánsa nem változik, ha egyik oszlopához a másik skalárszorosát adjuk. Ha tehát az egyik oszlop felírható a többi lineáris kombinációjaként, akkor ehhez az oszlophoz véges sokszor egy másik oszlop skalárszorosát adva olyan mátrixot kapunk, amelyben ez az oszlop csupa zérus, és a determinánsa azonos a kiinduló mátrix determinánsával. Láttuk tehát, hogy ha az oszlopok lineárisan összefüggők, akkor a mátrix determinánsa zérus. Viszont ha az oszlopok lineárisan függetlenek, akkor Gauss–Jordan-eliminációval minden oszlop a bázisba vihető, és a determináns előjeltől eltekintve a pivot elemek szorzata. E szorzat nem lehet zérus, hiszen egyik pivot elem sem lehet zérus. Meggondoltuk tehát: *Egy mátrix determinánsa pontosan akkor zérus ha az oszlopok (sorok) lineárisan összefüggők.*

Particionált mátrix determinánsa. Tegyük fel, hogy A egy $(n+m) \times (n+m)$ méretű mátrix, amely

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

módon van particionálva, ahol $A_{1,1}$ egy $n \times n$, $A_{2,2}$ egy $m \times m$ méretű négyzetes mátrix, továbbá $A_{2,1}$ egy $m \times n$ méretű és $A_{1,2}$ egy $n \times m$ méretű részmátrix. Tegyük fel, hogy $A_{2,1}$ minden eleme zérus. Ekkor

$$|A| = |A_{1,1}| \cdot |A_{2,2}|.$$

Definíció alapján: Látható ugyanis, hogy ha vesszük az A egy bástyafelrakását, akkor ha kerül bánya az $A_{1,2}$ részmátrixra, akkor kerül bánya az $A_{2,1}$ részmátrixra is. Így a determináns definíciójához elég csak olyan $\pi \in S_{n+m}$ permutációt venni, amelyre $\pi(i) \leq n$, ha $i \leq n$ és $\pi(i) > n$, ha $i > n$. Egy ilyen π permutációt össze lehet állítani $\pi = (\sigma, \tau)$ alakban, ahol σ az $\{1, \dots, n\}$ halmaz permutációja és τ az $\{n+1, \dots, n+m\}$ halmaz permutációja. Ha π egy ilyen permutáció, akkor olyan inverziója nincs, aminek egyik bánya $A_{1,1}$ -ben van és másik bánya $A_{2,2}$ -ben van. Ez azt jelenti, hogy ha k a σ inverzióinak száma és l a τ inverzióinak száma, akkor a π inverzióinak száma éppen $k + l$. Tehát

$$\text{sgn } \pi = (-1)^{k+l} = (-1)^k (-1)^l = \text{sgn } \sigma \cdot \text{sgn } \tau.$$

Jelölje a bizonyítás végéig, S_n az $\{1, \dots, n\}$ halmaz, S_m az $\{n+1, \dots, n+m\}$ halmaz partíció csoportjait, továbbá S'_{n+m} az $\{1, \dots, n+m\}$ halmaz bijekciói közül azon π partíciókat, amelyek előállnak $\pi = (\sigma, \tau) \in$

⁹A determináns kiszámolásának ez a preferált és professzionális módja. Vessük össze a definíció művelet igényével mondjuk 3,4,5 dimenzió mellett.

$S_n \times S_m$ alakban. Ekkor

$$\begin{aligned}
 |A_{1,1}| \cdot |A_{2,2}| &= \left(\sum_{\sigma \in S_n} \operatorname{sgn} \sigma \prod_{i=1}^n a_{i,\sigma(i)} \right) \left(\sum_{\tau \in S_m} \operatorname{sgn} \tau \prod_{i=1}^m a_{n+i,\tau(n+i)} \right) \\
 &= \sum_{(\sigma, \tau) \in S_n \times S_m} \operatorname{sgn} \sigma \operatorname{sgn} \tau \prod_{i=1}^n a_{i,\sigma(i)} \prod_{i=n+1}^{n+m} a_{i,\tau(i)} = \sum_{\pi \in S'_{n+m}} \operatorname{sgn} \pi \prod_{i=1}^{n+m} a_{i,\pi(i)} \\
 &= \sum_{\pi \in S_{n+m}} \operatorname{sgn} \pi \prod_{i=1}^{n+m} a_{i,\pi(i)} = |A|.
 \end{aligned}$$

Ezt kellett indokolni. \square

Gauss–Jordan-elimináció alapján: Először is azt vegyük észre, hogy az $A_{2,1}$ nullmátrix léte miatt, ha A oszlopai lineárisan összefüggők, akkor az $A_{1,1}$ és $A_{2,2}$ mátrixok egyikének oszlopai is lineárisan összefüggők. Az állítás tehát teljesül, ha $|A| = 0$. Feltesszük, hogy A oszlopai lineárisan függetlenek. Végezzük Gauss–Jordan-eliminációt az A mátrixon de úgy, hogy az első n lépésekben $A_{1,1}$ részéből, utána $A_{2,2}$ részéből választjuk a pivot elemeket. Ez megtehető, hiszen $A_{1,1}$ oszlopai is lineárisan függetlenek. Legyenek a pivot elemek $r_1, \dots, r_n, r_{n+1}, \dots, r_m$ és az utolsó oszlop bevonása után maradt $n + m \times n + m$ méretű permutáció mátrix P . Mivel csak a balfelső $n \times n$ -es és a jobbalsó $m \times m$ részből választottunk pivot elemeket, ezért a P

mátrix $P = \begin{pmatrix} P_{1,1} & 0 \\ 0 & P_{2,2} \end{pmatrix}$ alakú, ahol $P_{1,1}$ egy $n \times n$ méretű, és $P_{2,2}$ egy $m \times m$ méretű permutáció mátrix.

Ha k – illetve l – sorcsere kell ahhoz, hogy $P_{1,1}$ -et – illetve $P_{2,2}$ -t – a diagonálisba transzformáljuk, akkor $k + l$ sorcserével P is a diagonális permutációba megy át. Mivel egy permutáció determinánsa a permutáció előjele, ezért $|P| = (-1)^{k+l} = (-1)^k (-1)^l = |P_{1,1}| |P_{2,2}|$. Világos, hogy $|A_1| = r_1 \dots r_n |P_{1,1}|$. No de, $A_{2,1}$ a nullmátrix, ezért az elimináció első n lépéseiben $A_{2,2}$ változatlan marad, így $|A_{2,2}| = r_{n+1} \dots r_{n+m} |P_{2,2}|$. Összefoglalva:

$$|A| = r_1 \dots r_n \cdot r_{n+1} \dots r_{n+m} |P| = (r_1 \dots r_n |P_{1,1}|) (r_{n+1} \dots r_{n+m} |P_{2,2}|) = |A_{1,1}| \cdot |A_{2,2}|. \quad \square$$

16.3. A mértékek jellemzése

16.16. állítás. Legyen V egy n -dimenziós vektortér, f egy V feletti mérték. Ha a $\{v_1, \dots, v_n\} \subseteq V$ vektorrendszer lineárisan összefüggő, akkor $f(v_1, \dots, v_n) = 0$.

Bizonyítás: Mivel f értéke nem változik ha egy változójához a másik skalárszorosát hozzá adjuk, ezért ha $v_k = \sum_{i \neq k} \alpha_i v_i$, akkor $f(v_1, \dots, v_n) = f(v_1, \dots, 0, \dots, v_n) = 0$. \square

16.17. állítás. Legyen V egy n -dimenziós vektortér, és f egy V feletti mérték. Ekkor ha van olyan e_1, \dots, e_n bázisa a térrnek melyekre $f(e_1, \dots, e_n) = 0$, akkor f a konstans zérő, azaz a triviális mérték.

Bizonyítás: Mivel f anti-szimmetrikus és minden permutáció előáll transzpozíciók szorzataként, ezért tetszőleges $\pi \in S_n$ mellett $|f(e_{\pi(1)}, e_{\pi(2)}, \dots, e_{\pi(n)})| = |f(e_1, \dots, e_n)| = 0$. De az egyes rögzített változókban való linearitás és az alternáló tulajdonság miatt minden v_1, \dots, v_n vektorrendszerhez léteznek β_π ($\pi \in S_n$) együtthatók, melyekre $f(v_1, \dots, v_n) = \sum_{\pi \in S_n} \beta_\pi f(e_{\pi(1)}, e_{\pi(2)}, \dots, e_{\pi(n)}) = 0$. \square

Foglaljuk össze az előző két állítást:

16.18. állítás. Legyen f egy nem triviális mérték a V vektortér felett. A $\{v_1, \dots, v_n\}$ vektorrendszer lineárisan összefüggőségének szükséges és elegendő feltétele, hogy $f(v_1, \dots, v_n) = 0$.

16.19. állítás. Tetszőleges (legalább egydimenziós) vektortér feletti mértékek vektortere egydimenziós vektortér.

Bizonyítás: Rögzítsük a tér egy $\{e_1, \dots, e_n\}$ bázisát. Láttuk, hogy a

$$d(v_1, \dots, v_n) = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^n e_i^* v_{\pi(i)}$$

függvény nem triviális mértéket definiál. Elegendő tehát megmutatnunk, hogy amennyiben D egy másik mérték V felett, úgy található $\delta \in \mathbb{F}$, amelyre $D = \delta d$. Legyen ezért $\delta = D(e_1, \dots, e_n)$, így

$$(D - \delta d)(e_1, \dots, e_n) = D(e_1, \dots, e_n) - \delta d(e_1, \dots, e_n) = \delta - \delta \cdot 1 = 0.$$

Ez viszont azt jelenti, hogy a $D - \delta d$ mérték az $\{e_1, \dots, e_n\}$ bázison eltűnik, ami csak úgy lehetséges ha $D - \delta d = 0$, ergo $D = \delta d$. \square

A fejezet legfontosabb gondolatához érkeztünk.

16.20. definíció-állítás (Lineáris transzformáció determinánsa). Legyen $A \in L(V)$ a V véges dimenziós vektortér egy lineáris transzformációja. Ehhez létezik egyetlen olyan $\delta(A) \in \mathbb{F}$ szám, amelyre minden f nem triviális V feletti mérték esetén az

$$f(Av_1, \dots, Av_n) = \delta(A)f(v_1, \dots, v_n)$$

azonosság minden $v_1, \dots, v_n \in V$ vektor mellett fenáll.

Ezt a $\delta(A)$ számot nevezzük az A lineáris transzformáció determinánsának.

Bizonyítás: Vegyünk egy f nem triviális mértéket V felett. Láttuk, hogy ilyen valóban létezik, hiszen a mértékek egydimenziós teret alkotnak. Világos, hogy a $(v_1, \dots, v_n) \mapsto f(Av_1, \dots, Av_n)$ függvény is egy mérték. No de, a mértékek egy egydimenziós térféle az elemei, ezért van egyetlen olyan $\delta_f(A) \in \mathbb{F}$ szám, amelyre minden v_1, \dots, v_n mellett $f(Av_1, \dots, Av_n) = \delta_f(A)f(v_1, \dots, v_n)$. Legyen most g egy másik nem zérus mérték. Ekkor $g = \gamma f$ valamely $\gamma \neq 0$ számmal. Persze g -re is igaz, amit eddig igazoltunk, így

$$g(Av_1, \dots, Av_n) = \delta_g(A)g(v_1, \dots, v_n), \text{ valamint } f(Av_1, \dots, Av_n) = \delta_f(A)f(v_1, \dots, v_n).$$

Ha a baloldali kifejezésben a g mértéket a γf mértékre cseréljük, majd osztunk a γ számmal, akkor

$$f(Av_1, \dots, Av_n) = \delta_g(A)f(v_1, \dots, v_n)$$

következik. Így $\delta_f(A)$ egyértelműsége miatt $\delta_f(A) = \delta_g(A)$. A $\delta_f(A)$ értéke tehát minden f nem triviális mérték mellett azonos, tehát azt $\delta(A)$ -val jelölve, kapjuk az egyetlen olyan számot, amelyre a kívánt azonosság fennáll. \square

16.21. állítás (Szorzattétel). Legyenek az $A, B \in L(V)$ lineáris transzformációk a V véges dimenziós vektortéren. Ekkor

$$\delta(A \circ B) = \delta(A)\delta(B).$$

Bizonyítás: A $\delta(A \circ B)$ definíciója szerint ez az egyetlen olyan szám, amelyikre minden v_1, \dots, v_n mellett

$$f(ABv_1, \dots, ABv_n) = \delta(A \circ B)f(v_1, \dots, v_n).$$

Másrészt a $\delta(A)$ és $\delta(B)$ értelmezése szerint

$$f(ABv_1, \dots, ABv_n) = \delta(A)f(Bv_1, \dots, Bv_n) = \delta(A)\delta(B)f(v_1, \dots, v_n)$$

E két utolsó sort összevetve kapjuk, hogy $\delta(A \circ B) = \delta(A)\delta(B)$. \square

16.22. állítás. Legyen $A \in L(V)$ egy lineáris transzformáció. Ekkor az A bármely bázisban felírt mátrixának a determinánsa azonos A -nak mint lineáris transzformációinak a determinánsával. Formálisan tetszőleges $\{e_1, \dots, e_n\}$ bázis mellett:

$$\delta(A) = |[A]_{\{e_1, \dots, e_n\}}|.$$

Bizonyítás: Jelölje $d = d_{\{e_1, \dots, e_n\}}$ a bázishoz tartozó nem triviális mértéket. Ekkor

$$|[A]| = d(Ae_1, \dots, Ae_n) = \delta(A)d(e_1, \dots, e_n) = \delta(A).$$

\square

Meggondoltuk tehát, hogy egy transzformáció mátrixának determinánsa nem függ a koordinátázástól, tehát más-más bázisokban felírt mátrixának determinánsa azonos. A továbbiakban az A lineáris transzformáció determinánsára az egyszerűség kedvéért az $|A|$ jelölést (is) használjuk. Ezek szerint A -nak $|A|$ determinánsa, és A -nak valamely bázisban felírt mátrixának $|[A]|$ determinánsa között nincs különbség.

Mátrixok szorzattétele Ha $[A], [B]$ két négyzetes mátrix, akkor A -val és B -vel jelölve a mátrixok generálta $\mathbb{F}^n \rightarrow \mathbb{F}^n$ lineáris transzformációkat

$$|[A][B]| = |[A \circ B]| = \delta(A \circ B) = \delta(A) \delta(B) = |[A]| \cdot |[B]|.$$

Meggondoltuk tehát, hogy mátrixok szorzatának determinánsa a mátrixok determinánsának szorzata.

Az inverz determinánsa Legyen most $A \in L(V)$ egy reguláris transzformáció. Ekkor $AA^{-1} = I$, így véve a determinánsokat, $|A| \cdot |A^{-1}| = 1$, azaz $|A^{-1}| = \frac{1}{|A|}$. Összefoglalva, az inverz determinánsa a determináns reciproka.

Ugyan korábban már két indoklást láttunk a particionált mátrix determinánsának kiszámítására (16.2), de érdemes még egyszer átgondolnunk, hogyan következik ez a mértékek segítségével.

Particionált mátrix determinánsa. Tegyük fel, hogy $[A]$ egy $(n+m) \times (n+m)$ méretű mátrix, amely

$$[A] = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

módon van particionálva, ahol $A_{1,1}$ egy $n \times n$, $A_{2,2}$ egy $m \times m$ méretű négyzetes mátrix, továbbá $A_{2,1}$ egy $m \times n$ méretű és $A_{1,2}$ egy $n \times m$ méretű részmátrix. Tegyük fel most is, hogy a bal alsó $A_{2,1}$ partíció minden eleme zérus. Legyen ezért $A \in L(\mathbb{F}^{n+m})$ az a transzformáció, amelynek az $\{e_1, \dots, e_n, e_{n+1}, \dots, e_m\}$ bázisban felírt mátrixa $[A]$. Világos, hogy az $M = \text{lin}\{e_1, \dots, e_n\}$ jelöléssel az M altér egy A -ra nézve invariáns altér, hiszen az $[A_{2,1}]$ partíció zérus. Definíció szerint a szóban forgó bázisban az A transzformáció M altérre való megszorításának mátrixára $[A|_M] = [A_{1,1}]$. Jelölje d az \mathbb{F}^{n+m} fent rögzített bázisához tartozó szokásos mértéket, azaz

$$d(v_1, \dots, v_n, v_{n+1}, \dots, v_{n+m}) = \sum_{\pi \in S_{n+m}} \text{sgn } \pi \prod_{j=1}^{n+m} e_j^* v_{\pi(j)}.$$

Tudjuk, hogy $d(Ae_1, \dots, Ae_n, Ae_{n+1}, \dots, Ae_{n+m}) = |[A]|$. Definiáljuk most az $L(M)$ feletti mértéket a következőképpen:

$$f(v_1, \dots, v_n) = d(v_1, \dots, v_n, Ae_{n+1}, \dots, Ae_{n+m})$$

Világos, hogy f valóban egy mérték $L(M)$ felett, és az $A|_M$ megszorított transzformáció determinánsának definíciója szerint

$$\begin{aligned} |[A]| &= f(Ae_1, \dots, Ae_n) = \det(A|_M) f(e_1, \dots, e_n) \\ &= \det(A|_M) d(e_1, \dots, e_n, Ae_{n+1}, \dots, Ae_{n+m}) = |[A_{1,1}]| \cdot \left| \begin{pmatrix} I & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} \right|. \end{aligned}$$

Meggondoltuk tehát, hogy ha a bal alsó partíció zérus, akkor

$$\left| \begin{pmatrix} A_{1,1} & A_{2,1} \\ 0 & A_{2,2} \end{pmatrix} \right| = |A_{1,1}| \cdot \left| \begin{pmatrix} I & A_{2,1} \\ 0 & A_{2,2} \end{pmatrix} \right|. \quad (\dagger)$$

Az egész eddigi gondolatot analóg módon ismételhetjük arra az esetre, mikor a jobb felső partíció zérus (és persze ekkor a bal alsó $A_{2,1}$ partíció akármilyen lehet). Kapjuk tehát, hogy

$$\left| \begin{pmatrix} A_{1,1} & 0 \\ A_{2,1} & A_{2,2} \end{pmatrix} \right| = |A_{2,2}| \cdot \left| \begin{pmatrix} A_{1,1} & 0 \\ A_{2,1} & I \end{pmatrix} \right|. \quad (\ddagger)$$

A fent kiemelt (\dagger) és (\ddagger) alakokból a (16.2) állítás is könnyen adódik. Csak arra kell még emlékeznünk, hogy a mátrixnak és a transzponáltjának azonos a determinánsa. Ugyanis:

$$\begin{aligned} \left| \begin{pmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} \right| &= |A_{1,1}| \cdot \left| \begin{pmatrix} I & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} \right| = |A_{1,1}| \cdot \left| \begin{pmatrix} I & 0 \\ A_{1,2}^T & A_{2,2}^T \end{pmatrix} \right| \\ &= |A_{1,1}| \cdot |A_{2,2}^T| \cdot \left| \begin{pmatrix} I & 0 \\ A_{1,2}^T & I \end{pmatrix} \right| = |A_{1,1}| \cdot |A_{2,2}|. \end{aligned}$$

Az utolsó lépésben a particionált mátrix egy olyan alsóháromszög mátrix, amelynek a diagonálisa csupa 1 számból áll, ezért a determinánsa is 1.

16.4. A determináns kifejtése

A kifejtési tétel arra való, hogy a determináns kiszámítását viaszavezessük 1-el kisebb méretű mátrixok determinánsának kiszámolására. Ezt a gondolatot egymásután alkalmazva tetszőleges mátrix determinánsa kiszámítható.

Ebben a fejezetben vannak azok az explicit formulák, amelyeket a fejezet bevezetőjében említettem.

16.23. definíció (minor, kofaktor, kofaktormátrix). Legyen $A \in \mathbb{F}^{n \times n}$ egy mátrix. E mátrix

- (i, k) -csonkoltja a mátrix i -edik sorának és k -adik oszlopának törlése után megmaradt $n - 1 \times n - 1$ méretű mátrix. Jelölése: $A_{i,k}$.
- (i, k) -minorja az (i, k) -csonkolt $A_{i,k}$ mátrix determinánsa.
- (i, k) -kofaktora a $(-1)^{i+k} |A_{i,k}|$ szám.
- kofaktormátrixa az az $n \times n$ méretű mátrix, amelynél az i -edik sor k -adik eleme az (i, k) -kofaktor.

16.24. állítás. Legyen $A \in \mathbb{F}^{n \times n}$ egy mátrix. Írjuk felül az A mátrix k -adik oszlopának elemeit 0-val, majd a k -adik oszlop i -edik helyére írunk 1-et. E felülírt mátrix determinánsa az eredeti mátrix (i, k) -kofaktora.

Bizonyítás: Világos, hogy a szóban forgó mátrix $k - 1$ oszlopcsere, majd $i - 1$ sorcsere után olyan alakra hozható, ahol a balfelső elem 1, alatta az első oszlop minden eleme zérus, és a jobb alsó $n - 1 \times n - 1$ méretű részmátrix az eredeti mátrix (i, k) -csonkoltja:

$$\begin{pmatrix} 1 & \begin{bmatrix} a_{i,1} & \dots & a_{i,n} \\ a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix} \\ \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} & \end{pmatrix}$$

A fenti mátrix determinánsa a particionált mátrix determinánsára kapott formula, a 126. oldal, szerint $1 \cdot |A_{i,k}|$. Meggondoltuk tehát, hogy ha az állításban felülírt mátrix determinánsát δ -val jelöljük, akkor $\delta(-1)^{k-1+i-1} = 1 \cdot |A_{i,k}|$ -t kapjuk. Ebből persze $\delta = (-1)^{i+k} |A_{i,k}|$. \square

16.25. állítás (oszlopok szerinti kifejtés). Legyen $A \in \mathbb{F}^{n \times n}$ mátrix. Ekkor

1. minden $1 \leq k \leq n$ oszlopra $|A| = \sum_{i=1}^n a_{i,k} (-1)^{i+k} |A_{i,k}|$,
2. minden $1 \leq k, j \leq n$ oszlopra $j \neq k$ mellett $0 = \sum_{i=1}^n a_{i,j} (-1)^{i+k} |A_{i,k}|$.

A fenti két formula tömörebben: minden $1 \leq k, j \leq n$ mellett

$$\sum_{i=1}^n a_{i,j} (-1)^{i+k} |A_{i,k}| = \delta_{j,k} |A|.$$

Bizonyítás: Tekintsük az \mathbb{F}^n vektortér természetes bázisát: $\{e_1, \dots, e_n\}$. Láttuk, hogy a $d = d_{\{e_1, \dots, e_n\}}$ mérték mellett

$$|A| = d(a_1, \dots, a_n),$$

ahol a_1, \dots, a_n jelöli rendre az A mátrix oszlopait. A k -adik oszlopvektorra $a_k = \sum_{i=1}^n a_{i,k} e_i$. Mivel d az összes nem k -adik változója rögzítése mellett a k -adik változójában lineáris, ezért

$$|A| = d(a_1, \dots, a_k, \dots, a_n) = d\left(a_1, \dots, \sum_{i=1}^n a_{i,k} e_i, \dots, a_n\right) = \sum_{i=1}^n a_{i,k} d(a_1, \dots, e_i, \dots, a_n).$$

No de, itt $d(a_1, \dots, e_i, \dots, a_n)$ az előző állításban felülírt mátrix determinánsa, azaz $(-1)^{i+k} |A_{i,k}|$. Éppen ezt kellett belátni az 1. pont indoklásához.

Másoljuk a mátrix j -edik oszlopát a k -adik oszlopra. Ennek a mátrixnak két oszlopa azonossá vált, ezért determinánsa zérus. Felírva erre a mátrixra a k -adik oszlop szerinti kifejtést, éppen a 2. pont azonosságát kapjuk. \square

A mátrix transzponáltjának determinánsa azonos a mátrix determinánsával. Így ha a transzponált mátrixra alkalmazzuk az oszlopok szerinti kifejtést, akkor a sorok szerinti kifejtést kapjuk.

A determináns kifejtésével az mátrix inverzére kapunk explicit formulát:

16.26. állítás. Legyen $[A]$ egy reguláris mátrix. Ekkor az inverz mátrix a kofaktormátrix transzponáltjának a determináns reciprokával vett szorzata.

Magyarra áttérve:

$$[A^{-1}]_{k,i} = \frac{1}{|A|} (-1)^{i+k} |A_{i,k}|.$$

Bizonyítás: Jelölje a M a kofaktormátrix transzponáltját, azaz $m_{k,i} = (-1)^{i+k} |A_{i,k}|$. Így az $M \cdot A$ szorzat mátrix k, j pozíciója

$$\sum_{i=1}^n m_{k,i} a_{i,j} = \sum_{i=1}^n (-1)^{i+k} |A_{i,k}| a_{i,j} = \delta_{j,k} |A|$$

a kifejtési tétel szerint. Ez azt jelenti, hogy $M \cdot A = |A|I$. Mivel A reguláris ezért a nem zérus determinánsával osztva kapjuk, hogy $(\frac{1}{|A|} M) A = I$, ami azt jelenti, hogy $A^{-1} = \frac{1}{|A|} M$. \square

16.27. állítás (Cramer-szabály). Tekintsünk egy $n \times n$ méretű inhomogén lineáris egyenletrendszeret. Tudjuk, hogy pontosan akkor van egyetlen megoldása minden jobboldal mellett, ha az együttható mátrix reguláris. Jelölje ezért A a reguláris együttható mátrixot, és B_k azt a mátrixot, amelyet úgy kapunk, hogy az egyenlet jobboldalát reprezentáló oszlopvektort az A együttható mátrix k -adik oszlopára írjuk. Ekkor az x megoldásvektor k -adik koordinátájára

$$x_k = \frac{|B_k|}{|A|}.$$

Bizonyítás: Kezdjük a B_k mátrixnak a k -adik oszlopá szerinti kifejtésével. $|B_k| = \sum_{i=1}^n b_i (-1)^{i+k} A_{i,k}$. Na most az $Ax = b$ ekvivalens az $x = A^{-1}b$ -vel, így a Cramer-szabály szerint

$$x_k = [A^{-1}]_k \cdot b = \sum_{i=1}^n [A^{-1}]_{k,i} \cdot b_i = \sum_{i=1}^n \left(\frac{1}{|A|} (-1)^{i+k} |A_{i,k}| \right) b_i = \frac{1}{|A|} \cdot |B_k|. \quad \square$$

16.5. A karakterisztikus polinom

16.28. definíció. Legyen $A \in L(V)$ egy lineáris transzformációja az \mathbb{F} test feletti V, n -dimenziós vektortérnek. Definiáljuk a $k : \mathbb{F} \rightarrow \mathbb{F}$ függvényt, mint $k(t) = |tI - A|$ minden $t \in \mathbb{F}$ mellett. Ekkor k egy pontosan n -ed fokú, normált polinom, amelyre

$$k(t) = t^n - \text{tr}(A)t^{n-1} + \cdots + (-1)^n |A|.$$

Ezt a k függvényt nevezzük az A transzformáció karakterisztikus polinomjának.

Bizonyítás: Rögzítsük a tér egy $\{e_1, \dots, e_n\}$ bázisát, majd írjuk fel az $tI - A$ transzformáció mátrixát ebben a bázisban. E mátrix determinánsa $k(t)$, ami minden konkrét $t \in \mathbb{F}$ mellett az \mathbb{F} testbeli szám. Mivel a transzformáció mátrixának determinánsa minden bázis mellett ugyanaz a szám, ezért $k(t)$ konkrét értéke a választott bázistól független.

A diagonálisban szereplő elemek $(t - a_{i,i})$ alakúak, és a t változó máshol nem szerepel a mátrixban. Ha a determináns definíciójára gondolunk, akkor nyilvánvaló, hogy $k(t)$ a t változó legfeljebb n -ed fokú polinomja.

E polinom konstans tagja $k(0) = | - A | = (-1)^n |A|$.

Ha $\pi \neq \text{id}$ egy permutáció, akkor π legalább két helyen különbözik az identikus permutációtól, ergo a π -hez tartozó bástyafelrakás legfeljebb $n - 2$ diagonális elemet taposhat. Látjuk tehát, hogy identitástól különböző permutáció mellett nincs t -nek $n - 2$ -nél magasabb fokú hatványa.

A diagonális bástyafelrakáshoz a

$$(t - a_{1,1})(t - a_{2,2}) \cdots (t - a_{n,n})$$

szorzat tartozik, amelynek első két legmagasabb hatványú tagja

$$t^n + t^{n-1} (-a_{1,1} \cdots -a_{n,n}) = t^n - \text{tr}(A) t^{n-1}. \quad \square$$

A spektrum pontjai nem csak a minimálpolinom zérus helyei, hanem a karakterisztikus polinom zérus helyei is.

16.29. állítás (karakterisztikus polinom gyökei). *Legyen $k(t)$ az $A \in L(V)$ lineáris transzformáció karakterisztikus polinomja. A $\lambda \in \mathbb{F}$ szám pontosan akkor sajátértéke A -nak, ha $k(\lambda) = 0$.*

Bizonyítás: A λ pontosan akkor sajátértéke A -nak, ha $\lambda I - A$ szinguláris, ami avval ekvivalens, hogy $0 = |\lambda I - A| = k(\lambda)$. \square

Most tegyük fel egy pillanatra, hogy \mathbb{C} feletti vektortér $A \in L(V)$ lineáris transzformációját vizsgáljuk. Minimálpolinomja az algebrai alaptétele szerint felbomlik

$$m(t) = (t - \lambda_1)^{m_1} \dots (t - \lambda_s)^{m_s}$$

gyöktényezői szorzatára, ahol $\sigma(A) = \{\lambda_1, \dots, \lambda_s\}$. Jelölje k_i a λ_i sajátérték algebrai multiplicitását, azaz $k_i = \nu((A - \lambda_i I)^{m_i})$. Tudjuk, hogy $tI - A$ Jordan-normálalakú mátrixa egy olyan felsőháromszög alakú mátrix, amelynek diagonálisában $t - \lambda_i$ kifejezés éppen k_i -szer szerepel. Mivel egy felsőháromszög alakú mátrix determinánsa a diagonális elemek szorzata, ezért a következő tételt gondoltuk meg.

16.30. állítás. *Tegyük fel, hogy V egy a \mathbb{C} test feletti vektortér és $A \in L(V)$ egy lineáris transzformáció. Jelölje $\sigma(A) = \{\lambda_1, \dots, \lambda_s\}$ a különböző sajátértékeket, és jelölje k_1, \dots, k_s rendre sajátértékek algebrai multiplicitásait. Ekkor az A transzformáció karakterisztikus polinomja*

$$k(t) = (t - \lambda_1)^{k_1} \dots (t - \lambda_s)^{k_s}.$$

Az $A - \lambda_i I$ a $\ker(A - \lambda_i I)^{m_i}$ vektortér felett egy m_i -ed rendben nilpotens transzformáció, de tudjuk hogy a nilpotencia rendje legfeljebb a tér dimenziója, ergo $m_i \leq \dim(\ker(A - \lambda_i I)^{m_i}) = \nu((A - \lambda_i I)^{m_i}) = k_i$. Ebből azonnal látjuk, hogy a minimálpolinom osztója a karakterisztikus polinomnak. Meggondoltuk tehát a Cayley – Hamilton-tételt abban a speciális esetben, mikor a test a komplex számtest.

A Cayley – Hamilton-tétel nem függ az algebrai alaptételektől, amint azt rögtön megmutatjuk.

16.31. lemma. *Legyen \mathbb{F} egy test, amelyben $1 + 1 \neq 0$. Tetszőlegesen választott $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{F}$ számok mellett*

jelölje $A = \begin{pmatrix} -\alpha_{n-1} & 1 & 0 & \dots & 0 \\ -\alpha_{n-2} & 0 & 1 & \dots & 0 \\ \vdots & \dots & \dots & \ddots & 0 \\ -\alpha_1 & 0 & 0 & \dots & 1 \\ -\alpha_0 & 0 & 0 & \dots & 0 \end{pmatrix}$ az $n \times n$ méretű mátrixot. Ekkor ennek karakterisztikus polinomja $k(t) = t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_1t + \alpha_0$.

Bizonyítás: Először is tekintsük a $tI - A = \begin{pmatrix} t + \alpha_{n-1} & -1 & 0 & \dots & 0 \\ \alpha_{n-2} & t & -1 & \dots & 0 \\ \vdots & \dots & \dots & \ddots & 0 \\ \alpha_1 & 0 & 0 & \dots & -1 \\ \alpha_0 & 0 & 0 & \dots & t \end{pmatrix}$ mátrixot. E mátrix mérete szerinti indukcióval bizonyítunk.

Ha $n = 1$, akkor a polinom $t + \alpha_0$ és a mátrixnak is csak egyetlen $t + \alpha_0$ eleme van.

Most tegyük fel, hogy igaz az állítás tetszőlegesen választott $n - 1$ szám mellett, és lássuk be n -re. A $tI - A$ mátrix determinánsát megkapjuk az utolsó sor szerint kifejtéssel. Tehát

$$k(t) = |tI - A| = \alpha_0 (-1)^{1+n} \left| \begin{pmatrix} -1 & 0 & \dots & 0 \\ t & -1 & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{pmatrix} \right| + t (-1)^{n+n} \left| \begin{pmatrix} t + \alpha_{n-1} & -1 & 0 & \dots & 0 \\ \alpha_{n-2} & t & -1 & \dots & 0 \\ \vdots & \dots & \dots & \ddots & -1 \\ \alpha_1 & 0 & 0 & \dots & t \\ \alpha_0 & 0 & 0 & \dots & t \end{pmatrix} \right|.$$

Itt az első $(n - 1) \times (n - 1)$ méretű determináns $(-1)^{n-1}$, mert ez egy alsóháromszög alakú mátrix a diagonálisában csupa -1 számmal. A második determináns értéke az indukciós feltevés az $n - 1$ darab $\alpha_1, \dots, \alpha_{n-1}$ számra. A fent kezdtet kiemelést folytatva tehát,

$$k(t) = \alpha_0 (-1)^{1+n} (-1)^{n-1} + t (t^{n-1} + \alpha_{n-1}t^{n-2} + \dots + \alpha_2t + \alpha_1) = t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_1t + \alpha_0. \quad \square$$

16.32. állítás (Cayley–Hamilton). Legyen V egy tetszőleges test feletti vektortér, és $A \in L(V)$ egy lineáris transzformáció. Ekkor A gyöke a karakterisztikus polinomjának, azaz $k(A) = 0$ ($\in L(V)$).

Bizonyítás: Legyen $m = \dim(V)$, és $v \in V$ egy nem zérus elem. Tekintsük a v által generált legszűkebb A -invariáns alteret, azaz $\text{lin}(v; A)$ -t. Ha ez n dimenziós ($1 \leq n \leq m$), és a v -hez tartozó kis minimálpolinom

$$p_v(t) = t^n + \alpha_{n-1}t^{n-1} + \cdots + \alpha_1t + \alpha_0,$$

akkor a $\text{lin}(v; A)$ invariáns altérben $\{A^{n-1}v, \dots, Av, v\}$ egy bázis. Ebben a bázisban $A|_{\text{lin}(v; A)}$ transzformáció mátrixa

$$A_{1,1} = \begin{pmatrix} -\alpha_{n-1} & 1 & 0 & \cdots & 0 \\ -\alpha_{n-2} & 0 & 1 & \cdots & 0 \\ \vdots & \cdots & \cdots & \ddots & 0 \\ -\alpha_1 & 0 & 0 & \cdots & 1 \\ -\alpha_0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

alakú. Az előző gondolat szerint $A_{1,1}$ karakterisztikus polinomja éppen $p_v(t)$.

Egészítük ki $\text{lin}(v; A)$ fenti bázisát az egész V vektortér bázisává. Ebben a bázisban felírva A mátrixa $[A] = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$ alakú, ahol $A_{1,1}$ azonos $A|_{\text{lin}(v; A)}$ fenti mátrixával, és mivel a $\text{lin}(v; A)$ egy invariáns altér, ezért az $A_{1,1}$ alatt elhelyezkedő $(m-n) \times n$ méretű $A_{2,1}$ mátrix minden eleme zérus. Egy ilyen módon partícionált mátrix determinánsa már könnyen számolható:

$$k(t) = |[tI - A]| = \left| \begin{pmatrix} tI_{1,1} - A_{1,1} & -A_{1,2} \\ -A_{2,1} & tI_{2,2} - A_{2,2} \end{pmatrix} \right| = |tI_{1,1} - A_{1,1}| \cdot |tI_{2,2} - A_{2,2}| = p_v(t) k_2(t),$$

ahol k_2 az $A_{2,2}$ karakterisztikus polinomja. Azt kaptuk, hogy p_v -nek többszöröse k , ezért $k(A)v = 0$. \square

Mivel egy lineáris transzformáció gyöke a saját karakterisztikus polinomjának, ezért a karakterisztikus polinom egy pontosan n -ed fokú polinom, amely a minimálpolinom többszöröse, azaz a minimálpolinom osztója a karakterisztikus polinomnak. Persze ebből is következik, hogy a minimálpolinom egy legfeljebb n -ed fokú polinom.

Skalárisszorzatos terek geometriája

A továbbiakban \mathbb{K} feletti vektorterekről van szó, ahol $\mathbb{K} = \mathbb{R}$ vagy $\mathbb{K} = \mathbb{C}$.

17.1. Definíciók

17.1. definíció (skalárisszorzatos-tér). Egy $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$ függvényt *skaláris szorzatnak* mondunk, ha teljesülnek az alábbi axiómák:

1. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ minden $x, y \in V$ mellett.
2. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ minden $x, y \in V$ és minden $\alpha \in V$ mellett.
3. $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$ minden $x_1, x_2, y \in V$ mellett.
4. $\langle x, x \rangle \geq 0$ és $\langle x, x \rangle = 0$ akkor és csak akkor, ha $x = 0$.

Ha a V vektortéren adott a fenti skaláris szorzat, akkor a $(V, \langle \cdot, \cdot \rangle)$ párt *skalárisszorzatos-térnek* mondjuk.

Az első három pont szerint a skaláris szorzat rögzített második változó melett az elsőben lineáris, azaz $\langle \sum_{j=1}^s \alpha_j x_j, y \rangle = \sum_{j=1}^s \alpha_j \langle x_j, y \rangle$. Hasonlóan, a skaláris szorzás operáció rögzített első változó mellett a második változó *konjugáltan lineáris* függvénye, azaz $\langle x, \sum_{k=1}^r \beta_k y_k \rangle = \sum_{k=1}^r \overline{\beta_k} \langle x, y_k \rangle$. Persze ezeket nagyon sokszor vegyítjük is, így a

$$\langle \sum_{j=1}^s \alpha_j x_j, \sum_{k=1}^r \beta_k y_k \rangle = \sum_{j=1}^s \alpha_j \langle x_j, \sum_{k=1}^r \beta_k y_k \rangle = \sum_{j=1}^s \alpha_j \sum_{k=1}^r \overline{\beta_k} \langle x_j, y_k \rangle = \sum_{j=1}^s \sum_{k=1}^r \alpha_j \overline{\beta_k} \langle x_j, y_k \rangle$$

formula adódik, amit úgy jegyezhetünk meg, hogy minden tagot minden taggal kell szorozni, de a másik változóból konjugáltan emeljük ki a test elemeit.

Belső szorzat. A legfontosabb példa skaláris szorzatra a *belső szorzat*. Ha adott a vektortér egy $\{v_1, \dots, v_n\}$ bázisa, akkor minden vektort egyértelműen meghatároznak a koordinátái, azaz $x = \sum_{j=1}^n \xi_j v_j$ és $y = \sum_{j=1}^n \eta_j v_j$. Ezek belső szorzatát

$$[x, y] = \sum_{j=1}^n \xi_j \eta_j$$

definiálja. Nagyon fontos észrevenni, hogy a belső szorzat a definíciója szerint függ a vektorok koordinátáitól, tehát függ a bázis megválasztásától.¹

A másik fontos példa \mathbb{R} feletti vektortérnél értelmezett. Jelölje $C[a, b]$ valamely $[a, b]$ korlátos és zárt intervallum feletti összes folytonos függvények vektorterét. Definiálja $f, g \in C[a, b]$ két folytonos függvény skaláris szorzatát $\int_a^b f(x) g(x) dx$. Könnyen látható, hogy a fenti függvény kielégíti a skaláris szorzat definíciójában követelteket. Figyeljünk a folytonosság szerepére!

¹Helyesebb lenne a szörnyen nehézkes $[\cdot, \cdot]_{\{v_1, \dots, v_n\}}$ jelölés.

17.2. definíció (merőlegesség, halmaz ortokomplementere). Legyen $(V, \langle \cdot, \cdot \rangle)$ egy skaláriszorzatos-tér.

1. Ha az $x, y \in V$ két pontjára $\langle x, y \rangle = 0$, akkor azt mondjuk hogy a két vektor merőleges egymásra. Jelölés: $x \perp y$.
2. Legyen most, $H \subseteq V$ egy részhalmaz. Definiálja $H^\perp = \{x \in V : x \perp a \text{ minden } a \in H\}$ a H halmaz merőlegesét, vagy ortokomplementérét.
3. Tekintsük a $H_1, H_2 \subseteq V$ halmazokat. Azt mondjuk, hogy ezek merőlegesek egymásra ha minden $u \in H_1$ és minden $v \in H_2$ esetén $u \perp v$. Jelölés: $H_1 \perp H_2$.

A zérus vektor minden vektorra merőleges, sőt a zérus vektor az egyedüli ilyenvektor, hiszen ekkor sajátmagára is merőleges, de saját magára merőleges vektor csak a zérus vektor van. Pont ezt jelenti az axiómák között az utolsó. A H^\perp az a részhalmaza V -nek, amely az összes H -ra merőleges vektorokból áll. Tehát $\{0\}^\perp = V$ és $V^\perp = \{0\}$. Persze $H \perp H^\perp$ csak tautológia.

17.3. definíció. Legyen V egy skaláriszorzatos-tér, $H \subseteq V$. Az $\{x_1, \dots, x_k\}$ vektorrendszer

1. *ortogonálisnak* nevezünk, ha minden $i \neq j$ esetén $x_i \perp x_j$.
2. *ortonormáltnak* nevezünk, ha minden i, j mellett $\langle x_i, x_j \rangle = \delta_{i,j}$.

A felépítés jelen szintjén nem olyan egyszerű egy tetszőleges skaláris szorzatos térben ortonormált rendszert adni. Persze az üres vektorrendszer ortogonális, de ortonormált is. Az látható, hogy egy ortogonális rendszert tetszőlegesen sok zérus vektorral kiegészítve ortogonális rendszert kapunk. Persze a zérus vektor egy ortonormált rendszerhez nem tartozhat elemként, de ha van a vektortérben egy $v \neq 0$ elem, akkor az egyelemű $\left\{ \frac{1}{\sqrt{\langle v, v \rangle}} v \right\}$ vektorrendszer egy ortonormált rendszert alkot. (Mivel csak egyetlen elem van, ezért bármely két különböző elem merőleges egymásra; valamint $\langle \frac{1}{\sqrt{\langle v, v \rangle}} v, \frac{1}{\sqrt{\langle v, v \rangle}} v \rangle = \frac{1}{\sqrt{\langle v, v \rangle}} \frac{1}{\sqrt{\langle v, v \rangle}} \langle v, v \rangle = 1$.)

Egynél több elemből álló ortonormált rendszert most nem tudunk mutatni, de később ennek oka egészen világos lesz. Látjuk majd, hogy egy n -dimenziós skaláriszorzatos térben mindenkorán n elemű ortonormált rendszert, de annál több elemű soha nincs. Ebből annyi, hogy n -nél több nemzérus vektor nem alkothat ortogonális rendszert, az már látható is:

17.4. állítás. *Egy zérus elemeket nem tartalmazó ortogonális rendszer lineárisan független rendszer is.*

Bizonyítás: Ha a $\{v_1, \dots, v_n\}$ rendszer lineárisan összefüggő, és $v_k = \sum_{j=1}^{k-1} \alpha_j v_j$, akkor

$$\langle v_k, v_k \rangle = \left\langle \sum_{j=1}^{k-1} \alpha_j v_j, v_k \right\rangle = \sum_{j=1}^{k-1} \alpha_j \langle v_j, v_k \rangle = 0.$$

Ez a skaláris szorzás definíciója szerint csak abban az esetben lehetséges, ha $v_k = 0$, ami ellentmond annak a feltételnek, hogy a rendszer a null vektort nem tartalmazza. \square

Speciálisan, egy ortonormált rendszer a zérus vektort nem tartalmazza, emiatt lineárisan független rendszer. Így tehát a Steinitz-lemma szerint, egy n -dimenziós vektortérben nem lehet több mint n elemű ortonormált rendszert mutatni. Később látni fogjuk, hogy n elemű ortonormált rendszer viszont mindenkorán konstruálható egy n -dimenziós vektortérben, de erre itt még várnunk kell.

Az alábbiakban a merőlegesség legnyilvánvalóbb tulajdonságait foglaljuk össze. Ezeket a továbbiakban hivatkozás nélkül használjuk.

17.5. állítás (Merőlegesség tulajdonságai). *Legyen most is V egy skaláriszorzatos-tér, $H, H_1 \subseteq V$. Ekkor*

1. $H \cap H^\perp \subseteq \{0\}$,
2. $H \subseteq H_1$ esetén $H_1^\perp \subseteq H^\perp$,
3. $H_1 \perp H$ akkor és csak akkor, ha $H_1 \subseteq H^\perp$,
4. $H \subseteq (H^\perp)^\perp$,

$$5. H^\perp = (\text{lin } H)^\perp,$$

6. H^\perp altér.

Bizonyítás: Ha $v \in H \cap H^\perp$, akkor $v \perp v$, ergo $v = 0$, ami az 1. pontot igazolja.

Ha $H \subseteq H_1$ és egy v vektor a H_1 minden pontjára merőleges, akkor persze ez H minden pontjára is merőleges, azaz $H_1^\perp \subseteq H$.

A 3. pont nyilvánvaló.

A meggondolt 3. pont szerint a minden fennálló $H \perp H^\perp$ éppen $H \subseteq H^{\perp\perp}$ tartalmazást jelenti.

Mivel $H \subseteq \text{lin } H$, ezért a 2. pont szerint $(\text{lin } H)^\perp \subseteq H^\perp$. Ha viszont $v \in H^\perp$ és $u_1, \dots, u_n \in H$, akkor ezek tetszőleges lineáris kombinációjára $\langle v, \sum_{j=1}^n \alpha_j u_j \rangle = \sum_{j=1}^n \alpha_j \langle v, u_j \rangle = 0$, hiszen $v \perp u_j$. Meggondoltuk tehát,

hogy v merőleges a H -beli elemekből képzett tetszőleges lineáris kombinációra, ergo $H^\perp \subseteq (\text{lin } H)^\perp$.

Világos, hogy $0 \in H^\perp$, ha $v_1, v_2 \in H^\perp$, akkor $\alpha_1 v_1 + \alpha_2 v_2 \in H^\perp$, azaz H^\perp minden H részhalmaz mellett egy altér. \square

17.2. Egyenlőtlenségek

17.6. állítás (Bessel). *Legyen $(V, \langle \cdot, \cdot \rangle)$ skaláriszorzatos-tér. Tegyük fel, hogy $\{x_1, \dots, x_k\}$ egy ortonormált rendszer. Ekkor minden $u \in V$ vektorra:*

$$1. \text{ fennáll a Bessel-egyenlőtlenség: } \sum_{i=1}^k |\langle u, x_i \rangle|^2 \leq \langle u, u \rangle;$$

$$2. \text{ Az } \hat{u} = u - \sum_{i=1}^k \langle u, x_i \rangle x_i \text{ vektorra } \hat{u} \perp x_j \text{ (} j = 1, \dots, k \text{)}, \text{ azaz } \hat{u} \in \{x_1, \dots, x_k\}^\perp.$$

Bizonyítás: Jelölje a bizonyítás erejéig $\alpha_j = \langle u, x_j \rangle$. Ekkor az $\hat{u} = u - \sum_{j=1}^k \alpha_j x_j$ vektorra

$$\begin{aligned} 0 \leq \langle \hat{u}, \hat{u} \rangle &= \langle u, u \rangle + \langle u, -\sum_{j=1}^k \alpha_j x_j \rangle + \langle -\sum_{j=1}^k \alpha_j x_j, u \rangle + \langle -\sum_{j=1}^k \alpha_j x_j, -\sum_{i=1}^k \alpha_i x_i \rangle \\ &= \langle u, u \rangle - \sum_{j=1}^k \bar{\alpha}_j \langle u, x_j \rangle - \sum_{j=1}^k \alpha_j \langle x_j, u \rangle + \sum_{j=1}^k \sum_{i=1}^k \alpha_j \bar{\alpha}_i \langle x_j, x_i \rangle \\ &= \langle u, u \rangle - \sum_{j=1}^k \bar{\alpha}_j \alpha_j - \sum_{j=1}^k \alpha_j \bar{\alpha}_j + \sum_{j=1}^k \alpha_j \bar{\alpha}_j = \langle u, u \rangle - \sum_{j=1}^k |\alpha_j|^2. \end{aligned}$$

Ez éppen a Bessel-egyenlőtlenség. A merőlegességre vonatkozó második állítás már sokkal egyszerűbb:

$$\langle \hat{u}, x_j \rangle = \langle u - \sum_{i=1}^k \alpha_i x_i, x_j \rangle = \langle u, x_j \rangle - \sum_{i=1}^k \alpha_i \langle x_i, x_j \rangle = \alpha_j - \bar{\alpha}_j = 0. \quad \square$$

17.7. állítás (Schwartz-egyenlőtlenség). *Legyen $(V, \langle \cdot, \cdot \rangle)$ skaláriszorzatos-tér, $u, v \in V$. Ekkor*

$$|\langle u, v \rangle|^2 \leq \langle u, u \rangle \langle v, v \rangle.$$

Bizonyítás: Ha $v = 0$, akkor minden oldal zérus. Ha $v \neq 0$, akkor a $\left\{ \frac{1}{\sqrt{\langle v, v \rangle}} v \right\}$ rendszer egyetlen elemű ortonormált rendszer. Alkalmazva a Bessel-egyenlőtlenséget $k = 1$ mellett

$$\frac{1}{\langle v, v \rangle} |\langle u, v \rangle|^2 = \left| \frac{1}{\sqrt{\langle v, v \rangle}} \langle u, v \rangle \right|^2 = \left| \langle u, \frac{1}{\sqrt{\langle v, v \rangle}} v \rangle \right|^2 \leq \langle u, u \rangle,$$

ami a bizonyítandó állítás. \square

17.8. definíció (skaláriszorzat indukálta norma). *Legyen $(V, \langle \cdot, \cdot \rangle)$ skaláriszorzatos-tér, jelölje minden $x \in V$ mellett $\|x\| = \sqrt{\langle x, x \rangle}$.*

17.9. definíció (norma indukálta metrika). Legyen $(V, \langle \cdot, \cdot \rangle)$ skaláriszorzatos-tér, tetszőleges $x, y \in V$ mellett jelölje $d(x, y) = \|x - y\|$.

17.10. állítás. Egy V skaláriszorzatos-térben teljesül a Schwartz-egyenlőtlenség, azaz $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$, valamint a skaláriszorzatos-tér kielégíti a normált tér axiómákat, azaz

1. $\|x\| = 0$ akkor és csak akkor, ha $x = 0$,
2. $\|\alpha x\| = |\alpha| \|x\|$,
3. $\|x + y\| \leq \|x\| + \|y\|$,

továbbá egy normált tér kielégíti a metrikus tér axiómákat, azaz

1. $d(x, y) = 0 \iff x = y$,
2. $d(x, y) = d(y, x)$,
3. $d(x, z) \leq d(x, y) + d(y, z)$.

Bizonyítás: Egyedül a 3. normált tér axióma nem teljesen evidens. Emlékezzünk arra, hogy egy $\Re z \leq |z|$ minden $z \in \mathbb{K}$ szám mellett. Így

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

Mindkét oldal gyököt véve, kapjuk a háromszög-egyenlőtlenséget. □

17.3. Pont és altér távolsága

Ebben a szakaszban azt mutatjuk meg, hogy ha a V skaláriszorzatos térnek M egy olyan altere, amelyben van véges ortonormált bázis, akkor a tér tetszőleges V pontjának, van M -hez legközelebbi pontja. Mi több, ilyen legközelebbi pontból csak egyetlen egy van.

Ehhez először a jól ismert Pythagoras-tételt kell megfogalmaznunk.

17.11. állítás (Pythagoras). Ha $u \perp v$, akkor $\|u + v\|^2 = \|u\|^2 + \|v\|^2$. Hasonlóan, ha $\{u_1, \dots, u_k\}$ ortogonális rendszer, akkor $\|\sum_{i=1}^k u_i\|^2 = \sum_{i=1}^k \|u_i\|^2$.

Bizonyítás: Világos, hogy ha $\langle u, v \rangle = 0$, akkor $\|u + v\|^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \|u\|^2 + \|v\|^2$. Most tegyük fel, hogy k -nál kisebb elemszámú ortogonális rendszerre igaz az állítás és lássuk be k -ra. Világos, hogy $\langle \sum_{j=1}^{k-1} u_j, u_k \rangle = 0$. Így a már bizonyított állítás és az indukciós feltevés szerint

$$\left\| \sum_{j=1}^k u_j \right\|^2 = \left\| \left(\sum_{j=1}^{k-1} u_j \right) + u_k \right\|^2 = \sum_{j=1}^{k-1} \|u_j\|^2 + \|u_k\|^2 = \sum_{j=1}^{k-1} \|u_j\|^2 + \left\| \sum_{j=1}^{k-1} u_j \right\|^2 = \sum_{j=1}^k \|u_j\|^2. \quad \square$$

17.12. állítás (legközelebbi pont). Legyen $\{x_1, \dots, x_k\}$ ortonormált rendszer és jelölje $M = \text{lin} \{x_1, \dots, x_k\}$. Ekkor minden $u \in V$ vektorhoz létezik egyetlen $u^* \in M$ vektor, amelyre

$$d(u, u^*) = \inf \{d(u, w) : w \in M\}.$$

Erre az M altérbeli u^* vektorra teljesül az $u^* = \sum_{i=1}^k \langle u, x_i \rangle x_i$ egyenlőtlenség.

Ez azt jelenti, hogy minden $u \in V$ mellett a $\sum_{i=1}^k \langle u, x_i \rangle x_i$ vektor ez egyetlen olyan M -beli vektor, amelyre

$$d\left(u, \sum_{i=1}^k \langle u, x_i \rangle x_i\right) = \min \{d(u, w) : w \in M\}.$$

Ezt az $\|u - \sum_{j=1}^k \langle u, x_j \rangle x_j\|$ számot nevezzük, az u vektor és az M altér távolságának.

Bizonyítás: minden $w \in M$ mellett az $\hat{u} = u - \sum_{j=1}^k \langle u, x_j \rangle x_j$ vektorra $\hat{u} \perp (\sum_{j=1}^k \langle u, x_j \rangle x_j - w)$, hiszen $\sum_{j=1}^k \langle u, x_j \rangle x_j - w \in M$. Így a Pythagoras-tétel szerint

$$\begin{aligned} d(u, w)^2 &= \|u - w\|^2 \\ &= \|u - \sum_{j=1}^k \langle u, x_j \rangle x_j + \sum_{j=1}^k \langle u, x_j \rangle x_j - w\|^2 = \|\hat{u} + \left(\sum_{j=1}^k \langle u, x_j \rangle x_j - w \right)\|^2 \\ &= \|\hat{u}\|^2 + \left\| \sum_{j=1}^k \langle u, x_j \rangle x_j - w \right\|^2. \end{aligned}$$

A jobboldali első kifejezés $\|\hat{u}\|^2$ független $w \in M$ -től, ezért $d(u, w)^2 \geq \|\hat{u}\|^2$. Itt szigorú reláció áll fenn, minden $w \in M$ -re, amelyre $w \neq \sum_{j=1}^k \langle u, x_j \rangle x_j$, és persze egyenlőség áll fenn, ha $w = \sum_{j=1}^k \langle u, x_j \rangle x_j$. Ez azt jelenti, hogy

$$\min \{d(u, w) : w \in M\} = \|\hat{u}\| = \|u - \sum_{j=1}^k \langle u, x_j \rangle x_j\| = d\left(u, \sum_{j=1}^k \langle u, x_j \rangle x_j\right),$$

és az $u^* = \sum_{j=1}^k \langle u, x_j \rangle x_j$ az egyetlen olyan M -beli pont, amelyre az $\|u - u^*\| = \min \{\|u - w\| : w \in M\}$ egyenlőség fennáll. Ezt kellett belátni. \square

Láttuk tehát, hogy ha az M altérben található véges ortonormált bázis, akkor tetszőleges ponthoz van egy és csak egy legközelebbi pontja az M altérnek. Ez a korábbi tapasztalataink szerint egy kicsit sem meglepő. Az már sokkal érdekesebb, hogy mivel csak egy legközelebbi pont van, ezért ha az M altérben egy másik ortonormált bázist veszünk kiindulásul, akkor az azzal képzett $\sum_{j=1}^k \langle u, x_j \rangle x_j$ vektor is a ugyanazt a vektort, az u -hoz legközelebbi M -beli vektort adja. Ez formálisan nézve már meglepő, hiszen az

$$\sum_{j=1}^k \langle u, x_j \rangle x_j$$

vektor felírása szerint függ az $\{x_1, \dots, x_k\} \subseteq M$ ortonormált bázis megválasztásától, ugyanis a fent kiemelt képletben szerepelnek a bázis elemei. Viszont tudjuk, hogy a fenti vektor az u -hoz legközelebb eső vektor a M -nek, amiből persze csak egy van, emiatt a fenti kiemelt M -beli vektor nem függ az M altér ortonormált bázisának megválasztásától.

17.13. definíció (merőleges vetület, merőleges vetítés). Legyen $M = \text{lin} \{x_1, \dots, x_k\}$, ahol $\{x_1, \dots, x_k\}$ ortonormált rendszer.

Láttuk, hogy a $\sum_{i=1}^k \langle u, x_i \rangle x_i \in M$ vektor független az $\{x_1, \dots, x_k\}$ ortonormált rendszer megválasztásától, hiszen ez vektor az u vektorhoz legközelebb eső vektor a M altérnek. Definiálja $p_M : V \rightarrow V$ az alábbi függvényt

$$p_M(u) = \sum_{i=1}^k \langle u, x_i \rangle x_i.$$

A $p_M(u) \in M$ vektort nevezzük az u -nak M altérre eső merőleges vetületének. A p_M függvény neve: *merőleges vetítés az M altérre*.

Mivel a skaláriszorzás rögzített második változó mellett, az első változó lineáris függvénye, ezért egy rögzített $x \in V$ vektor mellett az $u \mapsto \langle u, x \rangle x$ függvény egy $V \rightarrow V$ lineáris operáció. Mivel p_M ilyenek véges összege, azért a fent definiált merőleges vetítés a V vektortér egy lineáris transzformációja.

17.4. Ortogonalizáció

A merőleges vetítés operáció bevezetésének feltétele volt, hogy ha az M altérre vonatkozó merőleges vetítést akarjuk használni, akkor léteznie kell egy ortonormált bázisnak az M altérben. Az eddigi tételeinkben ez a téTEL feltevése volt. Most megmutatjuk, hogy egy skaláriszorzatos térnek minden véges dimenziós altére rendelkezik ezzel a tulajdonsággal, így egy véges dimenziós altérre eső merőleges vetítés minden definiált.

Először is érdemes újra fogalmazni a Bessel-egyenlőtlenségről szóló 17.6 gondolatot, a bevezetett merőleges vetítés és hosszúság fogalom felhasználásával. Figyeljünk arra, hogy az M -beli ortonormált bázis léte, még mindig az állítás feltévése.

17.14. állítás (Bessel). *Legyen $(V, \langle \cdot, \cdot \rangle)$ skalárriszorozatos-tér. Tegyük fel, hogy $\{x_1, \dots, x_k\}$ egy ortonormált rendszer, $M = \text{lin} \{x_1, \dots, x_k\}$ és $p_M \in L(V)$ a merőleges vetítés az M alterre. Ekkor minden $u \in V$ mellett*

1. $\|p_M(u)\| \leq \|u\|$,
2. $u - p_M(u) \in M^\perp$.

Bizonyítás: Világos, hogy $\|p_M(u)\|^2 = \|\sum_{j=1}^k \langle u, x_j \rangle x_j\|^2 = \sum_{j=1}^k \|\langle u, x_j \rangle x_j\|^2 = \sum_{j=1}^k |\langle u, x_j \rangle|^2 \leq \langle u, u \rangle = \|u\|^2$ a 17.6. állítás szerint. Hasonlóan $u - p_M(u) \in \{x_1, \dots, x_k\}^\perp = \text{lin} \{x_1, \dots, x_k\}^\perp = M^\perp$. \square

17.15. állítás (Gram – Schmidt-ortogonalizáció). *Legyen a V skalárriszorozatos-térben $\{y_1, \dots, y_k\}$ egy lineárisan független vektorrendszer. Ehhez létezik olyan $\{x_1, \dots, x_k\}$ ortonormált rendszer, amelyre minden $i = 1, \dots, k$ mellett $x_i \in \text{lin} \{y_1, \dots, y_i\}$.*

Bizonyítás: Teljes indukció, a lineárisan független rendszer elemszáma szerint. Ha $k = 1$, akkor $y_1 \neq 0$, így az $x_1 = \frac{1}{\|y_1\|} y_1$ jelöléssel az $\{x_1\}$ egy egyelemű ortonormált rendszer, amelyre $x_1 \in \text{lin} \{y_1\}$ is teljesül.

Tegyük fel, hogy igaz az állítás k -nál kisebb elemszámú lineárisan független rendszerre és lássuk be $k \geq 2$ -re. Jelölje $M = \text{lin} \{y_1, \dots, y_{k-1}\}$. Az indukciós feltevés szerint létezik $\{x_1, \dots, x_{k-1}\}$ ortonormált rendszer, amelyre még az $\{x_1, \dots, x_{k-1}\} \subseteq M$ tartalmazás is fennáll. No de, $\dim(M) = k-1$, ezért $\{x_1, \dots, x_{k-1}\}$ nem csak lineárisan független, de maximális lineárisan független rendszer M -ben, ergo generátorrendszer is, ezért

$$\text{lin} \{x_1, \dots, x_{k-1}\} = M$$

is teljesül. Találtunk tehát az M altérben egy ortonormált bázist, így evvel az ortonormált bázissal felírt merőleges vetítés értelmesen definiált. Legyen ez p_M , azaz minden $u \in V$ vektorra $p_M(u) = \sum_{j=1}^{k-1} \langle u, x_j \rangle x_j$. Tekintsük az

$$\hat{y}_k = y_k - p_M(y_k)$$

vektort. Ha ez a zérus vektor lenne, akkor $y_k \in M$ lenne, ami ellentmondana az $\{y_1, \dots, y_k\}$ rendszer lineárisan függetlenségének. Ezek szerint az

$$x_k = \frac{1}{\|\hat{y}_k\|} \hat{y}_k$$

vektorra $\|x_k\| = 1$, $x_k \in M^\perp$, ezért az $\{x_1, \dots, x_k\}$ vektorrendszer valóban ortonormált.

Már csak azt kell igazolnunk, hogy minden $j \leq k$ mellett $x_j \in \text{lin} \{y_1, \dots, y_j\}$. Ez a $j \leq k-1$ esetben az indukciós feltevés miatt van igy. A $j = k$ esetre $p_M(y_k) \in M = \text{lin} \{y_1, \dots, y_{k-1}\}$, emiatt $\hat{y}_k \in \text{lin} \{y_1, \dots, y_k\}$, amiből persze egy \mathbb{K} testbeli elemmel való szorzás után már $x_k \in \text{lin} \{y_1, \dots, y_k\}$ is következik. \square

A fenti bizonyítás kiemelt soraiból leolvasható az ú.n. *Gram – Schmidt-ortogonalizációs algoritmus*. Ha adott az $\{y_1, \dots, y_n\}$ lineárisan független vektorok rendszere, akkor abból egy $\{x_1, \dots, x_n\}$ ortonormált rendszer konstruálható. Az első lépés:

$$x_1 = \frac{1}{\|y_1\|} y_1.$$

Ha az első $k-1$ darab vektort már definiáltuk úgy, hogy $\{x_1, \dots, x_{k-1}\}$ ortonormált rendszer, akkor a k -adik vektort a következőképpen kapjuk. Legyen $\hat{y}_k = y_k - p_M(y_k)$, ahol $p_M(y_k) = \sum_{j=1}^{k-1} \langle y_k, x_j \rangle x_j$, majd

$$x_k = \frac{1}{\|\hat{y}_k\|} \hat{y}_k.$$

A Gram – Schmidt-ortogonalizációról szóló állítás azonnali következménye az ortonormált bázisok egzisztenciáját biztosító állítás:

17.16. állítás (o.n.b). *Egy véges dimenziós skalárriszorozatos-térnek minden van ortonormált bázisa, azaz olyan $\{x_1, \dots, x_n\}$ ortonormált rendszere, amely generátorrendszer is.*

Ha $\{x_1, \dots, x_n\}$ egy ortonormált bázis, akkor minden $v \in V$ mellett $v = \sum_{j=1}^n \langle v, x_j \rangle x_j$.

Bizonyítás: Tudjuk, hogy V vektortérnek van $\{y_1, \dots, y_n\}$ bázisa. Vegyük egy ilyen bázist, tehát lineárisan független rendszert, majd alkalmazzuk a 17.15. állítást. Így kapunk egy $\{x_1, \dots, x_n\} \subseteq V$ ortonormált rendszert. No de, $\dim(V) = n$ és az ortonormált rendszer elemei közt a zérus vektor nem szerepel, így az ortonormált rendszer lineárisan független rendszer is. Mivel annak $n = \dim(V)$ darab eleme van, ezért az maximális lineárisan független rendszer V -ben, ergo bázis is. Konstruáltunk tehát egy ortonormált bázist.

Jelölje most P_V egy $\{x_1, \dots, x_n\}$ ortonormált bázis definíálta merőleges vetítést az egész V vektortérre, mint egy altérre. Világos, hogy P_V az identitás függvény, hiszen minden vektorra igaz, hogy saját maga a hozzá legközelebbi V -beli pont. Így minden $v \in V$ mellett $v = p_V(v) = \sum_{j=1}^n \langle v, x_j \rangle x_j$. \square

Ez utóbbi gondolat szerint, ha egy $\{x_1, \dots, x_n\}$ ortonormált bázis rögzítünk, akkor tetszőleges $x \in V$ vektornak ebben a bázisban felírt koordináta-vektora

$$[x]_{\{x_1, \dots, x_n\}} = \begin{pmatrix} \langle x, e_1 \rangle \\ \vdots \\ \langle x, e_n \rangle \end{pmatrix}$$

Ugynéz, kicsit eltérő szintaxisossal, hogy az ortonormált bázishoz tartozó i -edik duális lineáris funkcionál, amit korábban x_i^* -al jelöltünk, az $x \mapsto \langle x, x_i \rangle$ függvény.

17.5. Projekciós téTEL

Láttuk tehát, hogy egy véges dimenziós altérnek minden van ortonormált bázisa. Ezért ha egy véges dimenziós altérre való merőleges vetítést akarunk használni, nem kell kikötnünk, hogy legyen az altérben egy ortonormált bázis, hiszen ilyen minden van. Persze ahogyan bázis is sok van, ugyanúgy ortonormált bázisból is nagyon sok lehet. Meggondoltuk viszont, hogy függetlenül attól, hogy az $\{x_1, \dots, x_k\} \subseteq M$ vagy az $\{y_1, \dots, y_k\} \subseteq M$ ortonormált bázis segítségével írjuk fel az M altérre vonatkozó merőleges vetítést a vetület minden azonos, nevezetesen a ponthoz az M -beli legközelebbi pont:

$$\sum_{j=1}^k \langle u, x_j \rangle x_j = P_M(u) = \sum_{j=1}^k \langle u, y_j \rangle y_j.$$

Ennek fényében először is érdemes újra és újra fogalmazni a Bessel-egyenlőtlenségről szóló 17.14. gondolatot. Figyeljünk arra, hogy az M -beli ortonormált bázis léte nem szerepel a feltételek között, hanem helyette az altér véges dimenziós volta szerepel.²

17.17. állítás (Bessel). *Legyen $(V, \langle \cdot, \cdot \rangle)$ skaláriszorzatos-tér, és M egy véges dimenziós altere. Ekkor minden $u \in V$ -hez*

1. létezik egyetlen $u^* \in M$ vektor, amelyre $d(u, u^*) = \min \{d(u, w) : w \in M\}$. Erre a vektorra teljesül az $u^* = p_M(u) = \sum_{i=1}^k \langle u, x_i \rangle x_i$ egyenlőség, az M altér tetszőlegesen választott $\{x_1, \dots, x_k\}$ ortonormált bázisa mellett;
2. $\|p_M(u)\| \leq \|u\|$;
3. $u - p_M(u) \in M^\perp$.

Bizonyítás: Mivel M a V skaláriszorzatos tér egy véges dimenziós altere, ezért az M altérnek van ortonormált bázisa, amint azt a 17.16. állításban Gram–Schmidt-ortogonalizációt láttuk. Legyen ez $\{x_1, \dots, x_k\}$. Írjuk fel ezzel az ortonormált rendszerrel a 17.12. legközelebbi pontról szóló, és a 17.14. Bessel-egyenlőtlenségről szóló állításokat. \square

17.18. állítás (projekciós-tétel). *Legyen V egy skaláriszorzatos-tér, és M egy véges dimenziós altere. Ekkor*

$$M \oplus M^\perp = V.$$

²Majd mikor funkcionálanalízist és mértékelméletet tanulunk látjuk, hogy a feltétel még tovább gyengíthető, arra az esetre, mikor M a V Hilbert-tér egy zárt altere.

Bizonyítás: Az M és az M^\perp egymástól diszjunkt alterek. Tetszőleges $u \in V$ vektorra fennáll az

$$u = p_M(u) + (u - p_M(u))$$

nyilvánvaló egyenlőség. Itt az első vektorra $p_M(u) \in M$, és a második vektorra $u - p_M(u) \in M^\perp$ a 17.17. állítás szerint. Ez azt jelenti, hogy $M + M^\perp = V$ azonosság is teljesül, ahol $M + M^\perp$ az M és az M^\perp ortokomplementének a Minkowski-összege. \square

17.19. állítás. Legyen V egy skalárisszorzatos-tér. Ekkor

1. minden $M \subseteq V$ véges dimenziós altér mellett $M = M^{\perp\perp}$,
2. ha a $H \subseteq V$ részhalmazra a $\text{lin } H$ egy véges dimenziós altér, akkor $\text{lin } H = H^{\perp\perp}$.

Bizonyítás: Láttuk, hogy $M \subseteq M^{\perp\perp}$ minden halmazra is igaz. Legyen a fordított tartalmazás igazolásához, $u \in M^{\perp\perp}$. A projekciós téTEL miatt $u = u_1 + u_2$, ahol $u_1 \in M$ és $u_2 \in M^\perp$. Így

$$\langle u_2, u_2 \rangle = \langle u - u_1, u_2 \rangle = \langle u, u_2 \rangle - \langle u_1, u_2 \rangle = 0.$$

Itt a jobboldali első szám azért zérus, mert $u \in M^{\perp\perp}$ és $u_2 \in M^\perp$, a második meg azért zérus, mert $u_1 \in M$ és $u_2 \in M^\perp$. Azt kaptuk tehát, hogy $u_2 = 0$, amiből már látszik is, hogy $u = u_1 \in M$.

Tetszőleges $H \subseteq V$ mellett $\text{lin } H$ egy véges dimenziós altér, így a már igazolt állítás miatt $\text{lin } H = (\text{lin } H)^{\perp\perp}$. No de, már korábban is láttuk, hogy $H^\perp = (\text{lin } H)^\perp$, így

$$\text{lin } H = (\text{lin } H)^{\perp\perp} = ((\text{lin } H)^\perp)^\perp = (H^\perp)^\perp = H^{\perp\perp}.$$

\square

17.20. állítás. Legyen most már V egy véges dimenziós skalárisszorzatos-tér. Tegyük fel, hogy $M, N \subseteq V$ olyan alterek, amelyekre $M + N = V$ és $M \perp N$. Ekkor $N = M^\perp$.

Bizonyítás: Világos, hogy M és N diszjunkt alterek is, ezért $M \oplus N = V = M \oplus M^\perp$ a projekciós téTEL szerint. ekkor persze N és M^\perp izomorf alterek, amelyekre $N \subseteq M^\perp$ tartalmazás is teljesül. Ez csak $N = M^\perp$ módon teljesülhet. \square

17.6. Ortonormált rendszer teljessége

17.21. definíció (T.O.N.R.). Legyen $\{e_1, \dots, e_k\}$ egy ortonormált rendszer a V skalárisszorzatos-térben. Azt mondjuk, hogy ez egy teljes ortonormált rendszer, ha

$$\langle u, e_j \rangle = 0 \quad \forall j = 1, \dots, k \implies u = 0$$

implikáció teljesül.

Először is vegyük észre, hogy a fenti implikáció ekvivalens az $\{e_1, \dots, e_k\}^\perp = \{0\}$ feltétellel.

17.22. állítás. Legyen V egy skalárisszorzatos-tér, és tekintsünk egy a továbbiakban rögzített $\{e_1, \dots, e_n\}$ egy ortonormált rendszerét. Az alábbi feltévések ekvivalensek.

1. Az $\{e_1, \dots, e_n\}$ vektorrendszer teljes ortonormált rendszer;
2. Az $\{e_1, \dots, e_n\}$ vektorrendszer ortonormált bázis;
3. minden $x \in V$ mellett $x = \sum_{j=1}^n \langle x, e_j \rangle e_j$;
4. minden $x, y \in V$ mellett $\langle x, y \rangle = \sum_{j=1}^n \langle x, e_j \rangle \overline{\langle y, e_j \rangle}$;
5. minden $x \in V$ mellett $\|x\|^2 = \sum_{j=1}^n |\langle x, e_j \rangle|^2$.

Bizonyítás: Körbe bizonyítunk:

1. \Rightarrow 2. Mivel egy ortonormált rendszer egy zérus elemet nem tartalmazó ortogonális rendszer, ezért lineárisan független. Jelölje $H = \{e_1, \dots, e_n\}$. Világos, hogy $\text{lin } H$ véges dimenziós, ezért a projekciós tétel előző következménye miatt $\text{lin } H = H^{\perp\perp}$. No de, a teljesség szerint $H^{\perp} = \{0\}$, ergo

$$\text{lin } H = H^{\perp\perp} = 0^{\perp} = V.$$

Ez azt jelenti, hogy a H vektorrendszer egy lineárisan független generátorrendszer, ergo bázis.

2. \Rightarrow 3. Már beláttuk a 17.16. állítás második felében.

3. \Rightarrow 4. Jelölje $x \in V$ mellett $\xi_j = \langle x, e_j \rangle$ és $y \in V$ mellett $\eta_j = \langle y, e_j \rangle$. A 3. feltevés szerint

$$\langle x, y \rangle = \left\langle \sum_{j=1}^n \xi_j e_j, \sum_{k=1}^n \eta_k e_k \right\rangle = \sum_{j=1}^n \sum_{k=1}^n \xi_j \overline{\eta_k} \langle e_j, e_k \rangle = \sum_{j=1}^n \sum_{k=1}^n \xi_j \overline{\eta_k} \delta_{j,k} = \sum_{j=1}^n \xi_j \overline{\eta_j} = \sum_{j=1}^n \langle x, e_j \rangle \overline{\langle y, e_j \rangle}.$$

4. \Rightarrow 5. Az $x = y$ speciális esetben 4. szerint

$$\|x\|^2 = \langle x, x \rangle = \sum_{j=1}^n \langle x, e_j \rangle \overline{\langle x, e_j \rangle} = \sum_{j=1}^n |\langle x, e_j \rangle|^2.$$

5. \Rightarrow 1. Ha fennáll 5. és $v \in \{e_1, \dots, e_n\}^{\perp}$, akkor minden j mellett $\langle v, e_j \rangle = 0$, ezért

$$\|v\|^2 = \sum_{j=1}^n \langle v, e_j \rangle \overline{\langle v, e_j \rangle} = \sum_{j=1}^n \langle v, e_j \rangle \overline{0} = 0.$$

emiatt $v = 0$ is fennáll. Ez azt jelenti, hogy $\{e_1, \dots, e_n\}^{\perp} = \{0\}$, ergo az $\{e_1, \dots, e_n\}$ ortonormált rendszer teljes. \square

A fejezet elején láttunk konkrét példát véges dimenziós skalárisszorzatos-térre: egy rögzített bázisban felírt belsőszorzatot. Persze ebben a skalárisszorzatban, a kiinduló bázis ortonormált bázissá is válik. Most azt látjuk, hogy a fenti példán kívül nem is lehet más skalárisszorzatot definiálni. minden skalárisszorzat ugyanis, valamely bázishoz tartozó belső szorzat. Pontosan ezt állítja a fenti 4-es tulajdonság.

Ezt a tényt a fontossága miatt külön is kiemeljük:

17.23. következmény. minden véges dimenziós skalárisszorzatos-térben a skalárisszorzat nem más, mint erre a skalárisszorzatra nézve, de tetszőleges ortonormált bázishoz felírt belső szorzat.

18. fejezet

Az adjungált operátor bevezetése

Az egész fejezetben V egy \mathbb{K} feletti vektortér.

18.1. Riesz-reprezentáció véges dimenziós skaláriesszorzatos-térben

18.1. állítás (Riesz-reprezentáció). *Legy V egy véges dimenziós skaláriesszorzatos-tér. Ekkor minden $\varphi : V \rightarrow \mathbb{K}$ lineáris funkcionálhoz létezik egyetlen $z_\varphi \in V$ vektor, amelyre minden $x \in V$ mellett*

$$\varphi(x) = \langle x, z_\varphi \rangle$$

Ezt az egyetlen z_φ vektort nevezzük a φ lineáris funkcionált reprezentáló vektornak.

Bizonyítás: Először az unicitást lássuk be. Tegyük fel, hogy $z_1, z_2 \in V$ is reprezentáló vektorok, azaz minden $x \in V$ mellett

$$\langle x, z_1 \rangle = \varphi(x) = \langle x, z_2 \rangle.$$

Ekkor $(z_1 - z_2) \perp x$ minden $x \in V$ -re tehát $z_1 - z_2 = 0$, ergo $z_1 = z_2$.

Az egzisztencia igazolásához legyen $\{v_1, \dots, v_n\}$ egy ortonormált bázisa a V skaláriesszorzatos-térnek. Tudjuk – lásd a 17.16. állítást –, hogy minden $x \in V$ vektor $x = \sum_{j=1}^n \langle x, v_j \rangle v_j$ alakú. Emiatt

$$\varphi(x) = \varphi\left(\sum_{j=1}^n \langle x, v_j \rangle v_j\right) = \sum_{j=1}^n \langle x, v_j \rangle \varphi(v_j) = \sum_{j=1}^n \langle x, \overline{\varphi(v_j)} v_j \rangle = \langle x, \sum_{j=1}^n \overline{\varphi(v_j)} v_j \rangle.$$

Ha tehát bevezetjük a $z_\varphi = \sum_{j=1}^n \overline{\varphi(v_j)} v_j$ jelölést, akkor készen is vagyunk. \square

A bizonyításból az is látszik, hogy amennyiben $\{v_1, \dots, v_n\}$ a tér akármelyik ortonormált bázisa, akkor a reprezentáló vektort a

$$z_\varphi = \sum_{i=1}^n \overline{\varphi(v_i)} v_i$$

formula szolgáltatja. Mivel csak egyetlen reprezentáló vektor van, ezért ha például $\{x_1, \dots, x_n\}$ egy másik ortonormált bázisa a térnek, akkor

$$\sum_{i=1}^n \overline{\varphi(v_i)} v_i = z_\varphi = \sum_{i=1}^n \overline{\varphi(x_i)} x_i.$$

18.2. Az adjungált operáció

18.2. definíció (adjungált, adjungált azonosság). *Legyen $(V, \langle \cdot, \cdot \rangle_V)$ és $(W, \langle \cdot, \cdot \rangle_W)$ véges dimenziós skaláriesszorzatos-tér. Rögzített $A \in L(V, W)$ lineáris operátor és rögzített $w \in W$ vektor mellett tekintsük a*

$$\varphi_w = \langle \cdot, w \rangle_W \circ A$$

függvényt. Világos, hogy $\varphi_w : V \rightarrow \mathbb{K}$ lineáris funkcionál V -n, így a Riesz-reprezentációs tétel szerint létezik egyetlen $w^* \in V$, amelyre $\varphi_w(v) = \langle v, w^* \rangle_V$ minden $v \in V$ mellett teljesül.

Jelölje most A^* a $w \mapsto w^*$ függvényt, tehát $A^* : W \rightarrow V$. Ekkor minden $v \in V$ és $w \in W$ mellett fennáll az ú.n. adjungált azonosság:

$$\langle Av, w \rangle_W = \varphi_w(v) = \langle v, w^* \rangle_V = \langle v, A^*w \rangle_V. \quad (\dagger)$$

Ezt az $A^* : W \rightarrow V$ függvényt nevezzük az A lineáris transzformáció adjungáltjának.

18.3. állítás. Legyenek V, W skaláriszorzatos-terek, és $A \in L(V, W)$ lineáris operátor. Ekkor az A^* adjungált az egyetlen $W \rightarrow V$ függvény, amelyre a (\dagger) adjungált azonosság fennáll.

Bizonyítás: Legyen $B : W \rightarrow V$ olyan függvény, amelyre szintén fennáll adjungált azonosság, azaz

$$\langle v, Bw \rangle = \langle Av, w \rangle = \langle v, A^*w \rangle$$

fennáll minden $v \in V$ és $w \in W$ mellett. Mivel a skaláriszorzás rögzített első változó mellett a második változóban additív, ezért $\langle v, (A^* - B)w \rangle = 0$ fennáll minden $v \in V$ és $w \in W$ mellett. Speciálisan ez igaz $v = (A^* - B)w$ -re is, ergo minden $w \in W$ -re $\langle v, (A^* - B)w \rangle = 0$. Ez éppen azt jelenti, hogy $A^* = B$. \square

Látni fogjuk, hogy a (\dagger) adjungált azonosság szinte fontosabb mint a definíció. A definíció arra való, hogy megmutassuk: adott $A \in L(V, W)$ lineáris transzformációhoz van olyan $A^* : W \rightarrow V$ függvény, amelyre minden $v \in V$ és minden $w \in W$ mellett fennáll az adjungált azonosság:

$$\langle Av, w \rangle = \langle v, A^*w \rangle. \quad (\dagger)$$

Az adjungált unicitása, és összes fontos tulajdonsága az adjungált azonosságon alapul.

Az adjungált függvény linearitása. A 18.1. Riesz-reprezentációs tétel során meggondoltuk, hogy ha tetszőlegesen adott egy $\{v_1, \dots, v_n\} \subseteq V$ ortonormált bázis, akkor

$$A^*w = \sum_{j=1}^n \overline{\varphi_w(v_j)} v_j = \sum_{j=1}^n \overline{\langle Av_j, w \rangle}_W v_j = \sum_{j=1}^n \langle w, Av_j \rangle_W v_j.$$

Ebből már adódik, hogy $A^* : W \rightarrow V$ lineáris operátor. Ugyanis a skaláriszorzat rögzített második változó mellett az első változóban lineáris, emiatt egy rögzített j esetén az $w \mapsto \langle w, Av_j \rangle$ a w változó lineáris függvénye. A fenti kiemelt sor szerint A^* ilyen függvények összegeként maga is lineáris.

Alternatív indoklást kapunk az adjungált operáció linearitására az (\dagger) adjungált azonosság háromszori felírásával. Itt kihasználjuk a tényt, hogy a skaláriszorzás rögzített első változó mellett a második változó konjugáltan lineáris függvénye.

$$\begin{aligned} \langle u, A^*(\alpha_1 w_1 + \alpha_2 w_2) \rangle &= \langle Au, \alpha_1 w_1 + \alpha_2 w_2 \rangle = \\ \bar{\alpha}_1 \langle Au, w_1 \rangle + \bar{\alpha}_2 \langle Au, w_2 \rangle &= \bar{\alpha}_1 \langle u, A^*w_1 \rangle + \bar{\alpha}_2 \langle u, A^*w_2 \rangle = \langle u, \alpha_1 A^*w_1 \rangle + \langle u, \alpha_2 A^*w_2 \rangle = \\ &= \langle u, \alpha_1 A^*w_1 + \alpha_2 A^*w_2 \rangle. \end{aligned}$$

Mivel ez minden $u \in V$ és minden $w_1, w_2 \in W$ esetén fennáll, ezért az

$$A^*(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 A^*w_1 + \alpha_2 A^*w_2$$

azonosság is teljesül.

18.4. állítás. Legyen V és W véges dimenziós skaláriszorzatos-tér. Rögzítsünk egy $\{v_1, \dots, v_m\}$ ortonormált bázist V -ben, és rögzítsünk egy $\{w_1, \dots, w_n\}$ ortonormált bázist W -ben. Legyen $A \in L(V, W)$ egy lineáris operátor és $A^* \in L(W, V)$ az adjungáltja. Ekkor A -nak a fenti bázisokban rögzített mátrixa azonos A^* mátrix sorainak és oszlopainak felcseréléssel, majd konjugálásával kapott mátrixszal. Formálisan:

$$[A]_{i,j} = \overline{[A^*]_{j,i}}$$

minden $j = 1, \dots, m$ és minden $i = 1, \dots, n$ esetén.

Bizonyítás: $[A]_{i,j} = \langle Av_j, w_i \rangle = \langle v_j, A^*w_i \rangle = \overline{\langle A^*w_i, v_j \rangle} = \overline{[A^*]_{j,i}}$, a 17.16. állítás, a (\dagger) adjungált azonosság, és a skaláriszorzás tulajdonságai szerint. \square

Most összefoglaljuk az adjungálás operáció legfontosabb tulajdonságait, amelyeket későbbi számolásaink során külön hivatkozás nélkül használunk.

18.5. állítás. Legyenek V és W véges dimenziós skaláriszorzatos terek, $A, B \in L(V, W)$ lineáris operátorok. Ekkor

1. $(A + B)^* = A^* + B^*$,
2. $(\alpha A)^* = \bar{\alpha} A^*$,
3. $(A^*)^* = A$

Ha $I \in L(V)$ jelöli az identitás transzformációt, továbbá $A, B \in L(V)$, akkor

4. $I^* = I$,
5. $(AB)^* = B^*A^*$.

6. Az A lineáris operátor pontosan akkor reguláris ha A^* adjungált operáció is az. Ebben az esetben az adjungált operáció inverze éppen az inverz operáció adjungáltja, magyarul $(A^*)^{-1} = (A^{-1})^*$.

Bizonyítás: Az első öt tulajdonság teljesen nyilvánvaló az adjungált mátrixának felírásából. Csak azért, hogy hangsúlyozzuk az adjungált azonosság hasznosságát a mátrixokra való hivatkozások nélküli indoklást is adunk.

1., 2.

$$\begin{aligned} \langle u, (\alpha A + B)^* w \rangle &= \\ \langle (\alpha A + B) u, w \rangle &= \alpha \langle A u, w \rangle + \langle B u, w \rangle = \alpha \langle u, A^* w \rangle + \langle u, B^* w \rangle = \langle u, \bar{\alpha} A^* w \rangle + \langle u, B^* w \rangle = \\ &= \langle u, (\bar{\alpha} A^* + B^*) w \rangle \end{aligned}$$

Ez minden $\alpha \in \mathbb{K}$, $u \in V$ és minden $w \in W$ mellett, tehát $(\alpha A + B)^* = \bar{\alpha} A^* + B^*$.

3. $\langle A u, w \rangle = \langle u, A^* w \rangle = \overline{\langle A^* w, u \rangle} = \overline{\langle w, A^{**} u \rangle} = \langle A^{**} u, w \rangle$.
4. $\langle u, I v \rangle = \langle I u, v \rangle = \langle u, I^* v \rangle$.
5. $\langle u, (AB)^* v \rangle = \langle AB u, v \rangle = \langle Bu, A^* v \rangle = \langle u, B^* A^* v \rangle$.
6. A pontosan akkor reguláris, ha létezik B , amelyre $AB = I$, ami ekvivalens $B^* A^* = I$ -vel, ami ekvivalens avval, hogy A^* reguláris. Világos, hogy ebben az esetben $(A^*)^{-1} = B^* = (A^{-1})^*$. \square

Felépítésünk sarokköve a skaláriszorzatos terek rangtétele. Legfontosabb következménye, hogy zérus vektoron kívül nincs olyan vektor, amely egyszerre A képteréhez is, de A^* magteréhez is hozzátartozik.

18.6. állítás (skaláriszorzatos Rang-tétel). Legyenek V, W skaláriszorzatos terek, $A \in L(V, W)$ lineáris operáció. Ekkor $(\text{Im } A)^\perp = \ker A^*$ és $(\ker A)^\perp = \text{Im } A^*$. Emiatt

$$\begin{aligned} \text{Im } A \oplus \ker A^* &= W; & \text{Im } A \perp \ker A^*; \\ \ker A \oplus \text{Im } A^* &= V; & \ker A \perp \text{Im } A^*, \end{aligned}$$

így $\rho(A) = \rho(A^*)$.

Bizonyítás: Az adjungált azonosság szerint egy $w \in W$ vektorra az alábbi feltevések ekvivalensek

$$\begin{aligned} w \in (\text{Im } A)^\perp &\iff \langle Av, w \rangle = 0 \forall v \in V \iff \langle v, A^* w \rangle = 0 \forall v \in V \iff A^* w \in V^\perp \iff A^* w = 0 \\ &\iff w \in \ker A^*, \end{aligned}$$

ami igazolja az $(\text{Im } A)^\perp = \ker A^*$ azonosságot.

A projekciós tételet alkalmazva: $\text{Im } A = (\text{Im } A)^{\perp\perp} = (\ker A^*)^\perp$. Ez persze az A^* adjungált operátorra is igaz, ergo $\text{Im } A^* = (\ker A)^\perp$.

Újra felírva a projekciós tételel: $W = \text{Im } A \oplus (\text{Im } A)^\perp = \text{Im } A \oplus \ker A^*$ valamint $V = \ker A \oplus (\ker A)^\perp = \ker A \oplus \text{Im } A^*$.

A direkt összegek dimenziójára vonatkozó állítás és a Rang–defektus-tételnek az adjungált operációra vonatkozó alakja szerint $\rho(A) = \dim(W) - \nu(A^*) = \rho(A^*)$. \square

18.3. Önadzungált transzformációk

18.7. definíció. Legyen V egy skalárisszorzatos-tér, és $A \in L(V)$ egy lineáris transzformáció. Azt mondjuk, hogy az A transzformáció *önadzungált*, ha $A^* = A$.

18.8. állítás. A skalárisszorzatos-tér egy A lineáris transzformációja pontosan akkor önadzungált, ha minden $u, w \in V$ mellett

$$\langle Au, v \rangle = \langle u, Av \rangle.$$

Ha felírjuk a transzformáció mátrixát egy ortonormált bázisban, akkor messziről látszik, hogy a transzformáció önadzungált vagy sem.

18.9. állítás. A skalárisszorzatos-tér egy A lineáris transzformációja pontosan akkor önadzungált, valamely ortonormált bázisban felírt mátrixára

$$[A]_{i,j} = \overline{[A]_{j,i}}.$$

Persze a fenti tulajdonság, ha egy ortonormált bázis mellett teljesül, akkor A önadzungált, ezért minden más ortonormált bázisban is igaz marad.

A skalárisszorzatos-terek rangtétele, – tehát a 18.6. állítás – önadzungált transzformációk mellett azt jelenti, hogy a $\ker A$ és az $\text{Im } A$ egymásra merőleges alterek, amelyek direkt összege az egész tér. Később látjuk majd, hogy ez nem csak önadzungált transzformációkra igaz, hanem transzformációk egy sokkal szélesebb osztályára¹ is teljesül.

18.4. Unitér transzformációk

18.10. definíció. Legyen V egy skalárisszorzatos-tér, és $B \in L(V)$ egy lineáris transzformáció. Azt mondjuk, hogy B unitér, ha B reguláris és $B^{-1} = B^*$.

18.11. állítás. $A B \in L(V)$ transzformációra tett alábbi feltevések ekvivalensek.

1. B unitér;
2. minden $u, v \in V$ mellett $\langle Bu, Bv \rangle = \langle u, v \rangle$.
3. B egy ortonormált bázist ortonormált bázisra képez.

Bizonyítás: Körben bizonyítunk:

1. \Rightarrow 2. $\langle Bu, Bv \rangle = \langle u, B^*Bv \rangle = \langle u, B^{-1}Bv \rangle = \langle u, v \rangle$.
2. \Rightarrow 3. Legyen u_1, \dots, u_n egy ortonormált bázis. Ennek B képére $\langle Bu_j, Bu_k \rangle = \langle u_j, u_k \rangle = \delta_{j,k}$. Ez azt jelenti, hogy a $\{Bu_1, \dots, Bu_n\}$ egy ortonormált rendszer. A tér n dimenziós, hiszen $\{u_1, \dots, u_n\}$ egy bázis. Így $\{Bu_1, \dots, Bu_n\}$ egy maximális lineárisan független rendszerként maga is bázis.
3. \Rightarrow 1. Vegyük V -ben egy $\{u_1, \dots, u_n\}$ ortonormált bázist. Feltevésünk szerint a $\{Bu_1, \dots, Bu_n\}$ egy ortonormált rendszer. Így minden j, k mellett

$$\langle u_k, (B^*B - I)u_j \rangle = \langle u_k, B^*Bu_j \rangle - \langle u_k, u_j \rangle = \langle Bu_k, Bu_j \rangle - \langle u_k, u_j \rangle = \delta_{k,j} - \delta_{k,j} = 0.$$

Ez azt jelenti, hogy az $(B^*B - I)u_j$ vektor merőleges az $\{u_1, \dots, u_n\}$ bázis minden elemére. Az ortonormált rendszer teljessége szerint ez csak úgy lehetséges, ha minden j indexre $B^*Bu_j = Iu_j$. No de, a lineáris transzformációk bázison egyértelműen meghatározottak, ezért $B^*B = I$. Ez azt jelenti, hogy B reguláris és $B^{-1} = B^*$.

□

Egy ortonormált bázisban felírva a B transzformáció mátrixát, nagyon könnyen észrevehető, hogy unitér transzformációval állunk szemben. Ugyanis B mátrixának j -edik oszlopa a j -edik bázis elem képe. Így ortonormált bázisból kiindulva a transzformáció pontosan akkor unitér, ha a felírt $[B]$ mátrix oszlopai egy ortonormált rendszert alkotnak az eredeti bázisban felírt belső szorzattal.

¹Lásd a 19.4. állítást

18.12. állítás. Legyen V egy skaláriSSzorzos-tér és $B \in L(V)$ egy lineáris transzformáció. A B transzformáció pontosan akkor unitér, ha tetszőleges ortonormált bázisban felírt mátrixának oszloprendszer egy ortonormált rendszer.

Egyszerűbben kifejezve, ha B ortonormált bázisban felírt mátrixának elemei $\beta_{i,j}$, akkor a B unitér voltának szükséges és elegendő feltétele, hogy minden i, j -re $\sum_{k=1}^n \beta_{k,i} \overline{\beta_{k,j}} = \delta_{i,j}$. Ultra fontos következmény az alábbi:

18.13. állítás. Tegyük fel, hogy a V skaláriSSzorzos-tér, egy A lineáris transzformációjának mátrixát felírtuk két ortonormált bázisban. Legyen $\{u_1, \dots, u_n\}$ a régi ortonormált bázis és $\{e_1, \dots, e_n\}$ az új ortonormált bázis. Jelölje B az áttérés lineáris transzformációt, tehát $Bu_j = e_j$ minden $j = 1, \dots, n$ mellett. Ekkor

$$[A]_{\text{új}} = [\overline{B}]^T [A]_{\text{régi}} [B].$$

Normális transzformációk diagonalizálhatósága

19.1. definíció (normális lineáris transzformáció). A V skaláriszorzatos-tér egy $A \in L(V)$ lineáris transzformációját *normálisnak* nevezzük, ha

$$N^*N = NN^*$$

tehát, ha a transzformáció kommutál az adjungáltjával.

Nyilvánvaló példák normális transzformációkra:

1. Ha N örnadjungált;
2. Ha N reguláris és $N^{-1} = N^*$;¹
3. Ha létezik olyan q polinom, amelyre $N^* = q(N)$;
4. Ha N egy ortonormált bázisban diagonalizálható, akkor ebben a bázisban felírt matrixszára, a mátrix kommutál az adjungált mátrixszal, így az eredeti transzformáció is normális.

Ki fog derülni a fejezetben, hogy a fenti példák közül a 3. és a 4. le is fedi valamennyi normális transzformációt. Az unitér transzformációk olyanok, amelyek \mathbb{R} feletti esetben általában nem diagonalizálhatók.

19.2. állítás. *Egy normális transzformáció polinomja is normális.*

Bizonyítás: Először gondoljuk meg, hogy $(N^j)^*$ és N^k kommutálnak minden $j, k \geq 0$ mellett. Ez indukcióval könnyű:

$$\begin{aligned} (N^j)^*N^k &= (N^*)^jN^k = (N^*)^{j-1}N^*NN^{k-1} = (N^*)^{j-1}NN^*N^{k-1} \\ &= N(N^*)^{j-1}N^{k-1}N^* = N(N^{j-1})^*N^{k-1}N^* = NN^{k-1}(N^{j-1})^*N^* = N^k(N^*)^j = N^k(N^j)^*. \end{aligned}$$

Most legyen $p \in \mathbb{K}[t]$ egy plinom. Ha $p(t) = \sum_{j=0}^m \alpha_j t^j$, akkor

$$\begin{aligned} (p(N))^* p(N) &= \left(\sum_{j=0}^m \alpha_j N^j \right)^* \left(\sum_{k=0}^m \alpha_k N^k \right) = \left(\sum_{j=0}^m \bar{\alpha}_j (N^j)^* \right) \left(\sum_{k=0}^m \alpha_k N^k \right) \\ &= \sum_{j=0}^m \sum_{k=0}^m \bar{\alpha}_j \alpha_k (N^j)^* N^k = \sum_{k=0}^m \sum_{j=0}^m \alpha_k \bar{\alpha}_j N^k (N^j)^* \\ &= \left(\sum_{k=0}^m \alpha_k N^k \right) \left(\sum_{j=0}^m \bar{\alpha}_j (N^j)^* \right) = \left(\sum_{k=0}^m \alpha_k N^k \right) \left(\sum_{j=0}^m \alpha_j N^j \right)^* = p(N) (p(N))^*. \end{aligned}$$

Ezt kellett belátni. □

19.3. állítás. *Legyen N egy normális transzformáció. Ekkor $\ker N^* = \ker N$.*

¹Az ilyen transzformációt nevezzük *unitér* transzformációnak.

Bizonyítás: Legyen $u \in V$. Ekkor

$$\|Nu\|^2 = \langle Nu, Nu \rangle = \langle u, N^*Nu \rangle = \langle u, NN^*u \rangle = \langle N^*u, N^*u \rangle = \|N^*u\|^2.$$

Ez azt jelenti, hogy $Nu = 0$ akkor és csak akkor áll fenn, ha N^*u is fennáll. \square

Korábban láttuk, hogy tetszőleges lineáris transzformáció mellett $(\text{Im } A)^\perp = \ker A^*$. Normális transzformáció esetén ez még egyszerűbb:

19.4. állítás (Normális transzformációk rangtétele). *Legyen V skaláriszorzatos-tér és $N \in L(V)$ egy normális lineáris transzformáció. Ekkor*

$$\text{Im } N \oplus \ker N = V \quad \text{és} \quad \text{Im } N \perp \ker N.$$

Tudjuk, hogy $\lambda \in \sigma(N)$ akkor és csak akkor, ha $\ker(N - \lambda I) \neq \{0\}$. Mivel normális transzformáció mellett az $N - \lambda I$, az N -nek polinomja lévén, maga is normális, ezért az $\ker(N - \lambda I)$ és $\ker(N^* - \bar{\lambda}I)$ azonos alterek. Azt gondoltuk meg tehát, hogy normális N mellett tetszőleges u vektorra $Nu = \lambda u$ akkor és csak akkor igaz, ha $N^*u = \bar{\lambda}u$.

19.5. állítás. *Legyen N egy normális transzformáció a V skaláriszorzatos-téren. Ekkor*

1. *$A \lambda \in \sigma(N)$ akkor és csak akkor, teljesül, ha $\bar{\lambda} \in \sigma(N^*)$;*
2. *Minden $\lambda \in \mathbb{K}$ mellett $\ker(N - \lambda I) = \ker(N^* - \bar{\lambda}I)$;*
3. *Legyen $\lambda \neq \mu \in \mathbb{K}$. Ekkor $\ker(N - \lambda I) \perp \ker(N - \mu I)$. Ha tehát λ és μ az N transzformáció különböző sajátértékei, akkor az ezekhez tartozó sajátalterek egymásra merőlegesek.*
4. *Speciálisan, ha $N = N^*$, azaz N önadjungált transzformáció, akkor $\sigma(N) \subseteq \mathbb{R}$, azaz önadjungált transzformációknak valósak a sajátértékei.*
5. *Speciálisan, ha N reguláris és $N^{-1} = N^*$, azaz N unitér transzformáció, akkor minden $\lambda \in \sigma(N) \subseteq \mathbb{R}$, esetén $|\lambda| = 1$, azaz unitér transzformációknak a sajátértékei 1 abszolútértékűek.*

Bizonyítás: Az első két állítást már meggondoltuk.

Legyen $u \in \ker(N - \lambda I)$ és $v \in \ker(N - \mu I) = \ker(N^* - \bar{\mu}I)$. Ekkor persze $Nu = \lambda u$ és $N^*v = \bar{\mu}v$. Így

$$(\lambda - \mu) \langle u, v \rangle = \langle \lambda u, v \rangle - \langle u, \bar{\mu}v \rangle = \langle Nu, v \rangle - \langle u, N^*v \rangle = \langle u, N^*v \rangle - \langle u, N^*v \rangle = 0.$$

No de, a sajátértékek nem azonosak, ezért $u \perp v$.

Legyen most N önadjungált, és $\lambda \in \sigma(N)$ egy sajátérték. Van tehát olyan $u \neq 0$ vektor, amelyre $Nu = \lambda u$ és $N^*u = \bar{\lambda}u$. Ekkor

$$(\lambda - \bar{\lambda}) u = Nu - N^*u = Nu - Nu = 0.$$

No de, az $u \neq 0$ feltétel miatt $\lambda - \bar{\lambda} = 0$, tehát $\lambda \in \mathbb{R}$.

Legyen most N unitér, és $\lambda \in \sigma(N)$ egy sajátérték. Van tehát olyan $u \neq 0$ vektor, amelyre $Nu = \lambda u$ és $N^*u = \bar{\lambda}u$. Persze $N^{-1}Nu = N^{-1}(\lambda u) = \lambda N^{-1}u$, ezért

$$\left(\frac{1}{\lambda} - \bar{\lambda}\right) u = N^{-1}u - N^*u = Nu - Nu = 0.$$

No de, az $u \neq 0$ feltétel miatt $\frac{1}{\lambda} - \bar{\lambda} = 0$, tehát $\lambda\bar{\lambda} = |\lambda|^2 = 1$. \square

Eddig minden gondolat \mathbb{K} feletti skaláriszorzatos-téren volt. A komplex és a valós diagonalizálhatóság közös magja a következő állítás.

19.6. állítás. *Legyen V egy \mathbb{K} feletti skaláriszorzatos-tér, és $N \in L(V)$ egy normális transzformáció. Ekkor N minimálpolinomja különböző, normált, irreducibilis polinomok elsőfokú hatványainak szorzata.*

Bizonyítás: Tegyük fel, hogy $p^k|m$, ahol m a minimálpolinom, p egy irreducibilis polinom, és $k \geq 2$. Láttuk, hogy bevezetve a $V_1 = \ker p^k(N)$, és $A_1 = N|_{V_1}$ jelölésekkel, az A_1 minimálpolinomja p^k . Tehát ha $B = p(N)$, akkor $B|_{V_1}$ transzformáció k -ad rendben nilpotens, létezik tehát $v \in V_1$, amelyre $B^{k-1}v \neq 0$, de $B^k v = 0$. Mivel $k - 1 \geq 1$, ezért találtunk $x = B^{k-1}v \in \text{Im } p(N)$ nem zérus vektort, amelyre $x \in \ker(p(N))$. Ez ellentmond a normális transzformációk rangtételenek. \square

Normális transzformációk diagonalizálhatósága komplex skalárisszorzatos-tereken

A fenti gondolat azonnali következménye a normális transzformációk diagonalizálhatósága \mathbb{C} felett.

19.7. állítás. Legyen V egy \mathbb{C} feletti skalárisszorzatos-tér. Ekkor minden N normális transzformáció ortonormált bázisban diagonalizálható. Ez azt jelenti, hogy létezik a térfelület N sajátvektoraiból álló ortonormált bázisa. Ebben a bázisban felírva N mátrixát egy diagonális mátrixot kapunk, amelynek diagonálisában N sajátértékei vannak.

Bizonyítás: Az algebra alaptétele szerint irreducibilis \mathbb{C} feletti polinom csak elsőfokú polinom lehet. Így a minimálpolinom, a normalitást kihasználva a fenti (19.6.) gondolat alapján,

$$m(t) = (t - \lambda_1) \dots (t - \lambda_s).$$

Ezért a transzformációk redukálására vonatkozó 11.2. állítás szerint

$$V = \ker(N - \lambda_1 I) \oplus \dots \oplus \ker(N - \lambda_s I).$$

Most végezzünk Gram – Schmidt-ortogonalizációt minden egyes sajátaltérben külön-külön, majd egyesítsük a kapott ortogonális rendszereket az egész V skalárisszorzatos-tér ortogonális bázisává. Így N sajátvektoraiból álló ortonormált bázist kapunk. \square

Összefoglalva a komplex test feletti esetet: *Egy transzformáció pontosan akkor normális, ha ortonormált bázisban diagonalizálható.*

Szimmetrikus transzformációk diagonalizálhatósága valós skalárisszorzatos-tereken

A valós test feletti eset fontosságát hangsúlyozandó, az önadjungált transzformációkra külön elnevezést vezetünk be.

19.8. definíció (szimmetrikus transzformáció). Egy \mathbb{R} feletti skalárisszorzatos-tér egy $A \in L(V)$ lineáris transzformációját *szimmetrikusnak* mondjuk, ha az önadjungált, azaz $A^* = A$.

Egy kicsit sem nyilvánvaló, hogy miért lenne egy szimmetrikus transzformáció sajátértéke. Rögtön kiderül, hogy minden sajátértéke.

Ha $\dim(V) = 1$, akkor minden $A : V \rightarrow V$ transzformációra igaz, hogy ha $v \neq 0$ vektor, akkor van ilyen $\lambda \in \mathbb{R}$, hogy $Av = \lambda v$. Azt kaptuk tehát, hogy egydimenziós téren értelmezett minden lineáris transzformációnak van sajátértéke, és minden zérustól különböző vektor sajátvektor is.

Most nézzük csak 2 dimenzióban. Ha A mátrixa egy ortonormált bázisban $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ akkor A karakterisztikus polinomja $k(t) = (t - a)(t - c) - b^2 = t^2 - (a + c)t + ac - b^2$. Ez polinom diszkriminánsa $(a - c)^2 + 4b^2 \geq 0$. Ha a diszkrimináns éppen zérus, akkor $b = 0, a = c$, ergo $A = aI$, így A tetszőleges bázisban diagonális. Ha a diszkrimináns pozitív, akkor két különböző sajátértéke van A -nak, amely egy 2 dimenziós vektortérben van értelmezve, ergo diagonalizálható. Meggondoltuk tehát, hogy egy 2 dimenziós téren értelmezett szimmetrikus transzformációnak van sajátvektorokból bázisa.

19.9. állítás. Legyen V egy \mathbb{R} feletti skalárisszorzatos-tér. Ekkor minden A szimmetrikus transzformáció ortonormált bázisban diagonalizálható. Ez azt jelenti, hogy létezik a térfelület A sajátvektoraiból álló ortonormált bázisa. Ebben a bázisban felírva A mátrixát egy diagonális mátrixot kapunk, amelynek diagonálisában A sajátértékei vannak.

Bizonyítás: Az algebra alaptétele szerint irreducibilis \mathbb{R} feletti polinom csak első- vagy másodfokú polinom lehet. Mivel A szimmetrikus, ezért persze normális is. Így a minimálpolinom, a transzformáció normalitását használó 19.6. állítás alapján,

$$m = p_1 \dots p_s$$

alakú, ahol p_1, \dots, p_s egymástól különböző, normált, irreducibilis, legfeljebb másodfokú polinomok. Ezért a transzformációk redukálására vonatkozó 12.4. állítás szerint a térfelület előáll mint legfeljebb két dimenziós invariáns altérei direkt összege. minden ilyen legfeljebb 2 dimenziós invariáns altérben van A sajátvektoraiból álló bázis, amelyeket egyesítve az egész V vektortér egy egy olyan bázisát kapjuk, amelynek minden eleme az A sajátvektora. Ezért valamely $r \geq s$ mellett

$$V = \ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_r I).$$

Most végezzünk Gram–Schmidt-ortogonalizációt minden egyes fenti $A - \lambda_j I$ sajátaltérben külön-külön, majd egyesítsük a kapott ortogonális rendszereket az egész V skaláriszorzatos-tér ortogonális bázisává. Így A sajátvektoraiból álló ortonormált bázist kapunk. \square

A valós számtest feletti esetet is összefoglaljuk: *Egy transzformáció pontosan akkor szimmetrikus, ha ortonormált bázisban diagonalizálható.*

Ortogonalális projekciók

A fejezetben visszatérünk a skaláriszorzatos terek geometriai tulajdonságainak vizsgálatához. Emlékezzünk arra, hogy bevezettük a merőleges vetítés transzformációt. Egy $M \subseteq V$ altér esetén $P_M \in L(V)$ az a transzformáció, amely egy $v \in V$ vektorhoz hozzárendeli a v -hez legközelebb eső M -beli vektort. Megmutattuk, hogy ilyen vektor van és csak egy van az M altérben. Ez jelöljük $P_M v$ -vel. Azt is láttuk, hogy $v - P_M v \in M^\perp$, így a tér előáll $V = M \oplus M^\perp$ ortogonalis direktösszeg alakban.

Végül is elmondhatjuk, hogy a tér minden v vektorra egyértelműen felbomlik egy M -beli és egy M^\perp vektor összegére, és ebben a felbontásban az M -beli komponens a $P_M v$, a v -nek M -re eső merőleges vetülete. Egy v vektornak pontosan akkor a zérus vektor az M -beli komponense, ha $v \in M^\perp$ teljesül, ami éppen azt jelenti, hogy $\ker P_M = M^\perp$. Az is világos, hogy $\text{Im } P_M = M$, emiatt $\ker P_M \perp \text{Im } P_M$, és $P_M^2 = P_M$.

Eddig is többször használtuk már, hogy $P|_M = \text{id}$. Emiatt ha egy $v \in V$ vektorra $v = v_1 + v_2$, ahol $v_1 \in M$ és $v_2 \in M^\perp$, akkor $P_M v = v_1$ és $P_{M^\perp} v = v_2$, ergo $P_M + P_{M^\perp} = \text{id}$.

Érdemes még meggondolni, hogy milyen szép egy ortogonalis projekció mátrixa, ha megfelelő bázist választunk. Legyen $\{x_1, \dots, x_k\}$ egy M -beli teljes ortonormált rendszer, és $\{x_{k+1}, \dots, x_n\}$ egy M^\perp beli teljes ortonormált rendszer. Világos, hogy ha ezeket egyesítjük, akkor az egész V vektortérnek ortonormált bázisát kapjuk, ahol $P_M x_j = x_j$ minden $j = 1, \dots, k$ mellett és $P_M x_j = 0$ valamennyi $j = k+1, \dots, n$ esetén. Ez azt jelenti, hogy P_M mátrixában ebben az ortonormált bázisban olyan diagonális mátrix, ahol a diagonális első k eleme 1, minden más elem zérus.¹ Ez a mátrix a transzponáltjának konjugáltja, ergo P_M egy önadzungált transzformáció.²

Most azokat a tulajdonságokat keressük, amelyek karekterizálják a merőleges vetítéseket.

20.1. definíció. Azt mondjuk, hogy a $P \in L(V)$ transzformáció egy *projekció*, ha az idempotens, tehát fennáll a $P^2 = P$ azonosság.

Vegyük észre, hogy P pontosan akkor projekció, ha $P|_{\text{Im } P} = \text{id}$.

20.2. állítás. *Egy V skaláriszorzatos-tér minden P projekciójára*

$$\ker P \oplus \text{Im } P = V.$$

Bizonyítás: Világos, hogy minden $v \in V$ vektor mellett $v = Pv + (v - Pv)$. Itt $Pv \in \text{Im } P$, és $P(v - Pv) = Pv - P^2v = Pv - Pv = 0$, ergo $v - Pv \in \ker P$.

Ha $x \in \ker P \cap \text{Im } P$, akkor $x = Pv$ valamely v vektorral, és $0 = Px = \text{id}(x) = x$.

Megmutattuk tehát, hogy $\ker P$ és $\text{Im } P$ olyan diszjunkt alterek, amelyek Minkowski-összege V . \square

20.3. definíció. A V skaláriszorzatos-tér egy $P \in L(V)$ lineáris transzformációját *ortogonalis projekciónak* mondunk, ha

1. $P^2 = P$ és

¹A gyakorlatokon azt s meggyondoltuk, hogy ha $\{e_1, \dots, e_n\}$ ortonormált bázis V -ben, és $\{x_1, \dots, x_k\}$ egy teljes ortonormált rendszer M -ben, akkor a $\sum_{j=1}^k [x_j] \cdot \overline{[x_j]^T}$ diadikus szorzatok összege adja P_M mátrixát az $\{e_1, \dots, e_n\}$ bázisban. Itt az $[x_j]$ oszlopvektor az x_j vektor koordinátavektora a fent rögzített bázisban.

²A $P_M = P_M^*$ azonosság mátrixmentes indoklása: legyen $u = u_1 + u_2$, és $v = v_1 + v_2$, ahol $u_1, v_1 \in M$, valamint $u_2, v_2 \in M^\perp$. Ekkor $\langle P_M u, v \rangle = \langle u_1, v_1 + v_2 \rangle = \langle u_1, v_1 \rangle$, és hasonlóan $\langle u, P_M v \rangle = \langle u_1 + u_2, v_1 \rangle = \langle u_1, v_1 \rangle$. Ergo $\langle P_M u, v \rangle = \langle u, P_M v \rangle$.

$$2. \ker P \perp \text{Im } P.$$

Ezek szerint az ortogonális projekció olyan speciális projekciók, amelyekre nem csak a

$$\ker P \oplus \text{Im } P = V$$

direktösszeg alakú felbontás áll fenn, hanem azontúl az

$$\ker P \oplus \text{Im } P = V, \quad \ker P \perp \text{Im } P$$

ortogonális direktösszeg alakú felbontás is igaz. Persze ekkor $(\text{Im } P)^\perp = \ker P$ is fennáll. Gondolhatunk ezért úgy is az ortogonális projekció definíciójára, hogy a két tulajdonság a P transzformációnak az $\text{Im } P$ altérén és annak ortokomplementérén való viselkedést írja elő. Tehát ha P egy ortogonális projekció, akkor

$$P|_{\text{Im } P} = \text{id}, \quad \text{és } P|_{(\text{Im } P)^\perp} = 0.$$

Láttuk, hogy tetszőleges $M \subseteq V$ altér mellett a P_M merőleges vetítés egy példa ortogonális projekcióra. A következő állítás szerint nincs más típusú ortogonális projekció, mint valamely M altérre eső merőleges vetítés.

20.4. állítás. *Legyen V egy skaláriszorzatos-tér, és $P \in L(V)$ egy ortogonális projekció. Jelölje $M = \text{Im } P$. Ekkor*

$$P = P_M.$$

Bizonyítás: Mivel P egy projekció, ezért $P|_M = \text{id}$, és mivel P egy ortogonális projekció, ezért $P|_{M^\perp} = 0$. Így minden $x \in V$ mellett az $x = P_M x + P_{M^\perp} x$ azonosságra a P transzformációt alkalmazva azt kapjuk, hogy

$$Px = P \underbrace{(P_M x)}_{\in M} + P \underbrace{(P_{M^\perp} x)}_{\in M^\perp} = P_M x + 0 = P_M x. \quad \square$$

20.5. állítás. *Legyen $A \in L(V)$ egy lineáris transzformáció a V skaláriszorzatos-tér felett. Az A pontosan akkor ortogonális projekció, ha $A^* = A$ és $A^2 = A$.*

Bizonyítás: Láttuk korábban hogy egy P_M merőleges vetítés önadjugált is, idempotens is. No de, minden ortogonális projekció P_M alakú, valamely M altér megválasztása mellett, ezért A -ra is igaz, hogy önadjugált és idempotens.

Megfordítva, ha A önadjugált, akkor a skaláriszorzatos rang tétel szerint $(\ker A)^\perp = \text{Im } A^* = \text{Im } A$, amiből persze $\ker A \perp \text{Im } A$ már könnyen következik. \square

Mivel minden P ortogonális projekció $P = P_M$ alakú, ezért a Bessel-egyenlőség ortogonális projekciókra is igaz:

20.6. állítás (Bessel). *Minden $P \in L(V)$ ortogonális projekcióra és minden $v \in V$ vektorra $\|Pv\| \leq \|v\|$, és $\|Pv\|^2 = \langle v, Pv \rangle$.*

Bizonyítás: Jelölje $M = \text{Im } P$. Láttuk, hogy $P = P_M$ ezért $\|Pv\| = \|P_M v\| \leq \|v\|$. Mivel P önadjugált és idempotens, ezért $\|Pv\|^2 = \langle Pv, Pv \rangle = \langle v, P^* Pv \rangle = \langle v, P^2 v \rangle = \langle v, Pv \rangle$. \square

20.7. állítás. *Legyen $P, Q \in L(V)$ ortogonális projekció. A $QP = 0$ pontosan akkor teljesül, ha $\text{Im } P \perp \text{Im } Q$.*

Bizonyítás: Legyen $u \in \text{Im } P$ és $v \in \text{Im } Q$. Ekkor

$$\langle u, v \rangle = \langle Pu, Qv \rangle = \langle Q^* Pu, v \rangle = \langle QPu, v \rangle = \langle 0, v \rangle = 0.$$

Evvel igazoltuk, hogy $QP = 0$ esetén $\text{Im } P \perp \text{Im } Q$.

Megfordítva, tetszőleges $u \in V$ mellett $Pu \in \text{Im } P$ és $QPu \in \text{Im } Q$. Így

$$\|QPu\|^2 = \langle QPu, QPu \rangle = \langle Q^* QPu, Pu \rangle = \langle Q^2 Pu, Pu \rangle = \langle QPu, Pu \rangle = 0. \quad \square$$

20.8. definíció (ortogonális projekciók merőlegessége). Az Q és P ortogonális projekciókat egymásra merőlegesnek mondjuk, ha $QP = 0$ fennáll. Jelölés: $Q \perp P$.

Világos tehát, hogy $QP = 0$, $\text{Im } P \perp \text{Im } Q$ és $PQ = 0$ ekvivalens megfogalmazásai $Q \perp P$ -nak, amennyiben $P, Q \in L(V)$ ortogonális projekciók. Az $M_1, M_2 \subseteq V$ alerek esetén $P_{M_1} \perp P_{M_2}$ pontosan akkor, ha $M_1 \perp M_2$.

20.9. állítás. Legyenek P_1, \dots, P_r ortogonális projekciók. Jelölje $P = \sum_{j=1}^r P_j$ ezek összegét. A P összeg pontosan akkor ortogonális projekció, ha $P_i \perp P_j$ minden $i \neq j$ mellett.

Bizonyítás: Tegyük fel először, hogy $P_i \perp P_j$ minden $i \neq j$ mellett és legyen P az összeg. Ekkor

$$P^* = \sum_{j=1}^r P_j^* = \sum_{j=1}^r P_j = P, \quad \text{és} \quad P^2 = \sum_{k=1}^r \sum_{j=1}^r P_k P_j = \sum_{k=1}^r P_k^2 = \sum_{k=1}^r P_k = P.$$

Azt kaptuk tehát, hogy a P összeg önadjungált és idempotens, ergo egy ortogonális projekció.

Megfordítva, most azt tegyük fel, hogy P ortogonális projekció. Megmutatjuk, hogy tetszőleges $j \neq i$ -re $P_i P_j = 0$. Legyen ezért $u \in V$ rögzítve, és jelölje valamely rögzített j index mellett $x_j = P_j u$. Ekkor

$$\|x_j\|^2 \geq \|Px_j\|^2 = \langle x_j, Px_j \rangle = \langle x_j, \sum_{i=1}^r P_i x_j \rangle = \sum_{i=1}^r \langle x_j, P_i x_j \rangle = \sum_{i=1}^r \|P_i x_j\|^2 \geq \|P_j x_j\|^2 = \|x_j\|^2.$$

Ezek szerint az utolsó egyenlőtlenség is egyenlőség, ergo minden $i \neq j$ esetén $\|P_i x_j\| = 0$, ergo $P_i x_j = 0$, ergo $P_i (P_j u) = 0$. \square

20.1. Ortogonális projekciók lineáris kombinációja

Olyan nem zérus ortogonális projekciók lineáris kombinációt vizsgáljuk, amelyek összege az identikus transzformáció. Mivel az identikus transzformáció egyben projekció is, ezért az előző tételel használva azt kapjuk, hogy az összeadandó projekciók még egymásra merőlegesek is.

20.10. állítás (komplex együtthatók). Legyen V egy \mathbb{K} feletti skaláriszorzatos-tér. Tegyük fel, hogy $P_1, \dots, P_r \in L(V)$ olyan ortogonális projekciók, amelyekre

1. $P_j \neq 0$ minden $j = 1, \dots, r$ mellett.

2. $I = \sum_{j=1}^r P_j$.

Legyenek a $\{\lambda_1, \dots, \lambda_r\} \subseteq \mathbb{K}$ (nem feltétlenül különböző) számok tetszőlegesen rögzítve, és jelölje $N = \sum_{j=1}^r \lambda_j P_j$. Ekkor az N egy normális transzformáció, amelynek spektrumára $\sigma(N) = \{\lambda_1, \dots, \lambda_r\}$.

Bizonyítás: A feltétel szerint az ortogonális projekciók összege az identikus transzformáció, tehát az összeg is egy ortogonális projekció. Emiatt a P_j projekciók egymásra páronként merőlegesek, így

$$N^* N = \sum_{j=1}^r |\lambda_j|^2 P_j = N N^*.$$

Most megmutatjuk, hogy $\text{Im } P_j \subseteq \ker(N - \lambda_j I)$. Legyen $u_j \in \text{Im } P_j$. Világos, hogy minden $k \neq j$ mellett $P_k(u_j) = P_k(P_j(u_j)) = P_k P_j(u_j) = 0$. Így

$$N u_j = \left(\sum_{k=1}^r \lambda_k P_k \right) u_j = \lambda_j P_j u_j = \lambda_j u_j.$$

A $P_j \neq 0$ feltétel szerint $\ker(N - \lambda_j I) \supseteq \text{Im } P_j \neq \{0\}$, ezért $N - \lambda_j I$ szinguláris. Mivel ez minden $j = 1, \dots, r$ mellett fennáll, ezért minden j mellett a λ_j egy sajátértéke N -nek.

Most azt gondoljuk meg, hogy nincs más sajátérték, azaz ha $\sigma \neq \lambda_j$ minden $j = 1, \dots, r$ mellett, akkor σ nem sajátérték. Megmutatjuk tehát, hogy $\ker(N - \sigma I) = \{0\}$. Ehhez vegyük észre, hogy $\sum_{j=1}^r P_j = I$ szerint

$\sum_{j=1}^r \text{Im } P_j = V$ komplexus azonosság teljesül, és mivel N normális, ezért $\ker(N - \lambda_j I) \perp \ker(N - \sigma I)$ is teljesül minden $j = 1, \dots, r$ mellett. Így

$$V = \sum_{j=1}^r \text{Im } P_j \subseteq \sum_{j=1}^r \ker(N - \lambda_j I) \subseteq (\ker(N - \sigma I))^{\perp}.$$

Ha tehát $v \in \ker(N - \sigma I)$, akkor v merőleges saját magára, ergo $v = 0$. Ezt kellett belátni. \square

20.11. állítás. *Tegyük fel, hogy olyan P_1, \dots, P_r ortogonális projekcióra, amelyek összege az identikus transzformáció, és a tetszőlegesen megválasztott $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ számokra $N = \sum_{j=1}^r \lambda_j P_j$. Ekkor minden $q \in \mathbb{K}[t]$ polinom mellett $q(N) = \sum_{j=1}^r q(\lambda_j) P_j$.*

Bizonyítás: Először gondoljuk meg azt, hogy minden $l \geq 0$ mellett $N^l = \sum_{j=1}^r \lambda_j^l P_j$. Ha $l = 0$, akkor az ortogonális projekciók összege az identitás feltevést kapjuk. Ha az állítás igaz l -ig, akkor $l + 1$ -re:

$$N^{l+1} = N^l N = \left(\sum_{j=1}^r \lambda_j^l P_j \right) \left(\sum_{h=1}^r \lambda_h P_h \right) = \sum_{j=1}^r \sum_{h=1}^r \lambda_j^l \lambda_h P_j P_h = \sum_{j=1}^r \lambda_j^{l+1} P_j.$$

Itt azt használtuk ki, hogy amennyiben az ortogonális projekciók összege is egy ortogonális projekció, akkor $P_j P_h = \delta_{j,h} P_j^2 = \delta_{j,h} P_j$. No de, ha a q polinom $q(t) = \sum_{l=0}^m \alpha_l t^l$ alakú, akkor

$$q(N) = \sum_{l=0}^m \alpha_l N^l = \sum_{l=0}^m \alpha_l \left(\sum_{j=1}^r \lambda_j^l P_j \right) = \sum_{l=0}^m \sum_{j=1}^r \alpha_l \lambda_j^l P_j = \sum_{j=1}^r \sum_{l=0}^m \alpha_l \lambda_j^l P_j = \sum_{j=1}^r \left(\sum_{l=0}^m \alpha_l \lambda_j^l \right) P_j = \sum_{j=1}^r q(\lambda_j) P_j.$$

Ezt kellett belátni. \square

Valós együtthatós lineáris kombináció

Most nézzük a valós együtthatós lineáris kombinációk esetét.

20.12. állítás. *Legyen V egy \mathbb{K} feletti skalárisszorzatos-tér. Tegyük fel, hogy $P_1, \dots, P_r \in L(V)$ olyan ortogonális projekciók, amelyekre*

1. $P_j \neq 0$ minden $j = 1, \dots, r$ mellett.
2. $I = \sum_{j=1}^r P_j$.

Legyenek a $\{\lambda_1, \dots, \lambda_r\} \subseteq \mathbb{R}$ (nem feltétlenül különböző) valós számok tetszőlegesen rögzítve, és jelölje $N = \sum_{j=1}^r \lambda_j P_j$. Ekkor az N egy önadjungált transzformáció, amelynek spektrumára $\sigma(N) = \{\lambda_1, \dots, \lambda_r\}$.

Bizonyítás: Mivel az itt szereplő λ_j számok valósak, ezért azonosak a konjugáltjukkal. Így

$$N^* = \sum_{j=1}^r \overline{\lambda_j} P_j^* = \sum_{j=1}^r \lambda_j P_j = N.$$

\square

Egy abszolútértékű együtthatókkal képzett lineáris kombináció

Most nézzük az olyan lineáris kombinációk esetét, amikor az együtthatók 1 abszolútértékű számok.

20.13. állítás. *Legyen V egy \mathbb{K} feletti skalárisszorzatos-tér. Tegyük fel, hogy $P_1, \dots, P_r \in L(V)$ olyan ortogonális projekciók, amelyekre*

1. $P_j \neq 0$ minden $j = 1, \dots, r$ mellett.
2. $I = \sum_{j=1}^r P_j$.

Legyenek a $\{\lambda_1, \dots, \lambda_r\} \subseteq \mathbb{K}$ (nem feltétlenül különböző), 1 abszolútértekű számok tetszőlegesen rögzítve, és jelölje $N = \sum_{j=1}^r \lambda_j P_j$. Ekkor az N egy unitér transzformáció, amelynek spektrumára $\sigma(N) = \{\lambda_1, \dots, \lambda_r\}$.

Bizonyítás: N spektruma a zérust nem tartalmazza, tehát N reguláris. Persze

$$NN^* = \sum_{j=1}^n |\lambda_j|^2 P_j = \sum_{j=1}^n P_j = I$$

ami azt jelenti, hogy $N^{-1} = N^*$. □

21. fejezet

Spektrális felbontások

AZ IDENTITKUS TRANSZFORMÁCIÓ felbontása a közös magja a spektrális felbontásoknak:

21.1. állítás. *Tegyük fel, hogy V egy \mathbb{K} feletti skaláriszorzatos-tér, és tegyük fel, hogy az M_1, \dots, M_r alterekre*

$$V = M_1 + \dots + M_r, \quad \text{ahol } M_j \perp M_k \text{ minden } j \neq k.$$

Jelölje P_j az M_j altérre való merőleges vetítést. Ekkor

$$I = \sum_{j=1}^r P_j$$

és persze $P_j \perp P_k$ minden $j \neq k$ mellett.

Bizonyítás: minden $s \in V$ vektor előáll $s = \sum_{j=1}^r s_j$ alakban, ahol $s_j \in M_j$. Felhasználva, hogy $P_j|_{M_j} = \text{id}_{M_j}$ és $\sum_{\substack{k=1 \\ k \neq j}}^r M_k \subseteq M_j^\perp = \ker P_j$ minden $k \neq j$ mellett

$$\left(\sum_{j=1}^r P_j \right) s = \sum_{j=1}^r P_j s = \sum_{j=1}^r P_j \left(s_j + \sum_{\substack{k=1 \\ k \neq j}}^r s_k \right) = \sum_{j=1}^r \left(P_j s_j + P_j \left(\sum_{\substack{k=1 \\ k \neq j}}^r s_k \right) \right) = \sum_{j=1}^r P_j s_j = \sum_{j=1}^r s_j = Is. \quad \square$$

Emlékezzünk arra, hogy ha $\{\lambda_1, \dots, \lambda_r\}$ az N normális transzformáció különböző sajátértékei, akkor bevezetve az $M_i = \ker(N - \lambda_i I)$ jelölést

$$M_1 \oplus \dots \oplus M_r \subseteq V \quad \text{ahol } M_j \perp M_k \text{ minden } j \neq k.$$

Még azt is láttuk, hogy $\mathbb{K} = \mathbb{C}$ esetben itt egyenlőség is van, de azt is tudjuk, hogy az egyenlőség még a $\mathbb{K} = \mathbb{R}$ és $N^* = N$ esetben is fennáll.

21.1. Komplex eset

A spektrális felbontási téteteleket fogalmazzuk meg abban az esetben, mikor az alaptestre $\mathbb{K} = \mathbb{C}$. A legáltalánosabb eset a normális transzformációkra vonatkozik, majd felírjuk az önadjungált és az unitér transzformációk speciális esetét.

Normális transzformáció

21.2. állítás (spektrális felbontás). *Legyen V egy \mathbb{C} feletti skaláriszorzatos-tér, és $N \in L(V)$ egy normális lineáris transzformáció. Jelölje $\sigma(N) = \{\lambda_1, \dots, \lambda_k\}$ az N különböző sajátértékeit. Láttuk, hogy N diagonalizálható, ezért $k \geq 1$. Jelölje P_j a $\ker\{N - \lambda_j I\}$ altérre való merőleges vetítést. Ekkor*

1. *Minden $j = 1, \dots, r$ mellett $P_j \neq 0$;*

2. $I = \sum_{j=1}^r P_j$;
3. $N = \sum_{j=1}^r \lambda_j P_j$;
4. $N^* = \sum_{j=1}^r \bar{\lambda}_j P_j$.

Bizonyítás: Láttuk, hogy N diagonalizálható, azaz létezik a térféle N sajátvektoraiból álló bázisa, ezért

$$M_1 \oplus \cdots \oplus M_r = V \quad \text{itt } M_j \perp M_k \text{ minden } j \neq k, \quad (\dagger)$$

ahol $M_j = \ker(N - \lambda_j I)$.

Világos, hogy minden j mellett létezik $u_j \neq 0$, hogy $Nu_j = \lambda_j u_j$, így $M_j \neq \{0\}$, ergo $P_j \neq 0$.

Láttuk korábban, hogy (\dagger) -ből $I = \sum_{j=1}^r P_j$ is következik.

A 3. azonossághoz azt vegyük észre, hogy tetszőleges $v \in V$ mellett $P_j v \in M_j$ a λ_j -hez tartozó sajátaltér egy eleme, így $NP_j v = \lambda_j P_j v$. Tehát

$$Nv = N(Iv) = N\left(\sum_{j=1}^r P_j v\right) = \sum_{j=1}^r NP_j v = \sum_{j=1}^r \lambda_j P_j v = \left(\sum_{j=1}^r \lambda_j P_j\right) v.$$

A 4. azonosság a 3. tulajdonságból az adjungálás tulajdonságai és $P^* = P$ figyelembe vételével adódik. \square

Az eddigiek összefoglalásaként (20.10 és 20.11) kapjuk az alábbi szükséges és elegendő feltételét annak, hogy egy komplex mátrix ortonormált bázisban diagonalizálható.

21.3. állítás (Normalitás karakterizációja a spektrális felbontással). *Legyen V egy \mathbb{C} feletti skaláriszorzatos-tér és $N \in L(V)$ egy lineáris transzformáció. Az N transzformáció pontosan akkor normális, ha létezik $r \geq 1$ egész, léteznek P_1, \dots, P_r ortonormális projekciók és léteznek $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ komplex számok, amelyekre*

1. $P_j \neq 0$, minden $j = 1, \dots, r$;
2. $I = \sum_{j=1}^r P_j$;
3. $N = \sum_{j=1}^r \lambda_j P_j$.

Ha N normális, akkor $\sigma(N) = \{\lambda_1, \dots, \lambda_r\}$, $N^ = \sum_{j=1}^r \bar{\lambda}_j P_j$, és tetszőleges $q \in \mathbb{C}[t]$ komplex együtthatós polinom mellett $q(N)$ is normális, továbbá $q(N) = \sum_{j=1}^r q(\lambda_j) P_j$.*

Önadzungált transzformáció

Először azt kell látnunk, hogy komplex vektortér esetében a normális transzformációk közül éppen azok az önadzungáltak, amelyeknek a – nem üres – spektruma valós.

21.4. állítás. *Legyen V egy a komplex számtest feletti skaláriszorzatos-tér. $A \in L(V)$ egy lineáris transzformáció. Az A pontosan akkor önadzungált, ha A normális és $\sigma(A) \subseteq \mathbb{R}$*

Bizonyítás: Az, hogy önadzungált transzformáció normális és a (lehetséges, hogy üreshalmaz) spektruma valós, még \mathbb{K} feletti skaláriszorzatos-tér mellett is igaz. (19.5). Megfordítva, ha egy komplex vektortéren A normális, akkor van spektrális felbontása:

$$A = \sum_{j=1}^r \lambda_j P_j,$$

ahol $\{\lambda_1, \dots, \lambda_r\}$ a spektrum elemei. Mivel ezek valós számok, ezért

$$A^* = \sum_{j=1}^r \bar{\lambda}_j P_j = \sum_{j=1}^r \lambda_j P_j = A.$$

\square

Az önadzungált transzformációk spektrális felbontása a normális transzformációk spektrális felbontásának speciális esete.

21.5. állítás (spektrális felbontás). *Legyen V egy \mathbb{C} feletti skalárisszorzatos-tér, és $A \in L(V)$ egy önadjungált lineáris transzformáció. Jelölje $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$ az A különböző sajátértékeit. Láttuk, hogy $k \geq 1$, és $\sigma(A) \subseteq \mathbb{R}$. Jelölje P_j a $\ker\{A - \lambda_j I\}$ altérre való merőleges vetítést. Ekkor*

1. *Minden $j = 1, \dots, r$ mellett $P_j \neq 0$;*
2. *$I = \sum_{j=1}^r P_j$;*
3. *$A = \sum_{j=1}^r \lambda_j P_j$.*

Összefoglalásként (20.12 és 20.11) kapjuk az alábbi szükséges és elegendő feltételét annak, hogy egy komplex mátrix ortonormált bázisban diagonalizálható, és a diagonális elemei valós számok.

21.6. állítás (Önadjungáltság karakterizációja a spektrális felbontással). *Legyen V egy \mathbb{C} feletti skalárisszorzatos-tér és $A \in L(V)$ egy lineáris transzformáció. Az A transzformáció pontosan akkor önadjungált, ha létezik $r \geq 1$ egész, léteznek P_1, \dots, P_r ortogonális projekciók és léteznek $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ valós számok, amelyekre*

1. *$P_j \neq 0$, minden $j = 1, \dots, r$;*
2. *$I = \sum_{j=1}^r P_j$;*
3. *$A = \sum_{j=1}^r \lambda_j P_j$.*

Ha A önadjungált, akkor $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$, és tetszőleges $q \in \mathbb{R}[t]$ valós együtthatós polinom mellett $q(A)$ is önadjungált, továbbá $q(A) = \sum_{j=1}^r q(\lambda_j) P_j$.

Unitér transzformáció

Először azt kell látnunk, hogy komplex vektortér esetében a normális transzformációk közül éppen azok az unitér transzformációk, amelyekre a – nem üres – spektruma minden pontja egységnyi abszolútértékű.

21.7. állítás. *Legyen V egy a komplex számtest feletti skalárisszorzatos-tér. $A \in L(V)$ egy lineáris transzformáció. Az A pontosan akkor unitér, ha A normális és minden $\lambda \in \sigma(A)$ sajátértékre $|\lambda| = 1$.*

Bizonyítás: Még \mathbb{K} feletti skalárisszorzatos-tér mellett is igaz, hogy egy unitér transzformáció normális, és a (lehetséges, hogy üreshalmaz) spektruma minden pontja egységnyi abszolútértékű: Ugyanis a λ spektrumpontra $\lambda \neq 0$, és valamely $u \neq 0$ vektor mellett $Au = \lambda u$, $A^*u = \bar{\lambda}u$. Így

$$\left(\frac{1}{\lambda} - \bar{\lambda}\right)u = \frac{1}{\lambda}u - \bar{\lambda}u = A^{-1}u - A^*u = 0.$$

Ebből már $u \neq 0$ miatt $|\lambda| = 1$ következik.

Megfordítva, mivel komplex tér felett vagyunk, ezért A -nak normális volta miatt van $A = \sum_{j=1}^r \lambda_j P_j$ spektrális felbontása, ahol $\lambda_1, \dots, \lambda_r$ az A sajátértékei. Tehát A ortogonális projekciók egységnyi abszolútértékű együtthatókkal képzett lineáris kombinációja, amiről korábban már láttuk, hogy minden unitér transzformációt eredményez. \square

Az unitér transzformációk spektrális felbontása is a normális transzformációk spektrális felbontásának speciális esete.

21.8. állítás (spektrális felbontás). *Legyen V egy \mathbb{C} feletti skalárisszorzatos-tér, és $A \in L(V)$ egy unitér lineáris transzformáció. Jelölje $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$ az A különböző sajátértékeit. Láttuk, hogy $k \geq 1$, $|\lambda_j| = 1$ minden $j = 1, \dots, r$ mellett. Jelölje P_j a $\ker\{A - \lambda_j I\}$ altérre való merőleges vetítést. Ekkor*

1. *Minden $j = 1, \dots, r$ mellett $P_j \neq 0$;*
2. *$I = \sum_{j=1}^r P_j$;*
3. *$A = \sum_{j=1}^r \lambda_j P_j$.*

$$4. A^* = \sum_{j=1}^r \frac{1}{\lambda_j} P_j.$$

Összefoglalásként (20.13 és 20.11) kapjuk az alábbi szükséges és elegendő feltételét annak, hogy egy komplex mátrix ortonormált bázisban diagonalizálható, és a diagonalis elemei egy abszolútértekű komplex számok.

21.9. állítás (Unitér karakterizáció a spektrális felbontással). *Legyen V egy \mathbb{C} feletti skalárisszorzatos-tér és $A \in L(V)$ egy lineáris transzformáció. Az A transzformáció pontosan akkor unitér, ha létezik $r \geq 1$ egész, léteznek P_1, \dots, P_r ortonormális projekciók és léteznek $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ egységnnyi abszolútértekű komplex számok, amelyekre*

$$1. P_j \neq 0, \text{ minden } j = 1, \dots, r;$$

$$2. I = \sum_{j=1}^r P_j;$$

$$3. A = \sum_{j=1}^r \lambda_j P_j.$$

Ha A unitér, akkor $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$, és $A^{-1} = \sum_{j=1}^r \frac{1}{\lambda_j} P_j = A^$ is fennáll.*

21.2. Valós eset

Mivel valós test feletti skalárisszorzatos-tér felett még az is előfordulhat, hogy egy normális transzformációnak még sajátértéke sincs, ezért nem várható általános eredmény normális transzformációk diagonalizálhatóságára.

Szimmetrikus transzformáció

Láttuk, hogy szimmetrikus transzformációk diagonalizálhatók, ezért a komplex önjellegű esettel analóg állítás igaz valós vektortér esetére.

21.10. állítás (spektrális felbontás). *Legyen V egy \mathbb{R} feletti skalárisszorzatos-tér, és $A \in L(V)$ egy szimmetrikus lineáris transzformáció. Jelölje $\sigma(A) = \{\lambda_1, \dots, \lambda_k\} \subseteq \mathbb{R}$ az A különöző sajátértékeit. Láttuk, hogy $k \geq 1$. Jelölje P_j a $\ker\{A - \lambda_j I\}$ altérre való merőleges vetítést. Ekkor*

$$1. \text{ minden } j = 1, \dots, r \text{ mellett } P_j \neq 0;$$

$$2. I = \sum_{j=1}^r P_j;$$

$$3. A = \sum_{j=1}^r \lambda_j P_j.$$

Bizonyítás: Láttuk, hogy A diagonalizálható¹, azaz létezik a térnak A sajátvektoraiból álló bázisa, ezért

$$M_1 \oplus \dots \oplus M_r = V \quad \text{itt } M_j \perp M_k \text{ minden } j \neq k, \quad (\dagger)$$

ahol $M_j = \ker(N - \lambda_j I)$.

Világos, hogy minden j mellett létezik $u_j \neq 0$, hogy $Au_j = \lambda_j u_j$, így $M_j \neq \{0\}$, ergo $P_j \neq 0$.

Láttuk korábban, hogy (\dagger) -ből $I = \sum_{j=1}^r P_j$ is következik.

A 3. azonossághoz azt vegyük észre, hogy tetszőleges $v \in V$ mellett $P_j v \in M_j$ a λ_j -hez tartozó sajátaltér egy eleme, így $AP_j v = \lambda_j P_j v$. Tehát

$$Av = A(Iv) = A\left(\sum_{j=1}^r P_j v\right) = \sum_{j=1}^r AP_j v = \sum_{j=1}^r \lambda_j P_j v = \left(\sum_{j=1}^r \lambda_j P_j\right) v. \quad \square$$

Összefoglalásként (20.10 és 20.11) kapjuk az alábbi szükséges és elegendő feltételét annak, hogy egy valós szimmetrikus mátrix ortonormált bázisban diagonalizálható.

¹Innen a bizonyítás szó szerint azonos a komplex tér feletti normális transzformáció spektrálfebontásának igazolásával.

21.11. állítás (Szimmetrikus karakterizáció spektrális felbontással). Legyen V egy valós skalárisszorzatos-tér és $A \in L(V)$ egy lineáris transzformáció. Az A transzformáció pontosan akkor szimmetrikus, ha létezik $r \geq 1$ egész, léteznek P_1, \dots, P_r ortogonális projekciók és léteznek $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ valós számok, amelyekre

1. $P_j \neq 0$, minden $j = 1, \dots, r$;
2. $I = \sum_{j=1}^r P_j$;
3. $A = \sum_{j=1}^r \lambda_j P_j$.

Ha A szimmetrikus, akkor $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$, és tetszőleges $q \in \mathbb{R}[t]$ valós együtthatós polinom mellett $q(A)$ is szimmetrikus, továbbá $q(A) = \sum_{j=1}^r q(\lambda_j) P_j$.

21.3. Operátornorma

A spektrális felbontási téTEL egy fontos következménye, hogy egy normális transzformációt egyértelműen meghatározza a spektruma, és a sajátaltereire való ortogonális projekciók. Ez lehetőséget ad egy tetszőleges A lineáris operáció normájának meghatározására az A^*A spektrumának ismeretében.

Ebben a szakaszban feltesszük, hogy az operátornorma definíciója és alaptulajdonságai ismertek. Összefoglalom az általunk használt legfontosabb tényeket. Ha $A \in L(V, W)$ véges dimenziós normált téren értelmezett és a W normált térbé képező lineáris operáció, akkor

$$\max \{\|Ax\| : \|x\| \leq 1\} = \min \{K \geq 0 : \|Ax\| \leq K\|x\| \quad \forall x \in V\}$$

számok azonosak és végesek. A közös érték az $\|A\|$ operátornorma. Ha az $L(V, W)$ lineáris operációk vektortér ellátjuk az operátornormával, akkor egy normált teret kapunk. A $W = V$ speciális esetben $A, B \in L(V)$ -re $\|AB\| \leq \|A\|\|B\|$ egyenlőtlenség is fennáll.

A továbbiakban V egy véges dimenziós \mathbb{K} feletti skalárisszorzatos-tér.

21.12. definíció. Legyen $A \in L(V)$ egy lineáris transzformáció, amelynek spektruma nem üres. Ekkor az $r(A) = \max \{|\lambda| : \lambda \in \sigma(A)\}$ számot az A transzformáció spektrálisugarának mondjuk.

Világos, hogy $r(A)$ a legkisebb olyan sugár, amellyel a komplex síkra rajzolt origó középpontú kör tartalmazza az A valamennyi sajátértékét. Ezzel ekvivalens azt mondani, hogy a spektrálisugár a sajátértékek origótól vett távolságainak maximuma.

Ha például $\sigma(A) \subseteq \mathbb{R}$, azaz valós sajátértékek vannak², akkor a $\lambda_{\max} = \max \sigma(A)$ és $\lambda_{\min} = \min \sigma(A)$ jelölések értelmesek, és $r(A) = \max \{|\lambda_{\min}|, |\lambda_{\max}|\}$. Ha még azt is feltesszük, hogy minden sajátérték nem negatív, akkor $r(A) = \lambda_{\max}$.

21.13. állítás. Legyenek V, W ugyanazon \mathbb{K} test feletti skalárisszorzatos-terek. Legyen $A \in L(V, W)$ egy lineáris operáció. Ekkor az A^*A egy önadjungált transzformáció, így $\sigma(A^*A) \neq \emptyset$, A^*A valamennyi sajátértéke nem negatív valós szám, ezért $r(A^*A) = \max \sigma(A^*A)$.

Bizonyítás: Világos, hogy $(A^*A)^* = A^*(A^*)^* = A^*A$, ergo A^*A valóban egy önadjungált lineáris transzformációja a V vektortérnek. Láttuk, hogy A^*A még diagonalizálható is, tehát spektruma nem üres. Azt is láttuk, hogy önadjungált transzformáció minden sajátértéke valós. No de, ha $\lambda \in \sigma(A^*A)$ és $x \in V$ egy nem zérus vektor, amelyre $A^*Ax = \lambda x$, akkor

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle A^*Ax, x \rangle = \langle Ax, Ax \rangle = \|Ax\|^2.$$

Ebből persze $\lambda \geq 0$ azonnal következik. □

21.14. állítás. Legyen mint az előző állításban $A \in L(V, W)$. Definálja $Q(v) = \langle A^*Av, v \rangle$ a $Q : V \rightarrow [0, \infty)$ nem negatív, valós függvényt. Legyen $\lambda_{\max} = \max \sigma(A^*A)$, továbbá legyen $v_{\max} \in V$ az A^*A transzformációnak a λ_{\max} sajátértékéhez tartozó olyan sajátvektora, amelyre $\|v_{\max}\| = 1$. Ekkor

$$\max \{Q(v) : v \in V, \|v\| \leq 1\} = \lambda_{\max} = Q(v_{\max}).$$

²Például ha A önadjungált, akkor így van.

Bizonyítás: Tudjuk, hogy A^*A önadjungált, ezért minden \mathbb{R} , minden \mathbb{C} felett diagonalizálható. Létezik tehát u_1, \dots, u_n ortonormált bázis V -ben, amelynek minden vektorra sajátvektora A^*A -nak. Ha $v = \sum_{j=1}^n \alpha_j u_j$ olyan vektor, amelyre $\|v\|^2 = \sum_{j=1}^n |\alpha_j|^2 = 1$, akkor

$$\begin{aligned} Q(v) &= \langle A^*A \left(\sum_{j=1}^n \alpha_j u_j \right), \left(\sum_{k=1}^n \alpha_k u_k \right) \rangle = \sum_{j=1}^n \sum_{k=1}^n \alpha_j \bar{\alpha}_k \langle A^*A u_j, u_k \rangle \\ &= \sum_{j=1}^n \sum_{k=1}^n \alpha_j \bar{\alpha}_k \lambda_j \langle u_j, u_k \rangle = \sum_{j=1}^n \alpha_j \bar{\alpha}_j \lambda_j = \sum_{j=1}^n |\alpha_j|^2 \lambda_j \leq \lambda_{\max} \sum_{j=1}^n |\alpha_j|^2 = \lambda_{\max}. \end{aligned}$$

Itt az u_i sajátvektorok tartoznak a λ_i sajátértékekhez. Így $\sup \{Q(v) : v \in V, \|v\| \leq 1\} \leq \lambda_{\max}$.

No de, $Q(v_{\max}) = \langle A^*A v_{\max}, v_{\max} \rangle = \lambda_{\max} \langle v_{\max}, v_{\max} \rangle = \lambda_{\max}$, amiből már az is látszik, hogy $\lambda_{\max} = Q(v_{\max}) \leq \sup \{Q(v) : v \in V, \|v\| \leq 1\}$. \square

Összefoglalásként kapjuk az operátornorma és a spektrálsugár közti kapcsolatot:

21.15. állítás. Legyen $A \in L(V, W)$ egy lineáris operáció. Ekkor

$$\|A\|^2 = r(A^*A) = \max \sigma(A^*A).$$

Bizonyítás: Az $x \mapsto x^2$ függvény a számegyenes nem negatív része felett szigorúan monoton növő, ezért

$$\|A\|^2 = (\sup \{\|Av\| : \|v\| \leq 1\})^2 = \sup \{\|Av\|^2 : \|v\| \leq 1\} = \sup \{Q(v) : \|v\| \leq 1\} = \max \sigma(A^*A). \quad \square$$

Ha ezt egy önadjungált transzformációra írjuk fel, akkor még egyszerűbb állítást kapunk.

21.16. állítás. Legyen $A \in L(V)$ egy önadjungált lineáris transzformáció a V véges dimenziós, \mathbb{K} feletti skaláriszszor-zatos-téren. Ekkor

$$\|A\| = r(A) = \max \{|\min \sigma(A)|, |\max \sigma(A)|\}.$$

Bizonyítás: Mivel A önadjungált, ezért A -nak van spektrál felbontása, és ha ez $A = \sum_{j=1}^r \lambda_j P_j$, akkor az A^2 spektrál felbontása $A^2 = \sum_{j=1}^r \lambda_j^2 P_j$, ahol a $\lambda_1, \dots, \lambda_r$ valós számok az A spektrumának elemei. Ez azt jelenti, hogy A^2 minden sajátértéke négyzete A valamelyik sajátértékének. Így

$$r(A^2) = \max \{\lambda_j^2 : j = 1, \dots, r\} = (\max \{|\lambda_{\min}|, |\lambda_{\max}|\})^2 = r(A)^2.$$

No de, újra használva az $A^* = A$ feltevést

$$\|A\|^2 = r(A^*A) = r(A^2) = r(A)^2. \quad \square$$

21.17. állítás. Tetszőleges $A \in L(V, W)$ lineáris operáció mellett

1. Teljesül az úgynévezett C^* -azonosság:

$$\|A\|^2 = \|A^*A\|.$$

2. A transzformációnak és az adjungáltjának azonos a normája, azaz

$$\|A\| = \|A^*\|.$$

3. Az $A^*A \in L(V)$ és az $AA^* \in L(W)$ transzformációk spektrálsugara azonos, azaz

$$r(A^*A) = r(AA^*).$$

Bizonyítás: Mivel A^*A önadjungált, ezért

$$\|A^*A\| = r(A^*A) = \|A\|^2,$$

ami éppen a C^* -azonosság.

Felhasználva, hogy transzformációk szorzatának normája legfeljebb a normák szorzata, a C^* -azonosság szerint $\|A\|^2 \leq \|A^*\| \cdot \|A\|$, amiből $\|A\| \leq \|A^*\|$ következik. No de, ez minden lineáris operációra igaz, speciálisan az A^* adjungáltra is. Emiatt az $\|A^*\| \leq \|(A^*)^*\| = \|A\|$ egyenlőtlenség is teljesül igazolva, hogy egy lineáris operációknak és adjungáltjának azonos a normája.

Alkalmazva az eddig igazoltatakat

$$r(A^*A) = \|A\|^2 = \|A^*\|^2 = r(AA^*).$$

□

Fontos megérteni az utolsó azonosság jelentőségét. Az A^*A és az AA^* más-más vektortereken vannak értelmezve. Képzeljük el, hogy V egy nyolc dimenziós tér és W egy két dimenziós tér, $A \in L(V, W)$. Ekkor A mátrixa egy 2×8 méretű mátrix, ezért A^*A mátrixa egy 8×8 méretű mátrix. Hasonlóan AA^* egy 2×2 méretű mátrix. Ha az A operátor normáját kell meghatározunk, akkor nagyon nem mindegy, hogy egy 8×8 méretű-, vagy egy 2×2 méretű önadjungált mátrix legnagyobb sajátértékét kell megtalálnunk.

Legvégül a C^* -azonosság és következményei egy sorban

$$\|A\|^2 = \|A^*A\| = r(A^*A) = r(AA^*) = \|AA^*\| = \|A^*\|^2.$$

Függelékek

A. függelék

A komplex számokról

Az algebra alaptétele, és a komplex számtest egyértelműsége. Elsősorban (Ebbinghaus és tsai. 1991) és (Derkson 2003) alapján

A.1. Lineáris algebrai megközelítés

Ha $\{e_1, \dots, e_n\}$ bázisa egy \mathbb{C} feletti V vektortérnek és $m(t) \in \mathbb{C}[t]$ egy pontosan n -ed fokú, normált polinom, például $m(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_{n-1} t^{n-1} + t^n$, akkor definiálja az $A \in L(V)$ lineáris transzformációt

$$A(e_k) = \begin{cases} e_{k+1} & , \text{ ha } 1 \leq k \leq n-1 \\ -\sum_{j=0}^{n-1} \alpha_j e_{j+1} & , \text{ ha } k = n. \end{cases}$$

Világos, hogy $A^j e_1 = e_{j+1}$ tetszőleges $0 \leq j \leq n-1$ mellett. Ebből azonnal következik, hogy

1. $\{e_1, A e_1, A^2 e_1, \dots, A^{n-1} e_1\}$ rendszer lineárisan független,
2. $\{e_1, A e_1, A^2 e_1, \dots, A^{n-1} e_1, A^n e_1\}$ rendszer már lineárisan összefüggő, hiszen $A^n e_1 = A A^{n-1} e_1 = A e_n = -\sum_{j=0}^{n-1} \alpha_j A^j e_1$.

Látkuk tehát, hogy m a legalacsonyabb fokú normált polinom, melyre $m(A) e_1 = 0$, így az e_1 vektorhoz tartozó kis minimálpolinom éppen m . Mivel a tér n -dimenziós, ezért m egyben minimálpolinomja is A -nak.

Persze egy lineáris transzformációnak egy szám pontosan akkor sajátértéke, ha az a minimálpolinomjának gyöke, emiatt az algebra alaptétele a következőképpen is fogalmazható:

Minden komplex vektortér feletti lineáris transzformációnak van sajátvektora.

Jelölje $\rho(A)$ az A transzformáció rangját, azaz a képtér dimenzióját, és $\nu(A)$ a defektust, azaz a magtér dimenzióját. Jól ismert, hogy ha $A \in L(V)$ egy lineáris transzformáció és $p(t) \in \mathbb{F}[t]$ egy polinom, akkor $\ker p(A)$ és $\text{Im } p(A)$ is A -ra invariáns alterek. Két lineáris trafóról azt mondjuk, hogy *kommutálnak*, ha $A_1 A_2 = A_2 A_1$. Könnyen látható, hogy ha A_1 és A_2 kommutálnak, akkor tetszőleges két p, q polinom mellett $p(A_1)$ és $q(A_2)$ is kommutálnak.

A.1. lemma. *Legyenek az $A_1, A_2 \in L(V)$ kommutáló lineáris transzformációk, valamint $p, q \in \mathbb{F}[t]$ polinomok. Ekkor $\ker p(A_1)$ és $\text{Im } p(A_1)$ is invariáns alterek $q(A_2)$ -re.*

Bizonyítás: Elég megmutatni, hogy $\ker(A_1)$ -re és $\text{Im } (A_1)$ -re invariáns A_2 , hiszen $p(A_1)$ és $q(A_2)$ is kommutálnak. No de, az $A_1 A_2 x = A_2 A_1 x = A_2 0 = 0$ szerint a magra, és $u = A_1 v$ jelöléssel az $A_2 u = A_2 A_1 v = A_1 A_2 v$ azonosságból a képre vonatkozó állítás következik. \square

A.2. lemma. *Legyen a $d > 1$ pozitív egész rögzítve. Tegyük fel, hogy az \mathbb{F} test rendelkezik avval a tulajdonsággal, hogy minden az \mathbb{F} feletti d -vel nem osztható dimenziós vektortér tetszőleges lineáris trafójának van sajátvektora. Ekkor minden olyan \mathbb{F} feletti vektortérre, amelynek dimenziója d -vel nem osztható igaz, hogy bármely két kommutáló lineáris transzformációjának van közös sajátvektora is.*

Bizonyítás: A tér dimenziója szerinti indukció. Egy egydimenziós tér minden nem nulla vektora sajátvektora tetszőleges lineáris transzformációjának, így persze bármely két egydimenziós téren értelmezett lineáris transzformációnak is van közös sajátvektora. Tegyük fel, hogy az állítás igaz minden legfeljebb n -dimenziós vektortérre és tekintsünk egy olyan \mathbb{F} feletti vektorteret, amely éppen n -dimenziós és d nem osztója n -nek. Jelölje A_1 és A_2 a szóban forgó két lineáris transzformációt. A feltétel szerint mondjuk A_1 -nek van sajátvektora, így valamely $\mu \in \mathbb{F}$ mellett $\nu(A_1 - \mu I) > 0$. Ha az $\nu(A_1 - \mu I) = n$, akkor a tér minden vektora sajátvektora A_1 -nek, így mivel A_2 -nek is van sajátvektora, ezért ez közös sajátvektoruk is. Ha $\nu(A_1 - \mu I) < n$, akkor $\nu(A_1 - \mu I) + \rho(A_1 - \mu I) = n$ miatt az $K = \ker(A_1 - \mu I)$ és a $L = \text{Im}(A_1 - \mu I)$ valódi alterek dimenziójának egyike nem osztható d -vel. No de A_2 és A_1 invariáns K -ra is és L -re is az előző lemma miatt, így alkalmazhatjuk az indukciós feltevést K és L közül a d -vel nem osztható dimenziós altérre, amely garantálja az A_1 és A_2 közös sajátvektorát. \square

Mivel a karakteristikus polinom gyökei a sajátértékek, és mivel egy n -dimenziós téren értelmezett lineáris transzformációk pontosan n -edfokú a karakteristikus polinomja, ezért a Bolzano-tétel szerint egy \mathbb{R} feletti páratlan dimenziós vektortér lineáris transzformációjának van sajátvektora. A fenti lemma tehát kommutáló transzformációk esetében közös sajátvektort garantál páratlan dimenziójú valós vektortér felett.

A.3. állítás. *Egy páratlan dimenziós komplex vektortér minden lineáris transzformációjának van sajátvektora.*

Bizonyítás: Legyen V a \mathbb{C} feletti vektortér, n páratlan szám a dimenziója, $A \in L(V)$ a transzformáció. Jelölje $\mathcal{H} = \{A \in L(V) : A = A^*\}$ az önadjungált transzformációkat. Világos, hogy \mathcal{H} egy valós, n^2 dimenziós vektortér. minden $C \in L(V)$ lineáris transzformáció előáll

$$C = \frac{C + C^*}{2} + i \frac{C - C^*}{2i}$$

alakban, ahol persze $\frac{1}{2}(C + C^*)$ és $\frac{1}{2i}(C - C^*)$ is önadjungált transzformációk. A továbbiakban rögzített $A \in L(V)$ mellett jelölje $L_1, L_2 : \mathcal{H} \rightarrow \mathcal{H}$ függvényeket

$$L_1(B) = \frac{AB + BA^*}{2} \text{ és } L_2(B) = \frac{AB - BA^*}{2i}$$

Világos, hogy L_1 és L_2 lineáris transzformációk a \mathcal{H} valós vektortéren, amelyekre minden $B \in \mathcal{H}$ mellett

$$AB = L_1(B) + iL_2(B).$$

E két operátor felcserélhető, hiszen tetszőleges $B \in \mathcal{H}$ mellett, ugyanis

$$\begin{aligned} L_1 \circ L_2(B) &= \frac{1}{4i} (A(AB - BA^*) + (AB - BA^*)A^*) = \frac{1}{4i} (A^2B + BA^{*2}) \\ L_2 \circ L_1(B) &= \frac{1}{4i} (A(AB + BA^*) - (AB + BA^*)A^*) = \frac{1}{4i} (A^2B + BA^{*2}) \end{aligned} .$$

Alkalmazhatjuk az n^2 páratlan dimenziós valós vektortérre az előző lemmát. Létezik $B \in \mathcal{H}$ nem a konstans zéró transzformáció és létezik α_1, α_2 valós szám, melyekre $L_1(B) = \alpha_1 B$ és $L_2(B) = \alpha_2 B$. Tehát ha valamely $v \in V$ vektorra $Bv \neq 0$, akkor

$$A(Bv) = \alpha_1 Bv + i\alpha_2 Bv = (\alpha_1 + i\alpha_2)Bv,$$

ergo $\alpha_1 + i\alpha_2$ sajátértéke, és Bv sajátvektora A -nak. \square

Minden komplex számnak van gyöke, ezért minden legfeljebb másodfokú komplex együtthatós polinomnak van zérushelye. Az algebra alaptételével ekvivalens állítás tehát, hogy egy komplex vektortér felett minden lineáris transzformáció minimálpolinomjának van legfeljebb másodfokú faktora.¹ Az is nyilvánvaló, hogy minden egész szám egyértelműen áll $2^k n$ alakban, ahol n páratlan.

A következő állítás tehát az algebra alaptételének egy ekvivalens megfogalmazása.

¹Azaz, van legfeljebb két dimenziós nem triviális invariáns altere.

A.4. állítás. *Tekintsünk egy V komplex vektorteret, amelynek dimenziója $2^k n$ alakú, ahol n páratlan egész. Ekkor V minden lineáris transzformációja minimálpolinomjának van legfeljebb másodfokú faktora.*

Bizonyítás: A k szerinti indukció. A $k = 0$ esetben az előző állítás szerint van sajátvektor is, tehát elsőfokú faktora is van a minimálpolinomnak. Tegyük fel, hogy igaz az állítás minden k -nál kisebb szám mellett. E feltétel azt jelenti, hogy minden olyan vektortérre igaz az állítás, – így az algebra alaptétele – melynek dimenziója 2^l páratlan szorosa $l < k$ mellett, azaz amelynek dimenzióját a $d = 2^k$ szám nem osztja. Alkalmazva a lemmát azt kapjuk, hogy ilyen dimenziójú vektortér kommutáló lineáris transzformációinak van közös sajátvektora is. Legyen tehát V dimenziója $2^k n$ alakban felírva, ahol n páratlan.

Rögzítsünk a térek egy bázisát, és jelölje $\mathcal{S} \subseteq L(V)$ azon lineáris transzformációk összességét, amelyeknek mátrixá az itt rögzített bázisban szimmetrikus. Ez egy komplex vektortér, amelyre

$$\dim(\mathcal{S}) = \frac{2^k n (2^k n + 1)}{2} = 2^{k-1} n (2^k n + 1) = 2^{k-1} n',$$

ahol n' páratlan. Alkalmazhatjuk tehát a komplex \mathcal{S} vektortérre az indukciós feltevést. Ehhez, rögzített $A \in L(V)$ lineáris transzformáció mellett, vezessük be az $L_1, L_2 : \mathcal{S} \rightarrow \mathcal{S}$ függvényeket.

$$L_1(B) = AB + BA^T \text{ és } L_2(B) = ABA^T.$$

Könnyű számolatás mutatja, hogy az L_1 és L_2 lineáris transzformációk kommutálnak:

$$\begin{aligned} (L_1 \circ L_2)B &= A(ABA^T) + (ABA^T)A^T = A(ABA^T + BA^T A^T) = \\ &= A(AB + BA^T)A^T = (L_2 \circ L_1)B. \end{aligned}$$

Létezik tehát közös sajátvektora az L_1 és L_2 transzformációknak, ergo létezik $\lambda, \mu \in \mathbb{C}$ komplex szám, és létezik nem az azonosan zérus $B \in \mathcal{S}$ lineáris transzformáció, amelyekre $L_1(B) = \lambda B$ és $L_2(B) = \mu B$, azaz $AB + BA^T = \lambda B$ és $ABA^T = \mu B$. Ebből

$$A\lambda B = A(AB + BA^T) = A^2B + ABA^T = A^2B + \mu B.$$

Ha tehát $u = Bv \neq 0$, akkor

$$A^2u - \lambda Au + \mu u = 0.$$

Ha u nem sajátvektora A -nak, akkor a $p(t) = t^2 - \lambda t + \mu$ egy nem zérus vektor kis minimálpolinomja, ezért a p polinom osztója a minimálpolinomnak. \square

A.2. Analízis megközelítés

Az algebra alaptételének (A.2. téTEL) bizonyításához felhasználunk néhány az elemi analízisből jól ismert állítást. Ezek közül a lényegesebbek az alábbiak:

1. Az origó középpontú zárt kör a sík kompakt részhalmaza.
2. Kompakt halmazon folytonos függvény felveszi minimumát.
3. minden komplex számnak van legalább egy k -adik gyöke ($k > 0$).
4. Komplex síkon differenciálható függvények folytonosak is.

A főtételt könnyen megérthetjük, ha áttekintjük a felé vezető utat. Két lényeges pontot kell látnunk. Az első (A.6. állítás) kompaktsági megondolás, polinomok növekedési ütemére (A.5. lemma) támaszkodva. Ez utóbbi lemma talán önmagában is érdekes, hiszen azt állítja, hogy egy polinom legalább a fokszáma nagyságrendjében növekszik. A másik döntő lépés (A.8. állítás) az Argand-féle becslésen (A.7. lemma) nyugszik. Ez a holomorf függvényekre vonatkozó nyílt leképezés tételnek az itt éppen elegendő speciális esete.

A.5. lemma. *Legyen $f : \mathbb{C} \rightarrow \mathbb{C}$ nem a konstans nulla, komplex n -edfokú polinom. Ekkor létezik olyan $r > 0$ valós szám, hogy minden $z \in \mathbb{C}$, $|z| > r$ esetén $|f(z)| > \frac{1}{2} |a_n| |z^n|$.*

Bizonyítás: Nyilván $f(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$ alakú, ahol $a_n \neq 0$. Világos, hogy $z \neq 0$ esetén

$$f(z) = z^n \left(a_n + a_{n-1} \frac{1}{z} + \dots + a_1 \frac{1}{z^{n-1}} + a_0 \frac{1}{z^n} \right).$$

Legyen $h : \mathbb{C} \rightarrow \mathbb{C}$ komplex polinom a következőképpen definiálva:

$$h(w) := a_{n-1}w + a_{n-2}w^2 + \dots + a_1w^{n-1} + a_0w^n.$$

Ekkor minden $z \in \mathbb{C}, z \neq 0$ mellett

$$f(z) = z^n (a_n + h(1/z)). \quad (\text{A.1})$$

A h folytonos 0-ban, és $h(0) = 0$, így létezik olyan $\delta > 0$ valós szám, melyre $w \in \mathbb{C}, |w| < \delta$ esetén $|h(w)| < \frac{1}{2} |a_n|$. Így ha $|z| > 1/\delta$, akkor $|1/z| < \delta$, amiből következik, hogy

$$|h(1/z)| < \frac{1}{2} |a_n|.$$

A háromszög-egyenlőtlenség és (A.1) miatt az $r := 1/\delta$ választással minden $z \in \mathbb{C}, |z| > r$ mellett

$$|f(z)| = |z^n| |a_n + h(1/z)| \geq |z^n| (|a_n| - |h(1/z)|) > |z^n| \left(|a_n| - \frac{1}{2} |a_n| \right) = \frac{1}{2} |a_n| |z^n|. \quad \square$$

A.6. állítás. Legyen $f : \mathbb{C} \rightarrow \mathbb{C}$ komplex polinom. Ekkor létezik $c \in \mathbb{C}$ komplex szám, melyre

$$|f(c)| = \inf \{|f(z)| : z \in \mathbb{C}\}.$$

Bizonyítás: Most úgy válasszuk meg az r pozitív valós számot, hogy egyszer az előző lemma, másrészt az $\frac{1}{2} |a_n| r^n > |a_0|$ feltétel is teljesüljön. Ekkor persze minden $z \in \mathbb{C}, |z| > r$ esetén

$$|f(z)| > \frac{1}{2} |a_n| |z^n| > |a_0| = |f(0)|$$

is teljesül. Ez azt jelenti, hogy ha bevezetjük az $\alpha := \inf \{|f(z)| : z \in \mathbb{C}, |z| \leq r\}$ jelölést, akkor minden $z \in \mathbb{C}, |z| > r$ esetén $|f(z)| \geq |f(0)| \geq \alpha$. Persze ha $|z| \leq r$ ez utóbbi akkor is teljesül, ezért

$$\alpha \leq \inf \{|f(z)| : z \in \mathbb{C}\}.$$

Mivel a fordított irányú egyenlőtlenség triviális, azt kapjuk, hogy

$$\inf \{|f(z)| : z \in \mathbb{C}, |z| \leq r\} = \inf \{|f(z)| : z \in \mathbb{C}\}.$$

De láttuk, hogy $\{z \in \mathbb{C} : |z| \leq r\} \subseteq \mathbb{C}$ a komplex számsík kompakt halmaza, így az f polinom és az abszolútérték-függvény folytonossága miatt létezik $c \in \mathbb{C}, |c| \leq r$, amelyre

$$|f(c)| = \alpha = \inf \{|f(z)| : z \in \mathbb{C}, |z| \leq r\} = \inf \{|f(z)| : z \in \mathbb{C}\}. \quad \square$$

A.7. lemma (Argand). Legyen $k \in \mathbb{N}, k > 0$ egész és $b \in \mathbb{C}, b \neq 0$ komplex szám, valamint $g : \mathbb{C} \rightarrow \mathbb{C}, g(0) = 0$ olyan függvény, amely a $0 \in \mathbb{C}$ pontban folytonos. Tekintsük a következőképpen definiált $h : \mathbb{C} \rightarrow \mathbb{C}$ leképezést:

$$h(z) := 1 + bz^k + z^k g(z).$$

Ekkor létezik $z \in \mathbb{C}$ komplex szám, melyre $|h(z)| < 1$.

Bizonyítás: Azt fogjuk megmutatni, hogy található $d \in \mathbb{C}$ és $t \in \mathbb{R}, t \in (0, 1)$ melyekre $|h(dt)| < 1$.

$$h(dt) = 1 + bd^k t^k + d^k t^k g(dt).$$

Válasszuk d komplex számot úgy, hogy $bd^k = -1$ teljesüljön, azaz d legyen a $-1/b$ komplex szám egyik k -adik gyöke. Ekkor

$$h(dt) = 1 - t^k + d^k t^k g(dt).$$

Amiből

$$|h(dt)| \leq 1 - t^k + t^k |d^k g(dt)| = 1 - t^k (1 - |d^k g(dt)|).$$

Ebből látszik, hogy elegendő megválasztani $t \in (0, 1)$ -et olyan módon, hogy $|d^k g(dt)| < 1$. Ez pedig nyilván megtehető $g(0) = 0$ és g -nek a 0 pontban feltett folytonossága miatt. \square

A.8. állítás. Legyen $f : \mathbb{C} \rightarrow \mathbb{C}$ legalább elsőfokú polinom. Ekkor minden $c \in \mathbb{C}, f(c) \neq 0$ komplex számhoz létezik olyan $\hat{c} \in \mathbb{C}$ komplex szám, melyre

$$|f(\hat{c})| < |f(c)|.$$

Bizonyítás: Tekintsük a

$$h(z) := \frac{f(z+c)}{f(c)}$$

legalább elsőfokú polinomot. Vegyük észre, hogy $h(0) = 1$, így e polinom

$$h(z) = 1 + a_k z^k + \dots + a_n z^n$$

alakú, ahol $a_k \neq 0$ valamely $k \geq 1$ -re, hiszen az f és ebből következően a h polinom nem konstans. Tovább alakítva:

$$h(z) = 1 + a_k z^k + z^k (a_{k+1} z + \dots + a_n z^{n-k}).$$

Világos, hogy a fenti h polinomra alkalmazható az előző lemma, így létezik $u \in \mathbb{C}$, melyre $|h(u)| < 1$. Ez viszont azt jelenti, hogy

$$\left| \frac{f(u+c)}{f(c)} \right| < 1$$

amiből $\hat{c} := u + c$ választással kapjuk, hogy

$$|f(\hat{c})| = |f(u+c)| < |f(c)|. \quad \square$$

Az algebra alaptétele. Legyen $f(t) \in \mathbb{C}[t]$ nem konstans polinom. Ekkor f -nek van gyöke a komplex számok körében.

Bizonyítás: Láttuk (A.6. állítás), hogy van olyan $c \in \mathbb{C}$ melyre

$$|f(c)| = \inf \{|f(z)| : z \in \mathbb{C}\}.$$

Ha $f(c) \neq 0$ lenne, akkor lenne (A.8. állítás) $\hat{c} \in \mathbb{C}$ melyre

$$|f(\hat{c})| < |f(c)| = \inf \{|f(z)| : z \in \mathbb{C}\}$$

is teljesülne, ami nem lehetséges. \square

A.3. A komplex számok egyértelműsége

Az alábbiakban azt fogjuk megvizsgálni, hogy lehet-e a sík pontjain a komplex számok bevezetésénél megadott szorzástól eltérő módon bevezetni szorzás műveletet úgy, hogy ez a valós számokon már megszokott szorzás kiterjesztése legyen, és a sík ellátva a szokásos összeadással valamint evvel a szorzásnak nevezett műveettel test legyen. Meg fogjuk mutatni, hogy ez nem lehetséges. Sőt azt is látni fogjuk, hogy az \mathbb{R} -től illetve a \mathbb{C} -től eltekintve nincs véges dimenziós \mathbb{R} feletti vektortér, amely test lesz olyan összeadásnak illetve szorzásnak nevezett műveettel, amely az \mathbb{R} -ben szokásos összeadás és szorzás kiterjesztése.

A.9. állítás. Legyen \mathbb{G} az \mathbb{R} valós számtestet tartalmazó olyan test, amelyben az $(+, \cdot)$ műveletek az \mathbb{R} -ben szokásos $(+, \cdot)$ műveletek kiterjesztései, valamint amely kétdimenziós vektortér \mathbb{R} feletti. Ekkor \mathbb{G} test izomorf \mathbb{C} -vel.

Bizonyítás: Meg fogjuk mutatni, hogy létezik $w \in \mathbb{G}$ melyre $w^2 = -1$. Ekkor készen is leszünk, mert nyilván $w \notin \mathbb{R}$ így $\text{lin}\{1, w\} = \mathbb{G}$ amiből könnyen látható, hogy az $\alpha + w\beta \mapsto \alpha + i\beta$ megfeleltetés izomorfizmus \mathbb{G} és \mathbb{C} között.

Legyen $v \in \mathbb{G} \setminus \mathbb{R}$. Ilyen v létezik, mivel \mathbb{G} kétdimenziós. Világos, hogy az $\{1, v, v^2\}$ vektorrendszer lineárisan összefüggő, hiszen három vektor egy kétdimenziós vektortérben. Így léteznek $\alpha, \beta \in \mathbb{R}$ valós számok melyekre $v^2 = \alpha + \beta v$. Most legyen $\gamma := \frac{-\beta}{2}$. Ekkor nyilván $v^2 = \alpha - 2\gamma v$. Most tekintsük a $(v + \gamma)^2$ kifejezést.

$$(v + \gamma)^2 = v^2 + 2\gamma v + \gamma^2 = \alpha + \gamma^2$$

Azt kaptuk tehát, hogy létezik $r \in \mathbb{R}$ valós szám melyre $(v + \gamma)^2 = r$. De vegyük észre, hogy ha $r \geq 0$ lenne, akkor $t = \sqrt{r} \in \mathbb{R}$ jelöléssel $(v + \gamma)^2 = t^2$ következne, amiből pedig $v \in \mathbb{R}$ következtetésre juthatnánk ellentétben a v -re kiindulásul tett feltétellel. Ebből már világos, hogy ha w -t

$$w := \frac{v + \gamma}{\sqrt{|r|}} \in \mathbb{G}$$

módon definiáljuk akkor $w^2 = \frac{r}{|r|} = -1$, hiszen $r < 0$. \square

A.10. definíció. Legyen \mathbb{G} az \mathbb{R} valós számtestet tartalmazó olyan test amelyben az $(+, \cdot)$ műveletek az \mathbb{R} -ben szokásos $(+, \cdot)$ műveletek kiterjesztései. Az $x \in \mathbb{G}$ elemet \mathbb{R} felett algebraiknak nevezzük, ha létezik nem konstans zéró valós együtthatós polinom, melynek x gyöke.

A.11. állítás. Legyen \mathbb{G} az \mathbb{R} valós számtestet tartalmazó olyan test amelyben az $(+, \cdot)$ műveletek az \mathbb{R} -ben szokásos $(+, \cdot)$ műveletek kiterjesztései, valamint amely véges dimenziós vektortér \mathbb{R} felett. Ekkor \mathbb{G} minden eleme algebrai \mathbb{R} felett.

Bizonyítás: Legyen \mathbb{G} dimenziója n és $v \in \mathbb{G}$ tetszőleges vektor. Tekintsük az

$$\{1, v, v^2, \dots, v^n\}$$

vektorrendszert. Ez nyilván lineárisan összefüggő, hiszen $n + 1$ vektor egy n dimenziós vektortérben. Így van nem triviális 0 -t adó lineáris kombinációja. Azaz léteznek $\alpha_0, \alpha_1, \dots, \alpha_n$ nem csupa nulla valós számok melyekre

$$\sum_{i=0}^n \alpha_i v^i = 0$$

De ez pont azt jelenti, hogy ha a p polinomot

$$p(x) := \sum_{i=0}^n \alpha_i x^i$$

módon definiáljuk, akkor $p(v) = 0$. \square

A.12. állítás (Weierstrass). Legyen \mathbb{G} az \mathbb{R} valós számtestet tartalmazó olyan test, amelyben az $(+, \cdot)$ műveletek az \mathbb{R} -ben szokásos $(+, \cdot)$ művelet kiterjesztései, valamint amelynek minden eleme algebrai \mathbb{R} felett. Ekkor két eset lehetséges: Vagy \mathbb{G} test-izomorf \mathbb{R} -rel, vagy \mathbb{G} test-izomorf \mathbb{C} -vel.

Bizonyítás: Tegyük fel, hogy \mathbb{G} tartalmaz \mathbb{R} -től különböző v vektort. Ekkor azt kell megmutatni, hogy \mathbb{G} izomorf \mathbb{C} -vel.

Először azt mutatjuk meg, hogy

$$v^2 \in \text{lin}\{1, v\} \tag{A.2}$$

Mivel v algebrai \mathbb{R} felett, ezért létezik p valós együtthatós polinom melyre $p(v) = 0$. De láttuk, hogy valós együtthatós polinom első- és másodfokú tényezők szorzatára bomlik (1.66.), ami azt jelenti, hogy van olyan q első, vagy másodfokú valós együtthatós polinom melyekre $q(v) = 0$. Ha q első fokú lenne az azt jelentené, hogy található $\alpha, \beta \in \mathbb{R}$ valós számok melyekre $\alpha + \beta v = 0$, azaz $v \in \mathbb{R}$ lenne. Tehát q pontosan másodfokú. Ez viszont azt jelenti, hogy léteznek $\{\alpha, \beta, \gamma\} \subset \mathbb{R}$ valós számok $\gamma \neq 0$, melyekre $\alpha + \beta v + \gamma v^2 = 0$. Ebből persze

$$v^2 = \frac{-\alpha}{\gamma} + \frac{-\beta}{\gamma}v$$

már könnyen következik, bizonyítva (A.3)-et.

Most megmutatjuk, hogy \mathbb{G} tartalmaz \mathbb{C} -vel izomorf testet. Tekintsük a

$$K := \text{lin}\{1, v\}$$

kétdimenziós alterét \mathbb{G} -nek. Világos, hogy ennek test voltához elegendő megmutatni, hogy a \mathbb{G} -beli műveletekre zárt. Az összeadásra való zártsgá triviális a K altér mivoltából, a szorzásra való zártsgá pedig (A.3) következménye. Így tehát K két dimenziós test kiterjesztése \mathbb{R} -nek, ami azt jelenti (A.9.), hogy K test

izomorf \mathbb{C} -vel.

Utoljára meggematjuk, hogy \mathbb{G} minden w eleme \mathbb{C} -beli is. Mivel w algebrai \mathbb{R} felett, ezért létezik p valós együtthatós polinom, melyre $p(w) = 0$. De tekinthetjük p -t a komplex számtest feletti polinomnak is, így p felbomlik elsőfokú komplex polinomok szorzatára (1.64). Ez viszont azt jelenti, hogy létezik olyan c komplex szám, melyre $w - c = 0$, tehát $w \in \mathbb{C}$ valóban teljesül. \square

B. függelék

A Frobenius-normálalak

A kis minimálpolinom fogalmát általánosítjuk.

B.1. definíció (irányító polinom). Legyen $A \in L(V)$ egy lineáris transzformáció, $M \subseteq V$ egy az A -ra nézve invariáns altér, és $v \in V$ egy rögzített vektor. Tekintsük az $\mathbb{F}[t]$ polinomgyűrű következő részhalmazát.

$$S(v; M) = \{p \in \mathbb{F}[t] : p(A)v \in M\}.$$

Látható, hogy $S(v; M)$ egy ideálja $\mathbb{F}[t]$ -nek. A fenti ideált generáló normált polinomot nevezzük az A transzformáció v -t M -be vivő *irányító polinomjának* és $p_{v; M}$ módon jelöljük.

Világos, hogy ha $M_1 \subseteq M_2$ invariáns alterek, akkor $S(v; M_1) \subseteq S(v; M_2)$, ezért $p_{v; M_2}|p_{v; M_1}$. Speciálisan, ha $M = \{0\}$ triviális altér, akkor $p_{v; \{0\}}$ egybeesik a p_v kis minimálpolinommal. Meggondoltuk tehát, hogy minden M invariáns altér mellett a $p_{v; M}$ irányító polinom osztója a p_v kis minimálpolinomnak $p_{v; M}|p_v$, ergo az irányító polinom is legfeljebb $\dim(V)$ -ed fokú. Ebből az is nyilvánvaló, hogy ha egy irányító polinomra $p_{v; M}(A)v = 0$ teljesül, akkor $p_v|p_{v; M}$ is fennáll, tehát a v -hez tartozó irányító polinom egybeesik a v -hez tartozó kisminimálpolinommal.

Látható az is, hogy $p_{v; M}(t) = 1$ akkor és csak akkor teljesül, ha $v \in M$.

Most elegendő feltételezni, hogy két különböző vektorhoz tartozó irányító polinom azonos legyen.

B.2. állítás. Legyen $M \subseteq V$ egy A -invariáns altér, és $v, w \in M$ rögzített vektorok. Ha $v - w \in M$ fennáll, akkor $p_{v; M} = p_{w; M}$ is teljesül.

Bizonyítás: Mivel M egy A -invariáns altér, ezért $p(A)(v - w) \in M$ minden p polinom mellett. No de

$$p(A)(v - w) = p(A)v - p(A)w,$$

így $p(A)v \in M$ pontosan akkor teljesül, ha $p(A)w \in M$ is fennáll. Ez azt jelenti, hogy az $S(v; M)$ és az $S(w; M)$ főideálok azonosak, ergo a generáló polinom is ugyanaz. \square

B.3. definíció (megengedhető altér). Legyen $M \subseteq V$ egy A -invariáns altér. Azt mondjuk, hogy M megengedhető altér, ha minden $p \in \mathbb{F}[t]$ polinomra és minden $v \in V$ vektorra, a $p(A)v \in M$ tartalmazásból következik, hogy létezik olyan $w \in M$ vektor is, amelyre $p(A)v = p(A)w$.

B.4. állítás. Ha egy M invariáns altérnek van invariáns direktkiegészítője, akkor M megengedhető altér.

Bizonyítás: Legyen tehát $M \oplus N = V$ valamely N invariáns altér mellett, és tegyük fel, hogy valamely $u \in V$ esetén $p(A)u \in M$. Az u vektor előáll $u = v + w$ alakban, ahol $v \in M$ és $w \in N$. Persze

$$p(A)u - p(A)v = p(A)w.$$

A bal oldal minden két vektor a M -beli, a jobb oldali vektor N -beli az N direktkiegészítő invarianciája miatt. Így $p(A)w = 0$. Találtunk tehát $v \in M$ vektort, amelyre $p(A)u = p(A)v$. \square

Példaképpen nézzünk egy m -ed rendben nilpotens $B \in L(V)$ lineáris transzformációt. Válasszunk egy $v \in V$ vektort, amelyre $B^{m-1}v \neq 0$. Ekkor $\text{lin}(v; B)$ egy megengedhető altér. Legyen ugyanis valamely $x \in V$ vektorra $Bx \in \text{lin}(v; B)$. Ekkor Bx előáll

$$Bx = \alpha_0 v + \alpha_1 Bv + \dots + \alpha_{m-1} B^{m-1}v$$

alakban. Na most $B^{m-1}v \neq 0$, és m a nilpotencia rendje, ergo $v \notin \text{Im } B$. No de, a fent kiemelt azonosságban a bal oldal, de az első tag kivételével a jobb oldal valamennyi tagja is az $\text{Im } B$ eltér egy-egy eleme, ezért $\alpha_0 v \in \text{Im } B$. Ebből persze $\alpha_0 = 0$ következik, ergo Bx előáll mint egy $\text{lin}(v; B)$ -beli vektor képe:

$$Bx = B(\alpha_1 v + \dots + \alpha_{m-1} B^{m-2}v).$$

Ez azt jelenti, hogy $\text{lin}(v; B)$ valóban egy megengedhető altér.

A nilpotens transzformációk felbontásáról szóló 14.6. állítás általánosításaként meg fogjuk mutatni, hogy tetszőleges lineáris transzformáció mellett is igaz, hogy egy megengedhető altérnek minden van invariáns direktkiegészítője.

A triviális – de nagyon fontos – példa megengedhető altérre az $M = \{0\}$ altér, persze tetszőleges lineáris transzformáció mellett.

A megengedhető altér definíciója kicsit élesíthető az irányító polinom használatával. Ha minden irányító polinomra teljesül a megengedhető altér definíciójában előírt tulajdonság, akkor már minden más polinomra is fennáll.

B.5. állítás. Legyen M egy A -invariáns altér a V vektortérben. Az M pontosan akkor megengedhető, ha minden $w \in V$ vektorhoz létezik $v \in M$, amelyre $p_{w,M}(A)w = p_{w,M}(A)v$.

Bizonyítás: Legyen p tetszőleges olyan polinom, amelyre $p(A)w \in M$, valamely $w \in M$ mellett. Persze $p \in S(w; M)$. Ennek az ideálnak a generáló eleme $p_{w,M}$, ergo létezik h polinom, amelyre $p(t) = h(t)p_{w,M}(t)$. Így a feltevés szerint valamely $v \in M$ mellett

$$p(A)w = h(A)(p_{w,M}(A)w) = h(A)(p_{w,M}(A)v) = p(A)v. \quad \square$$

B.6. lemma. Legyen M egy megengedhető altér az $A \in L(V)$ lineáris transzformációra nézve és $v \notin M$ egy vektor. Ekkor létezik olyan $w \in V$ vektor, amelyre

1. $p_w = p_{w,M} = p_{v,M}$;
2. $\text{lin}(w; A) \cap M = \{0\}$;
3. $M + \text{lin}(w; A) = M + \text{lin}(v; A)$.

Bizonyítás: Jelölje $p_{v,M}$ a v vektort M -be vivő irányító polinomot. Mivel M megengedhető, ezért létezik $u \in M$, amelyre $p_{v,M}(A)v = p_{v,M}(A)u$. Legyen $w = v - u$.

1. Ekkor $w - v \in M$, ezért $p_{w,M} = p_{v,M}$. Persze $p_{v,M}(A)w = 0$, ezért $p_{w,M}(A)w = 0$ is fennáll, amiből már következik, hogy $p_{w,M} = p_w$.
2. Most legyen $x \in M$, amelyre $x \in \text{lin}(w; A)$. Ekkor létezik $f \in \mathbb{F}[t]$ polinom, amelyre $x = f(A)w$. Az irányító polinom definíciója szerint $p_{w,M}f$. Már láttuk, hogy $p_{w,M} = p_w$, így f voltaképpen a p_w kis minimálpolinom többszöröse, ezért $f(A)w = 0$, ergo $x = 0$.
3. A harmadik egyenlőséghez:

$$\begin{aligned} M + \text{lin}(w; A) &= \{m + f(A)w : m \in M, f \in \mathbb{F}[t]\} = \{(m - f(A)u) + f(A)v : m \in M, f \in \mathbb{F}[t]\} \\ &\subseteq \{m + f(A)v : m \in M, f \in \mathbb{F}[t]\} = M + \text{lin}(v; A). \end{aligned}$$

Itt a tartalmazás azért áll fenn, mert M egy invariáns altér, így $f(A)u \in M$. Az ellenkező irányú tartalmazás a fentivel analóg:

$$\begin{aligned} M + \text{lin}(v; A) &= \{m + f(A)v : m \in M, f \in \mathbb{F}[t]\} = \{(m + f(A)u) + f(A)w : m \in M, f \in \mathbb{F}[t]\} \\ &\subseteq \{m + f(A)w : m \in M, f \in \mathbb{F}[t]\} = M + \text{lin}(w; A). \end{aligned}$$

\square

B.7. lemma. Legyen M a V vektortér egy valódi altere, amely az A lineáris transzformációra nézve megengedhető. Válasszuk meg $a v \in V$ vektort olyan módon, hogy $\deg p_{v;M}$ a lehető legnagyobb legyen, azaz minden $x \in V$ mellett

$$\deg p_{v;M} \geq \deg p_{x;M}.$$

Jelölje

$$\bar{M} = M + \text{lin}(v; A)$$

invariáns alteret. Tegyük fel, hogy egy $w \in V$ mellett

$$p_{w;\bar{M}}(A)w = m + g(A)v \quad (\dagger)$$

valamely $m \in M$ vektor és valamely $g \in \mathbb{F}[t]$ polinom mellett. Ekkor alkalmass $m' \in M$ mellett

$$p_{w;\bar{M}}|g \quad \text{és} \quad p_{w;\bar{M}}(A)m' = m.$$

Bizonyítás: Először gondoljuk meg, hogy valóban megválasztható egy v vektor úgy, hogy $\deg p_{v;M}$ maximális legyen. Mivel $M \neq V$, ezért van $x \in V \setminus M$. minden ilyen x vektorra $0 < \deg p_{x;M} \leq \dim(V)$. Ezért bármely ilyen x meg is felel v -nek, amelynél $\deg p_{x;M}$ a lehető legnagyobb.

Mivel invariáns alterek Minkowski-összege invariáns, ezért \bar{M} valóban egy az M alteret szigorúan tartalmazó invariáns alter.

Az egyszerűbb jelölés kedvéért legyen $f = p_{w;\bar{M}}$ a w -t \bar{M} -be vivő irányító polinom. A maradékos osztás tétele szerint léteznek olyan h, r polinomok, amelyekre

$$g = h \cdot f + r, \quad \text{ahol} \quad \deg r < \deg f.$$

Definiálja $u = w - h(A)v$. Világos, hogy $u - w \in \bar{M}$ hiszen $v \in \bar{M}$, így

$$p_{u;\bar{M}} = p_{w;\bar{M}} = f.$$

Az $f(A)$ transzformációt az u definíciójára alkalmazva, majd felhasználva a lemma (\dagger) feltételét azt kapjuk, hogy

$$\begin{aligned} f(A)u &= f(A)w - f(A)h(A)v = m + g(A)v - f(A)h(A)v \\ &= m + (g(A) - f(A)h(A))v = m + r(A)v. \quad (\ddagger) \end{aligned}$$

Innen azonnal látszik, hogy összesen azt kell megmutatnunk, hogy az r a konstans zérus polinom. Ekkor nyilván $f|g$, és $f(A)u = m$. Mivel $m \in M$ és M egy megengedhető alter, ezért persze létezik $m' \in M$, amelyre $f(A)m' = m$.

Az $r(t) = 0$ megmutatásához jelölje $p = p_{u;\bar{M}}$ az u -t \bar{M} -be vivő irányító polinomot. Mivel $M \subset \bar{M}$, ezért $p_{u;\bar{M}}|p_{u;M}$, tehát a bevezetett jelölésekkel $f|p$. Létezik tehát valamely q nem zérus polinom, amelyre $p = fq$. Alkalmazzuk a $q(A)$ transzformációt (\ddagger) azonosságára.

$$p(A)u = q(A)f(A)u = q(A)m + q(A)r(A)v.$$

Itt $p(A)u \in M$ és $q(A)m \in M$, ergo $(q \cdot r)(A)v \in M$ is fennáll. Na most, ha r nem a konstans zérus, akkor $q \cdot r$ többszöröse a $p_{v;M}$ irányító polinomnak, ami pedig v konstrukciója szerint nem kisebb fokú, mint $p_{u;M}$. Így

$$\deg(q \cdot r) \geq \deg p_{v;M} \geq \deg p_{u;M}.$$

Innen persze $\deg q + \deg r \geq \deg p \geq \deg f + \deg q$ következik, ami képtelenség az r konstrukciója szerint. Ezt kellett belátni. \square

Első alkalmazásként lássuk, hogy egy megengedhető alterhez minden található olyan invariáns alter, amelyet hozzáadva újra egy megengedhető alteret kapunk. Ez szolgál a ciklikus felbontási téTEL rekurziójának alapjaként.

B.8. állítás. Legyen M egy megengedhető alter az $A \in L(V)$ lineáris transzformációra nézve. Válasszunk egy maximális fokszámú M -be vivő vezető polinommal rendelkező $v \in V$ vektort. Ekkor az $M = M + \text{lin}(v; A)$ alter is megengedhető.

Bizonyítás: Legyen $w \in V$, és a B.5. állításnak megfelelően tekintsük a $p_{w;\bar{M}}(A) w \in M + \text{lin}(v;A)$ vektort. Világos, hogy valamely $m \in M$ és valamely g polinom mellett

$$p_{w;\bar{M}}(A) w = m + g(A) v.$$

A B.7. lemma szerint valamely $m' \in M$ mellett $p_{w;\bar{M}}(A) m' = m$ és van olyan h polinom, amelyre $g = p_{w;\bar{M}} \cdot h$. Ezeket helyettesítve kapjuk, hogy

$$p_{w;\bar{M}}(A) w = p_{w;\bar{M}}(A) m' + p_{w;\bar{M}}(A)(h(A) v) = p_{w;\bar{M}}(A)(m' + h(A) v).$$

Itt $m' \in M \subseteq \bar{M}$ és $v \in \bar{M}$ miatt az argumentumban szereplő $m' + h(A) v$ vektor \bar{M} -ban van. A B.5. állítás szerint éppen ezt kellett belátni. \square

B.9. állítás (Ciklikus felbontás). Legyen V egy véges dimenziós vektortér az \mathbb{F} test felett, $A \in L(V)$ egy lineáris transzformáció. Tegyük fel, hogy adott egy $M_0 \subset V$ megengedhető altér, amelyre $M_0 \neq V$. Ekkor létezik r pozitív egész, és léteznek $w_1, \dots, w_r \in V$ vektorok, amelyekre

1. $A w_1, \dots, w_r$ egyike sem zérus;
2. $V = M_0 \oplus \text{lin}(w_1;A) \oplus \dots \oplus \text{lin}(w_r;A)$;
3. $A w_1, \dots, w_r$ vektorok kis minimálpolinomjai olyanok, hogy minden osztója az előzőnek, azaz ha $p_k = p_{w_k}$ a w_k vektorhoz tartozó kis minimálpolinom, akkor minden $k = 2, \dots, r$ mellett $p_k | p_{k-1}$.

Bizonyítás: Válasszunk egy $v_1 \notin M_0$ vektort, amelyre minden $x \in V$ mellett

$$\deg p_{v_1;M_0} \geq \deg p_{x;M_0}.$$

Alkalmazzuk a B.6. lemmát az M_0 megengedhető altérre és a v_1 vektorra. Azt kapjuk, hogy létezik $w_1 \notin M_0$ vektor, amelyre

1. $p_{w_1} = p_{w_1;M_0} = p_{v_1;M_0}$;
2. $\text{lin}(w_1;A) \cap M_0 = \{0\}$;
3. $M_0 + \text{lin}(w_1;A) = M_0 + \text{lin}(v_1;A)$.

Azt is láttuk, hogy v_1 maximalitási konstrukciója szerint $M_0 + \text{lin}(v_1;A)$ is megengedhető altér. Legyen

$$M_1 = M_0 + \text{lin}(v_1;A) = M_0 \oplus \text{lin}(w_1;A).$$

Ha $M_1 = V$, akkor $r = 1$ választással készen is vagyunk.

Ha $M_1 \neq V$, akkor megismételjük a fenti eljárást az M_1 megengedhető altérrel: Válasszunk tehát egy $v_2 \notin M_1$ vektort, amelyre minden $x \in V$ mellett

$$\deg p_{v_2;M_1} \geq \deg p_{x;M_1}.$$

Alkalmazzuk a B.6. lemmát az M_1 megengedhető altérre és a v_2 vektorra. Azt kapjuk, hogy létezik $w_2 \notin M_1$ vektor, amelyre

1. $p_{w_2} = p_{w_2;M_1} = p_{v_2;M_1}$;
2. $\text{lin}(w_2;A) \cap M_1 = \{0\}$;
3. $M_1 + \text{lin}(w_2;A) = M_1 + \text{lin}(v_2;A)$.

Tudjuk azt is, hogy v_2 maximalitási konstrukciója szerint $M_1 + \text{lin}(v_2;A)$ is egy megengedhető altér V -nek. Legyen

$$M_2 = M_1 + \text{lin}(v_2;A) = M_1 \oplus \text{lin}(w_2;A).$$

Most meg kell mutatnunk, hogy $p_{w_2} | p_{w_1}$. Ehhez először azt vegyük észre, hogy v_1 helyett w_1 vektorra is teljesül a

$$\deg p_{w_1;M_1} \geq \deg p_{x;M_1}$$

maximalitási feltétel minden $x \in V$ mellett. Másodsor vegyük észre, hogy triviálisan teljesül az

$$p_{w_2;M_1}(A)w_2 = 0 + p_{w_1;M_0}(A)w_1$$

egyenlőség, hiszen minden két oldal zérus. A B.7. lemmát alkalmazzuk és kapjuk, hogy $p_{v_2;M_1}|p_{v_1;M_0}$, ami a mi jelölésekkel azt jelenti, hogy $p_2|p_1$. Ha $M_2 = V$, akkor $r = 2$ választással készen is vagyunk.

Ha $M_2 \neq V$, akkor megismételjük a fenti eljárást M_1 helyett az M_2 megengedhető altérrel, stb.

Mivel minden egyes lépésben az M_k altér dimenziója nő, ezért van olyan r egész, amelyre $M_r = V$. Így az $M_0; \text{lin}(w_1; A), \dots, \text{lin}(w_r; A)$ alterek olyanok, hogy minden diszjunkt az előzőök összegétől, ezért a direktösszeg értelmes és

$$M_0 \oplus \text{lin}(w_1; A) \oplus \dots \oplus \text{lin}(w_r; A) = M_r = V$$

és minden $k \geq 2$ mellett $p_k|p_{k-1}$. \square

Korábban már láttuk, hogy ha egy invariáns altérnek van invariáns direktkiegészítője, akkor az megengedhető is. A ciklikus felbontás szerint ha M_0 megengedhető, akkor annak van $\text{lin}(w_1; A) \oplus \dots \oplus \text{lin}(w_r; A)$ invariáns direktkiegészítője. Igazoltuk tehát a következő állítást:

B.10. állítás. Legyen $A \in L(V)$ egy lineáris transzformáció, és $M_0 \subseteq V$ egy invariáns altér az A transzformációra nézve. Az M -nek pontosan akkor van invariáns direktkiegészítője, ha M egyben megengedhető is.

A ciklikus felbontási tételeben az első p_1 polinom különösen nevezetes. Amennyiben $M_0 = \{0\}$, akkor p_1 az A minimálpolinomja, amint azt rögtön megértjük.

Általános esetben tekintsük az összes olyan polinomot, amelyre $p(A)v \in M_0$ minden $v \in V$ mellett:

$$S(M_0) = \{p \in \mathbb{F}[t] : p(A)v \in M_0, \text{ minden } v \in V\}.$$

Világos, hogy az M_0 invarianciája szerint $S(M_0)$ egy nemzérus ideál. Jelölje s_{M_0} az $S(M_0)$ normált generátor elemét.

Világos, hogy ha visszatérünk az $M_0 = \{0\}$ esethez, akkor $s_{\{0\}}$ éppen az A lineáris transzformáció minimálpolinomja.

B.11. állítás. Legyen V egy véges dimenziós vektortér az \mathbb{F} test felett, $A \in L(V)$ egy lineáris transzformáció. Legyen $M_0 \subset V$ egy az A -ra nézve invariáns altér, és legyenek $w_1, \dots, w_r \in V$ vektorok, amelyekre és ezek p_k kis minimálpolinomjaira fenállnak a következők.

1. $A w_1, \dots, w_r$ egyike sem zérus;
2. $V = M_0 \oplus \text{lin}(w_1; A) \oplus \dots \oplus \text{lin}(w_r; A)$;
3. minden $k = 2, \dots, r$ mellett $p_k|p_{k-1}$.

Ekkor $p_1 = s_{M_0}$, azaz p_1 a fenti $S(M_0)$ ideál normált generáló eleme. Speciálisan, ha $M_0 = \{0\}$, akkor p_1 az A minimálpolinomja.

Bizonyítás: Először megmutatjuk, hogy minden $v \in V$ mellett $p_1(A)v \in M$. Legyen tehát $v \in V$ rögzítve. Ekkor a 2. feltétel szerint

$$v = m_0 + \sum_{k=1}^r f_k(A)w_k$$

valamely f_k polinomokra, és $m_0 \in M_0$ vektorra. Alkalmazzuk minden két oldalra a $p_1(A)$ transzformációt. Így

$$p_1(A)v = p_1(A)m_0 + \sum_{k=1}^r p_1(A)f_k(A)w_k.$$

No de minden k -ra $p_k|p_1$, és $p_k(A)w_k = 0$, azért a szumma minden tagja zérus, így M_0 invarianciáját használva kapjuk, hogy $p_1(A)v \in M = p_1(A)m_0 \in M$. Megmutattuk tehát, hogy $s_{M_0}|p_1$.

Világos, hogy $s_{M_0}(A)w_1 \in M_0 \cap \text{lin}(w_1; A) = \{0\}$, hiszen a $\text{lin}(w_1; A)$ egy az A transzformációra nézve invariáns altér. Így $s_{M_0}(A)w_1 = 0$, amiből már $p_1|s_{M_0}$ is következik. \square

Ezek szerint megmutattuk, hogy a minimálpolinom minden valamely vektorhoz tartozó kis minimálpolinom. Mitöbb ez a vektor a maximális fokú kis minimálpolinomhoz tartozó vektor. Megmutattuk tehát a következőt.

B.12. állítás. Legyen $A \in L(V)$ egy lineáris transzformáció. Legyen $v \in V$ olyan vektor, amelyhez tartozó kis minimálpolinom lehető legnagyobb fokszámú, azaz

$$\deg p_v \geq \deg p_x$$

minden $x \in V$ mellett. Ekkor a p_v kis minimálpolinom egyben a transzformáció minimálpolinomja is.

Persze ebből az állításból is látszik, hogy a transzformáció minimálpolinomjának fokszáma legfeljebb a tér dimenziója.

Az eddigiek től független következő lemma invariáns alereknek egy polinom által képzett direktképéről szól.

B.13. lemma. Legyen az $A \in L(V)$ lineáris transzformáció és az $f \in \mathbb{F}[t]$ polinom rögzítve. Ekkor

1. minden $v \in V$ mellett $\text{lin}(v; A) = \text{lin}(f(A)v; A)$;
2. Ha V előáll mint invariáns alerek $V = V_1 \oplus \dots \oplus V_k$ direktösszegének alakjában, akkor ezek f képére az $f(A)V = f(A)V_1 \oplus \dots \oplus f(A)V_k$ direktösszeg alakú előállítás is teljesül;
3. Ha az u és v vektorok kis minimálpolinomja azonos, akkor az $f(A)u$ és $f(A)v$ vektoroknak is ugyanaz a kis minimálpolinomja.

Bizonyítás: Az első állítás egyszerűen azért igaz, mert a transzformáció polinomjai kommutálnak egymással.

A második állításhoz csak azt kell észrevenni, hogy ha invariáns alerek direktösszege értelmes, akkor ezek $f(A)$ transzformációval képzett direktösszege is értelmes.

A harmadik állításhoz, mint 10.7. definícióban jelölje $J_{A,v}$ azon p polinomok halmazát, amelyekre $p(A)v = 0$ teljesül. A feltétel szerint $J_{A,u} = J_{A,v}$. Na most $p \in J_{A,f(A)v}$ pontosan akkor, ha $p \cdot f \in J_{A,v}$, ami persze ugyan az, mint $p \cdot f \in J_{A,u}$, ami már a $p \in J_{A,f(A)u}$ feltétellel ekvivalens. Meggondoltuk tehát, hogy $J_{A,f(A)v} = J_{A,f(A)u}$ is fennáll, ami éppen azt jelenti, hogy a fenti ideálok generáló elemei is azonosak. \square

B.14. állítás (Ciklikus felbontás egyértelműsége). Legyen $M_0 \subset V$ egy A -invariáns altér a V vektortérben. Tegyük fel, hogy adott az r, s pozitív egész, adottak az $v_1, \dots, v_r \in V$, és a $w_1, \dots, w_s \in V$ vektorok a hozzájuk tartozó p_1, \dots, p_r és a q_1, \dots, q_s kis minimálpolinomokkal. Tegyük fel, hogy

1. v_1, \dots, v_r egyike sem zérus;
2. $V = M_0 \oplus \text{lin}(v_1; A) \oplus \dots \oplus \text{lin}(v_r; A)$;
3. $p_k | p_{k-1}$ minden $k = 2, \dots, r$ mellett.

Hasonlóan a w_1, \dots, w_s vektortokra is tegyük fel, hogy

1. w_1, \dots, w_s egyike sem zérus;
2. $V = M_0 \oplus \text{lin}(w_1; A) \oplus \dots \oplus \text{lin}(w_s; A)$;
3. $g_k | g_{k-1}$ minden $k = 2, \dots, s$ mellett.

Ekkor $r = s$, $p_k = g_k$ minden $k = 1, \dots, r$ mellett és minden k indexre $\text{lin}(v_k; A)$ és $\text{lin}(w_k; A)$ izomorf alerek.

Bizonyítás: Tudjuk, hogy $p_1 = q_1$, hiszen a feltételekből következik, hogy p_1 és q_1 is az a legalacsonyabb fokszámú normált polinom, amely a tér minden elemét az M_0 altérbe viszi. Ebből kiindulva azt mutatjuk meg, hogy minden $k = 1, \dots, r$ mellett $k \leq s$ és $q_k = p_k$. Nézzük ezt k szerinti indukcióval.

Ha $k = 1$, akkor persze $k \leq s$ automatikusan teljesül és éppen az imént láttuk, hogy $p_1 = q_1$ is fennáll.

Most tegyük fel, hogy valamely $k < r$ mellett $k \leq s$ és $p_j = q_j$ minden $j = 1, \dots, k$. Meg kell mutatnunk, hogy ebből $k + 1 \leq s$ és $p_{k+1} = q_{k+1}$ is következik. Mivel $k < r$, és a $\text{lin}(w_j; A)$ alerek izomorfak a $\text{lin}(v_j; A)$ alerekkel, ezért

$$\dim(M_0 \oplus \text{lin}(w_1; A) \oplus \dots \oplus \text{lin}(w_k; A)) = \dim(M_0 \oplus \text{lin}(v_1; A) \oplus \dots \oplus \text{lin}(v_k; A)) < \dim(V),$$

ezért $k < s$, ergo $k + 1 \leq s$ valóban fennáll. Legyen most az egyszerűbb jelölés kedvéért $f = p_{k+1}$. Ekkor alkalmazva az előző B.13. lemma első két pontját $f(A)(\text{lin}(v_j; A)) = \text{lin}(f(A)v_j; A)$ és ez a $\{0\}$ altér, ha $j \geq k + 1$, hiszen az ilyen j indexek mellett $p_j|f$. Így az is világos, hogy

$$\begin{aligned} f(A)(V) &= f(M_0) \oplus \text{lin}(f(A)v_1; A) \oplus \dots \oplus \text{lin}(f(A)v_k; A) \text{ és} \\ f(A)(V) &= f(M_0) \oplus \text{lin}(f(A)w_1; A) \oplus \dots \oplus \text{lin}(f(A)w_k; A) \oplus \\ &\quad \text{lin}(f(A)w_{k+1}; A) \oplus \dots \oplus \text{lin}(f(A)w_s; A). \end{aligned}$$

A lemma harmadik pontja és az indukciós feltevés szerint a $f(A)v_j$ és az $f(A)w_j$ vektoroknak azonos a kis minimálpolinomjuk a $j = 1, \dots, k$ mellett, így az $\text{lin}(f(A)v_j; A)$ és a $\text{lin}(f(A)w_j; A)$ alterek izomorfak. Ez csak úgy lehetséges, ha

$$\text{lin}(f(A)w_{k+1}; A) \oplus \dots \oplus \text{lin}(f(A)w_s; A) = \{0\}$$

teljesül, speciálisan $f(A)w_{k+1} = 0$, ergo $q_{k+1}|f$, ergo $q_{k+1}|p_{k+1}$. Jelölje most $f = q_{k+1}$ és ismételjük meg a fenti gondolatot a v -vel és w -vel jelölt vektorok felcseréléssel. Így

$$\begin{aligned} f(A)(V) &= f(M_0) \oplus \text{lin}(f(A)w_1; A) \oplus \dots \oplus \text{lin}(f(A)w_k; A) \text{ és} \\ f(A)(V) &= f(M_0) \oplus \text{lin}(f(A)v_1; A) \oplus \dots \oplus \text{lin}(f(A)v_k; A) \oplus \\ &\quad \text{lin}(f(A)v_{k+1}; A) \oplus \dots \oplus \text{lin}(f(A)v_r; A), \end{aligned}$$

ami fenti érvveléssel párhuzamosan csak akkor lehetséges, ha

$$\text{lin}(f(A)v_{k+1}; A) \oplus \dots \oplus \text{lin}(f(A)v_r; A) = \{0\}$$

is teljesül. Speciálisan $f(A)v_{k+1} = 0$, ergo $p_{k+1}|f$, ergo $p_{k+1}|q_{k+1}$ is fennáll.

Az indukció szerint tehát $k = r$ mellett is igaz az állításunk, tehát azt mutattuk meg, hogy $r \leq s$ és minden $k = 1, \dots, r$ mellett $p_k = q_k$. Az állítás feltételeinek szimmetriáját használva az egész eddigi érvvelés megismételhetnénk, a v -vel és w -vel jelölt vektor rendszer felcseréléssel. Ekkor azt kapjuk, hogy $s \leq r$ is fennáll, tehát $r = s$ valóban teljesül, és persze minden $j = 1, \dots, r$ mellett $p_j = q_j$, amiből $\dim(\text{lin}(v_j; A)) = \deg p_j = \deg q_j = \dim(\text{lin}(w_j; A))$ is következik. Ezt kellett belátni. \square

B.15. állítás (Cayley–Hamilton II.). Legyen $A \in L(V)$ az \mathbb{F} test feletti V véges dimenziós vektortér lineáris transzformációja. Jelölje $m(t), k(t) \in \mathbb{F}[t]$ a transzformáció minimálpolinomját és a karakterisztikus polinomját. Ekkor

1. a minimálpolinom osztója a karakterisztikus polinomnak;
2. a minimálpolinom és a karakterisztikus polinom irreducibilis osztói azonosak;
3. ha a minimálpolinomnak és a karakterisztikus polinomnak a primtényezős előállítása

$$m(t) = f_1^{\alpha_1}(t) \cdot \dots \cdot f_r^{\alpha_r}(t) \quad \text{és} \quad k(t) = f_1^{\beta_1}(t) \cdot \dots \cdot f_r^{\beta_r}(t),$$

akkor minden $j = 1, \dots, r$ mellett a $\nu(f_j^{\alpha_j}(A))$ defektus és az $\deg f_j$ fok közötti összefüggés:

$$\beta_j = \frac{\nu(f_j^{\alpha_j}(A))}{\deg f_j}.$$

Bizonyítás: A ciklikus felbontási tételek értelmében a tér előáll

$$V = \text{lin}(v_1; A) \oplus \dots \oplus \text{lin}(v_s; A)$$

alakban ahol p_j a v_j vektorhoz tartozó kis minimálpolinom. Tudjuk, hogy $p_1 = m$ a minimálpolinom, és minden $j = 2, \dots, s$ mellett $p_j|p_{j-1}$.

1. Az is világos, hogy ha megszorítjuk a transzformációt a $\text{lin}(v_j; A)$ invariáns altérre, akkor ennek a megszorított transzformációnak a minimálpolinomja egybeesik a karakterisztikus polinomjával, és pedig ez éppen a p_j kis minimálpolinom. A determináns definíciója miatt a karakterisztikus polinomja A -nak az egyes invariáns alterekben vett karakterisztikus polinomok szorzata, azaz

$$k(t) = m(t) \cdot p_2(t) \cdot \dots \cdot p_s(t). \quad (\dagger)$$

Így persze a minimálpolinom valóban osztója a karakterisztikus polinomnak.

2. Ebből azonnal nyilvánvaló, hogy a minimálpolinom minden osztója a karakterisztikus polinomnak is osztója. Most tegyük fel, hogy $f(t)$ irreducibilis polinom osztója a karakterisztikus polinomnak. Mivel irreducibilis polinom prim tulajdonságú is, ezért (\dagger) felírást alkalmazva $f(t) | p_j(t)$ valamely j -re, no de a ciklikus felbontási tétele szerint bármelyik j mellett is $p_j(t) | m(t)$.

Meggondoltuk tehát, hogy $m(t)$ és $k(t)$ primtényezős felbontásában azonosak az irreducibilis polinomok, így ezek multiplicitásában lehet csak különbség, és persze $\alpha_j \leq \beta_j$.

3. A minimálpolinom faktorizációja szerint is szétesik a tér invariáns alterek direktösszegére. Jelölje $V_j = \ker f_j^{\alpha_j}(A)$ minden $j = 1, \dots, r$ mellett, így $V = V_1 \oplus \dots \oplus V_r$, és ha az A transzformációt visszaszorítjuk a V_j invariáns altérre, akkor ennek a megszorított transzformációnak $f_j^{\alpha_j}(t)$ lesz a minimálpolinomja. A 2. állítást már igazoltuk, így $A|V_j$ transzformáció minimálpolinomja és karakterisztikus polinomja is csak azonos irreducibilis polinomokat tartalmazhat, ezért a megszorított transzformáció karakterisztikus polinomja

$$f_j^{\gamma_j}(t), \quad \text{ahol} \quad \gamma_j \deg f_j = \dim(V_j) = \nu(f_j^{\alpha_j}(A)),$$

hiszen egy karakterisztikus polinom foka mindenkor a tér dimenziója. Újra alkalmazva, hogy a karakterisztikus polinom az egyes invariáns alterekre szorított transzformációk karakterisztikus polinomjainak szorzata azt kapjuk, hogy

$$k(t) = f_1^{\gamma_1}(t) \cdot \dots \cdot f_r^{\gamma_r}(t).$$

Persze a primtényezős előállítás egyértelmű, ezért minden szóba jövő j indexre $\beta_j = \gamma_j$. \square

C. függelék

Kvadratikus alakok definitségének egy naiv targyalása

A kvadratikus alaknak mondunk egy speciális többváltozós függvényt.

C.1. definíció. Legyen $A \in \mathbb{R}^{n \times n}$ egy szimmetrikus mátrix. Amint az szokásos jelölje $a_{k,j}$ a mátrix i, j pozíciójának elemét. Definiálja a $Q_A : \mathbb{R}^n \rightarrow \mathbb{R}$ függvényt

$$Q(x_1, x_2, \dots, x_n) = \sum_{k=1}^n \sum_{j=1}^n a_{k,j} x_k x_j.$$

Ekkor Q_A az A szimmetrikus mátrixhoz tartozó kvadratikus alak.

Például $n = 1$ és az A mátrix egyetlen eleme az $a \in \mathbb{R}$ szám, akkor $Q(x_1) = ax_1^2$. Ha $n = 2$ mellett $A \in \mathbb{R}^{2 \times 2}$, akkor

$$Q(x_1, x_2) = a_{1,1}x_1^2 + a_{1,2}x_1x_2 + a_{2,1}x_2x_1 + a_{2,2}x_2^2.$$

Gyakorláscléppen írjuk fel minden három változós kvadratikus alakok lehetségesek.

C.2. állítás. Legyen $A \in \mathbb{R}^{n \times n}$ egy szimmetrikus mátrix. Ekkor az A által meghatározott kvadratikus alakra

$$Q(x) = \langle x, Ax \rangle$$

tetszőleges $x \in \mathbb{R}^n$.

Ebből az alakból már világos, hogy az $A + B$ összeghez tartozó kvadratikus alak azonos az A -hoz és a B -hez tartozó kvadratikus alakok összegével, és hasonlóan az αA mátrixhoz tartozó kvadratikus alakot úgy is megkaphatjuk, hogy az A -hoz tartozó kvadratikus alakot mint n -változós függvényt szorozzuk az α számmal. Formálisabban:

C.3. állítás. Legyenek $A, B \in \mathbb{R}^{n \times n}$ szimmetrikus mátrixok és $\alpha, \beta \in \mathbb{R}$. Ekkor

$$Q_{\alpha A + \beta B} = \alpha Q_A + \beta Q_B.$$

A nemzérus kvadratikus alakokat az értékkészletük alapján osztályozzuk.

C.4. definíció. Legyen $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ egy nemzérus kvadratikus alak. Azt mondjuk, hogy ez

pozitív definit, ha $Q(x) > 0$ minden $x \in \mathbb{R}^n, x \neq 0$;

pozitív szemidefinit, ha $Q(x) \geq 0$ minden $x \in \mathbb{R}^n$, de létezik egy $z \in \mathbb{R}^n, z \neq 0$ vektor, amelyre $Q(z) = 0$;

negatív definit, ha $Q(x) < 0$ minden $x \in \mathbb{R}^n, x \neq 0$;

negatív szemidefinit, ha $Q(x) \leq 0$ minden $x \in \mathbb{R}^n$, de létezik egy $z \in \mathbb{R}^n, z \neq 0$ vektor, amelyre $Q(z) = 0$;

indefinit, ha vannak olyan $x, y \in \mathbb{R}^n$ vektorok, amelyekre $Q(x) > 0$ és $Q(y) < 0$.

A szimmetrikus diádok szerepe

Emlékezzünk, hogy $n \times n$ méretű szimmetrikus diádnak nevezünk egy \mathbb{R}^n beli vektornak oszlopakú és soraik mátrixszorozatát. Kicsit formálisabban az $A = a \cdot a^T$ tetszőleges $a \in \mathbb{R}^{n \times 1}$ mellett. A diádok tehát speciális szimmetrikus mátrixok. Látható, hogy egy szimmetrikus mátrix pontosan akkor diád, ha az 1 rangú szimmetrikus mátrix.

A szimmetrikus diádok azért érdekesek számunkra, mert éppen az ezekhez tartozó kvadratikus alakok a teljes négyzetek.

C.5. állítás. *Egy kvadratikus alak pontosan akkor teljes négyzet, ha az őt reprezentáló szimmetrikus mátrix egy diád. Pontosabban: Ha $A = a \cdot a^T$, ahol $a = (a_1, a_2, \dots, a_n)$, akkor $Q_A(x_1, x_2, \dots, x_n) = (a_1 x_1 + a_2 x_2 + \dots + a_n x_n)^2$.*

C.1. Szimmetrikus mátrix diád felbontása

Be fogjuk bizonyítani, hogy

C.6. állítás. *Minden nem zérus szimmetrikus mátrix előáll mint szimmetrikus diádok lineáris kombinációja. Sőt, ha egy szimmetrikus mátrix rangja r , akkor az előáll mint r db. szimmetrikus diád lineáris kombinációja, ahol a diádokat alkotó vektorok rendszere lineárisan független rendszert alkot.*

A bizonyítás úgy folyik, hogy konkrét algoritmust adunk a felbontásra.

Az algoritmus első mozzanata annak észrevétele, hogy egy szimmetrikus mátrixból leválasztható egy szimmetrikus diád olyan módon, hogy a maradék szimmetrikus mátrixhoz tartozó kvadratikus alak már egy változóval kevesebbét tartalmazzon:

C.7. állítás. *Tekintsünk egy $A \in \mathbb{R}^{n \times n}$ szimmetrikus mátrixot. Tegyük fel, hogy a balfelső elemre $a_{1,1} \neq 0$. Jelölje $d = \frac{1}{a_{1,1}}a_1$, ahol a_1 a mátrix első oszlopa A . Ekkor*

$$R = A - a_{1,1}d \cdot d^T$$

egy olyan szimmetrikus mátrix, ahol az első oszlop és első sor kizárolag zérus elemeket tartalmaz.

Bizonyítás: Az persze nyilvánvaló, hogy szimmetrikus mátrixok egy lináris kombinációja is szimmetrikus marad. Jelölje $\delta_k = \frac{a_{k,1}}{a_{1,1}} = \frac{a_{1,k}}{a_{1,1}}$ a d vektor k -adik koordinátáját. Ekkor

$$r_{k,j} = a_{k,j} - a_{1,1}\delta_k\delta_j = a_{k,j} - a_{1,1}\frac{a_{k,1}}{a_{1,1}}\delta_j = a_{k,j} - a_{k,1}\delta_j. \quad (\dagger)$$

Ha $j = 1$, akkor itt $r_{k,1} = a_{k,1} - a_{k,1}\delta_1 = 0$. Látjuk, hogy az R mátrix első oszlopa csupa zérust tartalmaz. Íme, az R egy szimmetrikus mátrix, ergo az első sorban is csak zérus elemek vannak. \square

Illusztrációként nézzük a következő kvadratikus alakot:

$$Q(z_1, z_2, z_3) = 2z_1^2 + \frac{3}{2}z_3^2 + 2z_1z_2 - 4z_1z_3 + 2z_2z_3.$$

Ezt a kvadratikus alakot az $A = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & 1 \\ -2 & 1 & \frac{3}{2} \end{pmatrix}$ szimmetrikus mátrix reprezentálja, legyen tehát $d = \begin{bmatrix} 1 \\ \frac{1}{2} \\ -1 \end{bmatrix}$.

Így

$$R = A - 2d \cdot d^T = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & 1 \\ -2 & 1 & \frac{3}{2} \end{pmatrix} - 2 \begin{bmatrix} 1 \\ \frac{1}{2} \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \frac{1}{2} & -1 \end{bmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 2 \\ 0 & 2 & -\frac{1}{2} \end{pmatrix}$$

No persze $A = 2d \cdot d^T + R$. Áttérve a kvadratikus alakokra $Q(z_1, z_2, z_3) = 2(z_1 + \frac{1}{2}z_2 - z_3)^2 + Q_R(z_1, z_2, z_3)$, ahol Q_R már nem függ a z_1 változótól, hiszen $Q_R(z_1, z_2, z_3) = -\frac{1}{2}z_2^2 - \frac{1}{2}z_3^2 + 4z_2z_3$.

Sikeresen megszabadultunk az első változótól, elegendő tehát a $\begin{pmatrix} -\frac{1}{2} & 2 \\ 2 & -\frac{1}{2} \end{pmatrix}$ mátrix reprezentálta kétváltozós

kvadratikus alakkal foglalkozni, amelynek változóit most z_2 és z_3 jelöli. Válasszunk le egy újabb szimmetrikus diádot:

$$\begin{pmatrix} -\frac{1}{2} & 2 \\ 2 & -\frac{1}{2} \end{pmatrix} - \begin{pmatrix} -\frac{1}{2} \end{pmatrix} \begin{bmatrix} 1 \\ -4 \end{bmatrix} \cdot [1 \ -4] = \begin{pmatrix} 0 & 0 \\ 0 & \frac{15}{2} \end{pmatrix}.$$

Lefordítva ezt a kvadratikus alakok nyelvére azt kaptuk, hogy $Q_R(z_1, z_2, z_3) = -\frac{1}{2}(z_2 - 4z_3)^2 + \frac{15}{2}z_3^2$. Összefoglalva tehát

$$Q(z_1, z_2, z_3) = 2 \left(z_1 + \frac{1}{2}z_2 - z_3 \right)^2 - \frac{1}{2}(z_2 - 4z_3)^2 + \frac{15}{2}z_3^2.$$

Az A mátrix diádfelbontását is leolvashatjuk a fenti számolásból:

$$A = 2 \begin{bmatrix} 1 \\ \frac{1}{2} \\ -1 \end{bmatrix} \cdot [1 \ \frac{1}{2} \ -1] - \frac{1}{2} \begin{bmatrix} 0 \\ 1 \\ -4 \end{bmatrix} \cdot [0 \ 1 \ -4] + \frac{15}{2} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot [0 \ 0 \ 1].$$

Voltaképpen készen is lennénk, mert ez a módszer minden olyan esetben megadja a szimmetrikus diádokra való felbontást, mikor az algoritmus kezdő és minden további lépésein a maradék kvadratikus alak mátrixának balfelső eleme nem zérus.¹

A maradék kvadratikus alak mátrixa, az $R = A - a_{1,1}d \cdot d^T$, a fentinél sokkal hatékonyabban számolható. Nézzük rá az állítás igazolásában kiemelt (†) azonosságra de úgy, hogy gondoljunk közben a Gauss–Jordan-eliminációra:

$$r_{k,j} = a_{k,j} - a_{k,1} \delta_j. \quad (\dagger)$$

Tudjuk, hogy az R maradék első oszlopa és első sora zérus. Ezért a kérdés csak az, hogy e mátrix jobb alsó $n - 1 \times n - 1$ méretű részének mik az elemei. A fenti formulát kell ehhez alkalmazni, mikor $k > 1$ és $j > 1$. Ismerjük fel, hogy (†) éppen az eliminációs algoritmus számolási szabálya abban az esetben, amikor a balfelső elem a pivot elem és a pivot elemet nem tartalmazó sorok elemeit számoljuk. Ez azt jelenti, hogy az $A - a_{1,1}d \cdot d^T$ művelet diád szorzás és mátrix kivonás helyett Gauss–Jordan-eliminációval is számolható. Bebizonyítottuk hát a következő állítást.

C.8. állítás. Legyen $A \in \mathbb{R}^{n \times n}$ egy olyan szimmetrikus mátrix, amelyre $a_{1,1} \neq 0$. Legyen R az első diád leválasztása után maradó kvadratikus alak mátrixa, azaz

$$R = A - a_{1,1}d \cdot d^T,$$

ahol d az A mátrix első oszlopának az $a_{1,1}$ reciprokával vett szorzata. Tudjuk, hogy R első oszlopa és sora csak zérus elemeket tartalmaz, viszont e mátrix a jobb alsó $(n - 1 \times n - 1)$ méretű része megkapható az A mátrixon végzett Gauss–Jordan-eliminációval, olyan módon hogy a balfelső $a_{1,1}$ elemet választjuk pivot elemnek, és az elimináció első sorát eldobjuk.

Ennek fényében az előző példa számolása radikálisan egyszerűsíthető. Nézzük újra.

$$\begin{array}{c|ccc} & z_1 & z_2 & z_3 \\ \hline \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} & 1 & -2 \\ 1 & 0 & 1 \\ -2 & 1 & 3/2 \\ \hline \delta & \frac{1}{2} & -1 \end{array} \Rightarrow \begin{array}{c|ccc} & z_1 & z_2 & z_3 \\ \hline \begin{bmatrix} -1/2 \\ 2 \\ -1/2 \end{bmatrix} & 2 & -1/2 \\ 2 & \delta & -4 \\ \hline \end{array} \Rightarrow \begin{array}{c|ccc} & z_1 & z_2 & z_3 \\ \hline & & & \boxed{15/2} \end{array}$$

Ezután csak értelmeznünk kell a fenti eliminációt. A második táblázat az elő diád leválasztása után maradt kvadratikus alak mátrixa. A leválasztott diád az első táblázat segédsorából olvasható le, és a diád szorzója az első táblázatban választott pivot elem. Tehát

$$\begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & 1 \\ -2 & 1 & \frac{3}{2} \end{pmatrix} = 2 \begin{bmatrix} 1 \\ 1/2 \\ -1 \end{bmatrix} \cdot [1 \ 1/2 \ -1] + \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1/2 & 2 \\ 0 & 2 & -1/2 \end{pmatrix}.$$

¹Evvel a problémával most ne foglalkozzunk később ki fog derülni, hogy könnyen megszabadulhatunk e kellemetlennek tűnő feltétetetlől.

A harmadik táblázat az előző maradékból leválasztott kvadratikus alak mátrixa. A leválasztott diád a második táblázat segédsorából olvasható le, és a szorzó a második táblázat pivot eleme. Tehát

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1/2 & 2 \\ 0 & 2 & -1/2 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix} \cdot [0 \ 1 \ -4] + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 15/2 \end{pmatrix}$$

Világos, hogy az utolsó maradék $15/2$ szerese a $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ vektor által meghatározott diádnak, tehát a diád felbontás a fenti eliminációs táblázatból azonnal leolvasható:

$$A = 2 \begin{pmatrix} 1 \\ \frac{1}{2} \\ -1 \end{pmatrix} \cdot [1 \ \frac{1}{2} \ -1] - \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix} \cdot [0 \ 1 \ -4] + \frac{15}{2} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot [0 \ 0 \ 1]$$

A diád felbontás a kvadratikus alakra nézve azt jelenti, hogy előállítottuk azt olyan teljes négyzetek lineáris kombinációjaként, ahol a teljes négyzeteknek megfelelő diádokat generáló vektorok rendszere linárisan független rendszert ad. Konkrétan itt

$$2z_1^2 + \frac{3}{2}z_3^2 + 2z_1z_2 - 4z_1z_3 + 2z_2z_3 = 2 \left(z_1 + \frac{1}{2}z_2 - z_3 \right)^2 - \frac{1}{2} (z_2 - 4z_3)^2 + \frac{15}{2}z_3^2.$$

Gyakorlat. Írjuk fel a $Q(x_1, x_2, x_3) = x_1^2 + 3x_2^2 + 6x_3^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3$ kvadratikus alakot teljesnégyzetek lineáris kombinációjaként. Az elimináció lépései:

$$\begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ \hline \boxed{1} & 1 & 1 & 1 \\ 1 & 3 & 3 & 3 \\ 1 & 3 & 6 & 6 \\ \hline \delta & 1 & 1 & 1 \end{array} \Rightarrow \begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ \hline \boxed{2} & 2 & 2 & 5 \\ 2 & 2 & 5 & 5 \\ \hline \delta & 1 & 1 & 1 \end{array} \Rightarrow \begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ \hline & & & \\ & & & \\ & & & \\ \hline & & & \boxed{3} \end{array}$$

Ez azt jelenti, hogy $Q(x_1, x_2, x_3) = (x_1 + x_2 + x_3)^2 + 2(x_2 + x_3)^2 + 3x_3^2$. Az is világos, hogy a három pivot elem pozitivitása szerint a kvadratikus alak pozitív definit, hiszen az

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_2 + x_3 &= 0 \\ x_3 &= 0 \end{aligned}$$

homogén lineáris egyenletrendszernek, az együttható mátrix oszlopainak lináris függetlensége miatt, csak triviális megoldása van!

Zérus elem a bal felső sarokban

Hangsúlyoznunk kell, hogy az algoritmusban a pivot elem megválasztása nem opcionális. A soron következő diád leválasztásához minden a balfelső elemet kell választanunk. Persze kérdés mi történik, ha a balfelső elem zérus? Ennek illusztrációjaként nézzük az alábbi kvadratikus alakot.

$$Q(x_1, x_2, x_3) = 3x_2^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3$$

Persze el sem tudjuk kezdeni az algoritmust, hiszen a balfelső elem zérus. Azt vegyük észre, hogy át tudjuk címkézni a változókat úgy, hogy a kapott kvadratikus alakban az első változónak legyen négyzete. Ehhez csak az x_1 és az x_2 változókat kell felcserélnünk. Vezessük be tehát a $z_1 = x_2$, $z_2 = x_1$ és $z_3 = x_3$ jelöléseket. Ezek szerint

$$3x_2^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3 = 3z_1^2 + 2z_2z_1 + 2z_2z_3 + 6z_1z_3$$

Így viszont már indulhat is az elimináció:

$$\begin{array}{|ccc|} \hline & z_1 & z_2 & z_3 \\ \hline 3 & 1 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 0 \\ \hline \delta & 1/3 & 1 \end{array} \Rightarrow \begin{array}{|ccc|} \hline & z_1 & z_2 & z_3 \\ \hline -1/3 & 0 \\ 0 & -3 \\ \hline \delta & 0 & 0 \end{array} \Rightarrow \begin{array}{|ccc|} \hline & z_1 & z_2 & z_3 \\ \hline & & & -3 \\ \hline \end{array}$$

A teljes négyzetek lineáris kombinációját leolvasva:

$$3z_1^2 + 2z_2z_1 + 2z_2z_3 + 6z_1z_3 = 3\left(z_1 + \frac{1}{3}z_2 + z_3\right)^2 - \frac{1}{3}z_2^2 - 3z_3^2 = 3\left(\frac{1}{3}x_1 + x_2 + x_3\right)^2 - \frac{1}{3}x_2^2 - 3x_3^2,$$

ami a diádfelbontásra nézve azt jelenti, hogy

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 0 \end{pmatrix} = 3 \begin{pmatrix} 1/3 \\ 1 \\ 1 \end{pmatrix} \cdot [1/3 \ 1 \ 1] - \frac{1}{3} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot [1 \ 0 \ 0] - 3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot [0 \ 0 \ 1].$$

Ha ezt az átcímkézési trükköt megértettük, akkor még egyszerűbb jelölésekkel is alkalmazhatunk. Végül is csak az volt a lényeg, hogy az első két változót felcserélünk. Jelezhetjük ezt a táblázatban a változók indexével is. Ugyanez méggyakrabban:

$$\begin{array}{|ccc|} \hline & x_1 & x_2 & x_3 \\ \hline 0 & 1 & 1 \\ 1 & 3 & 3 \\ 0 & 3 & 0 \\ \hline \end{array} \Rightarrow \begin{array}{|ccc|} \hline & x_2 & x_1 & x_3 \\ \hline 3 & 1 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 0 \\ \hline \delta & 1/3 & 1 \end{array} \Rightarrow \begin{array}{|ccc|} \hline & x_2 & x_1 & x_3 \\ \hline -1/3 & 0 \\ 0 & -3 \\ \hline \delta & 0 & 0 \end{array} \Rightarrow \begin{array}{|ccc|} \hline & x_2 & x_1 & x_3 \\ \hline & & & -3 \\ \hline \end{array}$$

A változók indexére ügyelve kell leolvasnunk a teljesnégyzeteket, és a legegszerűbb, ha a teljesnégyzetekből olvassuk le a diád felbontást.

Már csak azt kell meggyondolnunk, hogyan kaptuk az első táblázatból a másodikat? Azt kell látnunk, hogy az i -edik és a j -edik változó felcserélése avval jár, hogy a mátrixban felcseréljük az i -edik sort a j -edik sorral és az i -edik oszlopot a j -edik oszloppal. Persze, hiszen a mátrixban minden az i, k és a j, k pozícióban álló szám helyet cserél, és hasonlóan minden a k, i és a k, j pozícióban álló szám is helyet cserél minden k mellett. A többi elem viszont nem változik. Az első rész az i és j indexű sorok felcserélését jelenti, míg a második rész az i és j indexű oszlopok kicsereleléséből áll. A többi sort és oszlopot nem éri változás.

Látjuk tehát, hogy ha a balfelső elem zérus, de a diagonálisnak van nem zérus eleme, akkor a most tanult felcseréléssel a nem zérus elem a balfelső pozícióba hozható, így az algoritmusunk továbbra is működik.

A diagonálisban minden elem zérus

Most nézzük azt az esetet, amikor a diagonális minden eleme zérus, de van nem zéró elem.

Először is azt vegyük észre, hogy ha a feladat a kvadratikus alak definitségének meghatározása, akkor készen is vagyunk mert indefinite kvadratikus alakkal állunk szemben. Ha ugyanis a k, l pozíció tartalmaz nem zérus elemet, akkor a kvadratikus alak összes többi változóját állítuk zérusra és a k -adik és az l -edik változót vizsgáljuk. Így a kvadratikus alakunk az $2a_{k,l}x_kx_l$ kifejezésre egyszerűsödik, amely persze indefinite, hiszen az értékkészlete negatív és pozitív elemet egyaránt tartalmaz.

A teljesség kedvéért megmutatjuk, hogy az algoritmusunk ebben az esetben is használható. Ha az $a_{k,l} \neq 0$, akkor vezessük be a

$$z_l = x_l + x_k \quad z_j = x_j, \text{ ha } j \neq l$$

változókat. Ekkor $k \neq l$ -et is figyelembe véve

$$z_kz_l = x_k(x_l + x_k) = x_kx_l + x_k^2 = z_k^2 + x_kx_l, \text{ majd } j \neq k, j \neq l \text{ mellett:}$$

$$z_jz_l = x_j(x_l + x_k) = x_jx_l + x_jx_k = x_jx_l + z_jz_k.$$

Ez azt jelenti, hogy a x_kx_l kifejezést a $(z_kz_l - z_jz_k)$ -re kell cserélnünk, és minden további x_jx_l alakú kifejezést a $(z_jz_l - z_kz_k)$ kifejezéssel kell kicserelnünk. A többi változó indexe nem változik. Így pontsan egy négyzetet

kapunk, tehát a mátrix diagonálisában a k -adik sorban jelenik meg egy nem zérus elem. Ahogyan azt az előzőekben már megértettük ezt az elemet kell a balfelső pozícióba csempészni, és az algoritmusunk már működik is.

C.2. Definitség osztályozása a teljes négyzetek segítségével

Miután elkészült a kvadratikus alakot reprezentáló mátrix diádfelbontása, a kvadratikus alak definitsége az alábbiak szerint osztályozható. Legyen tehát Q egy n változós nemzérus kvadratikus alak.

Ha a diák felbontás

- tartalmaz pozitív és negatív pivot elemet, akkor Q *indefinit*.

Most tegyük fel, hogy Q felbontása pontosan n teljes négyzetet tartalmaz. Ha

- az összes pivot elem pozitív, akkor Q *pozitív definit*;
- az összes pivot elem negatív, akkor Q *negatív definit*;

Ha Q felbontása n -nél kevesebb teljes négyzetből áll, akkor nagyon hasonló karakterizációt kapunk, de a mátrix szinguláris, ezért csak szemidefinitseg jöhét szóba. Tehát ha

- az összes pivot elem pozitív, akkor Q *pozitív szemidefinit*;
- az összes pivot elem negatív, akkor Q *negatív szemidefinit*;

Irodalom

- Dancs, István és Csaba Puskás (2001). *Vektorterek*. Aula kiadó 2001, Budapest, ISBN:963 9345 53 9, BCE Catalogue: bcek.379187. (hiv. old. 5).
- Derkzen, Harm (2003). „The Fundamental Theorem of Algebra and Linear Algebra”. *The American Mathematical Monthly* 110.7, 620–623. old. ISSN: 00029890, 19300972. URL: <http://www.jstor.org/stable/3647746> (hiv. old. 171).
- Ebbinghaus, H.-D. és tsai. (1991). *Numbers*. 123. köt. Graduate Texts in Mathematics. With an introduction by K. Lamotke, Translated from the second 1988 German edition by H. L. S. Orde, Translation edited and with a preface by J. H. Ewing, Readings in Mathematics. Springer-Verlag, New York, xviii+395. old. ISBN: 0-387-97497-0. doi: 10.1007/978-1-4612-1005-4. URL: <http://dx.doi.org/10.1007/978-1-4612-1005-4> (hiv. old. 25, 171).

Tárgymutató

- 2 karakteristikájú test, 122
Abel-csoport, 12
abszolútérték, 27
adjungált, 145
adjungált azonosság, 145
affin halmaz, 61, 67
affin kombináció, 61
algebrai struktúra, 11
alternáló operátor, 122
altér, 34
anti-szimmetrikus operátor, 122
asszociatív, 11
áttérés transzformáció, 81
bástyafelrakás, 119
bázis, 53
belő szorzat, 135
Bessel-egyenlőtlenség, 137, 140, 141, 156
Bezout-azonosság, 17, 19
Cayley – Hamilton, 106, 117, 185
 C^* -azonosság, 166
 C^* -azonosság, 28
co-dimenzió, 69
Cramer-szabály, 131
csoport, 11
defektus, 77
determináns, 125, 128
determinánsok szorzattétele, 128
diád, 23, 52
diagonalizálhatóság, 93, 94
diagonális alakú mátrix, 94
diagonális mátrix, 24
direkt kiegészítő, 65, 69
direkt összeg, 63
diszjunkt alterek, 63
disztributív, 12
duális bázis, 123, 141
egyszerűsítési szabály, 12
együttható-mátrix, 37
elemi mátrix, 43
Euklideszi-algoritmus, 19
faktortér, 68
feszítőrang, 52
formális algebrai kifejezés, 14
félcsoport, 11
fődiagonális, 24
főegyüttható, 14
főideál, 13, 88, 91, 92
főideál-gyűrű, 13, 88, 91
Gauss – Jordan-elimináció, 25, 36–39, 41–44
generált altér, 34
generált ideál, 13
generátorrendszer, 35
geometriai multiplicitás, 95
Gram – Schmidt-ortogonalizáció, 140, 153, 154
gyökök multiplicitása, 17
gyűrű, 12
gyűrű-izomorfizmus, 79
homogén lineáris egyenletrendszer, 37
háromszög-egyenlőtlenség, 28
 i komplex szám, 26
identitás mátrix, 23
ideál, 13
inhomogén lineáris egyenletrendszer, 37
invariáns altér, 83
invertálható mátrix, 43
inverz, 12
inverzió, 121
irreducibilis polinom, 20
irányító polinom, 179
izomorf vektorterek, 58
izomorfizmus, 25, 58
Jordan-blokk, 115
Jordan-bázis, 117
Jordan-normálalak, 99, 115, 132
karakterisztikus polinom, 117, 131
kis minimálpolinom, 88, 92, 93, 98, 101, 107, 133, 171
kofaktor, 130
kofaktormátrix, 130
kommutatív, 12
kommutál, 44, 87, 171, 173
kommutáló mátrixok, 24, 25, 44
kommutáló transzformációk, 87

komplex n -edik egységgöök, 29
 komplex szám argumentuma, 29
 komplex szám konjugáltja, 27
 komplex szám képzetes része, 27
 komplex szám normálalakja, 27
 komplex szám trigonometrikus alakja, 29
 komplex szám valós része, 27
 komplex számok, 26
 komplex számok mátrix reprezentációja, 26
 komplex számtest feletti vektortér, 99
 konjugáltan lineáris, 135
 koordináta, 58
 kötött változó, 38
 Kronecker-delta, 23
 legkisebb közös többszörös, 16, 92
 legközelebbi pont, 138
 legnagyobb közös osztó, 16
 legszűkebb invariáns altér, 83
 lineáris burok, 34
 lineáris funkcionál, 57
 lineáris kombináció, 22, 34
 lineáris operáció, 57
 lineáris operáció mátrixa, 76
 lineáris transzformáció, 57
 lineáris transzformáció determinánsa, 128
 lineáris transzformációk szorzata, 78
 lineárisan független rendszer, 46
 lineárisan összefüggő rendszer, 45
 mátrixok szorzata, 22
 mátrixok összege, 21
 maximális lineárisan független rendszer, 47
 megengedhető altér, 179
 merőleges, 136
 minimális generátorrendszer, 47
 minimálpolinom, 91
 Minkowski-összeg, 61, 62, 67, 98
 minor, 130
 Moivre-formula, 29
 mátrix determinánsa, 125
 mérték, 123, 127
 művelet, 11
 n-lineáris operátor, 122
 neutrális elem, 11
 neutrális elemes félcsoport, 11
 nilpotens transzformáció, 99, 110
 normális transzformáció, 151, 157
 normált polinom, 14
 o.n.b, 140
 operátornorma, 165
 ortogonális projekció, 155
 ortokomplementer, 136
 oszlopvektor, 22
 oszlopvektortér, 40
 önadjungált transzformáció, 148, 158
 páratlan permutáció, 121
 páros permutáció, 121
 permutáció csoport, 119
 permutáció előjele, 121
 permutáció mátrix, 120
 permutációk, 12
 pivot elem, 37, 38, 41, 126
 polinom, 13
 polinom foka, 14
 polinom osztója, 15
 polinom többszöröse, 15
 projekció, 155
 projekciós-tétel, 141
 prím polinom, 20
 Pythagoras-tétel, 138
 rang, 55, 77
 rang-defektus-tétel, 147
 Rang-defektus-tétel, 70
 reducibilis polinom, 20
 reguláris mátrix, 78
 relatív prím polinomok, 16
 Riesz-reprezentáció, 145
 sajátaltér, 85, 88
 sajátvektor, 85, 93
 sajátérték, 85, 93
 sajátérték algebrai multiplicitása, 117, 132
 sajátérték geometriai dimenziója, 116
 Schwartz-egyenlőtlenség, 28
 skaláris szorzat, 135
 skaláriszorzatos-tér, 135
 sorvektor, 22
 sorvektortér, 40
 spektrum, 85
 spektrálisugár, 165
 Steinitz-lemma, 35, 51–54, 84, 88, 91
 szabad változó, 38
 szimmetrikus diád, 23
 szinguláris, 93
 szorzattétel, 128
 szám és mátrix szorzata, 22
 triviális altér, 34
 triviális lineáris kombináció, 47
 triviális megoldás, 39
 unitér, 151
 unitér transzformáció, 148, 158
 vektortér, 22, 33
 vektortér altere, 34
 véges dimenziós, 53
 végesen generált vektortér, 35, 53