Verzió információk

References	GitHub: March 7, 2020
Branch	GitHub: March 7, 2020
Dirty	
Hash	GitHub: March 7, 2020
Author Iso Date	GitHub: March 7, 2020
First Tag Describe	GitHub: March 7, 2020
Reln	GitHub: March 7, 2020
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, 2020. március 7-én

Tartalomjegyzék

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

$$\varphi\colon H^n \to H$$

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

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

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

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

$$a \star (b \star c) = (a \star b) \star c.$$ ⌟

Lefordítva ez azt jelenti, hogy

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

1.3. definíció (neutrális elem)**.** Az $(S, \star)$ félcsoportban az $s \in S$ elem *balról* (*jobbról*) *neutrális*, ha $s \star t = t$ ($t \star s = t$) 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 $s_1 = s_1 \star s_2 = s_2$. .

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

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

$$g \star g' = e = g' \star g. \qquad (\dagger)$$

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

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

Bizonyítás: Tegyük fel, hogy g', g'' inverz elemei g-nek. Azt mutatjuk meg, hogy ha g' baloldali– és g'' jobboldali inverze g-nek, akkor a két elem megegyezik:

$$g' = g' \star e = g' \star \big(g \star g''\big) = \big(g' \star g\big) \star g'' = e \star g'' = g''. \quad .$$

Példaként gondoljuk meg, hogy a $H \to H$ függvény halmaza a kompozíció művelettel neutrális elemes félcsoport, és a $H \to H$ kölcsönösen egyértelmű függvények halmaza a kompozíció művelettel csoportot alkotnak. Ez utóbbi csoportot mondjuk *permutáció csoportnak*.

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

$$a \star c = b \star c \implies a = b. \quad ⌟$$

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

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

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:

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

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 nemzérus[1] elemnek van inverze[2], és $0 \neq 1$[3]. ⌟

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

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

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

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 szerint ez éppen azt jelenti, hogy $-a = (-1) \cdot a$. .

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 gyűrűje[4], 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*

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

[1] értsd: minden elemnek, amely a $+$ műveletre nézve neutrális elemtől különbözik
[2] értsd: a $\cdot$ szorzás neutrális elemére mint egységelemre nézve
[3] érts: az összeadásra nézve és a szorzásra nézve képzett neutrális elemek nem azonosak.
[4] lásd kicsit később

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

$$0 = 0b' = (ab)\, b' = a\left(bb'\right) = a1 = a. \qquad \blacksquare$$

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

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

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

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

A generált ideál fogalma nagyon fontos.

1.16. definíció-állítás (generált ideál)**.** Legyen adott a kommutatív, egységelemes $(R, +, \cdot)$ gyűrűben véges sok $a_1, \ldots, a_r$ 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 $J(a_1, a_2, \ldots, a_r)$ módon.

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

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

az $a_1, \ldots, a_r$ elemeket tartalmazó legszűkebb ideál. Nevezzük ezt az ideált az $a_1, \ldots, a_r$ elemek *generálta ideálnak* is. $\lrcorner$

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

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

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

A fenti definícióban az *algebrai kifejezés* szó arra utal, hogy az $\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t^n$ 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 összes együtthatói azonosak. Ez szemben áll avval, hogy ha a polinomokra mint függvényekre tekintenénk, akkor a helyettesítési értékek egyenlősége jelentené a két polinom azonos voltát. A formális szó tehát azt jelenti, hogy nem mint függvényre gondolunk, hanem egyszerűen az adott $\alpha_0, \ldots, \alpha_n$ rögzített elemek – ezeket mondjuk együtthatóknak –, által előírt műveletekre. Az az előírás ugyanis, hogy tetszőleges $t \in \mathbb{F}$ mellett hajtsuk végre az

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

műveletsort. A műveletsorról és nem annak eredményéről van szó.

Két műveletsor akkor azonos, ha ugyanazok a műveletsort meghatározó $(\alpha_0, \alpha_1, \ldots, \alpha_n)$[5] együtthatók.[6] A jelölések megértése is fontos. $p(t) \in \mathbb{F}[t]$ semmi mást nem jelent, minthogy $p(t)$ egy polinom. Persze a polinom nem keverendő össze a helyettesítési értékével, hiszen az egyik egy algebrai kifejezés-együttes, 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 kiértékelnünk a polinomot ilyenkor $p(s)$ jelöli azt a testbeli elemet, amelyet t helyett s-et téve az előírt műveletek kiértékelése után kapunk. A szövegkörnyezetben mindig világosnak kell lennie, hogy $p(t)$ a polinomot jelenti, vagy egy konkrét t-re kiértékelt testbeli elemet.

1.18. definíció (polinom foka)**.** Legyen $p(t) \in \mathbb{F}[t]$ egy polinom. Azt mondjuk, hogy a n nem negatív egész szám e *polinom fokszáma*, ha n a legnagyobb nemzérus együttható indexe. A legnagyobb nemzérus 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 végzünk. 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.

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

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:

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

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

1. $\deg(pq) = \deg p + \deg q$;

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

[6] Persze 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 16. oldalon.

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. .

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: $p|f$. ⌟

Világos, hogy egy p polinom összes többszörösei – tehát azok, amelyeknek p osztója – ideált 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 $p|q$ és $q|p$ akkor a két polinom csak konstans szorzóban különbözik egymástól. Ha például a két polinom még normált is, akkor $p|q$ és $q|p$ csak $p = q$ esetben lehetséges. A polinomgyűrű ideáljaira fokuszálva, 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 közt is működik a maradékos osztás, ahogyan azt az egész számok közt 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

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

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 unicitás 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 figyelve

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

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. ▪

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. Így $\mathbb{F}[t]$ 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 meggondoltuk, 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. ▪

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 $p_1, \ldots, p_k$ polinomok.

1. A d polinom a *legnagyobb közös osztója* az adott polinomoknak, ha
   a) $d|p_j$ minden $j = 1, \ldots, k$-ra,
   b) ha $d_1|p_j$ minden $j = 1, \ldots, k$ mellett akkor $d_1|d$ is fennáll,
   c) d normált.

   A $p_1, \ldots, p_k$ 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 $j = 1, \ldots, k$-ra,
   b) ha $p_j|d_1$ minden $j = 1, \ldots, k$ 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 többszörösük.* ⌟

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 mindketten normáltak, ezért e konstans csak 1 lehet. .

1.27. állítás. *Bármely $p_1, \dots, p_r \in \mathbb{F}[t]$ nem zérus polinomoknak létezik egyetlen legnagyobb közös osztójuk. A d 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. .

A következő állítás azonosságát Bezout–azonosságnak mondjuk.

1.28. állítás (Bezout–azonosság)**.** *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 mindig kiemelhető, 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. .

1.30. definíció (gyök multiplicitása)**.** Legyen t_0 gyöke a $p(t)$ polinomnak. Azt mondjuk, hogy a k pozitív egész e 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 nullosztó mentessége és a gyöktényező kiemelhetősége miatt

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

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

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. ▪

A maradékos osztás tételének 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 többszörösét és legkisebb közös osztóját, akkor ugyan ezt mát meg tudnánk tenni véges sok 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. Véges sok lépésben végrehajtható, ezért a fentinél sokkal egyszerűbben működő algoritmust ad. Emlékezzünk 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*

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

Mivel minden egyes lépésben csökken a fokszám, ezért van olyan $s > 0$, 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 $p_{-1}, p_0, p_1, \ldots, p_s, p_{s+1}$ polinomokat kapunk, amelyekre minden $i = 0, \ldots, s$ mellett a (†) azonosság fennáll, és $\deg p_{s+1} = -\infty$, azaz $p_{s+1} = 0$.

A (†) azonosság szerint, minden $i = 0, 1, \ldots, s$ 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 (†) azonosságot értelmezhetjük úgyis, 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 $i = 0, \ldots, s$-re végül is $J(p_{i-1}, p_i) = J(p_i, p_{i+1})$. Alkalmazva ezt minden $i = 0, \ldots, s$ mellett

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

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 $p|qr$. Ekkor $p|r$.* ⌟

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 $p|r$. ▪

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. Ha s polinomom közös osztójuk, akkor ds is közös osztója p-nek és q-nak, amiből $ds|s$ következik. A fokszámokat összehasonlítva ez csak akkor lehetséges, ha s konstans polinomom, azaz $s|1$ valóban fennáll.

Most megmutatjuk, hogy az m legkisebb közös többszörösre

$$m = dr_1r_2. \qquad (\dagger)$$

Világos, hogy dr_1r_2 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 a nullosztómentesség szerint

$$r_1s_1 = r_2s_2.$$

Na most, a fent kiemelt azonosság szerint szerint $r_1|r_2s_2$, ahol r_1 és r_2 relatív prímek. Ebből azonnal kapjuk, hogy $r_1|s_2$. No de $s = qs_2 = dr_1s_2$, amiből már látszik, hogy s egy többszöröse a dr_1r_2 polinomnak. Ez éppen (†) azonosságot jelenti. Innen $d \cdot m = d(dr_1r_2) = (dr_1)(dr_2) = p \cdot q$ már nyilvánvaló. ▪

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 mindig zérus, és a hányados 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 felvetett 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, méghozzá lényegében csak egyféleképpen 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 tétel a polinomok gyűrűjében is igaz marad. Természetesen konkrét algoritmust nem adunk, hiszen ilyen még számokra sem igen van.[7]

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. ▪

Érdemes látni, hogy ha p irreducibilis, és f egy tetszőleges polinom, akkor vagy $p|f$ vagy p és f relatív prímek. Ugyanis ha g egy közös osztójuk, akkor p iredducibilitása miatt, $\deg g = 0$, vagy $\deg g = \deg p$. Ez utóbbi esetben g a p konstans szorosa, ergo $p|f$, az előbbi eset pedig éppen azt jelenti, hogy p és f relatív prímek.

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

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 $n \geq 2$. Induljunk ki tehát abból, hogy

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

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

Az állítás fordítva is igaz, de azt a gyakorlatokra hagyjuk. 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, irreduciblis $q_1, \ldots, q_s \in \mathbb{F}[t]$ 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 $p_1, \ldots, p_s$ polinomok alkalmas átindexelése után $p_j = q_j$, minden $j = 1, \ldots, s$ mellett.

Bizonyítás: Az 1.37. állításban már tisztáztuk az egzisztenciális részt, így már csak az unicitás maradt, amit a felbontandó polinom fokszáma szerinti indukcióval végzünk most is. Ha a polinom elsőfokú, akkor az unicitás 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 $p_1, \ldots, p_s, q_1, \ldots, q_r$ 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

[7] Ha ilyen lenne senki nem tudna interneten két számla közt pénzt mozgatni. Lásd például: `https://en.wikipedia.org/wiki/RSA_(cryptosystem)`

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

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

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

Az egész szakaszt összefoglalhatjuk így is:

1.40. állítás. *Minden legalább elsőfokú normált polinomhoz léteznek, méghozzá sorrendjüktől eltekintve egyértelműen léteznek $q_1, \ldots, q_s$ normált, irreducibilis polinomok, és $n_1, \ldots, n_s$ 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 nevezzünk, 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 jelöli az i-edik sor j-edik elemét. Sokszor használjuk az $A_{i,j} = A_i^j$ 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.

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

Az azonos típusú mátrixok közt 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 $i = 1, \ldots, m$ és $j = 1, \ldots, n$. 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 ellenőrizhetőek 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.

Most mátrixok szorzatát 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 $i = 1, \ldots, m$ és $j = 1, \ldots, n$. ⌟

Az összeadás és szorzás műveleteket a disztributivitás kapcsolja össze:

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

$$(A + B)\, C = AC + BC.$$ ⌟

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 szorzata. Magyarul: minden $1 \le i \le m$ és $1 \le j \le n$ 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 $1 \le j \le n$ 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 $1 \le i \le n$ 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:*

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

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

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

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

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*

$$A\,(BC) = (AB)\,C. \qquad \lrcorner$$

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 mindkét oldalon azonos méretű konkrétan $m \times n$ méretű mátrixok szerepelnek.

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

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

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

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

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

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 $\left(\mathbb{F}^{n\times n}, +, \cdot\right)$ algebrai struktúra egy egységelemes gyűrű.* $\lrcorner$

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 e nélkül is nagyon egyszerűen látszik.[8] Gyakorlatként próbáljunk magasabb n számok mellett is a kommutativitás és a nullosztómentesség hiányára példát találni.

Nagyon fontos látni, hogy a kommutativitás hiánya, az eddigiektől 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. Egy mátrixot *diagonális alakúnak* mondunk, ha minden nem zérus eleme a fő diagonálisában 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, ehhez már szükség van a lineáris függetlenség fogalmára vagy a Gauss-Jordan eliminációs algoritmusra, 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 köztük *művelettartó bijekció*, azaz létezik

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

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 közt 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 mindkét eleme azonos, a valós testtel izomorf testet alkot.

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

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.[9] Ezek közül számunkra a legfontosabb a következő részhalmaz.

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

[9] Karakterizációjukat lásd: (Ebbinghaus és tsai. 1991)-ben.

1.53. definíció-állítás (komplex számtest)**.** Jelölje két tetszőleges $a, b \in \mathbb{R}$ valós szám mellett $M_{a,b}$ az 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

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 kommutativitása 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

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

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. .

Ebben a $\mathbb{C}$ testben az $M_{0,-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ elem olyan, hogy a négyzete a szorzásra nézve reprodukáló elemnek 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 számoktestnek 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,[10] akkor azt kapjuk, hogy minden komplex szám

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

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

$$(a + ib) + (c + id) = (a + c) + i(b + d) \tag{†}$$

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

$$(a + ib) \cdot (c + id) = (ac - bd) + i(ad + bc) \tag{‡}$$

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 (†), és a szorzás műveletet definiálja (‡). Az így kapott algebrai struktúra test, amely izomorf a komplex számtest mátrix reprezentációjával. Az izomorfizmust a

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

művelettartó bijekció hozza létre. ⌟

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

1.8. A komplex számok abszolút értéke

A valós számokra jól ismert abszolút érték 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* $\bar{z} = a - ib$. ⌟

1.57. állítás. *Minden $z \in \mathbb{C}$ komplex szám mellett*

$$\bar{\bar{z}} = z, \quad z + \bar{z} = 2\Re z, \quad z - \bar{z} = 2i\Im z, \quad z \in \mathbb{R} \iff z = \bar{z}, \quad z\bar{z} = (\Re z)^2 + (\Im z)^2 \geq 0,$$

és bármely két $z, w \in \mathbb{C}$ komplex szám esetén

$$\overline{z + w} = \bar{z} + \bar{w}, \quad \overline{zw} = \bar{z}\bar{w}, \quad \overline{z - w} = \bar{z} - \bar{w}, \quad \overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}} \quad \textit{feltéve, hogy } w \neq 0. \qquad ⌟$$

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

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

Az $a^2 - b^2 = (a + b)(a - b)$ azonosság szerint, ha $a \geq b \geq 0$ akkor $a^2 \geq b^2$ is fennáll.

[10] Kicsit pontosabban: a valós számok 2×2-es mátrix reprezentációját használjuk.

1.58. definíció (komplex szám abszolút értéke)**.** 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 *abszolút értékének*. Jelölés: $|z| = \sqrt{z\bar{z}}$. ⌟

Fontos látni, hogy 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 abszolút értékének definíciójával. Világos, hogy $|\Re z|^2 \leq (\Re z)^2 + (\Im z)^2 = z\bar{z}$, emiatt

$$\Re z \leq |z|.$$

Az abszolút érték 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 egyenlőtlenségnek* nevezik. Egy kevésbé népszerű, de ekvivalens alakja

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

Vegyük észre, hogy a valós abszolút érték függvény tulajdonságait sem használtuk, és a fenti tulajdonságokat valós esetre is újra igazoltuk.

1.9. A komplex számok trigonometrikus alakja

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

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

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, mintha 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

$$z = |z|\,(\cos\varphi + i \sin\varphi) \tag{†}$$

Ezt a φ számot nevezzük a z komplex szám *argumentumának*, és $\arg z$ módon jelöljük. A (†) alak a komplex szám *trigonometrikus alakja*.

Két nem zérus komplex szám egyenlősége azt jelenti, hogy az abszolút értékük azonos, és az argumentumaik 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 egységkörének egy pontja. A cos és sin függvény definíciója szerint, ha φ jelöli az origót az (x, y) ponttal a körcikk peremén mért ív hosszát, akkor $x = \cos\varphi$ és $y = \sin\varphi$. .

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

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

Ekkor a két szám szorzatának abszolútértéke az abszolútértékek szorzata és a szorzat argumentuma az argumentumok összege, azaz

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

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

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

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 (egységgyök)**.** Legyen az $n \in \mathbb{N}$ természetes szám rögzítve. Tetszőleges $k = 0, 1, \ldots, n-1$ mellett jelölje

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

az úgynevezett *n-edik komplex egységgyököket*. $\lrcorner$

Az n-edik komplex egységgyökök a komplex számsík pontosan n különböző pontját alkotják, és az n-edik hatványuk 1, azaz és $(\epsilon_k^{(n)})^n = 1$. 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 $k = 0, \ldots, n-1$ 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. *Ha $p(t) \in \mathbb{C}[t]$ 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.* $\lrcorner$

Bizonyítás: Legyen p egy legalább másodfokú komplex polinom. Az algebra alaptétele miatt van gyöke, és tudjuk hogy a gyöktényező mindig kiemelhető. 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. ∎

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

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

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

Persze ez azt jelenti, hogy egy n-ed fokú komplex polinomnak mindig 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 első fokú 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 polinom gyűrűre 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 $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$. A konjugálás tulajdonságai szerint egy z komplex számra

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

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 az állítás semmit mondó. Ha most z egy nem valós komplex gyök, akkor $p(t) = (t - z)(t - \bar{z})\, h(t)$, 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. ▪

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 $x_1, \ldots, x_r \in \mathbb{R}$ valós számok, léteznek $\alpha_1, \beta_1, \ldots, \alpha_s, \beta_s$ valós együttható párok, és léteznek $n_1, \ldots, n_r, m_1, \ldots, m_s$ pozitív egészek úgy, hogy*

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

Itt $r, s \geq 0$, de $r + s > 0$, és $n_1 + \cdots + n_r + 2(m_1 + \cdots + m_s) = \deg p$, 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 tételt. A p így előáll mint első fokú, esetleg komplex polinomok szorzata. A gyököket osszuk két része. Legyenek $x_1, \ldots, x_r$ a különböző valós gyökök, amelyek multiplicitása rendre $n_1, \ldots, n_r$. Legyenek $z_1, \overline{z_1}, \ldots, z_s, \overline{z_s}$ a különböző nem valós de komplex gyökök, ahol $m_1, \ldots, m_s$ rendre a konjugált gyök párok multiplicitása. Így

$$p(t) = (t - x_1)^{n_1} \cdots (t - x_r)^{n_r} \cdot \left((t - z_1)(t - \bar{z_1})\right)^{m_1} \cdots \left((t - z_s)(t - \bar{z_s})\right)^{m_s}.$$

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

A fenti állításból két dolog azonnal látszik. Az első, hogy *valós számtest felett minden legalább harmadfokú polinom reducibilis*, a második pedig, hogy *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 $+ : V \times V \to V$ két változós művelet (ezt összeadásnak nevezzük) és adott egy $\cdot : \mathbb{F} \times V \to V$ szintén kétváltozós művelet (ezt skalárral való szorzásnak, vagy számmal való mondjuk). A V-t az $\mathbb{F}$ test feletti *vektortérnek* nevezzük, ha $(V, +)$ egy Ábel-csoport, és a számmal valós szorzás műveletre teljesülnek az alábbi axiómák:

1. Minden $\alpha \in \mathbb{F}$ és minden $u, v \in V$ mellett $\alpha \cdot (u + v) = \alpha \cdot u + \beta \cdot v$,
2. Minden $\alpha, \beta \in \mathbb{F}$ és minden $u \in V$ mellett $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot v$,
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 való szorzás $\cdot : \mathbb{F} \times V \to V$ 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 való szorzással vektorteret alkot az $\mathbb{F}$ test felett. Ha $n = 1$, akkor kapjuk a 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 közt szokásos összeadás művelettel, 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 felett.

Rögzítenünk kell magunkban, hogy a vektortér definíciójában a test is fontos szerepet játszik. Más test felett ugyan az Ábel-csoport már egy másik vektorteret alkot. 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 Ábel-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 $\alpha \star a = a^{\alpha}$. [1]

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 mindig kijelölni, mikor egy vektortérre hivatkozunk, de ha világos, hogy mi a vektortérbeli elemek közt az additív művelet, és hogy mit jelent egy test elemeivel szorozni, akkor elhagyjuk a műveletek kijelölését. A legpontosabb, – de persze a legkörülményesebb – 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 alterei

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 művelettel, 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* $\alpha v \in M$. ⌟

Tetszőleges V vektortérre a $\{0\}$ és maga V mindig 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 mindig 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 $\operatorname{gen} H$ 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 $H \subseteq \operatorname{gen} H$,*
2. *Ha $H \subseteq K \subseteq V$, akkor* $\operatorname{gen} H \subseteq \operatorname{gen} K$,
3. *Minden $H \subseteq V$ mellett* $\operatorname{gen}(\operatorname{gen} H) = \operatorname{gen} H$.

A $\operatorname{gen} H$ *a H halmazt tartalmazó alterek közt a legszűkebb. Így $H \subseteq V$ pontosan akkor altér, ha* $\operatorname{gen} H = H$. ⌟

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

$$\alpha_1 v_1 + \cdots + \alpha_r v_r$$

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

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

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

Definíció szerint nulla darab vektor lineáris kombinációja a vektortér zérus eleme, tehát $\operatorname{lin} \emptyset = \{0\}$. A 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*

$$\operatorname{lin} H = \operatorname{gen} H.$$ ⌟

[1] Itt $a \star \alpha$ mit jelent?

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

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

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

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

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

A V vektorteret *végesen generáltnak* mondunk, ha létezik véges generátorrendszere. ⌟

A generátorrendszer cseréről szóló 2.9. lemmának kiemelten fontos szerepe van felépítésünkben. Egyrészt 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 generátorrendszere 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

$$a = \sum_{j=1}^{m} \alpha_j x_j$$

alakú, akkor x_k helyére betéve, a fent kifejezett formulát és bevezetve a $\delta = \frac{\alpha_k}{\eta_k}$ jelölést, azt kapjuk hogy:

$$a = \alpha_k x_k + \sum_{\substack{j=1 \\ j \neq k}}^{m} \alpha_j x_j =$$

$$= \alpha_k \left(\frac{1}{\eta_k} y + \sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{-1}{\eta_k} \eta_j x_j \right) + \sum_{\substack{j=1 \\ j \neq k}}^{m} \alpha_j x_j = \frac{\alpha_k}{\eta_k} y + \sum_{\substack{j=1 \\ j \neq k}}^{m} \left(\alpha_j - \frac{\alpha_k}{\eta_k} \eta_j \right) x_j =$$

$$= \delta y + \sum_{\substack{j=1 \\ j \neq k}}^{m} (\alpha_j - \delta \eta_j) x_j.$$

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

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

együtthatókkal, akkor ugyanez a vektor a módosított vektorrendszerből is kifejezhető, méghozzá az

$$\Big(\underbrace{\alpha_1 - \delta\eta_1}_{1.}, \underbrace{\alpha_2 - \delta\eta_2}_{2.}, \dots, \underbrace{\alpha_{k-1} - \delta\eta_{k-1}}_{k-1.}, \underbrace{\delta}_{k.}, \underbrace{\alpha_{k+1} - \delta\eta_{k+1}}_{k+1.}, \dots, \underbrace{\alpha_m - \delta\eta_m}_{m.} \Big)$$

együtthatókkal. ∎

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 valamilyen 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, a hol 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

	y	a
x_1	η_1	α_1
$\vdots$	$\vdots$	$\vdots$
x_k	$\boxed{\eta_k}$	α_k
$\vdots$	$\vdots$	$\vdots$
x_m	η_m	α_m
	δ	$\frac{\alpha_k}{\eta_k}$

$\Longrightarrow$

	y	a
x_1	0	$\alpha_1 - \eta_1\delta$
$\vdots$	$\vdots$	$\vdots$
y	1	δ
$\vdots$	$\vdots$	$\vdots$
x_m	0	$\alpha_m - \eta_m\delta$

(2.1)

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

	y	a	b	c
x_1	η_1	α_1	β_1	γ_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	$\boxed{\eta_k}$	α_k	β_k	γ_k
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_m	η_m	α_m	β_m	γ_m
	δ	$\frac{\alpha_k}{\eta_k}$	$\frac{\beta_k}{\eta_k}$	$\frac{\gamma_k}{\eta_k}$

$\Longrightarrow$

	y	a	b	c
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	δ_a	δ_b	δ_c
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_m	0	$\alpha_m - \eta_m\delta_a$	$\beta_m - \eta_m\delta_b$	$\gamma_m - \eta_m\delta_c$

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 jeleztük 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.

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

Itt az $n, m \in \mathbb{N}$, az $a_{i,j} \in \mathbb{F}$ testbeli számok előre adottak minden $i = 1, \ldots, m$ és minden $j = 1, \ldots, n$ mellett. Adottak még az egyenletek jobb oldalát képező $b_i \in \mathbb{F}$ számok. A feladat megoldása annyit tesz, hogy keressük az $x_1, \ldots, x_n$ ismeretlenek összes olyan értékét az $\mathbb{F}$ testből, amelyre fenti egyenletek mind teljesülnek. Ha a jobboldali számokra $b_i = 0$ minden $i = 1, \ldots, n$ mellett, akkor *homogén lineáris egyenletrendszerről* 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 $k = 1, \ldots, m$ mellet kiszámoljuk, akkor éppen egy Gauss-Jordan eliminációt hajtunk végre az $a_{i,j}$ pivot[2] elem 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. $L_1, \ldots, L_m$ jelöli az együttható-mátrix sorait A jobb oldali mátrix az elimináció eredménye, a megfelelő sorműveletekkel az egyes sorok címkéiben:

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

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 újabb oszlopra. Az algoritmus tehát akkor áll meg, ha minden oszlopot becseréltünk, vagy van ugyan nem becserélt oszlop de annak minden koordinátája zérus az eredeti generátorrendszer nem cserélt helyein. Ezt utóbbi mondatot egyszerűbben úgy fogalmazhatjuk, hogy ha a táblázat minden nem zérus sora egy becserélt vektorhoz tartozik, akkor véget ér az algoritmus.

A lineáris egyenletrendszerek nyelvére átültetve ez azt jelenti, hogy az algoritmus célja a változók eliminálása, de persze egy sorban csak egy eliminált változó szerepelhet. Az algoritmus akkor áll meg, ha a táblázatnak minden nem zérus sorában van eliminált változó.

[2] sarok elem

Az algoritmus utolsó táblájából a homogén lineáris egyenletrendszer általános megoldása könnyen leolvasható: Dobjuk el a zérus sorokat, és foglalkozzunk a maradék táblázattal. Minden sor egy és csak egy eliminált változót tartalmaz. A többi változó, ha ilyen van egyáltalán egy nem eliminált változóhoz tartozik. A nem eliminált változók tetszőleges értéket felvehetnek, majd minden sorban ebből már egyértelműen kifejezhető az eliminált változók értéke.

Innen érthető, hogy a szokásos szóhasználat szerint, az eliminált változókat sokszor *kötött változónak* nevezzük, és a többi változót *szabad változónak* mondjuk. A következő tétel 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. ▪

Az teljesen nyilvánvaló, hogy a kötött változók száma és a szabad változók száma azonos a változók számával. 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 $x_1, \ldots, x_n$ pontosan akkor elégíti ki a lineáris egyenletrendszert, 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 egyenletrendszert. A Gauss-Jordan algoritmus lehet például a következő:

	a	b	c
	$\boxed{1}$	4	7
	2	5	8
	3	6	8
	δ	4	7

$\Longrightarrow$

	a	b	c
a	1	4	7
	0	$\boxed{-3}$	-6
	0	-6	-13
	0	δ	2

$\Longrightarrow$

	a	b	c
a	1	0	-1
b	0	1	2
	0	0	$\boxed{-1}$
	0	0	δ

$\Longrightarrow$

	a	b	c
a	1	0	0
b	0	1	0
c	0	0	1

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:

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

Az elimináció lehet a következő:

	a_1	a_2	a_3	a_4	a_5
	$\boxed{1}$	3	4	5	-1
	-2	1	-1	4	-5
	2	1	3	0	3
	3	1	4	-1	5
	δ	3	4	5	-1

$\Longrightarrow$

	a_1	a_2	a_3	a_4	a_5
a_1	1	3	4	5	-1
	0	$\boxed{7}$	7	14	-7
	0	-5	-5	-10	5
	0	-8	-8	-16	8
	0	δ	1	2	-1

$\Longrightarrow$

	a_1	a_2	a_3	a_4	a_5
a_1	1	0	1	-1	2
a_2	0	1	1	2	-1
	0	0	0	0	0
	0	0	0	0	0

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 mindig 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.* ⌟

Mátrix faktorizáció

Érdemes a Gauss-Jordan eliminációs algoritmust az egyenletrendszerek megoldásától függetlenül is szemlélni. Ide másolom az előző feladat megoldásának táblázatát:

	a_1	a_2	a_3	a_4	a_5
a_1	1	0	1	-1	2
a_2	0	1	1	2	-1
	0	0	0	0	0
	0	0	0	0	0

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

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

hiszen minden oszlop 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 generátorrendszere. 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 létezik r-elemű generátorrendszere. 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 sorvektorterének létezik r-elemű generátorrendszere. 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 ismeretlent 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 $Az = Ax - Ax_0 = b - b = 0$.

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

Egy nem nyilvánvaló következmény, hogy akármelyik partikuláris megoldással toljuk is el az inhomogén rendszer megoldását mindig ugyan azt 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 valamilyen 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 oszlop vektorterének. Egészítsük 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 zéró 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 oszlop vektorokhoz 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.

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

A Gauss-Jordan algoritmus:

	a_1	a_2	a_3	a_4	a_5	b
	$\boxed{1}$	3	4	5	-1	6
	-2	1	-1	4	-5	-5
	2	1	3	0	3	7
	3	1	4	-1	5	10
	δ	3	4	5	-1	6

$\Longrightarrow$

	a_1	a_2	a_3	a_4	a_5	b
a_1	1	3	4	5	-1	6
	0	7	7	14	-7	7
	0	-5	-5	-10	$\boxed{5}$	-5
	0	-8	-8	-16	8	-8
	0	-1	-1	-2	δ	-1

$\Longrightarrow$

	a_1	a_2	a_3	a_4	a_5	b
a_1	1	2	3	3	0	5
	0	0	0	0	0	0
a_5	0	-1	-1	-2	1	-1
	0	0	0	0	0	0

a rendszernek tehát van megoldása például $x_1 = 5$, $x_5 = -1$, $x_2 = 0$, $x_3 = 0$, $x_4 = 0$. 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. *Az $Ax = b$ inhomogén lineáris egyenletrendszernek pontosan akkor van egyetlen 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.

A következő állításban azt gondoljuk meg, hogy ha kevesebb ismeretlene mint sora van egy rendszernek, akkor megválasztható olyan jobb oldali vektor, amellyel az inhomogén lineáris egyenletrendszernek nincs megoldása.

2.17. 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. Mikor az A mátrixra alkalmazott Gauss-Jordan algoritmus véget ér, akkor minden nem zérus sorban pontosan egy eliminált ismeretlen van. Eszerint minden egyes sor vagy zéró sor, vagy van benne eliminált ismeretlen. Mivel kevesebb ismeretlen van mint sor, ezért legalább egy zéró sornak lennie kell. Ha a k-adik sor csak nullákat tartalmaz, akkor az a $b \in \mathbb{F}^m$ vektor, amelynek a k-adik koordinátája nem zérus, de minden más koordináta zérus meg is felel. .

Inverz mátrix

Gondoljunk arra, hogy milyen módszerrel automatizálhatnánk a Gauss-Jordan algoritmust. Mivel sormű-veletekről van szó, természetes gondolat, hogy az $A \in \mathbb{F}^{m\times n}$ mátrix transzformálásához $m \times m$ méretű mátrixokat használjunk, 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 generátorrendszere, 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 eliminálтuk az j-edik ismeretlent, tehát a megoldás j-edik koordinátáját tartalmazza az $m + k$-adik oszlop első eleme.

Ha tehát úgy cseréljük fel a táblázat sorait, hogy a bal oldali $m \times m$ 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 felcseréltünk, ami egy 3) típusú balszorzást jelent. Ha az 1) vagy 2) vagy 3) típusú mátrixot *elemi mátrix-nak* 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

$$BA = I.$$

2.18. definíció (nem szinguláris mátrix)**.** Egy $A \in \mathbb{F}^{m\times m}$ négyzetes mátrixot *invertálhatónak* vagy *reguláris-nak* vagy *nem szinguláris-nak* 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. ⌟

2.19. 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 generátorrendszere. 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 $A_1A_2 = A$. Azt kaptuk tehát, hogy

$$A_1\,(A_2B) = (A_1A_2)\,B = AB = I.$$

Ebből az következik, hogy tetszőleges $b \in \mathbb{F}^m$ vektor mellett az $x = A_2Bb \in \mathbb{F}^r$ vektor megoldása az $A_1x = b$ 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

$$A^{-1} = A^{-1}I = A^{-1}\,(AB) = \left(A^{-1}A\right)B = IB = B.$$.

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 22. 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 tudjunk végrehajtani az eliminációs algoritmusban. Egy lehetséges megoldás a következő:

$$\begin{array}{cccc|cccc}
\boxed{1} & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 3 & 1 & 2 & 0 & 1 & 0 & 0 \\
2 & 3 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 2 & 1 & 0 & 0 & 0 & 1 \\
\hline
\delta & 1 & 1 & 0 & 1 & 0 & 0 & 0
\end{array}
\Rightarrow
\begin{array}{cccc|cccc}
1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 3 & 1 & 2 & 0 & 1 & 0 & 0 \\
0 & 1 & -1 & 0 & -2 & 0 & 1 & 0 \\
0 & -1 & 1 & \boxed{1} & -1 & 0 & 0 & 1 \\
\hline
0 & -1 & 1 & \delta & -1 & 0 & 0 & 1
\end{array}
\Rightarrow$$

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

$$\begin{array}{cccc|cccc}
1 & 0 & 0 & 0 & -3 & -\frac{1}{2} & \frac{3}{2} & 1 \\
0 & 0 & 1 & 0 & 3 & \frac{1}{4} & -\frac{5}{4} & -\frac{1}{2} \\
0 & 1 & 0 & 0 & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} \\
0 & 0 & 0 & 1 & -3 & 0 & 1 & 1
\end{array}
\Rightarrow
\begin{array}{cccc|cccc}
1 & 0 & 0 & 0 & -3 & -\frac{1}{2} & \frac{3}{2} & 1 \\
0 & 1 & 0 & 0 & 1 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} \\
0 & 0 & 1 & 0 & 3 & \frac{1}{4} & -\frac{5}{4} & -\frac{1}{2} \\
0 & 0 & 0 & 1 & -3 & 0 & 1 & 1
\end{array}$$

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:

	a_1	a_2	a_3	a_4	e_1	e_2	e_3	e_4
e_1	$\boxed{1}$	1	1	0	1	0	0	0
e_2	0	3	1	2	0	1	0	0
e_3	2	3	1	0	0	0	1	0
e_4	1	0	2	1	0	0	0	1
	δ	1	1	0	1	0	0	0

$\Rightarrow$

	a_2	a_3	a_4	e_1	e_2	e_3	e_4
a_1	1	1	0	1	0	0	0
e_2	3	1	2	0	1	0	0
e_3	1	-1	0	-2	0	1	0
e_4	-1	1	$\boxed{1}$	-1	0	0	1
	-1	1	δ	-1	0	0	1

$\Rightarrow$

	a_2	a_3	e_1	e_2	e_3	e_4
a_1	1	1	1	0	0	0
e_2	5	$\boxed{-1}$	2	1	0	-2
e_3	1	-1	-2	0	1	0
a_4	-1	1	-1	0	0	1
	-5	δ	-2	-1	0	2

$\Rightarrow$

	a_2	e_1	e_2	e_3	e_4
a_1	6	3	1	0	-2
a_3	-5	-2	-1	0	2
e_3	$\boxed{-4}$	-4	-1	1	2
a_4	4	1	1	0	-1
	δ	1	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{2}$

$\Rightarrow$

	e_1	e_2	e_3	e_4
a_1	-3	$-\frac{1}{2}$	$\frac{3}{2}$	1
a_3	3	$\frac{1}{4}$	$-\frac{5}{4}$	$-\frac{1}{2}$
a_2	1	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{2}$
a_4	-3	0	1	1

$\Rightarrow$

	e_1	e_2	e_3	e_4
a_1	-3	$-\frac{1}{2}$	$\frac{3}{2}$	1
a_2	1	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{2}$
a_3	3	$\frac{1}{4}$	$-\frac{5}{4}$	$-\frac{1}{2}$
a_4	-3	0	1	1

2.3. Lineárisan független rendszerek

2.20. definíció (lineárisan összefüggő vektorrendszer)**.** Egy véges $\{y_1, \ldots, y_n\}$ vektorrendszert *lineárisan összefüggőnek* mondunk, ha van olyan vektora, amely kifejezhető a többi vektor lineáris kombinációjaként. ⌟

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

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

Nyilvánvaló példa, hogy ha a vektorrendszer a 0 vektort tartalmazza, akkor lineárisan összefüggő. Hasonlóan, ha ugyanaz a vektor többször szerepel a vektorrendszerben, a rendszer szintén összefüggő. Olyan vektorrendszerre, amelyre nem igaz a lineárisan összefüggőség definíciója példa egy egyedüli nem zérus vektor, de az $\{\}$ üres rendszer is.

2.21. állítás. *Legyen $\{y_1, \ldots, y_n\}$ 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 $y_1, \ldots, y_n$ vektorok lineáris kombinációja,*
   azaz formálisabban: 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 $\alpha_1, \ldots, \alpha_n$ skalárok, amelyekkel*

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

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

$$(\alpha_1, \ldots, \alpha_{k-1}, 0, \alpha_{k+1}, \ldots, \alpha_n) \qquad (0, \ldots, 0, 1, 0, \ldots, 0)$$

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

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

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

Ezt kellett belátni. ∎

Fontos észrevétel a következő.

2.22. állítás. *Minden, valamely lineárisan összefüggő vektorrendszert tartalmazó vektorrendszer maga is lineárisan összefüggő.* ⌟

2.23. definíció. Egy nem feltétlen véges vektorrendszert lineárisan összefüggőnek nevezünk, ha van véges részrendszere, amely lineárisan összefüggő. ⌟

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

Így egy nem véges vektorrendszer akkor lineárisan független, ha minden véges részrendszere is az. Egy véges vektorrendszer lineárisan függetlenségét, pedig a következő egymással ekvivalens állítások karakterizálják.

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

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

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

Bizonyítás: Nyilvánvaló a lineáris összefüggés karakterizációjából. ∎

A lehető legszűkebb 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 generátorrendszerek közül a lehető legszűkebbet keresni.

2.26. 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 generátorrendszert *minimális generátorrendszernek* mondunk, ha nem lehet szűkíteni úgy, hogy generátorrendszer maradjon. ⌟

2.27. állítás. *Legyen $\{x_1, \dots, x_m\}$ egy vektorrendszere a V vektortérnek. Az alábbi feltevések ekvivalensek.*

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

Bizonyítás: Az alábbi lépéseket követjük.

1.⇒2. Ha a vektorrendszer nem lenne generátorrendszer is, akkor a lineáris burkán kívül lenne egy vektor. Ezt a vektorrendszerhez illesztve, a vektorrendszer egy valódi lineárisan független bővítését kapjuk, ami ellentmond a maximalitás feltételének.

3.⇒2. Ha a vektorrendszer egyik eleme a többi elem lineáris kombinációja, akkor azt az elemet elhagyva is generátorrendszert kapunk, ami ellentmond a minimalitás feltételének.

2.⇒1. Ha nem lenne maximális a lineárisan független tulajdonságra nézve, akkor létezne egy vektor a lineáris burkán kívül is, ami ellentmond a generátorrendszer tulajdonságnak.

2.⇒3. Mivel a vektorrendszer egyik eleme, sincs a többi lineáris burkában, ezért egyetlen elemet sem hagyhatunk el a generátorrendszer tulajdonság megtartásával, ami azt jelenti, hogy ez egy minimális generátorrendszer. ▪

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.

2.28. 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 $\exists k \quad 2 \leq k \leq n : y_k \in \operatorname{lin}\{y_1, \dots, y_{k-1}\}$ ⌟

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:

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

2.29. 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 y 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 együtthatója 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. ▪

3. fejezet

A Steinitz-lemma

FÜGGETLEN- ÉS GENERÁTORRENDSZEREK elemszáma közti kapcsolat vezet a a bázis és a dimenzió fogalmához. A félév legfajsúlyosabb állítása, az egyszerű bizonyítás 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 alkalmas 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 alkalmas átindexelésével az

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

vektorrendszer is generátorrendszer.

Ilyen k biztosan van 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 — kicserélhető, 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. .

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 feszítő rangja r, ha r a legkisebb olyan pozitív egész, amelyre A előáll

$$A = BC$$

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

[1] Úgy 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 közt.

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 sorrendszerek 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 sorrendszere egy lineárisan független rendszer e V vektortérben, és C sorai pedig egy generátorrendszert 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. ▪

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 $B \in \mathbb{F}^{r\times m}$ mátrix, amelyre $AB = A$ 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 $A \in \mathbb{F}^{n\times m}$ mátrix, amelynek feszítőrangja r előáll*

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

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

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

rendszerek 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 $B = B_1B_2$ előállítás, ahol $B_1 \in \mathbb{F}^{n\times s}, B_2 \in \mathbb{F}^{s\times r}$, ahol $s < r$. Ekkor persze

$$A = BC = (B_1B_2)\,C = B_1\,(B_2C)$$

Itt B_1 mátrixnak s oszlopa van, a B_2C mátrixnak s sora van, ami ellentomond annak, hogy A feszítőrangja r. A sorrendszer lineárisan függetlensége is hasonlóan adódik. ▪

3.2. Dimenzió

A Steinitz-lemma következményeképpen:

3.4. következmény. Egy vektortérben bármely két véges egyszerre lineárisan független és egyszerre 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 generátorrendszere. ⌟

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

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

Bizonyítás: Tekintsünk egy véges generátorrendszert. Ha minden elem kívül esik a többi elem lineáris burkában, akkor a rendszer lineárisan független, és készen is vagyunk. Ha van olyan elem, amely a többi elem lineáris burkában van, akkor dobjuk el ezt az elemet, és tekintsük, a most már eggyel kevesebb elemből álló vektorrendszert. Világos, hogy ez is generátorrendszer marad.

Folytassuk az eljárást. Mivel véges sok vektor van az eredeti generátorrendszerben az algoritmus előbb-utóbb megáll, ami azt jelenti, hogy olyan generátorrendszert kapunk, ahol már minden elem a többi lineáris burkán kívül van, ergo lineárisan független. ▪

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 vektortérről azt mondjuk, hogy n dimenziós, vagy n a dimenzió száma, ha a vektortérben van n elemű bázis.* ⌟

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.

Az eddigiek összefoglalásaként is tekinthető a következő állítás.

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

[2] Ez 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 generátorrendszere, 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. ▪

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 generátorrendszere. Legyen $\{y_1, \dots, y_n\}$ egy lineárisan független. A Steinitz-lemma szerint ez a rendszer kiterjeszthető egy

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

generátorrendszerré, ami persze bázis is. Ezt kellett belátni. ▪

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

A független rendszerek cseréjéről (2.29.) 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 altér minden altere is végesen generált. Ha $M \subseteq V$ egy altér és V végesen generált, akkor* $\dim M \leq \dim V$. ⌟

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

Van tehát V-ben is M-ben is bázis. Mivel az M-beli bázis egyben lineárisan független rendszer V-ben, ezért a Steinitz-lemma szerint $\dim M \leq \dim V$. ▪

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 $A : V \to W$ függvényt *lineáris operációnak* 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 $V \to W$ 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 $x \in \mathbb{F}^n$ oszlop vektor mellett az

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

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

$$x \mapsto x \cdot A$$

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

4.2. definíció-állítás. Legyen $A \in L(V, W)$ egy lineáris operáció. Ennek értékkészlete egy altér, amelyet $\operatorname{Im} A$ 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:

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

4.3. állítás. *Egy lineáris leképezés pontosan akkor injektív, ha* $\ker A = \{0\}$*. Egy injektív lineáris leképezés inverz függvénye egy* $\operatorname{Im} A \to V$ *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 generátorrendszere 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 $A : V \to W$ függvényt *izomorfizmusnak* mondunk, ha A lineáris bijekció. Az ugyanazon test feletti V, W vektortereket egymással *izomorf vektortérnek* mondjuk, ha létezik $A : V \to W$ 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 $\Psi : \mathbb{F}^n \to V$ 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. ▪

4.7. definíció (koordináta)**.** Legyen $\{e_1,\dots,e_n\}$ egy bázisa a V vektortérnek. Tudjuk, hogy minden $v \in V$ vektor egyetlen egyféleképpen, 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 mindkét térnek.

Ha $\dim V = \dim W = n$, akkor mindkét vektortér izomorf $\mathbb{F}^n$ vektortérrel, ergo egymással is izomorfak. ▪

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 vektort koordináta-vektora. A 2.9., a 2.29. és a 3.11. lemmák összefoglalásaként kapjuk, hogy milyen módon változnak egy rögzített vektor koordinátái, ha a bázisban egy elemet megváltoztatunk.

4.9. lemma (bázis transzformáció)**.** *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 $\{x_1,\dots,x_{k-1},y,x_{k+1},\dots,x_m\}$ vektorrendszer 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 $a = \sum_{\substack{j=1,\\ j\neq k}}^{n} (\alpha_j - \eta_j\delta)\, x_j + \delta y_k$, ahol $\delta = \frac{\alpha_k}{\eta_k}$.*

Egy táblázatban összefoglalva:

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

⌟

5. fejezet

Alterek 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* az

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

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

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

az α szám és H_1 halmaz szorzatát. ⌟

Világos, hogy ha például H_1 üres, akkor $H_1 + H_2$ is, és a 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. $\alpha(A + B) = \alpha A + \beta B$;
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 $A + (-A) \neq \{0\}$, 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.* ⌟

[1] Egy $A \subseteq V$ halmazt *affin halmaznak* szokás mondani, ha zárt az 1 összegű lineáris kombinációra (*affin kombinációra*) nézve, tehát ha a $\lambda A + (1 - \lambda) A \subseteq A$ 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 altér $v \in V$ 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 halmazelméleti közös része. Felmerül, hogy mi a legszűkebb olyan altér, amely mindkét előre megadott alteret tartalmazza. A válasz kézenfekvő, hiszen éppen ez a generált altér fogalma. Ha $H_1, H_2 \subseteq V$ tetszőleges halmazok, akkor $\operatorname{lin}\{H_1 \cup H_2\}$ a H_1 és H_2 halmazokat tartalmazó legszűkebb altér. Amennyiben altereket tartalmazó legszűkebb alteret keresünk, akkor többet is állíthatunk:

5.4. állítás. *Legyenek $M, N \subseteq V$ alterek. Ekkor $\operatorname{lin}(M \cup N) = M + N$, tehát az alterek Minkowski-összege is altér, sőt ez az M és N altereket tartalmazó legszűkebb altér.* ⌟

Bizonyítás: Ha R egy altér, amelyre $M \subseteq R$ és $N \subseteq R$ is fennáll, akkor $M + N \subseteq R + R \subseteq R$, emiatt $M + N \subseteq \operatorname{gen}(M \cup N)$. Másrészt világos, hogy $M + N$ egy az $M \cup N$ halmazt tartalmazó altér, Így azt kaptuk, hogy $\operatorname{gen}(M \cup N) \subseteq M + N$. Figyelembe véve, hogy $\operatorname{gen}(M \cup N) = \operatorname{lin}(M \cup N)$, készen is vagyunk. ▪

A következő állítás szerint, ha előre adott két altér, akkor ezek összegének dimenziója mindig azonos a két alteret 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 generátorrendszere 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

$$\{m_1, \dots, m_s, p_1, \dots, p_r, n_1, \dots, n_t\} \tag{†}$$

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

$$\sum_{j=1}^{s} \alpha_j m_j + \sum_{j=1}^{r} \gamma_j p_j + \sum_{j=1}^{t} \beta_j n_j = 0. \tag{‡}$$

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 (‡) 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, $\{p_1, \dots, p_r, n_1, \dots, n_r\}$ lineárisan független, ezért itt minden együttható zérus. Speciálisan $\beta_j = 0$ minden $j = 1, \dots, t$ mellett. Ezt (‡)-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 $\alpha_1 = \dots = \alpha_s = \gamma_1 = \dots = \gamma_t = 0$.

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

[2] Figyeljünk 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.

5.6. definíció (direkt összeg)**.** Legyenek az $M_1, \dots, M_s \subseteq V$ alterek a V vektortér alterei ($s \geq 2$). Azt mondjuk, hogy az $N \subseteq V$ altér 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 $N = M_1 + M_2$.

5.7. 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 $v_j \in M_j, j = 1, \dots, s$ vektor, amelyre*

$$v = \sum_{j=1}^{s} v_j.$$ ⌟

Bizonyítás: Tegyük fel, hogy M direkt összege az $M_1, \dots, M_s$ altereknek. Minden szóba jövő k-ra $M_k \subseteq \sum_{j=1}^{s} M_j = M$. Mivel $M = M_1 + \dots + M_s$, ezért minden $v \in M$-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 $u_j \in V_j$ minden $j = 1, \dots, s$ mellett. Ekkor minden rögzített $k = 1, \dots, s$ mellett

$$u_k - v_k = \sum_{\substack{j=1 \\ j\neq k}}^{s} (v_j - u_j).$$

No de, a baloldali vektor M_k-beli, míg a jobboldali vektor $\sum_{\substack{j=1 \\ j\neq k}}^{s} M_j$-beli. A feltétel szerint ilyen vektor egyedül a zérus vektor, amiből $u_k = v_k$ következik. Mivel ezt minden $k = 1, \dots, s$ mellett elismételhetjük, ezért az előállítás egyértelműségét be is láttuk.

Megfordítva, most azt tegyük fel, hogy a két feltétel fennáll. Ekkor 1. szerint $M_1 + \dots + M_s \subseteq N + \dots + N = N$, és 2. szerint $N \subseteq M_1 + \dots + M_s$, amiből már $\sum_{j=1}^{s} M_j = N$ következik is. Most képzeljük el, hogy valamely k mellett $v_k \neq 0, v_k \in M_k$ és $v_k = \sum_{\substack{j=1 \\ j\neq k}}^{s} v_j$, ahol $v_j \in M_j$ minden $j = 1, \dots, s$ mellett. Ekkor

$$v_k - \sum_{\substack{j=1 \\ j\neq k}}^{s} v_j = 0 = \sum_{j=1}^{s} 0$$

az N vektortér zérus elemének két különböző előállítása, hiszen a k-adik vektortérbeli elem a baloldalon a v_k nem zérus vektora M_k-nak, míg a jobboldalon a zérus vektor szerepel M_k-ból (is). ▪

Érdemes észrevenni, hogy a fenti állítás 2. feltétele nem más mint az alábbi két feltétel együttesének tömör megfogalmazása:

2a. $N = M_1 + \dots + M_s$,

2b. $\sum_{j=1}^{s} v_j = 0$, $v_j \in M_j$ *esetén minden* $j = 1, \dots, s$ *mellett* $v_j = 0$.

Így a következő állítást is meggondoltuk:

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. *$M_k \subseteq N$ minden $k = 1, \dots, s$ mellett;*

2a. *$N = M_1 + \dots + M_s$;*

2b. *$\sum_{j=1}^{s} v_j = 0$, $v_j \in M_j$ esetén minden $j = 1, \dots, s$ mellett $v_j = 0$.* ⌟

Ha tehát csak az $M_1, \dots, M_s$ alterek adottak, akkor annak szükséges és elegendő feltétele, hogy ezeknek a direktösszege értelmezhető legyen éppen a fenti 2b feltétel. Ebben az esetben az $N = \sum_{j=1}^{s} M_j$ altér lesz az $M_1, \dots, M_s$ alterek

$$N = M_1 \oplus \dots \oplus M_s$$

direktösszege.

Megint 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 alterei N-nek, hogy N minden eleme előáll, de csak egyféleképpen 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 indirekt feltevés szerint

$$\dim N = \dim M + \dim M_s = \sum_{j=1}^{s-1} \dim M_j + \dim M_s = \sum_{j=1}^{s} \dim M_j. \quad .$$

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 generátorrendszere, és az elemeinek száma az 5.9. állítás szerint éppen dim N. Azt kaptuk tehát, hogy ez egy minimális generátorrendszere N-nek, ergo egy bázis.

A fenti állítás konkrét alkalmazhatóságához nagy segítség a most következő észrevétel.

5.10. állítás. *Adottak az $M_1, \dots, M_s \subseteq V$ alterek, ahol $s \geq 2$. Az alábbi feltételek ekvivalensek:*

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

2. *Minden $k = 2, \dots, s$ mellett $M_k \cap \left(\sum_{j=1}^{k-1} M_j \right) = \{0\}$.*

⌟

Bizonyítás: Mivel $k \geq 2$ mellett $\sum_{j=1}^{k-1} M_j \subseteq \sum_{\substack{j=1 \\ j \neq k}}^{s} M_j$, ezért 1. $\Rightarrow$ 2. nyilvánvaló.

Most azt mutatjuk meg, hogy ha 1. feltétel nem igaz, akkor 2. feltétel sem áll. Tegyük fel tehát, hogy van olyan $1 \leq k \leq s$, amelyre $M_k \cap \left(\sum_{\substack{j=1 \\ j \neq k}}^{s} M_j \right) \neq \{0\}$. Ha $k = s$, akkor készen is vagyunk. Legyen emiatt a továbbiakban $k < s$. Eszerint léteznek $v_j \in M_j (j = 1, \dots, s)$ vektorok, amelyekre

$$0 \neq v_k = \sum_{j=1}^{k-1} v_j + \sum_{j=k+1}^{s} v_j.$$

Legyen t a legnagyobb olyan szám, amelyre $k \leq t \leq s$ és $v_t \neq 0$. Ha $k = 1$, akkor a jobb oldali első szumma üres, emiatt $t > 1$. Ha viszont $k > 1$, akkor $1 < k \leq t$. Mindkét esetben látjuk tehát, hogy $t > 1$. Ekkor az

$$v_k = \sum_{j=1}^{k-1} v_j + \sum_{j=k+1}^{t} v_j.$$

azonosságot átrendezve kapjuk, hogy

$$-v_t = \sum_{j=1}^{k-1} v_j - v_k + \sum_{j=k+1}^{t-1} v_j.$$

Itt a bal oldali nem zérus vektor M_t-beli, a jobb oldali vektor $\sum_{j=1}^{t-1} M_j$-beli, ergo valamely $1 < t \leq s$ mellett találtunk az $M_t \cap \left(\sum_{j=1}^{t-1} M_j\right)$ altérben is egy nem zérus elemet. ▪

Az 1. feltétel előnye, hogy ennek alapján teljesen világos, hogy az $M_1 \oplus \cdots \oplus M_s$ direkt összeg nem függ az alterek sorrendjétől, 2. előnye pedig, hogy ha képeznünk kell az $M_1, \ldots, M_s$ alterek direkt összegét, akkor rekurzívan járhatunk el: ha a korábbi lépésben ellenőriztük, hogy az $M_1 \oplus \cdots \oplus M_{n-1}$ értelmes, akkor a következő lépésben csak azt kell ellenőriznünk, hogy M_n diszjunkt az első $n-1$ altér direkt összegétől.

5.3. Direkt kiegészítő

5.11. 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.* ⌟

Egy altérnek sok-sok direkt kiegészítője lehet. Például tekintsük a folytonos függvények $C[0,1]$ vektorterét $\mathbb{R}$ 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 $g(0) \neq 0$, az g által generált egy dimenziós altér direkt kiegészítője az N altérnek.[3]

[3] Ha véges dimenziós példára vágyunk, helyettesítsük $C[0,1]$-et a legfeljebb n-ed fokú $\mathbb{R}$ feletti polinomok $n+1$ dimenziós vektorterével.

6. fejezet

Vektortér faktortere

A VEKTORTÉR KONSTRUKCIÓK végéhez érkeztünk. Láttuk, hogy alterek közösrésze, 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 $V = \cup_{x \in V} M_x$; 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 úgyis 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 $v \in M_x$. ▪

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

$$M_{x_1} + M_{x_2} = (x_1 + M) + (x_2 + M) = x_1 + x_2 + M + M = (x_1 + x_2) + M = M_{x_1 + x_2}, \qquad (\dagger)$$

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 osztállyal.*

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

$$\alpha * M_x = M_{\alpha \cdot x} \tag{‡}$$

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 osztály lesz, sőt *egy vektorhoz tartozó ekvivalencia osztály α szorosa azonos e vektor α-szorosához tartozó ekvivalencia osztállyal*.

6.4. állítás (faktortér)**.** *Legyen V egy az $\mathbb{F}$ test feletti vektortér, és $M \subseteq V$ egy rögzített altér. Legyen $\sim$ a V vektortér elemein értelmzett reláció, amelyre $x \sim y$, ha $x - y \in M$ ekvivalencia reláció. 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 a Minkowski–összeget és a skalárral való szorzást:

$$M_{x_1} + M_{x_2} = M_{x_1 + x_2}, \qquad \text{és} \qquad \alpha * M_x = M_{\alpha \cdot x}.$$

Ekkor az itt bevezetett $(V/M, +, *)$ struktúra az $\mathbb{F}$ test feletti vektorteret alkot. Ezt a vektorteret nevezzük a V vektortér M 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 közt $M_0 = M$ 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 † azonosságot.

A $*$ szorzás és az összeadás kapcsolatát leíró aximák ellenőrzéséhez használjuk (†) és (‡) 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. ▪

6.5. állítás. *Legyen M a V vektortér egy altere. Definiálja $\varphi : V \to V/M$ az $x \mapsto M_x$ függvényt.*

1. *Ekkor $\varphi : V \to V/M$ 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 $\Phi : N \to V/M$ egy izomorfizmus.* ⌟

Bizonyítás: 1. A (†) és (‡) 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 injektiv 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. ▪

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, izomorfiától eltekintve csak egyetlen egy létezik. De van-e, mindig 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 $\operatorname{codim} M = k$ módon jelöljük. ⌟

6.7. állítás. *Tegyük fel, hogy M a V vektortér olyan altere, amelynek van véges dimenziós direkt kiegészítője. Ekkor*

$$\operatorname{codim} M = \dim(V/M).$$ ⌟

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 $\dim V/M = \dim N = \operatorname{codim} M$. ▪

6.8. állítás. *Legyenek V, W vektorterek, továbbá $A \in L(V, W)$ egy olyan lineáris operáció, amelyre az* $\operatorname{Im} A \subseteq W$ *végesen generált. Ekkor a* $\ker$ *A altérnek van végesen generált direkt kiegészítője, és*

$$\operatorname{codim}(\ker A) = \dim(\operatorname{Im} A).$$ ⌟

Bizonyítás: Legyen $\operatorname{Im} A$ egy bázisa $\{Ax_1, \dots, Ax_r\}$. Definiálja $N = \operatorname{lin}\{x_1, \dots, x_r\}$. 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.

Ha $x \in V$ tetszőleges vektor, akkor $Ax \in \operatorname{Im} A$, 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 $\operatorname{codim}(\ker A) = \dim N = r = \dim(\operatorname{Im} A)$. ▪

Magától értetődik a következő észrevétel.

6.9. 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.* ⌟

Nyilvánvaló következmény még a rang-defektus–tétel is.

6.10. 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* Im *A is véges dimenziós, és*

$$\dim(\ker A) + \dim(\operatorname{Im} A) = \dim V.$$ ⌟

Ha V véges dimenziós és $A \in L(V, W)$ egy lineáris operáció, akkor láttuk, hogy

$$\dim(\operatorname{Im} A) = \operatorname{codim}(\ker A) = \dim(V/\ker A).$$

Tudjuk, hogy azonos dimenziójú vektorterek egy mással izomorfak, emiatt $\operatorname{Im} A$ izomorf $V/\ker A$ faktortérrel. A következő állítás ennek a gondolatnak a nem véges dimenziós esetre vonatkozó általánosítása, de véges dimenzióban is érdekes, hiszen konkrét izomorfiát mutatunk.

6.11. 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 $\operatorname{Im} A$ vektortérrel.* ⌟

Bizonyítás: Jelölje $M = \ker A$ a V vektortér alterét. Definiálni szeretnénk egy $\Omega : \operatorname{Im} A \to V/M$ 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 $y \in \operatorname{Im} A$ mellett

$$y \mapsto M_x, \text{ ahol } Ax = y$$

jól definiálja az $\Omega : \operatorname{Im} A \to V/M$ 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

$$\Omega(y) = M_x = H,$$

amivel azt igazoltuk, hogy minden ekvivalencia osztály egy $\operatorname{Im} A$ 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. ▪

Rang-tétel mégegszer

6.12. definíció (rang, oszloprang, sorrang, feszítőrang)**.** Egy véges vektorrendszer *rangján* a vektorrendszer generálta altér dimenzióját értjük. Egy mátrix *oszloprangján* a mátrix oszlopai mint vektorrendszer rangját értjük. Egy mátrix *sorrangján* a mátrix sorai mint vektorrendszer rangját értjük. Ha $A \in \mathbb{F}^{n\times m}$ nemzérus mátrix, akkor legkisebb olyan r számot, amelyre létezik $B \in \mathbb{F}^{n\times r}$ és $C \in \mathbb{F}^{r\times m}$ mátrix úgy, hogy $A = BC$ szorzatfelbontás teljesül, az A mátrix *feszítőrangjának* nevezzük.

Jelölések: Ha $\{x_1, \dots, x_m\}$ a szóban forgó vektorrendszer, akkor

$$\operatorname{rank}\{x_1, \dots, x_m\} = \dim \operatorname{lin}\{x_1, \dots, x_m\}$$

Ha $A \in \mathbb{F}^{n\times m}$ egy mátrix, akkor

$$\operatorname{crank} A = \operatorname{rank}\left\{[A]^j : j = 1, \dots, m\right\}, \quad \operatorname{rrank} A = \operatorname{rank}\{[A]_k : k = 1, \dots, n\},$$

továbbá srank A jelöli a feszítőrangját A-nak. ⌟

6.13. állítás. (*Rang-tétel*) *Tetszőleges test feletti tetszőleges mátrix mellett, a fent bevezetett három rang-koncepció azonos.*

Formálisabban: Minden $A \in \mathbb{F}^{n\times m}$ mellett

$$\operatorname{crank} A = \operatorname{srank} A = \operatorname{rrank} A$$ ⌟

Bizonyítás: Induljunk ki a feszítőrang fogalmából. Legyen $r = \operatorname{srank} A$, és

$$A = BC, \tag{†}$$

ahol $B \in \mathbb{F}^{n\times r}, C \in \mathbb{F}^{r\times m}$. Azt mutatjuk meg, hogy ekkor B oszloprendszere minimális generátorrendszere A oszlop-vektorterének és C sorrendszere minimális generátorrendszere A sor-vektorterének.

Vegyük észre, hogy $\operatorname{srank} A \leq \operatorname{crank} A$. Ugyanis ha az A mátrix oszlop-vektorterének veszünk egy tetszőleges választott $b_1, \dots, b_k$ generátorrendszerét, akkor van $B \in \mathbb{F}^{n\times k}$ és $C \in \mathbb{F}^{k\times m}$ mátrix, hogy $A = BC$. Az $r = \operatorname{srank} A$ szám az ilyen k számok legkisebbike, tehát valóban $r \leq \operatorname{crank} A$.
Most tekintsük a (†)-ben rögzített szorzatot. Jelölje W a B mátrix és V az A mátrix oszlopvektorterét. Mivel BC oszlopai B oszlopainak lineáris kombinációja, ezért A minden oszlopa beleesik W-be, így az A oszlopainak lineáris burka is részhalmaza W-nek, azaz

$$V \subseteq W.$$

A B mátrixnak r darab oszlopa van, tehát $\dim W \leq r$. Látjuk tehát, hogy

$$\dim W \leq r \leq \operatorname{crank} A = \dim V,$$

ami csak úgy lehetséges, hogy $V = W$. A B oszlopai tehát V-nek is generátorrendszerét alkotják, és r minimalitása szerint egy elem sem elhagyható a generátorrendszer tulajdonság elvesztése nélkül.

A sorokra vonatkozó indoklás analóg. Először is $\operatorname{srank} A \leq \operatorname{rrank} A$. Ugyanis ha az A mátrix sorvektorterének veszünk egy tetszőleges választott $c_1, \dots, c_k$ generátorrendszerét, akkor van $B \in \mathbb{F}^{n\times k}$ és $C \in \mathbb{F}^{k\times m}$ mátrix, hogy $A = BC$. Az $r = \operatorname{srank} A$ szám az ilyen k számok legkisebbike, tehát valóban $r \leq \operatorname{rrank} A$.
Most tekintsük a (†)-ben rögzített szorzatot. Jelölje W a C mátrix és V az A mátrix sorvektorterét. Mivel

BC sorai C sorainak lineáris kombinációja, ezért A minden sora W-be esik, így az A sorainak lineáris burka is részhalmaza W-nek, azaz

$$V \subseteq W.$$

A C mátrixnak r sora van, tehát $\dim W \leq r$. Látjuk tehát, hogy

$$\dim W \leq r \leq \operatorname{crank} A = \dim V,$$

ami csak úgy lehetséges, hogy $V = W$. A C oszlopai tehát V-nek is generátorrendszerét alkotják, és r minimalitása szerint egy elem sem elhagyható a generátorrendszer tulajdonság elvesztése nélkül.

Ezt kellett belátni. ▪

6.14. definíció (mátrix rangja)**.** Mivel a sorrang, az oszloprang, a feszítőrang minden mátrix mellett azonos, ezért a továbbiakban a közös értékre a *rang* szót is használjuk.[1] ⌟

6.15. *megjegyzés.* Érdemes a rang-tétel következő összegzését megjegyezni. Legyen $A \in \mathbb{F}^{n\times m}$ mátrix, amelynek r a rangja. Ekkor létezik $A = BC$ felbontása, ahol $B \in \mathbb{F}^{n\times r}, C \in \mathbb{F}^{r\times m}$. Ez a felbontás persze nem egyértelmű, hiszen A oszlop-vektorterének nagyon sok bázisa van. Viszont minden ilyen felbontásban B oszloprendszere az A oszlop-vektorterének, míg C sorrendszere az A sorvektorterének minimális generátorrendszerét, ergo bázisát alkotja. ⌟

Következményképpen érdemes meggondolni a mátrix és inverzének felcserélhetőségére vezető állítást.

6.16. állítás. *Legyenek $A, B \in \mathbb{F}^{n\times n}$ négyzetes mátrixok, amelyekre AB=I. Ekkor BA=I is teljesül.* ⌟

Bizonyítás: Az identitás mátrix rangja nyilván n. E mátrix feszítőrangjának definíciójára gondolva, az előző megjegyzés szerint A oszlopai $\mathbb{F}^n$ lineárisan független rendszerét alkotják. Vegyük észre, hogy a mátrix szorzás asszociativitását is kihasználva

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

Na most, ha $BA \neq I$ lenne, akkor a $BA - I$ mátrixnak lenne egy olyan nem zérus oszlopa, melynek elemeivel mint együtthatókkal képzett lineáris kombinációja az A oszlopainak a zéró vektort eredményezi. Ez persze ellentmond az A oszloprendszer lineáris függetlenségének, tehát $BA = I$ valóban fennáll.[2] ▪

6.17. definíció-állítás (invertálható mátrix)**.** Legyen $A \in \mathbb{F}^{n\times n}$ egy négyzetes mátrix. Az alábbi feltételek egymással ekvivalensek.

1. Van egyetlen olyan $B \in \mathbb{F}^{n\times n}$ mátrix, amelyre $AB = I$,
2. Van olyan $B \in \mathbb{F}^{n\times n}$ mátrix, amelyre $AB = I$,
3. Van egyetlen olyan $B \in \mathbb{F}^{n\times n}$ mátrix, amelyre $BA = I$,
4. Van olyan $B \in \mathbb{F}^{n\times n}$ mátrix, amelyre $BA = I$,
5. $\operatorname{rank} A = n$,
6. A oszlopai lineárisan független rendszer alkotnak,
7. A sorai lineárisan független rendszert alkotnak.

Ha a fenti feltételek egyike (ergo mindegyike) fennáll, akkor azt mondjuk, hogy A mátrix *invertálható*. Szinonimaként használjuk még a *nemszinguláris*, vagy az *reguláris* szavakat. Ha egy mátrix nem invertálható, akkor *szinguláris*nak nevezzük.

Egy invertálható négyzetes mátrix esetén azt az egyetlen B mátrixot, amelyre $AB = I$ fennáll az A inverzének mondjuk, és $A^{-1} = B$-vel jelöljük. Világos, hogy

$$AA^{-1} = I = A^{-1}A, \quad (A^{-1})^{-1} = A.$$ ⌟

Bizonyítás: A 2., 4., 5., 6., 7. állítások ekvivalenciája nyilvánvaló az előzőek szerint. Ha $AB = I = AC$, akkor $A\,(B - C) = 0$ így A oszloprendszere lineáris függetlensége miatt $B = C$. Ezzel 2. $\Rightarrow$ 1. implikációt is beláttuk. Az 1. és 3. feltevések ekvivalenciája az előző állítás miatt teljesül. ▪

[1] Lásd: (Wardlaw 2005)

[2] A feszítőrang fogalmának ismerete nélküli – talán még elemibb – bizonyítás: (Paparella 2017)

Elemi fogalmak mégegyszer

Tegyük fel, hogy, hogy ismerjük az alábbi fogalmakat:

1. Vektorrendszer lineáris függetlensége;
2. Altér, lineáris burok, generátorrendszer;
3. Mátrix szorzás,
4. Gauss-Jordan elimináció. Pontosan azt tesszük fel, hogy ha Q egy négyzetes mátrix, melynek oszlopai lineárisan függetlenek, akkor az elemi sor műveletekkel az identikus mátrixszá transzformálható. Mivel az elemi sorműveletek, bal-szorzások alkalmas mátrixszal, azt kapjuk, hogy létezik P mátrix, amelyre $PQ = I$.

Amit határozottan kerülünk, és azt tesszük fel, hogy nem ismerjük,

1. Bázisok fogalma,
2. Különböző bázisok azonos számossága,
3. determináns.

Külön probléma a mátrix rangjának definíciója. Nem definiálhatom, mint az oszlop vagy sorvektortér dimenzióját, hiszen a felépítés jelen szintjén még nincs bázis. Természetesen azt sem tudjuk még, hogy a maximális lineárisan független sor- vagy oszloprendszer választástól függetlenül mindig azonos elemszámú. A mátrix feszítő rangja viszont definiálható.

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

$$A = BC$$

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

Világos, hogy tetszőleges nemzérus négyzetes mátrixra ez jól definiált és $1 \leq r \leq \min\{n, m\}$. A rangtételnek a szokásosnál egy kicsit erősebb formáját lehet megfogalmazni a dimenzió fogalmának bevezetése nélkül, ami a lenti 3. állítás.

6.19. állítás. *Az alábbi állítások egymás következményei:*

1. *Homogén lineáris egyenletrendszernek, amelynek több ismeretlene van, mint egyenlete mindig létezik nem triviális megoldása.*
2. *Lineárisan független vektorrendszer elemszáma nem nagyobb mint egy generátorrendszer elemszáma.*
3. *Minden nemzérus mátrixban a maximális lineárisan független oszloprendszerek és maximális lineárisan független sorrendszerek azonos elemszámúak, és ez a szám egybeesik a mátrix feszítőrangjával.*
4. *Legyen $A, B \in \mathbb{F}^{n\times n}$ négyzetes mátrixok. Ekkor $AB = I$ esetén $BA = I$ is fennáll.* ⌟

1. ⇒*2.:* Legyen $\{y_1, \dots, y_n\}$ egy generátorrendszer, és $\{x_1, \dots, x_m\}$ olyan vektorrendszer a vektortérben, ahol $m > n$. Meg kell mutatnunk, hogy ez utóbbi egy lineárisan összefüggő. Világos, hogy minden $1 \leq k \leq m$ mellett

$$x_k = \sum_{j=1}^{n} \alpha_{j,k} y_j.$$

Egyelőre tetszőleges $\xi_1, \dots, \xi_m$ együtthatók mellett

$$\sum_{k=1}^{m} \xi_k x_k = \sum_{k=1}^{m} \sum_{j=1}^{n} \xi_k \alpha_{j,k} y_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{m} \alpha_{j,k} \xi_k \right) y_j \tag{6.1}$$

Tekintsük az $(\alpha_{j,k})$ együtthatók generálta homogén lineáris egyenletrendszert. Itt $j = 1, \dots, n$ és $k = 1, \dots, m$. Mivel $m > n$, ezért az ismeretlenek száma több mint az egyenletek száma. Létezik tehát nem triviális megoldás, azaz léteznek nem mind nulla $\xi_1, \dots, \xi_m$ számok, amelyekre minden $j = 1, \dots, n$ esetén

$$\sum_{k=1}^{m} \alpha_{j,k} \xi_k = 0.$$

Találtunk tehát az $\{x_1, \dots, x_m\}$ vektorrendszernek egy nem triviális, de a zéró vektort eredményező, lineáris kombinációját (6.1). ∎

2.⇒*3.:* 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. A 2. állítás 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 sorrendszere egy lineárisan független rendszer e V vektortérben, és C sorai pedig egy generátorrendszert alkotnak ugyanebben a V vektortérben. A 2. állítás szerint $r_w \leq r$.

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

3.⇒*4.:* Tegyük fel, hogy $AB = I$. Az identitás mátrix rangja nyilván n. A 3. állítás miatt a feszítőrang is n. Ha A oszlopai nem lennének lineárisan függetlenek, akkor lenne $A = CD$ felbontás, ahol $C \in \mathbb{F}^{n \times r}$ és $D \in \mathbb{F}^{r \times n}$ valamely $r < n$ mellett. Ekkor persze $I = AB = (CD)\,B = C\,(DB)$ is teljesülne, ahol $DB \in \mathbb{F}^{r \times n}$ ellentmondva az identitás mátrix feszítőrangja definíciójának. Világos, hogy

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

Figyelembe véve, hogy A oszlopai lineáris függetlenek, ez csak úgy lehetséges, ha $BA - I$ a zéró mátrix, ergo $BA = I$. ∎

4.⇒1.: Legyen $A \in \mathbb{F}^{n\times m}$ a homogén lineáris egyenletrendszer együttható mátrixa, ahol n az egyenletek száma, m az ismeretlenek száma. Azt kell megmutatnunk, hogy az oszloprendszer lineárisan összefüggő. Ha független lenne, akkor egészítsük ki e mátrixot alulról $m-n$ darab csupa nullákat tartalmazó sorral. Mivel $m > n$ ezért a kiegészített $Q \in \mathbb{F}^{m\times m}$ mátrix legalsó sora csak nullát tartalmaz. Mivel A oszlopai lineárisan függetlenek, ezért Q oszlopai is azok. Emiatt létezik $P \in \mathbb{F}^{m\times m}$ mátrix, amelyre $PQ = I$. A 3. állitás szerint $I = QP$ is teljesül, ami azt jelenti, hogy I utolsó sora a csupa nullákat tartalmaz, ami ellentmondás. Beláttuk tehát, hogy A oszlopai lineárisan összefüggőek, azaz az eredeti egyenletrendszernek van nem triviális megoldása. ▪

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 $\mathbb{F}^m \to \mathbb{F}^n$ 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–tétel 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, $\rho(A) = \dim(\operatorname{Im} A)$ 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 az 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 elő áll, mit a By vektor A képe.

Megfordítva, ha A szürjektív, akkor a) szerint injektív is, van tehát $V \to V$ inverze. De lineáris függvény inverze is lineáris, így a $B = A^{-1}$ jelölést bevezetve találtunk $B \in L(V)$ transzformációt, amelyre $B \circ A = I$.

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. ∎

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

$$\{A_{i,j} : i = 1, \ldots, n; j = 1, \ldots, m\}$$

mátrixok rendszere bázisa az $\mathbb{F}^{n\times m}$ térben. Így persze $\dim \mathbb{F}^{n\times m} = n \cdot m$.

Most áttérünk 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 $\{v_1, \ldots, v_m\}$ egy bázis, és rögzítsünk W-ben egy tetszőleges m elemű $\{w_1, \ldots, w_m\}$ vektorrendszert.*

Ekkor létezik egyetlen egy $A \in L(V, W)$ lineáris operáció, amelyre $Av_j = w_j$ minden $j = 1, \ldots, m$ esetén. ⌟

Bizonyítás: Definiáljuk az $A : V \to W$ függvényt a következőképpen. Minden $v \in V$ egyértelműen áll elő mint $v = \sum_{j=1}^{m} \alpha_j v_j$ alakban. Egy ilyen v mellett legyen

$$A(v) = \sum_{j=1}^{m} \alpha_j w_j.$$

Világos, hogy $A : V \to W$ 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 $x_1 = \sum_{j=1} \alpha_j v_j$ és $x_2 = \sum_{j=1} \beta_j v_j$. 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 $v_j = 0v_1 + \ldots + 1v_j + \ldots + 0v_m$, tehát az A függvény definíciója értelmében

$$Av_j = w_j$$

valóban fennáll minden $j = 1, \ldots, m$ 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 $\{e_1, \ldots, e_m\} \subseteq V$ és az $\{f_1, \ldots, f_n\} \subseteq W$ bázisok rögzítve. Definiáljuk valamely rögzített $i \in \{1, \ldots, n\}$ és valamely rögzített $j \in \{1, \ldots, m\}$ 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 $k = 1, \ldots, m$. Ekkor az $\{A_{i,j} : i = 1, \ldots, n; j = 1, \ldots, m\}$ 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 generátorrendszere 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 j-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, egy 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. ∎

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 $\{y, \dots, y_n\}$ 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:

$$\{\underbrace{A_{1,1}, A_{2,1}, \dots, A_{n,1}}_{j=1}, \underbrace{A_{1,2}, A_{2,2}, \dots, A_{n,2}}_{j=2}, \underbrace{A_{1,3}, A_{2,3}, \dots, A_{n,3}}_{j=3}, \underbrace{\dots, \dots, \dots, \dots, \dots}_{j=4,\dots,m-1}, \underbrace{A_{1,m}, A_{2,3}, \dots, A_{n,m}}_{j=m}\}$$

Láttuk, hogy a fent 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

$$[A] = \begin{pmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \vdots \\ \alpha_{n,1} \\ \alpha_{1,2} \\ \alpha_{2,2} \\ \vdots \\ \alpha_{n,2} \\ \alpha_{1,3} \\ \alpha_{2,3} \\ \vdots \\ \alpha_{n,3} \\ \vdots \\ \vdots \\ \alpha_{1,m} \\ \alpha_{2,m} \\ \vdots \\ \alpha_{n,m} \end{pmatrix} \tag{†}$$

Mint minden vektortér így az $L(V,W)$ is izomorf a koordináta-terével, ergo $L(V,W)$ izomorf az $\mathbb{F}^{n \cdot m}$ vektortérrel.

Érdemes a koordinátatér elemeit, azaz az $n \cdot m$ elemből álló oszlopvektorokat n koordinátánként megtörni, és evvel egy $n \times m$-es mátrixba rendezni. Ezt a jelölést alkalmazva az A lineáris operációnak 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 $\{f_1, \dots, f_m\}$ bázisok mellett. ⌟

Meggondoltuk tehát, hogy a kapcsolat lineáris operáció és mátrixa között azonos avval a kapcsolattal, ami általában egy vektor és koordinátái közt van.

Nagyon fontos, hogy egy lineáris operáció mátrixát fel tudjuk írni, emiatt összefoglalom az eddigieket.

7.5. állítás. *Tegyük fel, hogy $A \in L(V,W)$ egy lineáris operáció. Rögzítsük 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 $[A]_{\substack{\{e_1,\dots,e_m\} \\ \{f_1,\dots,f_n\}}}$ mátrix,* [1] *amelynek j-edik oszlopa az Ae_j vektor $\{y_1, \dots, y_n\}$ bázisban felírt koordinátája. Formálisabban:*

$$[A]^{j}_{\substack{\{e_1,\dots,e_m\} \\ \{f_1,\dots,f_n\}}} = [Ae_j]_{\{y_1,\dots,y_n\}}. \qquad ⌟$$

[1] Ha világos, hogy mely bázisokat rögzítjük akkor a nehézkes $[A]_{\substack{\{e_1,\dots,e_m\} \\ \{f_1,\dots,f_n\}}}$ jelölés helyett csak $[A]$-t írunk. Persze mindig tudnunk kell, hogy az A lineáris operátor mely bázisokban felírt mátrixáról van szó.

Speciálisan, ha $V = W$, akkor $A : V \to V$ lineáris transzformációról beszélünk. Amikor egy lineáris transzformáció mátrixáról van szó, akkor az mindig 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 $[A]_{\substack{\{e_1,\dots,e_m\}\\\{e_1,\dots,e_m\}}}$ mátrixára: [2]

$$[A]^{j}_{\substack{\{e_1,\dots,e_m\}\\\{e_1,\dots,e_m\}}} = [Ae_j]_{\{e_1,\dots,e_m\}}\,.$$

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

$$[\alpha A + \beta B] = \alpha\,[A] + \beta\,[B]\,. \tag{7.2}$$

Izomorf struktúrák közt nem érdemes különbséget tenni, viszont vigyáznunk kell arra, hogy olyan fogalmakat definiáljunk, 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 operátornak 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*

$$[Ax]_{\{f_1,\dots,f_n\}} = [A]_{\substack{\{e_1,\dots,e_n\}\\\{f_1,\dots,f_n\}}} \cdot [x]_{\{e_1,\dots,e_m\}}.$$

Emiatt $\operatorname{Im} A$ *és* $\operatorname{Im}[A]$ *egymással izomorf alterek, hasonlóan* $\ker A$ *és* $\ker[A]$ *egymással izomorf vektorterek, így dimenziójuk is azonos. Konkrétan A lineáris operátornak 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

$$[Ax]_{\{f_1,\dots,f_n\}} = \left[\sum_{j=1}^{n} \xi_j Ae_j\right]_{\{f_1,\dots,f_n\}} = \sum_{j=1}^{n} \xi_j [Ae_j]_{\{f_1,\dots,f_n\}} = \sum_{j=1}^{n} \xi_j [A]^{j}_{\substack{\{e_1,\dots,e_n\}\\\{f_1,\dots,f_n\}}} =$$
$$[A]_{\substack{\{e_1,\dots,e_n\}\\\{f_1,\dots,f_n\}}} \cdot [x]_{\{f_1,\dots,f_n\}}.$$

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 Im A és 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 $\rho(A) = \rho([A])$. ▪

7.3. Lineáris operátorok szorzata

Emlékezzünk arra, hogy a lineáris operáció mátrixát úgy kaptuk, hogy a koordinátáit n elemenként megtörve az oszlop vektort egy mátrixszá rendeztük át. Ennek az átrendezésnek az igazi értelme, hogy ilyen módon a mátrix szorzás művelet a lineáris operátorok kompozíciójával kapcsolódik össze.

[2] A nehézkes $[A]_{\substack{\{e_1,\dots,e_m\}\\\{e_1,\dots,e_m\}}}$ 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 mátrixának szorzata. Formálisabban:

$$[C]^{\{e_1,\dots,e_m\}}_{\{f_1,\dots,f_n\}} = [A]^{\{z_1,\dots,z_r\}}_{\{f_1,\dots,f_n\}} \cdot [B]^{\{e_1,\dots,e_m\}}_{\{z_1,\dots,z_r\}}.$$ ⌟

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 jobb oldal összehasonlítása is értelmes.

Már csak azt kell meggondolni, hogy a bal oldali mátrix minden eleme azonos a jobb oldali 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

$$\sum_{i=1}^{n} \gamma_{i,j} f_i = Ce_j = A(Bej) = \sum_{k=1}^{r} \beta_{k,j} Az_k = \sum_{k=1}^{r} \beta_{k,j} \left(\sum_{i=1}^{n} \alpha_{i,k} f_i \right) =$$
$$\sum_{k=1}^{r} \sum_{i=1}^{n} \beta_{k,j} \alpha_{i,k} f_i = \sum_{i=1}^{n} \sum_{k=1}^{r} \alpha_{i,k} \beta_{k,j} f_i = \sum_{i=1}^{n} \left(\sum_{k=1}^{r} \alpha_{i,k} \beta_{k,j} \right) f_i.$$

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

$$[C]_{i,j} = \gamma_{i,j} = \sum_{k=1}^{r} \alpha_{i,k} \beta_{k,j} = ([A] \cdot [B])_{i,j}.$$ ▪

Azért, hogy teljes legyen az analógia a lineáris operátorok kompozíciója és a mátrixok szorzása műveletek közt, a lineáris operátorok $A \circ B$ kompozícióját 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

$$[AB] = [A][B]. \tag{7.3}$$

7.9. definíció (lineáris operátor hatványai)**.** A lineáris transzformációk a bevezetett szorzás művelettel egységelemes gyűrűt alkotnak, ahol az egységelem az $I : V \to V$ 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 operációnak akár melyik 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 A pontosan akkor izomorfizmus, ha $[A]$ reguláris és $\left[A^{-1}\right] = [A]^{-1}$.* ⌟

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 $[A]^{-1} = [B] = \left[A^{-1}\right]$. ▪

A (7.2.) és a (7.3.) szerint egy lineáris operátorhoz hozzárendeli valamely báziskban felírt mátrixát egy olyan bijekció, ami tartja a gyűrű műveleteket. Az $L(V)$ és a $F^{\dim V \times \dim V}$ tehát olyan egységelemes (nem kommutatív) gyűrűk, amelyek közt van a gyűrű műveleteket tartó bijekció (gyűrű-izomorfizmus).

8. fejezet

Általános bázis-transzformáció

EGY VEKTOR 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 $\{e_1, \ldots, e_n\}$ bázist régi bázisnak, és nevezzük az $\{f_1, \ldots, f_n\}$ 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 transzformációnak 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 operációnak a régi bázis páron felírt mátrixát, akkor hogyan számítható A-nak az valamely új bázis páron felírt mátrixa?

8.1. definíció. Legyenek az $\{e_1, \ldots, e_n\}$ és $\{f_1, \ldots, f_n\}$ bázisok rögzítve. Tekintsük azt a B lineáris transzformációt, amelyre $Be_j = f_j$ teljesül minden $j = 1, \ldots, n$ mellett. Ezt a lineáris transzformációt nevezzük az $\{e_1, \ldots, e_n\}$ bázisról az $\{f_1, \ldots, f_n\}$ 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 oszlopa a $Be_j = f_j$ vektornak a régi bázisban felírt koordinátája. Mivel B szürjektív, ergo injektiv 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 $\{e_1, \ldots, e_m\}$ a régi bázis, és $\{f_1, \ldots, f_n\}$ az új bázis. Jelölje B a régi bázisról az új bázisra való áttérést Ekkor tetszőleges $x \in V$ vektor mellett*

$$[x]_{új} = [B]^{-1}_{régi} \cdot [x]_{régi}\,. \qquad ⌟$$

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

$$[x]_{\text{régi}} = \left[\sum_{j=1}^{n} \alpha_j f_j\right]_{\text{régi}} = \sum_{j=1}^{n} \alpha_j \, [f_j]_{\text{régi}} = \sum_{j=1}^{n} \alpha_j \, [B]^{j}_{\text{régi}} = [B]_{\text{régi}} \cdot [x]_{\text{új}}\,.$$

A kívánt azonosságot kapjuk, ha balról szorozzuk mindkét oldalt $[B]^{-1}_{\text{régi}}$ inverz mátrixszal. ∎

[1] Könnyen 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 bázis párban

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 tér áttérés lineáris transzformációját és $D \in L(W)$ a W tér áttérés transzformációját. Ekkor*

$$[A]_{\textit{új}} = [D]^{-1}_{\textit{régi}}\,[A]_{\textit{régi}}\,[B]_{\textit{régi}}\,. \qquad \lrcorner$$

Bizonyítás: Azt mutatjuk meg, hogy a bal oldali mátrix és a jobb oldali szorzat mátrix megfelelő oszlopai megegyeznek. A j-edik oszlopra:

$$\begin{aligned}[A]^j_{\text{új}} = [A]^j_{\substack{\{v_1,\dots,v_m\}\\ \{w_1,\dots,w_n\}}} &= [Av_j]_{\{w_1,\dots,w_n\}} \\ &= [D]^{-1}_{\{f_1,\dots,f_n\}} \cdot [Av_j]_{\{f_1,\dots,f_n\}} = [D]^{-1}_{\{f_1,\dots,f_n\}} \cdot [A]_{\substack{\{e_1,\dots,e_m\}\\ \{f_1,\dots,f_n\}}} \cdot [v_j]_{\{e_1,\dots,e_m\}} \\ &= [D]^{-1}_{\{f_1,\dots,f_n\}} \cdot [A]_{\substack{\{e_1,\dots,e_m\}\\ \{f_1,\dots,f_n\}}} \cdot [B]^j_{\{e_1,\dots,e_m\}} \\ &= [D]^{-1}_{\text{régi}}\,[A]_{\text{régi}}\,[B]^j_{\text{régi}} = \left([D]^{-1}_{\text{régi}}\,[A]_{\text{régi}}\,[B]_{\text{régi}}\right)^j.\end{aligned}$$

Közben használtuk a mátrix szorzás művelet asszociativitását, és azt a tényt, hogy tetszőleges $[E]$, $[F]$ mátrixok mellett $([E]\,[F])^j = [E]\left([F]^j\right)$. .

8.3. Lineáris transzformáció mátrixa az új bázisban

8.4. állítás. *Legyen $A \in L(V)$ egy lineáris transzformáció. Ekkor*

$$[A]_{\textit{új}} = [B]^{-1}_{\textit{régi}}\;[A]_{\textit{régi}}\,[B]_{\textit{régi}}\,. \qquad \lrcorner$$

Bizonyítás: A lineáris operátorokra vonatkozó állítás speciális esete, mikor $W = V$ és $D = B$. .

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*

$$[B]_{\textit{régi}} = [B]_{\textit{új}} \quad \textit{és persze} \quad \left[B^{-1}\right]_{\textit{régi}} = \left[B^{-1}\right]_{\textit{új}}\,. \qquad \lrcorner$$

9. fejezet

Invariáns alterek

9.1. definíció. Legyen $A \in L(V)$ egy lineáris transzformációja a V vektortérnek, és $M \subseteq V$ a tér 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 altér A-invariáns, vagy azt, hogy az M altér invariáns az A transzformációra nézve.

Úgy is fogalmazhatnánk, hogy az M altér akkor invariáns altér, ha az A transzformációnak az M-re való $A|_M$ megszorítása az M altér egy lineáris transzformációja, azaz

$$A|_M \in L(M).$$

Nyilvánvaló példa invariáns alterekre M és $\{0\}$. Tetszőleges A lineáris transzformáció mellett $\ker A$ és $\operatorname{Im} A$ mindig invariáns alterek. Ugyanis, ha $u \in \ker A$, akkor $A(Au) = A0 = 0$, tehát $Au \in \ker A$. Az $\operatorname{Im} A$ altér esete még egyszerűbb: A tér minden elemének képe $\operatorname{Im} A$-ban van, speciálisan persze ez $\operatorname{Im} A$ elemeire is igaz.

A célunk, hogy a teret előállítsuk lehetőleg minél alacsonyabb dimenziós terek direktösszegeként.

9.2. definíció (legszűkebb invariáns altér)**.** Legyen $A \in L(V)$ egy lineáris transzformáció. Világos, hogy akárhány A-ra nézve invariáns altér metszete is A-invariáns altér, így egy $H \subseteq V$ halmazt tartalmazó *legszűkebb A-invariáns altér* a H halmazt tartalmazó A-invariáns alterek közös része. Formálisan

$$\operatorname{lin}(H; A) = \bigcap_{\substack{H \subseteq N \\ N \text{altér} \\ N \text{invariáns}}} N. \qquad ⌟$$

A legérdekesebb eset, amikor H egy elemű halmaz. A $H = \{v\}$ esetben a kissé nehézkes $\operatorname{lin}(\{v\}; A)$ helyett egyszerűbben $\operatorname{lin}(v; A)$-t írunk.

9.3. állítás. *Egy $A \in L(V)$ lineáris transzformációra és egy $v \in V$ vektorra*

$$\operatorname{lin}(v; A) = \operatorname{lin}\{v, Av, A^2v, \dots\}. \qquad ⌟$$

Bizonyítás: Mivel $\operatorname{lin}(v; A)$ egy v-t tartalmazó A-invariáns altér, ezért $\{v, Av, \dots, A^kv, \dots\} \subseteq \operatorname{lin}(v; A)$, amiatt

$$\operatorname{lin}\{v, Av, A^2v, \dots\} \subseteq \operatorname{lin}(v; A).$$

Most vegyük észre, hogy $\operatorname{lin}\{v, Av, A^2v, \dots\}$ is egy v-t tartalmazó A-invariáns altér és $\operatorname{lin}(v; A)$ ilyenek közt a legszűkebb, ezért

$$\operatorname{lin}(v; A) \subseteq \operatorname{lin}\{v, Av, A^2v, \dots\}. \qquad .$$

9.4. lemma. *Tegyük fel, hogy egy $A \in L(V)$ lineáris transzformációra és egy $v \in V$ vektorra a*

$$\left\{v, Av, \dots, A^kv\right\}$$

lineárisan összefüggő ($k \geq 1$). *Ekkor minden* $n \geq k$ *mellett*

$$A^n v \in \operatorname{lin}\left\{v, Av, \dots, A^{k-1}v\right\}. \qquad \lrcorner$$

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

$$\left\{v, Av, \dots, A^{t-1}v\right\} \text{ lineárisan független} \tag{†}$$

de

$$\left\{v, Av, \dots, A^{t-1}v, A^t v\right\} \text{ lineárisan összefüggő.} \tag{‡}$$

Világos, hogy $1 \leq t \leq k$. Jelölje $N = \operatorname{lin}\left\{v, Av, \dots, A^{t-1}v\right\}$.

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 (‡) 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 (†) 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

$$A^{t+m+1}v = \sum_{j=0}^{t-1} \alpha_j A^{j+1} v = \sum_{j=1}^{t} \alpha_{j-1} A^j v = \left(\sum_{j=1}^{t-1} \alpha_{j-1} A^j v\right) + \alpha_{t-1} A^t v \in N + N = N. \qquad .$$

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*

$$\left\{v, Av, \dots, A^{k-1}v\right\} \textit{ lineárisan független} \tag{†}$$

de

$$\left\{v, Av, \dots, A^{k-1}v, A^k v\right\} \textit{ lineárisan összefüggő.} \tag{‡}$$

A fenti (†) *rendszer bázisa* $\operatorname{lin}(v; A)$*-nak.* $\lrcorner$

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 $\left\{v, Av, A^2 v\right\}$ rendszert. Ha ez lineárisan összefüggő, akkor $k = 2$-vel készen vagyunk. Ha ez lineárisan független, akkor tekintsük a $\left\{v, Av, A^2 v, A^3 v\right\}$ vektorrendszert. 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 $\operatorname{lin}(v; A)$ egy v-t tartalmazó A-invariáns altér, ezért

$$\left\{v, Av, \dots, A^{k-1}v\right\} \subseteq \operatorname{lin}(v; A).$$

Az előző lemma szerint

$$\left\{v, Av, A^2 v, \dots\right\} \subseteq \operatorname{lin}\left\{v, Av, \dots, A^{k-1}v\right\}$$

ezért

$$\operatorname{lin}(v; A) = \operatorname{lin}\left\{v, Av, A^2 v, \dots\right\} \subseteq \operatorname{lin}\left\{v, Av, \dots, A^{k-1}v\right\} \subseteq \operatorname{lin}(v; A).$$

Ez azt jelenti, hogy a (†) vektorrendszer egy lineárisan független generátorrendszere a $\operatorname{lin}(v; A)$ térnek, tehát valóban bázisa is. .

A $\lin(v;A)$ tehát a fenti k-ra egy k-dimenziós altér, amelynek bázisa $\left\{A^{k-1}v,\dots,Av,v\right\}$. A játék kedvéért írjuk fel az $A|_{\lin(v;A)}$ transzformáció mátrixát ebben a bázisban:

$$\begin{array}{r|ccccc} & A^kv & A^{k-1}v & A^{k-2}v & \dots & Av \\ \hline A^{k-1}v & \alpha_{k-1} & 1 & 0 & \dots\dots & 0 \\ A^{k-2}v & \alpha_{k-2} & 0 & 1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ Av & \alpha_1 & 0 & 0 & \dots & 1 \\ v & \alpha_0 & 0 & 0 & \dots & 0 \end{array}, \text{ ahol } A^kv=\sum_{j=0}^{k-1}\alpha_jA^jv.$$

9.1. Transzformációk sajátértéke

A célunk, hogy a lehető legalacsonyabb dimenziós invariáns altereket találjunk. Így persze az 1 dimenziós 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 $\dim(\lin(v;A))=1$. 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

$$p(A) = \sum_{j=0}^{n} \alpha_j A^j. \qquad \lrcorner$$

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 $r(A) = p(A)\, q(A)$.* $\lrcorner$

Bizonyítás: Az első állítás azért teljesül, mert $\left(\alpha A^k + \beta A^k\right) = (\alpha + \beta) A^k$.

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. .

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.* $\lrcorner$

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* $\operatorname{Im} p(A)$ *is invariáns alterek.* $\lrcorner$

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 $v \in \operatorname{Im} p(A)$, 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 $Av \in \operatorname{Im} p(A)$. ∎

Speciálisan ez igaz az $t - \lambda$ polinomra is, amikor λ egy sajátértéke A-nak. Tehát a λ sajátértékhez tartozó $\ker(A - \lambda I)$ sajáltaltér egy 1 dimenziós 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 csak a $\{0\}$ elemet tartalmazó ideálnak van egyetlen legkisebb fokú és normált eleme, ami egyben az ideál generáló eleme is.

10.1. Kis minimál polinom

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 $\mathbb{F}[t]$ polinom gyűrű 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 polinom gyűrűnek.

Jelölje $n = \dim V$ és tekintsük az $n+1$ elemű $\{v, Av, \ldots, A^n v\}$ vektorrendszert. Mivel a Steinitz-lemma szerint $n+1$ vektor egy n-dimenziós vektortérben lineárisan összefüggő, ezért van $\alpha_0, \ldots, \alpha_n \in \mathbb{F}$ nem mind 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 $-\infty < \deg p \le n$. ∎

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 $J_{A,v}$. Ez a $J_{A,v}$ legkisebb fokú, normált polinomja. Ezt a polinomot nevezzük az A transzformáció, v vektorhoz tartozó *kis minimál polinomjának*. ⌟

10.8. állítás. *Ha n az A transzformáció v vektorhoz tartozó kis minimál polinomja, akkor*

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 $n|p$,*
4. *$\deg n \le \dim V$.* ⌟

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, es erre a polinomra

$$J_{A,v} = J(n) = \{hn : h \in \mathbb{F}[t]\}.$$

Ezt kellett belátni. ▪

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 minimál polinom.* ⌟

Bizonyítás: Ha maga $\deg p = 1$, akkor készen vagyunk, hiszen nem zérus vektornak minimálpolinomja legalább első fokú. Ha $\deg p > 1$ és p nem egyezne az n kis minimál polinommal, akkor $\deg n < \deg p$ lenne, és mivel n generálja a $J_{A,v}$ ideált, ezért $p = nh$ alakú lenne, ahol h is legalább első fokú. Így p két legalább első fokú polinom szorzata, azaz reducibilis lenne. ▪

A következő állítás módszert ad a kis minimál polinom meghatározására, amely mindig használható.

10.10. állítás. *Legyen $A \in L(V)$ és $v \neq 0$. Mivel V egy véges dimenziós altér, ezért létezik $1 \leq k \leq \dim V$, hogy $\{v, Av, \ldots, A^{k-1}v\}$ lineáris független, de $\{v, Av, \ldots, A^{k-1}v, A^k v\}$ lineárisan összefüggő. Ekkor léteznek $\alpha_0, \ldots, \alpha_{k-1} \in \mathbb{F}$ számok, amelyekre*

$$A^k v = \sum_{j=0}^{k-1} \alpha_j A^j v$$

Az A operátor ezen v vektorhoz tartozó kis minimál polinomja

$$n(t) = t^k - \alpha_{k-1}t^k - \cdots - \alpha_1 t - \alpha_0.$$ ⌟

Bizonyítás: Láttuk, hogy $\{v, Av, \ldots, A^{k-1}v\}$ bázisa $\operatorname{lin}(v; A)$ altérnek. Világos, hogy $A^k v \in \operatorname{lin}(v; A)$, emiatt a kívánt előállítás valóban létezik. Ezt átrendezve kapjuk, hogy n valóban olyan normált polinom, amelyre $n(A)v = 0$ fennáll. No de k-nal alacsonyabb fokú ilyen polinom csak a konstans zérus polinom lehet, hiszen $\{v, Av, \ldots, A^{k-1}v\}$ lineárisan független. Azt láttuk tehát, hogy n a legalcsonyabb fokú nem zérus eleme $J_{A,v}$-nek, tehát n generálja a főideált. ▪

10.2. Minimál polinom

10.11. állítás. *Legyen $A \in L(V)$ a V véges dimenziós vektortér egy lineáris transzformációja. Tekintsük az $\mathbb{F}[t]$ polinom gyűrű 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 polinom gyűrűnek.

Jelölje $n = \dim V$ és tekintsük az $n^2 + 1$ elemű $\left\{I, A, \ldots, A^{n^2}\right\}$ rendszerét az $L(V)$ vektortérnek. Mivel a Steinitz-lemma szerint $n^2 + 1$ vektor egy n^2-dimenziós vektortérben lineárisan összefüggő, ezért van $\alpha_0, \ldots, \alpha_{n^2} \in \mathbb{F}$ 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 $-\infty < \deg p \leq n^2$. ▪

10.12. definíció (minimál polinom)**.** 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 főideál-gyűrű, van tehát egyetlen normált polinomja J_A-nek, amely generálja J_A. Ez a J_A legkisebb fokú, normált polinomja, amit A transzformáció *minimál polinomjának* nevezünk. ⌟

Ha $V \neq \{0\}$, ergo $\dim V \geq 1$, akkor a J_A ideálnak nincs nulladfokú polinomja, hiszen $p(t) = c, (c \neq 0)$ mellett $p(A) = cI \neq 0$, tehát legalább egy dimenziós tér egy lineáris transzformációjának a minimál polinom legalább első fokú.

10.13. állítás. *Ha $m \in \mathbb{F}[t]$ polinom akkor és csak akkor az $A \in L(V)$ transzformáció minimál polinomja, 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 minimál polinom fokszámára:[1] $\deg m \leq (\dim V)^2$. ⌟

Bizonyítás: Definíció szerint m azon p polinomok közül, amelyek normáltak, és $p(A) = 0$, a legalacsonyabb fokú. Láttuk hogy ilyen polinom csak egy van, es erre a polinomra

$$J_A = J(m) = \{hm : h \in \mathbb{F}[t]\}.$$

Ezt kellett belátni. ▪

10.14. állítás. *Legyen V legalább 1 dimenziós, és tegyük fel, hogy $p(A) = 0$ valamely normált, irreducibilis p polinomra. Ekkor p az A minimál polinomja.* ⌟

Bizonyítás: Ha maga $\deg p = 1$, akkor készen vagyunk, hiszen a minimálpolinom legalább első fokú. Ha $\deg p > 1$ és p nem egyezne az m minimál polinommal, akkor $\deg m < \deg p$ lenne, és mivel m generálja a J_A ideált, ezért $p = mh$ alakú lenne, ahol h is legalább első fokú. Így p két legalább első fokú polinom szorzata, azaz reducibilis lenne. ▪

A következő állítás módszert ad a minimál polinom meghatározására.

10.15. 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 $\{I, A, \ldots, A^{k-1}\}$ lineáris független, de $\{I, A, \ldots, A^{k-1}, A^k\}$ lineárisan összefüggő. Ekkor léteznek $\alpha_0, \ldots, \alpha_{k-1} \in \mathbb{F}$ számok, amelyekre*

$$A^k = \sum_{j=0}^{k-1} \alpha_j A^j.$$

Ekkor az A operátor minimál polinomja.

$$m(t) = t^k - \alpha_{k-1}t^k - \cdots - \alpha_1 t - \alpha_0.$$ ⌟

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 $\{I, A, \ldots, A^{k-1}\}$ lineárisan független. Azt láttuk tehát, hogy m a legalcsonyabb fokú nem zérus eleme J_A-nek, tehát az m polinom generálja a főideált. ▪

A minimál polinom algoritmikus meghatározásához a kis minimál polinom is használhatók.

10.16. állítás. *Legyen az $A \in L(V)$ lineáris transzformáció minimál polinomja m. Tegyük fel, hogy az $\{e_1, \ldots, e_n\}$ egy bázisa V-nek, és $p_1, \ldots, p_n$ rendre a bázis elemekhez tartozó kis minimál polinomok. Ekkor a $p_1, \ldots, p_n$ polinomok legkisebb közös többszöröse az m minimal polinom.* ⌟

Bizonyítás: Tetszőleges $v \in V$ mellett, ha p_v a kis minimál polinom, akkor $p_v|m$, hiszen $m(A)$ a konstans zérus transzformáció. Emiatt m egy közös többszöröse a p_j kis minimál polinomoknak. Most tegyük fel, hogy egy p polinom többszöröse a p_j kis minimál polinomoknak. 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. ▪

[1] Kis vártatva kiderül, hogy $\deg m \leq \dim V$ is igaz.

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

$$J(d) = \cap_{i=1}^{n} J(p_i) = \cap_{i=1}^{n} J_{A,e_i} = J_A = J(m).$$

No de, a bal oldali ideál d generáló eleme – az 1.26. állítás szerint – a $p_1, \ldots, p_n$ polinomok legkisebb közös többszöröse. Mivel egy főidálnak csak egy normált generáló eleme van, ezért a d legkisebb közös többszörös azonos a jobb oldali ideált generáló m minimálpolinommal. .

10.3. Sajátvektorok és diagonalizálhatóság

10.17. állítás. *Tegyük fel, hogy p legalább elsőfokú osztója az A transzformáció minimál polinomjának. Ekkor $p(A)$ szinguláris.* ⌟

Bizonyítás: Jelölje m a minimál polinomot, 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. .

10.18. állítás (sajátérték és minimál polinom)**.** *Legyen $A \in L(V)$ lineáris transzformációja a V véges dimenziós vektortérnek, és $\lambda \in \mathbb{F}$ egy szám, valamint m az A minimál polinomja. 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 $3 \implies 4 \implies 5 \implies 6 \implies 5 \implies 3$ utat érdemes követni. Az utolsó lépéshez mutattuk meg az előző állítást. .

Az egyik legfontosabb definícióhoz érkeztünk.

10.19. definíció (diagonalizálható transzformáció)**.** Az $A \in L(V)$ lineáris transzformációt *diagonalizálhatónak* mondjuk, ha van a térnek olyan bázisa, amelyben a transzformáció mátrixa diagonális alakú, azaz a főátlóján 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 bázis vektorra $Ae_j = \lambda e_j$ 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.

10.20. állítás (diagonalizálhatóság)**.** *Az $A \in L(V)$ egy lineáris transzformáció pontosan akkor diagonalizálható, ha van térnek csupa sajátvektorokból álló bázisa.*

Ebben az esetben a transzformációnak a $\{v_1, \ldots, v_n\}$ sajátvektorokban felírt mátrixának 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 egy dimenziós, 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 sajáteltérre $\ker A = \operatorname{lin}\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$, míg a $\lambda = 2$ sajátértékhez tartozó sajátaltérre $\ker(A - 2I) = \operatorname{lin}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$. 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.

10.21. 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

$$\sum_{j=1}^{s-1} \lambda_s \alpha_j v_j = \lambda_s v_s = A v_s = \sum_{j=1}^{s-1} \alpha_j A v_j = \sum_{j=1}^{s-1} \alpha_j \lambda_j v_j,$$

ami a 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 sajátértékekről van szó, ezért minden szóba jövő j mellett $\alpha_j = 0$. Ebből $v_s = 0$ következik, ami ellentmond annak, hogy v_s egy sajátvektor. ▪

Egy n-dimenziós térben n-elemű lineárisan független rendszer generátorrendszer is, így azonnali következmény a diagonalizálhatóság egy elegendő feltétele:

10.22. 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:

10.23. állítás. *Legyen $A \in L(V)$ lineáris transzformáció. Ekkor A-nak legfeljebb* $\dim V$ *darab különböző sajátértéke lehet.* ⌟

10.24. definíció (geometriai multiplicitás)**.** Ha $A \in L(V)$ egy lineáris transzformáció, és $\lambda \in \sigma(A)$ annak egy sajátérétke, akkor az $A - \lambda I$ sajátaltér defektusát, tehát az $\ker(A - \lambda I)$ altér dimenzióját, a λ sajátérték *geometriai multiplicitásának* mondjuk. ⌟

10.25. 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 $M_1 \oplus \dots \oplus M_s$ direktösszeg.* ⌟

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 találnánk különböző sajátértékekhez tartozó sajátvektorok egy lineárisan összefüggő rendszerét, ami a 10.21. állítás szerint nem lehetséges. ▪

Meggondoltuk tehát, hogy ha páronként különböző $\{\lambda_1, \dots, \lambda_s\} \subseteq \sigma(A)$ sajátértékekből indulunk ki, és egyesítjük a $\ker(A - \lambda_j I)$ sajátalterek 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.

10.26. állítás. *A különböző sajátértékekhez tartozó sajátelterek bázisainak egyesítése a sajátalterek direktösszegének egy bázisa.* ⌟

10.27. 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. $\ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_s I) = V$*,*
3. *Minden vektor előáll mint sajátvektorok összege,*
4. *Az A diagonalizálható lineáris transzformáció.* ⌟

Bizonyítás: Mivel a $\ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_s I) \subseteq V$ tartalmazás mindig fennáll, ezért a 2. feltétel ekvivalens avval, hogy

$$\sum_{j=1}^{s} \dim(\ker(A - \lambda_j I)) = \dim(\ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_s I)) = \dim V,$$

ami éppen a geometriai multiplicitásra vonatkozó feltétel. Így az első két feltevés ekvivalenciáját megértettük.

Mivel a $\ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_s I) \subseteq V$ tartalmazás mindig 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 $\ker(A - \lambda_j I)$ sajátelterek 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. ▪

11. fejezet

Transzformációk redukálása

11.1. állítás. *Tekintsünk egy $A \in L(V)$ lineáris transzformációt, amelynek minimál polinomja $m \in \mathbb{F}[t]$. 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 direktösszegére, 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 minimál polinomja p és A_2 minimál polinomja 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

$$f(A)\,p(A)\,x + g(A)\,q(A)\,x = Ix = x. \qquad (\dagger)$$

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 direktösszege.

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 direktösszegre 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 $m|hq$. Mivel $p|m$, 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 minimál polinomja. Az A_2 minimál polinomja q állítás igazolása a fentivel analóg. ∎

11.2. állítás. *Tekintsünk egy $A \in L(V)$ lineáris transzformációt, amelynek minimál polinomja $m \in \mathbb{F}[t]$. Tegyük fel, hogy m előáll mint a páronként relatív prím normált polinomok*

$$m = p_1 p_2 \dots p_n$$

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 minimál polinomja minden szóba jövő* $i = 1, \ldots, n$ *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 $p = p_1, \ldots, p_{n-1}$. Világos, hogy p és p_n relatív prímek, hiszen ha d irreducibilis osztója p-nek és p_n-nek, akkor d prím tulajdonsága szerint $d|p_i$ valamely $i < n$-re, tehát $d|p_i$ és $d|p_n$. 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 minimál polinomja az $A|\ker p(A)$-nak és persze p_n minimál polinomja A_n-nek.

Alkalmazzuk most az indukciós feltevést a $\ker p(A)$ vektortérre. Ott az $A_{|\ker p(A)}$ lineáris transzformáció p minimál polinomja előáll mint $n-1$ páronként relatív prím polinom szorzata:

$$p = p_1, \ldots, p_{n-1}.$$

Világos tehát, hogy

1. $\ker p(A) = V_1 \oplus \cdots \oplus V_{n-1}$ és
2. p_i az A_i minimál polinomja minden $i = 1, \ldots, n-1$ mellett.

Teljesül tehát a $V_1, \ldots, V_{n-1}, V_n$ alterekre, hogy mind diszjunkt – az itt adott sorrendben – az előzőek összegétől, és a Minkowski-összegük az egész V vektortér. Az 5.10. állítás szerint tehát $V = V_1 \oplus \cdots \oplus V_n$. ▪

11.1. Sajátvektorok, minimálpolinom és diagonalizálhatóság

Alkalom nyílik, hogy a lineáris transzformáció diagonalizálhatóságát karakterizáló a 10.27. állítást tovább bővítsük, a minimál polinom szerepének hangsúlyozásával.

11.3. állítás (diagonalizálhatóság)**.** *Legyen* $A \in L(V)$ *lineáris transzformáció. Jelölje* $\{\lambda_1, \ldots, \lambda_s\} = \sigma(A)$ *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 minimál polinomja*

$$m(t) = \prod_{j=1}^{s} (t - \lambda_j).$$ ⌟

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 pozició 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 minimál polinom gyöke, ezért a fenti polinom a legalacsonyabb fokú normált polinom, amelynek gyöke A. [1]

[1] Egy másik érv: $t - \lambda_j$ a λ_j sajátértékhez tartozó sajátvektor kis minimál polinomja. Mivel a sajátvektorok egy bazist alkotnak, ezért ezen polinomok legkisebb közös többszöröse a minimál polinom. Persze dim V darab elsőfokú polinom 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

$$V = \ker p_1(A) \oplus \cdots \oplus \ker p_j(A) = \ker(A - \lambda_1 I) \oplus \cdots \oplus \ker(A - \lambda_s I)$$

Ezt kellett belátni. ▪

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 minimál polinomja 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 minimál polinomja $p_i^{m_i}$ minden $i = 1, \dots, r$ mellett.* ⌟

A konklúzió tehát az, hogy elég olyan transzformációkkal foglalkoznunk, amelyek minimál polinomja 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 minimál polinomjára

$$m(t) = (t - \alpha)^n$$

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 *nillpotens* fogalmához. Ha felírjuk valamely bázisban egy ilyen B nillpotens 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 gondaloton alapul a *Jordan-normálak* fogalma.

12. fejezet

Redukálás irreducibilis minimál polinom esetén

A LEGEGYSZERŰBB 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* $\dim \operatorname{lin}(v; A) = k$;
2. *minden $v \in V$ vektor és minden $K \subseteq V$ invariáns altér mellett* $\operatorname{lin}(v; A) \cap K = \{0\}$ *vagy* $\operatorname{lin}(v; A) \subseteq K$;
3. *a V altér nulla dimenziós, vagy k dimenziós, vagy k dimenziós A invariáns alterek direktösszege. Pontosabban, ha* $\dim V > 0$, *akkor létezik $r \geq 1$ szám, és léteznek $v_1, \ldots, v_r$ vektorok, amelyekre*

$$\operatorname{lin}(v_1; A) \oplus \cdots \oplus \operatorname{lin}(v_r; A) = V. \qquad \lrcorner$$

Bizonyítás (1.): Legyen adott $v \neq 0$ vektor mellett p_v a v-hez tartozó kis minimál polinom. Mivel $m(A)\, v = 0$, ezért $p_v | m$. No de m irreducibilis, ezért $m = p_v$. Ekkor viszont

$$k = \deg m = \deg p_v = \dim \operatorname{lin}(v; A). \qquad .$$

Bizonyítás (2.): Ha $v = 0$, akkor az állítás nyilvánvaló. A továbbiakban emiatt $v \neq 0$. Most tegyük fel, hogy $x \in \operatorname{lin}(v; A) \cap K$ és $x \neq 0$. Ekkor

$$\operatorname{lin}(x; A) \subseteq \operatorname{lin}(v; A) \cap K \subseteq \operatorname{lin}(v; A)$$

No de, a bal és jobb oldali altér azonos dimenziós alterek, emiatt fent mindenütt egyenlőség van. Speciálisan $\operatorname{lin}(v; A) = \operatorname{lin}(v; A) \cap K$, ami éppen azt jelenti, hogy $\operatorname{lin}(v; A) \subseteq K$. .

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 $V_1 = \operatorname{lin}(v_1; A)$. Ha $V = V_1$, akkor V egy k-dimenziós vektortér.

Ha $V \neq V_1$, akkor van van $v_2 \in V \setminus V_1$. Mivel V_1 egy invariáns altér, ezért a már igazolt állítás szerint bevezetve a $V_2 = \operatorname{lin}(v_2; A)$ jelölést $V_2 \cap V_1 = \{0\}$. Értelmes tehát venni e két invariáns altér direktösszegét. Ha $V = V_1 \oplus V_2$, akkor V előállt két k dimenziós invariáns alterének direktösszegeként.

Ha $V \neq V_1 \oplus V_2$, akkor van van $v_3 \in V \setminus (V_1 \oplus V_2)$. Mivel $V_1 \oplus V_2$ egy invariáns altér, ezért a már igazolt állítás szerint bevezetve a $V_3 = \operatorname{lin}(v_3; A)$ jelölést $V_3 \cap (V_1 \oplus V_2) = \{0\}$. Értelmes tehát venni e három invariáns altér direktösszegét, hiszen a V_1, V_2, V_3 alterek ebben a sorrendben véve olyanok, hogy mind diszjunkt az előzőek összegétől. Ha $V = V_1 \oplus V_2 \oplus V_2$, akkor V előállt három k dimenziós invariáns alterének direktösszegeként.

Ha $V \neq V_1 \oplus \cdots \oplus V_t$, valamely $t \geq 2$ mellett, akkor van van $v_{t+1} \in V \setminus (V_1 \oplus \cdots \oplus V_t)$. Mivel $V_1 \oplus \cdots \oplus V_t$ egy invariáns altér, ezért a már igazolt állítás szerint bevezetve a $V_{t+1} = \operatorname{lin}(v_{t+1}; A)$ jelölést $V_{t+1} \cap (V_1 \oplus \cdots \oplus V_t) = \{0\}$. Értelmes tehát venni ezen $t+1$ invariáns altér direktösszegét, hiszen a $V_1, V_2, V_3, \ldots, V_t, V_{t+1}$ alterek ebben a sorrendben véve olyanok, hogy mind diszjunkt az előzőek összegétől. Ha $V = V_1 \oplus \cdots \oplus V_{t+1}$, akkor V előállt $t+1$ darab k dimenziós invariáns alterének direktösszegeként.

Az eljárás előbb utóbb a V vektortér véges dimenziós volta miatt megáll. .

Következmény

Világos, hogy $\left\{A^{k-1}v_1, A^{k-2}v_1, v_1, \ldots, Av_1, v_1\right\}$ bázisa a $\operatorname{lin}\{v_1; A\}$ invariáns altérnek. Ha ebben a bázisban felírjuk a transzformáció mátrixát, akkor az első oszlopban vannak a minimálpolinom együtthatóinak ellentettjei, a diagonális feletti 1-ek, vannak és minden más elem zérus:

$$\begin{pmatrix} -\alpha_{k-1} & 1 & 0 & \ldots & 0 \\ -\alpha_{k-2} & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\alpha_1 & 0 & 0 & \ldots & 1 \\ -\alpha_0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Meggondoltuk tehát, hogy ha a transzformáció m minimálpolinomja irreducibilis és

$$m(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{k-1}t^{k-1} + t^k$$

alakú, akkor a térnek van olyan bázisa, amelyben a transzformáció mátrixa a fenti mátrix diagonális elrendezésű r darab másolatából áll, ahol $rk = \dim V$.

Házi feladatként gondoljuk meg, hogy minden nem konstans $m(t) \in \mathbb{F}[t]$ polinomhoz létezik egy $\mathbb{F}$ feletti vektortér, és azon egy lineáris transzformáció, amelynek minimálpolinomja éppen m. (Segítség: Nézzük és csodáljuk a fenti mátrixot.)

Illusztráció

Bontsuk lehető legalacsonyabb invariáns alterek direktösszegére 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 látjuk, hogy nincs valós sajátérték. Keressük tehát a minimálpolinomot a bázis egyes elemeihez tartozó kis minimálpolinomok meghatározásával.

	Au_1	A^2u_1
u_1	-2	1
	$\boxed{-2}$	2
	-1	1
	-1	1
	δ	-1

	A^2u_1
u_1	-1
Au_1	-1
	0
	0

$\{u_1, Au_1\}$ lineárisan független, de $A^2u_1 + Au_1 + u_1 = 0$, ezért $p_1(t) = t^2 + t + 1$.

	Au_2	A^2u_2
	0	0
u_2	-1	0
	0	0
	$\boxed{-1}$	1
	δ	-1

	A^2u_2
	0
u_2	-1
	0
Au_2	-1

$\{u_2, Au_2\}$ lineárisan független, de $A^2u_2 + Au_2 + u_2 = 0$, ezért $p_2(t) = t^2 + t + 1$.

	Au_3	A^2u_3
	$\boxed{3}$	-3
	3	-3
u_3	1	-2
	3	-3
	δ	-1

	A^2u_3
Au_3	-1
	0
u_3	-1
	0

$\{u_3, Au_3\}$ lineárisan független, de $A^2u_3 + Au_2 + u_3 = 0$, ezért $p_3(t) = t^2 + t + 1$.

	Au_4	A^2u_4
	0	0
	$\boxed{1}$	-1
	0	0
u_4	0	-1
	δ	-1

	A^2u_4
	0
Au_4	-1
	0
u_4	-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 bázis transzformációra azt kaptuk, hogy

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 & 1 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} -2 & 0 & 3 & 0 \\ -2 & -1 & 3 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & -1 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & -2 & 1 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}. \qquad \blacksquare$$

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 irreducibilis polinom. Ekkor:*

1. *Minden $v \neq 0$, $v \in \ker p(A)$ mellett $\dim \operatorname{lin}(v; A) = \deg p$;*
2. *A $\ker p(A)$ altér nulla dimenziós, vagy k dimenziós, vagy k dimenziós invariáns alterek direktösszege. Pontosabban fogalmazva ha $\nu(A) > 0$, akkor létezik $r \geq 1$ szám, és léteznek $v_1, \ldots, v_r$ vektorok, amelyekre*

$$\operatorname{lin}(v_1; A) \oplus \cdots \oplus \operatorname{lin}(v_r; A) = \ker p(A).$$ ⌟

Összefoglalhatjuk azt az esetet, mikor a minimálpolinom különböző irreducibilis polinomok szorzata. Ez éppen a 11.2. állítás esete.

12.3. állítás. *Tekintsünk egy $A \in L(V)$ lineáris transzformációt, amelynek minimál polinomja $m \in \mathbb{F}[t]$. Tegyük fel, hogy m előáll mint a különböző normált, irreduciblis polinomok elsőfokú hatványainak*

$$m = p_1 p_2 \ldots p_s$$

szorzata. Ekkor V előáll mint néhány – de legalább egy-egy darab – $\deg p_1, \deg p_2, \ldots, \deg p_s$ dimenziós minimális invariáns alterének direktösszege. ⌟

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 minimál polinom legalább lesőfokú 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 direktösszege. Mivel ez minden j mellett igaz, ezért $\ker p_j(A)$ felbontását a V felbontásába helyettesítve készen is vagyunk. $\blacksquare$

A komplex és a valós eset

Most az tegyük fel, hogy a V vektortér a $\mathbb{C}$ komplex számtest feletti vektortér. Ekkor az $A \in L(V)$ lineáris transzformáció minimál polinomja egy komplex együtthatós polinom. Az algebra alaptétele szerint irreducibilis normált polinom csak elsőfokú, azaz $t - \lambda$ alakú lehet, tehát az előző állítás feltétele most abba megy át, hogy A minimál polinomja

$$m(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_s)$$

alakú, ahol $\lambda_1, \dots, \lambda_s$ az m különböző gyökei. Tudjuk viszont, hogy a minimál polinom gyökei a sajátértékek, ezért a feltételt egyszerűbben úgy is fogalmazhatjuk, a minimál polinom gyökei egyszeresek, vagy ami ugyanaz: a sajátértékek a minimál polinom egyszeres gyökei.

12.4. állítás. *Tegyük fel, hogy V egy $\mathbb{C}$ feletti vektortér és $A \in L(V)$ egy lineáris transzformáció. A pontosan akkor diagonalizálható, ha minimál polinomja gyökei egyszeres multiplicitásúak.* ⌟

Bizonyítás: Ha a sajátértékek a minimál polinomnak egyszeres multiplicitású gyökei, akkor fennállnak az előző tétel feltételei. Így V előáll mint néhány egy dimenziós invariáns alterének direktösszege, ami azt jelenti, hogy van sajátvektorokból álló bázisa, ergo diagonalizálható.

Megfordítva, ha A diagonalizálható, akkor a diagonálisban a sajátértékei vannak. Ha $\{\lambda_1, \dots, \lambda_s\}$ a diagonális különböző elemei, akkor az A minimál polinomja nyilvánvalóan

$$m(t) = \prod_{j=1}^{s} (t - \lambda_j)$$

alakú. Ez persze olyan polinom, amelyben minden gyöktényező csak egyszer szerepel. .

Mivel a valós test feletti irreducibilis polinomok első– vagy másodfokúak, ezért a valós esetben is szép állítást kapunk. [1]

12.5. állítás. *Tegyük fel, hogy V egy $\mathbb{R}$ feletti vektortér és $A \in L(V)$. A V pontosan akkor bontható fel egy vagy két dimenziós minimális A-invariáns alterek direktösszegére, ha az A transzformáció m minimál polinomjának felbontásában minden irreducibilis polinom az első hatványon szerepel.* ⌟

Bizonyítás: A komplex esettel analóg. .

[1] Ha nem is annyira szépet mint a komplex esetben.

13. fejezet

A minimál polinom fokszámáról

A MINIMÁL POLINOM definiálásakor csak annyit láttunk, hogy legfeljebb n^2 fokú polionom mindig 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 minimál polinom 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 mindegyikére $B^m v_j = 0$, de a*

$$\left\{B^{m-1}v_1, B^{m-1}v_2, B^{m-1}v_3, \dots, B^{m-1}v_r\right\}$$

vektorrendszer lineárisan független. Ekkor a

$$\left\{\begin{array}{ccccc} v_1 & v_2 & v_3 & \dots & v_r \\ Bv_1 & Bv_2 & Bv_3 & \dots & Bv_r \\ B^2v_1 & B^2v_2 & B^2v_3 & \dots & B^2v_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^tv_1 & B^tv_2 & B^tv_3 & \dots & B^tv_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^{m-1}v_1 & B^{m-1}v_2 & B^{m-1}v_3 & \dots & B^{m-1}v_r \end{array}\right\}$$

vektorrendszer is lineárisan független. ⌟

Bizonyítás: Tekintsük ezen vektorok egy

$$\sum_{j=0}^{m-1}\sum_{k=1}^{r}\alpha_{j,k}B^j x_k = 0$$

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

$$\sum_{j=t}^{m-1}\sum_{k=1}^{r}\alpha_{j,k}B^j x_k = 0$$

Erre alkalmazva a B^{m-t-1} transzformációt azt kapjuk, hogy

$$0 = B^{m-t-1}0 = \sum_{j=t}^{m-1}\sum_{k=1}^{r}\alpha_{j,k}B^{j+m-t-1}x_k = \sum_{k=1}^{r}\alpha_{t,k}B^{t+m-t-1}x_k = \sum_{k=1}^{r}\alpha_{t,k}B^{m-1}x_k.$$

No de az $\left\{B^{m-1}x_1, \dots, B^{m-1}x_k\right\}$ egy lineárisan független rendszer, amiből már következik, hogy minden $\alpha_{t,k} = 0$, ami ellentmondásban van a t index definíciójával. ∎

13.2. állítás. *Legyen $B \in L(V)$ lineáris transzformációra $B^m = 0$. Ekkor* $\dim V \geq m\rho\left(B^{m-1}\right)$. ⌟

Bizonyítás: Létezik tehát olyan $\left\{B^{m-1}v_1, B^{m-1}v_2, \ldots, B^{m-1}v_r\right\}$ 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 $mr = m\rho\left(B^{m-1}\right)$ dimenziós. ▪

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 mind a minimálpolinom fokszáma, mind 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

$$\dim \operatorname{lin}(v; A) = \deg p = k,$$

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

$$\operatorname{lin}\left\{B^{m-1}v; A\right\} = \operatorname{lin}\left\{B^{m-1}v, AB^{m-1}v, A^2B^{m-1}v, \ldots, A^{k-1}B^{m-1}v\right\}$$

pontosan k dimenziós. Persze $\left\{B^{m-1}v, B^{m-1}Av, B^{m-1}A^2v, \ldots, B^{m-1}A^{k-1}v\right\} \subseteq \operatorname{Im} B^{m-1}$, ergo

$$\rho\left(B^{m-1}\right) \geq k.$$

Alkalmazva B-re az imént igazolt 13.2 tételt kapjuk a kívánt $\dim V \geq m\rho\left(B^{m-1}\right) \geq m \cdot k$ becslést. ▪

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 tétel szerint minden egyes i index mellett $m_i \cdot k_i \leq \dim(\ker p_i^{m_i}(A))$. Világos, hogy

$$\deg m = \sum_{i=1}^{r} k_i \cdot m_i \leq \sum_{i=1}^{r} \dim(\ker p_i^{m_i}(A)) = \dim V. \quad ▪$$

Később ki fog derülni, hogy a transzformáció karakterisztikus polinomjának, amely pontosan $\dim V$-ed fokú, is mindig gyöke a transzformáció. Ez az úgynevezett Cayley-Hamilton–tétel.

[1] Sőt még azt is láttuk, hogy néhány – lehet, hogy csak egy – k dimenziós invariáns altér direktösszege.

14. fejezet

Nilpotens transzformációk

14.1. Hatvány függvény alakú minimálpolinom

14.1. definíció (nillpotens transzformáció)**.** Egy $A \in L(V)$ lineáris transzformációt *nillpotensnek* mondjuk, ha létezik $k \in \mathbb{N}$, melyre $A^k = 0$. Ha A egy nillpotens transzformáció, akkor azt a legkisebb m számot, melyre $A^m = 0$ a *nillpotencia rendjének* nevezzük. ⌟

Például egy 5 dimenziós téren könnyen definiálhatunk első-, másod-, harmad-, negyed, és ötöd-rendű nilpotens transzformációkat. De van-e mondjuk hatod-rendű nillpotens transzformáció ezen öt dimenziós vektortéren? Az első észrevétel adja a negatív választ.

14.2. állítás. *Legyen B egy m-ed rendben nillpotens operátor. Ekkor létezik $v \in V$ vektor, melyre $B^{m-1}v \neq 0$. Minden ilyen v vektorra a*

$$\left\{v, B, B^2v, \cdots, B^{m-1}v\right\}$$

m elemű vektorrendszer lineárisan független.

Emiatt ha a V vektortérnek van m-edrendben nillpotens lineáris transzformációja, akkor $\dim V \geq m$*, azaz a nillpotencia 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 $\left\{B^{m-1}v\right\}$ egy elemet tartalmazó rendszer lineárisan független, ezért a 13.1. lemma szerint a tételbeli rendszer is lineárisan független. ▪

14.3. állítás. *Egy lineáris transzformáció pontosan akkor nillpotens, 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 nillpotens. Legyen $p(t) = t^m$. Ekkor $p \in J_A$, így a minimál polinom p osztója. Másrészt a 14.2. állítás szerint van a térnek olyan vektora, amelyhez tartalmazó kis minimál polinom is p. Így p osztója a minimál polinomnak, 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 nillpotens. ▪

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 alteret tartalmazó direkkiegészítője, azaz létezik $K \subset W$ altér, amelyre $V_2 \subseteq K$ és $V_1 \oplus K = V$.* ⌟

Bizonyítás: Jelölje $V = V_1 + V_2$. Világos, hogy V altér W-ben. Legyen L a direktkiegészítője, 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 $V_1 + K = W$. ▪

14.5. lemma. *Legyen a W véges dimenziós vektortérnek $H, K_0, \overline{K}$ altere. Tegyük fel, hogy*

1. $H \cap K_0 = \{0\}$;
2. $H + \overline{K} = W$;
3. $K_0 \subseteq \overline{K}$.

Ekkor létezik K altere W-nek, melyre

1. $K_0 \subseteq K \subseteq \overline{K}$ *és*
2. K *direkt kiegészítője H-nak, azaz* $H \oplus K = W$. ⌟

Bizonyítás: Világos, hogy $H \cap \overline{K} \subseteq \overline{K}$ és $K_0 \subseteq \overline{K}$ diszjunkt alterek $\overline{K}$-ban, hiszen

$$\left(H \cap \overline{K}\right) \cap K_0 \subseteq H \cap K_0 = \{0\}.$$

Alkalmazzuk az előző (14.4) lemmát a $\overline{K}$ altérben. Létezik tehát $K_0 \subseteq K \subseteq \overline{K}$ altér $\overline{K}$-ban, amelyre

$$\left(H \cap \overline{K}\right) \oplus K = \overline{K}.$$

Most megmutatjuk, hogy H és K diszjunkt alterek:

$$H \cap K = H \cap \left(K \cap \overline{K}\right) = \left(H \cap \overline{K}\right) \cap K = \{0\}.$$

Most azt mutatjuk meg, hogy $H + K = W$. Ugyanis

$$W = H + \overline{K} = H + \left(H \cap \overline{K} + K\right) = \left(H + H \cap \overline{K}\right) + K = H + K.$$

Ez éppen azt jelenti, hogy K direkt kiegészítője H-nak. ▪

14.6. állítás (nillpotens operátorok felbontása)**.** *Legyen $B \in L(W)$ egy m-ed rendben nillpotens lineáris transzformáció. Ekkor minden olyan $v \in V$ vektorhoz, amelyre $B^{m-1}v \neq 0$, a* $\operatorname{lin}\{v; B\}$ *invariáns altérnek van invariáns altér direkt kiegészítője.*

Formálisabban: létezik $K \subseteq W$ invariáns altér, amelyre $\operatorname{lin}\left\{v, Bv, \ldots, B^{m-1}v\right\} \oplus K = W$. ⌟

Bizonyítás: A nillpotens 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 egy dimenziós altérnek van direkt kiegészítője.

Tegyük fel, hogy igaz az állítás minden vektortér legfeljebb $m-1$-ed rendben nillpotens transzformációjára. Legyen tehát $m > 1$ és B egy a W vektortéren értelmezett m-ed rendben nillpotens lineáris transzformáció. Rögzítsünk egy $v \in W$ elemet, amelyre $B^{m-1}v \neq 0$. Tekintsük az $\operatorname{Im} B$ invariáns alteret. Világos, hogy $B|_{\operatorname{Im} B}$ egy lineáris transzformáció az $\operatorname{Im} B$ vektortéren. Az is világos, hogy $B|_{\operatorname{Im} B}$ egy $m-1$-rendben nillpotens lineáris transzformáció, hiszen minden $u \in \operatorname{Im} B$ mellett $B^{m-1}u = 0$. Azt is vegyük észre, hogy ezek szerint $v \notin \operatorname{Im} B$.

Alkalmazhatjuk tehát az indukciós feltevést $B|_{\operatorname{Im} B} \in L(\operatorname{Im} B)$ mellett a Bv vektorra. Persze $B^{m-2}Bv = B^{m-1}v \neq 0$. Létezik tehát $K_0 \subseteq \operatorname{Im} B$ a B-re is invariáns altér, amelyre

$$\operatorname{lin}\left\{Bv, B^2v, \ldots, B^{m-1}v\right\} \oplus K_0 = \operatorname{Im} B.$$

Most megmutatjuk, hogy $\operatorname{lin}\{v; B\} \cap K_0 = \{0\}$. Ugyanis, ha $x = \sum_{k=0}^{m-1} \alpha_k B^k v \in K_0 \subseteq \operatorname{Im} B$, akkor $\alpha_0 v \in \operatorname{Im} B$, ami csak úgy lehetséges, hogy $\alpha_0 = 0$. Ezek szerint $x \in \operatorname{lin}\{Bv; B\}$ altérnek melynek direkt kiegészítője K_0. Ez persze csak úgy lehetséges, hogy $x = 0$.

Definiálja

$$\overline{K} = \{x \in W : Bx \in K_0\}.$$

Mivel K_0 egy altér, ezért $\overline{K}$ is az. Mivel a K_0 altér B-invariáns, azért teljesül a $K_0 \subseteq \overline{K}$ tartalmazás.

Most megmutatjuk, hogy $\operatorname{lin}\{v; B\} + \overline{K} = W$. Válasszunk egy $u \in W$ vektort. Persze $Bu \in \operatorname{Im} B$, 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 $B\left(u - \sum_{k=0}^{m-2} \alpha_{k+1} B^k v\right) \in K_0$, ami persze $\overline{K}$ definícióját figyelembe véve azt jelenti, hogy $u - \sum_{k=0}^{m-2} \alpha_{k+1} B^k v \in \overline{K}$. Előállítottuk tehát az u vektort egy $\operatorname{lin}\{v; B\}$-beli és egy $\overline{K}$ beli összegeként.

Alkalmazhatjuk tehát a 14.5. lemmát. Így létezik $K_0 \subseteq K \subseteq \overline{K}$ altér, melyre $\operatorname{lin}\{v; B\} \oplus K = W$. Persze ha $u \in K \subseteq \overline{K}$, 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. ▪

Ahhoz, hogy megkapjuk az egész vektorteret v-invariáns altérként vagy ilyenek direktösszegeként, 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 nillpotens. 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

$$\operatorname{lin}\left\{v_2, Bv_2, \ldots, B^{n_2-1}v_2\right\} \oplus K_2 = K.$$

Itt persze $\dim K_2 < \dim K$, hiszen a baloldali első invariáns altér legalább egydimenziós. Az első két lépést összefoglalva:

$$\operatorname{lin}\{v_1; B\} \oplus \operatorname{lin}\{v_2; B\} \oplus K_2 = W,$$

Az eljárást folytatva minden lépésben legalább eggyel csökken a kiegészítő invariáns altér dimenziója. Végül a vektortér előáll néhány, mondjuk r darab B-re invariáns altér direktösszegeként:

$$\operatorname{lin}\{v_1; B\} \oplus \operatorname{lin}\{v_2; B\} \oplus \ldots \operatorname{lin}\{v_r; B\} = W.$$

14.7. állítás (nillpotens transzformáció normmálakja)**.** *Legyen $B \in L(W)$ egy m-ed rendben nilpotens transzformáció. Ekkor léteznek olyan $v_1, v_2 \ldots, v_r \in W$ vektorok és léteznek olyan $m = n_1 \geq n_2 \geq \cdots \geq n_r \geq 1$ pozitív egészek, amelyekre*

$$\operatorname{lin}\left\{v_1, Bv_1, \ldots, B^{n_1-1}v_1\right\} \oplus \operatorname{lin}\left\{v_2, Bv_2, \ldots, B^{n_2-1}v_2\right\} \oplus \ldots \oplus \operatorname{lin}\left\{v_r, Bv_r, \ldots, B^{n_r-1}v_r\right\} = W.$$

Emiatt a

$$\left\{B^{n_1-1}v_1, \ldots, Bv_1, v_1\right\} \cup \left\{B^{n_2-1}v_2, \ldots, Bv_2, v_2\right\} \cup \ldots \cup \left\{B^{n_r-1}v_r, \ldots, Bv_r, v_r\right\}$$

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 (első) mellék diagonális elemei nem nullák. A mellék diagonális elemei 1-esek. Minden más elem zérus. Ezt a mátrixot nevezzük a B nillpotens transzformáció normálalakjának. ⌟

14.3. Egyértelműség

Azt mutatjuk meg, hogy a normálalakban $r = \nu(B)$, és az $n_1, n_2, \ldots, n_r$ számok is a B nilpotens transzformáció altal egyertelműen meghatározottak. Ez azt jelenti, hogy minden nillpotens 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. *$\left\{v, Av, \ldots, A^{m-1}v\right\}$ lineárisan független, így $\operatorname{lin}\{v; A\} = \operatorname{lin}\left\{v, Av, \ldots, A^{m-1}v\right\}$;*
2. *Minden $0 \leq l \leq m$ mellett $\nu\left(A^l | \operatorname{lin}\{v; A\}\right) = l$ és $\rho\left(A^l | \operatorname{lin}\{v; A\}\right) = m - l$.*

⌟

Bizonyítás: Ha $l = m$, akkor mindkét állítás triviálisan teljesül. Legyen tehát $0 \leq l < m$. Az A^l transzformáció az $\left\{A^{m-l}v, \ldots, A^{m-1}v\right\}$ lineárisan független vektorrendszert nullára viszi, így $l \leq \nu\left(A^l | \operatorname{lin}\{v; A\}\right)$. A maradékot, a $\left\{v, \ldots, A^{m-l-1}v\right\}$ vektorrendszert pedig az $\left\{A^l v, \ldots, A^{m-1}v\right\}$ lineárisan független rendszerre képezi. Így $m - l \leq \rho\left(A^l | \operatorname{lin}\{v; A\}\right)$, amiből

$$l \leq \nu\left(A^l | \operatorname{lin}\{v; A\}\right) = m - \rho\left(A^l | \operatorname{lin}\{v; A\}\right) \leq l.$$

Ezt kellett belátni. ▪

14.9. lemma. *Legyen $A \in L(V)$ és tegyük fel, hogy a V vektortér előáll a K_1, K_2 invariáns alterei direktösszegeként, 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 nillpotens felbontás egyértelműsége)**.** *Legyen $A \in L(V)$ egy lineáris transzformáció. Tegyük fel, hogy valamely $\{v_1, \dots, v_r\}$ vektorrendszerre és valamilyen $m_1 \geq m_2 \geq \dots \geq m_r$ pozitív számokra $A^{m_k-1}v_k \neq 0$, de $A^{m_k}v_k = 0$ fennáll minden $k = 1, \dots, r$, továbbá*

$$V = \operatorname{lin}\{v_1; A\} \oplus \dots \oplus \operatorname{lin}\{v_r; A\}.$$

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 $A^{n_k-1}v_k \neq 0$, de $A^{n_k}v_k = 0$ fennáll minden $k = 1, \dots, s$ mellett, és

$$V = \operatorname{lin}\{w_1; A\} \oplus \dots \oplus \operatorname{lin}\{w_s; A\}.$$

Ekkor

1. *A nillpotens lineáris transzformációja V-nek;*
2. *Ha m jelöli a nillpotencia 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, ha észre vesszük, hogy a rendezettség szerint minden $k = 1, \dots, r$ mellett $A^{m_1}|\operatorname{lin}\{v_k; A\} = 0$.

A harmadik állításhoz:

$$\nu(A) = \sum_{k=1}^{r} \nu(A|\operatorname{lin}\{v_k; A\}) = \sum_{k=1}^{r} 1 = r.$$

Ugyanígy a másik direktösszeg 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 direktösszeg felbontásban a t-nél magasabb indexű tagokat elhagyva

$$\rho(A^{n_t}) \geq \left(\sum_{k=1}^{t-1} \rho(A^{n_t}|\operatorname{lin}\{v_k; A\})\right) + \rho(A^{n_t}|\operatorname{lin}\{v_t; A\}) = \left(\sum_{k=1}^{t-1}(m_k - n_t)\right) + m_t - n_t.$$

Hasonlóan, a második direktösszeg felbontásban a t-edik, és a t-nél magasabb indexű rangok zérók, így

$$\rho(A^{n_t}) = \sum_{k=1}^{t-1} \rho(A^{n_t}|\operatorname{lin}\{w_k; A\}) = \sum_{k=1}^{t-1}(n_k - n_t) = \sum_{k=1}^{t-1}(m_k - n_t),$$

ami ellentmond $m_t > n_t$ feltételnek. ▪

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 osztója, ezért csak $m(t) = t^4$ lehetséges, ezért B transzformáció 4-ed rendben nillpotens. Éppen most számoltuk ki, hogy $B^3u_1 \neq 0$, ezért az u_1-et tartalmazó legszűkebb B invariáns altér az egész W, tehát

$$\operatorname{lin}\left\{u_1, Bu_1, B^2u_1, B^3u_1\right\} = W$$

és a $\left\{B^3u_1, Bu_1, Bu_1, u_1\right\}$ 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

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

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 minimálpoliniom $p_4(t) = t^3$.
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 tételben szereplő invariáns direktfelbontást. Mivel a B rangja ránézésre 2, így a magtere 2 dimenziós, ergo két invariáns altér direktösszege W, amiből az egyik 3 dimenziós, így a másik csak egy dimenziós lehet, ezért azt csak a magtér egyik eleme generálhatja! A három dimenziós altér lehet például a $\operatorname{lin}\{u_4; B\}$, hiszen éppen az imént láttuk, hogy $B^2u_4 \neq 0$. A fenti 14.6 Tétel éppen azt állítja, 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 bármely olyan eleme a ker A magtérnek, amely lineárisan független B^2u_4-től. Például u_1.[1] Ilyen módon

$$\operatorname{lin}\left\{u_4, Bu_4, B^2u_4\right\} \oplus \operatorname{lin}\{u_1\} = W,$$

és az $\left\{B^2u_4, Bu_4, u_4, u_1\right\}$ vektorok olyan bázisát adják a térnek, melyben B mátrixa

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

alakú.

14.5. A nillpotens felbontási tétel nélkül?

Legyen $B \in L(V)$ egy m-edrendben nillpotens operátor. Világos, hogy $\nu(B) \geq 1$, hiszen egyébként B valamennyi hatványa is reguláris maradna.

Válasszuk ki a ker A egy $\{e_1, \ldots, e_r\}$ bázisát, ahol a rövidség kedvéért $r = \nu(A)$. Most minden $i = 1, \ldots, r$ 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 $m \geq m_1 \geq m_2 \geq \cdots \geq m_r$, 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 $i = 1, \ldots, r$ 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 kiválaszott bázis elemeit, majd föléjük a megfelelő csökkenő B hatványokat.[2]

$$\begin{array}{ccccc} v_1 & \vdots & \vdots & \ldots & \vdots \\ Bv_1 & v_2 & \vdots & \ldots & \vdots \\ B^2v_1 & Bv_2 & v_3 & \ldots & v_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^{m_1-k}v_1 & B^{m_2-k}v_2 & B^{m_3-k}v_3 & \ldots & B^{m_r-k}v_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B^{m_1-1}v_1 & B^{m_2-1}v_2 & B^{m_3-1}v_3 & \ldots & B^{m_r-1}v_r \end{array}$$

[1] Vagy bármi, ami $\alpha u_1 + \beta u_2$ alakú, ahol $\alpha \neq \beta$.

[2] Tipográfiai 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 lejjebb van mint v_1. Képzeljük úgy a táblázatot, mint monoton fogyó elemszámú oszlopok, alulra zárt összeségét, úgy hogy a legnagyobbal kezdem, stb.

Felmerül, hogy a fenti vektorrendszer bázis, evvel a nillpotens felbontási tételt megkerülve kapnánk a nillpotens normálalak igazolását.

Sajnos nem feltétlen bázis a fenti vektorrendszer. A lineárisan függetlenség a korábbi technikával igazolható – tegyük ezt meg! –, de a rendszer nem mindig generátorrendszer.

Tekintsük például a

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

mátrixát. Jelölje most $\{e_1, e_2, e_3, e_4\}$ a térnek 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 térnek.

A nillpontens felbontási tétel éppen azt mondja, hogy a ker B-nek van olyan alkalmasan megválasztott bázisa, amelyre a fenti konstrukcióban kapott vektorrendszer a térnek generátorrendszere, 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 dimneziójával egyezik meg. Az így kapott rendszer persze bázis lesz.

Függelékek

A. függelék

A komplex számokról

Az algebra alaptétele, és a komplex számtest egyértelműsége. Elsősorban (Ebbinghaus és tsai. 1991) és (Derksen 2003) alapján

A.1. Lineáris algebrai megközelítés

Ha $\{e_1, \ldots, e_n\}$ bázisa egy $\mathbb{C}$ feletti V vektortérnek és $m(t) \in \mathbb{C}[t]$ egy pontosan n-ed fokú, normált polinom, például $m(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n-1} t^{n-1} + t^n$, 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. $\left\{e_1, Ae_1, A^2 e_1, \ldots, A^{n-1} e_1\right\}$ rendszer lineárisan független,
2. $\left\{e_1, Ae_1, A^2 e_1, \ldots, A^{n-1} e_1, A^n e_1\right\}$ rendszer már lineárisan összefüggő, hiszen $A^n e_1 = AA^{n-1} e_1 = Ae_n = -\sum_{j=0}^{n-1} \alpha_j A^j e_1$.

Látjuk 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 $\operatorname{Im} p(A)$ is A-ra invariáns alterek. Két lineáris traféról 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 $\operatorname{Im} p(A_1)$ is invariáns alterek $q(A_2)$-re.* ⌟

Bizonyítás: Elég megmutatni, hogy $\ker(A_1)$-re és $\operatorname{Im}(A_1)$-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. ∎

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 egy dimenziós tér minden nem nulla vektora sajátvektora tetszőleges lineáris transzformációjának, így persze bármely két egy dimenziós 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 $L = \mathrm{Im}(A_1 - \mu I)$ 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. ▪

Mivel a karakterisztikus polinom gyökei a sajátértékek, és mivel egy n-dimenziós téren értelmezett lineáris transzformációnak pontosan n-edfokú a karakterisztikus 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 $L_1, L_2 : \mathcal{H} \to \mathcal{H}$ 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

$$L_1 \circ L_2(B) = \frac{1}{4i}\left(A(AB - BA^*) + (AB - BA^*)A^*\right) = \frac{1}{4i}\left(A^2B + BA^{*2}\right)$$
$$L_2 \circ L_1(B) = \frac{1}{4i}\left(A(AB + BA^*) - (AB + BA^*)A^*\right) = \frac{1}{4i}\left(A^2B + BA^{*2}\right) \quad .$$

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. ▪

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.

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.* ⌟

[1] Azaz, van legfeljebb két dimenziós nem triviális invariáns altere.

Bizonyítás: A k szerinti indukció. A $k = 0$ esetben az előző állítás szerint van sajátvektor is, tehát első fokú 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 térnek egy bázisát, és jelölje $\mathcal{S} \subseteq L(V)$ azon lineáris transzformációk összességét, amelyeknek mátrixa az itt rögzített bázisban szimmetrikus. Ez egy komplex vektortér, amelyre

$$\dim \mathcal{S} = \frac{2^k n \left(2^k n + 1\right)}{2} = 2^{k-1} n \left(2^k n + 1\right) = 2^{k-1} n',$$

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 $L_1, L_2 : \mathcal{S} \to \mathcal{S}$ függvényeket.

$$L_1(B) = AB + BA^T \text{ és } L_2(B) = ABA^T.$$

Könnyű számolgatás mutatja, hogy az L_1 és L_2 lineáris transzformációk kommutálnak:

$$(L_1 \circ L_2) B = A\left(ABA^T\right) + \left(ABA^T\right) A^T = A\left(ABA^T + BA^T A^T\right) = A\left(AB + BA^T\right) A^T = (L_2 \circ L_1) B.$$

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

$$A\lambda B = A\left(AB + BA^T\right) = A^2 B + ABA^T = A^2 B + \mu B.$$

Ha tehát $u = Bv \neq 0$, akkor

$$A^2 u - \lambda A u + \mu u = 0.$$

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. ∎

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 meggondolás, 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 $f : \mathbb{C} \to \mathbb{C}$ 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 $f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$ alakú, ahol $a_n \neq 0$. Világos, hogy $z \neq 0$ esetén

$$f(z) = z^n \left(a_n + a_{n-1} \frac{1}{z} + \ldots + a_1 \frac{1}{z^{n-1}} + a_0 \frac{1}{z^n} \right).$$

Legyen $h : \mathbb{C} \to \mathbb{C}$ komplex polinom a következőképpen definiálva:

$$h(w) := a_{n-1} w + a_{n-2} w^2 + \ldots + a_1 w^{n-1} + a_0 w^n.$$

Ekkor minden $z \in \mathbb{C}$, $z \neq 0$ mellett

$$f(z) = z^n \left(a_n + h(1/z) \right). \tag{A.1}$$

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

$$|f(z)| = |z^n| \, |a_n + h(1/z)| \geq |z^n| \left(|a_n| - |h(1/z)| \right) > |z^n| \left(|a_n| - \frac{1}{2} |a_n| \right) = \frac{1}{2} |a_n| \, |z^n|. \qquad \blacksquare$$

A.6. állítás. *Legyen $f : \mathbb{C} \to \mathbb{C}$ komplex polinom. Ekkor létezik $c \in \mathbb{C}$ komplex szám, melyre*

$$|f(c)| = \inf \{ |f(z)| : z \in \mathbb{C} \}. \qquad \lrcorner$$

Bizonyítás: Most úgy válasszuk meg az r pozitív valós számot, hogy egyrészt 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

$$|f(c)| = \alpha = \inf \{ |f(z)| : z \in \mathbb{C}, \ |z| \leq r \} = \inf \{ |f(z)| : z \in \mathbb{C} \}. \qquad \blacksquare$$

A.7. lemma (Argand)**.** *Legyen $k \in \mathbb{N}$, $k > 0$ egész és $b \in \mathbb{C}$, $b \neq 0$ komplex szám, valamint $g : \mathbb{C} \to \mathbb{C}$, $g(0) = 0$ olyan függvény, amely a $0 \in \mathbb{C}$ pontban folytonos. Tekintsük a következőképpen definiált $h : \mathbb{C} \to \mathbb{C}$ leképezést:*

$$h(z) := 1 + b z^k + z^k g(z).$$

Ekkor létezik $z \in \mathbb{C}$ komplex szám, melyre $|h(z)| < 1$. $\lrcorner$

Bizonyítás: Azt fogjuk megmutatni, hogy található $d \in \mathbb{C}$ és $t \in \mathbb{R}$, $t \in (0,1)$ melyekre $|h(dt)| < 1$.

$$h(dt) = 1 + b d^k t^k + d^k t^k g(dt).$$

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

$$|h(dt)| \leq 1 - t^k + t^k \left| d^k g(dt) \right| = 1 - t^k \left(1 - \left| d^k g(dt) \right| \right).$$

Ebből látszik, hogy elegendő megválasztani $t \in (0,1)$ -et olyan módon, hogy $\left| d^k g(dt) \right| < 1$. Ez pedig nyilván megtehető $g(0) = 0$ és g -nek a 0 pontban feltett folytonossága miatt. $\blacksquare$

A.8. állítás. *Legyen $f : \mathbb{C} \to \mathbb{C}$ 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

$$h(z) = 1 + a_k z^k + \ldots + a_n z^n$$

alakú, ahol $a_k \neq 0$ valamely $k \geq 1$ -re, hiszen az f és ebből következően a h polinom nem konstans. Tovább alakítva:

$$h(z) = 1 + a_k z^k + z^k \left(a_{k+1} z + \ldots + a_n z^{n-k} \right).$$

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

$$|f(\hat{c})| = |f(u+c)| < |f(c)|.$$.

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. .

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 művelettel 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 művelettel, 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 $\mathbb{R}$ felett. 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 $\operatorname{lin}\{1, w\} = \mathbb{G}$ amiből könnyen látható, hogy az $\alpha + w\beta \longmapsto \alpha + i\beta$ 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 $r < 0$. .

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 algebrainak 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

$$\left\{1, v, v^2, \ldots, v^n\right\}$$

vektorrendszert. Ez nyilván lineárisan összefüggő, hiszen $n+1$ vektor egy n dimenziós vektortérben. Így van nem triviális 0 -t adó lineáris kombinációja. Azaz léteznek $\alpha_0, \alpha_1, \ldots \alpha_n$ 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 $p(v) = 0$. .

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

$$v^2 \in \operatorname{lin}\{1, v\} \tag{A.2}$$

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

$$K := \operatorname{lin}\{1, v\}$$

kétdimenziós alterét $\mathbb{G}$-nek. Világos, hogy ennek test voltához elegendő megmutatni, hogy a $\mathbb{G}$-beli műveletekre zárt. Az összeadásra való zártság triviális a K altér mivoltából, a szorzásra való zártság 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 megmutatjuk, 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. ▪

Irodalom

Dancs, István és Csaba Puskás (2001). *Vektorterek*. Aula kiadó 2001, Budapest, ISBN:963 9345 53 9, BCE Catalogue: bcek.379187. (hiv. old. 3).

Derksen, 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. 111).

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. 22, 111).

Paparella, Pietro (2017). „A Short and Elementary Proof of the Two-Sidedness of the Matrix Inverse". *The College Mathematics Journal* 48.5, 366–367. old. DOI: 10.4169/college.math.j.48.5.366. eprint: https://doi.org/10.4169/college.math.j.48.5.366. URL: https://doi.org/10.4169/college.math.j.48.5.366 (hiv. old. 62).

Wardlaw, William P. (2005). „Row Rank Equals Column Rank". *Mathematics Magazine* 78.4, 316–318. old. DOI: 10.1080/0025570X.2005.11953349. eprint: https://doi.org/10.1080/0025570X.2005.11953349. URL: https://doi.org/10.1080/0025570X.2005.11953349 (hiv. old. 62).

Tárgymutató