



Fried Ervin

ALGEBRA I.

Elemi és lineáris algebra

Nemzeti Tankönyvkiadó

ALGEBRA I.

Elemi és lineáris algebra

Fried Ervin

ALGEBRA I.

Elemi és lineáris algebra

Felsőoktatási tankönyv

Megjelent az Oktatási Minisztérium támogatásával
a Felsőoktatási Pályázatok Irodája által lebonyolított
felsőoktatási tankönyv-támogatási program keretében

Bírálok

Dr. Csákány Béla
egyetemi tanár

Dr. Pálffy Péter Pál
egyetemi tanár

A mű más kiadványokban való részleges vagy teljes felhasználása, utánközlése, illetve sokszorosítása a jogtulajdonos engedélye nélkül tilos!

ISBN 963 19 1176 4

© Fried Ervin, Nemzeti Tankönyvkiadó Rt., Budapest 2000, jogutód Fried Katalin 2013

TARTALOM

Előszó	9
I. RÉSZ. ELEMI ALGEBRA	11
ELSŐ FEJEZET	
Komplex számok	13
1. A természetes számoktól a valós számokig	13
2. A komplex számok bevezetése	23
3. A komplex számok geometriai bevezetése	29
4. A komplex számok trigonometrikus alakja	33
MÁSODIK FEJEZET	
Mátrixok	43
1. A mátrix definíciója	43
2. Műveletek a mátrixokkal	46
3. Permutációk	53
4. A determináns	56
5. A determináns kifejtése	62
6. Speciális mátrixok	65
HARMADIK FEJEZET	
Egyhatározatlanú polinomok	71
1. Az egyhatározatlanú polinomok fogalma	71
2. Maradékos osztás és oszthatóság	78
3. Polinomideálok és a legnagyobb közös osztó	82
4. Polinomok egyértelmű felbontása	85
5. Polinomok kompozíciója, behelyettesítés	87
6. Polinomfüggvény, interpoláció	92
7. A legfeljebb negyedfokú polinomok gyökeinek meghatározása	95
8. Az algebra alaptételének ekvivalens alakjai	102

9. Racionális és egész együtthatós polinomok	107
10. Euklideszi gyűrűk	114
NEGYPEDIK FEJEZET	
Többhatározatlanú polinomok	119
1. A többhatározatlanú polinomok fogalma	119
2. Kompozíció, maradékos osztás, oszthatóság többhatározatlanú polinomokra	124
3. Egyhatározatlanú polinomok deriváltja és többszörös gyökei	127
4. Szimmetrikus és alternáló polinomok	132
5. Lineáris egyenletrendszerek megoldása	142
II. RÉSZ. LINEÁRIS ALGEBRA	145
ELSŐ FEJEZET	
Vektorterek	147
1. A vektortér fogalma és elemi tulajdonságai	147
2. Lineáris kombináció és lineáris függés	154
3. Lineáris összefüggés és függetlenség	157
4. Generátorrendszer és bázis	160
MÁSODIK FEJEZET	
Vektortér-konstrukciók	165
1. Alterek, lineáris alakzatok	165
2. Faktorterek	171
3. Direkt összeg és direkt szorzat	174
HARMADIK FEJEZET	
Lineáris leképezések	181
1. Homogén lineáris leképezések értelmezése	181
2. Lineáris leképezések elemi tulajdonságai	184
3. A lineáris leképezések tere	189
4. Lineáris leképezések szorzása	191
5. Lineáris függvények és a duális tér	197
NEGYPEDIK FEJEZET	
Koordinatizálás	200
1. Vektorok koordinátái és leképezések mátrixa	200
2. Áttérés új bázisra	207
3. Mátrix rangja és inverze	209

ÖTÖDIK FEJEZET

Bihomomorfizmusok	216
1. Bilineáris leképezések, bilineáris formák	216
2. Bilineáris függvények mátrixa	221
3. Homogén koordináták, kvadratikus alakok a valós térben	223
4. Kvadratikus alakok négyzetösszeggé transzformálása	227
5. Bilineáris függvények és kvadratikus alakok a komplex térben	230

HATODIK FEJEZET

Euklideszi terek	236
1. A valós euklideszi tér	236
2. A valós euklideszi terek geometriája	239
3. A komplex euklideszi tér	245

HETEDIK FEJEZET

Az euklideszi tér lineáris transzformációi	246
1. Lineáris transzformációk polinomja	246
2. Lineáris transzformációk invariáns alterei az euklideszi térben	252
3. Szimmetrikus és önadjungált transzformációk	256
4. Ortogonális és unitér transzformációk	258
5. Kvadratikus alakok az euklideszi térben	264

NYOLCADIK FEJEZET

A karakterisztikus polinom	267
1. A determináns	267
2. Polinom mátrixok normálalakja, karakterisztikus polinom	272
3. Mátrixpolinomok, invariáns faktorok	281
4. A Jordan-féle normálalak	285

KILENCEDIK FEJEZET

Determinánsok alkalmazása	291
1. Lineáris egyenletrendszerek megoldása	291
2. Homogén lineáris egyenletrendszerek	293
3. A rezultáns	294
4. Lineáris egyenletrendszerek közelítő megoldása	298
5. A Cramer-szabály	299
6. Kvadratikus alakok jellegének a megállapítása	300

TIZEDIK FEJEZET

Tenzorok	305
1. A tenzorszorzat	305
2. A tenzorszorzat elemi tulajdonságai	311
3. Mátrix-előállítások, tenzor koordinátái	318
4. A tenzoralgebra, szimmetrikus és antiszimmetrikus tenzorok	320
5. Alkalmazások	324
Betűrendes mutató	327

ELŐSZÓ

Ez a tankönyv elsődlegesen az Eötvös Loránd Tudományegyetem elsőéves matematikus és alkalmazott matematikus hallgatói számára készült, e szakoknak a tematikáját követi; az elemi algebrai és a lineáris algebrai ismeretek a matematika szinte minden területén és az alkalmazásokban is nélkülözhetetlenek. Emellett a lineáris algebra szükségszerűen absztrakt tárgyalása jó átmenetet nyújt a tervezett második kötetben szereplő algebrai struktúrákhoz is.

Ma már Magyarországon (is) sok egyetemi szintű jegyzet és tankönyv foglalkozik az elemi és lineáris algebra tárgyalásával. Ezek mindegyike más felfogásban tárgyalja a fenti tananyagot, ezért nem lehet e tankönyveket rangsorolni; tulajdonképpen jól kiegészítik egymást. Ez a tankönyv az 1977-ben megjelent *Klasszikus és lineáris algebra* c. tankönyvem pótlására készült, amelynek legutóbbi kiadása is elfogyott. Tekintettel arra, hogy az idézett tankönyvhöz képest lényeges változtatásokat éreztem szükségesnek (többek között szerettem volna egységesíteni az ugyancsak nem kapható *Általános algebra* c. tankönyvvel), ezért nem tartottam jónak a fenti tankönyv újabb — lényegében változatlan — kiadását. Nem változtattam a könyv „szellemén”, a tananyagot is főleg bővítettem. A tételek bizonyításában a leglényegesebb változás az, hogy az ottani formalizmust igyekeztem elkerülni, arra törekedve, hogy a definíciók ne „ügyesek”, hanem a lényeget jobban megmutatók legyenek.

A kötet két részre oszlik. Az első rész tárgya a klasszikus vagy elemi algebra. A középiskolában tanult számfogalom átisméltése és néhány általános algebrai fogalom (elnevezés) bevezetése után a komplex számok ismertetése következik. Ezek után a mátrixok, majd a determináns bevezetésére kerül sor. E résznek a befejezéseként az egy- és többváltozatos polinomokat tárgyaljuk. Itt alapvető szempont a fogalmak minél tisztább, minél precízebb bevezetése. Csak ezután kerülhet sor az érdemi tárgyalásra.

A második rész a lineáris algebra. A lineáris algebra eredetileg elsősorban a lineáris egyenletrendszerekkel foglalkozott. Ehhez a mátrixok és ezekhez kapcsolódva a koordináták szolgáltatták a módszert. E felfogással szemben nagy változást jelentett a tömör jelölésmód, amelyben a vektorterek és a lineáris leképezések jutottak szóhoz. Ennek megfelelően a fogalmak geometriai jelentést kaptak; ezáltal sokkal világosabbá váltak. Éles ellentétként a fogalmak absztraktabbak lettek, ami az elvontabb tárgyalásmódot tette szükségessé. A fentebb említett tankönyvhöz képest igyekeztem ezen enyhíteni, ahol tudtam (mind a definíciókban, mind a tárgyalás sorrendjében). A könyvben ■ jelöli a bizonyítások, illetve definíciók és □ jelöli a megjegyzések végét.

Remélem, hogy ezt a tankönyvet sikerrel használhatják más szakok és más egyetemek hallgatói is. Elsősorban természetesen azokra a hallgatókra gondolok, akik matematikus szakra járnak. Úgy vélem, hogy egyéb matematikát tanuló egyetemi hallgatók is tanulhatnak e könyvből; mindenekelőtt algebrai módszereket.

Természetesen egyetlen könyv (de az internet sem) sem pótolhatja az élő előadás élményét. A matematikát csak úgy lehet megtanulni, ha (lehetőleg aktívan) nyomon követjük a gondolkodásmódot, az esetleges hibákat; és a tételeket, a fogalmakat és a bizonyításokat *in statu nascendi* (a születés pillanatában) láthatjuk. Semmi sem pótolhat egy vitát az előadóval. Az írott segédanyagra az ismeretek felfrissítésekor van szükség. Ettől függetlenül célszerűnek tartom azt, hogy a tankönyv a közölt tananyagon kívül lehetőleg gondolkozni is tanítson és magyarázzon. Természetesen ehhez szükséges, hogy a fogalmak, tételek és a bizonyítások (eltekintve néhány hosszadalmas és mechanikus bizonyítástól) mind megtalálhatóak legyenek a tankönyvben.

A szereplő fogalmak és tételek az elméleti és az alkalmazott matematika legkülönbözőbb területeiről származnak. Ezeknek a fogalmaknak a motivációjáról azért mondtam le, mert ez az egész tárgyalást igen hosszadalmassá és esetleg érthetlenebbé tenné. Megmaradtam az algebrai keretek között, és az alkalmazásokra való utalást a megfelelő szaktárgyakra hagytam.

Köszönetnyilvánítás. E könyv készítésében hálával tartozom azoknak, akik velem a matematikai gondolkozásmódot megismertették és megszerettették. Így NEUKOMM GYULA gimnáziumi tanáromnak, GEHÉR ISTVÁN egyetemi diáktársamnak, FUCHS LÁSZLÓ, RÉNYI ALFRÉD, PÉTER RÓZSA és mindenekfelett TURÁN PÁL egyetemi tanárainak. Hálával tartozom diákjaimnak és tanítványaimnak, akik állandó javító célzattal bírálták munkáimat; és akiktől ugyancsak nagyon sokat tanultam. Ezeknek a diákoknak a száma olyan nagy, hogy őket felsorolva óhatatlanul kimaradna jó néhány, akiket nem szeretnék megbántani. Ezért inkább egyetlen nevet sem írok ide; ők úgyis tudják, hogy róluk van szó. Hálával tartozom algebrista kollégáimnak, akik jelenlétükkel erősítették a magyar algebrista közösséget.

Vannak, akik a könyv közvetlen megjelenését is elősegítették. Hálával és köszönettel tartozom két lektoromnak, CSÁKÁNY BÉLÁNAK és PÁLFY PÉTER PÁLNAK, akik magukra vállalták az átnézés keserveit, számos értelemzavaró hibától mentve meg a könyvet. Ha valami benne maradt, az nem az ő munkájukat, hanem az enyémet minősíti.

Hálával tartozom a könyv előállításában való részvételéért FRIED KATALINNAK a szerzésért, a Nemzeti Tankönyvkiadóban PALOJTAY MÁRIÁNAK és BALASSA ZSÓFIÁNAK, akik a könyvet gondozták. Hálával tartozom az anyagi háttér biztosításáért a SZÉCHENYI PROFESSZORI ÖSZTÖNDÚNAK, valamint a T 023186 és T 029525 számú OTKA-nak. Végül, de nem utolsósorban hálával tartozom feleségemnek, HAY ERZSÉBETNEK, az erkölcsi háttér biztosításáért, türelméért és a könyv átolvasásában nyújtott segítségével.

Budapesten a 2000. évben

Fried Ervin

I. rész

ELEMI ALGEBRA

ELSŐ FEJEZET

KOMPLEX SZÁMOK

1. A természetes számoktól a valós számokig

A komplex számok vizsgálata előtt tekintsük át a valós számok alapvető algebrai tulajdonságait. A valós számokat — mint középiskolai tananyagot — tárgyalásaink során ismereteknek tételezzük fel. Ennek megfelelően a felsorolt algebrai tulajdonságokat sem fogjuk bizonyítani. (A későbbiek során ezekre bizonyos értelemben majd sor kerül.) A valós számoknak nagyon sok fontos tulajdonsága van; itt azonban csak olyanokra lesz szükségünk, amelyekben csupán e számok közötti műveletek és relációk szerepelnek. Ezeket nevezzük algebrai tulajdonságoknak.

Tárgyalni fogjuk az összeadás, kivonás, szorzás és osztás, valamint a hatványozás, gyökvonás és logaritmálás alapvető azonosságait; továbbá az oszthatósági és rendezési relációkat.

Bármely két a és b valós számnak létezik az $a + b$ összege és az ab ($a \cdot b$, vagy más jelöléssel $a \times b$) szorzata. a és b az összeg *tagjai*, illetve a szorzat *tényezői*. Mindkét műveletre, az összeadásra és a szorzásra érvényes a *kommutativitás* (*felcserélhetőség*) és az *asszociativitás* (*társíthatóság*); azaz bármely a, b, c valós számokra:

$$a + b = b + a, \quad ab = ba, \quad (a + b) + c = a + (b + c), \quad (ab)c = a(bc).$$

A szorzás az összeadásra nézve *disztributív* (*szétoosztó*), azaz $c(a + b) = ca + cb$, tetszőleges a, b, c valós számok esetében. Ebből következik az $(a + b)(c + d) = ac + ad + bc + bd$, speciálisan az $(a + b)(a + b) = aa + ab + ab + bb$ összefüggés.

Ezeket a műveleteket *direkt műveleteknek* is szokták nevezni.

Az összeg és az egyik tag ismeretében egyértelműen meghatározható a másik tag, ennek a meghatározását *kivonásnak* nevezzük. Ha $a + b = c$, akkor a $b = c - a$ jelölést használjuk, c a *kisebbitendő*, a a *kivonandó* és b a *különbség*. Bármely a valós számra a $0 = a - a$ szám ugyanaz, ezt *nullának* nevezik. 0 azzal jellemezhető, hogy tetszőleges b valós szám esetén $0 + b = b$. A $-a = 0 - a$ szám csak a -tól függ, ez az *a ellentettje* (vagy *negatívja*, vagy *additív inverze*). Érvényesek az

$$(a + b) - c = a + (b - c), \quad (a - b) + c = a - (b - c), \quad (a - b) - c = a - (b + c), \quad a + (-b) = a - b$$

összefüggések. A kivonást nem kell új műveletnek tekinteni, mert az összeadással meghatározható. A kivonást *inverz műveletnek* is szokták nevezni.

A disztributivitást felhasználva kapjuk, hogy $(a - b)c = ac - bc$, $(a + b)(a - b) = aa - bb$; továbbá $(-a)b = a(-b) = -ab$, $(-a)(-b) = ab$, valamint $0a = 0$. Ez utóbbi tulajdonsággal egyedül a 0 rendelkezik.

Ugyancsak inverz művelet az osztás, amikor a szorzat és az egyik tényező ismeretében keressük a másikat. Az osztás nem mindig végezhető el:

0-val nem lehet osztani!

Ha ugyanis létezne $\frac{a}{0}$, akkor $a = 0 \left(\frac{a}{0} \right) = 0$ volna, bármely a valós számra.

Ha $c = ab$, akkor $b = \frac{c}{a}$ (vagy $b = c/a$); itt c az *osztandó*, a az *osztó* és b a *hányados*. $a \neq 0$ esetén $1 = \frac{a}{a}$ mindig ugyanaz, 1 neve *egy*; és jellemezhető azzal, hogy bármely b valós számra $1b = b$ igaz. Ha $a \neq 0$, akkor $\frac{1}{a}$ az a *reciproka* vagy *multiplikatív inverze*. Az osztás, valamint az összeadás és a kivonás között $c \neq 0$ esetén érvényesek az alábbi kapcsolatok:

$$\frac{a+b}{c} = \frac{a}{c} + \frac{b}{c}, \quad \frac{a-b}{c} = \frac{a}{c} - \frac{b}{c}, \quad \frac{-a}{c} = \frac{a}{-c} = -\frac{a}{c}, \quad \frac{-a}{-c} = \frac{a}{c}.$$

Ezek után speciális (ismert) számköröket fogunk nézni.

A természetes számok. Az emberekben eredendően „természetesen” kialakult szám-fogalom. Ezek azok a számok, amelyeket az 1-ből számlálással nyerhetünk: 1, 2, 3, 4, ... soha abba nem hagyva a számlálást és mindig újabb és újabb számokat nyerve. Ezt az alábbi módon fogalmazhatjuk meg pontosabban:

1. 1 természetes szám.
2. Minden n természetes számnak van egy n' rákövetkezője.
3. $n' \neq 1$.
4. Ha $n' = k'$, akkor $n = k$.

Ezekkel még nem írtuk le teljesen a természetes számokat, mert a természetes számoknak alapvető tulajdonsága a *teljes indukció*. Ez azt mondja ki, hogy:

Ha

1. az 1 számnak megvan egy T tulajdonsága és
2. valahányszor egy természetes számnak megvan e tulajdonsága, akkor a rákövetkezőnek is megvan;

úgy minden természetes szám rendelkezik e tulajdonsággal.

A teljes indukció segítségével *definiálhatjuk* az összeadást és a szorzást, valamint a *kisebb* relációt is.

Megjegyzések

1. Noha 0 sem történelmileg, sem „érzelmileg” nem természetes szám, matematikai szempontból sokszor hasznos annak tekinteni. Ez különösen akkor van így, ha egy teljes indukciós bizonyításnál a két lépés bizonyítása lényegében ugyanúgy történhet, míg az állítás a 0 esetére szinte bizonyítást sem igényel.

2. A teljes indukcióval való definiálás az valójában nem bizonyítás, de ez is megtehető. Ebben az esetben *rekurzioról* beszélünk.

3. Nem definiáltuk a „tulajdonság”-ot. Ilyenképpen szinte minden értelmetlen állítás bizonyítható volna. Valójában csak olyan tulajdonságokat engedhetünk meg, amelyeket „logikai formulával” lehet megadni. Például ilyen az, hogy bármely számot a rákövetkezővel szorozva páros számot kapunk.

4. Miután „tudjuk”, mik a természetes számok, ezért nem definiálhatjuk őket, mert csak úgy lehetne definiálni, hogy valami sokkal bonyolultabbat használunk fel. Ezért a fentiekben csupán azt fogalmaztuk meg, hogy a természetes számoknak melyek az alaptulajdonságaik. Ha ebben egyetértünk, akkor ugyanazt a matematikát „űzzük”. A fenti *axiómarendszert* bevezetjük nevéről *Peano-axiómáknak* nevezzük. \square

A természetes számok körében mindig elvégezhető az összeadás és a szorzás, de általában sem a kivonás, sem az osztás nem végezhető el. A természetes számok *halmazát* \mathbb{N} -nel fogjuk jelölni. Azokat az 1-től különböző természetes számokat, amelyek felírhatók két 1-től különböző természetes szám szorzataként, *összetett számoknak* nevezzük, amelyek nem írhatók így fel, azoknak a neve *prímszám*. A *számelmélet alaptétele* kimondja, hogy minden összetett szám egyértelműen (azaz a tényezők sorrendjétől eltekintve) felbontható prímszámok szorzatára. (Ha „egytényezős” szorzatot is megengedünk, akkor csak az 1-et kell kizárnunk.)

Az egész számok. Noha történetileg a pozitív racionális, sőt az irracionális számokat is korábban ismerték, egyszerűbb előbb az egész számokat tárgyalni. Ezek a számok a természetes számokból úgy nyerhetők, hogy a fenti műveleteken kívül a kivonást is megengedjük. Az így kapott „új” számok tehát 0, -1 , -2 , -3 , -4 , \dots ; egy új szám a 0, és minden természetes számhoz „ugyanaz”, de egy $-$ jelet téve eléje. Könnyen (de hosszadalmasan) belátható, hogy a műveleteket a „megszokott” módon értelmezve, azokra a már tárgyalt összefüggések igazak. Az egész számok halmazát \mathbb{Z} -vel fogjuk jelölni; ez a halmaz az összeadáson és szorzáson kívül még a kivonásra is zárt (azaz ezek a műveletek sem vezetnek ki e számkörből). Ha az egész számok körében csak az összeadást és a kivonást nézzük, akkor erre a \mathbb{Z}^+ jelölést fogjuk használni.

Érdemes megtanulni a következő elnevezéseket:

Ha egy számhalmaz az összeadásra és a kivonásra is *zárt* (vagyis ezek a műveletek nem vezetnek ki a halmazból), akkor azt mondjuk, hogy e számhalmaz az *összeadásra nézve csoport*. Ugyanígy, ha egy számhalmaz a szorzásra és az osztásra zárt, akkor ez a *szorzásra nézve csoport*. Ha csak az összeadásra (szorzásra) való zártaságot tudjuk, akkor az összeadásra (szorzásra) vonatkozó *félcsoportról* beszélünk.

\mathbb{N} az összeadásra és a szorzásra nézve is félcsoport. Ha egy számhalmaz az összeadásra nézve csoport és a szorzásra nézve félcsoport, akkor ez az *összeadásra és a szorzásra nézve gyűrű*. Ilyen például \mathbb{Z} .

\mathbb{Z} -ben két relációt értelmezünk, az egyik az *oszthatóság*, a másik a *rendezés*.

Az $a \mid b$ reláció azt fejezi ki, hogy a a b -nek *osztója*, illetve b az a -nak többszöröse, ami azt jelenti, hogy létezik egy olyan c egész szám, amire $b = ac$. E relációnak a következő alapvető tulajdonságai vannak:

Az $a \mid b$, $a \mid (-b)$, $(-a) \mid b$, $(-a) \mid (-b)$ feltételek ekvivalensek.

Az oszthatóság *reflexív* ($a \mid a$), *antiszimmetrikus* (ha $a \mid b$ és $b \mid a$, akkor vagy $a = b$, vagy $a = -b$) és *transzítív* (ha $a \mid b$ és $b \mid c$, akkor $a \mid c$).

Ha $a \mid b$ és $a \mid c$, akkor $a \mid (b + c)$ és $a \mid (b - c)$; továbbá tetszőleges d egész szám esetén $a \mid (bd)$.

Az a egész szám pontosan akkor osztója minden egész számnak, ha $a = 1$ vagy $a = -1$. Egy a egész számnak pontosan akkor osztója minden egész szám, ha $a = 0$. (Noha $\frac{0}{0}$ értelmetlen, a $0 \mid 0$ reláció igaz.)

Az a és b egész számoknak létezik $d = (a, b)$ *legnagyobb közös osztójuk* és $c = [a, b]$ *legkisebb közös többszörösük*. Ezekre $d \mid a$, $d \mid b$, $a \mid c$, $b \mid c$ teljesül úgy, hogy az $u \mid a$, $u \mid b$, illetve $a \mid v$, $b \mid v$ feltételekből $u \mid d$, illetve $c \mid v$ következik. Ha feltesszük, hogy c is, d is vagy természetes szám, vagy 0, akkor ezek a számok egyértelműen meghatározottak. Ha $(a, b) = 1$, akkor *relatív prím* számokról beszélünk.

A másik reláció a *rendezés*, ezt a kivonással definiáljuk:

a *nagyobb*, mint b (jelben $a > b$), ha $a - b$ természetes szám. Ekkor azt is mondjuk, hogy b *kisebb*, mint a ($b < a$). Ha $a > 0$, akkor a *pozitív*, ha $a < 0$, akkor a *negatív*.

A rendezés *irreflexív* ($a > a$ soha sem teljesül), *antiszimmetrikus* ($a > b$ és $b > a$ egyszerre nem teljesülhet, és *transzítív* (ha $a > b$ és $b > c$, akkor $a > c$). (Vigyázat! Az itt szereplő antiszimmetria nem egészen ugyanaz, mint az előző; ez azért van, mert itt az egyenlőséget is kizárjuk.)

A rendezés *teljes*, azaz bármely a és b számokra $a > b$, $a = b$ és $a < b$ közül pontosan az egyik igaz. A „hozzáadás” és a pozitív számmal való szorzás *monoton*, azaz $a > b$ és tetszőleges c , továbbá pozitív d mellett $a + c > b + c$ és $ad > bd$. (Ha $d = 0$, akkor $ad = bd$, ha $d < 0$, akkor $ad < bd$ következik.)

Használatos ezzel kapcsolatban még a *nagyobb-egyenlő*, illetve *kisebb-egyenlő* ($a \geq b$, illetve $b \leq a$), ami azt jelenti, hogy vagy $a > b$, vagy $a = b$. Ennek a relációnak a tulajdonságai az előzőeből leolvashatók. Ha $a \geq b$, akkor ezek *maximuma*, illetve *minimuma*: $\max(a, b) = a$, illetve $\min(a, b) = b$.

Az a és $-a$ közül csak egyik lehet pozitív. Ezt az a abszolút értékének nevezzük és $|a|$ -kel jelöljük. Külön definíció az, hogy $|0| = 0$. Az abszolút értékre az alábbiak teljesülnek:

$$|ab| = |a| \cdot |b|, \quad |a + b| \leq |a| + |b|, \quad |a - b| \geq \big||a| - |b|\big|.$$

Az abszolút érték segítségével megfogalmazható a *maradékos osztás*:

Bármely a egész és bármely 0-tól különböző b egész számokhoz léteznek olyan q és r egész számok, amelyekre:

$$a = bq + r \quad \text{és} \quad |r| < |b|. \quad (\text{Negatív } r \text{ is megengedett.})$$

A maradékos osztás biztosítja, hogy elvégezhető az úgynevezett *euklideszi algoritmus*, amelynek segítségével előállítható két szám legnagyobb közös osztója és bizonyítható a számelmélet alaptétele.

A *racióális számok*. Ezeket a számokat úgy kapjuk, hogy még a nemnulla egész számmal való osztást is megengedjük. A racionális számok tehát $\frac{a}{b}$ alakú *törtek*. Itt a a tört *számlálója*, b a tört *nevezője*. Az őket elválasztó vízszintes vonal neve *törtvonal*. Különböző törtek is jelölhetik ugyanazt a racionális számot: $\frac{a}{b} = \frac{c}{d}$ pontosan akkor, ha $ad = bc$. Ezek körében elvégezhető az összeadás, kivonás, szorzás és a nemnulla racionális számmal való osztás (amit : jelöl):

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \quad \frac{a}{b} - \frac{c}{d} = \frac{ad - bc}{bd}, \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \quad \frac{a}{b} : \frac{c}{d} = \frac{ad}{bc},$$

amennyiben a fellépő nevezők egyike sem 0. Itt is érvényesek az egész számokra látott azonosságok. A racionális számok körében elvégezhető az osztás is (persze 0-val nem!). Ilyen esetben a számgyűrűt *számtestnek* nevezzük. A racionális számtestet \mathbb{Q} -val fogjuk jelölni. Ha a racionális számokat mint additív csoportot nézzük, akkor erre a \mathbb{Q}^+ jelölést fogjuk használni. Vigyázat! A matematikai analízisben \mathbb{Q}^+ a pozitív racionális számokat jelöli.

A racionális számok körében is beszélhetünk rendezésről; erre is hasonlóak érvényesek, mint az egészek körében. Jó tudni, hogy $\frac{a}{b} > 0$ akkor és csak akkor igaz, ha $ab > 0$. Az egész számok körében bármely egész számnál volt „közvetlenül” kisebb is és nagyobb is. A racionális számoknál ez nincs így, bármely két különböző racionális szám között van újabb: Ha $r < s$, akkor $r < \frac{r+s}{2} < s$.

Érdeemes megjegyezni, hogy a pozitív racionális számok a szorzásra nézve csoportot alkotnak.

Megjegyzések

1. A racionális szó nem azt jelenti, hogy ésszerű, hanem azt, hogy viszonyszám. A régi görögök csak a „mérőrúd” egész számú többszöröseinek a mértékét tekintették számnak; a racionális számok csak úgy jelentkeztek náluk, mint két ilyen távolságnak az aránya.

2. Érdeemes észrevenni, hogy csak a természetes számok esetében soroltunk fel axiómákat, a többi számfajtát ezekből építettük fel. A Peano-axiómarendszerről senki sem tudja bebizonyítani, hogy ellentmondásmentes. Ha az újabb, bővebb számfajtákat is axiómákkal definiálnánk, akkor egyre kellemetlenebb helyzetbe jutnánk. Ezen úgy segítünk, hogy „tudva”, mik azok az egészek, illetve racionálisak, ezekre a természetes számok segítségével egy „modell”-t építünk fel. Ha tehát a természetes számok körében nincs ellentmondás, akkor ezekben a bővebb számkörökben sincs.

Általában azokat az axiómarendszereket fogadjuk el ellentmondásmenteseknek, amelyekre létezik véges modell.

3. A törtszámok között nincsenek ott az egész számok, mert „más az alakjuk”. Ezzel szemben vannak e számkörben olyan számok, amelyeket úgy tekinthetünk, mintha ők egész számok lennének. Nevezetesen az $\frac{a}{1}$ helyett azt képzelhetjük, hogy az a egész szám áll. A velük való műveletek pontosan úgy végezhetők, mint az egészekkel. \square

A valós számok. A valós számokat nem származtathatjuk a racionális számokból „algebrai módon”. Ezek szemléletesen a következőképpen kaphatók. A rendezés tulajdonságai alapján a racionális számokat úgy képzelhetjük el, hogy ezek egy egyenesen, a *számegyenesen* helyezkednek el. Közöttük azonban „kimaradnak” pontok. Ezt a tulajdonságot pontosabban a következőképpen fejezhetjük ki: Legyen P és Q a racionális számok két olyan részhalmaza, hogy bármely racionális szám e két részhalmaz közül pontosan az egyiknek eleme; továbbá, ha p egy P -beli és q egy Q -beli szám, akkor $p < q$. Ezen felül még azt is feltesszük, hogy P -ben nincs legnagyobb racionális szám. (Ha m ilyen volna, akkor P helyett vegyük azt a P_1 halmazt, amelyet az m elhagyásával nyerünk, míg Q helyett azt a Q_1 halmazt, amelyet úgy kapunk, hogy Q -hoz hozzávesszük m -et is. P_1 és Q_1 is eleget tesznek a kirótt feltételeknek, de P_1 -ben nincs legnagyobb szám, hiszen bármely két különböző racionális szám között van tőlük különböző.) Minden ilyen úgynevezett *szelet* meghatároz pontosan egy valós számot, amely minden P -beli számnál nagyobb, de egyik Q -belinél sem. Ez a valós szám pontosan akkor „tekinthető” racionálisnak, ha Q -nak van legkisebb eleme; méghozzá ekkor pontosan ezt a racionális számot jellemeztük. Ezek alapján a valós számokra is értelmezhető a rendezés, amire a már szerepelt tulajdonságok teljesülnek. A valós számokról már említettük, hogy számtestet alkotnak. Ezt a számtestet \mathbb{R} fogja jelölni. Ha a valós számok additív csoportját nézzük, akkor az \mathbb{R}^+ jelölést fogjuk használni. Itt is vigyázni kell, mert analízisben ez a pozitív valós számok halmazát jelöli. Itt is érvényesek a már tárgyalt műveleti azonosságok, valamint a rendezésnek és az abszolút értéknek a műveletekkel kapcsolatos tulajdonságai. (Valójában éppen ez az elv teszi lehetővé az összeadás és a szorzás definícióját.)

1-nél nagyobb n természetes számmal való szorzás megegyezik a másik tényező n példányban vett összegével: $n \cdot a = \overbrace{a + \dots + a}^{n\text{-szer}}$. Külön értelmezendő az $1 \cdot a = a$ és a $0 \cdot a = 0$. Negatív egész számmal való szorzás is értelmezhető; ha n természetes szám, akkor $(-n) \cdot a = n \cdot (-a)$. Ennek analógiájára definiálható a *hatványozás*.

Ha $n > 1$ természetes szám, akkor legyen $a^n = \overbrace{a \cdot \dots \cdot a}^{n\text{-szer}}$. Legyen $a^1 = a$ és $a^0 = 1$.

Ha n természetes szám, akkor legyen $a^{-n} = \frac{1}{a^n}$ (feltéve, hogy $a \neq 0$).

Igen fontos speciális eset az $n = 2$. Az a^2 számot az a szám *négyzetének* nevezzük. Egyedül a 0 négyzete 0, minden más valós szám négyzete pozitív. Fontos összefüggés, hogy $(-a)^2 = a^2$.

Az $n = 3$ esetben is szokásos külön elnevezés: a^3 az a szám *köbe*. Könnyen belátható, hogy különböző számok köbe is különböző, pozitívaké pozitív, negatívaké negatív.

Az $a \mapsto a^2$ megfeleltetés neve *négyzetre emelés*; az $a \mapsto a^3$ megfeleltetésé *köbre emelés*.

Ha $a^b = c$, akkor a neve *alap*, b neve *kitevő* és c neve *hatvány*. Az eddigiekben a kitevő egész szám volt; más kitevő esetén mindig feltesszük, hogy az alap pozitív szám.

Ha n természetes szám, akkor az $a^n = b$ kapcsolatot $a = \sqrt[n]{b}$ is jelöli. Az a pozitivitása következtében itt a egyértelműen meghatározott. E jelölésnél *gyökvonásról* beszélünk; b a gyök *alapja*, n a *gyökkitevő* és a a *gyök*.

Az $n = 1$ esetben nem szoktak gyökvonásról beszélni. Az $n = 2$ esetben a gyökkitevőt elhagyva $a = \sqrt[2]{b}$ helyett \sqrt{b} szerepel. Ebben az esetben *négyzetgyökvonásról* beszélünk. Mint a négyzetre emelésnél láttuk, valós szám négyzete nem lehet negatív; így negatív valós számnak nincs négyzetgyöke (a valós számok körében).

Legyen $r = \frac{p}{q}$ tetszőleges racionális szám, ahol feltehető, hogy $q > 0$. Definíció szerint legyen $a^r = \sqrt[q]{a^p}$. Az alap pozitivitásából következik, hogy a racionális hatvány egyértelműen meghatározott pozitív szám. Pozitív alap esetén a racionális kitevőjű hatványra az alábbiak teljesülnek:

$$a^{r+s} = a^r \cdot a^s, \quad a^{rs} = (a^r)^s, \quad (ab)^r = a^r \cdot b^r.$$

Ha $a < b$ és $r > 0$, akkor $a^r < b^r$; ha $r < s$ és $a > 1$, akkor $a^r < a^s$. Ezeket az összefüggéseket a hatványozás *monotonitásának* nevezzük. Ez a tulajdonság és a valós számoknak a racionális számokkal való leírása lehetőséget ad arra, hogy a valós kitevőjű a^b hatványt is definiálhassuk. Erre is érvényesek a már ismertetett összefüggések. Ez pozitív alapra speciális esetként tartalmazza a gyökvonást. Mivel a hatványozás nem kommutatív, ezért egy másik inverz művelete is van, a *logaritmálás*, amikor az alaptól és a hatványból határozzuk meg a kitevőt. Az $a^b = c$ esetben a következő elnevezések használatosak: a az *alap*, c a *logaritmálendő* vagy *numerus* és b a *logaritmus*. Ebben az esetben a $\log_a c = b$ jelölés használatos.

Halmaz, reláció, függvény. Ezek a fogalmak nem tartoznak ugyan a számfogalom körébe, de szerepelnek a középiskolai tananyagban, és a továbbiakban itt is szükség lesz rájuk. A halmazokról van szemléletes képünk. A halmazokat még annyira sem írjuk itt le, mint amennyire a természetes számokkal tettük a Peano-axiómáknál. Ennek az az oka, hogy a természetes szám fogalma és azoknak számlálás útján való nyerése valóban természetes; míg a halmazok „absztrakt” tulajdonságai sokkal nehezebben beláthatóak.

A halmazokat általában latin nagybetűkkel fogjuk jelölni. A halmazoknak „eredendő” tulajdonsága, hogy elemeik vannak. Ezeket az elemeket általában latin kisbetűkkel jelöljük. Azt, hogy x *elem*e az A halmaznak, úgy fogjuk jelölni, hogy $x \in A$. Van egyetlen halmaz, amelynek egyetlen eleme sincs; ezt úgy nevezik, hogy *üres halmaz*; ezt \emptyset jelöli. Azt mondjuk, hogy B *részhalmaza* A -nak ($B \subseteq A$), ha $b \in B$ esetén $b \in A$ is teljesül. Az A és B halmazok *direkt szorzatán* azt az $A \times B$ halmazt értjük, amelynek elemei azok az (a, b) párok, amelyekre $a \in A$ és $b \in B$. Több halmaznak is képezhető a direkt szorzata: $A \times B \times C$ az (a, b, c) hármassokból áll, ahol $a \in A$, $b \in B$, $c \in C$. Megemlítjük, hogy ezt sokszor ugyanannak fogjuk tekinteni, mint az $A \times (B \times C)$ vagy az $(A \times B) \times C$ halmazt; noha formálisan az előbbinek az elemei az $(a, (b, c))$ elemek, az utóbbinak pedig az $((a, b), c)$ alakú elemek.

Ha a halmazt elemei felsorolásával adjuk meg, akkor a felsorolt elemeket kapcsos zárójelek közé tesszük. Például az $\{1, 2, 5\}$ halmaz elemei a felsorolt három szám, míg az

$\{1, 2, 5, a, b\}$ halmazé ezeken kívül még az a és a b betű is. A felsorolásnál mindegy, hogy az elemeket milyen sorrendben írjuk, a halmaz nem változik. A halmaz elemei különbözőek, így az $\{1, 2, 5, 2, 2, 1\}$ halmaznak is csak három eleme van.

Ha a halmazt úgy akarjuk megadni, hogy egy halmaz elemeiből valamilyen „utasítással” választunk ki elemeket, akkor is kapcsos zárójelet használunk, de közöttük húzunk egy függőleges vonalat. Ettől balra szerepelnek az elemek és az, hogy honnan választjuk ki őket; míg jobbra az, hogy milyen módon történik a kiválasztás. Például az $A = \{i \in \mathbb{N} \mid 3 < i \leq 7\}$ azt jelenti, hogy a természetes számok közül azokat tekintjük, amelyek 3-nál nagyobbak, de 7-nél nem. Ez a halmaz tehát a $\{4, 5, 6, 7\}$.

Az A halmazon értelmezett *reláció* egy „kapcsolat” az A elemei között. Azt, hogy egy ϱ reláció fennáll az a és b elemek között, általában úgy jelöljük, hogy $a\varrho b$ (mint például $a < b$). Használatos még a $\varrho(a, b)$ jelölés is. Mivel itt két elem szerepel, ezért ilyenkor *kétváltozós relációról* beszélünk. Léteznek még *többsváltozós relációk* és *egyváltozós reláció* is.

Ugyancsak nem magyarázzuk, hogy mi a *függvény* vagy *leképezés*. Ha az f függvény az A halmazt képezi le a B halmazra, akkor ezt $f : A \rightarrow B$ vagy $A \xrightarrow{f} B$ fogja jelölni. Azt mondjuk, hogy $A = D(f)$ az f *értelmezési tartománya* és $B = R(f)$ az f *értékkészlete*. (Általában csak az $f(a)$ alakú elemek halmazát szokták értékkészletnek nevezni.) Ha f az A -beli a elemet a B -beli b elemre képezi le, ezt $b = f(a)$ vagy $f : a \mapsto b$ jelöli (figyeljük meg, hogy itt a nyílnek „talpa” van). (Néha a b elem kiírása nélkül csak $a \mapsto f(a)$ szerepel.)

Ha B minden eleme $f(a)$ alakú, akkor azt mondjuk, hogy $f : A \rightarrow B$ *szűrjektív*. Ha különböző A -beli elemek képe is különböző, azaz minden $b \in B$ elemhez legfeljebb egy olyan $a \in A$ található, amelyre $b = f(a)$, akkor *injektív* függvényről beszélünk. Ha mindkét feltétel teljesül, akkor a függvény *bijektív*. Bijektív függvényre igen fontos példa egy-egy halmaz *identitása*, az a függvény, amely a szóban forgó halmaz minden elemének önmagát felelteti meg. Természetesen az identitás(függvény) minden halmaznál más. Az A halmaz identitását 1_A fogja jelölni (erre tehát $D(1_A) = R(1_A)$ és ha $a \in A$, akkor $1_A : a \mapsto a$).

A függvények körében nagyon fontos művelet a *kompozíció*: Ha $f : A \rightarrow B$ és $g : B \rightarrow C$, akkor a $g \circ f$, vagy röviden $gf : A \rightarrow C$ függvény definíció szerint legyen az $a \mapsto g(f(a))$. Könnyen látható, hogy például a fenti f függvényre $f \circ 1_A = 1_B \circ f = f$.

Az identitások segítségével könnyen jellemezhetők a szereplő különféle függvényfajták. Legyen $f : A \rightarrow B$. Az f függvény pontosan akkor injektív, ha van olyan $g : B \rightarrow A$ függvény, amelyre $g \circ f = 1_A$, pontosan akkor szűrjektív, ha van olyan $g : B \rightarrow A$ függvény, amelyre $f \circ g = 1_B$, és pontosan akkor bijektív, ha van olyan $g : B \rightarrow A$ függvény, amely mindkét feltételt kielégíti. Az első esetben g szűrjektív, a második esetben injektív, a harmadik esetben pedig bijektív lesz. Az első két esetben a g függvény nem feltétlenül egyértelmű (csak ha f bijektív); a harmadik esetben viszont csak egy ilyen g függvény létezik. Ekkor ezt a függvényt $g = f^{-1}$ jelölheti és azt mondjuk, hogy g az f *inverz függvénye*.

Az algebrában alapvető jelentőségűek azok a leképezések, amelyek *művelettartók*. Ilyenekkel már találkoztunk is: Művelettartó az a leképezés, amely az a egész számnak

megfelelteti az $\frac{a}{1}$ törtet. Ugyancsak művelettartó leképezést nyertünk, amikor az r racionális számnak azt a P , Q részhalmazpárt (tehát valós számot) feleltettük meg, amelynél r a Q -nak a legkisebb eleme. (Gondoljuk meg, miképpen adjuk össze a valós számoknál említett részhalmazpárokat.) Ezek a leképezések teszik lehetővé, hogy az újabb számkörben „felfedezhessük” a régit.

A művelettartó leképezéseket *homomorfizmusoknak* nevezzük. A függvények tulajdonságának megfelelően beszélünk *injektív homomorfizmusról*, illetve *szürjektív homomorfizmusról*. Fontossága miatt a bijektív homomorfizmus külön nevet kapott, ezt *izomorfizmusnak* hívjuk. Könnyen belátható, hogy ha $f : A \rightarrow B$ izomorfizmus, akkor $f^{-1} : B \rightarrow A$ is az.

Megjegyzések

1. A relációkat tekinthetjük úgy, mint az $A \times A$, vagy általánosabban, mint az $A \times B$ részhalmazait; nevezetesen a ϱ relációval együtt tekinthetjük azokat az $(a, b) \in A \times B$ párokat, amelyekre $a\varrho b$ teljesül. Mivel „ennél több” egyetlen relációra sem mondható, ezért relációnak nevezhetjük az $A \times A$ — vagy általában tetszőleges direkt szorzat — részhalmazait.

2. A függvényeket viszont értelmezhetjük úgy, mint speciális relációkat. Egy $f : A \rightarrow B$ függvényt akkor „ismerünk”, ha ismerjük az összes $(a, f(a)) \in A \times B$ elemet. Világos, hogy ez egy φ reláció, amely azzal a tulajdonsággal rendelkezik, hogy ha (a, b_1) és (a, b_2) mindegyike a „relációhoz tartozik”, akkor $b_1 = b_2$.

3. Érdemes észrevenni, hogy ha az $f : A \rightarrow B$ és $g : B \rightarrow A$ függvényekre $g \circ f = 1_A$, akkor f biztosan injektív és g biztosan szürjektív. Ebben az esetben azt mondjuk, hogy g az f *balinverze* és f a g *jobb inverze*. Ha még $f \circ g = 1_B$ is teljesül, akkor mindkét függvény bijektív; egymás inverzei. \square

Szomma és produktum. (Noha középiskolában nem kötelező anyag, mégis szükségünk lesz két „rövidítésre”.)

Legyen $I = \{i \in \mathbb{N} \mid k \leq i \leq n\}$ és tegyük fel, hogy adott egy $f : I \rightarrow A$ függvény. Ilyen esetben $f(i)$ helyett használatos az a_i jelölés; az i neve *index* és I egy *indexhalmaz*.

k -től és n -től függően a következőképpen értelmezzük az

$$S = \sum_{i=k}^n a_i, \quad \text{illetve} \quad P = \prod_{i=k}^n a_i$$

jeleket:

1. Ha $k < n$: $S = a_k + \dots + a_n$ és $P = a_k \cdot \dots \cdot a_n$ ($n - k + 1$ tag, illetve tényező).

2. Ha $k = n$: $S = P = a_k$.

3. Ha $k = n + 1$: $S = 0$ és $P = 1$.

4. Ha $k > n + 1$: $S = -\left(\sum_{i=n+1}^{k-1} a_i\right)$, illetve $P = \left(\prod_{i=n+1}^{k-1} a_i\right)^{-1}$.

Az S , illetve P jeleket a következőképpen olvassák: szumma i egyenlő k -tól n -ig a_i , illetve produktum i egyenlő k -tól n -ig a_i . A fenti definícióval érvényes a $\sum_{i=k}^n a_i = \sum_{i=k}^r a_i + \sum_{i=r+1}^n a_i$ összefüggés, illetve ennek a produktumra vonatkozó analogonja.

Ha az $\{a_i \mid i \in \mathbb{N}\}$ számok közül csak véges sok különbözik 0-tól (illetve 1-től), akkor a megfelelő végtelen összeget, illetve szorzatot is értelmezhetjük: csak a 0-tól (illetve 1-től) különbözőket adjuk (illetve szorozzuk) össze.

Ha a k és n „határok” az adatokból világosak, akkor nem írjuk ki őket. Ugyancsak nem írjuk ki az indexeket sem, ha azok elhagyása nem okoz zavart.

A produktumnak igen fontos speciális esete a következő: $n! = \prod_{i=1}^n i$ (olv.: n faktoriális). A definícióból következik, hogy $0! = 1! = 1$; míg $n > 1$ esetén $n!$ az n -ig terjedő természetes számok szorzata.

Műveletek maradékokkal. Végezetül néhány példát adunk olyan halmazokra, amelyeknek az elemei nem számok, de a műveletek természetes módon definiálhatók rájuk. Továbbá az is igaz, hogy ezekre a műveletekre teljesülnek a számokra megismert műveleti szabályok. (De nem értelmezhetők e halmazon a tárgyalt relációk.)

Tekintsünk egy rögzített, 1-nél nagyobb m természetes számot. A vizsgált halmaz elemei az egész számok részhalmazai lesznek. Két egész szám akkor tartozzék ugyanabba a részhalmazba, ha a különbségük osztható m -mel. Az oszthatóság tulajdonságai alapján ekkor minden szám pontosan egy ilyen halmazba esik, amelyeket modulo m vett *maradékosztályoknak* nevezünk. Így ezeket bármely elemük egyértelműen meghatározza; jelölje az i -t tartalmazó maradékosztályt $[i]$.

Az oszthatósági tulajdonságokból könnyen látható, hogy az $[i + j]$, illetve $[i \cdot j]$ maradékosztály nem az i , illetve j számoktól, hanem csupán az $[i]$ és $[j]$ maradékosztályoktól függ. Éppen ezért ezek tekinthetők az $[i]$ és $[j]$ maradékosztályok összegének, illetve szorzatának: $[i] + [j] = [i + j]$ és $[i] \cdot [j] = [i \cdot j]$. A műveleti azonosságok teljesülnek, és elvégezhető a kivonás: $[i] - [j] = [i - j]$. Ezeket a maradékosztályokat az adott műveletekre nézve *gyűrűnek*; pontosabban *modulo m vett maradékosztály-gyűrűnek* nevezzük. Ezt a gyűrűt \mathbb{Z}_m jelöli. Ebben a gyűrűben pontosan akkor végezhető el az osztás, ha $m = p$ prímszám. Azt mondjuk, hogy a *modulo p vett maradékosztályok testet alkotnak*.

A \mathbb{Q} vagy az \mathbb{R} esetében a maradékosztályok szorzata problémát jelent, az összeadás és a kivonás viszont nem. Különösen fontos az az eset, amikor „modulo” 1 tekintjük e számokat, azaz két racionális számot akkor veszünk ugyanabba a „maradékosztály”-ba, ha különbségük egész szám.

2. A komplex számok bevezetése

A valós számkör sok feladat megoldásánál nem bizonyul elégségesnek. A középiskolai tanulmányok során is találkozhattunk olyan másodfokú egyenletekkel, amelyeknek nem volt gyökük a valós számok körében. Ez tulajdonképpen nem okoz gondot, mert ilyen esetben egyszerűen azt mondhattuk, hogy a kérdéses egyenletnek nincs gyöke. (Hiszen például az $x = x + 1$ egyenletnek semmilyen számkörben sem lehet gyöke.) Nagyon sokáig fel sem merült újabb számok bevezetésének a szükségessége avégett, hogy ezeknek a másodfokú egyenleteknek is legyen gyökük. Az újkor elején — amikor a harmadfokú egyenletek megoldását is megtalálták — megváltozott a helyzet. (Ezt majd a testbővítések tárgyalásánál értjük meg.) Azt tapasztalták, hogy bizonyos értelmetlen „valamik”-kel úgy számolva, mintha azok számok lennének, megkaphatjuk az egyenlet gyökeit. Ami még ennél is lényegesebb, ez éppen akkor fordult elő, amikor az egyenletnek három különböző valós gyöke volt. Más úton viszont nem lehetett ezeket a gyököket megtalálni.

Ennek a felfedezésnek a következménye, hogy ezeket a „valami”-ket kénytelenek voltak számoknak tekinteni. Ennek ellenére a matematikusok jó része sokáig idegenkedett az új, úgynevezett *komplex számoktól*. Azóta a komplex számok természetes megszokottá váltak és igen széles körben nyújtanak fontos segédeszközt.

A komplex számok bevezetésével olyan számkört akarunk találni, amelyben mind a négy „alapművelet” elvégezhető (tehát számtest), és még a negatív számok is felírhatóak négyzetként (azaz van *négyzetgyökük*).

Könnyen beláthatók a következők:

Ha ilyen számok vannak, akkor közöttük létezik egy olyan is, amelynek a négyzete -1 , jelölje ezt — pontosabban egy ilyet — a $*$ jel. Erre tehát $*^2 = -1$. Ha a műveleti azonosságok érvényben maradnak, akkor $*$ természetes kitevőjű hatványai rendre: $*$, -1 , $-*$, 1 ; ahonnan kezdve ismét előlről kezdődik a sor. Ennek megfelelően az új számkör elemei $a + b \cdot *$ alakúak (a és b számok — azaz valós számok).

Itt a következő kérdések merülnek fel: 1. Mind különbözőek-e ezek a számok? 2. Hogyan végzünk műveleteket ezekkel a számokkal? Pontosabban szólva eközben nem lépnék-e fel újabb számok, illetve nem jutunk-e ellentmondásra?

Ha $a + b* = c + d*$ és $b \neq d$, akkor a műveleti azonosságokból $* = \frac{a - c}{d - b}$ következne. Tekintettel arra, hogy ez valós szám, ennek a négyzete nem lehet -1 . A $b = d$ esetben viszont az $a = c$ összefüggést kapjuk, tehát a különböző formájú új számok valóban különbözőek.

A műveleti azonosságokból azonnal kapjuk, hogy a műveleteket a következőképpen kell végezni: $(a + b*) + (c + d*) = (a + c) + (b + d)*$, $(a + b*)(c + d*) = (ac - bd) + (ad + bc)*$. Elvégezhető a kivonás is, erre $(a + b*) - (c + d*) = (a - c) + (b - d)*$ adódik.

Az osztás elvégezhetőségét majd a „precíz” bevezetés után tárgyaljuk. Ehhez a komplex számokra egy *modellt* kell adni, amelyik a fenti vizsgálaton alapszik.

1.1. Definíció. Komplex számoknak nevezzük a valós számokból álló (a, b) számpárokat, amelyek egyenlőségét, összegét és szorzatát az alábbiakkal definiáljuk:

- (i) $(a, b) = (c, d)$ akkor és csak akkor, ha $a = c$ és $b = d$;
 (ii) $(a, b) + (c, d) = (a + c, b + d)$ és $(a, b)(c, d) = (ac - bd, ad + bc)$.

A (ii) alatt definiált két művelet közül az elsőt a komplex számok összeadásának, a másodikat a komplex számok szorzásának nevezzük. ■

1.1. Tétel. *A komplex számok körében mindkét művelet kommutatív és asszociatív, továbbá a szorzás az összeadásra nézve disztributív.*

Bizonyítás. A bizonyításnál minden esetben két komplex szám megegyezését kell belátni, amit az (i) összefüggés szerint tehetünk meg.

Mivel $(a, b) + (c, d) = (a + c, b + d)$ és $(c, d) + (a, b) = (c + a, d + b)$ jobb oldalai a valós számokra vonatkozó összeadás kommutativitása alapján megegyeznek, ezért a komplex számok összeadása is kommutatív. Az összeadás asszociativitásának a bizonyítása is egyszerű, bár hosszadalmasabb:

$$(a, b) + [(c, d) + (e, f)] = (a, b) + (c + e, d + f) = (a + (c + e), b + (d + f)),$$

$$[(a, b) + (c, d)] + (e, f) = (a + c, b + d) + (e, f) = ((a + c) + e, (b + d) + f),$$

és ezek az \mathbb{R} -beli asszociativitás alapján egyenlők.

A szorzás kommutativitását bizonyítjuk:

$$(a, b)(c, d) = (ac - bd, ad + bc) \text{ és } (c, d)(a, b) = (ca - db, cb + da),$$

amihez az \mathbb{R} -beli szorzás kommutativitásán kívül az összeadásé is kell. Hasonlóképpen bizonyítható a szorzás asszociativitása, de ez még hosszadalmasabb:

$$(a, b)[(c, d)(e, f)] = (a, b)(ce - df, cf + de) =$$

$$= (a(ce - df) - b(cf + de), a(cf + de) + b(ce - df)),$$

$$[(a, b)(c, d)](e, f) = (ac - bd, ad + bc)(e, f) =$$

$$= ((ac - bd)e - (ad + bc)f, (ac - bd)f + (ad + bc)e).$$

Végül a disztributivitás bizonyítása:

$$(a, b)[(c, d) + (e, f)] = (a, b)(c + e, d + f) =$$

$$= (a(c + e) - b(d + f), a(d + f) + b(c + e)),$$

$$(a, b)(c, d) + (a, b)(e, f) = (ac - bd, ad + bc) + (ae - bf, af + be) =$$

$$= (ac - bd + ae - bf, ad + bc + af + be).$$

Az egyenlőséget mindkét esetben a valós számokra vonatkozó azonosságok teljesülése biztosítja. ■

1.2. Tétel. *A komplex számok körében elvégezhető a kivonás, létezik nulla és minden komplex számnak létezik ellentettje.*

Bizonyítás. A kivonás elvégezhetősége azt jelenti, hogy adott (a, b) és (c, d) komplex számokhoz található olyan (x, y) komplex szám, amelyre $(a, b) + (x, y) = (c, d)$. A komplex számok összeadásának és egyenlőségének a definíciója alapján ez csak úgy lehet, ha $x = c - a$ és $y = d - b$. Azt, hogy ez valóban megfelel a követelménynek, az $(a, b) + (c - a, d - b) = (a + c - a, b + d - b) = (c, d)$ egyenlőség bizonyítja.

Ebből már világos, hogy a nulla komplex szám csak a $(0, 0)$ lehet, és valóban $(0, 0) + (a, b) = (0 + a, 0 + b) = (a, b)$. Hasonló módon adódik az is, hogy $-(a, b) = (-a, -b)$. ■

1.3. Tétel. *A komplex számok körében minden nullától különböző számmal egyértelműen lehet osztani; a komplex számok tehát számtestet alkotnak. Ezt a számtestet \mathbb{C} jelöli.*

Bizonyítás. Feladatunk tehát, hogy a $(c, d) \neq (0, 0)$ és a (a, b) komplex számokhoz olyan (x, y) komplex számot találjunk, amelyre $(c, d)(x, y) = (a, b)$ teljesül. A komplex számok szorzásának és egyenlőségének a definíciója alapján ez a

$$cx - dy = a \quad \text{és} \quad cy + dx = b$$

összefüggéshez vezet. Adjuk hozzá az első egyenlőség c -szeresét a második egyenlőség d -szereséhez, illetve vonjuk ki az első egyenlőség d -szeresét a második egyenlőség c -szereséből. Ekkor a

$$(c^2 + d^2)x = ac + bd \quad \text{és} \quad (c^2 + d^2)y = bc - ad$$

egyenlőségekhez jutunk. Mivel $(c, d) \neq (0, 0)$ és 0-tól különböző valós szám négyzete pozitív, ezért $c^2 + d^2 \neq 0$. Eszerint csak az lehetséges, hogy $x = \frac{ac + bd}{c^2 + d^2}$ és $y = \frac{bc - ad}{c^2 + d^2}$. Ha tehát elvégezhető az osztás, akkor csak ez lehet az eredmény. Egyszerű számolással belátható, hogy ez valóban megfelel, tehát az osztás elvégezhető. ■

Mivel a komplex számok (valós szám, valós szám) alakúak, ezért formailag nincsenek köztük a valós számok. A konstrukció alapján azonban világos, hogy az $(a, 0)$ alakú komplex szám pontosan az a valós számnak felel meg. Pontosabban szólva:

1.4. Tétel. *Az $a \mapsto (a, 0)$ függvény egy $\mathbb{R} \rightarrow \mathbb{C}$ injektív homomorfizmus. Ennek megfelelően a továbbiakban az $(a, 0)$ komplex szám helyébe a -t írhatunk.*

Bizonyítás. A megfeleltetés egyértelműsége azonnal következik a komplex számok egyenlőségének a definíciójából, hiszen, ha $(a, 0) = (b, 0)$, akkor $a = b$. A homomorfizmus azt jelenti, hogy művelettartó, azaz az összeg képe a képek összege és a szorzat képe a képek szorzata (ebből már azonnal következik, hogy hasonló igaz a különbségre és a hányadosra is). A komplex számok összeadásának és szorzásának a definíciójából

$$(a, 0) + (b, 0) = (a + b, 0) \quad \text{és} \quad (a, 0)(b, 0) = (ab, 0)$$

következik, ami az összeg- és szorzattartást bizonyítja. (A különbségre és a hányadosra vonatkozó állítás abból adódik, hogy ezeket az összeadással, illetve a szorzással egyértelműen lehet definiálni.) ■

1.5. Tétel. \mathbb{C} elemei egyértelműen felírhatók $a+b \cdot i$ alakban, ahol a és b valós számok és $i^2 = -1$, ahol i a $(0, 1) \in \mathbb{C}$ párt jelöli.

Bizonyítás. A szorzás definíciója szerint az $i = (0, 1)$ komplex számra $i^2 = (0, 1)(0, 1) = (0 - 1, 0 + 0) = (-1, 0) = -1$. A komplex számokra értelmezett műveletekből

$$a + b \cdot i = (a, 0) + (b, 0)(0, 1) = (a, 0) + (0 - 0, b + 0) = (a, b)$$

következik, amely a felírás egyértelműségét adja, de egyszersmind a felírhatóságot is, hiszen a komplex számok (a, b) alakúak. ■

Megjegyzés. A komplex számok bevezetésénél egészen sarkítva kellett foglalkozni a *modellalkotással*, noha erről már az előző „számkörbővítéseknél” is volt szó. Amikor még csupán a valós számoknak van értelmük, akkor — ha elfogadjuk is az $i^2 = -1$ tulajdonságú „új” szám létezését — nincs értelme annak, hogy ezt megszorozzuk egy valós számmal, illetve ehhez a szorzathoz még egy valós számot hozzá is adjunk. Éppen ezért egyszerre kellett az összes „valamit” tekinteni, értelmezni reájuk a műveleteket és észlelni, hogy ezek „lényegében” tartalmazzák a régieket is, azokra az új műveletek megegyeznek a régiekkel; és az új „valamik” úgy viselkednek, ahogy szeretnénk.

Az eljárás három részből áll:

1. *Elképzeljük, hogy milyen elemeket kapunk. A feltételezett azonosságokkal számolva megállapítjuk, hogy milyen alakúak lesznek az új elemek és miképpen kell velük műveleteket végeznünk.*
2. *Definiáljuk az új elemeket és a velük végzendő új műveleteket.*
3. *Ellenőrizzük, hogy amit kaptunk, megfelel-e kívánalmainknak. Rendszerint azt is meg kell nézni, hogy a régi elemek „megfelelően” ott vannak-e az újak között.*

A matematikai gondolatok szempontjából az első rész a legfontosabb, hiszen itt fogalmazódik meg, hogy mire van szükségünk, mit és hogyan akarunk tenni. Lehet, hogy meglepően hangzik, de itt egyáltalán nincs szükség a matematikai precizitásra. A matematikai precizitás ugyanis csak akkor szükséges, amikor elmondjuk — akár másoknak, akár magunknak — az eredményt, amit kaptunk. Ez a *heurisztika*, a „rájövés művészete”. Éppen ezért sok esetben célszerű ezt — és csak ezt — elmondani.

A második rész az, amit „száraz matematikának” is lehet nevezni. Itt nincs szükség „érzelmekre”, igen pontosan meg kell adni a definíciókat. Azt is mondhatjuk, hogy ez a „matematikai lényeg”.

A harmadik részben ugyancsak aprólékosan kell eljárni, hiszen itt dől el, hogy tényleg minden az előzetes elgondolás szerint történik-e. A számolások sokszor hosszadalmasak, megismétlődnek és unalmasak. Ezeket rendszerint rá lehet bízni azokra, akikkel az eredményt közöljük. □

A komplex szám fogalma a legtágabb „közismert” számfogalom. Minden „másfajta” számot úgy tekinthetünk, mint speciális tulajdonságú komplex számot. Éppen ezért a továbbiakban „szám” mindig azt fogja jelenteni, hogy komplex szám. Minden más esetben megmondjuk, milyen speciális esetre gondolunk.

A komplex számok bevezetésénél abból indultunk ki, hogy egy olyan * valamit keresünk, amelynek a négyzete -1 . Találtunk is ilyet, amit végül is i -vel jelöltünk. (Ezt a számot a matematikában *mindig* i jelöli.) Ezzel a tulajdonsággal viszont nemcsak i , hanem $-i$ is rendelkezik. *Nincs mód* annak az eldöntésére, hogy melyikük az i és melyikük a $-i$. Pontosabban szólva, ha i helyébe mindenütt $(-i)$ -t teszünk, az összes művelet „ugyanúgy” végezhető el. Ezzel egy időben célszerű néhány más fogalmat is definiálni:

1.2. Definíció. A \mathbb{C} komplex számtest i elemét *komplex egységnek* nevezzük.

A $z = a + bi$ komplex számnak $a = \Re(z)$ a valós és $b = \Im(z)$ a képzetes része. ■

1.6. Tétel. Az $a \in \mathbb{C} \rightarrow \mathbb{C}$ leképezés, amely $a z = a + bi$ komplex számhoz a konjugáltját, $a \bar{z} = a - bi$ komplex számot rendeli hozzá, egy izomorfizmus, amelyet kétszer alkalmazva az identitást nyerjük.

Emellett igazak a következők:

- (1) $S(z) = z + \bar{z}$ valós szám, amelyet z nyomának nevezünk.
- (2) $N(z) = z\bar{z}$ nemnegatív valós szám, amely pontosan akkor 0, ha $z = 0$. Ennek neve a z normája.

$z, w \in \mathbb{C}$ esetén $S(z + w) = S(z) + S(w)$ és $N(z \cdot w) = N(z) \cdot N(w)$.

$|z| = \sqrt{N(z)}$ neve a z abszolút értéke; $|z| = 0$ pontosan akkor, ha $z = 0$, egyébként az abszolút érték pozitív. $|z \cdot w| = |z| \cdot |w|$.

Bizonyítás. Tekintsük azt a $\Phi : \mathbb{C} \rightarrow \mathbb{C}$ függvényt, amelyre $\Phi : z \mapsto \bar{z}$. $\Phi^2(a + bi) = \Phi(a - bi) = a - (-b)i = a + bi$ alapján Φ -t kétszer alkalmazva valóban az identitást nyerjük. Erre úgy utalunk, hogy a konjugálás *involúció*. Ebből azonnal következik, hogy Φ bijekció: Ha $\Phi(z) = \Phi(w)$, akkor $z = \Phi^2(z) = \Phi^2(w) = w$; és bármely $z \in \mathbb{C}$ számra $z = \Phi(\Phi(z))$.

A művelettartáshoz legyen $z = a + bi$ és $w = c + di$. Ekkor $\bar{z} = a - bi$ és $\bar{w} = c - di$. Így

$$\bar{z} + \bar{w} = (a - bi) + (c - di) = (a + c) - (b + d)i = \overline{z + w};$$

$$\bar{z} \cdot \bar{w} = (a - bi) \cdot (c - di) = (ac - bd) - (ad + bc)i = \overline{z \cdot w}.$$

A konjugálás különbség- és hányadostartásának belátásához legyen $u = z - w$ és $v = \frac{z}{w}$ (itt $w \neq 0$). Ebből $z = u + w$, illetve $z = vw$ alapján — az összeg- és szorzattartás következtében $\bar{z} = \bar{u} + \bar{w}$, illetve $\bar{z} = \bar{v} \cdot \bar{w}$. (Az involúció tulajdonságai miatt a második esetben $\bar{w} \neq 0$.) A kivonás és az osztás egyértelműsége biztosítja a különbség- és hányadostartást.

Az (1) állítás abból adódik, hogy $S(z) = 2\Re(z)$, míg (2) abból, hogy $N(z) = (\Re(z))^2 + (\Im(z))^2$.

A nyomra és a normára vonatkozó állítás könnyen belátható abból, hogy a konjugálás összeg-, illetve szorzattartó. Valóban $S(z + w) = (z + w) + \overline{z + w} = (z + w) + \bar{z} + \bar{w} = S(z) + S(w)$. A szorzatra vonatkozó állítás bizonyítása teljesen hasonló módon történik. Az abszolút értékre vonatkozó állítás az előzőekből következik, tekintettel arra, hogy nemnegatív valós szám négyzetgyöke definíció szerint nemnegatív. ■

1.7. Tétel. Egy komplex szám pontosan akkor egyenlő konjugáltjával, ha valós; pontosan akkor egyenlő abszolút értékével, ha nemnegatív valós; pontosan akkor negatívja abszolút értékének, ha nempozitív valós.

Tetszőleges $z = a + bi$ komplex számhoz pontosan egy olyan $w = x + yi$ komplex szám van, amelyre $w^2 = z$, továbbá vagy $x > 0$, vagy $x = 0$ és ekkor $y \geq 0$. Ezt a

számat a z négyzetgyökének nevezzük és $w = \sqrt{z}$ -vel jelöljük; ily módon egy (egyértékű) függvényhez jutunk.

Bizonyítás. Az, hogy az $a+bi$ komplex szám megegyezik a konjugáltjával, azt jelenti, hogy $a+bi = a-bi$, ami a komplex számok egyenlőségének a definíciója szerint pontosan akkor áll fenn, ha $b = -b$, azaz, ha $b = 0$. Mivel egy komplex szám abszolút értéke nemnegatív valós, ezért a feltétel mindkét esetben szükséges. Ha $z = a \geq 0$, akkor a definícióból $|z| = \sqrt{(a^2+0)} = a$, míg a $z = a \leq 0$ esetben $|z| = \sqrt{(a^2+0)} = -a$.

A továbbiakhoz mindenekelőtt megnézzük, hogy milyen $w = x+yi$ elégítheti ki egyáltalában a $w^2 = z$ összefüggést. A négyzetre emelést elvégezve $z = (x^2 - y^2) + 2xyi$ adódik. A komplex számok egyenlőségének a definíciója alapján azt kapjuk, hogy

$$x^2 - y^2 = a \quad \text{és} \quad 2xy = b.$$

Ebből már meghatározható x és y lehetséges értéke, de az egyszerűbb számolás végett célszerű a második egyenlőség helyett az

$$a^2 + b^2 = N(z) = N(w^2) = (N(w))^2 = (x^2 + y^2)^2,$$

illetve az ebből adódó $x^2 + y^2 = \sqrt{a^2 + b^2}$ egyenlőséget venni figyelembe. Így az

$$x^2 = \frac{a + \sqrt{a^2 + b^2}}{2} \quad \text{és} \quad y^2 = \frac{-a + \sqrt{a^2 + b^2}}{2}$$

összefüggésekhez jutunk. Mivel $a^2 + b^2 \geq a^2$, ezért egyik tört sem negatív, tehát ilyen tulajdonságú x és y valós számok léteznek. Ez elvileg négy lehetőséget enged meg. Most célszerű figyelembe venni az elmellőzött $2xy = b$ egyenlőséget. Ha b pozitív, akkor a és b megegyező előjelűek, ha negatív, akkor különböző előjelűek; míg a $b = 0$ esetben valamelyikük 0. (Ez utóbbi esettől egyelőre tekintsünk el.) Ekkor mindkét esetben két-két lehetőség marad, s a két komplex szám egymás negatívja. Behelyettesítéssel meggyőződhetünk arról, hogy $w^2 = z$ mindkét esetben fennáll. Ha $b = 0$, akkor az $a > 0$ esetben $w = \pm\sqrt{a}$, míg az $a < 0$ esetben a $w = \pm\sqrt{-a} \cdot i$, végezetül az $a = 0$ esetben $w = 0$ adódik. Ez utóbbi esetben csak egyetlen szám négyzete z , a többi esetben pontosan két ilyen található; egymás negatívjai. A kirótt megszorítás pontosan azt teszi lehetővé, hogy a négyzetgyökvonást egyértékű függvénynek tekinthessük. ■

Megjegyezzük, hogy a „négyzetgyökfüggvény” nem rendelkezik a valós számoknál megszokott összes tulajdonsággal. Például

$$-1 = i^2 = \sqrt{-1} \cdot \sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1$$

nyilvánvaló ellentmondás. A definíciók alapján a hiba csak a középen álló egyenlőségben lehet. Eszerint egy szorzat négyzetgyöke nem feltétlen egyezik meg a tényezők négyzetgyökének a szorzatával. Azt lehetne gondolni, hogy ez azért történt, mert „rosszul” értelmeztük a négyzetgyökfüggvényt. Be lehet látni azonban, hogy nem az értelmezéssel van baj; a kívánt tulajdonság általában nem teljesülhet a komplex számok körében.

3. A komplex számok geometriai bevezetése

Mint már említettük, a komplex számok sokáig idegenszerűnek tűntek. Éppen ezért bevezetésükre többféle utat próbáltak meg, hogy lehetővé tegyék minél természetesebb leírásukat. Több szempontból is hasznos, ha itt az előző — úgynevezett algebrai — bevezetéstől lényegesen különböző geometriai bevezetést is megtárgyaljuk. Ennek egyik haszna az, hogy megismerhetjük a komplex számok egy másfajta „arculatát”, másrészt az itteni eredmények segítséget nyújtanak a komplex számok további tárgyalásához is.

A valós számok „betöltik” a számegeyenest. Éppen ezért a komplex számoknak már csak ezen kívül van hely. Ez azt jelenti, hogy a komplex számokat a síkban vagy a térben ábrázolhatjuk. Ha megtartjuk a már szereplő valós számpár alakot, akkor célszerű a síkbeli ábrázolás. (Azt az algebrai struktúrák tárgyalásánál látni fogjuk, hogy térbeli ábrázolás „józan” feltételek mellett lehetetlen.)

1.8. Tétel. *A $z = a + bi$ komplex számnak megfelelően a sík (a, b) koordinátájú pontját, illetve a $(0, 0)$ pontból e pontba mutató \mathbf{z} vektort, mindkét esetben bijekciót kapunk.*

Bizonyítás. Mindkét állítás azonnal következik abból, hogy mind az $a + bi$ komplex szám, mind az (a, b) koordinátájú pont, mind a $(0, 0)$ pontból e pontba mutató vektor egyértelműen meghatározott az a és b valós számokkal; továbbá minden a és b valós szám meghatároz egy komplex számot, egy pontot, valamint az ebbe a pontba mutató vektort. ■

E tétel következményeként felváltva használhatjuk a komplex szám, a pont és a vektor elnevezést, mindig ugyanarra fogunk gondolni. Amennyiben ki akarjuk hangsúlyozni, hogy most a „vektoros szemléletet” helyezzük előtérbe, akkor — mint ahogy a vektorok jelölésénél általában teszik — vastagított betűket használunk. Így \mathbf{z} és \mathbf{z} ugyanazt a komplex számot jelöl(het)i, de ez az utóbbi jelölés hangsúlyozza azt, hogy most vektorként kezeljük.

Fontos megjegyezni a következőket: Egy vektor nem változik meg azzal, hogy párhuzamosan eltoljuk. Ennek megfelelően a $(0, 0)$ pontból az (a, b) pontba mutató vektor ugyanazt a komplex számot jeleníti meg, mint a (c, d) pontból az $(a + c, b + d)$ pontba mutató vektor.

A következőkben a komplex számokról felhasználjuk az ismert azonosságokat, az összeg és a szorzat kommutativitását és asszociativitását, valamint a disztributivitást. Célnk megvizsgálni, miképpen lehet a síkbeli ábrázolásnál a műveletek elvégzését leírni.

1.9. Tétel. *A komplex számok összeadása a vektorösszeadás. Két komplex szám szorzatának a hossza megegyezik a tényezők hosszának a szorzatával. Van egy olyan \mathbf{e} vektor, amelyre bármely \mathbf{z} komplex szám esetén $\mathbf{e}\mathbf{z} = \mathbf{z}$. Az \mathbf{e} -vel párhuzamos vektorok felelnek meg a valós számoknak. Egy valós számmal való szorzás minden vektort vele párhuzamos vektorba visz. Ha a valós szám pozitív, akkor a kapott vektor az eredetivel egyenlő állású is.*

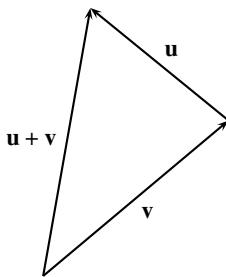
Bizonyítás. $(a+bi)+(c+di) = (a+c)+(b+d)i$ alapján tényleg a vektorösszeadás felel meg az összeadásnak. A Pitagorasz-tétel alapján egy \mathbf{z} vektor hossza éppen a megfelelő z komplex szám abszolút értéke; s az $|zw| = |z| \cdot |w|$ összefüggés bizonyítja a második állítást. Az \mathbf{e} vektor az, amelynek végpontja $(1, 0)$. Az a valós számnak megfelelő $\mathbf{a} = (a, 0)$ vektor valóban párhuzamos \mathbf{e} -vel; és más vektor nem párhuzamos vele. A $c(a, b) = (ca, cb)$ összefüggés bizonyítja az utolsó két állítást. ■

A komplex számok geometriai bevezetésénél az 1.9. Tételben megfogalmazottakon kívül elég azt feltenni, hogy a szorzás is kommutatív és asszociatív; továbbá az összeadásra nézve disztributív. (Az összeadás kommutativitása és asszociativitása, valamint a kivonás egyértelmű elvégezhetősége annak a következménye, hogy ez a vektorösszeadás.)

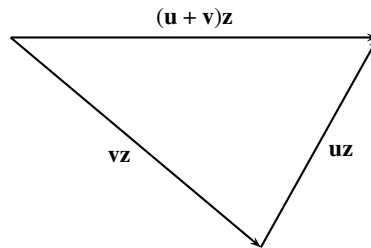
Az itt megfogalmazottakat nem fogjuk minden esetben hangsúlyozni, de a bizonyítások során ezeket bárki ellenőrizheti.

1.10. Tétel. *Egy rögzített $\mathbf{z} \neq 0$ komplex számmal való szorzásnál minden olyan háromszög, amelynek egyik csúcsa az origóban van, egy vele hasonló háromszögbe megy át, amelynek egyik csúcsa ugyancsak az origóban van. A hasonlóságnál mindegyik csúcsnak e csúcs \mathbf{z} -szerese felel meg.*

Bizonyítás. Legyen az origóból az eredeti háromszög másik két csúcsába mutató vektor \mathbf{w} és \mathbf{v} és legyen $\mathbf{u} = \mathbf{w} - \mathbf{v}$, azaz $\mathbf{w} = \mathbf{u} + \mathbf{v}$. A disztributivitás alapján a szorzás után kapott vektorokra $\mathbf{wz} = \mathbf{uz} + \mathbf{vz}$ teljesül. Az abszolút értékre vonatkozó feltétel alapján az új háromszög oldalai hosszának az eredeti háromszög megfelelő oldalai hosszához való aránya mindig $|z|$ lesz; tehát a mondott hasonlóság valóban fennáll. ■

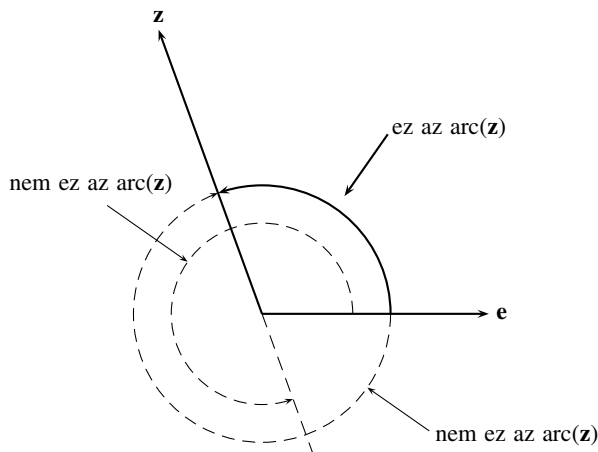


1. ábra



2. ábra

1.3. Definíció. Egy $\mathbf{z} \neq 0$ komplex szám $\text{arc}(\mathbf{z})$ arkuszán vagy irányyszögén azt a 360° -nál kisebb nemnegatív szöget értjük, amellyel az \mathbf{e} vektort pozitív (az óramutató járásával ellenkező) irányba elforgatva a \mathbf{z} -vel egyállású vektort kapunk (tehát \mathbf{z} -nek pozitív számszorosát). 0-nak *nincs* arkusza. ■



3. ábra

Megjegyezzük, hogy arkusz helyett egyes könyvekben használatos az argumentum elnevezés is.

1.11. Tétel. *Nemnulla komplex szám pozitív valós számszorosának az arkusza megegyezik az eredeti szám arkuszával. Minden nemnulla komplex szám egyértelműen felírható egy pozitív valós számnak és egy egységnyi abszolút értékű komplex számnak a szorzataként, amelynek az arkusza megegyezik az eredeti szám arkuszával.*

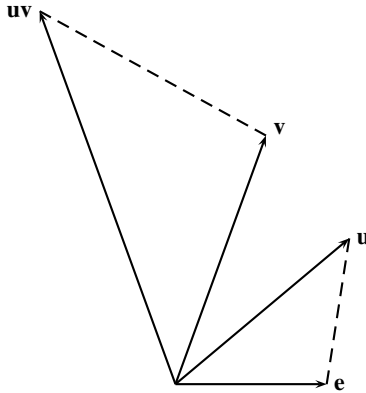
Bizonyítás. Az első állítás nem más, mint az 1.9. Tétel utolsó állításának átfogalmazása. Legyen $z \neq 0$ egy komplex szám. Ekkor $r = |z|$ pozitív valós szám; ennek s inverze tehát ugyancsak pozitív valós szám. Legyen $z_0 = sz$, amely az első állítás szerint z -vel egyirányú. Ennek abszolút értéke $|z_0| = s|z| = sr = 1$, és $z = rz_0$ biztosítja a felírhatóságot. Legyen most $z = tw$ a feltételeknek megfelelő felírás. Az abszolút érték egyértelműségének a következtében $t = r$. A fennálló $rw = rz_0$ felírást s -sel szorozva a $w = srw = srz_0 = z_0$ egyenlőséghez jutunk; ami pontosan az egyértelműséget jelenti. ■

1.12. Tétel. *Két nemnulla komplex szám szorzatának arkusza megegyezik a tényezők arkuszának (modulo 360° vett) összegével:*

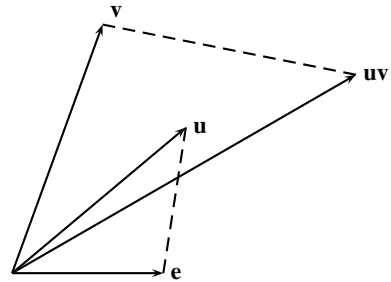
$$\text{arc}(zw) = \begin{cases} \text{arc}(z) + \text{arc}(w), & \text{ha } \text{arc}(z) + \text{arc}(w) < 360^\circ \\ \text{arc}(z) + \text{arc}(w) - 360^\circ, & \text{ha } \text{arc}(z) + \text{arc}(w) \geq 360^\circ \end{cases}$$

Bizonyítás. Legyen $r = |z|$ és $s = |w|$. Mint láttuk, vannak olyan egységnyi abszolút értékű u és v komplex számok, amelyekre $z = ru$ és $w = sv$; továbbá $\text{arc}(u)$ és $\text{arc}(z)$, valamint $\text{arc}(v)$ és $\text{arc}(w)$ megegyeznek. Ezen felül, a $zv = rsuv$ felírásból az is következik, hogy $\text{arc}(zv)$ és $\text{arc}(uv)$ is egyenlőek. Ezért elegendő a tételt olyan u és v komplex számokra bizonyítani, amelyek abszolút értéke 1.

Legyen \mathbf{u} arkusza φ és \mathbf{v} arkusza ψ . Az \mathbf{e} és \mathbf{u} meghatározta háromszöget \mathbf{v} -vel szorozva a \mathbf{v} és \mathbf{uv} által meghatározott háromszöget nyerjük. Mivel ez az előbbihez hasonló, ezért a \mathbf{v} és \mathbf{uv} vektorok bezárta szög ugyancsak φ . Azt azonban egyelőre nem tudjuk, hogy a szorzással kapott háromszög körüljárási iránya megegyezik-e az eredeti háromszögével, vagy éppen ellenkező. Más szóval nem tudjuk, hogy az \mathbf{uv} arkusza (tehát \mathbf{e} -vel bezárt szöge) $\psi + \varphi$ avagy $\psi - \varphi$. A kommutativitás miatt \mathbf{u} és \mathbf{v} szerepe felcserélhető, azaz ugyanígy azt is kapjuk, hogy ez a szög vagy $\varphi + \psi$, vagy $\varphi - \psi$.



4. ábra

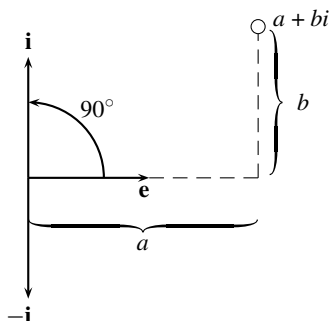


5. ábra

Ha valamelyik esetben az összeget kapjuk, akkor természetesen ez adódik a másik esetben is (nincs kizárva, hogy ez megegyezzen a különbséggel). Az az eset marad, amikor mindkét esetben a különbséget kapjuk. Mivel a két szög ekkor is egyenlő, ezért különbségük $2(\varphi - \psi) = (\varphi - \psi) - (\psi - \varphi)$ egész számú többszöröse 0° -nak; tehát $\varphi - \psi$ egész számú többszöröse 180° -nak. Feltehető tehát, hogy ez a különbség vagy 180° , vagy 360° . Az első esetben $\mathbf{uv} = -\mathbf{e}$, a második esetben $\mathbf{uv} = \mathbf{e}$. Az is következik, hogy az első esetben $\mathbf{v} = -\mathbf{u}$ és a második esetben $\mathbf{v} = \mathbf{u}$. Ezekből mindkét esetben a $\mathbf{u}^2 = \mathbf{e}$ egyenlőség következik. Mivel $\mathbf{e} = \mathbf{e}^2$, ezért a műveleti azonosságokat felhasználva azt kapjuk, hogy $(\mathbf{u} + \mathbf{e})(\mathbf{u} - \mathbf{e}) = 0$. Abszolút értékekre térve az adódik, hogy a két tényező valamelyikének az abszolút értéke — és így maga ez a tényező 0. Eszerint vagy $\mathbf{u} = -\mathbf{e}$ vagy $\mathbf{u} = \mathbf{e}$. Eszerint $\varphi = 180^\circ$ vagy $\varphi = 0^\circ$; és hasonló értékek adódnak ψ -re is. Ezekben az esetekben viszont a két szög különbsége és összege megegyezik, ezért a szorzat arkusza megegyezik az arkuszok összegével. ■

Az 1.12. Tételből azonnal következik, hogy az \mathbf{e} -re merőleges egységnyi hosszúságú \mathbf{i} komplex szám négyzete $-\mathbf{e}$, tehát ez az i komplex szám. Persze $(-\mathbf{i})^2 = -\mathbf{e}$ is igaz, bármelyiküket tekinthetjük i -nek. Általában azt tekintjük i -nek, amelyiknek az arkusza 90° .

Vegyük figyelembe, hogy ezeket a — már ismert — eredményeket tisztán az 1.9. Tételben szereplő és e tétel bizonyítása után közvetlenül említett műveleti azonosságok



6. ábra

alapján nyertük. Ebből a bevezetésből kiindulva az $a + bi$ komplex számot úgy nyerhetjük, hogy ezt feleltetjük meg az $\mathbf{a} + \mathbf{b} \cdot \mathbf{i}$ vektornak.

4. A komplex számok trigonometrikus alakja

A komplex számok bevezetésénél négy jellemző (valós szám) adattal találkoztunk. Ezek a komplex szám valós és képzetes része, abszolút értéke és arkusza. (Ez utóbbi csak akkor, ha a komplex szám nem 0.) Ezek az adatok nem függetlenek; közöttük az alábbi összefüggések állnak fenn:

1.13. Tétel. Legyen a z komplex számra $\Re(z) = a$, $\Im(z) = b$, $|z| = r$ és $z \neq 0$ esetén $\arg(z) = \varphi$. Ekkor

$$(i) \quad r = \sqrt{a^2 + b^2}, \quad \text{továbbá} \quad \cos \varphi = \frac{a}{r} \quad \text{és} \quad \sin \varphi = \frac{b}{r}$$

$$(ii) \quad a = r \cdot \cos \varphi \quad \text{és} \quad b = r \cdot \sin \varphi.$$

Bizonyítás. Az első egyenlőség azonnal következik a Pitagorasz-tételből, a többiek pedig a szögfüggvények általános értelmezéséből. ■

A (ii) alatti összefüggés azt adja meg, miképpen kaphatjuk meg a valós és a képzetes részt az abszolút érték és az arkusz ismeretében. Az (i) alatti összefüggésben az első egyenlőség azt adja meg, miképpen kaphatjuk meg az abszolút értéket a valós és a képzetes rész ismeretében; majd ezt is felhasználva miképpen kapható meg a komplex szám arkusza. Azért szerepel két szögfüggvény is, mert az arkuszt egyikük sem határozza meg egyértelműen. Ha viszont az egyik szögfüggvényt ismerjük, akkor az arkusz meghatározásához elég már a másik szögfüggvény előjelét tudni. A következőben lehetőséget mutatunk az arkusz egyértelmű kiszámíthatóságára:

1.14. Tétel. *A $z = a + bi \neq 0$ komplex szám arkusza egyértelműen meghatározott a $\varphi = 2 \cdot \arctan \frac{b}{r+a} = 2 \cdot \arctan \frac{r-a}{b}$ összefüggéssel (a $b = 0$ esetben a második összefüggés nem használható).*

Bizonyítás. Az $r^2 - a^2 = b^2$ összefüggésből következik, hogy a két tört megegyezik, kivéve a $b = 0$ és az $r = -a$ esetet. Így $b = 0$ esetén a második összefüggés nem használható, és $r = -a$ esetén az első sem. Ekkor azonban a tangensfüggvény értelmezéséből azonnal következik, hogy $\varphi = 180^\circ$. A továbbiakban legyen $b \neq 0$, amiből $r \neq -a$ is következik. Mint tudjuk, létezik olyan $w = x + yi$ komplex szám, amelyre $z = w^2$. Ekkor az 1.12. Tétel alapján $2 \cdot \arcc(w) = \arcc(z)$, azaz vagy $\arcc(w) = \frac{\varphi}{2}$, vagy $\arcc(w) = \frac{\varphi}{2} + 180^\circ$. Tekintettel arra, hogy az arkusz 360° -nál kisebb és a tangensfüggvény periódusa 180° , ezért mindkét esetben igaz a $\varphi = 2 \cdot \arctan(\operatorname{tg}(\arcc(w)))$ összefüggés. Használjuk fel a komplex szám négyzetgyökének meghatározásánál kapott egyenlőségeket (itt is ugyanazok a betűk szerepelnek): $x^2 + y^2 = r$, $2x^2 = a + r$ és $2xy = b$. Ebből $\operatorname{tg}(\arcc(w)) = \frac{y}{x} = \frac{2xy}{2x^2} = \frac{b}{r+a}$, mint állítottuk. ■

Az alábbiakban csak 0-tól különböző komplex számokkal foglalkozunk.

1.15. Tétel. *Minden $z \neq 0$ komplex szám egyértelműen felírható*

$$z = r \cdot (\cos \varphi + i \cdot \sin \varphi)$$

alakban, ahol r pozitív valós szám és $\varphi < 360^\circ$ nemnegatív szög.

E felírásban $r = |z|$ és $\varphi = \arcc(z)$.

Bizonyítás. Az 1.13. Tétel alapján a fenti választással valóban teljesül a

$$z = a + bi = r \cdot \cos \varphi + (r \cdot \sin \varphi)i = r \cdot (\cos \varphi + i \cdot \sin \varphi)$$

összefüggés. Tegyük fel, hogy valamilyen r és φ választással igaz a kívánt összefüggés. Ebből azonnal következnek az $a = \Re(z) = r \cdot \cos \varphi$ és a $b = \Im(z) = r \cdot \sin \varphi$ egyenlőségek.

Az 1.13. Tételben mondottak szerint tehát $|z| = \sqrt{a^2 + b^2}$; továbbá $\cos(\arcc(z)) = \cos \varphi$ és $\sin(\arcc(z)) = \sin \varphi$, a φ -re vonatkozó korlátozás szerint tehát $\arcc(z) = \varphi$. ■

1.4. Definíció. A $z \neq 0$ komplex számnak pozitív r -rel felírt

$$z = r \cdot (\cos \varphi + i \cdot \sin \varphi)$$

alakját a komplex szám trigonometrikus alakjának nevezzük. ■

Megjegyezzük, hogy a trigonometrikus alak annak ellenére egyértelmű, hogy a szögre nem tettünk semmiféle megkötést. Viszont a szög természetesen nem egyértelmű. A trigonometrikus függvények elemi tulajdonságaiból azonnal következik, hogy φ -nek pontosan azok a szögek felelnek meg, amelyek z arkuszától 360° egész számú többszörösében térnek el.

1.16. Tétel (MOIVRE). Legyen $z \neq 0$ és $w \neq 0$ komplex számok trigonometrikus alakja:

$$z = r \cdot (\cos \varphi + i \cdot \sin \varphi) \quad \text{és} \quad w = s \cdot (\cos \psi + i \cdot \sin \psi).$$

Ekkor zw , illetve z^n ($n \in \mathbb{N}$) trigonometrikus alakja:

$$zw = rs(\cos(\varphi + \psi) + i \cdot \sin(\varphi + \psi)), \quad \text{illetve} \quad z^n = r^n(\cos(n\varphi) + i \cdot \sin(n\varphi)).$$

Bizonyítás. Az 1.6. Tétel szerint a szorzat abszolút értéke megegyezik a tényezők abszolút értékeinek a szorzatával. Az 1.12. Tétel szerint a szorzat arkusza egyenlő a tényezők arkuszainak az összegével. A trigonometrikus alak egyértelműségéből tehát következik a tétel első egyenlősége. A $w = z$ választással kapjuk a második egyenlőséget az $n = 2$ esetben. Ebből teljes indukcióval következik a második egyenlőség minden n természetes számra. ■

Következmény. Bármely φ szögre $\cos^2 \varphi + \sin^2 \varphi = 1$. Tetszőleges φ és ψ szögekre érvényes:

$$\cos(\varphi + \psi) = \cos \varphi \cdot \cos \psi - \sin \varphi \cdot \sin \psi, \quad \sin(\varphi + \psi) = \sin \varphi \cdot \cos \psi + \cos \varphi \cdot \sin \psi.$$

Tetszőleges φ szögre érvényesek az alábbiak:

$$\begin{aligned} \cos(2\varphi) &= \cos^2 \varphi - \sin^2 \varphi \quad \text{és} \quad \sin(2\varphi) = 2 \cdot \sin \varphi \cdot \cos \varphi, \\ \cos(3\varphi) &= \cos^3 \varphi - 3 \cdot \cos \varphi \cdot \sin^2 \varphi = 4 \cdot \cos^3 \varphi - 3 \cdot \cos \varphi, \\ \sin(3\varphi) &= 3 \cdot \cos^2 \varphi \cdot \sin \varphi - \sin^3 \varphi = 3 \cdot \sin \varphi - 4 \cdot \sin^3 \varphi. \end{aligned}$$

Bizonyítás. Az első összefüggés azonnal adódik abból, hogy a $z = \cos \varphi + i \cdot \sin \varphi$ komplex szám abszolút értéke 1.

A Moivre tételében szereplő két komplex számot válasszuk úgy, hogy $r = s = 1$ legyen. Ekkor a komplex számok szorzatát algebrai alakban kiszámolva:

$$\begin{aligned} zw &= (\cos \varphi + i \cdot \sin \varphi)(\cos \psi + i \cdot \sin \psi) = \\ &= (\cos \varphi \cdot \cos \psi - \sin \varphi \cdot \sin \psi) + (\sin \varphi \cdot \cos \psi + \cos \varphi \cdot \sin \psi) \cdot i \end{aligned}$$

adódik. Moivre tétele szerint viszont a szorzat: $zw = \cos(\varphi + \psi) + i \cdot \sin(\varphi + \psi)$. Az algebrai alak egyértelműségéből kapjuk az első két összefüggést.

Az összeg hatványai alapján kiszámolva z^2 -et és z^3 -t kapjuk, hogy:

$$\begin{aligned} (\cos \varphi + i \cdot \sin \varphi)^2 &= (\cos^2 \varphi - \sin^2 \varphi) + i \cdot (2 \sin \varphi \cos \varphi), \\ (\cos \varphi + i \cdot \sin \varphi)^3 &= (\cos^3 \varphi - 3 \cos \varphi \sin^2 \varphi) + i \cdot (3 \cos^2 \varphi \sin \varphi - \sin^3 \varphi). \end{aligned}$$

Ezeket hasonlítsuk össze a Moivre-tételben kapott összefüggésekkel. Az összehasonlításból a komplex számok egyenlőségének a definícióját figyelembe véve azonnal következik a kétszeres szögekre vonatkozó összefüggés és a háromszoros szögekre vonatkozó első két összefüggés is. Ez utóbbiakat tovább alakíthatjuk a $\cos^2 \varphi + \sin^2 \varphi = 1$ összefüggés felhasználásával; így adódnak a háromszoros szögekre vonatkozó második összefüggések. ■

Megjegyzés. Az eljárást tovább folytatva hasonló összefüggéseket kapunk a többszörös szögekre.

Noha kerülő úton, de ténylegesen bizonyítottuk a szögösszegekre és a többszörös szögekre vonatkozó összefüggéseket; ugyanis ezeket előzetesen még burkoltan sem használtuk fel. Csupán a forgásszögek trigonometrikus függvényeinek az elemi tulajdonságaira támaszkodtunk. Ennek a bizonyításnak az egyik „szépsége” az, hogy nem kell eseteket megkülönböztetni aszerint, hogy a szögek, és azok összege kisebb-e vagy nagyobb-e 90° -nál, illetve 180° -nál. A trigonometrikus kifejezéseknek a komplex számokra való „átfordítási” lehetőségét a komplex számok geometriai tárgyalása tette lehetővé. A fenti trigonometrikus összefüggések már az 1.12. Tételben „el voltak bújtatva”. A másik „szépség” az, hogy ennél a bizonyításnál világosan látszik, miért ilyen alakúak az összefüggések. \square

A komplex számoknak tetszőleges (akár komplex kitevőjű) hatványa is értelmezhető. Nekünk itt csak egész kitevőjű hatványokra (és később gyökvonásra) lesz szükségünk.

1.5. Definíció. Tetszőleges $z \neq 0$ komplex számra $z^0 = 1$, és bármely n természetes szám esetén $z^{-n} = \frac{1}{z^n}$. \blacksquare

1.17. Tétel. Legyenek z és w , mint az 1.16. Tételben. Ekkor hányadosuk trigonometrikus alakja

$$\frac{z}{w} = \frac{r}{s}(\cos(\varphi - \psi) + i \cdot \sin(\varphi - \psi));$$

és tetszőleges n egész szám esetén z^n trigonometrikus alakja

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

Bizonyítás. Legyen a $\frac{z}{w} = u$ komplex szám trigonometrikus alakja: $t(\cos \chi + i \cdot \sin \chi)$. Moivre tétele szerint $z = wu$ trigonometrikus alakja $st(\cos(\psi + \chi) + i \cdot \sin(\psi + \chi))$. A trigonometrikus alak egyértelműsége alapján ez azt jelenti, hogy $st = r$ és az arkuszok különbsége: $\varphi - (\psi + \chi)$ a 360° egész számú többszöröse. Ebből viszont az következik, hogy $t = \frac{r}{s}$, és az, hogy $(\varphi - \psi) - \chi$ egész számú többszöröse 360° -nak.

A hatványozásra vonatkozó azonosságot bizonyítottuk pozitív kitevőre. Ez az azonosság triviálisan igaz az $n = 0$ esetben, mert 1 abszolút értéke 1 és arkusza 0° . Tekintsünk most egy negatív kitevőt, amelyet $0 - n$ alakba írhatunk, ahol n természetes szám. Ezt a felírást és a hányados trigonometrikus alakban való előállítását felhasználva kapjuk, hogy:

$$z^{-n} = \frac{1}{r^n}(\cos(0^\circ - n\varphi) + i \cdot \sin(0^\circ - n\varphi)) = r^{-n} \cos((-n)\varphi) + i \cdot \sin((-n)\varphi). \quad \blacksquare$$

Végezetül a gyökvonással és ennek egy fontos speciális esetével foglalkozunk.

1.18. Tétel. Legyen a $z \neq 0$ komplex szám trigonometrikus alakja

$$z = r(\cos \varphi + i \cdot \sin \varphi).$$

Ekkor a $w^n = z$ összefüggést pontosan azok a komplex számok elégítik ki, amelyeknek a trigonometrikus alakja

$$w = s(\cos \psi + i \cdot \sin \psi),$$

ahol $s = \sqrt[n]{r}$ és $\psi = \frac{\varphi}{n} + k \cdot \frac{360^\circ}{n}$, valamilyen k egész számmal.

A kapott w komplex számok között pontosan n darab különböző van. Ezek mindegyike megkapható például úgy, hogy k értékének rendre a $0, 1, \dots, n-1$ egész számokat választjuk.

Bizonyítás. Mindenekelőtt megmutatjuk, hogy csak a fenti w komplex számok jöhetnek szóba. Legyen w egy, a kívánalmaknak eleget tevő komplex szám. Tekintettel arra, hogy $z \neq 0$, ezért w is különbözik 0-tól; tehát létezik trigonometrikus alakja: írjuk fel ezt a tételben megadott jelölésekkel. Moivre tétele alapján a hatványozást elvégezve a következő összefüggéshez jutunk:

$$s^n = r \quad \text{és} \quad n\psi - \varphi = k \cdot 360^\circ, \quad \text{ahol } k \text{ tetszőleges egész.}$$

Mivel s csak pozitív lehet, ezért egyértelműen meghatározott: $s = \sqrt[n]{r}$. A szögekre most kapott összefüggésből n -nel való osztás után valóban a tételben megfogalmazott összefüggés adódik. Fordítva: a kiszámított s abszolút értékkel és ψ arkusszal trigonometrikus alakban felírt bármelyik w komplex szám n -edik hatványa nyilvánvalóan megegyezik z trigonometrikus alakjával.

Most még azt kell belátni, hogy $w^n = z$ tulajdonságú komplex szám valóban n darab van; s ezek mindegyike megegyezik a felsoroltak valamelyikével. A $w^n = z$ tulajdonságú komplex számok mindegyikének s az abszolút értéke, ezért csak azt kell megnézni, mikor egyenlő, illetve mikor különböző az arkuszuk. Legyenek $\psi_k = \frac{\varphi}{n} + k \cdot \frac{360^\circ}{n}$ a fellépő arkuszok. Két ilyennek a különbsége $\psi_k - \psi_j = (k - j) \cdot \frac{360^\circ}{n}$. Ez pedig pontosan akkor egész számú többszöröse 360° -nak, ha $k - j$ egész számú többszöröse n -nek. Világos, hogy a $k = 0, 1, \dots, n-1$ számok közül egyetlen pár különbsége sem osztható n -nel. Másrészt viszont tetszőleges k számnak az n -nel való osztási maradéka a fentiek között van, és így a különbségük n -nel osztható. ■

A gyökvonásnak most a $z = 1$ speciális esetét fogjuk vizsgálni:

1.6. Definíció. Azokat a w komplex számokat, amelyekre egy adott n természetes szám mellett $w^n = 1$, n -edik komplex egységgyököknek nevezzük. ■

1.19. Tétel. Az n -edik komplex egységgyököket a következő alakban adhatjuk meg:

$$\varepsilon_k = \cos\left(k \cdot \frac{360^\circ}{n}\right) + i \cdot \sin\left(k \cdot \frac{360^\circ}{n}\right),$$

ahol $k = 0, 1, \dots, n-1$. Emellett ugyanezen k értékekre érvényes az $\varepsilon_k = (\varepsilon_1)^k$ összefüggés.

Bizonyítás. A felírás azonnal következik az előző tételből, tekintettel arra, hogy 1 trigonometrikus alakja $1 \cdot (\cos 0^\circ + i \cdot \sin 0^\circ)$. Az $\varepsilon_k = (\varepsilon_1)^k$ egyenlőség ebből a Moivre-tétel alkalmazásával adódik. ■

1.20. Tétel. Az ε_k n -edik egységgyökre az alábbi négy állítás ekvivalens:

- (1) ε_k természetes kitevőjű hatványaként minden n -edik egységgyök előáll.
- (2) k az n -hez relatív prím.
- (3) ε_k -nak n -nél alacsonyabb természetes kitevőjű hatványa nem 1.
- (4) ε_k -nak n darab különböző természetes kitevőjű hatványa van.

Bizonyítás. Mindenekelőtt gondoljuk meg a bizonyítás menetét. Mivel azt kell belátni, hogy a négy állítás bármelyikéből következik a másik három, ez összesen $4 \times 3 = 12$ bizonyítást jelentene. Ezek közül néhányat el lehet hagyni. Azt fogjuk bebizonyítani ugyanis, hogy mindegyik állításból következik az utána szereplő és a legutolsóból az első. Ez a fenti tizenkét apróbb állítás közül csupán négynek a bizonyítását jelenti. Mégis, ezáltal mindegyik állítás bizonyítást nyer. Hiszen bármelyik állításból kiindulva bizonyított a következő, majd az utána következő és így tovább. Ha az utolsóhoz elérünk, akkor ebből ismét az első igazságára következtethetünk. Bármelyikből indultunk is ki, ebből legfeljebb három lépés után bármelyik másik igazolást nyer. Ezt a bizonyítási „eljárás”-t *ciklikus bizonyításnak* nevezik.

Lássuk tehát a tétel ciklikus bizonyítását:

(1)-ből következik (2): Ha ε_k természetes kitevőjű hatványaként minden n -edik egységgyök előáll, akkor előáll ε_1 is, tehát van olyan j természetes szám, amire $\varepsilon_k^j = \varepsilon_1$. A komplex számok trigonometrikus alakját és a Moivre-tételt felhasználva ez azt jelenti, hogy $kj \cdot \frac{360^\circ}{n}$ és $\frac{360^\circ}{n}$ egymástól 360° egész számú többszörösében térnek el. Ez azt jelenti, hogy $\frac{kj-1}{n} = \frac{kj}{n} - \frac{1}{n}$ egész szám, azaz $kj-1$ egész számú többszöröse n -nek. Létezik tehát olyan ℓ egész szám, amelyre $kj-1 = n\ell$, azaz $kj - n\ell = 1$. Ez pedig csak úgy lehet, ha k és n relatív prímek, hiszen bármely közös osztójuk 1-nek is osztója; és 1 egyetlen prímszámmal sem osztható.

(2)-ből következik (3): Feltételünk szerint k az n -hez relatív prím. Definíció szerint

$$\varepsilon_k = \cos\left(k \cdot \frac{360^\circ}{n}\right) + i \cdot \sin\left(k \cdot \frac{360^\circ}{n}\right).$$

Moivre tétele alapján tetszőleges j természetes számra

$$(\varepsilon_k)^j = \cos\left(kj \cdot \frac{360^\circ}{n}\right) + i \cdot \sin\left(kj \cdot \frac{360^\circ}{n}\right)$$

adódik. Azt kell belátnunk, hogy $0 < j < n$ esetén ez nem lehet 1. Ez pontosan azt jelenti, hogy $kj \cdot \frac{360^\circ}{n}$ nem lehet egész számú többszöröse 360° -nak, vagyis $\frac{kj}{n}$ nem lehet egész szám; amit úgy fogalmazhatunk, hogy n nem lehet osztója kj -nek. Tekintettel arra, hogy k és n relatív prímek, az egyértelmű prímtenyezős felbontás szerint az oszthatóság csak akkor állhat fenn, ha n a j -nek osztója. Ez viszont a $0 < j < n$ feltétel miatt lehetetlen.

(3)-ból következik (4): Feltétel szerint $(\varepsilon_k)^j \neq 1$, ha $0 < j < n$. Tekintsük ε_k -nak két különböző kitevőjű hatványát és nézzük meg, mikor lehetnek ezek egyenlőek. Az $u < v$

természetes számokra $(\varepsilon_k)^u = (\varepsilon_k)^v$ akkor teljesül, ha a $v - u$ természetes számra $(\varepsilon_k)^{v-u} = 1$. Amennyiben $v - u < n$, akkor ez feltétel szerint lehetetlen. Így az $(\varepsilon_k)^j$ számok mindegyike különböző, ha $0 \leq j < n$. Eszerint találtunk n darab különböző hatványt. Tekintettel arra, hogy ezek mindegyike n -edik egységgyök, amelyeknek a száma összesen n , ezért több különböző hatvány nem létezik.

(4)-ből következik (1): Mivel ε_k minden hatványa n -edik egységgyök és e hatványok száma feltételünk szerint n , ezért hatványai között minden n -edik egységgyök ott van, mert n -edik egységgyök is pontosan n van. ■

1.7. Definíció. Azokat az n -edik egységgyököket, amelyekre az 1.20. Tétel valamelyik feltétele teljesül, primitív n -edik egységgyököknek nevezzük.

1.21. Tétel. Minden ε n -edik egységgyökhöz létezik n -nek pontosan egy olyan k osztója, amelyre ε k -adik primitív egységgyök. A primitív n -edik egységgyökök száma $\varphi(n)$ (φ az Euler-féle φ -függvényt jelöli, azaz $\varphi(n)$ az n -nél nem nagyobb n -hez relatív prím természetes számok száma).

Bizonyítás. Ha ε egy n -edik egységgyök, akkor tekintsük azokat a j természetes számokat, amelyekre $\varepsilon^j = 1$. Ilyen biztosan van, hiszen $n = j$ megfelel. Mivel a természetes számok minden halmazában van legkisebb szám, ezért ezek között is található egy ilyen; jelölje ezt k . A k minimalitása következtében $\varepsilon^j \neq 1$, ha $0 < j < k$; az 1.20. Tétel (3) pontja szerint tehát ε primitív k -adik egységgyök.

A maradékos osztás alapján az n -hez és a pozitív k -hoz található olyan q és r egész szám, amelyre $n = kq + r$ és $|r| < k$. A hatványozás azonosságait felhasználva ebből a következőket kapjuk:

$$1 = \varepsilon^n = \varepsilon^{kq+r} = (\varepsilon^k)^q \cdot \varepsilon^r = 1^q \cdot \varepsilon^r = 1 \cdot \varepsilon^r = \varepsilon^r.$$

Mivel 1-nek a reciproka is 1, ezért $\varepsilon^{|r|} = 1$. Ebből k minimalitása és $|r| < k$ folytán $|r| = 0$ és így $r = 0$ következik. Ez pedig azt jelenti, hogy $n = kr$, azaz k valóban osztója n -nek. A k egyértelműsége ugyancsak az 1.20. Tétel (3) pontjából következik.

Végezetül, az 1.20. Tétel (2) pontja szerint a primitív n -edik egységgyökök száma megegyezik az n -nél nem nagyobb, n -hez relatív prím természetes számok számával. Már pedig ez éppen az Euler-függvény definíciója. ■

Az 1.21. Tétel mutatja meg, hogy miért volt szükség az 1.20. Tételben felsorolt állítások ekvivalenciájára; ugyanis itt a bizonyítás során hol az egyik, hol a másik tulajdonságot használtuk.

Feladatok

1. Gyűrűt, illetve testet alkotnak-e a következő számhalmazok a komplex számok összeadására és szorzására: a) az $(a, 0)$ alakúak, b) a $(0, b)$ alakúak, c) azok az (a, b) alakúak, amelyekre: a, b egészek; a, b racionálisak; a racionális és b egy racionális szám $\sqrt{3}$ -szorosa (vagy $\sqrt{5}$ -szöröse stb.), d) a egész és b páros, illetve egy egész szám $\sqrt{3}$ -szorosa (vagy $\sqrt{5}$ -szöröse stb.).

2. Legyen N rögzített egész, a és b pedig tetszőleges egész számok. Milyen p természetes szám esetében alkotnak gyűrűt az $\left(\frac{a}{p}, \frac{b}{p} \cdot \sqrt{N}\right)$ alakú (komplex) számok?

3. A nemnulla komplex számok alábbi részhalmazai közül melyek alkotnak félcsoportot a szorzásra:

a) az összes; b) azok, amelyeknek az abszolút értéke 1; c) azok, amelyeknek az abszolút értéke nem 1; d) azok, amelyeknek az abszolút értéke nagyobb, mint egy rögzített $r > 0$ valós szám, ahol $r > 1$; $r = 1$; $r < 1$, e) azok, amelyeknek az abszolút értéke kisebb, mint egy rögzített $r > 0$ valós szám, ahol $r > 1$; $r = 1$; $r < 1$.

4. Bizonyítsuk be, hogy az 1 abszolút értékű komplex számok (illetve az egységgyökök) a szorzásra nézve egy olyan csoportot alkotnak, amely izomorf a valós (illetve a racionális) számok additív csoportjával modulo 1.

5. Bizonyítsuk be, hogy bármely c valós számra $S(c \cdot z) = c \cdot S(z)$ ($S(z)$ a z nyoma); határozzuk meg az összes olyan $F: \mathbb{C} \rightarrow \mathbb{R}$ függvényt, amelyre $F(z)$ *additív* (azaz $F(z + w) = F(z) + F(w)$) és $F(c \cdot z) = c \cdot F(z)$.

6. Bizonyítsuk be, hogy ha két egész szám mindegyike előáll két egész szám négyzetösszegeként, akkor ugyanez igaz szorzatukra is; míg hányadosuk (ha az osztó nem nulla) két racionális szám négyzetösszege lesz.

7. Legyen $a + bi = (x + yi)^2$. Bizonyítsuk be, hogy x és y előjele a következőképpen függ b -től: Ha $b = 0$, akkor x és y valamelyike is 0, ha $b > 0$, akkor x és y megegyező előjelűek, ha $b < 0$, akkor x és y különböző előjelűek.

8. Határozzuk meg $2i$ négyzetgyökét.

9. Bizonyítsuk be, hogy $\Re(z)$ és $\Im(z)$ additív függvények.

10. Fejezzük ki $\Re(zw)$ -t és $\Im(zw)$ -t $\Re(z)$, $\Im(z)$, $\Re(w)$ és $\Im(w)$ segítségével.

11. Határozzuk meg $\Re(i \cdot z)$ -t és $\Im(i \cdot z)$ -t $\Re(z)$ és $\Im(z)$ segítségével.

12. Fejezzük ki $\Re(z)$ -t és $\Im(z)$ -t $S(z)$ és a műveletek segítségével.

13. A sík milyen transzformációját adja meg a $z \mapsto z + a$ függvény, ahol a tetszőleges adott komplex szám?

14. A sík milyen transzformációját adja meg a $z \mapsto a \cdot z$ függvény, ahol a tetszőleges adott komplex szám; ha $a = 0$, ha $|a| = 1$, ha $a > 0$ valós?

15. Milyen síktranszformációt hoz létre a $z \mapsto \bar{z}$ függvény?

16. Írjuk le az összes olyan síktranszformációt, amely előállítható $z \mapsto a \cdot z + b$ úgynevezett *lineáris függvényként*, ahol $a \neq 0$ és a, b tetszőlegesen adott komplex számok.

17. Tetszőlegesen adott a, b komplex számok ($a \neq 0$) esetén határozzuk meg, milyen transzformációt írnak le a $z \mapsto a \cdot \bar{z} + b$ úgynevezett *másodfajú lineáris függvények*.

Megjegyezzük, hogy a lineáris és másodfajú lineáris függvények esetében, ha $a \neq 0$, akkor a függvényt *regulárisnak* nevezik, míg az $a = 0$ esetben *szinguláris* függvényről beszélünk. Mi a továbbiakban csak a reguláris esetet nézzük, éppen ezért nem tesszük ki a „reguláris” jelzőt.

18. Bizonyítsuk be, hogy a lineáris függvények a kompozícióra nézve csoportot alkotnak, de a kompozíció művelete nem kommutatív (azaz léteznek olyan $f, g : \mathbb{C} \rightarrow \mathbb{C}$ lineáris függvények, amelyekre $f \circ g \neq g \circ f$). (Tetszőleges H nem üres halmaz egy adott \circ műveletre csoport, ha ez a művelet asszociatív, és bármely $a, b \in H$ elemhez létezik olyan [egyértelmű] $u, v \in H$, amelyekre $a \circ u = b$ és $v \circ a = b$.)

19. Bizonyítsuk be, hogy a lineáris és a másodfajú függvények a kompozícióra nézve csoportot alkotnak. Igaz-e ez akkor is, ha csak a másodfajúakat tekintjük?

20. Mi a feltétele annak, hogy két komplex szám — mint vektor — egymásra merőleges legyen? (Többféleképpen is megfogalmazható.)

21. Legyen a egy origón át nem menő egyenesnek az origóhoz legközelebbi pontja. Bizonyítsuk be, hogy ezen az egyenesen pontosan a $z = a(1 + \mu \cdot i)$ alakú számok vannak rajta, ahol $\mu \in \mathbb{R}$. Hogyan adhatók meg egy origón átmenő egyenes pontjai?

22. Adott $a \neq b$ komplex számokra hol helyezkednek el a síknak azok a z pontjai, amelyekre $z - b = (z - a)\mu \cdot i$, ahol $\mu \in \mathbb{R}$?

23. Legyen $a \neq 0$ adott komplex és $\lambda \neq 0$ valós szám, továbbá $\mu \in \mathbb{R}$. Hol helyezkednek el a síkon a $z = \frac{a \cdot (1 - \lambda \mu \cdot i)}{(1 - \mu \cdot i)}$ alakú számok? Mi a geometriai értelme annak, hogy $\lambda > 0$, illetve $\lambda < 0$?

24. Milyen geometriai transzformációt hoz létre az az *inverzió*nak nevezett függvény, amely minden $z \neq 0$ komplex számhoz konjugáltjának reciprokát, azaz $\frac{1}{\bar{z}}$ -t rendeli hozzá? Mutassuk meg, hogy az inverzió egy involúció.

25. Mibe visz át az inverzió egy origón átmenő egyenest?

26. Mibe visz át az inverzió egy az origón át nem menő egyenest? (Használjuk a 21. és 23. feladatok eredményét.)

27. Mibe visz át az inverzió egy origón átmenő, illetve egy origón át nem menő kört?

28. Lineáris törtfüggvénynek nevezzük a $z \mapsto \frac{az + b}{cz + d}$ alakú függvényeket, ahol a, b, c, d tetszőleges, az $ad \neq bc$ feltételnek eleget tevő komplex számok. Határozzuk meg a lineáris törtfüggvények által megadott geometriai transzformációk tulajdonságait. Állítsuk elő a lineáris törtfüggvényeket egyszerűbb függvények kompozíciójaként. (Miért van szükség az $ad \neq bc$ feltételre?)

29. Bizonyítsuk be, hogy a lineáris törtfüggvények a kompozícióra nézve csoportot alkotnak.

30. Csoportot alkotnak-e a $z \mapsto \frac{a\bar{z} + b}{c\bar{z} + d}$ alakú úgynevezett *másodfajú lineáris törtfüggvények* a kompozícióra, ahol a, b, c, d tetszőleges, az $ad \neq bc$ feltételnek eleget tevő komplex számok. Mivel kell még e halmazt kiegészíteni, hogy csoportot kapjunk?

31. Adott a komplex síkon az \mathbf{e} egységvektor és a pozitív forgásirány. Miképpen változik meg az $f(\mathbf{z})$ függvény geometriai képe, ha helyette az $f(\mathbf{z} + \mathbf{a}) - \mathbf{a}$ függvényt tekintjük, ahol \mathbf{a} rögzített komplex szám. Milyen változást jelent az, ha az eredeti függvénynek az $\frac{f(\mathbf{a} \cdot \mathbf{z})}{\mathbf{a}}$ függvényt feleltetjük meg, abban az esetben, ha $|\mathbf{a}| = 1$, illetve ha \mathbf{a} pozitív valós szám?

32. Milyen változást jelent az előző feladatbeli függvény geometriai képében, ha az $f(\mathbf{z}) \mapsto \overline{f(\overline{\mathbf{z}})}$ megfeleltetést tekintjük?

33. Állítsuk elő a sík tetszőleges hasonlósági transzformációját komplex függvények segítségével.

34. Bizonyítsuk be, hogy bármely egybevágósági síktranszformáció előállítható legfeljebb három tükrözés egymásutánjaként.

35. Mi a kapcsolat az $r \cdot (\cos \varphi - i \cdot \sin \varphi)$ és az $r \cdot (\cos \varphi + i \cdot \sin \varphi)$ számok között? Mi az $r \cdot (\cos \varphi - i \cdot \sin \varphi)$ trigonometrikus alakja?

36. A komplex számokra vonatkozó négyzetgyökvonást felhasználva bizonyítsuk be, hogy $\cos\left(\frac{\varphi}{2}\right) = \sqrt{\frac{1 + \cos \varphi}{2}}$. Bizonyítsuk be, hogy $\sin\left(\frac{\varphi}{2}\right) = \sqrt{\frac{1 - \sin \varphi}{2}}$, ha $\sin \varphi$ pozitív, és ennek negatívja, ha $\sin \varphi$ negatív ($0^\circ \leq \varphi \leq 360^\circ$).

37. Bizonyítsuk be, hogy a komplex számok összeadására érvényes az úgynevezett háromszögegyenlőtlenség; azaz tetszőleges z és w komplex számokra teljesül: $|z + w| \leq |z| + |w|$.

38. Bizonyítsuk be, hogy tetszőleges z és w komplex számokra teljesül: $|z + w| \geq ||z| - |w||$.

39. Mit tudunk mondani $\arccos(1 + z)$ -ről $\arccos(z)$ ismeretében?

40. Mit tudunk mondani $\arccos(z + w)$ -ről $\arccos(z)$ és $\arccos(w)$ ismeretében?

41. Tekintsük azt az egyenest, amelynek az origóhoz legközelebbi pontja az i komplex szám. Mibe viszi ezt az egyenest a négyzetre emelés mint geometriai transzformáció?

42. Az előző feladatbeli egyenesre alkalmazzuk a köbre, illetve a negyedik hatványra való emelést mint geometriai transzformációt. Mutassuk meg, hogy hurkolt görbét kapunk, amelyik egy pontban metszi önmagát. Milyen lesz e görbék képe?

43. Mi történik, ha az előző feladatban magasabb hatványt vizsgálunk? Mitől függ a görbe önmagával való metszéspontjainak a száma?

44. Tekintsük a 41. feladatban szereplő egyenest, és vizsgáljuk a képét, ha négyzetre emelés helyett a négyzetgyökvonást nézzük. Milyen alakú lesz a kép, ha a gyökkitevőt növeljük?

45. Bizonyítsuk be, hogy ha n és k relatív prímek, akkor minden nk -adik primitív egységgyök előáll egy n -edik és egy k -adik primitív egységgyök szorzataként; és minden ilyen szorzat egy nk -adik primitív egységgyököt állít elő.

46. Mit állíthatunk az előző feladatban, ha n és k nem relatív prímek?

47. Bizonyítsuk be, hogy minden primitív egységgyök előáll mint prímhatalványokhoz tartozó primitív egységgyökök szorzata.

48. Határozzuk meg a $\varrho = \frac{-1 - \sqrt{-3}}{2}$ komplex szám abszolút értékét és arkuszát; továbbá ϱ^2 algebrai alakját.

MÁSODIK FEJEZET

MÁTRIXOK

1. A mátrix definíciója

A matematikában — de általában az életben is — gyakoriak az olyan rendszerek, amelyeknek a „jellemzéséhez” több szám szükséges. (Például: egy tetraéder jellemezhető az élei hosszával, egy elektromos hálózat a csomópontok közötti részek ellenállásának nagyságával, egy „gazdasági helyzet” azzal, hogy minden egyes termékből mennyi áramlik egyik „ágazat”-ból a másikba stb.)

Ezekben az esetekben természetesen nemcsak az számít, hogy milyen adatok szerepelnek, hanem az is, hogy ezek az adatok „egymáshoz képest” hogyan helyezkednek el. Ezzel lehet ugyanis meghatározni, hogy egy-egy adat mit jellemez. Ilyen esetekben a szóban forgó alakzathoz tartozó mátrixról beszélünk.

A mátrixoknak nagyon különböző „alakjuk” lehet; annak megfelelően, hogy a szereplő számokat hogyan helyeztük el. Ezt az elhelyezést általában a célszerűség diktálja. A leggyakrabban téglalap alakú mátrixokat használnak. Az alábbiakban mi is csak ilyenekkel foglalkozunk, sőt a definíciót is úgy adjuk meg, hogy csupán ezeket tekintjük mátrixoknak.

Mátrixoknak tekintjük tehát adatoknak téglalap alakban való elrendezését. E téglalapokról feltesszük, hogy véges sok soruk és véges sok oszlopuk van. A mátrixban szereplő adatokat a mátrix elemeinek nevezzük. Ezek az adatok általában számok, de a későbbiekben más adatok is előfordulnak, például függvények vagy esetleg újabb mátrixok is. Az elemeket azzal határozzuk meg, hogy egy-egy elemről megmondjuk, melyik sorban és melyik oszlopban áll. Az itt levő sor- és oszlopszámot az elemhez írjuk kettős indexként. (Ha félreértésre ad okot, akkor az indexeket vesszővel választjuk el egymástól.) $a_{i,j}$ valamely mátrixnak az i -edik sorában és j -edik oszlopában levő elemet jelöli. Például $a_{3,5}$ azt az elemet jelöli, amelyik a mátrix harmadik sorában és ötödik oszlopában áll.

A mátrixokat rendszerint zárójelbe tesszük, azért, hogy az egész táblázatot egyetlen egységként kezelhessük. Az irodalomban többféle zárójel használatos. Szokás az is, hogy feltüntetik a soroknak is és az oszlopoknak is a számát. Így egy mátrix — általában — a következőképpen írható fel:

$$\begin{array}{lcl}
 & \begin{array}{c} \text{1. oszlop} \\ \text{2. oszlop} \\ \vdots \\ \text{j. oszlop} \\ \vdots \\ \text{n. oszlop} \end{array} & \\
 \begin{array}{l} \text{1. sor} \rightarrow \\ \text{2. sor} \rightarrow \\ \vdots \\ \text{i. sor} \rightarrow \\ \vdots \\ \text{k. sor} \rightarrow \end{array} & \rightarrow & \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kj} & \dots & a_{kn} \end{bmatrix}
 \end{array}$$

Mi a mátrixokat szögletes zárójelbe tettük. Használatos ezen kívül a kerek zárójel vagy két-két függőleges vonal. Amennyiben nem akarjuk a mátrixokat ilyen „terjedelmesen” felírni, rövidítést használunk. Ilyen esetben a jellemző adatokat kell feltüntetni. Azt szokták megadni tehát, hogy a mátrixnak hány sora és hány oszlopa van, továbbá azt, hogy miképpen határozható meg általában valamely sor egy-egy eleme. Ezt az alábbi módon tesszük (feltüntettük az egyéb mátrixjelölési módokat is) :

$$[a_{ij}]_{kn}, \quad \text{vagy} \quad (a_{ij})_{kn}, \quad \text{vagy} \quad \|a_{ij}\|_{kn}.$$

Ha valamilyen módon adott a sorok és oszlopok száma, vagy ezek az adatok pillanatnyilag nem lényegesek, akkor ezeket nem is írjuk ki. Így a mátrixjelölés a következő lesz:

$$[a_{ij}], \quad \text{vagy} \quad (a_{ij}), \quad \text{vagy} \quad \|a_{ij}\|.$$

A fentiekben a mátrixokat csak szemléletesen definiáltuk. A szemléletes definícióra mindenképpen szükség van, mert csak így lehet látni, hogy miről van szó. Másrészt a mátrixokkal való „manipulálás” szükségessé teszi ezeknek lehetőleg minél precízebb definícióját is.

Láttuk, hogy a mátrixot azzal tudtuk meghatározni, hogy megadtuk sorainak és oszlopainak a számát, valamint azt, hogy egy-egy előírt helyen milyen szám áll. Ez a szám tehát a *helynek a függvénye*. A „hely” pedig nem más, mint egy számpár. Így a mátrixot egy olyan „függvény”-nek tekinthetjük, amely egy természetes számokból álló számpárhoz egy számot rendel hozzá. Azt kell még megnézni, hogy milyen számpárok léphetnek fel. Ez pedig nyilvánvaló, hiszen pontosan azok a számpárok léphetnek fel, amelyekben az első helyen k -nál, a második helyen pedig n -nél nem nagyobb természetes szám áll. Speciálisan egyetlen sort vagy egyetlen oszlopot, sőt egyetlen elemet is egy-egy mátrixnak tekinthetünk.

A következőkben az alábbi jelöléseket fogjuk precízen definiálni. Az A mátrix sorainak a számát $s(A)$, oszlopainak a számát $o(A)$ fogja jelölni (s a sor és o az oszlop rövidítése). iA jelöli a mátrix i -edik sorát, A_j a j -edik oszlopát és iA_j a mátrix i -edik sorának a j -edik elemét (pontosabban azt az egyelemű mátrixot, amelynek ez az egyetlen eleme).

2.1. Definíció. Jelölje $I_n = \{i \in \mathbb{N} \mid i \leq n\}$ az n -nél nem nagyobb természetes számok halmazát, és legyen \mathbb{K} egy tetszőleges számtest. Egy $A : I_k \times I_n \rightarrow \mathbb{K}$ függvényt \mathbb{K} -elemű vagy \mathbb{K} feletti mátrixnak nevezünk. (Ha \mathbb{K} a komplex, a valós, illetve a racionális számtest, akkor — megfelelően — komplex, valós, illetve racionális elemű mátrixokról beszélünk.)

A fenti mátrixról azt mondjuk, hogy $s(A) = k$ sora és $o(A) = n$ oszlopa van. Ha $s(A) = o(A)$, akkor négyzetes (kvadrátikus) mátrixról beszélünk. ■

Ezzel elmondtuk, hogy a mátrix elemei honnan valók (a \mathbb{K} számtestből), és elmondtuk, hogy hány sora és oszlopa van. Megmondtuk továbbá, hogy hogyan fogjuk jelölni az A mátrix sorainak és oszlopainak a számát. Természetesen **a mátrixra továbbra is úgy kell gondolni, mint az elemeinek egy téglalap alakú elhelyezésére.** A fentebb leírt mátrix esetében tehát $a_{ij} = A(i, j)$, azaz a „függvény” értéke az (i, j) helyen. Tulajdonképpen $A((i, j))$ volna a precíz jelölés, de nem okozhat félreértést, ha a kétszeres zárójelek helyett egyszeres zárójeleket írunk (a továbbiakban is).

A következőkben formailag is értelmezni fogjuk a mátrix sorait, oszlopait és elemeit:

2.2. Definíció. Legyen A tetszőleges mátrix.

Az A mátrix i -edik során azt az ${}_iA$ mátrixot értjük, amelyre $s({}_iA) = 1$, $o({}_iA) = o(A)$, és ${}_iA(1, j) = A(i, j)$. (Itt természetesen az i index már rögzített, csak a j változik 1-től n -ig.)

Az A mátrix j -edik oszlopán azt az A_j mátrixot értjük, amelyre $s(A_j) = s(A)$, $o(A_j) = 1$, és $A_j(i, 1) = A(i, j)$. (Itt viszont a j index rögzített és az i változik 1-től k -ig.)

Ha $s(A) = 1$, akkor A sormátrix, ha $o(A) = 1$, akkor A oszlopmátrix.

Sormátrix, illetve oszlopmátrix helyett gyakorta használatos a sorvektor, illetve oszlopvektor kifejezés. ■

Az $A = [a_{ij}]_{k,n}$ mátrix esetében tehát

$${}_iA = [a_{i1} \quad a_{i2} \quad \dots \quad a_{ij} \quad \dots \quad a_{in}] \quad \text{és} \quad A_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{kj} \end{bmatrix}.$$

Itt is hangsúlyozzuk, hogy a formális definíció helyett elég azt tudni, hogy ${}_iA$ az A mátrix i -edik sorának, míg A_j az A mátrix j -edik oszlopának a rövid jelölése.

A mátrixok esetében is az a célszerű, ha az egyenlőségnél nem követeljük meg az értékkészletek megegyezését, vagyis ha például egy valós elemű mátrix minden eleme racionális, akkor nem tekintjük különbözőnek attól a mátrixtól, amelyik „ugyanaz”, csak éppen racionális elemű mátrixnak értelmeztük.

2.3. Definíció. Ha az A és B mátrixoknak ugyanaz az értelmezési tartománya ($D(A) = D(B)$), akkor egyező alakúaknak nevezzük őket. Ha ezen felül még minden $(i, j) \in D(A)$ párra $A(i, j) = B(i, j)$ is teljesül, akkor a két mátrixot egyenlőnek tekintjük. ■

Ez tulajdonképpen a formális megfogalmazása annak, hogy mindkét mátrixban ugyanannyi sor és ugyanannyi oszlop van, és a megfelelő helyen álló elemek megegyeznek.

2.1. Tétel. *Tetszőleges A mátrixra ${}_i(A_j) = ({}_iA)_j$.*

Bizonyítás. Definíció szerint mind az ${}_iA$ mátrixnak, mind az ${}_i(A_j)$ mátrixnak egy sora van. Mivel $s({}_iA)_j = s({}_iA)$, ezért az $s({}_iA)_j$ mátrixnak is egy sora van. Hasonlóképpen látható be, hogy mindkét szóban forgó mátrixnak egy oszlopa van. Ezért az egyenlőséghez elég azt megnézni, hogy e mátrixok ugyanazt az értéket veszik fel az $(1, 1)$ helyen:

$${}_i(A_j)(1, 1) = (A_j)(i, 1) = A(i, j) = ({}_iA)(1, j) = ({}_iA)_j(1, 1). \quad \blacksquare$$

A tételben szereplő két mátrixnak egyetlen eleme, $A(i, j)$, természetesen az A mátrix i -edik sorában álló j -edik elem, ami a tétel szerint ugyanaz, mint a j -edik oszlop i -edik eleme. Ezt mondja ki az alábbi

2.4. Definíció. Az ${}_iA_j = {}_i(A_j) = ({}_iA)_j$ mátrix egyetlen eleme az A mátrix i -edik sorának j -edik eleme. \blacksquare

Természetesen szemléletesen tudjuk, hogy mi egy mátrix sora és oszlopa, illetve egy eleme. A fenti „precízkedésre” azért van szükség, mert a mátrixokkal való műveletek során elég bonyolult kifejezések adódnak; és ezeknek az egyenlőségét könnyebb formálisan belátni.

2. Műveletek a mátrixokkal

A mátrixok haszna és a velük való műveletek „értelme” a lineáris algebra tárgyalásánál válik világossá. A mátrixok tárgyalása viszont enélkül is lehetséges, és több szempontból célszerű. Egyelőre elégedjünk meg azzal, hogy a mátrixokat a számok „általánosításainak” tekinthetjük. Pontosabban szólva — ahogy az előző pontban láttuk — az egyelemű mátrixokat „szinte” azonosaknak tekinthetjük a számokkal. Tekintettel arra, hogy a számokkal különböző műveleteket tudunk végezni, ezért lehetőség van arra, hogy a műveleteket általában a mátrixokra kiterjeszthessük. E kiterjesztésnél természetesen vigyázni kell arra, hogy speciális esetben — vagyis egyelemű mátrixokra — a művelet ugyanaz legyen, mint a számokra.

Mint említettük, a mátrixokkal végzett műveleteknek a célszerűségét majd a későbbiekben fogjuk látni. Egyelőre nehéz volna e műveleteket indokolni, mint ahogy a számokkal végzett műveletek általánosítása sem tekinthető kielégítő célnak. Így azután megérthető, hogy a mátrixokkal való műveletek közül például a szorzásnak kétféle általánosítása is lesz. Minden egyes műveletet egyébként a 2.3. Definíció figyelembevételével fogunk meghatározni. Fontos megfigyelni, hogy e műveletekkel kapcsolatos bizonyításoknál *nem*

szerepel az osztás és (a 2.8. Tétel második egyenlőségétől, valamint a 2.10. Tétel szlopaiban felsorolt második és negyedik egyenlőségtől eltekintve) nincs szükség a szorzás kommutativitására sem.

2.5. Definíció. Legyen A egy \mathbb{K} feletti mátrix és $c \in \mathbb{K}$. Ha $A = [a](= [a]_{11})$, akkor legyen $c[a] = [ca]$; egyébként legyen cA az a mátrix, amelyre $D(cA) = D(A)$ és ${}_i(cA)_j = c \cdot {}_iA_j$. ■

Megjegyezzük, hogy a definíció második felének is van értelme, mert egyelemű mátrixokra a definíció első felében értelmeztük a c elemmel való szorzást.

2.2. Tétel. *Tetszőleges \mathbb{K} feletti A mátrix, valamint $c, d \in \mathbb{K}$ esetében teljesülnek a $(cd)A = c(dA)$ és $1A = A$ egyenlőségek.*

Bizonyítás. A szereplő mátrixok nyilván mind egyező alakúak.

$${}_i((cd)A)_j = (cd){}_iA_j = c(d{}_iA_j) = c{}_i(dA)_j = {}_i(c(dA))_j$$

bizonyítja az első állítást, és ${}_i(1 \cdot A)_j = 1 \cdot {}_iA_j = {}_iA_j$ a másodikat. ■

A mátrixoknál az egyenlőséget úgy értelmeztük mint függvények egyenlőségét. Hasonló módon értelmezzük a mátrixok összegét is. Tekintettel arra, hogy csak olyan függvények (pl. valós függvények) összegét értelmezhetjük, amelyeknek ugyanaz az értelmezési tartományuk, ezért ugyanezt kikötjük a mátrixokra is.

2.6. Definíció. Ha az A és B mátrixokra $D(A) = D(B)$, akkor e mátrixok $A + B$ összegét a $D(A + B) = D(A)$ és ${}_i(A + B)_j = {}_iA_j + {}_iB_j$ összefüggéssel definiáljuk. ■

Megjegyzés. A definíció szerint a két mátrix összege tehát ugyanolyan alakú, mint amilyenek az eredeti mátrixok voltak, továbbá egy-egy helyen álló elem megegyezik az adott két mátrix megfelelő helyen álló elemeinek az összegével. A definíció csak egyező alakú mátrixok összegét adja meg. Ha a két mátrix alakja nem egyezik meg, akkor nem értelmezzük e két mátrix összegét. (Ilyen esetekben később sem adunk definíciót.) □

2.3. Tétel. *Legyenek A, B és C egyező alakú mátrixok. Ekkor*

$$A + B = B + A \quad \text{és} \quad A + (B + C) = (A + B) + C,$$

azaz az összeadás kommutatív és asszociatív. Ezenfelül tetszőleges $x, y \in \mathbb{K}$ mellett teljesülnek az alábbi egyenlőségek:

$$(x + y)A = xA + yA \quad \text{és} \quad x(A + B) = xA + xB.$$

Bizonyítás. A 2.5. és 2.6. Definíció alapján a felírt műveletek mindegyike elvégezhető, és a kapott mátrixok értelmezési tartománya mindegyik esetben ugyanaz. Jelölje az adott mátrixok i -edik sorának j -edik elemét rendre a_{ij} , b_{ij} és c_{ij} . Bizonyítandó, hogy mind a négy egyenlőségben a bal oldali és a jobb oldali mátrix i -edik sorának j -edik eleme megegyezik. Ez az alábbi négy egyenlőségre vezethető vissza:

$$\begin{aligned} a_{ij} + b_{ij} &= b_{ij} + a_{ij} & \text{és} & & a_{ij} + (b_{ij} + c_{ij}) &= (a_{ij} + b_{ij}) + c_{ij}, \\ (x + y)a_{ij} &= xa_{ij} + ya_{ij} & \text{és} & & x(a_{ij} + b_{ij}) &= xa_{ij} + xb_{ij}. \end{aligned}$$

Az első két egyenlőség a számok összeadásának a kommutativitása és asszociativitása miatt áll fenn. Az utolsó két egyenlőség pedig a szorzásnak az összeadásra vonatkozó disztributivitásából következik. ■

A mátrixok körében elvégezhető a kivonás is — bizonyos esetekben (éppen akkor, amikor a két mátrix alakja megegyezik). Emellett bármilyen, rögzített alakú mátrixok között van olyan mátrix, amely úgy viselkedik, mint a számok között a 0. Azt mondhatjuk, hogy minden alakhoz külön-külön „0-mátrix” létezik. Ennek ellenére, ha ez nem okoz zavart, a „0-mátrixot” az alaktól függetlenül egységesen fogjuk jelölni.

2.4. Tétel. *Ha A és B egyező alakú mátrixok, akkor létezik egy olyan egyértelműen meghatározott C mátrix, amelyre $A = B + C$. Létezik továbbá egy olyan — ugyancsak egyértelmű — O mátrix, amelyre $A = A + O$. A C mátrixot az A és B mátrixok különbségének nevezzük és $(A - B)$ -vel jelöljük. A O mátrix neve nullmátrix. Használatos még $O - A$ helyett a $-A$ jelölés, ez A ellentettje vagy negatívja.*

$cA = O$ csak úgy lehetséges, hogy vagy $c = 0$, vagy $A = O$, és ekkor teljesül is.

Bizonyítás. Először lássuk a kivonást. Nyilvánvaló, hogy csak $D(C) = D(B)$ lehetséges. Az összeadás definíciójából következik, hogy csak az a C mátrix felel meg, amelyre ${}_i C_j = {}_i A_j - {}_i B_j$. Másrészt világos, hogy ez a mátrix valóban kielégíti a feltételt. Ebből már következik az is, hogy az O mátrix egyértelmű és létezik; csupán az lehetne a baj, hogy minden mátrixra más és más adódna. Tekintettel azonban arra, hogy ${}_i A_j = {}_i A_j$, ezért a különbségmátrix minden eleme 0. Ezzel a nullmátrix egyértelműsége is bizonyítást nyert.

Az utolsó állítás nyilvánvaló. ■

Megjegyezzük még, hogy a 2.5. Definícióból tüstént adódik a $-A = (-1) \cdot A$ összefüggés is. Azt is könnyen beláthatjuk, hogy bármely n természetes számra az n -tagú $A + \dots + A$ összeg megegyezik $n \cdot A$ -val és $0 \cdot A = O$ is érvényes.

A továbbiakban a mátrixok szorzásával foglalkozunk. A mátrixszorzás azonban elég bonyolult, ezért célszerű, hogy előkészületül csak bizonyos speciális mátrixok szorzatát értelmezzük.

2.7. Definíció. Legyen az A és B mátrixokra $s(A) = o(B) = 1$ (tehát A sormátrix és B oszlopmátrix), és $o(A) = s(B)$ (tehát elemeik száma egyező). Ekkor a két mátrix $A \cdot B$ szorzatán azt az egyelemű mátrixot értjük, amelynek eleme $\sum_i {}_1 A_i \cdot {}_i B_1$. ■

E definíció szerint tehát az

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \quad \text{és} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

mátrixok szorzata $A \cdot B = a_1 \cdot b_1 + a_2 \cdot b_2 + \dots + a_n \cdot b_n$. Látható, hogy a szorzást csak akkor végezhetjük el, ha a két mátrixnak ugyanannyi eleme van. Abban a speciális esetben,

amikor a két mátrix mindegyike egyetlen elemből áll, akkor éppen a számok szorzására jutunk vissza.

A mátrixok szorzása valójában a lineáris függvényrendszerek behelyettesítésének felel meg. Így A az $y = a_1x_1 + \dots + a_nx_n$ függvényt, míg B az $x_i = b_i$ függvényrendszert jellemzi. Behelyettesítve, az $y = a_1b_1 + \dots + a_nb_n$ összefüggéshez jutunk, ahol egy konstans tag lép fel, ami éppen az AB mátrix egyetlen eleme.

Különböző elemszámú sor és oszlop szorzatát nem definiáljuk.

Most pedig rátérünk a mátrixok szorzatának az általános definíciójára.

2.8. Definíció. Tegyük fel, hogy az A és B mátrixokra $o(A) = s(B)$ teljesül. Ekkor e két mátrix $A \cdot B$ szorzatát a következőképpen értelmezzük:

$$s(AB) = s(A), \quad o(AB) = o(B) \quad \text{és} \quad {}_i(AB)_j = ({}_iA)(B_j). \quad \blacksquare$$

Megjegyzés. A feltétel szerint az A mátrix i -edik sorára és a B mátrix j -edik oszlopára teljesülnek a 2.7. Definíció feltételei, vagyis a szorzatmátrix elemeit valóban definiáltuk. A definíciót szóban úgy fogalmazhatjuk meg, hogy a szorzat i -edik sorának j -edik eleme az első tényező i -edik sorának és a második tényező j -edik oszlopának a szorzata. \square

A mátrixszorzást általában is jól motiválja lineáris függvények behelyettesítése. Ha

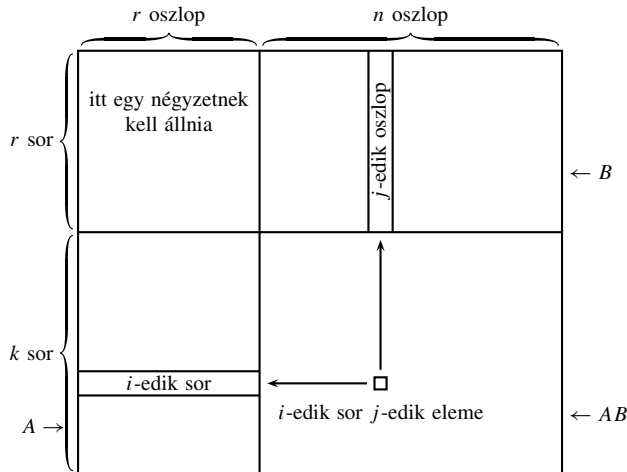
$x_i = \sum_j a_{i,j}$ és $y_j = \sum_k b_{j,k}z_k$, akkor $x_i = \sum_k \left(\sum_j a_{i,j}b_{j,k} \right) z_k$. Ezt a kapcsolatot a lineáris algebra tárgyalásánál majd részletesebben is bemutatjuk.

Technikailag az alábbi módon győződhetünk meg könnyen arról, hogy két — konkrétan megadott — mátrixot az adott sorrendben össze lehet-e szorozni, és ha igen, akkor hogyan kaphatjuk meg a szorzat elemeit:

Helyezzük el a két mátrixot úgy, hogy soraik vízszintesen legyenek; továbbá az első tényező „jobb oldali felső sarka” és a második tényező „bal oldali alsó sarka” ugyanoda kerüljön. A két mátrixot akkor és csak akkor szorozhatjuk össze, ha az a síkrész, amely az első tényező „fölött” és a második tényező „előtt” van, egy négyzet. Ebben az esetben a szorzatmátrixot az első tényező „után” és a második tényező „alatt” levő részben „helyezhetjük el”.

A szorzatmátrix bármelyik elemét a következőképpen írhatjuk fel: megnézzük, hogy a szóban forgó elem „előtt” melyik sor áll az első tényezőben és azt, hogy a szóban forgó elem „fölött” melyik oszlop áll a második tényezőben. A kapott sornak és a kapott oszlopnak a 2.7. Definícióban adott szorzata fogja megadni a kívánt elemet. Az is világos, hogy a szorzatnak annyi sora van, mint az első tényezőnek, és annyi oszlopa, mint a második tényezőnek.

A következő ábrán látható az eljárás:



2.5. Tétel. Ha a mátrixszorzás elvégezhető, akkor asszociatív, azaz $o(A) = s(B)$ és $o(B) = s(C)$ esetén $A(BC) = (AB)C$.

Bizonyítás. A szorzás definíciója szerint $s(BC) = s(B)$ és $o(AB) = o(B)$, tehát mindkét oldal értelmezve van. A megfelelő elemek egyenlőségét a következőképpen láthatjuk be:

$$\begin{aligned}
 i(A(BC))_j &= (iA)(BC)_j = \sum_r (iA_r)(r(BC)_j) = \sum_r (iA_r)((rB)(C)_j) = \\
 &= \sum_r (iA_r) \left(\sum_s (rB_s)(sC_j) \right) = \sum_r \sum_s iA_r \cdot rB_s \cdot sC_j = \\
 &= \sum_s \left(\sum_r (iA_r)(rB_s) \right) (sC_j) = \sum_s i(AB)_s (sC_j) = \\
 &= i(AB)(C)_j = i((AB)C)_j.
 \end{aligned}$$

■

2.6. Tétel. A mátrixok szorzása nem kommutatív.

Bizonyítás. Azt, hogy valamely azonosság nem érvényes, úgy kell érteni, hogy nem mindig teljesül. Tulajdonképpen ezt nagyon egyszerűen beláthatjuk. Ha ugyanis egy olyan A és egy olyan B mátrixot tekintünk, amelyre $o(A) = s(B)$, de $s(A) \neq o(B)$ (ilyenek nyilván léteznek), akkor az AB szorzat definiálva van, de a BA szorzat nincs. Felmerül azonban az a lehetőség, hogy talán amikor az AB és BA szorzat mindegyike értelmezve van, akkor megegyeznek. Hogy ennek sem kell mindig teljesülni, azt úgy láthatjuk be, hogy a mátrixok választásánál $s(A) = o(B)$ is teljesüljön, de ez a szám ne legyen egyenlő $s(B)$ -vel. Ekkor létezik mindkét szorzat, de nem lehetnek egyenlők, hiszen $D(AB)$ és $D(BA)$ is különböznek. Erre megint lehet azt az ellenvetést tenni, hogy esetleg $D(AB) = D(BA)$ esetén teljesül a kommutativitás.

Avégett, hogy ennek a lehetetlenségét is belássuk, olyan A és B mátrixokat kell megadni, hogy $s(A) = s(B) = o(A) = o(B)$, és a két szorzat ennek ellenére sem egyenlő. Lássunk erre egy konkrét példát, amikor a sorok és az oszlopok száma is kettő. Legyen

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{és} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Ekkor könnyű számolással adódik, hogy

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{és} \quad BA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad \blacksquare$$

2.7. Tétel. *Ha egy szorzat valamelyik tényezője nullmátrix, akkor a szorzat is az; azonban a szorzat lehet úgy is nullmátrix, hogy egyik tényező sem az.*

Bizonyítás. A most megadott példa bizonyítja a tétel második állítását, az első állítás pedig nyilvánvaló. \blacksquare

2.8. Tétel. *Ha létezik az AB szorzat, akkor tetszőleges c számra teljesül a $c(AB) = (cA)B = A(cB)$ összefüggés.*

Bizonyítás. Számmal való szorzásnál a mátrix értelmezési tartománya nem változik, ezért mindhárom esetben ugyanazt az értelmezési tartományt kapjuk. A mátrixok elemeit megfelelően a_{ir} -rel, illetve b_{rj} -vel jelölve az elemek egyenlőségéhez azt kell megmutatni, hogy

$$c \left(\sum_r a_{ir} b_{rj} \right) = \sum_r (c a_{ir}) b_{rj} = \sum_r a_{ir} (c b_{rj});$$

ami viszont azonnal következik a disztributivitásból és az asszociativitásból, valamint a kommutativitásból. \blacksquare

2.9. Tétel. *A mátrixszorzás az összeadásra vonatkozóan disztributív. Vagyis, ha az A , B , C és D mátrixokra $D(A) = D(B)$ és $o(C) = s(A)$, $o(A) = s(D)$, akkor érvényesek az alábbi egyenlőségek:*

$$C(A + B) = CA + CB \quad \text{és} \quad (A + B)D = AD + BD.$$

Bizonyítás. A feltételekből azonnal következik, hogy bármelyik egyenlőségnél mindkét oldalon ugyanolyan alakú mátrix áll. A mátrixok elemeit megfelelően a , b , c és d betűkkel jelölve az i -edik sor j -edik elemét kiszámítva, ezek egyenlősége a következőhöz vezet:

$$\begin{aligned} \sum_r c_{ir}(a_{rj} + b_{rj}) &= \sum_r c_{ir}a_{rj} + \sum_r c_{ir}b_{rj}, \\ \sum_s (a_{is} + b_{is})d_{sj} &= \sum_s a_{is}d_{sj} + \sum_s b_{is}d_{sj}. \end{aligned}$$

Ezek mindegyike azonnal következnek az összeadás azonosságából és a szorzásnak az összeadásra vonatkozó disztributivitásából. (A polinomoknál a behelyettesítés műveletének a tárgyalása alkalmából látjuk majd, hogy a „két disztributivitás” független egymástól.) ■

A 2.9. Tételből azonnal adódik az alábbi

Következmény. *A K testből vett elemű, n sorú és n oszlopú mátrixok K_n halmaza a fent értelmezett összeadásra és szorzásra nézve egy nemkommutatív gyűrűt alkot, amelyet a K feletti n -szer n -es teljes mátrixgyűrűnek nevezünk.* ■

A mátrixoknál fontos szerepet játszanak az alábbi fogalmak:

2.9. Definíció. Az A négyzetes mátrix fődiagonálisán az ${}_i A_i$ elemeket és $S(A)$ nyomán a $\sum_i {}_i A_i$ összeget értjük.

Ha egy négyzetes mátrixban a fődiagonális elemeinek kivételével minden elem 0, akkor diagonális mátrixról beszélünk. Ha emellett a fődiagonális elemei egyenlők, akkor ezt a mátrixot skalármátrixnak nevezzük. ■

Megjegyzés. Az A mátrix nyomát jelölik még $Sp(A)$ -val és $Tr(A)$ -val is. □

A mátrixok körében még két igen fontos műveletet vezetünk be. Ezek egyike sem mond semmit a számok esetében; egyikük a mátrixnak a fődiagonálisra való „tükrözése”. A másik művelet a konjugálás általánosításának tekinthető. Ezt úgy végezzük, hogy a mátrixot tükrözzük a fődiagonálisára, majd minden egyes elemének a konjugáltját vesszük.

2.10. Definíció. Legyen A tetszőleges mátrix. Az A mátrix A^\dagger transzponáltján azt a mátrixot értjük, amelyre $s(A^\dagger) = o(A)$, $o(A^\dagger) = s(A)$, továbbá ${}_i(A^\dagger)_j = {}_j A_i$. Az A mátrix adjungáltján pedig azt az A^* mátrixot értjük, amelyre $D(A^*) = D(A^\dagger)$, és ${}_i(A^*)_j = \overline{{}_j(A^\dagger)_i}$. ■

2.10. Tétel. *Ha az alábbi egyenlőségekben létezik a bal oldal, akkor létezik a jobb oldal is, és fennáll az egyenlőség:*

$$\begin{aligned} (A^\dagger)^\dagger &= A, & (A^*)^* &= A; \\ (c \cdot A)^\dagger &= c \cdot (A^\dagger), & (c \cdot A)^* &= \bar{c} \cdot A^*; \\ (A + B)^\dagger &= A^\dagger + B^\dagger, & (A + B)^* &= A^* + B^*; \\ (AB)^\dagger &= B^\dagger A^\dagger, & (AB)^* &= B^* A^*. \end{aligned}$$

Bizonyítás. A transzponáltakra vonatkozó első három állítás nyilvánvalóan teljesül. A negyedik állításból is világos annyi, hogy a jobb oldali mátrix létezik és ugyanolyan alakú, mint a bal oldali. A bal oldali szorzatmátrix i -edik sorának a j -edik eleme nem más, mint az AB mátrix j -edik sorának az i -edik eleme. Az elemeket megfelelően betűzve ez az elem $\sum_r a_{jr} b_{ri}$ lesz. A jobb oldalon álló mátrix i -edik sorának j -edik elemét úgy kapjuk, hogy B^\dagger i -edik sorának és A^\dagger j -edik oszlopának a szorzatát vesszük. Ez pedig, a transzponált

definícióját figyelembe véve, nem más, mint $\sum_r b_{ri} a_{jr}$. A két összeg pedig nyilvánvalóan megegyezik. ■

Az adjungáltakra vonatkozó azonosságok — a konjugálás tulajdonságait felhasználva — közvetlenül adódnak a transzponáltakra vonatkozó azonosságokból.

Megjegyzés. Bizonyítás közben láttuk, hogy sormátrix transzponáltja oszlopmátrix, az oszlopmátrix pedig sormátrix. Pontosabban szólva fennáll: $(A^\dagger)_i = ({}_i A)^\dagger$ és ${}_j (A^\dagger) = (A_j)^\dagger$.

Megemlítjük még a következőket. Ha az A mátrixra $A^\dagger = A$, illetve $A^* = A$ teljesül, akkor a mátrixot *szimmetrikusnak*, illetve *önadjungáltnak* nevezzük. □

3. Permutációk

A permutációkkal tulajdonképpen a kombinatorika foglalkozik, nem pedig az algebra. Itteni tárgyalásukra a később bevezetésre kerülő determináns definiálásánál van szükség. Bizonyos számok — vagy egyéb elemek — permutációján ezeknek más sorrendben való felírását értik. Mi e helyett a „statikus” szemlélet helyett azt az „eljárást” fogjuk tekinteni, amikor (vagy ahogyan) ezt a felírást végezzük. Éppen ezért a mozgást jobban kifejező *permutálást* fogjuk használni elnevezésként.

2.11. Definíció. Rögzített n természetes szám esetén permutálásnak nevezzük az $I_n = \{1, 2, \dots, n\}$ halmaznak vagy bármely n elemű halmaznak önmagára való kölcsönösen egyértelmű leképezését, azaz bijekcióját. ■

Megjegyzés. Egy σ permutálás tehát egy olyan függvény, amelyre $D(\sigma) = R(\sigma) = I_n$, azaz mind az értelmezési tartománya, mind az értékkészlete az 1-től n -ig terjedő természetes számok halmaza, ahol n egy rögzített természetes szám. Ezenfelül még annak is kell teljesülnie, hogy σ az I_n különböző elemeihez különböző elemeket rendel hozzá, továbbá mindig lehet találni olyan elemet, amelyhez a σ egy előírt elemet rendel hozzá.

Érdeemes figyelni arra a tényre, hogy egy véges halmaz önmagára való leképezése esetében mind az injektivitásból, mind a szürjektivitásból következik, hogy a leképezés bijektív. □

Ha $\sigma : I_n \rightarrow I_n$ egy permutálás, akkor $\sigma(1), \sigma(2), \dots, \sigma(n)$ a megfelelő permutáció.

A következő tétel a permutálásoknak számunkra alapvető tulajdonságait rögzíti:

2.11. Tétel. Egy n elemű A halmaz permutálásainak a száma $n!$. Az A halmaz permutálásai a függvénykompozícióra nézve csoportot alkotnak; azaz két permutálás szorzata (kompozíciója) is permutálás, ez a szorzás asszociatív, létezik „egységelem” és minden elemnek van inverze.

Bizonyítás. Először nézzük a csoporttulajdonságok (azaz a kompozíció felsorolt tulajdonságainak) bizonyítását. Legyenek $\sigma, \tau : A \rightarrow A$ bijektívek. Ha $a \in A$, akkor $\tau(a) \in A$, és így $\sigma\tau(a) = \sigma(\tau(a)) \in A$, tehát $\sigma\tau : A \rightarrow A$. Ha $a, b \in A$ és $a \neq b$, akkor τ injektivitása miatt $\tau(a) \neq \tau(b)$, és így σ injektivitása miatt $\sigma\tau(a) \neq \sigma\tau(b)$; tehát $\sigma\tau$ injektív.

Mivel σ szürjektív, ezért tetszőleges $a \in A$ elemhez létezik olyan $b \in A$ elem, amire $a = \sigma(b)$. A τ szürjektivitása miatt ehhez a b elemhez viszont van olyan c elem, amelyre $b = \tau(c)$. Ezért $a = \sigma\tau(c)$, tehát $\sigma\tau$ is szürjektív. Ez az injektivitással együtt pontosan azt jelenti, hogy $\sigma\tau$ is bijektív.

A szorzat asszociativitásához nincs is szükség másra, csak arra, hogy a szorzás az „egymás után végzés”. Legyen $\sigma, \tau, \varrho : A \rightarrow A$ és $a \in A$. Ekkor $((\sigma\tau)\varrho)(a) = (\sigma\tau)(\varrho(a)) = \sigma(\tau(\varrho(a))) = \sigma((\tau\varrho)(a)) = (\sigma(\tau\varrho))(a)$, ami a függvényegyenlőség definíciója szerint pontosan az asszociativitást jelenti.

A szorzás „egységeleme” egy olyan $\iota : A \rightarrow A$ elem, amelyre bármely $\sigma : A \rightarrow A$ esetén $\sigma\iota = \iota\sigma = \sigma$. Ennek a feltételnek nyilván megfelel az identitás, azaz az $\iota(a) = a$ tulajdonságú leképezés. Ez olyan bijekció, amely kielégíti a fenti kívánalmakat. (Más permutálás nyilván nem is felelne meg.)

Mivel σ bijektív, ezért $\sigma^{-1} : \sigma(a) \mapsto a$ is bijekció, amire $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \iota$ és $(\sigma^{-1})^{-1} = \sigma$ teljesülnek.

Most rátérünk az első állítás bizonyítására:

Általában azt bizonyítjuk be, hogy n -elemű A halmaznak bármely n -elemű B halmazra való bijekcióinak a száma $n!$. Állításunkat n -re vonatkozó teljes indukcióval bizonyítjuk.

1. Az $n = 1$ esetben egyetlen lehetőség van; és valóban $1! = 1$.

2. $n - 1$ -ről n -re. Tekintsük az $A = \{a_1, a_2, \dots, a_n\}$ és a $B = \{b_1, b_2, \dots, b_n\}$ halmazokkal együtt az összes $\sigma : A \rightarrow B$ bijekció halmazát. Bármely ilyen leképezés injektív módon képezi le az A halmaz $A' = \{a_1, a_2, \dots, a_{n-1}\}$ részhalmazát B -be. Nem B -re, mert egy elemnek biztosan ki kell maradnia. Jelöljük B_i -vel azt a halmazt, amelyet a B -ből a b_i elhagyásával nyerünk. Mivel A' -nek $n - 1$ eleme van, ezért a fenti σ leképezések az A' -t mindig valamelyik B_i -re képezik le. Az indukciós feltevés szerint ezeknek a bijekcióknak a száma minden egyes i esetén $(n - 1)!$. Ezek a bijekciók minden ilyen esetben pontosan egyféleképpen terjeszthetők ki az A -t B -be vivő bijekciókká, nevezetesen úgy, hogy az a_n elemet minden esetben a megfelelő b_i elemre képezzük le. Mivel a B -ből kihagyható elemek száma n , ezért az így előállított bijekciók száma $n \cdot (n - 1)! = n!$. Világos, hogy így minden $A \rightarrow B$ bijekció valóban elő is áll. ■

2.12. Definíció. Azt mondjuk, hogy a σ permutációnál a $\sigma(i), \sigma(j)$ pár inverzióban áll, ha $i < j$ és $\sigma(i) > \sigma(j)$. A σ permutációnál inverzióban álló párok számát a σ inverziószámának nevezzük és $I(\sigma)$ -val jelöljük.

A σ permutációt párosnak, illetve páratlannak nevezzük aszerint, hogy $I(\sigma)$ páros-e vagy páratlan. Ezt a tulajdonságot σ paritásának nevezzük. Ha σ és τ mindegyike páros vagy mindegyike páratlan, akkor egyenlő paritásúaknak nevezzük őket; egyébként különböző paritásúaknak. ■

2.12. Tétel. Legyen $P(\sigma) = \prod \{(\sigma(j) - \sigma(i)) \mid i, j \in I_n; i < j\}$ és $Q(\sigma) = \frac{P(\sigma)}{P(\iota)}$.

Ekkor $Q(\sigma) = 1$, ha σ páros és $Q(\sigma) = -1$, ha σ páratlan, azaz $Q(\sigma) = (-1)^{I(\sigma)}$.

Bizonyítás. Definíció szerint $Q(\sigma)$ az összes $\frac{\sigma(j) - \sigma(i)}{\iota(j) - \iota(i)}$, azaz az összes $\frac{\sigma(j) - \sigma(i)}{j - i}$ alakú tört szorzata, ahol $i < j$ és $i, j \in I_n$.

Ez azt jelenti, hogy a szorzat számlálójában is és a nevezőjében is az 1-től n -ig terjedő számok különbségei állnak; mégpedig minden különbség pontosan egyszer szerepel, mert σ bijektív. Így a szorzat számlálójának és nevezőjének megegyezik az abszolút értéke, tehát a hányados vagy $+1$, vagy -1 . Ha a $\sigma(j)$ és a $\sigma(i)$ pár inverzióban áll, akkor a megfelelő $\frac{\sigma(j) - \sigma(i)}{j - i}$ tört negatív, egyébként pozitív. Eszerint $Q(\sigma)$ annyi negatív tényezőt tartalmaz, amennyi a σ inverzióinak a száma; tehát pontosan akkor -1 a szorzat, ha σ páratlan. $Q(\sigma) = (-1)^{I(\sigma)}$ triviálisan igaz. ■

2.13. Tétel. Legyen $\sigma, \tau \in S_n$. Ekkor $Q(\sigma\tau) = Q(\sigma) \cdot Q(\tau)$; azaz Q multiplikatív. $n \geq 2$ esetén az I_n páros és páratlan permutációinak a száma megegyezik.

Bizonyítás. A $Q(\sigma)$ értéke nem függ attól, hogy a $\frac{\sigma(j) - \sigma(i)}{j - i}$ törtet milyen sorrendbe írjuk. Ez azt jelenti, hogy nem változik meg $Q(\sigma)$ értéke, ha I_n elemeire egy tetszőleges τ permutálást alkalmazunk:

$$Q(\sigma) = \frac{P(\sigma\tau)}{P(\iota\tau)} = \frac{P(\sigma\tau)}{P(\iota)} \cdot \frac{P(\iota)}{P(\tau)} = \frac{Q(\sigma\tau)}{Q(\tau)}.$$

Ez bizonyítja a $Q(\sigma\tau) = Q(\sigma) \cdot Q(\tau)$ összefüggést.

Legyen π az a permutálás, amelyre $\pi(1) = 2$, $\pi(2) = 1$ és $\pi(i) = i$, ha $i > 2$. A $Q(\pi)$ szorzat-előállításában azoknak a törteknek az értéke, amelyeknek a nevezőjében a különbség mindegyik tagja nagyobb, mint 2, az 1 lesz. Azoknak a törteknek az értéke, amelyeknek a nevezőjében a különbségnek csak az egyik tagja nagyobb, mint 2, az mindig pozitív lesz (ezek vagy $\frac{i-2}{i-1}$, vagy $\frac{i-1}{i-2}$ alakúak, ahol $i > 2$). A megmaradó egyetlen tényező $\frac{1-2}{2-1} = -1$. Eszerint π páratlan permutáció. Tekintettel arra, hogy $\pi\pi = \iota$, ezért a $\sigma \mapsto \sigma\pi$ megfeleltetés páros permutáláshoz páratlant, páratlanhoz párosat rendel hozzá; bijektív módon. ■

2.14. Tétel. Legyen $n \geq 2$, $\sigma \in S_n$. Ekkor $Q(\sigma^{-1}) = Q(\sigma)$ és bármely $\tau \in S_n$ esetén $Q(\tau\sigma\tau^{-1}) = Q(\sigma)$. Ha σ transzpozíció — azaz két elemet cserél fel és a többi önmagába viszi —, akkor σ páratlan.

Bizonyítás. $1 = Q(\iota) = Q(\sigma^{-1}\sigma) = Q(\sigma^{-1}) \cdot Q(\sigma)$ bizonyítja az első állítást. Q multiplikativitása és e tétel első állítása alapján $Q(\tau\sigma\tau^{-1}) = Q(\tau)Q(\sigma)Q(\tau^{-1}) = Q(\tau)Q(\tau^{-1})Q(\sigma) = Q(\sigma)$.

Legyen σ egy transzpozíció, amelyre például $\sigma(i) = j$, $\sigma(j) = i$ és minden más szóba jövő k esetén $\sigma(k) = k$. Tekintsük most az előző tétel bizonyításában szereplő π

transzpozíciót, amelyre tehát $\pi(1) = 2$, $\pi(2) = 1$ és $\ell > 2$ esetén $\pi(\ell) = \ell$. Ott azt is láttuk, hogy π páratlan permutálás. Legyen most $A = \{3, \dots, n\}$ és B az I_n -nek az A részhalmaza, amely az i és j elhagyásával keletkezik. Defináljuk az I_n τ permutálását a következőképpen: $\tau(1) = i$, $\tau(2) = j$ és képezze le τ az A halmazt a B -re bijektív módon — ez lehetséges, mert mindkét halmaznak $n - 2$ eleme van. Ekkor $\tau\pi\tau^{-1}(i) = \tau\pi(1) = \tau(2) = j$ és $\tau\pi\tau^{-1}(j) = \tau\pi(2) = \tau(1) = i$. B többi elemére $\tau^{-1}(k) \in A$ miatt $\pi\tau^{-1}(k) = \tau^{-1}(k)$ és így $\tau\pi\tau^{-1}(k) = k$, vagyis $\tau\pi\tau^{-1} = \sigma$. A tétel második állítása szerint $Q(\sigma)$ és π egyenlő paritásúak; tehát σ páratlan. ■

4. A determináns

A determinánsra a matematika majd minden ágában szükség van. Az algebraiban első-sorban elsőfokú egyenletrendszerek megoldásánál alkalmazzák. Tulajdonképpen a determináns fogalmával szinte mindenütt találkozunk, ahol négyzetes mátrixok szerepelnek — azaz olyan mátrixok, amelyeknél a sorok és oszlopok száma megegyezik. A determináns nem más, mint egy négyzetes mátrixhoz hozzárendelt szám. A determináns meghatározásánál már szükségünk van arra, hogy a mátrix elemeire igaz a szorzás kommutativitása; sőt az inverz mátrix meghatározásánál az osztás elvégezhetőségére is. A lineáris algebra tárgyalásánál látjuk majd a determináns „fogalmi” jelentését is. Itt most egy elemi definíciót adunk:

2.13. Definíció. Legyen az A négyzetes mátrixra $s(A) = o(A) = n$ és ${}_i A_j = a_{ij}$. Az A mátrix determinánsán az

$$|A| = \sum_{\sigma \in S_n} (-1)^{I(\sigma)} a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$$

számot értjük. Ha $o(A) = n$, akkor n -edrendű determinánsról beszélünk.

Az $a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$ szorzatokat a determináns tagjainak nevezzük, és $(-1)^{I(\sigma)}$ e tagok előjele. Az összegben szereplő tagokat előjeles tagoknak nevezzük. ■

Megjegyzések

1. Valós számok esetében egy mátrix determinánsának egy tagja lehet pozitív vagy negatív. E tag előjele nem ettől függ.

2. Egy n -edrendű mátrix determinánsának tehát annyi tagja van, amekkora S_n ; azaz $n!$. Egy-egy tagban (az előjeltől eltekintve) pontosan n darab tényező szerepel.

A 2.13. Definíció a következőket mondja a mátrix determinánsának az elkészítéséről:

Válasszunk ki a mátrix minden egyes sorából egy-egy elemet úgy, hogy ezek különböző oszlopokban legyenek, majd ezeket az elemeket szorozzuk össze. Ha a sorindexeket természetes sorrendben vesszük, akkor megnézzük, hogy az oszlopindexek permutációja páros-e vagy páratlan, és ennek megfelelően a szorzatot változatlanul hagyjuk, vagy megszorozzuk (-1) -gyel. Tegyük meg ezt minden lehetséges módon és a kapott szorzatokat adjuk össze, ez az összeg lesz a mátrix determinánsa.

Mint mondtuk, ez a definíció teljesen elemi, de igen bonyolult. Ezen a módon — általában — igen nehéz volna kiszámítani egy négyzetes mátrix determinánsát. Éppen ezért olyan számolási szabályokat fogunk megadni, amelyek a determináns kiszámítását egyszerűbbé teszik. A lineáris algebra tárgyalásánál majd látni fogjuk, hogy a determináns „fogalmi” definíciójánál pontosan ezek a tulajdonságok játszanak szerepet.

3. Vegyük észre, hogy a determináns tagjaiban levő tényezőket bármilyen sorrendben írhatjuk. Célszerű és szükséges az előjel meghatározása abban az esetben, ha a tagot nem úgy írjuk fel, hogy a sorindexek természetes sorrendben vannak. Legyen $a_{\tau(1)\varrho(1)} \cdot a_{\tau(2)\varrho(2)} \cdot \dots \cdot a_{\tau(n)\varrho(n)}$ a determináns egy kiszemelt tagja, amelyik megegyezik az $a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$ taggal. Ebben a szorzatban az $a_{i\sigma(i)}$ tényező valamelyik $a_{\tau(j)\varrho(j)}$ tényezővel egyező indexű, azaz $i = \tau(j)$ és $\sigma(i) = \varrho(j) = \varrho\tau^{-1}(i)$, vagyis $\sigma = \varrho\tau^{-1}$, tehát $(-1)^{I(\sigma)} = (-1)^{I(\varrho\tau^{-1})} = (-1)^{I(\varrho\tau)}$. Ennek megfelelően a 2.13. Definíció a következővel ekvivalens:

2.13. Definíció változat. Legyen az A négyzetes mátrixra $s(A) = o(A) = n$ és ${}_i A_j = a_{ij}$. Az A mátrix determinánsán az

$$|A| = \sum (-1)^{I(\varrho)+I(\tau)} a_{\tau(1)\varrho(1)} \cdot a_{\tau(2)\varrho(2)} \cdot \dots \cdot a_{\tau(n)\varrho(n)}$$

számot értjük, ahol $\varrho\tau^{-1}$ végigfut S_n elemein. ■

4. Ha az A mátrix elemekkel van megadva, akkor determinánsánál nem tesszük ki a mátrixot jelölő szimbólumot, hanem az elemeket csupán két függőleges vonal közé tesszük. Így, attól függően, hogy a determináns rendjét is jelölni akarjuk vagy sem, az $[a_{ij}]_{nn}$ mátrix determinánsára vagy az $|a_{ij}|_n$, vagy az $|a_{ij}|$ jelölést használjuk. $|A|$ helyett használni fogjuk még a $\det(A)$ jelölést is.

5. $n \leq 3$ esetén az n -edrendű determináns viszonylag könnyen meghatározható:

Az elsőrendű determinánsnak egyetlen tagja van, amelyet nem kell (-1) -gyel szorozni, mert az egyetlen elemű halmaznak nem létezik olyan permutációja, amelyben inverzió lenne. Ebben az esetben a determináns a mátrix egyetlen eleme lesz.

A másodrendű determinánsnak $2! = 2$ tagja van. Az oszlopindexek sorrendje vagy 1, 2, vagy 2, 1; az elsőhöz tartozó permutáció páros, a másodikhoz tartozó páratlan. Így a másodrendű determináns a következőképpen kapható:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \cdot d - b \cdot c.$$

A harmadrendű determináns $3! = 6$ tagjánál az oszlopindexek lehetséges sorrendje (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2) és (3, 2, 1). Az első, negyedik és ötödik esetben a megfelelő permutáció páros, a többiben páratlan. Tehát ezt a determinánst így kaphatjuk:

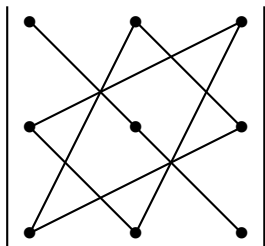
$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1.$$

Ez az összefüggés viszonylag még könnyen megjegyezhető. Évéggett írjuk le a harmadik oszlop után az első, majd ez után a második oszlopot is még egyszer. Egy olyan mátrixot kapunk, amelynek három sora és öt oszlopa van:

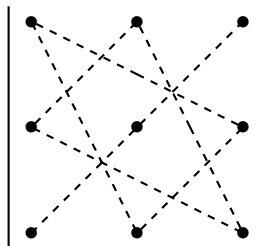
$$\begin{vmatrix} a_1 & a_2 & a_3 & a_1 & a_2 \\ b_1 & b_2 & b_3 & b_1 & b_2 \\ c_1 & c_2 & c_3 & c_1 & c_2 \end{vmatrix}$$

Könnyen ellenőrizhető, hogy a tagokat a következőképpen kaphatjuk meg. Az egy vonallal összekötött három-három elemet mindig összeszorozzuk. Ha a vonalat folyamatosan jelöltük, a szorzatot változtatlanul hagyjuk, ha viszont a vonalat szaggatottan jelöltük, a szorzatot még (-1) -gyel is meg kell szorozzuk. Ezt az összefüggést, illetve kiszámítási módot *Sarrus-szabálynak* nevezik. Bizonyos rutin után már nincs szükség az első két oszlop ismételt leírására. A következő sémákon az elemeket csupán pontokkal jelöljük. Az egy vonallal vagy egy „háromszöggel” összekötött elemeket kell összeszorozni. A folyamatos vonal esetében pozitív, a szaggatott vonal esetén negatív előjellel kell ellátni a szorzatokat.

pozitív előjellel:



negatív előjellel:



Legalább negyedrendű determinánsok esetében a számolásokhoz semmiféle „könnyű szabály” nincs. \square

A következő tételben a determinánsok alaptulajdonságait fogalmazzuk meg, vagyis azokat a tulajdonságokat, amelyeket közvetlenül a determináns definíciójából vezetünk le. A későbbi tulajdonságok mind olyanok lesznek, amelyeket az itt levezetett tulajdonságokból fogunk bizonyítani, anélkül, hogy a determináns definíciójára ismét hivatkoznánk.

2.15. Tétel. *Bármely n természetes szám esetén az n -edrendű determinánsokra igazak az alábbiak:*

a) *Ha egy négyzetes mátrix valamely sorát vagy oszlopát megszorozzuk egy c számmal, akkor a kapott mátrix determinánsa az eredeti mátrix determinánsának a c -szerese lesz.*

b) *Ha egy négyzetes mátrix sorait permutáljuk, akkor páros permutálás esetén nem változik a determináns, páratlan permutálás esetén pedig előjelet vált. Speciálisan, ha két sort felcserélünk, akkor a kapott mátrix determinánsa az eredeti mátrix determinánsának a negatívja lesz.*

c) *Tegyük fel, hogy az A' , A'' és A négyzetes mátrixokhoz található olyan i természetes szám, amelyre $j \neq i$ mellett igaz a $j A' = j A'' = j A$ összefüggés, továbbá $i A' + i A'' = i A$ (vagyis az első két mátrix i -edik sorának az összege a harmadik mátrix i -edik sora; míg a többi sorok mindhárom mátrixban megegyeznek). Ekkor $\det(A') + \det(A'') = \det(A)$.*

d) *Ha egy négyzetes mátrixban a fődiagonális minden eleme $+1$ és a többi elem mind 0 , akkor a mátrix determinánsa 1 .*

e) *Egy négyzetes mátrix transzponáltjának a determinánsa megegyezik az eredeti mátrix determinánsával.*

Bizonyítás. *a)* Legyen $A = [a_{ij}]$ az eredeti mátrix és $B = [b_{ij}]$ az a mátrix, amit úgy kapunk, hogy az A mátrix egy sorát vagy egy oszlopát c -vel szorozzuk. A 2.13. Definíció szerint ez utóbbi mátrix determinánsa az előbbtől abban különbözik, hogy mindenütt az a betű helyett b betű áll, azonban az indexek megegyeznek. Tekintettel arra, hogy a szorzatok minden sorból és minden oszlopból pontosan egy elemet tartalmaznak, az utóbbi mátrix determinánsában bármelyik tagot vesszük is ki, e tagban pontosan egy tényező lesz az eredetiben megfelelő tag megfelelő tényezőjének a c -szerese, a többi tényező változatlan marad. Tehát a második mátrix determinánsának minden tagja az első mátrix megfelelő tagjának a c -szerese lesz. Így az egész determináns is az eredeti mátrix determinánsának a c -szerese lesz.

b) Írjuk fel A determinánsát a 2.13. Definícióban adott módon:

$$|A| = \sum_{\sigma \in S_n} (-1)^{I(\sigma)} a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}.$$

Tegyük fel, hogy a sorokra a π permutálást alkalmaztuk. Ez azt jelenti, hogy a kapott $B = [b_{ij}]$ mátrixban a $\pi(i)$ -edik sorban szerepel az A mátrix i -edik sora. Azaz, a B -mátrix i -edik sora az eredeti mátrix $\pi^{-1}(i)$ -edik sora lesz. Így B determinánsa:

$$\begin{aligned} |B| &= \sum_{\sigma \in S_n} (-1)^{I(\sigma)} b_{1\sigma(1)} \cdot b_{2\sigma(2)} \cdot \dots \cdot b_{n\sigma(n)} = \\ &= \sum_{\sigma \in S_n} (-1)^{I(\sigma)} a_{\pi^{-1}(1)\sigma(1)} \cdot a_{\pi^{-1}(2)\sigma(2)} \cdot \dots \cdot a_{\pi^{-1}(n)\sigma(n)} = \\ &= \sum_{\sigma \in S_n} (-1)^{I(\sigma) + I(\pi^{-1})} a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}, \end{aligned}$$

felhasználva a 2.13. Definícióban a 3. megjegyzésében szereplő változatát. Az $I(\pi^{-1}) = I(\pi)$ összefüggés alapján tehát $|B| = (-1)^{I(\pi)} |A|$, mint állítottuk. Tekintettel arra, hogy két sor cseréje páratlan permutálás, ezért ekkor valóban előjelet vált a determináns.

c) A *b)* pont alapján feltehető, hogy $i = 1$. Legyen $A = [a_{ij}]$, $A' = [b_{ij}]$ és $A'' = [c_{ij}]$. Tekintsük $|A|$ -ban a σ permutáláshoz tartozó előjeles tagot. Ez $a_{1,\sigma(1)} \cdot a_{1,\sigma(1)}^*$ alakú, ahol a második tényező $a_{1,\sigma(1)}^* = a_{2,\sigma(2)} \cdot a_{n,\sigma(n)}$. E tényezőben nem szerepel a mátrix első sorából vett elem. Hasonlóképpen áll elő a másik két mátrixnak a σ permutáláshoz tartozó előjeles tagja, ezek $b_{1,\sigma(1)} \cdot b_{1,\sigma(1)}^*$, illetve $c_{1,\sigma(1)} \cdot c_{1,\sigma(1)}^*$ alakúak, ahol a $b_{1,\sigma(1)}^*$, illetve a $c_{1,\sigma(1)}^*$ szorzatokat hasonlóképpen határoztuk meg, mint az $a_{1,\sigma(1)}^*$ szorzatot. A mátrixok elemei a második sortól kezdve megegyeznek, így $a_{1,\sigma(1)}^* = b_{1,\sigma(1)}^* = c_{1,\sigma(1)}^*$, míg $a_{1,\sigma(1)} = b_{1,\sigma(1)} + c_{1,\sigma(1)}$. Ez azt jelenti, hogy az A mátrix determinánsának minden tagja előáll a másik két mátrix megfelelő tagjának összegeként, és így ez a determináns valóban egyenlő a másik két determináns összegével.

d) A szereplő mátrix determinánsában azoknak a tagoknak mindegyike 0, amelyekben egyetlen tényezőben is olyan a_{ij} elem áll, amelyre $i \neq j$, hiszen valamelyik tényező 0. Az egyetlen megmaradt tag $a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}$, ami az identikus permutáláshoz tartozik, és így előjele +1, mert az identikus permutáció páros.

e) Az eredeti $A = [a_{ij}]$ mátrix transzponáltja $[b_{ij}] = [a_{ji}]$. Itt a σ permutáláshoz tartozó tag $(-1)^{I(\sigma)} b_{1\sigma(1)} \cdot b_{2\sigma(2)} \cdot \dots \cdot b_{n\sigma(n)} = (-1)^{I(\sigma)} a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \dots \cdot a_{\sigma(n)n}$. Ez pedig a 2.13. Definíciónak a 3. megjegyzésében szereplő változata alapján megegyezik (előjelét tekintve is!) az A mátrix determinánsában fellépő $(-1)^{I(\sigma)} a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \dots \cdot a_{n\sigma(n)}$ taggal; tehát a két determináns valóban egyenlő. ■

1. Kiegészítés. A 2.15. Tétel b) és c) pontja akkor is igaz marad, ha sor helyett oszlopot tekintünk.

Bizonyítás. A szereplő mátrixokat transzponálva, a bizonyított tétel alapján a megfelelő összefüggések sorokra vonatkozó állításokba menvén át, igazak lesznek. A tétel e) pontja szerint ekkor az eredeti összefüggések is érvényesek. ■

2. Kiegészítés. Ha egy A négyzetes mátrix valamelyik sorában vagy oszlopában minden elem 0, akkor determinánsa 0.

Bizonyítás. Ha a szóban forgó sort 0-val szorozzuk, akkor a mátrix nem változik. A 2.15. Tétel a) pontja alapján determinánsa az eredeti determináns 0-szorosa lesz, azaz $|A| = 0 \cdot |A| = 0$. Az oszlopokra vonatkozó állítás ebből azonnal adódik az 1. Következmény alapján. ■

A determináns meghatározásánál igen fontos az alábbi

2.16. Tétel. Ha egy mátrixban két sor (vagy oszlop) megegyezik, akkor determinánsa 0. Ha egy mátrix valamely sorához (vagy oszlopához) hozzáadjuk egy tőle különböző sor (vagy oszlop) számszorosát, akkor determinánsa nem változik.

Bizonyítás. A 2.15. Tétel e) pontja alapján elegendő csak a sorokra vonatkozó állítást bizonyítani. Legyen az A (négyzetes) mátrix determinánsa $-d$. Ha A -nak két sorát felcseréljük, akkor egy olyan A' mátrixot kapunk, amelynek determinánsa a 2.15. Tétel b) pontja szerint d . Ha azt a két sort cseréljük fel, amelyik megegyezik, akkor $A' = A$ miatt $d = -d$ adódik. Ez csak a $d = 0$ esetben lehet.

Tekintsünk most egy tetszőleges (négyzetes) A mátrixot. Ha e mátrix i -edik sorához hozzáadjuk a j -edik sorának c -szeresét ($j \neq i$), akkor a kapott A' mátrix determinánsa a 2.15. Tétel c) pontja alapján két mátrix determinánsának az összegével lesz egyenlő. E két mátrix egyike az eredeti A mátrix. A másik mátrix — jelölje ezt A'' — úgy adódik az A -ból, hogy az A mátrix i -edik sora helyén a j -edik sornak a c -szerese szerepel. Az $|A''|$ determináns a következőképpen határozhatjuk meg. Legyen A''' az a mátrix, amelyet úgy kapunk az A mátrixból, hogy annak j -edik sorában is az i -edik sor szerepel. E tétel első állítása alapján $|A'''| = 0$. A 2.15. Tétel a) pontja alapján $|A''| = c \cdot |A'''| = 0$. Így valóban $|A'| = |A|$. ■

1. Kiegészítés. A 2.16. Tételben leírt eljárással, sor-, illetve oszlopcserevel, valamint egy sorának vagy oszlopának (-1) -gyel való szorzásával minden négyzetes mátrix diagonális alakra hozható úgy, hogy közben determinánása nem változik.

Bizonyítás. Legyen $A = [a_{ij}]$ négyzetes n -edrendű mátrix. Tegyük fel, hogy a k -nál kisebb indexű sorokban és oszlopokban a fődiagonális elemein kívül mindenütt 0 áll. A fenti eljárással el fogjuk érni, hogy ez a k -adik sorra és oszlopra is teljesüljön. Ezzel legfeljebb n lépésben megkapjuk a kívánt diagonális alakot. (A $k = 0$ eset adja az „induló” lépést.)

1. Ha minden $i \geq k$ indexre az i -edik sorban és oszlopban csak 0 áll, akkor a mátrix már diagonális alakban van.

2. Ha van olyan $i, j \geq k$ index, amire $a_{ij} \neq 0$, akkor a következőképpen járhatunk el: Mindenekelőtt célszerű a legkisebb ilyen i , majd ezen belül a legkisebb ilyen j indexet venni. Ha $i > k$, akkor felcseréljük az i -edik és k -adik sort, s ha $j > k$, felcseréljük a j -edik és k -adik oszlopot. Ezáltal a mátrix első $k - 1$ sora és oszlopa nem változott meg; tehát ott továbbra is minden diagonálison kívüli elem 0. Attól függően, hogy egy vagy két lépést kellett-e megtennünk, a determináns előjelet váltott, illetve nem változott. Az előbbi esetben (például) a k -adik sort (-1) -gyel szorozzuk. A mátrix első $k - 1$ sora és oszlopa most sem változott meg; tehát ott továbbra is minden diagonálison kívüli elem 0. A most kapott determináns viszont megegyezik az eredetivel. Jelöljük a most kapott mátrix elemeit b_{ij} -vel.

3. Minden $i > k$ indexre az i -edik sorból vonjuk ki a k -adik sor (b_{ik}/b_{kk}) -szorosát. Ezáltal a k -adik oszlopban csak a diagonális elem fog különbözni 0-tól, és a k -adik sor nem változik meg. Most minden $j > k$ indexre vonjuk ki a k -adik sor (b_{kj}/b_{kk}) -szorosát a j -edik oszlopból. Ezután már a k -adik sorban is csak a fődiagonális elem lesz 0-tól különböző. Mivel ezek a lépések a determinánst nem változtatják meg, ezért valóban elértük a bizonyítás elején kitűzött célt. ■

Megjegyzés. A fenti Kiegészítés módot ad a determináns viszonylag gyors kiszámolására, amely programozható is, és számítógépen gyorsan elvégezhető. □

2. Kiegészítés. Ha egy négyzetes mátrixban a fődiagonális alatt (vagy felett) álló minden elem 0, akkor a mátrix determinánása megegyezik a fődiagonális elemeinek szorzatával.

Bizonyítás. Ha az $A = [a_{i,j}]_n$ mátrixban a fődiagonális alatt álló minden elem 0 ($a_{i,j} = 0$, ha $i > j$), akkor tekintsük a fődiagonális elemeinek $a_{1,1}, \dots, a_{n,n}$ sorozatát. Tegyük fel, hogy ebben a sorozatban minden $i < r$ indexre $a_{i,i} \neq 0$. Vonjuk ki az első oszlop $\frac{a_{1,j}}{a_{1,1}}$ -szeresét a j -edik oszlopból minden $j > 1$ indexre. Ezután a második oszlop $\frac{a_{2,j}}{a_{2,2}}$ -szörösét a j -edik oszlopból minden $j > 2$ indexre és így tovább az $(r - 1)$ -edik oszlopig. A 2.16. Tétel szerint a kapott $B = [b_{i,j}]$ mátrix determinánása megegyezik az eredeti A mátrix determinánásával. Az eljárás alapján a B mátrix első $r - 1$ sorában a

fődiagonális minden eleme 0. Ezért az r -edik oszlopban $b_{i,r} = 0$, ha $i < r$. Az $i \geq r$ esetben az A mátrixra kirótt feltétel következtében $b_{i,r} = a_{i,r}$, hiszen az ezeken a helyeken álló elemekből mindig 0-t vontunk ki.

Ha $a_{r,r} = 0$, akkor a B mátrix r -edik oszlopának minden eleme 0. A 2.15. Tétel 2. Kiegészítése szerint tehát $|B| = 0$. Ekkor viszont a fődiagonális elemeinek a szorzata is 0, hiszen e szorzat egyik tényezője is 0. Amennyiben a fődiagonális egyetlen eleme sem 0, akkor az eljárást folytatva egy olyan B diagonális mátrixot kapunk, amelyben a fődiagonálison kívül minden elem 0, és a fődiagonális elemei ugyanazok, mint az A mátrixban. A 2.15. Tétel a) és d) pontja alapján ennek determinánsa éppen a fődiagonális elemeinek a szorzata. ■

5. A determináns kifejtése

Az első-, másod- és harmadrendű determináns meghatározási módját ismerjük, ezért lehetséges volna az n -edrendű determinánst rekurzívan meghatározni. Ezt valóban meg is lehet tenni. Előkészületül azonban néhány fogalomra lesz szükségünk.

2.14. Definíció. Ha egy mátrix egy sorát vagy egy oszlopát elhagyjuk, a kapott mátrixot az eredeti mátrix részmátrixának nevezzük. Ha egy mátrixnak egyetlen sora és egyetlen oszlopa van, akkor nincs részmátrixa. Egy mátrix részmátrixának nevezzük az összes olyan mátrixot, amelyke a fenti eljárással az eredeti mátrixból nyerhetőek. ■

Megjegyzés. Az eljárás első lépése az alábbi módon írható le részletesebben. Tetszőleges A mátrixból a következőképpen készítünk egy A' mátrixot: Rögzítünk egy i természetes számot, és $j < i$ esetén legyen ${}_j A' = {}_j A$, ha pedig $j > i$, akkor legyen ${}_j A' = {}_{j+1} A$. Hasonló vonatkozik egy oszlop elhagyására. A második rész azt mondja ki, hogy ezt a lépést a kapott új mátrixra tovább folytatva mindig az eredeti mátrix egy részmátrixát kapjuk. Másképpen fogalmazva: egy mátrix részmátrixait úgy nyerhetjük, hogy bizonyos sorait és/vagy oszlopait elhagyjuk.

Nem mondhatjuk azonban azt, hogy bizonyos sorokat vagy oszlopokat megtartunk, mert a „megmaradt” soroknak nem feltétlenül marad meg minden eleme. □

Induljunk ki egy A négyzetes mátrixból. Legyen ennek a determinánsa $|A|$, i -edik sorának a j -edik eleme pedig a_{ij} . Ha az A mátrix i -edik sorát és j -edik oszlopát elhagyjuk, akkor ismét egy négyzetes mátrixot kapunk. Ezeknek a továbbiakban fontos szerepük van.

2.15. Definíció. Legyen $A_{[ij]}$ az a mátrix, amely az A mátrixból az i -edik sor és a j -edik oszlop elhagyásával keletkezik. Az A mátrix ${}_i A_j$ eleméhez tartozó aldeterminánsnak nevezzük az $A_{ij} = (-1)^{i+j} \cdot |A_{[ij]}|$ számot. ■

Megjegyzések

1. Az aldetermináns tehát nem egy részmátrix determinánsa, hanem vagy részmátrix determinánsa, vagy ennek negatívja — az elhagyott oszlop és sor indexétől függően.

2. Könnyű meghatározni, hogy egy aldetermináns megegyezik-e a megfelelő részmátrix determinánsával, vagy annak ellentettje lesz-e. Erre szolgál az úgynevezett sakktáblaszabály. Képzeld el az eredeti mátrixot mint egy sakktáblát, amelynek n sora és n oszlopa van (az igazi sakktáblánál $n = 8$). Ezt a sakktáblát úgy képezhetjük, hogy az elemeket az egyes mezőkbe írjuk. A mezőket azonban nem „sötét”-re és „világos”-ra színezzük, hanem a felhasznált „színek” legyenek a „plusz” és „mínusz” jelek. Ezenfelül még azt is fel kell tenni, hogy az első sor első elemének megfelelő mezőbe a $+$ jelet írjuk. Így a következő „színezést” kapjuk:

+	−	+	−	+	−	...
−	+	−	+	−	...	
+	−	+	−	...		
⋮						

Amelyik mezőbe a $+$ jel került, az ahhoz tartozó aldetermináns megegyezik a megfelelő részmátrix determinánsával. Amelyik mezőben a $−$ jel van, az pedig e mátrix determinánsának az ellentettje lesz. \square

2.17. Tétel. Ha az $A = [a_{ij}]$ négyzetes mátrix i -edik sorában vagy j -edik oszlopában az a_{ij} -től különböző minden elem 0, akkor e mátrix $|A|$ determinánsára $|A| = a_{ij} \cdot A_{ij}$ teljesül.

Bizonyítás. Az $a_{ij} = 0$ esetben az állítás triviálisan igaz. Az $a_{ij} \neq 0$ esetben úgy bizonyítunk, hogy az állítást egyre egyszerűbb esetekre vezetjük vissza. Eközben a mátrixot megváltoztatjuk, azonban a megfelelő determinánsok nem változnak.

1. Feltehető, hogy sem az i -edik sorban, sem a j -edik oszlopban nincs a_{ij} -től eltekintve 0-tól különböző elem. Valóban, az i -edik sor (vagy a j -edik oszlop) megfelelő többszörösét a többi sorhoz (vagy oszlophoz) hozzáadva ez elérhető. Tudjuk, hogy eközben a determináns nem változott. Másrészt az i -edik sor j -edik eleméhez tartozó részmátrixszal az eljárás alatt nem történik semmi, ezért determinánsa sem változhat. Végül pedig a szóban forgó sor- és oszlopindexek változatlanok lévén, a $(-1)^{i+j}$ tényező sem változik.

2. Feltehető, hogy $i = j = 1$. A sorok és oszlopok szimmetrikus helyzetének következtében elég, ha azt bizonyítjuk, hogy $i = 1$ feltehető. Cseréljük fel az i -edik sort rendre az $(i-1)$ -edikkel, az $(i-2)$ -edikkel, \dots , a másodikkal, az elsővel. E sorcserénél a részmátrix nem változik, és így nem változik determinánsa sem. Nem változik meg az a_{ij} elem sem. Az eredeti mátrixban viszont $i-1$ számú sorcserét végeztünk, és így az új mátrix determinánsa az eredeti mátrixénak a $(-1)^{i-1}$ -szerese lesz. Tekintettel arra, hogy az eljárásnál az i -edik sorból az első lett, az aldeterminánsnál fellépő $(-1)^{i+j}$ faktorból $(-1)^{1+j}$ lett. E kettőnek a hányadosa pedig ugyancsak $(-1)^{i-1}$. Azaz, az eredetileg felírt összefüggés akkor és csak akkor áll fenn, ha az új mátrixban igaz a megfelelő összefüggés.

3. Feltehető, hogy a mátrixok diagonális alakúak — ekkor pedig az állítás igaz. Alkalmazzuk ugyanis a 2.16. Tétel utáni 1. Kiegészítést a részmátrixra. Ha eközben valamelyik sor vagy oszlop minden eleme 0 lesz, akkor ez a sor vagy oszlop az eredeti mátrixban is 0-vá válik; így $A = 0$ és $A_{ij} = 0$. Ha ez az eset nem fordul elő, akkor a részmátrixot is

és az eredeti mátrixot is diagonális alakra hoztuk. Mivel eljárás közben mind a részmátrixban, mind az eredetiben ugyanannyi sorcserét, illetve ugyanannyi oszlopserét végeztünk, a bizonyítandó egyenlőség mindkét oldalát (-1) -nek ugyanannyiadik hatványával szorozva helyes egyenlőséghez jutunk; bizonyítva a kívánt állítást. ■

2.18. Tétel (Kifejtési tétel). *Ha egy négyzetes mátrix valamely sorának vagy oszlopának elemeit megszorozzuk a hozzájuk tartozó aldeterminánssal, és a kapott szorzatokat összeadjuk, akkor a mátrix determinánsát kapjuk. Vagyis az $A = [a_{ij}]$ mátrix $|A|$ determinánsára:*

$$|A| = \sum_j a_{ij} \cdot A_{ij} = \sum_i a_{ij} \cdot A_{ij}.$$

Bizonyítás. Ismét elég, ha a sorokra vonatkozó állítást bizonyítjuk. Defináljuk — rögzített i esetén — az $A^{(j)}$ mátrixot úgy, hogy i -edik sorának j -edik eleme a_{ij} legyen, e sor többi eleme pedig 0. A mátrix többi sora egyezzen meg az A mátrix megfelelő sorával. Ekkor a 2.15. Tétel c) pontjának ismételt alkalmazásával azt nyerjük, hogy:

$$|A| = |A^{(1)}| + |A^{(2)}| + \dots + |A^{(n)}|.$$

A 2.17. Tétel szerint viszont $|A^{(j)}| = a_{ij} \cdot A_{ij}$, ami bizonyítja a tételt. ■

Kiegészítés (ferde kifejtés). *Ha egy négyzetes mátrix valamely sorának (oszlopának) elemeit megszorozzuk egy másik sor (oszlop) megfelelő eleméhez tartozó aldeterminánssal, akkor ezek összege 0 lesz, vagyis:*

$$0 = \sum_j a_{ij} \cdot A_{rj} = \sum_i a_{ij} \cdot A_{is}, \quad \text{ahol } r \neq i, \quad s \neq j.$$

Bizonyítás. Szintén elég, ha a sorokra szorítkozunk. Tegyük fel, hogy az i -edik sor elemeit a j -edik sor megfelelő elemeihez tartozó aldeterminánssal szorozzuk. Tekintsünk most egy mátrixot, amely az adottól abban különbözik, hogy e mátrix minden sora az eredetinek a megfelelő sorával egyenlő, kivéve a j -ediket, amely az eredeti mátrix i -edik sorával egyezik meg. E mátrix determinánsáról már tudjuk, hogy 0. Ha viszont e mátrix determinánsát „kifejtjük a j -edik sor szerint”, akkor éppen a fenti összeget kapjuk; ami bizonyítja állításunkat. ■

6. Speciális mátrixok

A determinánsok felhasználásával mód nyílik arra, hogy bizonyos mátrixok közötti speciális összefüggéseket megfogalmazhassunk. Előkészületül azonban szükségünk lesz egy — a későbbiekben is igen fontos — fogalomra, amelynek a segítségével megállapíthatjuk, hogy egy négyzetes mátrix determinánsa 0 vagy sem. (Itt csak egy „elvi” megállapításról van szó. Gyakorlatban a már ismertetett eljárással határozhatjuk meg a determinánst.)

2.16. Definíció. Legyenek $A^{(1)}, \dots, A^{(r)}$ azonos alakú mátrixok, továbbá c_1, \dots, c_r tetszőleges számok e mátrixok közös értékkészlet-tartományából. A

$$c_1 A^{(1)} + \dots + c_r A^{(r)}$$

mátrixokat az adott mátrixok lineáris kombinációinak nevezzük. Ha a tekintett számok mindegyike 0, akkor triviális lineáris kombinációról beszélünk. Minden más esetben nem triviális lineáris kombinációról. ■

Világos, hogy mátrixoknak a triviális lineáris kombinációja (ha készíthető lineáris kombináció) a nullmátrixot adja. Előfordulnak azonban olyan speciális esetek, amikor a mátrixok egy nem triviális lineáris kombinációja is a nullmátrixot eredményezi. Például az $A = [1, 2, 3]$, $B = [4, 5, 6]$, $C = [7, 8, 9]$ mátrixokra $1 \cdot A + (-2) \cdot B + 1 \cdot C = [0, 0, 0]$. Éppen ez adja a következő definíciónak az értelmét:

2.17. Definíció. Ha bizonyos mátrixoknak van olyan nem triviális lineáris kombinációja, amely a nullmátrixot adja, akkor e mátrixokat lineárisan összefüggőknek nevezzük. Ha ilyen lineáris kombináció nincs, akkor lineárisan független mátrixokról beszélünk. ■

2.19. Tétel. *Egy négyzetes mátrix determinánsa akkor és csak akkor 0, ha sorai (oszlopai) lineárisan összefüggők.*

Bizonyítás. Tekintsük az $A = [a_{ij}]$ négyzetes mátrixot. A szimmetria miatt itt is elegendő, ha csak a sorokkal foglalkozunk. Először tegyük fel azt, hogy e mátrix sorai lineárisan összefüggenek. Más szóval léteznek olyan c_1, \dots, c_n számok, amelyek közül nem mind 0, hogy a $c_1 \cdot ({}_1A) + \dots + c_n \cdot ({}_nA)$ mátrix minden eleme 0. A feltétel miatt van olyan i index, hogy $c_i \neq 0$. Szorozzuk ekkor meg az eredeti mátrix i -edik sorát c_i -vel, majd az így kapott mátrix i -edik sorához adjuk hozzá, minden i -től különböző j -re, a mátrix j -edik sorának c_j -szeresét. A kapott B mátrixra egyrészt $|B| = c_i \cdot |A|$, a 2.15. Tétel *a)* pontja és a 2.16. Tétel alapján. Másrészt e mátrixnak egy sora — a feltétel szerint — csupa 0-ból áll, amiből az következik, hogy determinánsa 0. Számok szorzata csak úgy lehet 0, ha valamelyik tényező 0, de $c_i \neq 0$, ezért $|A| = 0$.

Tegyük most azt fel, hogy az adott mátrix determinánsa 0. Állításunkat a determináns rendjére vonatkozó teljes indukcióval bizonyítjuk. Ha a determináns rendje 1, akkor a mátrixnak egyetlen eleme van, amely a feltétel alapján 0. A $c_1 = 1$ választással olyan nem triviális lineáris kombinációhoz jutunk, amely a nullmátrixot adja.

Tegyük fel, hogy az állítás igaz minden $(n-1)$ -edrendű determinánsra. Ha az adott A mátrix első sorában minden elem 0, akkor $c_2 = \dots = c_n = 0$, $c_1 = 1$ választással a $c_1 \cdot ({}_1A) + \dots + c_n \cdot ({}_nA)$ lineáris kombináció a nullmátrixot adja, és $c_1 \neq 0$ miatt ez nem triviális lineáris kombináció. Ha a mátrix első sorában nem minden elem 0, akkor feltehető, hogy ez éppen az első elem. A sorok összefüggőségén vagy függetlenségén ugyanis nem változtat az oszlopok cseréje (az összefüggés azt jelenti, hogy *minden* oszlopban 0 keletkezik a megfelelő lineáris kombinációnál), ezért feltehető, hogy $a_{11} \neq 0$. Az is világos, hogy a lineáris összefüggés, illetve függetlenség megmarad, ha az első sort elosztjuk a_{11} -gyel. Ezáltal a mátrix determinánsa is a_{11} -ed részére változik, tehát 0 marad.

Vonjuk most ki az i -edik sorból az első sor a_{i1} -szeresét. A kapott A' mátrixban az első oszlop első eleme 1, a többi pedig 0. Az eljárás során a determináns nem változik meg, tehát továbbra is $|A'| = 0$. Tekintsük most az első sor és első oszlop elhagyásával keletkező B mátrixot. A 2.17. Tétel alapján $0 = |A| = 1 \cdot |B|$, azaz $|B| = 0$. A teljes indukciós feltevés miatt léteznek olyan b_2, \dots, b_n számok, amelyek között van 0-tól különböző és $b_2 \cdot {}_2B + \dots + b_n \cdot {}_nB = 0$. Mivel az A' mátrix első oszlopában a második sortól kezdve mindenütt 0 áll, ezért $b_2 \cdot {}_2A' + \dots + b_n \cdot {}_nA' = 0$ is teljesül. A fenti összefüggésbe a fennálló ${}_iA' = {}_iA - a_{i1} \cdot {}_1A$ egyenlőségeket ($i > 1$) behelyettesítve

$$b_2 \cdot {}_2A + \dots + b_n \cdot {}_nA - (a_{21} + \dots + a_{n1}) \cdot {}_1A = 0$$

adódik. Ez az A mátrix sorainak egy lineáris kombinációja, de nem a triviális lineáris kombináció, hiszen feltételünk szerint a b_2, \dots, b_n számok között van 0-tól különböző. ■

2.20. Tétel. *Ha az n sorú négyzetes A mátrix determinánsa 0, akkor található hozzá olyan ugyancsak n sorú négyzetes B mátrix, amelynek nem minden eleme 0, de BA nullmátrix.*

Bizonyítás. A 2.19. Tétel szerint léteznek olyan c_i számok, amelyek között van 0-tól különböző, hogy $\sum_i c_i ({}_iA)$ a nullmátrixot adja. E mátrix j -edik eleme a tagok j -edik elemének az összege, azaz $\sum_i c_i ({}_iA_j)$; ami feltétel szerint 0-val egyenlő. Defináljuk most a B mátrixot úgy, hogy ${}_jB_j = c_j$. Ekkor B nem a nullmátrix, hiszen van 0-tól különböző eleme. (Az is feltehető, hogy B -nek éppen n sora van, tehát négyzetes mátrix, ez azonban a bizonyításnál tulajdonképpen nem is lényeges.) Most a BA mátrix i -edik sorának j -edik elemét határozzuk meg:

$${}_i(BA)_j = ({}_iB)(A_j) = \sum_k c_k \cdot ({}_kA_j) = 0. \quad \blacksquare$$

További előkészületül bevezetjük az úgynevezett Kronecker féle δ -függvényt.

2.18. Definíció. Jelölje δ_{ij} azt a természetes számpárokon értelmezett függvényt, amelynek értéke 1, ha i és j megegyezik, míg $i \neq j$ esetében a függvényérték 0. ■

Megjegyzés. A függvény analóg módon értelmezhető tetszőleges halmaz esetén, azaz 1 az értéke azonos elemekből álló párra, és 0 az értéke különböző elemekből álló párra. □

2.19. Definíció. Azokat az I négyzetes mátrixokat, amelyekre ${}_i I_j = \delta_{ij}$, egység mátrixoknak nevezzük. Ha azt is jelölni akarjuk hogy a mátrixnak n sora van, akkor az I_n jelölést használjuk. ■

2.21. Tétel. Ha egy I egység mátrixra létezik az AI , illetve az IB szorzat, akkor $AI = A$, illetve $IB = B$. Ezzel a tulajdonsággal csak az egység mátrixok rendelkeznek.

Bizonyítás. Az első állításnál a szimmetria miatt elég az AI -re vonatkozó részt belátni.

$${}_i(AI)_j = \sum_k ({}_i A_k)({}_k I_j) = \sum_k ({}_i A_k)\delta_{kj} = {}_i A_j,$$

mint állítottuk.

Ha valamely E mátrix hasonló tulajdonsággal rendelkezik, akkor a megfelelő sor-számú I egység mátrixszal szorozva az EI szorzat I a feltétel szerint, és E a most bizonyítottak alapján. Tehát e két mátrix valóban egyenlő. ■

2.20. Definíció. Ha az A négyzetes mátrixhoz található olyan B , illetve C mátrix, amelyekre BA , illetve AC egység mátrix, akkor ezeket — megfelelően — az A mátrix bal, illetve jobb oldali inverz mátrixának — vagy röviden balinverzének, illetve jobbinverzének — nevezzük. Ha $B = C$, akkor inverz mátrixról beszélünk. ■

Megjegyzés. Noha egyelőre nem lesz szükségünk rá, tetszőleges téglalap alakú mátrix esetében definiálható a balinverz és a jobbinverz. Ha A -nak k sora és n oszlopa van, akkor a BA szorzatnak (ha létezik) n oszlopa van. Ezért ez csak az I_n egység mátrix lehet. Ha $BA = I_n$, akkor B az A -nak balinverze. (Ekkor természetesen B -nek n sora és k oszlopa van.) Megfelelően, ha $AC = I_k$, akkor C az A -nak jobbinverze. A lineáris algebra tárgyalásánál látni fogjuk, hogy $k \neq n$ esetben csak egyik oldali inverz létezhet (ha $k < n$, akkor $AC = I_k$ lehetséges). □

2.22. Tétel. Ha az A négyzetes mátrix determinánsa nem 0, akkor létezik inverz mátrixa.

Bizonyítás. Legyen az $A = [a_{ij}]$ mátrix a determinánsa 0-tól különböző. Ekkor a következőképpen fogjuk megkonstruálni az A -nak a $B = [b_{ij}]$ inverz mátrixát: legyen $b_{ij} = A_{ji} \cdot a^{-1}$, ahol A_{ji} a j -edik sor i -edik eleméhez tartozó aldetermináns. Ebből:

$${}_i(AB)_j = \sum_k a_{ik} \cdot A_{jk} \cdot a^{-1}.$$

A jobb oldalon a kifejtési tétel szerint $a \cdot a^{-1} = 1$ áll az $i = j$ esetben, míg a ferde kifejtés szerint 0-t kapunk, ha $i \neq j$. Így B az A -nak jobbinverze. Az oszlopok szerinti kifejtést használva ugyanígy adódik, hogy B balinverze is A -nak, ami bizonyítja a tételt. ■

2.23. Tétel. Ha az A négyzetes mátrix determinánsa 0, akkor nem létezik sem balinverze, sem jobbinverze.

Bizonyítás. Azt mutatjuk meg, hogy az A -nak nincs jobbinverze. A balinverzre vonatkozó állítást hasonlóan (illetve a transzponálás felhasználásával) bizonyíthatjuk. Tekintsük az A mátrixhoz a 2.20. Tételben konstruált B mátrixot, amelynek minden egyes sora ugyanaz volt. $BA = 0$ alapján bármely C mátrixot is tekintünk, $B(AC)$ is nullmátrix lesz (amennyiben a szorzás elvégezhető). Így $\sum_k ({}_i B_k)({}_k AC_j) = 0$, ami azt jelenti, hogy AC

sorainak a B valamely sora elemeivel vett lineáris kombinációja a nullmátrixot adja. Tekintettel arra, hogy B nem a nullmátrix volt, ezért a szóban forgó lineáris kombináció nem a triviális. Ebből pedig a 2.19. Tétel alapján az következik, hogy az AC mátrix determinánsa 0. Az egységmátrix determinánsa a 2.15. Tétel d) pontja szerint 1 (tehát nem 0), ezért a szorzat nem lehet az egységmátrix. ■

Megjegyzés. Tulajdonképpen többet is bebizonyítottunk, mint amit állítottunk. Nevezetesen, csak annyit használtunk fel, hogy az A négyzetes mátrixhoz létezik olyan B mátrix, amelyik nem nullmátrix, de a BA szorzat nullmátrix. Már ebből következik tehát, hogy az A mátrixnak nem létezik jobbinverze. Ebből azonban az is következik, hogy az A mátrix sorainak létezik olyan nem triviális lineáris kombinációja, amely a nullmátrixot adja, vagyis a szóban forgó mátrix determinánsa 0.

A 2.23. Tétel szerint, ha az A négyzetes mátrixnak létezik balinverze, akkor determinánsa nem lehet 0. Ebből viszont a 2.22. Tétel alapján következik, hogy létezik inverze, ami megegyezik a balinverzével. □

Eredményeinket a következőképpen foglalhatjuk össze:

A négyzetes mátrixok közül pontosan azok a nulldeterminánsúak, amelyeknek nem létezik sem bal-, sem jobbinverzük, illetve található hozzájuk olyan mátrix, amelyik nem nullmátrix, de az eredetit szorozva vele mégis a nullmátrixot kapjuk (ilyen található „bal oldalról” is és „jobb oldalról” is).

A nem nulldeterminánsú mátrixok pontosan azok, amelyeknek létezik inverzük, illetve egy mátrixszal balról vagy jobbról szorozva csak úgy kaphatunk nullmátrixot, ha ez a tényező is nullmátrix.

Ezeknek az állításoknak a bizonyítását nem közöljük. Az eddig bizonyított tételekből azonban csupán a műveletekre vonatkozó összefüggések felhasználásával bizonyíthatók.

Megjegyzés. Láttuk, hogy ha két négyzetes mátrix valamelyikének a determinánsa 0, akkor a szorzatuknak (ha létezik) szintén 0 lesz a determinánsa. Ennél több is igaz: nevezetesen, ha két négyzetes mátrix összeszorozható, akkor szorzatuk determinánsa megegyezik a tényezők determinánsának a szorzatával. Ennek a bizonyítása itt nagyon hosszadalmas lenne; a lineáris algebra tárgyalása közben majd adunk erre egy egyszerű bizonyítást. □

Feladatok

Az 1–25. feladatokban azonos méretű négyzetes mátrixok szerepelnek:

1. Milyen feltétel mellett igaz $(A + B)^2 = A^2 + 2AB + B^2$; $(A + B)(A - B) = A^2 - B^2$; $(A + B)^2 = A^2 + B^2$?

2. Mutassuk meg, hogy ha $AB = A$ és $BA = B$, akkor mindkét mátrix *idempotens* — azaz $A^2 = A$, illetve $B^2 = B$.

3. Mutassuk meg, hogy lehet $A^2 = A$ és $B^2 = B$ és $AB = BA = O$.

4. Bizonyítsuk be, hogy ha A^{-1} és B^{-1} léteznek, akkor $(AB)^{-1} = B^{-1}A^{-1}$.

5. Mutassuk meg, hogy A akkor és csak akkor *involúció*, azaz $A^2 = I$, ha $(I + A)(I - A) = O$.

6. Bizonyítsuk be, hogy szimmetrikus A és B mátrixokra AB pontosan akkor szimmetrikus, ha $AB = BA$.

7. Bizonyítsuk be, hogy ha A szimmetrikus, akkor bármely P mátrixra létezik PAP^\dagger és az szimmetrikus.

8. Bizonyítsuk be, hogy A és B pontosan akkor felcserélhetőek, ha $A - cI$ és $B - cI$ felcserélhetőek (c egy szám).

9. Mutassuk meg, hogy $A + A^\dagger$ és AA^\dagger mindig szimmetrikus, és $A + A^*$ és AA^* mindig önadjungált.

10. Feleltessük meg az $a+bi$ komplex számnak az $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ valós elemű mátrixot. Bizonyítsuk be, hogy ez egy injektív homomorfizmus, amelynél a nyom a nyomba és a norma a determinánsba megy át.

11. Tekintsük a komplex elemű $\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$ mátrixot. Bizonyítsuk be, hogy ezek az összeadásra és a szorzásra egy „nemkommutatív” testet alkotnak (a szorzás nem kommutatív).

13. Bizonyítsuk be, hogy ha A *nilpotens* (azaz van olyan n természetes szám, amelyre $A^n = O$), akkor $(I - A)$ -nak létezik inverze.

14. Egy mátrix *felső (alsó) háromszög mátrix*, ha a fődiagonális alatti (feletti) elemek mindegyike 0. (Az a_{ij} elem a fődiagonális alatt van, ha $i > j$, a fődiagonális felett, ha $i < j$.) Bizonyítsuk be, hogy a felső (alsó) háromszög mátrixok az összeadásra és szorzásra egy nemkommutatív gyűrűt alkotnak.

15. Nevezzünk egy felső (alsó) háromszög mátrixot szigorúan felső (alsó) háromszög mátrixnak, ha a fődiagonálisban is csupa 0 áll. Bizonyítsuk be, hogy ezek is gyűrűt alkotnak; és ebben a gyűrűben minden n -tényezős szorzat O (n a mátrixok sorszáma).

16. Bizonyítsuk be, hogy ha A idempotens, akkor $B = I - A$ is az; és $AB = BA = O$.

17. Bizonyítsuk be, hogy ha egy mátrix minden mátrixszal felcserélhető, akkor skalár mátrix. Bizonyítsuk be, hogy diagonális mátrixokkal felcserélhető mátrixok pontosan a diagonális mátrixok.

18. Határozzuk meg az $AB - BA$ nyomát. Bizonyítsuk be, hogy $AB - BA = I$ lehetetlen.

19. Bizonyítsuk be, hogy A^{-1} akkor és csak akkor létezik, ha A -val lehet „egyszerűsíteni”, azaz az $AB = AC$ feltételből mindig következik $B = C$.

20. Egy A mátrixot *antiszimmetrikusnak* nevezzünk, ha $A^\dagger = -A$. Határozzuk meg az antiszimmetrikus mátrixok diagonális elemeit.

21. Bizonyítsuk be, hogy minden A önadjungált mátrix felírható $A + i \cdot B$ alakban, ahol A szimmetrikus, B antiszimmetrikus (valósak) és i a komplex egység.

22. Bizonyítsuk be, hogy ha A -nak létezik inverze, akkor $(A + B)A^{-1}(A - B) = (A - B)A^{-1}(A + B)$.

23. Bizonyítsuk be, hogy egy páratlan sorszámú antiszimmetrikus (valós elemű) mátrix determinánsa 0.

24. Bizonyítsuk be, hogy egy önadjungált mátrix determinánsa valós.

25. Legyenek az A, B, C négyzetes mátrixok méretei megegyezőek (mint ahogy a feladatok elején már feltettük). Bizonyítsuk be, hogy a $P = \begin{bmatrix} A & O \\ C & B \end{bmatrix}$ mátrixra $|P| = |A| \cdot |B|$.

26. Bizonyítsuk be, hogy ha az A_i mátrixok méretei, illetve a B_j mátrixok méretei egymás között azonosak és léteznek az $A_i B_j$ szorzatok, akkor

$$[A_1 \quad \dots \quad A_r] \cdot \begin{bmatrix} B_1 \\ \vdots \\ B_r \end{bmatrix} = A_1 B_1 + \dots + A_r B_r.$$

Definiáljuk a végtelen mátrixokat a következőképpen:

Definíció. Jelölje \mathbb{N} a természetes számok halmazát, és legyen \mathbb{K} egy tetszőleges számtest. Egy $A : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{K}$ függvényt \mathbb{K} -elemű, vagy \mathbb{K} feletti végtelen mátrixnak nevezünk.

A fenti mátrixról azt mondjuk, hogy végtelen sok sora és végtelen sok oszlopa van. Itt négyzetes (kvadrátikus) mátrixról nincs értelme beszélni. ■

27. Bizonyítsuk be, hogy az $A = [a_{i,j}]$ és $B = [b_{i,j}]$ végtelen mátrixok összegét az $A + B = C = [c_{i,j}] = [a_{i,j} + b_{i,j}]$ összefüggéssel; a $d \cdot A$ szorzatot a $d \cdot A = [d \cdot a_{i,j}]$ összefüggéssel definiálva, a műveletekkel a végtelen mátrixokra ugyanolyan azonosságok teljesülnek, mint a végesekre.

28. Mutassuk meg, hogy végtelen mátrixok esetében a $D = A \cdot B$ szorzatra a $D = [d_{i,j}] = \left[\sum_t a_{i,t} b_{t,j} \right]$ definíció értelmetlen.

Nevezzünk egy $A = [a_{i,j}]$ végtelen mátrixot

1. sorvégesnek, ha minden i -hez van olyan j_i , hogy $j > j_i$ esetén $a_{i,j} = 0$,
2. oszlopvégesnek, ha minden j -hez van olyan i_j , hogy $i > i_j$ esetén $a_{i,j} = 0$,
3. sor- és oszlopvégesnek, ha sorvéges és oszlopvéges,
4. végesnek, ha van olyan n , hogy $i, j > n$ esetén $a_{i,j} = 0$.

29. Bizonyítsuk be, hogy a fent definiált négyfajta mátrix esetén a definiált összeadásnak is és a számmal való szorzásnak is ugyanilyen fajta mátrix az eredménye.

30. Bizonyítsuk be, hogy a fent definiált négyfajta mátrix esetén a definiált szorzás elvégezhető és ugyanilyen fajta mátrix az eredménye.

31. Milyen mátrixot kapunk, ha egy sorvéges mátrixot és egy véges mátrixot szorzunk?

32. Legyen A sorvéges, B oszlopvéges; mit tudunk mondani az AB , illetve a BA mátrixról?

HARMADIK FEJEZET

EGYHATÁROZATLANÚ POLINOMOK

1. Az egyhatározatlanú polinomok fogalma

A polinomok bevezetése az egyenletek tárgyalásánál vált szükségessé. Az egyenletben szereplő ismeretlenről az egyenlet megoldása előtt csak azokat a tulajdonságokat lehet feltenni, amelyek minden számra teljesülnek. Valami „általános” vagy „határozatlan” számra lenne szükség, amellyel úgy lehetne számolni, mintha szám lenne, és a végén a helyére akármelyik számot beírhatjuk.

Példaként nézzük az $x^2 - 3x + 2 = 0$ egyenlet megoldását. Ennek az egyenletnek a bal oldalát felírhatjuk $(x - 2)(x - 1)$ alakban. Más szóval $x^2 - 3x + 2 = (x - 2)(x - 1)$. Ez utóbbi felírás egy „azonosság”, vagyis az egyenlőség mindig teljesül, bármilyen számot írunk is az x helyébe. Behelyettesítés után persze a két tényező szorzata csak úgy lehet 0, ha valamelyik tényező az; vagyis az eredeti egyenletnek csak 1 vagy 2 lehet a megoldása.

Egy ilyen x „határozatlan” számmal tehát műveleteket lehetne végezni; azaz tetszőleges „konkrét” a_0, a_1, \dots, a_n számokkal képezhetnénk az $a_0 + a_1x + \dots + a_nx^n$ alakú kifejezéseket, majd ezeknek esetleg a hányadosát is. Ez utóbbiakat most nem fogjuk tárgyalni. A fenti „kifejezéseknél” az okoz gondot, hogy értelmetlen egy „formális” határozatlant és egy konkrét számot összeszorozni, vagy a határozatlant hatványozni, illetve ilyeneket összeadni.

A polinomok $a_0 + a_1x + \dots + a_nx^n$ alakú „kifejezések”, ahol az a_i -k „valamilyen” számok. Tulajdonképpen számok helyett tekinthetnénk például mátrixokat is. Általánosabban ezek egy \mathcal{R} gyűrű elemei. Az „ x ” nem ismeretlen, mert nem egyenletet akarunk megoldani; de nem is változó, mert nem függvényeket adtunk meg. Az x -et itt *határozatlannak* fogjuk hívni.

A polinom a függvény „előkészítő” alakja; a határozatlan „tulajdonsága” az, hogy „bármit behelyettesíthetünk”. A „bármilyen” azt jelenti, hogy amit behelyettesítünk, az általában egy nagyobb S gyűrű eleme. (Egy egész együtthatós polinomba behelyettesítjük a „gyökeket”, amelyek általában nem is racionális számok.) Emellett azt „képzeld”, hogy a műveleti összefüggések megmaradnak.

Lehetséges volna a polinomokat úgy értelmezni, ahogyan azokat elképzeljük. A minél pontosabb bevezetésnél viszont itt is annak az általános algebrai elvnek az alkalmazása a célszerű, amit a komplex számoknál már láttunk. Előbb adjuk meg egyszerre az összes szóba jövő kifejezést, ezekre értelmezünk műveleteket úgy, ahogy szeretnénk, — majd ezután igazoljuk, hogy a megadott „valamik” között ott van a kívánt határozatlan is — sőt mi több, ugyanúgy lehet vele számolni, mint ahogyan akartuk. Látható, hogy szinte ez az egyetlen célravezető módszer. Ha ugyanis a fenti típusú valamik léteznek, és rendelkeznek a kívánt tulajdonságokkal, akkor egyszerre is megadhatók. Ez az eljárás tehát célhoz vezet. (Megjegyezzük, hogy más módszert is alkalmazhatunk. Ha valamikről eleve tudjuk, hogy úgy viselkednek, mint a kérdéses kifejezések, akkor egyszerűen ezeket adjuk meg — mint egy modellt. Így például a polinomok bevezetésére alkalmas lenne bizonyos speciális végtelen négyzetes mátrixokat tekinteni, a mátrixműveletekre mint műveletekre nézve.)

A polinomok tárgyalásánál először figyelembe kell vennünk, hogy milyen számok az együtthatók. A számunkra legfontosabb eset az, amikor az együtthatókat valamilyen test szerinti (szám)gyűrűből vehetjük. Az x „határozatlan” polinomjai tehát e (szám)gyűrűből való együtthatókkal adhatók meg. Vagyis ezek a polinomok úgy tekinthetők, mint a (szám)gyűrűből vett elemek egy sorozata. Erről a sorozatról nem tehetjük fel, hogy adott számú elemből áll, mert e polinomok akármilyen „hosszúak” lehetnek. Ezért tehát célszerű a polinomokat végtelen sorozatokkal jellemezni. E végtelen sorozatok azonban nem lehetnek ténylegesen végtelen hosszúak, mert a behelyettesítés után kapott végtelen összegeknek általában nincs értelmük. Ezen úgy segíthetünk, hogy kikötjük: a sorozat elemei valahonnan kezdve mind 0-val egyenlőek. (Egy olyan „végtelen” összeg, amelyben csak véges sok tag különbözik 0-tól, már könnyen meghatározható, ez megegyezik a 0-tól különböző tagok összegével.)

A továbbiakban meg kell figyelni, hogy a határozatlanra valóban semmiféle megkötés nem állhat-e fenn, azaz két polinom csak úgy egyezhet-e meg, ha a megfelelő együtthatók is megegyeznek. (Különben a határozatlan „gyöke” volna a két polinom „különbségének”, vagyis nem írhatnánk helyébe akármilyen számot.) Azt is meg kell néznünk, hogy miképpen végezhetjük a polinomokkal a műveleteket — figyelembe véve a számokra vonatkozó műveleti azonosságokat. Mindezek ismeretében definiálhatjuk a műveleteket általában. (A műveletek definíciójának a célszerűségét nem mutatjuk meg, de meg lehet vizsgálni, hogy a műveleteket valóban ennek az alapján fogjuk definiálni.)

Avégett, hogy a polinomokat a szereplő számgyűrű elemeitől jól megkülönböztethessük, a bevezetésüknél a polinomokat vastagított betűkkel fogjuk jelölni.

3.1. Definíció. R (szám)gyűrűbeli együtthatós vagy R feletti egyhatározatlanú polinomoknak nevezzük az R elemeiből képezett olyan

$$\mathbf{f} = (a_0, a_1, \dots, a_n, \dots)$$

végtelen sorozatokat, amelyekben csak véges sok 0-tól különböző szám szerepel.

A sorozatban szereplő számok a polinom együtthatói.

A fenti \mathbf{f} és $\mathbf{g} = (b_0, b_1, \dots, b_n, \dots)$ R -beli együtthatós polinomokat akkor tekintjük egyenlőeknek, ha minden i indexre érvényes az $a_i = b_i$ összefüggés.

Az R -beli együtthatós polinomokra az alábbi műveleteket értelmezzük:

1. Ha $a \in R$ és $\mathbf{f} = (a_0, \dots, a_s, \dots)$ egy R feletti polinom, akkor ezek szorzata az $a\mathbf{f} = (a \cdot a_0, \dots, a \cdot a_s, \dots)$ polinom.
2. Ha $\mathbf{f} = (a_0, \dots, a_s, \dots)$ és $\mathbf{g} = (b_0, \dots, b_s, \dots)$ két R -beli együtthatós polinom, akkor ezek $\mathbf{u} = \mathbf{f} + \mathbf{g} = (c_0, \dots, c_s, \dots)$ összegét a $c_i = a_i + b_i$ összefüggéssel definiáljuk, ahol i végigfut a nemnegatív egész számokon.
3. A (2) alatt adott \mathbf{f} és \mathbf{g} polinomok $\mathbf{v} = \mathbf{f} \cdot \mathbf{g} = (d_0, d_1, \dots, d_r, \dots)$ szorzatát a

$$d_i = \sum_{j=0}^i a_j b_{i-j} = a_0 b_i + a_1 b_{i-1} + \dots + a_i b_0$$

összefüggéssel definiáljuk, ahol i végigfut a nemnegatív egész számokon. ■

Most és a továbbiakban feltesszük, hogy az 1 szám eleme az R gyűrűnek.

Mindenekelőtt az R -beli együtthatós polinomokra vonatkozó műveleti azonosságokat állapítjuk meg:

3.1. Tétel. Az R feletti polinomokra érvényesek a következő azonosságok:

I. Az összeadásra vonatkozó azonosságok:

- (1) Az összeadás kommutatív.
- (2) Az összeadás asszociatív.
- (3) Elvégezhető a kivonás (egyértelműen).

II. A szorzásra vonatkozó azonosságok:

- (1) A szorzás kommutatív.
- (2) A szorzás asszociatív.

III. A műveleteket összekötő azonosságok:

- (1) A szorzás az összeadásra nézve disztributív.
- (2) $a(\mathbf{f} + \mathbf{g}) = a\mathbf{f} + a\mathbf{g}$,
- (3) $(a + b)\mathbf{f} = a\mathbf{f} + b\mathbf{f}$,
- (4) $(ab)\mathbf{f} = a(b\mathbf{f})$,
- (5) $1\mathbf{f} = \mathbf{f}$,
- (6) $a(\mathbf{f}\mathbf{g}) = (a\mathbf{f})\mathbf{g} = \mathbf{f}(a\mathbf{g})$,

ahol \mathbf{f} és \mathbf{g} tetszőleges R -beli együtthatós polinomok, $a, b \in R$, és 1 az „egy” szám.

Bizonyítás. A polinomok egyenlőségének a definíciója alapján azt kell megmutatnunk, hogy a megfelelő egyenlőségek minden együtthatóra érvényesek.

Legyenek \mathbf{f} , \mathbf{g} és \mathbf{h} adott polinomok, amelyeknek az $(n+1)$ -edik helyen álló együtthatói rendre a_n , b_n és c_n . Ezeket fogjuk használni a bizonyításnál.

I.(1) $\mathbf{f} + \mathbf{g}$ és $\mathbf{g} + \mathbf{f}$ megfelelő együtthatói $a_n + b_n$ és $b_n + a_n$, amelyek a számok összeadásának a kommutativitása alapján megegyeznek.

I.(2) $(\mathbf{f} + \mathbf{g}) + \mathbf{h}$ és $\mathbf{f} + (\mathbf{g} + \mathbf{h})$ megfelelő együtthatói $(a_n + b_n) + c_n$ és $a_n + (b_n + c_n)$, amelyek ugyancsak megegyeznek a számokra vonatkozó asszociativitás alapján.

I.(3) $\mathbf{f} - \mathbf{g}$ definíció szerint az a polinom, amelyet \mathbf{g} -hez adva \mathbf{f} -et kapunk. E különbségpolinomnak az $(n+1)$ -edik együtthatója csak $a_n - b_n$ lehet; és ez nyilván meg is felel a feltételnek. Jegyezzük meg azt is, hogy az a \mathbf{o} polinom, amelynek minden együtthatója 0, rendelkezik azzal a tulajdonsággal, hogy $\mathbf{o} + \mathbf{f} = \mathbf{f}$, tetszőleges \mathbf{f} polinomra. Létezik még minden \mathbf{f} polinomnak egy egyértelmű negatívja, amelynek $(n+1)$ -edik együtthatója éppen $-a_n$.

II.(1) A megfelelő együtthatók egyenlősége azt jelenti, hogy

$$\sum_{i=0}^n a_i b_{n-i} = \sum_{i=0}^n b_i a_{n-i}.$$

A második összegben az összegzést i helyett $(n-i)$ -re végezve és a tényezőket felcserélve éppen a bal oldali összeget kapjuk.

II.(2) Legyen a három vizsgált polinom \mathbf{f} , \mathbf{g} és \mathbf{h} . Azt kell megmutatni, hogy az $(\mathbf{fg})\mathbf{h}$ és az $\mathbf{f}(\mathbf{gh})$ szorzat megegyezik. A megfelelő együtthatók itt sokkal bonyolultabban írhatók csak fel:

$$\sum_i \left(\sum_j a_j \cdot b_{i-j} \right) c_{n-i}, \quad \text{illetve} \quad \sum_i a_i \left(\sum_j b_j \cdot c_{n-i-j} \right).$$

Itt mind a bal, mind a jobb oldalon pontosan az összes olyan $a_i \cdot b_j \cdot c_k$ alakú szorzat szerepel, amelyekre $i+j+k=n$. Így a megfelelő együtthatók megegyeznek, tehát egyenlő a két polinom is.

III.(1) Az $(\mathbf{f} + \mathbf{g})\mathbf{h}$ és $\mathbf{fh} + \mathbf{gh}$ polinomok $(n+1)$ -edik együtthatója:

$$\sum_i (a_i + b_i) c_{n-i}, \quad \text{illetve} \quad \sum_i a_i c_{n-i} + \sum_i b_i c_{n-i}.$$

Ezek pedig a számokra vonatkozó disztributivitás alapján egyenlőek.

A III. (2)-től (5)-ig terjedő azonosságok érvényessége rendre az $a(a_n + b_n) = aa_n + ab_n$, $(a+b)a_n = aa_n + ba_n$, $(ab)a_n = a(ba_n)$ és $1 \cdot a_n = a_n$ egyenlőségekből következik.

Végül a III.(6) összefüggés az

$$a \left(\sum_i a_i b_{n-i} \right) = \sum_i (aa_i) b_{n-i} = \sum_i a_i (ab_{n-i})$$

egyenlőségekből nyerhető. ■

Formálisan tekintve, az R -beli együtthatós polinomok között nem szerepelnek az R -beli számok. „Valójában” viszont világos, hogy a „konstans polinomok” R -beli számoknak is tekinthetők. Ezt mondja ki precízen az alábbi

3.2. Tétel. Az $a \mapsto \mathbf{f}_a = (a, 0, \dots, 0, \dots)$ megfeleltetés egyértelműen és művelettartóan képezi le az R számgyűrűt az R -beli együtthatós polinomok gyűrűjébe (tehát ez egy injektív homomorfizmus), továbbá bármely \mathbf{f} R -beli együtthatós polinomra $a\mathbf{f} = \mathbf{f}_a \cdot \mathbf{f}$.

Mindezek alapján úgy tekinthetjük, hogy az R -beli együtthatós polinomok között ott vannak az R -beli számok is.

Bizonyítás. A megfeleltetés egyértelműsége nyilvánvaló, a művelettartás pedig következik a polinomok közötti összeadás és szorzás definíciójából. A polinomok szorzásából az utoljára említett azonosság is azonnal adódik. (A megfelelő együtthatók kiszámítását az olvasóra bízjuk.)

Ezek szerint az \mathbf{f}_a alakú polinomokat úgy tekinthetjük, mint az R -beli számokat. Sőt, még az is érvényes, hogy egy ilyen polinommal való szorzás ugyanazt adja, mint a megfelelő számmal való szorzás. Ez azt jelenti, hogy az R elemei helyébe a megfelelő polinomokat írva be a 3.1. Tétel összes azonossága is érvényben marad (csak azokat kell megnézni, amelyekben számok is szerepelnek): III.(2) és III.(3) a disztributivitásból, III.(4) és III.(6) az asszociativitásból és a kommutativitásból következik, a III.(5) pedig triviális. Ez jelenti pontosan azt, hogy úgy tekinthetjük, hogy az \mathbf{f}_a polinom helyén az a szám áll, vagyis a polinomok között ott vannak az R elemei is. ■

Megjegyzés. Tekintettel arra, hogy a számokat is polinomoknak tekinthetjük, ezért a továbbiakban már nincs szükség arra, hogy a polinomokat vastagított betűkkel jelöljük. □

3.3. Tétel. Jelölje x azt a polinomot, amelyben a második együttható 1, a többi pedig 0, azaz $x = (0, 1, 0, \dots, 0, \dots)$. Ekkor az $f = (a_0, a_1, \dots, a_n, \dots)$ polinom felírható az

$$f = a_0 + a_1 \cdot x + \dots + a_n x^n + \dots$$

alakban, ahol az a_i szám a 3.2. Tételben megadott polinomnak felel meg, a műveletek a polinomokra értelmezett műveletek.

A kapott polinomokat az x határozatlan R -beli együtthatós polinomjainak nevezzük.

(x hatványait a szokásos módon rekurzívan definiáljuk: $x^0 = 1$, $x^1 = x$, \dots , $x^{k+1} = x^k \cdot x$.)

Bizonyítás. A polinomok összeadásának a definíciója szerint minden polinom olyan polinomoknak az összegeként írható fel, amelyekben egyetlenegy 0-tól különböző együttható szerepel — hiszen a 0-tól különböző együtthatók száma véges. Így az előbbi f polinom $(0, \dots, 0, a_i, 0, \dots)$ alakú polinomok összege, ahol az a_i szám az $(i+1)$ -edik együttható. Ezt a polinomot is tovább alakíthatjuk, és $a_i \cdot (0, \dots, 0, 1, 0, \dots) = a_i \cdot g_i$ alakba írhatjuk, ahol g_i -ben az egyetlen 1-es az $(i+1)$ -edik helyen áll. Megmutatjuk, hogy g_i éppen az adott x polinom i -edik hatványa. Ezt az i -re vonatkozó teljes indukcióval bizonyítjuk.

Az $i = 0$ esetben a szereplő g_i polinomra $g_0 = 1 = x^0$ a 3.2. Tételbeli megfeleltetés és x^n definíciója szerint. $g_1 = x^1$ definíció szerint teljesül. Tegyük fel, hogy $x^j = g_j$. A hatványozás azonosságai alapján ekkor $x^{j+1} = x \cdot x^j = x \cdot g_j$. A 3.1. Definícióban a szorzásnál használt polinomokra legyen $f = x$, $g = g_j$ és $v = x^{j+1}$. Ekkor $a_i = 0$, ha $i \neq 1$ és $a_1 = 1$, továbbá $b_i = 0$, ha $i \neq j$ és $b_j = 1$. Ebből $d_n = 0$ következik, kivéve az $n = j+1$ esetet, amikor az összeg tagjai között fellép az egyetlen nem 0 szorzat: $a_1 \cdot b_j = 1 \cdot 1 = 1$. Tehát $d_{j+1} = 1$, vagyis $x^{j+1} = g_{j+1}$. ■

3.2. Definíció. Az R számgyűrűbeli együtthatós polinomokat a definiált összeadásra és szorzásra nézve R feletti polinomgyűrűnek nevezzük.

Az $x = (0, 1, 0, \dots, 0, \dots)$ jelölés esetén e polinomgyűrűre az $R[x]$ jelölést használjuk. ■

3.3. Definíció. Az $R[x]$ -beli

$$f = a_0 + a_1x + \cdots + a_i x^i + \cdots + a_n x^n$$

polinom esetén az a_i neve az i -edfokú tag együtthatója. A 0-adfokú tag együtthatója a polinom konstans tagja.

Ha a_n különbözik 0-tól, akkor ezt a számot a polinom főegyütthatójának nevezzük. Ebben az esetben azt mondjuk, hogy n a polinom foka.

A polinom fokára a $\text{gr}(f)$ jelölést fogjuk használni. (Szokásos még a $\deg(f)$ jelölés is.)

Ha a polinomnak nincs 0-tól különböző együtthatója, akkor a polinomnak nem definiálunk fokot.

Ha $\text{gr}(f) = 0$, akkor a polinomot konstansnak vagy konstans polinomnak nevezzük.

Ha $\text{gr}(f) = 1$, akkor azt mondjuk, hogy a polinom elsőfokú vagy lineáris.

Ha a polinom főegyütthatója 1, akkor normált polinomról beszélünk. ■

3.4. Tétel. Legyen $f, g \in R[x]$.

- (1) Ha mindkettő különbözik 0-tól és $\text{gr}(f) \geq \text{gr}(g)$, akkor vagy $f + g = 0$, vagy $\text{gr}(f + g) \leq \text{gr}(f)$. A $\text{gr}(f) \neq \text{gr}(g)$ esetben igaz a $\text{gr}(f + g) = \text{gr}(f)$ összefüggés.
- (2) $f \cdot g = 0$ akkor és csak akkor teljesül, ha f és g közül legalább az egyik 0. Ellenkező esetben $\text{gr}(f \cdot g) = \text{gr}(f) + \text{gr}(g)$. Ekkor a szorzat főegyütthatója a tényezők főegyütthatóinak a szorzata. Speciálisan, ha mindkét polinom normált, akkor szorzatuk is az.

Bizonyítás. Legyen először mindkét polinom 0-tól különböző, és legyen $\text{gr}(f) = n \geq k = \text{gr}(g)$. Más szóval a polinomok

$$f = a_0 + \cdots + a_n x^n \quad \text{és} \quad g = b_0 + \cdots + b_k x^k$$

alakban írhatók, ahol mind a_n , mind b_k különbözik 0-tól. Az $n = k$ esetben $f + g = (a_0 + b_0) + \cdots + (a_n + b_n)x^n$. Ha ez a polinom 0-tól különbözik, akkor együtthatói között is van 0-tól különböző. Az n -nél magasabb fokú tagok együtthatói mind 0-val egyenlőek, ezért a polinom foka legfeljebb n , ahogy állítottuk. Ha viszont $k < n$, akkor az összeg főegyütthatója biztosan a_n lesz, ami különbözik 0-tól. Ezzel az (1) alatti állítást igazoltuk.

(2) bizonyításához először is megjegyezzük, hogy ha a két polinom közül valamelyik 0, akkor a szorzat is az. Valóban, ekkor — figyelembe véve a szorzás definícióját — e szorzatban fellépő kéttényezős szorzatok mindegyikében legalább az egyik tényező 0, s így e szorzatok maguk 0-val egyenlőek. Ekkor viszont ezek összege is csak 0 lehet, tehát a szorzatpolinom minden együtthatója 0, ami éppen azt jelenti, hogy a szorzatpolinom is 0. Ha a két polinom egyike sem 0, akkor a bizonyítás elején megadott alakban felírva, a_n és b_k éppen e polinomok főegyütthatói. Ekkor a disztributivitás és az $x^i \cdot x^j = x^{i+j}$ összefüggés felhasználásával e két polinom szorzata

$$f \cdot g = a_0 b_0 + \cdots + (a_n b_k) x^{n+k}$$

alakú lesz, ahol a szorzat fel nem írt tagjaiban x kitevője $(n+k)$ -nál kisebb és pozitív. (Persze egy-egy x -hatvány együtthatója több kéttényezős szorzat összege lesz.) A főegyütthatók

0-tól különböznek, ezért szorzatuk is 0-tól különböző. Egyrészt tehát a szorzat főegyütthatója éppen az $(n+k)$ -adfokú tag együtthatója, vagyis a szorzat foka megegyezik a tényezők fokának az összegével. Másrészt azt is kaptuk, hogy a szorzat főegyütthatója a tényezők főegyütthatóinak a szorzata, amiből viszont azonnal következik, hogy normált polinomok szorzata is normált polinom. ■

A 3.2. Definícióban szerepelt a *polinomgyűrű* elnevezés. A számgyűrűkkel már a komplex számok bevezetésénél találkoztunk. A „gyűrű” szó arra utal, hogy a vizsgált rendszer elemeire milyen műveleteket értelmeltünk; és ezekre a műveletekre milyen azonosságok teljesülnek. Mindkét esetben két művelet szerepelt, amelyeket *összeadásnak* és *szorzásnak* nevezünk. Az ezekre vonatkozó azonosságokat a 3.1. Tételben soroltuk fel. Ezek szerint:

az összeadás is és a szorzás is kommutatív és asszociatív; elvégezhető a kivonás, és a szorzás az összeadásra nézve disztributív.

Ha egy ilyen tulajdonságú rendszerünk van, akkor általában *gyűrűről* beszélünk. Hasonló azonosságoknak tesznek eleget az azonos alakú négyzetes mátrixok is; de azok körében a szorzás nem kommutatív. Tekintettel arra, hogy a mátrixok is igen fontos szerepet játszanak a matematikában, ezért a gyűrűk esetében nem teszik fel a szorzás kommutativitását. Nagyon sok esetben viszont mégis fennáll a kommutativitás is. Ezt hangsúlyozandó *kommutatív gyűrűről* beszélünk. A mátrixok esetében előfordulhatott, hogy két négyzetes mátrix szorzata annak ellenére 0 volt, hogy a tényezők egyike sem volt az. Ilyen esetben *nullosztópárról* beszélünk. A polinomoknál láttuk, hogy ez nem fordulhatott elő; és a számoknál sem. Ha a gyűrűben nincsenek nullosztópárok, akkor *nullosztómentes* gyűrűről beszélünk. Ha a gyűrű kommutatív és nullosztómentes, akkor „majdnem” olyan, mint az egész számok gyűrűje; és a neve *integritási tartomány* (a latin *integer* szóból, amelynek jelentése *egész*).

A páros számokra is teljesülnek a fenti azonosságok, de ennek az integritási tartománynak sok kellemetlen, „furcsa” tulajdonsága van. Ennek az az oka, hogy a páros számok között nincs ott az 1. Ha a gyűrűben van az 1 tulajdonságaival rendelkező elem (azaz olyan 1, amelyre $1 \cdot x = x$ teljesül bármely gyűrűbeli x esetén), akkor egységelemes gyűrűről beszélünk.

Igen fontos számgyűrűk a számtestek. Ennek fontos általánosításai a *kommutatív testek*, amelyek olyan (egységelemes) integritási tartományok, amelyekben a nemnulla elemekkel osztani is lehet.

A 3.1. Tétel kiegészítése. *Ha R tetszőleges egységelemes integritási tartomány, akkor $R[x]$ is az.*

A bizonyítás ugyanúgy történik, mint a 3.1. Tételé. Azt kell megfigyelni, hogy ott az R -nek csak azokat a tulajdonságait használtuk ki, amelyek minden egységelemes integritási tartományban igazak. ■

2. Maradékos osztás és oszthatóság

Tudjuk, hogy az egész számok körében elvégezhető a maradékos osztás. Ennek a segítségével az egész számoknak igen sok fontos tulajdonságát lehet bizonyítani. Ott „durván megfogalmazva” az érvényes, hogy a maradék abszolút értéke mindig kisebb lesz, mint az osztóé. Egy K (szám)testbeli együtthatós polinomok esetében is hasonló tulajdonságot fogunk igazolni, de itt az abszolút érték szerepét a polinom foka veszi át.

3.5. Tétel. *Legyen K tetszőleges kommutatív test (számtest). A $K[x]$ polinomgyűrű tetszőleges f eleméhez és g 0-tól különböző eleméhez található olyan ugyancsak $K[x]$ -beli q és r polinomok, amelyekre:*

- (1) $f = q \cdot g + r$, ahol
- (2) vagy $r = 0$, vagy $\text{gr}(r) < \text{gr}(g)$. Ezt az eljárást maradékos osztásnak nevezzük.

Kiegészítés. Az f és g polinomok a q és r polinomokat egyértelműen meghatározzák.

Bizonyítás. Az f fokára vonatkozó teljes indukcióval bizonyítunk. Az $f = 0$, illetve a $\text{gr}(f) < \text{gr}(g)$ esetben a $q = 0$ és $r = f$ választással a kívánt polinomokat már meg is kaptuk.

Tegyük fel, hogy $\text{gr}(f) = n \geq \text{gr}(g)$, és f helyébe bármely n -nél alacsonyabb fokú polinomot írva találhatunk a követelményeknek megfelelő polinomokat. Feltevésünk alapján g -nek létezik foka; legyen ez k . A fokokra vonatkozó feltevés szerint $n - k$ nemnegatív, és így létezik az x^{n-k} polinom, amely természetesen ugyancsak $K[x]$ eleme. Ekkor $K[x]$ -ben van a $g_1 = x^{n-k} \cdot g$ polinom is, amelynek a foka — a 3.4. Tétel (2) pontja szerint — pontosan n . Legyen a az f -nek és b a g -nek, és így g_1 -nek a főegyütthatója. Ekkor a $g_2 = \frac{a}{b} \cdot g_1$ polinom főegyütthatója ugyancsak a lesz; továbbá e polinom foka is n . Így az $f_1 = f - g_2$ polinom n -edfokú tagjának az együtthatója 0, azaz e polinom foka n -nél kisebb. A teljes indukciós feltevés szerint tehát található olyan q_1 és r_1 $K[x]$ -beli polinomok, amelyekre teljesül az $f_1 = q_1 \cdot g + r_1$ összefüggés, továbbá vagy $r_1 = 0$, vagy $\text{gr}(r_1) < \text{gr}(g)$.

Ekkor a $q = \frac{a}{b} \cdot x^{n-k} + q_1$ és az $r = r_1$ polinomok megfelelnek a kívánt feltételnek. Valóban, az r definíciója alapján (2) teljesül, mert ez a feltétel r_1 -re teljesül. Tehát csak (1) bizonyítása van hátra. Azonban

$$q \cdot g + r = g_2 + q_1 \cdot g + r_1 = g_2 + f_1 = g_2 + (f - g_2) = f$$

bizonyítja ezt az összefüggést is.

A kiegészítés bizonyítása végett tegyük fel, hogy a feltételeknek mind a q_1, r_1 , mind a q_2, r_2 pár megfelel. Ebből azonnal következik a

$$q_1 g + r_1 = q_2 g + r_2, \quad \text{illetve a} \quad (q_1 - q_2)g = r_1 - r_2$$

egyenlőség. A második egyenlőség jobb oldalán álló polinom a 3.4. Tétel (1) pontja alapján vagy a 0 polinom, vagy az r_i -kre vonatkozó feltétel szerint foka kisebb, mint a g foka.

Használjuk most fel a 3.4. Tétel (2) pontját. Eszerint az egyenlőség bal oldalán álló polinom vagy 0, vagy a foka legalább akkora, mint a g polinom foka. Ez utóbbi lehetőség azonban ellentmond a jobb oldali polinom fokáról megállapítotttnak, ezért a szorzat csak 0 lehet. A maradékos osztás feltételei szerint $g \neq 0$, ezért $q_1 - q_2 = 0$. Mivel a bal oldalon 0 áll, ezért a jobb oldalon is 0-nak kell állnia: $r_1 - r_2 = 0$. Ez pedig valóban a q és az r egyértelműségét jelenti. ■

Megjegyzés. Ha a K test helyett egy R gyűrűt tekintünk, akkor is bizonyíthatunk a 3.5. Tétellel analóg tételt. Ekkor a g polinomról azt célszerű feltenni, hogy normált (ebből $g \neq 0$ is következik). A maradékos osztás egyébként pontosan akkor végezhető el, ha g főegyütthatója minden R -beli elemnek „osztója” (R -ben léteznek a megfelelő hányadosok.) □

A továbbiakban a polinomokra vonatkozó oszthatóságot vizsgáljuk. Ezt a fogalmat úgy lehetne értelmezni, hogy a 3.5. Tételben szereplő maradékos osztásnál legyen a maradék 0. Ezzel azonban bizonyos fokú korlátozást kapnánk, ugyanis ott az „osztó” csak 0-tól különböző lehetett. Márpedig az oszthatóságnál nem csak meg lehet, de célszerű is megengedni a 0-t is.

3.4. Definíció. Legyen $f, g \in K[x]$. Azt mondjuk, hogy f osztója g -nek (g többszöröse f -nek) — jelben $f \mid g$ —, ha létezik olyan $h \in K[x]$, amelyre $g = f \cdot h$. ■

Megjegyzés. Vegyük észre, hogy ennél a definíciónál nem csak azt nem használtuk ki, hogy K test, hanem azt sem, hogy polinomokról van szó. Világos, hogy ez a fogalom tetszőleges integritási tartományban definiálható. □

3.6. Tétel. *Ha a 3.4. Definícióban szereplő $g \neq 0$ és $f = 0$, akkor ilyen h nem létezik. Ha $f = g = 0$, akkor minden h polinom megfelel. Egyébként, ha van ilyen h polinom, akkor egyértelmű, olyannyira, hogy ha egy K -nál bővebb L testbeli együtthatós h polinomot találunk, akkor ez megegyezik az itteni polinommal.*

Bizonyítás. Az első két állítás nyilvánvalóan igaz. Tegyük most fel, hogy $f \neq 0$, és egy K -nál esetleg bővebb L testben létezik a feltételnek eleget tevő h polinom. $K \subseteq L$ miatt $f, g \in L[x]$ is igaz. Ebben a polinomgyűrűben tehát teljesül a $g = f \cdot h + 0$ felírás; amely megfelel az $L[x]$ -beli maradékos osztásnak. Mivel $f \neq 0$, a 3.5. Tétel szerint léteznek olyan $q, r \in K[x]$ polinomok, amelyekre $g = f \cdot q + r$, továbbá vagy $r = 0$, vagy $\text{gr}(r) < \text{gr}(f)$. Ez a felírás $L(x)$ -ben is tekinthető a maradékos osztás felírásának. A 3.5. Tétel kiegészítése szerint a maradékos osztás $L[x]$ -ben is egyértelmű, amiből $h = q$ és $r = 0$ következik. $q \in K[x]$ bizonyítja, hogy a „hányados” valóban $K[x]$ -ben van. ■

3.7. Tétel. *Legyenek $f, g, h \in K[x]$ és $c, d \in K$ ($c \neq 0, d \neq 0$). A $K[x]$ -beli oszthatóság rendelkezik az alábbi tulajdonságokkal:*

- (1) *Reflexív, azaz $f \mid f$.*
- (2) *Legyen $f_1 = cf$ és $g_1 = dg$. Ekkor $f_1 \mid g_1$ és $f \mid g$ ekvivalensek (ilyen esetben azt mondjuk, hogy f_1 és f , illetve g_1 és g asszociáltak).*

- (3) Az oszthatóság antiszimmetrikus abban az értelemben, hogy $f \mid g$ és $g \mid f$ együttes teljesülése esetén e polinomok egymás asszociáltjai, illetve ha mindkettő normált polinom, akkor egyenlők.
- (4) Az oszthatóság tranzitív, azaz $f \mid g$ és $g \mid h$ esetén $f \mid h$.
- (5) Ha $f \mid g$, akkor bármely h -ra $f \mid (gh)$.
- (6) Ha $f \mid g$ és $f \mid h$, akkor $f \mid (g + h)$ és $f \mid (g - h)$.

Bizonyítás. (1) azonnal következik abból, hogy $1 \cdot f = f$.

(2)-nél feltétel szerint létezik c^{-1} és d^{-1} , amiből $f = c^{-1} \cdot f_1$ és $g = d^{-1} \cdot g_1$ következik. Így az ekvivalenciához elég annak a bizonyítása, hogy $f \mid g$ esetén $f_1 \mid g_1$ is igaz. Legyen tehát $g = f \cdot h$, ekkor $g_1 = f_1 \cdot (c^{-1} \cdot d \cdot h)$, tehát a másik oszthatóság is teljesül.

(3)-ban feltétel szerint alkalmas $u, v \in K[x]$ polinomokra $f = g \cdot u$ és $g = f \cdot v$, amiből azt kapjuk, hogy $f = f \cdot (vu)$ és $g = g \cdot (uv)$. Az $f = g = 0$ esetben ezek természetes asszociáltak. Ha ezek bármelyike nem 0, akkor a szorzat fokára vonatkozó összefüggés (3.4. Tétel (2)) szerint uv foka 0, és ugyanannak alapján mindkettő konstans. Így f és g valóban asszociáltak.

(4)-ben feltételünk alapján vannak olyan $u, v \in K[x]$ polinomok, amelyekre $g = fu$ és $h = gv$, és így $h = f(uv)$, tehát $f \mid h$.

Ebből azonnal következik (5), hiszen $g \mid (gh)$.

(6)-nál feltétel szerint léteznek olyan $u, v \in K[x]$ polinomok, amelyekre $g = fu$ és $h = fv$. Ebből azt kapjuk, hogy $g + h = f(u + v)$ és $g - h = f(u - v)$; ami bizonyítja a kívánt oszthatóságot. ■

3.8. Tétel. Egy polinom akkor és csak akkor osztója minden polinomnak, ha osztója 1-nek. Ez akkor és csak akkor teljesül, ha a polinom egy nemnulla konstans.

A 0 polinom egyedül a 0 polinomnak osztója. A 0 polinomnak minden polinom osztója.

Bizonyítás. Ha egy polinom osztója minden polinomnak, akkor osztója speciálisan az 1-nek is. Ha egy polinom osztója az 1-nek, akkor a 3.4. Tétel (2) pontja szerint csak 0-tól különböző konstans lehet. Amennyiben a polinom egy 0-tól különböző a konstans, akkor az $f = a \cdot (a^{-1}f)$ összefüggés alapján osztója minden polinomnak. Ezzel az első három tulajdonság ekvivalenciáját kimutattuk.

Ha $f = 0 \cdot g$, akkor $f = 0$, tehát a 0 polinom valóban csak önmagának lehet osztója. Végül $0 = 0 \cdot f$ bizonyítja az utolsó állítást. ■

Az alábbiakban az oszthatóság szempontjából speciálisan viselkedő, de igen fontos polinomtípust értelmezzünk. Szorítkozunk egyelőre csak normált polinomokra. Ekkor az 1 polinomnak önmagán kívül nincs osztója, tehát osztóinak a száma 1. Minden más polinomnak legalább két osztója van, nevezetesen 1 és önmaga. Azok a polinomok játszanak szempontunkból fontos szerepet, amelyeknek nincs is több osztójuk. E polinomok — láthatóan — a természetes számok körében a prímszámoknak felelnek meg.

3.5. Definíció. Egy polinomot a K számtest felett felbonthatatlannak (irreducibilisnek) nevezünk, ha nem bontható fel két nála alacsonyabb fokú $K[x]$ -beli polinom szorzatára, és a polinom nem konstans. ■

Megjegyzés. Elsőfokú polinom mindig irreducibilis, hiszen az 1-nél alacsonyabb fokú polinom konstans, és konstansok szorzata nem lehet elsőfokú. □

3.6. Definíció. Azt mondjuk, hogy a nem konstans $f \in K[x]$ polinom — a K számtest felett — rendelkezik a prímtulajdonsággal, ha bármely $g, h \in K[x]$ polinomok esetében az $f \mid (gh)$ feltételből következik, hogy vagy $f \mid g$, vagy $f \mid h$. ■

Megjegyzés. Mindkét definíciónál fontos, hogy egy adott K számtestre vonatkozik. Ugyanis például az $x^2 - 2$ polinomról megmutatható, hogy a racionális számtest felett irreducibilis és rendelkezik a prímtulajdonsággal is, a valós számtest felett pedig egyik állítás sem igaz e polinomra. □

3.9. Tétel. Rögzített K számtest felett az alábbi állítások ekvivalensek:

- (1) Az f polinom irreducibilis.
- (2) Ha $f = uv$ nem konstans, akkor u és v valamelyike konstans.
- (3) Ha $f = uv$ nem konstans, akkor u és v valamelyike f -nek asszociáltja.

Bizonyítás. Tegyük fel, hogy f irreducibilis és legyen $f = u \cdot v$. Feltétel szerint u és v valamelyikének a foka megegyezik f fokával (nagyobb nem lehet), ezért a másik tényező 0-adfokú. Tehát (1)-ből következik (2).

Ha (2) teljesül az f polinomra és $f = u \cdot v$, akkor például $u = a$ konstans. f nem konstans, ezért $a \neq 0$, és így $v = a^{-1}f$. Tehát (2)-ből következik (3).

Ha f -re igaz (3) és $f = u \cdot v$, akkor a tényezők valamelyike f -nek asszociáltja, tehát nem lehet f -nél alacsonyabb fokú. ■

Kiegészítés. Minden prímtulajdonságú polinom irreducibilis.

Bizonyítás. Írjuk fel a prímtulajdonságú f polinomot $f = gh$ alakban. Ekkor f természetesen osztója a gh szorzatnak. A prímtulajdonság szerint tehát f osztója e két tényező valamelyikének — például g -nek. g eleve osztója az f -nek, ezért a 3.7. Tétel (2) pontjából azonnal adódik az itteni (3) összefüggés. ■

Megjegyzések

1. Annak a bizonyítása, hogy minden irreducibilis polinom prímtulajdonságú, sokkal hosszadalmasabb; ezt csak a következő pontban fogjuk bebizonyítani.

2. Tulajdonképpen minden 0-tól különböző konstans polinom is rendelkezik a tételben szereplő tulajdonságok mindegyikével. Ennek ellenére ezeket a polinomokat ki kellett zárni a később megfogalmazandó egyértelmű felbontási tétel érdekében. Azt is beláthatjuk, hogy a 0 polinom is rendelkezik a prímtulajdonsággal, azonban ez a polinom nem irreducibilis. □

3. Polinomideálok és a legnagyobb közös osztó

Célunk a 3.9. Tétel bizonyításának a befejezése, azaz annak a bizonyítása, hogy rögzített K számtest felett egy irreducibilis polinom mindig rendelkezik a prímtulajdonsággal.

Ennek érdekében és a bizonyítás egyszerűbbé tétele végett két fogalmat vezetünk be, a *polinomideál* és a polinomok *legnagyobb közös osztójának* a fogalmát. Ez utóbbi teljesen analóg a természetes számoknál jól ismert legnagyobb közös osztó fogalmával. Megfordítva, az előbbi fogalom megfelelőjét értelmezhetjük az egész számok esetére. Be lehet látni, hogy az itt bizonyításra kerülő tételek megfelelői igazak az egész számokra is, tehát lehetséges ugyanezekkel a módszerekkel az egész számokra vonatkozó úgynevezett „számelmélet alaptételét” bizonyítani.

3.7. Definíció. Egy $K[x]$ polinomgyűrű elemeinek valamely H részhalmazát e polinomgyűrű polinomideáljának — vagy röviden ideálnak — nevezzük, ha az alábbi tulajdonságok teljesülnek rá:

- (1) H nem üres.
- (2) ha $f, g \in H$, akkor $f - g \in H$ is fennáll.
- (3) Ha $f \in H$ és $u \in K[x]$, akkor $fu \in H$. ■

3.10. Tétel. Legyen H a $K[x]$ egy polinomideálja. Ekkor: $0 \in H$, ha $g \in H$, akkor $-g \in H$ és ha $f, g \in H$, akkor $f + g \in H$.

Az egyedül a 0-ból álló $\{0\}$ részhalmaz és $K[x]$ ideálok. Ezeket triviális ideáloknak nevezzük, a többi ideál neve valódi ideál.

Bizonyítás. A 3.7. Definíció (1) pontja szerint bármely H ideálban van egy f polinom. A (2) pont szerint $0 = f - f \in H$. Ha $g \in H$, akkor ebből ugyancsak a fenti definíció (2) pontja alapján $-g = 0 - g \in H$; amiből ugyanezt a pontot felhasználva $f + g = f - (-g) \in H$ következik.

$\{0\}$ és $K[x]$ triviálisan kielégítik a 3.7. Definícióban szereplő mindhárom feltételt. ■

3.11. Tétel. Ha $\{H_\lambda \mid \lambda \in \Lambda\}$ a $K[x]$ ideáljainak halmaza, akkor ezek H közös része is ideál. A $K[x]$ polinomjaiból álló bármely halmazhoz van egy legkisebb olyan ideál, amelyik ezek mindegyikét tartalmazza. Ezt e polinomhalmaz generálta ideálnak nevezzük.

Ha a polinomhalmaz véges: $\{f_1, \dots, f_n\}$, akkor az általuk generált ideált (f_1, \dots, f_n) jelöli. Ennek az ideálnak az elemei az $f_1 u_1 + \dots + f_n u_n$ alakú polinomok, ahol u_1, \dots, u_n a $K[x]$ tetszőleges elemei.

Az egy elemmel generálható ideált főideálnak nevezzük.

Bizonyítás. Ahhoz, hogy H ideál, azt kell ellenőrizni, hogy teljesülnek-e rá a 3.7. Definíció pontjai. Mivel minden egyes H_λ ideál, ezért 0 mindegyiknek eleme, így benne van közös részükben, azaz H -ban is. Ha $f, g \in H$, akkor mindkét polinom eleme minden

egyes H_λ -nak is. Így $f - g$ is eleme ezeknek az ideáloknak, tehát $f - g$ eleme a közös résznek. Amennyiben $f \in H$, akkor minden egyes H_λ -ban is benne van, tehát ezekben benne van fu is, bármely $u \in K[x]$ esetén, és így $fu \in H$ is igaz.

Tekintsük most a $K[x]$ -beli polinomoknak egy tetszőleges \mathcal{F} halmazát. Az \mathcal{F} halmazt biztosan tartalmazza egy ideál; nevezetesen maga $K[x]$. Vegyük most az összes olyan H_λ ideált, amely az \mathcal{F} halmazt tartalmazza. Ezeknek a H közös része természetesen tartalmazza az \mathcal{F} halmaz minden elemét. Mint láttuk, ez egy ideál, amelyik persze minden \mathcal{F} -et tartalmazó ideálban benne van; így a legkisebb \mathcal{F} -et tartalmazó ideál.

Tegyük fel, hogy $\mathcal{F} = \{f_1, \dots, f_n\}$. Az ideál definíciója és a 3.10. Tétel szerint az \mathcal{F} generálta $H = (f_1, \dots, f_n)$ ideálban benne vannak az összes $f_1u_1 + \dots + f_nu_n$ alakú polinomok, ahol u_1, \dots, u_n a $K[x]$ tetszőleges elemei. Azt kell már csak belátni, hogy ezek az elemek tartalmazzák \mathcal{F} elemeit és ideált alkotnak.

Az $u_i = 1$ és $u_j = 0$ ($j \neq i$) választással kapjuk, hogy $f_i \in H$.

Az $u_i = 0$ választással kapjuk, hogy $0 \in H$.

Ha $f = f_1u_1 + \dots + f_nu_n$ és $g = f_1v_1 + \dots + f_nv_n$, akkor $f - g = f_1(u_1 - v_1) + \dots + f_n(u_n - v_n) \in H$ alapján H zárt a kivonásra.

Végül, ha $f = f_1u_1 + \dots + f_nu_n$ és $v \in K[x]$, akkor $fv = f_1u_1v + \dots + f_nu_nv \in H$ biztosítja a harmadik feltétel teljesülését is. ■

3.12. Tétel. $K[x]$ -ben minden ideál főideál.

Bizonyítás. Legyen H a szóban forgó polinomgyűrű egy tetszőleges ideálja. Ha H egyedül a 0 polinomból áll, akkor triviálisan $H = (0)$. Tegyük fel tehát, hogy H -ban található 0-tól különböző polinom. A polinomok foka természetes szám, ezért H -ban létezik olyan d polinom, amelynek a foka minimális, vagyis H -ban d fokánál alacsonyabb fokú polinom nem létezik. Tekintsük most a H ideálnak egy tetszőleges f elemét. Mivel $d \neq 0$, a 3.5. Tétel szerint léteznek olyan q és r polinomok, amelyekre $f = dq + r$, és vagy $\text{gr}(r) < \text{gr}(d)$, vagy $r = 0$. A felírt egyenlőséget $r = f - dq$ alakba írhatjuk át. Az ideál definíciójában szereplő (3), illetve (2) tulajdonság miatt d -vel együtt dq is, továbbá dq -val és f -fel együtt $r = f - dq$ is a H ideálban van. A d -re kirótt feltétel szerint azonban $\text{gr}(r) < \text{gr}(d)$ lehetetlen, mert $r \in H$. Így csak $r = 0$ lehet, vagyis H minden eleme a d -nek többszöröse. Más szóval H minden eleme a (d) főideálban van. A 3.11. Tétel alapján tehát $H = (d)$. ■

3.8. Definíció. A $d \in K[x]$ polinomot az $f_1, \dots, f_n \in K[x]$ polinomok legnagyobb közös osztójának nevezzük, ha:

- (1) $d \mid f_1, \dots, d \mid f_n$ (azaz d közös osztó), és
- (2) ha $g \mid f_1, \dots, g \mid f_n$, akkor $g \mid d$ (azaz d oszthatóságban „legnagyobb”). ■

Megjegyezzük, hogy a definícióból nem következik eleve, hogy legnagyobb közös osztó mindig létezik.

3.13. Tétel. *A $K[x]$ polinomgyűrű bármely f_1, \dots, f_n polinomjainak létezik legnagyobb közös osztója, amely asszociáltaktól eltekintve egyértelmű. Mégpedig, ha az (f_1, \dots, f_n) ideált — mint főideált — a d polinom generálja, akkor e polinomok legnagyobb közös osztója d .*

Bizonyítás. A legnagyobb közös osztó létezéséhez természetesen elég azt bizonyítani, hogy az (f_1, \dots, f_n) ideálnak a d generátoreleme e polinomok legnagyobb közös osztója. A 3.12. Tétel alapján létezik olyan $d \in K[x]$ polinom, amelyre $(f_1, \dots, f_n) = (d)$ teljesül. Ebből $f_i \in (d)$ alapján azonnal következik, hogy d az adott polinomoknak közös osztója. Másrészt $d \in (f_1, \dots, f_n)$ alapján d felírható alkalmas u_1, \dots, u_n polinomokkal $d = f_1 u_1 + \dots + f_n u_n$ alakban. Ha most $g \mid f_i$, akkor a 3.7. Tétel (5) és (6) pontja szerint $g \mid d$ is igaz. Ezzel a legnagyobb közös osztó létezését bebizonyítottuk.

Az egyértelműség bizonyításához tegyük fel, hogy a szóban forgó polinomoknak d_1 is és d_2 is legnagyobb közös osztója. Ekkor a legnagyobb közös osztó definíciójából $d_1 \mid d_2$ is és $d_2 \mid d_1$ is következik. A 3.7. Tétel (3) pontja biztosítja, hogy ezek valóban asszociáltak. ■

Megjegyzés. Vegyük észre, hogy az egyértelműség fenti bizonyítása azt is szolgáltatja, hogy egy főideál generátoreleme asszociáltaktól eltekintve egyértelmű. □

3.9. Definíció. Az f és g polinomok legnagyobb közös osztóját $d = (f, g)$ -vel fogjuk jelölni. ■

Megjegyzés. Valójában ez a — szokásos — jelölés kétértelműségre ad lehetőséget. Az (f, g) szimbólum jelölhet ugyanis egy ideált is és egy polinomot is. „Enyhít” e kétértelműségen az a tény, hogy a szereplő polinom éppen a szereplő ideált generálja. Ennek ellenére célszerű figyelni, hogy egy-egy esetben a jelölés mit takar; ez viszont a teljes mondat értelméből általában következik. □

3.14. Tétel. *Ha $d \mid (f - g)$, akkor $(d, f) = (d, g)$.*

Bizonyítás. Mivel $d \mid (f - g)$ pontosan akkor igaz, ha $d \mid (g - f)$, ezért elég azt bizonyítani, hogy $(d, f) \subseteq (d, g)$, azaz, hogy $d, f \in (d, g)$. Ez a d polinomra eleve igaz. A feltétel szerint van olyan h polinom, amelyre $f - g = dh$, vagyis $f = g + dh$. Az ideáltulajdonságok szerint ez pedig eleme a (d, g) ideálnak. ■

Megjegyzés. A most bizonyított tétel arra szolgál, hogy két polinom legnagyobb közös osztóját konkrétan meghatározhassuk. Tekintsük az f és g polinomokat. Ha $f = g = 0$, akkor $(f, g) = 0$, és készen vagyunk. Ha $f \neq 0$ és $g = 0$, akkor $f \mid 0$ miatt $(f, g) = f$. □

Következmény: euklideszi algoritmus. *Ha $f, g \in K[x]$ nemnulla polinomok, akkor a következő algoritmussal határozhatjuk meg legnagyobb közös osztójukat: Legyen $r_0 = f$ és $r_1 = g$. Ha valamely $i > 0$ indexre az r_0, r_1, \dots, r_i polinomok már definiálva vannak és $r_i \neq 0$, akkor definiáljuk az r_{i+1} polinomot az*

$$r_{i-1} = q_i \cdot r_i + r_{i+1}$$

maradékossal. Ekkor létezik egy olyan n index, amelyre $r_{n+1} = 0$; és ebben az esetben r_n a két adott polinom legnagyobb közös osztója.

Bizonyítás. Tegyük fel, hogy az r_0, r_1, \dots, r_i nemnulla polinomokat már meghatároztuk. A maradékos osztás definíciója alapján, ha $r_{i+1} = 0$, akkor $\text{gr}(r_{i+1}) < \text{gr}(r_i)$. Ezért $\text{gr}(r_0), \text{gr}(r_1), \dots, \text{gr}(r_i), \text{gr}(r_{i+1})$ nemnegatív egész számok csökkenő sorozata, így az eljárás véges sok lépésben véget ér.

A 3.14. Tétel alapján az egymás utáni $r = r_1, r_2, \dots, r_n, r_{n+1} = 0$ maradékok sorában $(r_i, r_{i+1}) = (r_{i-1}, r_i)$ teljesül ($1 \leq i \leq n$), amiből induktilven $(r_i, r_{i+1}) = (f, g)$ következik. Mint a megjegyzésben láttuk, tetszőleges h polinomra $(h, 0) = h$, ezért valóban $r_i = (f, g)$.

Ezzel a módszerrel nemcsak meghatározható a legnagyobb közös osztó, hanem elő is állítható $d = fu + gv$ alakban. Ezt egymás után bizonyítjuk be a maradékok $r_0, r_1, r_2, \dots, r_n = d = (f, g)$ sorozatára. A definíció szerint $r_0 = f = f \cdot 1 + g \cdot 0$ és $r_1 = g = f \cdot 0 + g \cdot 1$. Az $u_0 = v_1 = 1$ és $v_0 = u_1 = 0$ választással tehát az $i \in \{0, 1\}$ esetben $r_i = f \cdot u_i + g \cdot v_i$ teljesül. Tegyük fel, hogy ez igaz minden $0 \leq i \leq j$ esetén, ha $j \geq 1$, és tekintsük az r_{j+1} polinomot. A képzési „szabály” szerint az $u_{j+1} = u_{j-1} - q_j u_j$ és $v_{j+1} = v_{j-1} - q_j v_j$ polinomokra:

$$\begin{aligned} r_{j+1} &= r_{j-1} - q_j r_j = (f u_{j-1} + g v_{j-1}) - q_j (f u_j + g v_j) = \\ &= f(u_{j-1} - q_j u_j) + g(v_{j-1} - q_j v_j) = f u_{j+1} + g v_{j+1}, \end{aligned}$$

ami az $n = j + 1$ esetre bizonyítja az előállíthatóságot. ■

3.15. Tétel. Minden irreducibilis polinom rendelkezik a prímtulajdonsággal.

Bizonyítás. Legyen a p irreducibilis polinom osztója az fg szorzatnak. Ha $p \mid f$, akkor készen vagyunk. Ha nem, akkor tekintsük a $d = (p, f)$ polinomot. Mivel $d \mid p$, ezért — p irreducibilitása alapján — d vagy konstans, vagy asszociáltja p -nek. Ez utóbbi azonban lehetetlen, hiszen ekkor $d \mid f$ miatt $p \mid f$ lenne. Ezért d konstans, és ha normált alakját tekintjük, akkor $d = 1$. Ekkor az $1 \in (p, f)$ feltételből következik olyan u és v polinomok létezése, amelyekre $1 = pu + fv$ teljesül. Ezt az egyenlőséget a g polinommal szorozva a $g = gpu + fgv$ összefüggéshez jutunk. A $p \mid fg$ feltételből az oszthatóság tulajdonságait figyelembe véve $p \mid g$ adódik, mint állítottuk. ■

4. Polinomok egyértelmű felbontása

A középfokú tanulmányoknál szereplő egyenletek gyökeinek a meghatározásánál lényeges volt a fellépő polinom foka. Ahogy a fokszám növekszik, úgy bonyolódik a megoldás megkeresése. Éppen ezért célszerű a szereplő egyenleteket minél alacsonyabb fokúakra visszavezetni. Ennek egyik módja az, hogy a polinomot tényezőkre bontjuk. Itt azonnal fellép az a kérdés, hogy vajon egyértelmű-e, illetve mennyire egyértelmű ez a felbontás. Lehet, hogy ez természetesnek tűnik, de egyáltalában nem az. A mátrixegyütthatós polinomok esetében például nem igaz a felbontás egyértelműsége (pedig a mátrixegyütthatós polinomok is fontos szerepet játszanak).

Számtestbeli együtthatós polinomokra viszont be fogjuk bizonyítani a felbonthatóságot és az egyértelműséget is. Az alábbiakban egy rögzített $K[x]$ -beli polinomokat vizsgálunk. A kapott eredmények viszont az adott testtől függetlenek, abban az értelemben, hogy az eredmény „típusa” mindig ugyanaz, csupán a felbontásban kapott polinomok változnak a testtől függően.

E pontban tehát célunk a polinomokra egy — a számelmélet alaptételével analóg — tételnek a bizonyítása.

3.16. Tétel. *A $K[x]$ -beli tetszőleges nem konstans polinomok lényegében egyértelműen bonthatók fel véges sok K felett irreducibilis polinom szorzatára.*

Megjegyzés. A számelmélet alaptételét csak a természetes számokra kimondva egy szám két felbontása persze „formailag” lehet különböző, de eltérés csak a tényezők sorrendjében lehet. Az egész számok esetében már az is megtörténhet, hogy egy-egy prímszám helyett a negatívja szerepel. A polinomoknál viszont bármilyen konstans tényező felléphet. Ettől el lehetne tekinteni, ha csak normált polinomokra szorítkoznánk. A későbbi esetekben a megfelelő tétel bizonyításánál viszont ezt már nem tudjuk megtenni. Nézzük meg tehát, mit értünk azon, hogy „lényegében egyértelmű”.

Legyen az f polinomnak két felbontása K felett irreducibilis faktorokra:

$$f = p_1 \cdot \dots \cdot p_r \quad \text{és} \quad f = q_1 \cdot \dots \cdot q_s.$$

A felbontás egyértelműségén a következőket értjük. Létezik egy olyan $\varphi : \{1, \dots, r\} \rightarrow \{1, \dots, s\}$ bijektív leképezés, amelyre a $q_{\varphi(i)}$ polinom a p_i polinomhoz asszociált. Speciálisan az is adódik, hogy mindkét felbontásban ugyanannyi tényező lép fel. Az alábbi bizonyításban érdemes megfigyelni, hogy ez az eredmény akkor is igaz, ha a második felbontásban nem az f , hanem annak egy g asszociáltja szerepel. Az is látható, hogy ha az f és g polinomok felbontásához található a fenti tulajdonságú φ bijekció, akkor ez a két polinom csakis egymás asszociáltja lehet. \square

Bizonyítás. Először azt mutatjuk meg, hogy ha $f \in K[x]$ tetszőleges nem konstans polinom, akkor felírható K felett irreducibilis polinomoknak a szorzataként. (Az „egy-tényezős” felbontást is felbontásnak nevezzük.) A bizonyítást a polinom fokára vonatkozó teljes indukcióval végezzük. Mindenekelőtt jegyezzük meg, hogy ha f irreducibilis, akkor $f = f$ azonnal megad egy felbontást. A 3.4. Tétel (2) pontjából adódik, hogy minden elsőfokú polinom irreducibilis, tehát az állítás igaz, ha $\text{gr}(f) = 1$. Legyen most $\text{gr}(f) = n$, és tegyük fel, hogy az n -nél alacsonyabb fokú polinomokra igaz az állítás. Ha f irreducibilis, akkor az előzetes megjegyzés szerint már megvan a felbontás. Amennyiben f nem irreducibilis, azaz felírható $f = gh$ alakban, ahol egyik tényező sem konstans, akkor a fokra vonatkozó összefüggésből az következik, hogy mindkét tényező foka kisebb, mint n . A teljes indukciós feltevés szerint tehát mindkét polinom felírható irreducibilis polinomok szorzataként. E két felbontásban szereplő irreducibilis faktorokat összeszorozva az f -nek egy irreducibilis faktorokra való felbontását nyerhetjük.

Megjegyezzük, hogy a tényezők száma nem lehet nagyobb a felbontott polinom fokánál. Ez is teljes indukcióval bizonyítható, felhasználva a 3.4. Tétel (2) pontját.

A következőkben az egyértelműséget vizsgáljuk.

A bizonyításhoz megjegyezzük, hogy egy irreducibilis polinomnak definíció szerint csupán egyetlen felbontása van, azaz bármely két felbontásnál a megfeleltetés ezt a polinomot felelteti meg önmagának. Ezek után az egyértelműséget is a polinom fokára vonat-

kozó teljes indukcióval bizonyítjuk. (Valójában az indukció az irreducibilis felbontásokban szereplő tényezők minimális száma szerint történik, de a fenti megfogalmazás sokkal egyszerűbb, és a bizonyítás fokszámmal is végigkövethető.)

Ha a polinom elsőfokú, akkor irreducibilis, és így az előzetes megjegyzés szerint igaz az egyértelműség. Tegyük most fel, hogy $\text{gr}(f) = n$, és az egyértelműség minden n -nél alacsonyabb fokú polinomra igaz ($n > 1$). Tekintsük f -nek a már felírt két, irreducibilis faktorokra való felbontását:

$$f = p_1 \cdot \dots \cdot p_r \quad \text{és} \quad f = q_1 \cdot \dots \cdot q_s.$$

Az első felbontás alapján $p_1 \mid f$; így p_1 osztója a második felbontásban megadott szorzatnak is. p_1 irreducibilis, tehát rendelkezik a prímtulajdonsággal. Ebből teljes indukcióval azonnal következik, hogy ha p_1 osztója egy többtényezős szorzatnak, akkor osztója a szorzat valamelyik tényezőjének is. Nem megy az általánosság rovására, ha feltesszük, hogy ez éppen az első tényező. Így $p_1 \mid q_1$. A tényezők irreducibilitása alapján $q_1 = c \cdot p_1$, ahol $c \in K$. Ha csak egyetlen tényező szerepel, akkor már készen vagyunk. Ha nem, akkor legyen $g = p_2 \cdot \dots \cdot p_r$. Ennek a polinomnak egy másik felbontása a következő: $g = (c \cdot q_2) \cdot \dots \cdot q_r$. A szorzat fokára vonatkozó összefüggés alapján $\text{gr}(g) < \text{gr}(f)$. Tehát a g polinomra igaz a felbontás egyértelműsége. Ez azt jelenti, hogy az f két felbontásában szereplő tényezőket a másodiktól kezdve megfeleltethetjük egymásnak a kívánt módon. Ezt a megfeleltetést kiterjesztjük a $q_1 = c \cdot p$ összefüggéssel az első tényezőkre is. ■

5. Polinomok kompozíciója, behelyettesítés

A polinomok körében még egy igen fontos műveletet értelmezünk, a kompozíciót. Ez nem más, mint egy polinomnak a másikba való „behelyettesítése”.

3.10. Definíció. Legyenek $f = a_0 + a_1x + \dots + a_nx^n$ és $g \in K(x)$ -beli polinomok. E polinomok $f \circ g$ kompozícióján az $a_0 + a_1 \cdot g + \dots + a_n \cdot g^n$ polinomot értjük. ■

3.17. Tétel. *Rögzített $K[x]$ -beli g polinom esetén $f \mapsto f \circ g$ a $K[x]$ -nek olyan $K[x]$ -be való homomorfizmusa (művelettartó leképezése), amelynél minden konstansnak önmaga felel meg.*

Formulával: Ha $c \in K$ és $u, v \in K[x]$, akkor

- (1) $c \circ g = c$,
- (2) $(u + v) \circ g = u \circ g + v \circ g$,
- (3) $(u \cdot v) \circ g = (u \circ g) \cdot (v \circ g)$.

Bizonyítás. Mivel f egyértelműen meghatározza az együtthatóit, ezért $f \mapsto f \circ g$ valóban egyértelmű megfeleltetés. Az (1) állítás azonnal következik a definícióból. A (2) és

(3) állítás is nyilvánvaló; minden olyan azonosság, ami az x -szel való számolásnál teljesül, igaz a g -vel való számolás esetében is. ■

Kiegészítés. Ha $f \mapsto f^*$ a $K[x]$ -nek olyan $K[x]$ -be való homomorfizmusa, amelynél minden konstansnak önmaga felel meg, akkor $f^* = f \circ g$, valamilyen alkalmas $g \in K[x]$ polinommal.

Bizonyítás. Legyen $g = x^*$. Ebből a művelettartás és a konstansra vonatkozó feltétel alapján azonnal következik az állítás. ■

Megjegyzések

1. Világos, hogy az $f \in K[x]$ polinomba nem csak $K[x]$ -beli polinomot helyettesíthetünk be, hanem tetszőleges olyan polinomot is, amelynek az együttthatói egy K -nál bővebb L testből valók. Ez azonban nem okoz gondot, hiszen eleve kiindulhatunk az L testből. Amennyiben ugyanis $K \subseteq L$, akkor $f \in L[x]$ is igaz; ezért x helyébe valóban bármelyik $L[x]$ -beli polinomot beírhatjuk.

2. Az $f \circ g$ kompozícióra — a függvénykompozíciónak megfelelően — szokásos az $f(g)$ jelölés is. Mivel $f \circ x = f$, ezért adódik a polinomokra szokásos $f(x)$ jelölés. □

3.18. Tétel. A polinomok kompozíciója asszociatív.

Bizonyítás. Azt kell megmutatni, hogy a $K[x]$ tetszőleges f , g és h elemeire $f \circ (g \circ h) = (f \circ g) \circ h$. Ezt az állítást az f fokára vonatkozó teljes indukcióval bizonyítjuk.

Ha f konstans, akkor az állítás nyilvánvalóan igaz, mert mindkét polinom f -fel egyenlő. Legyen $\text{gr}(f) = n > 0$, és tegyük fel, hogy az állítás igaz minden $(n - 1)$ -edfokú polinomra. Speciálisan, f -et felírhatjuk $f = a + f_1 \cdot x$ alakban, ahol a konstans és $\text{gr}(f_1) = n - 1$. A teljes indukciós feltevés szerint tehát igaz az $f_1 \circ (g \circ h) = (f_1 \circ g) \circ h$ összefüggés. A 3.17. Tételt felhasználva ebből

$$\begin{aligned} f \circ (g \circ h) &= (f_1 \circ (g \circ h)) \cdot (x \circ (g \circ h)) + a = \\ ((f_1 \circ g) \circ h) \cdot (g \circ h) + a \circ h &= ((f_1 \circ g) \cdot g + a) \circ h = (f \circ g) \circ h \end{aligned}$$

következik, amit éppen bizonyítani kellett. ■

Megjegyzés. Láthatjuk, hogy a polinomok az összeadás és a kompozíció műveletére nézve „majdnem” gyűrűt alkotnak. Persze ez a szorzás nem kommutatív, hiszen például $(2x) \circ (x + 1) = 2x + 2$, míg $(x + 1) \circ (2x) = 2x + 1$. Ez azért még nem olyan meglepő, mert a mátrixok szorzása sem kommutatív. Itt azonban más furcsaság is van: A 3.17. Tétel (2) pontja szerint érvényes az „egyk oldali” disztributivitás. A „másik oldali” disztributivitás viszont nem igaz. Például $x^2 \circ (x + x) = (2x)^2 = 4x^2$, míg $(x^2 \circ x) + (x^2 \circ x) = x^2 + x^2 = 2x^2$. Ez mutatja, hogy a mátrixműveletek esetében valóban szükséges volt mindkét disztributivitásnak a bizonyítása.

A 3.17. Tétel utáni második megjegyzésben láttuk alapján az x polinom úgy viselkedik, mint a kompozícióra vonatkozó „jobb oldali” egység. Eleve nem világos az, hogy x „bal oldali” egység is, hiszen a kompozíció művelete nem kommutatív. Ennek ellenére tetszőleges f polinomra igaz az $x \circ f = f$ összefüggés is. Itt is kereshetjük egyes polinomok „inverzét”, azaz adott f polinomhoz olyan g polinomot, amelyre az $f \circ g = g \circ f = x$ összefüggések (vagy ezek egyike) fennállnak.

Kimutatható, hogy a két egyelőség teljesülése ekvivalens; és pontosan az elsőfokú polinomok esetében létezik „inverz” (Ezt a harmadfokú polinomok gyökeinek a keresésénél használni is fogjuk.) \square

A 3.17. Tétel (3) pontja alapján az összeadás helyett választhatnánk a szorzást is; ez is „majdnem” gyűrű volna.

3.19. Tétel. *Kompozíciónál a polinomok foka összeszorozódik (ha itt kivételesen a 0 polinom fokát is 0-nak tekintjük).*

Bizonyítás. Ha f és g bármelyike konstans, akkor nyilván $f \circ g$ is az. Egyébként, ha g k -adfokú, akkor a 3.4. Tétel (2) állítása szerint, g^i foka ik . Ha tehát az f polinom n -edfokú, akkor a 3.4. Tétel (1) állítását figyelembe véve $f \circ g$ foka megegyezik g^n fokával; azaz nk , ahogyan ezt állítottuk. \blacksquare

Következmény. *A $K[x]$ -beli rögzített f polinom esetén akkor és csak akkor írható minden $K[x]$ -beli polinom $f \circ g$, illetve $g \circ f$ alakba, ha f elsőfokú.*

Bizonyítás. Ha minden $K[x]$ -beli polinom $f \circ g$ vagy $g \circ f$ alakú, akkor a 3.19. Tétel alapján $\text{gr}(f)$ osztója minden természetes számnak, tehát $f = ax + b$, ahol $a, b \in K$ és $a \neq 0$. Ha viszont ez teljesül, akkor a $h = a^{-1}x - a^{-1}b$ polinomra $f \circ h = h \circ f = x$. A 3.18. Tételt felhasználva ebből $g = f \circ (h \circ g) = (g \circ h) \circ f$ adódik. \blacksquare

3.11. Definíció. Tetszőleges $a \in K$ szám és $f \in K[x]$ polinom esetén az $f(a) = f \circ a$ számot az f polinom a helyen vett helyettesítési értékének nevezzük. \blacksquare

Megjegyzés. A definícióban szereplő elnevezés helyes, mert a 3.19. Tétel szerint $f(a)$ valóban szám. \square

3.20. Tétel. *Rögzített $a \in K$ esetén az $f \mapsto f(a)$ megfeleltetés $K[x]$ -et úgy képezi le művelettartó módon K -ba, hogy K minden elemének önmaga felel meg.*

Ha $K[x]$ -et úgy képezzük le művelettartó módon K -ba, hogy K minden elemének önmaga felel meg, akkor létezik olyan $a \in K$, hogy az f -nek éppen $f(a)$ felel meg.

Bizonyítás. Az első állítás azonnal következik a 3.17. Tételből, tekintettel arra, hogy $f(a) \in K$.

A második állítást a 3.17. Tétel Kiegészítéséből kapjuk, figyelembe véve, hogy most $x^* \in K$. \blacksquare

Mint a kompozíciónál, látszólag itt is megszorítást jelent az, hogy a $K[x]$ -beli polinomokba csak K -beli számot helyettesíthetünk be (egy racionális együtthatós polinom gyökeit kereshetjük a valós számok körében is). A kompozícióhoz hasonlóan itt is megtehetjük, hogy a polinom együtthatóit eleve abból a bővebb testből valónak gondoljuk, amelyikből a behelyettesítendő számokat vesszük.

Most a maradékos osztásnak a helyettesítési értékkel való szoros kapcsolatát mutatjuk meg.

3.21. Tétel. Ha az f és az $x - c$ $K[x]$ -beli polinomokra elvégezzük a maradékos osztást, akkor az $f = g \cdot (x - c) + f(c)$ összefüggéshez jutunk.

Bizonyítás. Elsőfokú polinommal végezve a maradékos osztást a maradék konstans, ezért azt kapjuk, hogy $f = g \cdot (x - c) + d$, valamilyen K -beli d -vel. A behelyettesítésnek a 3.20. Tételben kimondott tulajdonságai alapján ebből $f(c) = g(c) \cdot (c - c) + d = d$ következik. ■

3.22. Tétel. Legyen $f = a_0 + a_1x + \dots + a_nx^n$, továbbá $f = g \cdot (x - c) + f(c)$ a $g = b_0 + b_1x + \dots + b_{n-1}x^{n-1}$ polinommal. Ekkor a g együtthatóit és az $f(c)$ helyettesítési értéket rekurzíven a következőképpen határozhatjuk meg:

- (1) $b_{n-1} = a_n$;
- (2) ha már b_i ismert, akkor $b_{i-1} = b_i \cdot c + a_i$;
- (3) ha már b_0 ismert, akkor $f(c) = b_0 \cdot c + a_0$.

Bizonyítás. Az $f = g \cdot (x - c) + f(c)$ összefüggésből átrendezéssel a

$$g \cdot x + f(c) = g \cdot c + f$$

összefüggéshez jutunk. A polinomok egyenlősége szerint ez azt jelenti, hogy minden szóba jövő i -re az x^i együtthatója a két oldalon megegyezik. Az x^n együtthatója a bal oldalon b_{n-1} , a jobb oldalon pedig a_n . $0 < i < n$ esetén x^i együtthatója a bal oldalon b_{i-1} , a jobb oldalon pedig $b_i \cdot c + a_i$. Végül a konstans tag a bal oldalon $f(c)$, a jobb oldalon pedig $b_0 \cdot c + a_0$. ■

Megjegyzés. A fenti tétel módot ad mind a hányadospolinom együtthatóinak, mind a helyettesítési értéknek a gyors rekurzív kiszámítására. (Ha a hatványokat szoroznánk az együtthatókkal, majd ezeket összeadnánk, akkor $2n - 1$ számú szorzásra és n darab összeadásra volna szükség, és csupán a helyettesítési értéket kapnánk meg. Látható, hogy a fenti eljárásnál csupán n darab szorzás és n darab összeadás szükséges.) A kiszámítás megkönnyítése érdekében az adatokat egy táblázatba szőtták beírni. A táblázatot és a számolási módszert *Horner-elrendezésnek* nevezik.

Ez az eljárás a következő:

Először sorba leírjuk egymás után a főegyütthatótól kezdve az f polinom együtthatóit:

a_n	\dots	a_{i+1}	a_i	\dots	a_1	a_0

Ezután a következő sor elejére beírjuk a behelyettesítendő c számot, és a_n alá a_n -t, ez lesz egyúttal b_{n-1} :

	a_n	\dots	a_{i+1}	a_i	\dots	a_1	a_0
c	$b_{n-1} = a_n$	\dots					

Ha a táblázat második sora valameddig már ki van töltve, akkor az első üres helyre beírjuk az előtte álló szám c -szeresének és a hely felett álló számnak az összegét:

$$\begin{array}{c|c|c|c|c|c|c|c}
 & a_n & \dots & a_{i+1} & a_i & \dots & a_1 & a_0 \\
 c & b_{n-1} = a_n & \dots & b_i & b_{i-1} = b_i \cdot c + a_i & & &
 \end{array}$$

Végül az utolsó hely kitöltésénél éppen az $f(c)$ helyettesítési értékhez jutunk:

$$\begin{array}{c|c|c|c|c|c|c|c}
 & a_n & \dots & a_{i+1} & a_i & \dots & a_1 & a_0 \\
 c & b_{n-1} = a_n & \dots & b_i & b_{i-1} = b_i \cdot c + a_i & & b_0 & f(c) = b_0 \cdot c + a_0
 \end{array}$$

Felhívjuk a figyelmet arra, hogy az f polinom együtthatóinak a leírásakor ne feledkezzünk meg a 0 együtthatókról, amelyek az f polinomalakjában nincsenek leírva.

A Horner-féle elrendezésnek nagy előnye, hogy egymás utáni sorokban közvetlenül elvégezhetők a helyettesítések más számmal is. Ha emellett valamikor — egy behelyettesítés után — az itt kapott hányadospolinomba akarunk behelyettesíteni, akkor ugyanilyen módon ezt is megtehetjük; egyszerűen „megfelelkezve” arról, ami e fölött a sor fölött áll, és eltekintve az ebben a sorban az utolsó helyen álló $f(c)$ helyettesítési értékről. \square

A továbbiakban éppen olyan esetekről lesz szó, amikor ez fontos lehet.

3.12. Definíció. A $c \in K$ számot az $f \in K[x]$ polinom gyökének nevezzük, ha $f(c) = 0$. \blacksquare

3.23. Tétel. Egy $f \in K$ polinomnak akkor és csak akkor gyöke a $c \in K$ szám, ha a polinom osztható $(x - c)$ -vel.

Egy K felett irreducibilis f polinomnak akkor és csak akkor van K -ban gyöke, ha a polinom elsőfokú.

0-tól különböző polinomnak nem lehet több gyöke, mint amekkora a polinom foka.

Egy polinomnak akkor és csak akkor gyöke minden K -beli szám, ha ez a 0 polinom.

Egyéb konstans polinomnak nincs gyöke.

Bizonyítás. Az $f = g \cdot (x - c) + f(c)$ felírásból következik, hogy $(x - c)$ akkor és csak akkor osztója f -nek, ha $f(c)$ -nek is osztója. Elsőfokú polinom egy konstansnak akkor és csak akkor lehet osztója, ha ez a konstans 0, ezért az oszthatóság pontosan akkor teljesül, ha $f(c) = 0$.

Ha az f irreducibilis polinomnak c gyöke, akkor $f = g \cdot (x - c)$ csak úgy lehetséges, hogy g konstans, f tehát valóban elsőfokú. Ha pedig f elsőfokú, akkor 0-tól különböző a -val $ax + b$ alakú ($a, b \in K$), és e polinomnak nyilván gyöke a $(-b)/a$ szám.

Legyen most $f = p_1 \cdot \dots \cdot p_r$ irreducibilis faktorokra való felbontás. Ekkor tetszőleges K -beli c -re teljesül az $f(c) = p_1(c) \cdot \dots \cdot p_r(c)$ összefüggés. Tehát f bármelyik gyöke valamelyik p_i -nek is gyöke lesz. Egy irreducibilis polinomnak legfeljebb egy gyöke lehet, ezért f -nek legfeljebb r darab gyöke van. Másrészt, f foka legalább r , mert egy szorzat foka megegyezik a tényezők fokának az összegével, és minden egyes tényező legalább elsőfokú.

Végezetül tekintsünk egy polinomot, amelynek minden K -beli szám gyöke. Bármely számtest tartalmazza az összes természetes számot, ezért e polinomnak végtelen sok gyöke van — tehát nem lehet foka. Így a szóban forgó polinom csak a 0 polinom lehet. Ennek valóban gyöke minden K -beli szám. Egyéb konstans polinomnak nyilván nem lehet gyöke. ■

Megjegyzés. Említettük a bevezető részben, hogy a modulo p (p prímszám) vett maradékosztályok testet alkotnak. Ezek véges testek. A testelmélet tárgyalásakor le fogjuk írni az összes véges testet. Ezekre a fenti tétel utolsó állításának első része nem igaz. Ekkor ugyanis abból, hogy a test minden eleme gyöke egy polinomnak, nem következik, hogy a polinomnak végtelen sok gyöke van. Be lehet látni, hogy ha egy véges testnek n eleme van, akkor e test minden eleme gyöke az $x^n - x$ polinomnak. (A tétel bizonyításánál nem volt szükség arra, hogy a szóban forgó test számtest legyen, csak arra támaszkodtunk, hogy végtelen sok eleme van. Ennek megfelelően a tételnek ez a része pontosan a végtelen testeket jellemzi.) Egyébként a testelmélet tárgyalásánál be fogjuk látni, hogy léteznek véges testek; minden q prímhatalványra lényegében egyetlenegy. □

3.13. Definíció. A K -beli c számot a K -beli együtthatós f polinom legalább r -szeres gyökének nevezzük, ha $(x - c)^r$ osztója f -nek. Ha $(x - c)^{r+1}$ nem osztója f -nek, akkor azt mondjuk, hogy c (pontosan) r -szeres gyök. Ebben az esetben az r számot a c gyök multiplicitásának nevezzük. Azt mondjuk, hogy c többszörös gyök, ha valamilyen $r > 1$ természetes számra r -szeres gyök. ■

A 3.23. Tételből következik, hogy egy polinom többszörös gyöke a polinomnak gyöke; és e tétel bizonyítása azt is adja, hogy egy 0-tól különböző polinomnak akkor sem lehet a fokánál több gyöke, ha azokat multiplicitással számoljuk.

6. Polinomfüggvény, interpoláció

Eddig a polinomok „formális” tulajdonságait vizsgáltuk, amelyek igen hasznosak a polinomok kezelésénél. Valójában a polinomok speciális függvényekként „keletkeztek”; s úgy tekinthetők, mint e függvények „tükröképei”. Az alábbiakban a polinomokat „visszaforldítjuk”; és megvizsgáljuk az általuk létrehozható függvényeket.

3.14. Definíció. Tetszőleges $f \in K[x]$ polinomhoz rendeljük hozzá azt az $f^*: K \rightarrow K$ függvényt, amelyre bármely K -beli a esetén $f^*: a \mapsto f(a)$. Az f^* neve polinomfüggvény. ■

3.24. Tétel. A 3.14. Definícióbeli $f \mapsto f^*$ megfeleltetés művelettartó az alábbi értelemben:

- (1) $(f + g)^* = f^* + g^*$ (függvény-összeadás);
- (2) $(fg)^* = f^*g^*$ (függvénytársorzás);
- (3) $(f \circ g)^* = f^*(g^*)$ (függvénykompozíció: közvetett függvény képzése).

Bizonyítás. Ha φ és ψ a K testnek (mint halmaznak) önmagára való leképezései, akkor ezek összegét, szorzatát és kompozícióját — megfelelően —

$(\varphi + \psi)(a) = \varphi(a) + \psi(a)$, $(\varphi \cdot \psi)(a) = \varphi(a) \cdot \psi(a)$, illetve $(\varphi \circ \psi)(a) = \varphi(\psi(a))$ definiálja, tetszőleges $a \in K$ esetén.

Ebből az első két összefüggés azonnal következik. A harmadik az

$$f^*(g^*(a)) = f(g(a)) = f \circ (g \circ a) = (f \circ g) \circ a = (f \circ g)(a) = (f \circ g)^*(a)$$

összefüggésből adódik. ■

A 3.24. Tétel alapján a polinomfüggvényekkel éppen olyan egyszerűen számolhatunk, mint a polinomokkal. Éppen ezért a polinomfüggvényekkel való számolás sokkal „gazdaságosabb”, mint ha valamilyen bonyolult függvénnyel számolnánk. Ez az egyszerűség azt is jelenti, hogy a számolás sokkal kevesebb időt vesz igénybe. Ennek érdekében sokszor igen célszerű tetszőleges függvényeket polinomfüggvényekkel „helyettesíteni”, „közelíteni”. Természetesen ennek „ára” van; nevezetesen a közelítési hiba miatt az eredmények sem lesznek „teljesen pontosak”. Tekintettel azonban arra, hogy a valóságban előforduló mérések esetében már az adatok sem pontosak, ennek következtében egy „kikövetkeztetett” függvényt nem is ismerünk, ezért konkrét számolások esetében a keletkezett hiba viszonylag kicsi lehet. Gépi számolásoknál az egyszerűség miatt általában ilyen módszert használnak.

A közelítés egyik módja az alábbi: Megnézzük egy adott függvényről, hogy bizonyos helyeken milyen értékeket vesz fel, és ezután olyan polinomot keresünk, amely az adott helyeken ugyanazokat az értékeket veszi fel, mint a megadott függvény. Ezt az eljárást interpolációnak nevezzük.

Az alábbiakban bebizonyítjuk az interpoláció elvégezhetőségét.

3.25. Tétel. Legyenek a_0, a_1, \dots, a_n a K test különböző és b_0, b_1, \dots, b_n a K test tetszőleges elemei. Ekkor létezik olyan f úgynevezett interpolációs polinom, amelyre $f(a_i) = b_i$ teljesül, ahol $i = 0, 1, \dots, n$.

Ha azt is feltesszük, hogy az interpolációs polinom foka legfeljebb n , akkor f egyértelműen meghatározott.

Tetszőleges g polinom akkor és csak akkor lesz a $g(a_i) = b_i$ ($i = 0, 1, \dots, n$) feltételeket kielégítő interpolációs polinom, ha a fenti f polinommal

$$g = f + h \cdot (x - a_0) \cdot (x - a_1) \cdot \dots \cdot (x - a_n)$$

teljesül valamely h polinomra.

Bizonyítás. Az utolsó állítás egész könnyen adódik. Ha g és f mindegyike interpolációs polinom, akkor különbségük az a_i helyeken a 0 értéket veszi fel, azaz osztható mindegyik az $(x - a_i)$ polinommal. E tényezők feltétel szerint különbözőek, és mivel irreducibilisek (így prímtulajdonságúak) is, ezért $g - f$ osztható szorzatukkal. Az állítás megfordítása behelyettesítéssel azonnal belátható.

Ezzel beláttuk azt is, hogy két interpolációs polinom vagy megegyezik, vagy különbözőségük legalább $(n+1)$ -edfokú. Két különböző, n -nél nem nagyobb fokú polinom különbsége viszont legfeljebb n -edfokú, ezért igaz az egyértelműsége vonatkozó állítás is.

Az interpolációs polinom létezését az n -re vonatkozó teljes indukcióval bizonyítjuk.

Ha $n = 0$, akkor $f = b_0$ egy olyan interpolációs polinom, amely mindenütt a b_0 értéket veszi fel.

Tegyük fel, hogy az állítás igaz $(n-1)$ -re, és legyen f olyan, n -nél alacsonyabb fokú polinom, amelyre $f(a_i) = b_i$ ($i = 0, 1, \dots, n-1$). Tetszőleges $c \in K$ mellett a $g = f + c \cdot (x - a_0) \cdot \dots \cdot (x - a_{n-1})$ polinomra $g(a_i) = b_i$ ($i = 0, 1, \dots, n-1$), mint láttuk. A $c = [(a_n - a_0) \cdot \dots \cdot (a_n - a_{n-1})]^{-1} \cdot (b_n - f(a_n))$ választással nyilván teljesül $g(a_n) = b_n$ is. (Az inverzképzés elvégezhető, mert az a_i elemek feltétel szerint páronként különbözőek voltak.) ■

Megjegyzés. Az interpolációs polinomnak a most ismertetett rekurzív meghatározási módszere a *Newton-féle interpoláció*. Ezt akkor célszerű használni, amikor egy függvényt vizsgálunk, és ezt egyre több „hely” felvételével akarjuk meghatározni.

Olyan esetben, amikor a helyek több feladatnál ugyanazok, csak a felvett értékek változnak, célszerűbb a *Lagrange-féle interpolációt* használni. Ez a következőképpen történik. Legyen $g = (x - a_0) \cdot \dots \cdot (x - a_n)$, és tekintsük a $g_i = \frac{g}{x - a_i}$ polinomokat, amelyekre $g_i(a_j) = 0$, ha $j \neq i$, míg $g_i(a_i) = c_i \neq 0$. Tekintsük az $\ell_i = \frac{g_i}{c_i}$ úgynevezett *interpolációs alappolinomokat*. Ezek azzal a tulajdonsággal rendelkeznek, hogy $\ell_i(a_j) = 0$, ha $j \neq i$, míg $\ell_i(a_i) = 1$. Ekkor $f = b_0 \ell_0 + \dots + b_n \ell_n$ éppen az interpolációs polinom lesz. □

3.26. Tétel. *Tetszőleges K számtest esetén a 3.14. Definícióban megadott $f \mapsto f^*$ megfeleltetés injektív, de nem szürjektív.*

Bizonyítás. Az injektivitás azt jelenti, hogy különböző polinomoknak különböző polinomfüggvény felel meg. Legyen ennek a bizonyítása végett $f^* = g^*$. A 3.24. Tételt felhasználva azt kapjuk, hogy a $h = f - g$ polinomra $h^* = f^* - g^*$ a nullafüggvény, azaz h^* minden $(K$ -beli) számot 0-ba visz. Ez azt jelenti, hogy ha $a \in K$, akkor $h(a) = 0$. K -nak végtelen sok eleme van és h -nak legfeljebb annyi gyöke van, amennyi a foka, ezért $h = 0$, tehát $f = g$.

Ahhoz, hogy a megfeleltetés nem szürjektív, azt kell megmutatni, hogy van olyan függvény, amelyik nem polinomfüggvény. Mint láttuk, egy polinomnak legfeljebb annyi gyöke van, amennyi a polinom foka. Így, ha egy polinomnak végtelen sok gyöke van, akkor ez a 0 polinom. Például az a függvény, amelyik minden 0-tól különböző K -beli helyen 0-t vesz fel, ha polinomfüggvény lenne, 0-ban is 0-t venne fel. Így, ha a függvény 0-ban az 1 értéket veszi fel, nem lehet polinomfüggvény. ■

Megjegyzés. A bizonyításból világos, hogy a K testnek az a tulajdonsága volt lényeges, hogy végtelen sok eleme van. A következőkben megmutatjuk, hogy ez a feltétel valóban szükséges. Bebizonyítjuk, hogy véges testek esetében fordított a helyzet, azaz a szóban forgó megfeleltetés nem injektív, hanem szürjektív. A szürjektivitás azonnal következik abból, hogy a testnek véges sok eleme van.

Így e helyeken tetszőleges függvényértékeket előírva az interpolációs polinomfüggvény megegyezik az előre felvett függvénnyel. Másrészt, vegyük figyelembe, hogy $K[x]$ -nek végtelen sok eleme van, míg a K testen csak véges sok függvény értelmezhető. Létezniök kell tehát olyan különböző polinomoknak, amelyeknek ugyanaz a polinomfüggvény felel meg. \square

7. A legfeljebb negyedfokú polinomok gyökeinek meghatározása

A következőkben megvizsgáljuk, miképpen lehet polinomok gyökeit meghatározni. A nulladfokú polinomok esetében a válasz érdektelen. Elsőfokú polinomokról pedig nyilvánvaló, hogy pontosan egy gyökük van, amelyik ugyanannak a számtestnek az eleme, mint ahonnan az együtthatók valók.

Az az eljárás, amely az elsőfokú polinomoknál már a megoldást szolgáltatja, a magasabbfokú polinomok esetében bizonyos „redukciót” tesz lehetővé. A gyöktényezők nem változnak akkor, ha a polinomot egy nemnulla konstanssal szorozzuk, ezért elegendő, ha normált polinomok gyökeit keressük. Először a másodfokú polinomok esetében nézzük meg, miképpen lehet a gyököket meghatározni. (Ez ugyan középiskolai anyag, de itt a komplex számtest is rendelkezésünkre áll, és a tárgyalásmód is különbözik az ottanitól.)

Legyen $x^2 + px + q \in K[x]$. Ha e polinomnak van gyöke, akkor a 3.23. Tétel szerint osztható a megfelelő gyöktényezővel, tehát a polinom reducibilis. Ez egy másodfokú polinom, ezért két elsőfokú polinom szorzata. Az egyik tényező és a szorzat normáltsága miatt a másik tényező is normált. Végeredményben tehát az

$$x^2 + px + q = (x - a) \cdot (x - b)$$

felbontáshoz jutunk. Összegezve:

Ha az $x^2 + px + q \in K[x]$ polinomnak van K -ban gyöke, akkor felírható az $(x - a) \cdot (x - b)$ úgynevezett gyöktényezős alakban. A polinom összes gyökei a és b ; ha ezek egyenlők, akkor a polinomnak egy kétszeres gyöke van.

A két elsőfokú tényezőt összeszorozva, a polinomok egyenlősége alapján a következőket nyerjük:

Ha az $x^2 + px + q$ polinom gyökei a és b , akkor

$$a + b = -p \quad \text{és} \quad a \cdot b = q.$$

Ezeket a gyökök és együtthatók közötti összefüggésnek nevezzük.

Az $(a + b)^2 - (a - b)^2 = 4ab$ összefüggésből azonnal következik, hogy $D = (a - b)^2 = p^2 - 4q$. Ezt a D számot a másodfokú polinom *diszkriminánsának* (meghatározó) nevezzük. A D szám azt határozza meg, hogy a polinomnak vannak-e többszörös gyökei. A *diszkrimináns akkor és csak akkor 0, ha a polinomnak többszörös gyöke van, amely ekkor eleme az együtthatókat tartalmazó testnek.*

Valóban, ha a két gyök megegyezik, akkor különbségük négyzete természetesen 0.

Fordítva, ha $D = 0$, akkor a polinom $x^2 + px + q = x^2 + px + \frac{p^2}{4} = \left(x + \frac{p}{2}\right)^2$ alakban írható fel, ami éppen az előbbi állítást adja. A továbbiakban feltesszük, hogy K a komplex számtest és $D \neq 0$. A komplex számok körében elvégezhető a négyzetgyökvonás, ezért létezik a \sqrt{D} komplex szám. Feltehető, hogy a gyököket éppen úgy jelöltük, hogy $a - b = \sqrt{D}$. Ebből, és az $a + b = -p$ összefüggésből azonnal következik, hogy a szóban forgó polinom két gyöke csak

$$a = \frac{-p + \sqrt{D}}{2} \quad \text{és} \quad b = \frac{-p - \sqrt{D}}{2}$$

lehet. Azt, hogy ezek valóban gyökök, behelyettesítéssel láthatjuk be. (Például Horner-elrendezéssel.) Következtethetünk azonban a 3.23. Tétel felhasználásával is. A fenti a és b számokra ugyanis $a + b = -p$ és $4ab = (a + b)^2 - (a - b)^2 = p^2 - D = 4q$ alapján $x^2 + px + q = (x - a)(x - b)$ adódik.

Ezzel a következőket láttuk be:

A komplex együtthatós $x^2 + px + q$ polinomnak két gyöke van, $\frac{-p + \sqrt{D}}{2}$ és $\frac{-p - \sqrt{D}}{2}$, ahol $D = p^2 - 4q$ a polinom diszkriminánsa. A két gyök pontosan akkor egyenlő, ha a diszkrimináns 0.

A másodfokú polinomok gyökeinek a meghatározásával már akkor is foglalkoztak, amikor még a komplex számokat nem ismerték. Ezért is, meg a valós számok fontossága miatt is célszerű, ha külön megvizsgáljuk a valós együtthatós polinomokat. Amennyiben a polinom D diszkriminánsa pozitív, a fenti két gyök mindegyike valós szám. Ha viszont a diszkrimináns negatív szám (a $D = 0$ esetet már megvizsgáltuk), akkor a két gyök egyike sem lesz valós. Mivel ebben az esetben a két gyök valós része megegyezik, képzetes részük pedig egymás ellentettje, ezért a két gyök egymásnak konjugáltja. Eredményünket a következőképpen foglalhatjuk össze:

Ha a valós együtthatós $x^2 + px + q$ polinom D diszkriminánsa pozitív, akkor e polinomnak a fenti két különböző valós szám a gyöke; a $D = 0$ esetben $-\frac{p}{2}$ a polinom kétszeres (valós) gyöke; negatív diszkrimináns esetén a polinomnak valós gyöke nincs, de a komplex számok körében két különböző gyöke van, amelyek egymás konjugáltjai.

Most a harmadfokú polinomok gyökeinek a meghatározására térünk rá. Ha egy $K[x]$ -beli harmadfokú polinom gyökeit keressük, ezek — általában — nem lesznek K -beli számok. Várható azonban, hogy a másodfokú polinom esetéhez hasonlóan itt is fellép egy K -beli tag (a $-p/2$ -nek megfelelően). Ez abból látható, hogy a gyökökhöz egy rögzített K -beli számot hozzáadva, a kapott számok egy ugyancsak $K[x]$ -beli polinomnak lesznek gyökei (ha az eredeti polinom f , és a gyökökhöz a -t adunk, akkor a kapott polinom $f \circ (x - a)$ lesz). Ez a tag igen „kényelmetlen” a megoldás során. Ennek kiküszöbölésére hasonlóképpen járhatunk el, mint a másodfokú polinomoknál, ha a gyököket a teljes négyzetté való kiegészítéssel keressük. (A különbség csak az, hogy a másodfokú polinomoknál

ezáltal már lényegében meg is kapjuk a megoldást, míg itt csupán ekkor kezdetjük el a gyökök keresését.)

3.27. Tétel. Minden $f \in K[x]$ harmadfokú polinomhoz található olyan $x^3 + px + q \in K[x]$ polinom, amelynek gyökeihez egy rögzített $d \in K$ hozzáadásával az f gyökeit nyerjük.

Bizonyítás. Itt is elegendő, ha normált polinomokra szorítkozunk. Legyen $f = x^3 + ax^2 + bx + c$. Ha ezt a polinomot teljes köbbsé szeretnénk kiegészíteni, akkor — az első két tagot figyelembe véve — ez a köb csak $\left(x + \frac{a}{3}\right)^3$ lehet. Így olyan g polinomot kell találni, amelyre $g \circ \left(x + \frac{a}{3}\right) = f$. Felhasználva a kompozíció asszociativitását és az $\left(x + \frac{a}{3}\right) \circ \left(x - \frac{a}{3}\right) = x$ összefüggést $g = f \circ \left(x - \frac{a}{3}\right) = \left(x - \frac{a}{3}\right)^3 + a\left(x - \frac{a}{3}\right)^2 + b\left(x - \frac{a}{3}\right) + c = x^3 + \left(b - \frac{a^2}{3}\right)x + \left(c - \frac{ab}{3} + \frac{2a^3}{27}\right)$ adódik. Ez valóban a kívánt alakú polinom. Az $f = g \circ \left(x + \frac{a}{3}\right)$ összefüggésből azonnal adódik, hogy $d = \frac{a}{3}$. ■

Megjegyzés. A tétel szerint, ha a kapott „hiányos” harmadfokú polinomnak vannak gyökei, akkor vannak gyökei az eredeti polinomnak is. Az állítás megfordítására általában nincs szükségünk, mert be fogjuk látni, hogy a hiányos harmadfokú polinomnak mindig vannak gyökei. Nem ez a helyzet akkor, ha a gyököket nem a komplex számtestben, hanem egy szűkebb számtestben keressük. A bizonyításból világos azonban, hogy az eredeti polinomnak pontosan annyi gyöke van egy adott számtestben, mint amennyi a kapott hiányos polinomnak. □

A következőkben hiányos harmadfokú polinomok gyökeinek a meghatározására térünk rá. A $p = 0$ esetben világos, hogy a polinom gyökeit köbgyökvonással megkaphatjuk. Most arra szeretnénk rámutatni, hogy milyen jelenség okozza, hogy p általában különbözik 0-tól. Itt csak egy példával tudjuk ezt megvilágítani, de a felsőbb tanulmányok alapján megmutatható, hogy ennek így kell lennie.

A $\sqrt[3]{2}$ valós szám gyöke egy harmadfokú racionális együtthatós polinomnak, az $(x^3 - 2)$ -nek. E szám négyzete, $\sqrt[3]{4}$ szintén gyöke egy harmadfokú racionális együtthatós polinomnak, $(x^3 - 4)$ -nek. Megmutatjuk, hogy e két szám összege, $\alpha = \sqrt[3]{2} + \sqrt[3]{4}$ ugyancsak egy harmadfokú racionális együtthatós polinom gyöke. Az

$$(u + v)^3 = u^3 + v^3 + 3u^2v + 3uv^2 = u^3 + v^3 + 3uv(u + v)$$

összefüggés alapján $\alpha^3 = 2 + 4 + 3 \cdot 2 \cdot \alpha$. Így α gyöke az $x^3 - 6x - 6$ polinomnak, amely ugyancsak racionális együtthatós. (Azt is láthatjuk, hogy a kapott polinom hiányos.)

Világos, hogy a fenti eljárás sokkal általánosabban is elvégezhető. A fenti példa azt is sugalmazza, hogy az adott $x^3 + px + q$ polinom gyökeit is célszerű $u + v$ alakban keresni, ahol mind u , mind v egy-egy K -beli szám köbgyöke. Általában még ez sem érhető el; éppen ezért a gyököket a komplex számtestben keressük.

A már látott $(u + v)^3 = u^3 + v^3 + 3uv(u + v)$ összefüggés alapján $u + v$ gyöke az $x^3 - 3uvx - (u^3 + v^3)$ polinomnak. Amennyiben tehát ez megegyezik az $x^3 + px + q$

polinommal, akkor $u + v$ ennek a polinomnak is gyöke lesz. E két polinom egyenlőségének szükséges és elégséges feltétele az, hogy:

$$u \cdot v = -\frac{p}{3} \quad \text{és} \quad u^3 + v^3 = -q$$

legyen. Célszerű az első feltételt

$$u^3 \cdot v^3 = -\frac{p^3}{27}$$

alakba írni. Ekkor ugyanis ebből és a második feltételből — a másodfokú polinom gyökeinek és együtthatóinak az összefüggése alapján — azt kapjuk, hogy u^3 és v^3 pontosan az

$$x^2 + qx - \frac{p^3}{27}$$

polinom gyökei. A két gyök szerepe szimmetrikus, ezért feltehető, hogy

$$u^3 = -\frac{q}{2} + \sqrt{D} \quad \text{és} \quad v^3 = -\frac{q}{2} - \sqrt{D},$$

ahol

$$D = \left(-\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3.$$

D -t a (hiányos) harmadfokú polinom *diszkriminánsának* nevezzük. (Magasabb szempontok figyelembevételével célszerűbb $108 \cdot D$ -t nevezni diszkriminánsnak.)

3.28. Tétel. Az $x^3 + px + q \in K[x]$ polinomnak akkor és csak akkor van többszörös gyöke, ha diszkriminánsa 0. Ebben az esetben multiplicitással számolva a polinomnak három gyöke van, amelyek elemei a K számtestnek.

Bizonyítás. Mivel a polinom harmadfokú, ezért kétszeres gyök esetén elsőfokú tényező szorzatára bontható. Így, ha a a polinomnak kétszeres gyöke, akkor

$$x^3 + px + q = (x - a) \cdot (x - a) \cdot (x - b).$$

A szorzást elvégezve, a polinomok egyenlősége alapján

$$2a + b = 0, \quad a^2 + 2ab = p, \quad a^2b = -q$$

adódik. b -t az első egyenlőségéből kifejezve és a másik két egyenlőségbe behelyettesítve azt kapjuk, hogy

$$p = -3a^2 \quad \text{és} \quad q = 2a^3.$$

Így

$$D = \left(-\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 = (a^3)^2 + (-a^2)^3 = a^6 - a^6 = 0,$$

ahogyan azt állítottuk.

Tegyük most fel, hogy $D = 0$. A $p = 0$ esetben ebből $q = 0$ következik, és ekkor 0 a polinomnak háromszoros gyöke. Ha $p \neq 0$, akkor azt mutatjuk meg, hogy a $p = -3a^2$ és $q = 2a^3$ összefüggésekből nyerhető a kétszeres gyöke a polinomnak. Legyen tehát ebben

az esetben $a = -\frac{3q}{2p}$. Ebből:

$$a^2 = \frac{9q^2}{4p^2} = \left(\frac{q}{2}\right)^2 \cdot \left(\frac{3}{p}\right)^2 = -\left(\frac{p}{3}\right)^3 \cdot \left(\frac{3}{p}\right)^2 = -\frac{p}{3},$$

és

$$a^3 = \left(-\frac{p}{3}\right) \cdot \left(-\frac{3q}{2p}\right) = \frac{q}{2}.$$

Ezzel a $p = -3a^2$ és a $q = 2a^3$ összefüggésekhez jutottunk. Ezekből közvetlenül adódik, hogy:

$$x^3 + px + q = x^3 - 3a^2x + 2a^3 = (x - a) \cdot (x - a) \cdot (x + 2a).$$

$3q$ és $2p$ is K -beliek, ezért a gyökök valóban elemei K -nak (a $p = 0$ esetben ez $0 \in K$ miatt nyilvánvaló). ■

3.29. Tétel. Az $x^3 + px + q \in K[x]$ polinom gyökei pontosan azok az $a + b$ alakú számok, amelyekre:

$$a^3 = -\frac{q}{2} + \sqrt{D}, \quad b^3 = -\frac{q}{2} - \sqrt{D} \quad \text{és} \quad D = \left(-\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3;$$

továbbá az a és b számok szorzata $-\frac{p}{3}$. A fenti polinomnak pontosan három gyöke van.

Bizonyítás. A $D = 0$ esetben eredményünk az előző tételből adódik. Ha $D \neq 0$ és $p = 0$, akkor $D = \left(-\frac{q}{2}\right)^2$ miatt $q \neq 0$. Ekkor $\sqrt{D} = \pm \frac{q}{2}$ (valamelyik rögzített előjellel). Ebből $a^3 = -\frac{q}{2} \pm \frac{q}{2}$ és $b^3 = -\frac{q}{2} \mp \frac{q}{2}$ adódik. Ezek egyike 0, másikuk $-q$, így $ab = 0$ igaz. A $p \neq 0$ esetben a polinom $x^3 + q$, és valóban $-q + q = 0$. Tekintettel arra, hogy $q \neq 0$, ezért $\sqrt[3]{-q}$ -ra valóban három különböző érték adódik.

A továbbiakban a $D \neq 0$, $p \neq 0$ esettel foglalkozunk.

Válasszuk az a -ra szóba jövő három érték valamelyikét, és tegyük fel, hogy éppen ezt jelöltük a -val. A b -re szóba jövő három érték b , ϱb és $\varrho^2 b$, ahol ϱ az egyik harmadik primitív egységgyök ($\varrho = -\frac{1}{2} + \frac{\sqrt{3}}{2}$ vagy $\varrho = -\frac{1}{2} - \frac{\sqrt{3}}{2}$). Ezeket a -val szorozva az ab , ϱab és $\varrho^2 ab$ számokhoz jutunk. E három szám mindegyikének a köbe $-\frac{p^3}{27}$. A $p \neq 0$ feltétel miatt a most kapott három szorzat különböző, így pontosan az egyikük egyezik meg $-\frac{p}{3}$ -mal, mivel ennek is $-\frac{p^3}{27}$ a köbe. Nem megy az általánosság rovására, ha feltesszük, hogy éppen azt a számot jelöltük b -vel, amelyikre $ab = -\frac{p}{3}$. A gyökökre szóba jövő három lehetőség tehát $a + b$, $\varrho a + \varrho^2 b$ és $\varrho^2 a + \varrho b$. A megfelelő gyöktényezők szorzatát

kiszámolva a következőkhöz jutunk:

$$\begin{aligned} & (x - (a + b)) \cdot (x - (\varrho a + \varrho^2 b)) \cdot (x - (\varrho^2 a + \varrho b)) = \\ & = (x - (a + b)) \cdot (x^2 + (a + b)x + (a^2 - ab + b^2)) = \\ & = x^3 - 3abx - (a^3 + b^3) = x^3 + px + q. \end{aligned}$$

Ezáltal az $x^3 + px + q$ polinomot három elsőfokú tényezőre bontottuk fel. Ennek a polinomnak a gyökei tehát éppen e három polinom gyökei. ■

Megjegyzés. A kapott eredményt a következőképpen szokták felírni:

$$\varrho^i \cdot \sqrt[3]{-\frac{q}{2} + \sqrt{\left(-\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \varrho^{2i} \cdot \sqrt[3]{-\frac{q}{2} - \sqrt{\left(-\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}},$$

ahol $i = 0, 1, 2$. A harmadfokú polinom gyökeire adott fenti kifejezés a *Cardano-képlet*.

Természetesen $\left(-\frac{q}{2}\right)^2$ helyett mindenütt szerepelhetne $\left(\frac{q}{2}\right)^2$ is. Azért használtuk mégis a $\left(-\frac{q}{2}\right)^2$ kifejezést, mert a köbgyök előtt is $-\frac{q}{2}$ áll. Így (talán) könnyebb emlékezni a képletre. □

Most rátérünk annak a fontos esetnek a vizsgálatára, amikor a polinom együtthatói valósak.

3.30. Tétel. *Ha a valós együtthatós $x^3 + px + q$ polinom D diszkriminánsa negatív, akkor a polinomnak három különböző valós gyöke van, ha a diszkrimináns pozitív, akkor pontosan egy valós gyök és egy konjugált komplex gyökpár lép fel. A diszkrimináns előjele a megfelelő esetnek szükséges feltétele is.*

Bizonyítás. Legyen először D pozitív. Ekkor a 3.29. Tételben szereplő u^3 és v^3 számok mindegyike valós. $D \neq 0$ miatt feltehető, hogy pl. $u \neq 0$, és ekkor választható valósnak is. Mivel az uv szorzat valós, ezért a választott valós u -hoz tartozó v is valós. Tehát a polinom $u + v$ gyöke is valós. A másik két gyök összege — például $\varrho^2 + \varrho = -1$ alapján — éppen ennek a gyöknek a negatívja, azaz valós. E két utóbbi gyök különbsége $(\varrho - \varrho^2)(u - v)$. Itt az első tényező $\sqrt{-3}$, vagyis tiszta képzetes szám. A második tényező valós és különbözik 0-tól, egyébként a 3.27. Tétel szerint $D = 0$ volna. Így e két gyök különbsége tiszta képzetes, és mivel összegük valós, a két gyök valóban egymás konjugáltja.

Negatív D esetén u^3 és v^3 egymásnak konjugáltjai. Választható tehát olyan u, v értékpár, amelyek ugyancsak konjugáltjai egymásnak. Ekkor szorzatuk természetesen valós lesz. Ahhoz azonban, hogy másféle párosításból nem indulhatunk ki, azt is be kell látni, hogy e szorzat nem lehet 0. Ennek lehetetlensége például úgy látható be, hogy $p = 0$ esetén D nem lehetne negatív. Mivel a két primitív harmadik egységgyök is konjugáltja egymásnak, ezért a $\varrho u, \varrho^2 v$, valamint a $\varrho^2 u, \varrho v$ párokban is konjugált komplex számok szerepelnek. Így a három fellépő összeg mindegyike valós, és $D \neq 0$ miatt különböző.

A feltétel szükségessége következik abból, hogy valós együtthatós polinom diszkriminánsa is valós, és egy valós számra a pozitivitás, negativitás és a 0-val való egyenlőség

feltétele közül pontosan az egyik teljesül. (Vigyázni kell azonban arra, hogy egyetlen valós számot kétszer felsorolva ne tekintsük konjugált komplex gyökpárnak.) ■

Megjegyzés. Valós együtthatós polinomoknál az a „legtermészetesebb” eset, amikor mindhárom gyök valós, és ezek különbözőek is. Ekkor mind az együtthatók, mind a gyökök a valós számtesten belül maradnak. A Cardano-féle formula azonban éppen ebben az esetben csak a komplex számok segítségével szolgáltatja a polinom gyökeit. Csupán a valós számok ismeretében, a fellépő „értelmetlen” komplex számokkal a „számolási szabályok” szerint dolgozva meg lehetett kapni a gyököket, amelyeknek ismét semmi közük sem volt a „segítséggül vett értelmetlen dolgokhoz”. Így vált szükségessé a komplex számoknak mint egy bővebb számkörnek a bevezetése. A harmadfokú valós együtthatós polinomoknál a negatív diszkrimináns esetét *casus irreducibilisnek* nevezik. (A testelméleti vizsgálatoknál válik majd világossá, hogy ez a „kellemetlenség” nem kerülhető el.) □

A továbbiakban a negyedfokú polinomok gyökeinek a meghatározását — illetve ilyen polinomok lineáris faktorokra való felbontását — tűzzük ki célul. Ha azonban figyelmesen átnézzük eddigi eredményeinket, azt látjuk, hogy a fokszám növekedésével egyre jobban növekszik a megoldás „bonyolultsága”. Ez még fokozódik a negyedfokú polinomok esetében. Így — bár itt is létezik „megoldóképlet” — csupán egy eljárást fogunk adni a polinom gyökeinek a meghatározására. Ennek az eljárásnak a segítségével minden konkrét esetben meg lehet határozni a polinom gyökeit.

Az ismert azonosság alapján elegendő, ha megmutatjuk, hogy megvalósítható az

$$x^4 + ax^3 + bx^2 + cx + d = \left(x^2 + \frac{a}{2}x + u\right)^2 - (px + q)^2$$

átalakítás — alkalmas u , p és q számokkal. A polinomok egyenlőségének alapján a fenti összefüggés ekvivalens a

$$b = \frac{a^2}{4} + 2u - p^2, \quad c = au - 2pq, \quad d = u^2 - q^2$$

egyenlőségek teljesülésével. Az első és a harmadik egyenlőségéből is és a másodikból is kifejezhetjük $4p^2q^2$ -t. Ezeket egyenlővé téve eredményül

$$(8u + a^2 - 4b) \cdot (u^2 - d) = (au - c)^2$$

adódik. Ebből azt kapjuk, hogy u -nak választhatjuk a

$$8x^3 - 4bx^2 + (2ac - 8d)x - (a^2d - 4bd + c^2)$$

polinom bármelyik gyökét. Egy harmadfokú polinomnak mindig van gyöke, ezért ilyen u -t találhatunk. Az u ismeretében, a $p^2 = \frac{a^2}{4} + 2u - b$ és a $q^2 = u^2 - d$ összefüggések alapján p és q meghatározható. Persze a meghatározás nem egyértelmű. Bármelyiküknek vehetjük a negatívját is. Mégsem kapunk négy megoldást, mert $2pq = au - c$ következtében csak két pár jöhet szóba. (Ez annak felel meg, hogy $px + q$ vagy $-px - q$ az a kifejezés, amit négyzetre akarunk emelni, azaz a felbontásban szereplő két másodfokú tényező milyen sorrendben lép fel.) Ezzel bizonyítást nyert a

3.31. Tétel. *Bármely komplex együtthatós negyedfokú polinom felbontható négy elsőfokú tényező szorzatára.*

Megjegyzések

1. Belátható, hogy a kapott harmadfokú polinom más-más gyökét véve az történik, hogy a szereplő négy elsőfokú tényezőt más-más lehetséges módon párosítjuk össze egymással. Ez a testelméleti vizsgálatoknál — pontosabban szólva a Galois-elmélet tárgyalásánál válik világossá.

2. Az algebrai vizsgálatoknak sokáig az volt az egyik célja, hogy megoldóképletet, de legalábbis a fenti módon leírt megoldási eljárást találjanak magasabb fokú polinomokra is. (Ez ugyan már egyáltalán nem volna használható vagy áttekinthető.) Bizonyítást nyert azonban, hogy ilyen megoldóképlet nem létezik. Ezt is a Galois-elmélet tárgyalásánál fogjuk belátni. \square

8. Az algebra alaptételének ekvivalens alakjai

A legfeljebb negyedfokú polinomok gyökeinek meghatározása alapján azt gondolhatjuk, hogy minden polinomot fel lehet bontani elsőfokú komplex együtthatós polinomok szorzatára. Azt nem állíthatjuk, hogy minden polinomnak annyi gyöke van, amennyi a foka, hiszen egy-egy elsőfokú faktor többször is előfordulhat. Az ebből származó „zavarokat” az alábbi módon lehet kiküszöbölni.

3.15. Definíció. Ha a $K[x]$ -beli f polinomnak a_1, \dots, a_r az összes különböző K -beli gyökei, s ezek multiplicitása rendre k_1, \dots, k_r , akkor azt mondjuk, hogy az f polinomnak a K számtestben összesen $k_1 + \dots + k_r$ gyöke van. \blacksquare

A fenti definíció figyelembevételével azt lehet sejteni, hogy érvényes az alábbi tétel:

Az algebra alaptétele: *A komplex számok körében minden polinomnak annyi gyöke van, mint amennyi a foka.*

Valójában e tételnek nem sok köze van az algebrához, mert ma már nem az az algebrai vizsgálatok kiindulópontja, hogy tekintünk egy polinomot és vesszük annak a gyökeit. Maga a tétel sem bizonyítható algebrai eszközökkel, csupán függvényteni módon, mert tulajdonképpen itt bizonyos típusú függvények viselkedéséről van szó (a polinomfüggvények is ilyenek). E tételt azonban igen sokféle formában ki lehet mondani, és ezeknek az ekvivalenciája már algebrai módszerekkel bizonyítható is. Azt természetesen nem mondhatjuk, hogy bizonyos, a polinomokra vonatkozó állítások a komplex számtestben ekvivalensek, mert ezek bármelyike *igaz* a komplex számtest esetében. Éppen ezért a tételt úgy fogalmazzuk meg, hogy ezek az állítások tetszőleges számtest esetében ugyanazt jelentik. Így, ha a komplex számokra ezeknek az állításoknak bármelyikét be tudjuk bizonyítani, akkor már az itt bizonyításra kerülő tétel következtében a többi állítás is igazolást nyer.

Egyébként az *algebra alaptétele* „nagyon sok” számtestre igaz.

3.32. Tétel. *Bármely K számtest esetén a $K[x]$ -beli polinomokra ekvivalensek az alábbi állítások:*

- (1) *Minden legalább elsőfokú polinomnak van K -ban gyöke.*
- (2) *Minden legalább elsőfokú polinomnak van elsőfokú faktora.*

- (3) Minden irreducibilis polinom elsőfokú.
 (4) Minden legalább elsőfokú polinom felbontható elsőfokú polinomok szorzatára.
 (5) Ha f normált n -edfokú polinom, akkor felírható $f = (x - a_1) \cdot \dots \cdot (x - a_n)$ alakban.
 (6) Ha f normált n -edfokú polinom, akkor felírható $f = (x - b_1)^{k_1} \cdot \dots \cdot (x - b_r)^{k_r}$ alakban, ahol a b_i számok különbözőek és $k_1 + \dots + k_r = n$.
 (7) Bármely n -edfokú polinomnak pontosan n gyöke van.

Azokat a (szám)testeket, amelyekben e feltételek igazak, algebrailag zárt számtesteknek nevezik.

Bizonyítás. A tételt ciklikusan bizonyítjuk be, azaz kimutatjuk, hogy a fenti feltételek bármelyikéből következik a sorrendben következő, és az utolsóból következik az első. Ezzel természetesen bizonyítást nyer az is, hogy bármelyikből következik az összes többi.

(1)-ből következik (2): Ha a gyöke az f polinomnak, akkor a 3.23. Tétel szerint f osztható az $x - a$ elsőfokú polinommal.

(2)-ből következik (3): Az f irreducibilis polinom legalább elsőfokú, így (2) szerint $f = (x - a) \cdot g$ alakú. Az irreducibilitás definíciója szerint e két tényező valamelyike konstans. Mivel az első faktor nem az, ezért a második lesz konstans. Tehát f valóban elsőfokú polinom.

(3)-ból következik (4): A 3.16. Tétel szerint tetszőleges K -beli legalább elsőfokú polinom felbontható irreducibilis tényezők szorzatára. (3) alapján viszont ezek elsőfokúak.

(4)-ből következik (5): A feltétel szerint az f polinom felírható elsőfokú polinomok szorzataként. E polinomok elsőfokú tagjainak az együtthatóit kiemelve az f polinomot egy konstansnak és normált elsőfokú polinomoknak a szorzatára bontottuk fel. Az f normált-ságából következik, hogy a szereplő konstans 1. A szorzásnál a fokok összeadódnak, ezért a tényezők száma megegyezik f fokával. Így éppen a kívánt felbontást kaptuk.

(5)-ből következik (6): A kiindulásul vett felírásból a kívántat úgy kapjuk meg, hogy egybefoglaljuk a megegyező tényezőket. A kitevők összegére vonatkozó állítás ismét abból következtethető, hogy a szorzat foka megegyezik a fokok összegével.

(6)-ból következik (7): Az f polinomnak (6) alapján megadott felírásából következik, hogy a 3.23. Tétel miatt e polinomnak pontosan a felsorolt r darab szám lesz a gyöke. Felhasználva a 3.13. és 3.15. Definíciókat, azt kapjuk, hogy a gyökök száma valóban megegyezik a polinom fokával.

(7)-ből következik (1): Minden polinomnak pontosan annyi gyöke van, amekkora a foka, ezért egy legalább elsőfokú polinomnak legalább egy gyöke van. ■

Megjegyzés. Vegyük figyelembe, hogy a 3.32. Tételben felsorolt feltételek csak algebrailag zárt (szám)testekben igazak. Mint említettük, a komplex számok testéről belátható, hogy ilyen. A feltételek viszont minden számtest esetében ekvivalensek. □

A komplex számok nehezebb kezelhetősége miatt sokszor célszerűbb az algebra alaptételét a valós számokra megfogalmazni. Természetesen ez nem azt jelenti, hogy az algebra

alaptétele a valós számokra is érvényes volna, hanem azt, hogy a valós számokra kimondható egy olyan tétel, amely ekvivalens az algebra alaptételével. Ennek is többféle, ugyan-csak ekvivalens megfogalmazása van. Mindenekelőtt ezeknek az ekvivalenciáját fogjuk kimutatni. Az előbbiekhöz hasonlóan itt sem mondhatjuk ki az ekvivalenciát a valós szám-
testre, csupán tetszőleges számtestre vonatkozó ekvivalenciáról beszélhetünk. Tekintettel arra, hogy az algebra alaptétele nagyon sok számtestre teljesül, ezért — a később szereplő 3.35. Tétel alapján — hasonló igaz erre a tételre is.

3.33. Tétel. *Bármely V számtest esetén $V[x]$ -beli polinomokra ekvivalensek az alábbi állítások:*

- (1) *Minden legalább elsőfokú polinomnak van legfeljebb másodfokú faktora.*
- (2) *Minden irreducibilis polinom legfeljebb másodfokú.*
- (3) *Minden legalább elsőfokú polinom felbontható legfeljebb másodfokú irreducibilis polinomok szorzatára.*

Azokat a (szám)testeket, amelyekben a fenti feltételek teljesülnek, valósan zárt (szám)testeknek nevezik.

Bizonyítás. Mindenekelőtt bebizonyítjuk az első három feltétel ekvivalenciáját.

(1)-ből következik (2): Az f irreducibilis polinomnak feltétel szerint van egy g , legfeljebb másodfokú faktora. Az $f = gh$ felbontásban az első tényező nem konstans, tehát f irreducibilitása alapján a második tényező az. Így f és g foka megegyezik, tehát f legfeljebb másodfokú.

(2)-ből következik (3): A 3.16. Tétel szerint bármely legalább elsőfokú polinom felbontható irreducibilis polinomok szorzatára. Ezek azonban a (2) feltétel alapján legfeljebb másodfokúak.

(3)-ból következik (1): A feltételben biztosított felbontásbeli bármelyik tényező az eredeti polinomnak egy legfeljebb másodfokú faktora lesz. ■

Megjegyzés. Itt is hasonló a helyzet, mint a 3.32. Tétel esetében. E feltételek csak a valósan zárt testek esetében teljesülnek, de bármely számtest esetében ekvivalensek. Azt is jó tudni, hogy a valós számtest valósan zárt (hiszen éppen ez a „valósan zárt” elnevezés alapja). □

Kiegészítés. *A 3.33. Tételben leírt feltételek teljesülése esetén minden páratlan fokú polinomnak van V -ben gyöke, minden V -beli szám vagy egy V -beli szám négyzete, vagy egy V -beli szám négyzetének az ellentettje, és V -beli elemek négyzetének az összege V -beli elem négyzete.*

Bizonyítás. Mivel polinomok szorzatának a foka megegyezik a tényezők fokának az összegével, ezért egy páratlan fokú polinom felbontásában nem lehet minden tényező másodfokú. A feltétel szerint viszont ekkor a polinomnak van elsőfokú faktora, ami a 3.23. Tétel figyelembevételével azt jelenti, hogy van V -beli gyöke.

Legyen most $a \in V$, és tekintsük az $x^4 - a$ polinomot. Ha ennek van V -ben gyöke, akkor a egy V -beli szám negyedik hatványa, tehát egyszersmind egy V -beli szám négyzete

is. Egyébként a feltétel alapján e polinom felbontható két irreducibilis másodfokú polinom szorzatára, amelyekről nyilván feltehető az is, hogy normáltak:

$$x^4 - a = f(x) \cdot g(x).$$

Ebből $(-x)^4 = x^4$ alapján kapjuk az eredeti polinomnak irreducibilis normált faktorokra való két felbontását:

$$x^4 - a = f(x) \cdot g(x) = f(-x) \cdot g(-x).$$

Az egyértelműség miatt itt két eset lehetséges. Ha $g(x) = g(-x)$, akkor $f(x) = f(-x)$ is fennáll. Ebből azt kapjuk, hogy $f(x) = x^2 - b$ és $g(x) = x^2 - c$ alakú ($b, c \in V$). A kapott

$$x^4 - a = (x^2 - b) \cdot (x^2 - c)$$

egyenlőségből formailag pontosan az $y^2 - a$ polinom $V[y]$ -beli felbonthatósága következik, bizonyítva, hogy a egy V -beli szám négyzete. A $g(x) = f(-x)$ esetben, 0-t behelyettesítve $-a = f(0)^2$ adódik, és így a egy V -beli szám négyzetének a negatívja.

A harmadik állításnak a bizonyításához elég belátni, hogy $a, b \in V$ esetén van olyan $c \in V$, amelyre $a^2 + b^2 = c^2$. A $b = 0$ esetben $c = a$ megfelel. Egyébként tekintsük az $f(x) = x^4 - 2ax^2 + a^2 + b^2$ polinomot. Ha ennek van V -ben gyöke, akkor ennek a gyöknek a négyzete gyöke az $x^2 - 2ax + a^2 + b^2$ polinomnak. Ezért $f(x)$ D diszkriminánsának a $\Delta = \sqrt{D}$ négyzetgyökeire $\Delta = \sqrt{(2a)^2 - 4(a^2 + b^2)} = 2b\sqrt{-1} \in V$, azaz $\sqrt{-1} = \frac{\Delta}{2b} \in V$ ($b \neq 0$). Mint láttuk, V -ben minden elem vagy a negatívja négyzet, tehát ebben az esetben minden elem négyzet, azaz $a^2 + b^2$ is. Ha az $f(x)$ polinomnak nincs elsőfokú faktora, akkor (3) szerint felbomlik két másodfokú irreducibilis tényező szorzatára. Létezik tehát $V[x]$ -beli irreducibilis tényezőkre való

$$x^4 - 2ax^2 + a^2 + b^2 = h(x) \cdot g(x)$$

felbontás. Ebből az $x \mapsto -x$ helyettesítéssel a $h(-x) \cdot g(-x)$ faktorokra való felbontás adódik. h és g irreducibilitásának a következtében két eset lehet. Ha $h(-x) = h(x)$ (és $g(-x) = g(x)$), akkor a fentebb látottakhoz hasonlóan az $x^2 - 2ax + a^2 + b^2$ reducibilitása adódik, amikor is ugyanaz a végkövetkeztetés, mint az előző esetben. Egyébként itt is $(x^2 - cx + d) \cdot (x^2 + cx + d)$ alakú felbontást nyerünk; amiből $a^2 + b^2 = d^2$ következik. ■

Megjegyzések

1. Annak a bizonyítása, hogy a Kiegészítésben adott feltételből következnek a 3.33. Tételbeliek, komolyabb algebrai segédeszközöket igényel. Aki tudja, hogy az algebra alaptételének a függvény-tani bizonyítása elég hosszadalmas, annak számára ez várható is, hiszen ennek a feltételnek a valós számtestre való teljesülése az analízis eszközeivel igen egyszerűen bizonyítható. Valóban, a feltétel első része azonnal következik a Bolzano-tételből — felhasználva a polinomok folytonosságát —, a másik két feltétel pedig abból, hogy pontosan a nemnegatív számok állnak elő négyzetként.

2. A Kiegészítésben szereplő második és harmadik feltétel „trükkös” bizonyítása természetessé válik a 3.36. Tétel felhasználásával. Ez a tétel köti ugyanis össze az algebrailag zárt és a valósan zárt testeket. Ezt felhasználva viszont a három tétel együttes bizonyítása válna szükségessé, ami a bizonyítási folyamatot bonyolítaná el.

3. A Kiegészítésben szereplő harmadik feltétel egyáltalán nem tűnik természetesnek. Hiszen a valós számtest esetén világosan „látható”, hogy páratlan fokú polinomnak van valós gyöke, és minden pozitív valós szám egy valós szám négyzete. Mégis szükség van erre. Ez a feltétel burkoltan

azt mondja ki, hogy pozitív számok összege is pozitív. Létezik olyan maximális számtest, amelyiknek eleme $\sqrt{-3}$, de nem eleme $\sqrt{-1}$. Erre a számtestre igaz e pont első két állítása, de a harmadik nem. Éppen ezért e számtest felett akármilyen $n \in \mathbb{N}$ -re van olyan irreducibilis polinom, amelynek a foka 2^n . (Ennek a belátása egyáltalán nem egyszerű.) \square

A továbbiakban a 3.32. és 3.33. Tételek egybevetése lesz a célunk. Mindenekelőtt az szükséges, hogy a V és a K számtest között olyan kapcsolat álljon fenn, mint a valós és a komplex számtest között.

3.34. Tétel. *Tegyük fel, hogy az i képzetes egység nem eleme a V számtestnek. Ekkor annak a legszűkebb — $V(i)$ -vel jelölt — számtestnek, amelyik mind a V számtestet, mind az i számot tartalmazza, az elemei egyértelműen $a + bi$ alakban írhatók, ahol $a, b \in V$. Az $a + bi \mapsto a - bi$ megfeleltetés rendelkezik a komplex konjugálás tulajdonságaival.*

Bizonyítás. Az $a + bi$ alakú számok — a műveleti zártság alapján — benne vannak minden olyan számtestben, amely mind a V -t, mind az i -t tartalmazza. A felírás egyértelműsége azonnal következik abból, hogy $i \notin V$. Egyszerű számolással belátható, hogy az adott alakú számok gyűrűt alkotnak. A megadott leképezés tulajdonságait szó szerint ugyanúgy bizonyíthatjuk, mint a komplex számok esetén a konjugálás tulajdonságait. Ebből pedig az $N = a^2 + b^2$ jelöléssel az

$$\frac{a + bi}{c + di} = \frac{(a + bi) \cdot (c - di)}{(c + di) \cdot (c - di)} = \frac{ac + bd}{N} + \frac{bc - ad}{N}$$

összefüggés alapján következik, hogy a megadott alakú számok testet alkotnak, mivel V is számtest. \blacksquare

Megjegyzés. Természetesen ez nem ugyanaz, mint a komplex számok bevezetése. A V számtestről azt sem kell feltenni, hogy a 3.32. Tétel állításai teljesülnek rá. Léteznek ugyanis olyan V számtestek is, amelyek nem tartalmazzák i -t, de az elemeik között nemcsak valós számok vannak. Például racionális a és b esetében az $a + b\sqrt{-3}$ alakú számok számtestet alkotnak, e számtest nem tartalmazza i -t, de elemei között vannak nem valós számok is. \square

3.35. Tétel. *Legyen $K = V(i)$, és rendeljük hozzá a K -beli együtthatós $f = a_0 + a_1x + \dots + a_nx^n$ polinomhoz az $\overline{f} = \overline{a_0} + \overline{a_1}x + \dots + \overline{a_n}x^n$ polinomot, ahol $a \mapsto \overline{a}$ a 3.34. Tételbeli megfeleltetést jelöli. Ekkor: $\overline{\overline{f}} = f$; $\overline{f + g} = \overline{f} + \overline{g}$; $\overline{f \cdot g} = \overline{f} \cdot \overline{g}$; $\overline{f \circ g} = \overline{f} \circ \overline{g}$; $\overline{f(a)} = \overline{f}(\overline{a})$. Továbbá: $f + \overline{f}, f \cdot \overline{f} \in V[x]$ és $\overline{f} = f$ akkor és csak akkor, ha $f \in V[x]$. Végül $\text{gr}(\overline{f}) = \text{gr}(f)$.*

Bizonyítás. Az első négy összefüggés — a definíciók alapján — egyszerűen kiszámítható. Az ötödik összefüggés — azaz $\overline{f(a)} = \overline{f}(\overline{a})$ — az előzőnek speciális esete. A definícióból azonnal látható, hogy $\overline{f} = f$ pontosan akkor teljesül, ha f együtthatói mind V -ből valók. Ebből pedig azonnal adódik az is, hogy \overline{f} -nak és f -nek az összege is és a szorzata is V -beli együtthatós. Az utolsó állítás ugyancsak nyilvánvaló. \blacksquare

3.36. Tétel. *Legyen $K = V(i)$. A V testre akkor és csak akkor teljesülnek a 3.33. Tételbeli tulajdonságok, ha a K testre teljesülnek a 3.32. Tételbeliek.*

Bizonyítás. Tegyük fel először, hogy a K számtestre érvényesek a 3.32. Tételben kimondott tulajdonságok, és legyen $f \in V[x]$ legalább elsőfokú polinom. A 3.32. (2) tulajdonság alapján f -nek van $K[x]$ -ben egy g elsőfokú faktora. E faktorról azt is feltehetjük, hogy normált. Ha $\bar{g} = g$, akkor a 3.6. Tétel szerint az $\frac{f}{g}$ hányados $V[x]$ -beli, és így f -nek van $V[x]$ -ben egy elsőfokú faktora. Ellenkező esetben a normáltság alapján \bar{g} és g nem egymás konstansszorosai, és így $\bar{g} \cdot g$ is osztója f -nek. Tekintettel arra, hogy ez a szorzat a 3.35. Tétel szerint V -beli együtthatós és nyilván másodfokú, ezért — ugyancsak a 3.6. Tétel következtében — létezik f -nek a $V[x]$ -ben egy másodfokú faktora. Ezzel tehát beláttuk, hogy V -re teljesül a 3.33. Tétel (1) pontja.

Tegyük most fel, hogy a V testre teljesülnek a 3.33. Tételben megadott feltételek. Mindenekelőtt megmutatjuk, hogy ekkor minden V -beli együtthatós másodfokú polinom felbontható két K -beli együtthatós elsőfokú polinom szorzatára. Nyilván elegendő, ha ezt normált polinomokra bizonyítjuk. Legyen a vizsgált normált másodfokú polinom diszkriminánsa D . A 3.33. Tétel Kiegészítése alapján vagy D , vagy $-D$ egy V testbeli elem négyzete. Ebből pedig azonnal következik az állítás, ha a másodfokú polinomoknak a felbontását a vizsgált valós együtthatós esetre végigvisszük. Legyen most f tetszőleges K -beli együtthatós legalább elsőfokú polinom. Az $f \cdot \bar{f}$ szorzat a 3.34. Tétel szerint V -beli együtthatós, és nyilván legalább elsőfokú. Így a 3.33. Tétel (3) pontja szerint felbontható legfeljebb másodfokú faktorok szorzatára. Ezek a másodfokú faktorok az előbb bizonyítottak alapján tovább bonthatók $K[x]$ -ben elsőfokú faktorokra. A polinomok egyértelmű felbontására vonatkozó tétel szerint ebből adódik az f -nek elsőfokú faktorokra való felbontása, és így teljesül a 3.32. Tétel (4) feltétele. ■

9. Racionális és egész együtthatós polinomok

Az eddigiekben lényegében a valós és komplex együtthatós polinomokat vizsgáltuk. Térjünk most át a harmadik „fontos” számtest, a racionális számtest feletti polinomok vizsgálatára. Mindenekelőtt azt mutatjuk meg, hogy a felbontási és gyökmeghatározási kérdések visszavezethetők egész együtthatós polinomok vizsgálatára.

A racionális együtthatós polinomok körében érvényes az egyértelmű faktorizáció, hiszen bármely számtest feletti polinomgyűrűben létezik maradékos osztás, aminek a segítségével az egyértelmű faktorizáció bizonyítható. Látni fogjuk, hogy az egyértelmű faktorizáció az egész együtthatós polinomok esetében is létezik. Ezt viszont nem tudjuk a maradékos osztás segítségével bizonyítani, mert a maradékos osztás itt nem végezhető el. Ennek a közvetlen belátása igen kellemetlen volna, mert azt kellene megmutatni, hogy a maradéknak „valamije” csökken; ez a valami természetes számmal volna mérhető. Ehelyett azt mutatjuk meg, hogy a maradékos osztásnak egyik következménye itt nem teljesül.

A maradékos osztás segítségével bizonyítottuk, hogy minden (polinom)ideál főideál. Nos ez az egész együtthatós polinomok esetében nem igaz. Tekintsük azokat az egész

együtthatós polinomokat, amelyeknek a konstans tagja páros. Világos, hogy két ilyennek a különbsége is és egy ilyennek a polinomszorosa is ilyen, tehát ezek egy I ideált alkotnak. Tekintsünk egy (f) főideált (ahol $f \in \mathbb{Z}[x]$), ami I -t tartalmazza. Ekkor $2, x \in (f)$ miatt léteznek olyan $u, v \in \mathbb{Z}[x]$ polinomok, amelyekre $2 = u(x) \cdot f(x)$ és $x = v(x) \cdot f(x)$. Mivel az első szorzat foka 0, ezért a tényezők foka is csak 0 lehet, vagyis $f(x) = c \in \mathbb{Z}$ konstans. A második egyenlőség szerint $x = v(x) \cdot c$; ami csak akkor lehet, ha $c = 1$ vagy $c = -1$. Ezek a polinomok viszont nem elemei az I ideálnak.

Ezért kell az $\mathbb{Z}[x]$ -beli polinomok esetére az egyértelmű felbontást „kerülő úton” bizonyítani.

3.16. Definíció. Egy egész együtthatós polinomot primitív polinomnak nevezünk, ha főegyütthatója pozitív, és együtthatóinak legnagyobb közös osztója 1. ■

Megjegyezzük, hogy a primitív polinom definíciójánál általában nem teszik fel azt, hogy a főegyüttható pozitív, de ez a megkötés — az egyértelműség érdekében — hasznos.

3.37. Tétel. Minden racionális együtthatós, 0-tól különböző polinom egyértelműen felbontható egy racionális számnak és egy primitív polinomnak a szorzatára.

Bizonyítás. Legyen $f(x)$ egy tetszőleges racionális együtthatós polinom, és legyen b az $f(x)$ együtthatói nevezőinek a szorzata. Ekkor $b \cdot f(x)$ egész együtthatós polinom. Ha kiemeljük e polinom együtthatóinak az a legnagyobb közös osztóját, akkor egy olyan $g(x)$ egész együtthatós polinomot kapunk, amelyben az együtthatók legnagyobb közös osztója 1. A kiemelt szám helyett esetleg $(-a)$ -t kiemelve elérhető az is, hogy $g(x)$ primitív polinom legyen. Így $f(x)$ az $\frac{a}{b}$ (vagy $-\frac{a}{b}$) racionális számnak és a $g(x)$ primitív polinomnak a szorzata.

Tekintsük egy racionális együtthatós polinomnak két ilyen felírását:

$$\frac{a}{b} \cdot (a_0 + a_1x + \cdots + a_nx^n) = \frac{c}{d} \cdot (c_0 + c_1x + \cdots + c_nx^n).$$

$b \cdot d$ -vel szorozva azt kapjuk, hogy bármely i indexre $ad \cdot a_i = cb \cdot c_i$. Ebből az egyenlőségből adódik, hogy a bal oldalon, illetve a jobb oldalon álló számok legnagyobb közös osztója megegyezik. A primitív polinomok definíciója alapján tehát ad és bc vagy megegyezik, vagy ellentétes előjelű. Ez utóbbi azonban lehetetlen, mert a_n és c_n egyike sem negatív és $ada_n = bcc_n$. Tehát $\frac{a}{b} = \frac{c}{d}$. Ebből a szereplő primitív polinomok egyértelműsége is következik. ■

A továbbiakban az egész együtthatós polinomok felbontásával foglalkozunk. Ehhez itt is szükségünk van az oszthatóság fogalmára. Az oszthatóságot ugyanúgy értelmezzük, mint tetszőleges polinomok esetében tettük, csupán most a hányadostól is megköveteljük, hogy egész együtthatós legyen.

3.17. Definíció. A $g(x)$ egész együtthatós polinom osztója az $f(x)$ egész együtthatós polinomnak, ha létezik olyan egész együtthatós $h(x)$ polinom, amelyre $f(x) = g(x) \cdot h(x)$. ■

Be fogjuk látni, hogy az egész együtthatós polinomok esetében nem osztója bármely konstans minden polinomnak. Azt is kimutatjuk, hogy a prímszámok prímtulajdonsága az egész együtthatós polinomok körében is teljesül.

3.38. Tétel. *Egy egész szám pontosan akkor osztója egy egész együtthatós polinomnak, ha minden együtthatójának osztója. Ha egy prímszám osztója két egész együtthatós polinom szorzatának, akkor osztója valamelyik tényezőnek is.*

Bizonyítás. A c egész szám definíció szerint akkor osztója az $f = a_0 + a_1x + \dots + a_nx^n$ polinomnak, ha létezik olyan $g = b_0 + b_1x + \dots + b_nx^n$ polinom, amelyre $c \cdot g = f$. Ez ekvivalens azzal a feltétellel, hogy $a_i = cb_i$ ($i = 0, 1, \dots, n$), ami pontosan azt jelenti, hogy c osztója az f polinom minden együtthatójának.

A második állítás bizonyítására legyen a két polinom

$$f(x) = a_0 + a_1x + \dots + a_nx^n \quad \text{és} \quad g(x) = b_0 + b_1x + \dots + b_kx^k,$$

továbbá ezek szorzata $h(x) = c_0 + c_1x + \dots + c_ix^i + \dots$. Legyen most p egy olyan prímszám, amelyik osztója a szorzatpolinomnak, azaz minden egyes c_i -nek. Ha p osztója az $f(x)$ polinom mindegyik együtthatójának, akkor az állítás igaz. Amennyiben p nem osztója az $f(x)$ mindegyik együtthatójának, akkor van egy legkisebb indexű együttható, amelynek nem osztója. Legyen ez az index u (azaz $p \mid a_0, \dots, p \mid a_{u-1}$, de $p \nmid a_u$). E feltétel mellett azt fogjuk bizonyítani, hogy p osztója a $g(x)$ polinomnak. Induktíven bizonyítunk (nem teljes indukcióval, mert a polinomnak véges sok együtthatója van). Tegyük fel, hogy p osztója a $g(x)$ polinom minden v -nél kisebb indexű együtthatójának, és azt bizonyítjuk be, hogy b_v -nek is osztója. (Speciálisan a $v = 0$ esetben „kezdőlépésként” azt kapjuk, hogy b_0 -nak osztója a p .) Tekintsük a szorzatpolinom $u + v$ indexű együtthatóját:

$$c_{u+v} = a_0b_{u+v} + \dots + a_{u-1}b_{v+1} + a_ub_v + a_{u+1}b_{v-1} + \dots + a_{u+v}b_0.$$

A bal oldalon álló szám a szorzatpolinomra vonatkozó feltétel szerint osztható p -vel. A jobb oldali összeg első u tagja az $f(x)$ együtthatóira, az összeg utolsó v tagja pedig a $g(x)$ együtthatóira kirótt feltétel alapján osztható p -vel. Egy p -vel osztható számból két p -vel osztható számot kivonva ugyancsak p -vel osztható számot nyerünk, ezért a_ub_v is osztható p -vel. A prímtulajdonság alapján tehát a szorzat valamelyik tényezője is osztható p -vel. Az $f(x)$ együtthatóira kirótt feltétel alapján ez csak a b_v lehet, mint állítottuk. ■

A fenti eredményből a primitív polinomok egy igen fontos tulajdonsága adódik:

3.39. Tétel (Gauss lemmája). *Két primitív polinom szorzata is primitív polinom.*

Bizonyítás. Ha két egész együtthatós polinom szorzata nem primitív, akkor osztható egy 1-nél nagyobb egész számmal, és így ennek egy prímosztójával. A 3.38. Tétel alapján ez a prímszám osztója valamelyik tényezőnek is, ami azt jelenti, hogy ez a tényező nem primitív. Tehát, amennyiben mindkét tényező primitív, akkor a szorzat is csak primitív lehet. ■

3.40. Tétel. *Ha egy primitív polinom irreducibilis az egész együtthatós polinomok körében, akkor irreducibilis a racionális együtthatós polinomok körében is.*

Bizonyítás. Legyen az $f(x)$ primitív polinom a $g(x)$ és $h(x)$ racionális együtthatós polinomok szorzata. A 3.37. Tétel szerint e polinomok felírhatók $g(x) = r \cdot g_0(x)$, illetve $h(x) = s \cdot h_0(x)$ alakban, alkalmas racionális számokkal és primitív polinomokkal. A felbontásból $f(x) = (r \cdot s) \cdot (g_0(x) \cdot h_0(x))$ adódik, ahol a zárójelben álló polinom a 3.39. Tétel alapján primitív. A 3.37. Tételben levő egyértelműség miatt ekkor $f(x) = g_0(x) \cdot h_0(x)$ (és $r \cdot s = 1$). Az irreducibilitás alapján itt valamelyik tényező konstans. Ekkor pedig konstans a megfelelő racionális együtthatós polinom is, hiszen mind r , mind s racionális szám. ■

3.41. Tétel. *Ha egy egész együtthatós polinom a racionális együtthatós polinomok körében osztható egy primitív polinommal, akkor az egész együtthatós polinomok körében is osztható vele.*

Bizonyítás. Legyen $f(x)$ egy primitív polinom, amely osztója a $g(x)$ egész együtthatós polinomnak, azaz $g(x) = f(x) \cdot (r \cdot h(x))$, ahol r racionális szám, és $h(x)$ is primitív polinom (ez a 3.37. Tétel alapján tehető fel). A 3.39. Tétel szerint $f(x) \cdot h(x)$ primitív, és a 3.37. Tétel következtében r egész szám. ■

3.42. Tétel. *Az egész együtthatós polinomok körében minden irreducibilis primitív polinom rendelkezik a prímtulajdonsággal.*

Bizonyítás. Tegyük fel, hogy az $f(x)$ irreducibilis primitív polinom $\mathbb{Z}[x]$ -ben osztója a $g(x)$ és $h(x)$ egész együtthatós polinomok szorzatának. Így a 3.40. és a 3.9. Tételek alapján $f(x)$ osztója például $g(x)$ -nek a racionális együtthatós polinomok körében. Ebből viszont a 3.41. Tételt alkalmazva kapjuk, hogy az oszthatóság az egész együtthatós polinomok körében is fennáll. ■

3.43. Tétel. *Az egész együtthatós polinomok körében érvényes az egyértelmű prímtegyezős felbontás; minden polinom (sorrendtől eltekintve) egyértelműen bontható fel prímszámoknak és irreducibilis primitív polinomoknak a szorzatára.*

Bizonyítás. A 3.37. Tétel alapján minden egész együtthatós polinom felbontható egy egész számnak és egy primitív polinomnak a szorzatára. Az egész számot a számelmélet alaptétele szerint felbonthatjuk prímszámok szorzatára. A 3.40. Tétel alapján a primitív polinomot felbonthatjuk irreducibilis primitív polinomok szorzatára. A 3.38. és a 3.42. Tételek következtében a felbontás egyértelműsége a 3.16. Tételhez hasonlóan bizonyítható. ■

A 3.37. Tétel következtében a racionális együtthatós irreducibilis polinomok vizsgálata visszavezethető az egész együtthatós irreducibilis polinomok vizsgálatára. Ez utóbbi azért könnyebb, mert figyelembe lehet venni az együtthatókra vonatkozó oszthatósági feltételeket. Az egész együtthatós polinomok körében számos olyan feltétel ismeretes, amely biztosítja az irreducibilitást. Ezek közül az úgynevezett „irreducibilitási kritériumok” közül itt

most egyet tárgyalunk, a *Schönemann–Eisenstein*-féle irreducibilitási kritériumot. A későbbiekben ugyanis csak ezt használjuk fel.

3.44. Tétel. *Ha egy egész együtthatós polinomhoz található egy olyan p prímszám, amelyre a következő feltételek teljesülnek:*

- (1) p nem osztója a főegyütthatónak,
- (2) p osztója az összes többi együtthatónak,
- (3) p^2 nem osztója a polinom konstans tagjának,

akkor a polinom irreducibilis.

Bizonyítás. Legyen az adott egész együtthatós polinom:

$$f(x) = a_n x^n + \cdots + a_1 x + a_0.$$

Tegyük fel, hogy $f(x)$ -et felbontottuk az ugyancsak egész együtthatós

$$g(x) = b_u x^u + \cdots + b_0 \quad \text{és} \quad h(x) = c_v x^v + \cdots + c_0$$

polinomok szorzatára. A (2) és (3) feltételeket figyelembe véve $a_0 = b_0 \cdot c_0$ következtében a b_0 és c_0 közül az egyik osztható p -vel, a másik pedig nem. (Ha mindkettő osztható lenne p -vel, akkor a szorzatuk osztható lenne p^2 -tel is.) A szimmetria alapján feltehető, hogy b_0 nem osztható p -vel, és c_0 többszöröse p -nek. Az (1) feltétel szerint p nem osztója az $f(x)$ polinomnak, és így nem lehet osztója a $h(x)$ polinomnak sem. Ez azt jelenti, hogy a $h(x)$ polinomnak van olyan együtthatója, amelyik nem osztható p -vel. Legyen c_i a legkisebb indexű ilyen együttható, azaz legyen c_0, c_1, \dots, c_{i-1} többszöröse p -nek, de c_i ne legyen p -vel osztható. Ekkor az

$$a_i = b_0 \cdot c_i + b_1 \cdot c_{i-1} + \cdots + b_i \cdot c_0$$

összefüggés alapján a következőket állapíthatjuk meg:

A jobb oldalon az első tag egyik tényezője sem osztható p -vel, ezért a prímtulajdonság szerint ez a tag sem osztható p -vel. A többi tagban a feltétel szerint a második tényezők mindig p -vel oszthatók, így e tagok, valamint e tagok összege is osztható p -vel. Egy p -vel osztható és egy p -vel nem osztható szám összege sem lehet p -vel osztható, ezért a_i nem osztható p -vel. A (2) feltétel szerint ebből viszont $i = n$ következik. Eszerint $n \leq v$, és a szorzat fokára vonatkozó összefüggés alapján $v \leq n$. Így $v = n$, azaz $g(x)$ konstans. Ezzel bebizonyítottuk az irreducibilitást. ■

Megjegyzés. Ha speciálisan normált egész együtthatós polinomokat tekintünk, akkor az (1) feltétel eleve teljesül, ezért csak a (2) és (3) feltételt kell ellenőrizni. □

Kiegészítés. *Tetszőleges n természetes számhoz létezik a racionális számtest felett irreducibilis n -edfokú polinom.*

Bizonyítás. Mint tudjuk, elegendő egész együtthatós polinomokat vizsgálni. Kimutatjuk, hogy például az $x^n - 2$ polinom mindig irreducibilis a racionális számtest felett. A $p = 2$ választással alkalmazhatjuk a 3.44. Tételt. A 3.40. Tétel alapján pedig következik az állítás. ■

3.45. Tétel. *Tetszőleges n természetes számhoz létezik olyan $F_n(x)$ egész együtthátós normált polinom, amelynek gyökei pontosan az n -edik primitív egységgyökök. Ezt a polinomot az n -edik körosztási polinomnak nevezik.*

Ha n prímszám, akkor $F_n(x)$ irreducibilis.

Bizonyítás. Tudjuk, hogy az $x^n - 1$ polinom gyökei pontosan az n -edik egységgyökök. Az $x^{(u+1)k} - 1 = x^{uk}(x^k - 1) + (x^{uk} - 1)$ összefüggés alapján teljes indukcióval be lehet bizonyítani azt, hogy $x^k - 1$ osztója $(x^n - 1)$ -nek, amennyiben k osztója n -nek. Ha $n = kq + r$ — ahol r nemnegatív, k -nál kisebb egész szám —, akkor az $x^n - 1 = x^r(x^{kq} - 1) + x^r - 1$ összefüggés felhasználásával a következőket nyerjük: a fenti típusú polinomokra a maradékos osztás ugyanúgy elvégezhető, mint a kitevőkre. Ennek alapján ilyen polinomok legnagyobb közös osztójára az $(x^n - 1, x^k - 1) = x^{(n,k)} - 1$ összefüggés teljesül. Ebből pedig az következik, hogy $x^k - 1$ pontosan akkor osztója az $x^n - 1$ polinomnak, ha $k \mid n$ teljesül. Tekintsük most az $n > 1$ természetes számot, és vegyük ennek összes n -nél kisebb d osztóját. Legyen ezekre a d osztókra $g_n(x)$ az összes $x^d - 1$ alakú polinom legkisebb közös többszöröse. A fenti polinomok primitívek, ezért $g_n(x)$ is primitív. Mivel $\mathbb{Z}[x]$ -ben érvényes az egyértelmű faktorizáció, ezért $g_n(x)$ osztója az $x^n - 1$ polinomnak, és így normált. Mivel polinomok szorzatának a főegyütthatója megegyezik a tényezők főegyütthatóinak a szorzatával, ezért $F_n(x) = \frac{x^n - 1}{g_n(x)}$ hányadosuk is normált.

A konstrukció alapján $F_n(x)$ -nek pontosan azok az n -edik egységgyökök a gyökei, amelyek semmilyen $0 < k < n$ esetén sem k -edik egységgyökök; ezek éppen a primitív n -edik egységgyökök. A fenti előállítást felhasználva, elemi számelméleti módszerekkel bebizonyítható, hogy $\text{gr}(F_n(x)) = \varphi(n)$, ahol φ az Euler-féle függvényt jelöli. Így e polinom minden gyöke egyszeres gyök, hiszen $\varphi(n)$ a különböző n -edik primitív egységgyökök száma. Az n -edik primitív egységgyököket ábrázoló pontok megszerkesztése ekvivalens a körív n egyenlő részre való felosztásával, ezért nevezik e polinomot az n -edik körosztási polinomnak.

Legyen most $n = p$ prímszám, ekkor $F_p(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + \dots + x + 1$. A 3.17. Tétel alapján tetszőleges $f(x)$ polinom irreducibilitása ekvivalens az $f(x+1)$ polinoméval, és így elegendő, ha kimutatjuk, hogy az $\frac{(x+1)^p - 1}{x}$ polinom irreducibilis. Ez utóbbi polinomot a binomiális tétel (4.13. Tétel) alapján az

$$x^{p-1} + \binom{p}{1}x^{p-2} + \dots + \binom{p}{i}x^{p-i-1} + \dots + \binom{p}{p-1}$$

alakban írhatjuk. Itt a főegyüttható 1, tehát nem osztható p -vel. Az összes többi együttható osztható p -vel (hiszen $p!$ osztható p -vel, de p -nél kisebb pozitív egész i -re sem $i!$, sem $(p-i)!$ nem osztható p -vel). Végül $\binom{p}{p-1} = p$ természetesen nem osztható p^2 -tel. Így a 3.44. Tétel bizonyítja az $F_p(x)$ körosztási polinom irreducibilitását. ■

Megjegyzés. Algebrai módszerekkel megmutatható, hogy $F_n(x)$ tetszőleges n esetén irreducibilis a racionális számtest felett. Ennek a bizonyítására az algebrai struktúrák tárgyalásakor kerül sor. \square

Végezetül a racionális együtthatós polinomok racionális gyökeivel foglalkozunk. Nyilván itt is elegendő egész együtthatós polinomokra szorítkozni.

3.46. Tétel. *Egy redukált alakban megadott tört csak akkor lehet gyöke egy egész együtthatós polinomnak, ha számlálója osztója a polinom konstans tagjának, nevezője pedig a polinom főegyütthatójának. Egy egész szám csak akkor lehet gyöke a polinomnak, ha osztója a polinom konstans tagjának. Normált egész együtthatós polinomnak minden racionális gyöke egész.*

Bizonyítás. Mindenekelőtt belátjuk, hogy a két utóbbi állítás azonnal következik az elsőből. Ha egy egész számot redukált tört alakban felírunk, akkor ez azt jelenti, hogy nevezője 1. Az 1 minden egész számnak osztója, ezért a főegyütthatóra vonatkozó feltétel mindig teljesül. A feltétel másik része pedig éppen az, ami a tételben szerepel. Ha pedig a polinom normált, akkor a feltétel szerint csak olyan racionális szám lehet gyöke, amelynek a nevezője — redukált alakban — osztója 1-nek, azaz csak egész szám lehet.

Az első állítás bizonyításához tegyük fel, hogy az

$$f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$$

egész együtthatós n -edfokú polinomnak gyöke a redukált alakban felírt $\frac{p}{q}$ tört (azaz p és q relatív prímek). Ekkor behelyettesítés és q^n -nel való szorzás után a

$$0 = q^n \cdot f\left(\frac{p}{q}\right) = a_0q^n + a_1q^{n-1}p + \cdots + a_{n-1}qp^{n-1} + a_np^n$$

összefüggéshez jutunk. A bal oldalon álló 0 osztható p -vel is és q -val is. A jobb oldalon az első tag kivételével a többiről tudjuk, hogy osztható p -vel, és az utolsó tag kivételével a többiről tudjuk, hogy osztható q -val. Az oszthatóság tulajdonságaiból azonnal következik tehát, hogy $q \mid (a_np^n)$ és $p \mid (a_0q^n)$. Mivel p és q relatív prímek, ezért — például az egyértelmű prímtenyezős felbontást figyelembe véve — azt kapjuk, hogy $q \mid a_n$ és $p \mid a_0$. ■

Megjegyzések

1. Természetesen — az irreducibilitási kritériumokhoz hasonlóan — itt is csak egy szükséges feltételt adtunk meg. Ennek a segítségével azonban véges számú próbálkozás után meg lehet határozni bármely adott egész együtthatós polinom összes racionális gyökét.

2. Az első állítás itteni bizonyításához csak középiskolai ismereteket használtunk fel. A 3.41. Tétel felhasználásával viszont sokkal egyszerűbben célt érhetünk. Ha ugyanis $\frac{p}{q}$ gyöke az $f(x)$ polinomnak, akkor $qx - p$ osztója $f(x)$ -nek $\mathbb{Q}[x]$ -ben. A feltételből azonnal kapjuk, hogy $qx - p$ primitív; az idézett tétel szerint tehát $\mathbb{Z}[x]$ -ben is fennáll az oszthatóság. Ez pedig pontosan azt jelenti, hogy p osztója a_0 -nak és q osztója a_n -nek. \square

10. Euklideszi gyűrűk

Az eddigiekben három olyan esettel találkoztunk, amikor érvényes volt az egyértelmű faktorizáció. Ezek közül kettőben a bizonyítás a maradékos osztáson alapult. Ezt a fontos esetet érdemes külön szemügyre venni, és a lényeges pontokat külön megfogalmazni.

3.18. Definíció. Elemek egy R halmazát gyűrűnek nevezzük, ha értelmezett a halmazban két kétváltozós művelet, az összeadás (jele $+$), és a szorzás (jele \cdot vagy egymás mellé írás), amelyekre a következők igazak:

(1) Az összeadás kommutatív és asszociatív; továbbá elvégezhető a kivonás, azaz egyértelműen megoldható az $a + x = b$ egyenlet, a megoldást $b - a$ jelöli. ($0 = a - a$.)

(2) A szorzás asszociatív és igaz mindkét disztributivitás.

Amennyiben a szorzás kommutatív, akkor kommutatív gyűrűről beszélünk; de ilyenkor is sokszor elhagyjuk a „kommutatív” jelzőt.

Ha egy gyűrűben $a \cdot b = 0$ csak akkor állhat fenn, ha vagy $a = 0$, vagy $b = 0$, akkor azt mondjuk, hogy a *gyűrű nullosztómentes*. Kommutatív nullosztómentes gyűrű neve *integritási tartomány*.

Ha az R gyűrűben van „egységelem”, azaz olyan e elem, hogy bármely $a \in R$ esetén $e \cdot a = a \cdot e = a$, akkor *egységelemes gyűrűről* beszélünk.

Egy R integritási tartományt *euklideszi gyűrűnek* nevezünk, ha nemnulla elemeire értelmezett egy φ úgynevezett *euklideszi norma*, amelynek értékei természetes számok és kielégítik a maradékos osztás tulajdonságot:

Minden $a, b \in R$ esetén, ha $b \neq 0$, léteznek olyan $q, r \in R$, amelyekre $a = b \cdot q + r$ teljesül és vagy $\varphi(r) < \varphi(b)$, vagy $r = 0$. ■

3.47. Tétel. *Az euklideszi gyűrű egységelemes, érvényes benne az egyértelmű faktorizáció. A minimális normájú elemek minden elemnek osztói.*

Bizonyítás. Legyen R euklideszi gyűrű a φ euklideszi normával. Mivel a norma értékkészlete a természetes számok halmaza, ezért van olyan $b \in R$, amelyre $\varphi(b)$ minimális. A maradékos osztás alapján léteznek olyan $q, r \in R$, amelyekre $b = bq + r$ és $\varphi(b)$ minimalitása következtében csak $r = 0$ lehetséges. Így $b = be$ alakú. Ezért tetszőleges $a \in R$ esetén $ba = bea$, azaz $b(a - ea) = 0$, s a nullosztómentesség miatt $a - ea = 0$, tehát $a = ea$. Ebből ugyanígy látható, hogy tetszőleges $c \in R$ esetén $c = ce$ is igaz, vagyis e egységelem. Ebből hasonló gondolatmenettel belátható, hogy minden minimális normájú elem osztója az egységelemnek; és így minden elemnek.

Az egyértelmű faktorizáció bizonyítása hasonlóképpen történhet, mint a testbeli együtthatós polinomok esetében, itt a polinomideálokkal analóg részhalmaz neve egyszerűen csak ideál. ■

3.48. Tétel. *Ha az R egységelemes integritási tartományban érvényes az egyértelmű faktorizáció, akkor a felette vett polinomgyűrűben is érvényes az egyértelmű faktorizáció.*

Bizonyítás. Lényegében ugyanúgy történik, mint az egész együtthatós polinomok esetében. Az egyetlen különbség az, hogy a primitív polinomok főegyütthatójáról nem is tehetjük fel a pozitivitást; hiszen annak tetszőleges gyűrűben nincs értelme. ■

Végezetül a polinomgyűrűk egy igen alapvető tulajdonságát mutatjuk meg, amely annak általánosítása, hogy minden racionális együtthatós polinom tekinthető valós együtthatós, és minden valós együtthatós tekinthető komplex együtthatós.

3.49. Tétel. *Legyenek R és S egységelemes integritási tartományok, és $\varphi : R \rightarrow S$ tetszőleges homomorfizmus. Ekkor létezik olyan egyértelmű $\psi : R[x] \rightarrow S$ homomorfizmus, amely az R elemein megegyezik φ -vel és x -et az S egy előre megadott s elemére képezi le.*

Bizonyítás. ψ egyértelműsége világos, hiszen $\sum a_i x^i$ képe csak $\sum a_i s^i$ lehet. Az, hogy ez valóban homomorfizmus, abból következik, hogy minden polinom egyértelműen meghatározza az együtthatóit. ■

Az, hogy ez a tulajdonság „jellemző” a polinomgyűrűre, a következőket jelenti:

3.50. Tétel. *Legyen R_1 az R -et tartalmazó integritási tartomány és $r \in R_1$. Ha tetszőleges S egységelemes integritási tartományhoz, $\varphi : R \rightarrow S$ homomorfizmushoz és S -beli s elemhez létezik olyan egyértelműen meghatározott $\psi : R_1 \rightarrow S$ homomorfizmus, amely az R elemein megegyezik φ -vel és $\psi(r) = s$, akkor az a $\psi_1 : R[x] \rightarrow R_1$ homomorfizmus, amely az $f(x) \in R[x]$ polinomhoz az $f(r)$ behelyettesítési értéket rendeli, izomorfizmus.*

Ha S -nek $R[x]$ -et, φ -nek az R identitását és s -nek x -et választjuk, akkor a kapott ψ homomorfizmus a ψ_1 inverze (azaz $\psi\psi_1$ az R_1 -nek és $\psi_1\psi$ az $R[x]$ -nek az identikus leképezése).

Bizonyítás. Tudjuk, hogy ezek a tulajdonságok teljesülnek az $R_1 = R[x]$ és $r = x$ választással.

Tegyük most fel azt, hogy az R -et tartalmazó R_1 integritási tartomány és ennek egy r eleme rendelkezik a kirótt tulajdonsággal. Válasszuk S -t $R[x]$ -nek, legyen φ az a homomorfizmus, amelyik R elemeihez önmagukat rendeli hozzá, és legyen $s = x$.

Feltétel szerint létezik (pontosan egy) olyan $\psi : R_1 \rightarrow R[x]$ homomorfizmus, amely R minden elemét önmagára képezi és $\psi(r) = x$. Tekintettel arra, hogy ψ és ψ_1 mindegyike homomorfizmus, ezért kompozíciójuk — ha létezik — ugyancsak az. Mivel $R_1 \xrightarrow{\psi} R[x]$ és $R[x] \xrightarrow{\psi_1} R_1$, ezért létezik mind az $\alpha = \psi_1\psi : R_1 \rightarrow R_1$, mind a $\beta = \psi\psi_1 : R[x] \rightarrow R[x]$ homomorfizmus.

Mivel ψ és ψ_1 mindegyike identikus R -en, ezért α és β is rendelkezik ezzel a tulajdonsággal. Ezen felül $\alpha(r) = \psi_1(\psi(r)) = \psi_1(x) = r$ és $\beta(x) = \psi(\psi_1(x)) = \psi(r) = x$. Ezekkel a tulajdonságokkal, megfelelően, az R_1 és az $R[x]$ identikus leképezései is rendelkeznek.

Alkalmazzuk most az egyértelműségi feltételt az R_1 gyűrűre. Eszerint pontosan egy olyan $\alpha : R_1 \rightarrow R_1$ leképezés van, amelyik R -en identitás és r -t önmagára képezi. Az

tehát, hogy az $\iota_{R_1} : R_1 \rightarrow R_1$ is és α is ilyen, csak úgy lehet, ha $\alpha = \iota_{R_1}$. Mivel $R[x]$ is rendelkezik az egyértelműségi tulajdonsággal, ezért $\beta : R[x] \rightarrow R[x]$ ugyancsak az identitás. Márpedig, ha mind $\psi_1\psi$, mind $\psi\psi_1$ identitás, akkor ψ és ψ_1 izomorfizmusok és egymás inverzei. ■

Feladatok

1. Tekintsük az $f = a_0 + a_1x + \dots + a_nx^n$ polinomhoz azt az $A_f = [a_{i,j}]$ sorvéges mátrixot, amelyre $a_{i,j} = a_{j-i}$, ha $j \geq i$ és $a_{i,j} = 0$ máskor. Bizonyítsuk be, hogy az $f \mapsto A_f$ megfeleltetés skalárszorost, összeget és szorzatot tartó és injektív.

2. A modulo m vett maradékosztályok közül melyek integritási tartományok ($m > 1$ természetes szám)?

3. Bizonyítsuk be, hogy ha az R gyűrűre $R[x]$ integritási tartomány, akkor R is az.

4. Mutassuk meg, hogy az egész számok gyűrűjében a maradékosztás nem egyértelmű.

5. Bizonyítsuk be, hogy azok az egész együtthatós polinomok, amelyeknek a konstans tagja adott $n > 1$ egész számmal osztható, ideált alkotnak. Mutassuk meg, hogy ez az ideál nem generálható egyetlen elemmel.

6. Bizonyítsuk be, hogy azok a racionális együtthatós polinomok, amelyeknek a konstans tagja páros egész szám, gyűrűt alkotnak. Létezik-e ebben a gyűrűben maradékosztás?

7. Legyen R tetszőleges egységelemes számgyűrű, és tegyük fel, hogy az $R[x]$ polinomgyűrűben létezik euklideszi algoritmus. Bizonyítsuk be, hogy ekkor R test.

8. Bizonyítsuk be, hogy az egész együtthatós polinomok gyűrűjében nem generálható n elemmel a $(2^n, 2^{n-1}x, \dots, 2x^{n-1}, x^n)$ ideál.

9. Legyen K test és $f \in K[x]$. Mi az f -re vonatkozó feltétele annak, hogy $K[x]$ minden eleme $g \circ f$, illetve annak, hogy $f \circ g$ legyen; alkalmas $g \in K[x]$ polinommal?

10. Adjunk a Horner-elrendezéshez hasonló eljárást magasabbfokú polinomokkal való osztás hányadosának és maradékának meghatározására.

11. Legyen K a modulo p vett maradékosztályok teste (p prímszám). Bizonyítsuk be, hogy az $f, g \in K[x]$ polinomoknak pontosan akkor felel meg ugyanaz a polinomfüggvény, ha $f - g$ osztható az $x^p - x$ polinommal.

12. Legyenek R, S integritási tartományok, és $\varphi : R \rightarrow S$ szürjektív homomorfizmus. Mutassuk meg, hogy φ egyértelműen kiterjeszthető egy olyan $\psi : R[x] \rightarrow S[x]$ ugyancsak szürjektív homomorfizmussá, amelyre $\psi(a) = \varphi(a)$, ha $a \in R$ és $\psi(x) = x$.

13. Legyen K a modulo p vett maradékosztályok teste (p prímszám). Adjunk meg olyan $f \in K[x]$ polinomot, amelyre $f(0) = 1$ és $f(a) = 0$, ha $a \neq 0$. Bizonyítsuk be, hogy tetszőleges $\varphi : K \rightarrow K$ függvényhez létezik olyan $f \in K[x]$ polinom, amelynek megfeleltetett polinomfüggvény megegyezik φ -vel.

14. Legyenek p_1, \dots, p_r különböző prímszámok és K_1, \dots, K_r rendre a megfelelő maradékosztályok teste modulo p_1, \dots, p_r . Bizonyítsuk be, hogy ezekben a testekben létezik „közös” interpoláció; azaz tetszőleges $\varphi_i : K_i \rightarrow K_i$ függvényekhez létezik olyan $f \in \mathbb{Z}[x]$ polinom, amelyhez az előző feladatban talált $f_i \in K_i[x]$ polinomnak megfelelően polinomfüggvény a φ_i . Mi a feltétele annak, hogy egy $g \in \mathbb{Z}[x]$ polinom rendelkezze ugyanazzal a tulajdonsággal?

15. Legyen K a modulo p vett maradékosztályok teste. Bizonyítsuk be, hogy az $x^p - x + 1$ polinom irreducibilis K felett.

16. Bizonyítsuk be, hogy $x^3 - x + 1$ irreducibilis a modulo 2 és a modulo 3 vett maradékosztályok teste felett is.

17. Legyen $p > 3$ prímszám és K a modulo p vett maradékosztályok teste. Bizonyítsuk be, hogy létezik olyan $a \in K$, amelyre $x^3 - x - a$ irreducibilis K felett.

18. Általánosítsuk az előbbi feladatot, $x^3 - x$ helyébe más f polinomot írva. Adjunk az f által meghatározott polinomfüggvényre vonatkozó szükséges és elégséges feltételt arra, hogy az f polinom rendelkezze ezzel a tulajdonsággal.

19. Legyen K a modulo p vett maradékosztályok teste, és legyen $f(x)$ irreducibilis harmadfokú polinom K felett. Bizonyítsuk be, hogy létezik olyan $a \in K$, amelyre az $f(x) - a$ polinomnak többszörös gyöke van K -ban.

20. Adjunk meg olyan racionális együtthatós harmadfokú $f(x)$ polinomot, amely irreducibilis a racionális számtest felett és egyetlen a racionális szám esetében sem lesz az $f(x) - a$ polinomnak többszörös gyöke (a komplex számtestben sem).

21. Bizonyítsuk be, hogy az $x^3 - 3x + 1$ polinom bármelyik gyöke bármely másik gyökének egész együtthatós polinomja (azaz, ha a, b a fenti polinom gyöke, akkor van olyan $f(x) \in \mathbb{Z}[x]$ polinom, amelyre $f(a) = b$).

22. Legyen K egy test és $f(x) \in K[x]$. Bizonyítsuk be, hogy ha $f(x)$ irreducibilis K felett, akkor tetszőleges K -beli b és nemnulla a és c esetén a $g(x) = c \cdot f(a \cdot x + b)$ polinom is irreducibilis.

23. Mutassuk meg, hogy az előző feladat állításának a bizonyítása egyúttal azt is adja, hogy ha $f(x)$ reducibilis a K felett, akkor $g(x)$ is az.

24. Legyen K tetszőleges (szám)test. Bizonyítsuk be, hogy a $\sum_{i=0}^n a_i x^i \in K[x]$ polinom ($a_0 \neq 0, a_n \neq 0$) pontosan akkor irreducibilis K fölött, ha a $\sum_{i=0}^n a_{n-i} x^i$ polinom az. Fogalmazzuk meg a Schönemann–Eisenstein-kritérium „duálisát”.

25. Mutassuk meg, hogy a 3.47. Tétel megfordítása nem igaz, azaz létezik olyan euklideszi gyűrű, amelyben van olyan elem, amelyik minden elemnek osztója, de a normája nem minimális.

26. Bizonyítsuk be, hogy tetszőleges n természetes számhoz létezik olyan euklideszi gyűrű, amelyben pontosan n darab irreducibilis elem van.

27. Legyen R a \mathbb{Q} -nak a \mathbb{Z} -t tartalmazó részgyűrűje. Bizonyítsuk be, hogy R is euklideszi gyűrű.

28. Bizonyítsuk be, hogy ha az S euklideszi gyűrűben az euklideszi algoritmus egyértelmű, akkor van olyan K részgyűrűje és olyan x eleme, hogy S az x elemnek K -beli együtthatós polinomjaiból áll.

29. Tetszőleges $f(x) \in \mathbb{Q}[x]$ polinomhoz tekintsük az $f^*(x) : \mathbb{Q} \rightarrow \mathbb{Q}$ polinomfüggvényt. A polinomok között vezessünk be két relációt:

1. $f \sim g$, ha van olyan $a, b \in \mathbb{Q}$ ($a \neq 0$), hogy $g = f \circ (ax + b)$.
2. $f \approx g$, ha f^* és g^* értékkészlete megegyezik.

Bizonyítsuk be, hogy mindkét reláció ekvivalenciareláció; és megegyeznek.

30. Bizonyítsuk be, hogy $\max(\text{gr}(f + g), \text{gr}(f - g)) = e \max(\text{gr}(f), \text{gr}(g))$.

31. Bizonyítsuk be, hogy $\text{gr}(f + g) = \text{gr}(f - g) = \max(\text{gr}(f), \text{gr}(g))$, ha $\text{gr}(f) \neq \text{gr}(g)$.

32. Legyen K test. Bizonyítsuk be, hogy $K[x]$ -nek azok az elemei, amelyeknek adott $c_1, \dots, c_r \in K$ elemek gyökei, egy polinomideált alkotnak.

33. Határozzuk meg, hogy a következő polinomoknak mely racionális számok lehetnek gyökei; és ezek közül melyek gyökök valóban:

1. $2x - 3, 2x + 3, x - 1, x + 1, x$;
2. $x^2 + 1, x^2 - 1, x^2 - 4, 4x^2 - 9, 6x^2 - 13x + 6, x^2 - 5x + 6$;
3. $x^4 + 2x^3 + x^2 - 2x - 2, x^4 - 5x^3 + 7x^2 - 5x + 6, 12x^4 - 56x^3 + 89x^2 - 56x + 12$;
4. $42x^2 - 105x + 42, 35x^3 - 210x^2 + 385 - 210$.

34. Az alábbi polinomokban határozzuk meg az a, b, c paraméterek értékét úgy, hogy azoknak

1. legyen racionális gyöke,
2. legyen minél több racionális gyöke,
3. legyen racionális gyöke, de ne legyen egész gyöke,
4. ne legyen racionális gyöke:

$x^2 + ax + 1,$	$x^2 + ax - 2,$	$2x^2 - ax + 2,$
$x^5 + ax + 1,$	$x^5 + ax - 2,$	$2x^5 + ax + 2,$
$x^5 + ax^2 + 1,$	$2x^5 + ax^2 + 2,$	$x^3 + ax^2 + bx + 1,$
$2x^3 + ax^2 + bx + 2,$	$x^5 + ax^4 + bx^3 + cx + 1,$	$6x^5 + ax^4 + bx^2 + cx + 6,$
$12x^5 + ax^4 + bx^3 - bx^2 + cx - 12.$		

35. Igaz-e az, hogy az egész együtthatós polinomok körében pontosan azok az f polinomok oszthatók maradékosan egy g polinommal, amelyeknek a főegyütthatója osztható g főegyütthatójával?

36. Bizonyítsuk be, hogy ha egy racionális együtthatós polinom minden egész helyen egész értéket vesz fel, akkor $n!$ -sal szorozva egész együtthatós polinomot kapunk. Állítsuk elő az ilyen tulajdonságú polinomokat rögzített racionális együtthatós polinomok polinomjaként.

37. Bizonyítsuk be, hogy az egész együtthatós polinomok gyűrűjében minden n természetes számhoz létezik főideálokból álló $I_1 \subset I_2 \subset \dots \subset I_n$ lánc, de végtelen lánc nem létezik.

NEGYEDIK FEJEZET

TÖBBHATÁROZATLANÚ POLINOMOK

1. A többhatározatlanú polinomok fogalma

Az egyhatározatlanú polinomokhoz hasonlóan lehetséges — és szükséges is — többhatározatlanú polinomokról beszélni. Ezek — szemléletesen — nem mások, mint a határozatlanokból képzett hatványszorzatok számszorosaik az összegei. Ebből az értelmezésből kiindulva azonban a műveleti azonosságok bizonyítása igen nehézkessé válna. Éppen ezért a definiálásnak egy másik módját választjuk.

Képzeljük el a többhatározatlanú polinomokat úgy, ahogy az előbb tekintettük. Legyen x a határozatlanok egyike. Vonjuk össze azokat a tagokat, amelyekben x -nek ugyanaz a hatványa szerepel. Ezt a hatványt kiemelve a „másik” tényező már nem tartalmazza ezt az x határozatlant, csak a többit. Így az x határozatlannak olyan polinomjait kapjuk, ahol az együttthatók ugyancsak polinomok. Ezek a polinomok azonban már csak eggyel kevesebb határozatlant tartalmaznak. Ez az eljárás módot ad a többhatározatlanú polinomok rekurzív definiálására.

Példaként tekintsük az $5x^2y^3 + 3x^3y + 7x^2y - 2x^3 + xy - 3x + y + 1$ polinomot. Ha itt az x határozatlan hatványai szerint csoportosítunk, akkor a $(3y - 2)x^3 + (5y^3 + y)x^2 + (y - 3)x + (y + 1)$ alakhoz jutunk. Hasonlóképpen csoportosíthatnánk y hatványai szerint is. A fenti polinomban két határozatlan is szerepel. Gondoljuk meg, hogy még „nem mondtuk meg”, mi a „kéthatározatlanú polinom”.

A fenti példában láttuk, hogy az együttthatók nem számok, hanem maguk is polinomok, egyhatározatlanú polinomok. Ezekről tudjuk, hogy ugyancsak gyűrűt alkotnak a közöttük értelmezett műveletekre. Előbb is találkoztunk már olyan polinomokkal, amelyeknek az együttthatóiról csak azt tettük fel, hogy gyűrűt alkotnak; ilyenek voltak az egész együttthatós polinomok. Tekinthesünk azonban olyan polinomokat, amelyeknek az együttthatói egy rögzített maradékosztály-gyűrű elemei. Ez az általános eljárás azért célszerű és eredményes, mert ki fogjuk mutatni, hogy a kapott „többhatározatlanú polinomok” összessége szintén gyűrűt alkot a természetes módon értelmezett műveletekre. Így rekurzív módon értelmezhetők az akármennyi határozatlanú polinomgyűrűk. Ekkor viszont az együttthatók már nem számok lesznek, hanem maguk is egy gyűrű elemei.

4.1. Definíció. Legyen R tetszőleges gyűrű. $R = R_0$ -t az R feletti nullhatározatlanú polinomgyűrűnek nevezzük. Ha R_{i-1} definiálva van ($i = 1, 2, \dots$), akkor R_i -t úgy definiáljuk mint az R_{i-1} feletti egyhatározatlanú polinomgyűrűt, és az R feletti i -határozatlanú polinomgyűrűnek nevezzük. E polinomgyűrű határozatlanját — általában — x_i -vel jelöljük. R_i helyett az $R_{i-1}[x_i]$ vagy az $R[x_1, \dots, x_i]$ jelölést is használjuk. ■

Megjegyzés. Ha kevés határozatlan szerepel, akkor különböző betűket használunk a jelölésükre, indexek nélkül. Így a kéthatározatlanú polinomgyűrűt például $R[x, y]$ jelöli. □

A definíció jogosságához szükséges az alábbi tétel:

4.1. Tétel. *Egy R gyűrű feletti n -határozatlanú polinomgyűrű ugyancsak gyűrű, tetszőleges nemnegatív n egész számra.*

Bizonyítás. Az n -re vonatkozó teljes indukcióval bizonyítjuk be az állítást: Az $n = 0$ esetben az állítás definíció szerint igaz. Tegyük fel, hogy az állítás igaz $(n - 1)$ -re, és bizonyítjuk n -re. A definíció alapján R_n az R_{n-1} feletti egyhatározatlanú polinomgyűrű, és így elég, ha az állítást egyhatározatlanú polinomgyűrűre bizonyítjuk.

Ez viszont tüstént adódik a 3.1. Tételből. ■

4.2. Definíció. Az $R_k = R_{k-1}[x_k]$ k -határozatlanú polinomgyűrű esetében a 3.3. Definíció fogalmai a következőképpen módosulnak:

i -edfokú tag együtthatója helyett x_k -ban i -edfokú tag együtthatója; konstans tag helyett x_k -ban konstans tag; polinom foka helyett polinom foka x_k -ban; a gr jel helyett gr_k jel (ha x_k helyett y a határozatlan, akkor gr_y); konstans helyett x_k -ban konstans; elsőfokú helyett x_k -ban elsőfokú; normált helyett x_k -ban normált. ■

Megjegyezzük, hogy a 3.4. Tételben kimondottaknak megfelelő eredmények minden további nélkül érvényesek a többhatározatlanú polinomokra is, ha a fogalmakat a 4.2. Definíciónak megfelelően módosítjuk.

4.2. Tétel. *Az R gyűrű feletti kéthatározatlanú $R[x, y]$ polinomgyűrű elemei egyértelműen felírhatók a $\sum a_{i,j} x^i y^j$ ($a_{i,j} \in R$) alakban, ahol az összeg véges sok tagból áll, és minden (i, j) nemnegatív egész számpár legfeljebb egyszer fordul elő. A polinomgyűrűnek eleme lesz minden fenti típusú összeg.*

Az $R[x, y]$ elemeire a

$$\left(\sum a_{i,j} x^i y^j \right) + \left(\sum b_{i,j} x^i y^j \right) = \sum (a_{i,j} + b_{i,j}) x^i y^j$$

és a

$$\left(\sum a_{i,j} x^i y^j \right) \cdot \left(\sum b_{i,j} x^i y^j \right) = \sum \left(\sum (a_{p,q} \cdot b_{i-p, j-q}) x^i y^j \right)$$

összefüggések, valamint a gyűrűket definiáló azonosságok egyértelműen meghatározzák az összeadást és a szorzást.

Bizonyítás. A felírhatóság azonnal következik abból, hogy az $(R[x])[y]$ elemei olyan polinomok, amelyekben az y^r együtthatója az x -nek R -beli együtthatós polinomja. Ha két ilyen polinom megegyezik, akkor az y^r -ek együtthatói rendre egyenlők. Ezek az együtthatók x -nek R -beli együtthatós polinomjai, ezért egyenlőségükből következik, hogy a megfelelő R -beli elemek megegyeznek.

Azt is be kell látni, hogy minden lehetőség előfordul. Tekintsük evégett az előre megadott $a_{i,j}$ elemekkel az $R[x]$ -beli $f_j = \sum_i a_{i,j} x^i$ polinomokat, valamint az ezekkel képzett $f(x, y) = \sum_j f_j y^j$ polinomot. Ez a polinom nyilván a kívánt kifejezést állítja elő.

A polinomokra már definiált összeadásból az összeadás azonosságait felhasználva adódik, hogy valóban a szereplő két polinom összegét írtuk fel. A szorzatra ez a szorzat azonosságából következik.

A szorzásnak az összeadásra vonatkozó disztributivitása biztosítja, hogy elegendő, ha a fenti típusú összegek egy-egy tagjának a szorzatát határozzuk meg, és a szorzat ebből már egyértelműen meghatározott. A szorzás kommutativitása alapján a fent megadott összefüggésből már egyértelműen meghatározható a szorzat. ■

Megjegyzés. Vegyük észre, hogy ezen a ponton visszaérkeztünk a polinomokról elképzelt képhez. Az itt tárgyalt módon viszont sikerült a polinomokra vonatkozó műveleti azonosságokat sokkal áttekinthetőbben bizonyítani. □

4.3. Tétel. *Az R gyűrű feletti $R[x, y]$ és $R[y, x]$ polinomgyűrű megegyezik.*

Bizonyítás. Azt kell megmutatni, hogy a két polinomgyűrűnek ugyanazok az elemei és a műveletek is ugyanazok.

A 4.2. Tétel alapján, a polinomokra vonatkozó azonosságok segítségével egyszerűen kiszámolható a két polinomgyűrű megegyezése. Itt, helyett a 3.49. Tételt felhasználva egy „elvi” bizonyítást adunk erre. Legyen $\varphi : R \rightarrow R[y, x]$ az a homomorfizmus, amely R elemein az identitás. A 3.49. Tétel szerint ennek létezik egy olyan $\psi : R[x] \rightarrow R[y, x]$ kiterjesztése, amely x -et x -be viszi. Ugyanezen ok miatt van olyan $\chi : R[x, y] \rightarrow R[y, x]$ kiterjesztése ψ -nek, amely y -t y -ba viszi. Hasonlóan definiálhatók a $\psi_1 : R[y] \rightarrow R[x, y]$ és $\chi_1 : R[y, x] \rightarrow R[x, y]$ homomorfizmusok.

Mivel mind a $\chi_1 \cdot \chi$, mind a $\chi \cdot \chi_1$ homomorfizmus identitás, ezért χ és χ_1 mindegyike izomorfizmus. Ez az izomorfizmus x -et x -be, y -t y -ba viszi, tehát a két polinomgyűrű valóban egyenlő. ■

Megjegyzés. Az a „kézenfekvő” kiterjesztés, hogy x -t y -ba, y -t x -be visszük, csak annyit bizonyít, hogy a két gyűrű izomorf (azaz létezik az egyiket a másikba vivő izomorfizmus).

Itt azonban többről van szó; a két gyűrűnek ugyanazok az elemei. (Hiszen ebben a gyűrűben például az xy és az yx elemek megegyeznek. A különbség csak annyi, hogy az első esetben y -nak olyan polinomjaként tekintjük, amelynek együtthatói x -nek a polinomjai; a másik esetben x -nek a polinomjaként, ahol az együtthatók y -nak a polinomjai. □

4.4. Tétel. Ha y_1, \dots, y_n az x_1, \dots, x_n határozatlanok egy tetszőleges sorrendje, akkor

$$R[y_1, \dots, y_n] = R[x_1, \dots, x_n].$$

Bizonyítás. Az n -re vonatkozó teljes indukcióval bizonyítunk. Az $n = 1$ esetben az állítás triviális, az $n = 2$ eset a 4.3. Tételben lett bizonyítva. Tegyük fel, hogy igaz az állítás, ha a határozatlanok száma kevesebb, mint $n > 2$. Az $x_n = y_n$ esetben az indukciós feltevés szerint $R[y_1, \dots, y_{n-1}] = R[x_1, \dots, x_{n-1}]$, és így a 4.2. Definíció következtében igaz az állítás. Amennyiben $x_n = y_p$ ($p < n$), akkor $\{y_1, \dots, y_{p-1}, y_{p+1}, \dots, y_n\} = \{x_1, \dots, x_{n-1}\}$, és így

$$\begin{aligned} R[y_1, \dots, y_{n-1}, y_n] &= R[y_1, \dots, y_{p-1}, y_{p+1}, \dots, y_n][y_p] = \\ &= R[x_1, \dots, x_{n-1}][x_n] = R[x_1, \dots, x_n], \end{aligned}$$

a polinomgyűrű definícióját és a teljes indukciós feltételt használva. ■

A 4.4. Tétel azt mutatja, hogy többhatározatlanú polinomok esetében egyik határozatlannak sincs a többihez képest kitüntetett szerepe. Ezzel nagy lépést tettünk meg a többhatározatlanú polinomoknak az eredetileg elképzelt alakú felírása felé, de úgy, hogy közben a szükséges összefüggéseket is bizonyítottuk. Az alábbiakban a kívánt felírást adjuk meg.

4.3. Definíció. Az $R[x_1, \dots, x_n]$ polinomgyűrűnek egy

$$p = a \cdot x_1^{i_1} \cdot \dots \cdot x_n^{i_n}$$

alakú elemét ($0 \neq a \in R$ és i_1, \dots, i_n nemnegatív egészek) egytagúnak nevezzük.

Az a neve a p egytagú együtthatója; i_1, \dots, i_n a p egytagúban fellépő kitevők, és $\text{gr}(p) = i_1 + \dots + i_n$ a p foka. Az $a = 1$ esetben a p -t normálnak nevezzük. ■

4.5. Tétel. Az $R[x_1, \dots, x_n]$ polinomgyűrű bármely nem nulla eleme egyértelműen felírható olyan egytagúak összegeként, amelyek mindegyike különböző normált egytagúak R -beli elemszerese. A felírásban szereplő egytagúakat a polinom tagjainak nevezzük.

Bizonyítás. A határozatlanok számára vonatkozó teljes indukcióval bizonyítunk. Egy határozatlan esetében az egyhatározatlanú polinomokra vonatkozó eredményből következik azonnal az állítás. (A 4.3. Tétel következtében kéthatározatlanú polinomokra is igaz az állítás.) Tegyük most fel, hogy az állítás igaz n -nél kevesebb határozatlanú polinomgyűrűkre, és legyen $R_1 = R[x_1, \dots, x_{n-1}]$. Ekkor az $R_2 = R_1[x_n] = R[x_1, \dots, x_n]$ polinomgyűrű elemei egyértelműen felírhatók az $f_0 + f_1 x_n + \dots + f_k x_n^k$ alakban, ahol az $f_j \in R_1$ polinomokról feltehető, hogy csak a 0-tól különböző tagokat írtuk ki. A teljes indukciós feltétel szerint ezekre igaz az állítás. Ebből közvetlenül adódik a felírhatóság R_2 -ben is. Az egyértelműség abból következik, hogy a felírásban x_n mindegyik hatványának az együtthatója egyértelmű, és ezeket az együtthatókat az indukciós feltétel szerint egyértelműen írhatjuk a kívánt alakba. ■

4.4. Definíció. Az $f \in R[x_1, \dots, x_n]$ polinom fokán az f tagjai fokának a maximumát értjük; az f fokát $\text{gr}(f)$ -fel jelöljük. A 0 polinomnak nem definiálunk fokot.

Egy polinomgyűrű valamely elemét homogén k -adfokúnak (vagy röviden homogénnek) nevezzük, ha minden tagja k -adfokú. ■

4.6. Tétel. *Egy polinomgyűrű minden 0-tól különböző eleme egyértelműen felbontható különböző fokú homogén polinomok összegére. Azonos fokú homogén polinomok összege vagy ugyanekkora fokú homogén polinom, vagy 0. Homogén polinomok szorzata homogén és foka a tényezők fokának összege.*

Bizonyítás. Tekintsük az $R_1 = R[x_1, \dots, x_n]$ polinomgyűrű feletti $R_1[t]$ polinomgyűrűt. A $p(x_1, \dots, x_n) = a \cdot x_1^{i_1} \cdot \dots \cdot x_n^{i_n}$ egytagú pontosan akkor i -edfokú, ha $p(tx_1, \dots, tx_n) = t^i p(x_1, \dots, x_n)$. Ennek megfelelően tetszőleges $f(x_1, \dots, x_n) \in R_1$ polinom pontosan akkor homogén i -edfokú, ha $f(tx_1, \dots, tx_n) = t^i f(x_1, \dots, x_n)$. Ez a tulajdonság egyértelműen jellemzi a homogén polinomokat, éppen ezért ezt is tekinthetjük definíciónak.

Tekintsünk egy felbontást:

$$f(x_1, \dots, x_n) = h_0(x_1, \dots, x_n) + h_1(x_1, \dots, x_n) + \dots + h_k(x_1, \dots, x_n);$$

ahol h_i i -edfokú homogén. A definíció szerint ebből

$$f(tx_1, \dots, tx_n) = h_0(x_1, \dots, x_n) + h_1(x_1, \dots, x_n)t + \dots + h_k(x_1, \dots, x_n)t^k$$

következik, ami azt jelenti, hogy a felbontásban szereplő i -edfokú homogén polinom csak az $f(tx_1, \dots, tx_n)$ polinom t -ben i -edfokú tagjának együtthatója lehet. Annak a bizonyítására, hogy ezek valóban homogén i -edfokú polinomok (minden szóba jövő i -re), tekintsük az

$$f(tx_1, \dots, tx_n) = f_0(x_1, \dots, x_n) + f_1(x_1, \dots, x_n)t + \dots + f_k(x_1, \dots, x_n)t^k$$

polinomot az $R_1[t, y]$ polinomgyűrűben. Most a következő két összefüggéshez jutunk:

Egyrészt x_j helyébe yx_j -t téve:

$$f(tyx_1, \dots, tyx_n) = f_0(yx_1, \dots, yx_n) + f_1(yx_1, \dots, yx_n)t + \dots + f_k(yx_1, \dots, yx_n)t^k.$$

Másrészt t helyébe ty -t téve:

$$f(tyx_1, \dots, tyx_n) = f_0(x_1, \dots, x_n) + f_1(x_1, \dots, x_n)ty + \dots + f_k(x_1, \dots, x_n)(ty)^k.$$

Ezeknek mint t polinomjainak az együtthatóit összevetve a kívánt összefüggéshez jutunk.

A másik két állítás triviális, figyelembe véve, hogy két azonos fokú homogén polinom összege valóban lehet 0 is; míg szorzatuk a nullosztómentesség miatt nem. ■

Megjegyzés. Homogén polinomok összege általában nem lesz homogén (ez azt jelentené, hogy minden polinom homogén). □

4.7. Tétel. *Két, 0-tól különböző polinom szorzatának foka megegyezik a tényezők fokának az összegével.*

Bizonyítás. Legyen

$$f = f_0 + \dots + f_r + \dots \quad \text{és} \quad g = g_0 + \dots + g_s + \dots$$

a két adott polinom felbontása homogén polinomok összegére ($\text{gr}(f_i) = \text{gr}(g_i) = i$, ha $f_i \neq 0$, illetve $g_i \neq 0$). Legyen ezek h szorzatának homogén polinomokra való felbontása $h = h_0 + \dots + h_i + \dots$ ($\text{gr}(h_i) = i$, ha $h_i \neq 0$). A polinomokra vonatkozó műveleti azonosságok alapján

$$h_i = f_0 g_i + f_1 g_{i-1} + \dots + f_{i-1} g_1 + f_i g_0.$$

Ha mármost $\text{gr}(f) = n$ és $\text{gr}(g) = k$, akkor f_n és g_k egyike sem 0, tehát szorzatuk sem az. Továbbá $i > n$ vagy $j > k$ esetén $f_i g_j = 0$. Ezért $h_{n+k} = f_n g_k$ különbözik 0-tól.

Másrészt, mivel $i > n$ vagy $j > k$ esetén $f_i g_j = 0$, ezért ha $i > n + k$, akkor $h_i = 0$ nyilvánvalóan teljesül, amiből következik az állítás. ■

2. Kompozíció, maradékos osztás, oszthatóság többhatározatlanú polinomokra

Az egyhatározatlanú polinomokhoz hasonlóan itt is beszélhetünk arról, hogy valamelyik határozatlan helyébe egy polinomot helyettesítünk; mint ahogyan ezt már a homogén polinomok esetében tettük. Ez szó szerint ugyanúgy történhet, mint az egyhatározatlanú polinomoknál, tekintettel arra, hogy a vizsgált polinom a kiszemelt határozatlanban egyhatározatlanú polinom. Sok esetben ez azonban nem elég, mert nemcsak az egyik, hanem a többi határozatlan helyébe is egy-egy polinomot kell írni. Éppen ezért esetünkben a „második helyen” nem egy, hanem annyi polinomot kell szerepeltetni, amennyi határozatlan a polinomgyűrűben szerepel.

A behelyettesítést nem lehet „külön-külön” elvégezni. Ha például az $x + y$ polinomban x helyébe is és y helyébe is $x + y$ -t helyettesítünk, akkor eredményül $2x + 2y$ adódik. Ha ezt két lépésben tennénk, akkor első lépésben $(x + y) + y = x + 2y$ adódik. Ha most írunk y helyébe $x + y$ -t, akkor a $3x + 2y$ polinomhoz jutunk. Az egyszerre való behelyettesítés az eljárást igen bonyolulttá teszi, ezért először egytagúakra értelmezzük a behelyettesítést, és csak azután általánosán.

4.5. Definíció. Legyen $p = x_1^{i_1} \cdot \dots \cdot x_n^{i_n}$ az $R[x_1, \dots, x_n]$ polinomgyűrű normált egytagúja és f_1, \dots, f_n tetszőleges polinomjai. Ekkor:

$$p \circ (f_1, \dots, f_n) = f_1^{i_1} \cdot \dots \cdot f_n^{i_n}.$$

Ha az f polinomnak egytagúakra való felbontása $f = \sum_i a_i \cdot p_i$ normált p_i egytagúakkal, akkor:

$$f \circ (f_1, \dots, f_n) = \sum_i a_i \cdot p_i \circ (f_1, \dots, f_n).$$

Végül, legyen $0 \circ (f_1, \dots, f_n) = 0$. ■

4.8. Tétel. Az $f \rightarrow f \circ (f_1, \dots, f_n)$ megfeleltetés az $R[x_1, \dots, x_n]$ polinomgyűrűnek önmagára való olyan művelettartó leképezése, amelynél konstansnak (0-adfokú polinomnak) önmaga felel meg.

Egy polinomgyűrűnek minden ilyen tulajdonsággal rendelkező leképezése kompozíció.

Bizonyítás. Legyen $f = \sum_i a_i p_i$ és $g = \sum_i b_i p_i$ a normált p_i egytagúakkal (feltehető, hogy ugyanazok az egytagúak szerepelnek; mert az a_i, b_i bármelyike 0 is lehet). Most $f + g = \sum_i (a_i + b_i) p_i$, így elég egyetlen egytagút nézni. Ha p tetszőleges normált egytagú, akkor a polinomgyűrűben érvényes azonosságok következtében tetszőleges $a, b \in R$ esetében

$$(ap + bp) \circ (f_1, \dots, f_n) = ((ap + bp) \circ f_1, \dots, (ap + bp) \circ f_n).$$

Ebből a 4.5. Definíció alapján azonnal következik, hogy a megadott megfeleltetés összegtartó.

Legyen most az f és g polinomoknak a 4.5. Tételben megadott alakja $f = \sum_i a_i p_i$ és $g = \sum_j b_j q_j$, ahol $a_i, b_j \in R$, továbbá p_i és q_j minden szóba jövő i és j esetén normált egytagú. Ha a két polinom bármelyike is 0, akkor nyilvánvaló, hogy ennek a tényezőnek is és a szorzatnak is a 0 felel meg, tehát ebben az esetben igaz a szorzattartás. Egyébként $f \cdot g$ az $fg = \sum_{i,j} a_i b_j p_i q_j$ alakba írható. Az összegre vonatkozóan bizonyítottak alapján elég, ha kimutatjuk, hogy a szorzattartás igaz egytagúak esetében. Ez az utóbbi állítás viszont közvetlen következménye a műveleti azonosságoknak.

Legyen most, megfordítva, egy művelettartó leképezés adott, amelynél minden konstans önmagának felel meg. Tegyük fel, hogy e megfeleltetésnél az x_i határozatlanok az f_i polinom felel meg. A szorzattartás következtében ekkor tetszőleges p egytagúnak a $p \circ (f_1, \dots, f_n)$ felel meg, mert a konstansoknak önmaguk felelnek meg.

A 4.5. Tétel és az összegtartás következtében így minden 0-tól különböző f polinomnak $f \circ (f_1, \dots, f_n)$ lesz a képe. ■

Az n -határozatlanú polinomok ismételt kompozíciójáról minden további nélkül nem lehet beszélni. Ha azonban egyszerre mindig n darab polinomot tekintünk, akkor az egyhatározatlanúakhoz hasonló eredményt lehet kimutatni:

4.9. Tétel. Legyenek $\mathbf{f} = (f_1, \dots, f_n)$, $\mathbf{g} = (g_1, \dots, g_n)$, $\mathbf{h} = (h_1, \dots, h_n)$ stb. n -határozatlanú polinom- n -esek. Értelmezzük ezekre az összeadás, szorzás és kompozíció műveleteit az alábbi módon:

$$\mathbf{f} + \mathbf{g} = (f_1 + g_1, \dots, f_n + g_n), \quad \mathbf{f} \cdot \mathbf{g} = (f_1 g_1, \dots, f_n g_n), \quad \mathbf{f} \circ \mathbf{g} = (f_1 \circ \mathbf{g}, \dots, f_n \circ \mathbf{g}).$$

Ekkor mindhárom művelet asszociatív, az összeadás és a szorzás kommutatív, az összeadásnak elvégezhető az inverz művelete, a szorzás az összeadásra nézve mindkét oldalról disztributív és a kompozíció az összeadásra és a szorzásra nézve jobb oldalról disztributív.

A tétel bizonyítását az olvasóra bízuk. ■

4.6. Definíció. Legyen $f \in R[x_1, \dots, x_n]$ és $a_1, \dots, a_n \in R$. Az $f(a_1, \dots, a_n) = f \circ (a_1, \dots, a_n)$ elemet az f polinom $\mathbf{a} = (a_1, \dots, a_n)$ helyen vett helyettesítési értékének nevezzük. Ha a helyettesítési érték 0, akkor azt mondjuk, hogy \mathbf{a} az f polinom gyöke. ■

A maradékos osztás a többhatározatlanú polinomok körében általában nem végezhető el. Ha ugyanis azt próbáljuk elérni, hogy a maradék foka valamelyik határozatlanban csökkenjen, akkor a többi határozatlanban a maradék foka nagyobb lehet, mint az osztóban. Ha azonban ettől eltekintünk, akkor normált polinomokra elvégezhető a maradékos osztás.

4.10. Tétel. Ha az $R[x_1, \dots, x_n]$ polinomgyűrű g eleme x_n -ben normált, akkor a polinomgyűrű tetszőleges f eleméhez található a polinomgyűrűnek olyan q és r eleme, amelyre

$$f = g \cdot q + r \quad \text{és} \quad \text{gr}_n(r) < \text{gr}_n(g) \quad \text{vagy} \quad r = 0.$$

Bizonyítás. Azonnal következik a 3.5. Tétel utáni megjegyzésből. ■

Következmény. Legyen $R_{n-1} = R[x_1, \dots, x_{n-1}]$ és $a \in R_{n-1}$. Az $f \in R_{n-1}[x_n]$ polinomhoz akkor és csak akkor található olyan $q \in R_{n-1}[x_n]$ polinom, amelyre $f = q \cdot (x_n - a)$, ha $f \circ (x_1, \dots, x_{n-1}, a) = 0$.

Bizonyítás. A 4.10. Tételbeli g szerepét itt az $x_n - a$ polinom játssza. Ez x_n -ben elsőfokú, ezért r az x_n -ben konstans. A nyilvánvaló $(x_n - A) \circ (x_1, \dots, x_{n-1}, a) = 0$ összefüggésből a 4.8. Tétel szerint következik, hogy $f \circ (x_1, \dots, x_{n-1}, a) = r \circ (x_1, \dots, x_{n-1}, a)$. ■

Nem okoz külön nehézséget az oszthatóság definíciója többhatározatlanú polinomok esetében:

4.7. Definíció. Az $R_n = R[x_1, \dots, x_n]$ polinomgyűrű egy f eleme osztója a polinomgyűrű g elemének, ha létezik a polinomgyűrűben olyan h elem, amelyre $g = h \cdot f$. Azt is mondjuk, hogy g osztható f -fel, illetve többszöröse f -nek. (Az oszthatóságot itt is $f \mid g$ -vel jelöljük.) ■

4.11. Tétel. A 3.7. és 3.8. Tételek analogonjai érvényesek többhatározatlanú polinomokra is.

A tétel bizonyítását az olvasóra bízuk. ■

Többhatározatlanú polinomokra is definiálhatjuk az irreducibilis polinom fogalmát, amelyet csak úgy lehet két tényező szorzatára bontani, hogy valamelyik tényező minden polinomnak osztója. (Ez a tényező nyilván csak az R gyűrű egy eleme lehet.) Itt is értelmezhető a prímtulajdonság és az asszociáltság.

4.12. Tétel. Ha az R gyűrűben érvényes az egyértelmű faktorizáció, akkor az $R[x_1, \dots, x_n]$ polinomgyűrűben is érvényes. Speciálisan, test feletti polinomgyűrűben mindig érvényes az egyértelmű faktorizáció.

Bizonyítás. Azonnal következik a 3.48. Tétel n -szeri alkalmazásával. ■

Olyan esetben, amikor az n -határozatlanú polinomgyűrű egy f eleme rögzített i mellett, az x_i -n kívül minden határozatlanban konstans, akkor $f \circ (f_1, \dots, f_n)$ helyett azt írjuk, hogy $f \circ f_i$. Ez nem okoz kétértelműséget, mert az f polinomban egyedül az x_i szerepel, tehát csak az x_i helyébe írhatjuk be az f_i -t. De mégsem teljesen azonos az egyhatározatlanú polinomok kompozíciójával, mert az f_i -ben több határozatlan is szerepelhet. A 3.49. Tétel alapján ez is behelyettesítés. Az ilyen esetek gyakoriak. A homogén polinomoknál is szerepelt hasonló. Most példaképpen azt nézzük meg, amikor kéthatározatlanú polinomgyűrűben az $x^n \circ (x + y)$ -t határozzuk meg:

4.13. Tétel (Binomiális tétel).

$$(x + y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \dots + \binom{n}{i}x^{n-i}y^i + \dots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n,$$

$$\text{ahol } \binom{n}{i} = \frac{n!}{i! \cdot (n-i)!}.$$

Bizonyítás. A 4.6. Tétel biztosítja a felírást, csupán az együtthatókat kell meghatároznunk. Ezekre könnyen kaphatunk egy rekurziós összefüggést az $(x + y)^{n+1} = (x + y)^n \cdot (x + y)$ azonosságból. $(x + y)^0 = 1$ miatt $\binom{0}{0} = 1$; továbbá $n = 0$ -ra $\binom{1}{0} = \binom{1}{1} = 1$. A felírt azonosság bal oldalán $x^{n-i}y^i$ együtthatója $\binom{n}{i}$, míg a jobb oldalon (elvégezve a szorzást) $\binom{n-1}{i} + \binom{n-1}{i+1}$, ha $0 < i < n$; és 1, ha $i = 0$ vagy $i = n$. Ebből a kívánt összefüggés teljes indukcióval bizonyítható. ■

3. Egyhatározatlanú polinomok deriváltja és többszörös gyökei

Az analízisbeli vizsgálatoknál a gyökökről igen sokat elárul a derivált. Polinomok esetében — mint látni fogjuk — ezeknek az eredményeknek jó része a derivált formális tulajdonságaiból következik. E formális tulajdonságoknak az a nagy előnye, hogy olyan polinomgyűrűk esetében is érvényesek, amikor a gyűrűben nem beszélhetünk folytonosságról. Érvényesek tehát egész együtthatós, illetve racionális együtthatós polinomok esetében is. Érvényesek még a modulo p vett maradékosztályok (p prím) esetében is. Vannak azonban olyan testek, amelyek esetében nem érvényesek. Itt most a tételeket csak számtestekre mondjuk ki. A későbbiekben más esetekben is fel fogjuk használni; a testelmélet tárgyalásánál majd megmutatjuk, mi az érvényesség pontos feltétele.

Mint tudjuk, egy $f(x)$ függvénynek az a helyen vett $f'(a)$ deriváltját az $\frac{f(a+h) - f(a)}{h}$ hányadosnak a határértéke adja meg, ha h -val 0-hoz tartunk. Könnyen ellenőrizhető — és a későbbiekben látni is fogjuk —, hogy ha $f(x)$ polinomfüggvény, akkor a fenti különbségi hányados is polinomfüggvény. Márpedig h -nak egy polinomfüggvénye csak akkor tarthat 0-hoz h -nak 0-hoz tartása esetén, ha e függvény osztható h -val. Így az

$$f(a+h) = f(a) + f'(a) \cdot h + F(a, h) \cdot h^2$$

összefüggéshez jutunk ($F(x, y)$ kéthatározatlanú polinom). Ez az összefüggés alkalmas arra, hogy segítségével értelmezhesük polinomok deriváltját.

4.8. Definíció. Legyen $f \in R[x]$. Az $f \circ (x + y) \in R[x, y]$ polinomban mint y polinomjában az elsőfokú tag f' együtthatóját az f polinom deriváltjának nevezzük. ■

Megjegyzés. Könnyen látható, hogy a fenti $f \circ (x + y)$ polinomban mint y polinomjában a konstans tag f . A felírásból azt kapjuk, hogy:

$$(*) \quad f(x+y) = f_0(x) + f_1(x)y + f_2(x)y^2 + \dots$$

Most a konstans tagot úgy kapjuk, hogy x -et változatlanul hagyva y helyébe 0-t írunk. Azaz a konstans:

$$f(x) = f(x+0) = f_0(x) + f_1(x)0 + 0 + \dots$$

E megjegyzés alapján a következő előállításoz jutottunk:

$$f(x+y) = f(x) + f'(x)y + F(x, y)y^2, \quad \text{ahol} \quad f' \in R[x], F \in R[x, y]. \quad \square$$

A deriválásra vonatkozó azonosságokat írja le az alábbi tétel:

4.14. Tétel. *A derivált egyértelműen meghatározott. A deriváltra az alábbi összefüggések teljesülnek:*

(1) *Konstans deriváltja 0;*

$$(2) (f+g)' = f' + g';$$

$$(3) (fg)' = f'g + fg';$$

$$(4) (f \circ g)' = (f' \circ g) \cdot g';$$

$$(5) \left(\sum_i a_i x^i \right)' = \sum_i i \cdot a_i x^{i-1} \quad (a_i \in R).$$

Bizonyítás. Az első állítás abból következik, hogy a $(*)$ alatti kifejezésben y együtthatója egyértelműen meghatározott. Ennek következménye, hogy egy $(*)$ alakú előállításnál y együtthatója mindig a deriváltat szolgáltatja.

Az (1)-ben szereplő állítás azonnal következik a $c = c + 0y + 0y^2$ ($c \in K$) felírásból.

Legyen most a $(*)$ -ban adott felírásen kívül $g(x+y) = g(x) + g'(x)y + G(x, y)y^2$. Ebből, a kompozícióra vonatkozó összefüggések figyelembevételével a $h = f + g$ és $k = fg$ polinomokra:

$$\begin{aligned} h(x+y) &= f(x+y) + g(x+y) = (f + f'y + Fy^2) + (g + g'y + Gy^2) = \\ &= (f+g) + (f' + g')y + (F+G)y^2, \end{aligned}$$

illetve

$$\begin{aligned} k(x+y) &= f(x+y)g(x+y) = (f + f'y + Fy^2)(g + g'y + Gy^2) = \\ &= fg + (f'g + fg')y + (Fg + f'g' + fG + Fg'y + f'Gy + FGy^2)y^2 \end{aligned}$$

adódik, ami bizonyítja a (2) és (3) állítást.

A kompozícióra vonatkozó összefüggések alapján könnyen belátható, hogy tetszőleges g és u polinomokra

$$f(g+u) = (f \circ (x+y)) \circ (g, u) = f(g) + f'(g)u + F(g, u) \cdot u^2$$

következik. Az u helyébe írjuk a g -re vonatkozó felírásból adódó $g'y + Gy^2$ polinomot. Így a jobb oldalon y együtthatója $(f'(g)) \cdot g'$ lesz, ahogyan ezt állítottuk.

Végül, (5) bizonyításához a (2) figyelembevételével elegendő az $(ax^i)' = i \cdot ax^{i-1}$ ($a \in R$) bizonyítása. Ezt teljes indukcióval bizonyítjuk be. $i = 0$ esetén az állítás (1) alapján igaz. (3)-ból

$$(ax^{i+1})' = (ax^i \cdot x)' = (ax^i)' \cdot x + (ax^i) \cdot x'$$

következik. A teljes indukciós feltevés szerint $(ax^i)' = i \cdot ax^{i-1}$; az $x + y = x + 1 \cdot y + 0 \cdot y$ alapján pedig $x' = 1$. Ebből

$$(ax^{i+1})' = i \cdot ax^{i-1} \cdot x + (ax^i) \cdot 1 = (i+1) \cdot ax^i. \quad \blacksquare$$

Megjegyzés. A 4.14. Tételből azonnal következik, hogy $a \in R$ és $f \in R[x]$ esetén $(af)' = af'$. Valóban, (3) és (1) miatt $(af)' = a'f + af' = 0f + af' = f'$. \square

Következmény. Ha $f \in R[x]$, akkor $f' \in R[x]$.

Bizonyítás. Feltétel szerint f együtthatói R -ből valók. Az f' minden egyes együtthatója úgy áll elő, hogy f valamelyik együtthatóját megszorozzuk egy egész számmal, azaz ezt az együtthatót annyiszor adjuk össze, amennyi ez az egész szám. Mivel minden számgyűrű tartalmaz minden egész számot, ha $1 \in R$, ezért ezek mind R -beliek, azaz ugyanez áll f' együtthatóira is. \blacksquare

A derivált segítségével egy — legalábbis elvi — lehetőség adható meg a polinomok többszörös gyökeinek a meghatározására. Mindenekelőtt egy fogalomra van szükségünk.

4.9. Definíció. Az f egyhatározatlanú polinomnak és deriváltjának a normált legnagyobb közös osztóját $f^* = (f, f')$ -vel fogjuk jelölni. \blacksquare

4.15. Tétel. Ha K test, akkor $f \in K[x]$ esetén $f^* \in K[x]$. Továbbá:

- (1) ha $f = gh$ és $(g, h) = 1$, akkor $f^* = g^*h^*$;
- (2) ha $f = g^r$, akkor $f^* = g^{r-1}g^*$;
- (3) ha f irreducibilis, akkor $f^* = 1$.

Bizonyítás. Mivel két $K[x]$ -beli polinom legnagyobb közös osztója is $K[x]$ -beli polinom, ezért az első állítás következik a 4.14. Tétel következményéből.

Ha g -nek és h -nak nincs közös irreducibilis faktora, akkor tetszőleges $K[x]$ -beli u polinom esetén az egyértelmű irreducibilis faktorokra való felbontás következtében $(f, u) = (gh, u) = (g, u)(h, u)$.

Legyen most $u = f' = g'h + gh'$. A polinomideálok tulajdonságai alapján $(g, f') = (g, g'h + gh') = (g, g'h)$. A g -nek és h -nak nincs közös faktora, ezért $(g, g'h) = (g, g') = g^*$, hasonlóképpen $(h, f') = h^*$. Tehát $f^* = (f, f') = (g, f')(h, f') = g^*h^*$.

Az $f = g^r$ esetben $f' = r \cdot g^{r-1} \cdot g'$ — a 4.14. Tétel (4) pontja szerint. Mivel a legnagyobb közös osztó úgyis normált, az r konstans tényezőtől eltekinthetünk. A g^{r-1} közös osztó, ezért $f^* = (f, f') = g^{r-1}(g, g') = g^{r-1}g^*$.

Legyen végül f irreducibilis. A 4.14. Tétel (5) pontja szerint $\text{gr}(f') = \text{gr}(f) - 1$, ha f nem konstans. Az f és f' polinomok legnagyobb közös osztója f -nek, és mivel f irreducibilis, ezért ez a legnagyobb közös osztó vagy f , vagy 1. f Azonban nem lehet, mert ekkor f egy nála alacsonyabb fokú polinomnak (nevezetesen f' -nek) lenne osztója. Ezzel a (3) alatti állítást is bizonyítottuk. ■

4.16. Tétel. *Legyen $f \in K[x]$ és L egy K -t tartalmazó test. Tegyük fel, hogy f az $L[x]$ -ben felbontható*

$$f = e_1 \cdot (e_2)^2 \cdot \dots \cdot (e_n)^n$$

alakba, ahol bármely két különböző e_i és e_j relatív prím, továbbá egyetlen e_i sem osztható $L[x]$ -beli polinom négyzetével. Ekkor minden egyes e_i eleme $K[x]$ -nek és az f polinom ismeretében egyértelműen meghatározható.

Bizonyítás. A bizonyítás során többször felhasználjuk azt a polinomok oszthatóságánál tárgyalt tényt, hogy ha egy $K[x]$ -beli polinom egy ugyancsak $K[x]$ -beli polinomnak $L[x]$ -ben osztója, akkor a hánnyados $K[x]$ -ben van.

Az egyértelmű irreducibilis polinomokra való felbonthatóság alapján az f polinomot is felbonthatjuk $L[x]$ -beli irreducibilis polinomok szorzatára. Legyen e_i ezek közül azoknak a különböző irreducibilis tényezőknek a szorzata, amelyeknek az i -edik hatványa osztója f -nek, de az $(i + 1)$ -edik hatványa már nem.

Minden irreducibilis faktorhoz pontosan egy ilyen kitevő tartozik, ezért bármely két ilyen tényező valóban relatív prím egymáshoz. Ugyancsak a kitevő egyértelműségéből következik, hogy egyetlen e_i sem osztható $L[x]$ -beli polinom négyzetével.

Az e_i faktorok relatív prímek egymáshoz, ezért a 4.15. Tétel (1) pontja szerint f^* megegyezik az összes $(e_i^i)^*$ szorzatával. A 4.15. Tétel (2) pontja szerint ez viszont nem más, mint $e_i^{i-1} \cdot e_i^*$. Mivel e_i nem osztható $L[x]$ -beli polinom négyzetével, ezért relatív prím irreducibilis faktorok szorzatára bontható: $e_i = p_1 \cdot \dots \cdot p_r$. A 4.15. Tétel (1) pontja szerint $e^* = p_1^* \cdot \dots \cdot p_r^*$; és e tétel (3) pontja alapján e tényezők mindegyike 1. Eszerint

$$f^* = e_2 \cdot (e_3)^2 \cdot \dots \cdot (e_n)^{n-1}.$$

Legyen most $f_1 = f$, és $f_j = f_{j-1}^*$, ha $j > 1$. Ekkor:

$$f_i = e_i \cdot \dots \cdot (e_n)^{n+1-i} \quad (i \leq n).$$

A 4.15. Tétel ismételt alkalmazásával azt kapjuk, hogy az f_i polinomok mind $K[x]$ -beliek.

Mivel f_{i+1} osztója f_i -nek, ezért ezek $g_i = \frac{f_i}{f_{i+1}}$ hányadosa is $K[x]$ -beli. Ez a hányados

$$g_i = e_i \cdot \dots \cdot e_n.$$

Láthatjuk, hogy g_{i+1} is osztója g_i -nek; tehát hányadosuk, $e_i = \frac{g_i}{g_{i+1}}$ valóban eleme $K[x]$ -nek.

Az egyértelmű felbontás szerint az adott felírás $L[x]$ -ben egyértelmű. Most beláttuk, hogy L -től sem függ, tehát az egyértelműséget is bizonyítottuk. ■

Vegyük észre, hogy tulajdonképpen egy algoritmust is adtunk az e_i polinomok meghatározására. Ennél az algoritmusnál eleve nem lehet tudni, hogy mikor kell megállni. Tekintettel arra, hogy f^* nem lehet 0, és foka kisebb, mint f foka, ezért valamilyen k indexre $f_k = 1$ adódik. Ezután már minden $f_{k+i} = 1$ teljesül.

Megjegyezzük még azt is, hogy ha L -nek a komplex számtestet vesszük (és elfogadjuk, hogy a komplex számtestben minden polinom elsőfokú faktorokra bomlik), akkor látható, hogy az e_i polinom az f polinomnak pontosan azokat a gyökeit adja, amelyek az eredeti polinomnak i -szeres gyökei. Az e_i polinomnak már csak egyszeres gyökei vannak, ezért a polinom gyökeinek a meghatározása mindig visszavezethető olyan polinomok gyökeinek a meghatározására, amelyeknek csupa egyszeres gyökei vannak. (Ha az algebra alaptételét nem vesszük figyelembe, akkor is adódik az, hogy az e_i polinomoknak legfeljebb egyszeres gyökei vannak.)

Az egyhatározatlanú polinomok vizsgálatában fontos szerepük van a magasabb rendű deriváltaknak is.

4.10. Definíció. Legyen $f^{(0)} = f$ és $f^{(k+1)} = (f^{(k)})'$ ($k = 0, 1, \dots$). Az $f^{(k)}$ polinomot az f polinom k -adik deriváltjának nevezzük. ■

4.17. Tétel. $a \in R$ akkor és csak akkor k -szoros gyöke az $f \in R[x]$ polinomnak, ha

$$f^{(0)}(a) = f^{(1)}(a) = \dots = f^{(k-1)}(a) = 0, \quad \text{de} \quad f^{(k)}(a) \neq 0.$$

Bizonyítás. Legyen a valamilyen k -ra k -szoros gyöke az f polinomnak. Ekkor f felírható $f = (x - a)^k g$ alakban, ahol $g(a) \neq 0$. Így $f' = k(x - a)^{k-1}g + (x - a)^k g' = (x - a)^{k-1} g_1$, ahol $g_1(a) \neq 0$. Ebből az $f^{(i)}$ polinomokat rendre meghatározva azt kapjuk, hogy a feltétel szükséges.

A feltétel elégségeségét a következőképpen láthatjuk be. Ha a feltétel teljesül valamilyen k -ra, akkor $f(a) = 0$, és így $f = (x - a)^r g$ valamilyen r -re, ahol $g(a) \neq 0$. Ebből az előző megfontolás alapján azt kapjuk, hogy a az f -nek pontosan r -szeres gyöke, és így $r = k$. ■

4.18. Tétel. Legyen $f(x) = a_0 + a_1x + \dots + a_nx^n \in R[x]$. Ekkor $a_i = \frac{f^{(i)}(0)}{i!}$. Ha

$$c \in R, \text{ akkor } f(x) = b_0 + b_1(x - c) + \dots + b_n(x - c)^n, \text{ ahol } b_i = \frac{f^{(i)}(c)}{i!}.$$

Bizonyítás. Legyen először $a = 0$. Ekkor a kiinduló felírásban mindkét oldalon az i -edik deriváltat véve $f^{(i)}(x) = i! \cdot a_i + ((i+1)! \cdot a_{(i+1)} + \dots) \cdot x$ adódik, amiből $x = 0$ helyettesítéssel kapjuk, hogy $f^{(i)}(0) = i! \cdot a_i$.

Tekintsük most a $g(x) = f(x+a) = b_0 + b_1x + \dots + b_nx^n$ polinomot (a kompozíció fokára vonatkozó összefüggés szerint ez is n -edfokú). A most kapott eredmény szerint $b_i = \frac{g^{(i)}(0)}{i!}$. Az $x+a$ deriváltja 1, ezért g magasabb rendű deriváltjait kiszámítva, azt kapjuk, hogy $g^{(i)}(x) = f^{(i)}(x+a)$. Ebből $g^{(i)}(0) = f^{(i)}(0+a) = f^{(i)}(a)$. ■

4. Szimmetrikus és alternáló polinomok

Az egyhatározatlanú polinomok vizsgálata után most visszatérünk a többhatározatlanú polinomokhoz. A polinomok felírásánál sok esetben célszerű, ha jelöljük a szereplő határozatlanokat. Így $f \in R[x_1, \dots, x_n]$ esetén az $f = f(x_1, \dots, x_n)$ jelölést is használni fogjuk.

Megjegyzések

1. Amennyiben $R = S[y_1, \dots, y_k]$ polinomgyűrű, akkor az y_1, \dots, y_k határozatlanokat paramétereknek nevezik. Tulajdonképpen az, hogy melyik határozatlan paraméter, attól függ, hogy melyiket tekintjük annak.

Ugyancsak fontos lesz a továbbiakban a helyettesítéskor kapott elem jelölése.

Az $f(x_1, \dots, x_n) \in R[x_1, \dots, x_n]$ polinomnak az x_n határozatlan $a \in R$ helyen vett $f \circ (x_1, \dots, x_{n-1}, a)$ helyettesítési értékét $f(x_1, \dots, x_{n-1}, a)$ jelöli.

2. Amennyiben a határozatlanok közül bizonyosoknak a helyébe számokat akarunk helyettesíteni, és ettől függően akarjuk a polinomok viselkedését vizsgálni, akkor ezeket szokták paramétereknek tekinteni. □

A polinomok gyökeinek a meghatározásánál igen fontos szerepet játszik annak a fel-tárása, hogy ezek a gyökök milyen „polinomiális” kapcsolatban állnak egymással. Pontosabban szólva, azoknak a többhatározatlanú polinomoknak a megkereséséről van szó, amelyekbe a határozatlanok helyébe az eredeti polinom gyökeit beírva 0-t kapunk. Ezeknek a megkeresése egyes konkrét polinomok esetében az úgynevezett Galois-elmélethez vezet. Itt csupán olyan többhatározatlanú polinomok felírását tűzzük ki célul, amelyek minden egyes polinom esetében fontos szerepet töltenek be. Ezek az úgynevezett szimmetrikus és alternáló polinomok. (Meg lehet mutatni, hogy más polinomnak valóban nincs ilyen univerzális szerepe.)

Mindenekelőtt a 3.49. és 3.50. Tételek általánosítását tárgyaljuk.

3.49/A. Tétel. *Legyenek R és S egységelemes integritási tartományok, és $\varphi : R \rightarrow S$ tetszőleges homomorfizmus. Ekkor létezik olyan egyértelmű $\psi : R[x_1, \dots, x_n] \rightarrow S$ homomorfizmus, amely az R elemein megegyezik φ -vel és minden egyes x_i -t az S egy előre megadott s_i elemére képezi le.*

Bizonyítás. Alkalmazzuk a 3.49. Tételt rendre az $R[x_1]$, $(R[x_1])[x_2]$ stb. polinomgyűrűkre. ■

3.50/A. Tétel. Legyen R_1 az R -et tartalmazó integritási tartomány és $x_1, \dots, x_n \in R_1$. Ha tetszőleges S egységelemes integritási tartományhoz, $\varphi : R \rightarrow S$ homomorfizmushoz és S -beli s_1, \dots, s_n elemekhez létezik olyan egyértelmű $\psi : R_1 \rightarrow S$ homomorfizmus, amely az R elemein megegyezik φ -vel és $\psi(x_i) = s_i$, akkor $R_1 = R[x_1, \dots, x_n]$.

Bizonyítás. A 3.50. Tétel bizonyításához hasonlóan történik. ■

Tekintsünk egy $\pi : \{x_1, \dots, x_n\} \rightarrow \{x_1, \dots, x_n\}$ tetszőleges bijekciót. A polinomgyűrűk definíciója alapján ez kiterjeszthető egy homomorfizmussá; amely a fenti két tétel miatt izomorfizmus. A továbbiakban ezeket az izomorfizmusokat vizsgáljuk.

4.11. Definíció. Az $f(x_1, \dots, x_n) \in R_n = R[x_1, \dots, x_n]$ polinomot szimmetrikusnak nevezzük, ha bármely π -re

$$f(\pi(x_1), \dots, \pi(x_n)) = f(x_1, \dots, x_n). \quad \blacksquare$$

Mivel a π által megadott izomorfizmus művelettartó, ezért:

4.19. Tétel. R_n szimmetrikus polinomjai egy T_n részgyűrűt alkotnak. ■

Megjegyzés. Míg a homogén polinomoknál nem lényeges a határozatlanok száma, a szimmetrikusoknál igen. Ha ugyanis már egyetlen új határozatlant is hozzáveszünk, akkor egy nemkonstans szimmetrikus polinom elveszti a szimmetriáját. Így az $x^2 + y^2$ polinom akármennyi újabb határozatlan hozzávétele esetén homogén másodfokú marad, azaz e polinom homogén másodfokú polinomja az x, y, z, u, v határozatlanoknak is. Amennyiben viszont egy újabb z határozatlant hozzáveszünk, már elveszti a szimmetrikusságot. □

A szimmetrikus polinomok között két nagyon fontos fajta van.

4.12. Definíció. Az $s_k = s_k(x_1, \dots, x_n) = x_1^k + \dots + x_n^k$ polinomot e határozatlanok k -adik hatványösszegének nevezzük. ■

4.13. Definíció. Az adott határozatlanok összes i -elemű halmazából képezett szorzatok összegét e határozatlanok i -edik elemi szimmetrikus polinomjának nevezzük. Ez

$$\sigma_i = \sigma_i^{(n)} = \sigma_i(x_1, \dots, x_n) = \sum x_{j_1} \cdot \dots \cdot x_{j_i},$$

ahol j_1, \dots, j_i az $\{1, \dots, n\}$ halmaz összes i -elemű részhalmazán fut végig. ■

A hatványösszeznél az index akármilyen nagy lehet. Világos, hogy $s_0 = n$. Az elemi szimmetrikus polinomoknál az index nem lépheti túl a határozatlanok számát; illetve ilyen esetben ezt a polinomot 0-nak célszerű definiálni (az összegnek egyetlen tagja sincs). A σ_0 úgy tekinthető, mint egyetlen üres szorzat összege; éppen ezért a célszerű definíció $\sigma_0 = 1$.

Az elemi szimmetrikus polinomokra — fontosságuk miatt — két másik, definícióként tekinthető tulajdonságot adunk meg.

4.20. Tétel. Az

$$a_0 t^n + a_1 t^{n-1} + \dots + a_{n-1} t + a_n = (t + x_1) \cdot \dots \cdot (t + x_n)$$

polinomban $a_i = \sigma_i$ ($i = 0, 1, \dots, n$).

Az elemi szimmetrikus polinomok (és csak ezek) eleget tesznek az alábbi rekurzióknak:

$$\sigma_i(x_1, \dots, x_{n-1}, x_n) = \sigma_i(x_1, \dots, x_{n-1}) + \sigma_{i-1}(x_1, \dots, x_{n-1}) \cdot x_n,$$

ha $0 < i \leq n$; emellett definíció szerint $\sigma_0 = 1$ és $\sigma_i^{(n)} = 0$, ha $i > n$.

Bizonyítás. Először megmutatjuk, hogy a polinomban fellépő együttthatók valóban az elemi szimmetrikus polinomokkal egyeznek meg. Ehhez azt kell belátni, hogy t^{n-i} együttthatójában minden i -tényezős szorzat pontosan egyszer lép fel. A szimmetria miatt elég ezt az $x_1 \cdot \dots \cdot x_i$ szorzatról megmutatni. A disztributivitás alapján a fenti szorzat olyan összegként írható fel, amelynek a tagjai olyan szorzatok, amelyekben minden egyes kéttagú tényezőből vagy a t , vagy a megfelelő x_i szerepel. A kiszemelt szorzatot pontosan akkor kapjuk meg, ha az első i tényezőből vesszük az x_1, \dots, x_i tagokat, és a többiből a t -t.

Ezután megmutatjuk, hogy a szereplő együttthatók kielégítik a megadott rekurziót. Ez könnyen következik a

$$[(t + x_1) \cdot \dots \cdot (t + x_{n-1})] \cdot (t + x_n) = (t + x_1) \cdot \dots \cdot (t + x_n)$$

felírásból.

Végezetül megmutatjuk, hogy a megadott rekurzióból következik, hogy a fellépő polinomok valóban az elemi szimmetrikus polinomok.

$n = 0$ esetén az állítás triviális. Tegyük fel, hogy az állítás igaz $(n - 1)$ -re. A rekurzió szerint

$$\sigma_i(x_1, \dots, x_{n-1}, x_n) = \sigma_i(x_1, \dots, x_{n-1}) + \sigma_{i-1}(x_1, \dots, x_{n-1}) \cdot x_n.$$

Itt a feltevéseünk alapján $\sigma_i(x_1, \dots, x_{n-1})$ az x_1, \dots, x_{n-1} határozatlanokból képezett összes i -tényezős szorzatból áll. Hiányoznak még azok az i -tényezős szorzatok, amelyekben x_n szerepel. Ezeket viszont az $\sigma_{i-1}(x_1, \dots, x_{n-1}) \cdot x_n$ tagból kapjuk meg. ■

Célunk az alábbi tétel bizonyítása:

4.21. Tétel (A szimmetrikus polinomok alaptétele). Minden szimmetrikus polinom egyértelműen előáll mint az elemi szimmetrikus polinomok polinomja.

Pontosabban megfogalmazva: Defináljuk tetszőleges pozitív n egészre az alábbi homomorfizmusokat:

$$\Phi_n : R[y_1, \dots, y_n] \rightarrow R[x_1, \dots, x_n], \quad \text{ahol} \quad \Phi_n : y_i \mapsto \sigma_i.$$

Ekkor Φ_n injektív és $\text{Im}(\Phi_n) = T_n$. (Definíció szerint $\varphi : A \rightarrow B$ esetén $\text{Im}(\varphi) = \{\varphi(a) | a \in A\}$.)

Mint általában, itt is egy eljárást adunk arra, miképpen állíthatunk elő egy szimmetrikus polinomot az elemi szimmetrikusok polinomjaként. Az itt közölt bizonyítás teljesen precíz, emiatt kicsit bonyolultnak tűnik. Éppen ezért először egy konkrét példán mutatjuk be az eljárást.

Állítsuk elő a harmadik hatványok összegét az elemi szimmetrikus polinomokkal!

Ha csak egy határozatlan van: x_1 , akkor x_1^3 -t kell előállítani. Ehhez egyetlen határozatlan szükséges, y_1 ; és a megfeleltetés $y_1 \rightarrow \sigma_1(x_1)$, hiszen $x_1^3 = x_1^3$. Két határozatlan esetében $x_1^3 + x_2^3$ az előállítandó polinom, és ehhez y_1^3 már adott. Az ennek megfeleltetett $\sigma_1(x_1, x_2)^3 = (x_1 + x_2)^3 = x_1^3 + x_2^3 + 3(x_1 x_2)(x_1 + x_2)$ nem egyezik meg a kívánt polinommal. Az eltérés $3(x_1 x_2)(x_1 + x_2)$. Ez ismét szimmetrikus a két határozatlanban, de közben „valami” csökkent. Ez a szimmetrikus polinom $\sigma_2(x_1, x_2)\sigma_1(x_1, x_2)$, azaz az $y_2 y_1$ -nek megfelelő polinom. Ha most végezzük el a számolást, azt kapjuk, hogy $x_1^3 + x_2^3 = \sigma_1^3 - 3\sigma_1\sigma_2$. Ez két határozatlan esetében igaz, de három esetében már nem. Mindenesetre a kapott $y_1^3 - 3y_1 y_2$ polinomhoz már csak olyan tagok járulhatnak hozzá, amelyekben 2-nél nagyobb indexű y szerepel (ez itt nem világos, csak a bizonyítás folyamán fog kiderülni — éppen ezért van szükség a bizonyításra). Azt kaptuk tehát, hogy az $x_1^3 + x_2^3 + x_3^3$ polinomot „majdnem” előállítja az $y_1^3 - 3y_1 y_2$ polinom. y_i helyébe a $\sigma_i(x_1, x_2, x_3)$ polinomot behelyettesítve az $x_1^3 + x_2^3 + x_3^3$ polinomtól való eltérésre $3x_1 x_2 x_3 = 3\sigma_3(x_1, x_2, x_3)$ adódik. A megfelelő polinom tehát $y_1^3 - 3y_1 y_2 + 3y_3$. Itt is igaz az, hogy a további tagokban fel kell lépnie olyan y -nak, amelynek az indexe legalább 4. A megfelelő elemi szimmetrikus polinomok minden tagjában tehát legalább négy határozatlan szorzata szerepel, ezért nem lehet harmadfokú. Ez az előállítás tehát már akármennyi határozatlan esetében jó.

A teljes indukciós bizonyításhoz két további homomorfizmusra és ezek inverzeire lesz szükségünk:

$$\begin{aligned}
 \xi_n : R[x_1, \dots, x_n] &\rightarrow R[x_1, \dots, x_{n-1}]; & \xi_n(x_i) &= \begin{cases} 0, & \text{ha } i = n \\ x_i & \text{máskor} \end{cases} \quad \text{és} \\
 \xi_n^* : R[x_1, \dots, x_{n-1}] &\rightarrow R[x_1, \dots, x_n]; & \text{ahol } \xi_n^*(x_i) &= x_i \quad (0 < i < n). \\
 \eta_n : R[y_1, \dots, y_n] &\rightarrow R[y_1, \dots, y_{n-1}]; & \eta_n(y_i) &= \begin{cases} 0, & \text{ha } i = n \\ y_i & \text{máskor} \end{cases} \quad \text{és} \\
 \eta_n^* : R[y_1, \dots, y_{n-1}] &\rightarrow R[y_1, \dots, y_n]; & \text{ahol } \eta_n^*(y_i) &= y_i \quad (0 < i < n).
 \end{aligned}$$

A bizonyításhoz tekintsük a következő diagramot:

$$\begin{array}{ccc}
 R[y_1, \dots, y_n] & \xrightarrow{\Phi_n} & R[x_1, \dots, x_n] \\
 \eta_n \downarrow & & \downarrow \xi_n \\
 R[y_1, \dots, y_{n-1}] & \xrightarrow{\Phi_{n-1}} & R[x_1, \dots, x_{n-1}]
 \end{array}$$

Mivel a továbbiakban csak egy rögzített n esetre tekintjük a fenti diagramot, ezért az η és a ξ indexét elhagyjuk.

Lemma. *A (*)-ban valóban homomorfizmust definiáltunk, amelyekre mind $\xi\xi^*$, mind $\eta\eta^*$ identitás; a (**) alatti diagram kommutatív, ami azt jelenti, hogy $\xi\Phi_n = \Phi_{n-1}\eta$.*

Bizonyítás. A polinomgyűrűket definiáló tulajdonság szerint egy leképezés, amelynek az alapgyűrűre való megszorítása homomorfizmus, egyértelműen megadott a határozatlanok képével. Így (*)-ban tényleg homomorfizmusokat generáltunk; és a felírt szorzatok valóban az identitást adják. Hasonlóképpen (**) -ban elég annak a bizonyítása, hogy $\xi(\Phi_n(y_i)) = \Phi_{n-1}(\eta(y_i))$ teljesül minden i indexre. A 4.20. Tételt felhasználva:

$$\begin{aligned}\xi(\Phi_n(y_i)) &= \xi(\sigma_i(x_1, \dots, x_{n-1}, x_n)) = \sigma_i(x_1, \dots, x_{n-1}, 0) = \\ &= \sigma_i(x_1, \dots, x_{n-1}) = \Phi_{n-1}(y_i) = \Phi_{n-1}(\eta(y_i)),\end{aligned}$$

ha $i < n$; míg az $i = n$ esetben $\xi(\Phi_n(y_n)) = \Phi_{n-1}(\eta(y_n)) = 0$. ■

Az egyértelműség bizonyítása. $R[y_1, \dots, y_n]$ -ben n szerinti, ezen belül pedig fokszámra vonatkozó teljes indukcióval bizonyítunk. $n = 1$ esetén Φ_1 izomorfizmus, tehát injektív. Tegyük most fel, hogy az állítás igaz minden olyan esetben, amikor a határozatlanok száma kevesebb, mint n , és legyen $F \in R[y_1, \dots, y_n]$ -re $\Phi_n(F) = 0$. Ha F konstans y_n -ben, akkor $\eta^*(G)$ alakú, amiből

$$\Phi_{n-1}(G) = \Phi_{n-1}(\eta(\eta^*(G))) = \Phi_{n-1}(\eta(F)) = \xi(\Phi_n(F)) = \xi(0) = 0$$

alapján $G = 0$, és így $F = 0$ következik. Egyébként legyen $G = \eta(F)$; ekkor a diagram kommutativitása miatt

$$\Phi_{n-1}(G) = \Phi_{n-1}(\eta(F)) = \xi(\Phi_n(F)) = \xi(0) = 0,$$

amiből az indukciós feltevés szerint $G = 0$ következik. η definíciója szerint tehát $F = Hy_n$ alakú (hiszen G „lényegében” az F -nek y_n -ben konstans tagja) és a diagram kommutativitása alapján $\Phi_n(H)\Phi_n(y_n) = 0$. A nullosztómentesség és $\sigma_n^{(n)} \neq 0$ következtében csak $\Phi_n(H) = 0$ lehet. Ebből a fokra vonatkozó indukciós feltevés alapján $H = 0$ következik. ■

Az előállíthatóság bizonyításához szükségünk lesz az alábbi tételre:

4.22. Tétel. *Egy szimmetrikus polinomot homogén polinomok összegére bontva, ezek mindegyike szimmetrikus; minden szimmetrikus polinom felírható homogén szimmetrikus polinomok összegeként.*

Bizonyítás. Tekintsük a homogén polinomokra való felbontást definiáló

$$f(tx_1, \dots, tx_n) = f_0(x_1, \dots, x_n) + f_1(x_1, \dots, x_n)t + \dots + f_k(x_1, \dots, x_n)t^k$$

felírást. Az x_i határozatlanokra alkalmazva egy tetszőleges π bijekciót, az

$$\begin{aligned}f(t\pi(x_1), \dots, t\pi(x_n)) &= \\ &= f_0(\pi(x_1), \dots, \pi(x_n)) + f_1(\pi(x_1), \dots, \pi(x_n))t + \dots + f_k(\pi(x_1), \dots, \pi(x_n))t^k\end{aligned}$$

összefüggéshez jutunk, amelyből az együtthatók egyértelműsége következtében azonnal adódik a kívánt eredmény. ■

A felírhatóság bizonyításánál is a határozatlanok számára, majd ezen belül a „fokra” vonatkozó teljes indukciót fogunk használni. Minden további nélkül viszont nem tudjuk belátni, hogy az eljárásnál valamiféle fokszám csökken. Éppen ezért szükségünk lesz arra, hogy milyen alakú polinomok képe lesz homogén szimmetrikus.

4.23. Tétel. Az $a \cdot y_1^{k_1} \cdot \dots \cdot y_n^{k_n} \in R[y_1, \dots, y_n]$ monom (azaz egytagú) súlyán a $k_1 + 2k_2 + \dots + nk_n$ számot értjük.

Egy k súlyú monomot Φ_n egy k -adfokú homogén szimmetrikus polinomba képez. Egy $R[y_1, \dots, y_n]$ -beli F polinom Φ_n -nél vett képe pontosan akkor lesz homogén i -edfokú, ha minden benne szereplő monom súlya i .

Ha az F polinom minden tagja i súlyú, akkor i súlyú polinomnak nevezzük.

Bizonyítás. Tekintettel arra, hogy $\Phi_n(y_i) = \sigma_i$ i -edfokú homogén, ezért — lévén „a szorzat foka megegyezik a tényezők fokának az összegével” — azonnal adódik az első állítás.

Bontsuk most fel az F polinomot $F_0 + F_1 + \dots + F_k$ alakba, ahol F_i súlya i . Ezt úgy tehetjük meg, hogy egy-egy F_i -be összegyűjtjük a polinom i súlyú tagjait. Ekkor minden F_i képe homogén i -edfokú; és állításunk azonnal következik a 4.6. Tételből. ■

4.24. Tétel. Minden R_n -beli homogén i -edfokú szimmetrikus polinom előáll egy i súlyú F polinom Φ_n -nél vett képeként.

Bizonyítás. A bizonyítást n -re és ezen belül i -re vonatkozó teljes indukcióval végezzük. Ha $n = 1$, akkor minden polinom szimmetrikus és Φ_1 szűrjektív, hiszen izomorfizmus.

Tegyük most fel, hogy az állítás igaz minden olyan esetben, amikor a határozatlanok száma kevesebb, mint n . Ha a szereplő polinom homogén 0-adfokú, akkor egy $c \in R$ konstans és $\Phi_n(c) = c$ miatt igaz az állítás. A továbbiakban tehát egy $f(x_1, \dots, x_n)$ homogén k -adfokú szimmetrikus polinomot tekintünk, és feltesszük, hogy az előállítás érvényes minden olyan homogén szimmetrikus polinomra, amelynek a foka k -nál kisebb.

Tekintsük a $g(x_1, \dots, x_{n-1}) = \xi(f(x_1, \dots, x_n)) \in R_{n-1}$ polinomot. Most két esetet különböztetünk meg. Ha $g = 0$, akkor f osztható x_n -nel; s a szimmetrikusság miatt minden egyes x_i -vel. Ezek páronként relatív prímek lévén, szorzatuk is osztja f -et: $f = h \cdot \sigma_n$. Mivel h foka kisebb, mint f foka, ezért előáll $\Phi_n(H)$ alakban, ahol H súlya megegyezik h fokával. Mivel $H y_n$ súlya H súlyánál n -nel több és f foka h fokánál ugyancsak n -nel több, ezért a nyilvánvalóan adódó $\Phi_n(H y_n) = f$ előállításban teljesülnek a kirótt feltételek. A továbbiakban tehát feltehetjük, hogy $g \neq 0$. A 4.20. Tételből ekkor azonnal adódik, hogy g foka megegyezik f fokával.

A határozatlanok számára vonatkozó teljes indukció miatt van olyan $G(y_1, \dots, y_{n-1})$ polinom $R[y_1, \dots, y_{n-1}]$ -ben, amelyre $g = \Phi_{n-1}(G)$ és G súlya ugyancsak k . Legyen $G_1 = \eta^*(G)$, ekkor

$$\xi(\Phi_n(G_1)) = \Phi_{n-1}(\eta(G_1)) = \Phi_{n-1}(\eta(\eta^*(G))) = \Phi_{n-1}(G) = g$$

alapján azt kapjuk, hogy az $f_1 = \Phi_n(G_1)$, illetve az $f_2 = f - f_1$ polinomra $\xi(f_2) = \xi(f - f_1) = 0$ teljesül. Emellett G_1 és G súlya nyilván egyenlő. Ebből azonnal adódik az is, hogy f_1 és f megegyező fokúak; így f_2 és f foka is megegyezik¹. Mivel $\xi(f_2) = 0$,

¹ „Elvileg” $f_2 = 0$ is lehetséges volna, de ezt $g \neq 0$ kizárta.

ezért van olyan $H \in R[y_1, \dots, y_n]$ k súlyú polinom, amelyre $\Phi_n(H) = f_2$. Ekkor viszont $F = G_1 + H$ is k súlyú és $\Phi_n(F) = f$. ■

Az előállíthatóság bizonyítása. Azonnal következik a 4.22. és 4.24. Tételekből. ■

Különösen fontos a hatványösszegek meghatározása. Erre egy rekurzív formulát bizonyítunk:

4.25. Tétel (Newton-képletek). Az R_n polinomgyűrűben tetszőleges pozitív k egész számra érvényesek az alábbiak:

$$s_k - \sigma_1 s_{k-1} + \dots + (-1)^{k-1} \sigma_{k-1} s_1 + (-1)^k k \sigma_k = 0.$$

Megjegyezzük, hogy ha $k > n$, akkor $\sigma_{n+1} = \dots = \sigma_k = 0$.

Bizonyítás. Tekintsük az $R[x_1, \dots, x_n, t]$ polinomgyűrűben a

$$t^n - \sigma_1 t^{n-1} + \dots + (-1)^{n-1} \sigma_{n-1} t + (-1)^n \sigma_n = (t - x_1) \cdot \dots \cdot (t - x_n)$$

polinomot, ahol $\sigma_i = \sigma_i(x_1, \dots, x_n)$ e határozatlanok i -edik elemi szimmetrikus polinomja. Tekintettel arra, hogy a jobb oldal minden x_i helyettesítésre 0, ezért ugyanez áll a bal oldalra is:

$$x_i^n - \sigma_1 x_i^{n-1} + \dots + (-1)^{n-1} \sigma_{n-1} x_i + (-1)^n \sigma_n = 0.$$

Ezeket az egyenlőségeket minden i -re összeadva az

$$s_n - \sigma_1 s_{n-1} + \dots + (-1)^{n-1} \sigma_{n-1} s_1 + (-1)^n n \sigma_n = 0$$

összefüggéshez jutunk, amely az eredeti összefüggést adja a $k = n$ esetre. ■

Tekintsük az eredeti egyenlőség bal oldalán álló $f_{k,n} = f_{k,n}(x_1, \dots, x_n)$ polinomot. Ez a polinom szimmetrikus, mert szimmetrikusok szorzatainak összege. A homogén polinomok tulajdonsága alapján $f_{n,k}$ homogén is, és vagy azonosan 0, vagy k -adfokú.

Az előzőekben belátottak szerint $f_{n,n} = 0$. Ekkor tételünk triviálisan következik az alábbi segédtételel: ■

Segédétel. Legyen $g_{k,n} = g_{k,n}(x_1, \dots, x_n)$ e határozatlanok k -adfokú homogén szimmetrikus polinomja. Ha $g_{n,n} = 0$, akkor minden k -ra igaz a $g_{k,n} = 0$ is.

Bizonyítás. A segédtelet rögzített k mellett n -re vonatkozó teljes indukcióval bizonyítjuk.

$n = k$ esetén az állítás feltétel szerint igaz. Ebből $n < k$ esetére is következik az állítás az $x_{k-1} = \dots = x_n = 0$ helyettesítéssel.

Legyen most $n \geq k$ és tegyük fel, hogy $g_{k,n} = 0$. Ha $g_{k,n+1}(x_1, \dots, x_n, x_{n+1})$ -ben x_{n+1} helyébe 0-t írunk, akkor pontosan az x_{n+1} „mentes” tagokat, azaz $g_{k,n}(x_1, \dots, x_n)$ -t kapjuk. Ez az indukciós feltevés szerint 0. Mivel $g_{k,n+1}$ szimmetrikus polinom, ezért létezik olyan $F_{n+1}(y_1, \dots, y_n, y_{n+1})$ polinom, amelyre $F_{n+1}(\sigma_1, \dots, \sigma_n, \sigma_{n+1}) = g_{k,n+1}$, mint az x_1, \dots, x_n, x_{n+1} határozatlanok polinomja.

Tudjuk, hogy arra a ξ , illetve η homomorfizmusra, amelyekre $\xi(x_{n+1}) = 0$ és $\xi(x_i) = x_i$, ha $i \leq n$, illetve $\eta(y_{n+1}) = 0$ és $\eta(y_i) = y_i$, ha $i \leq n$; továbbá azokra a Φ_t homomorfizmusokra, amelyek minden $i \leq t$ mellett y_i -t az i -edik elemi szimmetrikus polinomra képezik, fennáll a $\xi \Phi_{n+1} = \Phi_n \eta$ összefüggés. Az indukciós feltevés azt jelenti, hogy $\xi \Phi_{n+1}(F_{n+1}) = 0$, így $\Phi_n \eta(F_{n+1}) = 0$. Tekintettel arra, hogy Φ — mint tudjuk — injektív, ezért $\eta(F_{n+1}) = 0$ is igaz. Eszerint F_{n+1} -nek y_{n+1} -ben konstans tagja 0, azaz F_{n+1} osztható y_{n+1} -gyel. Eszerint $g_{k,n+1}$ osztható σ_{n+1} -gyel. Ennek a polinomnak a foka nagyobb, mint n , ami legalább akkora, mint k , azaz a $g_{k,n+1}$ polinom foka; ami csak akkor lehet, ha ez utóbbi a 0 polinom. ■

Térjünk most rá az úgynevezett alternáló polinomok vizsgálatára.

4.14. Definíció. Az $R[x_1, \dots, x_n]$ polinomgyűrű egy $f(x_1, \dots, x_n)$ elemét alternáló polinomnak nevezzük, ha bármely két határozatlan felcserélve a $-f(x_1, \dots, x_n)$ polinomba megy át (előjelet vált). ■

Az alternáló polinomok vizsgálatához szükségünk van a polinommátrixokra, illetve ezek determinánsára:

4.15. Definíció. Polinommátrix az olyan mátrix, amelynek elemei polinomok. ■

A mátrixokra vonatkozó eredmények alapján a polinommátrixokkal ugyanúgy számolhatunk, mint akármilyen számelemű mátrixszal. Négyzetes polinommátrixoknak létezik determinánsa, amelyet ugyanúgy számolhatunk ki, mint a számelemű mátrixét. Csupán arra kell vigyázni, hogy osztást ne végezzünk. Speciálisan, ha egy négyzetes mátrix két sora megegyezik, akkor determinánsa 0.

Az alternáló polinomoknál alapvető szerepet játszik az úgynevezett Vandermonde-féle determináns.

4.16. Definíció. Az x_1, \dots, x_n határozatlanok Vandermonde-féle mátrixa az a mátrix, amelyben az i -edik sor j -edik eleme x_i -nek a $(j-1)$ -edik hatványa. Ennek a mátrixnak a determinánsát nevezzük Vandermonde-determinánsnak. Ezt a determinánst $V = V_n = V(x_1, \dots, x_n)$ jelöli.

Az úgynevezett n -edrendű Vandermonde-determináns tehát a következő:

$$|V(x_1, \dots, x_n)| = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}. \quad \blacksquare$$

4.26. Tétel. *A Vandermonde-determináns a határozatlanjainak alternáló polinomja.*

Bizonyítás. Két határozatlan felcserélésekor a szereplő mátrix két sora felcserélődik. A determináns tulajdonságai alapján ekkor a determináns előjelet vált, tehát alternáló polinom. ■

A 4.26. Tétel bizonyos értelemben megfordítható. Ehhez azonban szükségünk lesz az alternáló polinomok előzetes vizsgálatára.

4.27. Tétel. *Az $R[x, y]$ polinomgyűrű minden alternáló polinomja osztható $(y - x)$ -szel.*

Bizonyítás. A 4.10. Tétel szerint az $f(x, y) \in R[x, y]$ polinom akkor és csak akkor osztható $(y - x)$ -szel, ha $f(x, x) = 0$. Márpedig $f(x, y) = -f(y, x)$ alapján $f(x, x) = -f(x, x)$, amiből $f(x, x) = 0$ következik. (Látható, hogy a bizonyítás nem vihető végbe például, ha az R gyűrű a modulo 2 vett maradékosztályokból áll.) ■

4.28. Tétel. *Az $R[x_1, \dots, x_n]$ polinomgyűrű minden $f(x_1, \dots, x_n)$ alternáló polinomja*

$$f(x_1, \dots, x_n) = V(x_1, \dots, x_n) \cdot g(x_1, \dots, x_n)$$

alakba írható, ahol $g(x_1, \dots, x_n)$ egy egyértelműen meghatározott szimmetrikus polinom. Ezen felül az is igaz, hogy a V Vandermonde-féle determináns megegyezik az összes olyan különböző $x_j - x_i$ polinom szorzatával, amelyben $j > i$.

Bizonyítás. Legyen f az adott polinomgyűrű tetszőleges alternáló polinomja. A 4.27. Tétel miatt f osztható minden egyes $x_j - x_i$ polinommal, ahol $j > i$ feltehető. Mivel ezek irreducibilisek és különbözőek, ezért f osztható ezek $p = p(x_1, \dots, x_n)$ szorzatával is. Így az

$$f(x_1, \dots, x_n) = p(x_1, \dots, x_n) \cdot g(x_1, \dots, x_n)$$

felbontáshoz jutottunk. Tekintettel arra, hogy p egyértelműen meghatározott, az f polinomot megadva g is egyértelműen meghatározott.

Tekintsük először speciálisan az $f = V$ esetet. Ebből a $V = ph$ felbontáshoz jutunk — valamilyen alkalmas h polinommal. Azt fogjuk megmutatni, hogy $h = 1$.

A szorzatfelbontás alapján ehhez elég annak a bizonyítása, hogy mind V -ben, mind p -ben ugyanaz az x_n foka és ugyanaz az x_n -ben vett főegyüttható.

Állításunkat n -re vonatkozó teljes indukcióval bizonyítjuk.

Az $n = 1$ esetben mindkét polinomot vehetjük 1-nek. Látható, hogy ez megegyezik a speciális esetként adódó definíciókkal. Egyébként az $n = 2$ esetben is fennáll az egyenlőség, mindkét polinom $x_2 - x_1$. Tegyük fel most, hogy n -re igaz az állítás, továbbá az is igaz, hogy $V_n = p_n$ különbözik 0-tól. V_{n+1} -nek az utolsó sor szerinti kifejtéséből azt kapjuk, hogy ez x_{n+1} -nek legfeljebb n -edfokú polinomja, továbbá x_{n+1} együtthatója éppen a hozzá tartozó aldetermináns — azaz V_n . Ez különbözik 0-tól, ezért V_{n+1} az x_{n+1} -nek n -edfokú polinomja; speciálisan maga is 0-tól különböző polinom. p_{n+1} -ben n olyan tényező

van, amelyben x_{n+1} szerepel, nevezetesen az $x_{n+1} - x_n, \dots, x_{n+1} - x_1$ alakú tényezők. Ezek szorzatában az x_{n+1} -ben legmagasabb fokú tag nyilván x_{n+1}^n . A többi tényezőben x_{n+1} nem szerepel, azaz x_{n+1} -ben konstans. Ezért x_{n+1}^n együtthatója ezek szorzata lesz. Ezek a tényezők azonban pontosan a p_n -ben fellépő tényezők, ami azt jelenti, hogy ez az együttható éppen p_n . A teljes indukciós feltétel szerint (figyelembe véve, hogy p_{n+1} a V_{n+1} -nek osztója!) következik, hogy $p_{n+1} = V_{n+1}$.

Így eljutottunk az $f = V_n g$ összefüggéshez, ahol f egy alternáló polinom, és V a Vandermonde-determináns. Mint láttuk, a g -t az f egyértelműen meghatározza. Láttuk, hogy V maga is alternáló polinom. Ezért bármely két határozatlan felcserélve a $-f = -(-V_n)g_1$ összefüggéshez jutunk, ahol g_1 a g -ből a szóban forgó két határozatlan felcserélésével keletkezik. Ebből azonnal adódik, hogy $g_1 = g$, vagyis g szimmetrikus. ■

Feladatok

1. Legyen K tetszőleges test, $R = K[x_1, \dots, x_n]$ és $\mathbf{f} = (f_1, \dots, f_n)$ stb. $K[x]$ -beli polinomok. Tekintsük ezeket a 4.9. Tételben definiált műveleteket. Bizonyítsuk be, hogy ha csak az R legfelsőbb elsőfokú elemeit tekintjük, akkor a polinomok az összeadás és a kompozíció műveletekre nemkommutatív gyűrűt alkotnak. Melyek e gyűrűben az (kompozícióra nézve) invertálható elemek és melyek a nullosztók?

2. Mutassuk meg, hogy a fenti R polinomgyűrű egy $f(x_1, \dots, x_n)$ elemének lehet olyan R -beli (g_1, \dots, g_n) gyöke, amelyikre $g_i \notin K$.

3. Defináljuk a többhatározatlanú polinomgyűrű ideálját az egyhatározatlanú esethez hasonlóan. Mutassuk meg, hogy minden k természetes számhoz van olyan ideál, amelyik nem generálható k -nál kevesebb polinommal.

4. Bizonyítsuk be, hogy többhatározatlanú polinomokra is érvényes az interpolációs tétel:

Legyenek $\mathbf{a}_i = (a_{i,1}, \dots, a_{i,k})$ ($i = 1, \dots, n$) adottak úgy, hogy a K test $a_{i,j}$ elemeire adott r, s esetén van olyan j , hogy $a_{r,j} \neq a_{s,j}$; továbbá $\mathbf{b} = (b_1, \dots, b_n)$ tetszőleges. Ekkor van olyan $f(x_1, \dots, x_k)$ polinom, amelyre $f(a_{i,1}, \dots, a_{i,k}) = b_i$.

5. Legyen p_i adott prímszám és K_i a modulo p_i vett maradékosztályok teste.

a) Adjunk meg olyan $m(x, y, z) \in K_i[x, y, z]$ polinomot, amelyre tetszőleges $a, b \in K_i$ esetén $m(a, a, b) = m(a, b, a) = m(b, a, a) = a$ (többségi polinom). Mutassuk meg, hogy a racionális számtest esetén nem létezik többségi polinom.

b) Adjunk meg olyan $q(x, y, z) \in K_i[x, y, z]$ polinomot, amelyre tetszőleges $a, b, c \in K_i$ esetén $q(a, b, c) = \begin{cases} a, & \text{ha } a = b \\ c, & \text{ha } a \neq b \end{cases}$ (duális diszkriminátor). Mutassuk meg, hogy a racionális számtest esetén nem létezik duális diszkriminátor polinom.

c) Adjunk meg olyan $p(x, y, z) \in K_i[x, y, z]$ polinomot, amelyre tetszőleges $a, b, c \in K_i$ esetén $p(a, b, c) = \begin{cases} c, & \text{ha } a = b \\ a, & \text{ha } a \neq b \end{cases}$ (diszkriminátor). Mutassuk meg, hogy a racionális számtest esetén nem létezik diszkriminátor.

6. Legyenek p_1, \dots, p_r különböző prímszámok és K_1, \dots, K_r a megfelelő prímszámok szerinti maradékosztályok teste. Bizonyítsuk be, hogy léteznek a $\mathbb{Z}[x, y, z]$ polinomgyűrűben olyan $m(x, y, z), q(x, y, z), p(x, y, z)$ polinomok, amelyek egyszerre adják a megfelelő többségi polinomot, duális diszkriminátor polinomot és diszkriminátor polinomot a K_i testekben.

7. Legyen \mathbb{Q} a racionális számtest. Határozzuk meg az alábbi polinomok gyökeit \mathbb{Q} -ban, $\mathbb{Q}[t]$ -ben és $\mathbb{Q}[u, v]$ -ben (t, u, v határozatlanok \mathbb{Q} felett):

$$x - y, \quad x^2 - y^2, \quad x^2 - y^3, \quad x^2 - y^2 - 1, \quad x^2 - y^2 - z^2.$$

(A legutóbbinál tegyük fel, hogy a \mathbb{Q} -beli gyökök ismertek.)

8. Legyen \mathbb{C} a komplex számtest, és legyen $n > 2$ természetes szám. Bizonyítsuk be, hogy a \mathbb{C} feletti $x^n + y^n - z^n$ polinom $\mathbb{C}[t]$ -beli gyökei $(a \cdot f(t), b \cdot f(t), c \cdot f(t))$ alakúak, ahol $f(t) \in \mathbb{C}[t]$ tetszőleges és az $a, b, c \in \mathbb{C}$ számokra $a^n + b^n = c^n$ teljesül.

9. Bizonyítsuk be, hogy azok az n -határozatlanú polinomok, amelyeknek előre megadott elem- n -esek gyökei, ideált alkotnak.

10. Bizonyítsuk be, hogy egy többhatározatlanú polinomgyűrűben azok a polinomok, amelyekben nem lép fel n -nél kisebb fokú tag, — a 0-val együtt — ideált alkotnak.

5. Lineáris egyenletrendszerek megoldása

A „klasszikus” algebra feladatának a magasabb fokú egyenletek megoldása mellett évszázadokon keresztül a lineáris egyenletrendszerek megoldását tekintették. A továbbiakban ezt vizsgáljuk. Egy olyan eljárást adunk, amelynek segítségével eldönthetjük, hogy az adott egyenletrendszer megoldható-e. Amennyiben igen, akkor az eljárással meghatározhatjuk az egyenletrendszer összes gyökét. Ez az eljárás az úgynevezett *Gauss-féle elimináció*.

4.17. Definíció. Legyen K egy számtest, és az $f_i(x_1, \dots, x_n)$ polinomok ($i = 1, \dots, k$) legyenek elemei a K feletti $K[x_1, \dots, x_n]$ polinomgyűrűnek. A K számtestbeli elemekből álló (a_1, \dots, a_n) mátrixot (elemsorozatot) e polinomok közös gyökének vagy az $\mathcal{F} = \{f_1, \dots, f_k\}$ polinomrendszer gyökének nevezzük, ha minden i index mellett teljesül az $f_i(a_1, \dots, a_n) = 0$ egyenlőség.

A K feletti $\mathcal{G} = \{g_1, \dots, g_\ell\}$ polinomrendszer az \mathcal{F} polinomrendszer következménye (a K felett), ha az \mathcal{F} polinomrendszer minden gyöke a \mathcal{G} polinomrendszernek is gyöke.

Két polinomrendszert (a K felett) ekvivalensnek nevezünk, ha gyökeik megegyeznek, azaz, ha mindegyik következménye a másiknak. ■

Megjegyzés. Következménynek azt érezzük, amit valamiképpen le tudunk vezetni. A levezetés definíciója viszont igen komplikált volna. Ezért célszerűbb a fenti definíció. Ezután már megállapíthatunk olyan eljárást, amely egy egyenletrendszernek a következményét állítja elő. Ennek segítségével kaphatunk esetenként az eredetivel ekvivalens, de könnyebben megoldható egyenletrendszert. □

(A továbbiakban a polinomgyűrűt mindig rögzítettnek gondoljuk.)

4.29. Tétel. Adott $\mathcal{F} = \{f_1, \dots, f_k\}$ polinomrendszerből rögzített $i \in \{1, \dots, k\}$ és $h_j \in K[x_1, \dots, x_n]$ ($j \neq i$) mellett készítsünk egy új $\mathcal{G} = \{g_1, \dots, g_k\}$ polinomrendszert úgy, hogy $g_i = f_i$ és $g_j = f_j - h_j f_i$ ($j \neq i$).

Ekkor a két polinomrendszer ekvivalens.

Bizonyítás. Először megmutatjuk, hogy a \mathcal{G} polinomrendszer következménye az \mathcal{F} -nek, akkor is, ha g_i helyett az $f_i - h_i f_i$ polinomot vesszük, egy tetszőleges K feletti h_i polinommal. Ebből speciális esetként adódik a kívánt következmény, a $h_i = 0$ választással. Ha (a_1, \dots, a_n) gyöke \mathcal{F} -nek, akkor

$$g_j(a_1, \dots, a_n) = f_j(a_1, \dots, a_n) - h_j(a_1, \dots, a_n)f_i(a_1, \dots, a_n) = 0 - 0 = 0$$

alapján gyöke \mathcal{G} -nek is.

Az ekvivalenciához elég belátni, hogy \mathcal{F} is következménye \mathcal{G} -nek. Ez viszont azonnal adódik abból, hogy $f_i = g_i$ és $f_j = g_j - (-h_j)g_i$; aminek a következtében alkalmazható az előbbi eljárás. ■

4.30. Tétel. Legyenek $\ell_1, \dots, \ell_k \in K[x_1, \dots, x_n]$ olyan elsőfokú polinomok, amelyekre az ℓ_2, \dots, ℓ_k polinomok x_1 -ben konstansok. Ekkor az $\mathcal{L} = \{\ell_1, \dots, \ell_k\}$ polinomrendszer minden (a_1, a_2, \dots, a_k) megoldására (a_2, \dots, a_k) megoldása az $\mathcal{L}_1 = \{\ell_2, \dots, \ell_k\}$ polinomrendszernek megoldása. Amennyiben ℓ_1 az x_1 -ben elsőfokú, akkor az \mathcal{L}_1 minden egyes (a_2, \dots, a_k) megoldásához található olyan egyértelműen meghatározott a_1 , hogy az \mathcal{L} polinomrendszernek megoldása (a_1, a_2, \dots, a_k) .

Bizonyítás. Az első állítás nyilvánvaló, hiszen az egész rendszer minden megoldása gyöke egy részrendszernek is, de a szóban forgó részrendszer polinomjaiban x_1 nem szerepel. Fordítva is, ha van a részrendszernek egy adott típusú megoldása, akkor ezt az első polinomba behelyettesítve egy x_1 -ben elsőfokú K -beli együtthatós polinomot kapunk, amelynek pontosan egy a_1 gyöke van. Feltétel szerint (a_1, a_2, \dots, a_k) gyöke az \mathcal{L}_1 polinomrendszernek (hiszen ezek x_1 -től „függetlenek”), és a most belátottak alapján gyöke az \mathcal{L} polinomnak is. ■

4.31. Tétel. Az $\mathcal{L} = \{\ell_1, \dots, \ell_k\}$ lineáris polinomrendszernek abban az esetben, amikor minden egyes ℓ_i konstans, pontosan akkor van gyöke, ha minden i -re $\ell_i = 0$.

Ebben az esetben minden szóba jövő (a_1, \dots, a_n) mátrix gyök.

Bizonyítás. Egy konstans helyettesítési értéke mindig önmaga. Tehát csak akkor van gyök, ha mindig $\ell_i = 0$ teljesül. Ekkor viszont ez az összefüggés minden helyettesítés mellett fennáll. ■

A fentiek figyelembevételével a következő eljárást adhatjuk egy lineáris polinomrendszer gyökeinek a megkeresésére:

1. Ha a polinomrendszer minden polinomja konstans, akkor a 4.31. Tétel szerint a megoldhatóság feltétele az, hogy mindegyik polinom 0 legyen. Ebben az esetben minden szóba jövő értékrendszer megoldás lesz.

2. Ha a polinomrendszerben van olyan polinom, amelyik valamelyik határozatlanban elsőfokú, akkor feltesszük, hogy ez az első polinom (átrendezéssel ez elérhető), és a szóba jövő határozatlan (esetleges átjelölés után) x_1 . Most az első polinom megfelelő többszöröseit levonva a többi polinomból, elérhetjük, hogy az első polinom kivételével a többi x_1 -ben konstans. Az új rendszernek a 4.29. Tétel alapján ugyanazok a megoldásai, mint az eredetinek.

3. Az első polinomot elhagyva és az x_1 határozatlantól eltekintve megnézzük a maradék rendszer megoldhatóságát, és meghatározzuk a megoldásait. (Ez is a most ismertetendő eljárással történik, de rekurzíven elvégzettnek tekinthető, mert kevesebb a határozatlan.) A 4.30. Tétel alapján ebből az eredeti rendszer megoldásai egyértelműen meghatározhatók.

Ez az eljárás a gyakorlatban is igen fontos **Gauss-féle elimináció**. Ennek segítségével bármely lineáris egyenletrendszerről eldönthető, hogy megoldható-e; amennyiben igen, akkor az eljárás a megoldásokat is szolgáltatja.

Maga a módszer olyan lépésekből áll, amelyeknek az absztrakt vizsgálata elvezet a gyakorlatban és elméletben is ugyancsak fontos lineáris algebra tárgyalásához.

II. rész

LINEÁRIS ALGEBRA

ELSŐ FEJEZET

VEKTORTEREK

1. A vektortér fogalma és elemi tulajdonságai

A lineáris egyenletrendszerek megoldásánál láttuk, hogy az egész megoldási módszer azon múlt, hogy a lineáris polinomokat egy-egy számmal szoroztuk, és ezeket összeadtuk. A lineáris algebra tárgya olyan rendszerek vizsgálata, amelyekben az összeadás és a számmal való szorzás végezhető el. Tipikusan ilyen rendszer a sík vagy a tér vektorainak a rendszere. Éppen ebből ered a *vektortér* elnevezés is. Az összeadás és a számmal való szorzás „linearitását” tükrözi e fogalomnak a másik neve, a *lineáris tér*. Ezekben az esetekben kétféle fogalom szerepel, „amiket összeadunk” és „amiket szorzunk”, ezek neve *vektor* és „amikkel szorzunk”, ezek neve *skalár*.

Ezekén kívül azonban a matematika szinte minden ágában találkozhatunk vektorterekkel. A gyakorlati alkalmazásokban is igen sok helyen lép fel ez a fogalom. Ezekben az esetekben a lineáris algebra további fejezeteit, illetve ezeknek az eredményeit is rendszeresen felhasználják. Mindezek indokolják a vektorterek részletes, alapos tárgyalását. Számos gyakorlati és elméleti felhasználás esetében nem csak számok, hanem maradékosztályok is fellépnek. Éppen ezért a skalároknak nem csak számokat engedünk meg, hanem tetszőleges test elemeit.

Ha az egyenletrendszerek általános tárgyalására gondolunk, akkor láthatjuk, hogy sok esetben nem csak számmal, de polinomokkal is szorozhatunk. Ez azt mutatja, hogy bizonyos esetekben szükség lehet arra, hogy a skalárokat egy gyűrűből vegyük. Erre mutat az is, hogy egyes geometriai feladatok megoldásánál egész koordinátájú vektorok lépnek fel. Ennél talán fontosabb az a tény, hogy a vektorterek esetében vannak olyan fogalmak, amelyek nem igazán különböztethetők meg egymástól, de ha a skalárokat egy gyűrűből vesszük, akkor azonnal látható ezek különbözősége.

Az eddigiekben elsősorban a számok és polinomok szerkezeti tulajdonságait igyekeztünk leírni, rámutatva arra, hogy ezek a tulajdonságok nem kizárólagosak. Még a mátrixok esetében sem volt semmi olyan fogalom, amelyet nem lehetett volna a középiskolai tanulmányok folyamán megérteni. Itt azonban az a cél, hogy egy általános, absztrakt fogalmat vezessünk be. Ettől fogva semmi mást nem szabad felhasználni, csak a feltett vagy már

bebizonyított tulajdonságokat. Mindenekelőtt megismétlünk néhány elnevezést, amelyek már szerepeltek:

Ha bizonyos elemek körében elvégezhető az összeadás, kivonás és szorzás, amely műveletek rendelkeznek a már ismertetett fontos tulajdonságokkal, akkor *gyűrűről* beszélünk. Ha a szorzás kommutatív, akkor ez egy *kommutatív* gyűrű. Egy gyűrű elemeinek a halmaza nem lehet üres; mindig van benne „nullelem”, azaz egy olyan 0 elem, amelyet a gyűrű bármely a eleméhez adva ismét a -t kapunk. Ha két elem szorzata csak akkor lehet 0 , ha valamelyikük 0 , akkor a gyűrű *nulloosztómentes*. Nulloosztómentes kommutatív gyűrű neve *integritási tartomány*. (Mi azt is feltesszük, hogy legalább két eleme van.) Ha létezik benne „maradékos osztás” egyre csökkenő „euklideszi normával”, akkor ez egy *euklideszi gyűrű*. Ha van olyan elem (1) , amellyel bármely a elemet megszorozva az eredmény ismét a , akkor ezt *egységelemnek*; és a gyűrűt *egységelemesnek* nevezzük. Ha a gyűrűben elvégezhető az osztás, akkor ez *test*. A továbbiakban elegendő, ha az ezekre való utalásnál mindig olyan példákra gondolunk, amelyeket az eddigiek során már megismertünk.

1.1. Definíció. Elemek egy \mathcal{V} halmazát a K test feletti vektortérnek nevezzük, ha értelmezve van a \mathcal{V} elemei között egy művelet, amelyet vektorösszeadásnak nevezünk, értelmezve van a \mathcal{V} elemeinek a K elemeivel való szorzása, amelyet skalárral való szorzásnak nevezünk; és ezekre a műveletekre az alábbiak teljesülnek:

I. \mathcal{V} a vektorösszeadásra nézve kommutatív csoport, azaz bármely $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ elempár-hoz hozzá van rendelve egy $\mathbf{u} + \mathbf{v} \in \mathcal{V}$ összeg a következő tulajdonságokkal:

- (1) Az összeadás kommutatív, azaz $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ ($\mathbf{u}, \mathbf{v} \in \mathcal{V}$).
- (2) Az összeadás asszociatív, azaz $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ ($\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$).
- (3) Létezik nullelem, azaz olyan $\mathbf{o} \in \mathcal{V}$, amire $\mathbf{o} + \mathbf{u} = \mathbf{u}$, ha $\mathbf{u} \in \mathcal{V}$.
- (4) Minden elemnek létezik a nullelemre vonatkozó additív inverze, azaz ha $\mathbf{u} \in \mathcal{V}$, akkor van olyan $\mathbf{v} \in \mathcal{V}$, amire $\mathbf{v} + \mathbf{u} = \mathbf{o}$.

Megjegyezzük, hogy sem a nullelem, sem az inverz egyértelműségét nem követeltük meg; ezeket majd bizonyítani fogjuk. Ezért az is gondot okoz, hogy az inverznél melyik nullelemet kell venni. Mivel az egyértelműség teljesül, azért itt feltehetjük azt is, hogy valamelyikre vonatkozó inverzről van szó, azt is, hogy mindig ugyanarról, kinek hogy tetszik.

II. A K elemeivel szorozhatjuk a \mathcal{V} elemeit, pontosabban szólva a K minden a eleméhez tartozik \mathcal{V} -nek egy „egyváltozós művelet”-e, amelyet $a : \mathbf{x} \rightarrow a\mathbf{x}$ (vagy $\rightarrow a \cdot \mathbf{x}$) jelöl ($\mathbf{x} \in \mathcal{V}$); és amelyre a következő azonosságok teljesülnek:

- (1) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$.
- (2) $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.
- (3) $(ab)\mathbf{u} = a(b\mathbf{u})$.
- (4) $1 \cdot \mathbf{u} = \mathbf{u}$.

A K test feletti vektorterekre a ${}_K\mathcal{V}$ jelölés használatos. Ha a test eleve adott, vagy adottnak gondolt, akkor elegendő a \mathcal{V} jelölés is.

\mathcal{V} elemeit vektoroknak, K elemeit skalároknak nevezzük. ■

Teljesen hasonlóan értelmezhető egy adott gyűrű fölötti modulus:

1.1.a) Definíció. Elemek egy \mathcal{M} halmazát az R gyűrű feletti (bal oldali) R -modulusnak nevezzük, ha értelmezve van a \mathcal{V} elemei között egy művelet, amelyet összeadásnak nevezünk, értelmezve van az \mathcal{M} elemeinek az R elemeivel balról való szorzása; és ezekre a műveletekre az alábbiak teljesülnek:

I. \mathcal{M} az összeadásra nézve kommutatív csoport, azaz fennállnak az 1.1. Definícióbeli I. alatti feltételek.

II. Az R elemeivel szorozhatjuk az \mathcal{M} elemeket, pontosabban szólva az R minden r eleméhez tartozik \mathcal{M} -nek egy „egyváltozós művelet”-e, amelyet $r : \mathbf{x} \mapsto r\mathbf{x}$ (vagy $\mapsto r \cdot \mathbf{x}$) jelöl ($\mathbf{x} \in \mathcal{M}$); és amelyre a következő azonosságok teljesülnek:

$$(1) \quad r(\mathbf{u} + \mathbf{v}) = r\mathbf{u} + r\mathbf{v}.$$

$$(2) \quad (r + s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}.$$

$$(3) \quad (rs)\mathbf{u} = r(s\mathbf{u}).$$

$$(4) \quad 1 \cdot \mathbf{u} = \mathbf{u}.$$

Az R gyűrű feletti bal oldali modulusokra az ${}_R\mathcal{M}$ jelölés használatos. Ha a gyűrű eleve adott, vagy adottnak gondolt, akkor elegendő az \mathcal{M} jelölés is. ■

Megjegyzések

1. Az 1.1.a) Definíciónál a II.(4) feltétel teljesülése esetében *unitér* R -modulusról beszélünk. A lineáris algebra tárgyalásánál mi mindig unitér modulusokat fogunk tekinteni, akkor is, ha ezt nem mondjuk. Nem unitér modulusok esetén lehetséges az is, hogy mindig $r\mathbf{x} = \mathbf{o}$ teljesül. Ilyen esetben *triviális* R -modulusról beszélünk.

2. A fentihez hasonlóan tekinthetünk jobb oldali R -modulust is. Ekkor II.(3) helyett $\mathbf{u}(r) = (\mathbf{u}r)$ teljesül; ami csak kommutatív R esetén lesz biztosan „ugyanaz”, mint a „balszorzás”. Ebben az esetben az \mathcal{M}_R jelölést használjuk.

Az R elemeket itt is *skalároknak*, az \mathcal{M} elemeket itt is *vektoroknak* fogjuk nevezni.

3. Tekintettel arra, hogy itt nagyon sokféle „mennyiség” szerepel, ezért célszerű következetes jelölést használni. Ennek megfelelően a *gyűrűket és testeket* nagy latin betűkkel, *ezek elemeit* kis latin betűkkel, a *vektortereket és modulusokat* írott nagybetűkkel, *ezek elemeit* vastagított kisbetűkkel fogjuk jelölni. K, L általában testeket, R, S általában gyűrűket; \mathcal{U}, \mathcal{V} általában vektortereket, \mathcal{M}, \mathcal{N} általában modulusokat jelöl. Vektorterek és modulusok részhalmazait vastagított nyomtatott nagybetűkkel fogjuk jelölni (\mathbf{U}, \mathbf{V}) stb.

4. Általában az algebrai „struktúrák” vizsgálatánál meg szokták kívánni, hogy az elemeinek a halmaza nem üres. Itt az I.(3) követelménynek köszönhetően erre nincs szükség. □

Ismételten felhívjuk a figyelmet arra a különbségre, ami az elemi algebrai és a lineáris algebrai tárgyalás között fennáll. Az elemi algebrai eredmények esetében mindig lehetnek (sőt vannak!) a számoknak olyan tulajdonságai, amelyeknek az igazságára a figyelembe vett tulajdonságok igazságából nem tudunk következtetni. *Ilyen esetekben saját magunknak kell eldönteni, hogy egy ilyen állítást igaznak tételezzünk-e fel, vagy sem.* Sőt, mi több, egyes esetekben a szemléletünk alapján állíthatjuk valamiről, hogy igaz.

A vektorterek és a modulusok esetében viszont nem ez a helyzet! A fenti két definícióban ugyanis *pontosan megmondtuk, hogy mik a vektorterek, illetve a modulusok.* Ez azt jelenti, hogy az ezekre vonatkozó állítások esetében a bizonyításnál *csakis azokat a*

tulajdonságokat használhatjuk, amelyeket a definíciókban kimondtunk. Más szóval a bizonyítások nem lehetnek „szemléletesek”. Az egy más kérdés, hogy a tételek igazságának és a bizonyítás menetének az elképzelésénél támaszkodhatunk szemléletünkre, elsősorban a síkbeli és a térbeli vektorok tulajdonságaira.

Mivel a skalárok számok, polinomok, esetleg mátrixok lehetnek, ezért ezekre automatikusan nem érvényes a fenti megállapítás; azaz nincsenek előre megadott tulajdonságaik. Ennek ellenére törekedni fogunk arra, hogy a skalárokról is csak annyit tegyünk fel, hogy gyűrűt, illetve testet alkossanak.

A fenti definíciókban szereplő összefüggések a **vektortér axiómái**. Ezek, eltekintve azoktól, amelyek a nullelem, illetve az additív inverz létezését kívánják meg, mind **azonosságok**.

Ezek után lássunk a vektorterekre és a modulusokra néhány példát:

1. példa: A tér (vagy a sík) vektorai vektorteret alkotnak a valós számtest felett, ha az összeadást és a skalárral való szorzást a geometriában ismert módon értelmezzük.

2. példa: Az ugyanolyan alakú mátrixok vektorteret alkotnak a fölött a számtest fölött, amelyből az elemeik valók, ha az összeadást és a számmal való szorzást úgy értelmezzük, ahogy ezt a mátrixok tárgyalásánál tettük. Ha az elemek egy számgyűrűből vagy egy polinomgyűrűből valók, akkor ezek a mátrixok modulust alkotnak e gyűrű felett.

3. példa: Egy adott számtestbeli végtelen sorozatok is vektorteret alkotnak a fölött a számtest fölött, amelyből az elemei valók, ha a műveleteket hasonlóan értelmezzük, mint az előző példában. Ha az elemek egy gyűrűből valók, itt is modulust nyerünk.

4. példa: Adott testbeli együtthatós polinomok vektorteret alkotnak a fölött a test fölött, amelyből az együtthatók valók, ha a műveleteket a polinomoknál tárgyalt módon értelmezzük. Ugyanezen műveletekre nézve modulust alkotnak az egész együtthatós polinomok az egész számok gyűrűje fölött. Hasonlóképpen modulust kapunk egy többhatározatlanú polinomgyűrű felett, ha további határozatlanokat veszünk hozzá. Egy test feletti többhatározatlanú polinomok e test felett vektorteret alkotnak.

5. példa: Egy adott zárt (vagy nyílt) intervallumban értelmezett valós függvények vektorteret alkotnak a valós számtest felett, ha az összeadás a függvényösszeadás és a skalárral való szorzás a függvény számmal való szorzása. Hasonlóképpen vektorteret kapunk, ha a folytonos, vagy integrálható, vagy differenciálható függvényeket tekintjük. (De nem kapunk vektorteret, ha a monoton függvényeket nézzük.)

6. példa: A komplex számok vektorteret alkotnak a valós számtest felett, ha az összeadást mint komplex számok összeadását értelmezzük és a skalárral való szorzást mint a komplex számoknak a valós számokkal való szorzását.

7. példa: Mind a valós számok, mind a komplex számok vektorteret alkotnak a racionális számtest felett, ha az összeadást úgy értelmezzük, mint a számok összeadását, a skalárral való szorzást pedig úgy, mint a számok szorzását.

A további példák előtt szükségünk van egy gyűrűelméleti fogalomra és egy tételre:

Definíció. Azt mondjuk, hogy S részgyűrűje az R gyűrűnek, ha S minden eleme R -nek is eleme, és S gyűrű az R -beli műveletekre. Ezt a kapcsolatot $S \leq R$ jelöli. ■

Noha a modulusokat mindig egy rögzített gyűrű felett vizsgáljuk, mégis a modulus sok esetben „természetesen válik” egy másik gyűrű feletti modulusá:

Tétel. Legyen $\varphi : S \rightarrow R$ egy tetszőleges gyűrűhomomorfizmus, és \mathcal{M} egy R -modulus. Ekkor \mathcal{M} az

$$s \cdot \mathbf{x} = \varphi(s) \cdot \mathbf{x} \quad (s \in S) \quad (\mathbf{x} \in \mathcal{M})$$

definícióval S -modullá válik.

A két legfontosabb eset az, amikor S az R részgyűrűje (tehát φ injektív); és amikor S az R feletti polinomgyűrű, ekkor φ egy behelyettesítés.

Bizonyítás. A modulusokat definiáló feltételek teljesülése könnyen ellenőrizhető. ■

8. példa: Ha $K \leq L$ testek, akkor L vektortér K felett. Ha $S \leq R$ gyűrűk, akkor R modulus S felett.

9. példa: Ha K test, R gyűrű és $K \leq R$, akkor R vektortér a K felett, amennyiben K egységeleme az R -nek is egységeleme. Mindig ez a helyzet, ha R nullosztómentes.

10. példa: Legyen K test, ekkor a K feletti \mathcal{V} vektortér modulus a $K[x]$ polinomgyűrű felett.

11. példa: Legyen \mathcal{M} modulus az R gyűrű felett. Ekkor \mathcal{M} modulus az $R[x_1, \dots, x_n]$ polinomgyűrű felett.

12. példa: Legyen I polinomideálja a $K[x]$ polinomgyűrűnek (vagy ideálja egy tetszőleges gyűrűnek), ekkor I modulus e gyűrű felett.

13. példa: Legyen R egy egységelemes integritási tartomány és $\varphi : R \rightarrow S$ egy szürjektív homomorfizmus. Ekkor S egy R -modulus a $c \cdot \varphi(r) = \varphi(cr)$ definícióval ($c, r \in R$). Speciálisan a modulo m vett maradékosztályok \mathbb{Z}_m gyűrűje modulus az egész számok gyűrűje fölött.

14. példa: Tetszőleges G kommutatív csoport \mathbb{Z} -modulusnak tekinthető az $n \cdot g = \underbrace{g + \dots + g}_{n\text{-szer}}, 0 \cdot g = 0$ és $(-n) \cdot g = n \cdot (-g)$ definícióval ($n \in \mathbb{N}$). ■

Tekintettel arra, hogy igen nehézkes volna minden bizonyítást mindig az axiómákra visszavezetni, ezért mindenekelőtt néhány egyszerűbb következményt bizonyítunk be, amelyeket aztán a továbbiakban felhasználhatunk.

1.1. Tétel. Többtagú vektorösszeg eredménye nem függ a tagok társításától és sorrendjétől.

Bizonyítás. Azt fogjuk megmutatni, hogy akárhogyan társítva és felcserélve az $\mathbf{u}_1, \dots, \dots, \mathbf{u}_n$ vektorok összege mindig ugyanaz lesz. Nevezetesen azt mutatjuk meg, hogy minden ilyen összeg megegyezik egy előre kiválasztottal. Ezt az előre kiválasztott összeget úgy kapjuk, hogy minden részösszeghez a soron következő vektort adjuk hozzá:

$$\mathbf{v}_1 = \mathbf{u}_1, \mathbf{v}_2 = \mathbf{v}_1 + \mathbf{u}_2, \dots, \mathbf{v}_i = \mathbf{v}_{i-1} + \mathbf{u}_i, \dots, \mathbf{v}_n = \mathbf{v}_{n-1} + \mathbf{u}_n.$$

Állításunkat teljes indukcióval bizonyítjuk. Az $n = 1$ esetben nincs mit bizonyítani, és az $n = 2$ esetben az állítás azonnal következik a kommutativitásból.

Tegyük fel, hogy az állítás igaz minden olyan k -ra, amelyre $k < n$. Többtagú összeget nem definiáltunk, ezért a fenti vektorokból képezett minden \mathbf{w} összeg $\mathbf{x} + \mathbf{y}$ alakú, ahol \mathbf{e} két vektor az adott vektorok közül bizonyosaknak az összege és a tekintett vektorok mindegyike \mathbf{e} két tag közül pontosan az egyikben fordul elő. A kommutativitás miatt feltehetjük, hogy \mathbf{u}_n az \mathbf{y} -ban. Az $\mathbf{y} = \mathbf{u}_n$ esetben az indukciós feltétel szerint $\mathbf{x} = \mathbf{v}_{n-1}$, amiből $\mathbf{w} = \mathbf{v}_{n-1} + \mathbf{u}_n = \mathbf{v}_n$ következik. Egyébként az indukciós feltételből következik, hogy $\mathbf{y} = \mathbf{z} + \mathbf{u}_n$. Az asszociativitás alapján $\mathbf{w} = \mathbf{x} + (\mathbf{z} + \mathbf{u}_n) = (\mathbf{x} + \mathbf{z}) + \mathbf{u}_n$; és itt ismét az első esettel állunk szemben. ■

Megjegyzés. Az 1.1. Tétel alapján az összegekben minden zárójelezést el lehet hagyni, és ezzel a lehetőséggel általában élni is fogunk. Akkor fogjuk csak kitenni a zárójeleket, ha ezek megkönnyítik a bizonyítás követését. □

1.2. Tétel. Az I. (3) axiómában szereplő nullvektor és az I. (4) axiómában szereplő ellentett egyértelműen meghatározott.

Bizonyítás. Legyen \mathbf{o}' a vektortérnek egy olyan eleme, amely ugyancsak eleget tesz a kirótt követelménynek. Eszerint $\mathbf{o}' + \mathbf{o} = \mathbf{o}$. Tekintettel arra, hogy \mathbf{o} nullvektor, ezért $\mathbf{o}' + \mathbf{o} = \mathbf{o}'$ is igaz.

Legyen most \mathbf{v} és \mathbf{w} mindegyike a \mathbf{u} vektor additív inverze, azaz $\mathbf{v} + \mathbf{u} = \mathbf{w} + \mathbf{u} = \mathbf{o}$. Ebből

$$\mathbf{v} = \mathbf{v} + \mathbf{o} = \mathbf{v} + (\mathbf{u} + \mathbf{w}) = (\mathbf{v} + \mathbf{u}) + \mathbf{w} = \mathbf{o} + \mathbf{w} = \mathbf{w}$$

következik. ■

1.3. Tétel. A K test feletti \mathcal{V} vektortérben $c\mathbf{u} = \mathbf{o}$ ($c \in K$), ($\mathbf{u} \in \mathcal{V}$) pontosan akkor igaz, ha vagy $c = 0$, vagy $\mathbf{u} = \mathbf{o}$.

Modulusok esetében $c\mathbf{u} = \mathbf{o}$ akkor is lehet, ha $c = 0$ és $\mathbf{u} = \mathbf{o}$ egyike sem igaz.

Bizonyítás. Legyen $\mathbf{u} \in \mathcal{V}$ és \mathbf{v} a $0\mathbf{u}$ ellentettje. Rendre felhasználva az ellentett definícióját, a II. (1) azonosságot, az összeadás asszociativitását, majd ismét az ellentett és végül a nullvektor definícióját, a következőket kapjuk:

$$\mathbf{o} = \mathbf{v} + 0\mathbf{u} = \mathbf{v} + (0 + 0)\mathbf{u} = \mathbf{v} + (0\mathbf{u} + 0\mathbf{u}) = (\mathbf{v} + 0\mathbf{u}) + 0\mathbf{u} = \mathbf{o} + 0\mathbf{u} = 0\mathbf{u}.$$

Az előbbihez hasonlóan, de II. (1) helyett II. (2)-t használva

$$\mathbf{o} = \mathbf{w} + c\mathbf{o} = \mathbf{w} + c(\mathbf{o} + \mathbf{o}) = \mathbf{w} + (c\mathbf{o} + c\mathbf{o}) = (\mathbf{w} + c\mathbf{o}) + c\mathbf{o} = \mathbf{o} + c\mathbf{o} = c\mathbf{o}$$

adódik, ahol \mathbf{w} a $c\mathbf{o}$ ellentettje.

Tegyük most fel, hogy \mathcal{V} vektortér a K test felett, $c \in K$, $\mathbf{u} \in \mathcal{V}$ és $c\mathbf{u} = \mathbf{o}$. Ha $c = 0$, akkor igaz az állítás. Egyébként létezik olyan $d \in K$, amelyre $dc = 1$; hiszen K test. Most a II. (4), majd a II. (3) azonosság alapján az imént bizonyított egyenlőséget felhasználva azt kapjuk, hogy:

$$\mathbf{u} = 1\mathbf{u} = (dc)\mathbf{u} = d(c\mathbf{u}) = d\mathbf{o} = \mathbf{o}.$$

Világos, hogy modulusok esetében a megfelelő feltételek hiányában ez az eljárás értelmetlen. Egyébként a későbbiekben számos olyan példát fogunk látni, amikor egy modulusban $c\mathbf{u} = \mathbf{o}$, noha $c \neq 0$ és $\mathbf{u} \neq \mathbf{o}$. ■

Megjegyzés. A szereplő test nullelemét nem szabad összetéveszteni a vektortér vagy a modulus nullelemével. A fenti tétel azonban azt mutatja, hogy ezek a nullelemek elég hasonlóan viselkednek. \square

1.4. Tétel. Az \mathbf{u} vektor ellentettje megegyezik az \mathbf{u} vektor (-1) -szeresével, c -szeresének az ellentettje az ellentettjének a c -szeresével, illetve e vektor $(-c)$ -szeresével. Az \mathbf{u} vektort n -szer összeadva a kapott vektor $n \cdot \mathbf{u}$ (ha $n \in R$).

Bizonyítás. A II. (4) és II. (1) axiómák szerint:

$$(-1)\mathbf{u} + \mathbf{u} = (-1)\mathbf{u} + 1\mathbf{u} = ((-1) + 1)\mathbf{u} = 0\mathbf{u} = \mathbf{o},$$

felhasználva az 1.3. Tételt. Az 1.2. Tétel alapján ebből következik az első állítás.

A megfelelő azonosságokat és az 1.3. Tételt felhasználva nyerjük, hogy:

$$c(-\mathbf{u}) + c\mathbf{u} = c((- \mathbf{u}) + \mathbf{u}) = c\mathbf{o} = \mathbf{o} \quad \text{és} \quad (-c)\mathbf{u} + c\mathbf{u} = (-c + c)\mathbf{u} = 0\mathbf{u} = \mathbf{o}.$$

Az 1.2. Tételt felhasználva adódik ezekből a második állítás.

A harmadik állítás nyilvánvaló teljes indukcióval bizonyítható a II. (4) axiómából és az $(n+1)\mathbf{u} = n\mathbf{u} + 1\mathbf{u}$ összefüggésből. \blacksquare

Kiegészítés. $\mathbf{x} = \mathbf{a} + (-\mathbf{b})$ az $\mathbf{x} + \mathbf{b} = \mathbf{a}$ egyenlet egyértelmű megoldása. Ezt a megoldást $\mathbf{a} - \mathbf{b}$ fogja jelölni, és ezt az elemet az \mathbf{a} és \mathbf{b} elemek ebben a sorrendben vett különbségének nevezzük. A különbségre teljesülnek az alábbi azonosságok:

$$(\mathbf{a} + \mathbf{b}) - (\mathbf{c} + \mathbf{d}) = (\mathbf{a} - \mathbf{c}) + (\mathbf{b} - \mathbf{d}) = \mathbf{a} + \mathbf{b} + (-\mathbf{c}) + (-\mathbf{d}),$$

$$\mathbf{a} - \mathbf{o} = \mathbf{a} \quad \text{és} \quad \mathbf{o} - \mathbf{b} = -\mathbf{b}.$$

Bizonyítás. $\mathbf{x} = (\mathbf{x} + \mathbf{b}) + (-\mathbf{b}) = \mathbf{a} + (-1)\mathbf{b}$ bizonyítja az egyértelműséget. $(\mathbf{a} + (-\mathbf{b})) + \mathbf{b} = \mathbf{a} + ((-\mathbf{b}) + \mathbf{b}) = \mathbf{a} + \mathbf{o} = \mathbf{a}$ pedig azt adja, hogy valóban megoldást kaptunk. A felírt azonosságok bizonyítását az olvasóra bízuk. \blacksquare

Megjegyzés. Az 1.4. Tétel szerint $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-1)\mathbf{v}$. Így vektortereknél a kivonásra nincs szükség, mert ez mindig megadható skalárszoros és összeadás felhasználásával. Modulusok esetében ez általában nincs így. \square

Feladatok

1. Bizonyítsuk be, hogy a felsorolt példákban valóban vektorterek, illetve modulusok szerepelnek.

2. Legyen $R = K_n$ a K feletti $(n \times n)$ -es mátrixok gyűrűje. Bizonyítsuk be, hogy az R -ből vett elemekből képezett k hosszúságú (r_1, \dots, r_k) sormátrixok R felett egy bal oldali ${}_R\mathcal{M}$, illetve egy jobb oldali \mathcal{M}_R modulust alkotnak, ha az összeadás a (sor)mátrixösszeadás és az $s \in R$ elemmel való szorzás (sr_1, \dots, sr_k) , illetve $(r_1s, \dots, r_k s)$. Bizonyítsuk be, hogy ez egy úgynevezett kettősmodulus, azaz $r, s \in R$, $\mathbf{u} \in \mathcal{M}$ esetén $r(\mathbf{u}s) = (r\mathbf{u})s$. Ezt ${}_R\mathcal{M}_R$ jelöli.

3. Mutassuk meg, hogy az előző feladatban van olyan $\mathbf{u} \in \mathcal{M}$ és $r \in R$, hogy $r\mathbf{u} = \mathbf{u}s$ egyetlen R -beli s -re sem igaz.

4. Bizonyítsuk be, hogy a pozitív valós számok \mathbb{R} halmaza vektorteret alkot a \mathbb{Q} racionális számtest felett az $\mathbf{u} \cdot \mathbf{v}$ összeadásra és az $r : \mathbf{u} \rightarrow \mathbf{u}^r$ skalárral való szorzásokra nézve ($\mathbf{u}, \mathbf{v} \in \mathbb{R}$ és $r \in \mathbb{Q}$).

5. Bizonyítsuk be, hogy a pozitív racionális számok \mathbb{Z} -modulust alkotnak, ha a vektorösszeadás a szorzás és az n egész számmal való szorzás az n -edik hatvány.

6. Bizonyítsuk be, hogy a nemnulla valós számok modulust alkotnak az egészek felett, ha a vektorösszeadás a szorzás és az n egész számmal való szorzás az n -edik hatvány.

7. Bizonyítsuk be, hogy a komplex számok alábbi halmazai modulust alkotnak az egészek \mathbb{Z} gyűrűje felett, ha az összeadást $\mathbf{u} \cdot \mathbf{v}$ és a skalárral való szorzást $n : \mathbf{u} \rightarrow \mathbf{u}^n$ definiálja ($n \in \mathbb{Z}$).

1. A nemnulla komplex számok.
2. Az 1 abszolút értékű komplex számok.
3. A komplex egységgyökök.
4. A pozitív valós számok.

2. Lineáris kombináció és lineáris függés

Sok esetben előfordul, hogy vektorok (vagy más elemek) felsorolásakor egy-egy vektor többször is előfordul. Például egy mátrix sorai maguk is mátrixok. Ezeket felsorolva lehetséges, hogy ugyanaz a sor többször is szerepel, de más helyeken. Tekintettel arra, hogy egy halmazban minden elem csak egyszer szerepelhet, ezért ilyen esetben nem beszélhetünk vektorhalmazról. Másrészt, éppen az említett esetben bizonyos fokig teljesen lényegtelen, hogy milyen sorrendben vesszük a mátrix sorait, ezért a vektorsorozat elnevezés sem célszerű. Ennek ellenére kénytelenek leszünk valami hasonlót mondani:

Definíció. Adott $\mathbf{v}_1, \dots, \mathbf{v}_n$ vektorok esetén vektorrendszeréről beszélünk. Ezt a rendszert $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ jelöli. A vektorrendszer nem változik meg, ha elemeit permutáljuk.

Hasonlóan értelmezhető a végtelen vektorrendszer is: Ha adott egy Λ „indexhalmaz”, akkor $\{\mathbf{v}_\lambda \mid \lambda \in \Lambda\}$ vektorrendszer, ha nem vagyunk tekintettel a benne szereplő elemek sorrendjére. ■

A vektorterek és a modulusok esetében igen fontos fogalmat vezetünk be:

1.2. Definíció. Az $\mathbf{u}_1, \dots, \mathbf{u}_n$ vektoroknak a c_1, \dots, c_n skalárokkal képezett lineáris kombinációján a

$$c_1 \cdot \mathbf{u}_1 + \dots + c_n \cdot \mathbf{u}_n$$

kifejezést értjük. Ha

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

akkor azt mondjuk, hogy \mathbf{v} lineáris kombinációja az $\mathbf{u}_1, \dots, \mathbf{u}_n$ vektoroknak. Az üres halmaz lineáris kombinációja egyedül a nullvektor.

A lineáris kombinációban fellépő skalárokat e lineáris kombináció együtthatóinak nevezzük.

A csupa 0 együtthatókkal képezett lineáris kombináció neve triviális lineáris kombináció. Az összes többi nemtriviális lineáris kombináció. ■

Megjegyzések

1. A fenti összeg egyértelmű az összeadás asszociativitása és kommutativitása miatt.
2. Az üres halmaz lineáris kombinációjára vonatkozó feltétel annak felel meg, hogy az üres összeget is nullának tekintjük.

3. Természetesen egy rögzített lineáris kombináció esetén az együtthatók mindig a megfelelő „sorszámú” vektorhoz vannak hozzárendelve; azaz a $c\mathbf{u} + d\mathbf{v}$, valamint a $c\mathbf{v} + d\mathbf{u}$ lineáris kombinációk különböznek. Ez akkor is igaz, ha véletlenül $\mathbf{u} = \mathbf{v}$.

4. Noha a lineáris kombináció csak egy formális kifejezés, mégis sokszor azonosnak fogjuk tekinteni azzal a vektorral, amit ez a kifejezés előállít.

5. Látható, hogy a „trivialitás” csupán a skalárokon múlik. □

Vektorok lineáris kombinációit véve, lehet, hogy bizonyos együtthatók nullával egyenlők. Ezt felhasználva értelmessé tehető végtelen sok vektor lineáris kombinációja:

1.3. Definíció. Végtelen sok vektor lineáris kombinációin azokat a lineáris kombinációkat értjük, amelyekben majdnem minden együttható 0. („Majdnem minden” azt jelenti, hogy véges sok kivétellel.) ■

A vektorterek vizsgálatában alapvető az alábbi:

1.5. Tétel (előállítási tétel). *Lineáris kombinációk lineáris kombinációja az eredeti vektorok lineáris kombinációja.*

Bizonyítás. Mivel minden lineáris kombináció előáll vektorok skalárszorosainak az összegeként, ezért (az asszociativitás miatt) elegendő azt belátni, hogy lineáris kombinációk skalárszorosa is és összege is lineáris kombináció. Legyenek az adott vektorok $\mathbf{u}_1, \dots, \mathbf{u}_n$, és ezeknek két lineáris kombinációja:

$$\mathbf{v} = a_1 \cdot \mathbf{u}_1 + \dots + a_n \cdot \mathbf{u}_n \quad \text{és} \quad \mathbf{w} = b_1 \cdot \mathbf{u}_1 + \dots + b_n \cdot \mathbf{u}_n.$$

Ekkor, tetszőleges $c \in R$ esetén:

$$c \cdot \mathbf{v} = (c \cdot a_1) \cdot \mathbf{u}_1 + \dots + (c \cdot a_n) \cdot \mathbf{u}_n,$$

valamint

$$\mathbf{v} + \mathbf{w} = (a_1 + b_1) \cdot \mathbf{u}_1 + \dots + (a_n + b_n) \cdot \mathbf{u}_n$$

bizonyítja az állítást. ■

1.4. Definíció. Ha az \mathcal{U} vektortér egy \mathbf{v} eleme felírható egy \mathcal{U} -beli \mathbf{U} vektorrendszer elemeinek a lineáris kombinációjaként, akkor azt mondjuk, hogy \mathbf{v} (lineárisan) függ az \mathbf{U} vektorrendszertől. Ha az \mathcal{U} vektortér \mathbf{V} vektorrendszerének minden eleme lineárisan függ az \mathbf{U} vektorrendszertől, akkor azt mondjuk, hogy \mathbf{V} lineárisan függ az \mathbf{U} -tól.

Ha a V vektorrendszer függ az U vektorrendszertől és az U vektorrendszer is függ a V vektorrendszertől, akkor azt mondjuk, hogy a két vektorrendszer ekvivalens. ■

A lineáris függés három alapvető tulajdonságát mondja ki az alábbi

1.6. Tétel (Függési alaptétel).

A) Minden vektorrendszer lineárisan függ önmagától, egy vektorrendszer minden részrendszere függ az eredeti vektorrendszertől (reflexivitás).

B) Ha a V vektorrendszer függ az U vektorrendszertől és a W vektorrendszer függ a V vektorrendszertől, akkor a W vektorrendszer függ az U vektorrendszertől (transzitivitás).

C) Ha a v vektor nem függ az U vektorrendszertől, de függ az $U \cup \{u\}$ vektorrendszertől, akkor az u vektor függ az $U \cup \{v\}$ vektorrendszertől.

Bizonyítás. Ha $u \in U$, akkor az $u = 1 \cdot u$ felírás adja, hogy u függ az U vektorrendszer-től (a többi együtthatót 0-nak választottuk). Ez azt bizonyítja, hogy egy vektorrendszertől minden részrendszere függ, így önmaga is.

A B) pont azonnal következik az előállítási tételből az A) pont alapján.

Tekintsünk végül egy, a feltételeket kielégítő U vektorrendszert, valamint egy u és v vektort. A lineáris függés alapján léteznek olyan c_0, c_1, \dots, c_n skalárok és olyan $u_1, \dots, u_n \in U$ vektorok, amelyekre $v = c_0 u + c_1 u_1 + \dots + c_n u_n$ teljesül. Itt $c_0 \neq 0$, mert különben v függene az U vektorrendszertől. Ezért léteznek a $b_i = -\frac{c_i}{c_0}$ skalárok, és így

$u = \frac{1}{c_0} v + b_1 u_1 + \dots + b_n u_n$ ad egy megfelelő lineáris kombinációt. ■

Megjegyezzük, hogy a harmadik pont gyűrűk esetében nem igaz, hiszen gyűrűben nem minden osztás végezhető el.

Felhívjuk a figyelmet arra, hogy a továbbiakban a lineáris függés esetében csak a fent bizonyított három tulajdonságot fogjuk használni. Ez azt jelenti, hogy ezek a tulajdonságok tekinthetők a függés axiomatikus definíciójának. Ezzel a témával az úgynevezett matroid-elmélet foglalkozik.

Feladatok

1. Legyenek u, v egy vektorrendszer elemei. Készítsünk egy új vektorrendszert úgy, hogy u helyébe az $u + cv$ vektort tesszük. Bizonyítsuk be, hogy a két vektorrendszerből képezhető lineáris kombinációk halmaza megegyezik. Általában miképpen cserélhetjük ki egy vektorrendszer valamely elemét, ha azt akarjuk, hogy a két vektorrendszerből képezhető lineáris kombinációk halmaza megegyezzen?

2. Tekintsük az egész számok \mathbb{Z} gyűrűjét mint önmaga feletti modulust. Mutassuk meg, hogy erre nem teljesül a fenti C) pont.

3. Bizonyítsuk be, hogy ha egy (kommutatív) gyűrű feletti bármely modulusban érvényes a C) pont, akkor ez a gyűrű test.

4. Legyen R egy euklideszi gyűrű, \mathcal{M} egy R -modulus és \mathbf{u}, \mathbf{v} elemei az \mathcal{M} egy \mathbf{U} vektorrendszerének. Cseréljük ki a vektorrendszer \mathbf{u} elemét $c\mathbf{u} + d\mathbf{v}$ -vel. Mi a feltétele annak, hogy a kapott \mathbf{V} vektorrendszerből képezhető lineáris kombinációk halmaza megegyezzen az \mathbf{U} vektorrendszerből képezhető lineáris kombinációk halmazával?

5. Bizonyítsuk be, hogy a \mathbb{Q} racionális számtest modulus \mathbb{Z} felett (azaz \mathbb{Q} -t mint \mathbb{Z} - \mathbb{Q} -t tekintjük). Mutassuk meg, hogy \mathbb{Q} -ban nincs olyan véges „vektorrendszer”, amelyből képezett vektorok halmaza \mathbb{Q} minden elemét előállítaná. Adjunk meg ilyen tulajdonságú végtelen vektorrendszert.

6. Legyen \mathcal{M} részmodulusa $\mathbb{Z}\mathbb{Q}$ -nak. Bizonyítsuk be, hogy pontosan akkor létezik \mathcal{M} -beli elemeknek olyan véges rendszere, amelyeknek a lineáris kombinációiként \mathcal{M} minden eleme előáll, ha az \mathcal{M} -beli törtek nevezői egy korlát alatt maradnak.

7. Tekintsünk egy R egységelemes gyűrű feletti azonos alakú mátrixokat. Tudjuk, hogy ezek modulusot alkotnak R felett. Tekintsük ezek közül azokat, amelyekben egyetlen helyen szerepel egy 1-es, a többi helyen pedig 0 áll. Bizonyítsuk be, hogy ezek lineáris kombinációjaként minden szóban forgó mátrix előállítható.

3. Lineáris összefüggés és függetlenség

Az alábbiakban a lineáris függéshez hasonló, ugyancsak fontos fogalmat vezetünk be.

1.5. Definíció. Vektorok egy $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ rendszere lineárisan összefüggő, ha van olyan nemtriviális lineáris kombinációjuk, amelyre $\mathbf{v} = c_1 \cdot \mathbf{u}_1 + \dots + c_n \cdot \mathbf{u}_n = \mathbf{o}$. ■

Valójában nem a lineáris összefüggés az alapvető fogalom, hanem ennek az ellenkezője:

1.6. Definíció. Vektorok egy $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ rendszere lineárisan független, ha nem lineárisan összefüggő; azaz, ha egy $\mathbf{v} = c_1 \cdot \mathbf{u}_1 + \dots + c_n \cdot \mathbf{u}_n$ lineáris kombinációjuk csak akkor lehet \mathbf{o} , ha ez a triviális lineáris kombináció, azaz $c_1 = \dots = c_n = 0$. ■

Megjegyzés. Modulusok esetében hasznosabb a lineáris függetlenséget a vektorokkal is összekötve úgy definiálni, hogy a $c_1\mathbf{u}_1 = \dots = c_n\mathbf{u}_n = \mathbf{o}$ egyenlőségek teljesülését kívánjuk meg, de feltesszük, hogy az $\mathbf{u}_1, \dots, \mathbf{u}_n$ vektorok egyike sem \mathbf{o} . Testek esetében — mint láttuk — a két fogalom megegyezik. □

1.7. Tétel. Ha \mathbf{U} lineárisan összefüggő, és $\mathbf{V} \supseteq \mathbf{U}$, akkor \mathbf{V} is lineárisan összefüggő.

Bizonyítás. Mivel \mathbf{U} minden eleme \mathbf{V} -ben van, ezért az az \mathbf{U} -beli elemekkel képzett nemtriviális lineáris kombináció, amelyik \mathbf{o} -t adott, egyszersmind \mathbf{V} -belinek is tekinthető. ■

Ezt a tételt érdemes fordított megfogalmazásban is kimondani:

1.8. Tétel. Ha U lineárisan független, és $V \subseteq U$, akkor V is lineárisan független. ■

A lineáris függést és a lineáris összefüggést kapcsolja össze az

1.9. Tétel. Ha u függ V -től, akkor $\{u\} \cup V$ lineárisan összefüggő.

Bizonyítás. Feltétel szerint léteznek olyan $v_1, \dots, v_n \in V$ vektorok, és léteznek olyan c_1, \dots, c_n skalárok, amelyekre $u = c_1 v_1 + \dots + c_n v_n$. Ezt átírva azt kapjuk, hogy:

$$(-1) \cdot u + c_1 \cdot v_1 + \dots + c_n \cdot v_n = \mathbf{o};$$

ami egy nemtriviális lineáris kombináció, hiszen $-1 \neq 0$. ■

Ez a tétel általában nem fordítható meg. Vektorterek esetében is csak részleges megfordítás lehet:

1.10. Tétel. Egy vektortér minden összefüggő rendszerében van olyan vektor, amely a többitől függ.

Bizonyítás. Legyen az U vektorrendszer elemeire $c_1 u_1 + \dots + c_n u_n$ egy olyan nemtriviális lineáris kombináció, amely a \mathbf{o} -t adja. Az általánosság megszorítása nélkül feltehető, hogy $c_1 \neq 0$ (hiszen az összeg tagjait bármilyen sorrendben írhatjuk, és az indexezést is megváltoztathatjuk). Ekkor:

$$u_1 = \left(-\frac{c_2}{c_1}\right) \cdot u_2 + \dots + \left(-\frac{c_n}{c_1}\right) \cdot u_n$$

előállítja u_1 -et a többiek lineáris kombinációjaként. ■

Az előző tételnek egy „erősebb” változata az alábbi:

1.11. Tétel. Ha a független U rendszerhez hozzávéve az u vektort összefüggő rendszert kapunk, akkor u függ U -tól.

Bizonyítás. Feltétel szerint vannak olyan c, c_1, \dots, c_n skalárok, amelyeknek nem mindegyike 0, és U -beli u_1, \dots, u_n vektorok, amelyekre $cu + c_1 u_1 + \dots + c_n u_n = \mathbf{o}$. Az U függetlensége miatt $c = 0$ lehetetlen. Most már a bizonyítás az előző tételéhez hasonlóan fejezhető be. ■

Megjegyzés. Moduluszok esetében a fenti két tétel egyike sem igaz. Tekintsük például \mathbb{Z} -t mint önmaga feletti modulust. Itt $2 \cdot (3) + 3 \cdot (-2) = (0)$, azaz (3) és (-2) lineárisan összefüggenek, de világos, hogy egyik sem lineárisan függő a másiktól. □

A vektorterekben a lineáris függetlenségre vonatkozóan alapvető fontosságú az

1.12. Kicserélési tétel. Ha az U lineárisan független rendszer függ a V vektorrendszertől, akkor minden $u \in U$ -hoz van olyan $v \in V$, hogy $W = (U \setminus \{u\}) \cup \{v\}$ is lineárisan független.

Bizonyítás. Tekintsük az $U' = U \setminus \{u\}$ rendszert. Ez, mint egy lineárisan független rendszer része, ismét lineárisan független. Ha az $U' \cup \{v\}$ rendszer minden $v \in V$ esetén

összefüggő volna, akkor az 1.11. Tétel következtében U' és V ekvivalensek lennének. A lineáris függés tranzitivitása szerint ekkor u függene az U' rendszertől; ami ellentmond U függetlenségének. ■

1. Következmény. *Ha az U véges független vektorrendszer függ a V vektorrendszertől, akkor van olyan $W \subseteq V$ (független) vektorrendszer, amelyre U függ W -től és $|U| = |W|$.*

Bizonyítás. A lineáris függetlenség miatt U -nak csupa különböző elemei vannak: $U = \{u_1, \dots, u_k\}$. A kicserélési tétel alapján U elemeit rendre kicserélhetjük V különböző elemeire; és végül egy olyan $W \subseteq V$ független rendszert kapunk, amelynek ugyanannyi eleme van, mint U -nak. ■

2. Következmény. *Ekvivalens független rendszerek elemszáma megegyezik.* ■

Megjegyezzük, hogy elég annyit feltenni, hogy a két rendszer egyike független. Az első következmény alapján már adódik az egyenlőség.

3. Következmény. *Minden vektorrendszer tartalmaz maximális független rendszert. Egy vektorrendszer ekvivalens bármely maximális független részrendszerével. Ha két vektorrendszer ekvivalens és maximális független részrendszerük véges, akkor maximális független részrendszereik elemszáma megegyezik.*

Bizonyítás. Az első állítás véges vektorrendszerekre triviális. Végtelen esetekben csak azt tudjuk, hogy független vektorrendszerek növvő láncának egyesítése is független. *Egy hal-mazelméleti axióma szerint — amit mi elfogadunk — ebből következik, hogy van maximális független rendszer.* Legyen U a V maximális független részrendszere. Azt tudjuk, hogy U függ V -től. Legyen $v \in V$. Ha $v \in U$, akkor tudjuk, hogy v függ U -tól. Ez a függés a $v \notin U$ esetben az 1.6. Tételből következik. A további állítások nyilvánvalóak. ■

1. példa: A síkvektorok körében bármely két nem párhuzamos vektor lineárisan független rendszert alkot, de bármely három lineárisan összefüggő.

2. példa: A térvektorok körében bármely három nem egysíkú vektor lineárisan független rendszert alkot, de bármely négy lineárisan összefüggő.

3. példa: Tetszőleges vektortérben bármely két (k számú) vektor bármely három ($k+1$ darab) lineáris kombinációja lineárisan összefüggő rendszert alkot.

4. példa: A K test feletti polinomok K feletti vektorterében különböző fokú polinomokból álló rendszer mindig független.

5. példa: A $(0, 2\pi)$ intervallumban értelmezett függvények között független rendszert alkot $\{1, \sin(x), \cos(x), \sin(2x), \cos(2x)\}$, de $\{1, \sin(x), \cos(x), \sin^2(x), \cos^2(x)\}$ nem.

Feladatok

1. Bizonyítsuk be, hogy egy test feletti n hosszúságú sormátrixok körében azok a mátrixok, amelyekben egyetlen (de különböző) helyen 1 áll, és minden más helyen 0, lineárisan függetlenek, de bármely más mátrix ezektől lineárisan függ. Mutassuk meg, hogy ebben a vektortérben bármely $n + 1$ mátrix lineárisan összefüggő.

2. Bizonyítsuk be, hogy \mathbb{Z} -ban bármely két vektor lineárisan összefüggő.

3. Legyen R egy egységelemes integritási tartomány. Adjunk az elemek oszthatóságával megfogalmazható feltételt arra, hogy R felett bármely modulusban minden összefüggő rendszerben legyen olyan vektor, amely a többitől függ. Mutassuk meg, hogy a racionális számtestben végtelen sok ilyen részgyűrű van. Adjuk meg mindet.

4. Tegyük fel, hogy egy R gyűrű feletti tetszőleges modulusban bármely független vektorrendszerhez újat véve ez a vektor függ az eredeti rendszertől. Bizonyítsuk be, hogy ekkor ez a modulus vektortér.

5. Legyen K egy (végtelen) test. Mutassuk meg, hogy a $K[x]$ polinomgyűrű természetes módon modulus $K[x]$ felett. Mutassuk meg, hogy van olyan vektor, amelytől minden vektor lineárisan függ. Tekintsük most csak azokat a polinomokat, amelyeknek a konstans tagja 0. Bizonyítsuk be, hogy az eredeti műveletekre nézve ezek is modulust alkotnak $K[x]$ felett. Bizonyítsuk be, hogy ebben a modulusban bármely két vektor lineárisan összefüggő, de nincs olyan véges vektorrendszer, amelytől minden vektor függene.

4. Generátorrendszer és bázis

1.7. Definíció. U generátorrendszere az \mathcal{U} vektortérnek, ha a vektortér minden eleme előáll az U elemeinek lineáris kombinációjaként. ■

A generátorrendszer bizonyos értelemben fordítottan viselkedik, mint a független rendszer:

1.13. Tétel. *Egy vektortér valamely generátorrendszerét tartalmazó bármely vektorrendszer is generátorrendszer.*

Bizonyítás. Az állítás triviálisan igaz, hiszen már az eredeti vektorrendszer elemeinek lineáris kombinációi is előállítják a vektortér minden elemét. ■

Az 1.12. Tétel 3. Következményéből kapjuk:

1.14. Tétel. *Ha egy vektortérnek van véges generátorrendszere, akkor minden generátorrendszerének legalább annyi eleme van, mint akármelyik lineárisan független rendszerének.* ■

1.8. Definíció. Lineárisan független generátorrendszert bázisnak nevezünk. ■

Megjegyzés. Modulusok esetén a lineáris függetlenséget természetesen az 1.6. Definíció utáni megjegyzés szerint vesszük. □

Ugyancsak a 3. Következményből adódik:

1.15. Tétel. *Ha egy vektortérnek van n -elemű bázisa, akkor minden generátorrendszerének legalább n , minden független rendszerének legfeljebb n és minden bázisának pontosan n eleme van.* ■

1.9. Definíció. Ha egy vektortérnek van n -elemű bázisa, akkor a vektorteret n -dimenziósnek nevezzük. A \mathcal{V} vektortér dimenzióját $\dim(\mathcal{V})$ jelöli.

Ha a vektortérben nincs véges bázis, akkor végtelen dimenziós vektorterről beszélünk. ■

Könnyen belátható:

1.16. Tétel. *Egy vektortérben ekvivalensek az alábbiak:*

- (1) *A vektortérben van n -elemű független rendszer, de nincs n -nél több elemű.*
- (2) *A vektortérben van n -elemű független rendszer, de nincs $(n + 1)$ -elemű.*
- (3) *A vektortérben van n -elemű generátorrendszer, de nincs n -nél kevesebb elemű.*
- (4) *A vektortérben van n -elemű generátorrendszer, de nincs $(n - 1)$ -elemű.*
- (5) *A vektortérben van n -elemű független rendszer és van n -elemű generátorrendszer.*
- (6) *A vektortér n -dimenziós.* ■

A fentiekből azonnal következik például az, hogy ha egy vektortérben nincs két független elem, akkor van egyelemű generátorrendszere. Ez modulusokra nincs így. Tekintsük például \mathbb{Q} -t mint \mathbb{Z} -modulust. Itt bármely két elem lineárisan összefüggő, de egyáltalában nincs véges generátorrendszere.

Megjegyzés. Az 1.16. Tételben felsorolt hat tulajdonság „elvben” nem ekvivalens. (1)-ből (2) és (3)-ból (4) logikailag is következik. A megfordítás „elve” nem biztos, hogy igaz; hiszen ha nem tudjuk, mit jelentenek a fenti fogalmak, elképzelhető, hogy van n -elemű és $(n + 2)$ -elemű független rendszer, de $(n + 1)$ -elemű nincs. Ugyancsak világos, hogy (6)-ból következik (5), mert ha van egy olyan n -elemű rendszer, amely független és egyben generátorrendszer is, akkor ez a rendszer kielégíti az (5)-ben megkívántakat. A megfordítás itt sem nyilvánvaló, hiszen lehetne n -elemű független rendszer is, n -elemű generátorrendszer is, de elképzelhető, hogy ezek nem egyeznek meg.

Ennek megfelelően szokásos a dimenziót a lehető „leggyengébb” módon definiálni:

Egy vektortér dimenziója a lineárisan független elemek maximális száma.

Mi itt azért választottuk a „legerősebb” változatot, mert ez mutat rá a lényegre. □

1.17. Tétel. *Egy vektortér minden független rendszere része egy bázisnak. Egy vektortér minden generátorrendszere tartalmaz bázist. Ha az \mathcal{U} vektortérben \mathbf{L} lineárisan független rendszer, \mathbf{G} generátorrendszer és $\mathbf{L} \subseteq \mathbf{G}$, akkor van a térben olyan \mathbf{B} bázis, amelyre $\mathbf{L} \subseteq \mathbf{B} \subseteq \mathbf{G}$.*

Bizonyítás. Ha a harmadik állításban \mathbf{G} -t az egész térnek vesszük, akkor az első állítás adódik. Ha pedig \mathbf{L} -t választjuk az üres rendszernek, akkor a második állítást kapjuk. Elég tehát a harmadik állítást bizonyítani.

Tekintsük az \mathbf{L} -et tartalmazó és \mathbf{G} -ben benne levő független rendszereket. Ha az \mathcal{U} vektortér véges dimenziós, akkor ezek mindegyikének az elemszáma legfeljebb akkora, mint a tér dimenziója, azaz van közöttük maximális elemszámú. Végtelen dimenziós vektorterekre ez a már szóba került halmazelméleti axiómának a következménye.

Legyen tehát $\mathbf{L} \subseteq \mathbf{B} \subseteq \mathbf{G}$ olyan, hogy \mathbf{B} lineárisan független, de minden nála bővebb \mathbf{G} -beli elemekből álló rendszer lineárisan összefüggő. \mathbf{B} feltétel szerint lineárisan független; azt kell bizonyítani, hogy generátorrendszer. Ha $\mathbf{u} \in \mathbf{G}$, akkor \mathbf{B} maximalitása miatt \mathbf{u} függ \mathbf{B} -től; így \mathbf{G} függ \mathbf{B} -től. Mivel \mathbf{G} generátorrendszer, a függés tranzitivitásának következtében \mathbf{B} is az. ■

Megjegyzések

1. Lehet, hogy a vektortér *triviális*, azaz egyetlen eleme a nullvektor (könnyen látható, hogy ez valóban vektortér). Ebben az esetben az üres halmaz egy maximális független rendszer; ez a vektortér *nulldimenziós*. Mivel minden vektortérben létezik maximális független rendszer, amely bázis, ezért végtelen dimenziós vektortérben létezik végtelen bázis. A végtelen dimenziós vektorterek dimenzióit nem különböztetjük meg, ha csak a véges dimenziós vektorterekkel foglalkozunk. Ha azonban végtelen dimenziós vektortereket is vizsgálunk, akkor a vektortér dimenzióján egy bázisának a számosságát értjük. Bebizonyítható, hogy ez a bázistól független.

2. Modulusok esetében is lehet a modulus dimenziójáról beszélni, de ez nem egyértelmű. Ugyanis nem minden modulusnak van bázisa; és a bázisok elemszáma sem egyértelmű. Ilyen esetben célszerű a fent említett definíciót használni: dimenziója a lineárisan független elemek maximális száma. □

A későbbiekben igen részletesen fogunk foglalkozni vektorterek művelettartó leképezéseivel (homomorfizmusokkal). Most csak egy igen fontos ilyen tulajdonságú leképezést nézünk meg:

1.10. Definíció. A K test feletti \mathcal{V} vektorteret a K test feletti \mathcal{U} vektortérbe vivő $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ bijekciót izomorfizmusnak nevezzük, ha művelettartó, azaz tetszőleges $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ vektorokra és $c \in K$ elemekre $\varphi(\mathbf{v}_1 + \mathbf{v}_2) = \varphi(\mathbf{v}_1) + \varphi(\mathbf{v}_2)$ és $\varphi(c\mathbf{v}_1) = c\varphi(\mathbf{v}_1)$.

A K test feletti \mathcal{V} vektorteret a K test feletti \mathcal{U} vektortérrel izomorfnek nevezzük, ha létezik egy $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ izomorfizmus. Ezt a kapcsolatot $\mathcal{V} \cong \mathcal{U}$ vagy $\mathcal{V} \cong_K \mathcal{U}$ jelöli. ■

Megjegyzés. Könnyen látható, hogy a K test feletti \mathcal{V} vektorteret a K test feletti \mathcal{U} vektortérbe vivő $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ leképezés pontosan akkor művelettartó, ha tetszőleges $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ vektorokra és $c_1, c_2 \in K$ elemekre $\varphi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\varphi(\mathbf{v}_1) + c_2\varphi(\mathbf{v}_2)$.

Ebből a feltételből a $c_1 = c_2 = 1$, illetve a $c_1 = c$ és $c_2 = 0$ választással azonnal adódik a művelettartás. Másrészt viszont a művelettartásból kapjuk, hogy:

$$\varphi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = \varphi(c_1\mathbf{v}_1) + \varphi(c_2\mathbf{v}_2) = c_1\varphi(\mathbf{v}_1) + c_2\varphi(\mathbf{v}_2).$$

□

1.18. Tétel. *A K test feletti vektorterek izomorfizmusa ekvivalenciareláció (reflexív, szimmetrikus és tranzitív). Ennek megfelelően beszélhetünk izomorf vektorterekről.*

A $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ izomorfizmus generátorrendszert generátorrendszerbe, lineárisan független rendszert lineárisan független rendszerbe visz.

Két vektortér pontosan akkor izomorf, ha dimenzióik megegyeznek.

Bizonyítás. Legyen \mathcal{V} vektortér a K felett és definiáljuk a $\varphi : \mathcal{V} \rightarrow \mathcal{V}$ leképezést tetszőleges $\mathbf{v} \in \mathcal{V}$ vektorra a $\varphi(\mathbf{v}) = \mathbf{v}$ összefüggéssel. Ez triviálisan izomorfizmus, tehát \mathcal{V} izomorf önmagával.

Legyen $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ a K test feletti vektorterek egy izomorfizmusa. Mivel φ bijekció, ezért létezik $\psi : \mathcal{U} \rightarrow \mathcal{V}$ inverze. Ez azt jelenti, hogy bármely $\mathbf{v} \in \mathcal{V}$ vektorra $\psi\varphi(\mathbf{v}) = \mathbf{v}$ és bármely $\mathbf{u} \in \mathcal{U}$ vektorra $\varphi\psi(\mathbf{u}) = \mathbf{u}$. Azt kell megmutatni, hogy ha φ művelettartó, akkor ψ is az. Legyen $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}$ és $c_1, c_2 \in K$. Mivel φ bijektív, ezért van olyan $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$, amelyekre $\varphi(\mathbf{v}_1) = \mathbf{u}_1$ és $\varphi(\mathbf{v}_2) = \mathbf{u}_2$ teljesül. Mivel φ művelettartó, ezért $\varphi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$, amiből $\psi(c_1\mathbf{u}_1 + c_2\mathbf{u}_2) = \psi(\varphi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\psi(\mathbf{u}_1) + c_2\psi(\mathbf{u}_2)$ következik. Az izomorfizmus tehát szimmetrikus.

Ha $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ és $\psi : \mathcal{U} \rightarrow \mathcal{W}$ a K test feletti vektorterek egy-egy izomorfizmusa, akkor $\psi\varphi$ triviálisan bijektív és művelettartó; ami a tranzitivitást bizonyítja.

Legyen most $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ a K test feletti vektorterek egy izomorfizmusa. Tegyük fel először, hogy $\mathbf{v}_1, \mathbf{v}_2, \dots$ a \mathcal{V} -nek egy generátorrendszere, és legyen $\mathbf{u} \in \mathcal{U}$. A bijektivitás alapján van olyan $\mathbf{v} \in \mathcal{V}$, hogy $\mathbf{u} = \varphi(\mathbf{v})$. A generátorrendszer definíciója szerint létezik egy $\mathbf{v} = \sum_i c_i \mathbf{v}_i$ felírás (az összeg véges). Ebből azonnal kapjuk, hogy $\mathbf{u} = \varphi(\mathbf{v}) = \sum_i c_i \varphi(\mathbf{v}_i)$; azaz $\varphi(\mathbf{v}_1), \varphi(\mathbf{v}_2), \dots$ valóban generátorrendszere \mathcal{U} -nak.

Amennyiben $\mathbf{v}_1, \mathbf{v}_2, \dots$ a \mathcal{V} -nek egy lineárisan független rendszere és $\mathbf{u}_i = \varphi(\mathbf{v}_i)$, akkor tekintsünk egy $\sum_i c_i \mathbf{u}_i = \mathbf{o}$ lineáris kombinációt (itt \mathbf{o} az \mathcal{U} vektortér nulleleme).

Tekintettel arra, hogy tetszőleges $\mathbf{u} \in \mathcal{U}$ vektorra $\mathbf{o} + \mathbf{u} = \mathbf{u}$, ezért $\varphi(\mathbf{o}) + \varphi(\mathbf{u}) = \varphi(\mathbf{u})$, hiszen φ művelettartó. φ szürjektivitásából következik, hogy \mathcal{V} tetszőleges eleme $\varphi(\mathbf{u})$ alakú; így csak $\varphi(\mathbf{o}) = \mathbf{o} \in \mathcal{V}$ lehetséges, mert a kimutatott tulajdonsággal csak a \mathcal{V} nullvektora rendelkezik. Mivel φ izomorfizmus, ezért létezik egy ψ inverze. Ekkor $\psi(\mathbf{o}) = \mathbf{o} \in \mathcal{V}$,

amiből $\sum_i c_i \mathbf{v}_i = \varphi\left(\sum_i c_i \mathbf{u}_i\right) = \varphi(\mathbf{o}) = \mathbf{o} \in \mathcal{V}$ következik. Tekintettel arra, hogy az eredeti rendszer lineárisan független, ezért ez csak úgy lehet, hogy minden i -re $c_i = 0$; ami a $\mathbf{u}_1, \mathbf{u}_2, \dots$ vektorrendszer függetlenségét jelenti.

Ezek szerint izomorfizmusnál egy lineárisan független generátorrendszer képe is az; ami a bázis és a dimenzió definíciójának a következtében pontosan azt jelenti, hogy izomorf vektorterek dimenziója megegyezik.

Fordítva, tegyük fel, hogy az \mathcal{U} és a \mathcal{V} vektorterek izomorfak; azaz létezik ugyanannyi elemből álló $\{\mathbf{u}_i \mid i \in I\}$ és $\{\mathbf{v}_i \mid i \in I\}$ bázisuk, ahol az I indexhalmaz mindkét

esetben ugyanaz. Ekkor a $\varphi : \sum_i c_i \mathbf{u}_i \mapsto \sum_i c_i \mathbf{v}_i$ leképezésről könnyen kimutatható, hogy izomorfizmus. ■

Megjegyzések

1. A fenti definíció és bizonyítás szóról szóra átvihető modulusokra; kivéve az 1.18. Tétel legutolsó állítását, hiszen egy modulusnak nem feltétlenül van bázisa.

2. Izomorf vektorterek „belső szerkezetüket tekintve” azonosoknak vehetők. Ez azt jelenti, hogy ha egy vektortérben valamit definiálhatunk vagy bizonyítunk, akkor hasonló fogalom definiálható és megfelelő tétel bizonyítható a vele izomorf térben is. Ennek ellenére a két vektorteret nem szabad megegyezőnek tekinteni. □

Feladatok

1. Legyen k az $m \in \mathbb{N}$ számban fellépő különböző prímtényezők száma. Bizonyítsuk be, hogy \mathbb{Z}_m -ben mint \mathbb{Z} modulusban minden $1 \leq r \leq k$ természetes számra létezik r -elemű bázis.

2. Bizonyítsuk be, hogy ha egy R -modulusban van egyelemű és n -elemű bázis, akkor minden $1 \leq r \leq n$ esetén van r -elemű bázis.

3. Bizonyítsuk be, hogy ha egy R -modulusban létezik egyelemű bázis, akkor nem létezik végtelen bázis.

4. Bizonyítsuk be, hogy ha egy R -modulusnak van véges bázisa, akkor nincs végtelen.

5. Mutassuk meg, hogy léteznek olyan \mathbb{Z} -modulusok, amelyek egyetlen elemmel generálhatóak, de nem izomorfak.

6. Mutassuk meg, hogy léteznek olyan \mathbb{Z} -modulusok, amelyek mindegyikében bármely két elem lineárisan összefüggő, de nem izomorfak.

7. Hány különböző bázisa van a kételemű \mathbb{Q}_2 test feletti n -dimenziós \mathcal{V} vektortérnek?

8. Hány különböző bázisa van a p elemű \mathbb{Q}_p test feletti n -dimenziós \mathcal{V} vektortérnek (p prímszám)?

MÁSODIK FEJEZET

VEKTORTÉR-KONSTRUKCIÓK

1. Alterek, lineáris alakzatok

A tér geometriai vizsgálatánál igen fontos szerepet töltenek be a síkok és az egyenesek. Ezeket közösen lineáris alakzatoknak vagy lineáris részhalmazoknak nevezik. A lineáris alakzatok a lineáris algebrai tárgyalásoknál is igen fontos szerepet töltenek be. Mindekenélőtt azokkal a lineáris alakzatokkal foglalkozunk, amelyek „átmennek” az origón. Ezek hasonlóképpen jellemezhetők, mint a gyűrűk részgyűrűi. Természetesen itt is utalunk a megfelelő moduluselméleti fogalmakra. A lineáris alakzatok és az alterek olyasmi kapcsolatban vannak egymással, mint az $a \cdot x + b$ és az $a \cdot x$ alakú polinomok.

2.1. Definíció. Egy \mathcal{V} , K test feletti vektortér elemeinek egy \mathcal{U} halmazát a \mathcal{V} alterének nevezzük, ha \mathcal{U} vektortér a \mathcal{V} -ben értelmezett műveletekre. Ezt $\mathcal{U} \leq \mathcal{V}$ jelöli. Ha ki akarjuk hangsúlyozni, hogy K feletti altérrel van szó, akkor az $\mathcal{U} \leq_K \mathcal{V}$ jelölést használjuk.

Egy \mathcal{M} R -modulus elemeinek egy \mathcal{N} halmazát az \mathcal{M} részmodulusának nevezzük, ha \mathcal{N} R -modulus az \mathcal{M} -ben értelmezett műveletekre. Ezt $\mathcal{N} \leq \mathcal{M}$ jelöli. Ha ki akarjuk hangsúlyozni, hogy R feletti részmodulusról van szó, akkor az $\mathcal{N} \leq_R \mathcal{M}$ jelölést használjuk. ■

Megjegyzés. Tulajdonképpen ez a definíció kissé pongyola. Hiszen az \mathcal{U} halmaz akkor válik vektortérrel, ha a megfelelő műveleteket rajta értelmezzük. A \mathcal{V} -beli műveletek viszont az egész \mathcal{V} -n vannak értelmezve. A definíció úgy volna pontos, ha azt mondanánk, hogy a \mathcal{V} -beli műveleteket \mathcal{U} -ra megszorítva, ezekre a műveletekre nézve válik \mathcal{U} vektortérrel. □

Természetesen igen hosszadalmas volna mindig végignézni az összes axióma teljesülését. Annál is inkább, mert eleve nem tudhatjuk, hogy a „kerülő úton” definiált elemek (nullvektor, különbség) az altérben ugyanazok-e, mint az eredeti vektortérben. (Majd az algebrai struktúrák tárgyalásánál látni fogjuk, hogy az ilyen irányú eredmények nem szükségképpen igazak.) Éppen ezért mindenekelőtt megmutatjuk, hogy a fenti módon definiált elemek az altérben is ugyanazok; majd bebizonyítunk egy olyan tételt, amely megadja, mit érdemes megnézni annak eldöntésére, hogy a vektortér egy részhalmaza altér-e.

2.1. Tétel. *Altérben (vagy modulusban) a nullvektor, illetve az ellentett megegyezik az eredeti vektortérben tekintett nullvektorral, illetve ellentettel.*

Bizonyítás. Legyen \mathbf{o} az eredeti vektortérbeli és \mathbf{o}' az altérbeli nullvektor. Ekkor az 1.2. Tételbeli bizonyítást szóról szóra alkalmazva kapjuk, hogy $\mathbf{o}' = \mathbf{o}$. Ugyanebben a tételben a másik állítás bizonyítása szó szerint megadja, hogy az altér egy \mathbf{u} vektorának altérbeli \mathbf{w} és az eredeti térbeli \mathbf{v} inverze megegyezik. ■

2.2. Tétel. *A K feletti \mathcal{V} vektortér egy U részhalmaza pontosan akkor altér, ha az alábbi három feltétel bármelyike teljesül:*

- (1) *Nem üres és zárt a \mathcal{V} -beli összeadásra és a K elemeivel való szorzásra.*
 - (2) *Nem üres és U bármely két elemének minden lineáris kombinációja is benne van U -ban.*
 - (3) *Nem üres és U elemeinek bármely lineáris kombinációja is benne van U -ban.*
- Ezek a feltételek modulusok részmodulusaira is érvényesek.*

Bizonyítás. A feltételek szükségessége világos.

A megfordításhoz mindenekelőtt belátjuk, hogy a három feltétel valóban ekvivalens. Ha (1) teljesül, akkor tekintsük az $\mathbf{u}_1, \dots, \mathbf{u}_r \in U$ vektorokat és a $c_1, \dots, c_r \in K$ skalárokat. Ekkor a feltétel szerint $c_1\mathbf{u}_1, \dots, c_r\mathbf{u}_r \in U$ is igaz, amiből a feltétel alapján $c_1\mathbf{u}_1 + \dots + c_r\mathbf{u}_r \in U$ következik; tehát (3) is igaz. (2) a (3) feltételnek speciális esete, így következik (3)-ból. Ha (2) igaz, akkor $\mathbf{u} + \mathbf{v} = 1 \cdot \mathbf{u} + 1 \cdot \mathbf{v}$ és $c\mathbf{u} = c \cdot \mathbf{u} + 0 \cdot \mathbf{u}$ alapján adódik (1).

Tegyük most fel, hogy a kirótt feltételek teljesülnek. (1)-ből következik, hogy U -ban mind az összeadás, mind a skalárral való szorzás értelmezett. A I. (1) és I. (2), valamint a II. alatti axiómák mind teljesülnek, hiszen ezek az azonosságok \mathcal{V} -ben is igazak. Azt kell még belátni, hogy létezik nullvektor és minden vektornak létezik ellentettje. A 2.1. Tételre való tekintettel elég belátni, hogy $\mathbf{o} \in U$ és $\mathbf{u} \in U$ esetén $-\mathbf{u} \in U$ is igaz. Mivel U nem üres, ezért létezik egy $\mathbf{a} \in U$ vektor, és (1) miatt $\mathbf{o} = 0 \cdot \mathbf{a} \in U$. Ha $\mathbf{u} \in U$, akkor (2) alapján $-\mathbf{u} = (-1) \cdot \mathbf{u} \in U$. ■

2.3. Tétel. *A \mathcal{V} vektortér $\{\mathcal{U}_\lambda \mid \lambda \in \Lambda\}$ altereinek $\bigcap \{\mathcal{U}_\lambda \mid \lambda \in \Lambda\}$ közös része is altér. Ha a Λ indexhalmaz véges: $\Lambda = \{1, \dots, r\}$, akkor használatos az $\mathcal{U}_1 \cap \dots \cap \mathcal{U}_r$ jelölés is. (Esetenként a $\bigwedge \{\mathcal{U}_\lambda \mid \lambda \in \Lambda\}$, illetve $\mathcal{U}_1 \wedge \dots \wedge \mathcal{U}_r$ jelölést is használjuk.)*

Minden \mathcal{V} vektortérben van egy legkisebb altér, amely egyedül a nullvektorból áll. Minden vektortérben van egy legnagyobb altér; maga az adott vektortér. Ezeket az altereket triviális altereknek nevezzük; az összes többi altér neve valódi altér.

Ha U a \mathcal{V} vektortér tetszőleges részhalmaza, akkor létezik egy legkisebb altér, amelyik U -t tartalmazza. Ezt az $\langle U \rangle$ alteret az U generálta altérnek nevezzük. Ennek az altérnek az elemei éppen az U elemeinek a lineáris kombinációi. Az üres halmaz generálta $\langle \emptyset \rangle$ altér az egyedül a nullvektorból álló altér. U akkor és csak akkor generátorrendszer, ha $\langle U \rangle = \mathcal{V}$.

Vektortér alterének altere az eredeti vektortérnek is altere. Az alterekre használt \leq reláció reflexív, antiszimmetrikus és tranzitív, azaz részbenrendezés. Ha $\mathcal{A} \leq \mathcal{B}$ alterek és $\mathcal{A} \neq \mathcal{B}$, akkor az $\mathcal{A} < \mathcal{B}$ jelölést használjuk.

A \mathcal{V} vektortér $\{\mathcal{U}_\lambda \mid \lambda \in \Lambda\}$ altereinek generátumát $\bigvee \{\mathcal{U}_\lambda \mid \lambda \in \Lambda\}$, illetve, ha a Λ indexhalmaz véges: $\Lambda = \{1, \dots, r\}$, akkor $\mathcal{U}_1 \vee \dots \vee \mathcal{U}_r$ jelöli.

Bizonyítás. Legyen \mathcal{U}' az \mathcal{U}_λ alterek közös része. Mivel \mathbf{o} ezen alterek mindegyikében benne van, ezért $\mathbf{o} \in \mathcal{U}'$, tehát \mathcal{U}' nem üres. Ha $\mathbf{u}, \mathbf{v} \in \mathcal{U}'$, akkor minden egyes λ -ra $\mathbf{u}, \mathbf{v} \in \mathcal{U}_\lambda$. Mivel ezek mind alterek, ezért minden egyes λ -ra $c\mathbf{u} + d\mathbf{v} \in \mathcal{U}_\lambda$ is igaz, ahol $c, d \in K$. Ekkor viszont $c\mathbf{u} + d\mathbf{v} \in \mathcal{U}'$ is fennáll, hiszen ez a fenti altereknek a közös része. Így \mathcal{U}' valóban altér.

A legkisebb és a legnagyobb altérre vonatkozó állítások triviálisan igazak.

Legyen $\mathbf{U} \subseteq \mathcal{V}$. Eszerint van olyan altér, amely \mathbf{U} -t tartalmazza. Az összes ilyen altér \mathcal{U} közös része az előző pont miatt altér; és tartalmazza \mathbf{U} minden elemét, hiszen ezek a szereplő alterek mindegyikében benne vannak. Mivel \mathcal{U} minden \mathbf{U} -t tartalmazó altérnek része, ezért valóban a legkisebb ilyen altér. (Ha \mathbf{U} üres, akkor a generátum $\{\mathbf{o}\}$.) Mint a 2.2. Tétel bizonyításánál láttuk, $\langle \mathbf{U} \rangle$ tartalmazza a \mathbf{U} elemeiből alkotott összes lineáris kombináció \mathbf{U}' halmazát, vagyis $\mathbf{U}' \subseteq \langle \mathbf{U} \rangle$. Az 1.5. Tétel (előállítási tétel) szerint \mathbf{U}' zárt a lineáris kombinációk képzésére; a 2.2. Tétel szerint tehát altér. A generátum definíciója szerint tehát $\langle \mathbf{U} \rangle \subseteq \mathbf{U}'$. A generátum tehát valóban e lineáris kombinációkból áll. A generátorrendszer definíciója szerint tehát $\langle \mathbf{G} \rangle = \mathcal{V}$ pontosan akkor teljesül, ha \mathbf{G} generátorrendszer.

Az alterek összehasonlítására vonatkozó állítások nyilvánvalóak.

Az utolsó állítás valójában csak egy jelölés. ■

Megjegyzés. Az \mathbf{U} halmaz generálta $\langle \mathbf{U} \rangle$ alteret úgy definiáltuk mint a \mathbf{U} halmazt tartalmazó összes altér metszetét. Ezután beláttuk, hogy ez nem más, mint a \mathbf{U} halmaz elemeiből képezhető összes lineáris kombináció halmaza. Ha a generátummal foglalkozunk, akkor ez utóbbi megfogalmazásra feltétlenül szükség van. Felmerül a kérdés, hogy miért nem így definiáltuk; ekkor az adott definíciót meg sem kellett volna említeni. Erre az a válasz, hogy az adott definíció egy *foglalmi* definíció; ehhez csak arra van szükség, hogy az alterek metszete altér. Nem kell tudni azt, hogy miképpen képezzük a generátum elemeit. Éppen ezért az adott definíció teljesen általános; minden megfelelő esetben használható. □

2.4. Tétel. Legyen $\mathcal{U} \leq \mathcal{V}$. Ekkor \mathcal{U} minden bázisa kiegészíthető \mathcal{V} egy bázisává. Következésképpen $\dim(\mathcal{U}) \leq \dim(\mathcal{V})$. Ha \mathcal{U} -nak van véges bázisa, akkor $\mathcal{U} \neq \mathcal{V}$ esetén $\dim(\mathcal{U}) < \dim(\mathcal{V})$.

Bizonyítás. Az \mathcal{U} altér bármely bázisa lineárisan független; így — az 1.17. Tétel szerint — kiegészíthető \mathcal{V} egy bázisává. A dimenziókra vonatkozó első állítás ebből azonnal következik. Ha $\mathcal{U} \neq \mathcal{V}$, akkor a kiegészített bázisnak több eleme van, mint a kiindulásul vett \mathcal{U} -beli bázis. Véges esetben tehát a dimenzió valóban növekszik. ■

Megjegyzés. Az alterek segítségével világítható meg legjobban, hogy miért nem szabad az izomorf vektortereket egyenlőknek tekinteni. Ugyanis egy vektortérnek különböző alterei lehetnek izomorfak; de az „egész” térben látható, hogy különböznek. □

A „geometriai” térben az alterek olyan síkoknak, illetve egyeneseknek felelnek meg, amelyek az origón keresztülmennek. Fontosak azok a síkok, egyenesek is, amelyek nem mennek keresztül az origón. Ezeket úgy kaphatjuk meg, hogy egy origón átmenő alakzatot „párhuzamosan eltolunk”. Ezt a képet az alábbi módon fogalmazhatjuk meg:

2.2. Definíció. A \mathcal{V} vektortér egy \mathbf{A} halmazát lineáris alakzatnak nevezzük, ha létezik olyan $\mathbf{a} \in \mathcal{V}$ vektor és olyan \mathcal{U} altér, hogy $\mathbf{A} = \{\mathbf{a} + \mathbf{u} \mid \mathbf{u} \in \mathcal{U}\}$. Ebben az esetben \mathbf{A} az \mathcal{U} -nak az \mathbf{a} -val való eltoltja. ■

2.5. Tétel. Tegyük fel, hogy a K testnek kettőnél több eleme van. Ekkor az alábbi két feltétel bármelyike ekvivalens azzal, hogy \mathbf{A} lineáris alakzat:

(1) Ha $\mathbf{a}_1, \mathbf{a}_2 \in \mathbf{A}$ és $a, c_1, c_2 \in K$ elemekre $c_1 + c_2 = 1$, akkor $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 \in \mathbf{A}$.

(2) Tetszőleges $n \in \mathbb{N}$ esetén: ha $\mathbf{a}_i \in \mathbf{A}$ és $c_i \in K$ olyanok, hogy $\sum_{i=1}^n c_i = 1$, akkor

$$\sum_{i=1}^n c_i \mathbf{a}_i \in \mathbf{A}.$$

Ha a K test kételemű, akkor (1) helyébe az alábbi feltétel lép:

(1') Ha $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{A}$, akkor $\mathbf{a} - \mathbf{b} + \mathbf{c} \in \mathbf{A}$.

Bizonyítás. Mindenekelőtt belátjuk, hogy az (1) feltételből következik (1'), ha K -nak kettőnél több eleme van.

Legyen $c \notin \{0, 1\}$ tetszőleges eleme K -nak. Ekkor létezik olyan $d \notin \{0, 1\}$ K -beli elem, amelyre $\frac{1}{c} + \frac{1}{d} = 1$. Valóban $d = \frac{c}{c-1}$ nyilván megfelel. Ha \mathbf{A} -ra teljesül (1), akkor $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{A}$ esetén $\mathbf{u} = c\mathbf{a} + (1-c)\mathbf{b}$, $\mathbf{v} = (1-d)\mathbf{b} + d\mathbf{c} \in \mathbf{A}$ is igaz. Ekkor $\frac{1-c}{c} + \frac{1-d}{d} = \frac{1}{c} - 1 + \frac{1}{d} - 1 = 1 - 1 - 1 = -1$ alapján

$$\mathbf{a} - \mathbf{b} + \mathbf{c} = \frac{1}{c}\mathbf{u} + \frac{1}{d}\mathbf{v} \in \mathbf{A}$$

is teljesül.

Most megmutatjuk, hogy (2) és (1) ekvivalensek. (2)-ből (1) mint speciális eset következik. A megfordítás bizonyítására legyen $|K| > 2$. Most n -re vonatkozó teljes indukcióval bizonyítunk. $n = 1$ esetén az állítás triviálisan igaz; míg $n = 2$ esetén ez éppen az (1) alatti állítás. Tegyük fel, hogy az állítás $n > 2$ esetében igaz, és tekintsük a $\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbf{A}$ vektoroknak egy $\mathbf{a}' = \sum_{i=0}^n c_i \mathbf{a}_i$ lineáris kombinációját, ahol $d = \sum_{i=0}^n c_i = 1$. Ha az itt fellépő skalárok valamelyike 0, akkor az összegnek legfeljebb n tagja van, a teljes indukciós feltevés szerint tehát $\mathbf{a}' \in \mathbf{A}$. Ha $c_0 = 1$, akkor a $\mathbf{b} = (c_1 + 1)\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_n\mathbf{a}_n$ elem az indukciós feltétel szerint \mathbf{A} -beli, hiszen a tagok száma n és az együtthatók összege 1. Ekkor $\mathbf{a}' = \mathbf{a}_0 - \mathbf{a}_1 + \mathbf{b} \in \mathbf{A}$ az (1') feltétel szerint. Végül a $c \neq 1$ esetben legyen $\mathbf{b} = \sum_{i=1}^n \frac{c_i}{1-c_0} \mathbf{a}_i$.

Az együtthatókra vonatkozó feltételből $\sum_{i=1}^n c_i = 1 - c_0$ alapján a \mathbf{b} -ben fellépő együtthatók

összege 1, így a teljes indukciós feltétel miatt $\mathbf{b} \in \mathbf{A}$. Most $\mathbf{a}' = c_0 \mathbf{a}_0 + (1 - c_0) \mathbf{b}$, és így $\mathbf{a}' \in \mathbf{A}$, hiszen ennek az összegnek csak két tagja van és $c_0 + (1 - c_0) = 1$.

Ha K -nak két eleme van, akkor ismét teljes indukcióval bizonyítunk. Most is feltehető, hogy minden együttható 1 és $n > 3$. Az indukciós feltétel szerint $\mathbf{b} = \mathbf{a}_2 + \dots + \mathbf{a}_n \in \mathbf{A}$, amiből $\mathbf{a}' = \mathbf{a}_0 + \mathbf{a}_1 + \mathbf{b} \in \mathbf{A}$ következik, hiszen a kételemű testben $-1 = +1$.

Tegyük most fel, hogy \mathbf{A} lineáris alakzat, azaz bármely eleme $\mathbf{a} + \mathbf{u}$ alakú, ahol \mathbf{a} rögzített és $\mathbf{u} \in \mathcal{U}$ (\mathcal{U} rögzített altér). Legyen \mathbf{U} -nak két eleme $\mathbf{a} + \mathbf{u}$ és $\mathbf{a} + \mathbf{v}$; és a $c, d \in K$ elemekre teljesüljön $c + d = 1$. Ekkor

$$c(\mathbf{a} + \mathbf{u}) + d(\mathbf{a} + \mathbf{v}) = (c + d)\mathbf{a} + (c\mathbf{u} + d\mathbf{v}) \in \mathbf{A},$$

hiszen $c + d = 1$ és $c\mathbf{u} + d\mathbf{v} \in \mathcal{U}$.

Fordítva, tegyük fel, hogy \mathbf{A} -ra érvényes (1), és legyen $\mathbf{a} \in \mathbf{A}$. Tekintsük az $\mathbf{A}^* = \{\mathbf{u} - \mathbf{a} \mid \mathbf{u} \in \mathbf{A}\}$ halmazt. $\mathbf{a} \in \mathbf{A}$ miatt \mathbf{A}^* nem üres.

$c(\mathbf{u} - \mathbf{a}) = c(\mathbf{u} - \mathbf{a}) + \mathbf{a} - \mathbf{a} = (c\mathbf{u} + (1 - c)\mathbf{a}) - \mathbf{a}$ miatt $c(\mathbf{u} - \mathbf{a}) \in \mathbf{A}^*$, hiszen $c\mathbf{u} + (1 - c)\mathbf{a} \in \mathbf{A}$, mert $c + (1 - c) = 1$.

$(\mathbf{u} - \mathbf{a}) + (\mathbf{v} - \mathbf{a}) = (\mathbf{u} + \mathbf{v} - \mathbf{a}) - \mathbf{a}$ miatt $(\mathbf{u} - \mathbf{a}) + (\mathbf{v} - \mathbf{a}) \in \mathbf{A}^*$, mert $1 + 1 - 1 = 1$.

Így $\mathcal{U} = \mathbf{A}^*$ altér, amelyre nyilvánvalóan igaz, hogy \mathbf{A} az \mathcal{U} -nak \mathbf{a} -val való eltoltja. ■

Feladatok

1. Tekintsük a \mathcal{V} vektortér altereinek $\mathcal{L}(\mathcal{V})$ halmazán a \wedge és \vee operációkat mint kétváltozós műveleteket. Bizonyítsuk be, hogy ezek a műveletek kielégítik az alábbiakat:

1. Mindkét művelet idempotens: $\mathcal{A} \wedge \mathcal{A} = \mathcal{A} \vee \mathcal{A} = \mathcal{A}$.
2. Mindkét művelet kommutatív.
3. Mindkét művelet asszociatív.
4. Érvényes a két elnyelési tulajdonság: $(\mathcal{A} \wedge \mathcal{B}) \vee \mathcal{A} = (\mathcal{A} \vee \mathcal{B}) \wedge \mathcal{A} = \mathcal{A}$.
5. Érvényes a modularitási tulajdonság: Ha $\mathcal{A} \leq \mathcal{C}$, akkor tetszőleges \mathcal{B} altérre

$$(\mathcal{A} \vee \mathcal{B}) \wedge \mathcal{C} = \mathcal{A} \vee (\mathcal{B} \wedge \mathcal{C}).$$

6. Általában nem disztributív, azaz léteznek olyan $\mathcal{A}, \mathcal{B}, \mathcal{C}$ alterek, amelyekre

$$(\mathcal{A} \vee \mathcal{B}) \wedge \mathcal{C} \neq (\mathcal{A} \wedge \mathcal{C}) \vee (\mathcal{B} \wedge \mathcal{C}) \text{ és } (\mathcal{A} \wedge \mathcal{B}) \vee \mathcal{C} \neq (\mathcal{A} \vee \mathcal{C}) \wedge (\mathcal{B} \vee \mathcal{C}).$$

Definíció. Az n -dimenziós vektortér altereinek összességét $(n - 1)$ -dimenziós projektív térnek nevezzük. Ha az \mathcal{A}, \mathcal{B} alterekre $\mathcal{A} \leq \mathcal{B}$, akkor azt mondjuk, hogy \mathcal{A} rajta van a \mathcal{B} -n, illetve \mathcal{B} keresztülmegy \mathcal{A} -n.

Az egydimenziós alterek a projektív tér pontjai, a kétdimenziósak a projektív tér egyenesei, a háromdimenziósak neve projektív sík. ■

2. Bizonyítsuk be, hogy a projektív sík bármely két pontján a projektív síknak pontosan egy egyenese megy keresztül és bármely két egyenesének pontosan egy közös pontja van.

3. Bizonyítsuk be, hogy ha lineáris alakzatok közös része nem üres, akkor lineáris alakzat.

4. Nevezzük az $\mathbf{A} = \mathbf{a} + \mathcal{A}$ lineáris alakzatot a $\mathbf{B} = \mathbf{b} + \mathcal{B}$ lineáris alakzattal párhuzamosnak, ha $\mathcal{A} \leq \mathcal{B}$. Bizonyítsuk be, hogy ez a reláció reflexív és tranzitív, de nem szimmetrikus.

5. Nevezzünk két lineáris alakzatot párhuzamosnak, ha mindegyik párhuzamos a másikkal. Mutassuk meg, hogy ez egy ekvivalenciareláció.

6. Tekintsük az $a_1x_1 + \dots + a_nx_n + b$ lineáris polinomot. Ennek gyöke az n -elemű sormátrixok körében az $\mathbf{u} = [\xi_1, \dots, \xi_n]$ vektor, ha $a_1\xi_1 + \dots + a_n\xi_n + b = 0$. Mutassuk meg, hogy egy lineáris polinom gyökei lineáris alakzatot alkotnak. Milyenek azok a lineáris polinomok, amelyekhez az ily módon megfeleltetett lineáris alakzat altér?

7. Tekintsük a háromdimenziós vektorteret, ebben a projektív síkot és egy olyan lineáris alakzatot, amely egy síknak az eltoltja, de nem altér. Tekintsük ennek az eltoltnak a metszetét a projektív sík pontjaival és egyeneseivel. Bizonyítsuk be, hogy így a „naiv értelemben vett” projektív síkot kapjuk. Minek felel meg a „végtelen távoli” egyenes, illetve a „végtelen távoli” pontok?

8. Bizonyítsuk be, hogy egy számtest feletti vektortérnek vagy egy, vagy kettő, vagy végtelen sok altere van.

9. Jelölje \mathbb{Q}_p az egészek modulo p vett maradékosztályainak testét. Hány eleme és hány altere van a \mathbb{Q}_p feletti n -dimenziós vektortérnek? Hány eleme van a \mathbb{Q}_p feletti n -dimenziós projektív térnek?

10. Bizonyítsuk be, hogy az $\begin{bmatrix} a & b \\ b & a+b \end{bmatrix}$ alakú mátrixok ($a, b \in \mathbb{Q}$) egy négyelemű testet alkotnak a mátrixműveletekre. (Egyébként, ha p olyan prímszám, amelyik 5-tel osztva 2-t vagy 3-at ad maradékkal, és $a, b \in \mathbb{Q}_p$, akkor a fenti mátrixok egy p^2 elemű testet alkotnak. Más prímszám esetében nem alkotnak testet.)

11. Tételezzük fel, hogy a LOTTÓ egyik húzásánál minden kihúzott szám legfeljebb 21. Hány lottószelvényt kell kitölteni ahhoz, hogy mindenképpen legyen közöttük kéttalalatos?

12. Legyen \mathbf{A} a k -dimenziós \mathcal{U} altérnek \mathcal{U} -tól különböző eltoltja. Bizonyítsuk be, hogy \mathbf{A} -ban bármely $k+2$ vektor lineárisan összefüggő, de van $k+1$ olyan, amelyik lineárisan független.

Az alábbiakban R mindig egységelemes integritási tartományt jelöl. Emlékeztetünk arra, hogy egy gyűrűnek valamely nemüres részhalmazát részgyűrűnek nevezzük, ha zárt a gyűrűbeli műveletekre. R_R jelöli a gyűrűt mint önmaga feletti modulust.

13. Bizonyítsuk be, hogy R_R minden részmodulusa részgyűrű.

14. Mutassuk meg, hogy $\mathbb{Q}[x]$ -ben van olyan részgyűrű, amelyik nem részmodulus ($\mathbb{Q}[x]$ -modulus).

15. Mutassuk meg, hogy \mathbb{Q} -ban van olyan \mathbb{Z} -modulus, amelyik nem részgyűrű, és bizonyítsuk be, hogy minden részgyűrű \mathbb{Z} -modulus.

16. Jellemezzük a fellépő törtek nevezőivel a \mathbb{Q} részmodulusait és részgyűrűit.

17. Bizonyítsuk be, hogy R_R -ben (R egységelemes!) létezik egyelemű generátorrendszer; s ha R euklideszi gyűrű, akkor minden részmodulusa is egy elemmel generálható.

18. Bizonyítsuk be, hogy $\mathbb{Z}[x]$ -ben van $\mathbb{Z}[x]$ részmodulusainak olyan $\mathcal{U}_1 \geq \mathcal{U}_2 \geq \dots \geq \mathcal{U}_n \geq \dots$ végtelen sorozata, hogy \mathcal{U}_n nem generálható n -nél kevesebb elemmel, de minden részmodulus generálható véges sok elemmel.

19. Mutassuk meg, hogy az előző feladattal analóg eredmény bizonyítható a \mathbb{Q} feletti k -határozatlanú polinomgyűrűre tetszőleges $k \in \mathbb{N}$ esetén.

20. Adjunk meg olyan R gyűrűt (egységelemes integritási tartományt), amelyekre R_R ugyan egy elemmel generálható, de van olyan részmodulusa, amelyik nem generálható véges sok elemmel.

2. Faktorterek

Az alábbiakban egyetlen altér eltoltjait vizsgáljuk. Ezek hasonlóképpen viselkednek, mint az egész számoknak modulo m vett maradékosztályai rögzített m egész szám esetén.

A lineáris alakzat fogalmát a 2.2. Definícióban vezettük be. A maradékosztályokhoz hasonlóan \mathcal{U} két eltoltjának vagy nincs közös eleme, vagy pontosan ugyanazok az elemei. Ennek a bizonyításához felhasználjuk ezeknek egy jellemzését:

2.6. Tétel. *A \mathcal{V} vektortér \mathbf{u}, \mathbf{v} elemei pontosan akkor vannak az \mathcal{U} altér ugyanazon eltoltjában, ha $\mathbf{u} - \mathbf{v} \in \mathcal{U}$.*

Bizonyítás. Ha a két vektor eleme az \mathcal{U} altér \mathbf{a} -val való eltoltjának, akkor $\mathbf{u} = \mathbf{a} + \mathbf{u}'$ és $\mathbf{v} = \mathbf{a} + \mathbf{v}'$, ahol $\mathbf{u}', \mathbf{v}' \in \mathcal{U}$. Ekkor $\mathbf{u} - \mathbf{v} = (\mathbf{a} + \mathbf{u}') - (\mathbf{a} + \mathbf{v}') = \mathbf{u}' - \mathbf{v}' \in \mathcal{U}$; hiszen \mathcal{U} altér. Fordítva, tegyük fel, hogy $\mathbf{u} - \mathbf{v} = \mathbf{w} \in \mathcal{U}$, és legyen \mathbf{v} az \mathcal{U} -nak \mathbf{a} -val való eltoltjában, azaz $\mathbf{v} = \mathbf{a} + \mathbf{u}'$, ahol $\mathbf{u}' \in \mathcal{U}$. Ekkor $\mathbf{u} = \mathbf{v} + \mathbf{w} = \mathbf{a} + (\mathbf{u}' + \mathbf{w})$ ugyancsak eleme az \mathcal{U} altér \mathbf{a} -val való eltoltjának, hiszen $\mathbf{u}' + \mathbf{w} \in \mathcal{U}$, mert \mathcal{U} altér. ■

2.7. Tétel. *Az adott altér szerinti ugyanazon eltolthoz való tartozás ekvivalenciareláció (azaz reflexív, szimmetrikus és tranzitív). Ezt a relációt (a kongruenciákhoz hasonlóan) $\mathbf{u} \equiv \mathbf{v}(\mathcal{U})$ fogja jelölni. Az ekvivalenciaosztályok pontosan az \mathcal{U} eltoltjai. Ha \mathbf{A} és \mathbf{B} két ekvivalenciaosztály, akkor ezek vagy megegyeznek, vagy diszjunktak.*

Bizonyítás. A 2.6. Tétel szerint \mathbf{u} és \mathbf{v} pontosan akkor tartoznak ugyanazon eltolthoz, ha $\mathbf{u} - \mathbf{v} \in \mathcal{U}$. Ennek megfelelően a reflexivitás azt jelenti, hogy $\mathbf{u} - \mathbf{u} \in \mathcal{U}$; ez abból következik, hogy $\mathbf{0} \in \mathcal{U}$. A szimmetria azt jelenti, hogy ha $\mathbf{w} = \mathbf{u} - \mathbf{v} \in \mathcal{U}$, akkor $-\mathbf{w} = -(\mathbf{u} - \mathbf{v}) = \mathbf{v} - \mathbf{u} \in \mathcal{U}$ is igaz; ami abból következik, hogy egy altérben bármely elemének az ellentettje is benne van. A tranzitivitás azt jelenti, hogy ha $\mathbf{u} - \mathbf{v} \in \mathcal{U}$ és $\mathbf{v} - \mathbf{w} \in \mathcal{U}$ mindegyike teljesül, akkor igaz $\mathbf{u} - \mathbf{w} \in \mathcal{U}$ is. Ez pedig annak következménye, hogy egy altér bármely két elemével együtt ezek összegét is tartalmazza; márpedig $\mathbf{u} - \mathbf{w} = (\mathbf{u} - \mathbf{v}) + (\mathbf{v} - \mathbf{w})$.

Azt, hogy az ekvivalenciaosztályok pontosan az \mathcal{U} eltoltjai, a 2.6. Tétel mondta ki. Ha \mathbf{A} és \mathbf{B} két ekvivalenciaosztály, akkor legyen $\mathbf{a} \in \mathbf{A}$ és $\mathbf{b} \in \mathbf{B}$ tetszőlegesen választva. Ha $\mathbf{u} \in \mathbf{A} \cap \mathbf{B}$, akkor $\mathbf{u} \equiv \mathbf{a}$ és $\mathbf{a} \equiv \mathbf{b}$; s az ekvivalenciarelációk tulajdonságai alapján $\mathbf{a} \equiv \mathbf{b}$, tehát $\mathbf{A} = \mathbf{B}$. ■

Következmény. *Legyen \mathcal{U} a \mathcal{V} vektortér altere, \mathbf{A} az \mathcal{U} -nak egy eltoltja és $\mathbf{a} \in \mathbf{A}$. Ekkor \mathbf{A} az \mathcal{U} -nak \mathbf{a} -val való eltoltja (is). Az \mathbf{A} tetszőleges elemét e lineáris alakzat egy reprezentánsának nevezzük.*

Bizonyítás. Tekintettel arra, hogy tetszőleges $\mathbf{a}' \in \mathbf{A}$ esetén $\mathbf{u} = \mathbf{a}' - \mathbf{a} \in \mathcal{U}$, ezért $\mathbf{a}' = \mathbf{a} + \mathbf{u}$, vagyis \mathbf{A} minden eleme benne van az \mathbf{a} -val való eltoltban. A 2.7. Tétel alapján a két eltolt megegyezik. ■

2.8. Tétel. *A K test feletti \mathcal{V} vektortér \mathcal{U} altere által a \mathcal{V} elemein megadott fenti ekvivalenciareláció a műveletekkel kompatibilis:*

Legyen \mathbf{A} és \mathbf{B} az \mathcal{U} -nak két eltoltja, $\mathbf{a}, \mathbf{a}' \in \mathbf{A}$, $\mathbf{b}, \mathbf{b}' \in \mathbf{B}$ és $c \in K$. Ekkor $\mathbf{a} + \mathbf{b}$, valamint $\mathbf{a}' + \mathbf{b}'$ ugyanabban az eltoltban vannak; ugyanúgy, mint ahogy a $c\mathbf{a}$ és $c\mathbf{a}'$ is.

Bizonyítás. A 2.6. Tétel szerint feltételünk azt jelenti, hogy $\mathbf{a} - \mathbf{a}'$, $\mathbf{b} - \mathbf{b}' \in \mathcal{U}$. Ebből kell arra következtetni, hogy $(\mathbf{a} + \mathbf{b}) - (\mathbf{a}' + \mathbf{b}') \in \mathcal{U}$ és $c\mathbf{a} - c\mathbf{a}' \in \mathcal{U}$.

Az első összefüggés az $(\mathbf{a} + \mathbf{b}) - (\mathbf{a}' + \mathbf{b}') = (\mathbf{a} - \mathbf{a}') + (\mathbf{b} - \mathbf{b}')$ egyenlőségből, a második a $c\mathbf{a} - c\mathbf{a}' = c(\mathbf{a} - \mathbf{a}')$ egyenlőségből következik; tekintettel arra, hogy \mathcal{U} mind az összeadásra, mind a skalárral való szorzásra zárt. ■

A 2.8. Tétel lehetőséget ad arra, hogy egy alter eltoltjaival műveleteket végezzünk; s ezek az alterek a definiált műveletekre nézve maguk is vektorteret alkossanak. A műveletek definiálására két lehetőség is van, de egyik esetben sem világos az, hogy a definiált műveletek valóban „értelmesek”. Ezt mindkét esetben utólag kell belátni.

2.3. Definíció. Legyen \mathcal{V} vektortér a K test felett és legyen $\mathcal{U} \leq \mathcal{V}$.

1. változat: Ha \mathbf{A} az \mathcal{U} -nak \mathbf{a} -val való eltoltja és \mathbf{B} az \mathcal{U} -nak \mathbf{b} -vel való eltoltja, akkor legyen $\mathbf{A} + \mathbf{B}$ az \mathcal{U} -nak $(\mathbf{a} + \mathbf{b})$ -vel való eltoltja és $c \cdot \mathbf{A}$ az \mathcal{U} -nak $c \cdot \mathbf{a}$ -val való eltoltja ($c \in K$).

2. változat: Ha \mathbf{A} és \mathbf{B} az \mathcal{U} -nak két eltoltja, akkor legyen

$$\mathbf{A} + \mathbf{B} = \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in \mathbf{A}, \mathbf{b} \in \mathbf{B}\} \quad \text{és} \quad c \cdot \mathbf{A} = \begin{cases} \{c \cdot \mathbf{a} \mid \mathbf{a} \in \mathbf{A}\}, & \text{ha } c \neq 0 \\ \mathcal{U}, & \text{ha } c = 0. \end{cases} \quad \blacksquare$$

Megjegyzés. Az 1. változat esetében nem világos, hogy a definiált összeg és skalárszoros nem függ-e attól, hogy melyik reprezentánst szemeltük ki. A 2. változat esetében az nem világos, hogy a definiált összeg és skalárszoros valóban az \mathcal{U} alter eltoltja. □

2.9. Tétel. *A 2.3. Definíció 1. pontjában megadott műveletek eredménye nem függ a reprezentánstól. A 2. pontban megadott műveletek eredménye az \mathcal{U} alter eltoltja. A kétféleképpen megadott műveletek ugyanazt az összeget és skalárszorost adják.*

Bizonyítás. A 2.8. Tétel bizonyítása azonnal szolgáltatja az 1. változatban a műveletek egyértelműségét.

A 2. változatban legyen $\mathbf{a}, \mathbf{a}' \in \mathbf{A}$ és $\mathbf{b}, \mathbf{b}' \in \mathbf{B}$. Az előbb tekintetbe vett egyenlőségekből következik, hogy $\mathbf{A} + \mathbf{B}$ elemei is és $c \cdot \mathbf{A}$ elemei is egyetlen eltoltba esnek. $\mathbf{a} + \mathbf{b} + \mathbf{u} = (\mathbf{a} + \mathbf{u}) + (\mathbf{b} + \mathbf{o})$ bizonyítja, hogy $\mathbf{A} + \mathbf{B}$ az adott eltolt minden elemét tartalmazza ($\mathbf{u} \in \mathcal{U}$). $c \cdot \mathbf{a} + \mathbf{u} = c \cdot \left(\mathbf{a} + \frac{1}{c} \cdot \mathbf{u} \right)$ bizonyítja, hogy $c \cdot \mathbf{A}$ az adott eltolt minden elemét tartalmazza ($\mathbf{u} \in \mathcal{U}$), ha $c \neq 0$; míg a $c = 0$ eset triviális.

A két definíció ekvivalenciája nyilvánvaló. ■

Megjegyzés. A 2. változatban elég lenne azt megkövetelni, hogy $\mathbf{A} + \mathbf{B}$ tartalmazza az összes $\mathbf{a} + \mathbf{b}$ ($\mathbf{a} \in \mathbf{A}$, $\mathbf{b} \in \mathbf{B}$) alakú elemet és $c \cdot \mathbf{A}$ tartalmazza az összes $c \cdot \mathbf{a}$ ($\mathbf{a} \in \mathbf{A}$) alakú elemet. Ezekről előzőleg már beláttuk, hogy egyetlen eltoltba esnek; ezért a definíció teljesen „szabályos” volna. Emellett még arra sem lenne szükség, hogy a skalárszoros definíciójában eseteket kelljen megkülönböztetni. Azért választottuk mégis az „ügyetlenebb” formát, mert az jobban megmutatja, hogy miképpen adódik az összeg és a skalárszoros. \square

2.10. Tétel. Legyen \mathcal{V} vektortér a K test felett és legyen $\mathcal{U} \leq \mathcal{V}$. Az \mathcal{U} eltoltjai vektorteret alkotnak a K felett a 2.3. Definícióban megadott műveletekre nézve. Ezt a vektorteret a \mathcal{V} vektortér \mathcal{U} szerinti faktortérének nevezzük és $(\mathcal{V}/\mathcal{U})$ -val jelöljük.

A \mathcal{V}/\mathcal{U} és a $\mathcal{V}/\{\mathbf{o}\}$ faktortereket triviális faktortereknek nevezzük; az összes többinek a neve valódi faktortér.

A \mathcal{V} -beli \mathbf{a} vektornak megfelelően az őt tartalmazó \mathbf{A} eltoltat, ez az $\mathbf{a} \mapsto \mathbf{A}$ megfeleltetés a következő tulajdonsággal rendelkezik: $\mathbf{a} + \mathbf{b} \mapsto \mathbf{A} + \mathbf{B}$ és $c \cdot \mathbf{a} \mapsto c \cdot \mathbf{A}$ (ahol $\mathbf{b} \mapsto \mathbf{B}$, illetve $c \in K$). (Ezzel ekvivalens a $c \cdot \mathbf{a} + d \cdot \mathbf{b} \mapsto c \cdot \mathbf{A} + d \cdot \mathbf{B}$ definíció, ahol az előbbieken túl $d \in K$ is fel van téve.)

A tétel — megfelelően — modulusokra is igaz.

Bizonyítás. A 2.9. Tétel alapján a 2.3. pontban az eltoltakra valóban műveleteket értelmeztünk. Ezek a műveletek megfelelnek a K feletti vektortereknél bevezetett műveleteknek. A műveleti azonosságok azonnal látható módon következnek a \mathcal{V} -ben teljesülő megfelelő műveleti azonosságokból. A faktortér nullvektora \mathcal{U} . A megfeleltetésnél állított tulajdonságok is azonnal következnek a megfeleltetés megadásából. \blacksquare

2.11. Tétel. Ha $\{\mathbf{u}_1, \dots, \mathbf{u}_r\} \cup \mathcal{U} = \mathcal{V}$ és $\mathbf{u}_i \in \mathbf{U}_i \in \mathcal{V}/\mathcal{U}$, akkor $\mathbf{U}_1, \dots, \mathbf{U}_r$ a \mathcal{V}/\mathcal{U} generátorrendszere. Speciálisan ez az eset áll fenn, ha $\mathbf{u}_1, \dots, \mathbf{u}_r$ generátorrendszer.

Ha $\mathbf{U}_1, \dots, \mathbf{U}_r$ a \mathcal{V}/\mathcal{U} -ban független rendszer, akkor bármely $\mathbf{u}_i \in \mathbf{U}_i$ elemekre (minden egyes \mathbf{U}_i -ből pontosan egyet választva!) ezek nemtriviális lineáris kombinációja nem lehet eleme \mathcal{U} -nak, speciálisan, lineárisan függetlenek.

Bizonyítás. Legyen $\mathbf{U} \in \mathcal{V}/\mathcal{U}$ és tekintsünk egy tetszőleges $\mathbf{u} \in \mathbf{U}$ vektort. Feltétel szerint ez eleme az $\{\mathbf{u}_1, \dots, \mathbf{u}_r\} \cup \mathcal{U}$ generátumnak, azaz léteznek olyan $c_1, \dots, c_r \in K$ skalárok és olyan $\mathbf{u}_0 \in \mathcal{U}$ vektor, amelyre $\mathbf{u} = \mathbf{u}_0 + c_1 \mathbf{u}_1 + \dots + c_r \mathbf{u}_r$. A lineáris alakzatokra definiált műveletek alapján a jobb oldal eleme a $c_1 \mathbf{U}_1 + \dots + c_r \mathbf{U}_r$ eltoltnak; tehát $\mathbf{U}_1, \dots, \mathbf{U}_r$ valóban generátorrendszer.

Tegyük most fel, hogy $\mathbf{U}_1, \dots, \mathbf{U}_r$ egy független rendszer és legyen $\mathbf{u}_i \in \mathbf{U}_i$. Legyenek a $c_1, \dots, c_r \in K$ skalárok olyanok, hogy $\sum_{i=1}^r c_i \mathbf{u}_i \in \mathcal{U}$. Ekkor — a lineáris alakzatokra

definiált műveleteket figyelembe véve — azt kapjuk, hogy $\sum_{i=1}^r c_i \mathbf{u}_i \in \sum_{i=1}^r c_i \mathbf{U}_i = \mathcal{U}$. A fel-

tett lineáris függetlenség következtében itt minden egyes i indexre $c_i = 0$; amiből azonnal következik a kiszemelt vektorokra az állítás. \blacksquare

Feladatok

1. Legyen K test, \mathcal{V}_K véges dimenziós, és $\mathcal{U} \leq_K \mathcal{V}$. Bizonyítsuk be, hogy $\dim(\mathcal{V}/\mathcal{U}) = \dim(\mathcal{V}) - \dim(\mathcal{U})$.
2. Legyenek $\mathcal{U} \leq_K \mathcal{V}$ vektorterek. Bizonyítsuk be, hogy van olyan $\mathcal{W} \leq_K \mathcal{V}$ altér, hogy $\mathcal{W} \cap \mathcal{U} = \{\mathbf{o}\}$ és $\mathcal{W} \vee \mathcal{U} = \mathcal{V}$. Mutassuk meg, hogy \mathcal{W} általában nem egyértelmű.
3. Bizonyítsuk be, hogy egy vektortér minden faktortere izomorf egy alterével és viszont.
4. Mutassuk meg, hogy $\mathbb{Z}_{\mathbb{Z}}$ minden valódi részmodulusa izomorf $\mathbb{Z}_{\mathbb{Z}}$ -vel, de valódi faktormodulusai nem izomorfak $\mathbb{Z}_{\mathbb{Z}}$ -vel.
5. Mutassuk meg, hogy $\mathbb{Z}_{\mathbb{Z}}$ különböző részmodulusok szerinti faktormodulusai nem izomorfak egymással.
6. Mutassuk meg, hogy $\mathbb{Q}[x]_{\mathbb{Q}[x]}$ -nek vannak olyan különböző részmodulusai, hogy a szerinti vett faktormodulusok izomorfak.
7. Mutassuk meg, hogy ha R olyan egységelemes integritási tartomány, hogy R_R minden valódi faktormodulusa izomorf egy valódi részmodulusával, akkor R test.
8. Legyen $\mathcal{U} \leq_K \mathcal{V}$ és tekintsük azt a $\varphi : \mathcal{V} \rightarrow \mathcal{V}/\mathcal{U}$ leképezést, amely minden $\mathbf{v} \in \mathcal{V}$ elemnek megfelelteti az \mathcal{U} -nak a \mathbf{v} -t tartalmazó eltoltját. Bizonyítsuk be, hogy $\mathcal{W} \leq \mathcal{V}$ esetén $\varphi(\mathcal{W}) = \{\varphi(\mathbf{v}) \mid \mathbf{v} \in \mathcal{W}\} \leq \mathcal{V}/\mathcal{U}$.
9. Legyen \mathcal{U}, \mathcal{V} és φ , mint a 8. feladatban, és legyen $\mathcal{W}^* \leq \mathcal{V}/\mathcal{U}$. Jelölje $\widetilde{\mathcal{W}}^*$ az $\{\mathbf{u} \mid \varphi(\mathbf{u}) \in \mathcal{W}^*\}$ halmazt. Bizonyítsuk be, hogy $\mathcal{U} \leq \widetilde{\mathcal{W}}^* \leq \mathcal{V}$.
10. A 8. és a 9. feladat jelöléseit használva bizonyítsuk be, hogy $\varphi(\widetilde{\mathcal{W}}^*) = \mathcal{W}^*$ és $\varphi(\widetilde{\mathcal{W}}) = \mathcal{W} \vee \mathcal{U}$.
11. Bizonyítsuk be, hogy a 8., 9. és 10. feladat eredményei modulusokra is igazak.
12. Legyen $\mathcal{U} \leq_K \mathcal{W} \leq_K \mathcal{V}$. Bizonyítsuk be, hogy $(\mathcal{V}/\mathcal{U})/(\varphi(\mathcal{W})/\mathcal{U})$ és \mathcal{V}/\mathcal{W} izomorfak.
13. Bizonyítsuk be, hogy $\mathbb{Q}_{\mathbb{Z}}$ minden valódi faktormodulusa izomorf a komplex egységgyökök \mathbb{Z} -modulusának egy részcsoportjával (ez utóbbiakban a vektorok összege a szorzás és az n egész számmal való szorzás az n -edik hatvány).

3. Direkt összeg és direkt szorzat

Tekintsük a K feletti \mathcal{V} vektortér egy \mathcal{U} alterét. Ha a \mathcal{V}/\mathcal{U} faktortérnek adott egy $\mathbf{U}_1, \dots, \mathbf{U}_n, \dots$ bázisa, akkor láttuk, hogy minden egyes \mathbf{U}_i -ből egy-egy \mathbf{v}_i elemet kiválasztva, ezek függetlenek lesznek, sőt egyetlen nemtriviális lineáris kombinációjuk sem eshet \mathcal{U} -ba (2.11. Tétel). Ez azt jelenti, hogy az általuk generált \mathcal{W} altérnek és az \mathcal{U} altérnek egyetlen közös eleme a \mathbf{o} vektor. Ugyancsak a 2.11. Tételből kapjuk, hogy $\mathcal{U} \vee \mathcal{W} = \mathcal{U}$. Ezt, és az előbb kapott $\mathcal{U} \cap \mathcal{W} = \{\mathbf{o}\}$ összefüggést nézve semmiféle „eltérést” nem találhatunk a két altér között; a \mathcal{V} tér e két alteréből teljesen szimmetrikusan „épül fel”. Ez a

jelenség hasonlatos ahhoz, ahogy a síkbeli koordináta-rendszer a két koordinátatengelyből előáll. Ez a konstrukció általában is elvégezhető és igen hasznos.

2.4. Definíció. A K feletti \mathcal{V} vektortér \mathcal{A} és \mathcal{B} altereinek direkt összege — jelölésben $\mathcal{V} = \mathcal{A} \oplus \mathcal{B}$ —, ha $\mathcal{A} \vee \mathcal{B} = \mathcal{V}$ és $\mathcal{A} \wedge \mathcal{B} = \mathcal{O}$, ahol $\mathcal{O} = \{\mathbf{o}\}$. Itt \mathcal{A} és \mathcal{B} a direkt összeg tagjai. Az $\mathcal{A} = \mathcal{O}$ és $\mathcal{B} = \mathcal{V}$, illetve az $\mathcal{A} = \mathcal{V}$ és $\mathcal{B} = \mathcal{O}$ esetben triviális direkt összegről beszélünk, minden más esetben valódi direkt összegről.

Megjegyzés. A direkt összeg egyik tagja általában nem határozza meg a másikat; kivéve, ha ez a tag \mathcal{O} vagy \mathcal{V} . Ekkor a másik tag, megfelelően, \mathcal{V} , illetve \mathcal{O} . \square

A direkt összeg modulusok esetében hasonlóan értelmezhető. \blacksquare

Az elnevezést és a jelölést jobban megmagyarázza az alábbi

2.12. Tétel. \mathcal{V} akkor és csak akkor direkt összege \mathcal{A} és \mathcal{B} altereinek, ha \mathcal{V} minden eleme egyértelműen felírható $\mathbf{a} + \mathbf{b}$ alakban, ahol $\mathbf{a} \in \mathcal{A}$ és $\mathbf{b} \in \mathcal{B}$.

Ekkor az \mathcal{A} egy \mathbf{A} és a \mathcal{B} egy \mathbf{B} generátorrendszerének egyesítése \mathcal{V} -nek egy generátorrendszere; és amennyiben ezek mindegyike a megfelelő altérnek bázisa, akkor egyesítésük ugyancsak bázisa \mathcal{V} -nek.

A \mathcal{V} vektortér bármely \mathcal{A} alteréhez létezik olyan \mathcal{B} altere, hogy $\mathcal{V} = \mathcal{A} \oplus \mathcal{B}$.

Minden ilyen \mathcal{B} altér az \mathcal{A} -nak direkt kiegészítője.

Bizonyítás. A direkt összeg tulajdonság is és az előállíthatósági tulajdonság is két részre bontható. Azt fogjuk bebizonyítani, hogy — megfelelően — ezek ekvivalensek.

Ha $\mathcal{V} = \mathcal{A} \vee \mathcal{B}$, akkor e két altér generálja \mathcal{V} -t; azaz \mathcal{V} minden \mathbf{v} eleme előáll olyan lineáris kombinációként, amelyben a fellépő vektorok mindegyike vagy \mathcal{A} -ban, vagy \mathcal{B} -ben van. Tekintettel arra, hogy mind \mathcal{A} , mind \mathcal{B} zárt a lineáris kombináció képzésére, ezért ezeket összevonva egy $\mathbf{v} = \mathbf{a} + \mathbf{b}$ alakú előállítást nyerünk, ahol $\mathbf{a} \in \mathcal{A}$ és $\mathbf{b} \in \mathcal{B}$.

Fordítva, ha \mathcal{V} minden eleme felírható ilyen alakban, akkor $\mathcal{V} = \mathcal{A} \vee \mathcal{B}$ nyilvánvalóan igaz.

Tegyük most fel, hogy $\mathcal{A} \wedge \mathcal{B} = \mathcal{O}$, és legyen $\mathbf{v} = \mathbf{a}_1 + \mathbf{b}_1 = \mathbf{a}_2 + \mathbf{b}_2$, ahol $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{A}$ és $\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{B}$. Ebből azonnal kapjuk, hogy $\mathbf{a}_1 - \mathbf{a}_2 = \mathbf{b}_2 - \mathbf{b}_1$. Itt a bal oldal \mathcal{A} -nak, a jobb oldal \mathcal{B} -nek az eleme. Mivel e két altérnek egyetlen közös eleme \mathbf{o} , ezért valóban $\mathbf{a}_1 = \mathbf{a}_2$ és $\mathbf{b}_1 = \mathbf{b}_2$, mint állítottuk.

Fordítva, ha minden olyan elem, amely $\mathbf{a} + \mathbf{b}$ alakban írható ($\mathbf{a} \in \mathcal{A}$, $\mathbf{b} \in \mathcal{B}$), csak egyféleképpen írható ilyen alakba, akkor tekintsük az $\mathcal{A} \wedge \mathcal{B}$ -nek egy \mathbf{u} elemét. Mivel \mathbf{o} is eleme a közös résznek, ezért \mathbf{u} -nak $\mathbf{u} + \mathbf{o}$ is és $\mathbf{o} + \mathbf{u}$ is egy-egy megfelelő felírása. Az egyértelműség alapján ebből azonnal következik $\mathbf{u} = \mathbf{o}$; azaz a közös résznek \mathbf{o} az egyetlen eleme.

Legyen most \mathbf{A} az \mathcal{A} -nak és \mathbf{B} a \mathcal{B} -nek generátorrendszere. Ekkor az $\mathbf{A} \cup \mathbf{B}$ generálta altér tartalmazza az \mathcal{A} és \mathcal{B} alterek mindegyikét, tehát $(\mathcal{A} \vee \mathcal{B})$ -t is; így valóban generátorrendszer. Amennyiben mind \mathbf{A} , mind \mathbf{B} bázis, akkor tekintsük ezeknek az elemeknek egy $\mathbf{o} = \sum_i c_i \mathbf{a}_i + \sum_j d_j \mathbf{b}_j$ lineáris kombinációját, ahol az \mathbf{a}_i -k az \mathbf{A} -nak és a \mathbf{b}_j -k a \mathbf{B} -nek

különböző elemei. Mivel $\mathbf{a} = \sum_i c_i \mathbf{a}_i \in \mathcal{A}$ és $\mathbf{b} = \sum_j d_j \mathbf{b}_j \in \mathcal{B}$, ezért $\mathbf{a} = -\mathbf{b} \in \mathcal{A} \wedge \mathcal{B}$;

vagyis $\mathbf{a} = -\mathbf{b} = \mathbf{o}$. Tekintettel arra, hogy mind \mathbf{A} , mind \mathbf{B} bázis, ez azt jelenti, hogy minden i -re $c_i = 0$ és minden j -re $d_j = 0$, ami bizonyítja a kívánt lineáris függetlenséget.

Végezetül tekintsük a \mathcal{V}/\mathcal{A} faktortér egy \mathbf{U}_i bázisát. A 2.11. Tétel szerint minden egyes \mathbf{U} -ból kiválasztva egy \mathbf{u}_i elemet, az ezek generálta \mathcal{B} altérre teljesül a $\mathcal{V} = \mathcal{A} \oplus \mathcal{B}$ felírás. ■

Megjegyzések

1. Elég gyakran előfordul olyan eset, amikor $\mathcal{A} \vee \mathcal{B} = \mathcal{V}$, de $\mathcal{A} \wedge \mathcal{B} \neq \emptyset$. A \mathcal{V} elemei ekkor is előállnak egy \mathcal{A} -beli és egy \mathcal{B} -beli elem összegeként, de az előállítás nem egyértelmű. Ebben az esetben a $\mathcal{V} = \mathcal{A} + \mathcal{B}$ jelölést használjuk. Természetesen az is lehet, hogy az alterek generátuma nem az egész tér. Ekkor is alkalmazhatjuk az $\mathcal{A} + \mathcal{B}$ jelölést, ami most a \mathcal{V} -nek egy altéré. Vegyük figyelembe, hogy az $\mathcal{A} + \mathcal{B}$ altér ugyanazt jelenti, mint az $\mathcal{A} \vee \mathcal{B}$ altér. Ez utóbbi jelölésnek az a haszna, hogy más „algebrai struktúrák” esetében is értelmes; míg az előbbi jelölés vektortereknél és néhány „hasznos jellegű” algebrai struktúránál sokkal szemléletesebben mutatja, hogy a kapott altér elemei összeg alakban írhatók fel.

2. A bizonyítás menetéből könnyen látható, hogy az egyes alterek egy-egy független rendszerének az egyesítése a direkt összegben független rendszer. Ez a tulajdonság egyébként azzal ekvivalens, hogy a két altérnek egyedül a nullvektor közös eleme. □

Kiegészítés. $\mathcal{A} \oplus$ részleges művelet kommutatív: Ha $\mathcal{A} \oplus \mathcal{B}$ és $\mathcal{B} \oplus \mathcal{A}$ bármelyike létezik, akkor létezik a másik is, és megegyeznek.

Bizonyítás. Azonnal következik a definícióból. ■

Hasonlóképpen, mint ahogy a sík előáll a két koordinátatengely „direkt összegeként”, a tér is előáll mint a három koordinátatengely direkt összege. Ennek megfelelően értelmezhető a többtagú direkt összegként való előállítás. Ezt azonban célszerűbb a 2.12. Tételben adott előállításhoz megfelelően értelmezni; mint látni fogjuk, az eredeti definíció átvitele egyáltalában nem magától értetődő.

2.5. Definíció. A \mathcal{V} teret az $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_r$ alterei direkt összegének nevezzük, ha minden $\mathbf{v} \in \mathcal{V}$ vektor egyértelműen előáll $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_r$ alakban, ahol $\mathbf{u}_i \in \mathcal{U}_i$.

Ebben az esetben a $\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2 \oplus \dots \oplus \mathcal{U}_r$ vagy a $\bigoplus_{i=1}^r \mathcal{U}_i$ jelölést használjuk.

Végtelen tagú direkt összeget is definiálhatunk, de akkor ki kell kötni, hogy a fellépő vektorok között csak véges sok különbözik \mathbf{o} -tól. ■

2.13. Tétel. A \mathcal{V} vektortér akkor és csak akkor áll elő \mathcal{U}_i altereinek direkt összegeként, ha az \mathcal{U}_i alterek generálják \mathcal{V} -t és bármely \mathcal{U}_i altérnek a többiek generátumával való közös része egyedül a \mathbf{o} vektort tartalmazza.

Bizonyítás. A 2.12. Tételhez hasonlóan az előállíthatóság nyilvánvalóan ekvivalens azzal, hogy a szóban forgó alterek generálják \mathcal{V} -t. Azt fogjuk bizonyítani, hogy a következő két állítás ekvivalens:

1. Minden olyan vektor esetén, amely előállítható \mathcal{U}_i -beli vektorok összegeként, az előállítás egyértelmű.

2. Bármely \mathcal{U}_i altérnek a többiek generátumával való közös része egyedül a \mathbf{o} vektort tartalmazza.

Tegyük fel, hogy 1. igaz. Nem megy az általánosság rovására, ha az \mathcal{U}_1 alteret szemeljük ki. Legyen $\mathbf{v} \in \mathcal{U}_1 \wedge (\mathcal{U}_2 \vee \dots \vee \mathcal{U}_r)$, azaz $\mathbf{v} = \mathbf{u}_2 + \dots + \mathbf{v}_r$ ($\mathbf{u}_i \in \mathcal{U}_i$). Ekkor az $\mathbf{u}_1 = -\mathbf{v}$ vektorra $\mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_r = \mathbf{o}$. Mivel $\mathbf{o} \in \mathcal{U}_i$ minden i -re igaz, ezért $\mathbf{o}_1 + \mathbf{o}_2 + \dots + \mathbf{o}_r = \mathbf{o}$ is igaz (itt \mathbf{o}_i azt jelenti, hogy most éppen az \mathcal{U}_i altérből vettük a \mathbf{o} vektort). Az előállíthatóság egyértelműségének a következtében tehát $\mathbf{u}_i = \mathbf{o}$, speciálisan $\mathbf{v} = -\mathbf{u}_1 = \mathbf{o}$, mint állítottuk.

Tegyük most fel, hogy 2. igaz, és legyen

$$\mathbf{v} = \mathbf{u}_1 + \dots + \mathbf{u}_n = \mathbf{u}'_1 + \dots + \mathbf{u}'_n \quad (\mathbf{u}_i, \mathbf{u}'_i \in \mathcal{U}_i)$$

a \mathbf{v} vektornak két előállítása. Ekkor bármely i indexre $\mathbf{u}_i - \mathbf{u}'_i = \sum_{j \neq i} (\mathbf{u}'_j - \mathbf{u}_j)$. Itt a bal

oldal \mathcal{U}_i -ben van, a jobb oldal pedig az összes többi \mathcal{U}_j generátumában. A feltétel szerint tehát $\mathbf{u}_i - \mathbf{u}'_i = \mathbf{o}$, azaz bármely i indexre $\mathbf{u}'_i = \mathbf{u}_i$, tehát az előállítás egyértelmű. ■

Megjegyzés. Hasonlóan a két altér esetéhez, itt is előfordulhat, hogy a generátumot kell vizsgálni olyan esetben, amikor az összeg-előállítás nem feltétlen egyértelmű. Mivel ebben az esetben is „érthetőbb” az, hogy összeg szerepel, ezért itt is használni fogjuk $\bigvee_i \mathcal{U}_i$ helyett a $\sum_i \mathcal{U}_i$ jelölést (végtelen sok altér esetében is). □

Tekintettel arra, hogy a generátumképzésnél lényegtelen a sorrend és a társítás, ezért igaz a következő:

Kiegészítés. A többtagú direkt összeg kommutatív és asszociatív. ■

Megjegyzés. Világos, hogy az az értelmezés, miszerint bármelyik altérnek a többiek generátumával való közös része egyedül a nullvektorból áll, két altér esetében is igaz. Az a kézenfekvőnek tűnő általánosítás, hogy bármely két altérnek a metszete egyedül a nullvektort tartalmazza, nem ekvivalens a többtagú direkt összeg fent adott definíciójával. Gondoljuk meg azt, hogy a síkon akármennyi (de legalább 2) origón átmenő egyenest tekintünk, ezek generálják a síkot, és bármely kettőnek egyedül a nullvektor a közös eleme. Ennek ellenére az összeg-előállíthatóság távolról sem egyértelmű.

Azért tekintettük kétagú összeg esetében mégis az ott megadott definíciót elsődlegesen, mert ez nem használja a műveleteket. Többtagú összegnél viszont nem lett volna érthető, hogy miért az a „művelet nélküli” definíció, ami a 2.13. Tételben adódott. □

A térbeli koordináta-rendszerben minden pont jellemezhető koordinátaival, más szóval (a, b, c) alakban adható meg. Ez a teret kicsit másképpen állítja elő egydimenziós alterek segítségével. Az egydimenziós alterekben egyetlen koordináta létezik, így a fenti vektorhoz tartozó három vektor (a) , (b) és (c) . Ezek most nincsenek a térben, mert nem $(*, *, *)$ alakúak. A térbeli halmaz tulajdonképpen az egyes halmazoknak a direkt szorzata. Ezt az elképzelést tükrözi a következő:

2.6. Definíció. Az $\mathcal{U}_1, \dots, \mathcal{U}_r$ K feletti vektorterek $\mathcal{V} = \mathcal{U}_1 \times \dots \times \mathcal{U}_r$ direkt szorzatán a következőket értjük:

\mathcal{V} tartóhalmaza (alaphalmaza) az $\underline{\mathbf{u}} = (\mathbf{u}_1, \dots, \mathbf{u}_r)$ sorozatokból áll, ahol $\mathbf{u}_i \in \mathcal{U}_i$.

Ha $c \in K$, akkor $c \cdot \underline{\mathbf{u}} = (c \cdot \mathbf{u}_1, \dots, c \cdot \mathbf{u}_r)$.

Ha $\underline{\mathbf{v}} = (\mathbf{v}_1, \dots, \mathbf{v}_r)$ a tartóhalmaz egy másik (nem feltétlenül különböző) eleme, akkor $\underline{\mathbf{u}} + \underline{\mathbf{v}} = (\mathbf{u}_1 + \mathbf{v}_1, \dots, \mathbf{u}_r + \mathbf{v}_r)$.

A direkt szorzatra használatos még a $\prod_{i=1}^r \mathcal{U}_i$ jelölés is.

Ugyanazon K test feletti végtelen sok vektortér direkt szorzatát is hasonlóképpen definiáljuk, de a sorozatok végtelen hosszúak.

Az adott \mathcal{U}_i vektortereket a direkt szorzat komponenseinek nevezzük.

Modulusok esetében a direkt szorzat hasonló módon definiálható. ■

2.14. Tétel. *Adott test feletti vektorterek direkt szorzata a 2.6. Definícióban adott műveletekre nézve vektortér e test felett.*

Bizonyítás. A műveleteket definiáltuk. Az ezekre vonatkozó azonosságok teljesülése azonnal következik abból, hogy ezek komponensenként teljesülnek. A nullvektornak és az ellentettnek a létezését kell még belátni; de ez is egyszerű. A direkt szorzat nullvektora a $\underline{\mathbf{0}} = (\mathbf{0}_1, \dots, \mathbf{0}_r)$, ahol $\mathbf{0}_i$ az \mathcal{U}_i nullvektora. A $\underline{\mathbf{0}} + \underline{\mathbf{u}} = \underline{\mathbf{u}}$ összefüggés azonnal következik abból, hogy az i -edik komponensben $\mathbf{0}_i + \mathbf{u}_i = \mathbf{u}_i$ teljesül. A direkt szorzatban az $\underline{\mathbf{u}} = (\mathbf{u}_1, \dots, \mathbf{u}_r)$ elem ellentettjét $-\underline{\mathbf{u}} = (-\mathbf{u}_1, \dots, -\mathbf{u}_r)$ definiálja. A komponensenként fennálló $\mathbf{u}_i + (-\mathbf{u}_i) = \mathbf{0}_i$ összefüggésből azonnal következik az, hogy $\underline{\mathbf{u}} + (-\underline{\mathbf{u}}) = \underline{\mathbf{0}}$. ■

A direkt szorzat „formailag” nem egyenlő a direkt összeggel. Mégis elég szoros kapcsolat áll fenn közöttük:

2.15. Tétel. *Véges sok komponens esetén a direkt összeg és a direkt szorzat izomorfak.*

Végtelen sok komponens esetén a direkt összeg izomorf a direkt szorzat egy valódi alterével.

Bizonyítás. Legyen a K test feletti \mathcal{V} vektortér az \mathcal{U}_i vektortereinek a direkt összege, és legyen $\underline{\mathcal{V}} = \prod_i \mathcal{U}_i$. Defináljuk a $\varphi : \mathcal{V} \rightarrow \underline{\mathcal{V}}$ leképezést a következőképpen:

Ha $\mathbf{v} = \sum_i \mathbf{u}_i$, ahol $\mathbf{u}_i \in \mathcal{U}_i$, akkor legyen $\varphi(\mathbf{v}) = \underline{\mathbf{v}} = (\mathbf{u}_1, \dots, \mathbf{u}_r, \dots)$.

Ha adott egy $\mathbf{v}' = \sum_i \mathbf{u}'_i$ elem is, ahol $\mathbf{u}'_i \in \mathcal{U}_i$, valamint adottak a K -beli c, d elemek is, akkor a direkt összeg tulajdonságai szerint $\mathbf{w} = c \cdot \mathbf{v} + d \cdot \mathbf{v}' = \sum_i (c \cdot \mathbf{u}_i + d \cdot \mathbf{u}'_i)$, ahol

$c \cdot \mathbf{u}_i + d \cdot \mathbf{u}'_i \in \mathcal{U}_i$. Ekkor, a direkt szorzatban értelmezett műveletek alapján

$$\begin{aligned}\varphi(\mathbf{w}) &= (c \cdot \mathbf{u}_1 + d \cdot \mathbf{u}'_1, \dots, c \cdot \mathbf{u}_i + d \cdot \mathbf{u}'_i, \dots) = \\ &= c \cdot (\mathbf{u}_1, \dots, \mathbf{u}_i, \dots) + d \cdot (\mathbf{u}'_1, \dots, \mathbf{u}'_i, \dots) = c \cdot \varphi(\mathbf{v}) + d \cdot \varphi(\mathbf{v}'),\end{aligned}$$

azaz φ művelettartó.

Tekintettel arra, hogy a direkt összeg egy vektorában majdnem minden tag a nullvektor (azaz csak véges sok tag különbözik a nullvektortól), ezért a direkt szorzatban csak olyan vektorok fordulhatnak elő képként, amelyekben majdnem minden komponens a nullvektor (azaz csak véges sok komponens különbözik a nullvektortól). Mivel az \mathcal{U}_i alterek generálják \mathcal{V} -t, ezért minden ilyen vektort megkapunk. Az is világos, hogy az ilyen vektorok a \mathcal{V} vektortérnek egy \mathcal{V}^* alterét alkotják. Eszerint φ tulajdonképpen a \mathcal{V}^* -ra képezi le a \mathcal{V} -t. Ebben az értelemben a $\varphi : \mathcal{V} \rightarrow \mathcal{V}^*$ szürjektív és művelettartó. Ahhoz, hogy ez éppen a kívánt izomorfizmus, azt kell még belátni, hogy φ injektív. Tekintsük evégett a $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$ elemeket, amelyekre $\varphi(\mathbf{v}) = \varphi(\mathbf{v}')$. A művelettartás alapján ez azt jelenti, hogy a $\mathbf{w} = \mathbf{v} - \mathbf{v}'$ elemre $\varphi(\mathbf{w}) = \mathbf{0}$ teljesül. Ennek a vektornak az i -edik komponense \mathbf{o}_i , ami azt jelenti, hogy \mathbf{w} a megfelelő alterekbe eső nullvektorok összege, azaz $\mathbf{w} = \mathbf{0}$; hiszen a direkt összegben a felbontás egyértelmű. Eszerint $\mathbf{v}' = \mathbf{v}$, azaz a leképezés valóban izomorfizmus.

Azt kell még belátni, hogy véges sok vektortér direkt szorzata izomorf ezek direkt összegével. Ennek azonban ebben a megfogalmazásban nincs értelme. Ugyanis a direkt összegben szereplő vektorterek egyetlen vektortérnek az alterei; míg a direkt szorzat komponenseiről semmi ilyet nem tehetünk fel. Éppen ezért az állítást célszerű a következőképpen átfogalmazni:

Ha adottak a K feletti $\mathcal{U}_1, \dots, \mathcal{U}_r$ vektorterek, akkor ezek $\mathcal{U} = \mathcal{U}_1 \times \dots \times \mathcal{U}_r$ szorzatához található olyan \mathcal{V} , K feletti \mathcal{U} -val izomorf vektortér és ennek olyan $\mathcal{V}_1, \dots, \mathcal{V}_r$ alterei, amelyeknek \mathcal{V} a direkt összege, továbbá minden i indexre \mathcal{U}_i és \mathcal{V}_i izomorfak.

Az ennek megfelelő állítást végtelen sok vektortérre bizonyítjuk, azzal a megkötéssel, hogy az \mathcal{U} direkt szorzat helyett azt az \mathcal{U}^* alteret tekintjük, amely azokból a sorozatokból áll, amelyekben majdnem minden komponens a nullvektor. (Tudjuk, hogy véges sok vektortér esetében $\mathcal{U}^* = \mathcal{U}$.)

A bizonyításnál legyen egyszerűen $\mathcal{V} = \mathcal{U}^*$. A \mathcal{V}_i altér álljon azokból a sorozatokból, amelyekben az i -edik helyen bármi állhat, a többi helyen pedig a nullvektor. Az $\mathbf{u}_i \mapsto (\mathbf{o}_1, \dots, \mathbf{o}_{i-1}, \mathbf{u}_i, \mathbf{o}_{i+1}, \dots)$ megfeleltetés nyilván izomorfizmus és a jobb oldali sorozatok rögzített i indexre \mathcal{U}^* -nak alterét alkotják. Mivel ilyen sorozatok összegeként minden olyan sorozat (egyértelműen) előáll, amelyben majdnem minden elem a nullvektor, ezért teljesül a direkt összeg tulajdonság. ■

A direkt szorzat nem lehet kommutatív, legalábbis úgy nem, mint a direkt összeg. Hiszen az $\mathcal{U} \times \mathcal{V}$ elemei (\mathbf{u}, \mathbf{v}) alakú párok, míg a $\mathcal{V} \times \mathcal{U}$ elemei (\mathbf{v}, \mathbf{u}) alakúak ($\mathbf{u} \in \mathcal{U}$, $\mathbf{v} \in \mathcal{V}$). Algebrai struktúrák — így vektorterek — esetében az egyenlőség helyett elegendő az izomorfizmust tekinteni. (Ezt a 2.15. Tétel bizonyításánál is láthattuk.)

2.16. Tétel. *A direkt szorzat kommutatív és asszociatív, azaz:*

$$\mathcal{A} \times \mathcal{B} \cong \mathcal{B} \times \mathcal{A} \quad \text{és} \quad (\mathcal{A} \times \mathcal{B}) \times \mathcal{C} \cong \mathcal{A} \times (\mathcal{B} \times \mathcal{C}),$$

ahol $\mathcal{A}, \mathcal{B}, \mathcal{C}$ vektorterek a K test felett. (A tétel modulusokra is igaz.)

Bizonyítás. A két izomorfizmust létrehozó leképezés megadása kézenfekvő:

$$(\mathbf{a}, \mathbf{b}) \mapsto (\mathbf{b}, \mathbf{a}) \quad \text{és} \quad ((\mathbf{a}, \mathbf{b}), \mathbf{c}) \mapsto (\mathbf{a}, (\mathbf{b}, \mathbf{c})),$$

ahol $\mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}, \mathbf{c} \in \mathcal{C}$.

A megfeleltetés nyilvánvalóan bijektív és a művelettartás bizonyítása is triviális. ■

Feladatok

1. Legyen $\mathcal{V} = \mathcal{A} \oplus \mathcal{B}$. Igaz-e az, hogy \mathcal{V} minden bázisa előáll \mathcal{A} és \mathcal{B} egy-egy bázisának egyesítéseként?

2. Bizonyítsuk be, hogy ha egy vektortér *direkt felbonthatatlan*, azaz nem egyelemű és csak triviális módon írható fel direkt összegként, akkor 1-dimenziós; a K feletti direkt felbonthatatlan vektorterek mind izomorfak.

3. Bizonyítsuk be, hogy az alábbi \mathbb{Z} -modulusok mind direkt felbonthatatlanok, de nem izomorfak: \mathbb{Z}_p ($p \in \mathbb{N}$ prímszám), valamint \mathbb{Q} .

4. Igaz-e az, hogy ha egy R integritási tartományra minden direkt felbonthatatlan R -modulus izomorf, akkor R test?

5. Bizonyítsuk be, hogy egy K test feletti legalább kételemű vektortér felírható egydimenziós altérinek direkt összegeként.

6. Legyen $\{\mathbf{u}_n \mid n \in \mathbb{N}\}$ az \mathcal{U} vektortér egy bázisa. Bizonyítsuk be, hogy \mathcal{U} -nak vannak olyan \mathcal{U}_k ($k \in \mathbb{N}$) alterei (végtelen sok!), amelyekre $\bigoplus_{i=1}^{\infty} \mathcal{U}_i = \mathcal{U}$ mellett minden egyes k -ra $\mathcal{U}_k \cong \mathcal{U}$ teljesül.

7. Legyen $\{\mathbf{u}_n \mid n \in \mathbb{N}\}$ az \mathcal{U} vektortér egy bázisa, és legyen $\mathcal{V}_i = \langle \mathbf{u}_i \rangle$. Ekkor természetesen $\mathcal{U} = \bigoplus_{i=1}^{\infty} \mathcal{V}_i$. Bizonyítsuk be, hogy $\mathcal{V} = \prod_{i=1}^{\infty} \mathcal{V}_i \not\cong \mathcal{U}$.

8. A vektorterekre adott feladatok sorában láttuk, hogy a pozitív racionális számok modulust alkotnak \mathbb{Z} felett, ha a vektorösszeadás a szorzás és az n egész számmal való szorzás az n -edik hatvány. Bizonyítsuk be, hogy rögzített p prímszám egész kitevőjű hatványai is \mathbb{Z} -modulust alkotnak ezekre a műveletekre. Bizonyítsuk be, hogy az előbbi ez utóbbiaknak a direkt összege.

9. A vektorterekre adott feladatok sorában láttuk, hogy a nemnulla komplex számoknak bizonyos részalmodulust alkotnak \mathbb{Z} felett, ha a vektorösszeadás a szorzás és az n egész számmal való szorzás az n -edik hatvány.

1. Bizonyítsuk be, hogy az összes nemnulla komplex számokból álló számhalmaz mint \mathbb{Z} -modulus direkt összege az 1 abszolút értékűeknek és a pozitív valósaknak.
2. Bizonyítsuk be, hogy az egységgyökök \mathbb{Z} -modulusa (végtelen) direkt összege a p -hatványadik egységgyököknek, rögzített p prímszámokra. (Ezek olyan ε komplex számok, amelyekhez található olyan n természetes szám, hogy $\varepsilon^{(p^n)} = 1$.)

HARMADIK FEJEZET

LINEÁRIS LEKÉPEZÉSEK

1. Homogén lineáris leképezések értelmezése

Bármiféle matematikai struktúránál központi szerepet játszanak a „struktúratartó” leképezések (függvények). Algebrai struktúrák esetében a struktúratartás azt jelenti, hogy a leképezés művelettartó, azaz esetünkben összeget és skalárszorost tart. Ilyennel már találkoztunk itt is: ilyenek voltak az izomorfizmusok.

A lineáris algebraban a művelettartó leképezések különösen fontosak. Ezek adják meg ugyanis az elvi háttérrel a mátrixokkal való „manipulációkhoz” (mesterkedésekhez). A lineáris leképezésekre vonatkozó eredmények, annak ellenére, hogy egyszerűeknek látszanak, távolról sem azok. Ezeknek a bonyolultságát érzékeltethetjük azzal, hogy vektorterek helyett — időnként — modulusokat is nézünk.

A mátrixok esetében azok elemei egy rögzített „értelmezési tartomány”-ból kerültek ki; ennek megfelelően az alábbiakban egy *rögzített* K test feletti vektortereket (illetve egy rögzített gyűrű fölötti modulusokat) fogunk nézni. A rögzített test általában vagy a valós, vagy a komplex számtest lesz. Annak megmutatására, hogy a kapott eredmények még vektorterek esetében sem triviálisan igazak, több esetben utalunk majd a racionális, illetve véges testek feletti vektorterekre.

Bevezetésként megemlíjtük még, hogy az „igazi” függvények „kicsiben” lineárisan viselkednek, ami különös fontosságot ad a lineáris függvények tárgyalásának.

3.1. Definíció. Az (adott K test feletti) \mathcal{U} vektortérnek a \mathcal{V} vektortérbe való $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ művelettartó leképezését (homogén) lineáris leképezésnek vagy vektortér-homomorfizmusnak (K -homomorfizmusnak) nevezzük.

A leképezésekhez hasonlóan itt is használjuk az $\mathcal{U} \xrightarrow{\varphi} \mathcal{V}$ jelölést.

Ha hangsúlyozni akarjuk, hogy mely test feletti vektorterekről van szó, vagy a test előzőleg még nem került szóba, akkor a φ_K , $\varphi_K : \mathcal{U} \rightarrow \mathcal{V}$, $\varphi_K : \mathcal{U}_K \rightarrow \mathcal{V}_K$, $\mathcal{U} \rightarrow {}_K\mathcal{V}$, valamint az $\mathcal{U} \xrightarrow{\varphi_K} \mathcal{V}$, $\mathcal{U}_K \xrightarrow{\varphi_K} \mathcal{V}_K$ jelöléseket is használjuk.

Modulusok esetében is használjuk a homogén lineáris leképezés elnevezést; de elsősorban modulus-homomorfizmusról (R -modulusok esetében R -homomorfizmusról) fogunk beszélni. ■

Megjegyzés. Lineáris leképezés (avagy függvény) például az $a \cdot x + b$ függvény, ahol a, b rögzített valós számok. Ez a polinomfüggvény nem homogén polinomhoz tartozik; ellentétben az $a \cdot x$ függvénnyel. Mi elsődlegesen *homogén* lineáris függvényekkel fogunk foglalkozni; a „homogén” jelzőt csak „rövidítésként” hagyjuk el.

Emlékeztetünk arra, hogy a művelettartás definíció szerint azt jelenti, hogy ha $\mathbf{a}, \mathbf{b} \in \mathcal{U}$, akkor $\varphi(\mathbf{a} + \mathbf{b}) = \varphi(\mathbf{a}) + \varphi(\mathbf{b})$, továbbá ha $c \in K$, akkor $\varphi(c \cdot \mathbf{a}) = c \cdot \varphi(\mathbf{a})$. □

3.1. Tétel. $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ pontosan akkor homogén lineáris, ha lineáris kombinációt tart.

Bizonyítás. Ha $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ művelettartó, akkor az összeadástartást, majd a skalárral való szorzástartást felhasználva

$$\varphi(c \cdot \mathbf{a} + d \cdot \mathbf{b}) = \varphi(c \cdot \mathbf{a}) + \varphi(d \cdot \mathbf{b}) = c \cdot \varphi(\mathbf{a}) + d \cdot \varphi(\mathbf{b})$$

adódik, ahol $\mathbf{a}, \mathbf{b} \in \mathcal{U}$ és $c, d \in K$.

Fordítva, ha $\varphi(c \cdot \mathbf{a} + d \cdot \mathbf{b}) = c \cdot \varphi(\mathbf{a}) + d \cdot \varphi(\mathbf{b})$, akkor $c = d = 1$ választással az összegtartást kapjuk; s ha $d = 0$, akkor azt kapjuk, hogy φ megőrzi a skalárral való szorzást. ■

Következmény. Homogén lineáris leképezés a nullvektort nullvektorba s egy vektor ellentettjét a kép ellentettjébe viszi.

Bizonyítás. $\varphi(\mathbf{u}) = \varphi(\mathbf{u} + \mathbf{o}) = \varphi(\mathbf{u}) + \varphi(\mathbf{o})$ bizonyítja az első állítást. A második abból következik, hogy ha $\mathbf{a} + \mathbf{b} = \mathbf{o}$, akkor $\varphi(\mathbf{a}) + \varphi(\mathbf{b}) = \varphi(\mathbf{o}) = \mathbf{o}$. ■

Az alábbiakban néhány példán megmutatjuk, hogy milyen sok esetben találkozhatunk lineáris leképezésekkel. Nem fogjuk bebizonyítani, hogy ezek valóban lineáris leképezések; a bizonyításokat feladatnak hagyjuk.

1. példa: Tekintsük a „geometriai síkon” az alábbi leképezéseket:

1. Egy origón átmenő egyenesre való merőleges vetítés.
2. Egy origón átmenő egyenesre való tükrözés.
3. Az origóra való tükrözés.
4. Az origó körüli forgatás.
5. Az origóból való arányos nyújtás (centrális hasonlóság).

(Megjegyezzük, hogy ha az origó helyett egy másik pontot választunk, akkor is lineáris leképezést kapunk, de nem homogén lineárist.)

2. példa: Vetítsük a geometriai (háromdimenziós) teret egy origón átmenő síkra vagy egyenesre; tükrözzük egy origón átmenő síkra vagy egyenesre; forgassuk a teret egy origón átmenő egyenes körül stb.

3. *példa:* Mátrixokból álló vektorteret (nem feltétlen az összes mátrixok vektorterét) képezzük le mátrixokból álló vektortérre úgy, hogy az eredeti mátrixnak elhagyjuk bizonyos sorait és/vagy oszlopait.

4. *példa:* Tekintsük egy rögzített K test felett a $K[x]$ polinomgyűrűt mint K feletti vektorteret. Ennek elemeit fogjuk leképezni.

1. f -nek feleltessük meg f' -t, az f deriváltját; leképezés önmagára (ha $1 + 1 \neq 0$).
2. Rögzített $g \in K[x]$ mellett $f \mapsto f \cdot g$; leképezés önmagára.
3. Rögzített n -edfokú $g \in K[x]$ mellett f -nek feleltessük meg a g -vel való osztás maradékát; leképezés az n -nél alacsonyabb fokú polinomok vektorterébe.
4. Rögzített $c \in K$ mellett $f \mapsto f(c)$; leképezés K -ba.

5. *példa:* Tetszőleges komplex számnak feleltessük meg a valós részét, illetve a képzetes részét ($\mathbb{C} \rightarrow \mathbb{R}$).

6. *példa:* Valós számokból álló konvergens sorozatoknak feleltessük meg a határértéküket (vektorterek \mathbb{R} felett).

7. *példa:* Adott intervallumban deriválható függvények \mathbb{R} feletti vektorterét képezzük le az összes függvények vektorterébe az $f \mapsto f'$ leképezéssel.

8. *példa:* Adott $[a, b]$ intervallumban integrálható függvények \mathbb{R} feletti vektorterét képezzük le \mathbb{R} -re az $f \mapsto \int_a^b f$ leképezéssel.

9. *példa:* Legyenek $a \leq b \leq c \leq d$ valós számok. $\mathcal{U}_{\mathbb{R}}$ az $[a, d]$ intervallumban és $\mathcal{V}_{\mathbb{R}}$ a $[b, c]$ intervallumban értelmezett valós függvények tere. Tetszőleges $f \in \mathcal{U}$ képe legyen az az $f_1 \in \mathcal{V}$ függvény, amelyre tetszőleges $b \leq \xi \leq c$ mellett $f_1(\xi) = f(\xi)$.

Mint a függvényeknél általában, a lineáris függvényeknél is lehetséges ezeket többféleképpen megadni. A függvényeknél — így a lineáris függvényeknél is — alapvető kíváncsiság, hogy ne függjenek a megadás módjától, csak attól, hogy „mit” és „hova” képeznek. Még egy — látszólag technikai — probléma is fellép. Ennek megvilágítására tekintsük az x^2 függvényt. Ez a függvény a valós számokat a valós számokba képezi le. Egyszersmind a racionális számokat is leképezi a racionális számokba. Ugyanaz-e ez a két függvény? Hatásukat tekintve igen, de formailag mégis célszerű megkülönböztetni őket. Teljesen eléget teszünk a „lényegnek”, ha azt mondjuk, hogy az utóbbi az előbbinek megszorítása a racionális számokra. Még jobban bonyolódik a helyzet, ha azt is figyelembe vesszük, hogy a függvény nemnegatív számokra képez. Itt is célszerű megkülönböztetni a két esetet aszerint, hogy az összes számot vagy csak a pozitívakat tekintjük „célterületnek”.

3.2. Definíció. Az $\alpha_K : \mathcal{A}_1 \rightarrow \mathcal{A}_2$ és $\beta_K : \mathcal{B}_1 \rightarrow \mathcal{B}_2$ (homogén lineáris) leképezéseket pontosan akkor tekintjük egyenlőknek, ha $\mathcal{A}_1 = \mathcal{B}_1$, $\mathcal{A}_2 = \mathcal{B}_2$; továbbá tetszőleges $\mathbf{u} \in \mathcal{A}$ elemre $\alpha(\mathbf{u}) = \beta(\mathbf{u})$. ■

Megjegyzések

1. A definíció „jogosságához” be kell látni, hogy itt az értelmezett egyenlőség rendelkezik az egyenlőség szokásos és természetes tulajdonságaival, azaz ekvivalenciareláció. Ennek a ténynek az egyszerű belátását az olvasóra bízuk.

2. Már találkoztunk az izomorfizmussal, ami nem más, mint egy bijektív homogén lineáris leképezés. Természetesen — mint általában is — fontos szerepet töltenek be az injektív és szürjektív homogén lineáris leképezések. Valójában már ezekre is láttunk példákat.

Ha $\mathcal{U} \leq \mathcal{V}$, akkor — mint említettük — az \mathcal{U} -beli műveletek tulajdonképpen nem azonosak a \mathcal{V} -beliekkel, hanem azok megszorításai. Ennek megfelelően úgy tekinthetjük a helyzetet, mintha \mathcal{U} és \mathcal{V} „független” vektorterek lennének egy $\mu : \mathcal{U} \rightarrow \mathcal{V}$ *beágyazással*, azaz egy $\mu(\mathbf{u}) \mapsto \mathbf{u}$, nyilvánvalóan injektív homogén lineáris leképezéssel.

Ha $\mathcal{W} = \mathcal{V}/\mathcal{U}$, akkor a $\nu : \mathbf{v} \rightarrow \mathbf{V}$ megfeleltetés, ahol $\mathbf{v} \in \mathbf{V}$, szürjektív. Erről is beláttuk, hogy művelettartó, tehát ez egy szürjektív homogén lineáris leképezés.

Mindkét esetben „egyöntetűen” adtuk meg a leképezést, a kép „természetesen” adódott. Ilyen esetekben *természetes leképezésekről* beszélünk. A természetes leképezést lehet egész pontosan definiálni, de ehhez igen sok fogalomra van szükség. *Vektorterek vagy modulusok esetében akkor monduk egy leképezést természetesnek, ha bázis felhasználása nélkül adható meg.* A fenti leképezések ilyenek. Nem minden leképezés ilyen. Tekintsük például az $\mathbf{u}, \mathbf{v}, \mathbf{w}$ bázissal rendelkező \mathcal{V} vektorteret és ennek $\mathcal{W} = \langle \mathbf{w} \rangle$ alterét. Defináljuk a $\varphi : \mathcal{V} \rightarrow \mathcal{V}/\mathcal{W}$ leképezést a következőképpen: $\varphi(a \cdot \mathbf{u} + b \cdot \mathbf{v} + c \cdot \mathbf{w}) = a \cdot \mathbf{V} + b \cdot \mathbf{U}$, ahol \mathbf{U} a \mathcal{W} -nek \mathbf{u} -val és \mathbf{V} a \mathcal{W} -nek \mathbf{v} -vel való eltoltja. (Tehát tulajdonképpen két bázisvektort „felcserélünk”.) Ez nem természetes leképezés. (Ezt majd akkor tudjuk belátni, ha a „természetesség” fogalmát kicsit pontosabban tudjuk definiálni.) \square

Feladatok

1. Legyenek S, R egységelemes integritási tartományok és legyen $\phi : S \rightarrow R$ gyűrűhomomorfizmus. Bizonyítsuk be, hogy minden R -homomorfizmus tekinthető S -homomorfizmusnak.

2. Az előző feladatnál legyen $S = \mathbb{Z}$. Ez azt jelenti, hogy minden modulus \mathbb{Z} -modulus. A vektortér-axiómák közül mi fejezi ki azt, hogy a modulus \mathbb{Z} -modulus?

3. Adjunk meg „természetes” $\pi_j : \prod_{i=1}^n \mathcal{A}_i \rightarrow \mathcal{A}_j$ szürjektív homogén lineáris leképezéseket.

2. Lineáris leképezések elemi tulajdonságai

Mint említettük, a lineáris leképezések elméleti hátteret adnak a „mátrixmanipulációknak”. Ezeknek a kézzelfogható formába öntése előtt szükségünk van a lineáris leképezések „jó” megismerésére.

Az előzőekben találkoztunk már bijektív, injektív és szürjektív lineáris leképezésekkel. A legtöbb lineáris leképezés természetesen nem ilyen. (Miért is volna?) Szükséges azonban tudni, hogy egy lineáris leképezés „mennyire” tér el ezektől az esetektől. Az injektivitást az jellemzi, hogy különböző vektoroknak a képe is különböző. Az ettől való eltérést az jellemzi, hogy több vektornak is lehet a képe ugyanaz. Vektorterek (és modulusok) esetében ez az eset könnyen jellemezhető:

3.2. Tétel. *A φ lineáris leképezésnél két vektornak pontosan akkor ugyanaz a képe, ha különbségük képe a nullvektor.*

Bizonyítás. Legyen $\varphi_K : \mathcal{U} \rightarrow \mathcal{V}$, és legyenek $\mathbf{a}, \mathbf{b} \in \mathcal{U}$. A művelettartás alapján $\varphi(\mathbf{a} - \mathbf{b}) = \varphi(\mathbf{a}) - \varphi(\mathbf{b})$. A két vektornak pontosan akkor ugyanaz a képe, ha a jobb oldalon a \mathcal{V} nullvektora áll; azaz, ha a két vektor különbségének a képe a \mathcal{V} nullvektora. ■

A szürjektivitástól való eltérést az adja meg, hogy a „képtérben” mely vektorok nem lépnak fel. Ezt és az előző tételt használja fel az alábbi igen alapvető

3.3. Definíció. Legyen $\varphi : \mathcal{U} \rightarrow_K \mathcal{V}$. Jelölje $\text{Ker}(\varphi)$ azoknak az \mathcal{U} -beli vektoroknak a halmazát, amelyeket φ a $\mathbf{o}_{\mathcal{V}}$ -re képez le. Ezt a halmazt a φ magjának nevezzük. Jelölje $\text{Im}(\varphi)$ azoknak a \mathcal{V} -beli vektoroknak a halmazát, amelyekre φ leképez \mathcal{U} -beli vektort. Ezt a halmazt a φ képének nevezzük. Ezek formális definíciója tehát:

$$\text{Ker}(\varphi) = \{\mathbf{u} \in \mathcal{U} \mid \varphi(\mathbf{u}) = \mathbf{o}_{\mathcal{V}}\} \quad \text{és} \quad \text{Im}(\varphi) = \{\varphi(\mathbf{u}) \in \mathcal{V} \mid \mathbf{u} \in \mathcal{U}\}. \quad \blacksquare$$

Az alábbi tétel a mag és a kép alapvető tulajdonságát mondja ki:

3.3. Tétel. *Legyen $\varphi : \mathcal{U} \rightarrow_K \mathcal{V}$. Ekkor $\text{Ker}(\varphi) \leq \mathcal{U}$ és $\text{Im}(\varphi) \leq \mathcal{V}$. Ennek megfelelően a magtér és képtér elnevezést fogjuk használni.*

Bizonyítás. Legyen $\mathbf{a}, \mathbf{b} \in \text{Ker}(\varphi)$. Ekkor tetszőleges $a, b \in K$ esetén:

$$\varphi(a \cdot \mathbf{a} + b \cdot \mathbf{b}) = a \cdot \varphi(\mathbf{a}) + b \cdot \varphi(\mathbf{b}) = a \cdot \mathbf{o}_{\mathcal{V}} + b \cdot \mathbf{o}_{\mathcal{V}} = \mathbf{o}_{\mathcal{V}},$$

vagyis $\text{Ker}(\varphi)$ zárt a lineáris kombináció képzésére. Mivel tartalmazza a $\mathbf{o}_{\mathcal{U}}$ vektort, ezért nem üres, tehát altér. Tekintsük $\text{Im}(\varphi)$ két elemét, ezek definíció szerint $\varphi(\mathbf{a})$ és $\varphi(\mathbf{b})$ alakúak, ahol $\mathbf{a}, \mathbf{b} \in \mathcal{U}$. Mivel φ lineáris kombinációt tart, ezért a K test a, b elemeire

$$a \cdot \varphi(\mathbf{a}) + b \cdot \varphi(\mathbf{b}) = \varphi(a \cdot \mathbf{a} + b \cdot \mathbf{b}),$$

vagyis $\text{Im}(\varphi) \leq \mathcal{V}$, hiszen nem üres. ■

A következő tétel egy általános algebrai tételnek a vektorterekre vonatkozó speciális esete. Ez a tétel megmutatja, hogy a magtér és a képtér együttese „méri” az izomorfizmustól való eltérést.

3.4. Tétel. *Ha $\varphi : \mathcal{U} \rightarrow \mathcal{V}$, akkor $\text{Im}(\varphi) \cong \mathcal{U} / \text{Ker}(\varphi)$.*

Bizonyítás. Jelölje \mathcal{W} a φ magterét. Az $\text{Im}(\varphi)$ tetszőleges \mathbf{v} eleme feltétel szerint $\mathbf{v} = \varphi(\mathbf{u})$ alakú, ahol $\mathbf{u} \in \mathcal{U}$. Feleltessük meg ennek a \mathbf{v} elemnek azt a \mathcal{W} szerinti \mathbf{u} eltoltat, amelyik az \mathbf{u} elemet tartalmazza. Ha $\varphi(\mathbf{u}') = \mathbf{v}$ ugyancsak fennáll, akkor a 3.2. Tétel szerint $\mathbf{u}' - \mathbf{u} \in \text{Ker}(\varphi)$, és így mindketten ugyanabban a mellékosztályban vannak. Eszerint $\text{Im}(\varphi)$ minden elemének egyértelműen feleltettünk meg egy eltoltat.

Ha $\mathbf{v}' = \varphi(\mathbf{u}')$ és \mathbf{u} , valamint \mathbf{u}' ugyanabban az eltoltban vannak, akkor $\mathbf{u}' - \mathbf{u} \in \mathcal{W}$, és így $\mathbf{v}' = \varphi(\mathbf{u}') = \varphi(\mathbf{u}) = \mathbf{v}$; tehát a megfeleltetés injektív.

Mivel \mathcal{W} -nek minden eltoltja tartalmazza \mathcal{U} egy elemét, amit φ a \mathcal{V} -be képez, ezért a megfeleltetés szürjektív.

Legyenek U_1 és U_2 a \mathcal{W} -nek u_1 -gyel és u_2 -vel való eltoltjai. Ekkor tetszőleges $c, d \in K$ esetén

$$\varphi(c \cdot u_1 + d \cdot u_2) = c \cdot \varphi(u_1) + d \cdot \varphi(u_2) \quad \text{és} \quad c \cdot u_1 + d \cdot u_2 \in c \cdot U_1 + d \cdot U_2$$

bizonyítja a művelettartást. ■

Megjegyzés. Az első lépés itt az volt, hogy megmutattuk, hogy tényleg egy leképezést definiáltunk. Erre azért volt szükség, mert több elemnek is ugyanaz a φ -nél vett képe. Eleve nem tudhatjuk, hogy ezek mindegyike ugyanabban az eltoltban van-e. □

Ennek a tételnek fontos „mérhető” következménye:

3.5. Tétel. *Legyen $\varphi : \mathcal{U} \rightarrow \mathcal{V}$. Ha \mathcal{U} véges dimenziós, akkor*

$$\dim(\text{Ker}(\varphi)) + \dim(\text{Im}(\varphi)) = \dim(\mathcal{U}).$$

Bizonyítás. Ha \mathcal{U} véges dimenziós, akkor minden altere és faktortere is az. Mivel véges dimenziós vektorterek izomorfizmusából következik, hogy dimenziójuk megegyezik, ezért $\dim(\text{Im}(\varphi)) = \dim(\mathcal{U}/\mathcal{W})$, ahol $\mathcal{W} = \text{Ker}(\varphi)$. A kívánt egyenlőség bizonyításához tehát elég azt belátni, hogy $\dim(\mathcal{U}/\mathcal{W}) = \dim(\mathcal{U}) - \dim(\mathcal{W})$. Ezt tetszőleges véges dimenziós \mathcal{U} térre és ennek tetszőleges \mathcal{W} alterére igazoljuk (ez volt a faktorterekre vonatkozó 1. feladat). A faktorterekre vonatkozó 2.4. Tétel szerint \mathcal{W} minden w_1, \dots, w_r bázisa kiegészíthető \mathcal{U} egy $w_1, \dots, w_r, u_1, \dots, u_s$ bázisává. Ez generátorrendszer, tehát a 2.11. Tétel szerint a \mathcal{W} -nek az ezeket tartalmazó eltoltjai a faktortérnek egy generátorrendszerét alkotják. Mivel a w_1, \dots, w_r elemek mind \mathcal{W} -ben vannak, ezért az u_i elemeket tartalmazó U_1, \dots, U_s eltoltak is generátorrendszert alkotnak. Ha ezeknek egy nemtriviális lineáris kombinációja a faktortér nullvektora lenne, akkor az u_1, \dots, u_s vektorok megfelelő lineáris kombinációja eleme volna \mathcal{W} -nek, ami a $w_1, \dots, w_r, u_1, \dots, u_s$ vektorok függetlensége alapján lehetetlen. ■

Megjegyzés. Érdemes figyelni arra, hogy a tételben megadott formula nem tartalmazza kifejezetten a \mathcal{V} vektorteret. Ez csak annyiban játszik szerepet, hogy a képtér dimenziója nem lehet nagyobb a \mathcal{V} dimenziójánál. □

3.4. Definíció. Egy φ lineáris leképezés rangjának nevezzük az $r(\varphi) = \dim(\text{Im}(\varphi))$ számot. ■

3.6. Tétel. $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ akkor és csak akkor injektív, ha $\text{Ker}(\varphi) = \{\mathbf{o}_{\mathcal{U}}\}$ és akkor és csak akkor szürjektív, ha $\text{Im}(\varphi) = \mathcal{V}$.

Bizonyítás. Az első állítás azonnal következik a 3.2. Tételből; míg a második a szürjektivitás definíciójából. ■

3.7. Tétel. *Legyen $\mathcal{U} = \prod_{i=1}^r \mathcal{A}_i$. Legyen $\pi_i : \mathcal{U} \rightarrow \mathcal{A}_i$ az i -edik projekció, azaz*

legyen $\pi_i : (a_1, \dots, a_r) \mapsto a_i$. Ez egy szürjektív homomorfizmus, amelynél $\text{Ker}(\pi_1) \wedge \dots \wedge \text{Ker}(\pi_r) = \{\mathbf{o}_{\mathcal{U}}\}$. Ha \mathcal{U}_j jelöli az összes j -től különböző $\text{Ker}(\pi_i)$ közös részét, akkor $\mathcal{U} = \bigoplus_{i=1}^r \mathcal{U}_i$.

Fordítva, ha léteznek olyan $\sigma_i : \mathcal{V} \rightarrow \mathcal{A}_i$ szürjektív homomorfizmusok, amelyekre teljesül $\text{Ker}(\sigma_1) \wedge \dots \wedge \text{Ker}(\sigma_r) = \{\mathbf{o}_{\mathcal{V}}\}$, akkor létezik egy (természetes) $\mathcal{V} \rightarrow \mathcal{U}$ injektív homomorfizmus.

Bizonyítás. A π_i leképezések mindegyike szürjektív homomorfizmus, a direkt szorzat definíciója szerint. $\mathbf{u} = (\mathbf{a}_1, \dots, \mathbf{a}_r)$ pontosan akkor eleme $\text{Ker}(\pi_i)$ -nek, ha minden $\mathbf{a}_i = \mathbf{o}_i$. Ha ez minden i -re teljesül, akkor \mathbf{u} minden komponense nullvektor, tehát $\mathbf{u} = \mathbf{o}$. Definíció szerint \mathcal{U}_j -ben azok a sorozatok vannak, amelyekben $i \neq j$ esetén $\mathbf{a}_i = \mathbf{o}_i$; azaz azok a sorozatok, amelyekben a j -edik helyen bármi áll, a többi helyen pedig a nullvektor. Ezeknek az \mathcal{U} — mint láttuk — direkt összege.

Tekintsük most az adott σ_i szürjektív homomorfizmusokat. Tetszőleges $\mathbf{v} \in \mathcal{V}$ vektornak feleltessük meg az $\mathbf{u} = (\sigma_1(\mathbf{v}), \dots, \sigma_r(\mathbf{v}))$ vektort. Ez a megfeleltetés nyilván homomorfizmus. (Általában nem szürjektív!) E homomorfizmus magja azokból az elemekből áll, amelyek minden \mathcal{A}_i -ben a nullvektorra képeződnek le, azaz amelyek minden σ_i magjában benne vannak. Feltételünk szerint ezek közös része egyedül a nullvektorból áll. A 3.6. Tétel szerint tehát ez a megfeleltetés valóban injektív. ■

A következő tétel az interpolációhoz hasonló tulajdonságot ír le. Ez a tétel nyújt segítséget ahhoz, hogy a leképezéseket mátrixokkal tudjuk megjeleníteni.

3.8. Tétel. Legyen $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} vektortér egy generátorrendszere és legyenek $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathcal{V}$ tetszőleges vektorok. Ekkor legfeljebb egy olyan $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ lineáris leképezés van, amelynél $\varphi(\mathbf{u}_i) = \mathbf{v}_i$. Ha az \mathbf{U} rendszer lineárisan független, akkor létezik is ilyen leképezés.

Bizonyítás. Mivel \mathbf{U} generátorrendszer, ezért az \mathcal{U} vektortér bármely \mathbf{u} vektora felírható $\mathbf{u} = c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n$ alakban. Mivel φ lineáris kombinációt tart, ezért \mathbf{u} képe csak $\mathbf{v} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n$ lehet. Ha \mathbf{U} bázis, akkor a felírás egyértelmű; és így a kép is az. A lineáris kombinációkra vonatkozó elemi összefüggések alapján ez a megfeleltetés művelet-tartó. ■

Megjegyzés. Könnyen belátható, hogy a kérdéses lineáris leképezés létezése azon múlik, hogy \mathbf{U} független. Ekkor ugyanis kiegészíthető bázissá; és a fenti tétel alapján valóban létezik a kívánt tulajdonsággal rendelkező leképezés. Ha \mathbf{U} nem bázis, akkor több lehetőségünk van a „kiterjesztésre”, és így több ilyen leképezés is létezik. □

3.5. Definíció. Legyen $\mathcal{U}_1 \leq \mathcal{U}$ és $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ egy lineáris leképezés. A $\varphi_1 : \mathcal{U}_1 \rightarrow \mathcal{V}$ leképezést a φ megszorításának nevezzük, ha minden $\mathbf{u} \in \mathcal{U}_1$ esetén $\varphi_1(\mathbf{u}) = \varphi(\mathbf{u})$. Ekkor azt is mondjuk, hogy φ a φ_1 -nek kiterjesztése (\mathcal{U} -ra).

A megszorításra használatos a $\varphi_1 = \varphi \upharpoonright \mathcal{U}_1$ jelölés. ■

3.9. Tétel. Ha $\mathcal{U}_1 \leq \mathcal{U}$ és $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ egy lineáris leképezés és $\varphi_1 : \mathcal{U}_1 \rightarrow \mathcal{V}$ a megszorítása, akkor φ_1 is lineáris leképezés.

Legyen $\mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2$ és φ_i a $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ megszorítása \mathcal{U}_i -re ($i \in \{1, 2\}$). Ha $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ ($\mathbf{u}_i \in \mathcal{U}_i$), akkor $\varphi(\mathbf{u}) = \varphi_1(\mathbf{u}_1) + \varphi_2(\mathbf{u}_2)$.

Fordítva, ha adottak a $\varphi_i : \mathcal{U}_i \rightarrow \mathcal{V}$ ($i \in \{1, 2\}$) leképezések, akkor létezik pontosan egy olyan $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezés, amelyre $\varphi(\mathbf{u}) = \varphi_1(\mathbf{u}_1) + \varphi_2(\mathbf{u}_2)$, ahol $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ és $\mathbf{u}_i \in \mathcal{U}_i$. Ennek a leképezésnek az \mathcal{U}_i -re való megszorítása a φ_i ; ezt a $\varphi = \varphi_1 \oplus \varphi_2$ leképezést a φ_1 és φ_2 direkt összegének nevezzük.

Bizonyítás. Az első állítás azonnal következik abból, hogy egy leképezés pontosan akkor lineáris, ha lineáris kombinációt tart.

A második állításban felírt egyenlőséghez a következőképpen jutunk:

$$\varphi(\mathbf{u}) = \varphi(\mathbf{u}_1) + \varphi(\mathbf{u}_2) = \varphi_1(\mathbf{u}_1) + \varphi_2(\mathbf{u}_2).$$

Ha a $\varphi_i : \mathcal{U}_i \rightarrow \mathcal{V}$ ($i \in \{1, 2\}$) leképezések adottak, akkor azt kell megmutatni, hogy a tételben definiált φ leképezés valóban \mathcal{V} -re képez és lineáris. Az, hogy $\varphi : \mathcal{U} \rightarrow \mathcal{V}$, azonnal következik a φ definiálásából. A lineáris kombináció tartásához legyen $\varphi(\mathbf{u}) = \varphi_1(\mathbf{u}_1) + \varphi_2(\mathbf{u}_2)$ és $\varphi(\mathbf{v}) = \varphi_1(\mathbf{v}_1) + \varphi_2(\mathbf{v}_2)$ ($\mathbf{u}_1, \mathbf{v}_1 \in \mathcal{U}_1$ és $\mathbf{u}_2, \mathbf{v}_2 \in \mathcal{U}_2$); továbbá c, d skalárok. Ekkor:

$$\begin{aligned} \varphi(c\mathbf{u} + d\mathbf{v}) &= \varphi((c\mathbf{u}_1 + d\mathbf{v}_1) + (c\mathbf{u}_2 + d\mathbf{v}_2)) = \varphi_1(c\mathbf{u}_1 + d\mathbf{v}_1) + \varphi_2(c\mathbf{u}_2 + d\mathbf{v}_2) = \\ &= c\varphi_1(\mathbf{u}_1) + d\varphi_1(\mathbf{v}_1) + c\varphi_2(\mathbf{u}_2) + d\varphi_2(\mathbf{v}_2) = c\varphi_1(\mathbf{u}_1) + c\varphi_2(\mathbf{u}_2) + d\varphi_1(\mathbf{v}_1) + d\varphi_2(\mathbf{v}_2) = \\ &= c(\varphi_1(\mathbf{u}_1) + \varphi_2(\mathbf{u}_2)) + d(\varphi_1(\mathbf{v}_1) + \varphi_2(\mathbf{v}_2)) = c\varphi(\mathbf{u}) + d\varphi(\mathbf{v}). \end{aligned} \quad \blacksquare$$

3.6. Definíció. Ha a 3.9. Tételben $\mathcal{V} = \mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2$ és φ_1 az identitás (azaz $\varphi_1(\mathbf{u}_1) = \mathbf{u}_1$ minden $\mathbf{u}_1 \in \mathcal{U}_1$ vektorra), és φ_2 a „null-leképezés” (azaz $\varphi_2(\mathbf{u}_2) = \mathbf{0}_2$ minden $\mathbf{u}_2 \in \mathcal{U}_2$ vektorra), akkor azt mondjuk, hogy direkt összegük az \mathcal{U}_1 -re való projekció, vagy vetítés. \blacksquare

Feladatok

1. Legyen \mathbf{U} az \mathcal{U} -nak egy n -elemű részhalmaza. Bizonyítsuk be, hogy ha bármely \mathcal{V} vektortérnek bármely n -elemű \mathbf{V} részhalmazához legfeljebb egy olyan $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ lineáris leképezés van, amely az adott $\mathbf{u}_i \in \mathbf{U}$ elemek mindegyikét az előírt $\mathbf{v}_i \in \mathbf{V}$ elembe viszi, akkor \mathbf{U} generátorrendszer. Ha viszont mindig van ilyen φ , akkor \mathbf{U} lineárisan független.

2. Bizonyítsuk be a 3.8. Tételt végtelen dimenziós vektorterekre.

3. Mutassuk meg, hogy egy modulus-homomorfizmus is megszorítható bármely részmodulusra, de egy részmodulusnak nem minden homomorfizmusa terjeszthető ki az egész modulus egy homomorfizmusává.

4. Legyen R egységelemes integritási tartomány. Bizonyítsuk be, hogy ha bármely \mathcal{M} R -modulus bármely \mathcal{M}_1 részmodulusának $\varphi_1 : \mathcal{M}_1 \rightarrow \mathcal{N}$ R -homomorfizmusa kiterjeszthető egy $\varphi : \mathcal{M} \rightarrow \mathcal{N}$ R -homomorfizmusává, akkor R test.

3. A lineáris leképezések tere

Az algebrai vizsgálatoknál alapvető követelmény az, hogy a vizsgált „tárgyakkal” műveleteket tudjunk végezni. Mint említettük, a lineáris leképezések adják a mátrixok elvi „háttérét”; és az itteni műveletek „magyarázzák meg” a mátrixműveleteket. A lineáris leképezésekkel végezhető műveletek viszont igen természetesen adódnak; ezek a függvényekre adott műveletekkel azonosak.

3.10. Tétel. *Tetszőleges $\varphi_K : \mathcal{U} \rightarrow \mathcal{V}$ lineáris leképezés és $c \in K$ esetén a $(c\varphi) : \mathbf{u} \mapsto c\mathbf{u}$ összefüggéssel értelmezett $c\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezés lineáris. Ezt a leképezést a φ leképezés c -szeresének nevezzük.*

Bizonyítás. $c\varphi : \mathcal{U} \rightarrow \mathcal{V}$ triviális. A lineáris kombináció tartása a következőképpen látható be. Legyenek $\mathbf{a}, \mathbf{b} \in \mathcal{U}$ és $a, b \in K$. Ekkor:

$$\begin{aligned} (c\varphi)(a\mathbf{a} + b\mathbf{b}) &= c \cdot (\varphi(a\mathbf{a} + b\mathbf{b})) = c \cdot (a\varphi(\mathbf{a}) + b\varphi(\mathbf{b})) = \\ &= ca \cdot \varphi(\mathbf{a}) + cb \cdot \varphi(\mathbf{b}) = ac \cdot \varphi(\mathbf{a}) + bc \cdot \varphi(\mathbf{b}) = a \cdot (c\varphi)(\mathbf{a}) + b \cdot (c\varphi)(\mathbf{b}). \quad \blacksquare \end{aligned}$$

A függvényösszeadás a lineáris leképezésekre is értelmezhető; és kiderül, hogy az eredmény ismét lineáris leképezés. Arra kell csak vigyázni, hogy az értelmezési tartományok is megegyezzenek és az értékkészlet-tartományok is.

3.11. Tétel. *Legyenek $\varphi, \psi : \mathcal{U} \rightarrow \mathcal{V}$ lineáris leképezések. Ezek $(\varphi + \psi) : \mathbf{u} \mapsto \varphi(\mathbf{u}) + \psi(\mathbf{u})$ definícióval adott összege egy $(\varphi + \psi) : \mathcal{U} \rightarrow \mathcal{V}$ lineáris leképezés.*

Bizonyítás. Itt is csupán a művelettartást kell ellenőrizni. Legyen, evégett, $\mathbf{a}, \mathbf{b} \in \mathcal{U}$ és $a, b \in K$. Ekkor:

$$\begin{aligned} (\varphi + \psi)(a\mathbf{a} + b\mathbf{b}) &= \varphi(a\mathbf{a} + b\mathbf{b}) + \psi(a\mathbf{a} + b\mathbf{b}) = \\ &= a\varphi(\mathbf{a}) + b\varphi(\mathbf{b}) + a\psi(\mathbf{a}) + b\psi(\mathbf{b}) = a((\varphi + \psi)(\mathbf{a})) + b((\varphi + \psi)(\mathbf{b})). \quad \blacksquare \end{aligned}$$

Megjegyzés. Ellentétben az alterek esetével a leképezések összege nem egy „gyengített formája” a direkt összegnek. A direkt összeg esetén a „tagok” egy vektortér (vagy modulus) altereit képezik le; míg az összeg esetén az egész teret. \square

3.12. Tétel. *Az \mathcal{U}_K vektorteret a \mathcal{V}_K vektortérbe vivő homogén lineáris leképezések K -vektorteret alkotnak a 3.10. és 3.11. Tételekben értelmezett műveletekre nézve. Ezt a vektorteret $\text{Hom}(\mathcal{U}, \mathcal{V})$, vagy precízebben $\text{Hom}_K(\mathcal{U}, \mathcal{V})$ jelöli. (R -modulusokra $\text{Hom}_R(\mathcal{M}, \mathcal{N})$ a jelölés. Ez csak kommutatív gyűrűk esetén R -modulus, egyébként kommutatív csoport.)*

Bizonyítás. A fent idézett tételek szerint a $\text{Hom}(\mathcal{U}, \mathcal{V})$ elemeire mind az összeadás, mind a skalárral való szorzás definiálva van; és az eredmény eleme $\text{Hom}(\mathcal{U}, \mathcal{V})$ -nek. Azt kell belátni, hogy teljesülnek a vektorterekre kimondott axiómák. Ezek között két olyan van, amelyik bizonyos elemek létezését mondja ki, a többiek azonosságok teljesülését

kívánják meg. Először ezeket bizonyítjuk. A leképezések egyenlőségének az értelmezése szerint azt kell megmutatni, hogy az azonosság bal és jobb oldalán álló leképezés tetszőlegesen választott vektort ugyanabba visz.

I. Az összeadásra vonatkozó axiómák:

- (1) Kommutativitás: Ha $\alpha, \beta \in \text{Hom}(\mathcal{U}, \mathcal{V})$ és $\mathbf{u} \in \mathcal{U}$, akkor $(\alpha + \beta)(\mathbf{u}) = \alpha(\mathbf{u}) + \beta(\mathbf{u}) = \beta(\mathbf{u}) + \alpha(\mathbf{u}) = (\beta + \alpha)(\mathbf{u})$.
- (2) Asszociativitás: Ha $\alpha, \beta, \gamma \in \text{Hom}(\mathcal{U}, \mathcal{V})$ és $\mathbf{u} \in \mathcal{U}$, akkor $((\alpha + \beta) + \gamma)(\mathbf{u}) = (\alpha(\mathbf{u}) + \beta(\mathbf{u})) + \gamma(\mathbf{u}) = \alpha(\mathbf{u}) + (\beta(\mathbf{u}) + \gamma(\mathbf{u})) = (\alpha + (\beta + \gamma))(\mathbf{u})$.

II. A skalárral való szorzás axiómái. Itt $\alpha, \beta \in \text{Hom}(\mathcal{U}, \mathcal{V})$ és $a, b, c \in K$:

- (1) $((a + b)\alpha)(\mathbf{u}) = (a + b)\alpha(\mathbf{u}) = a\alpha(\mathbf{u}) + b\alpha(\mathbf{u}) = (a\alpha + b\alpha)(\mathbf{u})$.
- (2) $(c(\alpha + \beta))(\mathbf{u}) = c(\alpha + \beta)(\mathbf{u}) = c\alpha(\mathbf{u}) + c\beta(\mathbf{u}) = (c\alpha + c\beta)(\mathbf{u})$.
- (3) $((ab)\alpha)(\mathbf{u}) = (ab)\alpha(\mathbf{u}) = a(b\alpha)(\mathbf{u}) = (a(b\alpha))(\mathbf{u})$.
- (4) $(1 \cdot \alpha)(\mathbf{u}) = 1 \cdot \alpha(\mathbf{u}) = \alpha(\mathbf{u})$.

A null-leképezésre és az ellentettre vonatkozó azonosságok előtt definiálni kell ezeket.

Legyen $\omega : \mathcal{U} \rightarrow \mathcal{V}$ az a leképezés, amelyre tetszőleges $\mathbf{u} \in \mathcal{U}$ esetén $\omega(\mathbf{u}) = \mathbf{o}_{\mathcal{V}}$. Ez a leképezés triviálisan lineáris és az 1.1. Definíció I. (3) tulajdonsága a következőképpen adódik:

$$(\alpha + \omega)(\mathbf{u}) = \alpha(\mathbf{u}) + \omega(\mathbf{u}) = \alpha(\mathbf{u}) + \mathbf{o}_{\mathcal{V}} = \alpha(\mathbf{u}).$$

Végezetül az α leképezés ellentettjét definiálja az a $\xi : \mathcal{U} \rightarrow \mathcal{V}$ leképezés, amelyre $\xi(\mathbf{u}) = -\alpha(\mathbf{u})$ tetszőleges $\mathbf{u} \in \mathcal{U}$ mellett. Ez, mint egy lineáris leképezés skalárszorosa, szintén lineáris leképezés. Most az 1.1. Definíció I. (4) azonosságát bizonyítjuk:

$$(\alpha + \xi)(\mathbf{u}) = \alpha(\mathbf{u}) + \xi(\mathbf{u}) = \alpha(\mathbf{u}) + (-1)\alpha(\mathbf{u}) = \mathbf{o}_{\mathcal{V}} = \omega(\mathcal{B}u). \blacksquare$$

Feladatok

1. Legyen $\mathcal{U} = \bigoplus_{i=1}^n \mathcal{U}_i$, $\varphi : \mathcal{U} \rightarrow \mathcal{V}$, $\varphi_i : \mathcal{U}_i \rightarrow \mathcal{V}$ olyanok, hogy $\varphi = \bigoplus_i \varphi_i$. Definiálja $\varphi^{(i)}$ -t $\varphi^{(i)} = (\bigoplus_{j \neq i} \omega_j) \oplus \varphi_i$. Bizonyítsuk be, hogy $\varphi = \varphi^{(1)} + \dots + \varphi^{(n)}$.

2. Bizonyítsuk be, hogy az \mathcal{U} vektortérre vonatkozó alábbi két állítás ekvivalens:

- a) A $\mathbf{v} \in \mathcal{U}$ lineárisan függ az $\mathbf{u}_1, \dots, \mathbf{u}_r \in \mathcal{U}$ elemektől.
- b) Ha $\varphi \in \text{Hom}(\mathcal{U}, \mathcal{U})$ lineáris leképezésre $\varphi(\mathbf{u}_1) = \dots = \varphi(\mathbf{u}_r) = \mathbf{o}$, akkor $\varphi(\mathbf{v}) = \mathbf{o}$ is igaz.

3. Bizonyítsuk be, hogy

$$r(\alpha + \beta) \leq \dim(\text{Im}(\alpha) \vee \text{Im}(\beta)) \leq r(\alpha) + r(\beta).$$

Mutassunk példát arra, hogy $r(\alpha + \beta) \neq r(\alpha) + r(\beta)$, noha $\text{Im}(\alpha) \wedge \text{Im}(\beta) = \{\mathbf{o}\}$.

4. Bizonyítsuk be, hogy ha $c \neq 0$ skalár, akkor $r(c\varphi) = r(\varphi)$.

5. Legyenek $\beta, \alpha_1, \dots, \alpha_r \in \text{Hom}(\mathcal{U}, \mathcal{V})$. Bizonyítsuk be, hogy β pontosan akkor függ az $\{\alpha_1, \dots, \alpha_r\}$ rendszertől, ha tetszőleges $\mathbf{u} \in \mathcal{U}$ esetén $\beta(\mathbf{u})$ az $\{\alpha_1(\mathbf{u}), \dots, \alpha_r(\mathbf{u})\}$ rendszernek ugyanazon együtthatókkal képezett lineáris kombinációja. (Mutassuk meg, hogy ezt elég \mathcal{U} egy bázisának az elemeire megkövetelni.)

4. Lineáris leképezések szorzása

A függvényekhez hasonlóan a lineáris leképezések szorzását is a kompozíciójuk definiálja. Tudjuk, hogy a kompozíció nem kommutatív; sőt az összeadásra nézve általában nem is disztributív. Itt viszont — hála a linearitásnak — teljesül a disztributivitás. Fellép viszont egy másik probléma, nevezetesen az, hogy a kompozíció nem mindig végezhető el. Ehhez az szükséges, hogy az első tényező „onnan” képezzen, „ahova” a második képez.

3.13. Tétel. Legyen $\psi_K : \mathcal{U} \rightarrow \mathcal{V}$ és $\varphi_K : \mathcal{V} \rightarrow \mathcal{W}$. Ekkor ezeknek a $\varphi\psi : \varphi(\psi(\mathbf{u}))$ ($\mathbf{u} \in \mathcal{U}$) összefüggéssel definiált szorzata is lineáris leképezés. Erre $\varphi\psi : \mathcal{U} \rightarrow \mathcal{W}$ igaz.

Bizonyítás. $\varphi\psi$ természetesen leképezés, a két leképezés kompozíciója, és az utolsó állítás is triviálisan teljesül. Még a művelettartást kell bizonyítani. Legyenek $\mathbf{a}, \mathbf{b} \in \mathcal{U}$ és $a, b \in K$. Ekkor:

$$\begin{aligned} (\varphi\psi)(a\mathbf{a} + b\mathbf{b}) &= \varphi(\psi(a\mathbf{a} + b\mathbf{b})) = \varphi(a\psi(\mathbf{a}) + b\psi(\mathbf{b})) = \\ &= a\varphi(\psi(\mathbf{a})) + b\varphi(\psi(\mathbf{b})) = a(\varphi\psi)(\mathbf{a}) + b(\varphi\psi)(\mathbf{b}). \end{aligned}$$

Megjegyzések

1. Általában tehát az $\mathcal{A} \xrightarrow{\alpha} \mathcal{B}$ és a $\mathcal{C} \xrightarrow{\beta} \mathcal{D}$ leképezések nem szorozhatók össze. (Legalábbis nem a kompozícióval.) Erre csak akkor van lehetőség, ha $\mathcal{C} = \mathcal{B}$ (vagy, ha $\mathcal{D} = \mathcal{A}$), tehát az $\mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{D}$ esetben. Ekkor a szorzatra $\mathcal{A} \xrightarrow{\beta\alpha} \mathcal{D}$ adódik. (A másik esetben az $\alpha\beta$ szorzat létezik.) Az egyik szorzat létezéséből nem következik a másik szorzat létezése.

2. Furcsának tűnhet, hogy a leképezések szorzásánál először a második tényezőt kell alkalmazni és utána az elsőt. Ennek az az oka, hogy a függvényjelölésnél elől áll a függvény és utána az, amire a függvényt alkalmazni kell. Sok esetben természetesebb az úgynevezett „lengyel jelölés”, amit első-sorban logikában alkalmaznak. A jelölésmód valószínűleg annak köszönhető, hogy az indoeurópai nyelvekben elől áll a birtok és utána a birtokos. A magyar nyelvben fordított a helyzet, számunkra tehát mindenképpen a fordított jelölés volna világosabb (számunkra $\sqrt{2}$ nem négyzetgyöke a 2-nek, hanem 2-nek a négyzetgyöke, vagy $\sin(\pi)$ nem szinusz a π -nek, hanem π -nek a szinusza). \square

3.14. Tétel. A lineáris leképezések szorzása asszociatív, az összeadásra nézve disztributív és érvényesek a $c(\alpha\beta) = (c\alpha)\beta = \alpha(c\beta)$ egyenlőségek, ahol $c \in K$ és az α, β K -homomorfizmusoknak létezik az $\alpha\beta$ szorzata.

Bizonyítás. Az asszociativitás tetszőleges leképezések szorzatára teljesül. Az, hogy az $(\alpha\beta)\gamma$ szorzat mindkét oldalán lineáris leképezés van, abból következik, hogy két lineáris leképezés szorzata is lineáris leképezés. (Könnyen látható, hogy a fenti három-tényezős szorzatok pontosan akkor léteznek, ha az $\alpha\beta$ és $\beta\gamma$ szorzatok mindegyike létezik.)

A disztributivitás bizonyítása előtt megjegyezzük, hogy itt kétféle disztributivitás van. Mint a polinomoknál láttuk, az egyik oldali disztributivitásból nem következik a másik oldali. Ezen túlmenően általában lehet, hogy az $(\alpha + \beta)\gamma$ szorzat létezik, míg a $\gamma(\alpha + \beta)$ szorzat nem. Ennek megfelelően kétféle bizonyításra van szükség. Noha a két bizonyítás

formailag igen hasonló, a lépések helyességének az oka a két bizonyításban más és más. Legyenek tehát adottak a következő K -homomorfizmusok: $\gamma : \mathcal{U} \rightarrow \mathcal{V}$, $\alpha, \beta : \mathcal{V} \rightarrow \mathcal{W}$ és $\delta : \mathcal{W} \rightarrow \mathcal{I}$. Azt fogjuk bizonyítani, hogy

$$(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma \quad \text{és} \quad \delta(\alpha + \beta) = \delta\alpha + \delta\beta.$$

A leképezések megadásából azonnal következik, hogy az egyenlőségek mindkét oldala értelmes; és ezek lineáris leképezések. A leképezések egyenlőségéhez azt kell megmutatni, hogy az első egyenlőség mindkét oldala ugyanabba a vektorba viszi az \mathcal{U} vektortér tetszőleges \mathbf{u} elemét; a második esetben ugyanezt kell bizonyítani tetszőleges $\mathbf{v} \in \mathcal{V}$ vektorra.

$$\begin{aligned} ((\alpha + \beta)\gamma)(\mathbf{u}) &= (\alpha + \beta)(\gamma(\mathbf{u})) = \alpha(\gamma(\mathbf{u})) + \beta(\gamma(\mathbf{u})) = \\ &= (\alpha\gamma)(\mathbf{u}) + (\beta\gamma)(\mathbf{u}) = (\alpha\gamma + \beta\gamma)(\mathbf{u}). \end{aligned}$$

$$\begin{aligned} (\delta(\alpha + \beta))(\mathbf{v}) &= \delta((\alpha + \beta)(\mathbf{v})) = \delta(\alpha(\mathbf{v}) + \beta(\mathbf{v})) = \\ &= \delta(\alpha(\mathbf{v})) + \delta(\beta(\mathbf{v})) = (\delta\alpha + \delta\beta)(\mathbf{v}). \end{aligned}$$

A tételbeli utolsó két egyenlőséghez feltétel szerint az szükséges, hogy $\beta : \mathcal{U} \rightarrow \mathcal{V}$ és $\alpha : \mathcal{V} \rightarrow \mathcal{W}$. Ekkor tetszőleges $\mathbf{u} \in \mathcal{U}$ esetén a kapott $(c(\alpha\beta))(\mathbf{u})$, $(c\alpha)(\beta(\mathbf{u}))$ és $\alpha((c\beta)(\mathbf{u}))$ vektorok mindegyike $c(\alpha(\beta(\mathbf{u})))$, tehát a leképezések valóban megegyeznek. ■

Kiegészítés. A szorzás a direkt összeggel felcserélhető, azaz

$$(\alpha_1 \oplus \alpha_2) \cdot (\beta_1 \oplus \beta_2) = \beta_1 \cdot \alpha_1 \oplus \beta_2 \cdot \alpha_2,$$

ha a fenti (részleges) műveletek elvégezhetők.

Bizonyítás. Legyenek $\mathcal{U}, \mathcal{V}, \mathcal{W}$ vektorterek egy adott K test felett, $\mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2$, $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$. Legyenek továbbá $\alpha_i : \mathcal{U}_i \rightarrow \mathcal{V}_i$ és $\beta_i : \mathcal{V}_i \rightarrow \mathcal{W}$ lineáris leképezések ($i \in \{1, 2\}$), továbbá $\alpha = \alpha_1 \oplus \alpha_2$ és $\beta = \beta_1 \oplus \beta_2$. Azt kell belátni, hogy

$$\beta \cdot \alpha = \beta_1 \cdot \alpha_1 \oplus \beta_2 \cdot \alpha_2.$$

Feltétel szerint tetszőleges \mathcal{U} -beli \mathbf{u} vektor $\mathbf{u}_1 + \mathbf{u}_2$ alakba írható, ahol $\mathbf{u}_1 \in \mathcal{U}_1$ és $\mathbf{u}_2 \in \mathcal{U}_2$. Ekkor

$$\begin{aligned} \beta\alpha(\mathbf{u}) &= \beta\alpha(\mathbf{u}_1 + \mathbf{u}_2) = \beta(\alpha_1(\mathbf{u}_1) + \alpha_2(\mathbf{u}_2)) = \beta_1(\alpha_1(\mathbf{u}_1)) + \beta_2(\alpha_2(\mathbf{u}_2)) = \\ &= \beta_1\alpha_1 \oplus \beta_2\alpha_2(\mathbf{u}_1 + \mathbf{u}_2) = \beta_1\alpha_1 \oplus \beta_2\alpha_2(\mathbf{u}). \end{aligned}$$

3.7. Definíció. Az \mathcal{U}_K vektortérnek azt az $\iota = \iota_{\mathcal{U}} : \mathcal{U} \rightarrow \mathcal{U}$ leképezését, amely minden vektort önmagára képez, az \mathcal{U} vektortér identitásának vagy identikus leképezésének nevezzük. ■

Megjegyezzük, hogy minden vektortérnél az identitás triviálisan homogén lineáris.

3.15. Tétel. Minden α lineáris leképezéshez található pontosan egy olyan ι_1 , illetve ι_2 identitás, amelyre $\iota_1\alpha = \alpha\iota_2 = \alpha$.

Tetszőleges ι identitásra érvényesek az $\alpha\iota = \alpha$ és $\iota\beta = \beta$ egyenlőségek, amennyiben a bal oldalon levő szorzások elvégezhetők. Ezen összefüggések közül bármelyik jellemzi az identitásokat.

Legyen $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ és $\omega_1 : \mathcal{U}_1 \rightarrow \mathcal{U}$, valamint $\omega_2 : \mathcal{V} \rightarrow \mathcal{V}_2$ null-leképezések. Ekkor $\alpha\omega_1 : \mathcal{U}_1 \rightarrow \mathcal{V}$ és $\omega_2\alpha : \mathcal{V} \rightarrow \mathcal{V}_2$ mindegyike null-leképezés.

Bizonyítás. Legyen $\alpha : \mathcal{U}_2 \rightarrow \mathcal{U}_1$. Az α balról csak az \mathcal{U}_1 tér ι_1 identitásával, jobbról pedig csak az \mathcal{U}_2 tér ι_2 identitásával szorozható; így ezek az identitások egyértelműek. Ha $\mathbf{u} \in \mathcal{U}_2$, akkor $(\alpha\iota_2)(\mathbf{u}) = \alpha(\iota_2(\mathbf{u})) = \alpha(\mathbf{u})$, tehát $\alpha\iota_2 = \alpha$. Továbbá $(\iota_1\alpha)(\mathbf{u}) = \iota_1(\alpha(\mathbf{u})) = \alpha(\mathbf{u})$, tehát $\iota_1\alpha = \alpha$. Ha az ι identitásra $\iota\alpha$, illetve $\beta\iota$ értelmezett, akkor az előbbiek szerint $\iota\alpha = \alpha$, illetve $\beta\iota = \beta$.

Tegyük fel, hogy az ι leképezésre teljesül az, hogy $\alpha\iota = \alpha$, amennyiben a bal oldalon levő szorzás elvégezhető. Legyen ι_1 az ι -hoz tartozó bal oldali identitás. Ez azt jelenti, hogy $\iota_1\iota = \iota$. Másrészt viszont az ι tulajdonsága alapján $\iota\iota_1 = \iota_1$, azaz $\iota = \iota_1$, tehát ι identitás. A másik állítás szimmetrikusan bizonyítható.

A null-leképezésekre vonatkozó állítások triviálisan következnek abból, hogy egy homogén lineáris leképezés pontosan akkor null-leképezés, ha minden vektort nullvektorra képez. ■

Lineáris leképezések esetén az osztás általában nem végezhető el, még akkor sem, ha a vektorterek megfelelő „elrendezése” ezt lehetővé tenné. Az osztás elvégezhetőségének vannak gyengébb formái, illetve vannak akadályai. Mindenekelőtt megjegyezzük, hogy itt külön kell nézni bal oldali és jobb oldali osztást, hiszen a szorzás nem kommutatív.

3.8. Definíció. Azt mondjuk, hogy a τ , illetve σ homogén lineáris leképezés az α , illetve a β homogén lineáris leképezésnek bal oldali, illetve jobb oldali inverze, ha $\tau\alpha$, illetve $\beta\sigma$ identitás. ■

3.9. Definíció. Azt mondjuk, hogy az α leképezés balreguláris (a β jobbreguláris), ha az $\alpha\xi = \alpha\eta$ (a $\mu\beta = \nu\beta$) feltételből mindig következik $\xi = \eta$ ($\mu = \nu$). ■

3.10. Definíció. Tegyük fel, hogy $\alpha\beta = \omega$. Ha $\beta \neq \omega$, akkor α -t bal oldali nullosztónak nevezzük; ha $\alpha \neq \omega$, akkor β jobb oldali nullosztó. ■

A következő két tétel a fenti fogalmak közötti kapcsolatokra világít rá:

3.16. Tétel. A következő állítások ekvivalensek:

- (1) α injekció (injektív).
- (2) α -nak létezik balinverze.
- (3) α balreguláris.
- (4) α nem bal oldali nullosztó.

Bizonyítás. Legyen először $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ injektív, és legyen $\mathcal{V}_1 = \text{Im}(\alpha)$. A 2.12. Tétel alapján van olyan $\mathcal{V}_2 \leq \mathcal{V}$, amelyre $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$. Mivel $\text{Ker}(\alpha) = \{\mathbf{0}\}$, ezért az $\mathbf{u} \mapsto \alpha(\mathbf{u})$ egy $\mathcal{U} \rightarrow \mathcal{V}_1$ izomorfizmust hoz létre. Ennek létezik tehát egy $\tau_1 : \mathcal{V}_1 \rightarrow \mathcal{U}$

inverze. Legyen $\tau_2 : \mathcal{V}_2 \rightarrow \mathcal{U}$ az a (lineáris) leképezés, ami \mathcal{V}_2 minden elemét \mathbf{o} -ba viszi. A $\tau = \tau_1 \oplus \tau_2$ megfeleltetésre (\mathbf{o}_2 a \mathcal{V}_2 -beli nullvektor):

$$\tau(\alpha(\mathbf{u})) = \tau(\alpha(\mathbf{u}) + \mathbf{o}_2) = \tau_1(\alpha(\mathbf{u})) + \tau_2(\mathbf{o}_2) = \mathbf{u} + \mathbf{o} = \mathbf{u},$$

tehát α -nak létezik balinverze.

Legyen $\alpha\xi = \alpha\eta$, és tegyük fel, hogy α -nak létezik egy τ balinverze. Ebből kapjuk, hogy $\xi = \iota\xi = \tau\alpha\xi = \tau\alpha\eta = \iota\eta = \eta$, tehát α balreguláris.

Legyen most α balreguláris, és tegyük fel, hogy $\alpha\beta = \omega$. (Vigyázat, ω mindig a megfelelő null-leképezést jelöli!) Tekintettel a triviálisan teljesülő $\alpha\omega = \omega$ egyenlőségre $\alpha\beta = \alpha\omega$, s a balregularitás alapján $\beta = \omega$, azaz α nem bal oldali nullosztó.

Végezetül tegyük fel, hogy $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ nem bal oldali nullosztó. Legyen $\mathcal{U}_1 = \text{Ker}(\alpha)$, és legyen $\mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2$. Legyen $\sigma_1 : \mathcal{U}_1 \rightarrow \mathcal{U}$ az a leképezés, amely minden vektort önmagára képez (ez nyilván homogén lineáris). Legyen $\sigma_2 : \mathcal{U}_2 \rightarrow \mathcal{U}$ a null-leképezés. A $\sigma = \sigma_1 \oplus \sigma_2$ leképezésre és az $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ ($\mathbf{u}_i \in \mathcal{U}_i$) vektorra:

$$\alpha\sigma(\mathbf{u}) = \alpha\sigma(\mathbf{u}_1 + \mathbf{u}_2) = \alpha\sigma_1(\mathbf{u}_1) + \alpha\sigma_2(\mathbf{u}_2) = \alpha(\mathbf{u}_1) + \alpha(\mathbf{o}) = \mathbf{o} + \mathbf{o} = \mathbf{o}$$

alapján $\alpha\sigma = \omega$. A feltétel szerint tehát $\sigma = \omega$, így σ_1 minden vektort \mathbf{o} -ra képez, ami csak úgy lehet, ha $\mathcal{U}_1 = \{\mathbf{o}\}$, ami az injektivitással ekvivalens. ■

3.17. Tétel. A következő állítások ekvivalensek:

- (1) α szűrjekció (szűrjektív).
- (2) α -nak létezik jobbinverze.
- (3) α jobbreguláris.
- (4) α nem jobb oldali nullosztó.

Bizonyítás. Legyen $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ szűrjektív, $\mathcal{U}_1 = \text{Ker}(\alpha)$ és $\mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2$. A 2.11. Tétel alapján α -nak az \mathcal{U}_2 -re való megszorítása egy $\alpha_2 : \mathcal{U}_2 \rightarrow \mathcal{V}$ izomorfizmus. Ez azt jelenti, hogy \mathcal{V} minden eleme egyértelműen felírható $\mathbf{v} = \alpha(\mathbf{u}_2)$ alakban, ahol $\mathbf{u}_2 \in \mathcal{U}_2$; továbbá α_2 -nek létezik egy $\tau : \mathcal{V} \rightarrow \mathcal{U}_2$ inverze. Ekkor tetszőleges $\mathbf{v} \in \mathcal{V}$ elemre $\mathbf{v} = \alpha(\mathbf{u}_2)$, ahol $\mathbf{u}_2 \in \mathcal{U}_2$ és $\tau(\mathbf{v}) = \mathbf{u}_2$. Így $\alpha(\tau(\mathbf{v})) = \alpha(\mathbf{u}_2) = \mathbf{v}$, azaz $\alpha\tau$ a \mathcal{V} identitása, tehát τ az α jobbinverze.

Ha α -nak létezik egy τ jobbinverze, akkor tegyük fel, hogy a μ, ν lineáris leképezésekre $\mu\alpha = \nu\alpha$ teljesül. Ekkor $\mu = \mu\tau = \mu\alpha\tau = \nu\alpha\tau = \nu\tau = \nu$, azaz α jobbreguláris.

Amennyiben α jobbreguláris és $\beta\alpha = \omega$, akkor a jobbregularitás és a $\omega\alpha = \omega$ (a két null-leképezés itt sem ugyanaz!) összefüggés alapján kapjuk, hogy $\beta = \omega$; azaz α nem jobb oldali nullosztó.

Végezetül tegyük fel, hogy $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ nem jobb oldali nullosztó. Legyen $\mathcal{V}_1 = \text{Im}(\alpha)$ és $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$. Legyen $\sigma_1 : \mathcal{V}_1 \rightarrow \mathcal{V}$ a null-leképezés és $\sigma_2 : \mathcal{V}_2 \rightarrow \mathcal{V}$ a beágyazás (azaz $\sigma_2(\mathbf{v}_2) = \mathbf{v}_2$, ha $\mathbf{v}_2 \in \mathcal{V}_2$) és legyen $\sigma = \sigma_1 \oplus \sigma_2$. Most tetszőleges $\mathbf{u} \in \mathcal{U}$ vektorra

$$(\sigma\alpha)(\mathbf{u}) = \sigma(\alpha(\mathbf{u})) = (\sigma_1 \oplus \sigma_2)(\alpha(\mathbf{u})) = \sigma_1(\alpha(\mathbf{u})) + \sigma_2(\mathbf{o}) = \mathbf{o} + \mathbf{o} = \mathbf{o},$$

tehát σ a null-leképezés, amiből $\mathcal{V}_2 = \{\mathbf{o}\}$; és ezért α szűrjektitvása következik. ■

3.18. Tétel. *Ha egy $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ lineáris leképezésnek létezik bal oldali inverze is és jobb oldali inverze is, akkor ezek megegyeznek. Ebben az esetben a leképezés inverzéről beszélünk. Az α inverzét α^{-1} jelöli. Ez pontosan akkor igaz, ha α izomorfizmus.*

Ha létezik α^{-1} és β^{-1} és létezik a $\beta\alpha$ szorzat, akkor $(\beta\alpha)^{-1} = \alpha^{-1}\beta^{-1}$ és $(\alpha^{-1})^{-1} = \alpha$.

Bizonyítás. Legyen τ bal oldali és σ jobb oldali inverz, azaz $\tau\sigma = \iota_{\mathcal{U}}$ és $\alpha\sigma = \iota_{\mathcal{V}}$. Ekkor

$$\tau = \tau\iota_{\mathcal{V}} = \tau(\alpha\sigma) = (\tau\alpha)\sigma = \iota_{\mathcal{U}}\sigma = \sigma.$$

Ez nyilvánvalóan pontosan az izomorfizmusokra igaz.

Ha létezik α^{-1} és β^{-1} és létezik az $\alpha\beta$ szorzat, akkor

$$(\alpha^{-1}\beta^{-1})(\beta\alpha) = \alpha^{-1}(\beta^{-1}\beta)\alpha = \alpha^{-1}\iota_{\mathcal{V}}\alpha = \alpha^{-1}\alpha = \iota_{\mathcal{U}}$$

alapján $\alpha^{-1}\beta^{-1}$ a $\beta\alpha$ balinverze. Hasonlóképpen látható be, hogy ez jobbinverz is; így $\alpha^{-1}\beta^{-1} = (\beta\alpha)^{-1}$. Ugyancsak egyszerűen belátható az $(\alpha^{-1})^{-1} = \alpha$ összefüggés is. ■

Már szerepeltek olyan lineáris leképezések, amelyek egy vektorteret önmagába képeztek. Ebben a fontos speciális esetben külön nevet adunk a leképezésnek:

3.11. Definíció. Egy vektortérnek önmagába való lineáris leképezését lineáris transzformációnak nevezzük. ■

3.19. Tétel. *Az \mathcal{U}_K vektortér lineáris transzformációi az összeadásra és szorzásra nézve egy általában nemkommutatív gyűrűt alkotnak, amelyet \mathcal{U} endomorfizmusgyűrűjének nevezünk és $\text{End}(\mathcal{U})$ -val vagy $\text{End}_K(\mathcal{U})$ -val jelölünk.*

Bizonyítás. Azonnal következik a 3.14. Tételből. ■

Megjegyzés. Ha $\text{End}(\mathcal{U})$ műveleteihez hozzávesszük a skalárokkal való szorzást, valamint a 3.14. Tételben szereplő újabb azonosságot is, akkor egy ugyancsak fontos struktúrát kapunk, amelynek a neve *hiperkomplex rendszer*, vagy *algebra* a K test felett. □

3.20. Tétel. *Ha α egy véges dimenziós vektortér lineáris transzformációja, akkor a 3.16. és a 3.17. Tételekben adott tulajdonságok ekvivalensek.*

Bizonyítás. Legyen α az n -dimenziós \mathcal{U} vektortér lineáris transzformációja. Ekkor a 3.5. Tétel szerint

$$\dim(\text{Ker}(\alpha)) + \dim(\text{Im}(\alpha)) = n.$$

Ezért a $\dim(\text{Ker}(\alpha)) = 0$ és a $\dim(\text{Im}(\alpha)) = n$ feltételek ekvivalensek. Az első egyenlőség azt jelenti, hogy α magja egyedül a nullvektorból áll, azaz α injektív, míg a második azt, hogy α képe az egész tér, azaz α szürjektív. Ezek pedig az idézett tételekben az (1) helyen szereplő feltételek. ■

Ebből, a 3.18. Tételt is figyelembe véve, azonnal adódik a

Következmény. Ha egy véges dimenziós tér valamely transzformációjának létezik egyik oldali inverze, akkor ez már kétoldali inverz. ■

A gyakorlati alkalmazások szempontjából is igen fontos definíció következik:

3.12. Definíció. Legyen φ az \mathcal{U}_K tér egy lineáris transzformációja.

I. Ha minden $\mathbf{v} \in \mathcal{V} \subseteq \mathcal{U}$ vektorra $\varphi(\mathbf{v}) \in \mathcal{V}$, akkor \mathcal{V} -t a φ invariáns alterének nevezzük.

II. Egydimenziós invariáns altér generátoreleme a φ sajátvektora.

III. Ha \mathcal{V} nem csak a nullvektorból áll, és maximális olyan altér, amelynek minden eleme φ -nek sajátvektora, akkor \mathcal{V} neve sajátaltér.

IV. Ha a c skálárhoz van olyan \mathbf{u} sajátvektor, amelyre $\varphi(\mathbf{u}) = c \cdot \mathbf{u}$, akkor c -t a φ sajátértékének nevezzük. (c az \mathbf{u} -hoz tartozó sajátérték és \mathbf{u} egy c -hez tartozó sajátvektor.) ■

Megjegyzés. A sajátértékeket általában λ -val szokták jelölni. Mi ezt itt azért nem tesszük, mert a görög betűket a leképezések jelölésére használjuk. Az invariáns alterek, a sajátvektorok és a sajátértékek a lineáris algebrában és az alkalmazásokban is központi jelentőségűek; részletesebb vizsgálatukra később térünk rá. □

Feladatok

1. Legyen \mathbf{u}, \mathbf{v} „a” kétdimenziós \mathcal{U} tér egy bázisa, és definiálja az $\alpha, \beta : \mathcal{U} \rightarrow \mathcal{U}$ lineáris leképezéseket az $\alpha(\mathbf{u}) = \mathbf{o}$, $\alpha(\mathbf{v}) = \mathbf{u}$, illetve a $\beta(\mathbf{u}) = \mathbf{u}$, $\beta(\mathbf{v}) = \mathbf{o}$ összefüggés. Bizonyítsuk be, hogy pontosan egy-egy ilyen lineáris leképezés létezik; továbbá $\beta\alpha = \alpha$ és $\alpha\beta = \omega$.

2. Mutassuk meg, hogy egy lineáris leképezésnek létezhet olyan bal-, illetve jobbinverze, amelyik nem lineáris leképezés.

3. Mutassuk meg, hogy egy lineáris leképezésnek létezhet több bal-, illetve jobbinverze.

4. Bizonyítsuk be, hogy ha egy α leképezés mindkét oldali nullosztó, akkor létezik olyan $\beta \neq \omega$ leképezés, amelyre $\alpha\beta = \omega$ és $\beta\alpha = \omega$ is teljesül.

6. Adjunk meg olyan lineáris leképezést, amely „egyik oldali” nullosztó és létezik „másik oldali” inverze.

7. Mutassuk meg, hogy végtelen dimenziós terekben a 3.20. Tétel Következménye nem igaz.

8. Bizonyítsuk be, hogy ugyanahhoz a transzformációhoz tartozó invariáns alterek metszete is generátuma is invariáns altér.

9. Bizonyítsuk be, hogy $\alpha\beta = \omega$ pontosan akkor igaz, ha $\text{Im}(\beta) \leq \text{Ker}(\alpha)$.

10. Bizonyítsuk be, hogy egy sajátaltér minden vektora ugyanahhoz a sajátértékhez tartozik; és úgy is definiálható, hogy egy adott sajátértékhez tartozó sajátvektorok összessége.

11. Bizonyítsuk be, hogy $r(\alpha\beta\gamma) \leq r(\beta)$ (ha a szorzat létezik).

12. Bizonyítsuk be, hogy az előző feladatból „formailag” következik az, hogy $r(\varphi\psi) \leq \min(r(\varphi), r(\psi))$; és az is, hogy reguláris α, γ esetében $r(\alpha\beta\gamma) = r(\beta)$.

13. Bizonyítsuk be, hogy

$$r(\alpha\beta) \leq r(\beta) - \dim(\text{Im}(\beta) \wedge \text{Ker}(\alpha)) \leq \min(r(\alpha), r(\beta)).$$

14. Bizonyítsuk be, hogy $\varepsilon : \mathcal{U} \rightarrow \mathcal{U}$ pontosan akkor projekció, ha ε idempotens, azaz $\varepsilon^2 = \varepsilon$.

5. Lineáris függvények és a duális tér

A $\text{Hom}(\mathcal{U}, \mathcal{V})$ „kifejezés” úgy viselkedik, mint egy kétváltozós függvény, amely két K fölötti vektortérhez egy harmadikat rendel hozzá. (Az ilyenfajta függvényeket *funktoroknak* nevezik.) Ennek a hozzárendelésnek a „szabályait” a későbbiekben fogjuk vizsgálni. Egyelőre bizonyos speciális eseteket nézünk meg. Az $\mathcal{U} = \mathcal{V}$ esetet már láttuk; ekkor kaptuk a lineáris transzformációkat.

A következőkben azt nézzük meg, amikor a két szereplő vektortér egyike igen speciális. Ezen azt értjük, hogy a dimenziója kicsi. A legkisebb lehetséges dimenzió a 0. Világos, hogy $\text{Hom}(\mathcal{O}, \mathcal{V}) = \mathcal{O}$ és $\text{Hom}(\mathcal{V}, \mathcal{O}) = \mathcal{O}$. A következő lehetőség, amikor az egyik vektortér 1-dimenziós. Természetesen „rengeteg” 1-dimenziós vektortér van (rögzített K esetében is). Ki lehet azonban mutatni (éppen a funktorokra vonatkozó szabályok segítségével), hogy az eredmény mindig „ugyanaz” (izomorfiától eltekintve — sőt „természetes” izomorfiától eltekintve). Inkább kiválasztunk a sok egydimenziós tér közül egy olyat, amely egyszerűen és természetesen adódik; ez pedig maga a K test. Ennek a speciális választásnak még egy nagy előnye van. Egy tetszőleges 1-dimenziós vektortér esetén bármely $\mathbf{u} \neq \mathbf{o}$ vektor bázisa a térnek; és nincs mód arra, hogy egy „természetes” bázist válasszunk. Ezzel szemben a K esetében erre van mód, és ezt meg is tesszük:

K -ban mindig az 1 egységelemből álló $\{1\}$ halmazt tekintjük (természetes) bázisnak.

Az előzőeknek megfelelően két speciális esettel foglalkozunk, az egyik a $\text{Hom}(K_K, \mathcal{V})$, a másik a $\text{Hom}(\mathcal{U}, K_K)$. Valójában szólni kellene a legspeciálisabb esetről (ez a $\text{Hom}(K_K, K_K)$ eset), de ezt az első esettel együtt tárgyaljuk.

3.21. Tétel. *A $\mathbf{v} \in \mathcal{V}$ vektornak feleltessük meg azt a $\tilde{\mathbf{v}} \in \text{Hom}(K, \mathcal{V})$ leképezést, amelyre $\tilde{\mathbf{v}}(1) = \mathbf{v}$. Ez a leképezés egyértelmű, és a $\mathbf{v} \rightarrow \tilde{\mathbf{v}}$ megfeleltetés egy természetes $\mathcal{V} \rightarrow \text{Hom}(K, \mathcal{V})$ izomorfizmust hoz létre; amin azt értjük, hogy tetszőleges $\varphi : \mathcal{V} \rightarrow \mathcal{W}$ homogén lineáris leképezésre $\varphi(\tilde{\mathbf{v}}) = \varphi \circ \tilde{\mathbf{v}} = \varphi(B\mathbf{v})$.*

Bizonyítás. Tekintettel arra, hogy 1 a K bázisa, ezért, ha megmondjuk 1 képét, akkor ezzel egyértelműen meghatároztunk egy homogén lineáris leképezést. Ha $\varphi \in \text{Hom}(K, \mathcal{V})$, akkor $\varphi(1)$ meghatározza φ -t; így a $\mathbf{v} \rightarrow \tilde{\mathbf{v}}$ megfeleltetés szürjektív. Ha $\tilde{\mathbf{v}} = \omega$, akkor

$\mathbf{v} = \widetilde{\mathbf{v}}(1) = \omega(1) = \mathbf{o}$ következtében a megfeleltetés injektív; és így valóban bijekció. Legyenek $\mathbf{a}, \mathbf{b} \in \mathcal{V}$, $a, b \in K$. Ekkor

$$(\widetilde{a\mathbf{a} + b\mathbf{b}})(1) = a\mathbf{a} + b\mathbf{b} = a\widetilde{\mathbf{a}}(1) + b\widetilde{\mathbf{b}}(1) = (a\widetilde{\mathbf{a}} + b\widetilde{\mathbf{b}})(1)$$

alapján érvényes a művelettartás.

A természetesség a következőképpen látható be:

$$\widetilde{\varphi(\mathbf{v})}(1) = \varphi(\mathbf{v}) = \varphi(\widetilde{\mathbf{v}}(1)) = (\varphi \circ \widetilde{\mathbf{v}})(1).$$

Mivel a két leképezés a K_K bázisán megegyezik, ezért egyenlő a két leképezés is. ■

Megjegyzés. Ennek a tételnek egy nagyon hasznos „formális” következménye van. Az egyszerre megadott izomorfizmusok alapján a vektorokat egyszerre helyettesíthetjük leképezésekkel. Ez olyan, mintha vektorok nem is lennének, csupán homogén lineáris leképezések. Hiszen a bizonyított izomorfizmusok alapján a kapott leképezésekkel ugyanúgy számolhatunk, mint az eredeti vektorokkal. A $\varphi(\mathbf{v})$ -nek megfelelő $\widetilde{\varphi(\mathbf{v})}$ leképezés nem más, mint a φ és a $\widetilde{\mathbf{v}}$ leképezéseknek ebben a sorrendben vett szorzata. Ha a természetes izomorfizmus alapján „elfelejtkezünk” a föléhúzásról, akkor azt kapjuk, hogy $\varphi(\mathbf{v})$ nem más, mint a φ -nek és a \mathbf{v} -nek a szorzata. Ennek megfelelően $\varphi(\mathbf{v})$ helyett használni fogjuk a $\varphi\mathbf{v}$ jelölést is, amelyik úgy viselkedik, mint a szorzás; legalábbis az összeadásra nézve disztributív és asszociatív (ha elvégezhető).

A $\varphi(c\mathbf{v}) = c\varphi(\mathbf{v})$ összefüggést pedig úgy értelmezhetjük, hogy a skalárok minden leképezéssel felcserélhetők. Egyébként a skalároknak is megfeleltethetünk egy lineáris leképezést a 3.21. Tétel alapján: $c \in K$ skalárnak az a $\text{Hom}(K, K)$ -beli leképezés felel meg, amelyik 1-et c -re képezi; ez a c -vel való szorzás. □

Ezek után rátérünk a „másik” esetre, a $\text{Hom}(\mathcal{U}, K)$ vizsgálatára. Mivel itt K rögzített, ez úgy tekinthető, mint egy egyváltozós funktor (vagyis függvény).

3.13. Definíció. Az $\mathcal{U}^* = \text{Hom}(\mathcal{U}, K)$ vektorteret az \mathcal{U} duális terének nevezzük. Ennek elemeit duális vektoroknak vagy lineáris függvényeknek vagy lineáris operátoroknak nevezzük. ■

Megjegyzés. Az $\widetilde{\mathbf{u}}$ jelölést csak „házi használatra” vezettük be. Az \mathbf{u}^* jelöléssel gyakrabban találkozhatunk; de ez sem általános. □

3.22. Tétel. Ha \mathcal{U} véges dimenziós, akkor $\mathcal{U}^* \cong \mathcal{U}$. Ez az izomorfizmus azonban nem természetes izomorfizmus.

Bizonyítás. Tekintsük az \mathcal{U} vektortérnek egy $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázisát. A két vektortér izomorfizmusához elegendő bizonyítani, hogy egyenlő dimenziósak. Ehhez az kell, hogy \mathcal{U}^* -nak legyen n -elemű bázisa. Tekintsük az \mathcal{U}^* tér $\{\mathbf{u}_1^*, \dots, \mathbf{u}_n^*\}$ elemeit, amelyeket a következőképpen definiálunk: Legyen $\mathbf{u}_i^*(\mathbf{u}_j) = \delta_{i,j}$, ahol δ az úgynevezett Kronecker-féle δ , azaz $\delta_{i,j} = 1$, ha $i = j$ és $\delta_{i,j} = 0$, ha $i \neq j$. (Tehát $\mathbf{u}_i^*(\mathbf{u}_i) = 1$ és $\mathbf{u}_i^*(\mathbf{u}_j) = 0$, ha $j \neq i$.) Ezek a leképezések lineárisan függetlenek: Ha ugyanis $\varphi = c_1\mathbf{u}_1^* + \dots + c_n\mathbf{u}_n^* = \omega$,

akkor $\varphi(\mathbf{u}_i) = \underbrace{0 + \dots + 0}_{(i-1)\text{-szer}} + c_i + \underbrace{0 + \dots + 0}_{(n-i)\text{-szer}} = \omega(\mathbf{u}_i) = 0$ alapján $c_i = 0$; tehát a fenti lineáris kombináció triviális.

Legyen $\varphi \in \mathcal{U}^*$ egy tetszőleges leképezés és legyen $\varphi(\mathbf{u}_i) = d_i$. Tekintsük most a $\psi = d_1 \mathbf{u}_1^* + \dots + d_n \mathbf{u}_n^*$ leképezést. Erre $\psi(\mathbf{u}_i) = \overbrace{0 + \dots + 0}^{(i-1)\text{-szer}} + d_i + \overbrace{0 + \dots + 0}^{(n-i)\text{-szer}} = d_i = \varphi(\mathbf{u}_i)$ alapján $\psi = \varphi$, hiszen $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} egy bázisa. Így a felvett leképezések \mathcal{U}^* -nak generátorrendszerét is alkotják. ■

Megjegyzés. A 3.22. Tétel alapján a $\sum_i c_i \mathbf{u}_i \rightarrow \sum_i c_i \mathbf{u}_i^*$ megfeleltetés izomorfizmus. De biztosan nem természetes, hiszen, ha \mathcal{U} -ban egy másik bázist veszünk fel, akkor egy másik izomorfizmust kapunk. Ettől még lehetséges volna az, hogy létezik egy természetes izomorfizmus a két tér között; de bebizonyítható, hogy nem így van.

Azt is meg lehet mutatni, hogy ha \mathcal{U} nem véges dimenziós, akkor a két vektortér nem is izomorf.

Érdekes megemlíteni viszont az \mathcal{U} és az $\mathcal{U}^{**} = (\mathcal{U}^*)^*$ közötti „természetes” kapcsolatot. Legyen $\mathbf{u} \in \mathcal{U}$, és feleltessük meg ennek azt az $\widehat{\mathbf{u}} \in \mathcal{U}^{**}$ leképezést, amely tetszőleges $\varphi \in \mathcal{U}^*$ leképezést $\varphi(\mathbf{u})$ -ba visz ($\widehat{\mathbf{u}}(\varphi) = \varphi(\mathbf{u})$). Ez a megfeleltetés véges dimenziós vektorterekre természetes izomorfizmus; a végtelen dimenziós esetben is természetes, de csupán injekció. □

Feladatok

1. Legyen $\{\mathbf{u}, \mathbf{v}\}$ az \mathcal{U} \mathbb{Q} feletti vektortér egy bázisa és $\{\mathbf{u}^*, \mathbf{v}^*\}$ az ehhez tartozó bázisa az \mathcal{U}^* duális térnek. Keressük meg \mathcal{U}^* -nak azt a bázisát, amelyet akkor kapunk, ha \mathcal{U} -ban az $\mathbf{e} = \mathbf{u} + 3\mathbf{v}$ és $\mathbf{f} = 2\mathbf{e} + \mathbf{v}$ vektorokból álló bázist választjuk.

2. Bizonyítsuk be, hogy az $\widehat{\mathbf{u}}(\varphi) = \varphi(\mathbf{u})$ megfeleltetéssel adott leképezés valóban természetes; véges dimenziós esetben izomorfizmus és végtelen dimenziós esetben nem izomorfizmus, csak injektív.

3. Tekintsük a \mathbb{Z} -modulusok $\mathfrak{H} = \{\mathbb{Z}, \mathbb{Z}_n, \mathbb{Q}, \mathbb{Z} \oplus \mathbb{Z}, \mathcal{E}, \mathcal{E}_p\}$ halmazát, ahol \mathcal{E} az egységgyökök halmaza és \mathcal{E}_p a p -hatványadik egységgyökök halmaza és a vektorösszeadás a szorzás. Határozzuk meg a $\text{Hom}(\mathcal{M}, \mathcal{N})$ \mathbb{Z} -modulusokat, ha $\mathcal{M}, \mathcal{N} \in \mathfrak{H}$.

4. Legyen $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ izomorfizmus. Bizonyítsuk be, hogy $\text{Hom}(\mathcal{U}, \mathcal{A}) \cong \text{Hom}(\mathcal{V}, \mathcal{A})$ és $\text{Hom}(\mathcal{B}, \mathcal{U}) \cong \text{Hom}(\mathcal{B}, \mathcal{V})$ „természetes módon”.

5. Legyenek $\mathcal{U}, \mathcal{V}, \mathcal{U}_i$ -k és \mathcal{V}_j -k vektorterek egy adott test felett. Bizonyítsuk be, hogy $\text{Hom}(\mathcal{U}, \bigoplus_j \mathcal{V}_j) \cong \bigoplus_j \text{Hom}(\mathcal{U}, \mathcal{V}_j)$ és $\text{Hom}(\bigoplus_i \mathcal{U}_i, \mathcal{V}) \cong \prod_i \text{Hom}(\mathcal{U}_i, \mathcal{V})$.

NEGYEDIK FEJEZET

KOORDINATIZÁLÁS

1. Vektorok koordinátái és leképezések mátrixa

Mint már említettük, a vektorok a lineáris függvényeknek (vagy a homogén lineáris polinomoknak), a lineáris leképezések meg ilyenekből álló rendszernek a tömör jelölései. Ahhoz, hogy ezekkel számolni tudjunk, szükséges ezeknek számokkal való megadása. Azonnal megjegyezzük, hogy a „numerikus jellemzők” nem mindig számok, lehetnek például polinomok, maradékosztályok; általában tetszőleges gyűrű elemei. Ez a numerikus jellemző még mindig csak egy „közbülső állapot”; itt például az $(x, y) \mapsto 2x + 3y$ függvény helyett csupán a $(2, 3)$ számpár szerepel. Az egész fejezetben csak véges dimenziós vektorterekkel foglalkozunk; mert a tényleges gyakorlati alkalmazásban ezek fordulnak elő. Ennek ellenére célszerű gondolatban a végtelen dimenziós vektortereket és a modulusokat is figyelembe venni.

A vektorok koordinatizálásához szükség van a koordináta-rendszerre. Ezt egy bázissal lehet megadni. Ha más bázist veszünk fel, akkor természetesen megváltoznak a vektorok koordinátái. Még azt is fontos megjegyezni, hogy egy síkbeli „alakzat” nem változik meg, ha például a két koordinátatengelyt felcseréljük, de az alakzat által meghatározott függvény igen (más az $y = x^3$ és az $x = y^3$, vagyis $y = \sqrt[3]{x}$ függvény). Erre a jelenségre már utaltunk akkor, amikor említettük, hogy a lineáris kombinációknál figyelembe kell venni, hogy melyik skalár melyik vektornak az együtthatója. Eddig a bázisokat lényegében halmazként kezeltük; ezután viszont különbözőnek kell tekinteni azokat a bázisokat, amelyekben ugyanazok a vektorok szerepelnek, csak más sorrendben.

4.1. Definíció. Az \mathcal{U} vektortér $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és $\mathbf{U}' = \{\mathbf{u}'_1, \dots, \mathbf{u}'_n\}$ bázisa azonos, ha minden $i \in \{1, \dots, n\}$ mellett $\mathbf{u}_i = \mathbf{u}'_i$. Ezt $\mathbf{U} \equiv \mathbf{U}'$ fogja jelölni. A $\mathbf{U} = \mathbf{U}'$ jelölés továbbra is a két halmaz egyenlőségét fogja jelenteni. ■

A vektorok koordinátáinak, illetve a leképezések mátrixának a meghatározására hasznosak olyan formális szabályok, amelyek a számolásokat, de mindenekelőtt a bizonyítások menetét megkönnyítik. Előkészületül bizonyítjuk az alábbiakat:

4.1. Tétel. Legyen $U = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U}_K és $V = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ a \mathcal{V}_K vektortér egy-egy bázisa az $U^* = \{\mathbf{u}_1^*, \dots, \mathbf{u}_n^*\}$ és $V^* = \{\mathbf{v}_1^*, \dots, \mathbf{v}_k^*\}$ duális bázisokkal együtt. Ekkor:

- (1) $\alpha = \sum_i \{\tilde{\mathbf{u}}_i \mathbf{u}_i^* \mid \mathbf{u}_i \in U\} = \iota_U$ az \mathcal{U} vektortér identitása.
- (2) $\psi = \tilde{\mathbf{v}} \mathbf{u}^*$ a $\text{Hom}(\mathcal{U}, \mathcal{V})$ egy ≤ 1 rangú eleme ($\mathbf{v} \in \mathcal{V}$ és $\mathbf{u} \in \mathcal{U}$); továbbá $\text{Hom}(\mathcal{U}, \mathcal{V})$ minden ≤ 1 rangú eleme ilyen alakba írható. (Egy ilyen szorzatot diádnak nevezzük.)
- (3) Minden $\varphi: \mathcal{U} \rightarrow \mathcal{V}$ homomorfizmushoz léteznek olyan (nem egyértelműen meghatározott) $\mathbf{a}_1, \dots, \mathbf{a}_r \in \mathcal{V}$ és $\mathbf{b}_1, \dots, \mathbf{b}_r \in \mathcal{U}$ vektorok, amelyekre $\varphi = \tilde{\mathbf{a}}_1 \mathbf{b}_1^* + \dots + \tilde{\mathbf{a}}_r \mathbf{b}_r^*$, ahol $r = r(\varphi)$.
- (4) Ha adott $\mathbf{e}_1, \dots, \mathbf{e}_s \in \mathcal{V}$ és $\mathbf{f}_1, \dots, \mathbf{f}_s \in \mathcal{U}$ vektorokra $\varphi = \tilde{\mathbf{e}}_1 \mathbf{f}_1^* + \dots + \tilde{\mathbf{e}}_s \mathbf{f}_s^*$, akkor $s \geq r(\varphi)$.

Bizonyítás. (1): A duális bázis definíciója alapján $\alpha(\mathbf{u}_j) = \sum_i \tilde{\mathbf{u}}_i \mathbf{u}_i^* \mathbf{u}_j = \mathbf{u}_j = \iota(\mathbf{u}_j)$

($\mathbf{u}_j \in U$). Így a két leképezés \mathcal{U} egy bázisán megegyezik. Mivel mindegyikük lineáris, ezért a két leképezés egyenlő.

(2): $\mathcal{U} \xrightarrow{\mathbf{u}^*} K \xrightarrow{\tilde{\mathbf{v}}} \mathcal{V}$ miatt $\psi \in \text{Hom}(\mathcal{U}, \mathcal{V})$. Ha $\mathbf{u}' \in \mathcal{U}$, akkor $\mathbf{u}^* \mathbf{u}' = c \in K$ miatt $\psi(\mathbf{u}') = c\mathbf{v}$, tehát $\text{Im}(\psi) \leq \langle \mathbf{v} \rangle$, azaz $r(\psi) \leq 1$.

Legyen most $r(\psi) \leq 1$. Ha a rang 0, akkor $\psi = \omega$; és $\omega = \tilde{\mathbf{o}} \mathbf{o}^*$ egy megfelelő felírás. Egyébként van olyan $\mathbf{u}_1 \in \mathcal{U}$, amelyre $\mathbf{v} = \psi(\mathbf{u}_1) \neq \mathbf{o}$. Mivel $\mathbf{u} \notin \text{Ker}(\psi)$, továbbá $\dim(\text{Ker}(\psi)) = \dim(\mathcal{U}) - 1$, ezért $\text{Ker}(\psi)$ egy $\mathbf{u}_2, \dots, \mathbf{u}_n$ bázisát \mathbf{u}_1 -gyel kiegészítve \mathcal{U} -nak egy olyan U bázisát kapjuk, amelynek elemeit az \mathbf{u}_1 kivételével ψ \mathbf{o} -ba viszi. Tekintsük ennek a bázisnak az $\mathbf{u}^+_1, \dots, \mathbf{u}^+_n$ duális bázisát. Definíció szerint $\mathbf{u}^+_1 \mathbf{u}_1 = 1$ és $\mathbf{u}^+_1 \mathbf{u}_i = 0$, ha $i > 1$. Mivel $\mathbf{u} \mapsto \mathbf{u}^*$ izomorfizmus, ezért van olyan $\mathbf{u} \in \mathcal{U}$, amelyre $\mathbf{u}^* = \mathbf{u}^+_1$. Erre tehát $\tilde{\mathbf{v}} \mathbf{u}^*(\mathbf{u}_1) = \mathbf{v}$ és $\tilde{\mathbf{v}} \mathbf{u}^*(\mathbf{u}_i) = \mathbf{o}$, ha $i > 1$. Eszerint ψ és $\tilde{\mathbf{v}} \mathbf{u}^*$ ezen a bázison megegyeznek, tehát $\psi = \tilde{\mathbf{v}} \mathbf{u}^*$.

(3): Legyen $\dim(\mathcal{U}) = n$ és $r = r(\varphi)$, ekkor $\dim(\text{Ker}(\varphi)) = n - r$. A $\text{Ker}(\varphi)$ -nek egy $\mathbf{c}_{r+1}, \dots, \mathbf{c}_n$ bázisát egészítsük ki a $\mathbf{c}_1, \dots, \mathbf{c}_r$ elemekkel az \mathcal{U} egy \mathbf{C} bázisává. Legyen $\mathbf{C}^+ = \mathbf{c}^+_1, \dots, \mathbf{c}^+_n$ az ehhez a bázishoz tartozó duális bázis. Ez azt jelenti, hogy $\mathbf{c}^+_i \mathbf{c}_j = \delta_{i,j}$. \mathcal{U} és \mathcal{U}^* izomorfizmusa alapján léteznek olyan \mathbf{b}_i elemek, amelyek az eredeti bázishoz tartozó izomorfizmusnál \mathbf{c}^+_i -nak felelnek meg, azaz $\mathbf{b}_i^* = \mathbf{c}^+_i$ ($i = 1, \dots, n$). Ez azt jelenti, hogy $\mathbf{b}_i^* \mathbf{c}_j = \delta_{i,j}$. Legyenek továbbá $\mathbf{a}_1 = \varphi(\mathbf{c}_1), \dots, \mathbf{a}_r = \varphi(\mathbf{c}_r)$; és tekintsük a $\psi = \tilde{\mathbf{a}}_1 \mathbf{b}_1^* + \dots + \tilde{\mathbf{a}}_r \mathbf{b}_r^*$ elemet, amely nyilván eleme $\text{Hom}(\mathcal{U}, \mathcal{V})$ -nek. Erre, $1 \leq i \leq n$ esetén:

$$(\psi - \varphi)(\mathbf{c}_i) = \varphi(\mathbf{c}_i) - \sum_j \tilde{\mathbf{a}}_j \mathbf{b}_j^* \mathbf{c}_i = \begin{cases} \mathbf{a}_i - \mathbf{a}_i = \mathbf{o}, & \text{ha } i \leq r \\ \mathbf{o} - \mathbf{o} = \mathbf{o}, & \text{ha } i > r. \end{cases}$$

Így a két leképezés egy bázison megegyezik; tehát $\varphi = \psi$.

(4): Tegyük most fel, hogy adott $\mathbf{e}_1, \dots, \mathbf{e}_s \in \mathcal{V}$ és $\mathbf{f}_1, \dots, \mathbf{f}_s \in \mathcal{U}$ vektorokra $\varphi = \tilde{\mathbf{e}}_1 \mathbf{f}_1^* + \dots + \tilde{\mathbf{e}}_s \mathbf{f}_s^*$. Most tetszőleges $\mathbf{u} \in \mathcal{U}$ esetén $\varphi(\mathbf{u})$ előáll mint az $\mathbf{e}_1, \dots, \mathbf{e}_s$ vektoroknak a $\mathbf{f}_1^*(\mathbf{u}), \dots, \mathbf{f}_s^*(\mathbf{u})$ skalárokkal képezett lineáris kombinációja. Ez azt jelenti, hogy $\varphi(\mathbf{u}) \in \langle \mathbf{e}_1, \dots, \mathbf{e}_s \rangle$, azaz $\text{Im}(\varphi) \leq \langle \mathbf{e}_1, \dots, \mathbf{e}_s \rangle$. Mivel ennek az altérnek $\mathbf{e}_1, \dots, \mathbf{e}_s$ egy generátorrendszere, ezért dimenziója $\leq s$. Ezért ugyanez igaz ezen altér $\text{Im}(\varphi)$ alterére, tehát $s \geq r(\varphi)$. ■

4.2. Tétel. Legyen $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U}_K vektortérnek és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ a \mathcal{V}_K vektortérnek egy-egy bázisa. Ekkor $\{\tilde{\mathbf{v}}_j \mathbf{u}_i^* \mid i = 1, \dots, n; j = 1, \dots, k\}$ bázisa $\text{Hom}(\mathcal{U}, \mathcal{V}^*)$ -nek.

Bizonyítás. A 4.1. Tétel szerint a fenti homomorfizmusok mind elemei $\text{Hom}(\mathcal{U}, \mathcal{V}^*)$ -nek. Tekintsük a $\varphi = \sum_{i,j} c_{i,j} \tilde{\mathbf{v}}_j \mathbf{u}_i^*$ homomorfizmust ($c_{i,j} \in K$), és tegyük fel, hogy $\varphi = 0$.

Ekkor, rögzített $\mathbf{u}_p \in \mathcal{U}$ és $\mathbf{v}_q^* \in \mathcal{V}^*$ mellett

$$0 = \mathbf{v}_q^*(\varphi(\mathbf{u}_p)) = c_{q,p}((\mathbf{v}_q^* \tilde{\mathbf{v}}_q)(\mathbf{u}_p^*(\mathbf{u}_p))) = c_{q,p},$$

tehát a fenti homomorfizmusok függetlenek. Legyen most $\psi \in \text{Hom}(\mathcal{U}, \mathcal{V}^*)$ adott és tekintsük a K -beli $d_{i,j} = \mathbf{v}_j^* \psi \mathbf{u}_i^*(1)$ skalárokat. Azt kapjuk, hogy a

$$\sum_{i,j} d_{i,j} \tilde{\mathbf{v}}_j \mathbf{u}_i^* \quad 1 \leq i \leq n, \quad 1 \leq j \leq k$$

homomorfizmus minden \mathbf{u}_i bázisvektort ugyanabba visz, mint ψ , így megegyeznek; tehát az adott rendszer valóban bázis. ■

Ezután rátérhetünk az \mathcal{U} tér tetszőleges \mathbf{u} vektora koordinátáinak, illetve tetszőleges $\varphi : \mathcal{U} \rightarrow \mathcal{V}^*$ homomorfizmus mátrixának a meghatározására. Mint mondtunk, rögzítve van az \mathcal{U} tér egy \mathbf{U} bázisa, a \mathcal{V} tér egy \mathbf{V} bázisa a megfelelő \mathbf{U}^* és \mathbf{V}^* duális bázisokkal együtt.

4.2. Definíció. Egy $\mathbf{u} = c_1 \mathbf{u}_1 + \dots + c_n \mathbf{u}_n \in \mathcal{U}$ vektornak az \mathbf{U} bázisban felírt koordinátáin a c_1, \dots, c_n elemsorozatot értjük; c_i az i -edik koordináta. ■

4.3. Tétel. Az \mathbf{u} vektor i -edik koordinátája $\mathbf{u}_i^*(\mathbf{u})$.

Az \mathbf{u} vektornak az \mathbf{U} bázisban adott koordinátáit oszlopmátrixban felírva, ezt a következőképpen adhatjuk meg:

$$[\mathbf{u}] = \mathbf{U} [\mathbf{u}] = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^*(\mathbf{u}) \\ \vdots \\ \mathbf{u}_n^*(\mathbf{u}) \end{bmatrix} = [\mathbf{u}_i^*(\mathbf{u})]_{n,1}.$$

Rögzített bázis esetén a vektor és a koordináták egyértelműen meghatározzák egymást.

Bizonyítás. Az \mathbf{U} vektor i -edik koordinátájára adott összefüggés azonnal következik a duális bázis definíciójából.

A harmadik formula a jelzett oszlopmátrix tényleges felírása. A negyedik formula ugyanezt adja, a koordinátákat az előbbi összefüggés alapján behelyettesítve. Az utolsó formula ugyanennek a tömör jelölése, megadva a képzési szabályt és a mátrix sorainak és oszlopainak a számát. Az első és a második formula tulajdonképpen jelölés. Az első formula egyszerűen azt mondja, hogy az \mathbf{u} vektor mátrixáról beszélünk, míg a második formula azt is hozzáteszi, hogy az \mathbf{U} bázisban felírt mátrixról van szó.

Az utolsó állítás ismét triviális. ■

Megjegyzések

1. Az első helyen álló jelölés akkor célszerű, ha a bázis rögzített, és formális számolásokat akarunk végezni. A második helyen álló jelölés hasonló célokat szolgál, de olyan esetekben, amikor tekintettel kell lenni a bázisra; esetleg több bázisban is fel akarjuk írni az oszlopvektort. A harmadik jelölés a konkrét „numerikus” számolás esetén szükséges. A negyedik és ötödik jelölésre akkor van szükség, amikor meg akarjuk mondani, miképpen jöttek létre a koordináták.

2. Furcsának tűnhet, hogy a koordinátákat nem sorvektorral, hanem oszlopvektorral adjuk meg. Ez annak a következménye, hogy a függvényeket az argumentumok elé írjuk. \square

4.3. Definíció. Legyenek az \mathcal{U} és \mathcal{V} K -vektorterek adva az $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ bázisokkal; és legyen $\varphi : \mathcal{U} \rightarrow \mathcal{V}$. E leképezésnek a fenti bázisokban vett $[\varphi] = \mathbf{V}[\varphi]\mathbf{U}$ mátrixát $[\varphi] = [\mathbf{v}_i^*(\varphi(\mathbf{u}_j))]_{k,n}$ definiálja. \blacksquare

Megjegyzés. Természetesen φ mátrixa mindkét bázistól függ, ezért van szükség a két „felső” indexre. A formula megjegyzésének a megkönnyítésére megjegyezzük, hogy a φ „alkalmazásakor”, amire alkalmazzuk, az „utána van”, és az „elkészített” eredmény „előtte keletkezik”. Ezért van jobb oldalon az \mathcal{U} tér bázisa és bal oldalon a \mathcal{V} téré. Egyébként ezt a mátrixot részletesen kiírva azt kapjuk, hogy

$$\mathbf{V}[\varphi]\mathbf{U} = \begin{bmatrix} \mathbf{v}_1^*(\varphi(\mathbf{u}_1)) & \dots & \mathbf{v}_1^*(\varphi(\mathbf{u}_n)) \\ \vdots & \ddots & \vdots \\ \mathbf{v}_k^*(\varphi(\mathbf{u}_1)) & \dots & \mathbf{v}_k^*(\varphi(\mathbf{u}_n)) \end{bmatrix}. \quad \square$$

4.4. Tétel. Rögzített bázisok esetén a leképezés és a mátrixa egyértelműen meghatározzák egymást.

$[\tilde{\mathbf{v}}] = [\mathbf{v}]$, $[\mathbf{u}^*] = [\mathbf{u}]^\dagger$, $[\varphi]_j = [\varphi(\mathbf{u}_j)]$, ${}_i[\varphi] = [\mathbf{v}_i^* \varphi]$, ${}_i[\varphi]_j = [\mathbf{v}_i^* \varphi(\mathbf{u}_j)]$; a megfelelő, rögzített bázisokban. $[\iota] = [\delta_{i,j}]$ minden bázisban.

Bizonyítás. Mindenekelőtt megjegyezzük, hogy a K test bázisa természetesen $\{1\}$.

1. Definíció szerint $\mathbf{V}[\tilde{\mathbf{v}}]^{(1)} = [\mathbf{v}_i^* \tilde{\mathbf{v}}(1)]_{k,1} = [\mathbf{v}_i^* \mathbf{v}]_{k,1}$; ami nem más, mint a \mathbf{v} mátrixa.

2. Definíció szerint ${}^{(1)}[\mathbf{u}^*]\mathbf{U} = [1^* \mathbf{u}^*(\mathbf{u}_i)]_{1,n}$. Ennek a mátrixnak az elemei ugyanazok, mint az \mathbf{u} elemei, csak nem egy oszlopban, hanem egy sorban vannak elrendezve; így valóban $[\mathbf{u}^*] = [\mathbf{u}]^\dagger$.

3. Definíció szerint φ mátrixa az adott bázisokban $[\varphi] = [\mathbf{v}_p^*(\varphi(\mathbf{u}_q))]_{k,n}$, ahol p és q a „futó” sor-, illetve oszlopindex. Ennek a mátrixnak a j -edik oszlopát úgy kaphatjuk, hogy q helyébe a rögzített j indexet írjuk. Ekkor a $[\mathbf{v}_p^*(\varphi(\mathbf{u}_j))]_{k,1}$ egyoszlopú mátrixot kapjuk. A vektorok koordinátájára adott utolsó formula szerint ez nem más, mint a $\varphi(\mathbf{u}_j)$ mátrixa az adott bázisban.

4. Ismét induljunk ki a $[\varphi] = [\mathbf{v}_p^*(\varphi(\mathbf{u}_q))]_{k,n}$ felírásból. Most az i -edik sorra $[\mathbf{v}_i^*(\varphi(\mathbf{u}_q))]_{i,n}$ adódik. A 2. ponthoz hasonlóan kapjuk, hogy ez a $\mathbf{v}_i^* \varphi$ leképezés mátrixa.

5. Ez az eredmény azonnal adódik a 4. és 5. pont egymás utáni alkalmazásával.

6. Definíció szerint $[\iota]_{n,n} = [\mathbf{u}_i^*(\iota(\mathbf{u}_j))]_{n,n} = [\mathbf{u}_i^*(\mathbf{u}_j)]_{n,n} = [\delta_{i,j}]_{n,n}$ \blacksquare

Megjegyzés. Érdemes figyelni arra a 3. pont alapján adódó fontos tényre, hogy egy leképezés mátrixában a j -edik oszlopban pontosan a j -edik bázisvektor képének koordinátái állnak. \square

4.5. Tétel. *Rögzített bázisok esetén*

$$[c \cdot \varphi] = c \cdot [\varphi], \quad [\alpha + \beta] = [\alpha] + [\beta], \quad [\sigma \tau] = [\sigma][\tau],$$

ha a megfelelő műveletek elvégezhetők.

Bizonyítás. Az első állítás azonnal következik a $\mathbf{v}_i^*((c\varphi)(\mathbf{u}_i)) = c\mathbf{v}_i^*((\varphi)(\mathbf{u}_i))$ felírásból és a mátrixokra vonatkozó $[cF] = c[F]$ összefüggésből.

Legyen most $\alpha, \beta : \mathcal{U} \rightarrow \mathcal{V}$ az $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ bázisokkal. A megfelelő mátrixokban az i -edik sor j -edik eleme: $\mathbf{v}_i^*(\alpha + \beta)(\mathbf{u}_j)$, illetve $\mathbf{v}_i^*\alpha(\mathbf{u}_j)$ és $\mathbf{v}_i^*\beta(\mathbf{u}_j)$. A leképezésekkel végezhető műveletek alapján az első elem a másik kettő összege. A mátrixok definíciója szerint tehát valóban igaz az $[\alpha + \beta] = [\alpha] + [\beta]$ összefüggés.

Tekintsük végezetül az $\mathcal{U} \xrightarrow{\tau} \mathcal{V} \xrightarrow{\sigma} \mathcal{W}$ leképezéseket a megfelelő $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$, $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ és $\mathbf{W} = \{\mathbf{w}_1, \dots, \mathbf{w}_\ell\}$ bázisokkal együtt. A szorzatleképezés mátrixában az i -edik sor j -edik elemére a következőket kapjuk:

$$\begin{aligned} i[\sigma \tau]_j &= \mathbf{w}_i^* \sigma \tau \mathbf{u}_j = \mathbf{w}_i^* \sigma \iota_{\mathcal{V}} \tau \mathbf{u}_j = \mathbf{w}_i^* \sigma \left(\sum_t \mathbf{v}_t \mathbf{v}_t^* \right) \tau \mathbf{u}_j = \\ &= \sum_t (\mathbf{w}_i^* \sigma \mathbf{v}_t \mathbf{v}_t^* \tau \mathbf{u}_j) = \sum_t (\mathbf{w}_i^* \sigma \mathbf{v}_t) (\mathbf{v}_t^* \tau \mathbf{u}_j) = [\mathbf{w}_i^* \sigma][\tau(\mathbf{u}_j)] = i[\sigma][\tau]_j. \end{aligned}$$

Ezzel a szorzásra vonatkozó állítást is bizonyítottuk. ■

Megjegyzés. Tulajdonképpen a mátrixműveleteket a fenti összefüggések alapján lehetne definiálni. Ez rámutat arra, hogy miért így definiáltuk a mátrixokkal végezhető műveleteket. Egy kérdés azonban nyitva maradna ennél a definíciónál, nevezetesen az, hogy a műveletek nem függenek-e attól, hogy melyik bázisban írjuk fel a leképezések mátrixát. A bázistól való függetlenség azonban világos, hiszen a műveletek eredménye csak az előre megadott mátrixok elemeitől függ. □

A mátrixok esetében még egy „művelet” szerepelt, a transzponálás. Most azt fogjuk megnézni, hogy ez „miképpen hat vissza” a leképezésekre. Már előljáróban megjegyezzük, hogy az eredmény bázisfüggő, azaz a leképezés transzponáltja függ a felvett bázistól.

4.4. Definíció. Legyenek adva a K feletti \mathcal{U} és \mathcal{V} vektorterek az $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ bázisokkal. A $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezésnek a fenti bázisban felírt $\varphi^* : \mathcal{V} \rightarrow \mathcal{U}$ transzponáltján azt a leképezést értjük, amelynek a fenti bázisokban felírt mátrixára $[\varphi^*] = [\varphi]^\dagger$ teljesül. ■

Megjegyzés. Lehetséges, hogy különböző bázisok esetében is ugyanaz a transzponált. Annak a leírására, hogy az ilyen bázisok milyen kapcsolatban állnak egymással, az euklideszi tereknél kerül sor. □

4.6. Tétel. *Érvényesek az alábbiak:*

- (1) $\tilde{\mathbf{u}}^* = \mathbf{u}^*$.
- (2) $\varphi^* = \sum_{i,j} (\tilde{\mathbf{u}}_j \mathbf{v}_i^* \varphi \tilde{\mathbf{u}}_j \mathbf{v}_i^*)$.
- (3) $(c\varphi)^* = c\varphi^*$, $(\varphi^*)^* = \varphi$, $(\alpha + \beta)^* = \alpha^* + \beta^*$, $(\sigma\tau)^* = \tau^* \sigma^*$.
- (4) $r(\varphi^*) = r(\varphi)$.

Bizonyítás. (1) azonnal következik abból, hogy a 4.4. Tétel szerint $[\mathbf{u}^*] = [\mathbf{u}]^\dagger$, ami definíció szerint $[\tilde{\mathbf{u}}^*]$; és adott bázis esetén a leképezéseket egyértelműen meghatározza a mátrixuk.

(2) bizonyításához is azt mutatjuk meg, hogy a két leképezésnek ugyanaz a mátrixa az adott bázisban. Definíció szerint ${}_p[\varphi^*]_q = {}_q[\varphi]_p$. Másrészt a $\psi = \sum_{i,j} (\tilde{\mathbf{u}}_j \mathbf{v}_i^* \varphi \tilde{\mathbf{u}}_j \mathbf{v}_i^*)$

leképezésre világos, hogy $\psi : \mathcal{V} \rightarrow \mathcal{U}$, továbbá:

$$\begin{aligned} {}_p[\psi]_q &= \mathbf{u}_p^* \psi \mathbf{v}_q = \sum_{i,j} (\mathbf{u}_p^* \mathbf{u}_j \mathbf{v}_i^* \varphi \mathbf{u}_j \mathbf{v}_i^* \mathbf{v}_q) = \\ &= \sum_{i,j} \delta_{j,p} \mathbf{v}_i^* \varphi \mathbf{u}_j \delta_{i,q} = \mathbf{v}_q^* \varphi \mathbf{u}_p = {}_q[\varphi]_p. \end{aligned}$$

A (3) alatti összefüggések azonnal adódnak abból, hogy a megfelelő összefüggések a mátrixokra igazak és rögzített bázisok esetében a leképezések és mátrixuk egyértelműen meghatározzák egymást.

(4) bizonyításához írjuk fel φ -t diádok összegeként a 4.1. Tétel (3) pontjának megfelelően úgy, hogy a tagok száma megegyezzzék φ rangjával: $\varphi = \tilde{\mathbf{v}}_1 \mathbf{u}_1^* + \dots + \tilde{\mathbf{v}}_r \mathbf{u}_r^*$. A (3) alatti harmadik, negyedik, majd az első egyenlőség szerint:

$$\begin{aligned} \varphi^* &= (\tilde{\mathbf{v}}_1 \mathbf{u}_1^* + \dots + \tilde{\mathbf{v}}_r \mathbf{u}_r^*)^* = (\tilde{\mathbf{v}}_1 \mathbf{u}_1^*)^* + \dots + (\tilde{\mathbf{v}}_r \mathbf{u}_r^*)^* = \\ &= (\mathbf{u}_1^*)^* (\tilde{\mathbf{v}}_1)^* + \dots + (\mathbf{u}_r^*)^* (\tilde{\mathbf{v}}_r)^* = \tilde{\mathbf{u}}_1 \mathbf{v}_1^* + \dots + \tilde{\mathbf{u}}_r \mathbf{v}_r^*. \end{aligned}$$

A 4.1. Tétel (4) pontja szerint tehát $r(\varphi^*) \leq r(\varphi)$. Ezt az eredményt φ helyett φ^* -ra alkalmazva $(\varphi^*)^* = \varphi$ alapján kapjuk, hogy $r(\varphi) \leq r(\varphi^*)$. ■

Megjegyzés. Tulajdonképpen a (3) alatti összefüggések bizonyításához sincs szükség a mátrixokról tanultakra. A (2) alatti összefüggés alapján ezek számolással igazolhatók. □

Feladatok

1. Legyen K egy véges, q -elemű test és \mathcal{U}_K egy n -dimenziós vektortér (tudjuk, hogy q lehet prímszám és be lehet látni, hogy prímszámhatvány is; az is igaz, hogy más eset nem lehetséges).

1. Hány különböző bázisa van \mathcal{U} -nak?
2. Hány bázisban fordulnak elő ugyanazok a vektorok?
3. Milyen oszthatósági kapcsolatra következtethetünk a fentiekből?

2. Legyen $\varphi = \tilde{\mathbf{e}}_1 \mathbf{f}_1^* + \dots + \tilde{\mathbf{e}}_s \mathbf{f}_s^*$ a $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezés egy felírása, ahol $\mathbf{e}_1, \dots, \mathbf{e}_s \in \mathcal{V}$ és $\mathbf{f}_1, \dots, \mathbf{f}_s \in \mathcal{U}$. Bizonyítsuk be, hogy $s = r(\varphi)$ akkor és csak akkor teljesül, ha mind az \mathbf{e}_i , mind az \mathbf{f}_i vektorok függetlenek.

3. Mutassuk meg, hogy létezik olyan $\varphi = \tilde{\mathbf{e}}_1 \mathbf{f}_1^* + \dots + \tilde{\mathbf{e}}_s \mathbf{f}_s^*$ felírás, amelyben például az \mathbf{e}_i vektorok függetlenek, de az \mathbf{f}_i vektorok nem.

4. Tegyük fel, hogy a $\varphi = \tilde{\mathbf{e}}_1 \mathbf{f}_1^* + \dots + \tilde{\mathbf{e}}_s \mathbf{f}_s^*$ felírásban \mathbf{f}_s függ a többi \mathbf{f}_i -től: $\mathbf{f}_s = \sum_{i=1}^{s-1} c_i \mathbf{f}_i$.

Helyettesítsük be \mathbf{f}_s -nek ezt a kifejezését a φ -re adott formulába. Mutassuk meg, hogy megfelelő összevonások után φ -re egy „rövidebb” előállítást kapunk.

5. Tegyük fel, hogy egy leképezést előállítottunk diádok összegeként. Az előző feladat felhasználásával mutassuk meg, hogy ebből az előállításból konstruálhatunk egy olyan előállítást, amelyben a diádok száma megegyezik a leképezés rangjával.

6. Írjuk fel egy projekció mátrixát egy olyan bázisban, amelynek minden egyes eleme vagy a projekció magterében, vagy a projekció képterében van.

7. Legyen $\varphi \in \text{End}(\mathcal{U})$, $\mathcal{V} \leq \mathcal{U}$ a φ -nek invariáns altere, \mathbf{V} a \mathcal{V} -nek és \mathbf{U} az \mathcal{U} -nak egy \mathbf{V} -t tartalmazó bázisa. Milyen lesz φ mátrixa ebben a bázisban?

8. Határozzuk meg a $\varphi = \alpha \oplus \beta$ mátrixát alkalmas bázisban.

9. Legyenek \mathbf{u}_i -k az \mathbf{U} tér \mathbf{U} bázisának és \mathbf{v}_j -k a \mathcal{V} tér \mathbf{V} bázisának az elemei. Határozzuk meg ezekben a bázisokban a $\mathbf{v}_j \mathbf{u}_i^*$ és általában a $\mathbf{v} \mathbf{u}^*$ mátrixát. Milyen esetben tudjuk meghatározni az $\mathbf{u}^* \mathbf{v}$ mátrixát? Ebben az esetben határozzuk is meg.

10. Legyenek \mathcal{U} és \mathcal{V} nem feltétlenül véges dimenziós vektorterek a K test felett; és legyenek \mathbf{U} , \mathbf{V} a megfelelő vektorterek bázisai. Milyen feltételnek tesznek eleget e vektorterek elemeinek a koordinátái? Milyen mátrixok állnak elő $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezések mátrixaiként?

11. Fogalmazzuk meg, miképpen kell végezni a végtelen dimenziós vektorterek leképezéseinek vizsgálatánál keletkező mátrixokkal végzett műveleteket.

12. Mutassuk meg, hogy bázisokkal rendelkező \mathbb{Z} -modulusok leképezéseinek mátrixában az elemek \mathbb{Z} -beliek (azaz egész számok), és minden egész elemű véges mátrix előáll ilyen módon. Mit mondhatunk a végtelen mátrixok esetében?

13. Mutassuk meg, hogy ha egy R -modulusnak van egy \mathbf{G} generátorrendszere, akkor minden eleméhez hozzá tudunk rendelni egy mátrixot a tárgyalat módon. (Mikor egyértelmű ez a mátrix minden vektor esetében?) Állapítsuk meg, hogy milyen kapcsolatnak kell fennállnia két mátrix között, ha azok ugyanazt a vektort jellemzik. Bizonyítsuk be, hogy ez a kapcsolat „a priori” ekvivalenciareláció. Írjuk le „numerikusan” ezt a relációt.

14. Tekintsük a $\mathbb{Q}^+ \mathbb{Z}$ -nek a $G = \left\{ \mathbf{u}_n = \frac{1}{n!} \mid n \in \mathbb{N} \right\}$ generátorrendszerét. Írjuk fel az $\frac{a}{b} \in \mathbb{Q}$

törtnek a koordinátáit ebben a generátorrendszerben. Állapítsuk meg, hogy két ilyen koordináta mikor jellemzi ugyanazt a törtet. Határozzuk meg a fenti modulus endomorfizmusainak mátrixait; illetve e mátrixok ekvivalenciaosztályait az előző feladatnál tárgyalat ekvivalenciarelációnál.

2. Áttérés új bázisra

A különféle geometriai alakzatok felismeréséhez szükséges, hogy azok olyan koordináta-rendszerben legyenek felírva, amely lehetővé teszi geometriai jellemzőik felismerését. Átvitt értelemben ezek az alakzatok fizikai objektumok leírására is szolgálnak. A matematikai vizsgálatok más ágaiban is lényeges, hogy ezekre a kérdésekre válaszolni tudjunk. Egyelőre még nem tudjuk azt, hogy milyen alakzatokat kell vizsgálni, és az sem világos, hogy mely koordináta-rendszerben árulják el az alakzatok jellegüket. Mielőtt ezekre rátérnénk, szükség van arra, hogy miképpen térhetünk át új koordináta-rendszerre. Egyelőre azt nézzük meg, hogy miképpen változik meg egy vektor koordinátája vagy egy leképezés mátrixa akkor, ha új koordináta-rendszert — azaz új bázist — választunk.

Bármely leképezés esetében (tudjuk, hogy a vektorokat magukat is tekinthetjük leképezéseknek) két vektortér szerepel; az, amit leképezünk és az, ahova leképezünk. Ennek megfelelően mindkét vektortérben meg kell vizsgálni a bázisváltoztatás hatását. Egy dologra azért nagyon kell vigyázni; ha egyetlen vektortér lineáris transzformációit nézzük, akkor célszerű a két kiinduló bázist ugyanannak venni, és ennek megfelelően a bázist mindkét esetben ugyanúgy kell változtatni. Természetesen a K alaptestben a bázis mindig a rögzített $\{1\}$.

A bizonyítás során néha zavaró a sokféle jelölés; éppen ezért eleve rögzítünk minden előforduló jelölést, amelyeket az új bázisra való áttérésnél használni fogunk.

Jelölés. Legyen adva a K test feletti \mathcal{U} és \mathcal{V} vektortér, megfelelően az $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és a $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ bázisokkal. Vegyünk fel a két vektortérben egy-egy új bázist, \mathcal{U} -ban az $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bázist és \mathcal{V} -ben az $\mathbf{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_k\}$ bázist.

Legyen $\sigma : \mathcal{U} \rightarrow \mathcal{U}$ az a transzformáció, amelyre $\sigma(\mathbf{u}_j) = \mathbf{e}_j$ és $\tau : \mathcal{V} \rightarrow \mathcal{V}$ az a transzformáció, amelyre $\tau(\mathbf{v}_i) = \mathbf{f}_i$ ($1 \leq j \leq n$ és $1 \leq i \leq k$). Ezeket a transzformációkat kísérő transzformációknak fogjuk nevezni.

Legyen $\mathbf{U}^* = \{\mathbf{u}_1^*, \dots, \mathbf{u}_n^*\}$ az \mathbf{U} -nak, $\mathbf{V}^* = \{\mathbf{v}_1^*, \dots, \mathbf{v}_k^*\}$ a \mathbf{V} -nek, $\mathbf{E}^+ = \{\mathbf{e}_1^+, \dots, \mathbf{e}_n^+\}$ az \mathbf{E} -nek és $\mathbf{F}^+ = \{\mathbf{f}_1^+, \dots, \mathbf{f}_k^+\}$ az \mathbf{F} -nek megfelelő duális bázis.

4.7. Tétel. A fenti jelölést használva legyen $\varphi : \mathcal{U} \rightarrow \mathcal{V}$. Ekkor

$$\mathbf{F}_{[\varphi]} \mathbf{E} = \mathbf{V}_{[\tau^{-1}]} \mathbf{V} \cdot \mathbf{V}_{[\varphi]} \mathbf{U} \cdot \mathbf{U}_{[\sigma]} \mathbf{U}.$$

Bizonyítás. Definíció szerint $\mathbf{F}_{[\varphi]} \mathbf{E} = [\mathbf{f}_i^+(\varphi(\mathbf{e}_j))]_{k,n}$.

Itt $\mathbf{e}_j = \sigma(\mathbf{u}_j)$. Az \mathbf{f}_i^+ meghatározásához vegyük figyelembe, hogy $\mathbf{f}_i = \tau(\mathbf{v}_i)$. Ebből azonnal kapjuk, hogy $\mathbf{f}_i^+ \tau(\mathbf{v}_s) = \mathbf{f}_i^+ \mathbf{f}_s = \delta_{i,s}$. Ez azt jelenti, hogy a \mathbf{V} bázison $\mathbf{f}_i^+ \tau$ és \mathbf{v}_i^* ugyanúgy hatnak; tehát megegyeznek. Ezért $\mathbf{f}_i^+ = \mathbf{v}_i^* \tau^{-1}$ (τ egy bázist bázisba visz, tehát létezik inverze). Behelyettesítve a kérdéses mátrixra $[\mathbf{v}_i^* \tau^{-1}(\varphi(\sigma(\mathbf{e}_j)))]_{k,n}$ adódik, ami nem más, mint a $\tau^{-1} \varphi \sigma$ leképezésnek az eredeti bázisokban felírt mátrixa. Mivel adott bázisokban a szorzat mátrixa megegyezik a tényezők mátrixának a szorzatával, ezért valóban:

$$\mathbf{F}_{[\varphi]} \mathbf{E} = \mathbf{V}_{[\tau^{-1}]} \mathbf{V} \cdot \mathbf{V}_{[\varphi]} \mathbf{U} \cdot \mathbf{U}_{[\sigma]} \mathbf{U}.$$

■

Speciális esetként adódik az alábbi

Kiegészítés. Az $\mathcal{U} = \mathcal{V}$ esetben (ekkor feltétel szerint $\mathbf{V} = \mathbf{U}$, $\mathbf{F} = \mathbf{E}$ és $\tau = \sigma$):

$$\mathbf{E}_{[\varphi]} \mathbf{E} = \mathbf{U}_{[\sigma^{-1}]} \mathbf{U} \cdot \mathbf{U}_{[\varphi]} \mathbf{U} \cdot \mathbf{U}_{[\sigma]} \mathbf{U}.$$

Az $\mathcal{U} = K$ esetben (ekkor feltétel szerint $\mathbf{U} = \mathbf{E} = \{1\}$ és $\sigma = \iota$):

$$\mathbf{F}[\tilde{\mathbf{v}}] = \mathbf{V}[\tau^{-1}] \mathbf{V} \cdot \mathbf{V}[\tilde{\mathbf{v}}], \quad \text{azaz} \quad \mathbf{F}[\mathbf{v}] = \mathbf{V}[\tau^{-1}] \mathbf{V} \cdot \mathbf{V}[\mathbf{v}].$$

A $\mathcal{V} = K$ esetben (ekkor feltétel szerint $\mathbf{V} = \mathbf{F} = \{1\}$ és $\tau = \sigma$):

$$[\mathbf{u}^*] \mathbf{E} = [\mathbf{u}^*] \mathbf{U} \cdot \mathbf{U}_{[\sigma]} \mathbf{U}. \blacksquare$$

Megjegyzések

1. Írjuk fel σ mátrixát az \mathbf{E} bázisban:

$$\mathbf{E}_{[\sigma]} \mathbf{E} = \mathbf{U}_{[\sigma^{-1}]} \mathbf{U} \cdot \mathbf{U}_{[\sigma]} \mathbf{U} \cdot \mathbf{U}_{[\sigma]} \mathbf{U}.$$

Ez nem más, mint $\mathbf{U}_{[\sigma^{-1} \sigma \sigma]} \mathbf{U} = \mathbf{U}_{[\sigma]} \mathbf{U}$, azaz a kísérő transzformációk mátrixa az új bázisban is ugyanaz, mint a régiben.

2. Az új bázisra való áttérésnél szerepel egy $[\tau^{-1}]$ alakú mátrix. Mivel a szorzat mátrixa egyenlő a mátrixok szorzatával, ezért $[\tau^{-1}][\tau] = [\tau^{-1} \tau] = [\iota]$. Azt már tudjuk, hogy $[\iota]$ mátrixában a fődiagonális minden eleme 1, s a többi elem 0. Ez a megfelelő méretű egységmátrix, és ezért $[\tau^{-1}]$ a $[\tau]$ mátrix inverze: $[\tau]^{-1}$. Az inverz mátrix meghatározásával a következő pontban foglalkozunk. \square

Feladatok

1. Legyenek $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ az \mathcal{U} vektortér bázisai. Legyen $\mathbf{e}_i = \sum_j c_{i,j} \mathbf{u}_j$. Írjuk fel a σ kísérő transzformáció mátrixát.

2. Legyenek $\mathbf{E} = \mathbf{U}$ (nem $\mathbf{E} \equiv \mathbf{U}$) az \mathcal{U} tér bázisai. Írjuk fel a σ kísérő transzformáció mátrixát.

3. Határozzuk meg az új bázisról a régire való áttérés kísérő transzformációinak a mátrixát.

4. Hogyan változnak meg a fent tárgyalt tételek, ha a vektorterek nem feltétlenül véges dimenziósak?

5. Tekintsük a modulo 30 vett egész számokat mint \mathbb{Z} -modulust (\mathbb{Z}_{30}). Bizonyítsuk be, hogy a következő „vektorrendszerek” bázist alkotnak: $\{1\}$, $\{6, 25\}$, $\{10, 21\}$, $\{15, 16\}$, $\{6, 10, 15\}$. Írjuk fel az n „vektor” koordinátáit e bázisokban. Határozzuk meg az ezekhez tartozó kísérő transzformációk mátrixát (a megfelelő bázisban). Bizonyítsuk be, hogy ennek a modulusnak minden homomorfizmusa $\varphi_n : k \rightarrow n \cdot k$ alakú ($n, k \in \mathbb{N} \cup \{0\}$). Határozzuk meg, hogy miképpen változnak a mátrixok az új bázisra való áttérésnél.

3. Mátrix rangja és inverze

Ha egy leképezés mátrixa egy alkalmas bázisban esetleg diagonális alakú, akkor világos, hogy az a leképezés a bázisvektorok irányában hogyan hat; amiből a leképezés teljes hatása könnyen meghatározható. Általában a leképezések nem ilyenek, de majd látni fogjuk, hogy minden leképezéshez tartozik olyan bázis, amelyikben a leképezés mátrixa eléggé elárulja a leképezés hatását. Az új bázisra való áttérésnél általában egyenként veszünk fel egy-egy új bázisvektort (hasonlóképpen, mint a kicserélési tételnél). Éppen ezért hasznos, ha megnézzük, milyen alakú az ilyen leképezések mátrixa.

Láttuk, hogy ha $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ a \mathcal{V} vektortér egy-egy bázisa, akkor a $\{\tilde{\mathbf{v}}_j \mathbf{u}_i^* \mid 1 \leq i \leq n, 1 \leq j \leq k\}$ vektorok a $\text{Hom}(\mathcal{U}, \mathcal{V})$ egy bázisát alkotják. A $\mathcal{V} = \mathcal{U}$ esetben $\mathbf{V} \equiv \mathbf{U}$. Ekkor a fenti báziselemekre $\varepsilon_{i,j} = \tilde{\mathbf{u}}_j \mathbf{u}_i^*$ adódik. Ezek a leképezések a következő „szabály” szerint szorozhatók össze: $\varepsilon_{i,j} \cdot \varepsilon_{p,q} = \delta_{j,p} \cdot \varepsilon_{i,q}$. Ezek mindegyikének 1 a rangja. Az $i = j$ esetben a négyzete önmaga, az $i \neq j$ esetben a négyzete a null-leképezés. Ezeknek az eseteknek az általánosítására is szükségünk van:

4.5. Definíció. Az $\alpha \in \text{End}(\mathcal{U})$ elemet nilpotensnek nevezzük, ha van olyan $n \in \mathbb{N}$, amelyre $\alpha^n = 0$. Az $\alpha \in \text{End}(\mathcal{U})$ elemet idempotensnek nevezzük, ha $\alpha^2 = \alpha$. ■

4.8. Tétel. Ha $\alpha \in \text{End}(\mathcal{U})$ nilpotens, akkor $\iota - \alpha$ is és $\iota + \alpha$ is invertálható; az $\alpha^2 = 0$ esetben $\iota - \alpha$ és $\iota + \alpha$ egymás inverzei.

Bizonyítás. Ha $\alpha^k = 0$, akkor triviális számolással kapjuk, hogy

$$(\iota - \alpha) \cdot (\iota + \alpha + \dots + \alpha^{k-1}) = \iota - \alpha^k = \iota.$$

A második állítás abból következik, hogy $\alpha^k = 0$ esetén $(-\alpha)^k = 0$ is igaz. A fenti bizonyításból az $\alpha^2 = 0$ esetben $(\iota - \alpha)(\iota + \alpha) = \iota$ következik. ■

4.9. Tétel. Tegyük fel, hogy az $\text{End}(\mathcal{U})$ valamely $\varepsilon \neq 0$ elemére $\varepsilon^2 = \varepsilon$. Ekkor $\iota + c\varepsilon$ pontosan akkor invertálható, ha $c \neq -1$. Ekkor $(\iota + c\varepsilon)^{-1} = \iota - \frac{c}{1+c}\varepsilon$.

Bizonyítás. A $c \neq -1$ esetben létezik $d = \frac{c}{1+c}$, és így

$$(\iota + c\varepsilon)(\iota - d\varepsilon) = \iota + (c - d - cd)\varepsilon = \iota + (c - d(1+c))\varepsilon = \iota.$$

Amennyiben $c = -1$, akkor $\varepsilon \neq 0$ miatt létezik olyan $\mathbf{u} \in \mathcal{U}$ vektor, amelyre $\mathbf{v} = \varepsilon(\mathbf{u}) \neq 0$. Erre $\varepsilon = \varepsilon^2$ miatt

$$(\iota - \varepsilon)(\mathbf{v}) = \iota(\mathbf{v}) - \varepsilon(\mathbf{v}) = \mathbf{v} - \varepsilon^2(\mathbf{u}) = \mathbf{v} - \varepsilon(\mathbf{u}) = \mathbf{v} - \mathbf{v} = 0,$$

ami azt jelenti, hogy $\text{Ker}(\iota - \varepsilon) \neq \{0\}$; tehát $\iota - \varepsilon$ nem invertálható. ■

4.6. Definíció. Az \mathcal{U} vektortérnek a rögzített \mathbf{U} bázishoz tartozó $\iota + c\varepsilon_{i,j}$ alakú endomorfizmusait elemi transzformációknak nevezzük, kivéve, ha $i = j$ és $c = -1$. ■

Megjegyzések

1. Az elemi transzformációk valójában a mátrixok elemi átalakításainak felelnek meg, mint azt látni fogjuk. A célunk az, hogy a vektortérben elemről elemre újabb és újabb bázist keressünk, míg végre a leképezés mátrixa kívánalmainknak megfelelő alakot ölt. A vizsgált bázisokhoz tartozó elemi transzformációk minden lépésnél megváltoznak. Tekintettel arra, hogy az adott transzformáció mátrixa is „ugyanolyan módon változik” minden lépésnél, ezért e mátrixoknak megfelelő elemi átalakításoknak a „mátrixleírása” mindig ugyanolyan lesz.

2. A 4.8. és a 4.9. Tételek alapján egy elemi transzformációnak az inverze is elemi transzformáció; mégpedig ugyanolyan „jellegű”, mint az eredeti.

3. Az elemi transzformációknak két különböző típusa van, annak megfelelően, hogy $i = j$ vagy sem. Ezeket, illetve a megfelelő mátrixukat a 4.11. Tételben fogjuk jellemezni. \square

4.10. Tétel. Legyen $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} vektortér adott bázisa, és $\sigma = \iota + c\varepsilon_{i,j}$ egy ezen bázishoz tartozó elemi transzformáció. Ekkor a $\sigma(\mathbf{U}) = \{\sigma(\mathbf{u}_1), \dots, \sigma(\mathbf{u}_n)\}$ bázis úgy áll elő az eredetiből, hogy az i -edik bázisvektor c -szeresét hozzáadjuk a j -edik bázisvektorhoz; és a j -edik bázisvektort kivéve a többit változatlanul hagyjuk.

Bizonyítás. Az $\varepsilon_{i,j}(\mathbf{u}_t) = \tilde{\mathbf{u}}_j \mathbf{u}_i^*(\mathbf{u}_t) = \tilde{\mathbf{u}}_j(\delta_{i,t}) = \delta_{i,t} \mathbf{u}_j$ összefüggés alapján

$$\sigma(\mathbf{u}_t) = \mathbf{u}_t + c\delta_{i,t} \mathbf{u}_j = \begin{cases} \mathbf{u}_t, & \text{ha } t \neq i \\ \mathbf{u}_i + c\mathbf{u}_j, & \text{ha } t = i, \end{cases}$$

mint állítottuk. \blacksquare

Megjegyzés. Az $i = j$ esetben a tétel szerint az i -edik bázisvektort kell megszorozni $(1+c)$ -vel. Ha $c = -1$, akkor ennél a transzformációnál az i -edik bázisvektor $\mathbf{0}$ -ba megy át; tehát az eredeti bázis képe nem lesz bázis. \square

4.11. Tétel. Legyen $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} vektortér egy bázisa és legyen σ egy elemi transzformáció. Ekkor $\mathbf{U}[\sigma]^\mathbf{U}$ a következő alakú:

(1) Ha $i = j$, akkor a mátrix fődiagonálisában az i -edik elem $1 + c$, a fődiagonális többi eleme 1 és a mátrix többi eleme 0.

(2) Az $i \neq j$ esetben a fődiagonális minden eleme 1, a j -edik sor i -edik eleme c és a mátrix többi eleme 0.

Bizonyítás. Mint tudjuk, $[\iota + c\varepsilon_{i,j}] = [\iota] + c[\varepsilon_{i,j}]$. ι mátrixában a fődiagonális minden eleme 1 és a mátrix többi eleme 0. Továbbá:

$$p[\varepsilon_{i,j}]_q = \mathbf{u}_p^* \tilde{\mathbf{u}}_j \mathbf{u}_i^* \tilde{\mathbf{u}}_q(1) = \delta_{p,j} \cdot \delta_{i,q} = \begin{cases} 1, & \text{ha } p = j, q = i \\ 0 & \text{máskor,} \end{cases}$$

amiből azonnal következik az állítás. \blacksquare

4.12. Tétel. Legyen $\mathcal{W} \xrightarrow{\psi} \mathcal{U} \xrightarrow{\varphi} \mathcal{V}$, továbbá $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} vektortér egy bázisa és az $\varepsilon_{i,j} = \tilde{\mathbf{u}}_j \mathbf{u}_i^*$ jelöléssel $\sigma = \iota + c\varepsilon_{i,j}$ az \mathcal{U} egy elemi transzformációja.

Ezt az elemi transzformációt végrehajtva az új bázisban φ mátrixa úgy kapható az eredetiből, hogy annak i -edik oszlopához hozzáadjuk a j -ediknek a c -szeresét. A ψ esetében annak j -edik sorából kivonjuk a i -ediknek a d -szeresét, ahol $\sigma^{-1} = \iota - d\varepsilon_{i,j}$.

Ezekben az esetekben azt mondjuk, hogy a mátrixon elemi átalakítást végeztünk.

Bizonyítás. A φ mátrixa az új bázisban ugyanaz, mint $\varphi\sigma$ a régiben. Tekintsük ennek a mátrixnak az r -edik oszlopát:

$$\begin{aligned} [\varphi\sigma]_r &= [\varphi\sigma\mathbf{u}_r] = [\varphi\mathbf{u}_r] + c[\varphi\mathbf{u}_j\mathbf{u}_i^*\mathbf{u}_r] = [\varphi\mathbf{u}_r] + c\delta_{i,r}[\varphi\mathbf{u}_j] = \\ &= \begin{cases} [\varphi]_r, & \text{ha } r \neq i \\ [\varphi]_i + c[\varphi]_j, & \text{ha } r = i. \end{cases} \end{aligned}$$

ψ esetében az új bázisban mátrixa megegyezik $\tau\psi$ mátrixával a régi bázisban, ahol $\tau = \sigma^{-1}$. A 4.6. Definíció utáni második megjegyzés szerint τ ugyanolyan jellegű, mint σ . A fenti bizonyításhoz hasonlóan:

$$\begin{aligned} s[\tau\psi] &= [\mathbf{u}_s^*\tau\psi] = [\mathbf{u}_s^*] - d[\mathbf{u}_s^*\mathbf{u}_j\mathbf{u}_i\psi] = [\mathbf{u}_s^*] - d\delta_{s,j}[\mathbf{u}_i^*] = \\ &= \begin{cases} s[\psi], & \text{ha } s \neq j \\ j[\psi] - d_i[\psi], & \text{ha } s = j. \end{cases} \end{aligned}$$

Ezzel állításunkat bizonyítottuk. ■

Megjegyzés. A mátrixoknak kétféle elemi átalakítása van. Az $i \neq j$ esetben egy sorhoz (oszlophoz) egy *tőle különböző* sor (oszlop) skalárszorosat adjuk hozzá. Az $i = j$ esetben viszont egy sort (oszlopot) szorzunk egy nemnulla skalárral.

A determinánsok tárgyalásánál láttuk, hogy négyzetes mátrixok esetében az első fajta elemi átalakítás nem változtatja meg a mátrix determinánsát; míg a második fajtnál a determináns a szereplő skalárral szorozódik. (Ez egyébként akkor is igaz, ha a skalár 0.) □

Láttuk, hogy milyen kapcsolat áll fenn egy leképezés mátrixai között. Most azt nézzük meg, hogy mi mondható azokról a leképezésekről, amelyeknek a mátrixa megegyezik. Erre azért van szükség, mert általában a mátrixok adóttak.

4.13. Tétel. Ha ${}^V[\alpha]{}^U = {}^F[\beta]{}^E$, akkor vannak olyan σ, τ invertálható homomorfizmusok, amelyekre $\beta = \tau\alpha\sigma^{-1}$.

Bizonyítás. Eleve nem tesszük fel azt sem, hogy α és β ugyanazokat a vektortereket és ugyanoda képezi le. Legyen \mathbf{V} a \mathcal{V} vektortérnek, \mathbf{U} az \mathcal{U} vektortérnek, \mathbf{F} az \mathcal{F} vektortérnek és \mathbf{E} az \mathcal{E} vektortérnek a bázisa. Mivel a mátrixok egyező alakúak, ezért $|\mathbf{V}| = |\mathbf{F}|$ és $|\mathbf{U}| = |\mathbf{E}|$, tehát a \mathcal{V} és az \mathcal{F} , valamint az \mathcal{U} és az \mathcal{E} vektorterek egyenlő dimenziósak. Léteznek tehát $\sigma : \mathcal{U} \rightarrow \mathcal{E}$ és $\tau : \mathcal{V} \rightarrow \mathcal{F}$ izomorfizmusok. Azt is feltehetjük, hogy ezek az izomorfizmusok a felvett bázisokat viszik egymásba; még hozzá úgy, hogy a képelemnek ugyanaz az indexe, mint az eredetié. A jobb áttekinthetőség végett a fellépő leképezéseket egy úgynevezett *diagrammal* ábrázoljuk:

$$\begin{array}{ccc} \mathcal{U} & \xrightarrow{\alpha} & \mathcal{V} \\ \sigma \downarrow & & \downarrow \tau \\ \mathcal{E} & \xrightarrow{\beta} & \mathcal{F} \end{array}$$

Mivel a \mathcal{V} és az \mathcal{F} terek különbözöknek tekinthetők, mindkét térben jelölhetjük ugyanúgy a duális bázist. E két mátrix megegyezése azt jelenti, hogy a bázisok megfelelő elemeire $\mathbf{v}_i^*\alpha\mathbf{u}_j = \mathbf{f}_i^*\beta\mathbf{e}_j$. Vegyük most figyelembe, hogy $\mathbf{f}_i = \tau\mathbf{v}_i$ és $\mathbf{e}_j = \sigma\mathbf{u}_j$. A $\delta_{i,k} = \mathbf{f}_k^*\mathbf{f}_i = \mathbf{f}_k^*\tau\mathbf{v}_i$ összefüggésből következik, hogy $\mathbf{f}_k^*\tau = \mathbf{v}_k^*$ (hiszen $\mathbf{f}_k^*\tau$ \mathcal{V} -t képezi le az alaptestbe).

Ez utóbbit $\mathbf{v}_k^* = \mathbf{f}_k^* \tau^{-1}$ alakba írva azt nyerjük, hogy $\mathbf{v}_k^* \alpha \mathbf{u}_j = \mathbf{v}_k^* \tau^{-1} \beta \sigma \mathbf{u}_j$ teljesül minden indexpárra. A bázisvektorok lineáris függetlensége alapján ebből $\alpha = \tau^{-1} \beta \sigma$ következik, ami ekvivalens az eredeti állítással. ■

4.14. Tétel. *Ha az A mátrixra $A = [\alpha]$, akkor $r(\alpha)$ csak az A -tól függ. Ezt a számot az A rangjának nevezzük és $r(A)$ -val jelöljük. Egy mátrix rangja nem változik meg, ha a mátrixon elemi transzformációt hajtunk végre.*

Bizonyítás. Mindkét állítás azon alapszik, hogy ha $\alpha\beta$ létezik, akkor a szorzat rangja legfeljebb akkora, mint az egyes tényezőké.

Legyen $\mathcal{U} \xrightarrow{\beta} \mathcal{V} \xrightarrow{\alpha} \mathcal{W}$. Mivel $(\alpha\beta)(\mathbf{u}) = \alpha(\beta(\mathbf{u})) \in \text{Im}(\alpha)$, azaz $\text{Im}(\alpha\beta) \subseteq \text{Im}(\alpha)$, ezért $r(\alpha\beta) \leq r(\alpha)$. Ha $\sum_i c_i \beta(\mathbf{u}_i) = \mathbf{o}$, akkor $\sum_i c_i \alpha(\beta(\mathbf{u}_i)) = \mathbf{o}$ is igaz, vagyis $\text{Im}(\alpha\beta)$ -ban a lineárisan független elemek maximális száma legfeljebb annyi, mint $\text{Im}(\beta)$ -ban, azaz $r(\alpha\beta) \leq r(\beta)$.

Ebből könnyen bizonyítható, hogy ha a szorzat egyik tényezője invertálható, akkor a szorzat rangja egyenlő a másik tényező rangjával.

Legyen $\varphi = \alpha\psi\beta$, ahol α, β invertálhatóak. Az előbb belátottak szerint $r(\varphi) \leq r(\psi)$. Az invertálhatóságot felhasználva $\psi = \alpha^{-1}\varphi\beta^{-1}$ adódik. Itt is következik az előbb belátottak szerint az, hogy $r(\psi) \leq r(\varphi)$; tehát a két rang megegyezik.

Ezek után rátérhetünk a tétel két állításának a bizonyítására:

Ha „alkalmas” bázisokban $A = [\alpha] = [\beta]$, akkor a 4.13. Tétel szerint léteznek olyan σ, τ invertálható homomorfizmusok, amelyekre $\beta = \tau\alpha\sigma^{-1}$. Az előrebocsátottak szerint tehát $r(\alpha) = r(\beta)$.

A 4.12. Tétel szerint, ha egy mátrixon elemi átalakítást végzünk, akkor ez annak felel meg, hogy a „megfelelő” leképezést egy invertálható leképezéssel szorozzuk. Az előrebocsátottak szerint tehát a mátrix rangja nem változik meg. ■

4.15. Tétel. *Elemi átalakítások sorozatával elérhető, hogy egy mátrix két sorát (oszlopát) felcseréljük.*

Bizonyítás. Két sor (oszlop) felcserélése azt jelenti, hogy az eredeti bázisban (illetve ezek egyikében) két bázisvektort felcserélünk. Elég tehát azt megmutatni, hogy ez elérhető elemi transzformációkkal. Mivel itt a bázisnak csupán két eleme szerepel, ezért elég csak ezeket figyelembe venni. Legyen e két bázisvektor \mathbf{u} és \mathbf{v} . Az

$$\begin{aligned} (\mathbf{u}, \mathbf{v}) &\rightarrow (\mathbf{u} + \mathbf{v}, \mathbf{v}) \rightarrow (\mathbf{u} + \mathbf{v}, \mathbf{v} - (\mathbf{u} + \mathbf{v})) = (\mathbf{u} + \mathbf{v}, -\mathbf{u}) \rightarrow \\ &\rightarrow (\mathbf{u} + \mathbf{v} + (-\mathbf{u}), -\mathbf{u}) = (\mathbf{v}, -\mathbf{u}) \rightarrow (\mathbf{v}, (-1)(-\mathbf{u})) = (\mathbf{v}, \mathbf{u}) \end{aligned}$$

sorozat minden egyes lépése elemi transzformáció; és végül a két bázisvektor felcserélődik. ■

4.16. Tétel. *Részmátrix rangja nem lehet nagyobb, mint az eredeti mátrix rangja. Egy mátrix rangja megegyezik lineárisan független oszlopainak (sorainak) maximális számával. Egy mátrix rangja r , ha van olyan r sorú négyzetes részmátrixa, amelynek a determinánsa nem 0, de bármely $s > r$ esetén az összes s sorú négyzetes részmátrixának a determinánsa 0.*

Bizonyítás. Legyen $A = [\alpha]$, ahol $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ és $\mathbf{u}_1, \dots, \mathbf{u}_n$ az a bázis, amelyben a fenti mátrixfelírás megvalósul. A rangja az $\alpha(\mathbf{u}_i)$ vektorok generálta altér dimenziója. Ha e vektorok közül egyet elhagyunk, akkor a dimenzió nem nőhet. Ez az eredeti mátrix egy oszlopának az elhagyását jelenti. Egy sor elhagyására vonatkozóan ezen kívül még azt is figyelembe kell venni, hogy a transzponált mátrix rangja megegyezik az eredetivel. Az első állítás most már a részmátrix definíciójából következik.

Mivel $r(A) = \dim(\text{Im}(\alpha))$, ezért a második állításnak az oszlopokra vonatkozó része is igaz; a sorokra vonatkozó résznél ismét figyelembe kell venni a transzponált rangjára vonatkozó eredményt.

Ha az A mátrix rangja r , akkor van r darab független oszlopa. A többi elhagyva olyan mátrixot kapunk, amelynek r számú oszlopa van, és ugyanennyi a rangja is. Hasonlóképpen járhatunk el a sorokkal is; ekkor egy olyan négyzetes mátrixhoz jutunk, amelynek sorszáma is, oszlopszáma is és rangja is r . A determinánsról tanultak alapján e mátrix determinánsa nem lehet 0, mert sorai függetlenek. Ha $s > r$, akkor bármely $n - s$ sor és oszlop elhagyásával olyan négyzetes mátrixhoz jutunk, amelynek rangja $< s$, tehát sorai összefüggenek. Ugyancsak a determinánsra bizonyítottak alapján ebből következik, hogy e mátrix determinánsa 0. ■

4.17. Tétel. *Elemi transzformációkkal minden mátrixból olyan $[a_{i,j}]$ mátrix nyerhető, amelyben a „fődiagonális” első r eleme 1, a mátrix többi eleme 0 (azaz $a_{i,j} = 1$, ha $i = j \leq r$ és $a_{i,j} = 0$ máskor); és itt r a mátrix rangja.*

Bizonyítás. Ha a mátrix minden eleme 0, akkor készen vagyunk. Ha van a mátrixban nemnulla elem, akkor a 4.15. Tétel szerint elemi átalakításokkal elérhető, hogy ez az első sor első eleme legyen. Az első sort ennek az elemnek a reciprokéval megszorozva az új mátrixban az első sor első eleme 1 — ez is elemi átalakítás. Ugyancsak elemi átalakításokkal (az első sor, illetve oszlop megfelelő többszöröseinek a többi sorból, illetve oszlopból való kivonásával) elérhető, hogy az első sorban és oszlopban minden további elem 0 legyen. Az eljárást most azzal a részmátrixszal folytatjuk, amelyik az eredetiből az első sor és oszlop elhagyásával keletkezik. Ennek a mátrixnak minden elemi átalakítása természetesen az eredetinek is elemi átalakítása. Az eljárás addig folytatható, amíg a megmaradó részmátrixnak (ha ilyen van) minden eleme 0 lesz. Ilyen módon egy, a tételben előírt alakú mátrixot kapunk, amelynek a rangja a 4.14. Tétel szerint megegyezik az eredeti mátrix rangjával. E mátrixban az első r sorból és oszlopból álló négyzetes mátrix determinánsa 1. Ha más sor vagy oszlop (is) szerepel, akkor a determináns 0, hiszen egy sorának (és oszlopának) minden eleme 0. A 4.16. Tétel szerint tehát a mátrix rangja r . ■

A most tárgyalt eljárás tulajdonképpen alkalmas egy négyzetes mátrix determinánsának a megállapítására; amennyiben az előforduló osztásokat nem végezzük el. Ekkor egy diagonális mátrixot kapunk. A fenti lépéseknél a determináns (mint ott bizonyítottuk) nem változik meg; csupán a sorcserénél vált előjelet. Ezt figyelembe véve megkaphatjuk a determinánst.

A fenti eljárás segítségével elég egyszerűen meghatározhatjuk egy négyzetes mátrix inverzét is:

4.18. Tétel. *Legyen A egy négyzetes mátrix. Ha determinánsa nem 0, akkor A csak az oszlopokkal végzett elemi átalakításokkal átvihető az I identitásmátrixba. Ha egy-szersmind az identitásmátrixszal is sorban elvégezzük ezeket az elemi átalakításokat, akkor ebből pontosan az A mátrix inverzéhez jutunk.*

Természetesen hasonló eredményt kaphatunk, ha az átalakításokat csak a sorokkal végezzük.

Bizonyítás. Legyen — alkalmas bázisokban — $A = [\alpha]$, ahol $\alpha : \mathcal{U} \rightarrow \mathcal{U}$. A 4.17. Tétel szerint léteznek olyan $\sigma_1, \dots, \sigma_s$ és τ_1, \dots, τ_r elemi transzformációk, amelyekre a $\beta = \tau \cdot \alpha \cdot \sigma$ transzformáció mátrixa az identitásmátrix, azaz $\beta = \iota$, ahol $\tau = \tau_r \cdot \dots \cdot \tau_1$ és $\sigma = \sigma_1 \cdot \dots \cdot \sigma_s$. A kapott $\iota = \tau \cdot \alpha \cdot \sigma$ egyenlőséget balról τ inverzével és jobbról τ -val szorozva azt kapjuk, hogy $\iota = \alpha \cdot \sigma \cdot \tau$.

Ez azt jelenti, hogy $\sigma \cdot \tau$ az α inverze. Ezt úgy írhatjuk, hogy $\alpha^{-1} = \iota \cdot \sigma \cdot \tau$.

Az $\iota = \alpha \cdot \sigma_1 \cdot \dots \cdot \sigma_s \cdot \tau_r \cdot \dots \cdot \tau_1$ összefüggés azt mondja, hogy az A mátrixból az oszlopokkal a megfelelő elemi átalakításokat elvégezve (először a σ_1 -hez tartozót stb.) az identitásmátrixhoz jutunk. Ezzel az első állítást már igazoltuk is.

Az $\alpha^{-1} = \alpha \cdot \sigma_1 \cdot \dots \cdot \sigma_s \cdot \tau_r \cdot \dots \cdot \tau_1$ felírás alapján azt kapjuk, hogy ha az identitásmátrixból indulunk ki, akkor ugyanez az elemi átalakítás sorozat az A mátrix inverzét szolgáltatja. ■

A gyakorlatban sok módszer ismeretes az inverz mátrix meghatározására. Ezek „sikere” azon múlik, hogy az eredeti mátrix milyen „feltételek”-nek tesz eleget. Itt most egy olyan eljárást mutatunk meg, amelyik a diádfelbontáson alapszik. Az eljárás alap gondolata — leképezésekre megfogalmazva — a következő:

4.19. Tétel. *Legyen $\alpha : \mathcal{U} \rightarrow \mathcal{U}$ egy invertálható transzformáció és legyen rögzítve egy bázis. Ekkor eldönthető, hogy milyen \mathbf{u}, \mathbf{v} esetén létezik inverze az $\alpha + \widetilde{\mathbf{v}}\mathbf{u}^*$ transzformációnak és az inverz egyszerűen meghatározható.*

Bizonyítás. Az $\alpha = \iota$ esetben a 4.9. Tételhez hasonlóan járhatunk el. Legyen $\mathbf{u}^* \widetilde{\mathbf{v}}(1) = c$. Ha $c = -1$, akkor nincs inverz, egyébként az inverz $\iota - d \widetilde{\mathbf{v}}\mathbf{u}^*$, ahol $d = \frac{c}{1+c}$. Ennek a bizonyítása ugyanúgy történik, mint a 4.9. Tételé.

Az általános esetben az $\alpha + \widetilde{\mathbf{v}}\mathbf{u}^* = \alpha(\iota + \alpha^{-1} \widetilde{\mathbf{v}}\mathbf{u}^* = \alpha(\iota + \widetilde{\alpha^{-1}(\mathbf{v})}\mathbf{u}^*))$ összefüggésből adódó

$$(\alpha + \widetilde{\mathbf{v}}\mathbf{u}^*)^{-1} = (\iota + \widetilde{\alpha^{-1}(\mathbf{v})}\mathbf{u}^*)^{-1} \alpha^{-1}$$

egyenlőség alapján kaphatjuk meg az inverzet. ■

Megjegyzés. Ez az eljárás „diagonálisan domináns” mátrixok esetén használható. Ez azt jelenti, hogy a mátrix $A = D + T$ alakú, ahol D egy invertálható diagonális mátrix és T -nek kicsi a rangja. Ekkor ugyanis a diagonális mátrix inverze triviálisan adódik (a megfelelő diagonális elemeknek kell venni az inverzét); és a 4.19. Tételbeli lépést kell annyiszor alkalmazni, amekkora a T rangja.

A gyakorlatban — közelítő adatok esetén — a „nagyon kis elemű diádot” el is hanyagolhatjuk. □

Feladatok

1. Írjuk fel az $\varepsilon_{i,j} = \tilde{\mathbf{u}}_j \mathbf{u}_i^*$ leképezések mátrixát az adott \mathbf{U} bázisban.
2. Milyen alakú az $\varepsilon_{i,j} = \tilde{\mathbf{u}}_j \mathbf{u}_i^*$ leképezések mátrixa tetszőleges bázisban?
3. Bizonyítsuk be, hogy ha egy mátrixnak csak a soraival végzünk elemi átalakítást, akkor az a következő alakra hozható: Minden i sorindexhez létezik egy $f(i) \geq i$ oszlopindex, úgy, hogy $a_{i,j} = 0$, ha $j < f(i)$; $a_{i,f(i)} = 1$, $f(i+1) > f(i)$.
4. A sík derékszögű forgatásának a mátrixa semmilyen bázisban sem lehet diagonális. Elemi transzformációkkal viszont diagonális alakra hozható. Nem ellentmondás ez?!
5. Legyen α az \mathcal{U} vektortér egy lineáris transzformációja és \mathcal{V}_i ($i = 1, \dots, r$) az α invariáns alterei, amelyekre $\mathcal{U} = \mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_r$. Milyen alakú lesz α mátrixa egy olyan bázisban, amelynek minden eleme valamelyik \mathcal{V}_i -ben van; mégpedig úgy, hogy először a \mathcal{V}_1 -be eső, majd a \mathcal{V}_2 -be eső stb. bázisvektorokat soroljuk fel?
6. Írjuk fel egy idempotens transzformáció mátrixát „alkalmas” bázisban.
7. Írjuk fel egy nilpotens transzformáció mátrixát „alkalmas” bázisban.

ÖTÖDIK FEJEZET

BIHOMOMORFIZMUSOK

1. Bilineáris leképezések, bilineáris formák

A számok szorzásánál a disztributivitás azt fejezi ki, hogy a szorzás — mint kétváltozós függvény — bármelyik tényezőt rögzítve a másik tényezőben „additív”. Hasonló jelenséget figyelhetünk meg a vektortér axiómáinál is: $(a+b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ és $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$. Ez a tulajdonság fennáll a mátrixok szorzatára is: ha létezik az $A_i B_j$ mátrixszorzat, akkor $(A_1 + A_2)B_j = A_1 B_j + A_2 B_j$ és $A_i(B_1 + B_2) = A_i B_1 + A_i B_2$. A mátrixokra bizonyított $c(AB) = (cA)B = A(cB)$ tulajdonság alapján azt is látjuk, hogy a szorzás egyik tényezőjét rögzítve a másik tényezőben lineáris függvényt kapunk. Világos, hogy ez a megelőző két esetben is igaz. A lineáris függvényeket lineáris leképezéseknek vagy homomorfizmusoknak is nevezzük. A homomorfizmus szó a legáltalánosabb; ez azt jelenti, hogy nem számít, milyen műveletek vannak értelmezve; a leképezés legyen művelettartó. Ennek megfelelően itt is felsoroljuk az „általános” elnevezést is.

5.1. Definíció. Legyenek adottak a K test feletti $\mathcal{U}, \mathcal{V}, \mathcal{W}$ vektorterek. Egy $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ leképezést bihomomorfizmusnak vagy bilineáris leképezésnek nevezünk, ha bármelyik változóját bárhogyan rögzítve a másikban lineáris leképezést kapunk. A $\mathcal{W} = K$ esetben bilineáris függvényről fogunk beszélni.

Ha a K test helyett egy R gyűrű feletti modulusokat nézünk, akkor csak a bihomomorfizmus elnevezést használjuk.

Bihomomorfizmusok egyenlőségét függvényegyenlőségként definiáljuk.

A bihomomorfizmusokat vastagított latin nagybetűkkel fogjuk jelölni. ■

Egy bihomomorfizmus *nem* a direkt szorzat homomorfizmusa. Ez azonnal látható, ha a bihomomorfizmus jelentését formulákkal megfogalmazzuk:

5.1. Tétel. $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ függvény akkor és csak akkor bilineáris leképezés, ha *tetszőleges* $\mathbf{u}, \mathbf{u}' \in \mathcal{U}$, $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$ és $c \in K$ esetén érvényesek az alábbiak:

- (1) $\mathbf{B}(\mathbf{u} + \mathbf{u}', \mathbf{v}) = \mathbf{B}(\mathbf{u}, \mathbf{v}) + \mathbf{B}(\mathbf{u}', \mathbf{v})$,
- (2) $\mathbf{B}(\mathbf{u}, \mathbf{v} + \mathbf{v}') = \mathbf{B}(\mathbf{u}, \mathbf{v}) + \mathbf{B}(\mathbf{u}, \mathbf{v}')$,
- (3) $\mathbf{B}(c \cdot \mathbf{u}, \mathbf{v}) = \mathbf{B}(\mathbf{u}, c \cdot \mathbf{v}) = c \cdot \mathbf{B}(\mathbf{u}, \mathbf{v})$. ■

5.2. Tétel. Az $\mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bihomomorfizmusok K -vektorteret alkotnak az alábbi műveletekre:

- (1) $(c\mathbf{B}) : (\mathbf{u}, \mathbf{v}) \mapsto c \cdot (\mathbf{B}(\mathbf{u}, \mathbf{v}))$,
- (2) $(\mathbf{A} + \mathbf{B}) : (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{A}(\mathbf{u}, \mathbf{v}) + \mathbf{B}(\mathbf{u}, \mathbf{v})$.

Bizonyítás. A lineáris leképezésekre vonatkozó skalárszoros és összeg definíciója alapján világos, hogy mindkét esetben bilineáris leképezést nyerünk. A vektortér-axiómák teljesülésének triviális kiszámolását az olvasóra bízuk. ■

A két változó természetesen „nem cserélhet helyet”, annak ellenére, hogy a $\mathbf{B}'(\mathbf{v}, \mathbf{u}) = \mathbf{B}(\mathbf{u}, \mathbf{v})$ is bihomomorfizmus; hiszen ez a $\mathcal{V} \times \mathcal{U}$ vektorteret képezi \mathcal{W} -be. Ez a helycsere még akkor sem lehetséges, ha az \mathcal{U} és \mathcal{V} vektorterek megegyeznek. Abban az — egyébként igen fontos — speciális esetben, amikor ez lehetséges, ezekre a bilineáris leképezésekre külön elnevezést használunk:

5.2. Definíció. Az $\mathbf{S} : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{W}$ bihomomorfizmust szimmetrikusnak nevezzük, ha minden $\mathbf{u}, \mathbf{v} \in \mathcal{U}$ esetén $\mathbf{S}(\mathbf{u}, \mathbf{v}) = \mathbf{S}(\mathbf{v}, \mathbf{u})$ teljesül.

A $\mathbf{T} : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{W}$ bihomomorfizmust antiszimmetrikusnak nevezzük, ha minden $\mathbf{u}, \mathbf{v} \in \mathcal{U}$ esetén $\mathbf{T}(\mathbf{u}, \mathbf{v}) = -\mathbf{T}(\mathbf{v}, \mathbf{u})$ teljesül. ■

Később majd látni fogjuk, hogy az antiszimmetrikus leképezések legalább olyan fontosak, mint a szimmetrikusak. Az első erre utaló jelet az alábbi tételben fogalmazhatjuk meg:

5.3. Tétel. Egy $\mathbf{T} : \mathcal{U} \times \mathcal{U} \rightarrow \mathcal{W}$ bihomomorfizmusra az alábbiak ekvivalensek:

- (1) \mathbf{T} antiszimmetrikus.
- (2) Ha $\mathbf{u} \in \mathcal{U}$, akkor $\mathbf{T}(\mathbf{u}, \mathbf{u}) = \mathbf{o}$.
- (3) A függvényérték nem változik, ha valamelyik argumentumhoz hozzáadjuk a másik argumentum egy skalárszorosát, azaz $\mathbf{u}, \mathbf{v} \in \mathcal{U}$ és $c \in K$ esetén

$$\mathbf{T}(\mathbf{u} + c \cdot \mathbf{v}, \mathbf{v}) = \mathbf{T}(\mathbf{u}, \mathbf{v} + c \cdot \mathbf{u}) = \mathbf{T}(\mathbf{u}, \mathbf{v}).$$

Bizonyítás. Ha \mathbf{T} antiszimmetrikus, akkor $\mathbf{T}(\mathbf{u}, \mathbf{u}) = -\mathbf{T}(\mathbf{u}, \mathbf{u})$, amiből $\mathbf{T}(\mathbf{u}, \mathbf{u}) = \mathbf{o}$ következik.

Ha (2) igaz, akkor a bilinearitás alapján

$$\mathbf{T}(\mathbf{u} + c \cdot \mathbf{v}, \mathbf{v}) = \mathbf{T}(\mathbf{u}, \mathbf{v}) + c \cdot \mathbf{T}(\mathbf{v}, \mathbf{v}) = \mathbf{T}(\mathbf{u}, \mathbf{v}) \quad \text{és}$$

$$\mathbf{T}(\mathbf{u}, \mathbf{v} + c \cdot \mathbf{u}) = \mathbf{T}(\mathbf{u}, \mathbf{v}) + c \cdot \mathbf{T}(\mathbf{u}, \mathbf{u}) = \mathbf{T}(\mathbf{u}, \mathbf{v}).$$

(3)-ból a 4.15. Tétel bizonyításához hasonlóan kapható az antiszimmetria:

$$\begin{aligned} \mathbf{T}(\mathbf{u}, \mathbf{v}) &= \mathbf{T}(\mathbf{u} + \mathbf{v}, \mathbf{v}) = \mathbf{T}(\mathbf{u} + \mathbf{v}, \mathbf{v} - (\mathbf{u} + \mathbf{v})) = \mathbf{T}(\mathbf{u} + \mathbf{v}, -\mathbf{u}) = \\ &= \mathbf{T}(\mathbf{u} + \mathbf{v} - \mathbf{u}, -\mathbf{u}) = \mathbf{T}(\mathbf{v}, -\mathbf{u}) = -\mathbf{T}(\mathbf{v}, \mathbf{u}). \end{aligned}$$

■

5.4. Tétel. Legyen \mathcal{M} az $\mathcal{U} \times \mathcal{U}$ -t a \mathcal{W} -be vivő bihomomorfizmusok tere. Ennek egy \mathcal{S} , illetve \mathcal{I} alterét alkotják a szimmetrikus, illetve az antiszimmetrikus bihomomorfizmusok. Ezekre teljesül: $\mathcal{M} = \mathcal{S} \oplus \mathcal{I}$.

Bizonyítás. Az, hogy \mathcal{S} és \mathcal{T} mindegyike altér, az világos. Az is nyilvánvaló, hogy (ha K -ban $1 + 1 \neq 0$, akkor) a két altérnek egyetlen közös eleme az azonosan \mathbf{o} leképezés. A direkt összeg felbontáshoz most már csak azt kell megmutatni, hogy bármely bilineáris leképezés felírható egy szimmetrikus és egy antiszimmetrikus leképezés összegeként. Tekintsük evégett a \mathbf{B} -vel együtt azt a \mathbf{B}^* leképezést, amelyre $\mathbf{B}^*(\mathbf{u}, \mathbf{v}) = \mathbf{B}(\mathbf{v}, \mathbf{u})$. Világos, hogy ez is bihomomorfizmus és ugyancsak eleme \mathcal{M} -nek. Legyen most:

$$\mathbf{S} = \frac{1}{2} \cdot (\mathbf{B} + \mathbf{B}^*) \quad \text{és} \quad \mathbf{T} = \frac{1}{2} \cdot (\mathbf{B} - \mathbf{B}^*).$$

Világos, hogy $\mathbf{B} = \mathbf{S} + \mathbf{T}$. Triviális számolással belátható az is, hogy \mathbf{S} szimmetrikus és \mathbf{T} antiszimmetrikus. ■

5.5. Tétel. Legyen $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bihomomorfizmus, $\alpha : \mathcal{U}' \rightarrow \mathcal{U}$, $\beta : \mathcal{V}' \rightarrow \mathcal{V}$ és $\gamma : \mathcal{W} \rightarrow \mathcal{W}'$ homomorfizmusok. Ekkor a $\gamma(\mathbf{B}(\alpha(\mathbf{u}'), \beta(\mathbf{v}')))$ formulával megadott leképezés egy $\mathcal{U}' \times \mathcal{V}' \rightarrow \mathcal{W}'$ bihomomorfizmus.

Bizonyítás. A leképezések kompozíciója alapján világos, hogy egy $\mathcal{U}' \times \mathcal{V}' \rightarrow \mathcal{W}'$ leképezést kaptunk. Ha \mathbf{u}' rögzített, akkor $\alpha(\mathbf{u}')$ sem változik. Ekkor \mathbf{B} bihomomorfizmus volta miatt három homomorfizmus kompozícióját kapjuk. Hasonló eredmény adódik \mathbf{v}' rögzítése esetében is. ■

A homomorfizmusokhoz hasonlóan a bihomomorfizmusok is előírhatók egy „bázispáron”:

5.6. Tétel. Legyen $\mathbf{U} = \{\mathbf{u}_i \mid i \in I\}$ az \mathcal{U} vektortérnek és $\mathbf{V} = \{\mathbf{v}_j \mid j \in J\}$ a \mathcal{V} vektortérnek egy-egy bázisa. Ekkor tetszőleges $\{\mathbf{w}_{i,j} \in \mathcal{W} \mid (i, j) \in I \times J\}$ vektorrendszerhez létezik pontosan egy olyan $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bihomomorfizmus, amelyre $\mathbf{B}(\mathbf{u}_i, \mathbf{v}_j) = \mathbf{w}_{i,j}$.

Bizonyítás. Az egyértelműség azonnal következik a bilinearitásból, tekintettel arra, hogy minden vektor (egyértelműen) előállítható a bázisvektorok lineáris kombinációjaként.

Ez a \mathbf{B} leképezés tehát csak úgy definiálható, hogy:

$$\mathbf{B} \left(\sum_{i \in I_0} c_i \mathbf{u}_i, \sum_{j \in J_0} d_j \mathbf{v}_j \right) = \sum_{i \in I_0} \sum_{j \in J_0} c_i d_j \mathbf{w}_{i,j},$$

ahol $I_0 \subseteq I$ és $J_0 \subseteq J$ valamilyen véges részhalmazok. Mivel \mathbf{U} és \mathbf{V} bázisok (tehát független rendszerek), ezért a fenti összefüggés valóban egy leképezést definiál. Világos, hogy itt akár az első, akár a második argumentumot rögzítjük, a kapott leképezés pontosan úgy van megadva, mint ahogy azt a lineáris leképezések esetében tettük. ■

A \mathcal{W} vektortérnek is tekintve egy bázisát azonnal adódik a következő:

Kiegészítés. Ha $\mathbf{W} = \{\mathbf{w}_m \mid m \in M\}$ a \mathcal{W} vektortér egy bázisa, és ebben a bázisban $\mathbf{w}_{i,j} = \sum_{m \in M_0} e_{i,j,m} \mathbf{w}_m$, ahol $M_0 \subseteq M$ is egy véges részhalmaz, akkor

$$\mathbf{B} \left(\sum_{i \in I_0} c_i \mathbf{u}_i, \sum_{j \in J_0} d_j \mathbf{v}_j \right) = \sum_{i \in I_0} \sum_{j \in J_0} \sum_{m \in M_0} c_i d_j e_{i,j,m} \mathbf{w}_m. \blacksquare$$

Következmény. Minden $\mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bilineáris leképezés egyértelműen meghatározható $\mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvényekkel.

Bizonyítás. Világos, hogy a fenti kiegészítésben szereplő formula egyértelműen adott a $\mathbf{B}_m \left(\sum_{i \in I_0} c_i \mathbf{u}_i, \sum_{j \in J_0} d_j \mathbf{v}_j \right) = \sum_{i \in I_0} \sum_{j \in J_0} \sum_{m \in M_0} c_i d_j e_{i,j,m}$ leképezésekkel. E leképezések mindegyike az $\mathcal{U} \times \mathcal{V}$ vektorteret képezi le K -ba. Ezek a leképezések $\mathbf{w}_m^*(\mathbf{B}(\mathbf{u}, \mathbf{v}))$ alakúak, ahol \mathbf{w}_m^* a megfelelő duális bázis elemein fut végig. Az 5.5. Tétel szerint tehát bihomomorfizmusok. \blacksquare

Megjegyzés. Ez az eljárás csak akkor vihető át maradéktalanul modulusokra, ha a \mathcal{W} modulusnak létezik bázisa. \square

Tekintsük az 5.6. Tételben szereplő \mathbf{U} és \mathbf{V} bázisokat. Legyenek $\mathbf{x} = \sum_i x_i \mathbf{u}_i$ és $\mathbf{y} = \sum_j y_j \mathbf{v}_j$ adott vektorok. Legyen továbbá a $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvényre

$$\mathbf{B}(\mathbf{u}_i, \mathbf{v}_j) = b_{i,j}. \text{ Ekkor a } \mathbf{B} \text{ bilineáris függvényre } \mathbf{B}(\mathbf{x}, \mathbf{y}) = \sum_i \sum_j b_{i,j} x_i y_j \text{ teljesül. Itt } x_i$$

és y_j tetszőleges K -beli elemek lehetnek; vagyis „változók”. Ennek megfelelően tekinthetjük őket határozatlanoknak, amikor is egy $n + k$ határozatlanú polinomot nyerünk. E polinomokban az x_i -knek (y_j -knek) tetszőleges rögzített értéket adva, a többi határozatlanban (homogén) lineáris polinomokat nyerünk.

5.3. Definíció. Az x_i és y_j határozatlanokban felírt K -beli együtthatós $\sum_i \sum_j b_{i,j} x_i y_j$ polinomot bilineáris formának vagy bilineáris alaknak nevezzük. \blacksquare

Megjegyzés. Tekintsük a K számtest feletti $K[x_1, \dots, x_n, y_1, \dots, y_k]$ polinomgyűrűben a $\sum c_{i,j} x_i y_j$ alakú „bilineáris” polinomokat. Ezek a polinomok a polinomokkal végezhető műveletekre nézve vektorteret alkotnak a K test felett. Tekintsük az $\mathbf{u}_1, \dots, \mathbf{u}_n$ bázisú \mathcal{U} és a $\mathbf{v}_1, \dots, \mathbf{v}_k$ bázisú \mathcal{V} vektortereket.

Tetszőleges $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvény egyértelműen definiálható a $\mathbf{B}(\mathbf{u}_i, \mathbf{v}_j) = c_{i,j}$ egyenlőségekkel — és ezek az egyenlőségek mindig megadnak egy bilineáris függvényt. Az ezekkel a konstansokkal megadott bilineáris függvénynek megfelelően a $\sum c_{i,j} x_i y_j$ alakú bilineáris polinomot egy izomorfizmust nyerünk a fenti vektortér és a bilineáris függvények vektortere között (a bizonyítás triviális). \square

Feladatok

1. Mutassuk meg, hogy az 5.3. Tétel nem igaz, pontosabban szólva nem így igaz; és küszöböljük ki a hiányosságot.

2. Mutassuk meg, hogy az 5.6. Tételben, ha valamelyik bázis helyett generátorrendszert mondunk, akkor az egyértelműség igaz, de a bihomomorfizmus nem feltétlenül létezik. Ha viszont mindkét rendszer független, de valamelyikük nem bázis, akkor több ilyen tulajdonságú bihomomorfizmus is létezik.

3. Legyenek $\mathcal{U}_i, \mathcal{V}, \mathcal{W}_j$ vektorterek ugyanazon test felett és legyen $\mathcal{M}_{i,j}$ az $\mathcal{U}_i \times \mathcal{V} \rightarrow \mathcal{W}_j$ bilineáris leképezések vektortere. Miképpen konstruálhatjuk meg az $\mathcal{M}_{i,j}$ vektorterek segítségével

$$\begin{aligned} \text{a } \left(\bigoplus_i \mathcal{U}_i \right) \times \mathcal{V} &\rightarrow \left(\bigoplus_j \mathcal{W}_j \right), & \left(\bigoplus_i \mathcal{U}_i \right) \times \mathcal{V} &\rightarrow \left(\prod_j \mathcal{W}_j \right), & \left(\prod_i \mathcal{U}_i \right) \times \mathcal{V} &\rightarrow \left(\bigoplus_j \mathcal{W}_j \right), \\ \left(\prod_i \mathcal{U}_i \right) \times \mathcal{V} &\rightarrow \left(\prod_j \mathcal{W}_j \right) \end{aligned}$$

bihomomorfizmusokból álló vektortereket?

4. Bizonyítsuk be, hogy bármely $\mathbf{B} : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{C}$ modulus-bihomomorfizmusra és bármely $\mathbf{a} \in \mathcal{A}$ és $\mathbf{b} \in \mathcal{B}$ elemre $\mathbf{B}(\mathbf{a}, \mathbf{o}_{\mathcal{B}}) = \mathbf{B}(\mathbf{o}_{\mathcal{A}}, \mathbf{b}) = \mathbf{o}_{\mathcal{C}}$.

5. Legyen $\mathbf{B} : \mathbb{Q}_{\mathbb{Z}} \times \mathbb{Q}_{\mathbb{Z}} \rightarrow \mathbb{Q}_{\mathbb{Z}}$ bihomomorfizmus. Hány nemnulla racionális számpáron kell megadni a \mathbf{B} értékét, hogy ezzel \mathbf{B} egyértelműen meghatározott lehessen? Ennek alapján fogalmazzuk meg az 5.6. Tételt \mathbb{Z} -modulusokra.

6. Mennyi — az összeadásra nézve disztributív — szorzatot értelmezhetünk \mathbb{Z} -n, \mathbb{Q} -n és \mathbb{Z}_n -en? Ezek közül mennyi asszociatív? Van-e közöttük nemkommutatív?

7. Tekintsük a komplex egységgyökök \mathcal{E} halmazát mint \mathbb{Z} -modulust (vektorösszeadás a szorzás, skalárral való szorzás a hatványozás). Határozzuk meg az összes $\mathcal{E} \times \mathcal{E} \rightarrow \mathcal{E}$ bihomomorfizmust.

8. Legyen L a K testnek részteste. Ekkor minden K -vektortér L -vektortér is. Legyenek $\mathcal{U}, \mathcal{V}, \mathcal{W}$ K -vektorterek. Milyen kapcsolat áll fenn a $\mathbf{B}_K : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ K -bihomomorfizmusok és a $\mathbf{B}_L : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ L -bihomomorfizmusok vektorterei között?

9. Milyen feltétel mellett lesz két bilineáris leképezés szorzata is bilineáris?

10. Legyenek $\mathbf{u} \in \mathcal{U}$ és $\mathbf{v} \in \mathcal{V}$ vektorterek. Bizonyítsuk be, hogy $\mathbf{A}(\mathbf{u}, \mathbf{v}) = \widetilde{\mathbf{u}}\mathbf{v}^*$, és $\mathcal{U} = \mathcal{V}$ esetén $\mathbf{B}(\mathbf{u}, \mathbf{v}) = \mathbf{v}^*\widetilde{\mathbf{u}}$ is bilineáris leképezések. Hova képeznek?

11. Legyen $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ homogén lineáris leképezés. Bizonyítsuk be, hogy $\mathbf{C}(\mathbf{u}, \mathbf{v}) = \mathbf{u}^*\varphi(\mathbf{v})$ bihomomorfizmus. Honnan hova képez?

12. Legyen \mathbf{B}^* az 5.4. Tételben leírt megfeleltetés, azaz $\mathbf{B}^*(\mathbf{u}, \mathbf{v}) = \mathbf{B}(\mathbf{v}, \mathbf{u})$. Bizonyítsuk be, hogy \mathbf{B} akkor és csak akkor szimmetrikus (antiszimmetrikus), ha $\mathbf{B}^* = \mathbf{B}$ ($\mathbf{B}^* = -\mathbf{B}$).

13. Tekintsük a háromdimenziós térbeli vektoriális szorzatot. Bizonyítsuk be, hogy ez egy antiszimmetrikus bilineáris leképezés.

2. Bilineáris függvények mátrixa

Láttuk, hogy bármely $\mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bilineáris leképezés megadható $\mathcal{U} \times \mathcal{V} \rightarrow K$ speciális bilineáris leképezések egy rendszerével. (Mint az 5.1. Definícióban mondtuk, a K -ba történő bilineáris leképezéseket bilineáris függvényeknek fogjuk nevezni.)

5.4. Definíció. Legyen $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U}_K és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ a \mathcal{V}_K vektortér egy-egy rögzített bázisa. Ekkor a $B = [\mathbf{B}] = [\mathbf{B}]^{\mathbf{U}, \mathbf{V}} = [\mathbf{B}(\mathbf{u}_i, \mathbf{v}_j)]_{n,k}$ mátrixot a $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvény adott bázisokban felírt mátrixának nevezzük. ■

5.7. Tétel. *Rögzített bázisok esetén tetszőleges $B = [b_{i,j}]_{n,k}$ mátrixhoz pontosan egy olyan $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvény létezik, amelyre $[\mathbf{B}] = B$.*

Bizonyítás. Az 5.6. Tétel szerint az adott $B = [b_{i,j}]_{n,k}$ mátrixhoz létezik pontosan egy olyan \mathbf{B} bilineáris függvény, amelyre $\mathbf{B}(\mathbf{u}_i, \mathbf{v}_j) = b_{i,j}$ teljesül minden szóba jövő index-párra, ahol $1 \leq i \leq n$ és $1 \leq j \leq k$. Az 5.4. Definíció alapján $[\mathbf{B}] = B$. ■

5.8. Tétel. *Adott $\mathbf{B} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvényhez pontosan egy olyan $\beta : \mathcal{V} \rightarrow \mathcal{U}$ lineáris leképezés található, amelyre az adott bázisokban $[\mathbf{B}] = [\beta]$. Ezen a módon minden bilineáris leképezés megkapható.*

Bizonyítás. Tekintettel arra, hogy a bilineáris függvények is és a homogén lineáris leképezések is kölcsönösen és egyértelműen meghatározhatók mátrixukkal, ezért elegendő (és szükséges) azt belátni, hogy a megadott mátrixok azonos méretűek.

A \mathbf{B} mátrixában a sorindexek száma annyi, mint \mathcal{U} dimenziója és az oszlopindexek száma \mathcal{V} dimenziójával egyenlő. Eszerint $[\mathbf{B}]$ -nek n sora és k oszlopa van. Mivel β egy k -dimenziós teret képez le egy n -dimenziósra, ezért β mátrixában k oszlop és n sor van; mint állítottuk. ■

Kiegészítés. *A fenti megfeleltetésben $\mathbf{B} = \mathbf{u}^* \beta(\mathbf{v})$, ahol \mathbf{u}^* az \mathbf{U} bázist az \mathbf{U}^* duális bázisba képező izomorfizmusnál az \mathbf{u} vektor képe.*

Bizonyítás. Világos, hogy $\mathbf{u}^* \beta(\mathbf{v})$ mind \mathbf{u} -ban, mind \mathbf{v} -ben lineáris, ha a másik argumentumot rögzítjük. Elég tehát annak a belátása, hogy a két bilineáris függvénynek az adott bázisokban ugyanaz a mátrixa. \mathbf{B} mátrixa a szereplő bázisokban — definíció szerint — $[\mathbf{B}(\mathbf{u}_i, \mathbf{v}_j)]$. Egy tetszőleges $\varphi : \mathcal{E} \rightarrow \mathcal{F}$ leképezés mátrixa az \mathcal{E} -beli \mathbf{E} és \mathcal{F} -beli \mathbf{F} bázisokban olyan, hogy a j -edik oszlopban az \mathbf{e}_j képe áll; mégpedig az i -edik helyen $\mathbf{f}_i^* \varphi(\mathbf{e}_j)$. Esetünkben ez $\mathbf{u}_i^* \beta(\mathbf{v}_j)$. A mátrixok egyenlősége szerint ez ugyancsak $[\mathbf{B}(\mathbf{u}_i, \mathbf{v}_j)]$. ■

Bilineáris függvények, illetve bilineáris leképezések esetében is szükséges tudni azt, miképpen változik meg a mátrix, ha új bázisra térünk át. Itt is használni fogjuk azokat a jelöléseket, amelyek a lineáris transzformációk esetében szerepeltek:

Jelölés. Legyen adva a K test feletti \mathcal{U} és \mathcal{V} vektortér, megfelelően az $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és a $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ bázisokkal. Vegyünk fel a két vektortérben egy-egy új bázist, \mathcal{U} -ban az $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bázist és \mathcal{V} -ben az $\mathbf{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_k\}$ bázist.

Legyen $\sigma : \mathcal{U} \rightarrow \mathcal{U}$ az a transzformáció, amelyre $\sigma(\mathbf{u}_j) = \mathbf{e}_j$ és $\tau : \mathcal{V} \rightarrow \mathcal{V}$ az a transzformáció, amelyre $\tau(\mathbf{v}_i) = \mathbf{f}_i$ ($1 \leq j \leq n$ és $1 \leq i \leq k$). Ezeket a transzformációkat kisértő transzformációknak fogjuk nevezni.

Legyen $\mathbf{U}^* = \{\mathbf{u}_1^*, \dots, \mathbf{u}_n^*\}$ az \mathbf{U} -nak, $\mathbf{V}^* = \{\mathbf{v}_1^*, \dots, \mathbf{v}_k^*\}$ a \mathbf{V} -nek, $\mathbf{E}^+ = \{\mathbf{e}_1^+, \dots, \mathbf{e}_n^+\}$ az \mathbf{E} -nek és $\mathbf{F}^+ = \{\mathbf{f}_1^+, \dots, \mathbf{f}_k^+\}$ az \mathbf{F} -nek megfelelő duális bázis.

5.9. Tétel. Legyen $[\mathbf{B}]^{\mathbf{U}, \mathbf{V}} = \mathbf{V}[\beta]^{\mathbf{U}}$. Ekkor — a fenti jelölést használva — $[\mathbf{B}]^{\mathbf{E}, \mathbf{F}} = \mathbf{V}[\sigma^* \beta \tau]^{\mathbf{U}}$.

Bizonyítás. A bilineáris függvény mátrixának a definíciója szerint \mathbf{B} új bázisban felírt mátrixában az i -edik sor j -edik eleme

$$\mathbf{B}(\mathbf{e}_i, \mathbf{f}_j) = \mathbf{e}_i^* \beta(\mathbf{f}_j) = (\sigma(\mathbf{v}_i))^* \beta(\tau(\mathbf{u}_j)) = \mathbf{v}_i^* \sigma^* \beta \tau \mathbf{u}_j = \mathbf{v}_i^* (\sigma^* \beta \tau) \mathbf{u}_j,$$

ami pontosan a $\mathbf{V}[\sigma^* \beta \tau]^{\mathbf{U}}$ mátrix i -edik sorának a j -edik eleme. ■

5.10. Tétel. Rögzített bázisok esetén $\mathbf{B}(\mathbf{u}, \mathbf{v}) = [\mathbf{u}]^\dagger [\mathbf{B}][\mathbf{v}]$.

Bizonyítás. Az 5.8. Tételben szereplő β homomorfizmusra:

$$\mathbf{B}(\mathbf{u}, \mathbf{v}) = [\mathbf{B}(\mathbf{u}, \mathbf{v})] = [\mathbf{v}^* \beta \mathbf{u}] = [\mathbf{v}^*][\beta][\mathbf{u}] = [\mathbf{v}]^\dagger [\mathbf{B}][\mathbf{u}]. \quad \blacksquare$$

5.11. Tétel. A $\mathbf{B}^*(\mathbf{v}, \mathbf{u}) = \mathbf{B}(\mathbf{u}, \mathbf{v})$ egyenlőséggel definiált függvény is bilineáris, amelyre $[\mathbf{B}^*] = [\mathbf{B}]^\dagger$.

Bizonyítás. Az első állítás triviális; a második azonnal adódik a bilineáris függvény mátrixának a definíciójából. ■

Feladatok

1. Mutassuk meg, hogy az 5.8. Tételben szereplő $\beta : \mathcal{V} \rightarrow \mathcal{U}$ lineáris leképezéshez hasonlóan $\beta^* : \mathcal{U} \rightarrow \mathcal{V}$ is egyértelműen meghatározza \mathbf{B} -t.

2. Hogyan változik meg egy bilineáris függvény mátrixa, ha valamelyik bázis elemein egy elemi transzformációt végzünk?

3. Mutassuk meg, hogy tetszőleges bilineáris függvény esetén a megfelelő vektortereken választhatók olyan bázisok, amelyekben a függvény mátrixa „diagonális”. (Ha nem négyzetes a mátrix, akkor ez azt jelenti, hogy a mátrix $a_{i,j}$ eleme 0, ha $i \neq j$.)

4. Milyenek azok a bilineáris függvények, amelyeknél valamelyik vektortérnek az alaptestet választjuk?

5. Legyenek σ és τ , mint az 5.8. Tétel utáni jelölésben, és \mathbf{B} , valamint β , mint a Tételben. Milyen σ és τ esetében lesz a $\mathbf{B} \rightarrow \beta$ megfeleltetés az új bázisban is ugyanez? Mutassunk olyan példát is, amikor a megfeleltetés megváltozik.

3. Homogén koordináták, kvadratikus alakok a valós térben

5.5. Definíció. Legyen \mathcal{U} vektortér a K test felett és $\mathbf{A} : \mathcal{U} \times \mathcal{U} \rightarrow K$ egy bilineáris függvény. A $\mathbf{Q}(\mathbf{x}) = \mathbf{A}(\mathbf{x}, \mathbf{x})$ kifejezést kvadratikus alaknak (kvadratikus formának) nevezzük. ■

Megjegyzés. Az 5.4. Definíció utáni megjegyzésben szereplő megfeleltetés itt is figyelembe vehető. Ekkor a $\mathbf{Q}(\mathbf{x})$ kvadratikus alaknak egy $\sum c_{i,j} x_i x_j$ homogén másodfokú (kvadratikus) polinom felel meg. Itt is világos, hogy a megfeleltetés mindkét irányban létrehozható, csak az nem igaz, hogy egyértelmű. A fellépő másodfokú polinomok ugyanis nem különböznek, hiszen például $x_1 x_2 = x_2 x_1$. Ennek megfelelően az is lehetséges, hogy az \mathbf{A} és \mathbf{B} bilineáris függvények esetében $\mathbf{A}(\mathbf{x}, \mathbf{x}) = \mathbf{B}(\mathbf{x}, \mathbf{x})$. □

5.12. Tétel. Az $\mathbf{A} : \mathcal{U} \times \mathcal{U} \rightarrow K$ és $\mathbf{B} : \mathcal{U} \times \mathcal{U} \rightarrow K$ bilineáris függvények pontosan akkor hozzák létre ugyanazt a kvadratikus alakot, ha a különbségük antiszimmetrikus. Minden kvadratikus függvényhez pontosan egy olyan szimmetrikus függvény van, amely ezt a kvadratikus alakot hozza létre.

A kvadratikus alakok a skalárral való szorzásra és az összeadásra nézve vektorteret alkotnak. Az a Φ megfeleltetés, amelyik minden bilineáris függvénynek az általa létrehozott kvadratikus alakot felelteti meg, egy szürjektív homomorfizmus, amelynek magtere az antiszimmetrikus függvényekből áll és amelyet a szimmetrikusok alterére megszorítva izomorfizmust kapunk.

Bizonyítás. Világos, hogy a Tétel első része következik a második részből.

Tekintsük a Φ megfeleltetést. Ha \mathbf{A} és \mathbf{B} bilineáris függvények, akkor a $\mathbf{C} = a\mathbf{A} + b\mathbf{B}$ bilineáris függvényre $\mathbf{C}(\mathbf{x}, \mathbf{x}) = a\mathbf{A}(\mathbf{x}, \mathbf{x}) + b\mathbf{B}(\mathbf{x}, \mathbf{x})$. Ezzel igazoltuk, hogy Φ szürjektív homomorfizmus. Az 5.3. Tétel szerint \mathbf{B} pontosan akkor antiszimmetrikus, ha $\mathbf{Q}(\mathbf{x}) = \mathbf{B}(\mathbf{x}, \mathbf{x}) = \mathbf{o}$. Az 5.4. Tétel szerint a bilineáris függvények tere a szimmetrikusoknak és az antiszimmetrikusoknak a direkt összege; ezért Φ -nek a szimmetrikusokra való megszorítása valóban izomorfizmus. ■

Megjegyzések

1. Az előzőekből világosan látható, hogy igen szoros a kapcsolat a koordináta-geometria és a vektorterek között. Egy helyen azonban „ellentmondás” lép fel. Tekintsük például az $x - 2y = 3$ egyenlet meghatározta egyenest. Ez az egyenes *nem* lineáris altere a kétdimenziós vektortérnek; mert

a vektorterekben minden altér tartalmazza a 0 vektort. Két lehetőség kínálkozik a probléma megoldására. Az egyik az, ha a fenti egyenest úgy tekintjük, mint az $x - 2y$ altérnek azt a mellékosztályát, amelyik a $(3, 0)$ ponton megy át. Ez tulajdonképpen jó, csak technikailag kissé ügyetlen.

A másik lehetőség a következő:

A kétdimenziós síkot úgy ágyazzuk be a háromdimenziós térbe, hogy minden egyes pontjának a harmadik koordinátája 1 legyen. Így a fenti egyenlet helyett az $x - 2y - 3z = 0$ síknak és a $z = 1$ síknak a metszetéről van szó. Mivel a második sík rögzített, ezért az első sík teljesen meghatározza a szoban forgó egyenest. Sőt, mi több, meghatározza az egyenes „végtelen távoli pontját” is (vagyis az egyenes irányát). Nevezetesen a $z = 0$ esetben kapott $x = 2y$ lesz ez az irány. Ilyen módon a kétdimenziós síkot „kiegészítettük” az első pont feladatai között már tárgyalt kétdimenziós projektív síkká.

Ez az eljárás tetszőleges alakzatnál automatikusan elvégezhető. Tekintsük például az $x^2 + 2y^2 = 4$, illetve az $x^2 - 2y^2 = 4$ egyenletekkel adott ellipszist, illetve hiperbolát. Ezek az $x = \frac{x_1}{x_3}$ és $y = \frac{x_2}{x_3}$ helyettesítéssel (és x_3^2 -nel való szorzás után) a következőbe mennek át: $x_1^2 + 2x_2^2 - 4x_3^2 = 0$, illetve $x_1^2 - 2x_2^2 - 4x_3^2 = 0$. Látható, hogy a kapott egyenletek bal oldalán homogén (másodfokú) polinomok állnak. Ha itt végezzük el az $x_3 = 0$ helyettesítést, akkor azt kapjuk, hogy az első alakzatnak nincs végtelen távoli pontja, míg a másodiknak kettő van. Ezért ellipszis az első és ezért hiperbola a második.

Világos, hogy a fenti koordináták mindegyikét tetszőleges számmal szorozva az eredeti alakzatnak ugyanazt a pontját kapjuk.

A fenti eljárás az idézett feladatoknál már tárgyalt következő képet sugallja:

Az n -dimenziós tér minden pontját az $(n + 1)$ -dimenziós vektortér egy-egy egyenese jellemzi. Az egyenesek jellemzése síkokkal stb. történik. Tekintettel arra, hogy ebben a modellben a „végtelen távoli alakzatok” is helyet kapnak, ezért nem az euklideszi teret, hanem az úgynevezett *projektív teret* modelleztük. Ennek megfelelően a projektív sík modelljében az elemek a rögzített 0 ponton átmenő alakzatok; a pontokat az egyenesek, az egyeneseket a síkok helyettesítik.

Azt is láthatjuk, hogy a kapott polinom az x_i -k rögzítése esetén az y_j -knek, míg ezek rögzítése esetén az előbbieknél lesz homogén lineáris polinomja. Amiket kaptunk, azok tehát bilineáris formák.

2. A vektoros „felfogás” tömörsége főleg akkor lesz hasznos, amikor az új bázisra való áttérést fogjuk vizsgálni. A megértést viszont elősegíti az, ha a fenti kapcsolatokat bilineáris és kvadratikusság formákon is végigkövetjük. Tekintsük a $b(x_1, \dots, x_n, y_1, \dots, y_n) = \sum b_{i,j} x_i y_j$ bilineáris formát.

Itt az $y_i \rightarrow x_i$ helyettesítéssel egy $q(x_1, \dots, x_n) = \sum b_{i,j} x_i x_j$ kvadratikusságot nyerünk. A nyert kvadratikusság alak pontosan akkor lesz a triviális 0 polinom, ha minden i, j párra $b_{i,j} = -b_{j,i}$, vagyis a fenti bilineáris formában $x_i y_j$ és $x_j y_i$ együtthatói egymásnak negatívjai. Ez azt jelenti, hogy az x_i -ket a megfelelő y_i -kkel felcserélve a bilineáris forma a negatívjába megy át; azaz antiszimmetrikus. Ha a $b(x_1, \dots, x_n, y_1, \dots, y_n)$ polinom egyben szimmetrikus is, azaz az x_i -ket a megfelelő y_i -kkel felcserélve a bilineáris forma nem változik, akkor ez a polinom csak az azonosan 0 polinom lehet; hiszen ez azt jelenti, hogy $b_{i,j} = b_{j,i}$ is igaz. (Vigyázat! A fenti polinom szimmetrikussága nem ugyanaz, mint a szimmetrikus polinomok esetében!)

Amennyiben a kvadratikusság adott, akkor nem tudunk különbséget tenni az $x_i x_j$ és az $x_j x_i$ között. Ennek megfelelően egy kvadratikusságot csak úgy tudunk megadni, hogy $q(x_1, \dots, x_n) = \sum_{i \leq j} c_{i,j} x_i x_j$. Ehhez a kvadratikussághoz viszont könnyen megtalálhatjuk azt a

$b(x_1, \dots, x_n, y_1, \dots, y_n) = \sum b_{i,j} x_i y_j$ szimmetrikus bilineáris formát, amelyhez ez a kvadratikusság tartozik, nevezetesen, ha az együtthatókat úgy választjuk, hogy $b_{i,j} = b_{j,i} = \frac{1}{2} c_{i,j}$ legyen.

3. Noha nem tartozik az algebra tárgykörébe, de érdemes szót ejteni az alábbiakról. Az úgynevezett „analitikus sokaságok”-nak minden egyes pontban számos jellemző tulajdonságuk van. Mindezenelőtt a „ponthoz nagyon közel” van valamilyen „irányuk”, amit a derivált fejez ki. Ez egy „lineáris valami”; aminek a leírása (homogén) lineáris leképezésekkel történhet. Ez a lineáris leképezés lényegében leírja az „érintőteret”. A másik tulajdonság a következőképpen kapható. Mozgassuk el az érintőteret önmagával párhuzamosan egy kicsit, majd kezdjük visszafele tolni. Ezek a terek a vizsgált alakzatot egy 1-gyel kisebb dimenziós valamiben metszik; amelyeknek a formája egyre inkább hasonlítani kezd egymáshoz, ha a teret visszatoljuk az érintőtérbe. A határesetként kapott alakzat kvadratikusan, és megadja a sokaság „jellegét” a pontban.

Durván szólva egy „tisztességes” felület minden pontban vagy lapos, mint egy sík, vagy dudorodik, mint egy ellipszoid, vagy olyan, mint egy úgynevezett nyeregfelület stb.

Ennek megfelelően a „lineáris formák” (azaz homogén lineáris polinomok) mellett igen fontos szerepet játszanak a *kvadratikusan formák* (azaz homogén másodfokú polinomok) is. A kvadratikusan formák vizsgálata látszólag nem lineáris algebra. A már látott egyszerű, de fontos és hasznos trükkkel azonban minden kvadratikusan forma „átjátszható” a lineáris algebraiba. Nevezetesen azzal, hogy „a határozatlanok felét átnevezzük”. □

A valós test vagy annak résztestei esetében beszélhetünk egy kvadratikusan alak által felvett értékek előjeléről. Így például az x, y változók esetében $x^2 + y^2$ csak akkor nem pozitív, ha $x = y = 0$ (ekkor sem negatív!), az x^2 sohasem lehet negatív, míg az $x^2 - y^2$ polinom akár pozitív, akár negatív értékeket is felvehet. Ennek megfelelően a következőképpen osztályozzuk a kvadratikusan alakokat:

5.6. Definíció. Az \mathbb{R} feletti $Q(x)$ kvadratikusan alak, illetve a Q -t definiáló szimmetrikus bilineáris függvény kvadratikusan karakterét (kvadratikusan jellegét) a következőképpen definiáljuk:

- (i) Pozitív definit, ha $x \neq 0$ esetén $Q(x) > 0$.
- (i') Negatív definit, ha $-Q(x)$ pozitív definit.
- (ii) Pozitív szemidefinit, ha $Q(x) \geq 0$.
- (ii') Negatív szemidefinit, ha $-Q(x)$ pozitív szemidefinit.
- (iii) Indefinit minden más esetben.

A kvadratikusan karakter definícióját a későbbiekben még finomítani fogjuk. ■

A kvadratikusan alakok kvadratikusan jellegének felismeréséhez szükséges a kvadratikusan alakot alkalmas bázisban felírni. Ehhez mindenelőtt a kvadratikusan alakot jellemző mátrixra van szükségünk; valamint arra, miképpen változik meg egy kvadratikusan alak mátrixa akkor, ha új bázist választunk.

5.7. Definíció. Legyen $Q(x)$ kvadratikusan alak az \mathcal{U} vektortéren. Ekkor Q -nak az U bázisban megadott $[Q]^U$ mátrixán a Q kvadratikusan alakot definiáló $S(u, v)$ szimmetrikus bilineáris függvénynek az adott bázisban felírt $[S]^{U,U}$ mátrixát értjük. ■

5.13. Tétel. Legyenek U és E az \mathcal{U} vektortér bázisai, és σ az a transzformáció, amely U -t, az elemek sorrendjét is megtartva, E -be viszi. Ekkor tetszőleges Q kvadratikusan alak esetében $[Q]^E = [\sigma]^\dagger [Q]^U [\sigma]$.

Bizonyítás. Az 5.5. Definíció alapján azonnal következik az 5.9. Tételből. ■

Feladatok

1. A kvadratikus alakok vektorterének alterét alkotják-e az adott kvadratikus karakterű kvadratikus alakok?
2. Hogyan változik meg egy kvadratikus alak mátrixa, ha a bázison egy elemi transzformációt végzünk?
3. Bizonyítsuk be, hogy egy \mathbb{R} feletti \mathbf{Q} kvadratikus alak akkor és csak akkor definit, ha $\mathbf{x} \neq \mathbf{0}$ esetén $\mathbf{Q}(\mathbf{x}) \neq 0$.
4. Bizonyítsuk be, hogy \mathbb{Q} felett van olyan \mathbf{Q} indefinit kvadratikus alak, amelyre $\mathbf{x} \neq 0$ esetében $\mathbf{Q}(\mathbf{x}) \neq 0$.
5. A valós számokból álló (a_1, \dots, a_n, \dots) végtelen sorozatok \mathbb{R} felett vektorteret alkotnak a komponensenkénti műveletekre. Mutassuk meg, hogy a $\sum_1^{\infty} c_i, ja_i^2$ kifejezés kvadratikus alak, de nincs kvadratikus karaktere.
6. Mutassuk meg, hogy a valós számokból álló azon (a_1, \dots, a_n, \dots) végtelen sorozatok, amelyekben csak véges sok elem nem 0, az előző feladatban adott vektortérnek alterét alkotják. Mutassuk meg, hogy itt a $\sum_1^{\infty} c_{i,j}a_i^2$ kvadratikus alaknak létezik kvadratikus karaktere.
7. Tekintsük azokat a valós számokból álló (a_1, \dots, a_n, \dots) végtelen sorozatokat, amelyekre a $\sum_1^{\infty} a_i^2$ sor konvergens. Bizonyítsuk be, hogy ezek a sorozatok vektorteret alkotnak \mathbb{R} felett. Mutassuk meg, hogy itt is léteznek olyan kvadratikus alakok, amelyeknek létezik kvadratikus karakterük.
8. Legyen \mathbf{Q} és \mathbf{P} két kvadratikus alak. Mi mondható a $\mathbf{Q} + \mathbf{P}$ kvadratikus karakteréről a tagok kvadratikus karakterének ismeretében?
9. Bizonyítsuk be, hogy bármely kvadratikus alak előáll két pozitív definit kvadratikus alak különbségeként.
10. Adott egy kvadratikus alak mátrixa. Mondjunk olyan feltételeket, amelyek alapján ránézésre eldönthetjük, hogy ez a kvadratikus alak biztosan indefinit, illetve biztosan nem definit.
11. Egy kvadratikus alak mátrixának a determinánsa 0. Mit jelent ez?
12. Bizonyítsuk be, hogy egy bilineáris függvény mátrixának a rangja nem változik meg, ha új bázisra térünk.

4. Kvadratikus alakok négyzetösszegé transzformálása

Mint említettük, a kvadratikus alakok jellegének a megállapításához a kvadratikus alakokat olyan bázisban szükséges felírni, amelyben ez a jelleg világosan látható. Ebben a pontban kizárólag a valós test feletti vektorterekkel foglalkozunk.

5.14. Tétel. *A valós test feletti véges dimenziós \mathcal{U} térben minden $\mathbf{Q}(\mathbf{x})$ kvadratikus alak négyzetösszegé transzformálható, azaz létezik olyan $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázis, amelyben tetszőleges $\mathbf{x} = \sum_i x_i \mathbf{u}_i$ vektor esetén $\mathbf{Q}(\mathbf{x}) = \sum_i c_i x_i^2$, ahol $c_i = \mathbf{Q}(\mathbf{u}_i)$.*

A Tételt szimmetrikus bilineáris függvényekre átfogalmazva bizonyítjuk be; ehhez szükségünk van egy fogalomra:

5.8. Definíció. Legyen $\mathbf{A} : \mathcal{U} \times \mathcal{U} \rightarrow K$ tetszőleges szimmetrikus bilineáris függvény. Legyenek $\mathbf{u}, \mathbf{v} \in \mathcal{U}$, és $\mathcal{V}, \mathcal{W} \leq \mathcal{U}$.

Azt mondjuk, hogy \mathbf{u} és \mathbf{v} \mathbf{A} -ortogonálisak (vagy röviden ortogonálisak), ha $\mathbf{A}(\mathbf{u}, \mathbf{v}) = 0$. Ezt a relációt $\mathbf{u} \perp_{\mathbf{A}} \mathbf{v}$ (vagy röviden $\mathbf{u} \perp \mathbf{v}$) jelöli. Azt mondjuk, hogy \mathbf{u} (\mathbf{A} -)ortogonális \mathcal{V} -re (jelben $\mathbf{u} \perp \mathcal{V}$, ha $\mathbf{u} \perp \mathbf{v}$, minden $\mathbf{v} \in \mathcal{V}$ esetében. Azt mondjuk, hogy \mathcal{V} (\mathbf{A} -)ortogonális \mathcal{W} -re, ha minden $\mathbf{v} \in \mathcal{V}$ esetén $\mathbf{v} \perp \mathcal{W}$.

Tetszőleges \mathcal{V} esetében legyen $\mathcal{V}^\perp = \{\mathbf{u} \in \mathcal{U} \mid \mathbf{u} \perp \mathcal{V}\}$.

Az \mathcal{U} vektortér egy $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n, \dots\}$ bázisát (\mathbf{A} -)ortogonálisnak nevezzük, ha $i \neq j$ esetén $\mathbf{u}_i \perp \mathbf{u}_j$. ■

5.15. Tétel. *Az 5.8. Definíció jelöléseivel a következők igazak:*

Az ortogonalitás szimmetrikus reláció. $\mathbf{u} \perp \mathcal{V}$ akkor és csak akkor teljesül, ha \mathbf{u} ortogonális \mathcal{V} egy bázisára. Igazak továbbá az alábbiak:

- (1) $\mathcal{V} \leq \mathcal{W}$ esetén $\mathcal{W}^\perp \leq \mathcal{V}^\perp$.
- (2) $\mathcal{V}^{\perp\perp} \geq \mathcal{V}$.
- (3) $\mathcal{V}^{\perp\perp\perp} = \mathcal{V}^\perp$.
- (4) Ha \mathbf{A} a \mathcal{V} -n (pozitív) definit, akkor $\mathcal{V}^\perp \cap \mathcal{V} = \{0\}$.
- (5) Ha \mathcal{V} még véges dimenziós is, akkor $\mathcal{U} = \mathcal{V}^\perp \oplus \mathcal{V}$.
- (6) Ez utóbbi esetben \mathcal{V} -nek létezik ortogonális bázisa (azaz olyan bázis, amelynek elemei páronként ortogonálisak egymásra).

Bizonyítás. Az ortogonalitás szimetriája tüstént következik \mathbf{A} szimmetrikusságából.

Ha $\mathbf{u} \perp \mathcal{V}$, akkor persze \mathbf{u} ortogonális \mathcal{V} minden bázisára. Fordítva, ha \mathbf{u} bizonyos vektorokra ortogonális, akkor a bilinearitás folytán ortogonális e vektorok tetszőleges lineáris kombinációjára is; és így ortogonális az ezen vektorok által kifeszített altérre is.

(1): Ha $\mathbf{x} \in \mathcal{W}^\perp$, akkor $\mathbf{x} \perp \mathbf{w}$ minden $\mathbf{w} \in \mathcal{W}$ esetén, így $\mathbf{x} \perp \mathbf{v}$ minden $\mathbf{v} \in \mathcal{V}$ esetén is, azaz $\mathbf{x} \in \mathcal{V}^\perp$.

(2): Legyen $\mathbf{v} \in \mathcal{V}$. Ekkor, minden $\mathbf{x} \in \mathcal{V}^\perp$ esetén $\mathbf{x} \perp \mathbf{v}$, azaz $(\mathcal{V}^\perp)^\perp$ definíciója szerint $\mathbf{v} \in \mathcal{V}^{\perp\perp}$.

(3): Legyen $\mathcal{W} = \mathcal{V}^\perp$, $\mathcal{S} = \mathcal{W}^\perp$, $\mathcal{T} = \mathcal{S}^\perp$. (2)-t alkalmazva \mathcal{V} helyett \mathcal{W} -re, azt kapjuk, hogy $\mathcal{T} \supseteq \mathcal{W}$. (2) szerint $\mathcal{S} \supseteq \mathcal{V}$, amiből (1) alapján $\mathcal{W} \supseteq \mathcal{T}$ következik.

(4): Legyen $\mathbf{v} \in \mathcal{V}^\perp \wedge \mathcal{V}$. Ekkor $\mathbf{v} \perp \mathbf{v}$, azaz $Q(\mathbf{v}) = 0$, ami a pozitív definités alapján csak akkor lehet, ha $\mathbf{v} = 0$.

(5) és (6): E két tulajdonságot egyszerre bizonyítjuk, a \mathcal{V} -nek a k dimenziójára vonatkozó teljes indukcióval. A $k = 1$ esetben (6) triviálisan igaz. Tegyük fel, hogy (6) igaz a $k = n$ esetben és legyen $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ a \mathcal{V} egy ortogonális bázisa. Azt kell megmutatni, hogy tetszőleges $\mathbf{u} \in \mathcal{U}$ felírható $\mathbf{u} = \mathbf{v} + \mathbf{w}$ alakban, ahol $\mathbf{v} \in \mathcal{V}$ és $\mathbf{w} \in \mathcal{V}^\perp$. Legyen:

$$\mathbf{v} = \frac{\mathbf{A}(\mathbf{u}, \mathbf{v}_1)}{\mathbf{A}(\mathbf{v}_1, \mathbf{v}_1)} \cdot \mathbf{v}_1 + \dots + \frac{\mathbf{A}(\mathbf{u}, \mathbf{v}_n)}{\mathbf{A}(\mathbf{v}_n, \mathbf{v}_n)} \cdot \mathbf{v}_n.$$

(Itt $\mathbf{A}(\mathbf{v}_i, \mathbf{v}_i) \neq 0$ a definités miatt.) Nyilván $\mathbf{v} \in \mathcal{V}$; továbbá a bilinearitás és a bázisvektorok ortogonalitása alapján $\mathbf{A}(\mathbf{v}, \mathbf{v}_i) = \mathbf{A}(\mathbf{u}, \mathbf{v}_i)$. Ebből azonnal következik, hogy a $\mathbf{w} = \mathbf{u} - \mathbf{v}$ vektor ortogonális a felírt bázisra; így, a tétel második állítása szerint ortogonális az egész altérre.

Eszerint (5) igaz, ha $\dim(\mathcal{V}) = n$. Tekintsünk most egy — a feltételeknek megfelelő — $(n+1)$ -dimenziós \mathcal{V} alteret és annak egy n -dimenziós \mathcal{W} alterét a $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ ortogonális bázissal. Ez a bázis \mathcal{V} -ben lineárisan független, ezért kiegészíthető egy \mathbf{u} vektorral a \mathcal{V} egy bázisává. Az indukciós feltétel alapján — (5)-öt felhasználva — ehhez található egy olyan \mathbf{w} vektor, amellyel \mathbf{V} -t kiegészítve a \mathcal{V} egy ortogonális bázisát kapjuk. ■

Kiegészítés (SCHMIDT-féle ortogonalizáció). Ha $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} vektortér egy bázisa, akkor a \mathbf{Q} kvadratikus alakhoz található olyan $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ ortogonális bázis, amelyre $\langle \mathbf{v}_1, \dots, \mathbf{v}_i \rangle = \langle \mathbf{u}_1, \dots, \mathbf{u}_i \rangle$ teljesül minden $1 \leq i \leq n$ esetén.

Bizonyítás. Az 5.15. Tétel (6) pontjának bizonyításánál válasszuk \mathcal{V} -nek rendre az \mathbf{U} bázis első i számú vektora által generált alteret. Ekkor a konstruált \mathbf{V} bázisok rendre rendelkezni fognak a kívánt tulajdonsággal. ■

5.16. Tétel. Legyen \mathbf{A} egy szimmetrikus bilineáris függvény a véges dimenziós \mathcal{U} téren. Ekkor a tér felbomlik a páronként ortogonális \mathcal{R} , \mathcal{S} és \mathcal{T} alterek összegére, ahol \mathbf{A} -nak az \mathcal{R} -re vonatkozó megszorítása pozitív definit, a \mathcal{T} -re vonatkozó megszorítása negatív definit, és az \mathcal{S} -re vonatkozó megszorítása azonosan 0.

Tehetlenségi tétel (SYLVESTER): A fenti alterek dimenziói — r , s és t — nem függenek az alterektől, csak a szimmetrikus bilineáris függvénytől (illetve a kvadratikus alaktól!).

Bizonyítás. Legyen \mathcal{R} egy maximális olyan altér, amelyen \mathbf{A} pozitív definit. A maximalitás miatt tetszőleges $\mathbf{w} \in \mathcal{R}^\perp$ vektorhoz van olyan $\mathbf{v} \in \mathcal{R}$, hogy az $\mathbf{u} = \mathbf{v} + \mathbf{w} \neq 0$ vektorra $\mathbf{A}(\mathbf{u}, \mathbf{u}) \leq 0$ teljesül. Ebből:

$$0 \geq \mathbf{A}(\mathbf{u}, \mathbf{u}) = \mathbf{A}(\mathbf{v}, \mathbf{v}) + 2\mathbf{A}(\mathbf{v}, \mathbf{w}) + \mathbf{A}(\mathbf{w}, \mathbf{w}) = \mathbf{A}(\mathbf{v}, \mathbf{v}) + \mathbf{A}(\mathbf{w}, \mathbf{w}) \geq \mathbf{A}(\mathbf{w}, \mathbf{w});$$

és így \mathbf{A} az \mathcal{R}^\perp altéren negatív szemidefinit. Analóg módon választhatunk \mathcal{R}^\perp -ben egy maximális olyan \mathcal{I} alteret, ahol \mathbf{A} negatív definit. Most a fenti eljárással a \mathcal{I} -re ortogonális \mathcal{R}^\perp -beli vektorokra egy olyan \mathcal{S} altér adódik, amelyen \mathbf{A} pozitív szemidefinit. Tekintettel arra, hogy \mathcal{S} -en \mathbf{A} negatív szemidefinit is, ezért identikusan 0.

Ezzel megkaptuk a kívánt három páronként ortogonális alteret. Legyen ezeknek a dimenziója rendre r, s, t . Tegyük fel, hogy találtunk három hasonló tulajdonsággal rendelkező \mathcal{R}' , \mathcal{S}' és \mathcal{I}' alteret, amelyeknek a dimenziója megfelelően r', s', t' . (Ilyen újabb altérhármass mindig van is, kivéve, ha a fenti alterek közül valamelyik az egész tér. Az \mathcal{S} azonban abszolút egyértelmű és megegyezik \mathcal{U}^\perp -sel.)

Ha n az egész tér dimenziója, akkor $r + s + t = r' + s' + t' = n$. Mivel $(\mathcal{R} + \mathcal{S}) \cap \mathcal{I}'$ csak a nullvektort tartalmazhatja, ezért $r + s + t' \leq n = r + s + t$, azaz $t' \leq t$. Ugyanígy látható be az is, hogy $t \leq t'$; tehát $t' = t$. Hasonlóképpen kaphatjuk az $r' = r$ egyenlőséget is; amiből már $s' = s$ is következik. ■

Most bizonyítjuk be az 5.14. Tételt, illetve e tételnek az alábbi kiegészítését:

5.14/B. Tétel. Az n -dimenziós tér bármely $\mathbf{Q}(\mathbf{x})$ kvadratikus alakjához található olyan $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázis, amelyben a kvadratikus alak négyzetösszegé transzformálódik. Ez azt jelenti, hogy léteznek olyan c_1, \dots, c_n skalárok, hogy az $\mathbf{x} = x_1\mathbf{u}_1 + \dots + x_n\mathbf{u}_n$ helyen a $c_1x_1^2 + \dots + c_nx_n^2$ értéket veszi fel (tehát $c_1x_1^2 + \dots + c_nx_n^2$ alakúvá „válík”).

A c_i -k között a pozitívak, negatívak és nullák száma nem függ a bázistól.

Bizonyítás. Tekintsük a \mathbf{Q} kvadratikus alakhoz tartozó \mathbf{A} szimmetrikus bilineáris függvényt. Az 5.15. Tétel szerint az eredeti \mathcal{U} vektortér felbomlik páronként ortogonális $\mathcal{R}, \mathcal{S}, \mathcal{I}$ alterek direkt összegére úgy, hogy \mathbf{A} az első altéren pozitív definit, a másodikon azonosan 0 és a harmadikon negatív definit. Az 5.15. Tétel szerint az \mathcal{R} és a \mathcal{I} altereken létezik \mathbf{A} -ortogonális bázis; míg az \mathcal{S} altéren minden bázis ilyen. E három független rendszer egyesítése az eredeti altérnek egy \mathbf{A} -ortogonális bázisát adja. Ez a bázis az \mathbf{A} -ortogonalitás miatt rendelkezik a kívánt tulajdonsággal.

Ha az alterek dimenziója — megfelelően — r, s, t , akkor e számok éppen a pozitív, 0 és negatív együtthatók számát adják meg. Tegyük fel, hogy adott egy olyan bázis, amelyben a szóban forgó kvadratikus alak a kívánt alakot ölti és az első r darab pozitív, a következő s darab 0 és az utolsó t darab negatív. Ekkor az ezek által kifeszített páronként ortogonális altereken a bilineáris függvény rendre pozitív definit, azonosan 0, illetve negatív definit. A tehetetlenségi tétel szerint viszont ezek a dimenziók nem függhetnek a bázistól. ■

Megjegyzés. Az 5.14/B. Tétel adja azt a finomítást, amire a kvadratikus karakter elnevezésnél már utaltunk. A \mathbf{Q} kvadratikus alaknál a fellépő r, s, t számok invariánsak, és ezek a számok adják meg az adott kvadratikus alak kvadratikus karakterét. □

Feladatok

1. Legyen $\mathbf{A}(\mathbf{u}, \mathbf{v})$ tetszőleges bilineáris függvény az \mathcal{U} téren, amelyre az \mathbf{A} -ortogonalitás szimmetrikus. Következik-e ebből, hogy \mathbf{A} is szimmetrikus?
2. Legyen \mathbf{A} szimmetrikus, nem definit bilineáris függvény az \mathbb{R} feletti \mathcal{U} vektortéren. Bizonyítsuk be, hogy van olyan $\mathbf{u} \in \mathcal{U}$, amelyik önmagára \mathbf{A} -ortogonális.
3. Bizonyítsuk be, hogy az \mathbb{R} feletti \mathcal{U} vektortéren értelmezett \mathbf{A} szimmetrikus bilineáris függvényhez akkor és csak akkor van olyan bázis, hogy e bázis \mathbf{u} elemeire $\mathbf{A}(\mathbf{u}, \mathbf{u}) \neq 0$, ha bármely $\mathcal{V} \leq \mathcal{U}$ altérre $\mathcal{V}^{\perp\perp} = \mathcal{V}$.
4. Síkbeli görbék homogén koordinátái legyenek x, y, z . Milyen görbét jellemez egy pozitív definit $(x^2 + y^2 + z^2)$, egy szemidefinit $(x^2 + y^2)$ és egy indefinit $(x^2 - y^2)$ kvadratikusan alak?
5. Tetszőleges \mathbf{B} bilineáris függvényre definiáljuk az ortogonalitást a szimmetrikus bilineáris függvények esetéhez hasonlóan. Mutassuk meg, hogy az ortogonalitás szimmetriájából nem következik \mathbf{B} szimmetriája. Milyen \mathbf{B} esetében lesz minden vektor önmagára ortogonális?
6. Mutassuk meg, hogy az 5.15. Tételben mindig van olyan \mathcal{V} altér, amelyre $\mathcal{V}^{\perp} \wedge \mathcal{V} \neq \{0\}$, amennyiben \mathbf{A} nem definit.
7. Mutassuk meg, hogy ha \mathcal{U} végtelen dimenziós és \mathbf{A} pozitív definit, akkor létezik olyan $\mathcal{V} \leq \mathcal{U}$, amelyre $\mathcal{V}^{\perp} \oplus \mathcal{V} < \mathcal{U}$.
8. Bizonyítsuk be, hogy az 5.16. Tételben szereplő \mathcal{S} altér valóban egyértelmű és megegyezik \mathcal{U}^{\perp} -sel.
9. Legyen $\mathcal{U}' \leq \mathcal{U}$ és $\mathcal{R}', \mathcal{S}', \mathcal{I}'$ az \mathcal{U} -beli $\mathcal{R}, \mathcal{S}, \mathcal{I}$ alterekkel a \mathbf{Q} kvadratikusan alakhoz, illetve ennek \mathcal{U}' -re való megszorításához analóg módon definiált alterek; az 5.16. Tételnek megfelelően. Bizonyítsuk be, hogy általában nem igaz az, hogy $\mathcal{R}' \leq \mathcal{R}$ vagy $\mathcal{I}' \leq \mathcal{I}$. Bizonyítsuk be, hogy az $\mathcal{R}, \mathcal{S}, \mathcal{I}$ megválasztható úgy, hogy ez teljesüljön. Bizonyítsuk be, hogy $\mathcal{S}' \leq \mathcal{S}$ még ekkor sem mindig igaz. Bizonyítsuk be, hogy szemidefinit kvadratikusan alak esetében ez is elérhető.
10. Legyen \mathbf{Q} kvadratikusan alak a véges dimenziós \mathcal{U} vektortéren, és $\mathcal{R}, \mathcal{S}, \mathcal{I}$ az 5.16. Tételben konstruált három altér. Legyenek $\mathcal{R}_1 \leq \mathcal{R}, \mathcal{S}_1 \leq \mathcal{S}, \mathcal{I}_1 \leq \mathcal{I}$ és $\mathcal{U}_1 = \mathcal{R}_1 \oplus \mathcal{S}_1 \oplus \mathcal{I}_1$. Bizonyítsuk be, hogy ekkor létezik olyan $\mathcal{U}_2 \perp \mathcal{U}_1$ altér, amelyre $\mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2$.

5. Bilineáris függvények és kvadratikusan alakok a komplex térben

Mind az analízisben, mind a geometriában, leginkább az algebrai geometriában igen fontos szerepet játszanak a komplex számtest feletti vektorterek; illetve az ezekben értelmezett bilineáris és kvadratikusan alakok. A komplex tér esetében is fontos jellemzést ad a kvadratikusan alak kvadratikusan karaktere. Illetve csak adhatna, ha a kvadratikusan karakternek minden további nélkül volna értelme. Tekintsünk ugyanis egy tetszőleges \mathbf{A} bilineáris

függvényt. Erre $\mathbf{A}(i\mathbf{x}, i\mathbf{x}) = i^2 \mathbf{A}(\mathbf{x}, \mathbf{x}) = -\mathbf{A}(\mathbf{x}, \mathbf{x})$ adódik. Így nem beszélhetünk arról, hogy egy kvadratikus alak pozitív (szemi)definit lehetne. Ez ellen úgy „védekezhetünk”, hogy az egyik (az első) tényezőtől nem a szereplő skalárt, hanem annak a konjugáltját emeljük ki. Ennek a következtetés végigvételéhez néhány definícióra és néhány eddigi eredmény analógiájára van szükség.

Ezen vizsgálatok folyamán mindig a komplex számtest feletti véges dimenziós vektortereket fogunk tekinteni.

5.9. Definíció. A \mathbb{C} komplex test feletti \mathcal{U}, \mathcal{V} vektorterek esetében a $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezést másodfajú homogén lineárisnak nevezzük, ha additív ($\varphi(\mathbf{u} + \mathbf{v}) = \varphi(\mathbf{u}) + \varphi(\mathbf{v})$) és másodfajú homogén, azaz $c \in \mathbb{C}$ esetén $\varphi(c\mathbf{u}) = \bar{c}\varphi(\mathbf{u})$.

Amennyiben másodfajú homogén lineáris függvények is szerepelnek, akkor a homogén lineáris függvényeket elsőfajú homogén lineáris függvényeknek nevezzük. Azt fogjuk mondani, hogy az elsőfajú homogén lineáris leképezés jellege $+1$, a másodfajúé pedig -1 . ■

Megjegyzés. Természetesen tetszőleges test esetében is beszélhetünk másodfajú lineáris függvényekről. Ekkor is $c \rightarrow \bar{c}$ a komplex konjugálás. Az eredmények azért akkor is érvényben maradnak, ha ez a megfeleltetés az identitás. Valójában csak annyi szükséges, hogy a megfeleltetés összeg- és szorzattartó involúció ($\overline{\overline{c}} = c$) legyen. □

A továbbiakban mindenekelőtt az lesz a célunk, hogy a kétféle homogén lineáris leképezés közötti kapcsolatot minél jobban leírjuk.

5.17. Tétel. Legyenek $\mathcal{U} \xrightarrow{\varphi} \mathcal{V} \xrightarrow{\psi} \mathcal{W}$ homogén lineáris leképezések. Ekkor $\psi\varphi$ is homogén lineáris és jellege megegyezik ψ jellegének és φ jellegének a szorzatával.

Bizonyítás. Mivel mindkét leképezés összegtartó, ezért szorzatuk is az. Legyen $\varphi(c\mathbf{u}) = \mu(c)\varphi(\mathbf{u})$ és $\psi(d\mathbf{v}) = \nu(d)\psi(\mathbf{v})$. Ekkor $\psi\varphi(c\mathbf{u}) = \nu(\mu(c))\psi\varphi(\mathbf{u})$. Ha ν és μ mindegyike a konjugálás, vagy mindegyikük az identitás, akkor $\nu \circ \mu$ identitás, mert a konjugálás involutórius. Ha a kettő valamelyike az identitás, a másik meg a konjugálás, akkor $\nu \circ \mu$ a konjugálás. ■

A kapcsolat leírásához egy speciális fajta másodfajú homogén lineáris leképezést fogunk felhasználni.

5.10. Definíció. A $\varphi(\mathbf{u}) = \bar{\mathbf{u}}$ jelöléssel megadott $\varphi : \mathcal{U} \rightarrow \mathcal{U}$ másodfajú homogén lineáris leképezést konjugálásnak fogjuk nevezni, ha involutórius, azaz, ha $\overline{\overline{\mathbf{u}}} = \mathbf{u}$.

Összegtartó $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezés esetén a $\bar{\varphi} : \mathbf{u} \mapsto \varphi(\bar{\mathbf{u}})$ összefüggéssel definiált $\bar{\varphi} : \mathcal{U} \rightarrow \mathcal{V}$ leképezést a φ konjugáltjának nevezzük. ■

Megjegyzés. Természetesen a konjugálás minden vektortéren más és más. Sőt, mi több, egyetlen vektortéren is léteznek különböző konjugálások. □

5.18. Tétel. Egy homogén lineáris leképezés előtt vagy után tetszőleges konjugálást végrehajtva a kapott leképezés jellege az eredeti leképezés jellegének a negatívja. Minden homogén lineáris leképezés megkapható ilyen módon egy ellenkező jellegű homogén lineáris leképezésből; tetszőlegesen választott konjugálással.

Ha a $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ és a $\overline{\varphi} : \mathcal{U} \rightarrow \mathcal{V}$ valamelyike additív leképezés, akkor a másik is az. Teljesül a $\overline{\overline{\varphi}} = \varphi$ összefüggés. Ha valamelyikük homogén lineáris, akkor a másik is az; és különböző jellegűek.

Bizonyítás. Az első állítás azonnal adódik az 5.17. Tételből, tekintettel arra, hogy a konjugálás másodfajú.

A második állítás bizonyításához tekintsünk egy tetszőleges $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ homogén lineáris leképezést és legyen α az \mathcal{U} és β a \mathcal{V} vektortér egy-egy konjugálása. Definíció szerint ez azt jelenti, hogy mind $\alpha \circ \alpha$, mind $\beta \circ \beta$ az identitás. Ezért $\varphi = \varphi \circ \beta \circ \beta = \alpha \circ \alpha \circ \varphi$. Ezt $\varphi = (\varphi \circ \beta) \circ \beta = \alpha \circ (\alpha \circ \varphi)$ alakba írva, az első állítás alapján máris következik a második állítás.

Tekintsünk most egy $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ additív leképezést, és legyen $\nu : \mathcal{U} \rightarrow \mathcal{U}$ a $\nu(\mathbf{u}) = \overline{\mathbf{u}}$ összefüggéssel definiált konjugálás. Ekkor definíció szerint $\overline{\varphi} = \varphi \circ \nu$. A tétel eddig bizonyított állításaiból most már azonnal következnek a továbbiak. ■

Ezáltal a másodfajú lineáris leképezéseket visszavezettük az elsőfajúakra és egy „kitüntetett” típusú másodfajúra. Most ez utóbbinak adjuk meg a leírását:

5.19. Tétel. Legyen $\mathbf{U} = \{\mathbf{u}_i \mid i \in I\}$ az \mathcal{U} vektortér egy bázisa. Ekkor az $\mathbf{u} = \sum_i c_i \mathbf{u}_i$

vektorhoz hozzárendelve a $\overline{\mathbf{u}} = \sum_i \overline{c_i} \mathbf{u}_i$ vektort, konjugálást kapunk.

Fordítva, minden konjugáláshoz található olyan bázis, amelyben ez a fenti formát ölti.

Bizonyítás. Az első állítás bizonyítása csupa triviális számolásból áll, amelyeket az olvasóra bízunk. Szemléltetésül csak a legkevésbé triviálisat közöljük, amely szerint a hozzárendelés másodfajú.

$$\overline{c\mathbf{u}} = \overline{\sum_i c c_i \mathbf{u}_i} = \sum_i \overline{c c_i} \mathbf{u}_i = \overline{c} \sum_i \overline{c_i} \mathbf{u}_i = \overline{c} \cdot \overline{\mathbf{u}}.$$

Tekintsük a \mathbb{C} feletti \mathcal{U} vektortér egy φ konjugálását. Mint tudjuk, \mathcal{U} természetes módon vektortér az \mathbb{R} valós test felett. Tekintettel arra, hogy egy valós szám konjugáltja önmaga, ezért φ az $\mathcal{U}_{\mathbb{R}}$ vektortérnek homogén lineáris leképezése. Mivel φ involúció, ezért

$$\varphi(\mathbf{u} + \varphi(\mathbf{u})) = \mathbf{u} + \varphi(\mathbf{u}) \quad \text{és} \quad \varphi(\mathbf{u} - \varphi(\mathbf{u})) = -(\mathbf{u} - \varphi(\mathbf{u})).$$

Mármint világos, hogy $\mathcal{V} = \{\mathbf{v} \mid \varphi(\mathbf{v}) = \mathbf{v}\}$ és $\mathcal{W} = \{\mathbf{w} \mid \varphi(\mathbf{w}) = -\mathbf{w}\}$ mindegyike altér \mathbb{R} felett, mégpedig a +1 és -1 sajátértékekhez tartozó sajátaltér. E két altérnek nyilván egyedül a \mathbf{o} a közös része és $\mathbf{u} = \frac{1}{2} \cdot [(\mathbf{u} + \varphi(\mathbf{u})) + (\mathbf{u} - \varphi(\mathbf{u}))]$ alapján \mathcal{U} a generátumuk.

Ha $\varphi(\mathbf{u}) = c\mathbf{u}$ valamilyen valós c -re, akkor $\varphi(i\mathbf{u}) = -i\varphi(\mathbf{u}) = -ic\mathbf{u} = -c(i\mathbf{u})$. Ez azt jelenti, hogy az i -vel való szorzás egy bijekciót létesít a valós feletti \mathcal{V} és \mathcal{W} alterek között.

Tekintsük most a $\mathcal{V}_{\mathbb{R}}$ altér egy tetszőleges $\mathbf{U} = \{\mathbf{u}_t \mid t \in T\}$ bázisát. A fenti bijekció és $\mathcal{U} = \mathcal{V} \oplus_{\mathbb{R}} \mathcal{W}$ következtében \mathbf{U} generátorrendszere $\mathcal{U}_{\mathbb{C}}$ -nek. Ha ezek egy $\sum_t (c_t + i d_t) \mathbf{u}_t$ lineáris kombinációja a \mathbf{o} -vektort adja, akkor tekintsük a \mathcal{V} -beli $\mathbf{v} = \sum_t c_t \mathbf{u}_t$ és $\mathbf{w} = \sum_t d_t \mathbf{u}_t$

vektorokat. Feltétel szerint $\mathbf{v} + i\mathbf{w} = \mathbf{o}$. Mivel $i\mathbf{w} \in \mathcal{W}$ és $\mathcal{V} \wedge \mathcal{W} = \{\mathbf{o}\}$, ezért csak $\mathbf{v} = \mathbf{w} = \mathbf{o}$ lehet. Tekintettel arra, hogy \mathbf{U} bázis, ezért minden $t \in T$ indexre $c_t = d_t = 0$ teljesül; így \mathbf{U} az $\mathcal{U}_{\mathbb{C}}$ térnek is bázisa. Az $\mathbf{u} = \sum_t (c_t + id_t)\mathbf{u}_t$ elemre

$$\varphi(\mathbf{u}) = \varphi\left(\sum_t (c_t + id_t)\mathbf{u}_t\right) = \sum_t c_t\mathbf{u}_t - \sum_t id_t\mathbf{u}_t = \sum_t \overline{c_t + id_t}\mathbf{u}_t. \quad \blacksquare$$

5.11. Definíció. A komplex számtest feletti $\mathcal{U}, \mathcal{V}, \mathcal{W}$ vektorterekre az $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ leképezést bilineárisnak nevezzük, ha $\mathbf{A}(\mathbf{u}, \mathbf{v})$ rögzített \mathbf{u} esetén \mathbf{v} -ben elsőfajú és rögzített \mathbf{v} esetén \mathbf{u} -ban másodfajú. A $\mathcal{W} = \mathbb{C}$ esetben bilineáris függvényről fogunk beszélni.

Tetszőleges $\mathbf{A} : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{C}$ esetén $\mathbf{Q}(\mathbf{x}) = \mathbf{A}(\mathbf{x}, \mathbf{x})$ kvadratikus alak. \blacksquare

Megjegyzés. Itt is megállapítható egy szoros kapcsolat a „bilineáris és kvadratikus polinomokkal”. Ez esetben viszont meg kell változtatni a polinom fogalmát. Így bizonyos határozatlanok helyébe azok „konjugáltját” kell írni. Ez „elrontja” a kvadratikus alakok szimmetrikusságát. Az x határozatlannak kvadratikus alakja például az $x^2 + 3x\bar{x} - 5\bar{x}^2$. \square

A valós esetben kapott eredmények legnagyobb része szinte szó szerint átvihető a komplex esetre is; csupán a megfelelő helyen a konjugáltat kell venni. Van azonban néhány egészen elütő eset. A bilineáris függvények természetesen vektorteret alkotnak a komplex számtest felett; éppúgy, mint a kvadratikus alakok. Az is világos, hogy az $\mathbf{A}(\mathbf{u}, \mathbf{v}) \mapsto \mathbf{A}(\mathbf{x}, \mathbf{x})$ megfeleltetés homomorfizmus. A valós esettel ellentétben itt a következő igaz:

5.20. Tétel. Az $\mathbf{A}(\mathbf{u}, \mathbf{v}) \mapsto \mathbf{A}(\mathbf{x}, \mathbf{x})$ megfeleltetés bijektív.

Bizonyítás. Azt kell belátni, hogy egyedül az azonosan nulla bilineáris leképezésnek felel meg az azonosan nulla kvadratikus alak.

Mint a valós esetben láttuk, ha $\mathbf{A}(\mathbf{x}, \mathbf{x}) = 0$, akkor

$0 = \mathbf{A}(\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) = \mathbf{A}(\mathbf{u}, \mathbf{u}) + \mathbf{A}(\mathbf{u}, \mathbf{v}) + \mathbf{A}(\mathbf{v}, \mathbf{u}) + \mathbf{A}(\mathbf{v}, \mathbf{v}) = 0 + \mathbf{A}(\mathbf{u}, \mathbf{v}) + \mathbf{A}(\mathbf{v}, \mathbf{u}) + 0$
alapján $\mathbf{A}(\mathbf{v}, \mathbf{u}) = -\mathbf{A}(\mathbf{u}, \mathbf{v})$. Ebből

$$i \cdot \mathbf{A}(\mathbf{u}, \mathbf{v}) = \mathbf{A}(\mathbf{u}, i \cdot \mathbf{v}) = -\mathbf{A}(i \cdot \mathbf{v}, \mathbf{u}) = -(-i)\mathbf{A}(\mathbf{v}, \mathbf{u}) = (-i)\mathbf{A}(\mathbf{u}, \mathbf{v})$$

adódik, amiből valóban következik, hogy \mathbf{A} azonosan nulla. \blacksquare

Tekintsük az \mathcal{U} térnek egy rögzített $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és a \mathcal{V} térnek egy rögzített $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ bázisát. A bilinearitás következtében az $a_{i,j} = \mathbf{A}(\mathbf{u}_i, \mathbf{v}_j)$ számok, illetve az ezekből álló $[a_{i,j}]$ mátrix egyértelműen meghatározza a bilineáris függvényt; és az is világos, hogy adott számokból álló $[a_{i,j}]$ mátrix egyértelműen meghatároz egy bilineáris függvényt, amelyhez pontosan ez a mátrix tartozik.

Ha viszont adottak az $\mathbf{u} = \sum_i c_i \mathbf{u}_i$ és $\mathbf{v} = \sum_j d_j \mathbf{v}_j$ vektorok, akkor $\mathbf{A}(\mathbf{u}, \mathbf{v})$ nem

$$\sum_{i,j} a_{i,j} c_i d_j \text{ lesz, hanem } \sum_{i,j} a_{i,j} \bar{c}_i d_j.$$

Természetesen itt is tekinthetjük azt az $\alpha : \mathcal{V} \rightarrow \mathcal{U}$ homogén lineáris leképezést, amelynek a mátrixa $[a_{i,j}]$. Ebben az esetben viszont $\mathbf{u}^* \alpha(\mathbf{v})$ nem lesz minden nélküli egyenlő $\mathbf{A}(\mathbf{u}, \mathbf{v})$ -vel, ahogy ezt éppen most mutattuk meg. Ezen úgy segítünk, hogy a duális bázisra való áttérésnél az $\mathbf{u} = \sum_i c_i \mathbf{u}_i$ vektornak az $\mathbf{u}^* = \sum_i \overline{c_i} \mathbf{u}_i$ vektort feleltetjük meg.

Komplex tereknél mindig a fentiek szerint értelmezzük az \mathbf{u}^ -ot.*

Ennek megfelelően az $\mathbf{A}(\mathbf{u}, \mathbf{v})$ értékét nem az $[\mathbf{u}]^\dagger [\alpha][\mathbf{v}]$, hanem az $[\mathbf{u}]^* [\alpha][\mathbf{v}]$ mátrixszorzat adja meg; ahol $*$ a mátrixoknál tárgyalt *adjungáltat* (azaz a transzponált konjugáltját) jelöli.

Ennek megfelelően változik a „formula” az új bázisra való áttérésnél.

Még mindig gondot okoz a definitség definíciója. Az $x\overline{y}$ kvadratikus alak például az $x \rightarrow i, y \rightarrow 1$ helyettesítésnél az i értéket veszi fel, ami sem nem pozitív, sem nem negatív, sem nem 0. A definitségről tehát csak akkor beszélhetünk, ha a kvadratikus alak mindig valós értéket vesz fel.

5.21. Tétel. *Az $\mathbf{A}(\mathbf{u}, \mathbf{v})$ bilineáris függvényhez tartozó kvadratikus alak értékkészlete pontosan akkor áll csupa valós számból, ha $\mathbf{A}(\mathbf{v}, \mathbf{u}) = \overline{\mathbf{A}(\mathbf{u}, \mathbf{v})}$. Az ilyen bilineáris függvényeket Hermite-féléknek nevezik.*

Bizonyítás. Ha az \mathbf{A} bilineáris függvény Hermite-féle, akkor $\mathbf{A}(\mathbf{x}, \mathbf{x}) = \overline{\mathbf{A}(\mathbf{x}, \mathbf{x})}$ következtében a megfelelő kvadratikus alak valóban csak valós értékeket vesz fel.

Tetszőleges $\mathbf{A}(\mathbf{u}, \mathbf{v})$ bilineáris függvény esetében a $\mathbf{B}(\mathbf{u}, \mathbf{v}) = \overline{\mathbf{A}(\mathbf{v}, \mathbf{u})}$ ugyancsak bilineáris, hiszen a konjugálás miatt ez is \mathbf{u} -ban másodfajú és \mathbf{v} -ben elsőfajú. Ezért ezek $\mathbf{T}(\mathbf{u}, \mathbf{v}) = \mathbf{A}(\mathbf{u}, \mathbf{v}) - \mathbf{B}(\mathbf{u}, \mathbf{v})$ különbsége is bilineáris. Ha az adott bilineáris függvényhez tartozó kvadratikus alak csupa valós értéket vesz fel, akkor a $\mathbf{T}(\mathbf{u}, \mathbf{v})$ -hez tartozó kvadratikus alakra:

$$\mathbf{T}(\mathbf{x}, \mathbf{x}) = \mathbf{A}(\mathbf{x}, \mathbf{x}) - \overline{\mathbf{A}(\mathbf{x}, \mathbf{x})} = 0,$$

tetszőleges \mathbf{x} mellett. Az 5.20. Tétel szerint tehát $\mathbf{T}(\mathbf{u}, \mathbf{v})$ a nulla függvény, azaz \mathbf{A} valóban Hermite-féle. ■

5.12. Definíció. Hermite-féle bilineáris függvényekre és a hozzájuk tartozó kvadratikus alakokra a kvadratikus karaktert ugyanúgy értelmezzük, ahogy azt az 5.6. Definícióban tettük szimmetrikus bilineáris függvényekre és a megfelelő kvadratikus alakokra. ■

5.13. Definíció. Az \mathbf{A} -ortogonalitás és az \mathbf{A} -ortogonális bázis fogalmát hasonló módon definiáljuk, mint azt a valós esetre az 5.8. Definícióban, illetve az 5.15. Tételben tettük. ■

5.22. Tétel. *A komplex test feletti véges dimenziós \mathcal{U} térben minden $\mathbf{Q}(\mathbf{x})$ kvadratikus alak négyzetösszeggé transzformálható, azaz létezik olyan $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázis, amelyben tetszőleges $\mathbf{x} = \sum_i x_i \mathbf{u}_i$ vektor esetén $\mathbf{Q}(\mathbf{x}) = \sum_i c_i \overline{x_i} x_i$, ahol $c_i = \mathbf{Q}(\mathbf{u}_i)$.*

Bizonyítás. Hasonlóképpen történhet, mint az 5.14. Tételé valós esetben; illetve az Hermite-féle bilineáris függvényre is igaz a megfelelő tétel, mint a valós esetben a szimmetrikus bilineáris függvényekre. ■

5.23. Tétel. *Hermite-féle bilineáris függvényekre, illetve kvadratikus alakokra érvényes a tehetetlenségi tétel.*

Bizonyítás. A bizonyítás teljesen hasonló az 5.16. Tételben szereplő valós esetre vonatkozó bizonyításhoz. ■

Feladatok

1. Mutassuk meg, hogy a $\widehat{\varphi} : \mathbf{u} \mapsto \overline{\varphi(\mathbf{u})}$ leképezésre hasonlóak igazak, mint a $\overline{\varphi}$ leképezésre. Mutassuk meg, hogy $\widehat{\varphi}$ általában különbözik $\overline{\varphi}$ -től.

2. Mutassuk meg, hogy $\overline{\varphi \cdot \psi} = \overline{\varphi} \cdot \overline{\psi}$ nem feltétlenül teljesül.

3. Milyen feltételek esetén teljesül az előző feladatban szereplő egyenlőség?

4. Bizonyítsuk be, hogy minden bilineáris függvény egyértelműen felírható $\alpha + i \cdot \beta$ alakban, ahol α és β Hermite-féle bilineáris függvények.

5. Legyen $\Phi : \mathcal{U} \rightarrow \mathcal{U}^*$ másodfajú lineáris leképezés és $\alpha : \mathcal{U} \rightarrow \mathcal{U}$ lineáris transzformáció. Bizonyítsuk be, hogy $(\Phi(\mathbf{u}) \circ \alpha)(\mathbf{v})$ bilineáris függvény.

6. Tegyük fel, hogy az összes \mathbb{C} -vektortérben rögzítettünk egy-egy konjugálást. Bizonyítsuk be, hogy azok a $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ leképezések, amelyekre $\overline{\varphi} = \widehat{\varphi}$, vektorteret alkotnak a „természetes” műveletekre.

7. Tegyük fel, hogy az \mathcal{U} tér α, β transzformációira $\mathbf{u}^* \alpha \mathbf{v}$ és $\mathbf{u}^* \beta \mathbf{v}$ mindegyike Hermite-féle. Bizonyítsuk be, hogy $\alpha \beta = \beta \alpha$ esetén $\mathbf{u}^* \alpha \beta \mathbf{v}$ is az.

8. Bizonyítsuk be, hogy az $\mathcal{U} \times \mathcal{U} \rightarrow \mathbb{C}$ bilineáris függvények \mathbb{R} feletti vektorterében az Hermite-félék alteret alkotnak.

HATODIK FEJEZET

EUKLIDESZI TEREK

1. A valós euklideszi tér

A vektorterek eddigi vizsgálatában csak az úgynevezett *affin* tulajdonságok játszottak szerepet. (Ezért nevezik a vektortereket más szóval *affin* tereknek is.) Pontosabban szólva, eddig nem beszéltünk méretes vonatkozásokról. Nem tudtuk mérni a vektorok hosszát sem, csupán a párhuzamos vektorok arányáról beszélhettünk. Sok olyan kérdés van azonban, amelyeknél a távolságok vagy szögek is szerepelnek. Ezek nem csak a tényleges geometriai vizsgálatoknál fordulnak elő, hanem sok más esetben is lényeges szerepet játszanak.

A két- vagy háromdimenziós valós tér esetében a hosszúság és a szög segítségével definiált skaláris szorzat igen sok esetben hasznos segítséget nyújt feladatok megoldásához. Az itteni „eredendően absztrakt” tárgyalásmód alapján nincs lehetőségünk a hosszúság, illetve a szög közvetlen értelmezésére. Definiálhatjuk viszont közvetlenül a skaláris szorzatot, amely lehetőséget ad a hosszúság és a szög definiálására is. Az \mathbf{u} és a \mathbf{v} vektorok skaláris szorzatát a geometriai vizsgálatoknál úgy értelmezzük, hogy $|\mathbf{u}| \cdot |\mathbf{v}| \cdot \cos(\varphi)$, ahol $|\mathbf{u}|$, $|\mathbf{v}|$ a vektorok hossza és φ az általuk közbezárt szög. Ez a definíció esetünkben tehát szóba sem jöhet. Ismeretes azonban, hogy két vektor skaláris szorzata kifejezhető ezek koordinátaival. Egy vektor koordinátáinak a meghatározásához tehát egy bázisra van szükség. Ez azt jelenti, hogy a skaláris szorzat függ a bázistól.

6.1. Definíció (A változat). Legyen $U = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ a (valós) test feletti \mathcal{U} véges dimenziós vektortér egy bázisa. E tér $\mathbf{a} = \sum_i a_i \mathbf{u}_i$ és $\mathbf{b} = \sum_i b_i \mathbf{u}_i$ vektorainak e bázisban megadott skaláris vagy belső szorzatán az $(\mathbf{a}, \mathbf{b}) = \sum_i a_i b_i$ számot értjük.

Ha az \mathcal{U} téren egy skalárszorzat (azaz skaláris szorzat) van értelmezve, akkor azt mondjuk, hogy \mathcal{U} euklideszi tér erre a skalárszorzatra nézve. ■

(Megjegyezzük, hogy ha a valós számtest helyett más testet tekintünk, akkor a skaláris szorzat ebben a testben van. Bizonyos testeknél — például a komplex számtestnél — másképpen célszerű definiálni a skaláris szorzatot.)

Kiegészítés. Ha az \mathbf{u} és \mathbf{v} vektoroknak az U bázisban felírt mátrixai $[\mathbf{u}]$ és $[\mathbf{v}]$, akkor $(\mathbf{u}, \mathbf{v}) = [\mathbf{u}]^\dagger [\mathbf{v}]$. ■

6.1. Tétel. Az \mathcal{U} téren értelmezett bármely skalárszorzat pozitív definit bilineáris függvény; és minden pozitív definit bilineáris függvényhez létezik olyan bázis, amelyben ez a függvény skalárszorzattá válik.

Bizonyítás. Legyen $\mathbf{S}(\mathbf{u}, \mathbf{v}) = (\mathbf{u}; \mathbf{v})$ egy adott \mathbf{U} bázishoz tartozó skaláris szorzat. A szorzás kommutativitása alapján \mathbf{S} szimmetrikus. A szorzásnak az összeadásra vonatkozó disztributivitása alapján \mathbf{S} bilineáris. A $\mathbf{b} = \mathbf{a}$ esetben $\mathbf{S}(\mathbf{a}, \mathbf{a}) = \sum_i a_i^2$; ami nem lehet negatív, és 0 is csak akkor, ha $\mathbf{a} = \mathbf{o}$, ami bizonyítja, hogy a skalárszorzat valóban pozitív definit.

Legyen most \mathbf{A} tetszőleges pozitív definit bilineáris függvény az \mathcal{U} téren. Az ortogonalizációs tétel (5.14.) szerint létezik olyan $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázis, amelyben a megfelelő \mathbf{Q} kvadratikus alak négyzetösszeggé válik: $\mathbf{Q}(\mathbf{x}) = \sum_i c_i x_i^2$, ahol $\mathbf{x} = \sum_i x_i \mathbf{u}_i$; és a pozitív definitésg miatt mindegyik c_i pozitív. Válasszuk most a $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ bázist úgy, hogy $\mathbf{u}_i = \sqrt{c_i} \mathbf{v}_i$ legyen, ami a c_i -k pozitivitása alapján lehetséges. Ekkor ebben a bázisban $\mathbf{Q}(\mathbf{x}) = \mathbf{A}(\mathbf{x}, \mathbf{x}) = \sum_i y_i^2$, ahol $\mathbf{x} = \sum_i y_i \mathbf{v}_i$. Legyen \mathbf{S} az ezen bázishoz tartozó skalárszorzat. Az \mathbf{S} által meghatározott kvadratikus alak pontosan $\mathbf{Q}(\mathbf{x})$. Mivel a skalárszorzat is szimmetrikus bilineáris függvény és két szimmetrikus bilineáris függvényhez tartozó kvadratikus alak csak akkor egyenlő, ha az eredeti szimmetrikus bilineáris függvények is megegyeznek, ezért $\mathbf{A} = \mathbf{S}$. ■

A skaláris szorzatnak az \mathbf{A} változatban adott eredeti definíciója a 6.1. Tétel alapján megváltoztatható:

6.1. Definíció (B változat). A valós test feletti \mathcal{U} vektortéren adott bármely pozitív definit szimmetrikus bilineáris függvényt skaláris szorzatnak nevezzük. ■

Megjegyzések

1. A skaláris szorzat nem „igazi” művelet. Az eddigi műveletek mind olyanok voltak, amelyek vektorhoz vagy vektorokhoz rendelték vektort. A skaláris szorzat viszont nem vektort, hanem skalárt rendel a vektorokhoz. Ez inkább egy relációhoz hasonlít, csak itt nem az „igaz” vagy „hamis” értéket, hanem számértéket rendelünk hozzá egy párhoz.

2. A skaláris szorzat definíciójára adott \mathbf{B} változatnak többek között az az előnye, hogy végtelen dimenziós tereken is értelmezhető (l.: Schwarz-féle egyenlőtlenség). □

Euklideszi terek esetében is lehet homomorfizmusról, azaz homogén lineáris leképezésről beszélni, ez olyan leképezést jelent, amelyik a skaláris szorzatot is megtartja:

6.2. Definíció. Euklideszi terek esetében a $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ homomorfizmust euklideszi tér homomorfizmusnak nevezzük, ha skalárszorzattartó. Ha φ bijektív, akkor ez izomorfizmus. Ebben az esetben izomorf euklideszi terekről beszélünk. ■

6.2. Tétel. Euklideszi terek közti homomorfizmus mindig injektív. Két euklideszi tér pontosan akkor izomorf, ha mint vektorterek izomorfak.

Nem minden vektortér-izomorfizmus izomorfizmusa az euklideszi tereknek!

Bizonyítás. Legyen $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ euklideszi tér homomorfizmus. Ha $\mathbf{u} \in \text{Ker}(\varphi)$, akkor $\varphi(\mathbf{u}) = \mathbf{0} \in \mathcal{V}$, és a skalárszorzzattartás miatt $(\mathbf{u}; \mathbf{u}) = (\mathbf{0}; \mathbf{0}) = 0$, amiből $\mathbf{u} = \mathbf{0}$ következik, hiszen a skalárszorzat pozitív definit.

Ha φ izomorfizmus, akkor természetesen dimenziótartó is. Ha \mathcal{U} -nak és \mathcal{V} -nek megegyezik a dimenziója, akkor a skalárszorzatot definiáló bázisoknak ugyanannyi elemük van. Legyenek ezek — megfelelően — $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Ekkor a $\varphi : \mathbf{u}_i \mapsto \mathbf{v}_i$ megfeleltetéssel definiált homomorfizmus nyilván skalárszorzzattartó.

(Ha az \mathbf{u}_1 bázisvektort $2\mathbf{u}_1$ -be visszük és minden más bázisvektort önmagába, akkor ez nyilván izomorfizmus lesz, de nem tartja meg a skalárszorzatot.) ■

Egy bázis természetesen egyértelműen meghatároz egy skalárszorzatot. Fordítva azonban ez nem igaz; egy skalárszorzat, mint bilineáris függvény, több bázishoz is tartozhat. Ha adott egy tetszőleges pozitív definit szimmetrikus bilineáris függvény, akkor megadhatjuk az összes olyan bázist, amely ezt a bilineáris függvényt hozza létre skalárszorzatként:

6.3. Tétel. Egy $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázis pontosan akkor adja skalárszorzzatként az $\mathbf{S}(\mathbf{x}, \mathbf{y})$ szimmetrikus pozitív definit bilineáris függvényt, ha e függvényre nézve ortonormált; azaz $\mathbf{S}(\mathbf{u}_i, \mathbf{u}_j) = \delta_{i,j}$.

Bizonyítás. Az \mathbf{U} bázisához tartozó skalárszorzatra $(\mathbf{u}_i; \mathbf{u}_j) = \delta_{i,j}$ teljesül. Mivel egy bilineáris függvényt az 5.6. Tétel szerint egyértelműen meghatároz egy bázispáron felvett értékeinek rendszere, ezért ez a skalárszorzat ugyanaz, mint az adott \mathbf{S} bilineáris függvény.

Az, hogy a skalárszorzatot definiáló bázis e skalárszorzatra nézve ortonormált, a definíció alapján triviális. ■

6.3. Definíció. Az \mathcal{U} euklideszi tér egy \mathcal{V} altere euklideszi altér, ha az \mathcal{U} -n definiált skalárszorzatot \mathcal{V} -re megszorítva, \mathcal{V} alkalmas bázisában skalárszorzzattá válik. ■

6.4. Tétel. Egy euklideszi tér minden altere euklideszi altér.

Bizonyítás. Legyen \mathbf{S} az eredeti téren definiált skaláris szorzat. Ez, mint bilineáris függvény, az altéren is pozitív definit és szimmetrikus. A 6.3. Tétel szerint tehát ezen altér alkalmas bázisában skalárszorzzattá válik. ■

Feladatok

1. Álljon az \mathbb{R} feletti \mathcal{U}_0 vektortér az $\mathbf{a} = (a_0, a_1, \dots, a_n, \dots)$ végtelen sorozatokból. Legyen \mathcal{U}_1 az az altér, amelyben a sorozatoknak majdnem minden eleme 0. Tekintsük azokat az \mathbf{a} sorozatokat, amelyekre a $\sum_{i=0}^{\infty} (a_i)^2$ végtelen sor konvergens. Bizonyítsuk be, hogy ezek egy \mathcal{U} alteret alkotnak,

amelyre $\mathcal{U}_1 \leq \mathcal{U} \leq \mathcal{U}_0$. Bizonyítsuk be, hogy $\mathbf{a}, \mathbf{b} \in \mathcal{U}$ esetén a $\sum_{i=0}^{\infty} a_i b_i$ végtelen sor konvergens és

az $A(\mathbf{a}, \mathbf{b}) = \sum_{i=0}^{\infty} a_i b_i$ függvény pozitív definit — tehát skalárszorzatnak tekinthető. Bizonyítsuk be, hogy \mathcal{U} -nak van olyan \mathcal{V} altere, amelyre $\mathcal{V} + \mathcal{V}^{\perp} \neq \mathcal{U}$.

2. Bizonyítsuk be, hogy az előző feladat \mathcal{U} vektorterének nincs megszámlálható bázisa. Igaz-e, hogy ha egy \mathcal{U} vektortérnek van megszámlálható bázisa, akkor minden $\mathcal{V} \leq \mathcal{U}$ altérre teljesül: $\mathcal{V} + \mathcal{V}^{\perp} = \mathcal{U}$?

3. Melyek az \mathbb{R} feletti \mathcal{U} altéren értelmezett azon szimmetrikus bilineáris függvények, amelyekre minden ortogonális rendszer független?

4. Tekintsük a $[0, 1]$ intervallumon folytonos függvényeket. Bizonyítsuk be, hogy ezek az \mathbb{R} felett egy \mathcal{C} vektorteret alkotnak. Bizonyítsuk be, hogy e téren az $\int_0^1 fg$ integrál ($f, g \in \mathcal{C}$) pozitív definit szimmetrikus bilineáris függvény.

5. Az \mathcal{U} euklideszi tér tetszőleges \mathbf{u} eleméhez tekintsük a $\varphi_{\mathbf{u}} : \mathbf{v} \mapsto (\mathbf{u}; \mathbf{v})$ lineáris leképezést. Bizonyítsuk be, hogy az $\mathbf{u} \mapsto \varphi_{\mathbf{u}}$ leképezés \mathcal{U} -t a duális térbe vivő izomorfizmus. Mutassuk meg, hogy az $\mathbf{u}^* = \varphi_{\mathbf{u}}$ esetben a skalárszorzatot értelmező bázis képe pontosan a duális bázis. Mutassuk meg, hogy ennél a megfeleltetésnél pontosan az ortonormált bázisoknak felel meg a duális bázisuk.

6. Bizonyítsuk be, hogy véges testek feletti legalább háromdimenziós vektorterekben egy kvadratikus alak sohasem lehet „pozitív definit”, azaz mindig van önmagára merőleges vektor.

7. Bizonyítsuk be, hogy $p = 4k + 3$ alakú prímszám esetén a p elemű test feletti kétdimenziós térben az $x^2 + y^2$ kvadratikus alak csak az $x = y = 0$ esetben nulla.

2. A valós euklideszi terek geometriája

Annak a ténynek, hogy az egyező dimenziójú euklideszi terek izomorfak, két fontos következménye is van. Mindkettő a geometriai tér fontosságát mutatja.

Tegyük fel, hogy az \mathbb{R} feletti két- vagy háromdimenziós térben kell valamit bizonyítani. Ez a tér a 6.2. Tétel szerint izomorf a geometriai síkkal, illetve térral. Így módunkban van bármely állítást szemléletesen ellenőrizni. Még azt is megtehetjük, hogy egy geometriai eredményt az izomorfizmussal „átvizsgálunk” a szóban forgó euklideszi térben.

A másik lehetőséget a 6.4. Tétel adja. Tegyük fel, hogy egy euklideszi térben valamilyen állítást a dimenzió szerinti teljes indukcióval akarunk bizonyítani. Mivel a tér bármely háromdimenziós altere a geometriai háromdimenziós altérrel izomorf, ezért a bizonyítást bármelyik háromdimenziós altéren „elkezdhetjük”. Sőt, mi több, az esetleges sejtésünk helyességét — az előbbiek alapján — a geometriai térben valószínűsíthetjük.

Ezek után megtehetnénk, hogy a legfeljebb háromdimenziós geometriai térre vonatkozó eredményeket itt is igaznak fogadjuk el. Ezt azonban két ok miatt nem tesszük:

Egyrészt, a geometriai tér geometriájából nagyon kevés fogalomra lesz szükségünk, amiket itt geometriai szemlélet nélkül is definiálhatunk. Ezáltal egyébként a geometriai tér fogalmait is sokkal pontosabban tudjuk definiálni.

Másrészt, az itteni bizonyítások „absztrakt vázát” is jobban lehet látni; ami azt jelenti, hogy könnyebben átvihetők más, olyan esetekre is, ahol szemléletünk nem segít.

6.4. Definíció. Az $(\mathbf{u}; \mathbf{v})$ skalárszorzzattal megadott \mathcal{U} euklideszi tér egy \mathbf{u} vektorának hosszán az $|\mathbf{u}| = \sqrt{(\mathbf{u}; \mathbf{u})}$ nemnegatív számot értjük. A \mathbf{o} -tól különböző \mathbf{u} és \mathbf{v} vektorok φ hajlásszöge az a nemnegatív, 180° -nál nem nagyobb szög legyen, amelynek koszinuszára $|\mathbf{u}| \cdot |\mathbf{v}| \cdot \cos(\varphi) = (\mathbf{u}; \mathbf{v})$ teljesül.

A \mathbf{o} vektornak és tetszőleges \mathbf{u} vektornak nem definiálunk hajlásszöget. ■

A fenti definíció nyilvánvalóan tükrözi a geometriai fogalmakat. Az előzetes „elvéink”-nek megfelelően mégis be kell látni, hogy a definíció helyes — azaz a definiált fogalmak „léteznek”. Az abszolút érték létezik, mert a skalárszorzat nem negatív, tehát a négyzetgyökvonás elvégezhető, és definíció szerint a négyzetgyökök sem negatív.

A hajlásszög definíciójánál két problémával is szembe kell nézni. A „második” probléma az egyszerűbb; nevezetesen az, hogy adott koszinuszú szög a tekintett intervallumban pontosan egy van. Ezt akár geometriai, akár függvénytan úton beláthatjuk; és el is fogadjuk.

A másik probléma az, hogy létezik-e ilyen szög. Itt azt kell elfogadnunk, hogy minden számhoz, amelynek abszolút értéke legfeljebb 1, található olyan szög, amelynek pontosan ennyi a koszinusza. Ezzel arra az algebrai kérdésre redukáltuk a problémát, hogy igaz-e az $|\mathbf{u}| \cdot |\mathbf{v}| \geq |(\mathbf{u}; \mathbf{v})|$ összefüggés. Az abszolút értékek nemnegativitása miatt ez ekvivalens az alábbi, úgynevezett *Cauchy–Bunyakovszkij*-egyenlőtlenséggel:

$$(\mathbf{u}; \mathbf{u}) \cdot (\mathbf{v}; \mathbf{v}) \geq (\mathbf{u}; \mathbf{v})^2.$$

Ezt az egyenlőtlenséget az alábbiakban valamivel általánosabban bizonyítjuk, megmutatva, hogy mely bilineáris függvény esetében teljesül az analóg egyenlőtlenség.

6.5. Tétel. Egy valós \mathcal{U} téren értelmezett szimmetrikus bilineáris \mathbf{A} függvényre akkor és csak akkor igaz minden $\mathbf{a}, \mathbf{b} \in \mathcal{U}$ esetén az

$$\mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(\mathbf{b}, \mathbf{b}) \geq \mathbf{A}(\mathbf{a}, \mathbf{b})^2$$

egyenlőtlenség, ha \mathbf{A} szemidefinit.

Bizonyítás. Ha \mathbf{A} indefinit, akkor bármely \mathbf{A} -ortogonális bázisban léteznek olyan \mathbf{A} -ortogonális \mathbf{u} és \mathbf{v} vektorok, amelyekre $\mathbf{A}(\mathbf{u}, \mathbf{u})$ és $\mathbf{A}(\mathbf{v}, \mathbf{v})$ különböző előjelűek. Ezek szorzata negatív, míg az \mathbf{A} -ortogonalitás következtében $\mathbf{A}(\mathbf{u}, \mathbf{v}) = 0$; így a tételbeli egyenlőtlenség nem teljesül.

Legyen most \mathbf{A} szemidefinit. Mindenekelőtt belátjuk, hogy elég az állítást pozitív szemidefinit bilineáris függvényekre bizonyítani. Ha ugyanis \mathbf{A} negatív szemidefinit, akkor $-\mathbf{A}$ pozitív szemidefinit. Ha erre igaz az állítás, akkor

$$\mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(\mathbf{b}, \mathbf{b}) = (-\mathbf{A}(\mathbf{a}, \mathbf{a})) \cdot (-\mathbf{A}(\mathbf{b}, \mathbf{b})) \geq (-\mathbf{A}(\mathbf{a}, \mathbf{b}))^2 = (\mathbf{A}(\mathbf{a}, \mathbf{b}))^2$$

alapján az eredeti bilineáris függvényre is igaz.

Legyen most \mathbf{A} pozitív szemidefinit szimmetrikus bilineáris függvény. Ha $\mathbf{A}(\mathbf{a}, \mathbf{a}) = \mathbf{A}(\mathbf{b}, \mathbf{b}) = 0$, akkor a bilinearitást felhasználva a pozitív szemidefinitésgből tetszőleges c valós számra

$$0 \leq \mathbf{A}(\mathbf{a} + c\mathbf{b}, \mathbf{a} + c\mathbf{b}) = 2 \cdot c \cdot \mathbf{A}(\mathbf{a}, \mathbf{b})$$

következik. A $c = -\mathbf{A}(\mathbf{a}, \mathbf{b})$ választással ebből $\mathbf{A}(\mathbf{a}, \mathbf{b})^2 \leq 0$ adódik, ami csak úgy lehetséges, hogy $\mathbf{A}(\mathbf{a}, \mathbf{b}) = 0$; és ekkor az egyenlőtlenség valóban fennáll.

Feltehető tehát, hogy például $\mathbf{A}(\mathbf{a}, \mathbf{a}) > 0$. Mivel \mathbf{A} pozitív szemidefinit szimmetrikus bilineáris függvény, ezért tetszőleges c valós számra

$$0 \leq \mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(c\mathbf{a} + \mathbf{b}, c\mathbf{a} + \mathbf{b}) = c^2(\mathbf{A}(\mathbf{a}, \mathbf{a}))^2 + 2c \cdot \mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(\mathbf{a}, \mathbf{b}) + \mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(\mathbf{b}, \mathbf{b}) = (c \cdot \mathbf{A}(\mathbf{a}, \mathbf{a}) + \mathbf{A}(\mathbf{a}, \mathbf{b}))^2 + \mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(\mathbf{b}, \mathbf{b}) - (\mathbf{A}(\mathbf{a}, \mathbf{b}))^2$$

teljesül. A feltétel szerint $\mathbf{A}(\mathbf{a}, \mathbf{a}) \neq 0$, ezért választható $c = -\frac{\mathbf{A}(\mathbf{a}, \mathbf{b})}{\mathbf{A}(\mathbf{a}, \mathbf{a})}$. Ekkor a jobb oldalon álló első tag 0 lesz, és így

$$0 \leq \mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(\mathbf{b}, \mathbf{b}) - (\mathbf{A}(\mathbf{a}, \mathbf{b}))^2$$

adódik, mint állítottuk. ■

Kiegészítés (Schwarz-féle egyenlőtlenség). *Ha f és g az $[a, b]$ intervallumban integrálható függvények, akkor*

$$\left(\int_a^b fg \right)^2 \leq \int_a^b f^2 \cdot \int_a^b g^2.$$

Bizonyítás. Ismeretes, hogy az adott intervallumban négyzetesen integrálható függvények a szokásos műveletekre vektorteret alkotnak \mathbb{R} felett. Az ebben az intervallumban vett szorzatintegrál nyilvánvalóan szimmetrikus és bilineáris. Tekintettel arra, hogy ez nem lehet negatív (de 0 igen!), ezért ez a függvény pozitív szemidefinit. A 6.5. Tétel szerint tehát igaz az egyenlőtlenség. (Egyébként az, hogy két négyzetesen integrálható függvény összege is négyzetesen integrálható függvény, éppen a Schwarz-féle egyenlőtlenségből következik.) ■

Következmény. *Két vektor hajlásszöge pontosan akkor 90° , ha skalárszorzatuk 0, azaz a skalárszorzatra mint bilineáris függvényre nézve ortogonálisak.*

Ilyen vektorokat ortogonálisnak vagy egymásra merőlegesnek hívunk. Azt, hogy \mathbf{u} és \mathbf{v} merőlegesek, $\mathbf{u} \perp \mathbf{v}$ jelöli.

A nullvektor minden vektorra merőleges.

Bizonyítás. A hajlásszög pontosan akkor derékszög, ha koszinusa 0, azaz a szöget definiáló egyenlőség bal oldalán 0 áll. Ez pedig pontosan akkor teljesül, ha a jobb oldalon is 0 van, azaz a skalárszorzat 0. Az utolsó állítás a fenti kiegészítésből adódik. ■

6.6. Tétel (háromszög-egyenlőtlenség). *Az euklideszi tér tetszőleges \mathbf{a} és \mathbf{b} vektorára*

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|.$$

Bizonyítás. Mivel mindkét oldalon nemnegatív számok állnak, ezért elég bebizonyítani, hogy az egyenlőtlenség a négyzeteikre teljesül. Az abszolút érték definíciója szerint tehát a bizonyítandó állítás:

$$(\mathbf{a} + \mathbf{b}; \mathbf{a} + \mathbf{b}) \leq (\mathbf{a}; \mathbf{a}) + 2 \cdot |\mathbf{a}| \cdot |\mathbf{b}| + (\mathbf{b}; \mathbf{b}).$$

A bilineáris függvényekre vonatkozó azonosságok szerint ez az egyenlőtlenség az $(\mathbf{a}; \mathbf{b}) \leq |\mathbf{a}| \cdot |\mathbf{b}|$ egyenlőtlenséggel ekvivalens. Ez utóbbi pedig triviálisan következik a Cauchy–Bunyakovszkij-féle egyenlőtlenségből. ■

Megjegyzés. Ha az \mathbf{a} és \mathbf{b} vektorokat „egymáshoz fűzzük”, akkor egy olyan háromszöget kapunk, amelynek a harmadik oldala éppen az $\mathbf{a} + \mathbf{b}$ vektor. A 6.6. Tétel tehát pontosan azt mondja ki, hogy egy háromszög két oldala hosszának összege mindig nagyobb a háromszög harmadik oldala hosszánál. (Ha a két vektor párhuzamos, akkor nem kapunk háromszöget. Csak ekkor állhat egyenlőség; s ha a két vektor iránya is megegyezik, akkor valóban egyenlőséget kapunk.) □

A szögekre vonatkozó összefüggések közül csak olyanokkal foglalkozunk, amelyekben a merőlegesség szerepel. A merőlegesség a geometriában is alapvetőbb fogalom a szögnél. Ezt a lineáris algebrai vizsgálatoknál is láttuk; az ortogonalitást sokkal előbb értelmeztük, mint a szöget. Azt is láttuk, hogy az ortogonalitás definíciójához nem volt szükség olyan „bonyolult” fogalomra, mint a koszinuszfüggvény viselkedése. A továbbiakban érdemes a definíciókat és a tételeket a geometriai megfelelő fogalmakkal összevetni.

6.5. Definíció. Egy \mathbf{u} vektort egy \mathcal{V} altérre merőlegesnek (ortogonálisnak) nevezünk, ha minden, az altérbe eső vektorra merőleges. Ezt a relációt $\mathbf{u} \perp \mathcal{V}$ jelöli. ■

6.7. Tétel. Egy vektor akkor és csak akkor merőleges egy altérre, ha az altér valamely bázisának elemeire merőleges.

Bizonyítás. Lásd az **A**-ortogonalitásra vonatkozó megfelelő 5.15. Tétel bizonyítását. ■

6.6. Definíció. Az \mathcal{U} vektortér \mathcal{V} altere merőleges a \mathcal{W} altérre, ha a \mathcal{V} altér minden eleme merőleges a \mathcal{W} altérre. Ezt a relációt $\mathcal{V} \perp \mathcal{W}$ jelöli. ■

6.8. Tétel. A \mathcal{V} altér akkor és csak akkor merőleges a \mathcal{W} altérre, ha a \mathcal{V} altér valamely bázisának elemei merőlegesek a \mathcal{W} altérre. Az alterek merőlegessége szimmetrikus fogalom.

Bizonyítás. Lásd az **A**-ortogonalitásra vonatkozó megfelelő 5.15. Tétel bizonyítását. ■

Legyen \mathcal{V} az \mathcal{U} euklideszi tér altere. Az 5.15. Tétel miatt $\mathcal{U} = \mathcal{V} \oplus \mathcal{V}^\perp$. Ez azt jelenti, hogy minden $\mathbf{u} \in \mathcal{U}$ vektor egyértelműen felírható $\mathbf{u} = \mathbf{v} + \mathbf{w}$ alakban, ahol $\mathbf{v} \in \mathcal{V}$ és $\mathbf{w} \perp \mathcal{V}$. Ezt az euklideszi tereknél fontos tételt itt ismételtelen bizonyítjuk; megadva az előállítást is, ha adott \mathcal{V} egy bázisa. A következő tétel ennél többet is mond ki.

6.9. Tétel. *Ha \mathcal{V} altere az \mathcal{U} euklideszi térnek, akkor minden $\mathbf{u} \in \mathcal{U}$ egyértelműen előállítható $\mathbf{u} = \mathbf{v} + \mathbf{w}$ alakban, ahol $\mathbf{v} \in \mathcal{V}$ és $\mathbf{w} \perp \mathcal{V}$. Bármely $\mathbf{u} = \mathbf{v}' + \mathbf{w}'$ előállításnál a $\mathbf{v}' \in \mathcal{V}$ feltételből következik, hogy $|\mathbf{w}| \leq |\mathbf{w}'|$ és egyenlőség csak a $\mathbf{w}' = \mathbf{w}$ esetben lehet. \mathbf{v} az \mathbf{u} vektor \mathcal{V} -re való merőleges (ortogonális) vetülete.*

Bizonyítás. Legyen $\mathbf{V} = \{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ a \mathcal{V} egy ortonormált bázisa. Ha az \mathbf{u} vektorhoz van megfelelő \mathbf{v} vektor, akkor ez egyrészt $\sum_i c_i \mathbf{e}_i$ alakú, másrészt a $\mathbf{w} = \mathbf{u} - \mathbf{v}$ egyenlőséggel definiált vektorra $\mathbf{w} \perp \mathcal{V}$ teljesül. Ez pontosan akkor áll fenn, ha minden i indexre $\mathbf{w} \perp \mathbf{e}_i$, azaz $(\mathbf{w}; \mathbf{e}_i) = 0$ teljesül. A $\mathbf{w} = \mathbf{u} - \mathbf{v}$ egyenlőség és a skalárszorzat bilinearitása miatt ez azt jelenti, hogy minden i és j indexre igaz az $(\mathbf{u}; \mathbf{e}_i) = \sum_j c_j (\mathbf{e}_j; \mathbf{e}_i)$ összefüggés. Tekintettel arra, hogy a felvett bázis ortonormált volt, ezért $(\mathbf{e}_j; \mathbf{e}_i) = \delta_{i,j}$. Ezt behelyettesítve azt kapjuk, hogy $c_i = (\mathbf{u}; \mathbf{e}_i)$, azaz $\mathbf{v} = \sum_i (\mathbf{u}; \mathbf{e}_i) \mathbf{e}_i$. Ebből a felírt egyenlőséget felhasználva \mathbf{w} is megkapható.

Tekintsünk most egy $\mathbf{u} = \mathbf{v}' + \mathbf{w}'$ előállítást, ahol $\mathbf{v}' \in \mathcal{V}$. Ebből $\mathbf{w}' = \mathbf{u} - \mathbf{v}' = (\mathbf{v} + \mathbf{w}) - \mathbf{v}' = \mathbf{w} + (\mathbf{v} - \mathbf{v}')$ következik. Mivel $\mathbf{w} \perp (\mathbf{v} - \mathbf{v}')$, ezért skalárszorzatuk 0; így $(\mathbf{w}'; \mathbf{w}') = (\mathbf{w}; \mathbf{w}) + ((\mathbf{v} - \mathbf{v}'); (\mathbf{v} - \mathbf{v}'))$, azaz $|\mathbf{w}'|^2 = |\mathbf{w}|^2 + |\mathbf{v} - \mathbf{v}'|^2$. Ez azt jelenti, hogy $|\mathbf{w}'|^2 \geq |\mathbf{w}|^2$ és egyenlőség csak a $\mathbf{v} = \mathbf{v}'$ esetben lehet. ■

Megjegyezzük, hogy a \mathbf{v} és \mathbf{w} vektorok akkor is előállíthatók, ha kiindulásul nem ortonormált, hanem tetszőleges bázist veszünk, de ez esetben a leírás sokkal bonyolultabb. Itt is a bázisvektorpárok skaláris szorzata játszik szerepet. Ennek leírásához hasznos az alábbi mátrix, illetve ennek determinánsa:

6.7. Definíció. Az $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ vektorrendszerből elkészített $[(\mathbf{a}_i; \mathbf{a}_j)]$ mátrixot e rendszer Gram-féle mátrixának, és ennek $|(\mathbf{a}_i; \mathbf{a}_j)|$ determinánsát e rendszer Gram-féle determinánsának nevezik. ■

Természetesen az a vektortér, amelyikből a fenti vektorokat vettük, nem kell, hogy r -dimenziós legyen.

6.10. Tétel. *Egy vektorrendszer Gram-féle mátrixa akkor és csak akkor szinguláris (tehát Gram-féle determinánsa akkor és csak akkor 0), ha a vektorrendszer lineárisan összefüggő.*

Bizonyítás. A mátrixok tárgyalásánál láttuk, hogy egy négyzetes mátrix determinánsa pontosan akkor 0, ha oszlopai (vagy sorai) lineárisan összefüggenek. Egy r oszlopú M mátrix oszlopai akkor lineárisan összefüggők, ha a megfelelő oszlopokat alkalmas c_1, \dots, c_r skalárokkal szorozva és összeadva a csupa 0-ból álló oszlopot kapjuk, ahol nem mindegyik $c_i = 0$. Ezt úgy fogalmazhatjuk, hogy arra a \mathbf{v} vektorra, amelynek mátrixa $[\mathbf{v}]^\dagger = [c_1, \dots, c_r]$, az $M\mathbf{v} = \mathbf{o}$ összefüggés teljesül.

Tekintsük most azt az A mátrixot, amelynek j -edik oszlopában az \mathbf{a}_j koordinátái állnak ($1 \leq j \leq r$). A mátrixok szorzásának a definíciója szerint a szereplő Gram-féle

mátrix pontosan $A^\dagger A$. Az előrebocsátottak szerint azt kell tehát bizonyítani, hogy pontosan akkor létezik olyan $\mathbf{u} \neq \mathbf{o}$ vektor, amelyre $A\mathbf{u} = \mathbf{o}$, ha létezik olyan $\mathbf{v} \neq \mathbf{o}$ vektor, amelyre $A^\dagger A\mathbf{v} = \mathbf{o}$. Azt fogjuk kimutatni, hogy $\mathbf{v} = \mathbf{u}$ is igaz.

Ha $\mathbf{u} \neq \mathbf{o}$ és $A\mathbf{u} = \mathbf{o}$, akkor természetesen $A^\dagger A\mathbf{u} = \mathbf{o}$ is igaz. Fordítva, az $A^\dagger A\mathbf{v} = \mathbf{o}$ feltételből (ahol $\mathbf{v} \neq \mathbf{o}$) azt kapjuk, hogy $\mathbf{v}^\dagger A^\dagger A\mathbf{v} = \mathbf{v}^\dagger \mathbf{o} = 0$. A szorzat transzponáltjára vonatkozó azonosság szerint a bal oldalon $(A\mathbf{v})^\dagger A\mathbf{v}$ áll, ami nem más, mint $A\mathbf{v}$ -nek önmagával való skalárszorzata. A skalárszorzat pozitív definitisége miatt ebből $A\mathbf{v} = \mathbf{o}$ következik. ■

Mint láttuk, két euklideszi tér pontosan akkor izomorf egymással, ha mint vektorterek izomorfak. Az eltérés annyi, hogy sokkal kevesebb izomorfizmus létezik euklideszi terek esetében. Ezt akkor láthatjuk legvilágosabban, ha egy tér önmagával való izomorfizmusait nézzük. Tekintsük a síkot mint euklideszi teret, és legyenek \mathbf{e} tér egy ortonormált bázisának elemei az $\mathbf{e} = [1, 0]$ és $\mathbf{f} = [0, 1]$ vektorok. Ha a síkon felvesszünk két tetszőleges független vektort, akkor van olyan vektortér-transzformáció, amely az eredetit ezekbe viszi. Ha viszont azt is megkívánjuk, hogy a transzformáció megtartsa a skalárszorzatot, akkor már az \mathbf{e} vektor képe is csak egységnyi hosszúságú lehet; az \mathbf{f} vektor képére pedig csak két lehetőségünk van már. Ez mutatja, hogy valóban sokkal kevesebb transzformáció tartja meg a skalárszorzatot.

Feladatok

1. Írjuk fel az euklideszi tér két vektorát egy adott bázisban. Írjuk fel a Cauchy–Bunyakovszkij-egyenlőtlenséget a koordináták segítségével.
2. Mi a feltétele annak, hogy a Cauchy–Bunyakovszkij-egyenlőtlenségben egyenlőség álljon? Mi következik ebből a háromszög-egyenlőtlenségre?
3. Általánosítsuk az alterek merőlegességét úgy, hogy ez két, nem a nullvektorban metsző sík esetében a geometriai merőlegességet adja.
4. Bizonyítsuk be, hogy két lineáris alakzat elemeiből képezett különbségvektorok hosszának mindig létezik minimuma. Ennek alapján értelmezzük lineáris alakzatok távolságát. Érvényes-e itt is a háromszög-egyenlőtlenség?
5. Értelmezzük olyan alterek szögét, amelyek csak a nullvektorban metszik egymást.
6. A tér dimenziójára vonatkozó teljes indukcióval értelmezzük valódi módon metsző — de egymást nem tartalmazó — alterek szögét is, felhasználva a háromdimenziós geometriai térben két sík hajlásszögének a definícióját.
7. Bizonyítsuk be, hogy a Gram-féle determináns mindig nemnegatív.
8. Az $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ és $\{\mathbf{b}_1, \dots, \mathbf{b}_r\}$ vektorrendszerekre tekintsük a $|\langle \mathbf{a}_i, \mathbf{b}_j \rangle|$ determinánst. Bizonyítsuk be, hogy ez akkor és csak akkor 0, ha a két vektorrendszer valamelyike lineárisan összefüggő. Igaz-e, hogy ez sem lehet negatív?
9. Bizonyítsuk be, hogy tetszőleges A mátrix esetén igaz az $r(A^\dagger A) = r(A)$ egyenlőség.

3. A komplex euklideszi tér

A komplex számtest feletti vektorterekhez hasonlóan beszélhetünk a komplex számtest feletti euklideszi terekről. Erre analízisben, geometriában és algebrai geometriában szükség is van. A komplex számtest esetében is értelmezhető egy bázishoz tartozó skaláris szorzat. Ebben az esetben is gondot okoz a megfelelő kvadratikus alak „pozitivitása”. Itt is hasonlóképpen kell eljárni, mint a bilineáris függvények esetében általában.

6.8. Definíció. Legyen $U = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ a komplex test feletti \mathcal{U} véges dimenziós vektortér egy bázisa. E tér $\mathbf{a} = \sum_i a_i \mathbf{u}_i$ és $\mathbf{b} = \sum_i b_i \mathbf{u}_i$ vektorainak e bázisban megadott skaláris vagy belső szorzatán az $(\mathbf{a}; \mathbf{b}) = \sum_i \overline{a_i} b_i$ számot értjük.

Ha az \mathcal{U} téren egy skalárszorzat van értelmezve, akkor azt mondjuk, hogy \mathcal{U} euklideszi tér erre a skalárszorzatra nézve. ■

Kiegészítés. Ha az \mathbf{u} és \mathbf{v} vektoroknak az U bázisban felírt mátrixai $[\mathbf{u}]$ és $[\mathbf{v}]$, akkor $(\mathbf{u}; \mathbf{v}) = [\mathbf{u}]^* [\mathbf{v}]$. ■

6.11. Tétel. A skalárszorzat pozitív definit Hermite-féle bilineáris függvény. Minden pozitív definit Hermite-féle bilineáris függvényhez található olyan bázis, amelyben skalárszorzattá válik.

Ennek alapján a skalárszorzat itt is definiálható mint pozitív definit Hermite-féle bilineáris függvény.

Bizonyítás. A 6.1. Tétel bizonyítása a megfelelő fogalmak kicserélésével szóról szóra átvihető. ■

Megjegyzés. A valós euklideszi terek altereire vonatkozó eredmények bizonyításai is szó szerint átvihetők a komplex számtest feletti euklideszi terekre. □

6.12. Tétel. A Cauchy–Bunyakovszkij-egyenlőtlenség és a háromszög-egyenlőtlenség igaz a komplex euklideszi térben is.

Bizonyítás. A 6.5. Tétel bizonyítása kisebb változtatásokkal átvihető a komplex esetre is. Tekintettel arra, hogy az Hermite-féle bilineáris függvény nem szimmetrikus, ezért az egyenlőtlenséget az

$$\mathbf{A}(\mathbf{a}, \mathbf{a}) \cdot \mathbf{A}(\mathbf{b}, \mathbf{b}) \geq \mathbf{A}(\mathbf{a}, \mathbf{b}) \cdot \mathbf{A}(\mathbf{b}, \mathbf{a})$$

alakban kell megadni. Annak a bizonyítása, hogy ez csak szemidefinit függvényekre teljesülhet, pontosan úgy történhet, mint a valós esetben. Ugyancsak a valós esethez hasonlóan látható be az is, hogy elég pozitív szemidefinit függvényekkel foglalkozni; és ekkor az egyenlőtlenség teljesül, ha $\mathbf{A}(\mathbf{a}, \mathbf{a}) = \mathbf{A}(\mathbf{b}, \mathbf{b}) = 0$. A fennmaradó eset bizonyítása is hasonló, csak éppen nem használhatjuk a szimmetriát. Ennek megfelelően nem teljes négyzetre való kiegészítést kell alkalmazni, hanem ehelyett $(c \cdot \mathbf{A}(\mathbf{a}, \mathbf{a}) + \mathbf{A}(\mathbf{a}, \mathbf{b}))$ -nek és konjugáltjának a szorzata szerepel.

Ebből már adódik a háromszög-egyenlőtlenség is. ■

HETEDIK FEJEZET

AZ EUKLIDESZI TÉR LINEÁRIS TRANSZFORMÁCIÓI

1. Lineáris transzformációk polinomja

A lineáris transzformációk „geometriai viselkedésének” leírásánál fontos szerepet tölt be egy-egy transzformáció ismételt alkalmazása; illetve mindazok a transzformációk, amelyek e transzformációból a transzformációkra értelmezett műveletek segítségével előállíthatók.

7.1. Tétel. Legyen \mathcal{U} (véges dimenziós) vektortér a K test felett és $\alpha \in \text{End}(\mathcal{U})$ (az \mathcal{U} egy lineáris transzformációja). Ekkor létezik pontosan egy olyan $\Phi_\alpha : K[x] \rightarrow \text{End}(\mathcal{U})$, amelynél $\Phi_\alpha(c) = c\iota$ ($c \in K$ és ι az $\text{End}(\mathcal{U})$ identitása) és $\Phi_\alpha(x) = \alpha$.

A képként fellépő transzformációkat α polinomjainak nevezzük. Az $f(x) = c_0 + c_1x + \dots + c_r x^r$ polinom képét $f(\alpha) = c_0 + c_1\alpha + \dots + c_r\alpha^r$ jelöli. Ha $f(\alpha) = \omega$, akkor azt mondjuk, hogy α gyöke az $f(x)$ polinomnak. Azok a polinomok, amelyeknek α gyökük, ideált alkotnak. Ez az ideál nem 0. Ennek az ideálnak az $m(x) = m_\alpha(x)$ normált generátorelemét az α minimálpolinomjának nevezzük.

A $\dim(\mathcal{U}) = n$ esetben minden minimálpolinom foka legfeljebb n^2 .

Bizonyítás. A vektorterek elemi tulajdonságai alapján a $c \mapsto c\iota$ megfeleltetés izomorfizmus. A polinomgyűrűk alapvető tulajdonsága szerint ez kiterjeszthető az $x \mapsto \alpha$ feltétel mellett; feltéve, hogy az ezek által generált gyűrű kommutatív. Ez viszont így van, hiszen az identitás skalárszorosai és egy rögzített transzformáció hatványai egymással felcserélhetőek (a szorzásra nézve).

A továbbiakhoz elég azt belátni, hogy minden lineáris transzformáció gyöke egy nemnulla polinomnak, amelynek a foka legfeljebb n^2 . Evégett tekintsük az $\iota, \alpha, \dots, \alpha^{n^2}$ transzformációkat. Ezeknek a száma $n^2 + 1$. Tekintettel arra, hogy $\dim(\text{End}(\mathcal{U})) = n^2$, ezért e transzformációk lineárisan összefüggenek: $c_0\iota + c_1\alpha + \dots + c_{n^2}\alpha^{n^2}$, nem csupa 0 együtthatókkal. Ez azt jelenti, hogy az $f(x) = c_0 + c_1x + \dots + c_{n^2}x^{n^2}$ nemnulla polinomnak α gyöke. ■

A 7.1. Tétel egy igen fontos példát ad modulusra.

7.2. Tétel. *A 7.1. Tételben megadott Φ_α homomorfizmus által \mathcal{U} R -modulussá válik, ahol $R = K[x]$. Rögzített, de tetszőleges α esetén ${}_R\mathcal{U}$ végesen generált és torziómodulus. Ez utóbbi azt jelenti, hogy minden $\mathbf{u} \in \mathcal{U}$ elemhez található olyan $r \in R$ ($r \neq 0$), amelyre $r\mathbf{u} = \mathbf{o}$.*

Bizonyítás. Az első állítás azonnal következik a lineáris transzformációkra és a vektorokra vonatkozó azonosságokból. A kapott modulusnak van véges generátorrendszere, hiszen az \mathcal{U} tetszőleges bázisának lineáris kombinációjaként akkor is megkaphatunk minden elemet, ha csak a K elemeivel szorzunk. Mivel $m_\alpha(\mathbf{u}) = \mathbf{o}$ és $m_\alpha(x) \neq 0$, ezért minden vektorhoz univerzálisan választhatjuk az $r = m_\alpha(x)$ elemet. ■

Tekintettel arra, hogy a 7.2. Tételben az r univerzálisan választható, elég volna a továbbiakban csak ilyen tulajdonságú modulusokat vizsgálni. Azért teszünk fel a továbbiakban látszólag kevesebbet, mert az alábbi általános tételben is következik az „univerzális szorzó” léte. Jelenleg csak a fenti speciális R -modulusra vonatkozó eredményekre van szükségünk, de a bizonyítás az alábbi általános esetben is pontosan ugyanígy történik. Ezzel egyúttal arra is rávilágíthatunk, hogy „miért igaz” a tétel.

7.3. Tétel (főideálgyűrűk feletti végesen generált torziómodulusok alaptétele I. rész). *Legyen \mathcal{M} végesen generált torziómodulus az R főideálgyűrű felett. Ekkor létezik olyan $r \neq 0$ elem az R -ben, amelyre $r\mathbf{u} = \mathbf{o}$, tetszőleges $\mathbf{u} \in \mathcal{M}$ esetén. Az ezen tulajdonsággal rendelkező $r \in R$ elemek főideált alkotnak, amelynek egy generátorelemét az \mathcal{M} exponensének nevezzük.*

Ha az \mathcal{M} r exponense felírható páronként relatív prímek $r_1 \cdots r_k$ szorzataként, akkor \mathcal{M} felbontható részmodulusainak $\mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_k$ direkt összegére úgy, hogy \mathcal{M}_i is végesen generált torzió R -modulus és \mathcal{M}_i exponense r_i .

Ha az r_i elemek tovább már nem bonthatók relatív prím elemek szorzatára (azaz egyetlen felbonthatatlan elem hatványai), akkor a felbontás abszolút egyértelmű: bármely más felbontásban csak a tagok sorrendje változhat.

Az $\mathbf{u} \in \mathcal{M}$ generálta részmodulus $r(\mathbf{u})$ exponensét az \mathbf{u} elem rendjének nevezzük.

Bizonyítás. Mindenekelőtt felidézzük a definícióban használt gyűrűelméleti fogalmakat és szükséges eredményeket.

Az R gyűrű integritási tartomány, ha a szorzás kommutatív és nincsenek benne null-osztók. Egy $I \subseteq R$ nemüres részhalmaz ideál, ha zárt a kivonásra, összeadásra és a gyűrűbeli elemmel való szorzásra. Egy ideál főideál, ha egyetlen elem gyűrűbeli többszöröseiből áll (feltesszük, hogy a gyűrűben létezik egységelem). Egy gyűrű főideálgyűrű, ha benne minden ideál főideál (feltesszük, hogy integritási tartomány). Ha R főideálgyűrű, akkor bármely $a, b \in R$ esetében létezik olyan $x, y \in R$, amelyre az a, b generálta ideál d generátoreleme kielégíti az $ax + by = d$ feltételt (azaz $(a, b) = (d)$). Az a és b elemek relatív prímek, ha $(a, b) = (1)$. Főideálgyűrűben érvényes az egyértelmű faktorizáció. Ezt az eredményt egyelőre csak euklideszi gyűrű esetére tudjuk, de a későbbiek során belátjuk

főideálgyűrűkre is. Az egyértelmű faktorizációból következik, hogy ha a főideálgyűrű elemeinek $a_1, a_2, \dots, a_n \dots$ sorozatában minden elem osztója az előzőnek, akkor valahonnét kezdve ezek az elemek egymás asszociáltjai (tehát ugyanazt a főideált generálják); hiszen mindegyiknek legfeljebb annyi prímtényezője van, mint az előzőnek. A lineáris algebrai alkalmazásnál elég euklideszi gyűrűkre gondolni, mert a 7.1. Tételben megadott megfeleltetésnél polinomgyűrű szerepel, ami euklideszi gyűrű.

Ezek után nézzük a tétel bizonyítását:

Legyen $U = \{u_1, \dots, u_t\}$ az \mathcal{M} egy generátorrendszere. Mivel \mathcal{M} torziómodulus, ezért vannak olyan nemnulla $r_1, \dots, r_t \in R$ elemek, amelyekre $r_i u_i = \mathbf{o}$ ($1 \leq i \leq t$). Az $r = \prod_i r_i$ szorzat nem 0, mert R integritási tartomány, emellett $ru_i = \mathbf{o}$. Mivel U generátorrendszer és R kommutatív, ezért $ru = \mathbf{o}$ teljesül az R minden u elemére. A modulus axiómáiból következik, hogy az ilyen tulajdonságú $r \in R$ elemek ideált alkotnak; s mivel R főideálgyűrű, ezért ez az ideál főideál.

Tegyük most fel, hogy az \mathcal{M} modulus r exponense felbomlik a relatív prím a és b elemek szorzatára; így van olyan $x, y \in R$, amelyre $ax + by = 1$. Ez azt jelenti, hogy minden $u \in \mathcal{M}$ felírható $u = axu + byu$ alakban. Legyen $\mathcal{M}_a = \{bu \mid u \in \mathcal{M}\}$ és $\mathcal{M}_b = \{au \mid u \in \mathcal{M}\}$. Világos, hogy \mathcal{M}_a és \mathcal{M}_b részmodulusok, amelyekre a fenti felbontás alapján $\mathcal{M}_a + \mathcal{M}_b = \mathcal{M}$ következik.

Mivel $r = ab$ az \mathcal{M} exponense, ezért $a\mathcal{M}_a = \{abu \mid u \in \mathcal{M}\} = \{\mathbf{o}\}$ és $b\mathcal{M}_b = \{bau \mid u \in \mathcal{M}\} = \{\mathbf{o}\}$. Ha mármost $v \in \mathcal{M}_a \cap \mathcal{M}_b$, akkor $av = bv = \mathbf{o}$ miatt $v = 1v = x(av) + y(bv) = x\mathbf{o} + y\mathbf{o} = \mathbf{o}$, vagyis $\mathcal{M}_a \cap \mathcal{M}_b = \{\mathbf{o}\}$, tehát $\mathcal{M} = \mathcal{M}_a \oplus \mathcal{M}_b$.

Legyen a' az \mathcal{M}_a és b' az \mathcal{M}_b exponense. Mint láttuk, a' osztója a -nak és b' osztója b -nek. Másrészt $a'\mathcal{M}_a = \{\mathbf{o}\}$ miatt $a'b\mathcal{M} = \{\mathbf{o}\}$, és így $r = ab$ osztója $a'b$ -nek, amiből az következik, hogy a osztója a' -nek; vagyis \mathcal{M}_a exponense a . Hasonlóképpen \mathcal{M}_b -nek b az exponense.

Tekintettel arra, hogy U az \mathcal{M} generátorrendszere, ezért a $\{bu \mid u \in U\}$ elemek az \mathcal{M}_a egy generátorrendszerét alkotják. Tehát e részmodulusok végesen generáltak.

Ebből teljes indukcióval következik a többtagú összegfelbontásra vonatkozó állítás.

Az egyértelműség bizonyításánál tegyük fel, hogy egyetlen r_i sem bontható tovább relatív prím tényezők szorzatára. Legyen $\mathcal{M} = \mathcal{M}'_1 \oplus \dots \oplus \mathcal{M}'_{k'}$ egy tetszőleges, hasonló felbontás, ahol \mathcal{M}'_i exponense r'_i . Az R -beli egyértelmű faktorizáció miatt $k' = k$, és az r'_i elemek csak sorrendben különbözhetnek az r_i elemektől. Ezért feltehető, hogy $r'_i = r_i$. A továbbiakban már nem lesz szükség arra, hogy az r_i -k nem bonthatók fel relatív prím tényezők szorzatára.

Legyen $s_i = \frac{r}{r_i}$. Azt tudjuk, hogy r_i az \mathcal{M}_i exponense, és s_i a többi tag direkt összegének az exponense. Mivel r_i és s_i relatív prímekek, ezért r_i a többi tag direkt összegének egyetlen elemét sem viheti \mathbf{o} -ba, azaz $\mathcal{M}_i = \{u \in \mathcal{M} \mid r_i u = \mathbf{o}\}$. Ugyanez mondható el \mathcal{M}'_i -ről is. Tekintettel arra, hogy $\{u \in \mathcal{M} \mid r_i u = \mathbf{o}\}$ a részmodulustól független definíció, ezért a megfelelő részmodulusok valóban megegyeznek. ■

Következmény. Legyen \mathcal{U} (véges dimenziós) vektortér a K test felett és $\alpha \in \text{End}(\mathcal{U})$ (az \mathcal{U} egy lineáris transzformációja). Legyen továbbá $m(x) \in K[x]$ az α minimálpolinomja, és $m(x) = m_1(x) \cdot \dots \cdot m_t(x)$ e polinomnak $K[x]$ -beli páronként relatív prím faktorokra való felbontása.

Ekkor a tér felbomlik $\mathcal{M} = \mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_t$ direkt összegre, ahol minden egyes \mathcal{M}_i az α -nak invariáns altere. α -t az \mathcal{M}_i -re megszorítva, a kapott α_i minimálpolinomja pontosan $m_i(x)$. $\alpha = \alpha_1 \oplus \dots \oplus \alpha_t$. Ha minden egyes $m_i(x)$ faktor egyetlen irreducibilis polinomnak a hatványa, akkor az invariáns alterek (sorrendtől eltekintve) egyértelműen meghatározottak.

Legyen $\mathbf{U}_i = \{\mathbf{u}_{i,1}, \dots, \mathbf{u}_{i,n_i}\}$ ($1 \leq i \leq t$) és \mathbf{U} e bázisok sorrendben való egyesítése. Ha $A_i = [\alpha_i]$ a szereplő bázisban, akkor \mathcal{U} -nak a megadott bázisában $A = [\alpha]$ a következő alakú — úgynevezett blokkokra diagonalizált — mátrix:

Legyen $f(i) = \sum_{j < i} n_j$. Ekkor az A mátrixban az $(f(i) + 1)$ -től $f(i + 1)$ -ig terjedő

sorokban és oszlopokban pontosan az A_i mátrix áll; és minden további elem 0.

Bizonyítás. Mindenekelőtt megjegyezzük, hogy a részmodulusok pontosan az invariáns alterek.

Valóban, ha $\mathcal{V} \leq \mathcal{U}$ részmodulus, akkor α minden polinomja — speciálisan α is — önmagába viszi; így invariáns altér. Fordítva, ha \mathcal{V} invariáns altér, akkor α , így α minden hatványa és ezek skalárszorosának összegei is önmagába viszik; vagyis részmodulus.

A 7.1. Tétel szerint \mathcal{U} végesen generált torziómodulus $K[x]$ felett, és exponense az α transzformáció $m(x)$ minimálpolinomja. Tekintettel arra, hogy $K[x]$ -ben érvényes az euklideszi algoritmus, ezért alkalmazható a 7.3. Tétel. Eszerint tehát létezik egy $\mathcal{U} = \mathcal{U}_1 \oplus \dots \oplus \mathcal{U}_t$ direkt összegre való olyan egyértelmű felbontás, amelyben az i -edik komponens exponense az $m_i(x)$ polinom.

A 7.3. Tételből az is következik, hogy a megszorítással adódó α_i minimálpolinomja az $m_i(x)$ polinom. Az α -nak ezek direkt összegére való felbontása a 3.9. Tételből következik.

Az A mátrix előállítás azonnal következik abból, hogy az \mathcal{U}_i alterek invariánsak. ■

7.4. Tétel. Legyen az \mathcal{U} tér α lineáris transzformációjának az $m(x)$ minimálpolinomja egy K felett irreducibilis r -edfokú $p(x)$ polinom k -adik hatványa. Ekkor bármely $i \leq k$ természetes számhoz létezik a térben α -nak olyan invariáns altere, amelynek dimenziója pontosan $i \cdot r$.

Egy transzformáció minimálpolinomjának foka legfeljebb akkora, mint a tér dimenziója.

Bizonyítás. Legyen $\beta = p(\alpha)$. Feltétel szerint $\beta^k = \omega$, de $i < k$ esetén $\beta^i \neq \omega$. Ez azt jelenti, hogy van a térben olyan \mathbf{u} vektor, amelyre $\beta^k(\mathbf{u}) = \mathbf{o}$, de $\beta^{k-1}(\mathbf{u}) \neq \mathbf{o}$. Tekintsük az

$$\mathbf{u}_1 = \beta^0(\mathbf{u}) = \mathbf{u}, \mathbf{u}_2 = \beta^1(\mathbf{u}), \dots, \mathbf{u}_k = \beta^{k-1}(\mathbf{u})$$

vektorokat. Tekintsük továbbá minden $i < r$ mellett az

$$\mathbf{u}_i, \alpha(\mathbf{u}_i), \dots, \alpha^{r-1}(\mathbf{u}_i)$$

vektorokat is. Ezeket a vektorokat írjuk fel táblázatban:

$$\begin{array}{cccc} \mathbf{u}_1, & \alpha(\mathbf{u}_1), & \dots, & \alpha^{r-1}(\mathbf{u}_1) \\ \mathbf{u}_2, & \alpha(\mathbf{u}_2), & \dots, & \alpha^{r-1}(\mathbf{u}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k, & \alpha(\mathbf{u}_k), & \dots, & \alpha^{r-1}(\mathbf{u}_k) \end{array}$$

Azt fogjuk megmutatni, hogy ezek a vektorok lineárisan függetlenek, és bármelyik sortól kezdve e sorban és az összes alatta levő sorban fellépő vektorok α egy invariáns alterét generálják. Tekintsük az i -edik sortól kezdve az összes elemet. Ezek $\mathbf{u}_{s,j} = \alpha^s(\mathbf{u}_j)$, ahol $1 \leq s < r$ és $i \leq j \leq k$. Ha $s < r-1$, akkor $\alpha(\mathbf{u}_{s,j}) = \mathbf{u}_{s+1,j}$. Az $s = r-1$ esetben $\alpha(\mathbf{u}_{r-1,j}) = \alpha^r(\mathbf{u}_j)$ adódik.

Feltehető, hogy a $p(x)$ polinom normált: $p(x) = x^r + c_{r-1}x^{r-1} + \dots + c_1x + c_0$. Ebből azt kapjuk, hogy

$$\beta = p(\alpha) = \alpha^r + c_{r-1}\alpha^{r-1} + \dots + c_1\alpha + c_0,$$

azaz

$$\alpha^r(\mathbf{u}_j) = \mathbf{u}_{j+1} - (c_{r-1}\alpha^{r-1}(\mathbf{u}_j) + \dots + c_1\alpha(\mathbf{u}_j) + c_0\mathbf{u}_j).$$

Mivel a jobb oldalon álló vektorok mind a vizsgált altérben vannak, ezért $\alpha^r(\mathbf{u}_j)$ is oda esik (a $j = k$ esetben $\beta(\mathbf{u}_k) = \mathbf{o}$). Így a j -edik sor bármely vektorát α a szereplő altérbe viszi. Mivel ez minden sorra igaz, ezért az altér invariáns.

A fenti $\alpha^s(\mathbf{u}_j)$ alakú vektorok száma kr . Tekintettel arra, hogy ezek a vektorok generátorrendszert alkotnak, ezért lineáris függetlenségük bizonyítására elegendő ebben az altérben kr darab független vektort találni. Vegyük evégett az $\mathbf{u}, \alpha(\mathbf{u}), \dots, \alpha^{kr-1}(\mathbf{u})$ alakú vektorokat. Ha ezeknek a $\sum_{i=0}^{kr-1} c_i \alpha^i(\mathbf{u})$ lineáris kombinációja \mathbf{o} , akkor az $f(x) = \sum_{i=0}^{kr-1} c_i x^i$

polinomra $f(\alpha) : \mathbf{u} \mapsto \mathbf{o}$. Feltehető, hogy $f(x)$ a legalacsonyabbfokú olyan polinom, amelyre $f(\alpha)(\mathbf{u}) = \mathbf{o}$. Ha mármost $d(x)$ az $m(x)$ és az $f(x)$ legnagyobb közös osztója, akkor $d(\alpha) : \mathbf{u} \mapsto \mathbf{o}$ is igaz, mert $m(\alpha)$ minden vektort \mathbf{o} -ba visz. Az $f(x)$ választása folytán tehát $d(x) = f(x)$. A polinomok egyértelmű faktorizációja alapján csak $d(x) = p(x)^t$ lehet ($t \leq k$). Mivel $p(\alpha) = \beta$, ezért azt kaptuk, hogy $\beta^t(\mathbf{u}) = \mathbf{o}$. Feltételünk szerint ez csak a $t \geq k$ esetben lehet, amiből $gr(f(x)) \geq kr$ következik, ellentétben feltevésünkkel. ■

7.5. Tétel. *Az α transzformáció minimálpolinomjának gyökei pontosan az α sajátértékei.*

Bizonyítás. Ha c az α egy sajátértéke, akkor van olyan $\mathbf{u} \neq \mathbf{o}$ vektor, amelyre $\alpha(\mathbf{u}) = c\mathbf{u}$. Ebből következik, hogy tetszőleges $f(x)$ alaptestbeli együtthatós polinomra $f(\alpha)(\mathbf{u}) = f(c)\mathbf{u}$. Ha $f(x) = m(x)$ az α minimálpolinomja, akkor $\mathbf{o} = m(\alpha)(\mathbf{u}) = m(c)\mathbf{u}$, ami $\mathbf{u} \neq \mathbf{o}$ miatt csak úgy lehet, ha $m(c) = 0$; a sajátérték tehát gyöke a minimálpolinomnak.

Ha c gyöke az α transzformáció $m(x)$ minimálpolinomjának, akkor $m(x)$ osztható az $x - c$ polinommal. A 7.3. Tétel Következménye és a 7.4. Tétel alapján tehát van a térben az α -nak olyan invariáns altere, amely egydimenziós és α -t erre megszorítva a minimálpolinomja pontosan $x - c$. Ez azt jelenti, hogy ennek az altérnek az elemeit az $\alpha - c$

transzformáció a nullvektorba viszi. Ennek az alternak az \mathbf{u} generátoreleme nem a nullvektor (hiszen az altern egydimenziós) és $(\alpha - c\iota)(\mathbf{u}) = \mathbf{0}$, azaz $\alpha(\mathbf{u}) = c\mathbf{u}$. ■

Megjegyzés. A karakterisztikus polinom tárgyalásánál láttuk, hogy egy transzformáció karakterisztikus polinomjának a gyökei is a transzformáció sajátértékei. A későbbiekben majd látni fogjuk, hogy e két polinom között sokkal szorosabb kapcsolat áll fenn. □

7.6. Tétel. *A komplex számtest feletti véges dimenziós vektorterekben minden transzformációnak van sajátvektora; a valós számtest feletti véges dimenziós vektorterekben minden transzformációnak van legfeljebb kétdimenziós invariáns altére.*

Bizonyítás. Az első állítás azonnal következik a 7.5. Tételből, hiszen a komplex számtestben minden polinomnak van gyöke.

A valós esetben a 7.3. Tétel Következménye és a 7.4. Tétel alapján van a térben olyan invariáns altern, amelynek dimenziója megegyezik a minimálpolinom egy irreducibilis faktorának a fokával. Ebből azonnal következik az állítás, figyelembe véve, hogy a valós test felett minden nemkonstans polinom felbomlik első- vagy másodfokú polinomok szorzatára. ■

Feladatok

1. Legyenek $R, \mathcal{M}, \mathcal{M}_i, r, r_i$, mint a 7.3. Tételben. Mutassuk meg, hogy \mathcal{M}_i a legnagyobb olyan részmodulus, amelynek r_i az exponense. Mutassuk meg, hogy $j \neq i$ esetén az r_i -vel való szorzás bijekció az \mathcal{M}_j -n.

2. Legyen α az \mathbb{C} feletti n -dimenziós \mathcal{U} vektortér lineáris transzformációja. Mekkora lehet minimálisan és maximálisan az α invariáns altéréinek a száma?

3. Adott az n -dimenziós \mathcal{U} vektortér \mathbb{C} felett. Mely $f(x)$ polinomok esetén létezik hasonlóságtól eltekintve egyetlen olyan transzformáció, amelynek ez a minimálpolinomja? (Az α és β transzformációk hasonlóak, ha van olyan σ invertálható transzformáció, amelyre $\alpha = \sigma^{-1}\beta\sigma$. A hasonlóság azt jelenti, hogy a két transzformáció ugyanúgy hat, csak más bázisok esetén. Ebből következik, hogy hasonló transzformációk minimálpolinomjai is megegyeznek.)

4. Lehet-e egy transzformáció minimálpolinomjának többszörös gyöke?

5. Bizonyítsuk be, hogy adott elsőfokú polinomhoz csak egy olyan transzformáció létezik, amelynek ez a minimálpolinomja.

6. Írjuk le (a legalább háromdimenziós térben) azokat a transzformációkat, amelyeknek a minimálpolinomja $x^2 - 1$, $(x - 1)^2$, illetve $x^2 - x$. Mutassuk meg, hogy ezen polinomok bármelyike esetében található olyan transzformációk, amelyeknek ez a minimálpolinomjuk, de nem hasonlóak.

7. Bizonyítsuk be, hogy az n -dimenziós ${}_K\mathcal{U}$ vektortér esetében a tetszőlegesen adott $x^n - a_{n-1}x^{n-1} - \dots - a_1x - a_0 \in K[x]$ polinomhoz létezik olyan transzformáció, amelynek ez a minimálpolinomja.

8. Mutassuk meg, hogy a kétdimenziós valós térnek van olyan transzformációja, amelynek nincs sajátvektora.

9. Bizonyítsuk be, hogy a háromdimenziós valós térben viszont minden transzformációnak van sajátvektora.

10. Legyen $f(x) \in \mathbb{Q}[x]$ egy irreducibilis n -edfokú polinom és α a racionális test feletti n -dimenziós \mathcal{U} térnek egy olyan lineáris transzformációja, amelynek $f(x)$ a minimálpolinomja. Bizonyítsuk be, hogy $\text{End}(\mathcal{U})$ -nak az $\iota, \alpha, \dots, \alpha^{n-1}$ transzformációk által kifeszített altere a transzformációkkal végezhető műveletekre nézve testet alkot; s e testben az $f(x)$ polinomnak létezik gyöke.

11. Bizonyítsuk be, hogy ha az előző feladatban az $f(x)$ polinom nem irreducibilis, akkor nem kaphatunk testet.

2. Lineáris transzformációk invariáns alterei az euklideszi térben

Az euklideszi terek esetében a skalárszorzat segítségével igen egyszerűen definiálható a transzformációk adjungáltja és transzponáltja. Tulajdonképpen elég az adjungálttal foglalkozni, hiszen valós esetben ez automatikusan a transzponáltba megy át.

Tekintsük az $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázishoz tartozó skalárszorzatot, és írjuk fel az α transzformáció mátrixát ebben a bázisban. E mátrix j -edik oszlopában az $\alpha(\mathbf{u}_j)$ vektor koordinátái állnak; s ennek i -edik eleme $\mathbf{u}_i^*(\alpha(\mathbf{u}_j))$, ami nem más, mint $(\mathbf{u}_i; \alpha(\mathbf{u}_j))$. Így az α transzformáció \mathbf{U} bázisában felírt $[\alpha]^{\mathbf{U}} = A = [a_{i,j}]$ mátrixára $a_{i,j} = (\mathbf{u}_i; \alpha(\mathbf{u}_j))$ teljesül. Mármint, a skalárszorzat tulajdonságai szerint

$$(\mathbf{u}_i; \alpha^*(\mathbf{u}_j)) = \overline{a_{j,i}} = \overline{(\mathbf{u}_j; \alpha(\mathbf{u}_i))} = (\alpha(\mathbf{u}_i); \mathbf{u}_j).$$

Tekintettel arra, hogy a skalárszorzat bilineáris függvény, ezért tetszőleges \mathbf{u} és \mathbf{v} vektorokra igaz az

$$(\mathbf{u}; \alpha^*(\mathbf{v})) = (\alpha(\mathbf{u}); \mathbf{v})$$

összefüggés. Ez azt jelenti, hogy a transzformáció adjungáltját definiálhatjuk a skalárszorzattal. Ha tetszőleges ortonormált bázist veszünk fel, akkor is ugyanezt az adjungálást kapjuk, erről egyszerűen meggyőződhetünk a bázisvektorok behelyettesítésével. Az is világos, hogy a skalárszorzatot definiáló bázis skalárszorosát véve is ugyanez marad az adjungálás.

Megmutatjuk, hogy az adjungálás a fenti bázisváltoztatás erejéig egyértelműen meghatározza a skalárszorzatot.

Tekintsünk a téren két skalárszorzatot (pozitív definit szimmetrikus függvényt): $\mathbf{S}(\mathbf{u}, \mathbf{v})$ és $\mathbf{T}(\mathbf{u}, \mathbf{v})$. Az 5.8. Tétel szerint adott bázis esetén minden bilineáris függvényhez létezik olyan transzformáció, amelynek az adott bázisban ugyanaz a mátrixa, mint a megadott bilineáris függvényé. Eszerint léteznek olyan σ és τ transzformációk, amelyekre

$$\mathbf{S}(\mathbf{u}, \mathbf{v}) = \mathbf{T}(\sigma(\mathbf{u}), \mathbf{v}) \quad \text{és} \quad \mathbf{T}(\mathbf{u}, \mathbf{v}) = \mathbf{S}(\tau(\mathbf{u}), \mathbf{v})$$

teljesül. Ebből $\mathbf{S}(\mathbf{u}, \mathbf{v}) = \mathbf{S}(\tau\sigma(\mathbf{u}), \mathbf{v})$ következik, tehát $\tau\sigma = \iota$, vagyis τ és σ egymás inverzei. Jelölje S , illetve T , megfelelően, az adjungálást, azaz legyen

$$\mathbf{S}(\alpha(\mathbf{u}), \mathbf{v}) = \mathbf{S}(\mathbf{u}, \alpha^S(\mathbf{v})) \quad \text{és} \quad \mathbf{T}(\alpha(\mathbf{u}), \mathbf{v}) = \mathbf{T}(\mathbf{u}, \alpha^T(\mathbf{v})).$$

Ebből azt kapjuk, hogy:

$$\begin{aligned} \mathbf{T}(\mathbf{u}, \alpha^T(\mathbf{v})) &= \mathbf{T}(\alpha(\mathbf{u}), \mathbf{v}) = \mathbf{S}(\tau\alpha(\mathbf{u}), \mathbf{v}) = \\ &= \mathbf{S}(\mathbf{u}, \alpha^S\tau^S(\mathbf{v})) = \mathbf{T}(\sigma(\mathbf{u}), \alpha^S\tau^S(\mathbf{v})) = \mathbf{T}(\mathbf{u}, \sigma^T\alpha^S\tau^S(\mathbf{v})); \end{aligned}$$

amiből $\alpha^T = \sigma^T\alpha^S\tau^S$, illetve $\tau^T\alpha^T = \alpha^S\tau^S$ következik, figyelembe véve, hogy τ és σ egymás inverzei.

Ha a két adjungálás megegyezik, akkor ebből $\tau^S\alpha^S = \alpha^S\tau^S$ következik. Mivel α és így α^S is tetszőleges, ezért ez csak úgy lehet, ha τ^S és így τ is az identitásnak nemnulla skalárszorosa. ■

7.1. Definíció. A \mathbb{C} feletti \mathcal{U} (komplex) euklideszi tér α transzformációjának adjungáltján azt az α^* transzformációt értjük, amelyre $(\mathbf{u}; \alpha^*(\mathbf{v})) = (\alpha(\mathbf{u}); \mathbf{v})$ teljesül.

Az \mathbb{R} esetében (tehát valós euklideszi térben) transzponáltról beszélünk. ■

Ez a definíció nyilvánvalóan ekvivalens az eredetivel. Ugyancsak azonnal következik az alábbi

7.7. Tétel. Az \mathbf{U} bázishoz tartozó skalárszorzatot tekintve az α és α^* transzformációkra minden ortonormált bázisban teljesül az $[\alpha^*] = [\alpha]^*$ összefüggés. ■

Megjegyzés. A mátrixok tárgyalásánál láttuk, hogy $AB^* = B^*A^*$, ezért ez igaz a lineáris transzformációkra is. Az itteni definíció alapján ez most minden számolás nélkül adódik:

$$(\mathbf{u}; (\alpha\beta)^*(\mathbf{v})) = (\alpha\beta(\mathbf{u}); \mathbf{v}) = (\beta(\mathbf{u}); \alpha^*(\mathbf{v})) = (\mathbf{u}; \beta^*\alpha^*(\mathbf{v})). \quad \square$$

7.8. Tétel. Az \mathcal{U} euklideszi tér egy \mathcal{V} altere akkor és csak akkor invariáns altere α -nak, ha \mathcal{V}^\perp invariáns altere α^* -nak.

Bizonyítás. Legyen \mathcal{V} az α invariáns altere és tekintsük a \mathcal{V}^\perp egy tetszőleges \mathbf{w} elemét. Mivel \mathcal{V} az α -nak invariáns altere, ezért tetszőleges $\mathbf{v} \in \mathcal{V}$ vektorra $\alpha(\mathbf{v}) \in \mathcal{V}$ is igaz. A merőleges alterek definíciója szerint tehát $\alpha(\mathbf{v}) \perp \mathbf{w}$, azaz $(\alpha(\mathbf{v}); \mathbf{w}) = 0$. Az adjungált definíciójából $(\mathbf{v}; \alpha^*(\mathbf{w})) = 0$ következik, ami azt jelenti, hogy $\mathbf{v} \perp \alpha^*(\mathbf{w})$. Tekintettel arra, hogy \mathbf{v} a \mathcal{V} altern tetszőleges eleme volt, ezért $\alpha^*(\mathbf{w}) \perp \mathcal{V}$, azaz $\alpha^*(\mathbf{w}) \in \mathcal{V}^\perp$, mint állítottuk.

Ebből már azonnal következik az állítás megfordítása is, tekintettel arra, hogy $(\alpha^*)^* = \alpha$ és $(\mathcal{V}^\perp)^\perp = \mathcal{V}$. ■

7.2. Definíció. Az euklideszi tér egy α transzformációját normálisnak nevezzük, ha $\alpha^*\alpha = \alpha\alpha^*$. ■

A normális transzformációk leírásához fel fogunk használni egy mátrixokra vonatkozó, önmagában is érdekes eredményt:

Lemma. Tetszőleges A és B mátrixokra, ha létezik az AB és a BA szorzat, akkor $S(AB) = S(BA)$. Ha az A mátrixra $S(AA^*) = 0$, akkor $A = O$. ($S(M)$ az M nyoma.)

Bizonyítás. Legyen $A = [a_{i,j}]$ és $B = [b_{p,q}]$. Az AB mátrix i -edik sorának i -edik eleme $\sum_j a_{i,j} b_{j,i}$. A fődiagonális elemeinek az összege tehát $S(AB) = \sum_{i,j} a_{i,j} b_{j,i}$. Az $S(BA)$ -ra adódó érték: $\sum_{p,q} b_{p,q} a_{q,p}$, ami nyilvánvalóan megegyezik $S(AB)$ -vel.

$A B = A^*$ esetben azt kapjuk, hogy $S(AA^*) = \sum_{i,j} a_{i,j} \overline{a_{i,j}} = \sum_{i,j} |a_{i,j}|^2$. Ez pedig nem más, mint az A mátrix elemei abszolút értéke négyzeteinek az összege. Ezért az $S(AA^*) = 0$ esetben A minden eleme 0, ezért $A = O$. ■

7.9. Tétel. Egy euklideszi tér bármely α normális transzformációjához találhatók olyan minimális invariáns alterek, amelyek páronként ortogonálisak, direkt összegük az egész tér és α -t ezek bármelyikére megszorítva ismét normális transzformációt nyerünk.

Bizonyítás. A tételt a tér dimenziójára vonatkozó teljes indukcióval bizonyítjuk. Egy-dimenziós térre az állítás triviális.

Tegyük fel, hogy igaz a tétel minden n -nél kisebb dimenziójú térre, és legyen α az n -dimenziós \mathcal{U} euklideszi tér egy normális lineáris transzformációja. Ha α -nak nincsen valódi invariáns altere, akkor \mathcal{U} minimális invariáns altér, és a tétel triviálisan igaz.

A továbbiakban legyen \mathcal{V} az α -nak k dimenziós valódi invariáns altere (azaz $0 < k < n$). Legyen \mathbf{V} a \mathcal{V} egy ortonormált bázisa és \mathbf{W} a $\mathcal{W} = \mathcal{V}^\perp$ altéré. Világos, hogy $\mathbf{U} = \mathbf{V} \cup \mathbf{W}$ ortonormált bázis az \mathcal{U} térben. Írjuk fel α mátrixát ebben a bázisban úgy, hogy először \mathbf{V} vektorait soroljuk fel. Tekintettel arra, hogy \mathcal{V} invariáns altér, ezért \mathcal{V} -beli elemeket \mathcal{V} -beli elemekbe visz, ezért az első k oszlopban az első k sort kivéve csupa 0 áll.

Ez azt jelenti, hogy α mátrixa $\begin{bmatrix} A & B \\ O & C \end{bmatrix}$ alakú, ahol A és C négyzetes mátrixok (O a csupa 0 elemű mátrix). Ekkor viszont α^* mátrixa ugyanebben a bázisban $\begin{bmatrix} A^* & O^* \\ B^* & C^* \end{bmatrix}$. α és α^* felcserélhetősége alapján, a mátrixokat „blokkonként” összeszorozva

$$\begin{bmatrix} AA^* + BB^* & BC^* \\ CB^* & CC^* \end{bmatrix} = \begin{bmatrix} A^*A & A^*B \\ B^*A & B^*B + C^*C \end{bmatrix}$$

adódik. Az első k sorból és oszlopból álló blokkokat összehasonlítva, ebből azt kapjuk, hogy $AA^* + BB^* = A^*A$. Mivel egy mátrix nyoma egyértelmű és összeg nyoma megegyezik a nyomok összegével, ezért ebből $S(AA^*) + S(BB^*) = S(A^*A)$ következik.

A Lemma első állítása szerint $S(BB^*) = 0$, amiből a Lemma második állítása szerint $B = O^*$ következik (azért írunk itt O^* -ot, mert ez az eredeti mátrixban szereplő O mátrix tükörképe).

Ezek alapján α a \mathcal{W} altér elemeit is \mathcal{W} -be viszi, így \mathcal{W} is invariáns altere α -nak. Mivel ezek eleve ortogonálisak, ezért a teret valóban felbontottuk két egymásra merőleges valódi invariáns altér direkt összegére. α -t \mathcal{V} -re megszorítva, ennek mátrixa a \mathbf{V} bázisban A . α^*

esetében $B^* = O$ miatt erre A^* adódik. Mivel $B = O$, ezért $BB^* = O$, amiből $AA^* + BB^* = A^*A$ következtében $AA^* = A^*A$ adódik. Ez pedig pontosan azt jelenti, hogy α -nak \mathcal{V} -re való megszorítása normális. A két szorzatmátrix utolsó blokkjainak összehasonlításából hasonlóképpen kapjuk, hogy a \mathcal{W} -re való megszorítás is normális. ■

7.10. Tétel. *A komplex számtest feletti vektortér egy lineáris transzformációja akkor és csak akkor normális, ha alkalmas ortonormált bázisban mátrixa diagonálissá válik.*

A valós számtest feletti vektortérben bármely normális lineáris transzformációhoz van olyan ortonormált bázis, amelyben mátrixa legfeljebb kétszer kettes diagonális blokkokra bomlik.

Bizonyítás. A 7.9. Tétel szerint a tér felbomlik páronként ortogonális minimális invariáns alterek direkt összegére. A 7.6. Tétel szerint ezek a minimális invariáns alterek a komplex esetben egydimenziósak, a valós esetben pedig legfeljebb kétdimenziósak. Ha a bázisvektorokat ezekből az alterekből választjuk, akkor egy olyan ortonormált bázist kapunk, amelyben a transzformáció mátrixa éppen a kívánt alakot ölti.

A megfordításhoz, a komplex esetben, tekintsünk egy olyan α transzformációt, amelynek a mátrixa alkalmas ortonormált bázisban diagonálissá válik. Természetesen, ekkor α^* mátrixa is diagonális lesz. A diagonális mátrixok szorzási „szabály”-ának következtében a két mátrixtényező felcserélhető, hiszen a komplex számok szorzása kommutatív. Tekintettel arra, hogy rögzített bázis esetén a mátrixok egyértelműen meghatározzák a lineáris transzformációt, a két transzformáció valóban felcserélhető. ■

Feladatok

1. Az adjungált itteni definícióját felhasználva bizonyítsuk be az adjungálás elemi tulajdonságait; beleértve azt, hogy az $\alpha\alpha^* = \omega$ feltételből következik, hogy $\alpha = \omega$.

2. Egy euklideszi tér α és β transzformációit felcserélhetőeknek nevezzük, ha $\alpha\beta = \beta\alpha$. Bizonyítsuk be, hogy a térben létezik olyan közös invariáns altér, amelyekre ezeket megszorítva a megszorítások minimálpolinomja az eredeti minimálpolinomok egy-egy irreducibilis faktora.

3. Bizonyítsuk be, hogy az előző feladatban szereplő invariáns altér dimenziója legfeljebb a két irreducibilis faktor fokának a szorzata. Mutassuk meg, hogy a szorzat akkor is felléphet, ha a két faktor foka megegyezik.

4. Igaz-e, hogy a komplex euklideszi térben két felcserélhető normális transzformáció szorzata ismét normális?

5. Igaz-e, hogy a komplex euklideszi térben két felcserélhető normális transzformációnak megegyeznek a sajátvektorai?

6. Igaz-e, hogy ha a komplex euklideszi tér két normális lineáris transzformációjának megegyeznek a sajátvektorai, akkor a transzformációk felcserélhetőek? Igaz-e, hogy ebben az esetben a szorzat is normális?

7. Bizonyítsuk be, hogy egy legalább kétdimenziós komplex euklideszi térben minden transzformációhoz található tőle lineárisan független vele felcserélhető transzformáció.

8. Bizonyítsuk be, hogy egy kétdimenziós komplex euklideszi térben nem létezik három lineárisan független páronként felcserélhető lineáris transzformáció.

9. Bizonyítsuk be, hogy egy n -dimenziós komplex euklideszi térben létezik n darab lineárisan független páronként felcserélhető lineáris transzformáció.

10. Bizonyítsuk be, hogy ha α^* polinomja α -nak, akkor α normális transzformáció.

11. Az interpoláció segítségével mutassuk meg, hogy a komplex esetben, ha α normális, akkor α^* polinomja α -nak.

3. Szimmetrikus és önadjungált transzformációk

A normális transzformációknak két igen fontos speciális esete van. Először a könnyebben leírhatóval foglalkozunk. Ezt a kvadratikus alakok euklideszi térben való vizsgálatánál fogjuk felhasználni.

7.3. Definíció. A komplex, illetve valós számtest feletti euklideszi tér egy α lineáris transzformációját önadjungáltnak, illetve szimmetrikusnak nevezzük, ha $\alpha^* = \alpha$. ■

Mindenekelőtt megmutatjuk a transzformációknak a kvadratikus alakokkal való szoros kapcsolatát. Emlékeztetünk rá, hogy a valós számtest feletti vektorterekben a kvadratikus alakokat a szimmetrikus bilineáris függvényekkel definiáltuk; míg a komplex számtest esetében a kvadratikus karakter csak akkor vált értelmezhetővé, ha a bilineáris függvény Hermite-féle volt.

7.11. Tétel. *Az euklideszi tér α lineáris transzformációja akkor és csak akkor önadjungált, illetve szimmetrikus, ha az $\mathbf{A}(\mathbf{u}, \mathbf{v}) = \mathbf{u}^* \alpha \mathbf{v}$ bilineáris függvény Hermite-féle, illetve szimmetrikus.*

Bizonyítás. Mindenekelőtt tisztázni kell \mathbf{u}^* jelentését. Tekintettel arra, hogy a skalárszorzatot egy bázis határozza meg, ezért világos, hogy a duális tér kijelölt bázisa e bázisnak legyen a duálisa.

A bilineáris függvényekre vonatkozó feltétel mindkét esetben azt jelenti, hogy $\mathbf{A}(\mathbf{u}, \mathbf{v})$ és $\mathbf{A}(\mathbf{v}, \mathbf{u})$ egymás konjugáltjai. Az $\overline{\mathbf{A}(\mathbf{v}, \mathbf{u})} = \overline{\mathbf{v}^* \alpha \mathbf{u}} = \mathbf{u}^* \alpha^* \mathbf{v}$ összefüggés szerint e feltétel azzal ekvivalens, hogy $\mathbf{u}^* \alpha \mathbf{v} = \mathbf{u}^* \alpha^* \mathbf{v}$. Ebből a $\beta = \alpha^* - \alpha$ lineáris transzformációra $(\mathbf{u}; \beta \mathbf{v}) = \mathbf{u}^* \beta \mathbf{v} = 0$ következik. Most, a speciális $\mathbf{u} = \beta \mathbf{v}$ választással azt kapjuk, hogy $(\beta \mathbf{v}; \beta \mathbf{v}) = 0$. Mivel a skalárszorzat pozitív definit, ezért ez csak akkor lehet, ha $\beta \mathbf{v} = \mathbf{0}$ teljesül minden \mathbf{v} vektorra, azaz $\beta = \mathbf{0}$. Így $\alpha^* = \alpha$, mint állítottuk.

Fordítva, az $\alpha^* = \alpha$ egyenlőségből azonnal következik, hogy $\mathbf{u}^* \alpha \mathbf{v} = \mathbf{u}^* \alpha^* \mathbf{v}$, ami — mint láttuk — ekvivalens azzal, hogy α önadjungált, illetve szimmetrikus. ■

7.12. Tétel. *Önadjungált, illetve szimmetrikus transzformáció mátrixa minden ortonormált bázisban önadjungált, illetve szimmetrikus; s ha egy transzformáció mátrixa valamely ortonormált bázisban önadjungált, illetve szimmetrikus, akkor a transzformáció is önadjungált, illetve szimmetrikus.*

Bizonyítás. Azonnal következik a 7.7. Tételből, tekintettel arra, hogy α önmagával felcserélhető — így normális. ■

7.13. Tétel. *Önadjungált transzformáció sajátértékei valósak. Önadjungált transzformációhoz található olyan ortonormált bázis, amelyben mátrixa diagonálissá válik, és a diagonális elemei a transzformáció sajátértékei.*

Bizonyítás. Legyen c az α önadjungált transzformációnak egy \mathbf{u} sajátvektorhoz tartozó sajátértéke. A skalárszorzat pozitív definitiségének következtében van olyan d skalár, amelyre $(\mathbf{u}, \mathbf{u}) = d^2$. Mivel egy sajátvektor nem lehet nullvektor, ezért feltehető, hogy $d > 0$. Így $d\mathbf{e} = \mathbf{u}$ egyértelműen definiál egy \mathbf{e} vektort, amely ugyancsak a c sajátértékhez tartozik és abszolút értéke 1. Erre — $\alpha^* = \alpha$ alapján — :

$$c = (\mathbf{e}; c\mathbf{e}) = (\mathbf{e}; \alpha\mathbf{e}) = (\alpha^*\mathbf{e}; \mathbf{e}) = (\alpha\mathbf{e}; \mathbf{e}) = \overline{(\mathbf{e}; \alpha\mathbf{e})} = \bar{c} \cdot 1,$$

ami pontosan azt jelenti, hogy c valós.

A tétel további része a normális transzformációk leírását nyújtó 7.10. Tételből következik. ■

Szimmetrikus transzformáció a valós esetben is normális. Tekintettel azonban arra, hogy a valós test feletti polinomok között másodfokúak is lehetnek irreducibilisek, ezért normális transzformációknak lehet kétdimenziós minimális invariáns altère — és ez valóban elő is fordul. Szimmetrikus transzformációk esetében viszont kétdimenziós invariáns altér soha nem minimális:

7.14. Tétel. *Szimmetrikus transzformációhoz található a térben olyan ortonormált bázis, amelynek elemei a transzformáció sajátvektorai. E bázisban a transzformáció mátrixa diagonális.*

Bizonyítás. Az első állításhoz csupán azt kell belátni — mint előzetesen megjegyeztük —, hogy kétdimenziós invariáns altér nem lehet minimális.

Ezt sokféleképpen lehet bizonyítani, a legegyszerűbb talán az, ha felírjuk a transzformáció mátrixát egy ortonormált bázisban. A szimmetria miatt ez a mátrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ alakú, ahol a, b, c valós számok. E mátrixnak a karakterisztikus polinomja $(x - a)(x - c) - b^2 = x^2 - (a + c)x + ac - b^2$. E polinom diszkriminánsa $D = (a + c)^2 - 4ac + 4b^2 = (a - c)^2 + 4b^2$. Mivel ez nem negatív, ezért a karakterisztikus polinomnak létezik valós gyöke, amiről látuk, hogy a megfelelő transzformációnak sajátértéke. Ezért a transzformációnak van egydimenziós invariáns altère.

A második állítás most már azonnal adódik a 7.10. Tételből. ■

Feladatok

1. Határozzuk meg azoknak a transzformációknak a szerkezetét, amelyeknek az adjungáltja, illetve transzponáltja az eredetinek elsőfokú polinomja.
2. Bizonyítsuk be, hogy önadjungált, illetve szimmetrikus transzformációk összege is önadjungált, illetve szimmetrikus.
3. Bizonyítsuk be, hogy szimmetrikus transzformációk lineáris kombinációja is szimmetrikus. Mutassuk meg, hogy ez önadjungáltakra csak bizonyos megszorítással érvényes.
4. Mi a feltétele annak, hogy két önadjungált, illetve szimmetrikus transzformáció szorzata is önadjungált, illetve szimmetrikus legyen? Mi a feltétel háromtényezős szorzat esetén?
5. Milyen kapcsolatot jelent a komplex test feletti vektortér egy transzformációja és annak adjungáltja között az, hogy a transzformáció mátrixa egy ortonormált bázisban szimmetrikus?
6. Adjunk meg a \mathbb{Q} feletti kétdimenziós euklideszi térben olyan szimmetrikus transzformációt, amelynek nincs sajátvektora.
7. Bizonyítsuk be, hogy az euklideszi tér tetszőleges α transzformációjára $\alpha\alpha^*$ mindig önadjungált, illetve szimmetrikus. A normális transzformációknál szereplő lemma bizonyításában látottak felhasználásával mutassuk meg, hogy nem minden önadjungált, illetve szimmetrikus transzformáció írható fel a fenti alakban.

4. Ortogonális és unitér transzformációk

Most a normális transzformációk egy másik fontos típusát fogjuk vizsgálni.

7.4. Definíció. A komplex, illetve a valós tér egy lineáris transzformációját unitérnek, illetve ortogonálisnak nevezzük, ha adjungáltja, illetve transzponáltja megegyezik inverzével. ■

Megjegyzés. Mivel egy transzformáció inverze mindkétoldali inverz, ezért egy unitér, illetve egy ortogonális transzformáció normális. □

Az unitér, illetve ortogonális transzformáció fontosságát és a skaláris szorzattal való szoros kapcsolatát mutatja az alábbi:

7.15. Tétel. *Az euklideszi tér egy α lineáris transzformációja akkor és csak akkor unitér, illetve ortogonális, ha skalárszorzzattartó.*

Bizonyítás. Tekintsük az

$$(\alpha \mathbf{u}; \alpha \mathbf{v}) - (\mathbf{u}; \mathbf{v}) = (\mathbf{u}; \alpha^* \alpha \mathbf{v}) - (\mathbf{u}; \mathbf{v}) = (\mathbf{u}; (\alpha^* \alpha - \mathbf{I}) \mathbf{v})$$

egyenlőséget. Itt a bal oldalon akkor és csak akkor áll minden \mathbf{u} és \mathbf{v} esetén 0, ha α skalárszorzzattartó. A jobb oldalon biztosan 0 áll, ha α unitér, illetve ortogonális. Ha a jobb oldalon minden vektorpárra 0 áll, akkor 0-t kapunk az $\mathbf{u} = (\alpha^* \alpha - \mathbf{I}) \mathbf{v}$ esetben is. A skalárszorzat pozitív definitisége alapján ez azt jelenti, hogy $(\alpha^* \alpha - \mathbf{I}) \mathbf{v} = \mathbf{0}$ teljesül minden vektorra, azaz $\alpha^* \alpha - \mathbf{I} = \mathbf{0}$, vagyis α unitér, illetve ortogonális. ■

Ahhoz, hogy egy transzformáció unitér, illetve ortogonális legyen, ennél kevesebb is elég:

7.16. Tétel. *Az euklideszi tér egy transzformációja akkor és csak akkor unitér, illetve ortogonális, ha távolságtartó.*

Bizonyítás. A 7.15. Tétel alapján természetesen elég annak a kimutatása, hogy a távolságtartás ekvivalens a skalárszorlattartással. Ezt úgy fogalmazhatjuk meg, hogy az $(\alpha \mathbf{u}; \alpha \mathbf{v}) - (\mathbf{u}; \mathbf{v})$ bilineáris függvény pontosan akkor azonosan 0, ha az $(\alpha \mathbf{x}; \alpha \mathbf{x}) - (\mathbf{x}; \mathbf{x})$ kvadratikus alak is az. Ez viszont azonnal következik az 5.12., illetve az 5.21. Tételből, mert a vizsgált bilineáris függvény szimmetrikus, illetve Hermite-féle. ■

Kiegészítés. *Unitér, illetve ortogonális transzformációk szorzata és inverze is unitér, illetve ortogonális.*

Bizonyítás. A szorzatra vonatkozó állítást a következőképpen láthatjuk be formálisan, az $(\alpha\beta)^* = \beta^* \alpha^*$ felhasználásával:

$$(\alpha\beta)(\alpha\beta)^* = (\alpha\beta)\beta^* \alpha^* = \alpha(\beta\beta^*)\alpha^* = \alpha \iota \alpha^* = \alpha \alpha^* = \iota.$$

A 7.16. Tétel viszont nyújt egy fogalmi bizonyítást; eszerint ugyanis a két transzformáció mindegyike azzal jellemezhető, hogy távolságtartók. Ekkor viszont a szorzatuk is távolságtartó, ami éppen a kívánt tulajdonságot jelenti. Ugyanez a megfontolás bizonyítja az inverz transzformációra vonatkozó állítást is. ■

Itt is fontos tudni, hogy miképpen ismerhető fel egy unitér, illetve egy ortogonális transzformáció ortonormált bázisban felírt mátrixáról.

7.17. Tétel. *Egy unitér, illetve egy ortogonális transzformáció bármely ortonormált bázisban felírt mátrixában a mátrix i -edik oszlopa és j -edik oszlopa adjungáltjának, illetve transzponáltjának a mátrixszorzata $\delta_{i,j}$. Az ilyen mátrixokat unitér, illetve ortogonális mátrixoknak nevezik. Fordítva is igaz, minden unitér, illetve ortogonális mátrix előáll, mint egy unitér, illetve ortogonális transzformációnak ortonormált bázisban felírt mátrixa.*

Bármely unitér, illetve ortogonális mátrixban a sorokra is hasonló feltétel teljesül, mint az oszlopokra.

Bizonyítás. Tekintettel arra, hogy valós esetben az adjungált a transzponálttal egyezik meg, ezért elég az unitér esettel foglalkozni.

Legyen α unitér transzformáció és $\mathbf{e}_1, \dots, \mathbf{e}_n$ egy ortonormált bázis az \mathcal{U} euklideszi téren. A 7.15. Tétel szerint $\alpha \mathbf{e}_1, \dots, \alpha \mathbf{e}_n$ is ortonormált (ezért bázis). Ez azt jelenti, hogy $(\alpha \mathbf{e}_i; \alpha \mathbf{e}_j) = \delta_{i,j}$. A transzformáció mátrixát ebben a bázisban felírva, az teljesíti a mondott feltételeket.

Tegyük most fel, hogy a mátrix eleget tesz a kirótt feltételeknek és tekintsük az általa meghatározott α transzformációt. Vegyük azt a skalárszorzatot, amely ehhez az $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázishoz tartozik. Ekkor ezek egy ortonormált rendszert alkotnak. A mátrixra kirótt feltételek alapján $(\alpha \mathbf{e}_i; \alpha \mathbf{e}_j) = (\mathbf{e}_i; \mathbf{e}_j)$, hiszen $\alpha \mathbf{e}_i$ mátrixa éppen az adott mátrix i -edik oszlopa. Ebből — a 7.15. Tétel bizonyításában látottakhoz hasonlóan — azt kapjuk, hogy

$(\mathbf{e}_i; (\alpha^* \alpha - \iota) \mathbf{e}_j) = 0$. A bilinearitást felhasználva ebből $(\mathbf{u}; (\alpha^* \alpha - \iota) \mathbf{v}) = 0$ következik minden \mathbf{u}, \mathbf{v} vektorpárra, így csak $\alpha^* \alpha - \iota$ lehet, vagyis α valóban unitér.

A sorokra vonatkozó állítás azzal ekvivalens, hogy α^* is unitér. Ez viszont pontosan azt jelenti, hogy α^{-1} unitér, hiszen $\alpha^* = \alpha^{-1}$. Ez viszont azért igaz, mert egy skalárszor-zattartó leképezés inverze is skalárszor-zattartó. ■

Megjegyzések

1. A tételben szereplő i -edik oszlop és j -edik oszlop konjugáltjának a szorzata az egyik sorrendben egy mátrix, a másik sorrendben egy szám. Természetesen ebben a sorrendben összeszorozva adódhat csak $\delta_{i,j}$. Illetve, ha a másik sorrendben is egy számot kapunk, akkor a mátrixnak csak egyetlen eleme van, amikor az egész eset triviális.

2. A komplex esetben vigyázni kell arra, hogy két oszlop skalárszor-zatánál az egyiknek a konjugáltja szerepel. Ezt számolás közben gyakran elfelejtik, és ezáltal hamis eredményhez jutnak.

3. A 7.17. Tételből azonnal következik, hogy egy mátrix pontosan akkor „oszloportogonális”, ha „sorortogonális”. Ennek a fenti tétel nélküli bizonyítása még kétszer kettes mátrixok esetében is sok számolást igényel. Nagyobb mátrixoknál valószínűleg kilátástalan. □

7.18. Tétel. *Unitér, illetve ortogonális transzformáció sajátértékei egységnyi abszolút értékűek.*

Bizonyítás. Ha c sajátértéke az α unitér vagy ortogonális transzformációnak, akkor tekintsünk egy ehhez tartozó \mathbf{u} sajátvektort. A távolságtartás következményeként:

$$(\mathbf{u}; \mathbf{u}) = (\alpha \mathbf{u}; \alpha \mathbf{u}) = (c \mathbf{u}; c \mathbf{u}) = \bar{c} \cdot c (\mathbf{u}; \mathbf{u}).$$

Mivel $\mathbf{u} \neq \mathbf{o}$, hiszen sajátvektor, ezért $(\mathbf{u}; \mathbf{u})$ -val egyszerűsíthetünk, és a $\bar{c} \cdot c = 1$ eredményhez jutunk. ■

A továbbiakban külön kell tárgyalni az unitér és az ortogonális transzformációkat, mert az eredmények különböző típusúak.

7.19. Tétel. *Egy transzformáció akkor és csak akkor unitér, ha található hozzá olyan ortonormált bázis, amelyben mátrixa diagonális, és a diagonális elemei egységnyi abszolút értékűek.*

Bizonyítás. Mivel unitér transzformáció normális, ezért alkalmazható a 7.10. Tétel. A kapott diagonális mátrix elemei a transzformáció sajátértékei, amelyek a 7.18. Tétel szerint egységnyi abszolút értékűek.

A megfordítás triviális, hiszen a leírt tulajdonságú mátrix olyan transzformáció mátrixa, amely egy adott ortonormált bázist ortonormált bázisba visz, tehát unitér. ■

A valós euklideszi térben nem csak a normális, hanem speciálisan az ortogonális transzformáció mátrixa sem lesz alkalmas ortonormált bázisban diagonális. Lehetséges ugyanis, hogy a transzformációnak létezik kétdimenziós minimális invariáns altére. Mindezekelőtt ezeket a speciális ortogonális transzformációkat írjuk le.

7.20. Tétel. *Ha a kétdimenziós valós euklideszi tér egy α ortogonális transzformációjának nincs valódi invariáns altére, akkor létezik olyan φ szög, amelyre alkalmas ortonormált bázisban mátrixa*

$$\begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

alakú. Az ilyen transzformációkat φ szöggel pozitív irányba való elforgatásnak nevezzük. Az elforgatás ortogonális transzformáció, amelynek nincs valódi invariáns altere, kivéve, ha $\varphi = 180^\circ$ egész számú többszöröse.

Minden ilyen mátrixhoz ortogonális transzformáció tartozik.

Bizonyítás. Legyen α mátrixa egy ortonormált bázisban $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$. A 7.17. Tétel alapján fennállnak az

$$a^2 + b^2 = 1, \quad ac + bd = 0, \quad c^2 + d^2 = 1$$

összefüggések. Ha $b = 0$, illetve $c = 0$, akkor az első, illetve a második bázisvektor α -nak sajátvektora, tehát a tér nem minimális invariáns altér. Egyébként létezik olyan t valós szám, amelyre $c = tb$, mert $b \neq 0$. Ezt a felírt három egyenlőség közül a középsőbe beírva a $0 = atb + bd = b(at + d)$ egyenlőséget nyerjük. $b \neq 0$ miatt ebből azt kapjuk, hogy $d = -ta$. Így a másik két felírt egyenlőségből

$$1 = c^2 + d^2 = t^2 b^2 + t^2 a^2 = t^2 (b^2 + a^2) = t^2$$

következik. Eszerint vagy $t = 1$, vagy $t = -1$. Az első esetben a transzformáció mátrixa szimmetrikus, ezért α -nak van valódi invariáns altere. A második esetben $c = -b$ és $d = a$. Az $a^2 + b^2 = 1$ feltétel alapján létezik olyan $0 \leq \varphi < 360^\circ$, amelyre $a = \cos \varphi$ és $b = \sin \varphi$ (persze $\varphi = 180^\circ$ esetében ismét szimmetrikus transzformációt kapunk).

A „megfordítás” triviális. ■

Megjegyzés. Könnyen belátható, hogy a fenti transzformációnak minden bázisban ilyen alakú a mátrixa. □

Következmény. A valós euklideszi térben minden ortogonális transzformációhoz található olyan ortonormált bázis, amelyben legfeljebb kétszer kettes diagonális blokkokon kívül minden elem 0. A kétszer kettes diagonális blokkok olyan alakúak, mint amilyeneket a tételben leírtunk; az egyszer egyes blokkok eleme pedig vagy +1, vagy -1.

Bizonyítás. A normális transzformációkra vonatkozó 7.9. Tétel szerint az ortogonális transzformációkhoz létezik olyan ortonormált bázis, amelyben mátrixa „blokkdiagonális” és a blokkok mérete legfeljebb kétszer kettes. Ha a megfelelő altér a transzformációnak minimális invariáns altere, akkor a kétszer kettes diagonális blokk valóban olyan alakú, mint amelyet a tételben leírtunk. Ha a blokk egyszer egyes, akkor a blokk eleme természetesen sajátérték; a 7.18. Tétel szerint tehát valóban egységnyi abszolút értékű, s mivel valós szám, ezért vagy +1, vagy -1.

Világos, hogy minden ilyen mátrixhoz létezik olyan ortogonális transzformáció, amelynek alkalmas ortonormált bázisban pontosan ez a mátrixa. ■

A fenti leírás az ortogonális transzformációk számolásra alkalmas formáját mutatja be. Az alábbiakban ezeknek egy geometriai jellegű formáját írjuk le.

7.5. Definíció. Az \mathcal{U} euklideszi tér egy α transzformációját a \mathcal{V} altérre való tükrözésnek nevezzük, ha bármely $\mathbf{v} \in \mathcal{V}$ elemre $\alpha \mathbf{v} = \mathbf{v}$ és minden $\mathbf{w} \in \mathcal{W} = \mathcal{V}^\perp$ elemre $\alpha \mathbf{w} = -\mathbf{w}$. ■

7.21. Tétel. Minden tükrözés ortogonális transzformáció. Az \mathcal{U} egész térre való tükrözés az identitás. A $\{\mathbf{o}\}$ altérre való tükrözés középpontos tükrözés.

Az α transzformáció akkor és csak akkor tükrözés, ha szimmetrikus és ortogonális.

Bizonyítás. Tekintsünk egy olyan ortonormált bázist, amelynek minden eleme vagy \mathcal{V} -ben, vagy \mathcal{W} -ben van. Ebben a bázisban α mátrixa diagonális, és a diagonális minden eleme 1 abszolút értékű. α tehát ortogonális. Mivel a kapott mátrix szimmetrikus, ezért a transzformáció szimmetrikus.

Legyen most α szimmetrikus és ortogonális. Mivel szimmetrikus, tehát alkalmas ortonormált bázisban mátrixa diagonális. Mivel ortogonális, ezért a diagonális elemei egységnyi abszolút értékűek, azaz mind $+1$ vagy -1 . Ez a transzformáció tehát azon bázisvektorok által kifeszített altérre való tükrözés, amelyekhez tartozó oszlopban a diagonális eleme $+1$.

Az egész térre való tükrözés triviálisan az identitás. A $\{\mathbf{o}\}$ altérre való tükrözésnél minden vektor a negatívjába megy; ezt tehát joggal nevezhetjük középpontos tükrözésnek. ■

7.22. Tétel. A valós euklideszi tér minden ortogonális transzformációja vagy tükrözés, vagy két tükrözés szorzata.

Bizonyítás. Mindenekelőtt jegyezzük meg, hogy egy tükrözés is és ezért két tükrözés szorzata is ortogonális. A tétel állítása tehát úgy is fogalmazható, hogy csak ezen a módon nyerhetünk ortogonális transzformációt.

A tételt a tér dimenziójára vonatkozó teljes indukcióval bizonyítjuk. Ha a tér egydimenziós, akkor az α ortogonális transzformációnak van sajátvektora, amelyhez tartozó sajátérték $+1$ vagy -1 ; így α megfelelően vagy az identitás, vagy középpontos tükrözés.

Ha a tér kétdimenziós, akkor két esetet különböztetünk meg. Azt az esetet, amikor α -nak van sajátvektora, az indukciós lépésnél fogjuk tárgyalni. Egyébként van a térben olyan \mathbf{u} vektor, amelyre $\alpha\mathbf{u}$ nem párhuzamos \mathbf{u} -val, azaz e vektorok lineárisan függetlenek. Nyilván feltehető, hogy \mathbf{u} abszolút értéke 1. Mivel a tér kétdimenziós, ezért a lineárisan független $\{\mathbf{u}, \alpha\mathbf{u}\}$ rendszer a tér egy bázisa (persze nem feltétlenül ortonormált). Tekintsük az $\mathbf{e} = \mathbf{u} + \alpha\mathbf{u}$ és az $\mathbf{f} = \mathbf{u} - \alpha\mathbf{u}$ vektorokat. (Érdemes a bizonyítás elvégzése előtt megnézni, hogy ha α forgatás, akkor mit tudhatunk e két vektorról!) A skalárszorzat bilinearitását és szimmetriáját felhasználva:

$(\mathbf{u} + \alpha\mathbf{u}; \mathbf{u} - \alpha\mathbf{u}) = (\mathbf{u}; \mathbf{u}) + (\alpha\mathbf{u}; \mathbf{u}) - (\mathbf{u}; \alpha\mathbf{u}) - (\alpha\mathbf{u}; \alpha\mathbf{u}) = 1 + (\alpha\mathbf{u}; \mathbf{u}) - (\alpha\mathbf{u}; \mathbf{u}) - 1 = 0$,
azaz \mathbf{e} és \mathbf{f} ortogonálisak.

Definiáljuk a τ transzformációt az $\{\mathbf{u}, \alpha\mathbf{u}\}$ bázison úgy, hogy $\tau(\mathbf{u}) = \alpha\mathbf{u}$ és $\tau(\alpha\mathbf{u}) = \mathbf{u}$. Ekkor $\tau(\mathbf{e}) = \tau(\mathbf{u} + \alpha\mathbf{u}) = \alpha\mathbf{u} + \mathbf{u} = \mathbf{e}$ és $\tau\mathbf{f} = \tau(\mathbf{u} - \alpha\mathbf{u}) = \alpha\mathbf{u} - \mathbf{u} = -\mathbf{f}$. Eszerint τ az \mathbf{e} generálta altérre való tükrözés. Ekkor viszont a $\sigma = \tau\alpha$ transzformáció is ortogonális, mert két ortogonálisnak a szorzata. Erre $\sigma(\mathbf{u}) = \tau\alpha\mathbf{u} = \mathbf{u}$, vagyis \mathbf{u} a σ sajátvektora, tehát σ is tükrözés. Tekintettel arra, hogy minden tükrözésnek a négyzete az identitás, ezért $\alpha = \tau\tau\alpha = \tau\sigma$, vagyis α előáll két tükrözés szorzataként.

Ha a tér dimenziója nagyobb, mint 2, vagy a tér nem minimális invariáns altér, akkor a tér felbomlik α -nak két \mathcal{U}_1 és \mathcal{U}_2 egymásra merőleges valódi invariáns altere direkt összegére. Az indukciós feltevés miatt ezekre már igaz az állítás. Legyen α_i az α megszorítása az \mathcal{U}_i altérre. Mivel \mathcal{U}_1 és \mathcal{U}_2 invariáns alterek, ezért α_i ortogonális transzformáció az \mathcal{U}_i

altéren. Mivel valódi alterekről van szó, az indukciós feltétel alapján léteznek olyan σ_1, τ_1 , illetve σ_2, τ_2 tükrözések az \mathcal{U}_1 , illetve \mathcal{U}_2 altéren, amelyekre $\alpha_i = \tau_i \sigma_i$ ($i \in \{1, 2\}$). A 3.14. Tétel utáni kiegészítés szerint a szorzat felcserélhető a leképezések direkt összegével:

$$(\sigma_1 \oplus \sigma_2) \cdot (\tau_1 \oplus \tau_2) = (\sigma_1 \cdot \tau_1) \oplus (\sigma_2 \cdot \tau_2) = \alpha_1 \oplus \alpha_2 = \alpha.$$

A tétel bizonyításához ezek után már csak azt kell kimutatni, hogy $\sigma = \sigma_1 \oplus \sigma_2$ (és $\tau = \tau_1 \oplus \tau_2$) tükrözés. A tükrözés definíciója szerint van olyan $\mathcal{V}_1 \leq \mathcal{U}_1$, illetve $\mathcal{V}_2 \leq \mathcal{U}_2$, hogy az \mathcal{U}_1 -beli, illetve \mathcal{U}_2 -beli $\mathcal{W}_1 = \mathcal{V}_1^\perp$, illetve $\mathcal{W}_2 = \mathcal{V}_2^\perp$ alterekkel σ_i identitás \mathcal{V}_i -n és középpontos tükrözés \mathcal{W}_i -n ($i \in \{1, 2\}$). A szereplő négy altér páronként ortogonális, így $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$ és $\mathcal{W} = \mathcal{W}_1 \oplus \mathcal{W}_2$ is ortogonálisak és direkt összegük az egész tér. A leképezések direkt összegének definíciója szerint σ identitás a \mathcal{V} altéren és középpontos tükrözés a \mathcal{W} altéren, azaz valóban tükrözés. A τ -ra vonatkozó állítás hasonlóan látható be. ■

Feladatok

1. Láttuk, hogy egy véges dimenziós térben egy \mathbf{U} bázis egy $\Phi_{\mathbf{U}}$ bijekciót hoz létre a bilineáris függvények és a lineáris transzformációk között, megfeleltetve egymásnak azokat, amelyeknek a mátrixa e bázisban megegyezik: $\Phi_{\mathbf{U}} : \mathbf{A} \mapsto \alpha$ akkor és csak akkor, ha az \mathbf{U} bázisban $[\mathbf{A}] = [\alpha]$. Bizonyítsuk be, hogy $\Phi_{\mathbf{U}} = \Phi_{\mathbf{V}}$ pontosan akkor teljesül, ha \mathbf{U} és \mathbf{V} ugyanazt a skalárszorzatot definiálják.

2. Az n -dimenziós térben hipersíknak nevezik az $(n - 1)$ -dimenziós altereket. Nevezzük a hipersíkokra való tükrözést hipertükrözésnek. Bizonyítsuk be, hogy a valós számtest feletti n -dimenziós tér minden ortogonális transzformációja előáll mint legfeljebb $n + 1$ hipertükrözés szorzata.

3. Ha τ tükrözés az euklideszi téren, akkor $\tau^2 = \text{id}$. Következik-e ebből az utóbbi feltételből, hogy τ tükrözés?

4. Bizonyítsuk be, hogy bármely normális transzformáció előáll mint egy önadjungált és egy unitér, illetve egy szimmetrikus és egy ortogonális transzformáció szorzata. Mutassuk meg, hogy egy önadjungált és egy unitér, illetve egy szimmetrikus és egy ortogonális transzformáció szorzata nem feltétlen normális. Milyen feltétel mellett lesz az?

5. Mutassuk meg, hogy az alábbi mátrixok ortogonálisak, és írjuk fel diagonális alakjukat és minimálpolinomjukat:

$$\begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

6. Van-e a \mathbb{Q} feletti háromdimenziós térben olyan ortogonális transzformáció, amelynek nincs sajátvektora?

7. Van-e a \mathbb{Q} feletti nyolcdimenziós térben olyan ortogonális transzformáció, amelyiknek nincs valódi invariáns altére?

8. Bizonyítsuk be, hogy a $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ és az $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$ mátrixok hasonlóak.

9. Tekintsük a kétdimenziós \mathbb{C} feletti vektortér következő unitér lineáris transzformációit: $1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $I = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$, $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$. Bizonyítsuk be, hogy az ezek generálta altér a műveletekre egy nemkommutatív testet alkot. Határozzuk meg az e testhez tartozó mátrixok általános alakját.

10. Legyenek az $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ mátrix elemei egy R főideálgűrűből, és tegyük fel, hogy $t = ax + by$ az a és b „legnagyobb közös osztója”. Bizonyítsuk be, hogy van olyan B , az R felett invertálható mátrix, amelyre $AB = \begin{bmatrix} t & 0 \\ * & * \end{bmatrix}$ alakú. Általánosítsuk a feladatot.

11. Mutassuk meg, hogy egy \mathbb{Q} feletti kétszer kettes ortogonális mátrix minden sorában pontosan egy 0 áll, s a többi elem abszolút értéke 1.

12. Mutassuk meg, hogy van olyan négyszer négyes T ortogonális mátrix, amelyben minden elem abszolút értéke $\frac{1}{2}$.

13. Mutassuk meg, hogy ha A ortogonális mátrix, és az előző feladat T mátrixában minden elem helyébe $\frac{1}{2}A$ -t teszünk, akkor ugyancsak ortogonális mátrixot kapunk.

14. Legyen $\begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$ a valós kétdimenziós tér egy ortogonális transzformációjának a mátrixa adott ortonormált bázisban. Mennyiben változhat meg a mátrix, ha egy másik ortonormált bázist veszünk fel?

15. Legyen A a valós kétdimenziós euklideszi tér sajátvektor nélküli ortogonális transzformációjának valamely ortonormált bázisban felírt mátrixa. Tekintsük A -t egy unitér transzformáció mátrixának, és írjuk fel diagonális alakját.

5. Kvadratikus alakok az euklideszi térben

A kvadratikus alakokat a valós térben egyértelműen meghatározhattuk egy szimmetrikus bilineáris függvénnyel. A komplex tér esetében nem volt szükség speciális bilineáris függvényre, de a definitésről csak Hermite-féle függvények esetében tudtunk beszélni. Az euklideszi terekben mindkét függvénytípus egyértelműen meghatároz egy-egy speciális fajta lineáris transzformációt, nevezetesen egy szimmetrikus, illetve egy önadjungált transzformációt. Ezek segítségével két igen hasznos információt nyerhetünk az euklideszi terekben megadott kvadratikus alakokról. Mindenekelőtt egy fogalomra lesz szükségünk:

7.6. Definíció. Az n -dimenziós euklideszi térben azoknak az \mathbf{x} vektoroknak a halmazát, amelyeknek az abszolút értéke 1, az n -dimenziós tér egységsgömbjének nevezzük. ■

A fenti definíció független attól, hogy a valós vagy a komplex euklideszi térről beszélünk. A további eredmények is teljesen hasonlóak mind a valós, mind a komplex tér esetében. Az alábbi tételek kimondásában is és bizonyításában is azt tesszük fel, hogy a komplex számtest feletti vektorteret vizsgálunk. A valós számtest esetében ugyanígy történik a

bizonyítás; de azt is mondhatjuk, hogy az alábbi bizonyítások a valós számtest esetében „a valós számtest feletti tételek bizonyításává válnak”. Először a sajátértékek úgynevezett *extremális tulajdonságát* bizonyítjuk be:

7.23. Tétel. *Legyen α a komplex euklideszi tér egy önadjungált lineáris transzformációja. Ekkor az $(\mathbf{x}; \alpha\mathbf{x})$ kvadratikus alak az egész téren valós értékű, és az egységömbre megszorítva ott felveszi a minimumát. A minimum helye α egy sajátvektora, és a minimum a hozzá tartozó sajátérték.*

Bizonyítás. A 7.13. Tétel szerint az α önadjungált transzformációhoz található olyan $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ ortonormált bázis, amelynek elemei éppen a transzformáció sajátvektorai. Az (orto)normáltság szerint a bázis elemei az egységömbön vannak. Mivel sajátvektorok, ezért vannak olyan $\lambda_1, \dots, \lambda_n$ számok, amelyek a megfelelő sajátértékek (mint említettük, a sajátértékeket általában a λ betűvel jelölik). Ez azt jelenti, hogy $\alpha(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$ ($1 \leq i \leq n$). α -nak az \mathbf{E} bázisban felírt mátrixa diagonális, és a diagonális elemei éppen a λ_i számok, amelyek valósak, hiszen α önadjungált. A báziselemek sorrendjének esetleges cseréjével elérhető, hogy $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ legyen.

Ekkor a $\beta = \alpha - \lambda_1 \cdot \iota$ transzformáció ugyancsak önadjungált, mert két önadjungált transzformáció különbsége. Az is igaz, hogy β mátrixa a fenti \mathbf{E} bázisban szintén diagonális; továbbá a mátrix diagonálisának i -edik eleme $\lambda_i - \lambda_1$, ami nemnegatív. A kvadratikus alakok tehetetlenségi tétele alapján β pozitív szemidefinit. Ezért az egységömb bármely \mathbf{x} vektorára:

$$(\mathbf{x}; \alpha(\mathbf{x})) = (\mathbf{x}; (\beta + \lambda_1 \cdot \iota)(\mathbf{x})) = (\mathbf{x}; \beta(\mathbf{x})) + \lambda_1(\mathbf{x}; \mathbf{x}) \geq 0 + \lambda_1 \cdot 1 = \lambda_1.$$

Ezzel beláttuk, hogy λ_1 az $(\mathbf{x}; \alpha\mathbf{x})$ kvadratikus alaknak az egységömbön felvett minimuma. Ezt a minimumot a vizsgált kvadratikus alak valóban felveszi (például az \mathbf{e}_1 helyen, de esetleg máshol is). A felvett érték pontosan λ_1 , ami valóban sajátértéke az α transzformációnak. ■

7.24. Tétel (Főtengelytétel). *Az euklideszi térben bármely Hermite-féle bilineáris függvényhez tartozó kvadratikus alakhoz található olyan ortonormált bázis, amelyben a kvadratikus alak négyzetösszegé válik.*

Bizonyítás. A tekintett \mathbf{A} Hermite-féle bilineáris alakhoz a 7.11. Tétel szerint egyértelműen található olyan α önadjungált transzformáció, amelyre $\mathbf{A}(\mathbf{u}, \mathbf{v}) = (\mathbf{u}; \alpha(\mathbf{v}))$. A 7.13. Tétel szerint létezik olyan $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ ortonormált bázis, amelyben α mátrixa diagonális lesz. Mivel a bázis ortonormált, ezért benne $[\mathbf{A}] = [\alpha]$. Így az adott kvadratikus alak mátrixa diagonális; a bázisban \mathbf{A} négyzetösszegé transzformálódik. ■

Megjegyzés. A 7.23. és a 7.24. tételek alapján tetszőleges kvadratikus alaknak elvileg egyszerűen megtalálhatjuk a főtengelyeit. Először meghatározzuk a kvadratikus alaknak mint másodfokú függvénynek a minimumát, azzal a feltétellel, hogy a koordináták négyzetösszege 1 (feltételes szélsőérték). Ez nem más, mint a vizsgált kvadratikus alak legkisebb sajátértéke. A szemléletesség kedvéért feltesszük, hogy ez pozitív. Ekkor megkapjuk, hogy mi annak a testnek a legnagyobb(!) vastagsága, amit a megadott másodrendű felület határol (ez a legkisebb sajátértékhez tartozik).

Ezután meghatározzuk a kapott c sajátértékhez tartozó \mathbf{u} sajátvektort (illetve ezek egyikét). Ha A a kvadratikus alak mátrixa (eleve abban az oszlopvektorok meghatározta bázisban dolgozunk, amelynek elemeiben egyetlen 1 szerepel, a többi 0), akkor az \mathbf{u} sajátvektor koordinátáit az $(A - c \cdot I)(\mathbf{u}) = \mathbf{0}$ (\mathbf{u} koordinátaiban) lineáris egyenletrendszer szolgáltatja. Ez lesz egyúttal az egyik tengelyirány. (Lehetséges, hogy a test olyan, mint egy hosszabb tengelye körül megforgatott ellipszis. Ekkor több ilyen irány van, ebben az esetben választunk közülük páronként merőlegeseket. Ez a választás nem „kötelező”, de lehetséges.)

Ha bizonyos sajátvektorokat már megkaptunk, akkor a kvadratikus alakot megszorítjuk az ezekre ortogonális vektorokból álló altérre, és ott folytatjuk tovább az eljárást.

A fentiek alapján a főtengelek meghatározása nem okoz elvi nehézséget. Egyetlen dolog szorul magyarázatra; nevezetesen az, hogy miért abban az irányban a legvastagabb a test, amelyikhez a legkisebb sajátérték tartozik. Síkgörbékét tekintünk, azaz három homogén koordináta adott. A négyzetösszeggé való transzformálás után a mátrix diagonális lesz. Tekintettel arra, hogy a harmadik

koordinátát 1-nek kell választani, célszerű azt feltenni, hogy a mátrix $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & -1 \end{bmatrix}$ alakú, vagyis a

szóban forgó kifejezés $ax^2 + by^2 = 1$. Az is feltehető, hogy $a \geq b$. A sajátvektorok transzponáltjai: $\mathbf{e} = [1, 0, 0]$, $\mathbf{f} = [0, 1, 0]$, $\mathbf{g} = [0, 0, 1]$. A sajátértékek rendre $a, b, -1$. A kapott görbénél a tengelyek hossza: \mathbf{e} irányban $\frac{1}{\sqrt{a}}$ és \mathbf{f} irányban $\frac{1}{\sqrt{b}}$ — mármint, ha $b > 0$. Látható, hogy kisebb sajátértékhez valóban nagyobb vastagság járul. Annak a megállapítását, hogy a és b előjelétől $(+, 0, -)$ miképpen függ a görbe jellege, az olvasóra bízunk. \square

Feladatok

1. Határozzuk meg az alábbi kvadratikus alakok kvadratikus karakterét és főtengeleseit:

- a) xy ; b) $x^2 + 3y^2$; c) $x^2 - 3y^2$; d) $x^2 - 2xy + 3y^2$;
 e) $x^2 - 4xy + 3y^2$; f) $x^2 + 4xy + 8xz + 6y^2 - 4yz + 18z^2$;
 g) $x^2 + 6xy + 2xz + 4y^2 - 2yz + 2z^2$; h) $x^2 + 4xy + 6xz + 2xz + 2z^2 + 2zw + 4w^2$;
 i) $2xy + 8xz + 2xw + 2yz + 4yw + 2zw + 3w^2$; j) $2xy + 2xz + 2yz$; k) $2xy + 2yz + 2zw$.

2. Legyen U az n -dimenziós tér egy bázisa. Igaz-e az, hogy ha egy kvadratikus alak az U bizonyos „speciális” elemei generálta bármely altéren pozitív definit, akkor az egész téren az, ha e speciális rendszer: a) a kételeműek; b) az $(n-1)$ -eleműek; c) a legfeljebb $(n-1)$ -eleműek?

3. Bizonyítsuk be, hogy ha egy vektortérben adott egy pozitív definit és egy tetszőleges (szimmetrikus, illetve Hermite-féle) bilineáris függvény, akkor található olyan bázis, amelyben mindkét függvény négyzetösszeggé válik.

NYOLCADIK FEJEZET

A KARAKTERISZTIKUS POLINOM

1. A determináns

Négyzetes mátrixok determinánsával már az első részben, a mátrixok tárgyalásánál találkoztunk. A továbbiakban a determináns fogalmának a lineáris algebrai háttérét szeretnénk megvilágítani.

Emlékeztetünk arra, hogy egy négyzetes mátrix determinánsát meghatározhatjuk úgy, hogy csak a soraival, illetve csak az oszlopaival végzünk műveleteket. Számunkra célszerűbb az oszlopokat vizsgálni. Egy négyzetes mátrix determinánsát meg tudjuk határozni az alábbi „szabályok” felhasználásával:

- (1) Ha egy mátrix valamelyik oszlopát egy c skalárral szorozzuk, akkor a kapott mátrix determinánsa az eredeti mátrix determinánsának a c -szerese lesz.
- (2) Ha az A és a B mátrixban az i -edik oszlop kivételével a többi megegyezik, és a C mátrix i -edik oszlopában e két oszlop összege szerepel, s a többi oszlop ugyanaz, mint A megfelelő oszlopa, akkor $\det(C) = \det(A) + \det(B)$.
- (3) Ha egy négyzetes mátrix két oszlopát felcseréljük, akkor az új mátrix determinánsa az eredeti determináns negatívja lesz.
- (4) Ha egy négyzetes mátrix fődiagonálisának minden eleme 1, és a többi elem 0, akkor determinánsa 1.

(Megjegyezzük, hogy a determináns tárgyalásánál nem ezeket a tulajdonságokat mutattuk ki, de bebizonyítottuk, hogy ezek ekvivalensek az eredetileg adott tulajdonságokkal.)

Vegyük szemügyre ezeket a tulajdonságokat mint az oszlopvektorok tulajdonságát. Ha a mátrix sorainak száma n , akkor az n -dimenziós tér vektorairól van szó. Mivel az oszlopok száma is n , ezért a determináns úgy tekinthető, mint az n -dimenziós téren értelmezett n -változós függvény. Nézzük meg a fenti négy tulajdonság lineáris algebrai értelmét:

Az (1) és (2) tulajdonság azt mondja ki, hogy ha a mátrix $n - 1$ oszlopát rögzítjük, akkor a determináns a fennmaradó (vektor)változóban lineáris. A bilinearitáshoz hasonlóan erre úgy fogunk utalni, hogy a determináns vektorváltozóiban *multilineáris*, pontosabban *n-lineáris* függvény. A (3) tulajdonság egyike azoknak, amelyek az antiszimmetrikus bilineáris függvényeket jellemezték. Ennek alapján a determinánst *antiszimmetrikus n-lineáris*

függvénynek nevezzük. A (4) tulajdonságból, noha nagyon fontos, jelenleg csak annyit érdemes megjegyezni, hogy a függvény nemtriviális, azaz nem minden vektor- n -esre veszi fel a 0 értéket. (Az azonosan 0 függvény ugyanis szintén kielégíti az első három „axiómát”).

Érdekes megnézni ezeknek a tulajdonságoknak a geometriai jelentését is. Természetesen csak a valós számtest feletti két-, illetve háromdimenziós térrel foglalkozhatunk. A síkban két vektor egy paralelogrammát feszít ki. Világos, hogy ha a két vektor bármelyikét c -szeresre nyújtjuk, akkor a paralelogramma területe is c -szeresére növekszik (egyelőre maradjunk pozitív számoknál). A (2) feltétel a területnek azt a tulajdonságát mondja ki, hogy egyik oldaluknál „összeragasztott” paralelogrammák területe összeadódik. Nehezebb értelmezni a harmadik feltételt. Tegyük fel, hogy a két oszlopvektor \mathbf{a} és \mathbf{b} ebben a sorrendben, és mondjuk pozitív forgatás viszi \mathbf{a} -t a \mathbf{b} irányába és állásába. Tegyük fel azt is, hogy ekkor a determináns is pozitív. Ha most megcseréljük a két vektor sorrendjét, akkor az első vektort negatív forgatás viszi a második vektor irányába és állásába. Mint tudjuk, ekkor a determináns is előjelet vált. Ezt a helyzetet úgy is felfoghatjuk, hogy most a síkot a „másik oldaláról” nézzük. Ha a két vektor által kifeszített háromszöget tekintjük, akkor a „közös csúcs \rightarrow első vektor vége \rightarrow második vektor vége” ponthármában a forgásirány változik meg. Mindez azt mutatja, hogy érdemes előjeles területet tekinteni. Valójában persze nem tudjuk, hogy a determináns minek a területét adja meg, a paralelogrammáét vagy a háromszögét. Ami biztos, hogy ezekkel arányos. (Egyébként a paralelogrammáét, de ennek a bizonyítása nem az algebrának, hanem a geometriának a feladata.) Hasonló a helyzet a térben, ott a három vektor által kifeszített paralelepipedon előjeles térfogatát adja a determináns. A determináns tehát az n -dimenziós térben megadja egy test térfogatának a mértékét.

8.1. Definíció. A K test felett n -dimenziós \mathcal{U} téren értelmezett

$$\mu : \overbrace{\mathcal{U} \times \cdots \times \mathcal{U}}^{n\text{-szer}} \rightarrow K$$

n -változós függvényt n -lineárisnak nevezünk, ha bármely $n - 1$ változóját rögzítve a n -edikben (homogén) lineáris.

Egy, az n -dimenziós téren értelmezett n -lineáris μ függvényt mértéknek nevezzük, ha bármely $n - 2$ változóját rögzítve a megmaradó kettőben antiszimmetrikus. ■

8.1. Tétel. *Lineárisan összefüggő vektorokra bármely mérték 0.*

Bizonyítás. Tegyük fel először, hogy a vektorok között szerepel a \mathbf{o} is. Ekkor a $0 \cdot \mathbf{o} = \mathbf{o}$ alapján 0-t kiemelve a mérték 0-szorosra változik. Másrészt nem változik meg, hiszen a vektorok nem változtak meg. Így a mérték valóban 0. Ha mármost a szereplő vektorok lineárisan összefüggenek, akkor van olyan vektor, amelyhez a többiek megfelelő skalárszorosit hozzáadva a \mathbf{o} vektort kapjuk; tehát a mérték 0. Mint az antiszimmetrikus bilineáris függvényekre bizonyítottuk, az eljárás során a függvényérték nem változik meg; így az eredeti mértéke is 0. ■

8.2. Tétel. *Ha az \mathcal{U} vektortéren értelmezett μ mérték egy adott bázison 0, akkor azonosan 0.*

Bizonyítás. Legyen $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ a szóban forgó bázis. Az antiszimmetria miatt az $\{1, 2, \dots, n\}$ halmaz bármely σ permutációjára $\mu(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}) = \pm \mu(\mathbf{e}_1, \dots, \mathbf{e}_n) = 0$, a feltétel alapján. A multilinearitás miatt bármely n darab vektorra $\mu(\mathbf{u}_1, \dots, \mathbf{u}_n)$ előáll fenti típusú kifejezések skalárszorosainak összegeként; így valóban 0. ■

8.3. Tétel. *A K feletti n -dimenziós térben értelmezett n -lineáris függvények vektorteret alkotnak a K test felett, amelynek a mértékek egydimenziós alterét alkotják.*

Bizonyítás. A tétel első állítása nyilvánvalóan igaz. Az is világos, hogy ennek a vektortérnek alterét alkotják a mértékek, hiszen — mint a bilineáris függvényeknél láttuk — antiszimmetrikus lineáris függvények lineáris kombinációja is antiszimmetrikus lineáris függvény.

Mindenekelőtt belátjuk, hogy a mértékek altere nem nulldimenziós, azaz létezik nemtriviális mérték. (Ezt a bevezető részben már jeleztük.) Legyenek $\mathbf{a}_1, \dots, \mathbf{a}_n$ az \mathcal{U} vektortérnek e sorrendben felírt elemei. Rögzítsük a tér egy \mathbf{U} bázisát, és tekintsük azt az A mátrixot, amelyben a j -edik oszlop az \mathbf{U} bázisban felírt $[\mathbf{a}_j]$ mátrix. Defináljuk a μ függvényt a következőképpen: $\mu(\mathbf{a}_1, \dots, \mathbf{a}_n) = \det(A)$. A mátrix determinánsának a tulajdonságai következtében ez a függvény az oszlopoknak — és így az őket meghatározó vektoroknak — multilineáris antiszimmetrikus függvénye, tehát mérték. Tekintettel arra, hogy a bázis esetén a mátrix egy olyan skalármátrix, amelyben a diagonális minden eleme 1, ezért az ehhez tartozó függvényérték 1; így μ egy nemtriviális mérték.

Legyen most μ_0 egy nemtriviális mérték. A 8.2. Tétel szerint létezik olyan $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bázis, amelyre $\mu_0(\{\mathbf{e}_1, \dots, \mathbf{e}_n\}) = e \neq 0$. Ezért tetszőleges μ mérték esetén létezik olyan $c = c_\mu$ skalár, amelyre

$$\mu(\mathbf{e}_1, \dots, \mathbf{e}_n) = c \cdot e = c \cdot \mu_0(\mathbf{e}_1, \dots, \mathbf{e}_n).$$

Mivel a mértékek vektorteret alkotnak, ezért $\mu' = \mu - c\mu_0$ is mérték. A most látottak szerint

$$\mu'(\mathbf{e}_1, \dots, \mathbf{e}_n) = \mu(\mathbf{e}_1, \dots, \mathbf{e}_n) - c \cdot \mu_0(\mathbf{e}_1, \dots, \mathbf{e}_n) = 0.$$

A 8.2. Tétel szerint tehát μ' triviális mérték, vagyis

$$\mu(\mathbf{u}_1, \dots, \mathbf{u}_n) = c \cdot \mu_0(\mathbf{u}_1, \dots, \mathbf{u}_n)$$

teljesül minden $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ vektorrendszerre, azaz $\mu = c\mu_0$. ■

8.4. Tétel. *Legyen α az n -dimenziós \mathcal{U} vektortér lineáris transzformációja. Ekkor létezik olyan egyértelműen meghatározott $|\alpha|$ skalár, hogy bármely $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ vektorrendszerre és μ mértékre*

$$\mu(\alpha(\mathbf{u}_1), \dots, \alpha(\mathbf{u}_n)) = |\alpha| \mu(\mathbf{u}_1, \dots, \mathbf{u}_n).$$

Bizonyítás. Válasszunk egy tetszőleges μ mértéket. α linearitása alapján a fenti egyenlőség bal oldala a szereplő vektorokban n -lineáris és antiszimmetrikus, tehát mérték. Ugyancsak mérték szerepel a jobb oldalon is; mégpedig tetszőlegesen választott skalár tényező esetében is. Ha a mérték nem triviális, akkor tetszőleges $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bázis esetében egyenlőséget kapunk, az $|\alpha| = \mu(\alpha(\mathbf{e}_1), \dots, \alpha(\mathbf{e}_n)) \cdot (\mu(\mathbf{e}_1, \dots, \mathbf{e}_n))^{-1}$ választással. Ha egy másik, nemtriviális mértéket választunk, akkor a fenti egyenlőség mindkét oldala

ugyanazzal a skalárral szorozódik, tekintettel arra, hogy a mértékek tere egydimenziós. Eszerint $|\alpha|$ minden nemtriviális mértékre ugyanaz. Ez a skalár a triviális mértékre is megfelel; hiszen ekkor a fenti egyenlőség mindkét oldalán 0 áll. ■

8.5. Tétel. $|\alpha \cdot \beta| = |\alpha| \cdot |\beta|$.

Bizonyítás. Legyen μ a tér egy nemtriviális mértéke, és válasszuk az $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bázist úgy, hogy azon a mérték 1 legyen (ez nem lényeges, de így könnyebb a számolás). A 8.4. Tétel alapján

$$\begin{aligned} |\alpha\beta| &= |\alpha\beta| \cdot \mu(\mathbf{e}_1, \dots, \mathbf{e}_n) = \mu(\alpha\beta(\mathbf{e}_1), \dots, \alpha\beta(\mathbf{e}_n)) = \\ &= |\alpha| \cdot \mu(\beta(\mathbf{e}_1), \dots, \beta(\mathbf{e}_n)) = |\alpha| \cdot |\beta| \cdot \mu(\mathbf{e}_1, \dots, \mathbf{e}_n) = |\alpha| \cdot |\beta|. \end{aligned} \quad \blacksquare$$

8.6. Tétel. α akkor és csak akkor szinguláris, ha $|\alpha| = 0$.

Bizonyítás. Válasszunk egy nemtriviális mértéket és egy $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bázist. Mivel e bázison a mérték nem 0, ezért $|\alpha|$ akkor és csak akkor 0, ha $\mu(\alpha(\mathbf{e}_1), \dots, \alpha(\mathbf{e}_n)) = 0$. A 8.1. és a 8.2. Tételek szerint ez pontosan akkor teljesül, ha az $\alpha(\mathbf{e}_1), \dots, \alpha(\mathbf{e}_n)$ vektorok lineárisan összefüggenek. Ennek viszont az a feltétele, hogy $\text{Ker}(\alpha)$ nem triviális, azaz α szinguláris transzformáció. ■

8.7. Tétel. $|\alpha| = \det[\alpha]$, a tér bármely bázisában.

Bizonyítás. Tekintsünk először egy olyan $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ bázist, amelynek a mértéke 1. A 8.3. Tétel szerint az ebben a bázisban felírt mátrixra $\det(A) = \det[\alpha]$ egy mérték. Tekintettel arra, hogy az adott bázison ezek megegyeznek és a mértékek tere egydimenziós, ezért $|\alpha| = \det[\alpha]$ igaz az adott bázisban. Új bázisra való áttérésnél α mátrixa az új bázisban megegyezik a $\sigma^{-1}\alpha\sigma$ transzformációnak a régi bázisban felírt mátrixával. A 8.5. Tétel alapján:

$$|\sigma^{-1}\alpha\sigma| = |\sigma^{-1}| \cdot |\alpha| \cdot |\sigma| = |\sigma^{-1}| \cdot |\sigma| \cdot |\alpha| = |\sigma^{-1}\sigma\alpha| = |\alpha|. \quad \blacksquare$$

1. Következmény (a determinánsok szorzástétele). $\det(AB) = \det(A) \cdot \det(B)$.

Bizonyítás. Legyenek α és β olyan transzformációk az \mathcal{U} vektortérben, amelyeknek valamely bázisban felírt mátrixaira $[\alpha] = A$ és $[\beta] = B$. A transzformációk szorzatára $[\alpha\beta] = AB$ teljesül; amiből a 8.7. Tétel alapján $|\alpha\beta| = \det(AB)$ következik. Ugyanezt a tételt felhasználva kapjuk, hogy $|\alpha| = \det(A)$ és $|\beta| = \det(B)$. A 8.5. Tételből következik, hogy $\det(AB) = \det(A) \cdot \det(B)$. ■

2. Következmény. Ha $\det(A) \neq 0$, akkor $\det(A^{-1}) = \frac{1}{\det(A)}$.

Bizonyítás. Ha $\det(A) \neq 0$, akkor A reguláris, így létezik inverze. A determinánsok szorzástételéből $A \cdot A^{-1} = I$ alapján $\det(A) \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det(I) = 1$ következik. ■

Megjegyzések

1. A determináns itt tárgyalt bevezetése megmutatja a determináns igazi jelentését. Az n -dimenziós térben a térfogat nem egyértelmű dolog. Az megállapodás, hogy az egységnyi élhosszúságú (n -dimenziós) kocka térfogata legyen egységnyi. Az viszont bizonyítható, hogy a térfogat egy arányossági tényezőtől eltekintve egyértelműen meghatározott. Itt az analízisbeli bizonyítást „átugorva” formálisan értelmeztük a mértéket (a térfogatot) bizonyos testekre. Ez a definíció eleget tett azoknak a kíváncsagságoknak, amelyeket a térfogattól elvárunk. A fent bizonyított eredmények azt mutatják, hogy egy lineáris transzformáció bármely szóba jövő „valami”-nek a térfogatát ugyanolyan arányban változtatja meg; noha a térfogat nem is egyértelmű. Nos, ez a térfogat-változtatási arány a transzformáció mátrixának a determinánsával egyezik meg. Külön érdekesség az, hogy ez a determináns nem függ attól, hogy melyik bázisban tekintjük a transzformáció mátrixát. (Persze miért is függene, hiszen a transzformáció mátrixa csak arra szolgál — esetünkben —, hogy ezt az arányt ki tudjuk számítani.)

2. A determináns „elvi” definíciójánál szépséghibának tűnhet az, hogy a determináns létezését csak a mátrixok tárgyalásánál találhatók kellemetlen számolásos formában tudtuk bizonyítani. (Persze, ez szinte mindig így van, az egyértelműség könnyebben látható be, mint a létezés.) Létezik azonban egy „elvi” kiszámítás is. Gondoljuk meg, hogy egy síkidom területét vagy egy test térfogatát az „alapszor magasság” elvével lehet kiszámolni. Ezt a mértéknél is meg lehet tenni. Természetesen vigyázni kell arra, hogy az „alap” is és a „magasság” is előjeles. Ez az eljárás, illetve elképzelés lehetőséget ad a determináns (vagy a mérték) rekurzív definíciójára. A mátrixok tárgyalásánál szerepelt is ez a módszer. Nevezetesen a kifejtési tétel pontosan ezt az eljárást tükrözi. □

Feladatok

1. Hány dimenziós vektorteret alkotnak a K test felett a K feletti n -dimenziós térben értelmezett n -lineáris függvények?

2. Bizonyítsuk be, hogy pozitív definit kvadratikus alak mátrixának a determinánsa pozitív. Mutassuk meg, hogy a megfordítás nem igaz. Mely kvadratikus alakok mátrixának a determinánsa 0?

3. Bizonyítsuk be, hogy unitér transzformáció ortonormált bázisban vett mátrixának a determinánsa 1 abszolút értékű.

4. Bizonyítsuk be, hogy a valós tér ortogonális transzformációjának ortonormált bázisban vett mátrixának a determinánsa +1 vagy -1. Mutassuk meg, hogy a megfordítás nem igaz.

5. Egy mátrixot *permutációs mátrixnak* nevezünk, ha minden sorában és minden oszlopában egyetlen 1-es szerepel és a többi elem 0. (Mit permutál?) Bizonyítsuk be, hogy a permutációs mátrixok ortogonálisak. Határozzuk meg (elvben) minimálpolinomjukat, sajátértékeiket és sajátvektoraikat.

6. Legyenek A , B és U n -sorú négyzetes mátrixok és O az az n -sorú négyzetes mátrix, amelynek minden eleme 0. Tekintsük a következő blokkokra beosztott $2n$ -sorú négyzetes mátrixokat: $P = \begin{bmatrix} B & U \\ O & A \end{bmatrix}$ és $Q = \begin{bmatrix} A & U \\ AB & O \end{bmatrix}$. Bizonyítsuk be, hogy $\det(P) = \det(A) \cdot \det(B)$ és $\det(Q) = (-1)^n \det(BA) \cdot \det(U)$.

7. Az előző feladatban legyen $U = -I$. Mutassuk meg, hogy P elemi átalakításokkal olyan formájúvá alakítható át, mint Q . Ennek felhasználásával bizonyítsuk be elemi úton a determinánsok szorzástételét.

2. Polinommátrixok normálalakja, karakterisztikus polinom

Ismét visszatérünk a mátrixok vizsgálatára, célunk a karakterisztikus polinom részletes tanulmányozása. Az alábbiakban, ha nem mondunk mást, mátrixon rendszerint négyzet alakú mátrixot értünk. A K testbeli elemű A négyzetes mátrix karakterisztikus polinomja definíció szerint az $A - x \cdot I$ mátrix determinánsa, ahol I és A azonos méretűek és x határozatlan a K test felett. Az $A - x \cdot I$ mátrixot úgy tekinthetjük, mint egy olyan mátrixot, amelynek elemei K -beli együtthatós polinomok.

8.2. Definíció. A K test feletti polinommátrixnak nevezünk egy mátrixot, ha elemei $K[x]$ -beli polinomok. Az ilyen mátrixok jelölésére az $A(x)$ alakot fogjuk használni.

Egy R gyűrű feletti mátrixnak nevezünk egy mátrixot, ha elemei R -beliek. ■

A bevezető részben láttuk az ilyen mátrixokról, hogy hasonlóképpen kezelhetők, mint a testből vett elemű mátrixok. Mint ott említettük, az egyetlen kivétel az, hogy ebben az esetben nem feltétlenül képezhetjük a mátrix inverzét, akkor sem, ha a determinánsa nem 0.

A leképezések mátrixának a vizsgálatánál láttuk, hogy igen fontos szerepet játszanak az invertálható mátrixok, illetve az ilyenekkel való szorzás. Ott viszont csak olyan mátrixokkal foglalkoztunk, amelyeknek elemeit egy testből vettük. Minden további nélkül alkalmazhatóak volnának ezek az eljárások, ha tudnánk, hogy a szereplő R gyűrű benne van egy testben. Ez igaz akkor, ha R egységelemes integritási tartomány. A megfelelő testet hasonlóképpen kaphatjuk meg, mint ahogy az egész számokból felépítjük a racionális számtestet. Ezt a testet az adott integritási tartomány *hányadostestének* nevezzük. Az egyáltalán nem bonyolult bizonyításra a második kötetben kerül sor.

8.8. Tétel. *Tegyük fel, hogy az A (négyzetes) mátrix elemei egy R egységelemes integritási tartományból valók. A -nak akkor és csak akkor létezik inverze R felett, ha determinánsa R -beli egység. A $K[x]$ feletti $A(x)$ (négyzetes) polinommátrixnak akkor és csak akkor létezik polinommátrix inverze, ha determinánsa K -beli nemnulla elem.*

Bizonyítás. A determináns definíciójánál láttuk, hogy ha egy négyzetes mátrix determinánsa nem 0, akkor inverzének elemeit az eredeti mátrix elemeiből összeadás, kivonás, szorzás segítségével és a determinánssal való osztással kaphatjuk meg. Ez azt jelenti, hogy ha a mátrix determinánsa egység (azaz R minden elemének osztója), akkor az inverz mátrix elemei is R -beliek. Ha A -nak létezik inverze, akkor található hozzá olyan B mátrix, amelyre $AB = I$. A determinánsok szorzástétele alapján $\det(A) \cdot \det(B) = \det(I) = 1$, így $\det(A)$ osztója 1-nek, tehát egység.

A polinommátrixokra vonatkozó állítás ebből azonnal következik, hiszen testbeli együtthatós polinomok esetében pontosan a nemnulla konstansok az egységek. ■

Megjegyzés. A determinánsok szorzástételére létezik egy számolós bizonyítás, ahol nincs szükség a hányadostestre (l. az előző rész feladatsorát). Ez a bizonyítás viszont nem mutatja meg a determináns jelentését. □

A következőkben egységelemes integritási tartomány esetére átnézzük a mátrixokon végezhető elemi átalakításokat, amelyeket a rang és az inverz vizsgálatánál tárgyaltunk (a 4.5. Definíciótól a 4.17. Tételig). Ezekből az elemi átalakításokra lesz szükségünk, és azokra a mátrixokra, amelyekkel való szorzás ezt az elemi átalakítást létrehozza.

8.3. Definíció. Rögzített méretű négyzetes mátrixoknál $E_{i,j}$ jelöli azt a mátrixot, amelyben az i -edik sor j -edik eleme 1 és a többi elem 0. $I = \sum_i E_{i,i}$ az identitásmátrix. ■

8.9. Tétel. R feletti rögzített méretű négyzetes mátrixok esetén $E_{i,j} \cdot E_{p,q} = \delta_{j,p} E_{i,q}$.

Az alábbi három típusú mátrix invertálható:

- (1) $I + c \cdot E_{i,j}$ ($i \neq j$); inverze $I - c \cdot E_{i,j}$.
- (2) $I + c \cdot E_{i,i}$, ha $1 + c$ egység; inverze $I - \frac{c}{1+c} \cdot E_{i,i}$.
- (3) $I + E_{i,j} + E_{j,i} - E_{i,i} - E_{j,j}$; inverze önmaga.

E mátrixokat elemi átalakítómátrixoknak fogjuk nevezni.

Bizonyítás. Az első állítás triviálisan kiszámolható. Mindhárom további állítás bizonyítást nyert a mátrixok rangja tárgyalásánál; de egyszerű számolással megkapható az $E_{i,j} \cdot E_{p,q} = \delta_{j,p} E_{i,q}$ összefüggést felhasználva. ■

8.4. Definíció. Téglalap alakú mátrix esetén elemi oszlopátalakításnak nevezzük a következőket:

1. Ha egy oszlophoz egy tőle különböző oszlop c -szeresét hozzáadjuk.
2. Egy oszlopot egy c egységgel megszorozunk.
3. Két oszlopot felcserélünk.

Hasonló módon értelmezhetők az elemi sorátalakítások. ■

8.10. Tétel. Az elemi oszlopátalakítások pontosan úgy állnak elő, hogy a megfelelő méretű, azonos típusú elemi átalakítómátrixszal szorzunk jobbról. Az elemi sorátalakítások pontosan úgy állnak elő, hogy a megfelelő méretű, azonos típusú elemi átalakítómátrixszal szorzunk balról.

Bizonyítás. Hasonlóképpen történik, mint azokra a mátrixokra, amelyeknek az elemeit egy testből vettük. ■

Mátrixok elemi átalakítását a rang meghatározására használtuk. Ott az eredeti mátrixból egy olyan diagonális mátrixot nyertünk, amelyben az i -edik sor i -edik eleme valameddig 1 volt, és az összes további elem 0. Ez gyűrűk esetében nem érhető el még egyszer

egyes mátrixokra sem; ha például a mátrix egyetlen eleme az x határozatlan, akkor ebből a gyűrűműveletek segítségével soha nem kaphatunk nemnulla konstans. Ennek ellenére speciális gyűrűk esetében kaphatunk egy speciális diagonális alakot.

8.5. Definíció. Az R gyűrű feletti $A = [a_{i,j}]$ mátrixot normálalakúnak mondjuk, ha minden i -re $a_{i,i}$ osztója $a_{i+1,i+1}$ -nek és a mátrix többi eleme 0. ■

8.11. Tétel. *Euklideszi gyűrű feletti mátrix véges sok elemi átalakítással normálalakra hozható.*

Bizonyítás. Tegyük fel, hogy a mátrixnak k oszlopa, n sora van, és például $k \leq n$. Az állítást k -ra vonatkozó indukcióval bizonyítjuk. A $k = 0$ esetben a feltétel üres, így az állítás igaz. (Ez a feltétel valójában azt takarja, hogy az indukciós lépés magában foglalja speciális esetként a $k = 1$ esetet is.)

Az eredeti mátrixból elemi átalakításokkal egy olyan mátrixot fogunk készíteni, amelyben az első sor első eleme a mátrix minden elemének osztója. Ehhez egy többlépéses eljárást (algoritmust) adunk meg. Mivel R euklideszi gyűrű, ezért létezik olyan $\varphi : R \rightarrow \mathbb{N}$ euklideszi norma, amelyre tetszőleges $a, b \in R$ esetén, ha $b \neq 0$, található olyan $q, r \in R$, hogy $a = bq + r$ és vagy $r = 0$, vagy $\varphi(r) < \varphi(b)$.

Legyen a vizsgált mátrix $A = [a_{i,j}]$. Ha a mátrix minden eleme 0, akkor a mátrix normálalakú. Ha van nemnulla elem a mátrixban, akkor esetleges sor- és oszlopcserével elérhető, hogy $a_{1,1} \neq 0$. (A keletkezett új mátrixban szereplő elemek indexeit megváltoztatjuk aszerint, hogy azok a mátrix melyik helyén állnak. Tehát, ha az eredeti mátrix i -edik sorának j -edik eleme kerül az első sor első helyére, akkor ettől kezdve ezt az elemet nem $a_{i,j}$, hanem $a_{1,1}$ fogja jelölni.) Tekintettel arra, hogy φ értékkészlete természetes számokból áll, ezért esetleges újabb sor-, illetve oszlopcserékkel az is elérhető, hogy $\varphi(a_{1,1}) \leq \varphi(a_{i,j})$, tetszőleges i, j párra. Ez az eljárás első lépése.

Ha a mátrix első sorában van olyan $a_{1,j}$, amelyik nem osztható $a_{1,1}$ -gyel, akkor a maradékos osztás léte miatt van olyan $q, r \in R$, amelyre $a_{1,j} = q \cdot a_{1,1} + r$, ahol $r \neq 0$ miatt $\varphi(r) < \varphi(a_{1,1})$. Az első oszlop q -szorosát kivonva a j -edikből egy olyan elemi átalakítást végeztünk, amely után egy elemnek az euklideszi normája kisebb, mint $a_{1,1}$ -é. Ez az eljárás második lépése. Ezután megismételjük az első lépést, majd újra a másodikat, és így tovább. Tekintettel arra, hogy a kapott minimális normájú elem normája egyre csökken, és a norma természetes szám, ezért véges sok lépésben eljutunk egy olyan helyzethez, hogy a második lépés már nem végezhető el (hiszen az első mindig elvégezhető). Ez viszont akkor következik be, ha az első sor minden eleme osztható $a_{1,1}$ -gyel. Ha az első sor minden eleme osztható $a_{1,1}$ -gyel, akkor az első oszlop megfelelő többszörösét kivonva a többiből, az első sorban az első elemet kivéve minden elem 0 lesz; és ezek is elemi átalakítások.

Ezután következik a harmadik lépés, ami ugyanolyan, mint az első lépés, de az első sor helyett az első oszlopra végezzük. Ha itt az elemi átalakítás után találunk egy olyan elemet, amelynek az euklideszi normája kisebb, mint $a_{1,1}$ -é, akkor megint előlről kezdjük az eljárást. Ennek az ismétlése is csak véges sokszor következhet be, hiszen természetes számok minden csökkenő sorozata véges. Most már feltehető, hogy az első oszlop minden eleme osztható $a_{1,1}$ -gyel. (Itt is „kinullázhatnánk” az első oszlopot, de a továbbiak céljából ezt nem most tesszük.)

Tegyük most fel, hogy van a mátrixban olyan $a_{i,j}$ elem, amelyik nem osztható $a_{1,1}$ -gyel, azaz $a_{i,j} = qa_{1,1} + r$, ahol $\varphi(r) < \varphi(a_{1,1})$. A harmadik lépés következtében van olyan p , amelyre $a_{i,1} = pa_{1,1}$. Ezután két elemi átalakítást végzünk. Először kivonjuk az i -edik sorból az első sor $(p-1)$ -szeresét, majd a j -edik oszlopból az első oszlop q -szorosát. Az első átalakítás után az i -edik sor első helyén $a_{1,1}$ fog állni, míg a j -edik helyen álló elem nem változik, hiszen $a_{1,j} = 0$. A második elemi átalakítás után $a_{i,j}$ helyére r kerül. Most ismét visszatérhetünk az első lépéshez, és így tovább. Mivel természetes számok csökkenő sorozata véges, ezért véges sok lépésben befejeződik az eljárás, ami azt jelenti, hogy a mátrix minden eleme osztható $a_{1,1}$ -gyel.

Vegyük észre, hogy ha a kiindulásnál egy d elem a mátrix összes elemének osztója volt, akkor ez az elem az eljárás végén is osztója lesz minden elemnek.

A mátrix első sorát és első oszlopát elhagyva egy olyan mátrixhoz jutunk, amelynek $k-1$ oszlopa és $n-1$ sora van. A teljes indukciós feltétel szerint ez a mátrix elemi átalakításokkal a kívánt alakra hozható. Ezek az elemi átalakítások eleve tekinthetők az eredeti mátrix elemi átalakításainak. Ennek a mátrixnak az $a_{2,2}, \dots, a_{k,k}$ elemeire az indukciós feltétel szerint érvényesek az oszthatósági megkötések; és mivel $a_{1,1}$ a mátrix minden elemének osztója, ezért $a_{2,2}$ -nek is. Végül az első, majd a második, \dots , végül a k -adik sor megfelelő többszöröseit kivonva az utánuk következő sorokból olyan mátrixot kapunk, amelyben az $a_{i,i}$ -ktől különböző minden elem 0. ■

Mint mondtuk, számunkra itt csak az euklideszi gyűrűk fontosak. Valamivel összetettebb eljárással viszont egy általánosabb tételt bizonyíthatunk:

Kiegészítés. *A 8.11. Tétel akkor is igaz, ha az R gyűrű csak főideálgyűrű.*

Bizonyítás. Lényegében ugyanazt az eljárást folytatjuk, mint az euklideszi gyűrűk esetében. Az oszlop-, illetve sorcserék továbbra is invertálható mátrixszal való szorzással adódnak.

Tegyük fel, hogy a mátrix első sorának első két eleme a és b , amelyek d legnagyobb közös osztója $ax + by = d$ alakú, alkalmas $x, y \in R$ elemekkel, hiszen R főideálgyűrű. A mátrix többi elemét $*$ -gal jelölve tekintsük a következő szorzatot:

$$\begin{bmatrix} a & b & * & \dots & * \\ * & * & * & \dots & * \\ * & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \dots & * \end{bmatrix} \cdot \begin{bmatrix} x & -\frac{b}{d} & 0 & \dots & 0 \\ y & \frac{a}{d} & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

A második mátrix determinánsa $x \cdot \frac{a}{d} + y \cdot \frac{b}{d} = 1$, ezért invertálható. A szorzat első sorának első eleme $ax + by = d$, és a sor második eleme $-a \cdot \frac{b}{d} + b \cdot \frac{a}{d} = 0$.

Ezt az eljárást folytatva elérhető, hogy az első sor első eleme az eredeti mátrix első sorában levő elemeinek (egyik) legnagyobb közös osztója legyen, míg a sor többi eleme mind 0.

Most ugyanazt az eljárást végezzük az első oszloppal, majd ismét az első sorral és így tovább. Tekintettel arra, hogy a kapott mátrix első sorának első eleme mindig osztója lesz az előző mátrix első sora első elemének, ezért az eljárás véges sok lépésben véget ér, mint ez az egyértelmű faktorizációból következik.

Jelölje a kapott mátrix első sorában álló elemet $a_{1,1} (\neq 0)$; az első sor és az első oszlop többi eleme 0. Legyen $a_{i,j}$ az i -edik sor j -edik eleme. A i -edik sort az elsőhöz adva (ez is elemi átalakítás) az első sor j -edik eleme $a_{i,j}$. Ha ez nem osztható $a_{1,1}$ -gyel, akkor ismét folytathatjuk az eljárást. Tekintettel arra, hogy minden lépésben ismét osztót nyerünk, ezért az egész eljárás véges sok lépésben befejeződik az egyértelmű faktorizáció következtében. ■

A normálalakot többféleképpen állíthatjuk elő, ezért szükséges annak a belátása, hogy a normálalak nem függ az előállítás módjától. Evégett olyan invariáns adatokat kell találni, amelyek a normálalak előállításakor változatlanok maradnak, és meghatározzák a normálalakot.

8.6. Definíció. Az R euklideszi gyűrű feletti A mátrix i -edik $\Delta_i(A)$ determinánsosztóját a következőképpen definiáljuk:

$$\Delta_0(A) = 1$$

Ha $i > 0$, de nem nagyobb a mátrix sorai és oszlopai számánál, akkor tekintsük a mátrix i -edrendű négyzetes részmatrixainak a determinánsát, és legyen $\Delta_i(A)$ ezeknek a legnagyobb közös osztója.

Ha i nagyobb, mint a mátrix sorainak vagy oszlopainak a száma, akkor legyen $\Delta_i(A) = 0$. ■

8.12. Tétel. *Ha egy mátrixon elemi átalakítást végzünk, akkor a determinánsosztók asszociáltság erejéig nem változnak meg.*

Bizonyítás. A tétel triviálisan igaz, ha $i = 0$, illetve, ha nagyobb, mint a mátrix sorainak vagy oszlopainak a száma.

Az elemi átalakításnál azt vesszük figyelembe, hogy ezek egy invertálható mátrixszal való szorzással hozhatók létre. Azt kell tehát megmutatni, hogy ha P , illetve Q invertálható mátrixok, akkor $\Delta_i(PA) = \Delta_i(AQ) = \Delta_i(A)$, amennyiben a megfelelő szorzások elvégezhetők.

Tekintettel arra, hogy bármely A mátrix esetén $\det(A^\dagger) = \det(A)$, ezért $\Delta_i(A^\dagger) = \Delta_i(A)$. Ha tehát $\Delta_i(AQ) = \Delta_i(A)$ minden invertálható Q -ra és minden A -ra igaz, akkor P^\dagger invertálhatósága miatt

$$\Delta_i(PA) = \Delta_i((PA)^\dagger) = \Delta_i(A^\dagger P^\dagger) = \Delta_i(A^\dagger) = \Delta_i(A).$$

Ha Q invertálható, akkor Q^{-1} is az. Ezért elég belátni, hogy $\Delta_i(A)$ osztója $\Delta_i(AQ)$ -nak. Valóban, hiszen emiatt az is igaz, hogy $\Delta_i(AQ)$ osztója $\Delta_i(AQQ^{-1})$ -nek, így ezek asszociáltak.

Ennél — látszólag — többet bizonyítunk be, nevezetesen azt, hogy tetszőleges A és B mátrixokra, ha AB létezik, akkor $\Delta_i(A)$ osztója $\Delta_i(AB)$ -nek (i a szereplő sorok és

oszlopok számánál nem nagyobb természetes szám). Legyen $A = [a_{u,v}]$, $B = [b_{u,v}]$ és legyen n az A oszlopainak és B sorainak a száma. Ekkor az AB mátrixban a p -edik sor q -adik eleme:

$$a_{p,1} \cdot b_{1,q} + \cdots + a_{p,n} \cdot b_{n,q}.$$

Az AB mátrix egy i sorú részmátrixának egy oszlopában ilyen elemek állnak, ahol a sorindexek rendre p_1, \dots, p_i , az oszlopindex pedig változatlan. Jelölje \mathbf{a}_r azt az oszlop mátrixot, amelyben az s -edik helyen $a_{p_s,r}$ áll. Ekkor a szóban forgó mátrix megfelelő oszlopa:

$$\mathbf{a}_1 \cdot b_{1,q} + \cdots + \mathbf{a}_n \cdot b_{n,q}.$$

A determináns tulajdonsága alapján e mátrix determinánsa felbontható n -tagú összegre, ahol az összeg r -edik tagja $b_{r,q}$ -szorosa azon mátrix determinánsának, amelyben az r -edik oszlop \mathbf{a}_r , s a többi oszlop változatlan. Ezt az eljárást minden oszlopra elvégezve olyan összeget kapunk, amelynek a tagjai úgy állnak elő, hogy az A mátrix i sorú négyzetes mátrixának determinánsát szorozzuk a B mátrix elemeinek a szorzatával. Tekintettel arra, hogy $\Delta_i(A)$ osztója a fellépő determinánsok mindegyikének, ezért osztója a fellépő szorzatösszegeknek, amelyek pontosan az AB mátrix megfelelő aldeterminánsai. Ekkor viszont $\Delta_i(A)$ osztója ezek legnagyobb közös osztójának, $\Delta_i(AB)$ -nek is. ■

1. Következmény. *Euklideszi gyűrű feletti mátrixok normálalakja egyértelműen meghatározott.*

Bizonyítás. Tekintsük az eredeti mátrix $\Delta_0, \Delta_1, \dots, \Delta_n$ determinánsosztóit. A fenti tétel szerint ezek megegyeznek a normálalakban kapott determinánsosztókkal. A normálalak diagonálisában álló elemek ($a_0 = 1$), a_1, a_2, \dots, a_n . Ezek között szerepelhet 0 is, de az oszthatósági feltétel miatt az $a_i = 0$ esetben $a_{i+1} = 0$ is fennáll. Feltehető tehát, hogy r a legnagyobb index, amelyre $a_r \neq 0$ (ekkor r pontosan a mátrix rangja). Ha $i > r$, akkor minden i sorú négyzetes mátrixban szerepel olyan sor, amelynek minden eleme 0, tehát minden ilyen mátrix determinánsa 0, így ezek legnagyobb közös osztója, Δ_i is 0. Az $i \leq r$ esetben 0 lesz minden olyan négyzetes mátrix determinánsa, amelyben r -nél nagyobb indexű sor vagy oszlop szerepel. Hasonlóképpen belátható, hogy a fennmaradó esetekben is 0 a determináns, amennyiben a részmátrix nem szimmetrikus (tehát a fellépő oszlopindexek nem egyeznek meg a fellépő sorindexekkel). Ezek a mátrixok elhagyhatók, hiszen 0 minden elemnek többszöröse. A fennmaradó mátrixok diagonálisak, ezért determinánsuk a diagonális elemek szorzata. Legyenek ezek a_{j_1}, \dots, a_{j_i} , ahol $j_1 < \dots < j_i$. Ebből azonnal következik, hogy tetszőleges r mellett $r \leq j_r$. Az a_s elemekre belátott oszthatósági kapcsolatot alapján ezek a szorzatok mind oszthatók az $a_1 \cdot \dots \cdot a_i$ szorzattal. Tekintettel arra, hogy ez maga is egy fellépő determináns, ezért éppen ez az i -edik determinánsosztó. Ezek szerint:

$$\Delta_0 = a_0 = 1, \quad \Delta_1 = a_1, \quad \Delta_2 = a_1 \cdot a_2, \quad \dots, \quad \Delta_n = a_1 \cdot \dots \cdot a_n.$$

Ebből meghatározhatók a normálalak diagonálisának elemei. Ha $\Delta_i \neq 0$, akkor $a_{i+1} = \frac{\Delta_{i+1}}{\Delta_i}$, míg a $\Delta_i = 0$ esetben $a_{i+1} = 0$ is igaz. ■

2. Következmény. *Minden determinánsosztó osztója a következőnek. A fellépő hányadosokat elemi osztóknak nevezzük. Ha nincs hányados, akkor a megfelelő elemi osztó 0.*

Bizonyítás. Ez azonnal adódik az 1. Következményből, mint ahogy az is, hogy az elemi osztók pontosan az a_i elemek. ■

3. Következmény. *A fenti tétel és a két következmény akkor is igaz, ha a mátrixnak végtelen sok sora van.*

Természetesen ekkor nem a mátrix formális definíciójára gondolunk, hanem a sorokat egy halmaz elemeivel indexezzük.

Bizonyítás. A normálalak megállapításánál nem használtuk ki, hogy a sorok száma véges. Az egyértelműség bizonyításánál viszont szükség volt a determinánsosztókra, és arra, hogy invertálható mátrixszal való szorzásnál ezek nem változnak meg. Végtelen mátrixoknál ezek bonyodalmat okozhatnak. Előfordulhat ugyanis, hogy egyik sorból ki kell vonni egy másik sor „skalárszorását”, de a két sor között végtelen sok sor van. Ezen a következőképpen segíthetünk.

Azt kell belátni, hogy ha kétféleképpen állítjuk is elő a normálalakot, ugyanazokhoz az elemi osztókhoz jutunk. Mindkét eljárásban csak véges sok sor szerepelhet. Sorcserekkel hozzuk előre ezeket a sorokat. Ezek a harmadik típusú elemi átalakítások. Mint láttuk, ilyen elemi átalakításoknál a részmatrixok lényegében változatlanul maradnak. Ezután már szorítkozhatunk arra a véges részmatrixra, amelyben a számunkra lényeges szerepet betöltő sorok vannak. Most már a véges esetben használt eljárásból következik, hogy mindkét esetben ugyanazokhoz az elemi osztókhoz jutunk.

Megjegyezzük, hogy ekkor az utolsó eljárási lépés nem végezhető el, hiszen végtelen sok sorból kellene másik sor skalárszorását kivonni, ami véges sok lépésben lehetetlen. Ekkor csak azt mondhatjuk ki, hogy egy-egy oszlop elemei az ezen oszlop diagonális elemeinek a többszörösei. ■

Az alábbi következményeket főideálgyűrűkre mondjuk ki, de csak euklideszi gyűrűkre bizonyítjuk. A főideálgyűrűk esetében a számolás bonyolultabb.

4. Következmény. *Legyen $U = \{u_1, \dots, u_n\}$ az R főideálgyűrű feletti \mathcal{M} modulus egy bázisa (abban az értelemben, hogy elemeinek csak a csupa 0 együtthatókkal képezett lineáris kombinációja \mathbf{o}). Ekkor az \mathcal{M} tetszőleges \mathcal{N} részmodulusához található \mathcal{M} -ben olyan $V = \{v_1, \dots, v_n\}$ bázis és R -nek olyan a_1, \dots, a_n elemei, amelyek mindegyike osztója a következőnek, hogy a $w_1 = a_1 v_1, \dots, w_n = a_n v_n$ elemek közül a \mathbf{o} -tól különbözőek az \mathcal{N} egy (fentebbi értelemben vett) bázisát alkotják.*

Az a_1, \dots, a_n sorozat \mathcal{N} -nel (asszociáltság erejéig) egyértelműen meghatározott.

Bizonyítás. Tekintsük azokat a $[c_1, \dots, c_n]$ sorvektorokat, amelyekre $c_1 u_1 + \dots + c_n u_n \in \mathcal{N}$. Készítsük el azt a (egy olyan) végtelen sok sorból álló mátrixot, amelynek ezek a sorai. A kapott mátrixot normálalakra hozhatjuk. Az oszlopokkal végzett elemi átalakítások, mint láttuk, elemi bázistranszformációknak felelnek meg. Minden ilyen lépésnél \mathcal{M} -nek egy új bázisát kapjuk, és a sorok elemei most már \mathcal{N} ugyanazon elemének az új bázisban felírt koordinátái lesznek. A sorokkal végzett elemi átalakításnál mindig olyan sort kapunk, amelyik már valahol szerepelt a mátrixban, hiszen bármely részmodulus zárt a lineáris kombináció képzésére. Ezért a sorokkal csupán cserét kell végezni. Végezetül

egy olyan mátrixot kapunk, amelyben az első n sor által meghatározott mátrix diagonális, a diagonális elemei a_1, \dots, a_n , minden diagonális elem osztója a következőnek, ezek az elemek lényegében egyértelműek és minden elem többszöröse az ezen oszlopba eső diagonális elemnek.

Ekkor az oszlopoknak megfeleltetett \mathbf{v}_i báziselemek nyilván eleget tesznek a megfogalmazott tulajdonságoknak. ■

5. Következmény (főideálgyűrűk feletti végesen generált torziómodulusok alaptétele II. rész). *Legyen \mathcal{M} végesen generált torziómodulus az R főideálgyűrű felett. Ekkor \mathcal{M} -nek létezik olyan $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bázisa, és olyan $r_i \in R$ elemek, amelyek mindegyike egy felbonthatatlan elem hatványa, hogy $r_i = r(\mathbf{u}_i)$ (azaz r_i a \mathbf{u}_i rendje). Az r_i elemek rendszere egyértelműen meghatározott.*

Bizonyítás. A fenti bázis léte azt jelenti, hogy \mathcal{M} olyan tagok direkt összegére bontható, amelyek egyetlen elem R -beli elemszereseiből állnak, és ezeknek a részmodulusoknak az exponense egy-egy felbonthatatlan elem hatványa. A főideálgyűrűk feletti végesen generált torziómodulusok alaptétele I. rész szerint \mathcal{M} felbontható olyan tagok direkt összegére, amelyeknek az exponense egy felbonthatatlan elem hatványa. Ez a felbontás abszolút egyértelmű. Ezért most már csak azt kell belátni, hogy ezek a modulusok tovább bonthatók a fent meghatározott módon. Legyen \mathcal{M} egy olyan végesen generált modulus, amelynek az exponense p^k , ahol $p \in K$ felbonthatatlan, és legyen $\mathbf{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ az \mathcal{M} generátorrendszere. Tekintsük az R elemeiből készített n hosszúságú sormátrixokat. Ennek bázisát alkotják azok az \mathbf{u}_i mátrixok, amelyekben az i -edik helyen álló elem 1, a többi pedig 0 ($0 < i \leq n$). Ezzel egy végesen generált \mathcal{K} modulust kaptunk az R euklideszi gyűrű (illetve főideálgyűrű) felett, amelynek $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ egy bázisa. Mivel ez bázis, ezért létezik egy olyan egyértelműen meghatározott $\varphi : \mathcal{K} \rightarrow \mathcal{M}$ homomorfizmus, amelynél $\varphi(\mathbf{u}_i) = \mathbf{a}_i$. Mivel \mathbf{A} generátorrendszer, ezért φ szürjektív. Legyen \mathcal{L} ennek a homomorfizmusnak a magja. A 4. Következmény szerint \mathcal{K} -nak van olyan $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ bázisa és R -nek olyan r_1, \dots, r_n elemei, hogy a $\mathbf{w}_i = r_i \mathbf{v}_i$ elemek közül, amelyek nem \mathbf{o} , az \mathcal{L} egy bázisát alkotják. Ezekre az elemekre még az is igaz, hogy r_i osztója az r_{i+1} -nek.

Legyen $\varphi(\mathbf{v}_i) = \mathbf{b}_i$. Itt $r_i \mathbf{b}_i = \mathbf{o}$, mert $r_i \mathbf{v}_i \in \mathcal{L}$. Mivel \mathcal{M} exponense a p felbonthatatlan elem egy hatványa, és r_i osztója az exponensnek, ezért r_i is hatványa a p -nek. Az r_i -k oszthatósági viszonya alapján nagyobb indexű r_i a p -nek nagyobb kitevőjű hatványa. Előfordulhatnak 0 kitevőjű hatványok is. Ha például az első t darab ilyen, akkor \mathcal{M} előáll a $\{\mathbf{v}_{t+1}, \dots, \mathbf{v}_n\}$ független rendszer generálta részmodulus képeként is. Ezért eleve feltehető, hogy r_1 nem egység.

A 4. Következmény alapján a \mathbf{b}_i elemek egy lineáris kombinációja csak akkor lehet \mathbf{o} , ha minden egyes együttható osztható a megfelelő r_i -vel. Ezzel megkaptuk a kívánt felbontást, és az egyértelműség következik az r_i -k egyértelműségéből. ■

Megjegyzés. A 8.11. Tétel, illetve annak kiegészítése akkor is igaz marad, ha az R gyűrűről csak azt tesszük fel, hogy abban minden végesen generált ideál főideál. Ebben az esetben nem használható az egyértelmű faktorizáció, ugyanis az nem igaz (l. az alábbi példát). Elég viszont azt megmutatni, hogy véges sok lépéssel (elemi átalakítással) a „bal felső sarokba” vihető egy olyan elem, amelyik minden elemnek osztója. Azt véges sok lépésben elérhetjük, hogy az első sorban és az első oszlopban

az első elem kivételével minden elem 0 legyen — ugyanúgy, mint ahogy a kiegészítésben tettük. Teljes indukciót használva feltehető, hogy az első sor és első oszlop elhagyásával kapott mátrix véges sok lépéssel normálalakra hozható. Legyenek a az első sor első eleme és b_1, \dots, b_r a megmaradt mátrixban a „diagonális” elemek. (Azért van idézőjel, mert a mátrix nem feltétlenül négyzetes.) A második sort az elsőhöz adva, a bal felső kétszer kettes részmátrix $\begin{bmatrix} a & b \\ 0 & b \end{bmatrix}$ alakú lesz ($b = b_1$). Az

idézett kiegészítés jelöléseit használva a szorzatmátrix bal felső kétszer kettes mátrixra $\begin{bmatrix} d & 0 \\ b_y & \frac{ab}{d} \end{bmatrix}$ adódik; és a többi elem nem változik. Tekintettel arra, hogy $d \mid b_1$ és $b_1 \mid b_i$ ($i > 1$), ezért d osztója a diagonális minden elemének, a harmadiktól kezdve. De a másodiknak is osztója, hiszen az $\frac{a}{d} \cdot b$.

Természetesen felmerül a kérdés, hogy van-e olyan gyűrű, amelyben minden végesen generált ideál főideál, de van nem végesen generált ideál is. Ennek megmutatására tekintsük azokat a K test feletti „formális polinomokat”, amelyekben az x „határozatlan” „kitevője” tetszőleges nemnegatív racionális szám. Ezek $f(x) = a_1 x^{r_1} + \dots + a_n x^{r_n}$ alakú kifejezések, ahol $a_1, \dots, a_r \in K$ és $0 \leq r_1 < \dots < r_n$ racionális számok. Ha véges sok ilyen kifejezésünk van, akkor a fellépő kitevők mindegyike egy (nem egyértelmű) r racionális szám egész számú többszöröse. Ezek a polinomok úgy írhatók, mint az $y = x^r$ „igazi” polinomjai, és aszerint is lehet velük számolni. Ennek megfelelően bármely véges sok elem által generált ideál főideál. Létezik viszont olyan ideál, amelyik nem generálható véges sok elemmel. Tekintsük például azokat a polinomokat, amelyekben a „konstans tag” 0. Ezek nyilván egy I ideált alkotnak. Ez az ideál nem lehet főideál. Ez az ideál tartalmazza ugyanis az összes x^s alakú polinomot, ezért generátoreleme is csak x^s alakú lehetne. Mivel a generátorelem minden ilyen elemnek osztója, ezért csak $x^0 = 1$ lehetne a generátorelem, ami viszont nincs bent az ideálban.

Az ilyen gyűrűk feletti modulusokra nem feltétlenül igaz viszont a 8.12. Tétel utáni 4. Következmény (valójában a 8.11. Tétel következménye, amelynek a bizonyításához van szükség a 8.12. Tételre). Legyen R a fenti racionális kitevős polinomok gyűrűje, és tekintsük R -et mint önmaga feletti modulusot. ${}_R R$ végesen generált, 1 generátoreleme; sőt bázisa. ${}_R R$ -nek részmodulusa a fenti I ideál. Könnyen látható, hogy ${}_R R$ -nek egyelemű bázisa csak K -beli konstans lehet; ami ugyanazt jelenti, mint ha az 1-et tekintjük. Az 1 elemnek akár véges sok r_1, \dots, r_n elemmel képezett többszörösei (vagyis maguk ezek az elemek) csak akkor vannak I -ben, ha ezeknek a konstans tagja 0. Mint láttuk, ezek nem generálják I -t. (Egyébként R felett nem is lineárisan függetlenek.)

Ez azt mutatja, hogy a 8.12. Tétel utáni 3. Következmény sem lehet igaz. Ha a „végtelen $\times \times 1$ ”-es mátrix soraiba az I ideál elemeit írjuk, akkor ezt a mátrixot véges sok lépésben nem tudjuk normálalakra hozni. \square

Feladatok

1. Mutassuk meg, hogy például a $\mathbb{Q}[x, y]$ feletti $\begin{bmatrix} x & y \\ y & x \end{bmatrix}$ mátrix elemi átalakításokkal nem hozható normálalakra. Mutassuk meg, hogy ha ennek volna normálalakja, akkor az $\begin{bmatrix} 1 & 0 \\ 0 & x^2 - y^2 \end{bmatrix}$ volna.

2. Bizonyítsuk be, hogy tetszőleges integritási tartomány feletti modulusban azok az elemek, amelyeket alkalmas nemnulla gyűrűbeli elemmel szorozva \mathbf{o} -t kapunk, részmodulusot alkotnak. Bizonyítsuk be, hogy euklideszi gyűrű esetében ez a részmodulus végesen generált, ha az eredeti az.

3. Bizonyítsuk be, hogy euklideszi gyűrű feletti minden végesen generált modulusra igaz az alaptétel, és azoknak a direkt összeadandóknak a száma is egyértelműen meghatározott, amelyeket egyetlen nemnulla gyűrűbeli elemmel szorozva sem kapunk \mathbf{o} -t.

4. Legyen R a \mathbb{Q} feletti kéthatározatlanú polinomok gyűrűje. Adjunk példát olyan \mathcal{M} végesen generált R -modulusra, amelyben létezik részmodulusok $\mathcal{M}_1 > \mathcal{M}_2 > \dots > \mathcal{M}_n > \dots$ végtelen lánc, hogy mindegyik \mathcal{M}_i generálható i elemmel, de egyik sem generálható i -nél kevesebb elemmel.

6. Mutassuk meg, hogy minden véges kommutatív csoport felírható prímmhatványrendű Abel-csoportok direkt összegeként; és az adott rendek, valamint a rögzített rendű tagok száma egyértelműen meghatározott (**Véges Abel-csoportok alaptétele**).

3. Mátrixpolinomok, invariáns faktorok

A továbbiakhoz előkészületül szükségünk van két, négyzetes polinommátrixokra vonatkozó tételre:

8.13. Tétel. *Legyenek $a_1(x), \dots, a_n(x)$ rendre a normálalakban megadott $A(x)$ polinommátrix diagonálisának elemei. Az $A(x)$ -hez és az $f(x)$ polinomhoz akkor és csak akkor létezik olyan $B(x)$ polinommátrix, amelyre $B(x)A(x) = f(x)I$, ha $a_n(x)$ osztója $f(x)$ -nek.*

Bizonyítás. Ha az oszthatósági kapcsolat fennáll, akkor az oszthatóság tranzitivitása alapján minden diagonális elem osztója $f(x)$ -nek. Léteznek tehát olyan $b_i(x)$ polinomok, amelyekre $f(x) = b_i(x)a_i(x)$. Ha mármost $B(x)$ -nek azt a diagonális mátrixot választjuk, amelyben a diagonális elemek rendre éppen ezek a polinomok, akkor a fenti szorzat pontosan a kívánt skalármátrixot adja.

Amennyiben létezik a fenti szorzattulajdonságot kielégítő $B(x) = [b_{i,j}(x)]$ polinommátrix, akkor a szorzat utolsó sorának utolsó elemére $b_{n,n}(x)a_n(x) = f(x)$ adódik, vagyis teljesül az oszthatósági feltétel. ■

8.14. Tétel. *A tetszőleges $A(x)$ polinommátrixhoz és az $f(x)$ polinomhoz akkor és csak akkor létezik olyan $B(x)$ polinommátrix, amelyre $B(x)A(x) = f(x)I$, ha $A(x)$ normálalakjában az utolsó sor utolsó eleme osztója $f(x)$ -nek.*

Bizonyítás. (A jobb áttekinthetőség kedvéért a bizonyítás során a határozatlan nem jelöljük.) Tudjuk, hogy az A mátrix normálalakját PAQ szorzatként kaphatjuk meg, ahol P és Q invertálható polinommátrixok. Tegyük fel először, hogy $B \cdot A = f \cdot I$. Ha P és Q tetszőleges invertálható polinommátrixok, akkor

$$(Q^{-1}BP^{-1}) \cdot (PAQ) = Q^{-1}(BA)Q = f \cdot (Q^{-1}Q) = f \cdot I,$$

felhasználva azt, hogy az f polinom skalár, és így a Q^{-1} mátrixszal felcserélhető.

Amennyiben a PAQ mátrixhoz található olyan B' mátrix, amelyre $B'PAQ = f \cdot I$, akkor az előző egyenlőségsorozat alapján található megfelelő B mátrix az $A = P^{-1}(PAQ)Q^{-1}$ mátrixhoz is (nevezetesen $B = QB'P$ megfelel). ■

Ezek után rátérünk a mátrixpolinomok vizsgálatára. A kiindulási észrevétel az, hogy az x határozatlan egy skalár.

8.15. Tétel. *Tetszőleges $K[x]$ feletti $A(x)$ (négyzetes) polinommátrixhoz egyértelműen léteznek olyan A_0, A_1, \dots, A_r K feletti mátrixok ($A_r \neq O$), amelyekre*

$$A(x) = A_0 + A_1x + \dots + A_rx^r$$

úgynevezett mátrixpolinom alakba írható. Minden mátrixpolinom egy polinommátrix mátrixpolinom alakja. A mátrixpolinomok a polinomműveletekre egy nemkommutatív gyűrűt alkotnak.

Az $A_r = I$ esetben r -edfokú normált mátrixpolinomról beszélünk.

Bizonyítás. Legyen $A(x) = \left[\sum_k a_{i,j}^{(k)} x^k \right]$. Ekkor az $A_k = [a_{i,j}^{(k)}]$ mátrixok nyilván

megfelelőek. Az is világos, hogy az adott $A(x)$ polinommátrixhoz csak ezek a mátrixok felelnek meg. Továbbá, tetszőlegesen választott K -ból vett elemű A_0, A_1, \dots, A_r mátrixokból a fenti módon elkészített $A_0 + A_1x + \dots + A_rx^r$ mátrix polinommátrixszá alakítható. Tekintettel arra, hogy a skalárok a mátrixokkal felcserélhetők és a mátrixok egy nemkommutatív gyűrűt alkotnak, ezért az utolsó állítás is igaz. ■

A mátrixpolinomok is rendelkeznek azzal a tulajdonsággal, hogy a határozatlan helyébe bármit beírhatunk, és ezzel egy homomorfizmust hozunk létre. Természetesen itt a képek nem egy kommutatív, hanem egy nemkommutatív gyűrű elemei lesznek. Az viszont gondot jelent, hogy a polinomszorzás definíciójában a határozatlan felcserélhető az együttműködőkkel. Ezért itt a kommutatív gyűrűk esetében talált tételnél csak egy gyengébb eredmény igaz.

8.7. Definíció. A polinomokkal kapcsolatos fogalmakat hasonlóképpen értelmezzük, mint a skaláregyűthetős esetben. A fenti $A(x)$ polinom B helyen vett helyettesítési értéke:

$$A(B) = A_0 + A_1B + \dots + A_rB^r,$$

ahol B ugyanakkora méretű mátrix, mint $A(x)$. Az $A(B) = O$ esetben B -t az $A(x)$ gyökének nevezzük. ■

Mint említettük, a szorzással gondok vannak. Valóban, ha például $A(x) = A \cdot x$, és $B(x) = B$, akkor $D(x) = A(x) \cdot B(x) = ABx$. Ekkor tetszőleges M mátrixra $A(M) \cdot B(M) = AMB$, míg $D(M) = ABM$, amelyek általában nem egyeznek meg. A továbbiakban szükségünk lesz ugyan arra, hogy a szorzat helyettesítési értéke megegyezik a helyettesítési értékek szorzatával, de szerencsére csak olyan esetben, amikor az igaz:

8.16. Tétel. *Legyenek $A(x), B(x)$ mátrixegyűthetős polinomok, és legyen $C(x) = A(x) + B(x)$, továbbá $D(x) = A(x) \cdot B(x)$. Ekkor tetszőleges azonos alakú M mátrixra $C(M) = A(M) + B(M)$, és ha M felcserélhető a $B(x)$ egyűthetőséivel, akkor $D(M) = A(M) \cdot B(M)$.*

Bizonyítás. Az összeadásra vonatkozó azonosságnál a szorzásról csupán annak az összeadásra vonatkozó disztributivitását használjuk fel; az tehát igaz. A szorzásnál a kommutativitást akkor használjuk, amikor azt írjuk fel, hogy $A_i x^i \cdot B_j x^j = (A_i B_j) x^{i+j}$. Az

asszociativitás alapján ehhez annyi elég, hogy $x^i B_j = B_j x^i$. Ha tehát ez igaz M -re, azaz $M^i B_j = B_j M^i$, akkor érvényes a szorzatra vonatkozó állítás. Ez az egyenlőség viszont nyilván következik az $M B_j = B_j M$ egyenlőségből. ■

A maradékos osztásra itt is hasonló eredmény igaz, mint a tetszőleges kommutatív gyűrűk feletti polinomokra:

8.17. Tétel. *Tetszőleges $A(x)$ és normált $B(x)$ polinommátrixokhoz léteznek olyan $Q(x)$ és $R(x)$ polinommátrixok, amelyekre*

$$A(x) = Q(x)B(x) + R(x),$$

ahol vagy $R(x) = O$, vagy $R(x)$ foka kisebb, mint $B(x)$ foka.

Bizonyítás. Tekintettel arra, hogy $B(x)$ normált, ezért nem a O mátrix, így van foka. A bizonyítás hasonló módon történik, mint a kommutatív esetben, azaz rögzített $B(x)$ esetén $A(x)$ fokára való teljes indukcióval:

Ha $A(x)$ foka kisebb $B(x)$ fokánál, vagy $A(x) = O$, akkor választhatjuk a $Q(x) = O$ és $R(x) = A(x)$ mátrixokat.

Legyen $B(x) = B_0 + \dots + I \cdot x^k$, $A(x) = A_0 + \dots + A_n x^n$, ahol $n \geq k$, és tegyük fel, hogy a felírás érvényes minden olyan esetben, amikor az $A(x)$ foka kisebb, mint n . Tekintsük az $A'(x) = A(x) - A_n \cdot x^{n-k} \cdot B(x)$ polinommátrixot. A jobb oldalon álló első tag foka n . A jobb oldalon álló második tag foka is n , hiszen a szorzat foka a tényezők fokainak az összegével egyenlő. E második tag főegyütthatója is A_n (mint az első tagé), hiszen szorzásnál a főegyütthatók (a megfelelő sorrendben!) összeszorozódnak. Ezért az $A'(x)$ foka kisebb, mint n , tehát felírható $A'(x) = Q'(x)B(x) + R'(x)$ alakban. Ekkor viszont az eredeti $A(x)$ mátrixra is megkapható a kívánt felírás, nevezetesen a $Q(x) = A_n \cdot x^{n-k} + Q'(x)$ és $R(x) = R'(x)$ választással. ■

Megjegyzések

1. Tekintettel arra, hogy minden mátrix osztható a K test bármely 0-tól különböző elemével, ezért a 8.17. Tételben $B(x)$ helyett $c \cdot B(x)$ is vehető, ahol $c \in K$ és $c \neq 0$.
2. A másik oldali maradékos osztás is ugyanígy elvégezhető. A továbbiakban azonban csak erre lesz szükségünk. □

8.18. Tétel. *Az A mátrix akkor és csak akkor gyöke az $F(x)$ mátrixegyütthatós polinomnak, ha létezik olyan $Q(x)$ polinommátrix, amelyre*

$$F(x) = Q(x) \cdot (A - I \cdot x).$$

Bizonyítás. A 8.17. Tétel szerint tetszőleges $F(x)$ polinommátrixhoz létezik olyan $Q(x)$ polinommátrix és R konstans mátrix, amelyre

$$F(x) = Q(x) \cdot (A - I \cdot x) + R,$$

hiszen $A - I \cdot x$ egy normált elsőfokú (-1) -szerese (ez önmagába foglalja az $R = O$ esetet is). Tekintettel arra, hogy A felcserélhető a polinom együtthatóival, ezért a 8.16. Tétel szerint $F(A) = R$, hiszen $A - A = O$. Amennyiben tehát $F(A) = O$, akkor létezik a kívánt

felbontás. Ha viszont létezik a kívánt felbontás, akkor már a 8.16. Tételből következik, hogy $F(A) = O$. ■

8.19. Tétel. Legyen α az n -dimenziós tér egy lineáris transzformációja, és A ennek valamely bázisban felírt mátrixa. Az $A - I \cdot x$ mátrix normálalakja nem függ a bázistól. E normálalak diagonálisának $a_i(x)$ elemei (ezek neve megfelelően az i -edik invariáns faktor) ugyancsak függetlenek a bázistól. Az $a_n(x)$ elem az α transzformáció (és az A mátrix) minimálpolinomja. Ennek többszöröse a $\det(A - I \cdot x)$ determináns, amely megegyezik a transzformáció (és a mátrix) karakterisztikus polinomjával — és ez ugyancsak nem függ a bázistól. A karakterisztikus polinom n -edfokú. Az α transzformáció és az A mátrix gyöke a karakterisztikus polinomnak (CAYLEY–HAMILTON-tétel).

Bizonyítás. A 4.7. Tétel kiegészítése szerint α mátrixa egy másik bázisban $T^{-1}AT$ alakú. Ebben a bázisban a vizsgált polinommátrix

$$T^{-1}AT - I \cdot x = T^{-1}(A - I \cdot x)T,$$

amelynek a 8.12. Tétel szerint ugyanaz a normálalakja, mint az eredetinek. Ezért az invariáns faktorok valóban nem függenek a bázistól.

Skaláregyütthatós polinomokra $f(\alpha)$ mátrixa megegyezik az $f[A] = [f(A)]$ mátrixszal. A 8.18. Tétel szerint A akkor és csak akkor gyöke az $f(x)$ polinomnak, ha van olyan $Q(x)$ polinommátrix, amire $f(x) \cdot I = Q(x) \cdot (A - I \cdot x)$. A 8.13. Tétel szerint ez pontosan akkor következik be, ha $a_n(x)$ osztója $f(x)$ -nek, ami azt jelenti, hogy $a_n(x)$ az A (s az α) minimálpolinomja. A karakterisztikus polinom éppen az n -edik determinánsosztó, így ez sem függ a bázistól. Mivel ez éppen az elemi osztók szorzata, azért A ennek is gyöke. ■

Megjegyzés. Az, hogy egy mátrix gyöke a karakterisztikus polinomjának, viszonylag egyszerű számítással is igazolható. Ez a számítás viszont nem adja az elemi osztók tulajdonságait.

Célszerű viszont felhívni a figyelmet az alábbiakra: Az A mátrix gyöke ugyan a $\det(A - I \cdot x)$ polinomnak, de itt x helyébe formálisan A -t írva nem kapunk azonosan O -t, hiszen az $A - I \cdot x$ mátrixban az x helyébe A -t írva egy olyan mátrixhoz jutunk, amelynek a fődiagonálisában is mátrixok állnak, s az eredmény teljesen áttekinthetetlen. □

Feladatok

1. Bizonyítsuk be, hogy az $A - I \cdot x$ polinommátrix invariáns faktoraik között nem szerepelhet a 0 polinom.

2. Van-e olyan transzformáció, amelynek a karakterisztikus polinomja megegyezik minimálpolinomjával?

3. Határozzuk meg a komplex számtest feletti normális, önadjungált és unitér transzformációk, illetve a valós számtest feletti szimmetrikus és ortogonális transzformációk invariáns faktoraikat.

4. Hozzuk az $A - I \cdot x$ mátrixot normálalakra, ha:

1. A diagonális mátrix.
2. Az A mátrix egyetlen eleme 1, s a többi 0.

5. Legyen α a \mathbb{Q} feletti n -dimenziós vektortér lineáris transzformációja. Milyen k esetén lehet α^k az identitás, illetve az ω transzformáció?

6. Egy polinommátrix normálalakjában nincs két elem, amely csak konstansban térne el egymástól. Mit tudunk mondani a mátrix elemeinek a fokáról?

7. Melyek a kétdimenziós tér azon transzformációi, amelyek karakterisztikus polinomja nem egyezik meg minimálpolinomjával?

8. Bizonyítsuk be, hogy az n -dimenziós tér bármely transzformációjának legfeljebb $\frac{1}{2}(1 + \sqrt{8n+1})$ különböző invariáns faktora van.

9. Legyenek $\ell < k \leq n$ rögzített természetes számok. Adjunk meg az n -dimenziós térben olyan α transzformációt, amelynek minimálpolinomja $x^k - x^\ell$. Igaz-e, hogy bármely két ilyen transzformáció hasonló? (α és β hasonlóak, ha „ugyanúgy hatnak, de más bázison”; azaz, ha $\beta = \sigma^{-1}\alpha\sigma$ alakú.)

10. Legyen α az \mathcal{U} vektortér lineáris transzformációja. Tetszőleges $\mathbf{u} \in \mathcal{U}$ vektorhoz rendeljük hozzá azt a minimális fokú $g_{\mathbf{u}}(x)$ polinomot, amelyre $g_{\mathbf{u}}(\alpha) : \mathbf{u} \mapsto \mathbf{o}$. Bizonyítsuk be, hogy ez értelmes definíció. Mennyi ezeknek a polinomoknak (asszociáltaktól eltekintve) a száma? Milyen kapcsolatban állnak ezek a polinomok a minimálpolinommal?

11. Legyen α az \mathcal{U} vektortér lineáris transzformációja. Tetszőleges $g(x)$ polinomhoz rendeljük hozzá azt a $\mathcal{V} \subseteq \mathcal{U}$ halmazt, amelynek elemeire $g(\alpha)(\mathbf{v}) = \mathbf{o}$. Milyen kapcsolatot létesít ez és az előző példában megadott hozzárendelés a polinomgyűrű és a vektortér között?

12. Jellemezzük azokat a lineáris transzformációkat, amelyeknek a minimálpolinomja megegyezik karakterisztikus polinomjukkal.

4. A Jordan-féle normálalak

A komplex euklideszi tér normális lineáris transzformációinak mátrixa alkalmas bázisban diagonálissá vált. A valós esetben ez nem volt igaz: egy forgatás mátrixa általában nem lehet diagonális. Ennek oka az, hogy e transzformációnak nincs sajátvektora, mert karakterisztikus polinomja irreducibilis. Tekintettel arra, hogy a komplex számtest feletti tetszőleges polinomnak van komplex gyöke, ezért egy transzformációnak létezik sajátértéke, így sajátvektora. Ennek ellenére a komplex számtest feletti vektortérben vannak olyan transzformációk, amelyek mátrixa nem hozható diagonális alakra.

Erre mutatunk egy tipikus példát. Majd látni fogjuk, hogy más, ettől lényegesen eltérő eset nem is lehet. Legyen $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ a komplex számtest feletti \mathcal{U} vektortér egy bázisa, és definiáljuk e tér ϑ lineáris transzformációját a következőképpen: $i < n$ esetén legyen $\vartheta(\mathbf{u}_i) = \mathbf{u}_{i+1}$, és legyen $\vartheta(\mathbf{u}_n) = \mathbf{o}$. Világos, hogy e transzformációt n -szer alkalmazva az ω transzformációt kapjuk, azaz ϑ gyöke az x^n polinomnak. Ebből következik, hogy minimálpolinomja osztója ennek a polinomnak (tekintettel arra, hogy $\vartheta^{n-1} \neq \omega$,

ezért x^n pontosan a minimálpolinom). Mivel egy transzformáció minden sajátértéke gyöke a minimálpolinomnak, ezért ϑ -nak egyetlen sajátértéke van, a 0. Amennyiben ϑ mátrixa egy bázisban diagonálissá válik, akkor a diagonális elemei a sajátértékek, ezért csak a $\vartheta = \omega$ eset lehet, ami csak az $n = 1$ esetben lehetséges. E transzformáció mátrixát a fenti bázisban felírva, a mátrix i -edik oszlopának $(i + 1)$ -edik eleme 1 ($i < n$), s a többi elem 0.

E transzformációból (a legalább kétdimenziós térben) több olyan transzformáció is készíthető, amely nem hozható diagonális alakra. Tetszőleges c komplex szám esetén ilyen a $\vartheta + cI$ transzformáció is. Ha ugyanis ennek a mátrixa valamely bázisban diagonális volna, akkor e bázisban ϑ mátrixa is diagonális lenne, tekintettel arra, hogy $-cI$ mátrixa minden bázisban diagonális.

Ebből további hasonló jellegű transzformációkat tudunk készíteni. Legyen $\mathcal{U} = \mathcal{U}_1 \oplus \dots \oplus \mathcal{U}_k$, és legyen ϑ_i az \mathcal{U}_i altéren hasonlóképpen értelmezve, mint fent. Legyen továbbá ι_i az \mathcal{U}_i tér identitása, és c_i tetszőleges komplex szám. Ekkor az $\alpha_i = \vartheta_i + c_i \iota_i$ transzformációkra az $\alpha = \alpha_1 \oplus \dots \oplus \alpha_k$ transzformáció egyetlen bázisban sem hozható ennél „diagonálisabb” alakra. A továbbiakban azt mutatjuk meg, hogy ilyen alak alkalmas bázisban minden transzformációnál elérhető.

8.20. Tétel. *Legyen α a komplex számtest feletti \mathcal{U} vektortér egy lineáris transzformációja. Ekkor \mathcal{U} előáll α invariáns altereinek $\mathcal{U} = \mathcal{U}_1 \oplus \dots \oplus \mathcal{U}_k$ direkt összegeként úgy, hogy α -nak ezekre az alterekre való α_i megszorítására a következő igaz:*

Az \mathcal{U}_i altéren van olyan ϑ_i transzformáció és olyan \mathbf{u}_i vektor, hogy valamely n_i természetes számra az $\mathbf{u}_i, \vartheta_i \mathbf{u}_i, \dots, \vartheta_i^{n_i-1} \mathbf{u}_i$ vektorok az \mathcal{U}_i tér egy bázisát alkotják és $\vartheta_i^{n_i} \mathbf{u}_i = \mathbf{o}_i$; továbbá van olyan c_i komplex szám, hogy az \mathcal{U}_i altér ι_i identitásával $\alpha_i = \vartheta_i + c_i \iota_i$ teljesül.

Természetesen $\alpha = \alpha_1 \oplus \dots \oplus \alpha_k$.

Bizonyítás. Mint tudjuk, azzal a $\Phi : \mathbb{C}[x] \rightarrow \text{End}(\mathcal{U})$ homomorfizmussal, amely \mathbb{C} elemeinek a velük való szorzást és x -nek az α transzformációt felelteti meg, az \mathcal{U} vektortér modulussá válik $\mathbb{C}[x]$ felett. Mivel $\mathbb{C}[x]$ euklideszi gyűrű, ezért érvényes rá a 8.12. Tétel. Ennek 5. Következménye alapján létezik \mathcal{U} -nak α invariáns alterei direkt összegére való olyan $\mathcal{U} = \mathcal{U}_1 \oplus \dots \oplus \mathcal{U}_k$ felbontása, amelyben a megfelelő α_i megszorításokra a következő igaz. Az α_i -nek az $m_i(x)$ minimálpolinomja egyetlen irreducibilis $p_i(x)$ polinom hatványa ($m_i(x) = p_i(x)^{n_i}$), és van olyan \mathbf{u}_i elem, amelynek a rendje $m_i(x)$; továbbá ezek \mathcal{U} -nak — mint modulusnak — a bázisát alkotják. Ez azt jelenti, hogy $\sum_i a_i \mathbf{u}_i = \mathbf{o}$ csak akkor igaz, ha minden i -re $a_i \mathbf{u}_i = \mathbf{o}$ teljesül (itt $a_i \in \mathbb{C}[x]$). Ebből következik, hogy az $\alpha_i^j(\mathbf{u}_i)$ elemek ($j \geq 0$) generálják \mathcal{U}_i -t. Tekintettel arra, hogy \mathbb{C} felett csak az elsőfokú polinomok irreducibilisek, ezért $p_i(x) = x - c_i$ alkalmas $c_i \in \mathbb{C}$ számmal. (Több indexhez is tartozhat ugyanaz a c_i szám és ugyanaz az n_i kitevő.)

(A következő részben a könnyebb áttekinthetőség végett elhagyjuk az indexeket.) Tekintsük az \mathcal{U}_i tér $\vartheta_i = \alpha_i - c_i \iota_i$ transzformációját, ahol ι_i identikus transzformáció az \mathcal{U}_i -n. Mivel α minimálpolinomja $(x - c)^n$, ezért ϑ minimálpolinomja x^n . Eszerint $\vartheta^n = \omega$,

de $\vartheta^{n-1} \neq \omega$. Van tehát olyan \mathbf{u} vektor, amelyre $\vartheta^{n-1}(\mathbf{u}) \neq \mathbf{o}$, de $\vartheta^n(\mathbf{u}) = \mathbf{o}$. Ekkor a $\vartheta^j(\mathbf{u})$ vektorok ($0 \leq j < n$) függetlenek. A feltétel szerint e vektorok generátorrendszert alkotnak; azaz bázist. ■

Következmény. Jordan-féle normálalak. Legyen α a komplex számtest feletti n -dimenziós tér egy lineáris transzformációja, és legyenek α sajátértékei a c_1, \dots, c_k számok. Ekkor van a térnek olyan bázisa, amelyben mátrixa az alábbi alakot ölti:

$$\begin{bmatrix} A_1 & O & \dots & O \\ O & A_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & A_n \end{bmatrix}; A_i = \begin{bmatrix} A_{i,1} & O & \dots & O \\ O & A_{i,2} & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & A_{i,n_i} \end{bmatrix}; A_{i,j} = \begin{bmatrix} c_i & 0 & \dots & 0 & 0 \\ 1 & c_i & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & c_i & 0 \\ 0 & 0 & \dots & 1 & c_i \end{bmatrix},$$

ahol az első mátrix olyan blokkokból áll, mint amilyent a második mutat és ennek blokkjait a harmadikban láthatjuk. Az egyes A_i mátrixokban ugyanazok a sajátértékek szerepelnek, és az $A_{i,j}$ mátrixokról feltehető, hogy növekvő j mellett a méreteik nem nőnek.

A harmadik típusú mátrixokat Jordan-blokkoknak nevezzük.

Bizonyítás. A megfelelő bázist „blokkokból” állítjuk elő. Az első lépésben tekintsük az α minimálpolinomjának páronként relatív prím faktorokra való maximális felbontását. Az ezekhez tartozó invariáns alterekre megszorítva a kapott transzformációk minimálpolinomja egy-egy lineáris polinom hatványa lesz, minden esetben más és más lineáris polinomé (alaptétel I.). A kapott mátrix olyan lesz, mint az elsőnek felírt mátrix, függetlenül attól, hogy ezekben az alterekben milyen bázist veszünk fel.

A második lépésnél az alteret tovább bontjuk olyan alterekre, amelyeket modulusként egyetlen elem generál (alaptétel II.). Az előzőhöz hasonló mátrixfelbontást kapunk, azzal a különbséggel, hogy itt már mindig ugyanaz lesz a lineáris faktor. Ezeket az altereket (illetve a megfelelő „bázisblokkokat”) úgy rakhatjuk sorba, hogy a felsorolásnál a dimenziók ne nőjenek.

A harmadik lépésnél a 8.20. Tételben adott ϑ transzformáció segítségével kapjuk meg a mátrix alakját. ■

Megjegyzés. A Jordan-féle normálalak igen hasznos akkor, ha a mátrix (vagy a transzformáció) valamilyen polinomját akarjuk kiszámítani. A transzformációk direkt összegének a definíciójánál láttuk, hogy a direkt összeg az összeadással is és a szorzással is felcserélhető. Tetszőleges $f(x)$ polinomra fennáll tehát az $f(\alpha_1 \oplus \dots \oplus \alpha_r) = f(\alpha_1) \oplus \dots \oplus f(\alpha_r)$ összefüggés. Ez azt jelenti, hogy az $f(A)$ kiszámításánál csak a Jordan-blokkok polinomját kell meghatározni. Könnyen kiszámolható, hogy tetszőleges c konstansra

$$f(x+c) = f(c) + f'(c)x + \dots + \frac{f^{(i)}(c)}{i!}x^i + \dots,$$

ahol $f^{(i)}$ az i -edik deriváltat jelöli. Az $A = T + c \cdot I$ behelyettesítéssel ($T = [\vartheta]$) a következőt kapjuk:

$$f(A) = f(T + c \cdot I) = f(c) \cdot I + f'(c) \cdot T + \dots + \frac{f^{(i)}(c)}{i!} \cdot T^i + \dots$$

Látható, hogy itt az eredeti mátrixtól függetlenül egyetlen T mátrix hatványait kell meghatározni, majd ezeket kell megszorozni egy-egy A -tól függő skalárral, és a kapott mátrixokat összeadni.

A T -hez tartozó ϑ transzformáció a tekintetbe vett bázis minden elemét a következőbe viszi, s az utolsót \mathbf{o} -ba. ϑ -nak az i -edik hatványa tehát minden bázisvektort az i -vel nagyobb indexűbe visz, s ha az indexek „elfogytak”, akkor \mathbf{o} -ba. Ezért T^i mátrixában a j -edik oszlop $(i + j)$ -edik eleme 1, s az összes többi elem 0. Ezek a mátrixok „egymást nem zavarják”, s így az $f(A)$ könnyen meghatározható. Példaként felírunk egy ilyen mátrixot az ötdimenziós tér esetére:

$$\begin{bmatrix} f(c) & 0 & 0 & 0 & 0 \\ f'(c) & f(c) & 0 & 0 & 0 \\ \frac{f''(c)}{2} & f'(c) & f(c) & 0 & 0 \\ \frac{f'''(c)}{6} & \frac{f''(c)}{2} & f'(c) & f(c) & 0 \\ \frac{f^{(4)}(c)}{24} & \frac{f'''(c)}{6} & \frac{f''(c)}{2} & f'(c) & f(c) \end{bmatrix}$$

Érdemes észrevenni, hogy akármeekkora fokú polinom esetében is legfeljebb annyi tagot kell figyelembe venni, mint a tér dimenziója. \square

8.21. Tétel. A Jordan-féle normálalak egyértelműen meghatározott.

Bizonyítás. Induljunk ki egy α transzformáció „valamelyik” A Jordan-féle normálalakjából és határozzuk meg az $A - x \cdot I$ polinommátrix normálalakját. Tudjuk, hogy ez nem függ attól, hogy a transzformáció mátrixát melyik bázisban írtuk fel. Ha tehát a polinommátrix normálalakjából egyértelműen vissza tudunk következtetni a Jordan-féle normálalakra, akkor ez is egyértelműen meghatározott. Az $A - x \cdot I$ polinommátrix fődiagonalizálásában a $c_i - x$ polinomok állnak, ahol a c_i számok az α sajátértékei. Avégett, hogy az amúgy is sok paramétert tartalmazó bizonyítást valamivel áttekinthetőbbé tegyük, ezt a polinomot p_i -vel fogjuk jelölni.

Tegyük fel, hogy az eredeti A mátrixnak n sora és oszlopa van. Ez a mátrix felbomlik az A_1, \dots, A_k mátrixok direkt összegére, ahol az A_i mátrixnak n_i sora van és a fődiagonalisban mindenütt p_i áll. Itt $n = n_1 + \dots + n_k$.

Az A_i mátrix tovább bomlik az $A_{i,1}, \dots, A_{i,r_i}$ mátrixok direkt összegére, ahol az $A_{i,s}$ mátrixnak $n_{i,s}$ sora van, és ezekre $n_{i,1} \geq \dots \geq n_{i,r_i}$, és $n_i = n_{i,1} + \dots + n_{i,r_i}$.

Feladatunk annak a bizonyítása, hogy az itt szereplő n, k, n_i, n_{i,r_i} számok és a p_i polinomok egyértelműen visszakaphatók e mátrix normálalakjából.

Tegyük fel, hogy e mátrix t -edik determinánsosztóját akarjuk meghatározni. Ehhez szükségünk van a mátrix t -edrendű aldeterminánsaira. Tekintettel arra, hogy a 0 mindennek többszöröse, ezért számunkra csak a nemnulla determinánsok lényegesek.

Nézzük meg először, hogy két mátrix direkt összegében miképpen kaphatunk t -edrendű nemnulla aldeterminánst. Tegyük fel, hogy az első mátrixnak ℓ sora van és a másodiknak $n - \ell$. Legyen a t -edrendű mátrix olyan, hogy az első p sorának és első q oszlopának az indexe nem nagyobb, mint ℓ . Ha $p < q$, akkor ez a mátrix $\begin{bmatrix} U & V \\ O & W \end{bmatrix}$ alakú, ahol U és W négyzetes mátrixok, O -nak minden eleme 0, U -nak q sora és oszlopa van és

W -nek $t - q$; továbbá U -nak az utolsó $q - p$ sora csupa 0. Ekkor e mátrix determinánsa a $\det(U)$ és $\det(W)$ szorzata, s mivel az első tényező 0, így az egész szorzat az. Hasonló a helyzet a $p > q$ esetben is. Ezért csak akkor kaphatunk nemnulla determinánst, ha az ℓ -edrendű és az $(n - \ell)$ -edrendű két részmátrix egy-egy megfelelő méretű mátrixának a determinánsát szorozzuk össze.

Ebből induktíven azonnal következik, hogy hasonlóképpen állíthatók elő a t -edrendű részmátrixok determinánsai többtagú direkt összeg esetében is.

Nézzük meg most, hogy mivel lesznek egyenlők az egyes $A_{i,j}$ mátrixokhoz tartozó determinánsok. Itt egyetlen $n_{i,j}$ -rendű részmátrix van (maga az egész mátrix), ennek a determinánsa a fődiagonális elemeinek a szorzata, azaz p_i -nek az $n_{i,j}$ -edik hatványa. Ha $p < n_{i,j}$, akkor tekintsük azt a részmátrixot, amit úgy kapunk, hogy az első sort elhagyva vesszük a következő p sort, továbbá az első p darab oszlopot. Ebben a mátrixban a fődiagonális minden eleme 1, alatta pedig minden elem 0. Ezért ennek a mátrixnak a determinánsa 1. Tekintettel arra, hogy feladatunk ilyen determinánsok legnagyobb közös osztóját meghatározni, ezért csak ezt a két lehetőséget kell figyelembe venni.

Ha az összes blokkból elhagyunk egy sort, akkor olyan mátrixot kapunk, amely determinánsának minden „blokk-tényezője” 1. Ezért az ekkora méretű mátrixok determinánsának a legnagyobb közös osztója is 1. Ugyanez érvényes akkor is, ha ennél több sort hagyunk el.

Nézzük meg, mi történik akkor, ha ennél kevesebb sort hagyunk el. Tekintsük azt az esetet, amikor csak az A_i -ből hagyunk el néhány sort. Ha ugyanabból az $A_{i,j}$ -ből hagyunk el két sort, akkor a kapott determinánsban ez egy darab 1-es tényezőt jelent, míg, ha egy másiktól is, az két ilyen tényezőt hoz létre. Ez azt jelenti, hogy az utóbbi osztója az előzőnek. Mivel a legnagyobb közös osztót kell meghatározni, ezért elég azokat az eseteket nézni, amikor minden egyes blokkból csak egyetlen sort hagyunk el. De ezek között még további oszthatósági reláció áll fenn. Az A -nak a determinánsa $p_i^{n_i}$. Ha az $A_{i,j}$ -ből hagyunk el egy sort, akkor a kitevő $n_{i,j}$ -vel csökken. Ezek legnagyobb közös osztóját akkor nyerjük, ha a kitevő a lehető legnagyobb, azaz $n_i - n_{i,1}$.

Ezáltal egy olyan $(n - 1)$ -edrendű mátrixot nyertünk, amelyben p^i az $n_i - n_{i,1}$ kitevővel szerepel, míg az összes többi j -re p_j az n_j -ediken. Az $(n - 1)$ -edik determinánsosztóban, Δ_{n-1} -ben, ezek legnagyobb közös osztójában, minden p_i a lehető minimális kitevővel szerepel, azaz $n_i - n_{i,1}$ kitevővel. A következő determinánsosztó, Δ_{n-2} úgy kapható, hogy a következő méretű mátrixnak is elhagyjuk egy sorát, azaz a p_i kitevője most már $n_{i,2}$ -vel is csökken. Az eljárást tovább folytatva a megfelelő determinánsosztókban a kitevők egyre csökkennek.

Célszerű a determinánsosztók helyett az elemi osztókat figyelembe venni. Az n -edik elemi osztó $\frac{\Delta_n}{\Delta_{n-1}}$ alakja:

$$p_1^{n_{1,1}} \cdot p_2^{n_{2,1}} \cdot \dots \cdot p_k^{n_{k,1}}.$$

Hasonlóképpen az $(n - 1)$ -edik elemi osztó:

$$p_1^{n_{1,2}} \cdot p_2^{n_{2,2}} \cdot \dots \cdot p_k^{n_{k,2}},$$

és így tovább. Természetesen, ha valamelyik kitevő már nem lép fel, akkor a megfelelő hatvány itt már nem jelentkezik.

Ezek a polinomok már egyértelműen meghatározzák a lineáris faktorokat és a fellépő blokkok méreteit. ■

Feladatok

1. Bizonyítsuk be, hogy ha két transzformációnak a megfelelő invariáns faktori megegyeznek, akkor a transzformációk hasonlóak.

2. Mutassuk meg, hogy két transzformációnak akkor is megegyezhet a minimálpolinomja és a karakterisztikus polinomja, ha a transzformációk nem hasonlóak.

3. Mutassuk meg, hogy „éppen nem lehet látni” olyan nem hasonló transzformációkat, amelyeknek minimálpolinomjuk és karakterisztikus polinomjuk is megegyezik (azaz hány dimenziós valós térben lehetséges ez?).

4. Bizonyítsuk be, hogy egy transzformáció mátrixának minden bázisban ugyanannyi a nyoma.

5. Mi a kapcsolat egy transzformáció mátrixában a szimmetrikusan elhelyezkedő részmátrixok determinánsai és a karakterisztikus polinom között?

6. Legyen $f(x)$ tetszőleges polinom és A négyzetes mátrix. Mutassuk meg, hogy ha P és Q ugyanekkora méretű invertálható mátrixok, akkor $f(A) = P^{-1}f(PAQ)Q^{-1}$. Miképpen hasznosítható ez tetszőleges mátrix polinomjának a kiszámítására?

A továbbiakban csak adott méretű négyzetes mátrixokat tekintünk.

Azt mondjuk, hogy az $A_k = [a_{i,j}^{(k)}]$ mátrixok ($k \in \mathbb{N} \cup \{0\}$) sorozata tart (konvergál) az $A = [a_{i,j}]$ mátrixhoz (jelben $\lim_{k=0}^{\infty} A_k = A$), ha tetszőleges i, j indexpárra $\lim_{k=0}^{\infty} a_{i,j}^{(k)} = a_{i,j}$.

Azt mondjuk, hogy a $\sum_{k=0}^{\infty} A_k$ sor összege az A mátrix, ha a $B_k = B_0 + \dots + B_k$ mátrixok sorozata tart A -hoz.

Tetszőleges $f(x) = \sum_{k=0}^{\infty} c_k x^k$ hatványsor esetén formálisan tekinthetjük az $f(A) = \sum_{k=0}^{\infty} c_k A^k$ hatványsort. A hatványsor konvergencia az A helyen, ha ez a sor konvergencia.

7. Bizonyítsuk be, hogy bármely hatványsor konvergencia az A helyen, ha A -nak minden sajátértéke 0.

8. Adjunk az A sajátértékeire vonatkozó feltételt arra, hogy egy adott hatványsor az A helyen konvergencia legyen.

9. Defináljuk az e^A , $\sin A$ és $\cos A$ mátrixokat.

10. Határozzuk meg az $(I - A)^{-1}$ hatványsorát. Milyen A esetén konvergál ez az $I - A$ inverzéhez?

11. Milyen A mátrixokra létezik $\log(A)$?

KILENCEDIK FEJEZET

DETERMINÁNSOK ALKALMAZÁSA

1. Lineáris egyenletrendszerek megoldása

A lineáris egyenletrendszereket már tárgyaltuk a lineáris algebra bevezetésekor, pontosabban a bevezetése előtt. Akkor megmutattuk, miképpen lehet ezeket a Gauss-féle eliminációval megoldani. Noha ez a legegyszerűbb és legjobb megoldási eljárás, mégis érdemes a lineáris egyenletrendszereket a lineáris algebrai módszerek segítségével is megvizsgálni. Itt elsősorban azt szeretnénk megnézni, hogy milyen feltétel mellett oldható meg az egyenletrendszer.

A lineáris egyenletrendszereket a következő formában szokták megadni:

$$a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,k}x_k = b_1$$

$$a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,k}x_k = b_2$$

$$\vdots$$

$$(*) \quad a_{n,1}x_1 + a_{n,2}x_2 + \dots + a_{n,k}x_k = b_n,$$

ahol az $a_{i,j}$ -k és a b_i -k adott számok és az x_j -k ismeretlenek. Maga a felírás azt jelenti, hogy keressünk a ismeretlenek helyébe a szóban forgó egyenlőséget kielégítő számokat, az egyenletrendszer *megoldásait*.

Az $a_{i,j}$ *együtthatókat* és a b_i *konstansokat* valamilyen (szám)testből vesszük, és a megoldásokat is ebben a testben keressük. (Előfordulnak olyan esetek is, amikor test helyett gyűrű szerepel; például egész együtthatós egyenletrendszerek egész megoldásainak a keresése.) A megoldhatóság vizsgálata végett célszerűbb az egyes oszlopokat oszlopvektorok alakjában felírni:

$$\mathbf{a}_1 = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{n,2} \end{bmatrix}, \dots, \quad \mathbf{a}_k = \begin{bmatrix} a_{1,k} \\ a_{2,k} \\ \vdots \\ a_{n,k} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Ezek után az adott egyenletrendszer a következő alakot ölti:

$$(**) \quad x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_k\mathbf{a}_k = \mathbf{b}.$$

Ekkor úgy tekinthetjük, hogy az \mathbf{a}_i és \mathbf{b} vektorok egy adott n -dimenziós térnek az elemei. $A(*)$ egyenletrendszer megoldhatósága azt jelenti, hogy alkalmas x_i számokra teljesül $(**)$. Megfogalmazva:

9.1.A Tétel. $A(*)$ egyenletrendszer akkor és csak akkor oldható meg, ha a \mathbf{b} vektor benne van az \mathbf{a}_i vektorok generálta altérben. ■

Ezt a feltételt átírhatjuk vektorrendszerre. Eszerint a megoldhatóságnak az a feltétele, hogy \mathbf{b} függjön az $\mathbf{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ vektorrendszertől; más szóval az, hogy az \mathbf{A} és az $\mathbf{A} \cup \{\mathbf{b}\}$ vektorrendszerek rangja megegyezzen. Ezt a feltételt célszerű mátrixokra átforgalmazni. Legyen $A = [[\mathbf{a}_1], \dots, [\mathbf{a}_k]]$ és $B = [[\mathbf{a}_1], \dots, [\mathbf{a}_k], [\mathbf{b}]]$. A -t az *egyenletrendszer mátrixának* és B -t a *kibővített mátrixnak* nevezzük. (B úgy áll elő A -ból, hogy utolsó oszlopnak még odaírjuk a $[\mathbf{b}]$ oszlopmátrixot.)

9.1.B Tétel. $A(*)$ egyenletrendszer akkor és csak akkor oldható meg, ha az egyenletrendszer mátrixának és a kibővített mátrixnak megegyezik a rangja, azaz $r(A) = r(B)$. ■

Lehetőség van még tömörebb megfogalmazásra: Tekintsük a k -dimenziós \mathcal{U} és n -dimenziós \mathcal{V} vektortereket, és legyen $\alpha : \mathcal{U} \rightarrow \mathcal{V}$ az a lineáris leképezés, amely az \mathcal{U} vektortér $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ bázisának elemein a következőképpen hat: $\alpha(\mathbf{u}_i) = \mathbf{a}_i$. Ha van az egyenletrendszernek megoldása, ez azt jelenti, hogy az \mathcal{U} tér $\mathbf{x} = x_1\mathbf{u}_1 + \dots + x_k\mathbf{u}_k$ vektorára $\alpha(\mathbf{x}) = x_1\mathbf{a}_1 + \dots + x_k\mathbf{a}_k = \mathbf{b}$. Feladatunk tehát olyan \mathbf{x} vektort találni, amelyre $\alpha(\mathbf{x}) = \mathbf{b}$. Itt α is és \mathbf{b} is mátrixával adottak, $[\alpha] = A = [a_{ij}]$ és $[\mathbf{b}] = [b_i]$ (\mathbf{x} mátrixa $[\mathbf{x}] = [x_j]$). A megoldhatóság feltétele, hogy létezzék a most megfogalmazott feltételt kielégítő \mathbf{x} vektor:

9.1.C Tétel. $A(*)$ egyenletrendszer akkor és csak akkor oldható meg, ha a \mathbf{b} vektor benne van az α képterében, azaz $\mathbf{b} \in \text{Im}(\alpha)$. ■

Általában egy lineáris egyenletrendszernek több megoldása is van (megoldáson természetesen megoldásvektort értünk). Vegyük szemügyre, mi a feltétele annak, hogy a megoldás egyértelmű legyen. Ha \mathbf{x} is és \mathbf{y} is megoldások, akkor $\alpha(\mathbf{x} - \mathbf{y}) = \alpha(\mathbf{x}) - \alpha(\mathbf{y}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$, vagyis $\mathbf{x} - \mathbf{y} \in \text{Ker}(\alpha)$. Fordítva, tetszőleges $\mathbf{z} \in \text{Ker}(\alpha)$ és \mathbf{x} megoldás esetén világos, hogy $\mathbf{x} + \mathbf{z}$ is megoldás. Eszerint a megoldás egyértelműsége azt jelenti, hogy α magtere egyedül a nullvektorból álljon. Ez akkor következik be, ha α független vektorokat független vektorokba visz; speciálisan, ha a felvett \mathbf{U} bázis vektorainak az $\mathbf{a}_1, \dots, \mathbf{a}_k$ képei lineárisan függetlenek. A lineáris függetlenség viszont azt jelenti, hogy $r(A) = k$. Mátrixokkal megfogalmazva:

9.1.D Tétel. $A(*)$ egyenletrendszernek akkor és csak akkor van egyértelmű megoldása, ha $r(A) = r(B) = k$. (Van megoldás és a rang megegyezik az ismeretlenek számával.) ■

2. Homogén lineáris egyenletrendszerek

Mint fentebb láttuk, az $A[\mathbf{x}] = [\mathbf{b}]$ mátrixegyenlet megoldásait a következőképpen kaphatjuk meg. Tekintünk egy \mathbf{x}_0 megoldást, és ekkor minden megoldás $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$ alakú lesz, ahol $\mathbf{z} \in \text{Ker}(\alpha)$ (ahol A a α mátrixa). Más szóval \mathbf{z} az $A[\mathbf{x}] = \mathbf{0}$ úgynevezett *homogén lineáris egyenletrendszer* megoldása. A homogén lineáris egyenletrendszerek megoldásainak a vizsgálata más szempontokból is igen fontos. Ennek az egyenletrendszernek triviálisan megoldása a $\mathbf{0}$ vektor.

A $\mathbf{0}$ vektort a homogén lineáris egyenletrendszer triviális megoldásának nevezzük.

Homogén lineáris egyenletrendszer esetében éppen az a fontos kérdés, hogy mikor van a rendszernek triviálistól különböző megoldása. Mivel megoldás mindig létezik, ezért ennek a feltétele $r(A) < k$ (hiszen nagyobb nem is lehet):

9.2.A Tétel. *Homogén lineáris egyenletrendszernek akkor van nemtriviális megoldása, ha a mátrix rangja kisebb az ismeretlenek számánál.* ■

Természetesen ebben az esetben is meg kell adni a megoldásokat, ami — általában — nem egyszerű. Most két fontos speciális esetet nézünk.

Először a sajátvektorokról szólnunk. Ha adott egy A négyzetes mátrix, akkor tudjuk, hogy ennek sajátértékei a $\det(A - x \cdot I)$ karakterisztikus polinom gyökei. Ha c ennek egy gyöke, akkor ehhez a sajátértékhez tartozó összes \mathbf{x} sajátvektort úgy kaphatjuk, hogy megkeressük az $(A - c \cdot I)[\mathbf{x}] = [\mathbf{0}]$ homogén lineáris egyenletrendszer megoldásait. Ezzel a sajátvektorok meghatározását elvileg megoldottuk. (A gyakorlati meghatározásnál sok nehézséggel kell szembenézni.)

A másik fontos speciális esetben ugyancsak megegyezik az egyenletek és ismeretlenek száma. Ekkor az a feltétel, hogy α magja nem csak a nullvektorból áll, pontosan azt jelenti, hogy α szinguláris, azaz $\det(A) = 0$.

9.2.B Tétel. *Az $n = k$ esetben az $A[\mathbf{x}] = [\mathbf{0}]$ egyenletnek pontosan akkor van nemtriviális megoldása, ha $\det(A) = 0$.* ■

Tekintsük egy geometriai alakzatnak, például egy körnek az egyenletét: $a(x^2 + y^2) + bx + cy + d = 0$. Fel szeretnénk írni három adott ponton átmenő kör egyenletét. Legyenek ezek a pontok (x_0, y_0) , (x_1, y_1) és (x_2, y_2) . Emellett van még egy „futó” pont, amelynek a koordinátái (x, y) . Ez azt jelenti, hogy e négy pont bármelyikét behelyettesítve az egyenlet bal oldalába valóban 0-t kapunk. Feladatunk az a, b, c, d számok meghatározása. Most a, b, c, d az ismeretlenek. Ezeknek együtthatói rendre az $x^2 + y^2$, x , y és 1 polinomok, illetve az $x_i^2 + y_i^2$, x_i , y_i és 1 számok (ahol $i \in \{0, 1, 2\}$). A kapott homogén lineáris egyenletrendszernek pontosan akkor van megoldása, ha mátrixának determinánsa 0. Ezt felírva:

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_0^2 + y_0^2 & x_0 & y_0 & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \end{vmatrix} = 0$$

adódik. Vegyük észre, hogy ez egy egyenletet szolgáltat, ami pontosan a kör (x, y) pontjaira teljesül. Ez tehát éppen a kör egyenletét adja meg. Ha két pont egybeesik, akkor nem kapunk egyenletet (nem egyértelmű a kör). Ha a három pont egy egyenesen van, akkor kideríthető, hogy $a = 0$ adódik, azaz egy egyenest kapunk.

Érdemes megemlíteni, hogy hasonló a helyzet az interpolációnál is. Ezt az esetet később tárgyaljuk.

Feladatok

1. Másodfokú „görbék” általános koordináta-rendszerbeli egyenlete $ax^2 + bxy + cy^2 + dx + ey + f = 0$, ahol $a, b, c, d, e, f \in \mathbb{R}$, és ezek közül nem mind 0. Bizonyítsuk be, hogy bármely öt ponton keresztül pontosan egy másodfokú görbe fektethető.

2. Határozzuk meg a $(0, 0)$, $(0, 1)$, $(1, 0)$, illetve $(-1, 2)$, $(0, 1)$, $(1, 0)$ pontokon átmenő kör egyenletét.

3. A rezultáns

Tekintsünk az R egyértelmű faktorizációs tartomány (a továbbiakban így rövidítjük: EFT) feletti két polinomot: $f, g \in R[x]$. Azt akarjuk megnézni, hogy milyen feltétel mellett van ezeknek közös gyökük. Először ennek egy elég triviálisnak tűnő átfogalmazását adjuk meg, majd ezt fogjuk átírni a lineáris egyenletrendszerek megoldhatósága feltételével egy jól használható alakba.

9.3. Tétel. *Az R egyértelmű faktorizációs tartomány feletti $f(x), g(x)$ polinomoknak akkor és csak akkor van közös faktoruk, ha léteznek olyan $u(x), v(x)$ R feletti polinomok, amelyekre a következő teljesül:*

$$u(x) \cdot f(x) - v(x) \cdot g(x) = 0 \quad \text{és} \quad gr(u) < gr(g), \quad gr(v) < gr(f).$$

Bizonyítás. Legyen $d(x) = \text{lko}(f(x), g(x))$. Ha $d(x)$ nem konstans, akkor $u(x) = \frac{g(x)}{d(x)}$ és $v(x) = \frac{f(x)}{d(x)}$ kielégítik a kirótt feltételeket. Tegyük most fel — fordítva —, hogy ilyen u és v polinomok léteznek. Mivel R EFT, ezért $R[x]$ is az. Feltehető tehát, hogy u és v relatív prímek (hiszen osztásnál a fok csak csökkenhet!). Ez azt jelenti, hogy például $v(x)$ osztója $f(x)$ -nek a hányadostest felett; s az egyértelmű faktorizáció alapján R felett is. Eszerint létezik olyan $d(x)$ polinom, amelyre $f = d \cdot v$; s a fokokra vonatkozó feltétel miatt $d(x)$ nem konstans. Behelyettesítve a fenti egyenlőségbe, majd $v(x)$ -szel osztva a $g = d \cdot u$ egyenlőséghez jutunk; ami éppen azt mutatja, hogy a két adott polinomnak létezik közös faktora. ■

9.1. Definíció. Legyenek $f(x) = a_n x^n + \dots + a_1 x + a_0$ és $g(x) = b_k x^k + \dots + b_1 x + b_0$ R -beli együtthatós polinomok. Ezek rezultánsán az

$$R(f, g) = \begin{vmatrix} a_n & a_{n-1} & \dots & a_{n-k} & a_{n-k-1} & \dots & a_0 & 0 & \dots & 0 & 0 \\ 0 & a_n & \dots & a_{n-k-1} & a_{n-k} & \dots & a_1 & a_0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a_n & a_{n-1} & \dots & a_k & a_{k-1} & \dots & a_1 & a_0 \\ b_k & b_{k-1} & \dots & b_0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & b_k & \dots & b_1 & b_0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & b_{k-1} & b_{k-2} & \dots & b_0 & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & b_k & b_{k-1} & \dots & b_1 & b_0 \end{vmatrix}$$

determinánst értjük. ■

9.4. Tétel. Az f és g R EFT-beli polinomoknak akkor és csak akkor van közös gyökük, ha rezultánsuk 0.

Bizonyítás. Legyenek f és g , mint a 9.1. Definícióban. A 9.3. Tétel szerint a közös gyökök létezésének szükséges és elégséges feltétele alkalmas u és v polinomok létezése. Alkalmas x -hatvánnyal szorozva elérhető, hogy $gr(u) = k - 1$ és $gr(v) = n - 1$ legyen: $u(x) = u_{k-1}x^{k-1} + \dots + u_1x + u_0$, és $v(x) = v_{n-1}x^{n-1} + \dots + v_1x + v_0$. Írjuk most fel az $u(x) \cdot f(x) - v(x) \cdot g(x)$ polinomban x hatványainak az együtthatóját:

$$\begin{array}{lll} x^{n+k-1} \text{ együtthatója:} & a_n u_{k-1} & -b_k v_{n-1} \\ x^{n+k-2} \text{ együtthatója:} & a_{n-1} u_{k-1} + a_n u_{k-2} & -b_{k-1} v_{n-1} - b_k v_{n-2} \\ & \vdots & \vdots \\ x^{k-1} \text{ együtthatója:} & a_0 u_{k-1} + a_1 u_{k-2} + \dots & -b_{k-n}(?)v_{n-1} - \dots \\ x^{k-2} \text{ együtthatója:} & 0u_{k-1} + a_0 u_{k-2} + \dots & -b_{k-n-1}(?)v_{n-1} - \dots \\ & \vdots & \vdots \\ x \text{ együtthatója:} & a_0 u_1 + a_1 u_0 & -b_0 v_1 - b_1 v_0 \\ x^0 \text{ együtthatója:} & a_0 u_0 & -b_0 v_0 \end{array}$$

[(?) azt jelenti, hogy negatív index esetén ez a tag nem lép fel.]

Az a feltétel, hogy $u(x) \cdot f(x) - v(x) \cdot g(x) = 0$, azt jelenti, hogy léteznek olyan u_{k-1}, \dots, u_0 és v_{n-1}, \dots, v_0 nem csupa 0 elemek, amelyekre a fenti együtthatók mindegyike 0. A homogén lineáris egyenletrendszer megoldhatóságára látott feltétel szerint ez pontosan akkor következik be, ha ezekben mint ismeretlenekben tekintett egyenletrendszer determinánsa 0. Ez a determináns viszont éppen $R(f, g)$. ■

A rezultáns segítségével — legalábbis elvben — magasabbfokú egyenletrendszerek is megoldhatók. A fenti módszer azt adja, hogy ha van például két polinomunk két határozatlannal, akkor az egyik határozatlanra felírt rezultáns a másik határozatlanra egy olyan

egyenletet ad, amelynek minden gyöke esetén (és csak ekkor!) *lehet* közös gyök. Így, lépésenként egy-egy határozatlan kiküszöbölhető; de azon az áron, hogy a többi határozatlanban a polinomok foka „iszonyatosan” megnövekszik.

1. Következmény. *Legyenek a 9.1. Definícióban szereplő $f(x)$, illetve $g(x)$ polinom gyökei, megfelelően, $\alpha_1, \dots, \alpha_n$ és β_1, \dots, β_k . Ekkor:*

$$R(f, g) = a_n^n \cdot b_k^k \cdot \prod_{i=1}^n \prod_{j=1}^k (\alpha_i - \beta_j).$$

Bizonyítás. Mindenekelőtt megjegyezzük a következőket.

A testbővítéseknél látni fogjuk, hogy bármilyen integritási tartományból indulunk is ki, alkalmas testben a polinomnak létezik annyi gyöke (multiplicitással), amekkora a foka.

A rezultáns a két polinom együtthatóinak a polinomja. Ezek viszont polinomjai a főegyütthatónak és a gyököknek, hiszen $\frac{a_i}{a_n}$ a gyökök elemi szimmetrikus polinomjaival vagy azok negatívjával egyeznek meg (hasonló igaz $\frac{b_j}{b_k}$ esetében is).

Ennek megfelelően tekinthetjük a főegyütthatókat és a gyököket határozatlanoknak. Ha ezekre bebizonyítjuk az egyenlőséget, akkor bármely konkrét esetre is egyenlőséget kapunk, behelyettesítéssel.

Legyen tehát $f(t) = a_n \prod_{i=1}^n (t - x_i)$ és $g(t) = b_k \prod_{j=1}^k (t - y_j)$. Ebből azt kapjuk,

hogy $R(f, g) = P(a_n, b_k, x_1, \dots, x_n, y_1, \dots, y_k)$ polinomja e határozatlanoknak. A rezultáns tulajdonsága alapján akármelyik x_i helyébe bármelyik y_j -t írva a rezultáns 0-val lesz egyenlő. Mint a többhatározatlanú polinomok tárgyalásánál láttuk, ennek következménye, hogy a rezultáns (mint polinom) osztható minden egyes $x_i - y_j$ polinommal, mert ezek normáltak. Tekintettel arra, hogy ezek a polinomok relatív prímek is, ezért a rezultáns osztható ezek szorzatával is. A rezultáns első k sorában szereplő minden egyes polinom osztható az a_n határozatlannal és a többi sorban szereplő polinomok mindegyike osztható b_k -val. Mivel ezek is relatív prímek a többiekhez, ezért a rezultáns osztható ezek szorzatával, ami pontosan a jobb oldalon levő szorzat.

A rezultáns első k sorában minden egyes x_i legfeljebb az első hatványon szerepel. Ezért a rezultáns minden egyes x_i -ben legfeljebb k -adfokú és hasonlóképpen minden egyes y_j -ben legfeljebb n -edfokú tagot tartalmazhat. Így a bal oldal a jobb oldalnak konstansszoros (számszorosa). Ez a konstans alkalmas behelyettesítéssel meghatározható (bármely olyan behelyettesítéssel, amelyre a jobb oldal nem 0). Legyen például $a_n = b_k = y_1 = \dots = y_k = 1$ és $x_1 = \dots = x_n = 0$. Ekkor a szorzat értéke 1. A rezultánst szolgáltató mátrix pedig $\begin{bmatrix} I & O \\ A & B \end{bmatrix}$ alakú, ahol I identitásmátrix, O minden eleme 0 és B -ben a fődiagonálisban egyesek, felette nullák állnak. Ennek a determinánsa is 1, ami bizonyítja az egyenlőséget. ■

2. Következmény. Legyenek $f(x)$ és $g(x)$, mint az 1. Következményben. Ekkor:

$$R(f, g) = a_n^n \prod_{i=1}^n g(\alpha_i) = (-1)^{nk} b_k^k \prod_{j=1}^k f(\beta_j).$$

Bizonyítás. Az első egyenlőség bizonyítására bontsuk fel $g(x)$ -et tényezőire:

$$g(x) = b_k \prod_{j=1}^k (x - \beta_j).$$

Ezt behelyettesítve az első egyenlőség jobb oldalába pontosan a rezultánst kapjuk, ami bizonyítja az első egyenlőséget. A rezultánst meghatározó mátrixban $n \cdot k$ sorcserével az $R(g, f)$ rezultánst meghatározó mátrix adódik; így a második egyenlőség is is igaz. ■

3. Következmény. Legyen $f(x)$, mint a fentiekben, azzal a megkötéssel, hogy $a_n = 1$ (tehát normált). Legyen $g(x) = f'(x)$ az $f(x)$ deriváltja. Definiálja $D = D(f) = R(f, f')$ az $f(x)$ polinom diszkriminánsát. Ekkor $D(f) = \prod_{i \neq j} (\alpha_i - \alpha_j)$.

A diszkrimináns pontosan akkor 0, ha $f(x)$ -nek többszörös gyökei vannak.

Bizonyítás. Induljunk ki a 2. Következményben szereplő első egyenlőségből. Az $f(x) = \prod_{i=1}^n (x - \alpha_i)$ polinom deriváltját a szorzat deriválási szabálya alapján meghatározva azt kapjuk, hogy:

$$f'(x) = \sum_{i=1}^n (x - \alpha_1) \cdots (x - \alpha_{i-1}) \cdot (x - \alpha_{i+1}) \cdots (x - \alpha_n).$$

Az $R(f, f') = \prod_{i=1}^n f'(\alpha_i)$ egyenlőségből adódik a diszkrimináns előállítása.

A többszörös gyökökre vonatkozó állítás azonnal következik a szorzat-előállításból. ■

Megjegyzés. A diszkrimináns előáll mint $V(\alpha_1, \dots, \alpha_n)$ négyzete vagy annak negatívja ($V(x_1, \dots, x_n)$ a Vandermonde-mátrix). (Nem írjuk ki az előjelet, mert ez a szakirodalomban nem egységes.) A rezultánsalakban való előállítás lehetőséget ad a diszkriminánsnak az együtthatók segítségével való meghatározására, anélkül, hogy a szimmetrikus polinomok alaptételét felhasználnánk. □

Példák:

1. Legyen $f(x) = x^2 + ax + b$. Ekkor

$$D(f) = \begin{vmatrix} 1 & a & b \\ 2 & a & 0 \\ 0 & 2 & a \end{vmatrix} = \begin{vmatrix} 1 & a & b \\ 0 & -a & -2b \\ 0 & 2 & a \end{vmatrix} = -a^2 + 4b,$$

ami előjeltől eltekintve megegyezik a másodfokú polinom diszkriminánsával.

2. Legyen $f(x) = x^3 + px + q$ hiányos harmadfokú polinom. Ennek a diszkriminánsára a következő adódik:

$$\begin{vmatrix} 1 & 0 & p & q & 0 \\ 0 & 1 & 0 & p & q \\ 3 & 0 & p & 0 & 0 \\ 0 & 3 & 0 & p & 0 \\ 0 & 0 & 3 & 0 & p \end{vmatrix} = \begin{vmatrix} 1 & 0 & p & q & 0 \\ 0 & 1 & 0 & p & q \\ 0 & 0 & -2p & -3q & 0 \\ 0 & 0 & 0 & -2p & -3q \\ 0 & 0 & 3 & 0 & p \end{vmatrix} = \begin{vmatrix} -2p & -3q & 0 \\ 0 & -2p & -3q \\ 3 & 0 & p \end{vmatrix} =$$

$$= 4p^3 + 27q^2 = 108 \left(\left(-\frac{q}{2} \right)^2 + \left(\frac{p}{3} \right)^3 \right),$$

ami lényegében megegyezik a Cardano-képletben a gyök alatt álló kifejezéssel.

Feladatok

1. Határozzuk meg mechanikusan, hogy milyen a valós szám esetén van az $x^3 + x^2 + 3x + a$ és $x^2 + 3x + 2$ polinomoknak közös faktoruk.

2. Oldjuk meg az

$$\begin{aligned} (x-1) \cdot y^2 + (x+1) \cdot y - 2 &= 0 \\ (x-1) \cdot y^2 + x \cdot y - 1 &= 0 \end{aligned}, \quad \text{illetve} \quad \begin{aligned} (x-1) \cdot y^2 + (x+1) \cdot y - 1 &= 0 \\ (x-1) \cdot y^2 + x \cdot y - 1 &= 0 \end{aligned}$$

egyenletrendszereket.

3. Oldjuk meg az $x^2 + 1 = 2y$, $y^2 + 1 = 2z$, $z^2 + 1 = 2x$ egyenletrendszert a valós számtestben.

4. Lineáris egyenletrendszerek közelítő megoldása

A gyakorlati életben adódó lineáris egyenletrendszereknél az együtthatók nem pontosak. Az ismeretlenek „jobb” meghatározhatósága végett a szükségesnél több mérést végeznek. Ha k darab ismeretlen mennyiség szerepel, akkor minden mérés egy egyenletet jelent, ezért az egyenletek n száma nagyobb k -nál. (Persze túl sok mérés sokkal több számolást jelent, ezért csak egy kicsivel több mérést érdemes végezni, mondjuk $k + \log(k)$ számút.) Természetesen teljesen felesleges a méréseket ugyanolyan adatokkal végezni, azaz igyekezni kell e mérési adatok függetlenségére. Ha például az $ax + by = c$ egyenletből kell mérések útján x -et és y -t meghatározni, akkor úgy mérünk, hogy az a -kból és b -kból álló vektorok függetlenek legyenek. Tegyük fel, hogy három mérés esetén a következők adódtak c -re:

$$x + y = 7, \quad x + 2y = 9 \quad \text{és} \quad 2x + y = 13.$$

Ennek az egyenletrendszernek nincs megoldása, mert a második és harmadik egyenletből azt kapjuk, hogy $3x + 3y = 22$, ami nem ugyanaz, mint az első egyenletből adódó $3(x + y) = 21$.

Itt tehát a bal oldalon fellépő $\mathbf{a}_1, \dots, \mathbf{a}_n$ vektoroknak nem lesz a jobb oldalon álló \mathbf{b} vektor lineáris kombinációja. Emellett, a mérések függetlensége alapján feltehetjük, hogy az $\mathbf{a}_1, \dots, \mathbf{a}_n$ vektorok lineárisan függetlenek. Esetünkben az a cél, hogy a legjobban közelítő megoldást megtaláljuk. Evégett tekintsük az oszlopvektorokból álló n -dimenziós teret euklideszi térnek, amelyben az a bázis definiálja a skalárszorozást, amelynek vektoraiban egyetlen 1-es szerepel és a többi elem 0. Ebben a térben az $\mathbf{a}_1, \dots, \mathbf{a}_n$ vektorok egy alteret feszítenek ki, amelynek \mathbf{b} nem eleme. Tudjuk viszont, hogy \mathbf{b} -nek van erre az alterre egy \mathbf{a} vetülete, amelyik az összes altérbeli vektorok közül a legközelebb van \mathbf{b} -hez. Erre az $\mathbf{a} = \sum_i x_i \mathbf{a}_i$ vektorra az teljesül, hogy $\mathbf{b} - \mathbf{a}$ merőleges az alterre, vagyis minden egyes \mathbf{a}_i vektorra. A skalárszorzat felhasználásával ezt egy olyan egyenletrendszerre írtuk át, amelynek mátrixa egy Gram-féle mátrix:

$$\begin{aligned}
 (\mathbf{a}_1; \mathbf{a}_1)x_1 + (\mathbf{a}_1; \mathbf{a}_2)x_2 + \dots + (\mathbf{a}_1; \mathbf{a}_k)x_k &= (\mathbf{a}_1; \mathbf{b}) \\
 (\mathbf{a}_2; \mathbf{a}_1)x_1 + (\mathbf{a}_2; \mathbf{a}_2)x_2 + \dots + (\mathbf{a}_2; \mathbf{a}_k)x_k &= (\mathbf{a}_2; \mathbf{b}) \\
 (*) (*) & \\
 & \vdots \\
 (\mathbf{a}_k; \mathbf{a}_1)x_1 + (\mathbf{a}_k; \mathbf{a}_2)x_2 + \dots + (\mathbf{a}_k; \mathbf{a}_k)x_k &= (\mathbf{a}_k; \mathbf{b}).
 \end{aligned}$$

Mivel a \mathbf{a}_i vektorok függetlenek, ezért a Gram-féle determináns nem 0. Így az egyenletrendszer mátrixának a rangja k , s mivel a kibővített mátrix rangja ennél nem lehet nagyobb (természetesen kisebb sem), ezért az egyenletrendszernek létezik egyértelmű megoldása. A fenti számpéldánál a kapott egyenletrendszer: $6x + 5y = 42$, $5x + 6y = 38$, amelynek a megoldása $x = \frac{62}{11}$, $y = \frac{18}{11}$.

5. A Cramer-szabály

A Gram-féle mátrixnál egy olyan egyenletrendszer szerepel, amelyben az egyenletek és az ismeretlenek száma megegyezik és az egyenletrendszernek egyetlen megoldása van. (Nem nehéz belátni, hogy tetszőleges lineáris egyenletrendszerrel ezen kívül csak „lineáris paraméterekre” van szükség.)

Tegyük fel, hogy a (*) alatti egyenletrendszerrel $n = k$, és ez megegyezik az egyenletrendszer mátrixának a rangjával. Ez azt jelenti, hogy az A mátrix reguláris, tehát determinánsa nem 0. Tudjuk, hogy ekkor az egyenletrendszernek pontosan egy megoldása van, legyen ez (**) alakban:

$$\mathbf{b} = \mathbf{a}_1 x_1 + \dots + \mathbf{a}_n x_n.$$

A mértékek multilinearitása alapján ebből, tetszőleges szóba jövő i index esetén

$$\mu(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) = \mu(\mathbf{a}_1, \dots, \mathbf{a}_n) \cdot x_i$$

adódik, tekintettel arra, hogy lineárisan összefüggő vektorokra a mérték 0.

Mivel a mérték skalártényezőitől eltekintve egyértelmű, választhatjuk azt a mértéket, amely az I mátrix oszlopain éppen e mátrix determinánsát veszi fel. Ez a mérték minden mátrix esetében az oszlopvektorokhoz a mátrix determinánsát rendeli. Ezáltal a következőket bizonyítottuk be:

Ha $()$ -ban $n = k = r(A)$, akkor tekintsük az egyenletrendszer mátrixának a D determinánsát és legyen D_i annak a mátrixnak a determinánsa, amely az A mátrixból úgy adódik, hogy az i -edik oszlop helyébe $[b]$ -t írunk. Ekkor az egyenletrendszer megoldását $x_i = \frac{D_i}{D}$ adja.*

Ezt az eljárást nevezzük Cramer-szabálynak.

A Cramer-szabály hasznosan alkalmazható interpoláció esetében is. Itt adottak az a_0, a_1, \dots, a_n (különböző) helyek, és ezeken a b_0, b_1, \dots, b_n függvényértékek. Olyan $p(x)$ legfeljebb n -edfokú polinomot keresünk, amelyre $p(a_i) = b_i$. Ezek a $p(x)$ polinom $n + 1$ együtthatójára $n + 1$ egyenletet adnak, amelyek az együtthatókban lineárisak. Itt mind az egyenletek száma, mind az ismeretlenek száma $n + 1$; a Cramer-szabály alkalmazhatóságához az kell még, hogy az egyenletrendszer mátrixának a determinánsa ne legyen 0. Ez a mátrix viszont az a_0, a_1, \dots, a_n számokból képezett Vandermonde-mátrix, így determinánsa, a helyekre vonatkozó feltevésünk alapján, nem 0.

6. Kvadratikus alakok jellegének a megállapítása

A kvadratikus alakok jellegének a megállapítására elsősorban geometriai kérdéseknél van szükség. Ezért a következőkben feltesszük, hogy a valós számtest feletti vektorterekről van szó; bár a komplex számtest feletti vektorterekre is igazolhatók hasonló módon a megfelelő eredmények.

Tudjuk, hogy a kvadratikus alakok mátrixa szimmetrikus, és alkalmas bázisban diagonálissá válik. Azt is tudjuk, hogy a diagonálisban szereplő pozitív, negatív és 0 számok száma nem függ a bázistól. Az új bázisra való áttérés sok esetben eléggé bonyolult, ezért most olyan feltételeket szeretnénk találni, amelyek enélkül is megadják az adott kvadratikus alak kvadratikus karakterét.

Tudjuk, hogy egy vektortér új bázisát elemi transzformációk sorozatával hozhatjuk létre. Azt is tudjuk, hogy egy elemi transzformáció vagy azt jelenti, hogy egy bázisvektor skalárszorosát hozzáadjuk egy tőle különböző bázisvektorhoz, vagy azt, hogy egy bázisvektort egy nemnulla skalárral szorzunk, vagy azt, hogy két bázisvektort felcserélünk. (E harmadik külön említésére nincs szükség, de sok esetben kényelmesebbé teszi az eljárást.) Ezek az eljárások egy-egy mátrixon mindig egy elemi átalakítást jelentenek. A háromféle elemi átalakítás: egy oszlop skalárszorosát hozzáadjuk egy másik oszlophoz, egy oszlopot egy nemnulla skalárral szorzunk, vagy két oszlopot felcserélünk. Hasonló elemi átalakításokat végezhetünk a sorokkal is.

Amikor a lineáris transzformációkat vizsgáltuk, akkor elég volt ezeket vagy csak balról, vagy csak jobbról nézni, noha az átalakítás $A \mapsto S^{-1}AS$ alakú. Kvadratikus alakoknál $A \mapsto S^{\dagger}AS$ az átalakítás menete. Itt nem lehet a két oldalról való szorzást külön nézni, mert akkor „menet közben” elveszik a szimmetrikusság. Az elemi transzformációk és transzponáltjuk mátrixai közti összefüggés alapján kvadratikus alak mátrixával a következő elemi átalakítások végezhetők:

A kvadratikus alak mátrixában az i -edik sor c -szeresét hozzáadjuk a j -edik sorhoz és az i -edik oszlop c -szeresét hozzáadjuk a j -edik oszlophoz ($i \neq j$).

A kvadratikus alak mátrixában az i -edik sort is és az i -edik oszlopot is megszorozzuk egy $c \neq 0$ számmal.

A kvadratikus alak mátrixában az i -edik és j -edik oszlopot, valamint az i -edik és j -edik sort is felcseréljük.

Itt is azt fogjuk belátni (mint a mátrixok normálalakjánál), hogy a megfelelő diagonalizáció közben bizonyos jellemző adatok nem változnak meg; és ezek az adatok itt is a mátrixból elkészíthető determinánsok lesznek.

Legyen n a (négyzetes) szimmetrikus A mátrix sorainak a száma és legyen A_i az a (szimmetrikus) részmátrix, amelyet úgy kapunk, hogy az i -edik sor utáni sorokat és az i -edik oszlop utáni oszlopokat elhagyjuk. Ezután a következőképpen definiáljuk az A mátrix Δ_i főminorjait (vagy főaldeterminánsait): $\Delta_0 = 1$, és $i > 0$ esetére Δ_i az A_i determinánsa. (Nem tévesztendő össze a determinánsosztókkal!)

Az A mátrix főminorjai sorozatán a $\Delta_0, \Delta_1, \dots, \Delta_n$ sorozatot értjük.

9.5.A Tétel. *Egy kvadratikus alak akkor és csak akkor pozitív (negatív) definit, ha főminorjainak a sorozata jeltartó (jelváltó).*

Bizonyítás. Mindenekelőtt megjegyezzük, hogy egy kvadratikus alak negatívját véve mátrixa (-1) -gyel szorzódik. Ezáltal a páros rendű főminorok nem változnak, míg a páratlan rendűek előjelet váltanak. Így a tételt pozitív definit kvadratikus alakokra bizonyítva abból tüstént következik a tétel negatív definit kvadratikus alakokra is (mivel $\Delta_0 = 1 > 0$).

A bizonyításhoz arra is szükségünk lesz, hogy pozitív definit kvadratikus alak mátrixának a determinánsa pozitív. Ez igaz egy olyan bázisban, amelyben a kvadratikus alak mátrixa diagonálissá válik, hiszen a pozitív definittség miatt a diagonálisnak csak pozitív elemei lehetnek. Legyen A a kvadratikus alak mátrixa, ekkor egy új bázisban a mátrix $S^{\dagger}AS$ alakú, ahol S és így S^{\dagger} is invertálható. Ha ez egy diagonális alak, akkor

$$0 < \det(S^{\dagger}AS) = \det(S^{\dagger})\det(A)\det(S) = \det(S^{\dagger})\det(S)\det(A).$$

Tekintettel arra, hogy egy mátrix transzponáltjának a determinánsa megegyezik az eredeti mátrix determinánsával, ezért $\det(S^{\dagger})\det(S)$ pozitivitása következik, és így $\det(A)$ is pozitív.

Ezután diagonalizálni fogjuk a kvadratikus alak mátrixát úgy, hogy közben a főminorok sorozata ne változzék meg. Az eljárásban arra kell ügyelni, hogy egyik sorhoz (vagy

oszlophoz) se adjunk egy nagyobb indexű sort (vagy oszlopot), mert ezáltal a főminorok sorozata megváltozhatna. Éppen ezért csak a következő típusú lépést végezzük: $i < j$ esetén az i -edik sor c -szeresét hozzáadjuk a j -edik sorhoz és az i -edik oszlop c -szeresét hozzáadjuk a j -edik oszlophoz.

Tegyük fel, hogy a tételbeli két feltétel bármelyike teljesül, és az első k sorra és oszlopra már elvégeztük a diagonalizálást úgy, hogy minden k -nál kisebb indexre ezekben a sorokban és oszlopokban a fődiagonálison kívül minden elem 0, és a fődiagonális elemei pozitívak; továbbá egyetlen főminor sem változott meg. A $k = 1$ esetben ez eleve teljesül, mert $\Delta_0 = 1 > 0$. Tekintsük most A_k -t. Ennek a determinánsa mindkét esetben pozitív. Ha a kvadratikus alak pozitív definit, akkor minden altéren pozitív definit, az előzetesen belátott eredmény miatt tehát pozitív a determinánsa. Ha viszont azt tudjuk, hogy a főminorok sorozata jeltartó, akkor ez azért pozitív, mert a nulladik főminor pozitív. Tekintettel arra, hogy ez a pozitív determináns megegyezik a diagonális elemek szorzatával, ezért a diagonális k -adik eleme — mint két pozitív elem hányadosa — szintén pozitív. A kapott mátrix szimmetrikus. A szimmetria megmarad, ha minden k -nál nagyobb indexű sorból kivonjuk a k -adik sor megfelelő skalárszorosát és ugyanezen skalárral szorozva minden k -nál nagyobb indexű oszlopból a k -adik oszlop skalárszorosát. Ezt megtehetjük úgy, hogy most már a k -adik sorban és a k -adik oszlopban is csak a diagonális elem ne legyen 0; s mint láttuk, a diagonális k -adik eleme pozitív. Ahhoz, hogy az indukciós lépést alkalmazni tudjuk, már csak azt kell belátni, hogy a főminorok sorozata továbbra is változatlan. Ez viszont világos, hiszen az elemi átalakításoknál mindig a megfelelő mátrixokon belül maradtunk. Ebből azonnal következik a két feltétel ekvivalenciája. ■

A következőkben a szemidefinit kvadratikus alakokkal foglalkozunk. Természetesen itt is elegendő a pozitív szemidefinit kvadratikus alakokra szorítkozni. Pozitív szemidefinit kvadratikus alakoknál is meg lehet fogalmazni egy viszonylag egyszerű szükséges feltételt, ez azonban nem elégséges. Meg fogunk adni két elégséges feltételt is, ezek viszont nem olyan egyszerűek, mint a szükséges feltétel.

A pozitív szemidefinit kvadratikus alakok mátrixára is hasonló jellegű tétel igaz, mint a pozitív definit esetben:

9.5.B Tétel. *Pozitív szemidefinit kvadratikus alak mátrixának a determinánsa nemnegatív, s ha a kvadratikus alak nem pozitív definit, akkor a determináns 0.*

Bizonyítás. Mint láttuk, új bázisra való áttérésnél egy kvadratikus alak mátrixának a determinánsa pozitív számmal szorzódik. Ezért elég az állítást olyan bázisra bizonyítani, amelyben a kvadratikus alak mátrixa diagonális. A pozitív szemidefinitésg miatt a diagonális elemei nem lehetnek negatívak, s ha a kvadratikus alak nem pozitív definit, akkor szerepel közöttük 0 is. ■

9.5.C Tétel. *Pozitív szemidefinit kvadratikus alak mátrixa főminorjainak a sorozata valameddig jeltartó, és attól kezdve mindegyik 0.*

Bizonyítás. Tekintsük a $\mathbf{Q}(\mathbf{x})$ kvadratikus alak mátrixát valamelyik bázisban. Mivel $\mathbf{Q}(\mathbf{x})$ pozitív szemidefinit, ezért létezik olyan maximális s index, hogy $\mathbf{Q}(\mathbf{x})$ az első s számú bázisvektor generálta altéren pozitív definit ($s = 0$ is lehet). Így a pozitív definit kvadratikus alakokra vonatkozó tétel miatt a főminorok sorozatában az első $s+1$ jeltartó. Az $(s+1)$ -edik vektor, és bármely további vektor hozzávételével a kapott altéren $\mathbf{Q}(\mathbf{x})$ már *nem* pozitív definit. Az előző tétel szerint tehát a megfelelő determinánsok mindegyike 0. ■

Ebből a feltételből még nem következik az, hogy a kvadratikus alak valóban pozitív szemidefinit.

9.5.D Tétel. *Tegyük fel, hogy a $\mathbf{Q}(\mathbf{x})$ kvadratikus alakhoz tartozó mátrix első $r+1$ főminorjának $\Delta_0, \Delta_1, \dots, \Delta_r$ sorozata jeltartó, és minden olyan $r+1$, illetve $r+2$ sorból és oszlopból álló szimmetrikus mátrix determinánsa 0, amelyik az első r sorból és oszlopból álló A_r mátrixot tartalmazza. Ekkor $\mathbf{Q}(\mathbf{x})$ pozitív szemidefinit.*

Bizonyítás. Végezzük el a pozitív definit kvadratikus alakok esetében adott eljárást az első r bázisvektorra. Ezáltal sem a kvadratikus karakter, sem az érintett determinánsok nem változnak meg. Az eljárás végén a mátrix diagonálisában a pozitív a_1, \dots, a_r elemek állnak, míg az első r sor és első r oszlop összes többi eleme 0 lesz. Azt fogjuk megmutatni, hogy a feltételek teljesülése esetén a mátrix összes többi eleme is 0; a tehetetlenségi tétel alapján tehát a kvadratikus alak pozitív szemidefinit.

Legyen $i > r$, a_i a diagonális i -edik eleme, és vegyük hozzá A_r -hez az i -edik sort és oszlopot. A kapott mátrix a konstrukció következtében diagonális, és determinánsa a diagonális elemek $a_1 \cdot \dots \cdot a_r \cdot a_i$ szorzata. Feltétel szerint ez a szorzat 0, de az első r tényező nem az, ezért $a_i = 0$. Ezért a diagonális összes többi eleme 0.

Legyen most $r < i < j$, és legyen a mátrix i -edik sorának j -edik eleme b . Tekintsük ezután azt a mátrixot, amelyik A_r -en kívül az i -edik és j -edik sort és oszlopot is tartalmazza. Mivel a diagonális i -edik és j -edik eleme is 0, ezért az új mátrixban az utolsó két sor és oszlop alkotta részmátrix $\begin{bmatrix} 0 & b \\ b & 0 \end{bmatrix}$. A két utolsó sort felcserélve a determináns előjelet vált és a mátrix diagonálissá válik. Ennek megfelelően a kapott mátrix determinánsa $-a_1 \cdot \dots \cdot a_r \cdot b^2$. Feltétel szerint ez 0, de az első r tényező szorzata nem 0, ezért $b = 0$, mint állítottuk. ■

9.5.E Tétel. *Tegyük fel, hogy egy kvadratikus alak mátrixában az első $r+1$ főminor sorozata jeltartó, ahol r a mátrix rangja. Ekkor a kvadratikus alak pozitív szemidefinit.*

Bizonyítás. A rangra vonatkozó feltétel szerint minden olyan négyzetes részmátrix determinánsa 0, amelyikben r -nél több sor szerepel. Ezért alkalmazható az előző tétel. ■

Feladatok

1. Mutassuk meg, hogy nem csak pozitív definit kvadratikus alakok mátrixának a determinánsa pozitív.
2. Mutassuk meg, hogy nem csak szemidefinit kvadratikus alakok mátrixának a determinánsa 0.
3. Mutassuk meg, hogy nem csak pozitív szemidefinit kvadratikus alakok mátrixára teljesülhet, hogy a főminorok valameddig jeltartók, s attól kezdve mind 0. Adjuk meg a legkisebb méretű ellenpéldát.
4. Bizonyítsuk be, hogy pozitív szemidefinit kvadratikus alakok esetében a két adott elégséges feltétel nem szükséges.
5. A fenti tételek alapján fogalmazzunk meg szükséges és elégséges feltételt is arra, hogy egy kvadratikus alak indefinit.
6. Tekintsük egy kvadratikus alak mátrixának a fődiagonálisát. Milyen esetben mondhatjuk azonnal, hogy a kvadratikus alak *nem* definit, illetve *nem* szemidefinit?
7. Fogalmazzunk meg az előbbihez hasonló feltételt a másodrendű szimmetrikus aldeteminánsokra.

TIZEDIK FEJEZET

TENZOROK

1. A tenzorszorzat

A tenzorok elsősorban a fizikában és a geometriában fordulnak elő. Közöttük a három legismertebb a gradiens, a rotáció és a divergencia. Mindenekelőtt szeretnénk ezeket vázlatosan megismertetni:

Ha az n -dimenziós térben (fizikában minden változó paraméter egy dimenziót jelent) egy skalárfüggvény (például a hőmérséklet) pontról pontra változik, akkor vannak olyan irányok, amelyekben a változás a legerősebb. Ezt az irányt határozza meg a gradiens.

A másik két tenzor esetében vektorfüggvény változik pontról pontra (például szél iránya és erőssége). A rotáció azt mutatja, hogy valamely pontban van-e örvénylelés, illetve milyen irányú és mekkora. A divergencia a „térészbe” bemenő és kijövő „anyagmennyiség” vektori összegét méri. Ha ez a vektorösszeg valahol pozitív, akkor ott forrás van, ha negatív, akkor elnyelés.

Ezeket a „mennyiségeket”, noha nem függenek a koordinátáktól, koordináta-rendszerben szokták megadni. Az adatok jobb kiszámíthatósága és értékelése végett természetesen még alkalmas koordináta-rendszert kell találni. A koordináta-rendszer megváltoztatása esetén a fenti mennyiségeket megadó adatok bizonyos szabály szerint változnak meg. Ennek a naiv kifejezése az, hogy „a tenzorok olyan mennyiségek, amelyek koordinátatranszformáció esetén bizonyos (megadott) szabály szerint változnak”. Valójában a tenzorok úgy tekinthetők, mint bizonyos típusú multilineáris vektorfüggvények.

A következőkben a tenzorfogalom algebrai megalapozását fogjuk megadni. A fenti kapcsolatokról már nem fogunk beszélni, csupán az algebrai bevezetés a cél. Érdemes megemlíteni, hogy a tenzorok ma már az algebrán belül is számos helyen felhasználhatók.

A tenzorok esetében több vektortérhez rendelünk egy újabbat (bevezetésként csak két vektortérhez). Ezt a hozzárendelést nevezzük tenzorszorzásnak. A tenzorszorzat bevezetésére három lehetőség nyílik.

A „legföldrögzadtabb” az, amikor egyszerűen megadjuk a konstruált tér egy bázisát. Ez azért nem célszerű, mert itt elsikkad a hozzárendelés és ez a tér könnyen összetévesztendő mással. (Erre majd később rá fogunk mutatni.)

A legabsztraktabb az, amikor megadjuk a tenzorszorzat jellemző tulajdonságait, és azt mondjuk, hogy ha van valami, ami ezekkel a tulajdonságokkal rendelkezik, akkor azt nevezzük tenzorszorzatnak. Ennek a bevezetésnek az a hátránya, hogy nem lehet látni, miről van szó.

Mi a közélet fogjuk választani, megkonstruáljuk a tenzorszorzatot. Valójában az előző két „felfogásra” is szükség van, ezeket tárgyalás közben meg fogjuk mutatni.

Még egy dolgot szükséges megjegyezni. Tekintettel arra, hogy a tenzorszorzat, mint vektortér, mással is összetéveszthető, ezért párhuzamosan modulusok tenzorszorzatát is nézni fogjuk. A továbbiakban adott feladatok a modulusok esetében rámutatnak a tenzorszorzat viselkedési furcsaságaira.

Kiindulásul tekintsünk egy R egységelemes kommutatív gyűrűt és két rögzített R -modulust, \mathcal{U} -t és \mathcal{V} -t. Vizsgálni szeretnénk az $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bihomomorfizmusokat. Tekintettel arra, hogy ezek elég bonyolultak, ezért szeretnénk e bihomomorfizmusokat egyetlen rögzített bihomomorfizmus segítségével visszavezetni homomorfizmusokra. Az első gond ott lép fel, hogy a fenti \mathbf{A} leképezés *nem* a direkt szorzatról való homomorfizmus. Valóban, a bihomomorfizmus tulajdonságai alapján $\mathbf{A}(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = \mathbf{A}(\mathbf{u}_1, \mathbf{v}) + \mathbf{A}(\mathbf{u}_2, \mathbf{v})$, ha ez a direkt szorzat homomorfizmusa lenne, akkor $\mathbf{v} = \mathbf{o} + \mathbf{v}$ miatt $\mathbf{A}(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = \mathbf{A}(\mathbf{u}_1, \mathbf{o}) + \mathbf{A}(\mathbf{u}_2, \mathbf{v})$ is fennállna, azaz $\mathbf{A}(\mathbf{u}_1, \mathbf{v})$ nem függene \mathbf{v} -től; és hasonlóképpen \mathbf{u}_1 -től sem. Ezért a keresett homomorfizmus nem képezhet a direkt szorzatról.

Ennek a problémának a megoldásához a szabad modulusokra van szükségünk:

10.1. Definíció. Legyen R egy egységelemes kommutatív gyűrű, \mathcal{M} egy R -modulus. Ha \mathbf{M} az \mathcal{M} egy bázisa, akkor \mathcal{M} -et az \mathbf{M} generálta szabad modulusnak nevezzük. (Bázison most azt értjük, hogy elemeinek egy lineáris kombinációja csak akkor lehet \mathbf{o} , ha minden együttható 0.) ■

10.1. Tétel. Tetszőleges \mathbf{H} halmazhoz létezik a \mathbf{H} generálta $\mathcal{F}_R(\mathbf{H})$ szabad R -modulus.

Bizonyítás. Készítsük el a $\left\{ \sum_{\mathbf{h}} a_{\mathbf{h}} \mathbf{h} \mid \mathbf{h} \in \mathbf{H} \right\}$ formális összegeket, ahol véges sok $\mathbf{h} \in \mathbf{H}$ kivétellel minden $a_{\mathbf{h}} = 0$. Két ilyen $\left\{ \sum_{\mathbf{h}} a_{\mathbf{h}} \mathbf{h} \mid \mathbf{h} \in \mathbf{H} \right\}$ és $\left\{ \sum_{\mathbf{h}} b_{\mathbf{h}} \mathbf{h} \mid \mathbf{h} \in \mathbf{H} \right\}$ összeget akkor tekintünk egyenlőnek, ha minden $\mathbf{h} \in \mathbf{H}$ esetén $a_{\mathbf{h}} = b_{\mathbf{h}}$. Ezek összeadását és R -beli elemmel való szorzását komponensenként végezzük. Világos, hogy ezáltal egy R -modulust nyertünk, amelynek „lényegében” \mathbf{H} egy bázisa. ■

Ha a továbbiakban a bihomomorfizmusokat kiterjesztjük az $\mathcal{U} \times \mathcal{V}$ generálta szabad R -modulusra, akkor ott már megfogalmazhatjuk a bihomomorfizmus tulajdonságait. Első lépésként ezt a kiterjesztést tetszőleges függvényre tesszük meg:

10.2. Tétel. Legyen $\Phi : \mathbf{H} \rightarrow \mathcal{F}_R(\mathbf{H})$ az a leképezés, amely \mathbf{H} tetszőleges \mathbf{h} elemét önmagára mint $\mathcal{F}_R(\mathbf{H})$ báziselemére képezi le. Ekkor tetszőleges \mathcal{M} R -modulus esetén létezik egy bijekció az $\mathbf{A} : \mathbf{H} \rightarrow \mathcal{M}$ függvények és az $\hat{\mathbf{A}} : \mathcal{F}_R(\mathbf{H}) \rightarrow \mathcal{M}$ homomorfizmusok között, amelyet $\mathbf{A} = \hat{\mathbf{A}}\Phi$ ad meg.

Bizonyítás. Világos, hogy ha $\hat{\mathbf{A}} : \mathcal{F}_R(\mathbf{H}) \rightarrow \mathcal{M}$ egy homomorfizmus, akkor az $\mathbf{A} = \hat{\mathbf{A}}\Phi$ függvényre $\mathbf{A} : \mathbf{H} \rightarrow \mathcal{M}$. Ha viszont \mathbf{A} adott, akkor $\hat{\mathbf{A}}$ csak az az $\hat{\mathbf{A}}$ függvény lehet, amelyre $\hat{\mathbf{A}}(\Phi(\mathbf{h})) = \mathbf{A}(\mathbf{h})$. Ilyen homomorfizmus viszont egyértelműen létezik, hiszen egy bázison definiált tetszőleges függvény egyértelműen kiterjeszthető homomorfizmussá. A definícióból világos, hogy a két megfeleltetés egymás inverze. ■

Most azt fogjuk megnézni, milyen feltétel mellett lesz \mathbf{A} bihomomorfizmus.

10.2. Definíció. Tekintsük az $\mathcal{F}_R(\mathcal{U} \times \mathcal{V})$ szabad modulusnak azt az \mathcal{L} részmodulusát, amelyet a következő elemek generálnak:

$$\begin{aligned} (\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) - (\mathbf{u}_1, \mathbf{v}) - (\mathbf{u}_2, \mathbf{v}) \quad \text{és} \quad (\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) - (\mathbf{u}, \mathbf{v}_1) - (\mathbf{u}, \mathbf{v}_2), \\ (c \cdot \mathbf{u}, \mathbf{v}) - c \cdot (\mathbf{u}, \mathbf{v}) \quad \text{és} \quad (\mathbf{u}, c \cdot \mathbf{v}) - c \cdot (\mathbf{u}, \mathbf{v}); \end{aligned}$$

ahol $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u} \in \mathcal{U}$, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v} \in \mathcal{V}$ és $c \in R$. A fenti elemeket elemi linearitásoknak, az \mathcal{L} elemeit pedig a linearitásoknak fogjuk nevezni. ■

10.3. Tétel. \mathbf{A} akkor és csak akkor bihomomorfizmus, ha $\mathcal{L} \leq \text{Ker}(\hat{\mathbf{A}})$.

Bizonyítás. A bihomomorfizmus definíciója szerint \mathbf{A} akkor és csak akkor bihomomorfizmus, ha

$$\begin{aligned} \mathbf{A}(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) &= \mathbf{A}(\mathbf{u}_1, \mathbf{v}) + \mathbf{A}(\mathbf{u}_2, \mathbf{v}) \quad \text{és} \quad \mathbf{A}(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = \mathbf{A}(\mathbf{u}, \mathbf{v}_1) + \mathbf{A}(\mathbf{u}, \mathbf{v}_2), \\ \mathbf{A}(c \cdot \mathbf{u}, \mathbf{v}) &= c \cdot \mathbf{A}(\mathbf{u}, \mathbf{v}) \quad \text{és} \quad \mathbf{A}(\mathbf{u}, c \cdot \mathbf{v}) = c \cdot \mathbf{A}(\mathbf{u}, \mathbf{v}). \end{aligned}$$

Írjunk most \mathbf{A} helyébe $\hat{\mathbf{A}}\Phi$ -t. Ekkor az

$$\begin{aligned} \hat{\mathbf{A}}\Phi(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) &= \hat{\mathbf{A}}\Phi(\mathbf{u}_1, \mathbf{v}) + \hat{\mathbf{A}}\Phi(\mathbf{u}_2, \mathbf{v}) \\ \text{és} \quad \hat{\mathbf{A}}\Phi(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) &= \hat{\mathbf{A}}\Phi(\mathbf{u}, \mathbf{v}_1) + \hat{\mathbf{A}}\Phi(\mathbf{u}, \mathbf{v}_2), \end{aligned}$$

$$\hat{\mathbf{A}}\Phi(c \cdot \mathbf{u}, \mathbf{v}) = c \cdot \hat{\mathbf{A}}\Phi(\mathbf{u}, \mathbf{v}) \quad \text{és} \quad \hat{\mathbf{A}}\Phi(\mathbf{u}, c \cdot \mathbf{v}) = c \cdot \hat{\mathbf{A}}\Phi(\mathbf{u}, \mathbf{v})$$

egyenlőségekhez jutunk. Mivel Φ a szabad modulusba képez, ahol az $\mathcal{U} \times \mathcal{V}$ elemeinek már képezhetjük a formális lineáris kombinációit, ezért azt kapjuk, hogy

$$\hat{\mathbf{A}}(\Phi(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) - \Phi(\mathbf{u}_1, \mathbf{v}) - \Phi(\mathbf{u}_2, \mathbf{v})) = \mathbf{o},$$

$$\hat{\mathbf{A}}(\Phi(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) - \Phi(\mathbf{u}, \mathbf{v}_1) - \Phi(\mathbf{u}, \mathbf{v}_2)) = \mathbf{o},$$

$$\hat{\mathbf{A}}(\Phi(c \cdot \mathbf{u}, \mathbf{v}) - c \cdot \Phi(\mathbf{u}, \mathbf{v})) = \mathbf{o} \quad \text{és} \quad \hat{\mathbf{A}}(\Phi(\mathbf{u}, c \cdot \mathbf{v}) - c \cdot \Phi(\mathbf{u}, \mathbf{v})) = \mathbf{o}.$$

Így valóban $\mathcal{L} \leq \text{Ker}(\hat{\mathbf{A}})$, hiszen tartalmazza \mathcal{L} egy generátorrendszerét.

Tegyük most fel, hogy $\mathcal{L} \leq \text{Ker}(\widehat{\mathbf{A}})$. Ez azt jelenti, hogy $\widehat{\mathbf{A}}$ az alábbi elemek mindegyikét \mathbf{o} -ba viszi:

$$\begin{aligned} \Phi(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) - \Phi(\mathbf{u}_1, \mathbf{v}) - \Phi(\mathbf{u}_2, \mathbf{v}) & \quad \text{és} \quad \Phi(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) - \Phi(\mathbf{u}, \mathbf{v}_1) - \Phi(\mathbf{u}, \mathbf{v}_2), \\ \Phi(c \cdot \mathbf{u}, \mathbf{v}) - c \cdot \Phi(\mathbf{u}, \mathbf{v}) & \quad \text{és} \quad \Phi(\mathbf{u}, c \cdot \mathbf{v}) - c \cdot \Phi(\mathbf{u}, \mathbf{v}). \end{aligned}$$

Ennek alapján, az $\mathbf{A} = \widehat{\mathbf{A}}\Phi$ felhasználásával az

$$\begin{aligned} \mathbf{A}(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) &= \mathbf{A}(\mathbf{u}_1, \mathbf{v}) + \mathbf{A}(\mathbf{u}_2, \mathbf{v}) \quad \text{és} \quad \mathbf{A}(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = \mathbf{A}(\mathbf{u}, \mathbf{v}_1) + \mathbf{A}(\mathbf{u}, \mathbf{v}_2), \\ \mathbf{A}(c \cdot \mathbf{u}, \mathbf{v}) &= c \cdot \mathbf{A}(\mathbf{u}, \mathbf{v}) \quad \text{és} \quad \mathbf{A}(\mathbf{u}, c \cdot \mathbf{v}) = c \cdot \mathbf{A}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

összefüggést kapjuk, tehát \mathbf{A} valóban bihomomorfizmus. ■

Tekintsünk most egy tetszőleges \mathcal{U} R -modulust és annak egy \mathcal{V} részmodulusát. Mint tudjuk, létezik az \mathcal{U}/\mathcal{V} faktormodulus, amelynek elemei a \mathcal{V} eltoltjai. Az \mathcal{U} -beli \mathbf{u} vektort tartalmazó eltoltat $[\mathbf{u}]$ -val jelöltük. Láttuk, hogy ezek R -modulust alkotnak az $[\mathbf{u}] + [\mathbf{v}] = [\mathbf{u} + \mathbf{v}]$ és $c[\mathbf{u}] = [c\mathbf{u}]$ műveletekre. Azt is láttuk, hogy $\Psi : \mathbf{u} \mapsto [\mathbf{u}]$ homomorfizmus, amelynek magja \mathcal{V} .

10.4. Tétel. *Ha $\varphi : \mathcal{U} \rightarrow \mathcal{M}$ tetszőleges olyan homomorfizmus, amelynek magjára $\mathcal{V} \leq \text{Ker}(\varphi)$ teljesül, akkor létezik pontosan egy olyan $\varphi/\mathcal{V} : \mathcal{U}/\mathcal{V} \rightarrow \mathcal{M}$ homomorfizmus, amelyre $\varphi = (\varphi/\mathcal{V})\Psi$. Minden $\mathcal{U}/\mathcal{V} \rightarrow \mathcal{M}$ homomorfizmus egyértelműen felírható ilyen alakban. (Második izomorfizmustétel.)*

Bizonyítás. Definíálnunk kell egy, a feltételeknek eleget tevő $\psi : \mathcal{U}/\mathcal{V} \rightarrow \mathcal{M}$ függvényt. A $\varphi = \psi\Psi$ összefüggésből következik, hogy minden $\mathbf{u} \in \mathcal{U}$ mellett $\varphi(\mathbf{u}) = \psi(\Psi(\mathbf{u})) = \Psi([\mathbf{u}])$, ami megadja, hogy csak $\psi([\mathbf{u}]) = \varphi(\mathbf{u})$ lehet a definíció (ezzel egyébként a második állításban szereplő egyértelműséget is beláttuk). Meg kell mutatni, hogy az így definiált ψ valóban homomorfizmus. A művelettartás a faktormodulus definíciójából azonnal következik. Az ilyen esetekben szokásos módon csak az a kérdés, hogy ψ leképezés-e. Hiszen a leképezendő elemeket többféleképpen is megadhatjuk. Legyen $[\mathbf{u}] = [\mathbf{u}']$. Ez azt jelenti, hogy $[\mathbf{u} - \mathbf{u}'] = [\mathbf{o}]$, azaz $\mathbf{u} - \mathbf{u}' \in \mathcal{V}$. Tekintettel arra, hogy $\mathcal{V} \leq \text{Ker}(\varphi)$, ezért $\mathbf{u} - \mathbf{u}' \in \text{Ker}(\varphi)$, azaz $\varphi(\mathbf{u} - \mathbf{u}') = \mathbf{o}$, és így $\varphi(\mathbf{u}) = \varphi(\mathbf{u}')$. Ez pedig pont azt jelenti, hogy \mathcal{U}/\mathcal{V} elemeinek ψ -nél egyértelmű a képe. Tehát $\varphi/\mathcal{V} = \psi$ a megfelelő választás.

Ezzel tulajdonképpen a második állítást is beláttuk, hiszen adott ψ esetén $\varphi = \psi\Psi$ megfelel a követelményeknek. ■

10.3. Definíció. Legyenek adottak az \mathcal{U} és \mathcal{V} R -modulusok, és legyen \mathcal{L} az $\mathcal{F}_R(\mathcal{U} \times \mathcal{V})$ -beli linearitások halmaza (\mathcal{L} részmodulus). Jelölje $\mathcal{U} \otimes \mathcal{V}$ az $\mathcal{F}_R(\mathcal{U} \times \mathcal{V})/\mathcal{L}$ modulust. Az

$$\mathcal{U} \times \mathcal{V} \xrightarrow{\Phi} \mathcal{F}_R(\mathcal{U} \times \mathcal{V}) \xrightarrow{\Psi} \mathcal{U} \otimes \mathcal{V}$$

diagramban legyen $\mathfrak{T} = \Psi\Phi$. Ekkor $\mathcal{U} \otimes \mathcal{V}$ -t az \mathcal{U} és \mathcal{V} (ebben a sorrendben vett) tenzorszorzatának, és az $\mathcal{U} \times \mathcal{V} \xrightarrow{\mathfrak{T}} \mathcal{U} \otimes \mathcal{V}$ diagramot tenzorszorzat-diagramnak nevezzük. ■

10.5. Tétel. $\mathcal{U} \times \mathcal{V} \xrightarrow{\mathfrak{T}} \mathcal{U} \otimes \mathcal{V}$ bihomomorfizmus. Minden $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{M}$ bihomomorfizmushoz létezik pontosan egy olyan $\alpha : \mathcal{U} \otimes \mathcal{V} \rightarrow \mathcal{M}$ homomorfizmus, amelyre $\mathbf{A} = \alpha \mathfrak{T}$, azaz az

$$\begin{array}{ccc} \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathbf{A}} & \mathcal{M} \\ \mathfrak{T} \downarrow & & \downarrow \iota \\ \mathcal{U} \otimes \mathcal{V} & \xrightarrow{\alpha} & \mathcal{M} \end{array}$$

diagram kommutatív.

Bizonyítás. Definíció szerint $\mathfrak{T} = \Psi\Phi$, és $\text{Ker}(\Psi) = \mathcal{L}$. A 10.3. Tétel szerint tehát \mathfrak{T} bihomomorfizmus. A második állítás azonnal következik a 10.3. és 10. 4. Tételekből. ■

A most bizonyított tétel teljesen definiálja a tenzorszorzatot a következő értelemben:

10.6. Tétel. Legyen $\mathfrak{t} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{T}$ egy olyan bihomomorfizmus, amelyre a következő igaz:

Tetszőleges $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bihomomorfizmushoz létezik pontosan egy olyan $\varphi : \mathcal{T} \rightarrow \mathcal{W}$ homomorfizmus, amelyre $\mathbf{A} = \varphi \cdot \mathfrak{t}$, azaz az alábbi diagram kommutatív:

$$\begin{array}{ccc} \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{t}} & \mathcal{T} \\ \iota \downarrow & & \downarrow \varphi \\ \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathbf{A}} & \mathcal{W}. \end{array}$$

Ekkor léteznek olyan α és β izomorfizmusok, amelyek egymás inverzei és amelyekre

$$\begin{array}{ccc} \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{T}} & \mathcal{U} \otimes \mathcal{V} \\ \iota \downarrow & & \downarrow \alpha \\ \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{t}} & \mathcal{T} \end{array} \quad \begin{array}{ccc} \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{t}} & \mathcal{T} \\ \iota \downarrow & & \downarrow \beta \\ \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{T}} & \mathcal{U} \otimes \mathcal{V}. \end{array}$$

kommutatív diagramok. ■

Bizonyítás. A tenzorszorzat előbb bizonyított tulajdonsága és a tételben feltett tulajdonság alapján a fenti α és β homomorfizmusok léteznek; csak azt kell belátni, hogy egymás inverzei (ebből ugyanis azonnal következik az is, hogy izomorfizmusok). A fenti két diagramot „egymáshoz fűzve” az alábbi két diagramot nyerjük:

$$\begin{array}{ccc} \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{T}} & \mathcal{U} \otimes \mathcal{V} \\ \iota \downarrow & & \downarrow \beta\alpha \\ \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{T}} & \mathcal{U} \otimes \mathcal{V} \end{array} \quad \begin{array}{ccc} \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{t}} & \mathcal{T} \\ \iota \downarrow & & \downarrow \alpha\beta \\ \mathcal{U} \times \mathcal{V} & \xrightarrow{\mathfrak{t}} & \mathcal{T}. \end{array}$$

amelyek nyilván ugyancsak kommutatívak.

Világos, hogy a fenti két diagram oly módon is kommutatívvá tehető, hogy $\beta\alpha$ helyére az $\mathcal{U} \otimes \mathcal{V}$ -nek, az $\alpha\beta$ helyére pedig a \mathcal{T} -nek az identitását tesszük. Tekintettel arra, hogy a *kiegészítés egyértelmű*, ezért $\beta\alpha = \iota_{\mathcal{U} \otimes \mathcal{V}}$ és $\alpha\beta = \iota_{\mathcal{T}}$. ■

Nagyon fontos, hogy ebből nem csak az derül ki, hogy a tenzorszorzat izomorfizmus erejéig egyértelmű, hanem az is, hogy az izomorfizmus a „*tenzorszorzatra való leképezések*”-et is „egymásba viszi”.

A 10.6. Tétel alapján a tenzorszorzat definiálható volna a most bebizonyított tulajdonsággal, amelyet „*ko-univerzalitás*”-nak nevezünk. Ez egy sokkal egyszerűbben megfogalmazható tulajdonság, csak az a hibája, hogy nem következik belőle a tenzorszorzat létezése.

10.7. Tétel. Jelölje $\mathbf{u} \otimes \mathbf{v}$ az $(\mathbf{u}, \mathbf{v}) \in \mathcal{U} \times \mathcal{V}$ elem \mathfrak{T} -nél vett képét: $\mathbf{u} \otimes \mathbf{v} = \mathfrak{T}((\mathbf{u}, \mathbf{v}))$. Ezek az elemek $\mathcal{U} \otimes \mathcal{V}$ egy generátorrendszerét alkotják, és közöttük az alábbi összefüggések állnak fenn:

$$(\mathbf{u}_1 + \mathbf{u}_2) \otimes \mathbf{v} = \mathbf{u}_1 \otimes \mathbf{v} + \mathbf{u}_2 \otimes \mathbf{v}, \quad \mathbf{u} \otimes (\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{u} \otimes \mathbf{v}_1 + \mathbf{u} \otimes \mathbf{v}_2,$$

$$c \cdot (\mathbf{u} \otimes \mathbf{v}) = (c \cdot \mathbf{u}) \otimes \mathbf{v} = \mathbf{u} \otimes (c \cdot \mathbf{v}).$$

Bizonyítás. Definíció szerint az (\mathbf{u}, \mathbf{v}) párok $\mathcal{F}_R(\mathcal{U} \times \mathcal{V})$ egy bázisát alkotják, tehát generátorrendszert. Mivel bármely homomorfizmus minden generátorrendszert generátorrendszerbe visz, ezért az $\mathbf{u} \otimes \mathbf{v}$ elemek valóban generátorrendszert alkotnak. Bármelyik felírt egyenlőségben a baloldalt és a jobboldalt szereplő elemek olyan $\mathcal{F}_R(\mathcal{U} \times \mathcal{V})$ -beli elemek képei, amelyek ugyanazon \mathcal{L} szerinti eltoltban fekszenek, tehát Ψ -nél vett képük megegyezik. ■

Mint említettük, vektorterek esetében van a definícióra egy harmadik lehetőség. Régebben ezt használták; de sok olyan tulajdonságot nem lehetett látni belőle, amely a tenzorszorzatnak alapvetően fontos tulajdonsága. Mindenekelőtt nincs „beleépítve” a \mathfrak{T} bihomomorfizmus. Igen kevés előnyt jelent emellett, hogy „könnyebb”. Ezt a tulajdonságot írja le az alábbi:

10.8. Tétel. Ha $\{\mathbf{u}_i \mid i \in I\}$ az \mathcal{U} , $\{\mathbf{v}_j \mid j \in J\}$ a \mathcal{V} egy-egy bázisa, akkor az $\{\mathbf{u}_i \otimes \mathbf{v}_j \mid i \in I, j \in J\}$ elemek az $\mathcal{U} \otimes \mathcal{V}$ egy bázisát alkotják.

Bizonyítás. Mint láttuk, az összes $\mathbf{u} \otimes \mathbf{v}$ alakú elem $\mathcal{U} \otimes \mathcal{V}$ egy generátorrendszere. A bilinearitás alapján ezek mindegyike felírható az $\mathbf{u}_i \otimes \mathbf{v}_j$ -k lineáris kombinációjaként; tehát a tételbeli elemek generátorrendszert alkotnak. Mint a bihomomorfizmusok tárgyalásánál bizonyítottuk, létezik olyan $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bihomomorfizmus, amely a báziselempárokon tetszőlegesen előírt. Legyen $\{\mathbf{u}_i^* \mid i \in I\}$ és $\{\mathbf{v}_j^* \mid j \in J\}$ a megfelelő terek egy-egy duális bázisa. Adott $p \in I, q \in J$ indexre tekintsük most azt az $\mathbf{A}_{p,q} : \mathcal{U} \times \mathcal{V} \rightarrow R$ függvényt, amelyet $\mathbf{A}_{p,q} : (\mathbf{u}, \mathbf{v}) \mapsto (\mathbf{u}_p^*(\mathbf{u})) \cdot (\mathbf{v}_q^*(\mathbf{v}))$ definiál. Ez nyilvánvalóan bihomomorfizmus, amely az $(\mathbf{u}_p, \mathbf{v}_q)$ elemet R egységelemébe viszi; és minden más $(\mathbf{u}_i, \mathbf{v}_j)$ báziselemet 0-ba. Tekintsük azt az egyértelműen meghatározott $\varphi_{p,q} : \mathcal{U} \otimes \mathcal{V} \rightarrow R$ homomorfizmust,

amelyre $\mathbf{A}_{p,q} = \varphi_{p,q} \mathfrak{T}$. Legyen $\sum_{i,j} c_{i,j} \mathbf{u}_i \otimes \mathbf{v}_j = 0$. Ekkor:

$$0 = \varphi_{p,q} \left(\sum_{i,j} c_{i,j} \mathbf{u}_i \otimes \mathbf{v}_j \right) = \varphi_{p,q} \mathfrak{T} \left(\sum_{i,j} c_{i,j} (\mathbf{u}_i, \mathbf{v}_j) \right) = \mathbf{A}_{p,q} \left(\sum_{i,j} c_{i,j} (\mathbf{u}_i, \mathbf{v}_j) \right) = c_{p,q}.$$

Eszerint a kiválasztott elem bármely együtthatója 0; tehát a felsorolt elemek valóban lineárisan függetlenek.

Megjegyzések

1. A fenti eljárás mutatja a legabsztraktabb módszer fontosságát. Ez az eljárás adja az egyetlen közvetlen lehetőséget arra, hogy a tenzorszorzat bizonyos elemeinek a függetlenségét bebizonyítsuk.

2. A fentiekben az R gyűrűt rögzítettük. Egy-egy modulus több gyűrű fölötti modulus is lehet. Ennek megfelelően az \otimes jel helyett célszerű az \otimes_R jelet használni. \square

Feladatok

1. Bizonyítsuk be, hogy $\mathcal{U} \otimes \mathcal{V}$ minden eleme felírható $\mathbf{u} \otimes \mathbf{v}$ alakú elemek (együttható nélküli) összegeként.

2. Legyenek $\mathbf{u}_i \otimes \mathbf{v}_i \in \mathcal{U} \otimes \mathcal{V}$, ahol $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ vektorrendszerek. Bizonyítsuk be, hogy a $\sum_i \mathbf{u}_i \otimes \mathbf{v}_i$ felírható kevesebb $\mathbf{u} \otimes \mathbf{v}$ alakú elem összegeként, ha akár az \mathbf{U} , akár a \mathbf{V} rendszer lineárisan összefüggő.

3. Bizonyítsuk be, hogy vektorterek esetén, ha az előző feladatban \mathbf{U} és \mathbf{V} függetlenek, akkor a vizsgált összeg nem írható fel kevesebb tag összegeként.

4. Legyenek \mathcal{U} és \mathcal{V} egydimenziós vektorterek. Bizonyítsuk be, hogy $\mathcal{U} \otimes \mathcal{V} \cong \mathcal{U}$.

5. Bizonyítsuk be, hogy $\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C}$ és $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}$ nem izomorf \mathbb{R} -vektorterek.

2. A tenzorszorzat elemi tulajdonságai

10.9. Tétel. $R \otimes \mathcal{U} \cong \mathcal{U}$, a tenzorszorzás kommutatív és asszociatív, azaz $\mathcal{U} \otimes \mathcal{V} \cong \mathcal{V} \otimes \mathcal{U}$ és $(\mathcal{U} \otimes \mathcal{V}) \otimes \mathcal{W} \cong \mathcal{U} \otimes (\mathcal{V} \otimes \mathcal{W})$. A fenti izomorfizmusok természetesek.

Bizonyítás. Mindenekelőtt megmutatjuk, hogy az $\mathbf{u} \mapsto 1 \otimes \mathbf{u}$ leképezés természetes izomorfizmus.

Feltelessük meg a (c, \mathbf{u}) párnak $(c \in R, \mathbf{u} \in \mathcal{U})$ az \mathcal{U} -beli $c\mathbf{u}$ elemet. Ez nyilvánvalóan egy $\mathbf{A} : R \times \mathcal{U} \rightarrow \mathcal{U}$ bihomomorfizmus. Így van olyan $\varphi : R \otimes \mathcal{U} \rightarrow \mathcal{U}$ homomorfizmus, hogy $\mathbf{A} = \varphi \mathfrak{T}$.

$R \otimes \mathcal{U}$ -ban a $c \otimes \mathbf{u} = 1 \otimes c \cdot \mathbf{u}$ alakú elemek generátorrendszert alkotnak. Mivel ilyen alakú elemek összege is ilyen alakú, ezért $R \otimes \mathcal{U}$ minden eleme $1 \otimes \mathbf{u}$ alakú. Ha $1 \otimes \mathbf{u} \in \text{Ker}(\varphi)$, akkor $\mathbf{o} = \varphi(1 \otimes \mathbf{u}) = \varphi\mathfrak{T}((1, \mathbf{u})) = \mathbf{A}((1, \mathbf{u})) = \mathbf{u}$ miatt $1 \otimes \mathbf{u} = \mathbf{o}$. Ebből azonnal következik, hogy φ izomorfizmus; amelynek Ψ inverze éppen a megadott izomorfizmus.

Az, hogy a fenti $\Psi : \mathcal{U} \rightarrow R \otimes \mathcal{U}$ természetes (hasonlóan az $\mathcal{U} \rightarrow \text{Hom}(K, \mathcal{U})$ esethez), a következőket jelenti:

Legyen $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ és jelölje $1 \otimes \varphi$ azt a leképezést, amelyre $1 \otimes \varphi : 1 \otimes \mathbf{u} \mapsto 1 \otimes \varphi(\mathbf{u})$; ekkor $\Psi\varphi = (1 \otimes \varphi)\Psi$. Ez viszont azonnal látható.

A kommutativitás bizonyításához tekintsük a $\mathfrak{T}_1 : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{U} \otimes \mathcal{V}$ és a $\mathfrak{T}_2 : \mathcal{V} \times \mathcal{U} \rightarrow \mathcal{V} \otimes \mathcal{U}$ tenzorszorzatokat, valamint az $\mathbf{A}_1(\mathbf{u}, \mathbf{v}) = \mathbf{v} \otimes \mathbf{u}$ és az $\mathbf{A}_2(\mathbf{v}, \mathbf{u}) = \mathbf{u} \otimes \mathbf{v}$ összefüggéssel definiált leképezéseket. Ez utóbbiak bihomomorfizmusok, így léteznek olyan φ_1, φ_2 homomorfizmusok, amelyekre $\mathbf{A}_1 = \varphi_1 \mathfrak{T}_1$ és $\mathbf{A}_2 = \varphi_2 \mathfrak{T}_2$ teljesül. Ebből a szokásos módon következik, hogy φ_1, φ_2 egymás inverzei — tehát izomorfizmusok is. Itt is belátható a „természetesség”.

Az asszociativitás is hasonló elven látható be. Itt a bihomomorfizmus (vagy „trihomomorfizmus”?) az $((u, v), w) \mapsto \mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w})$ megfeleltetéssel — illetve ennek duálisával — definiálható. Ezek a leképezések bármely két változó rögzítése esetén a harmadikban lineárisak. Ebből kapható, hogy átvezethetők a megfelelő tenzorszorzaton, ami itt is az izomorfizmushoz vezet (amely természetesen természetes). ■

10.10. Tétel. *Tetszőleges \mathcal{A}, \mathcal{B} és \mathcal{C} vektorterek (sőt kommutatív R -re R -modulusok) esetén $\text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{C})$ és $\text{Hom}(\mathcal{A}, \text{Hom}(\mathcal{B}, \mathcal{C}))$ (természetes módon) izomorfak.*

Bizonyítás. Ha $\varphi \in \text{Hom}(\mathcal{A}, \text{Hom}(\mathcal{B}, \mathcal{C}))$, akkor φ minden $\mathbf{a} \in \mathcal{A}$ elemhez a $\text{Hom}(\mathcal{B}, \mathcal{C})$ egy $\varphi_{\mathbf{a}}$ elemét rendeli hozzá. Erre $\varphi_{\mathbf{a}}(\mathbf{b}) = \mathbf{c} \in \mathcal{C}$. Mivel φ is és $\varphi_{\mathbf{a}}$ is homomorfizmus, ezért a megfeleltetés $\mathcal{A} \times \mathcal{B}$ -nek egy \mathcal{C} -be való bihomomorfizmusa. Létezik tehát egy olyan egyértelmű $\Phi \in \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{C})$, amelyre $\varphi = \Phi \mathfrak{T}$. Ez a megfeleltetés „nyilván” bijektív.

A természetesség a következőket jelenti: Legyenek adottak az

$$\mathbf{A} : \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{C}) \rightarrow \text{Hom}(\mathcal{A}, \text{Hom}(\mathcal{B}, \mathcal{C})) \quad \text{és}$$

$$\mathbf{A}' : \text{Hom}(\mathcal{A}' \otimes \mathcal{B}', \mathcal{C}') \rightarrow \text{Hom}(\mathcal{A}', \text{Hom}(\mathcal{B}', \mathcal{C}'))$$

megfelelő izomorfizmusok. Ekkor tetszőleges

$$\alpha : \mathcal{A}' \rightarrow \mathcal{A}, \quad \beta : \mathcal{B}' \rightarrow \mathcal{B}, \quad \gamma : \mathcal{C} \rightarrow \mathcal{C}'$$

esetén „természetesen” értelmezhetők a

$$\text{Hom}(\alpha, \text{Hom}(\beta, \gamma)) : \text{Hom}(\mathcal{A}, \text{Hom}(\mathcal{B}, \mathcal{C})) \rightarrow \text{Hom}(\mathcal{A}', \text{Hom}(\mathcal{B}', \mathcal{C}')),$$

valamint a

$$\text{Hom}(\alpha \otimes \beta, \gamma) : \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{C}) \rightarrow \text{Hom}(\mathcal{A}' \otimes \mathcal{B}', \mathcal{C}')$$

homomorfizmusok; és ezekre a

$$\begin{array}{ccc} \text{Hom}(\mathcal{A}, \text{Hom}(\mathcal{B}, \mathcal{C})) & \xrightarrow{\mathbf{A}} & \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{C}) \\ \text{Hom}(\alpha, \text{Hom}(\beta, \gamma)) \downarrow & & \text{Hom}(\alpha \otimes \beta, \gamma) \downarrow \\ \text{Hom}(\mathcal{A}', \text{Hom}(\mathcal{B}', \mathcal{C}')) & \xrightarrow{\mathbf{A}'} & \text{Hom}(\mathcal{A}' \otimes \mathcal{B}', \mathcal{C}') \end{array}$$

diagram mindig kommutatív. Ennek a bizonyítását itt nem végezzük el. ■

Megjegyzés. Vektorterek esetében ebből következik, hogy:

$$(\mathcal{A} \otimes \mathcal{B})^* = \text{Hom}(\mathcal{A} \otimes \mathcal{B}, K) \cong \text{Hom}(\mathcal{A}, \text{Hom}(\mathcal{B}, K)) = \text{Hom}(\mathcal{A}, \mathcal{B}^*). \quad \square$$

10.11. Tétel. *Létezik egy $\mathbf{A} : \text{Hom}(\mathcal{A}, \mathcal{B}) \otimes \mathcal{C} \rightarrow \text{Hom}(\mathcal{A}, \mathcal{B} \otimes \mathcal{C})$ természetes homomorfizmus.*

Bizonyítás. Adott $\varphi \in \text{Hom}(\mathcal{A}, \mathcal{B})$ és $\mathbf{c} \in \mathcal{C}$ esetén feleltessük meg a (φ, \mathbf{c}) párnak azt a leképezést, amely az \mathcal{A} -beli \mathbf{a} elemet a $\varphi(\mathbf{a}) \otimes \mathbf{c}$ elembe viszi. Mivel ez bihomomorfizmus, ezért létezik a kívánt \mathbf{A} homomorfizmus. A „természetesség” az előző tételhez hasonlóan egy

$$\begin{array}{ccc} \text{Hom}(\mathcal{A}, \mathcal{B}) \otimes \mathcal{C} & \xrightarrow{\mathbf{A}} & \text{Hom}(\mathcal{A}, \mathcal{B} \otimes \mathcal{C}) \\ \text{Hom}(\alpha, \beta) \otimes \gamma \downarrow & & \text{Hom}(\alpha, \beta \otimes \gamma) \downarrow \\ \text{Hom}(\mathcal{A}', \mathcal{B}') \otimes \mathcal{C}' & \xrightarrow{\mathbf{A}'} & \text{Hom}(\mathcal{A}', \mathcal{B}' \otimes \mathcal{C}') \end{array}$$

alakú diagram kommutativitását jelenti, alkalmas

$$\alpha : \mathcal{A}' \rightarrow \mathcal{A}, \quad \beta : \mathcal{B} \rightarrow \mathcal{B}', \quad \gamma : \mathcal{C} \rightarrow \mathcal{C}'$$

homomorfizmusokkal. ■

A megfelelő izomorfizmusokat figyelembe véve véges dimenziós vektorterekre a következőket kapjuk:

Legyenek az $\mathcal{A}, \mathcal{B}, \mathcal{C}$ vektorterek egy-egy bázisának elemei az $\mathbf{a}_i, \mathbf{b}_j, \mathbf{c}_k$ vektorok. Az \mathcal{A}^* duális bázisának elemei az \mathbf{a}_i^* leképezések. Ekkor az első vektortérnek egy bázisa a $(\mathbf{b}_j \cdot \mathbf{a}_i^*) \otimes \mathbf{c}_k$ elemekből áll. \mathbf{A} ezeknek a $(\mathbf{b}_j \otimes \mathbf{c}_k) \cdot \mathbf{a}_i^*$ elemeket felelteti meg; ami viszont a második vektortér egy bázisa. Eszerint \mathbf{A} ebben az esetben izomorfizmus.

A $\mathcal{B} = K$ speciális esetben az $\mathcal{A}^* \otimes \mathcal{C} \cong \text{Hom}(\mathcal{A}, \mathcal{C})$ izomorfizmushoz jutunk; tekintettel arra, hogy $K \otimes \mathcal{C} \cong \mathcal{C}$. (Mivel itt minden izomorfizmus természetes, ezért a most kapott izomorfizmus is az.)

Mostani eredményünket összevetve azzal, amit az előbb kaptunk, véges dimenziós vektorterekre a következő adódik:

$$(\mathcal{A} \otimes \mathcal{B})^* \cong \text{Hom}(\mathcal{A}, \mathcal{B}^*) \cong \mathcal{A}^* \otimes \mathcal{B}^*.$$

10.12. Tétel. *Létezik egy $\mathbf{A} : \mathcal{A} \rightarrow (\mathcal{A}^*)^*$ természetes homomorfizmus.*

Bizonyítás. Feleltessük meg az \mathcal{A} -beli \mathbf{a} vektornak azt az α homomorfizmust, amely az $\mathbf{u}^* \in \mathcal{A}$ leképezésnek az $\mathbf{u}^*(\mathbf{a})$ elemet felelteti meg. Ez nyilván homomorfizmus; amelyről könnyen belátható, hogy természetes. ■

Ha a természetesen izomorf vektorterek között „nem teszünk különbséget”, akkor véges dimenziós vektorterek esetében minden egyes vizsgált funktornál (vagyis vektorteret vektortérbe vivő függvénynél) elég az eredeti tereket, a duális tereket és ezek tenzorszorzatát nézni.

10.4. Definíció. Tetszőleges $\mathcal{U}_1, \dots, \mathcal{U}_r, \mathcal{V}_1, \dots, \mathcal{V}_s$ K -vektorterek esetén az

$$\mathcal{U}_1^* \otimes \dots \otimes \mathcal{U}_r^* \otimes \mathcal{V}_1 \otimes \dots \otimes \mathcal{V}_s$$

elemeit tenzoroknak nevezzük. Azt mondjuk, hogy ezek a tenzorok az első r változóban kovariánsak; a többiben pedig kontravariánsak. ■

Megjegyzés. Érdemes figyelni arra, hogy egy tenzor éppen azokban a változókban kovariáns, amelyeket kontravariáns funktor hoz létre; míg azokban, amelyeket kovariáns funktor hoz létre, a tenzor kontravariáns. Ennek az az oka, hogy amikor a tenzorokat használni kezdték, még nem voltak funktorok, és a varianciát annak megfelelően tekintették, hogy új bázisra áttérve miképpen változnak meg a koordináták. □

Mátrixok tenzorszorzatát kétféleképpen is tekinthetjük. Egyrészt, mint egy olyan „szupermátrixot”, amely a tenzorszorzaton hat; másrészt, mint az eredeti mátrixok vektorterének a tenzorszorzatát. Ezek nem azonos fogalmak, de szoros kapcsolat van közöttük:

10.13. Tétel. *Létezik egy $\Lambda : \text{Hom}(\mathcal{A}, \mathcal{U}) \otimes \text{Hom}(\mathcal{B}, \mathcal{V}) \rightarrow \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{U} \otimes \mathcal{V})$ természetes homomorfizmus.*

Bizonyítás. Tetszőleges $(\alpha, \beta) \in \text{Hom}(\mathcal{A}, \mathcal{U}) \times \text{Hom}(\mathcal{B}, \mathcal{V})$ párnak feleltessük meg azt az $[\alpha, \beta] : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{U} \otimes \mathcal{V}$ függvényt, amelyre $[\alpha, \beta] : (\mathbf{a}, \mathbf{b}) \mapsto \alpha(\mathbf{a}) \otimes \beta(\mathbf{b})$. Könnyen látható, hogy ez egy bihomomorfizmus. Létezik tehát egy olyan $\{\alpha, \beta\} : \mathcal{A} \otimes \mathcal{B} \rightarrow \mathcal{U} \otimes \mathcal{V}$ homomorfizmus, amelyre $\{\alpha, \beta\} : \mathbf{a} \otimes \mathbf{b} \mapsto \alpha\mathbf{a} \otimes \beta\mathbf{b}$ és $[\alpha, \beta] = \{\alpha, \beta\} \cdot \mathfrak{T}_1$, ahol $\mathfrak{T}_1 : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{A} \otimes \mathcal{B}$ a tenzorhomomorfizmus.

Ezáltal létrehoztunk egy $\{\cdot\} : \text{Hom}(\mathcal{A}, \mathcal{U}) \times \text{Hom}(\mathcal{B}, \mathcal{V}) \rightarrow \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{U} \otimes \mathcal{V})$ leképezést. Ez is nyilvánvalóan bihomomorfizmus. Ezért felírható $\{\cdot\} = \Lambda \cdot \mathfrak{T}_2$ alakba, ahol $\mathfrak{T}_2 : \text{Hom}(\mathcal{A}, \mathcal{U}) \times \text{Hom}(\mathcal{B}, \mathcal{V}) \rightarrow \text{Hom}(\mathcal{A}, \mathcal{U}) \otimes \text{Hom}(\mathcal{B}, \mathcal{V})$ a tenzorhomomorfizmus.

Annak a bizonyítására, hogy ez a homomorfizmus természetes, mindenekelőtt meg kell mondani, hogy milyen homomorfizmusokat definiálunk a $\text{Hom}(\mathcal{A}, \mathcal{U}) \otimes \text{Hom}(\mathcal{B}, \mathcal{V})$ -ken és a $\text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{U} \otimes \mathcal{V})$ -ken. A megfelelő „varianciákat” figyelembe véve legyenek adottak a

$$\varphi : \mathcal{A} \leftarrow \mathcal{A}', \quad \psi : \mathcal{B} \leftarrow \mathcal{B}', \quad \sigma : \mathcal{U} \rightarrow \mathcal{U}', \quad \tau : \mathcal{V} \rightarrow \mathcal{V}'$$

homomorfizmusok. Ezek a következő homomorfizmusokat definiálják:

$$\zeta : \text{Hom}(\mathcal{A}, \mathcal{U}) \otimes \text{Hom}(\mathcal{B}, \mathcal{V}) \rightarrow \text{Hom}(\mathcal{A}', \mathcal{U}') \otimes \text{Hom}(\mathcal{B}', \mathcal{V}'),$$

$$\xi : \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{U} \otimes \mathcal{V}) \rightarrow \text{Hom}(\mathcal{A}' \otimes \mathcal{B}', \mathcal{U}' \otimes \mathcal{V}'),$$

ahol $\zeta : \alpha \otimes \beta \mapsto \sigma\alpha\varphi \otimes \tau\beta\psi$; ha $\pi : \mathbf{a} \otimes \mathbf{b} \mapsto \pi(\mathbf{a} \otimes \mathbf{b})$, akkor legyen $\xi(\pi) : \mathbf{a}' \otimes \mathbf{b}' \mapsto \Lambda(\sigma \otimes \tau)\zeta \Lambda(\varphi \otimes \psi)(\mathbf{a}' \otimes \mathbf{b}')$. Egyszerű számolással belátható, hogy $\Lambda\zeta = \xi\Lambda$. ■

Megjegyzés. A fent definiált Λ általában nemcsak hogy nem izomorfizmus, de se nem szűrjektív, se nem injektív.

Tekintsünk példaként \mathbb{Z} -modulusokat, azaz Abel-csoportokat. Legyenek $\mathcal{A} = \langle \mathbf{a} \rangle$, $\mathcal{B} = \langle \mathbf{b} \rangle$, $\mathcal{U} = \langle \mathbf{u} \rangle$, $\mathcal{V} = \langle \mathbf{v} \rangle$ ciklikus csoportok; az első kettő p -elemű, a második kettő p^2 -elemű. Ekkor $\text{Hom}(\mathcal{A}, \mathcal{U})$ egy α generálta ciklikus csoport, amelynek generátorelemére $\alpha(\mathbf{a}) = p\mathbf{u}$, és hasonlóképpen $\text{Hom}(\mathcal{B}, \mathcal{V})$ egy β generálta ciklikus csoport, amelynek generátorelemére $\beta(\mathbf{b}) = p\mathbf{v}$. $\mathcal{A} \otimes \mathcal{B}$ is ciklikus p -edrendű csoport $\mathbf{a} \otimes \mathbf{b}$ generátorelemmel, míg a ciklikus p^2 -rendű $\mathcal{U} \otimes \mathcal{V}$ generátoreleme $\mathbf{u} \otimes \mathbf{v}$. Az előzőekhez hasonlatosan a $\text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{U} \otimes \mathcal{V})$ is ciklikus p -edrendű csoport. A

$$\Lambda : \text{Hom}(\mathcal{A}, \mathcal{U}) \otimes \text{Hom}(\mathcal{B}, \mathcal{V}) \rightarrow \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{U} \otimes \mathcal{V})$$

homomorfizmus tehát egy ciklikus p -edrendű csoportot képez egy másik ciklikus p -edrendű csoportba. Így egyértelműen meghatározott az $\alpha \otimes \beta$ generátorelem képével. Erre a definíció alapján

$$(\alpha \otimes \beta)(\mathbf{a} \otimes \mathbf{b}) = \alpha(\mathbf{a}) \otimes \beta(\mathbf{b}) = (p\mathbf{u}) \otimes (p\mathbf{v}) = (p^2\mathbf{u}) \otimes \mathbf{v} = \mathbf{0} \otimes \mathbf{v} = \mathbf{0}$$

adódik, tehát $\Lambda(\alpha \otimes \beta) = \omega$, vagyis Λ sem szűrjektív, sem injektív. \square

Kiegészítés. Ha \mathcal{A}, \mathcal{B} véges dimenziós vektorterek, akkor Λ izomorfizmus.

Bizonyítás. Legyenek $\{\dots, \mathbf{a}_i, \dots\}$, $\{\dots, \mathbf{b}_p, \dots\}$, $\{\dots, \mathbf{u}_j, \dots\}$, $\{\dots, \mathbf{v}_q, \dots\}$, rendre az $\mathcal{A}, \mathcal{B}, \mathcal{U}, \mathcal{V}$ vektorterek bázisai. Tudjuk, hogy ekkor $\alpha_{i,j} : \mathbf{a}_k \mapsto \delta_{i,k} \cdot \mathbf{u}_j$, illetve $\beta_{p,q} : \mathbf{b}_r \mapsto \delta_{p,r} \cdot \mathbf{v}_q$ bázisok. Ezért bázist alkotnak az $\alpha_{i,j} \otimes \beta_{p,q}$ elemek is.

$\Lambda(\alpha_{i,j} \otimes \beta_{p,q}) : \mathbf{a}_k \otimes \mathbf{b}_r \mapsto \delta_{i,k} \delta_{p,r} \mathbf{u}_j \otimes \mathbf{v}_q$ következtében a fent kapott bázis képe is egy bázis lesz, mert véges dimenziós vektorterek tenzorszorzata is véges dimenziós.

A $\mathcal{B} = R$ speciális esetben egy $\text{Hom}(\mathcal{A}, \mathcal{U}) \otimes \mathcal{V} \rightarrow \text{Hom}(\mathcal{A}, \mathcal{U} \otimes \mathcal{V})$, az $\mathcal{U} = \mathcal{V} = R$ esetben pedig egy $\mathcal{A}^* \otimes \mathcal{B}^* \rightarrow (\mathcal{A} \otimes \mathcal{B})^*$ természetes homomorfizmust nyerünk, amelyek mindegyike izomorfizmust ad véges dimenziós vektorterekre.

Mivel a $\Lambda(\alpha \otimes \beta)$ úgy hat $\mathcal{A} \otimes \mathcal{B}$ -n, mint az $\alpha \otimes \beta$ természetes képe, ezért „azonosíthatjuk vele”. Ha tehát adottak

$$\mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{C} \quad \text{és} \quad \mathcal{U} \xrightarrow{\varphi} \mathcal{V} \xrightarrow{\psi} \mathcal{W},$$

akkor tekinthetjük a következőket:

$$\mathcal{A} \otimes \mathcal{U} \xrightarrow{\alpha \otimes \varphi} \mathcal{B} \otimes \mathcal{V} \xrightarrow{\beta \otimes \psi} \mathcal{C} \otimes \mathcal{W}.$$

A fentiek alapján $\beta \otimes \psi \cdot \alpha \otimes \varphi = \beta\alpha \otimes \psi\varphi$, és $1 \otimes 1 = 1$; ami azt jelenti, hogy a tenzorszorzat mindkét változójában kovariáns funktor. A „Hom” funktornál nem voltak ilyen problémák, mert a $\text{Hom}(\alpha, \varphi)$ alakú „valami”-nek nem volt eleve jelentése; ezért szabadon definiálhattuk.

Tekintsünk most egy $f : \mathcal{A} \rightarrow \text{Hom}(\mathcal{B}, \mathcal{C})$ homomorfizmust. Ez minden \mathcal{A} -beli \mathbf{a} -nak megfeleltet egy $f_{\mathbf{a}}$ homomorfizmust, amelyre $f_{\mathbf{a}}(\mathbf{b}) \in \mathcal{C}$. Mivel $f_{\mathbf{a}}$ homomorfizmus, ezért f rögzített \mathbf{a} -ra \mathcal{B} -ben lineáris. Mivel f homomorfizmus, ezért rögzített \mathbf{b} esetén $\mathbf{a} \mapsto f_{\mathbf{a}}(\mathbf{b})$ homomorfizmus. Így f egy bihomomorfizmusnak tekinthető.

Fordítva, bármely $\mathbf{A} : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{C}$ bihomomorfizmus egyértelműen definiál egy $\mathcal{A} \rightarrow \text{Hom}(\mathcal{B}, \mathcal{C})$ homomorfizmust, az $\mathbf{a} \mapsto (\mathbf{b} \mapsto \mathbf{A}(\mathbf{a}, \mathbf{b}))$ definícióval. Könnyen látható, hogy e két megfeleltetés egymás inverze.

Legyenek most adva az $\mathbf{A}, \mathbf{B} : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{C}$ bihomomorfizmusok, a hozzájuk tartozó $f, g : \mathcal{A} \rightarrow \text{Hom}(\mathcal{B}, \mathcal{C})$ homomorfizmusokkal. Az $\mathbf{A} + \mathbf{B} : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{C}$ bihomomorfizmushoz tartozó $h : \mathcal{A} \rightarrow \text{Hom}(\mathcal{B}, \mathcal{C})$ homomorfizmusra $h_{\mathbf{a}}(\mathbf{b}) = (\mathbf{A} + \mathbf{B})(\mathbf{a}, \mathbf{b}) = \mathbf{A}(\mathbf{a}, \mathbf{b}) + \mathbf{B}(\mathbf{a}, \mathbf{b}) = f_{\mathbf{a}}(\mathbf{b}) + g_{\mathbf{a}}(\mathbf{b})$; ami azt mutatja, hogy a megfeleltetés összegtartó (hasonlóan látható, hogy a skalárszorozást is megtartja). Így a megfeleltetés izomorfizmus.

Rögzített $\mathcal{A}, \mathcal{B}, \mathcal{C}$ esetén a $\mathfrak{T} : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{A} \otimes \mathcal{B}$ tenzorszorozatra tekintsük a $\varphi \rightarrow \mathbf{A} = \varphi \cdot \mathfrak{T}$ -vel definiált megfeleltetést, amely triviálisan homomorfizmus. A tenzorszorozat definíciója alapján szürjektív is és injektív is. Így ez egy izomorfizmus.

A fenti két izomorfizmust komponálva egy $\text{Hom}(\mathcal{A}, \text{Hom}(\mathcal{B}, \mathcal{C})) \rightarrow \text{Hom}(\mathcal{A} \otimes \mathcal{B}, \mathcal{C})$ (természetes) izomorfizmust nyerünk. ■

Az előzőek szembevető ellentétéként most megmutatjuk, hogy a duális térrel való izomorfizmus nem természetes izomorfizmus:

Tekintsük a K test feletti egydimenziós \mathcal{U} vektorteret, s legyen ennek egy báziseleme \mathbf{e} . A $\Phi : \mathcal{U} \rightarrow \mathcal{U}^*$ megfeleltetésnél $\mathbf{u}^* \in \mathcal{U}^*$ definíció szerint az a homomorfizmus, amelyre $\mathbf{u}^*(\mathbf{u}) = 1$. Az, hogy Φ természetes homomorfizmus (izomorfizmus), azt jelentené, hogy minden $\varphi : \mathcal{U} \rightarrow \mathcal{U}$ homomorfizmusához létezik egy olyan $\varphi^* : \mathcal{U}^* \rightarrow \mathcal{U}^*$ homomorfizmus, amelyre a

$$\begin{array}{ccc} \mathcal{U} & \xrightarrow{\varphi} & \mathcal{U} \\ \Phi \downarrow & & \downarrow \Phi \\ \mathcal{U}^* & \xrightarrow{\varphi^*} & \mathcal{U}^* \end{array}$$

diagram kommutatív. Legyen $\varphi(\mathbf{e}) = c\mathbf{e}$. Tekintettel arra, hogy $(c\mathbf{e})^*(c\mathbf{e}) = 1$, ezért ekkor $(\Phi\varphi(\mathbf{e}))(\mathbf{e}) = (c\mathbf{e})^*(\mathbf{e}) = \frac{1}{c}$, vagyis $\Phi\varphi(\mathbf{e}) = \frac{1}{c}\mathbf{e}^*$. Ha a diagram kommutatív volna, akkor $\varphi^*(\mathbf{e}^*) = \varphi^*\Phi(\mathbf{e}) = \frac{1}{c}\mathbf{e}^*$ lenne. Hasonlóan, $\psi(\mathbf{e}) = d\mathbf{e}$ esetében $\psi^*(\mathbf{e}^*) = \frac{1}{d}\mathbf{e}^*$ és $(\varphi + \psi)(\mathbf{e}) = (c + d)\mathbf{e}$ következtében $(\varphi + \psi)^*(\mathbf{e}^*) = \frac{1}{c + d}\mathbf{e}^*$ volna. Ez viszont ellentmond a disztributivitásból adódó $\Phi(\varphi + \psi) = (\varphi + \psi)^*\Phi$ összefüggésnek.

Az sem segítene, ha φ^* irányát ellenkezőre változtatnánk, mert a $\varphi = \omega$ választással a Φ vagy értelmetlenné válna, vagy az ω -homomorfizmusra képezne, amit viszont egyetlen homomorfizmus sem vihet át \mathbf{e}^* -ba.

Feladatok

1. Az előző rész feladatai között láttuk, hogy egydimenziós vektortérre \mathcal{U} izomorf $\mathcal{U} \otimes \mathcal{U}$ -val. Bizonyítsuk be, hogy nem létezik köztük természetes izomorfizmus.

2. Bizonyítsuk be, hogy rögzített $\mathbf{v} \in \mathcal{U}$ esetén $\mathbf{u} \mapsto \mathbf{u} \otimes \mathbf{v}$ egy $\mathcal{U} \rightarrow \mathcal{U} \otimes \mathcal{U}$ injektív homomorfizmus. Mutassuk meg, hogy nem természetes homomorfizmus.

3. Határozzuk meg a $\text{Hom}(\mathbb{Z}_n, \mathbb{Z}_k)$ csoportokat, ha $(n, k) = 1$.
4. Határozzuk meg a $\text{Hom}(\mathbb{Z}_n, \mathbb{Z}_k)$ csoportokat, ha $k|n$.
5. Határozzuk meg a $\text{Hom}(\mathbb{Z}_n, \mathbb{Z}_k)$ csoportokat, tetszőleges k és n esetén.
6. Határozzuk meg a $\text{Hom}(\mathbb{Z}_n, \mathbb{Z})$ csoportokat.
7. Határozzuk meg a $\text{Hom}(\mathbb{Z}, \mathbb{Z}_n)$ csoportokat.
8. Határozzuk meg a $\text{Hom}(\mathbb{Z}_n, \mathbb{Q})$ és $\text{Hom}(\mathbb{Q}, \mathbb{Z}_n)$ csoportokat.
9. Határozzuk meg a $\text{Hom}(\mathbb{Z}, \mathbb{Q})$ és $\text{Hom}(\mathbb{Q}, \mathbb{Z})$ csoportokat.
10. Legyen \mathbb{T} a komplex egységgyökök csoportja. Bizonyítsuk be, hogy $\mathbb{T} \cong \mathbb{Q}/\mathbb{Z}$.
11. Határozzuk meg a $\text{Hom}(\mathbb{Z}_n, \mathbb{T})$ csoportokat.
12. Legyen $\mathbb{G} = \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_k}$. Határozzuk meg $\text{Hom}(\mathbb{G}, \mathbb{T})$ -t.
13. Határozzuk meg: $\text{Hom}(\mathbb{Z}, \mathbb{T})$, $\text{Hom}(\mathbb{Q}, \mathbb{T})$, $\text{Hom}(\mathbb{T}, \mathbb{Z})$, $\text{Hom}(\mathbb{T}, \mathbb{Q})$.
14. Jelölje \mathbb{T}_{p^∞} a \mathbb{T} azon elemeinek halmazát, amelyek rendje a p prímszám egy hatványa. Bizonyítsuk be, hogy \mathbb{T}_{p^∞} részcsoport, és \mathbb{T} az összes ilyenek direkt összege.
15. A 11., 12. és 13. feladatok \mathbb{T} helyett a \mathbb{T}_{p^∞} csoporttal.
16. Határozzuk meg — sorrendben — a fenti \mathbb{H} csoportokra az $\text{End}(\mathbb{H})$ gyűrűket.
17. Határozzuk meg — sorrendben — a fenti \mathbb{H}_1 és \mathbb{H}_2 csoportokra a $\mathbb{H}_1 \otimes \mathbb{H}_2$ csoportokat.
18. Legyenek $\varphi : \mathcal{A} \rightarrow \mathcal{A}'$ és $\psi : \mathcal{U} \rightarrow \mathcal{U}'$ Abel-csoport- (modulus-)homomorfizmusok. Tekintsük a $\text{Hom}(\varphi, 1) : \text{Hom}(\mathcal{A}', \mathcal{U}) \rightarrow \text{Hom}(\mathcal{A}, \mathcal{U})$ és a $\text{Hom}(1, \psi) : \text{Hom}(\mathcal{A}, \mathcal{U}) \rightarrow \text{Hom}(\mathcal{A}, \mathcal{U}')$ indukált homomorfizmusokat. Milyen „jektívek” ezek, ha φ , illetve ψ injektív, illetve szürjektív?
Mit mondhatunk a $\varphi \otimes 1 : \mathcal{A} \otimes \mathcal{U} \rightarrow \mathcal{A}' \otimes \mathcal{U}$ homomorfizmusról?
19. Legyenek $\mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{C}$ Abel-csoport-homomorfizmusok és $1 = 1_{\mathcal{U}}$ az \mathcal{U} Abel-csoport identitása. Tegyük fel, hogy $\text{Im}(\alpha) = \text{Ker}(\beta)$. Igaz-e, hogy $\text{Im}(\text{Hom}(1, \alpha)) = \text{Ker}(\text{Hom}(1, \beta))$, $\text{Im}(\text{Hom}(\beta, 1)) = \text{Ker}(\text{Hom}(\alpha, 1))$, $\text{Im}(1 \otimes \alpha) = \text{Ker}(1 \otimes \beta)$?
20. Határozzuk meg az összes olyan \mathcal{A} Abel-csoportot, amelyre megfelelően fennállnak: $\text{Hom}(\mathcal{A}, \mathbb{T}) = 0$, $\text{Hom}(\mathbb{T}, \mathcal{A}) = 0$, $\text{Hom}(\mathcal{A}, \mathbb{T}_{p^\infty}) = 0$, $\text{Hom}(\mathcal{A}, \mathbb{Q}) = 0$, $\text{Hom}(\mathbb{Q}, \mathcal{A}) = 0$, $\text{Hom}(\mathbb{T}_{p^\infty}, \mathcal{A}) = 0$, $\mathcal{A} \otimes \mathbb{T} = 0$, $\mathcal{A} \otimes \mathbb{T}_{p^\infty} = 0$, $\mathcal{A} \otimes \mathbb{Q} = 0$.
21. Van-e olyan \mathcal{A} Abel-csoport, amelyre $\text{Hom}(\mathcal{A}, \mathcal{G}) = 0$ esetén $\mathcal{G} = 0$ adódik? Ha van ilyen, határozzuk meg az összeset.
22. Van-e olyan \mathcal{A} Abel-csoport, amelyre $\text{Hom}(\mathcal{G}, \mathcal{A}) = 0$ esetén $\mathcal{G} = 0$ adódik? Ha van ilyen, határozzuk meg az összeset.
23. Van-e olyan \mathcal{A} Abel-csoport, amelyre $\mathcal{G} \otimes \mathcal{A} = 0$ esetén $\mathcal{G} = 0$ adódik? Ha van ilyen, határozzuk meg az összeset.
24. Legyen \mathbf{U}^* a K feletti \mathcal{U} vektortér \mathbf{U} bázisának a duális bázisa. Bizonyítsuk be, hogy minden $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bihomomorfizmusoz van olyan $\alpha : \mathcal{V} \rightarrow \mathcal{U}$ homomorfizmus, amelyre $\mathbf{A}(\mathbf{u}, \mathbf{v}) = \mathbf{u}^*(\alpha(\mathbf{v}))$.

3. Mátrix-előállítások, tenzor koordinátái

Mindenekelőtt célszerű felidézni a bilineáris formákra és leképezésekre kapott mátrix-előállításokat:

Tekintsünk egy $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{W}$ bihomomorfizmust. Elegendő a $\mathcal{W} = K$ esetet nézni, amikor bilineáris függvényről beszélünk.

Tegyük fel, hogy $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ a \mathcal{V} vektortér egy-egy bázisa. Ha mármost $\mathbf{x} = \sum x_i \mathbf{u}_i$ és $\mathbf{y} = \sum y_j \mathbf{v}_j$ egy-egy tetszőleges vektor a megfelelő vektorterekben, akkor:

$$\mathbf{A}(\mathbf{x}, \mathbf{y}) = \sum_i \sum_j (x_i \cdot y_j) \cdot \mathbf{A}(\mathbf{u}_i, \mathbf{v}_j).$$

Definíció. Ha $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ az \mathcal{U} és $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ a \mathcal{V} vektortér egy-egy bázisa, akkor az $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvény e bázisban felírt mátrixának nevezzük az

$$F = [f_{i,j}] = [\mathbf{A}] = \mathbf{U}[\mathbf{A}]\mathbf{V} = [\mathbf{A}(\mathbf{u}_i, \mathbf{v}_j)]_{n,k}$$

mátrixot. ■

Az $\mathbf{A}(\mathbf{x}, \mathbf{y}) = \sum_i \sum_j (x_i \cdot y_j) \cdot f_{i,j}$ felírást az x_i és y_j határozatlanok polinomjának tekintve *bilineáris alakot* vagy *bilineáris formát* kapunk.

Tétel. Adott bázisok esetén minden $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvényhez létezik pontosan egy olyan $\varphi : \mathcal{V} \rightarrow \mathcal{U}$ homomorfizmus, amelyre $[\mathbf{A}] = [\varphi]$. Erre a homomorfizmusra $\mathbf{A}(\mathbf{x}, \mathbf{y}) = \mathbf{y}^* \varphi \tilde{\mathbf{x}}(1)$. ■

A fenti megfeleltetés természetesen „nem természetes”, azaz más bázis esetén más homomorfizmust kapunk!

Tétel. Rögzített bázisok esetén $\mathbf{A}(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]^\dagger [\mathbf{A}] [\mathbf{y}]$. ■

Az \mathcal{U} és \mathcal{V} terekben egy-egy új bázist, \mathbf{E} -t és \mathbf{F} -et felvéve, az $\mathbf{e}_i = \sigma(\mathbf{u}_i)$ és $\mathbf{f}_j = \tau(\mathbf{v}_j)$ kíséző transzformációkkal:

$$\mathbf{F}[\mathbf{A}]_j^{\mathbf{E}} = [\mathbf{e}_i]^\dagger [\varphi] [\mathbf{f}_j] = [\mathbf{u}_i]^\dagger [\sigma]^\dagger [\varphi] [\tau] [\mathbf{v}_j],$$

amiből adódik az alábbi

Tétel. A fenti módon új bázisokra való áttérésnél az \mathbf{A} bilineáris függvény mátrixa a következőképpen változik: $\mathbf{F}[\mathbf{A}]^{\mathbf{E}} = [\sigma]^\dagger \mathbf{V}[\mathbf{A}]^{\mathbf{U}} [\tau]$. ■

A homomorfizmusok és a bilineáris leképezések mátrixai között igen lényeges elvi különbség van.

Legyen adva az \mathcal{U} vektortérben egy $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ és a \mathcal{V} vektortérben egy $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ bázis. Tekintsünk most egy $\varphi : \mathcal{U} \rightarrow \mathcal{V}$ homomorfizmust és egy $\mathbf{A} : \mathcal{U} \times \mathcal{V} \rightarrow K$ bilineáris függvényt, amelyek mindegyikének $[a_{i,j}]$ a mátrixa.

Az, hogy a homomorfizmusnak ez a mátrixa, azt jelenti, hogy

$$\varphi = \sum_i \sum_j a_{i,j} \tilde{\mathbf{v}}_j \mathbf{u}_i^*$$

alakba írható fel. Az itt álló összeget úgy képzelhetjük el, hogy ez a $\mathcal{V}^* \otimes \mathcal{U}$ egy homomorfizmusánál jött létre; mégpedig úgy, hogy a $\tilde{\mathbf{v}}_j \otimes \mathbf{u}_i$ báziselemeknek a $\tilde{\mathbf{v}}_j \mathbf{u}_i^*$ homomorfizmusokat feleltetjük meg. Eszerint:

$$\varphi \text{ a } \sum_i \sum_j |a_{i,j}| \tilde{\mathbf{v}}_j \otimes \mathbf{u}_i \text{ képe.}$$

A lényegében a tenzorszorzatból származtatható. Nevezetesen ez „ekvivalens” azzal a leképezéssel, amelyik a $\mathbf{v}_j \otimes \mathbf{u}_i$ elemeket $a_{i,j}$ -be képezi. Ha tehát meg akarjuk kapni adott $\mathbf{x} = \sum_i x_i \mathbf{u}_i$ és $\mathbf{y} = \sum_j y_j \mathbf{v}_j$ esetén $\mathbf{A}(\mathbf{x}, \mathbf{y})$ -t, akkor ez úgy áll elő, mint

$$\sum_i \sum_j (x_i \cdot y_j) \mathbf{v}_j \otimes \mathbf{u}_i \quad \text{képe, ha} \quad \mathbf{v}_j \otimes \mathbf{u}_i \mapsto a_{i,j}.$$

Mindkét esetben tehát a tenzorszorzat egy elemének képéről van szó. (Az pillanatnyilag nem lényeges, hogy a két tenzorszorzat nem ugyanaz.) De amíg a homomorfizmusnál a mátrix elemei a fellépő tenzor együtthatói, addig a bilineáris függvénynél az együtthatók a megfelelő $\square \otimes \square$ -k képei. Ez az oka annak, hogy új bázisra való áttérésnél a bilineáris függvény mátrixa „mindkét változóban” kovariánsan transzformálódik; noha a tenzorszorzat egyik eleme sem a duális térből való. Az „igazi” tehát az, amit a homomorfizmusnál kapunk; és ez a fellépő együtthatókból készített mátrix. Ezt tükrözi az alábbi

10.5. Definíció. Tekintsük a

$$\mathcal{T} = \mathcal{U}_1^* \otimes \cdots \otimes \mathcal{U}_r^* \otimes \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_s$$

tenzorszorzatnak azt a bázisát, amelynek elemei:

$$\mathbf{u}_{1_i}^* \otimes \cdots \otimes \mathbf{u}_{r_i}^* \otimes \mathbf{v}_{1_j} \otimes \cdots \otimes \mathbf{v}_{s_j};$$

ahol az egyes komponensek a megfelelő vektorterek adott bázisvektorain futnak végig. A tenzorszorzat egy elemét e bázisban felírva a kapott együtthatókat a tenzornak az adott bázisokban felírt koordinátáinak nevezzük. ■

Nem okoz elvi nehézséget az alábbi tétel bizonyítása, de a hosszadalmas számolás (és a nem is egész pontos definíció) miatt elhagyjuk:

10.14. Tétel. *Ha az egyes vektorterekben új bázisokat vezetünk be, akkor a tenzorszorzatban a koordináták a kovariáns komponensekben kovariánsul; a kontravariánsokban kontravariánsul transzformálódnak.*

A „...variáns transzformálódás” nem volt definiálva, mert itt az együtthatók nem egy mátrixot alkotnak, és ezt kellene „szorozni” egy mátrixszal. Elégedjünk meg azzal, amit a homomorfizmusoknál láttunk.

4. A tenzoralgebra, szimmetrikus és antiszimmetrikus tenzorok

A következőkben egyetlen vektortéren értelmezett kontravariáns tenzorokat fogunk vizsgálni.

10.6. Definíció. Adott \mathcal{U} vektortér esetén az

$$\mathcal{U}^{(k)} = \overbrace{\mathcal{U} \otimes \cdots \otimes \mathcal{U}}^{k \text{ tényező}} \quad (k > 1)$$

elemeit k -adfokú tenzoroknak nevezzük. Az elsőfokú tenzorok az $\mathcal{U}^{(1)} = \mathcal{U}$ elemei; a nulldfokúak pedig $\mathcal{U}^{(0)} = K$ elemei. ■

Az eddigiekben láttuk, hogyan lehet tenzorokat skalárral szorozni és egyenlő fokú tenzorokat összeadni. A továbbiakban az lesz a célunk, hogy az összeadást különböző fokú tenzorokra is kiterjesszük; továbbá tenzorok szorzatát is definiáljuk. Ehhez szükségünk lesz az alábbira:

Lemma. Legyen $V = \{\dots, \mathbf{v}_i, \dots \mid i \in I\}$ a \mathcal{V} vektortér egy bázisa. A bázisvektorokra értelmezett tetszőleges asszociatív szorzás egyértelműen kiterjeszthető \mathcal{V} -nek egy asszociatív és disztributív szorzásává, amellyel \mathcal{V} algebrává válik.

Bizonyítás. Ha egy disztributív kiterjesztést akarunk, akkor ez csak

$$\left(\sum_i a_i \mathbf{v}_i \right) \left(\sum_j b_j \mathbf{v}_j \right) = \sum_i \sum_j (a_i b_j) \mathbf{v}_i \mathbf{v}_j$$

lehet. Így a kiterjesztés egyértelmű. Az azonosságok teljesülése a következőképpen látható be:

Legyen $\mathbf{a} = \sum_i a_i \mathbf{v}_i$, $\mathbf{b} = \sum_j b_j \mathbf{v}_j$ és $\mathbf{c} = \sum_k c_k \mathbf{v}_k$. Az

$$(\mathbf{a}\mathbf{b})\mathbf{c} = \sum_i \sum_j \sum_k (a_i b_j c_k) (\mathbf{v}_i \mathbf{v}_j) \mathbf{v}_k \quad \text{és} \quad \mathbf{a}(\mathbf{b}\mathbf{c}) = \sum_i \sum_j \sum_k (a_i b_j c_k) \mathbf{v}_i (\mathbf{v}_j \mathbf{v}_k)$$

összefüggés bizonyítja az asszociativitást. Az $(\mathbf{a} + \mathbf{b})\mathbf{c} = \mathbf{a}\mathbf{c} + \mathbf{b}\mathbf{c}$ disztributivitás belátásához tegyük fel, hogy az \mathbf{a} és \mathbf{b} vektorokban ugyanazok az indexek szerepelnek (a „hiányzó” bázisvektorok együtthatóit 0-nak vesszük). Ekkor azt kell tehát belátnunk, hogy

$$[(a_i + b_i)c_k] \cdot \mathbf{v}_i \mathbf{v}_k \quad \text{és} \quad [a_i c_k] \cdot \mathbf{v}_i \mathbf{v}_k + [b_i c_k] \cdot \mathbf{v}_i \mathbf{v}_k$$

megegyeznek, ami igaz. A másik oldali disztributivitás hasonlóképpen látható be.

A $(\mathbf{c}\mathbf{u})\mathbf{v} = \mathbf{u}(\mathbf{c}\mathbf{v}) = \mathbf{c}(\mathbf{u}\mathbf{v})$ bizonyítása triviális. ■

10.15. Tétel. Adott \mathcal{U} vektortér esetén tekintsük a

$$\mathcal{T}(=\mathcal{T}(\mathcal{U}))=\mathcal{U}^{(0)}+\mathcal{U}^{(1)}+\mathcal{U}^{(2)}+\dots+\mathcal{U}^{(n)}+\dots$$

direkt összeget. Defináljuk \mathcal{T} elemeinek a szorzatát mint az

$$(\mathbf{u}_1 \otimes \dots \otimes \mathbf{u}_r) \cdot (\mathbf{u}_{r+1} \otimes \dots \otimes \mathbf{u}_k) = \mathbf{u}_1 \otimes \dots \otimes \mathbf{u}_k$$

szorzás kiterjesztését. Ezáltal \mathcal{T} egy algebrává válik, amelyet az \mathcal{U} feletti tenzoralgebrának nevezünk.

Bizonyítás. Legyen $\mathbf{E} = \{\dots, \mathbf{e}_i, \dots\}$ az \mathcal{U} egy rögzített bázisa. Ekkor \mathcal{T} -nek egy bázisa lesz az összes $\mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r}$ tenzorok halmaza, ahol a tényezők száma 0 is lehet (ekkor az 1 a báziselem), és az indexek között azonosak is lehetnek.

A fent definiált szorzás egyértelműen meghatározza ezeknek a báziselemeknek a szorzatát. (Azért nem a báziselekkel definiáltuk a szorzást, hogy az ne függjön a bázistól.) Mivel ez a szorzás triviálisan asszociatív, ezért valóban egy algebrát nyertünk. ■

Noha a tenzori szorzat kommutatív, $\mathbf{u} \otimes \mathbf{v}$ és $\mathbf{v} \otimes \mathbf{u}$ mégis különbözőek. A kommutativitás ugyanis a két vektortér (természetes) izomorfiját jelenti; nem pedig azt, hogy az izomorfizmusnál egymásnak megfeleltetett elemek megegyeznek. Ennek ellenére elképzelhető, hogy egy tenzor szimmetrikus; például $\mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u}$ ilyen. Világos, hogy azonos fokú szimmetrikus tenzorok összege is ugyanilyen fokú szimmetrikus tenzor; különböző fokúaké viszont már nem az. Szimmetrikus tenzorok szorzata sem szimmetrikus. Éppen ezért a szimmetrikus tenzorokat másképpen célszerű definiálni. Azt mondjuk meg, hogy két tenzor (például $\mathbf{u} \otimes \mathbf{v}$ és $\mathbf{v} \otimes \mathbf{u}$) mikor számítnak azonosnak. Világos, hogy akkor, ha egyikből úgy állítható elő a másik, hogy bármelyik tagban felcserélünk egy-egy komponenst. Tekintettel arra, hogy ez mindig visszavezethető egymás melletti komponensek cseréjére, ezért célszerű az alábbi

10.7. Definíció. A $\mathcal{T}(\mathcal{U})$ -beli $\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u}$ elemeket antiszimmetriáknak nevezzük. Legyen \mathcal{A} az antiszimmetriák generálta (kétoldali!) ideál. Az $\mathcal{S}(\mathcal{U}) = \mathcal{T}(\mathcal{U})/\mathcal{A}$ faktoralgebra elemeit szimmetrikus tenzoroknak nevezzük. ■

A szimmetrikus tenzorok tehát mellékosztályok. Tekintettel arra, hogy a szimmetrikus tenzorokat nem sokat vizsgáljuk, ezért számukra nem vezetünk be külön jelölést; sőt még azt sem írjuk oda, hogy mi szerinti (\mathcal{A} -szerinti) mellékosztályok szerepelnek.

10.16. Tétel. Legyen $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ az \mathcal{U} egy bázisa. Ekkor a szimmetrikus tenzorok algebrájának egy bázisa az

$$\{\mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \mid i_1 \leq \dots \leq i_r\}$$

halmaz.

Bizonyítás. Mivel egy „egytagú” tenzorban bármely két egymás utáni komponenst felcserélve a két tenzor különbsége mindig egy antiszimmetria és bármely permutáció előáll egymás melletti elemek cseréjeként, ezért a fenti elemek egy generátorrendszert alkotnak.

Az, hogy lineárisan függetlenek, hasonlóan bizonyítható, mint ahogy az antiszimmetrikus esetben fogjuk tenni, csak sokkal bonyolultabb. Ezért ezt elhagyjuk. ■

Az antiszimmetrikus tenzorokat hasonló elvek szerint definiáljuk, mint a szimmetrikusokat.

10.8. Definíció. Legyen \mathcal{I} az \mathcal{U} vektortér tenzoralgebrája. Legyen \mathcal{J} az $\mathbf{u} \otimes \mathbf{u}$ alakú elemek generálta ideál. A $\mathcal{G} = \mathcal{G}(\mathcal{U}) = \mathcal{I}/\mathcal{J}$ algebrát az \mathcal{U} feletti Grassmann-algebrának nevezzük; és elemeit antiszimmetrikus tenzoroknak.

Az $\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_r$ elemnek a Grassmann-algebrabeli képét $\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_r$ jelöli. A Grassmann-algebrabeli szorzás neve külső szorzás.

Az $\mathcal{U}^{(i)}$ -nek \mathcal{G} -beli képét $\mathcal{G}^{(i)}$ -vel jelöljük. ■

10.17. Tétel. Legyen $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ az \mathcal{U} egy bázisa. Ekkor a következők teljesülnek:

- (1) Bármely \mathbf{u}, \mathbf{v} vektorra $\mathbf{v} \wedge \mathbf{u} = -\mathbf{u} \wedge \mathbf{v}$.
- (2) Bármely

$$\mathbf{A} : \overbrace{\mathcal{U} \times \cdots \times \mathcal{U}}^{r \text{ tényező}} \rightarrow \mathcal{V}$$

r -lineáris antiszimmetrikus leképezéshez egyértelműen létezik olyan $\varphi : \mathcal{G}^{(r)} \rightarrow \mathcal{V}$ homomorfizmus, amelyre

$$\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_r) = \varphi(\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_r).$$

- (3) A Grassmann-algebra egy bázisát alkotja az

$$\{\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_r} \mid i_1 < \cdots < i_r\}$$

almaz.

- (4) $\dim(\mathcal{G}) = 2^n$, $\dim(\mathcal{G}^{(k)}) = \binom{n}{k}$.

Bizonyítás. Az (1) állítás abból következik, hogy:

$$\mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u} = (\mathbf{u} + \mathbf{v}) \otimes (\mathbf{u} + \mathbf{v}) - \mathbf{u} \otimes \mathbf{u} - \mathbf{v} \otimes \mathbf{v} \in \mathcal{J}.$$

(2)-t a következőképpen bizonyítjuk:

Egy adott $\mathbf{A} : \mathcal{U}^r \rightarrow \mathcal{V}$ r -lineáris antiszimmetrikus függvényhez a tenzorszorzat tulajdonságai szerint létezik olyan egyértelmű $\psi : \mathcal{U}^{(r)} \rightarrow \mathcal{V}$ homomorfizmus, amelyre $\mathbf{A} = \psi \mathfrak{T}_r$; ahol $\mathfrak{T}_r : \mathcal{U}^r \rightarrow \mathcal{U}^{(r)}$ a tenzorszorzatot jellemző r -homomorfizmus. Az antiszimmetria miatt $\text{Ker}(\psi)$ tartalmazza $\mathcal{U}^{(r)} \cap \mathcal{J}$ elemeit. A második izomorfizmustétel szerint tehát létezik a megfelelő faktorizáció.

(3) bizonyításához először a következőket nézzük. \mathcal{I} -nek az \mathbf{E} által meghatározott bázisából azoknak az elemeknek az összege vagy a különbsége, amelyekben ugyanazok a

bázisvektorok (multiplicitással) szerepelnek, (1) miatt \mathcal{F} -be esik. Ezért elegendő olyan szorzatokat nézni, ahol az indexek nem csökkennek. A definíció szerint azok a tagok, amelyekben két egymás utáni index megegyezik, ugyancsak \mathcal{F} -be esnek; ezért a felsorolt elemek generátorrendszert alkotnak.

A lineáris függetlenséghez először azt látjuk be, hogy a $\mathcal{U}^{(n)}$ -beli $\mathbf{e}_1 \otimes \cdots \otimes \mathbf{e}_n \notin \mathcal{F}$. Ehhez (2) miatt elég azt megmutatni, hogy van olyan n -lineáris antiszimmetrikus függvény, amely nem azonosan 0 — illetve a fenti elemet például 1-be viszi. Ezt a tér dimenziójára vonatkozó teljes indukcióval bizonyítjuk. $n = 1$ esetén az állítás triviálisan igaz.

Tegyük fel, hogy az állítás igaz bármely $(n - 1)$ -dimenziós térre, és tekintsük az n -dimenziós tér fenti bázisát.

Definiálnunk kell egy $\mathfrak{d}_n(\mathbf{u}_1, \dots, \mathbf{u}_n)$ antiszimmetrikus n -lineáris függvényt. Ilyen függvény viszont létezik, hiszen bármely $\mu(\mathbf{u}_1, \dots, \mathbf{u}_n)$ mérték rendelkezik ezzel a tulajdonsággal.

Legyen most $\mathfrak{s} \in \mathcal{F}$ olyan, hogy a feltételezett bázisvektoroknak egy lineáris kombi-nációja. Azt fogjuk bizonyítani, hogy minden egyes együttható 0. Legyen evégett $c \cdot \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r}$ ($i_1 < \dots < i_r$) az \mathfrak{s} -nek egy olyan tagja, hogy minden más tagban legalább r tényező lép fel. Legyenek a kimaradt indexek $j_1 < \dots < j_s$ ($r + s = n$), és szorozzuk meg \mathfrak{s} -t $\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_s}$ -sel. Mivel minden más tagban e komponensek közül legalább az egyik fellép, ezért ez a szorzat $c \cdot \mathbf{e}_1 \otimes \cdots \otimes \mathbf{e}_n$ (vagy a negatívja). Az ideáltulajdonság szerint ez is \mathcal{F} -ben van, ami — mint előbb beláttuk — csak úgy lehet, ha $c = 0$. Ebből triviálisan következik, hogy $\mathfrak{s} = 0$. Így a felsorolt elemek valóban lineárisan függetlenek.

(4) triviálisan következik abból, hogy a felsorolt báziselemek annyian vannak, ahány részhalmaza egy n -elemű halmaznak van; illetve ahány k -elemű részhalmaza egy n -elemű halmaznak van. ■

Feladatok

1. Bizonyítsuk be, hogy a vektoriális szorzás a háromdimenziós valós téren értelmezett antiszimmetrikus bihomomorfizmus.

2. Mutassuk meg, hogy az n -dimenziós \mathcal{U} vektortér minden $f : \mathcal{U}^{n-1} \rightarrow \mathcal{U}$ antiszimmetrikus multihomomorfizmusához van olyan $\varphi : \bigwedge_{n-1} \mathcal{U} \rightarrow \mathcal{U}$ homomorfizmus, hogy $f(\mathbf{u}_1, \dots, \mathbf{u}_{n-1}) = \varphi(\mathbf{u}_1 \wedge \dots \wedge \mathbf{u}_{n-1})$.

5. Alkalmazások

10.9. Definíció. Legyen A egy mátrix, és hagyjuk el bizonyos sorait és oszlopait. Így egy részmatrixát kapjuk. Ha $I = \{i_1, \dots, i_r\}$ az elhagyott sorok indexei (növe sorrendben) és $J = \{j_1, \dots, j_s\}$ az elhagyott oszlopok indexei (növe sorrendben), akkor a kapott mátrixot az (I, J) -hez tartozó $A_{I,J}$ részmatrixnak nevezzük.

Ha a kapott mátrix négyzetes, akkor az (I, J) -hez tartozó részmatrix $D(A_{I,J})$ determinánsának a neve az eredeti mátrix (I, J) -aldeterminánsa.

Ha az eredeti mátrix is négyzetes, akkor legyen $S(I, J) = \sum_i (i_i + j_i)$ és $\text{sg}(I, J) = (-1)^{S(I,J)}$ az (I, J) párhoz és egyszersmind $A_{I,J}$ -hez, valamint determinánsához tartozó előjel. Az (I, J) -hez tartozó algebrai aldetermináns az $\text{sg}(I, J) \cdot D(A_{I,J})$ skalárt értjük. ■

10.18. Tétel (Laplace-kifejtés). Rögzítsük az $A = [a_{i,j}]_n$ négyzetes mátrix I -beli indexű sorait ($I = \{i_1 < \dots < i_r\}$) és J fusson végig az $N = \{1, \dots, n\}$ halmaz r -elemű részhalmazain. Jelölje továbbá I' és J' — megfelelően — az I és J N -beli komplemente-rét.

Szorozzunk meg minden $D(A_{I',J'})$ aldeterminánst a megfelelő (I, J) párhoz tartozó algebrai aldeterminánssal. Ezeknek a szorzatoknak az összege $D(A)$.

Bizonyítás. Mind a $D(A_{I',J'})$, mind a $D(A_{I,J})$ determinánsok tekinthetők az eredeti mátrix $\mathbf{u}_1, \dots, \mathbf{u}_n$ oszlopvektorai függvényének.

Ezt egy kicsit részletesebben megnézzük:

Legyen $\mathbf{e}_1, \dots, \mathbf{e}_n$ az oszlopmatrixok vektorterének az a bázisa, amelyben az \mathbf{e}_i mátrixnál az i -edik helyen 1 áll, s a többi helyen 0. Legyen $\mathbf{u} = \sum_i c_i \mathbf{e}_i$, és legyen $\pi(\mathbf{u}) = \sum_{i \in I} c_i \mathbf{e}_i$. Hasonlóan definiálható π' ; ahol I helyett I' áll. Ekkor $D(A_{I',J'})$ éppen a $\pi(\mathbf{u}_{j_1}), \dots, \pi(\mathbf{u}_{j_r})$ vektorokból mint oszlopmatrixból álló mátrixnak a determinánsa lesz. Hasonló teljesül $D(A_{I,J})$ -re is.

Készítsük el a

$$D_1 = D_1(\mathbf{u}_1, \dots, \mathbf{u}_n) = \sum_J \text{sg}(I, J) \cdot D(A_{I',J'}(\mathbf{u}_1, \dots, \mathbf{u}_n)) \cdot D(A_{I,J}(\mathbf{u}_1, \dots, \mathbf{u}_n))$$

függvényt. Vegyük észre, hogy az összeg bármely tagját tekintjük, minden egyes \mathbf{u}_i vektor a fellépő két determináns tényező közül pontosan az egyikben szerepel, mert J' és J egymás komplementerei. Ebből azonnal látható, hogy D_1 a szereplő vektoroknak multilineáris függvénye. Azt akarjuk belátni, hogy antiszimmetrikus; azaz ha a vektorok között azonosak vannak, akkor a függvényérték 0.

Tegyük fel, hogy $\mathbf{u}_j = \mathbf{u}_{j'}$. Azok a tagok, amelyekre $j, j' \in J$ vagy $j, j' \in J'$ teljesül, biztos egyenlők 0-val; mert a megfelelő tényezők valamelyike 0. Legyen tehát $j \in J$ és

$j' \in J'$. Minden egyes ilyen taghoz egyértelműen párosítható egy másik úgy, hogy itt az oszlopindexek K halmaza J -ből a j elhagyásával és j' hozzávételével keletkezik. (Ez már egyértelműen meghatározza K' -t is.) E két tag legfeljebb előjelben térhet el egymástól (éppen azt akarjuk bizonyítani, hogy különböző előjelűek, tehát összegük 0). Az eltérés egyik oka, hogy $\text{sg}(I, J)$ és $\text{sg}(I, K)$ nem biztosan egyenlők — eltérésük $(-1)^{j'-j}$. Emellett maga a két determináns is eltérhet előjelben; hiszen például az első tényezőben a j -edik oszlop azon a helyen áll, ahol nagyságrendben szerepel a j a J elemei között, míg a másik esetben a vele megegyező oszlop ott áll, ahol j' -nek kell nagyságrend szerint állni. Ugyanez a helyzet a másik determináns tényezőnél is. Az együttes előjelváltást tehát az adja meg, hogy mennyi cserét kell végeznünk, amíg a j' -edik oszlopot a j -edik helyre visszük. E cserék száma $j' - j - 1$; tehát a két tag valóban egymás negatívja.

Ezzel beláttuk, hogy D_1 a D -nek skalárszorosa. Mivel pedig az $\mathbf{e}_1, \dots, \mathbf{e}_n$ bázison nyilván egyenlők, ezért $D_1 = D$. ■

Tekintsük az n -dimenziós \mathcal{U} vektortér r -edik külső szorzatát. Ennek elemei $\mathbf{u}_1 \wedge \dots \wedge \mathbf{u}_r$ alakú tenzorok lineáris kombinációi. E tér egy bázisát alkotják azok az $\mathbf{e}_P = \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_r}$ alakú elemek, ahol $P = \{i_1 < \dots < i_r\}$ végigfut az $1, \dots, n$ halmaz összes rendezett r -eseinek a halmazán; és $\mathbf{e}_1, \dots, \mathbf{e}_n$ az \mathcal{U} egy bázisa.

Azt akarjuk megállapítani, hogy mik a fenti $\mathbf{u}_R = \mathbf{u}_1 \wedge \dots \wedge \mathbf{u}_r$ elem koordinátái ebben a bázisban. Tekintsük evégett az adott bázisnak az \mathbf{e}_P^* elemekből álló duális bázisát, majd rögzítsük P -t. Az \mathbf{u}_R elem „ P -edik” koordinátája $\mathbf{e}_P^*(\mathbf{u}_R)$. Ezt úgy tekinthetjük, mintha az $\mathbf{u}_1, \dots, \mathbf{u}_r$ elemekre először egy antiszimmetrikus r -homomorfizmust, majd egy homomorfizmust alkalmaztunk volna, amelynek az eredménye ugyancsak egy antiszimmetrikus r -homomorfizmus.

Legyen most a $\varphi_P : \mathcal{U} \rightarrow \mathcal{U}$ transzformáció úgy definiálva, hogy $\varphi_P(\mathbf{e}_i) = \mathbf{e}_i$, ha $i \in P$, és legyen $\varphi_P(\mathbf{e}_i) = 0$ egyébként. (Ezt nem csak r -elemű, hanem akárhány elemű P részhalmazra definiálhatjuk.) Tetszőleges $\mathbf{u} \in \mathcal{U}$ vektor egyértelműen felírható $\mathbf{u} = \varphi_P(\mathbf{u}) + \varphi_Q(\mathbf{u})$ alakba, ahol Q a P -nek komplementere az $\{1, \dots, n\}$ halmazban. Írjunk most fel minden az \mathbf{u}_R -ben szereplő \mathbf{u}_i vektort a fenti módon összegként, és alkalmazzuk erre az \mathbf{e}_P^* -ot. Ez az

$$\mathbf{e}_P^*(\varphi_{X_1}(\mathbf{u}_1) \wedge \dots \wedge \varphi_{X_r}(\mathbf{u}_r)) \quad (X_1, \dots, X_r \in \{P, Q\})$$

alakú tagok összege lesz. Egy ilyen tag biztosan 0, ha valamelyik $X_i = Q$, ezért amit kapunk, nem más, mint $\mathbf{e}_P^*(\varphi_P(\mathbf{u}_1) \wedge \dots \wedge \varphi_P(\mathbf{u}_r))$.

Ezt a „függvényt” most úgy tekinthetjük, mint ami az $\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_r}$ vektorok generálta altéren van értelmezve, ahol $\{i_1, \dots, i_r\} = P$. Ez a függvény nyilván r -lineáris és antiszimmetrikus; továbbá az $\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_r}$ bázison az 1 értéket veszi fel. Ez tehát pontosan a $[\varphi_P(\mathbf{u}_1) \dots \varphi_P(\mathbf{u}_r)]$ mátrix determinánása. E determinánst $\vartheta(\varphi_P(\mathbf{u}_R))$ -rel jelölve azt kapjuk, hogy:

$$\mathbf{u}_R = \sum_P \vartheta(\varphi_P(\mathbf{u}_R)) \cdot \mathbf{e}_P, \quad \text{ahol } P \subseteq \{1, \dots, n\}; \quad |P| = r.$$

10.19. Tétel (Cauchy–Binet-formula). Legyen A az U és V mátrixok szorzata. I és J az A mátrix sorainak és oszlopainak indexhalmaza, ahol $|I| = |J| = r$. Fusson végig P az U és V mátrix oszlopai, illetve sorai indexhalmazának r -elemű részhalmazain. Ekkor

$$\det(A_{I,J}) = \sum_P \det(U_{I,P}) \cdot \det(V_{P,J}).$$

Bizonyítás. Mivel az A mátrix i -edik sorának a j -edik eleme az U mátrix i -edik sorának és a V mátrix j -edik oszlopának a szorzata, ezért feltehető, hogy az A mátrixnak r sora és r oszlopa van; s a kérdéses determináns az A mátrix determinánsa.

Legyen n az U oszlopainak, illetve a V sorainak a száma. Ekkor ezeket

$$U = [\mathbf{u}_1, \dots, \mathbf{u}_r]^\dagger \quad \text{és} \quad V = [\mathbf{v}_1, \dots, \mathbf{v}_r]$$

alakú mátrixoknak tekinthetjük, ahol $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{v}_1, \dots, \mathbf{v}_r$ egy n -dimenziós vektortér elemei. Maga az A mátrix $[U^\dagger(\mathbf{v}_1), \dots, U^\dagger(\mathbf{v}_r)]$ alakban is írható. Ebből azonnal következik, hogy determinánsa a $\mathbf{v}_1, \dots, \mathbf{v}_r$ vektoroknak (rögzített U esetén) r -lineáris antiszimmetrikus függvénye. A dualitás miatt hasonló mondható az U mátrix sorvektorairól. Eszerint az

$$\mathbf{A}([\mathbf{u}_1, \dots, \mathbf{u}_r], [\mathbf{v}_1, \dots, \mathbf{v}_r]) = \det(A)$$

„kétváltozós” függvény keresztülvezethető a külső szorzaton:

$$\varphi(\mathbf{u}_1 \wedge \dots \wedge \mathbf{u}_r, \mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_r) = \det(A).$$

Legyen most \mathbf{e}_i az a bázisvektor, amelyben az i -edik helyen 1 áll, s a többi eleme 0 ($i = 1, \dots, n$). Legyen P és Q az $\{1, \dots, n\}$ két r -elemű részhalmaza. Az $U^\dagger = [\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_r}]$, $V = [\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_r}]$ esetben A az identitásmátrix, ha $P = Q$, és ezért $\det(A) = 1$; míg $P \neq Q$ esetén a kapott A mátrixnak van egy 0 sora, és ezért $\det(A) = 0$. Eszerint $\varphi(\mathbf{e}_P, \mathbf{e}_Q) = \delta_{P,Q}$.

A két változóban való bilinearitás miatt

$$\begin{aligned} \varphi(\mathbf{u}_1 \wedge \dots \wedge \mathbf{u}_r, \mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_r) &= \varphi\left(\sum_{P,Q} \partial(\varphi_P(\mathbf{u}_R)) \cdot \mathbf{e}_P, \partial(\varphi_Q(\mathbf{v}_R)) \cdot \mathbf{e}_Q\right) = \\ &= \sum_{P,Q} \partial(\varphi_P(\mathbf{u}_R)) \cdot \partial(\varphi_Q(\mathbf{v}_R)) \cdot \varphi(\mathbf{e}_P, \mathbf{e}_Q) = \sum_P \partial(\varphi_P(\mathbf{u}_R)) \cdot \partial(\varphi_P(\mathbf{v}_R)); \end{aligned}$$

ahol P és Q egymástól függetlenül végigfutják az $\{1, \dots, n\}$ halmaz r -elemű részhalmazait.

Ez pedig éppen a kívánt összefüggést adja. ■

Betűrendes mutató

- α minimálpolinomja 246
- A-ortogonális 227
- A-ortogonális bázis 234
- A-ortogonalitás 234
- abszolút érték, komplex számé 27
- additív 231
 - inverz 13
- adjungált mátrix 52
 - , transzformációé 253
- affin tér 236
- alap 18
 - , hatványé 18
- alaptétel, szimmetrikus polinomoké 128
- aldetermináns 62, 324
 - hoz tartozó előjel 324
 - , algebrai 324
- algebra alaptétele 102
- algebrai aldetermináns 324
- altér 165
 - , generált 166
 - , invariáns 196
 - , triviális 166
 - , valódi 166
- alterek direkt összege 175
- alternáló polinom 139
- altérre merőleges (ortogonális) altér 242
 - – vektor 242
 - való tükrözés 261
- antiszimmetria 16, 321
 - , oszthatóságé 80
- antiszimmetrikus bihomomorfizmus 217
 - mátrix 69
 - tenzor 322
- argumentum 31
- arkusz 30
- asszociált 79
- asszociativitás (társíthatóság), összegé, szorzaté 13
 - , tenzorszorzaté 311
- azonosságok 150
- bal (jobb) oldali inverz, leképezése 193
 - – nullosztó 193
- bal oldali modulus 149
 - – R -modulus 149
- balinverz, mátrixé 67
- balreguláris leképezés 193
- bázis 160, 161
 - , ortogonális 227
 - , természetes 197
- bihomomorfizmus 216
 - , antiszimmetrikus 217
 - , szimmetrikus 217
- bijektív függvény (leképezés) 20
- bilinéaris forma (alak) 216, 219
- bilinéaris függvény 216
 - – mátrixa 221
 - –, Hermite-féle 234
 - leképezés 216
- binomiális tétel 127
- blokkokra diagonalizált mátrix 249
- Cardano-képlet 100
- casus irreducibilis 101
- csoport 53
 - , kommutatív 148
 - , (összeadásra, szorzásra) nézve 15

- derivált, polinomé 128
- , –, k -adik 131
- determináns 56
- , Gram-féle 243
- rendje 56
- , Vandermonde-féle 139
- determinánsosztó 276
- diád 201
- diagonális mátrix 52
- diagram 211, 308
- , kommutatív 309
- dimenzió 161
- direkt kiegészítő 175
- direkt szorzat 178
- –, halmazoké 19
- diszkrimináns 95, 98, 297
- duális tér 198
- egész számok 15
- egyenletrendszer mátrixa 292
- , homogén lineáris 293
- megoldásai 291
- egyenlő mátrixok 45
- egyértelmű felbontás, polinomoké 85
- egyező alakú mátrixok 45
- egyhatározatlanú polinom 72
- egység, komplex 27
- egységelemes gyűrű 114
- egységgyömb 264
- egységgyökök, komplex 37
- , primitív n -edik 39
- egységmátrix 67
- egytagú polinom 122
- együthatható 72, 76, 291
- , x_k -ban i -edfokú tagé 120
- egyváltozós művelet 148
- ekvivalens polinomrendszer 142
- vektorrendszerek 156
- elem 19
- elemi átalakítás 210
- átalakítómátrixok 273
- linearitások 307
- oszlopátalakítás 273
- osztók 277
- sorátalakítások 273
- szimmetrikus polinom 133
- transzformáció 209
- elforgatás 261
- ellentett 13
- elsőfajú homogén lineáris függvény 231
- elsőfokú vagy lineáris polinom 76
- eltoltakkal való műveletek 172
- értékkészlet (függvényé) 20
- értelmezési tartomány (függvényé) 20
- euklideszi algoritmus 17, 84, 116
- altér 238
- gyűrű 114
- homomorfizmus 237
- norma 114, 148
- tér 236
- faktoriális 22
- faktortér 173
- felbonthatatlan (irreducibilis) 81
- félcsoport 15
- felső (alsó) háromszög mátrix 69
- fok, polinomé 76
- fődiagonális, négyzetes mátrixé 52
- főegyüthatható, polinomé 76
- főideál 82, 247
- főideálgyűrű 247
- főminor 301

- függvény 19, 20
 –, additív, lineáris, reguláris, szinguláris 40, 41
 –, bilineáris 216
 – (leképezés) inverz függvénye 20
 –, lineáris 198
 –, n -lineáris 268
- Gauss-féle elimináció 142
 Gauss-lemma 109
 generált altér 166
 generátorrendszer 160
 Gram-féle determináns 243
 – mátrix 243
 Grassmann-algebra 322
- gyök 126, 282
 –kitevő 19
 –, legalább és pontosan r -szerez 92
 –, mátrixpolinomé 282
 –, polinomé 91
 –, polinomrendszeré 142
 gyöktényezős alak 95
 gyökvonás 19
 gyűrű 22, 114
 –, egységelemes 114
 –, euklideszi 114
 – feletti mátrix 272
 –, kommutatív 114
 –, nullosztómentes 114
- halmaz 19
 –ok direkt szorzata 19
 hányados 14
 harmadfokú polinom, „hiányos” 97
 háromszög mátrix, felső (alsó) 69
 hatvány, hatványozás 18
 hatványösszeg, k -adik 133
 – monotonitása 19
 helyettesítési érték 89, 126, 282
 – –, mátrixpolinomé 282
- Hermite-féle bilineáris függvény 234
 homogén k -adfokú polinom 123
 homogén lineáris egyenletrendszer 293
 – – leképezés 181
 – – –, másodfajú 231
 – szimmetrikus polinom 136
 homomorfizmus (injektív, szürjektív) 21
 Horner-elrendezés 90
- ideál 82, 247
 –, polinomhalmaz generálta 82
 –, triviális és valódi 82
 idempotens 209
 identikus leképezés 192
 identitás 20, 192
 indefinit 225
 index 21
 –halmaz 21
 injektív függvény (leképezés) 20
 – homomorfizmus 21
 integritási tartomány 114, 247
 interpoláció, Lagrange-féle 94
 –, Newton-féle 94
 interpolációs alappolinomok 94
 – polinom 93
 invariáns altér 196
 – faktor 284
 inverz függvény 20
 – leképezés 195
 – mátrix 67
 inverzió 54
 inverziószám 54
 involúció 27
 irányszög 30
 irreducibilis (felbonthatatlan) 81
 irreflexivitás 16
 izomorfizmus 21, 162, 311
 –, természetes 197

- jobbinverz, mátrixé 67
- jobbreguláris leképezés 193
- Jordan-blokk 287
- Jordan-féle normálalak 287
- k -adfokú tenzor 320
- k -adik hatványösszeg 127
- kép, képtér 185
- kibővített mátrix 292
- kifejtési tétel 64
- kisebb-egyenlő 16
- kísérőtranszformáció 207
- kiterjesztés 187
- kitevő (hatvány alapja) 18
- kommutatív csoport 148
 - diagram 309
 - gyűrű 114
- kommutativitás (felcserélhetőség), összegé, szorzaté 13
- , tenzorszorzaté 311
- komplex egység 27
 - egységgyök 37
 - skalárszorzat 245
- komplex szám abszolút értéke 27
 - – konjugáltja, nyoma, normája 27
 - – valós és képzetes része 27
 - számok 23
 - – számteste 25
- kompozíció 20, 124
 - , polinomoké 87
- konjugálás 231
- konjugált, komplex számé 27
 - , leképezése 231
- konstans 291
 - polinom 76
 - tag 76
 - –, x_k -ban 120
- kontravariáns tenzor 314
- koordináták 202
- ko-univerzalitás 310
- kovariáns tenzor 314
- köbre emelés 18
- közös gyök, polinomoké 142
- Kronecker féle δ -függvény 66
- különbség 153
- külső szorzás 322
- kvadratikus alak (forma) 223
 - – mátrixa 225
- kvadratikus karakter 225, 234
- Lagrange-féle interpoláció 94
- legkisebb közös többszörös 16
- legnagyobb közös osztó 16
 - – –, polinomoké 83
- leképezés 20
 - , balreguláris 193
 - , bilineáris 216
 - c -szere 189
 - inverze 195
 - , jobbreguláris 193
 - konjugáltja 231
 - mátrixa 203
 - rangja 186
 - , természetes 184
 - transzponáltja 204
- leképezések direkt összege 187
 - összege 189
- lineáris alakzat 168
 - – egy reprezentánsa 171
- lineáris függés 155
 - függvény (operátor) 198
 - törtfüggvény (másodfajú) 41
 - transzformáció 195
 - vagy elsőfokú polinom 76

- lineáris kombináció 154
- –, mátrixoké 65
- –, nemtriviális 155
- –, triviális 65, 155
- –, végtelen sok vektoré 155
- lineáris leképezés 181
- – jellege 231
- lineárisan független 157
- – mátrixok 65
- lineárisan összefüggő 157
- – mátrixok 65
- linearitások 307
- logaritmálandó 19
- logaritmálás 19
- logaritmus 19
- alapja 19
- mag, magtér 185
- maradékos osztás 16, 78
- maradékosztály (műveletek) 22
- -gyűrű (modulo m vett) 22
- másodfajú homogén lineáris leképezés 231
- mátrix 43
- , adjungált 52
- , antiszimmetrikus 69
- balinverze 67
- , blokkokra diagonalizált 249
- , diagonális 52
- , egyenletrendszeré 292
- , Gram-féle 243
- inverze 67
- jobbinverze 67
- , kibővített 292
- , négyzetes (kvadratikus) 45
- , normálalakú 274
- , ortogonális 259
- , oszlopvéges 70
- , önadjungált 53
- , polinom- 139
- , skalár- 52
- , sor- és oszlopvéges 70
- , szimmetrikus 53
- , transzponált 52
- , unitér 259
- , Vandermonde-féle 139
- , véges 70
- , végtelen 70
- mátrixgyűrű, teljes 52
- mátrixok egyenlősége 45
- lineáris kombinációi 65
- , lineárisan független 65
- , lineárisan összefüggő 65
- összege 47
- szorzata 48, 49
- mátrixpolinom 282
- gyöke 282
- helyettesítési értéke 282
- , normált 282
- maximum 16
- megoldás, triviális 293
- megszorítás 187
- mérték 268
- minimum 16
- modell 23
- modulus, bal oldali 149
- , szabad 306
- homomorfizmus 182
- monotonitás 16
- multiplicitás 92
- multiplikatív inverz, valós számé 14
- művelet, egyváltozós 148
- művelettartó függvény (leképezés) 20, 181

- n -edik egységgyök, primitív 39
- n -lineáris függvény 268
- nagyobb-egyenlő 16
- negatív 13, 16
 - definit 225
 - szemidefinit 225
- négyzetes (kvadratikus) mátrix 45
 - – fődiagonálisa 52
- négyzetgyökvonás 19
- négyzetre emelés 18
- nemtriviális lineáris kombináció 155
- nevező (törté) 17
- Newton-féle interpoláció 94
- Newton-képletek 138
- nilpotens 209
- norma (komplex számé) 27
- normálalak, Jordan-féle 287
- normálalakú mátrix 274
- normális transzformáció 253
- normált mátrixpolinom 282
 - polinom 76, 122
- nullosztómentes gyűrű 114
- numerus 19
- nyom 52
- , komplex számé 27
- ortogonális (merőleges) 227
 - bázis 227
 - mátrix 259
 - transzformáció 258
 - vektorok 241
- oszlop mátrix 45
- oszlopvéges mátrix 20
- oszlopvektor 45
- osztandó 14
- oszthatóság 16, 79
- oszthatóság tulajdonságai, reflexivitása, antiszimmetriája, tranzitivitása 79
- osztó 14, 16, 79
- önadjungált mátrix 53
 - transzformáció 256
- összeg, mátrixoké 47
- páratlan permutáció 54
- paritás 54
- páros permutáció 54
- Peano-axiómák 15
- permutáció (páros, páratlan) 54
- permutálás 53
- polinom, alternáló 139
 - deriváltja 128
 - , egyhatározatlanú 72
 - , egytagú 122
 - , elsőfokú vagy lineáris 76
 - foka 76, 122
 - foka x_k -ban 120
 - főegyütthatója 76
 - gyöke 91
 - , homogén k -adfokú 123
 - k -edik deriváltja 131
 - , normált 76, 122
 - összes gyöke 102
 - , primitív 108
 - súlya 137
 - , szimmetrikus 133
 - , R -beli együtthatós 74
 - , x_k -ban elsőfokú 120
 - , x_k -ban konstans 120
 - , x_k -ban normált 120
- polinomfüggvény 92
- polinomgyűrű 75
 - , többhatározatlanú 120
- polinomhalmaz generálta ideál 82
- polinomideál 82
- polinom mátrix 139, 272
- polinomműveletek 73

- polinomok egyértelmű felbontása 85
 - kompozíciója 87
 - közös gyöke 142
 - legnagyobb közös osztója 83
 - oszthatósága, egész együtthatós 108
 - –, testbeli együtthatós 79
- polinomrendszer gyöke 142
 - következménye 142
- pozitív 16
 - definit 225
 - szemidefinit 225
- primitív polinom 108
- prímtulajdonság 81
- produktum 21
- projekció 186, 188
- projektív sík 169
- projektív tér 169, 224
- R*-beli együtthatós polinom 74
- R*-modulus 149
- racióális szám 17
- rang 212
- reciprok, valós számé 14
- reflexivitás 16
 - , oszthatóságé 79
- rekurzió 15
- reláció 19
 - , két- és többváltozós 20
- relatív prím 16
- rendezés (teljes) 16
- részgyűrű 150
- részhalmaz 19
- részmatrix 62, 324
- részmodulus 165
- rezultáns 294, 295
- sajátaltér 196
- sajátérték 196
- sajátvektor 196
- Schönemann–Eisenstein-tétel 111
- skaláris (belső) szorzat 236, 237
- skalármatrix 52
- skalárok 148
- skalárral való szorzás 148
- sor- és oszlopvéges mátrix 70
- sormatrix 45
- sorvéges mátrix 70
- sorvektor 45
- szabad modulus 306
- számegyenes 18
- számláló (törté) 17
- számtest 17
 - , komplex számoké 25
- szimmetrikus bihomomorfizmus 217
 - mátrix 53
 - polinom 133
 - polinomok alaptétele 134
 - tenzor 321
 - transzformáció 256
- szorzás skalárral 148
- szorzat, mátrixoké 48, 49
- szumma 21
- szürjektív függvény (leképezés) 20
 - homomorfizmus 21
- teljes indukció 14
 - mátrixgyűrű 52
- tenzor 314
 - , antiszimmetrikus 322
 - , *k*-adfokú 320
 - koordinátái 319
 - , szimmetrikus 321
- tenzoralgebra 321
- tenzorszorzat 308
 - -diagram 308
- természetes izomorfizmus 197, 311

- természetes leképezés 184
- többszörösen polinomgyűrű 120
- többszörös 79
 - gyök 92
- tört 17
 - nevezője 17
 - számlálója 17
- transzformáció adjungáltja 253
 - , elemi 209
 - , normális 253
 - , ortogonális 258
 - , önjungált 256
 - , szimmetrikus 256
 - , unitér 258
- transzformációk polinomja 246
- transzponált 204
 - mátrix 52
- transzitivitás 16
 - , oszthatóság 80
- trigonometrikus alak 34
- triviális altér 166
 - direkt összeg 175
 - faktortér 173
 - ideál 82
 - lineáris kombináció 65, 155
 - megoldás 293
 - vektortér 162
- unitér mátrix 259
 - transzformáció 258
- üres halmaz 19
- valódi altér 166
 - direkt összeg 175
- faktortér 173
- ideál 82
- valós szám (összege, szorzata) 13
- valós számok 18
 - – kivonása 13
 - – négyzete, köbe 18
- Vandermonde-determináns 139
 - -féle mátrix 139
- véges mátrix 70
- végesen generált torziómodulus 247
- végtelen dimenziós vektortér 161
 - mátrix 70
 - sok vektor lineáris kombinációja 155
 - tagú direkt összeg 176
- vektor merőleges vetülete 242
 - hossza 240
- vektorok 148
 - hajlásszöge 240
 - , ortogonális (merőleges) 241
- vektorösszeadás 148
- vektorrendszer 154
 - ek ekvivalenciája 156
- vektortér 147
 - , nulldimenziós 162
- vektortér-axiómák 150
- vektortér-homomorfizmus 181
- vetítés 188
- x_k -ban elsőfokú polinom 120
 - konstans polinom 120
 - konstans tag 120
 - normált polinom 120
- zárttság (összeadásra, kivonásra) 15