LINEÁRIS ALGEBRA

NAGY ATTILA

EGYETEMI JEGYZET

Budapesti Műszaki és Gazdaságtudományi Egyetem

Algebra Tanszék

2011

Ez a jegyzet a Budapesti Műszaki és Gazdaságtudományi Egyetemen a Matematika Alapszak keretén belül oktatandó "Lineáris algebra előadás" című tárgy általam tartott előadásainak anyagát tartalmazza.

Tartalomjegyzék

1. fejezet

Alapfogalmak

Ebben a fejezetben azon általános algebrai fogalmakat és ismereteket foglaljuk össze, amelyekre a jegyzetben utalást találunk.

Definíció 1.1 *Tetszőleges A és B halmazok (ebben a sorrendben vett) descartes-szorzatán értjük az*

$$A \times B = \{(a, b) : \ a \in A, b \in B\}$$

halmazt. Az $A \times A$ szorzatot A^2-tel fogjuk jelölni. Tetszőleges $n \geq 2$ egész szám esetén A^n jelöli az $A^{n-1} \times A$ halmazt. Az A^n elemei tulajdonképpen az A elemeiből képezett rendezett elem n-esek halmaza (n pozitív egész szám (A^1 az A-t jelöli)).

1.1. Algebrai struktúrák

Definíció 1.2 *Egy A halmazon értelmezett n-változós műveleten (n pozitív egész szám) az A^n halmaznak az A halmazba való egyértelmű leképezését értjük. Az $n = 2$ esetben a műveletet binér műveletnek is fogjuk nevezni. Ekkor, ha a szóban forgó műveletet f jelöli, akkor valamely $(x, y) \in A^2$ elempárhoz hozzárendelt f-szerinti A-beli elemet $f(a, b)$ helyett afb módon jelöljük (például: $a + b$, $a * b$ vagy $a \cdot b$).*

1.1.1. Az algebrai struktúra fogalma

Definíció 1.3 *Algebrai struktúrán olyan nem-üres A halmazt értünk, amelyen értelmezve van legalább egy művelet. Az A halmazt az algebrai struktúra* *alaphalmazának nevezzük. Az algebrai struktúrákat kétféleképpen fogjuk jelölni. Vagy* $(A;\Omega)$ *módon, ahol* A *jelöli az algebrai struktúra alaphalmazát* Ω *pedig az* A*-n értelmezett műveletek halmazát, vagy egyszerűen csak az alaphalmazzal. Például,* $(\mathbb{Z};+,\cdot)$ *jelöli az egész számok azon algebrai struktúráját, amelyen mind az összeadást, mind a szorzást figyelembe vesszük. A* $(\mathbb{Z};+)$ *jelölés arra utal, hogy most is az egész számok halmazát tekintjük, de csak az összeadásra koncentrálunk. A* $\mathbb{Z}$ *jelölés arra utal, hogy az egész számokról van szó, de nem részletezzük, hogy milyen műveletek kerülnek a figyelem középpontjába.*

Megjegyzés 1.1 *Két algebrai struktúráról azt fogjuk mondani, hogy azonos típusú algebrai struktúrák, ha a hozzájuk tartozó műveletek halmazai között létesíthető olyan bijekció, hogy az egymásnak megfeleltetett műveletek azonos változójúak. Az egymásnak megfeleltetett műveleteket azonos betűvel jelöljük.*

Definíció 1.4 *Legyenek* (A,Ω) *és* (B,Ω) *azonos típusú algebrai struktúrák. Az* A *halmaznak a* B *halmazba való valamely egyértelmű* φ *leképezétét az* (A,Ω) *algebrai struktúrának a* (B,Ω) *algebrai struktúrába való homomorfizmusának nevezzük, ha tetszőleges* $\omega\in\Omega$ *(*n*-változós) művelet esetén*

$$\varphi(\omega(a_1,\ldots,a_n))=\omega(\varphi(a_1),\ldots,\varphi(a_n))$$

teljesül tetszőleges $a_1,\ldots,a_n\in A$ *elemekre (ezen utóbbi feltétel helyett azt is szoktuk mondani, hogy* φ *kompatibilis az* ω *műveletre nézve).*

Definíció 1.5 *Legyenek* A *és* B *azonos típusú algebrai struktúrák. Azt mondjuk, hogy* A *beágyazható* B*-be ha megadható* A*-nak* B*-be való injektív homomorfizmusa. Ha ez a homomorfizmus szürjektív is (azaz* A*-t* B*-re képezi), akkor azt mondjuk, hogy* A *izomorf* B*-vel (ekkor persze* B *is isomorf* A*-val, ezért azt is szoktuk mondani, hogy* A *és* B *izomorf algebrai struktúrák).*

Megjegyzés 1.2 *Ebben a jegyzetben csak binér műveletekkel fogunk foglalkozni, ezért a továbbiakban a művelet kifejezés mindig egy binér műveletet fog jelenteni.*

1.1.2. A félcsoport és a csoport

Definíció 1.6 *Azt mondjuk, hogy egy* A *halmazon értelmezett* $*$ *művelet asszociatív, ha minden* $a,b,c\in A$ *elemhármas esetén*

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

teljesül.

Definíció 1.7 *Ha egy A halmazon értelmezett $*$ művelet esetén tetszőleges $a, b \in A$ elemekre teljesül az $a * b = b * a$ egyenlőség, akkor azt mondjuk, hogy a $*$ művelet kommutatív.*

Definíció 1.8 *Egy egyműveletes $(S; *)$ algebrai stuktúrát félcsoportnak nevezünk, ha a $*$ művelet asszociatív. Ha a $*$ művelet ezen felül még kommutatív is, akkor kommutatív félcsoportról beszélünk.*

Definíció 1.9 *Egy A halmazon értelmezett $*$ műveletet invertálhatónak nevezünk (az A halmazon), ha minden $a, b \in A$ elempárhoz megadhatók az A halmaz olyan x és y elemei, hogy teljesül az $a * x = b$ és $y * a = b$ egyenlőségek mindegyike.*

Megjegyzés 1.3 *Például az egész számok halmazán az összeadás invertálható, mert az $a + x = b$ (és $y + a = b$) egyenletek megoldhatók $\mathbb{Z}$-ben tetszőleges $a, b \in \mathbb{Z}$ esetén. Megjegyezzük, hogy ez nem igaz a szorzásra, mert az $ax = b$ (és $ya = b$) egyenleteknek nem minden $a, b \in \mathbb{Z}$ esetén van $\mathbb{Z}$-ben megoldása.*

Definíció 1.10 *Ha egy $(G; *)$ félcsoport esetén a $*$ művelet invertálható is G-n, akkor a $(G; *)$ algebrai stuktúrát csoportnak nevezzük. Ha a $*$ művelet kommutatív is, akkor kommutatív csoportról (vagy Abel-csoportról) beszélünk.*

Megjegyzés 1.4 *Legyen $(S; *)$ egy félcsoport és $a \in S$ tetszőleges elem. Az $a*a$ elemet jelölhetjük $2a$-val, de jelölhetjük a^2-tel is, annak mintájára, hogy a számoknál $a+a$ helyett $2a$-t, illetve $a \cdot a$ helyett a^2-t írunk. Az első esetben azt mondjuk, hogy additív írásmódot, a második esetben multiplikatív írásmódot használunk. Additív írásmód esetén a művelet jeleként a $+$ jelet, multiplikatív írásmód esetén a művelet jeleként a $\cdot$ jelet használjuk. Multiplikatív irásmód esetén (ha nem okoz félreértést) a művelet jelét elhagyjuk, s az $a \cdot b$ kifejezés helyett egyszerűen ab-t írunk.*

Definíció 1.11 *Egy $(S; *)$ félcsoport e elemét bal oldali neutrális elemnek nevezzük, ha minden $s \in S$ elemre $e * a = a$ teljesül. A jobb oldali neutrális elem fogalma a bal oldali neutrális elem fogalmának duálisa. Az e elemet az S félcsoport neutrális elemének nevezzük, ha e bal oldali és jobb oldali neutrális eleme S-nek.*

Megjegyzés 1.5 *Megjegyezzük, hogy a neutrális elem kifejezés helyett multiplikatív írásmód esetén az egységelem kifejezést, additív írásmód esetén pedig a nullelem kifejezést használjuk. Ezek szerint beszélhetünk bal oldali és jobb oldali egységelemekről, illetve nullelemekről, valamint egységelemről és nullelemről.*

Tétel 1.1 *Ha egy félcsoportnak van bal oldali és jobb oldali neutrális eleme, akkor azok egybeesnek. Következésképpen minden félcsoportnak legfeljebb egy neutrális eleme van.*

Bizonyítás. Ha f, illetve e jelöli egy $(S; *)$ félcsoport valamely bal oldali, illetve jobb oldali neutrális elemét, akkor $f * e = e$, mert f bal oldali neutrális elem, továbbá $f * e = f$, mert e jobb oldali neutrális elem. Tehát $e = f$. □

Definíció 1.12 *Egy neutrális elemes félcsoportot monoidnak nevezünk.*

Definíció 1.13 *Egy $(S; *)$ monoid a' elemét az $a \in S$ elem bal oldali inverzének nevezzük, ha $a' * a = e$, ahol e jelöli az S neutrális elemét. Az a elem jobb oldali inverzének fogalma a bal oldali inverz fogalmának duálisa (azaz $a'' \in S$ az $a \in S$ elem jobb oldali inverze, ha $a * a'' = e$). Az S monoid valamely a' elemét az a elem inverzének nevezzük, ha az az a elemnek jobb oldali és bal oldali inverze (azaz $a' * a = a * a' = e$).*

Tétel 1.2 *Ha a', illetve a'' egy $(S; *)$ monoid valamely a elemének bal oldali, illetve jobb oldali inverze, akkor $a' = a''$. Következésképpen, egy monoid minden elemének legfeljebb egy inverze van.*

Bizonyítás. A tétel jelöléseit használva,

$$a' = a' * e = a' * (a * a'') = (a' * a) * a'' = e * a'' = a'',$$

ahol e jelöli az S neutrális elemét. □

Megjegyzés 1.6 *Additív írásmód esetén egy a elem inverzét az a elem ellentettjének is nevezzük, és $-a$ módon jelöljük. Multiplikatív írásmód esetén egy a elem inverzére a definícióbeli eredeti kifejezést (inverz) használjuk, és azt a^{-1} módon jelöljük.*

Tétel 1.3 *Egy $(S; *)$ félcsoport akkor és csak akkor csoport, ha van neutrális eleme (azaz $(S; *)$ monoid) és minden elemének van inverze.*

Bizonyítás. Legyen $(S;*)$ olyan monoid, amelyben minden elemnek van inverze. Akkor a $*$ művelet invertálható S-en, hiszen az S-beli $a' * b$, illetve $b*a'$ elemek megoldásai az $a*x = b$, illetve $y*a = b$ egyenleteknek tetszőleges $a, b \in S$ esetén, ahol a' jelöli az a elem inverzét. Ezért - definíció szerint - S csoportot alkot a $*$ műveletre nézve.

Fordítva, tegyük fel, hogy $(S;*)$ csoport. Legyen $a \in S$ tetszőleges elem. Mivel a feltétel szerint a $*$ művelet invertálható S-en, ezért az $y * a = a$ egyenletnek van S-ben megoldása, azaz van olyan $e \in S$ elem, hogy $e * a = a$. Ugyancsak a $*$ művelet invertálhatóságát használva, tetszőleges $s \in S$ elemhez megadható S-nek olyan x eleme, hogy $a * x = s$. Ezért

$$e * s = e * (a * x) = (e * a) * x = a * x = s.$$

Tehát e az S bal oldali neutrális eleme. Hasonlóan igazolható, hogy S-nek van jobb oldali neutrális eleme. Ezért a Tétel 1.1 miatt S monoid. Jelölje e az S neutrális elemét. A $*$ művelet invertálhatósága miatt minden $a \in S$ elemhez megadhatók olyan $a', a'' \in S$ elemek, hogy $a * a'' = e = a' * a$. A Tétel 1.2 miatt $a" = a'$, amely az a elem inverze. Tehát az S monoid minden elemének van inverze. □

A továbbiakban félcsoporton multiplikatív félcsoportot fogunk érteni, és a művelet jelét nem fogjuk kiírni.

Definíció 1.14 *Egy S félcsoportot egyszerűsítése félcsoportnak nevezünk, ha minden $a, b, x \in S$ esetén az $ax = by$ és az $xa = ya$ feltételek bármelyikéből $a = b$ következik.*

Tétel 1.4 *Egy kommutatív S félcsoportot be lehet ágyazni egy csoportba akkor és csak akkor, ha S egyszerűsítéses.*

Bizonyítás. Ha S beágyazható egy G csoportba, akkor S szükségképpen egyszerűsítéses.

Fordítva, tegyük fel, hogy S kommutatív, egyszerűsítéses félcsoport. Az S^2 halmazon definiáljunk egy összeadást a következőképpen. Tetszőleges $(a, b), (c, d) \in S^2$ esetén legyen

$$(a, b) + (c, d) = (a + c, b + d).$$

Könnyen ellenőrízhető, hogy S^2 félcsoportot alkot erre a műveletre nézve. Az S^2 halmazon definiáljunk egy (binér) relációt a következőképpen. Az S^2 két eleme, (a,b) és (c,d) akkor és csak akkor álljanak relációban, ha $a+d=c+b$. Könnyen ellenőrízhető, hogy ez a reláció reflexív, szimmetrikus és tranzitív, azaz egy ekvivalenciareláció. Jelölje $[a,b]$ az (a,b) elempárt tartalmazó ekvivalenciaosztályt. Világos, hogy tetszőleges $a,b,c,d \in S$ elemek esetén

$$[a,b]+[c,d] \subseteq [a+c, b+d].$$

Ezért

$$[a,b]+[c,d]=[a+b, c+d]$$

jól definiált művelet az ekvivalenciaosztályok G halmazán. Ez a művelet (összeadás) asszociatív és kommutatív. Megmutatható, hogy az (a,a) ($a \in S$) alakú elemek egyetlen osztályt alkotnak, amely a G félcsoport neutrális eleme. Az is világos, hogy a $[b,a]$ osztály az $[a,b]$ osztály ellentettje. Ezért G kommutatív csoportot alkot a szóban forgó Összeadásra nézve. Könnyen ellenőrízhető, hogy a

$$\phi\ a \ \mapsto\ [2a, a]$$

leképezés az S félcsoportnak a G csoportba való beágyazása. □

Definíció 1.15 *Egy multiplikatív S félcsoport valamely 0 elemét jobbzéró elemnek nevezzük, ha minden $s \in S$ elemre $s0 = 0$ teljesül. A balzéró elem fogalma a jobbzéró elem fogalmának duálisa. Az S félcsoport egy elemét zéró elemnek nevezzük, ha az jobbzéró és balzéró eleme S-nek.*

Tétel 1.5 *Ha egy (multiplikatív) S félcsoportnak van balzéró és jobbzéró eleme, akkor azok egybeesnek. Következésképpen, minden félcsoportnak legfeljebb egy zéró eleme lehet.*

Bizonyítás. Ha 0, illetve $0'$ jelöli S egy balzéró, illetve jobbzéró elemét, akkor $0' = 0'0 = 0$. □

Definíció 1.16 *Egy A halmaz önmagába való (egyértelmű) leképezéseit az A halmaz transzformációinak nevezzük. Ha egy transzformáció bijektív (szürjektív és injektív), akkor azt az A halmaz egy permutációjának nevezzük. Egy n-elemű (n pozitív egész szám) halmaz permutációit n-edfokú permutációknak nevezzük; ezek halmazát S_n-nel fogjuk jelölni.*

Definíció 1.17 *Legyenek $(\cdot)\sigma$ és $(\cdot)\tau$ egy A halmaz permutációi. A σ és τ (ebben a sorrendben vett kompozícióján azt a $\sigma \circ \tau$ permutációt értjük, amely az A tetszőleges a eleméhez az $((a)\sigma)\tau$ elemet rendeli.*

Tétel 1.6 *Egy halmaz összes permutációinak halmaza a $\circ$ műveletre nézve csoportot alkot. Ennek a csoportnak az egységeleme az identikus leképezés, s egy elemének inverze az elem (mint bijektív leképezés) inverz leképezése.*

A továbbiakban az $\{1, \ldots, n\}$ halmaz permutációival foglalkozunk, ezért n-edfokú permutáció említésekor mindig ennek a halmaznak a permutációira gondolunk. Az n-edfokú permutációk S_n csoportját n-edfokú szimmetrikus csoportnak nevezzük.

Definíció 1.18 *Egy n-edfokú σ permutáció egy inverzióján olyan (i, j) rendezett párt értünk, amelyben szereplő i és j elemei az $\{1, \ldots, n\}$ halmaznak, és teljesítik a következő két feltételt: $i < j$, viszont $\sigma(i) > \sigma(j)$. A σ permutáció inverzióinak számát $I(\sigma)$-val fogjuk jelölni.*

Feladat 1.1 *Az*

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

permutáció inverziói: $(1, 2), (1, 3)$. Tehát $I(\sigma) = 2$.

Tétel 1.7 *Tetszőleges n-edfokú σ permutáció esetén*

$$I(\sigma) = I(\sigma^{-1}).$$

Definíció 1.19 *Egy (n-edfokú) permutációt párosnak nevezünk, ha inverzióinak száma páros. Ellenkező esetben páratlan permutációról beszélünk. Az identikus permutációt (ahol az inverziók száma 0) páros permutációnak tekintjük.*

Tétel 1.8 *Legyenek $x_1, \ldots, x_n$ tetszőleges, páronként különböző valós számok. Egy n-edfokú σ permutáció akkor és csak akkor páros, ha*

$$\prod_{1 \leq i < j \leq n}^{n} (x_j - x_i) = \prod_{1 \leq i < j \leq n}^{n} (x_{\sigma(j)} - x_{\sigma(i)}).$$

σ akkor és csak akkor páratlan, ha

$$\prod_{1 \leq i < j \leq n}^{n} (x_j - x_i) = - \prod_{1 \leq i < j \leq n}^{n} (x_{\sigma(j)} - x_{\sigma(i)}).$$

Tétel 1.9 *Az S_n csoport összes páros permutációinak A_n halmaza az S_n egy normális részcsoportja.*

Definíció 1.20 *Egy n-edfukú σ permutációt k-elemű ($k = 1, \ldots, n$) ciklusnak nevezünk, ha megadhatók olyan páronként különböző $i_1, \ldots, i_k \in \{1, \ldots, n\}$ elemek, úgy, hogy*

$$(i_j)\sigma = i_{j+1}, \quad j = 1, \ldots k-1,$$

$$(i_k)\sigma = i_1,$$

a többi elemet viszont σ fixen hagyja. A ciklus jelölése (a fentiek alapján): $(i_1 \ldots i_k)$. Két ciklusról azt mondjuk, hogy diszjunktak, ha az őket alkotó elemek halmazai diszjunktak. Egy kételemű ciklust transzpozíciónak nevezünk.

Feladat 1.2 *A*

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \quad \textit{és} \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix}$$

permutációk diszjunkt ciklusok, mivel $\sigma = (1\ 2)$ és $\tau = (3\ 4)$, valamint $\{1,2\} \cap \{3,4\} = \emptyset$.

Tétel 1.10 *Diszjunkt (n-edfokú) ciklusok szorzata nem függ a tényezők sorrendjétől.*

Tétel 1.11 *Minden n-edfokú permutáció előáll (n-edfokú) diszjunk ciklusok szorzataként. Ez az előállítás (a tényezők sorrendjétől eltekintve) egyértelmű.*

Megjegyzés 1.7 *Legyen $(R;+)$ egy kommutatív csoport. Ha az S_n csoport minden σ eleméhez hozzá van rendelve egy R-beli r_σ elem, akkor*

$$\Sigma_{\sigma \in S_n} r_\sigma = \Sigma_{\sigma \in S_n} r_{\sigma^{-1}}.$$

1.1.3. A gyűrű, a ferdetest és a test

Definíció 1.21 *Egy kétműveletes $(R;+,\cdot)$ algebrai struktúrát gyűrűnek nevezünk, ha az $(R;+)$ additív struktúra kommutatív csoport (ennek neutrális elemét a gyűrű nulleleménék nevezzük), az $(R;\cdot)$ multiplikatív struktúra félcsoport, valamint a szorzás $\cdot$ művelete mindkét oldalról disztributív az összeadás $+$ műveletére nézve, azaz minden $a, b, c \in R$ esetén teljesülnek az $a \cdot (b + c) = a \cdot b + a \cdot c$ és $(b + c) \cdot a = b \cdot a + c \cdot a$ egyenlőségek.*

Egy olyan gyűrűt, amelyben a szorzás kommutatív is, kommutatív gyűrűnek nevezzük.

Ha egy $(R : +, \cdot)$ gyűrűben az $(R; \cdot)$ multiplikatív félcsoportnak van neutrális eleme (amelyet a gyűrű egységelemének hívunk), akkor a gyűrűt egységelemes gyűrűnek nevezzük.

Megjegyzés 1.8 *Tetszőleges $(R; +)$ kommutatív csoportból képezhetünk gyűrűt, ha R-en a szorzást úgy definiáljuk, hogy bármely két elem szorzara legyen egyenlő $(R; +)$ neutrális elemével (azaz $(R; +)$ nullelemével). Az ilyen gyűrűt zérógyűrűnek nevezzük. Ennek speciális esete az a gyűrű, amelynek csak egy eleme van; az ilyen gyűrűt nullgyűrűnek nevezzük.*

Megjegyzés 1.9 *Könnyen ellenőrízhető, hogy $(\mathbb{Z}; +, \cdot)$, $(\mathbb{Q}; +, \cdot)$, $(\mathbb{R}; +, \cdot)$ egy-egy példa gyűrűre. Ezek mindegyike kommutatív és egységelemes.*

Tétel 1.12 *Ha 0 jelöli egy $(R; +, \cdot)$ gyűrű nullelemét (azaz 0 az $(R; +)$ additív csoport neutrális eleme), akkor minden $r \in R$ elemre $r0 = 0r = 0$ teljesül.*

Bizonyítás. Legyen $r \in R$ tetszőleges elem. Akkor

$$r0 = r(0 + 0) = r0 + r0.$$

Ha $r0$ ellentettjét hozzáadjuk az egyenlőség mindkét oldalához, akkor

$$0 = r0 + (-r0) = [r0 + r0] + (-r0) = r0 + [r0 + (-r0)] = r0 + 0 = r0$$

adódik.

A $0r = 0$ egyenlőség is hasonlóan igazolható. □

Megjegyzés 1.10 *Az előző tétel miatt egy R gyűrű tetszőleges r eleme esetén akkor és csak akkor van R-nek olyan x eleme, amelyre $0x = r$ (vagy $x0 = r$) teljesül, ha $r = 0$. Ezért egy R gyűrűben a szorzás akkor és csak akkor invertálható, ha a gyűrűnek egyetlen eleme van (azaz R nullgyűrű). Ezért a szorzás invertálhatósága nem fordulhat elő legalább kételemű gyűrűben. Viszont megvizsgálhatjuk, hogy a szorzás invertálható-e a gyűrű valamely olyan részhalmazán, amely nem tartalmazza a nullelemet. A fentiekben példaként szereplő három gyűrű esetén $(\mathbb{Q}; +, \cdot)$ és $(\mathbb{R}; +, \cdot)$ olyan gyűrűk, amelyekben a szorzás invertálható a $\mathbb{Q} - \{0\}$, illetve $\mathbb{R} - \{0\}$ monoidokon; pontosabban mondva, $((\mathbb{Q} - \{0\}); \cdot)$, illetve $((\mathbb{R} - \{0\}); \cdot)$ (multiplikatív) csoportok.*

Definíció 1.22 *Egy $(F; +, \cdot)$ gyűrűt ferdetestnek nevezünk, ha $((F - \{0\}); \cdot)$ csoport. Egy kommutatív ferdetestet testnek nevezünk.*

Megjegyzés 1.11 *A definícóból következik, hogy minden* $(F;+,\cdot)$ *ferdetestnek legalább két eleme van, az* $(F;+)$ *additív csoport* 0 *neutrális eleme (azaz* F *nulleleme) és az* $((F-\{0\});\cdot)$ *multiplikatív csoport* e *neutrális eleme (amely egyben az* F *egységeleme).*

Megjegyzés 1.12 $(\mathbb{Q};+,\cdot)$ *és* $(\mathbb{R};+,\cdot)$ *egy-egy példa testre. Nem kommutatív ferdetestre a ? fejezetben fogunk példát látni.*

Definíció 1.23 *Egy* R *gyűrű valamely* a *elemét bal oldali nullosztónak nevezzük, ha van olyan* $0 \neq b \in R$ *elem, amelyre* $ab = 0$ *teljesül. A jobb oldali nullosztó fogalma a bal oldali nullosztó fogalmának duálisa.*

Megjegyzés 1.13 *Tetszőleges nem-nullgyűrű nulleleme bal oldali és jobb oldali nullosztó. Könnyen igazolható, hogy egy gyűrű akkor és csak akkor nem tartalmaz nem-nulla bal oldali nullosztót, ha nem tartalmaz nem-nulla jobb oldali nullosztt́.*

Definíció 1.24 *Egy gyűrűt nullosztómentesnek nevezünk, ha nem tartalmaz nem-nulla bal oldali nullosztót (és így nem-nulla jobb oldali nullosztót sem).*

Definíció 1.25 *Egy kommutatív nullosztómentes gyűrűt integritástartománynak nevezünk.*

Definíció 1.26 *Legyen* $(R;+,\cdot)$ *egy egységelemes integritástartomány. Azt mondjuk, hogy valamely* $a \in R$ *elem osztója egy* $b \in R$ *elemnek, ha van olyan* $x \in R$ *elem, hogy* $ax = b$*. Azt a tényt, hogy* a *osztója* b*-nek,* $a|b$ *módon jelöljük.*

Megjegyzés 1.14 *Könnyen belátható, hogy tetszőleges egységelemes integritástartományon az oszthatóság reflexív és tranzitív.*

Definíció 1.27 *Egy egységelemes integritástartomány egységelemének osztóit egységeknek nevezzük.*

Megjegyzés 1.15 *Például az egész számok gyűrűjében (amely egységelemes integritástartomány) egységek:* $+1, -1$.

Megjegyzés 1.16 *Világos, hogy minden test integritástartomány, amelyben minden nem-nulla elem egység.*

Definíció 1.28 *Egy egységelemes integritástartomány a és b elemeiről akkor mondjuk, hogy asszociáltak (jelölése: $a \sim b$), ha $a|b$ és $b|a$.*

Megjegyzés 1.17 *Két egész szám akkor és csak akkor aszociált, ha egyenlőek vagy előjelben különböznek, azaz ha bármelyikük egyenlő a másiknak egy egységgel képezett szorzatával. Ez nem csak az egész számokra igaz; érvényes a következő tétel.*

Tétel 1.13 *Egységelemes R interitástartomány a és b elemei akkor és csak akkor asszociáltak, ha ha bármelyikük egyenlő a másiknak egy R-beli egységgel képezett szorzatával.*

Bizonyítás.

Definíció 1.29 *Egy egységelemes R integritástartomány valamely $a \neq 0$ elemét irreducibilis elemnek nevezzük, ha a nem egysége R-nek és bármely $b, c \in R$ elemek esetén az $a = bc$ egyenlőségből $a \sim b$ vagy $a \sim c$ következik.*

Megjegyzés 1.18 *Legyen a egy R integritástartomány irreducibilis eleme. Tegyük fel, hogy $a = bc$ valamely $b, c \in R$ elemekre.*

Ha $a \sim b$, akkor $a = axc$, azaz $0 = ae - axc = a(e - xc)$ teljesül valamely $x \in R$ elemre (itt e jelöli R egységelemét). Az R nullosztómentessége miatt ebből $xc = e$ következik, azaz c az R egyik egysége.

Hasonlóan igazolható, hogy ha $a = bc$ és $a \sim c$, akkor b az R integritástartomány egyik egysége.

Tétel 1.14 *Minden R integritástartományt be lehet ágyazni egy olyan testbe, melynek minden eleme kifejezhető R-beli a és $b \neq 0$ elemek segítségével ab^{-1} formában.*

Bizonyítás. □

Definíció 1.30 *Legyen n egy pozitív egész szám. Azt mondjuk, hogy n egy R gyűrű karakterisztikája, ha minden $r \in R$ elemre $nr = 0$ teljesül, és n a legkisebb az ezen feltételt teljesítő pozitív egészek között.*

Ha $nr = 0$ akkor és csak akkor teljesülhet egy R gyűrű valamely r elemére és valamely nemnegatív egész számra, ha $n = 0$ vagy $r = 0$, akkor azt mondjuk, hogy az R gyűrű karakterisztikája 0. Egy R gyűrű karakterisztikáját $\mathrm{char} R$-rel fogjuk jelölni.

Tétel 1.15 *Nullosztómentes gyűrű karakterisztikája* 0, 1, *vagy prímszám.*

Bizonyítás. Legyen R nullosztómentes gyűrű. Ha $R = \{0\}$, akkor $char R = 1$. Tegyük fel, hogy $|R| > 1$ és $char R \neq 0$. Akkor van olyan $m \geq 2$ pozitív egész szám és R-nek olyan $a \neq 0$ eleme, hogy $ma = 0$. Akkor minden $b \in R$ elemre

$$a(mb) = a(b + \cdots + b) = ab + \cdots + ab = (a + \cdots + a)b = (ma)b = 0,$$

amiből $a \neq 0$ miatt $mb = 0$ következik. Legyen $n \geq 2$ a legkisebb olyan pozitív egész szám, amelyre $nb = 0$ teljesül minden $b \in R$ elemre. Megmutatjuk, hogy n prímszám. Ellenkező esetben n felírható $n = pn'$ alakban, ahol p prímszám, n' pedig egy 1-nél nagyobb egész szám. Ekkor minden $b \in R$ elemre $0 = nb = p(n'b)$ következne. Mivel $n' < n$, ezért $n'b \neq 0$ minden $b \in R$-re (ugyanis, ha $n'b = 0$ valamely $b \in R$ elemre, akkor - mint ahogy azt fent már láttuk - $n'b = 0$ teljesülne minden $b \in R$-re), ezért $pb = 0$ minden $b \in R$-re, ami ellentmondás. Tehát n prímszám. □

Tétel 1.16 *Test karakterisztikája* 0 *vagy prímszám.*

Bizonyítás. Mivel minden test legalább két elemet tartalmazó nullosztómentes gyűrű, ezért az állítás a Tétel 1.15 miatt nyilvánvaló. □

1.2. A komplex számok

1.2.1. A komplex szám fogalma; algebrai alak

Az algebra elméletében alapvető szerepet játszik a számok és az egyenletek elmélete. A legegyszerűbb számfogalom a természetes szám fogalma. A közöttük értelmezett összeadás asszociatív és kommutatív, azaz a természetes számok $\mathbb{N}^+$ halmaza az összeadásra nézve kommutatív félcsoportot alkot. Az összeadás viszont nem invertálható $\mathbb{N}^+$-on, azaz az $a + x = b$ egyenleteknek nem minden $a, b \in \mathbb{N}^+$ esetén van megoldása $\mathbb{N}^+$-ban. Azonban további információt mondhatunk az $\mathbb{N}^+$-ban meg nem oldható $a + x = b$ ($a, b \in \mathbb{N}^+$) egyenletekről, ha az $\mathbb{N}^+$ félcsoportot be lehetne ágyazni olyan, ugyancsak kommutatív, additív G félcsoportba, amelyben az $a + x = b$ egyenletek minden $a, b \in G$ (és így minden $a, b \in \mathbb{N}^+$) esetén G-ben megoldhatók. Mivel az $\mathbb{N}^+$ additív félcsoport egyszerűsítéses és kommutatív, ezért a Tétel 1.4 szerint beágyazható egy kommutatív csoportba. Ez a csoport izomorf az egész számok $\mathbb{Z}$ additív csoportjával, amelyben az $a + x = b$ egyenleteknek minden $a, b \in \mathbb{Z}$ esetén van $\mathbb{Z}$-ben megoldása. Az egész számok a szokásos összeadásra és szorzásra nézve integritástartományt alkotnak; ez a struktúra nem test (az $ax = b$ egyenleteknek nem minden $a, b \in \mathbb{Z}$ esetén van $\mathbb{Z}$-ben megoldása). Az előzőek mintájára felvetődik annak igénye, hogy találjunk az egész számok gyűrűjét részgyűrűként tartalmazó F testet, mert abban már az $ax = b$ ($a \neq 0$) egyenletek mindegyike megoldható tetszőleges $a, b \in F$ (és így tetszőleges $a, b \in \mathbb{Z}$ elemek esetén). A Tétel 1.14 felhasználásával az egész számok gyűrűje beágyazható egy testbe; ez a test izomorf a racionális számok $\mathbb{Q}$ testével. Ebben a testben az összes $a + x = b$ és $cx = d$ alakú egyenletek megoldhatók ($a, b, c, d \in \mathbb{Q}$, $c \neq 0$). Azonban nem minden $f(x) = 0$ egyenlet oldható meg $\mathbb{Q}$-ban, ahol $f(x)$ egy $\mathbb{Q}$ feletti legalább elsőfokú polinom (azaz $f(x) = a_n x^n + \ldots + a_1 x + a_0$, amelyben n pozitív egés szám és $a_n, \ldots, a_0$ racionális számok ($a_n \neq 0$). Például az $x^2 - 2 = 0$ egyenletnek nincs $\mathbb{Q}$-ban megoldása. Itt nem részletezzük, de $\mathbb{Q}$ is beágyazható olyan testbe (a valós számok $\mathbb{R}$ testébe), amelyben már az $x^2 - 2 = 0$ egyenlet is megoldható. Viszont a valós számok testében sem oldhatók meg az $f(x) = 0$ alakú egyeneletek tetszőleges $\mathbb{R}$ feletti $f(x)$ polinomra. Ez a helyzet például a negatív diszkriminánsú másodfokú egyenletekkel. Ebben a fejezetben definiálunk egy testet, a komplex számok $\mathbb{C}$ testét, amelybe beágyazható résztestként a valós számok teste, s ebben a $\mathbb{C}$ testben már megoldhatók az $f(x) = 0$ alakú egyenletek tetszőleges $\mathbb{C}$ feletti legalább elsőfokú $f(x)$ polinom esetén (az ilyen tulajdonságú testeket algebrailag zárt testeknek nevezzük; ezzel a szóhasználattal élve azt mondhatjuk, hogy $\mathbb{C}$ algebrailag zárt test (lásd a ? fejezetet).

Konstrukció 1.1 *A $\mathbb{C} = \mathbb{R}^2$ halmazon definiáljunk két műveletet. Tetszőleges $(a, b), (c, d) \in \mathbb{C}$ elemek esetén legyen*

$$(a, b) + (c, d) = (a + c, b + d)$$

és

$$(a, b)(c, d) = (ac - bd, ad + bc).$$

Tétel 1.17 *A $(\mathbb{C} : +, \cdot)$ algebrai struktúra test.*

Bizonyítás. Mivel tetszőleges a, b, c, d, e, f valós számok esetén

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

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

ezért az összeadás asszociatív a $\mathbb{C}$ halmazon. Világos, hogy $(0, 0)$ a $(\mathbb{C}; +)$ félcsoport neutrális eleme (azaz a nulleleme), $(-a, -b) \in \mathbb{C}$ pedig az $(a, b) \in \mathbb{C}$ elem ellentettje. Tehát $(\mathbb{C}; +)$ csoport. Az is világos, hogy $(\mathbb{C}; +)$ kommutatív.

Mivel tetszőleges a, b, c, d, e, f valós számokra

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

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

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

teljesül, ezért a szorzás asszociatív a $\mathbb{C}$ halmazon. Világos, hogy $(1, 0)$ a $(\mathbb{C}; \cdot)$ multiplikatív félcsoport neutrális eleme (azaz az egységeleme), az $(\frac{a}{a^2+b^2}, \frac{b}{a^2+b^2}) \in \mathbb{C}$ pedig az $(a, b) \in \mathbb{C}$ elem inverze, feltéve, hogy $a^2 + b^2 \neq 0$ (azaz (a, b) nem a $\mathbb{C}$ nulleleme). Könnyen igazolható, hogy a szorzás kommutatív. Ezért $(\mathbb{C}; \cdot)$ olyan kommutatív félcsoport, amelyben $\mathbb{C} - \{(0, 0)\}$ kommutatív csoportot alkot a szorzásra nézve. Már csak azt kell bebizonyítani, hogy a szorzás disztributív az összeadásra nézve (a szorzás kommutativitása miatt elegendő csak az egyik oldali disztributivitást igazolni).

Legyenek a, b, c, d, e, f tetszőleges valós számok. Akkor

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

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

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

Igy a szorzás valóban disztributív az összeadásra nézve, s ezért a $(\mathbb{C}; +, \cdot)$ algebrai struktúra test. □

Definíció 1.31 *A $(\mathbb{C}; +, \cdot)$ test elemeit komplex számoknak, a $(\mathbb{C}; +, \cdot)$ testet a komplex számok testének nevezzük.*

Tétel 1.18 *A ϕ $a \mapsto (a, 0)$ leképezés a valás számok testének a komplex számok testébe való beágyazása.*

Bizonyítás. Tetszőleges a, b valós számok esetén

$$\phi(a + b) = (a + b, 0) = (a, 0) + (b, 0) = \phi(a) + \phi(b)$$

és

$$\phi(ab) = (ab, 0) = (a, 0)(b, 0) = \phi(a)\phi(b).$$

Igy ϕ homomorfizmus. Az világos, hogy ϕ injektív. Tehát ϕ a valós számok testének a komplex számok testébe való beágyazása. □

Megjegyzés 1.19 *Bevezetve az $a = (a, 0)$, $b = (b, 0)$ és $i = (0, 1)$ jelöléseket, tetszőleges $(a, b) \in \mathbb{C}$ elem felírható*

$$(a, b) = (a, 0) + (b, 0)(0, 1) = a + bi$$

alakban, továbbá tetszőleges $a, b, c, d \in \mathbb{R}$ esetén

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

és

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Ezért a komplex számokat definiálhatjuk úgy is, mint az $a + bi$ alakú kifejezések halmazát, amelyben a és b tetszőleges valós számok, i pedig minden valós számtól különböző szimbólum. Közöttük az előző képleteknek megfelelően értelmezzük az összeadást és a szorzást.

Megjegyzés 1.20 *A komplex számokat egy síkon (az un. Gauss-féle számsíkon) ábrázolhatjuk. Rögzítsünk a síkban egy derékszögű koordinátarendszert. Nevezzük a "vízszintes" tengelyt (az $(a, 0)$ koordinátájú pontok halmazát) valós tengelynek, a függőleges tengelyt (a $(0, b)$ koordinátájú pontok halmazát) pedig képzetes tengelynek. Mindkét tengely nemnegatív felén jelöljünk ki egy-egy egységpontot (az $1 = (1, 0)$ és az $i = (0, 1)$ pontokat). Az $a + bi$ komplex számot szemléltethetjük az (a, b) végpontú helyvektorral. Megjegyezzük, hogy két komplex szám összegét szemléltető helyvektor megegyezik az egyes komplex számokat szemléltető helyvektorok összegével.*

Tétel 1.19 *(A binomiális tétel) Egységelemes kommutatív gyűrű tetszőleges a és b elemeire, illetve tetszőleges n nemnegatív egész számra*

$$(a + b)^n = \Sigma_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

1.2.2. A komplex szám trigonometriai alakja

Mint azt már korábban tárgyaltuk, a komplex számokat a Gauss-számsíkon (derékszögű koordinátarendszer rögzítés után) egy-egy helyvektorral ábrázolhatjuk. Ha egy z komplexszámnak mefelelő helyvektor hossza r, a valós tengely nemnegatív felével bezárt pozitív irányú szöge pedig φ, akkor z felírható

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

alakban is. Ezt az alakot a z komplex szám trigonometriai alakjának, r-et a z komplex szám abszolút értékének, φ-t pedig a z argumentumának nevezzük.

Minden $z \neq 0$ komplex szám egyértelműen írható trigonometriai alakban. A $z = 0$ komplex számnál az argumentum nem egyértelmű.

Tétel 1.20 *Tetszőleges $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ és $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$ komplex számok esetén*

$$z_1 z_2 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)).$$

Tétel 1.21 *Tetszőleges $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ és $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2) \neq 0$ komplex számok esetén*

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)).$$

Tétel 1.22 *(Moivre-tétel) Tetszőleges $z = r(\cos\varphi + i\sin\varphi)$ komplex szám és tetszőleges n egész szám esetén*

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

($n \leq 0$ esetén $z \neq 0$).

Definíció 1.32 *Adott pozitív n egész szám esetén egy z komplex szám n-edik komplex gyökein értjük mindazon w komplex számokat, amelyekre $w^n = z$ teljesül.*

Tétel 1.23 *Adott n pozitív egész szám esetén az alábbiak igazak. A 0 komplex számnak csak egyetlen n-dik komplex gyöke van; ez a gyök önmaga. Minden $z = r(\cos\varphi + i\sin\varphi) \neq 0$ komplex számnak n darab páronként különböző n-dik komplex gyöke van; ezeket a gyököket a*

$$\sqrt[n]{z} = \sqrt[n]{r}(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}), \quad (k = 0, \ldots, n-1)$$

képlet szolgáltatja.

Definíció 1.33 *Az 1 komplex szám n-edik komplex gyökeit n-dik komplex egységgyököknek nevezzük.*

Megjegyzés 1.21 *Adott n pozitív egész szám esetén az n-edik komplex egységgyökök (a ? Tétel szerint)*

$$z_k = cos\frac{2k\pi}{n} + isin\frac{2k\pi}{n}$$

alakúak, ahol k lehetséges értékei $k = 0, 1, \ldots n-1$. A Moivre képlet szerint

$$z_k = (cos\frac{2\pi}{n} + isin\frac{2\pi}{n})^k = z_1^k$$

teljesül minden $k = 0, 1, \ldots n-1$ indexre, azaz a z_1 komplex szám k-dik ($k = 0, 1, \ldots, n-1$) hatványai rendre megadják az összes n-dik komplex egységgyököt.

Megjegyzés 1.22 *Ha w komplex n-edik egységgyök, akkor nem csak $w^n = 1$, hanem $w^{kn} = 1$ is teljesül minden pozitív egész k-ra. Azaz, ha w n-edik komplex egységgyök, akkor nk-adik komplex egységgyök is. Például, 1 és -1 nem csak komplex 4-edik egységgyökök, hanem 8-dik, 12-dik, ... egységgyök.*

Viszont 1 *és* -1 *nem csak* $4k$*-adik egységgyök* ($k = 1, \ldots$)*, hanem megadható* 4*-nél kisebb* n *is, hogy arra az* n*-re is* n*-dik egységgyökök;* -1 *esetében ilyen a* 2*, az* 1 *esetében pedig két ilyen* n *is van, a* 2 *és az* 1*. Viszont már nincs a* 2*-nél kisebb olyan* n*, amelyre teljesülne, hogy* -1 *komplex* n*-edik egységgyök.*

Definíció 1.34 *Egy komplex* n*-edik egységgyököt primitív* n*-edik egységgyöknek nevezünk, ha nem adható meg hozzá olyan* $1 \leq m \leq n-1$ *egész szám, amelyre komplex* m*-edik egységgyök lenne. Más szavakkal, egy* z *komplex számot primitív* n*-edik egységgyöknek nevezünk, ha* n *az a legkisebb pozitív egész szám, amelyre* $z^n = 1$ *teljesül.*

Megjegyzés 1.23 *Például* -1 *primitív második egységgyök, viszont az* 1 *komplex szám második egységgyök, de nem primitív (mert az első hatványa is egyenlő* 1*-gyel).*

Tétel 1.24 *Valamely* n*-dik komplex egységgyök akkor és csak akkor primitív* n*-edik egységgyök, ha a* k*-dik* ($k = 0, 1, \ldots, n-1$) *hatványai rendre megadják az összes komplex* n*-dik egységgyököt.*

Bizonyítás.

Tétel 1.25 *Minden* n *poziív egész számhoz létezik primitív* n*-edik egységgyök.*

Bizonyítás. Adott n pozitív egész szám esetén tekintsük a

$$z_1 = cos\frac{2\pi}{n} + isin\frac{2\pi}{n}$$

komplex számot. A Megjegyzés 1.21 és Tétel 1.24 szerint z_1 primitív n-dik egységgyök. $\square$

Tétel 1.26 *Valamely* w *primitív* n*-edik egységgyök esetén a* w^k ($k = 1, \ldots n-1$) *hatvány akkor és csak akkor primitív* n*-edik egységgyök, ha* k *és* n *relatív prímek.*

Következmény 1.1 *Tetszőleges* p *prímszám esetén, minden* 1*-től különböző komplex* p*-dik egységgyök primitív.*

1.3. Polinomok

Definíció 1.35 *Egy R gyűrű feletti, x határozatlannal képezett n-edfokú (n nemnegatív egész) polinomon olyan*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

formális összeget értünk, amelyben szerelő $a_n, \ldots, a_1, a_0$ un. együtthatók az R gyűrű tetszőleges elemei, de $a_n \neq 0$.

Megjegyzés 1.24 *Az a_0 alakú polinomokat konstans polinomoknak nevezzük. Csak egyetlen polinomnak nincs fokszáma: az azonosan nulla konstanspolinomnak.*

Egy f polinom fokszámát (ha f nem a 0 konstanspolinom) f^*-gal fogjuk jelölni.

A polinomokat szokásos a fenti forma helyett az

$$a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n$$

formában is írni.

Definíció 1.36 *Ugyanazon gyűrű feletti két polinomot akkor tekintünk egyenlőnek, ha azonos indexű együtthatóik rendre megegyeznek. Megjegyezzük, hogy a konstans nullapolinom csak önmagával egyenlő.*

Az világos, hogy kölcsönösen egyértelmű megfeleltetés létesíthető egy R gyűrű feletti polinomok és az R elemeiből képezhető azon sorozatok között, amely sorozatokban csak véges sok olyan elem van, amelyik nem az R nulleleme. Egy n-edfokú

$$a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n$$

polinomnak az

$$[a_0, a_1, \ldots, a_{n-1}, a_n, 0, 0, \ldots]$$

sorozat felel meg. Igy definiálhatjuk egy gyűrű feletti polinom fogalmát a következőképpen is.

Definíció 1.37 *Egy R gyűrű elemeiből képezett n-edfokú polinomon az R elemeiből alkotott olyan $[a_0, a_1, \ldots, a_{n-1}, a_n, \ldots]$ sorozatot értünk, amelyre $a_n \neq 0$ és $a_k = 0$ ($k = n+1, n+2, \ldots$) teljesül (0 az R gyűrű nullelemét jelöli). A polinomot reprezentáló sorozat i indexű elemét a polinom i-dik együtthatójának nevezzük.*

1.3.1. Műveletek polinomok között

Definíció 1.38 *Legyenek $f = [a_i]$, illetve $g = [b_j]$ egy R gyűrű feletti tetszőleges polinomok. Az f és g polinomok $f + g$ összegén azt a polinomot értjük, melynek k-dik együtthatója egyenlő az $a_k + b_k$ összeggel.*

Tétel 1.27 *Tetszőleges gyűrű feletti polinomok halmaza a polinomok összeadására nézve kommutatív csoportot alkot; ennek neutrális eleme a nullapolinom, s egy n-edfokú $[a_i]$ polinom ellentettje az n-edfokú $[-a_i]$ polinom.*

Bizonyítás. Nyílvánvaló.

Megjegyzés 1.25 *Tetszőleges f, g polinomok esetén*

$$f + g = 0$$

vagy

$$(f + g)^* = f^* \quad (ha \quad g = 0)$$

vagy

$$(f + g)^* = g^* \quad (ha \quad f = 0)$$

vagy

$$(f + g)^* \leq max\{f^*, g^*\}.$$

Definíció 1.39 *Legyenek $f = [a_i]$ illetve $g = [b_j]$ egy R gyűrű feletti tetszőleges polinomok. A két polinom fg szorzatán azt az $fg = [c_k]$ polinomot értjük, amelyben szereplő c_k együtthatók a*

$$\sum_{\substack{0 \leq i \leq n, 0 \leq j \leq m \\ i+j=k}} a_i b_j$$

szorzatösszeggel egyenlők.

Megjegyzés 1.26 *Két polinom, f és g szorzatának fokszáma, ha létezik, legfeljebb $f^* + g^*$. Nullosztómentes gyűrű feletti polinomok esetén az fg szorzat fokszáma (ha létezik) egyenlő $f^* + g^*$-gal.*

Tétel 1.28 *Tetszőleges R gyűrű feletti polinomok halmaza a polinomok előzőekben definiált összeadására és szorzására nézve gyűrűt alkot. Ha R kommutatív, akkor a polinomgyűrű is kommutatív. Ha R egységelemes, akkor a polinomgyűrű is egységelemes.*

Bizonyítás.

Definíció 1.40 *Egy R gyűrű feletti $f = [a_i]$ polinomnak az R valamely c elemével balról való szorzatán a $cf = [ca_i]$ polinomot, jobbról való szorzatán pedig az $fc = [a_n c]$ polinomot értjük.*

Megjegyzés 1.27 *Világos, hogy egy f polinomnak egy $c \in R$ elemmel balról [jobbról] való szorzata egyenlő a $c = [c, 0, \dots]$ konstans polinomnak és az f polinomnak a cf [fc] szorzatával.*

Megjegyzés 1.28 *Legyen R egységelemes gyűrű. Jelölje x az $[0, 1, 0, \dots]$ polinomot, ahol 1 az R egységeleme. Könnyen igazolható, hogy az x polinom n-edik hatványa az a polinom, amelynek $n + 1$-edik együtthatója 1, az összes többi pedig egyenlő a gyűrű nullelemével. Könnyen ellenőrízhető, hogy tetszőleges R feletti n-edfokú $f = [a_i]$ polinom*

$$f = \ldots + a_n x^n + \cdots + a_1 x + a_0 = \ldots + x^n a_n + \cdots + x a_1 + a_0$$

alakban írható.

Megjegyzés 1.29 *Az előző megjegyzés alapján egy R gyűrű feletti $f[a_i]$ polinomot*

$$f(x) = \ldots + a_n x^n + \cdots + a_1 + a_0$$

és

$$f(x) = \ldots + x^n a_n + \cdots + x a_1 + a_0$$

formában is írhatunk. A továbbiakban, a szokásnak inkább megfelelő

$$f(x) = a_n x^n + \cdots + a_1 x + a_0$$

alakot fogjuk használni.

Egy R gyűrű feletti polinomok halmazát $R[x]$ módon fogjuk jelölni.

Megjegyzés 1.30 *A fentiek alapján érthető, hogy egy R gyűrű feletti polinomokon*

$$a_n x^n + \ldots + a_1 x + a_0$$

alakú formális kifejezéseket is szoktak érteni, ahol $a_n, \dots, a_0$ az R gyűrű elemei, az x pedig az R elemeitől különböző "szimbólum". A fentiekben ennek a szimbólumnak konkrét tartalmat adtunk.

Megjegyzés 1.31 *Ha az $a_nx^n+\ldots+a_1x+a_0$, illetve az $x^na_n+\cdots+xa_1+a_0$ kifejezések esetén az x polinom helyére az $r \in R$ elemnek megfelelő $[r]$ konstans polinomot helyettesítjük, akkor konstans polinomokat kapunk. Ha az R gyűrű r elemét azonosítjuk a konastans $[r]$ polinommal, akkor beszélhetünk az $[a_n,\ldots,a_1,a_0]$ polinomnak az R gyűrű r eleméhez tartozó bal oldali, illetve jobb oldali helyettesítési értékeiről, mint az R gyűrű azon elemeiről, amelyeknek megfelelő konstans polinomok az $[a_n,\ldots,a_1,a_0]$ polinomnak az $[r]$ konstans polinomhoz tartozó bal oldali, illetve jobb oldali helyettesítési értékeiként adódnak. Más szavakkal, az R gyűrű feletti $[a_n,\ldots,a_1,a_0]$ polinomnak valamely $r \in R$ elemhez tartozó bal oldali helyettesítési értékén az R gyűrű $r^na_n+\cdots+ra_1+a_0$ elemét, jobb oldali helyettesítési értékén pedig az $a_nr^n+\cdots+a_1r+a_0$ elemét értjük.*

1.3.2. Polinomok maradékos osztása

Definíció 1.41 *Tetszőleges R gyűrű együtthatóival képezett $a(x)$ és $b(x)$ polinomok esetén $a(x)$-nek $b(x)$-szel balról való maradékos osztásán olyan R-beli együtthatós $q(x)$ és $r(x)$ polinomok meghatározását értjük, amelyekre fennáll az $a(x)=b(x)q(x)+r(x)$ egyenlőség, ahol $r(x)=0$ vagy $r(x)$ fokszáma kisebb a $b(x)$ fokszámánál. A jobbról való maradékos osztás fogalma a balról való maradékos osztás fogalmának duálisa.*

Megjegyzés 1.32 *Természetesen kommutatív gyűrű feletti polinomok gyűrűjében a balról, illetve jobbról való maradékos osztás fogalmai egybeesnek. Ekkor csak egyszerűen, maradékos osztásról beszélünk.*

Tétel 1.29 *Ha $b(x)=b_mx^m+\ldots+b_1x+b_0$ valamely egységelemes R gyűrű feletti olyan polinom, melyben szereplő b_m együtthatónak R-ben van inverze, akkor tetszőleges R feletti $a(x)$ polinomnak a $b(x)$ polinommal mindkét oldalról való maradékos osztása egyértelműen elvégezhető.*

Bizonyítás. Csak a balról való maradékos osztással foglalkozunk. A jobb oldali eset bizonyítása hasonló.

Ha $a(x)$ a nullapolinom vagy fokszáma kisebb, mint $b(x)$ fokszáma, akkor a $q(x)=0$ és $r(x)=a(x)$ választással fennáll az $a(x)=b(x)q(x)+r(x)$ egyenlőség és $r(x)=0$ (ha $a(x)=0$) vagy $r(x)$ fokszáma kisebb $b(x)$ fokszámánál (ha $a(x)\neq 0$).

Vizsgáljuk most azt az esetet, amikor az $a(x)$ polinom fokszáma nagyobb vagy egyenlő, mint $b(x)$ fokszáma. Megmutatjuk, hogy ekkor vannak olyan R

feletti $q_1(x)$ és $r_1(x)$ polinomok, amelyekre $a(x) = b(x)q_1(x) + r_1(x)$ teljesül, ahol $r_1(x)$ a nullapolinom vagy fokszáma kisebb az $a(x)$ fokszámánál. Ha az $a(x)$ polinom fokszáma n és a főegyütthatója a_n, akkor legyen

$$r_1(x) = a(x) - b(x)((b_m)^{-1}a_n x^{n-m}).$$

Az világos, hogy $r_1(x)$ vagy a nullapolinom vagy fokszáma kisebb az $a(x)$ fokszámánál. A

$$q_1(x) = b_m^{-1}a_n x^{n-m}$$

jelölést használva,

$$a(x) = b(x)q_1(x) + r_1(x)$$

teljesül, és $r_1(x)$ vagy a nullapolinom vagy fokszáma kisebb az $a(x)$ fokszámánál. Ha $r_1(x)$ a nullapolinom, akkor készen vagyunk a bizonyítással. Ellenkező esetben alkalmazzuk az előző gondolatmenetet $a(x)$ helyett $r_1(x)$-re: vannak olyan R feletti $q_2(x)$ és $r_2(x)$ polinomok, hogy

$$r_1(x) = b(x)q_2(x) + r_2(x)$$

teljesül, és $r_2(x)$ vagy a nullapolinom vagy fokszáma kisebb az $r_1(x)$ fokszámánál. Ekkor

$$a(x) = b(x)(q_1(x) + q_2(x)) + r_2(x).$$

Folytatva ezt az eljárást, megadható olyan j pozitív egész szám, hogy a j-dik lépésben megjelenő $r_j(x)$ polinom vagy a nullapolinom vagy a fokszáma kisebb (nem csak ar $r_{j-1}(x)$ polinom fokszámánál, hanem) a $b(x)$ polinom fokszámánál, továbbá

$$a(x) = b(x)(q_1(x) + \cdots + q_j(x)) + r_j(x).$$

A $q(x) = q_1(x) + \cdots + q_j(x)$ és $r(x) = r_j(x)$ jelölést használva kapjuk az

$$a(x) = b(x)q(x) + r(x)$$

egyenlőséget, amelyben szereplő $r(x)$ polinom vagy a nullapolinom vagy fokszáma kisebb a $b(x)$ polinom fokszámánál. Ezzel bebizonyítottuk, hogy $a(x)$-nek $b(x)$-szel balról való maradékos osztása elvégezhető. Megjegyezzük, hogy a jobbról való maradékos osztás elvégezhetőségének bizonyításakor a $q_1(x)$ polinom együtthatója (a baloldali esetben szereplő $b_m^{-1}a_n$ helyett) $a_n b_m^{-1}$; általában a $q_i(x)$ polinom együtthatójában b_m^{-1} jobboldali tényezőként jelenik meg.

A bizonyítás hátralévő részében az egyértelműséget bizonyítjuk. Tegyük fel, hogy valamely $a(x), b(x) \in R[x]$, $b(x) \neq 0$ polinomok esetén teljesülnek az

$$a(x) = b(x)q_1(x) + r_1(x)$$

és az

$$a(x) = b(x)q_2(x) + r_2(x)$$

egyenlőségek valamely $q_1(x), q_2(x), r_1(x), r_2(x) \in R[x]$ polinomokra, ahol $r_i(x) = 0$ vagy $r_i(x)$ fokszáma kisebb a $b(x)$ fokszámánál ($i = 1, 2$). Kivonva a második egyenlőséget az elsőből, az egyenlet rendezése után kapjuk az

$$b(x)(q_2(x) - q_1(x)) = r_1(x) - r_2(x)$$

egyenlőséget. A feltétel szerint $r_1(x) - r_2(x)$ egyenlő a nulla polinommal vagy fokszáma kisebb a $b(x)$ fokszámánál. A $q_2(x) - q_1(x)$ polinom vagy a nulla polinom vagy a $b(x)(q_2(x) - q_1(x))$ polinom fokszáma nem kisebb a $b(x)$ polinom fokszámánál. Ugyanis a második esetben, a $b(x)(q_2(x) - q_1(x))$ polinom főegyütthatója a $b(x)$, illetve a $q_2(x) - q_1(x)$ polinomok főegyütthatóinak szorzata, amely $q_2(x) - q_1(x) \neq 0$ miatt nem lehet nulla, mert $b(x)$ főegyütthatója invertálható. A fenti egyenlőség miatt ez csak akkor teljesülhet, ha az $r_1(x) - r_2(x)$ polinom is és a $q_2(x) - q_1(x)$ polinom is a nulla polinom. Ebből pedig $r_1(x) = r_2(x)$ és $q_1(x) = q_2(x)$ adódik, ami az egyértelműséget bizonyítja. □

Tétel 1.30 *Adott $\mathbb{F}$ test feletti tetszőleges $f(x) \in \mathbb{F}[x]$ polinomnak tetszőlegs $0 \neq g(x) \in \mathbb{F}[x]$ polinommal való maradékos osztása mindig elvégezhető, mégpedig egyértelműen.*

Bizonyítás. Az előző tétel miatt nyílvánvaló.

1.3.3. Polinomok zérushelyei

Definíció 1.42 *Egy $\mathbb{F}$ test feletti*

$$f(x) = a_n x^n + \cdots + a_1 x + a_0$$

polinom valamely $c \in \mathbb{F}$ elemhez tartozó helyettesítési értékén az $\mathbb{F}$ test

$$f(c) = a_n c^n + \cdots + a_1 c + a_0$$

elemét értjük. Ha $f(c)$ az $\mathbb{F}$ test nulleleme, akkor c-t az $f(x)$ polinom egy zérushelyének nevezzük.

Tétel 1.31 *Egy $\mathbb{F}$ test valamely c eleme esetén egy $f(x) \in \mathbb{F}[x]$ polinomnak az elsőfokú $g(x) = x - c$ polinommal való maradékos osztásánál maradékként az $f(c)$ helyettesítési érték adódik.*

Bizonyítás. A Tétel 1.29 miatt

$$f(x) = q(x)(x - c) + r(x)$$

teljesül valamely $\mathbb{F}$ feletti $q(x)$ és $r(x)$ polinomokra, ahol $r(x)$ vagy a nulla polinom, vagy az $r(x)$ fokszáma kisebb a $g(x) = x - c$ polinom fokszámánál. Mindkét esetben az adódik, hogy $r(x)$ egy konstans polinom; mondjuk $r(x) = r_0$. A fenti $f(x) = q(x)(x - c) + r(x)$ egyenlőségből

$$f(c) = q(c)(c - c) + r_0,$$

azaz

$$f(c) = r_0$$

adódik. □

Megjegyzés 1.33 *Tetszőleges F test feletti egyhatározatlanú polinomok $F[x]$ gyűrűje egységelemes integritástartomány. Igy $F[x]$-ben értelmezve van az oszthatóság fogalma. Az egységek a nem-nulla konstans polinomok, továbbá két polinom akkor és csak akkor asszociált, ha egymás nem-nulla konstans-szorosai.*

Tétel 1.32 *Egy $\mathbb{F}$ test valamely c eleme akkor és csak akkor zérushelye egy $f(x) \in \mathbb{F}[x]$ polinomnak, ha $f(x)$ osztható a $g(x) = x - c$ polinommal, azaz az $f(x)$ polinomnak a $g(x) = x - c$ polinommal való maradékos osztásánál maradḱként a nulla polinom adódik.*

Bizonyítás. Az előző tétel és a polinomok oszthatóságának definíciója alapján nyilvánvaló. □

Megjegyzés 1.34 *A polinomok maradékos osztásának módszerét alkalmazva, könnyen lehet ellenőrizni, hogy egy $\mathbb{F}$ test feletti legalább elsőfokú*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad (a_n \neq 0)$$

polinomnak egy $g(x) = x - c$ $(c \in F)$ polinommal való maradékos osztásánál fellépő

$$q(x) = b_{n-1} x^{n-1} + \cdots b_1 x + b_0$$

hányados polinom együtthatóira

$$b_{n-1} = a_n$$

és

$$b_{n-j} = cb_{n-j+1} + a_{n-j+1}, \quad (j = 2, \ldots, n),$$

a maradékra pedig

$$f(c) = cb_0 + a_0$$

teljesül. A fenti tényt a következő táblázatos elrendezéssel lehet szemlélteni.

a_n	$\ldots$	a_{n-j+1}	$\ldots$	a_1	a_0
$b_{n-1} = a_n$	$\ldots$	$b_{n-j} = cb_{n-j+1} + a_{n-j+1}$	$\ldots$	$b_0 = cb_1 + a_1$	$f(c) = cb_0 + a_0$

A táblázat második sorát úgy is kitölthetjük, hogy annak első oszlopába a_n-et írunk, s az i-dik ($i = 2, \ldots, n+1$) oszlopába pedig azt az elemet, amit úgy kaphatunk meg, hogy a második sor $i-1$-dik oszlopában álló elem c-szereséhez hozzáadjuk az i-dik oszlop első sorában álló elemet.

Egy $c \in \mathbb{F}$ elem akkor és csak akkor zérushelye az $f(x)$ polinomnak, ha a c felhasználásával elkészített fenti táblázat utolsó oszlopának második sorában álló elem nulla. Ekkor az $f(x)$ polinomnak a $g(x) = x - c$ polinommal való maradékos osztásánál fellépő hányados polinom együtthatóit rendre a táblázat második sorának első, második, ..., utolsó előtti oszlopában megjelenő elemek adják. Ezzel egy aránylag egyszerű módszert kaptunk annak eldöntésére, hogy egy $\mathbb{F}$ test valamely c eleme zérushelye-e egy legalább elsőfokú $f(x) \in \mathbb{F}[x]$ polinomnak; ha igen, akkor ez a módszer megadja az $f(x)$ polinomnak a $g(x) = x - c$ polinommal való maradékos osztásánál fellépő hányados polinom együtthatóit is. Az előzőekben ismertetett módszert Horner-módszernek szokták nevezni.

1.3.4. Polinomok legnagyobb közös osztója

1.3.5. Irreducibilis polinomok

Megjegyzés 1.35 *A Definíció* **??** *alapján világos, hogy egy F test feletti polinom (azaz az $F[x]$ egységelemes integritási tartomány valamely eleme)akkor és csak akkor irreducibilis, ha legalább elsőfokú és nem bontható fel két, nálánál kisebb fokszámú F feletti polinom szorzatára.*

A Tétel **??** eredményét alkalmazhatjuk egy $\mathbb{F}$ test feletti polinomok $\mathbb{F}[x]$ polinomgyűrűjére, mint speciális egységelemes integritástartományra. Igy kapjuk a következő tételt.

Tétel 1.33 *Tetszőleges F test feletti legalább elsőfokú polinomok mindegyike felbontható véges sok irreducibilis F feletti polinom szorzatára. Ez a felbontás a tényezők sorredjétől, illetve asszociálttól eltekintve egyértelmű. Ez utóbbi feltétel azt jelenti, hogy ha $f(x) = p_1(x) \cdots p_{(}x)$ és $f(x) = q_1(x) \cdots q_s(x)$ egy $f(x) \in F[x]$ polinomnak két felbontása F feletti irreducibilis polinomok szorzatára, akkor $r = s$ és a $p_i(x)$ $(i = 1, \ldots, r)$ polinomok halmaza, valamint a $q_i(x)$ $(i = 1, \ldots, r)$ polinomok halmaza között van olyan bijektív megfeleltetés, hogy az egymásnak megfeleltetett polinomok asszociáltak (azaz legfeljebb egységfaktorban különböznek egymástól).*

Bizonyítás.

A következőkben a polinomok irreducibilitását vizsgáljuk speciális testek, nevezetesen a komplex számok teste, a valós számok teste, valamint a racionális számok teste feletti polinomgyűrűben.

Tétel 1.34 *Egy komplex együtthatós polinom akkor és csak akkor irreducibilis, ha elsőfokú*

Bizonyítás. Az algebra alaptétele miatt nyilvánvaló. □

Tétel 1.35 *Minden komplex együtthatós, legalább elsőfokú*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \quad (a_n \neq 0)$$

polinomhoz megadhatók egyértelműen olyan páronként különböző $c_1, \ldots c_r$ komplex számok és olyan $k_1, \ldots, k_r$ pozitív egész számok, hogy $k_1 + \cdots + k_r = n$ és

$$f(x) = a_n(x - c_1)^{k_1} \cdots (x - c_r)^{k_r}.$$

Bizonyítás A Tételthirredfelbontas és Tétel 1.34 alapján nyilvánvaló. □

Tétel 1.36 *Ha egy c komplex szám zérushelye a valós együtthatójú*

$$f(x) = a_n x^n + \cdots + a_1 x + a_0$$

polinomnak. akkor c konjugáltja is zérushelye $f(x)$-nek.

Bizonyítás. Legyen c zérushelye az $f(x)$ polinomnak. Akkor

$$0 = \overline{0} = \overline{f(c)} = \overline{a_n c^n + \cdots + a_1 c + a_0} =$$

$$= \overline{a_n c^n} + \cdots + \overline{a_1 c} + \overline{a_0} = a_n \overline{c}^n + \cdots + a_1 \overline{c} + a_0 = f(\overline{c}),$$

felhasználva azt a tényt, hogy összeg konjugáltja egyenlő a tagok konjugáltjainak összegével, szorzat konjugáltja egyenlő a tényezők konjugáltjainak szorzatával, valamint azt a tényt, hogy a valós számok konjugáltja önmaga. Tehát c konjugáltja is zérushelye az $f(x)$ polinomnak. □

Tétel 1.37 *Minden valós együtthatós, legalább elsőfokú*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad (a_n \neq 0)$$

polinomhoz felírható elsőfokú valós együtthatójú vagy negatív diszkriminánsú másodfokú valós együtthatójú polinomok szorzatára, ez a felbontás a tényezők sorrendjétől és konstrans szorzótól eltekintve egyértelmű.

Bizonyítás. Legyen

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad (a_n \neq 0)$$

tetszőleges valós együtthatós, legalább elsőfokú polinom. A Tétel 1.35 szerint megadhatók egyértelműen olyan páronként különböző $c_1, \ldots c_r$ komplex számok és olyan $k_1, \ldots, k_r$ pozitív egész számok, hogy $k_1 + \cdots + k_r = n$ és

$$f(x) = a_n (x - c_1)^{k_1} \cdots (x - c_r)^{k_r}.$$

Ha $c_1, \ldots, c_r$ mindegyike valós szám, akkor nincs mit bizonyítani. Tegyük fel, hogy valamelyikük, például c_1, nem valós. A Téte 1.36 szerint van közöttük olyan, a c_1-től különböző szám, például a c_2, hogy $c_2 = \overline{c}_1$. Legyen c_1 algebrai alakja $c_1 = a + bi$. Akkor

$$g(x) = (x - c_1)(x - c_2) = (x - c_1)(x - \overline{c}_1) = (x - a - bi)(x - a + ib) = (x - a)^2 + b^2 = x^2 - 2ax + a^2 + b + 2$$

amely valós együtthatós polinom diszkriminánsa

$$D = 4a^2 - 4(a^2 + b^2) - 4b^2 < 0.$$

Tegyük fel, hogy $k_1 < k_2$. A valós együtthatós

$$f(x) = g^{k_1}(x)(x - c_2) \cdots (k - c_r)^{k_r}$$

polinomnak a valós együtthatós $g^{k_1}(x)$ polinommal való maradékos osztása $\mathbb{R}$ felett egyértelműen elvégezhető, azaz vannak olyan valós együtthatós $q(x)$ és $r(x)$ polinomok, hogy $f(x) = g^{k_1}(x)q(x) + r(x)$. Ez az előállítás $\mathbb{C}$ felett is érvényes, hiszen $f(x), g(x), q(x), r(x) \in \mathbb{C}[x]$. Viszont $\mathbb{C}$ felett érvényes az

$$f(x) = g^{k_1}(x)(x - c_2)\cdots(k - c_r)^{k_r}$$

előállítás is. Igy

$$q(x) = (x - c_2)\cdots(k - c_r)^{k_r}$$

és

$$r(x) = 0.$$

Ez viszont azt jelenti, hogy c_2 zérushelye a valós együtthatós $q(x)$ polinomnak, viszont c_2 konjugáltja, azaz c_1 nem. Ez ellentmondás. Igy $k_1 < k_2$ nem lehetséges. Hasonlóana igazolható, hogy a $k_1 > k_2$ feltétel is ellentmondásra vezet. Tehát $k_1 = k_2$. Igy

$$f(x) = g^{k_1}(x)h(x),$$

ahol

$$h(x) = (x - c_2)\cdots(k - c_r)^{k_r}$$

valós együtthatós polinom. Ha $h(x)$ konstans polinom, akkor készen vagyunk a bizonyítással. Ellenkező esetben a fenti gondolatmenetet $f(x)$ helyett a nála kisebb fokszámú $h(x)$ polinomra alkalmazhatjuk. Folytatva a gondolatmenetet, véges sok lépés után eljutunk a tétel állításához, azaz $f(x)$ felírható valós együtthatós elsőfokú vagy negat́iv diszkriminánsú másodfokú polinomok szorzatára.

Az egyértelműség a Tétel 1.33 folyamánya. □

Tétel 1.38 *Egy valós együthatós polinom akkor és csak akkor irreducibilis, ha elsőfokú vagy negatív diszkriminánsú másodfokú.*

Bizonyítás. A Tétel 1.37 felhasználásával nyilvánvaló. □

Definíció 1.43 *Egy egész együtthatós polinomot primitív polinomnak nevezünk, ha együthatói relatív prímek.*

Tétel 1.39 *Minden racionális együtthatós $f(x)$ polinom előjeltől eltekintve egyértelműen felírható*

$$f(x) = \frac{a}{b}g(x)$$

alakban, ahol a és b relatív prím egész számokat, $g(x)$ pedig primitív polinomo.

Bizonyítás. Jelölje b_0 az $f(x)$ polinom együtthatóiban szereplő nevezők legkisebb közös többszörösét. Ekkor az $f(x)$ polinom

$$f(x) = \frac{1}{b_0} f_0(x)$$

alakban írható, ahol $f_0(x)$ egész együtthatós polinom. Jelölje a_0 az $f_0(x)$ polinom együtthatóinak legnagyobb közös osztóját. Ha az $\frac{a_0}{b_0}$ tört tovább már nem egyszerűsíthető alakja $\frac{a}{b}$, akkor

$$f(x) = \frac{a}{b} g(x),$$

ahol $g(x)$ az a polinom, amelyre $a_0 g(x) = f_0(x)$ teljesül. Az világos, hogy $g(x)$ primitív polinom.

Az egyértelműség bizonyításához tegyük fel, hogy

$$\frac{a}{b} g(x) = f(x) = \frac{c}{d} h(x)$$

teljesül, ahol $(a,b) = 1$, $(c,d) = 1$, valamint a $g(x)$ és a $h(x)$ polinomok primitívek. Akkor

$$adg(x) = bch(x).$$

Mivel a osztója a bal oldalnak, ezért osztója a jobb oldalnak is. Mivel $h(x)$ együtthatói relatív prímek és $(a,b) = 1$, ezért $a = 1$ vagy a osztja c-t. Igy

$$dg(x) = bc'h(x)$$

adódik, ahol $ac' = c$. Hasonlóan, d osztja b-t, s így

$$g(x) = b'c'h(x)$$

, ahol $db' = b$. Mivel $g(x)$ primitív polinom, ezért a vele egyenlő $b'c'h(x)$ polinom is az, viszont $b'c'$ osztója a $b'c'h(x)$ polinom minden együtthatójának. Ezért $b'c' = \pm 1$, azaz $b' = \pm 1$ és $c' = \pm 1$. Igy $a = \pm c$ és $b = \pm d$. Tehát $\frac{a}{b} = \pm \frac{c}{d}$. Ebből már következik, hogy $h(x) = \pm g(x)$. □

Tétel 1.40 *(Gauss-lemma) Primitív polinomok szorzata is primitív.*

Tétel 1.41 *Egy egész együthatós polinom akkor és csak akkor irreducibilis $\mathbb{Q}$ felett, ha irreducibilis $\mathbb{Z}$ felett.*

Tétel 1.42 *(Schönemann-Eisenstein irreducibilitási kritérium) Ha az egész együtthatós, legalább elsőfokú*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

polinomhoz lehet találni olyan p prímszámot, hogy p az együtthatók közül nem osztja a_n-et, de osztja az összes többit, viszont p^2 nem osztója a_0-nak, akkor $f(x)$ irreducibilis $\mathbb{Q}$ (és persze $\mathbb{Z}$) felett.

1.3.6. Körosztási polinom

Definíció 1.44 *Legyenek $\xi_1, \ldots, \xi_{\varphi(n)}$ az n-edik primitív egységgyökök (itt φ az un. Euler-függvény). A*

$$\Phi_n(x) = \Pi_{i=1}^{\varphi(n)}(x - \xi_i)$$

polinomot n-edik körosztási polinomnak nevezzük.

Megjegyzés 1.36 *Az első néhány példa:*

$$\Phi_1(x) = x - 1,$$

$$\Phi_2(x) = x + 1,$$

$$\Phi_3(x) = x^2 + x + 1,$$

$$\Phi_4(x) = x^2 + 1.$$

Tétel 1.43 *Az n-dik körosztási polinom minden n-re egész együtthatós, és irreducibilis a racionális számok teste felett.*

1.3.7. Polinominterpoláció

1.3.8. A Cardano-képlet

Vizsgáljuk az

$$f(z) = a_3 z^3 + a_2 z^2 + a_1 z + a_0 \quad (a_3 \neq 0)$$

komplex együtthatós polinom zérushelyeit. A Tétel 1.35 szerint ennek a polinomnak 3 komplex gyöke van (ha a gyökök multiplicitását is figyelembe vesszük). A fenti polinom helyett vizsgálhatjuk az

$$\frac{1}{a_3} f(z)$$

polinom zérushelyeit, azaz az

$$y^3 + ay^2 + by + c = 0$$

alakú egyenletek gyökeit. A cél az, hogy megoldóképletet adjunk a gyökök előállítására. Könnyen ellenőrízhető, hogy az

$$y = x - \frac{a}{3}$$

helyettesïtéssel

$$x^3 + px + q = 0$$

alakú egyenletet kapunk. Elég csak ezen utóbbi egyenlethez megoldóképletet találnunk, ugyanis ezen egyenelet x_i $(i = 1, 2, 3)$ megoldásainak ismeretében az $y_i = x_i - \frac{a}{3}$ komplex számok az $y^3 + ay^2 + by + c = 0$ egyenlet megoldásai.

Legyen x_0 az

$$x^3 + px + q = 0$$

egyeneletnek egyik gyöke. Tekintsük az

$$f(u) = u^2 - x_0 u - \frac{p}{3}$$

polinomot. Jelölje α és β az f polinom két komplex gyökét. Viéte képletei szerint

$$\alpha + \beta = x_0$$

és

$$\alpha\beta = -\frac{p}{3}.$$

Mivel x_0 gyöke ax $x^3 + px + q = 0$ egyenletnek, ezért

$$0 = x_0^3 + px_0 + q = (\alpha + \beta)^3 + p(\alpha + \beta) + q =$$

$$= \alpha^3 + \beta^3 + (3\alpha\beta + p)(\alpha + \beta) + q = \alpha^3 + \beta^3 + q,$$

azaz

$$\alpha^3 + \beta^3 = -q.$$

Ez és a fenti $\alpha\beta = -\frac{p}{3}$ összefüggésből adódó

$$\alpha^3\beta^3 = -\frac{p^3}{27}$$

egyenlőség azt mutatja, hogy α^3 és β^3 gyöke a

$$z^2 + qz - \frac{p^3}{27} = 0$$

komplex együtthatós másodfokú egyenletnek. Ha alkalmazzuk a másodfokú egyenletekre érvényes megoldóképletet, akkor azt kapjuk, hogy

$$x_0 = \alpha + \beta = \sqrt[3]{-\frac{q}{2} + \sqrt[2]{\frac{q}{2}^2 + \frac{p}{3}^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt[2]{\frac{q}{2}^2 + \frac{p}{3}^3}}.$$

Legyen α_1 a

$$\sqrt[3]{-\frac{q}{2} + \sqrt[2]{\frac{q}{2}^2 + \frac{p}{3}^3}}$$

komplex gyök egyik lehetséges értéke. Akkor a mások két gyök kifejezhető

$$\alpha_2 = \epsilon\alpha_1, \quad \alpha_3 = \epsilon^2\alpha_1$$

formában, ahol ϵ egy harmadik egységgyök. Ha β_1 jelöli az

$$\sqrt[3]{-\frac{q}{2} - \sqrt[2]{\frac{q}{2}^2 + \frac{p}{3}^3}}$$

harmadik gyök azon értékét, amelyre

$$\alpha_1\beta_1 = -\frac{p}{3}$$

teljesül. Mivel a másik két gyökre $\beta_2 = \epsilon\beta_1$ és $\beta_3 = \epsilon^2\beta_1$ teljesül, ezért

$$\alpha_2\beta_3 = \epsilon\alpha_1\epsilon^2\beta_1 = \alpha_1\beta_1 = -\frac{p}{3}.$$

Tehát az α gyökkifejezés α_2 értékének a β gyökkifejezés β_3 értéke felel meg. Hasonlóképpen az α_3 értéknek a β_2 érték felel meg. Igy a vizsgát $x^3+px+q = 0$ harmadfokú egyenlet három gyöke felírható a következőképpen¿

$$\begin{array}{rcll} x_1 & = & \alpha_1 & +\beta_1 \\ x_2 & = & \epsilon\alpha_1 & +\epsilon^2\beta_1 \\ x_3 & = & \epsilon^2\alpha_1 & +\epsilon\beta_1 \end{array}.$$

1.3.9. Többváltozós polinomok

Definíció 1.45 *Egy R kommutatív gyűrű feletti n-változós polinomon véges sok tagot tartalmazó olyan formális összeget értünk, melynek tagjai*

$$ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

alakúak, amelyekben az $x_1, x_2, \ldots, x_n$ szimbólumok jelölik a változókat, az a együttható az R gyűrű tetszőleges eleme, a $[k_1, k_2, \ldots, k_n]$ kitevősorozat elemei pedig tetszőleges nemnegatív egész számok. Továbbá az is teljesül, hogy különböző tagokhoz különböző $[k_1, k_2, \ldots, k_n]$ kitevősorozatok tartoznak.

Definíció 1.46 *Két n-változós polinomot egyenlőnek tekintünk, ha tagjai között létesíthető olyan bijekció, hogy az egymásnak megfeleltetett tagok egymással egyenlőek, azaz együtthatóik és kitevősorozataik megegyeznek.*

A definíció szerint például a kétváltozós $f(x_1, x_2) = 3x_1^5x_2$ és $g(x_1, x_2) = 3x_2x_1^5$ polinomok egyenlőek, mert mindkét polinom egy-egy tagból áll, mindkét tag együtthatója 3, kitevősorozata pedig $[5, 1]$.

Definíció 1.47 *Egy n-változós polinom $ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ $(a \neq 0)$ tagjának fokszámán a $k_1 + \cdots + k_n$ összeget értjük. Egy n-változós polinom fokszámán tagjai fokszámának maximumát értjük.*

Definíció 1.48 *Két n-változós polinom összegén azt az n-változós polinomot értjük, melynek tagjai az összeadandó polinomok tagjaiból áll azzal a kiegészítéssel, hogy az azonos kitevősorozattal rendelkező tagokat összevonjuk, azaz, ha mindkét polinom tartalmazott ugyanazon $[k_1, k_2, \ldots, k_n]$ kitevősorozattal rendelkező a, illetve b együtthatójú tagot, akkor az összegben a két tag együtt,*

$$(a + b)x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

formában jelenik meg.

Definíció 1.49 *Az $ax_1^{k_1}\ldots x_n^{k_n}$ és $bx_1^{t_1}\ldots x_n^{t_n}$ egytagú n-változós polinomok szorzatán az*

$$abx_1^{k_1+t_1}\ldots x_n^{k_n+t_n}$$

egytagú n-változós polinomot értjük. Két tetszőleges n-változós $f(x_1, \ldots, x_n)$ és $g(x_1, \ldots, x_n)$ polinom szorzatán azt az n-változós polinomot értjük, amelyet úgy is megkaphatunk, hogy f minden tagját megszorozzuk g minden tagjával (mint egytagú polinomokat), majd az azonos kitevősorozattal rendelkező tagokat összevonjuk (a fentebb már részletezett formában).

Tétel 1.44 *Egy R kommutatív gyűrű feletti n-változós polinomok $R[x_1, x_2, \ldots, x_n]$ halmaza a polinomok összeadására és szorzására nézve kommutatív gyűrűt alkot.*

Definíció 1.50 *Egy többváltozós $f(x_1, \ldots, x_n)$ polinomot szimmetrikus polinomnak nevezünk, ha tetszőleges n-edfukú σ permutáció esetén*

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = f(x_1, \ldots, x_n),$$

azaz az f polinom nem változik akárhogy is permutáljuk benne a változókat.

Például, a kétváltozós $f(x_1, x_2) = x_1^3 x_2 + x_1 x_2^3$ polinom szimmetrikus, mert az x_1 és x_2 változóknak egyetlen nem-identikus permutációja van (ez az x_2, x_1) és

$$f(x_2, x_1) = x_2^3 x_1 + x_2 x_1^3 = x_1^3 x_2 + x_1 x_2^3 = f(x_1, x_2).$$

Tétel 1.45 *Az $R[x_1, \ldots, x_n]$ polinomgyűrű szimmetrikus polinomjai az $R[x_1, \ldots, x_n]$ gyűrű egy részgyűrűjét alkotják.*

Definíció 1.51 *n-változós elemi szimmetrikus polinomokon a*

$$\sigma_1 = x_1 + \cdots x_n,$$

$$\sigma_2 = x_1 x_2 + \cdots + x_1 x_n + x_2 x_3 + \cdots + x_2 x_n + \cdots + x_{n-1} x_n,$$

$$\vdots$$

$$\sigma_n = x_1 x_2 \cdots x_n$$

polinomokat értjük (lásd a Viete-képleteket).

Tétel 1.46 *(A szimmetrikus polinomok alaptétele) Tetszőleges F test feletti n-változós szimmetrikus polinomok mindegyike előáll az F feletti n-változós elemi szimmetrikus polinomok F-beli együtthatókkal képezett polinomjaként.*

2. fejezet

A vektortér

A vektor szó nem ismeretlen az olvasó számára. A síkbeli vektor fogalmával középiskolában már megismerkedtek. Szó volt a síkbeli vektorok összegéről és valós számokkal (azaz a valós számok testéhez tartozó elemekkel) való szorzásáról. Ismert tény, hogy a síkbeli vektorok halmazán definiált összeadás asszociatív, kommutatív és invertálható (azaz a síkbeli vektorok a vektorok összeadására nézve kommutatív csoportot alkotnak), továbbá tetszőleges $\underline{a}$ és $\underline{b}$ síkbeli vektorokra, valamint tetszőleges α, β valós számokra teljesülnek az alábbiak

$$\alpha(\underline{a}+\underline{b}) = \alpha\underline{a}+\alpha\underline{b},$$

$$(\alpha+\beta)\underline{a} = \alpha\underline{a}+\beta\underline{a},$$

$$(\alpha\beta)\underline{a} = \alpha(\beta\underline{a}) = \beta(\alpha\underline{a})$$

$$1\underline{a} = \underline{a}.$$

A matematikában számos egyéb helyen is találkozunk olyan konstrukcióval, ahol egy $(V;+)$ kommutatív csoport és egy $(F;+,\cdot)$ test elemei között értelezve van egy szorzás (skalárral való szorzás) úgy, hogy tetszőleges $\underline{a},\underline{b} \in V$ és $\alpha,\beta \in F$ elemekre teljesül a fenti négy feltétel (a negyedikben 1 az F test egységelemét jelöli). Ebben a fejezetben ilyen konstrukciókkal foglalkozunk.

2.1. A vektortér fogalma

Definíció 2.1 *Legyen $(V;+)$ egy kommutatív csoport és $(F;+,\cdot)$ egy test. Azt mondjuk, hogy V vektorteret alkot F felett, ha tetszőleges $(\alpha,a) \in F\times V$ elempárhoz hozzá van rendelve egyértelműen a V egy αa módon jelölt eleme* *úgy, hogy tetszőleges $\alpha, \beta \in F$ és $a, b \in V$ elemek esetén az alábbi egyenlőségek teljesülnek:*

$$\alpha(a+b) = \alpha a + \alpha b,$$

$$(\alpha+\beta)a = \alpha a + \beta a,$$

$$(\alpha\beta)a = \alpha(\beta a) = \beta(\alpha a),$$

$$1a = a,$$

ahol 1 *az* F *test egységelemét jelöli.*

Tétel 2.1 *Egy F test feletti V vektortérben $\alpha a = 0$ akkor és csak akkor teljesül, ha $a = 0$ vagy $\alpha = 0$.*

Bizonyítás. Először azt mutatjuk meg, hogy ha $a = 0$ vagy $\alpha = 0$, akkor $\alpha a = 0$.

Legyen $a \in V$ tetszőleges vektor. Akkor

$$0a = (0+0)a = 0a + 0a,$$

amiből

$$0a = 0$$

következik.

Legyen $\alpha \in F$ tetszőleges skalár. Akkor

$$\alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0,$$

amiből

$$\alpha 0 = 0$$

következik.

A fordított állítás bizonyításához tegyük fel, hogy $\alpha a = 0$ teljesül valamely $\alpha \in F$ skalárra és $a \in V$ vektorra. Elegendő megmutatni, hogy ha $\alpha \neq 0$, akkor $a = 0$. Tegyük fel, hogy $\alpha \neq 0$. Akkor

$$a = 1a = (\frac{1}{\alpha}\alpha)a = \frac{1}{\alpha}(\alpha a) = \frac{1}{\alpha}0 = 0,$$

felhasználva a vektortér definíciójában szereplő negyedik azonosságot és a fentebb már bizonyított $0a = 0$ egyenlőséget. □

Definíció 2.2 *Egy* F *test feletti* V *vektortér valamely* $W \neq \emptyset$ *részhalmazát a* V *vektortér egy alterének nevezzük ha* W *a* V *beli műveletre és a* V *elemeinek az* F*-beli skalárokkal való szorzására nézve vektorteret alkot* F *felett.*

Tétel 2.2 *Egy* F *test feletti* V *vektortér valamely* $W \neq \emptyset$ *részhalmaza akkor és csak akkor altere a* V *vektortérnek, ha minden* $\underline{a}, \underline{b} \in W$ *és tetszőleges* $\alpha, \beta \in F$ *skalárokra*

$$\alpha\underline{a} + \beta\underline{b} \in W$$

teljesül.

2.2. Lineáris kombináció, lineáris függetlenség

Definíció 2.3 *Egy F test feletti V vektortér*

$$a_1, \ldots, a_n$$

vektorainak az

$$\alpha_1, \ldots, \alpha_n \in F$$

skalárokkal képezett lineáris kombinációján az

$$\alpha_1 a_1 + \ldots + \alpha_n a_n \in V$$

vektort értjük. Akkor mondjuk, hogy egy $b \in V$ vektor előáll az

$$a_1, \ldots, a_n \in V$$

vektorok lineáris kombinációjaként, ha megadhatók olyan

$$\alpha_1, \ldots, \alpha_n \in F$$

skalárok, amelyekkel teljesül a

$$b = \alpha_1 a_1 + \ldots + \alpha_n a_n$$

egyenlőség.

Megjegyzés 2.1 *A Tétel 2.1 alapján tetszőleges $a_1, \ldots, a_n$ vektorok eseteén*

$$0a_1 + \ldots + 0a_n = 0,$$

azaz a 0 skalárokkal képezett lineáris kombináció eredményeként a nullvektort kapjuk. A vektorterek elméletében fontos szerepet játszanak azok a vektorok, amelyeknél nincs is más lineáris kombináció, amely eredményeként a nullvektor adódik.

Definíció 2.4 *Egy F test feletti V vektortér*

$$a_1, \ldots, a_n$$

vektorairól azt mondjuk, hogy lineárisan függetlenek, ha az

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$$

egyenlőség akkor és csak akkor teljesül, ha

$$\alpha_1 = 0, \ldots, \alpha_n = 0.$$

Ellenkező esetben azt mondjuk, hogy az $\{a_1, \ldots, a_n\}$ *vektorrendszer lineárisan függő. Ez utóbbi annyit jelent, hogy megadhatók olyan* $\alpha_1, \ldots, \alpha_n \in F$ *skalárok, amelyek között van legalább egy nem-nulla, s amelyekre fennáll az*

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$$

egyenlőség.

Megjegyzés 2.2 *Lineárisan független vektorrendszer bármely nem-üres részredszere is lineárisan független, mert ha az* $\{a_1, \ldots, a_n\}$ *vektorrendszernek* $\{a_{i_1}, \ldots, a_{i_k}\}$ *egy lineárisan függő részrendszere, azaz*

$$\alpha_{i_1} a_{i_1} + \cdots + \alpha_{i_k} a_{i_k} = 0$$

úgy is teljesül valamely $\alpha_{i_1}, \ldots, \alpha_{i_k}$ *skalárokkal, hogy közöttük van nullától különböző, akkor ezt a lineáris kombinációt úgy kiegészítve, hogy a benne nem szereplő* $\{a_1, \ldots, a_n\}$*-beli vektorokat is beírjuk 0 együtthatókkal, akkor az* $\{a_1, \ldots, a_n\}$ *vektorrendszernek olyan lineáris kombinációját kapjuk, amely a nullvektorral egyenlő, de a benne szereplő együtthatók között van olyan, amelyik nem nulla. Ez viszont ellentmond annak, hogy az* $\{a_1, \ldots, a_n\}$ *vektorrendszer lineárisan független.*

Megjegyzés 2.3 *A vektorok függetlenségének definíciójában (Definíció 2.4) csak véges sok vektorból álló vektorrendszer függetlenségét definiáltuk, mert főleg ezzel a fogalommal foglalkozunk, de megjegyezzük, hogy a vektorok függetlensége vektorok tetszőleges számosságú (nem-üres) részhalmazára is értelmezhető. Egy* F *test feletti vektortér vektoraiból álló tetszőleges (nem-üres) vektorrendszerét lineárisan függetlennek nevezzük, ha annak bármely (nem-üres) véges részredszere lineárisan független.*

Megjegyzés 2.4 *Lineárisan független* $a_1, \ldots, a_n$ *vektorrendszer nem tartalmazza a nullvektort, mert ha (például)* $a_1 = 0$ *lenne, akkor teljesülne az*

$$1a_1 + 0a_2 + \ldots 0a_n = 0$$

egyenlőség, amelyben az a_1 *együtthatója nem nulla.*

Megjegyzés 2.5 *Lineárisan független $a_1, \ldots, a_n$ vektorrendszer páronként különböző vektorokból áll, mert ha (például) $a_1 = a_2$ teljesülne, akkor teljesülne az*

$$1a_1 + (-1)a_2 + \ldots 0a_n = 0$$

egyenlőség, pedig abban van nem nulla együttható.

Tétel 2.3 *Egy egyelemű $\{a\}$ vektorrendszer akkor és csak akkor lineárisan független, ha az a vektor nem a nullvektor. Legalább kételemű vektorendszer akkor és csak akkor lineárisan független, ha a redszert alkotó vektorok egyikét sem lehet kifejezni a rendszerhez tartozó többi vektor lineáris kombinációjaként.*

Tétel 2.4 *Ha $a_1, \ldots, a_n$ egy F test feletti V vektortér lineárisan független vektorrendszere, akkor a V bármely vektora legfeljebb egyféleképpen állítható elő az $a_1, \ldots, a_n$ vektorok lineáris kombinációjaként.*

Bizonyítás. Legyen $a_1, \ldots, a_n$ egy F test feletti V vektortér lineárisan független vektorrendszere. Tegyük fel, hogy

$$\alpha a_1 + \ldots \alpha_n a_n = \beta_1 a_1 + \ldots + \beta_n a_n.$$

Akkor

$$(\alpha_1 - \beta_1)a_1 + \ldots + (\alpha_n - \beta_n)a_n = 0,$$

amiből

$$\alpha_i = \beta_i, \quad i = 1, \ldots, n$$

következik. □

Definíció 2.5 *Egy V vektortér $\{a_1, \ldots, a_n\}$ lineárisan független vektorrendszerről azt mondjuk, hogy maximálisan lineárisan független, ha V tetszőleges b vektora esetén a $\{b, a_1, \ldots, a_n\}$ vektorrendszer lineárisan függő, azaz az $\{a_1, \ldots, a_n\}$ vektorrendszert a V akármelyik vektorával egészíjük ki, lineárisan függő vektorrendszert kapunk.*

2.3. Generátorrendszer, bázis, dimenzió

Definíció 2.6 *Egy F test feletti V vektortér $\{\underline{a}_1, \ldots, \underline{a}_n\}$ vektorrendszerét a V vektortér egy generátorrendszerének nevezzük, ha V minden vektora elöáll az $\underline{a}_1, \ldots, \underline{a}_n$ vektorok lineáris kombinációjaként.*

Megjegyzés 2.6 *Egy vektortér generátorrendszere mindig tartalmaz legalább egy vektort.*

Megjegyzés 2.7 *A generátorrendszer fogalmát csak véges sok vektorból álló vektorrendszerre értelmeztük, de megjegyezzük, hogy tetszőleges számosságú (nemüres) vektorrendszerre is értelmezhető. Egy F test feletti vektortér vektoraiból álló tetszőleges (nem-üres) $\mathcal{G}$ vektorrendszerét a vektortér generátorrendszerének nevezzük, ha a V vektortér bármely $\underline{v}$ vektorához megadható $\mathcal{G}$-nek olyan véges részrendszere, amelyhez tartozó vektorok lineáris kombinációjaként előállítható a $\underline{v}$ vektor. Jó példa erre egy F test feletti egyváltozós polinomok vektortere, amelyben az*

$$1, x, \ldots, x^n, \ldots$$

polinomok generátorrendszert alkotnak.

Definíció 2.7 *Egy F test feletti V vektortér $\{\underline{a}_1, \ldots, \underline{a}_n\}$ generátorrendszerét minimális generátorrendszernek nevezzük, ha bárhogy is hagyunk el ebből a rendszerből vektorokat, akkor vagy az üres halmazt kapjuk, vagy a keletkezett vektorrendszer már nem generátorrendszer.*

Definíció 2.8 *Egy F test feletti V vektortér $\{\underline{a}_1, \ldots, \underline{a}_n\}$ vektorrendszerét a V vektortér egy bázisának nevezzük, ha V minden vektora elöáll egyértelműen az $\underline{a}_1, \ldots, \underline{a}_n$ vektorok lineáris kombinációjaként.*

Tétel 2.5 *Egy F test feletti V vektortér valamely nemüres vektorrendszere akkor és csak akkor bázis, ha független generátorrendszer.*

Megjegyzés 2.8 *A bázis fogalmát csak véges sok vektorból álló vektorrendszerre értelmeztük, de megjegyezzük, hogy tetszőleges számosságú (nemüres) vektorrendszerre is értelmezhető. Egy F test feletti vektortér vektoraiból álló tetszőleges (nem-üres) $\mathcal{B}$ vektorrendszerét a vektortér bázisának nevezzük, ha a V vektortér bármely $\underline{v}$ vektorához megadható $\mathcal{G}$-nek egy és csak egy olyan véges részrendszere, amelyhez tartozó vektorok lineáris kombinációjaként egyértelműen előállítható a $\underline{v}$ vektor. Egy F test feletti egyváltozós polinomok vektorterében az*

$$1, x, \ldots, x^n, \ldots$$

polinomok halmaza a vektortér egy bázisa. Tetszőleges számosság esetén is érvényes, hogy a bázis fogalma egybeesik a független generátorrendszer fogalmával.

Tétel 2.6 *Tetszőleges F test feletti $V \neq \{\underline{0}\}$ vektortér tetszőleges, nemüres $\mathcal{B}$ vektorrendszere esetén a következő feltételek egymással ekvivalensek.*

1. *$\mathcal{B}$ a V vektortér bázisa.*
2. *$\mathcal{B}$ a V vektortér maximális független rendszere.*
3. *$\mathcal{B}$ a V vektortér minimális generátorrendszere.*

Tétel 2.7 *(Steinitz-féle kicserélési tétel) Ha $\underline{a}_1, \ldots, \underline{a}_n$ egy generátorrendszere, $\underline{b}_1, \ldots, \underline{b}_m$ pedig egy lineárisan független rendszere egy F test feletti V vektortérnek, akkor bármely $\underline{b}_i$ vektorhoz van olyan $\underline{a}_j$ vektor, hogy a*

$$\underline{b}_1, \ldots, \underline{b}_{i-1}, \underline{a}_j, \underline{b}_{i+1} \ldots, \underline{b}_m$$

vektorrendszer lineárisan független.

Proof. A feltétel szerint $V \neq \{\underline{0}\}$. Vizsgáljuk először azt az esetet, amikor $m = 1$. Mivel $\underline{a}_1, \ldots, \underline{a}_n$ generátorrendszer, ezért van benne legalább egy nem nullvektor. Ezt a $\underline{b}_1$ helyére írva, lineárisan független rendszert kapunk. Tegyük fel, hogy $m > 1$. Ha valamelyikhez, mondjuk $\underline{b}_1$-hez nincs olyan $\underline{a}_j$ vektor, hogy az $\underline{a}_j, \underline{b}_2, \ldots, \underline{b}_m$ vektorrendszer lineárisan független lenne, azaz minden $\underline{a}_j$ $(j = 1, \ldots, n)$ vektor esetén az $\underline{a}_j, \underline{b}_2, \ldots, \underline{b}_m$ vektorrendszer lineárisan függő, akkor van olyan $\alpha_1\underline{a}_j + \alpha_2\underline{b}_2 + \cdots + \alpha_m\underline{b}_m$ lineáris kombinációjuk ami egyenlő a nullvektorral, de nem minden együttható nulla. Ekkor biztos, hogy $\alpha_1 \neq 0$. $m = 1$ esetén ez nyilvánvaló, $n > 1$ esetén pedig azért, mert a $\underline{b}_2, \ldots, \underline{b}_m$ vektorrendszer lineárisan független. Tehát minden $\underline{a}_j$ vektor előáll a $\underline{b}_2, \ldots, \underline{b}_m$ vektorok lineári kombinációjaként. Ebből az következik, hogy a $\underline{b}_2, \ldots, \underline{b}_m$ vektorrendszer a V egy generátorrendszere, és így a $\underline{b}_1$ vektor előáll a $\underline{b}_2, \ldots, \underline{b}_m$ vektorok lineáris kombinációjaként. A ?? **??**Tétel szerint ebből az következik, hogy a $\underline{b}_1, \ldots, \underline{b}_m$ vektorrendszer lineárisan függő. Mivel ez ellentmondás, ezért valamely $\underline{a}_j$ vektor esetén az $\underline{a}_j, \underline{b}_2, \ldots, \underline{b}_m$ vektorrendszer lineárisan független. □

Tétel 2.8 *(2. kicserélési tétel) Egy F test feletti V vektortérnek legyen $\underline{a}_1, \ldots, \underline{a}_n$ egy generátorrendszere, $\underline{b}_1, \ldots, \underline{b}_m$ pedig egy lineárisan független rendszere. Akkor bármely $\underline{b}_i$ vektorhoz van olyan $\underline{a}_j$ vektor, hogy az*

$$\underline{a}_1, \ldots, \underline{a}_{j-1}, \underline{b}_i, \underline{a}_{j+1} \ldots, \underline{a}_n$$

vektorrendszer a V generátorrendszere.

Bizonyítás. Mivel egy lineárisan független vektorrendszer nem tartalmazza a nullvektort, ezért tetszőleges $i = 1, \ldots, m$ index esetén $\underline{b}_i \neq \underline{0}$. Mivel az $\underline{a}_1, \ldots, \underline{a}_n$ vektorrendszer a V generátorrendszere, tetszőleges i index esetén megadhatók olyan α_j $(j = 1, \ldots, n)$ skalárok, hogy

$$\underline{b}_i = \alpha_1 \underline{a}_1 + \cdots + \alpha_n \underline{a}_n,$$

és $\alpha_j \neq 0$ valamely $j = 1, \ldots, n$ indexre. Az általánosság megszorítása nélkül feltehetjük, hogy $i = j = 1$. Akkor

$$\underline{a}_1 = \frac{1}{\alpha_1}\underline{b}_1 - \frac{\alpha_2}{\alpha_1}\underline{a}_2 - \cdots - \frac{\alpha_n}{\alpha_1}\underline{a}_n.$$

Legyen $\underline{c}$ a V tetszőleges vektora. Akkor

$$\underline{c} = \gamma_1 \underline{a}_1 + \cdots + \gamma_n \underline{a}_n =$$

$$= \gamma_1\left(\frac{1}{\alpha_1}\underline{b}_1 - \frac{\alpha_2}{\alpha_1}\underline{a}_2 - \cdots - \frac{\alpha_n}{\alpha_1}\underline{a}_n\right) + \gamma_2\underline{a}_2 + \cdots + \gamma_n\underline{a}_n =$$

$$= \frac{\gamma_1}{\alpha_1}\underline{b}_1 + \left(\gamma_2 - \alpha_2\frac{\gamma_1}{\alpha_1}\right)\underline{a}_2 + \cdots + \left(\gamma_n - \alpha_n\frac{\gamma_1}{\alpha_1}\right)\underline{a}_n.$$

Így $\underline{a}_1, \ldots, \underline{a}_{j-1}, \underline{b}_i, \underline{a}_{j+1} \ldots, \underline{a}_n$ a V egy generátorrendszere. □

Tétel 2.9 *Ha egy F test feletti vektortérnek van n elemű bázisa, akkor bármely bázisa n vektort tartalmaz.*

Definíció 2.9 *Valamely n pozitív egész szám esetén, egy F test feletti vektorteret n-dimenziós vektortérnek nevezünk, ha V ben van n-elemű bázis. Az egyetlen vektorból álló vektortér dimenziója 0 (mivel egyetlen elme van, a nullvektor, ami ugyan generátorrendszer, de nem bázis).*

Megjegyzés 2.9 *Egy vektortérben akkor és csak akkor van véges sok elemet tartalmazó bázis, ha a vektortérnek van véges sok vektorból álló generátrorrendszere. (ugyanis, minden véges generátorrendszer tartalmaz egy minimális generátorrendszert és minden véges bázis véges generátorrendszer).*

Definíció 2.10 *Egy vektorteret véges dimenziós vektortérnek nevezünk, ha van véges sok vektorból álló bázisa. Egy vektorteret végtelen dimenziós vektortérnek nevezünk, ha nincs véges generátorrendszere.*

2.4. Vektorrendszer rangja

Definíció 2.11 *Egy F test feletti V vektortér $a_1, \ldots, a_n$ vektoraiból álló vektorrendszer rangján a közöttük lévő lineárisan független vektorok számának maximumát értjük. Egy $\{a_1, \ldots, a_n\}$ vektorrendszer rangját $\rho(a_1, \ldots, a_n)$ módon fogjuk jelölni.*

Megjegyzés 2.10 *Az világos, hogy egy vektorrendszer rangja nem függ a benne szereplő vektorok sorrendjétől.*

Tétel 2.10 *Egy vektorrendszer rangja megegyezik az általuk kifeszített altér dimenziójával.*

Bizonyítás. Legyenek $a_1, \ldots, a_n$ egy F test feletti V vektortér tetszőleges vektorai. Legyen az $\{a_1, \ldots, a_n\}$ vektorrendszer rangja r ($r \leq n$). Jelölje m az általuk kifeszített W altér dimenzióját. Megjegyzés 2.10 miatt feltehetjük, hogy $a_1, \ldots, a_r$ lineárisan független (így p'aronként különböző) vektorok V-ben. Világos, hogy az $\{a_1, \ldots, a_r\}$ vektorrendszer lineárisan független W-ben is, így $r \leq m$. Az is nyílvánvaló, hogy az $a_{r+1}, \ldots, a_n$ vektorok mindegyike kifejezhető az $a_1, \ldots, a_r$ vektorok lineáris kombinációjaként, és így az $\{a_1, \ldots, a_r\}$ vektorrendszer a W altér egy generátorrendszere. Ebből $m \leq r$ adódik, ami a fent bizonyított $r \leq m$ egyenlőtlenséggel együtt az $r = m$ egyenlőséget eredményezi. □

Tétel 2.11 *Egy vektorrendszer rangja nem változik, ha a benne szereplő vektorok valamelyikét megszorozzuk egy nem nulla skalárral.*

Bizonyítás. Legyenek $a_1, \ldots, a_n$ egy F test feletti V vektortér tetszőleges vektorai. Legyen $0 \neq \alpha \in F$ tetszőleges. Megmutatjuk, hogy

$$\rho(a_1, \ldots, a_n) = \rho(\alpha a_1, \ldots, a_n).$$

Ebből már következik a tétel állítása.

Mivel $\alpha a_1 \in \langle a_1, \ldots, a_n \rangle$, ezért

$$\langle \alpha a_1, \ldots, a_n \rangle \subseteq \langle a_1, \ldots, a_n \rangle.$$

Legyen

$$a = \xi_1 a_1 + \ldots + \xi_n a_n \in \langle a_1, \ldots, a_n \rangle$$

tetszőleges elem. Akkor

$$a = (\frac{1}{\alpha}\xi_1)(\alpha a_1) + \ldots + \xi_n a_n$$

miatt

$$a \in \langle \alpha a_1, \ldots, a_n \rangle,$$

és ezért

$$\langle a_1, \ldots, a_n \rangle \subseteq \langle \alpha a_1, \ldots, a_n \rangle.$$

Tehát teljesül az

$$\langle a_1, \ldots, a_n \rangle = \langle \alpha a_1, \ldots, a_n \rangle$$

egyenlőség. Ebből már következik a

$$\rho(a_1, \ldots, a_n) = \rho(\alpha a_1, \ldots, a_n)$$

egyenlőség (lásd Tétel 2.10). □

Tétel 2.12 *Egy vektorrendszer rangja nem változik, ha a benne szereplő vektor valamelyikéhez hozzáadjuk egy tőle különböző, a vektorrendszerben szereplő vektor konstansszorosát.*

Bizonyítás. Legyenek $a_1, \ldots, a_n$ $(n \geq 2)$ egy F test feletti V vektortér tetszőleges vektorai. Legyen $\alpha \in F$ tetszőleges skalár. Megmutatjuk, hogy

$$\rho(a_1, a_2, \ldots, a_n) = \rho((a_1 + \alpha a_2), a_2, \ldots, a_n).$$

Ebből már következik a tétel állítása.

Mivel $a_1 + \alpha a_2 \in \langle a_1, \ldots, a_n \rangle$, ezért

$$\langle (a_1 + \alpha a_2), a_2, \ldots, a_n \rangle \subseteq \langle a_1, a_2 \ldots, a_n \rangle.$$

Legyen

$$a = \xi_1 a_1 + \xi_2 a_2 + \ldots + \xi_n a_n \in \langle a_1, a_2, \ldots, a_n \rangle$$

tetszőleges elem. Akkor

$$a = \xi_1 a_1 + (\alpha\xi_1 + \xi_2 - \alpha\xi_1)a_2 + \ldots + \xi_n a_n =$$

$$\xi_1(a_1 + \alpha a_2) + (\xi_2 - \alpha\xi_1)a_2 + \ldots + \xi_n a_n$$

miatt

$$a \in \langle (a_1 + \alpha a_2), a_2, \ldots, a_n \rangle,$$

és ezért
$$\langle a_1, \ldots, a_n \rangle \subseteq \langle (a_1 + \alpha a_2), a_2, \ldots, a_n \rangle.$$
Tehát teljesül az
$$\langle a_1, \ldots, a_n \rangle = \langle (a_1 + \alpha a_2), a_2, \ldots, a_n \rangle$$
egyenlőség. Ebből már következik a
$$\rho(a_1, \ldots, a_n) = \rho((a_1 + \alpha a_2), a_2, \ldots, a_n)$$
egyenlőség (lásd Tétel 2.10). □

2.5. A térbeli vektorok vektortere

Definíció 2.12 *Térbeli vektoron egy irányított szakaszt értünk.*

Megjegyzés 2.11 *Kicsit pontosabban, térbeli vektoron a tér pontjaiból képezett* (A, B) *rendezett elempárt értünk, ahol A-t a térbeli vektor kezdőpontjának, a* B *pontot pedig a végpontjának nevezzük.*

Megjegyzés 2.12 *A szabad vektor fogalmát a következőképpen is értelmezhetjük. Jelölje* M *a tér pontjainak halmazát. Az* M^2 *halmazon definiáljunk egy relációt a következőképpen:* (A, B) *és* (C, D) *rendezett pontpárok akkor és csak akkor legyenek relációban, ha megadható a térben olyan párhuzamos eltolás, amely az* A *pontot a* C *pontba, a* B *pontot pedig a* D *pontba viszi át. Világos, hogy ez a reláció reflexív, szimmetrikus és tranzitív, azaz egy ekvivalencia-reláció. Ennek osztályai a szabad vektorok.*

2.6. Rendezett elem n-esek vektortere

Definíció 2.13 *hhh*

3. fejezet

Mátrixok

Definíció 3.1 *Legyen R egy tetszőleges gyűrű, és jelöljön m, illetve n egy-egy pozitív egész számot. Az $I_m \times I_n$ halmaznak R-be való egyértelmű leképezéseit R feletti $m \times n$-típusú mátrixoknak nevezzük.*

Megjegyzés 3.1 *Ha a_{ij} jelöli R azon elemét, amelyet egy A mátrix az $(i,j) \in I_m \times I_n$ elempárhoz rendel, akkor az A mátrixot a következő módokon jelöljük:*

$$A, \quad [a_{ij}], \quad A_{m\times n}, \quad [a_{ij}]_{m\times n}, \quad \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Az a_{ij} elemre az "i-dik sor j-dik eleme" (vagy a "j-dik oszlop i-dik eleme") kifejezést szoktuk használni.

Ha $k = min\{m,n\}$, akkor az

$$[a_{11}, a_{22}, \dots, a_{kk}]$$

elemsorozatot az A mátrix főátlójának, az

$$[a_{1n}, a_{2,n-1}, \dots, a_{k,n-(k-1)}]$$

elemsorozatot az A mátrix mellékátlójának nevezzük.

Az

$$[a_{i1}, \dots, a_{in}] \in R^n$$

elemsorozatot ($i \in I_m$) az A mátrix i-dik sorának, az

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in R^m$$

elemsorozatot ($j \in I_n$) az A mátrix j-dik oszlopának nevezzük.

Definíció 3.2 *Két mátrixot akkor tekintünk egyenlőnek, ha azonos típusúak, és elemeik rendre megegyeznek.*

Megjegyzés 3.2 *Ha egy*

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

mátrix esetén bevezetjük az

$$[a_j] = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in R^m$$

jelölést ($j \in I_n$), akkor az A mátrix úgy is tekinthető mint egy n-elemű sorozat:

$$A = [[a_1], \dots, [a_n]].$$

Megjegyzés 3.3 *Ha A egy F test feletti $m \times n$-típusú mátrix, akkor az A mátrix i-dik ($i \in I_n$) sorát alkotó $[a_{i1}, \dots, a_{i,n}]$ elemsorozat az F^n vektortér (mint sorvektorok tere) egy elemének tekinthető, ezért az $[a_{i1}, \dots, a_{in}]$ elemsorozatot az A mátrix i-dik sorvektorának is szoktuk nevezni. Hasonlóan, az*

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in F^m$$

elemsorozatot ($j \in I_n$) az A mátrix j-dik oszlopvektorának is szoktuk nevezni.

3.1. Műveletek mátrixok között

Definíció 3.3 *Azonos, $m \times n$-típusú*

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

és

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

mátrixok összegén az ugyancsak $m \times n$-típusú

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

mátrixot értjük.

Megjegyzés 3.4 *Az összeadást tehát csak azonos típusú mátrixok között értelmezzük.*

Tétel 3.1 *Tetszőleges R gyűrű feletti $m \times n$-típusú mátrixok $R^{m \times n}$ halmaza a mátrixok (előzőekben definiált) összeadására nézve kommutatív csoportot alkot; benne a*

$$O = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

mátrix (az un. $m \times n$-típusú nullmátrix) a nullelem.

Bizonyítás.

Definíció 3.4 *Egy R gyűrű feletti $m \times n$-típusú A mátrix transzponáltján azt az A^T módon jelölt $n \times m$-típusú (R feletti) mátrixot értjük, amelyben az i-dik sor ($i = 1, \dots n$) megegyezik az A mátrix i-dik sorával.*

Tétel 3.2 *Egy R gyűrű feletti $m \times n$-típusú mátrixok esetén*

$$(A+B)^T = A^T + B^T.$$

Bizonyítás.

Definíció 3.5 *Egy R gyűrű feletti $m \times n$-típusú*

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

és egy $n \times k$-típusú

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{bmatrix}$$

mátrix szorzatán azt az $m \times k$-típusú C mátrixot értjük, amelyben az i-dik sor j-dik eleme a következőképpen van definiálva:

$$c_{ij} = \sum_{k=1}^{n} (a_{ik} b_{kj}) = a_{i1} b_{1j} + \ldots + a_{in} b_{nj}.$$

Megjegyzés 3.5 *Figyeljük meg, hogy az AB szorzat csak akkor van értelmezve, ha A oszlopainak száma megegyezik B sorainak számával.*

Megjegyzés 3.6 *Legyen R egységelemes gyűrű, n pedig egy tetszőleges pozitív egész szám. Jelölje $[e_i]$ $(i = 1, \ldots, n)$ azt az R feletti $n \times 1$-típusú mátrixot, amelyben az i-dik elem az R egységeleme, az összes többi pedig az R nulleleme. Akkor tetszőleges R feletti $m \times n$-típusú $A = [[a_1], \ldots, [a_n]]$ mátrix esetén*

$$A[e_i] = [a_i].$$

Tétel 3.3 *Azonos, $m \times n$-típusú A és B mátrixok akkor és csak akkor egyenlőek, ha $A[e_i] = B[e_i]$ teljesül minden $i = 1, \ldots, n$ indexre.*

Bizonyítás. Megjegyzés 3.6 alapján nyílvánvaló □

Tétel 3.4 *Tetszőleges gyűrű feletti A, B, C mátrixok esetén az $(AB)C$ szorzat akkor és csak akkor van értelmezve, ha értelmezve van az $A(BC)$ szorzat. Ha a szorzatok értelmezve vannak, akkor $(AB)C = A(BC)$.*

Bizonyítás. A □

Tétel 3.5 *Tetszőleges gyűrű feletti A, B, C mátrixok esetén az $A(B + C)$ mátrix akkor és csak akkor van értelmezve, ha értelmezve van az $AB + AC$ szorzatösszeg. Ha a szóban forgó mátrixok értelmezve vannak, akkor $A(B + C) = AB + AC$.*

Bizonyítás.

Tétel 3.6 *Tetszőleges R gyűrű feletti $n \times n$-típusú mátrixok $R^{n\times n}$ halmaza a mátrixok összeadására és szorzására nézve gyűrűt alkot. Ez a gyűrű általában nem kommutatív, még akkor sem, ha R kommutatív. Ha R egységelemes gyűrű, akkor az $R^{n\times n}$ mátrixgyűrű is az; benne az*

$$E = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}$$

mátrix (az un. egységmátrix) az egységelem.

Bizonyítás.

Definíció 3.6 *Egy $n \times n$-típusú mátrixot négyzetes mátrixnak nevezünk.*

Tétel 3.7 *Egy R gyűrű feletti azonos típusú négyzetes A és B mátrix esetén*

$$(AB)^T = B^T A^T.$$

Bizonyítás.

Definíció 3.7 *Egy egységelemes R gyűrű feletti $A \in R^{n\times n}$ négyzetes mátrix inverzén azt az $A^{-1} \in R^{n\times n}$ mátrixot értjük, amelyre $AA^{-1} = A^{-1}A = E$ teljesül. Ekkor azt is mondjuk, hogy A invertálható mátrix.*

Megjegyzés 3.7 *Ha egy (négyzetes) A mátrix invertálható, akkor az A inverze is invertálható (ugyanis az A^{-1} mátrix inverze az A mátrix).*

Tétel 3.8 *Ha az $A, B \in R^{n\times n}$ mátrixoknak van inverze, akkor az AB (és a BA szorzatnak is) van inverze, mégpedig $(AB)^{-1} = B^{-1}A^{-1}$ (és $(BA)^{-1} = A^{-1}B^{-1}$).*

Bizonyítás.

Tétel 3.9 *Egységelemes R gyűrű esetén az $R^{n\times n}$ mátrixgyűrű invertálható mátrixainak halmaza a mátrixok szorzására nézve csoportot alkot.*

Bizonyítás. Az invertálható mátrixok halmaza nem üres (mert tartalmazza az egységmátrixot) és zárt a szorzásra nézve (Tétel 3.8, így a szorzásra nézve egy monoidot alkot. A Megjegyzés 3.7 miatt ez a monoid csoport. □

3.2. A determináns

Definíció 3.8 *Legyen $\mathbb{R}$ egy kommutatív gyűrű. Egy R feletti $n \times n$-típusú*

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

mátrix determinánsán a $\mathbb{R}$ gyűrű

$$det(A) = \Sigma_{\sigma \in S_n} (-1)^{I(\sigma)} a_{1,\sigma(1)} \dots a_{n,\sigma(n)}$$

elemét értjük.

Megjegyzés 3.8 *Egy A négyzetes mátrix determinánsát $|A|$ módon is szoktuk jelölni. Ha fel akarjuk tüntetni az A matrix elemeit is, akkor egy*

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

mátrix determinánsát jelölhetjük még a következőképpen is:

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

Megjegyzés 3.9 *Egy kommutatív gyűrű feletti négyzetes mátrix determinánsa az R egy eleme, de szoktunk beszélni egy determináns oszlopairól, illetve sorairól. Ilyenkor a szóbanforgó mátrix oszlopaira, illetve soraira hivatkozunk.*

3.2.1. A determináns alaptulajdonságai

Tétel 3.10 *Kommutatív gyűrű feletti tetszőleges négyzetes A mátrix esetén*

$$det(A) = det(A^T),$$

azaz egy négyzetes mátrix determinánsa megegyezik transzponáltjának determinánsával.

Bizonyítás. Legyen A egy kommutatív gyűrű feletti $n \times n$-típusú mátrix. Jelölje B az A transzponáltját. Akkor a Tétel 1.7 és Megjegyzés 1.7 felhasználásával kapjuk, hogy

$$det(B) = \Sigma_{\sigma \in S_n} (-1)^{I(\sigma)} b_{1,\sigma(1)} \dots b_{n,\sigma(n)} =$$

$$\Sigma_{\sigma \in S_n} (-1)^{I(\sigma)} a_{\sigma(1),1} \dots a_{\sigma(n),n} =$$

$$\Sigma_{\sigma \in S_n} (-1)^{I(\sigma)} a_{\sigma(1),(\sigma^{-1} \circ \sigma)(1)} \dots a_{\sigma(n),(\sigma^{-1} \circ \sigma)(n)} =$$

$$\Sigma_{\sigma \in S_n} (-1)^{I(\sigma^{-1})} a_{1,\sigma^{-1}(1)} \dots a_{n,\sigma^{-1}(n)} =$$

$$\Sigma_{\sigma \in S_n}^{I(\sigma)} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} = det(A).$$

□

Tétel 3.11 *Egy mátrix akkor és csak akkor invertálható, ha reguláris.*

Tétel 3.12 *Azonos típusú, négyztes mátrixok szorzatának determinánsa egyenlő a tényezők determinánsainak szorzatával.*

Bizonyítás. Legyenek

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

és

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}$$

tetszőleges kommutatív gyűrű feletti mátrixok. Képezzük a $2n \times 2n$-típusú

$$\begin{bmatrix} A & 0 \\ -E & B \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} & 0 & \dots & 0 & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} & 0 & \dots & 0 & \\ -1 & \dots & 0 & b_{11} & \dots & b_{1n} & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & -1 & b_{n1} & \dots & b_{nn} & \end{bmatrix}$$

mátrixot. Alkalmazuk a Laplace-féle kifejtési tételt az első n sorra. Akkor

$$\begin{vmatrix} A & 0 \\ -E & B \end{vmatrix} = det(A)det(B).$$

Ha az

$$\begin{bmatrix} a_{11} & \dots & a_{1n} & 0 & \dots & 0 & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} & 0 & \dots & 0 & \\ -1 & \dots & 0 & b_{11} & \dots & b_{1n} & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & -1 & b_{n1} & \dots & b_{nn} & \end{bmatrix}$$

mátrix $(n+i)$-dik oszlopából $(i = 1, \dots, n)$ kivonjuk a j-dik oszlop $(j = 1, \dots, n)$ b_{ji}-szeresét, akkor a következő mátrixot kapjuk:

$$\begin{bmatrix} 0 & AB \\ -E & 0 \end{bmatrix}.$$

Ennek determinánsa egyrészt megegyezik az

$$\begin{bmatrix} A & 0 \\ -E & B \end{bmatrix}$$

mátrix determinánsával, másrészt viszont (alkalmazva a Laplace-féle kifejtési tételt az utolsó n oszlopára) egyenlő $det(AB)$-vel. Tehát $det(A)det(B) = det(AB)$. □

3.3. Mátrix rangja

Definíció 3.9 *Egy kommutatív gyűrű elemeiből képezett mátrix rangján a mátrixból kiválasztható nem-zérus értékű aldeterminánsok rendszámának maximumát értjük.*

Tétel 3.13 *Egy test elemeiből képezett mátrix rangja megegyezik az oszlopvektoraiból álló vektorrendszer rangjával (lásd Definíció 2.11).*

Bizonyítás. Legyen F egy test és

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

egy F feletti tetszőleges mátrix. Jelölje r a mátrix rangját. Legyen

$$D = \begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \dots & a_{i_1j_r} \\ a_{i_2j_1} & a_{i_2j_2} & \dots & a_{i_2j_r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_rj_1} & a_{i_rj_2} & \dots & a_{i_rj_r} \end{vmatrix}$$

az A mátrix olyan aldeterminánsa, amely nem egyenlő az F nullelemével. Megmutatjuk, hogy az A mátrix j_1-edik, j_2-edik, $\dots j_r$-edik oszlopvektoraiból álló vektorrendszer lineárisan független. Tegyük fel, hogy

$$\xi_1 a_{j_1} + \dots + \xi_r a_{j_r} = 0,$$

ahol $a_{j_1}, \dots, a_{j_r}$ jelöli az A mátrix j_1-dik, $\dots j_r$-edik oszlopvektorát, $\xi_1, \dots, \xi_r$ pedig F-beli skalárok. Akkor a $\xi_1, \dots, \xi_r$ skalárokra teljesül az

$$\begin{bmatrix} a_{i_1j_1} & a_{i_1j_2} & \dots & a_{i_1j_r} \\ a_{i_2j_1} & a_{i_2j_2} & \dots & a_{i_2j_r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_rj_1} & a_{i_rj_2} & \dots & a_{i_rj_r} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

egyenlőség. Mivel

$$0 \neq D = \begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \dots & a_{i_1j_r} \\ a_{i_2j_1} & a_{i_2j_2} & \dots & a_{i_2j_r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_rj_1} & a_{i_rj_2} & \dots & a_{i_rj_r} \end{vmatrix},$$

ezért a

$$B = \begin{bmatrix} a_{i_1j_1} & a_{i_1j_2} & \dots & a_{i_1j_r} \\ a_{i_2j_1} & a_{i_2j_2} & \dots & a_{i_2j_r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_rj_1} & a_{i_rj_2} & \dots & a_{i_rj_r} \end{bmatrix}$$

mátrix reguláris. A Tétel 3.11 miatt a B mátrixnak van B^{-1} inverze. Beszorozva ezzel az egyenlőséget balról, kapjuk, hogy

$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_r \end{bmatrix} = B^{-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Tehát $\xi_k = 0$ minden $k = 1, \dots, r$ indexre. Tehát az $a_{j_1}, \dots, a_{j_r}$ oszlopvektorok lineárisan függetlenek.

Legyen $j \notin \{j_1, \dots, j_r\}$ tetszőleges index. Megmutatjuk, hogy az A mátrix j-dik oszlopvektora előáll az A mátrix i_1-edik, i_2-edik, $\dots i_r$-edik oszlopvektorainak lineáris kombinációjaként. Legyen i tetszőleges sorindex $(1 \le i \le m)$. Képezzük az

$$A_i = \begin{bmatrix} a_{i_1j_1} & a_{i_1j_2} & \dots & a_{i_1j_r} & a_{i_1j} \\ a_{i_2j_1} & a_{i_2j_2} & \dots & a_{i_2j_r} & a_{i_2j} \\ \vdots & \vdots & \ddots & \vdots & \\ a_{i_rj_1} & a_{i_rj_2} & \dots & a_{i_rj_r} & a_{i_rj} \\ a_{ij_1} & a_{ij_2} & \dots & a_{ij_r} & a_{ij} \end{bmatrix}$$

mátrixot. Az világos, hogy $det(A_i) = 0$ minden $i = 1, 2, \dots m$ indexre. Ez azért igaz, mert $i \in \{i_1, i_2, \dots, i_r\}$ esetén az A_i mátrix két sora egymással megegyezik (és így $det(A_i) = 0$), $i \notin \{i_1, i_2, \dots, i_r\}$ esetén pedig $det(A_i)$ az A mátrix egy $r+1$-edrendű aldeterminánsa (és így $det(A_i) = 0$, mert az A mátrix rangja r). Jelölje D_k $(k = 1, \dots, r)$ az A_i mátrixban az utolsó sor a_{i,j_k} eleméhez tartozó előjeles aldeterminánst. (Megjegyezzük, hogy az A_i mátrix a_{ij} eleméhez tartozó előjeles aldetermináns egyenlő D-vel.) Az világos, hogy a D és D_k $(k = 1, \dots r)$ determinánsok függetlenek az i indextől. Az utolsó sor szerinti kifejtést alkalmazva,

$$0 = det(A_i) = a_{ij_1} D_1 + \dots + a_{i,j_r} D_r + a_{ij} D.$$

Mivel ez a képlet minden $i = 1, 2, \dots, m$ index esetén érvényes, ezért kapjuk a

$$0 = a_{j_1} D_1 + \dots + a_{j_r} D_r + a_j D$$

egyenlőséget, ahol $a_{j_1}, a_{j_2}, \dots a_{j_r}, a_j$ jelölik az A matrix j_1-edik, j_2-edik, …,j_r-edik, j-edik oszlopvektorát. Mivel $D \neq 0$ a feltétel miatt, ezért

$$a_j = \frac{D_1}{D} a_{j_1} + \ldots + \frac{D_r}{D} a_{j_r}.$$

Tehát az A mátrix j-dik oszlopvektora előáll az A mátrix j_1-edik, j_2-edik, … j_r-edik oszlopvektorainak lineáris kombinációjaként. Igy az A mátrix oszlopvektoraiból álló vektorrendszer rangja r. Ezzel a tételt bebizonyítottuk. □

Tétel 3.14 *Egy F test feletti n-dimenziós V vektortér bármely $\mathcal{B}$ bázisa esetén a V tetszőleges $\{a_1, \dots, a_m\}$ vektorrendszerének rangja megegyezik az $[[a_1]_\mathcal{B}, \dots, [a_m]_\mathcal{B}]$ mátrix rangjával.*

Bizonyítás. A Tétel 6.13 alapján nyílvánvaló □

Tétel 3.15 *Egy F test feletti $n \times n$-típusú mátrix oszlopvektorai (az F^n vektortérben) akkor és csak akkor lineárisan függetlenek ha a mátrix reguláris.*

Bizonyítás. A Tétel 3.13 miatt egy test feletti $n \times n$-típusú mátrix oszlopvektorai akkor és csak akkor lineárisan függetlenek, ha a mátrix rangja n, amely a rang definíciója szerint azt jelenti, hogy a mátrix determinánsa nem nulla, azaz a mátrix reguláris. □

Tétel 3.16 *Test elemeiből képezett mátrix oszlopvektoraiból álló vektorrendszer rangja megegyezik a sorvektoraiból álló vektorrendszer rangjával.*

Bizonyítás.

4. fejezet

Lineáris egyenletrendszerek

Definíció 4.1 *Lineáris egyenletrendszeren olyan egyenletrendszert fogunk érteni, amely véges sok elsőfokú egyenletből áll és véges sok ismeretlent tartalmaz.*

Az m egyenletből álló, n ismeretlent tartalmazó lineáris egyenletrendszer általános alakja

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

Megjegyzés 4.1 *Az fenti egyenletrendszerben szereplő a_{ij}-ket ($1 \leq i \leq m$, $1 \leq j \leq n$) az egyenletrendszer együtthatóinak, a b_i-ket ($1 \leq j \leq m$) az egyenletrendszer konstansainak, az x_j-ket ($1 \leq j \leq n$) pedig az egyenletrendszer ismeretlenjeinek nevezzük. A továbbiakban az együtthatókról, illetve a konstansokról feltesszük, hogy valamely F test elemei (ilyenkor azt is fogjuk mondani, hogy F feletti egyenletrendszerről van szó) az ismeretlenek pedig F elemeitől és egymástól is különböző szimbólumok.*

Egy lineáris egyenletrendszert homogén lineáris egyenletrendszernek nevezünk ha az egyenletrendszer konstansainak mindegyike az F test nulleleme; ellenkező esetben inhomogén lineáris egyenletrendszerről beszélünk.

Definíció 4.2 *Egy F test feletti m egyenletből álló, n-ismeretlent tartalmazó*

$$\begin{array}{ccccccc} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\dots & +a_{mn}x_n & = & b_m \end{array}$$

lineáris egyenletrendszer megoldásain olyan $[c_1, c_2, \dots, c_n] \in F^n$ rendezett elem n-eseket (vektorokat) értünk, amelyek esetén F-ben teljesül az

$$a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{1n}c_n = b_i$$

egyenlőség minden $i = 1, 2, \cdots, m$ indexre.

Definíció 4.3 *Egy*

$$\begin{array}{ccccccc} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\dots & +a_{mn}x_n & = & b_m \end{array}$$

lineáris egyenletrendszer együtthatóiból képezett

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

mátrixot a lineáris egyenletrendszer mátrixának, az

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

mátrixot pedig a lineáris egyenletrendszer kibővített mátrixának nevezzük. Ezeket, valamint az

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

jelöléseket használva, a lineáris egyenletredszer

$$A\underline{x} = \underline{b}$$

alakban is felírható; ezt az alakot a lineáris egyenletrendszer mátrixszorzatos alakjának nevezzük.

Ha $\underline{a}_j$ $(j = 1, \ldots, n)$ jelöli az A mátrix j-dik oszlopvektorát (amely eleme az F^m vektortérnek), akkor a fenti lineáris egyenletrendszer

$$x_1\underline{a}_1 + \cdots + x_n\underline{a}_n = \underline{b}$$

alakban is írható. Egy $[c_1, \ldots, c_n] \in F^n$ vektor akkor ás csak akkor megoldása a fenti egyenletrendszernek, ha a $\underline{b} \in F^m$ vektor előáll az $\underline{a}_1, \ldots, \underline{a}_n \in F^m$ vektoroknak a $c_1, \ldots c_n$ skalárokkal képezett lineáris kombinációjaként.

Definíció 4.4 *Egy* $\mathbf{A}\underline{x} = \underline{b}$ *lineáris egyenletrendszerhez tartozó homogén lineáris egyenletrendszeren (más szavakkal: a lineáris egyenletrendszer homogén részén) az* $\mathbf{A}\underline{x} = \underline{0}$ *homogén lineáris egyenletrendszert értjük.*

Egy homogén lineáris egyenletrendszernek az

$$x_1 = 0, \ldots, x_n = 0$$

elem n-es mindig megoldása; ezt az egyenletrendszer triviális megoldásának nevezzük.

Tétel 4.1 *Egy* m *egyenletből álló,* n *ismeretlent tartalmazó* $\mathbf{A}\underline{x} = \underline{0}$ *homogén lineáris egyenletrendszer megoldásainak halmaza az* F^n *vektortér egy altere.*

Bizonítás. Jelölje $H \subseteq F^n$ az $\mathbf{A}\underline{x} = \underline{0}$ homogén lineáris egyenletrendszer megoldásainak halmazát. $H \neq \emptyset$, mert a nullvektor (azaz a triviális megoldás) eleme H-nak. Legyenek

$$\underline{a}, \underline{b} \in H$$

tetszőleges vektorok. Akkor tetszőleges $\alpha, \beta \in F$ esetén

$$\mathbf{A}(\alpha\underline{a} + \beta\underline{b}) = \alpha(\mathbf{A}\underline{a}) + \beta(\mathbf{A}\underline{b}) = \underline{0}.$$

Tehát

$$\alpha\underline{a} + \beta\underline{b} \in H.$$

A Tétel **??** miatt H az F^n vektortér egy altere.

Tétel 4.2 *Jelölje I az $\mathbf{A}\underline{x} = \underline{b}$ lineáris egyenletrendszer megoldásainak, H pedig a lineáris egyenletrendszerhez tartozó homogén lineáris egyenletrendszer megoldásainak halmazát. Akkor $I = H + \underline{c}$, ahol $\underline{c}$ az $\mathbf{A}\underline{x} = \underline{b}$ egyenletrendszer egy (un. partikuláris) megoldása.*

Bizonyítás. Legyen $\underline{a} \in H$ tetszőleges. Akkor

$$\mathbf{A}(\underline{a} + \underline{c}) = (\mathbf{A}\underline{a}) + (\mathbf{A}\underline{c}) = \underline{0} + \underline{b} = \underline{b},$$

amiből

$$H + \underline{c} \subseteq I$$

következik.

Legyen $\underline{d} \in I$ tetszőleges. Akkor

$$\underline{d} = (\underline{d} - \underline{c}) + \underline{c}.$$

Mivel

$$\mathbf{A}(\underline{d} - \underline{c}) = (\mathbf{A}\underline{d}) - (\mathbf{A}\underline{c}) = \underline{b} - \underline{b} = \underline{0},$$

ezért $\underline{d} \in H + \underline{c}$, amiből

$$I \subseteq H + \underline{c}$$

következik. □

A következőkben szükséges és elégséges feltételt adunk arra, hogy egy F test feletti lineáris egyenletrendszer megoldható legyen F-ben, valamint módszert adunk a megoldásra (ha létezik).

4.1. A megoldhatóság mátrixrangos feltétele

Tétel 4.3 *Egy $\mathbf{A}\underline{x} = \underline{b}$ lineáris egyenletrendszer akkor és csak akkor oldható meg, ha az egyenletrendszer $\mathbf{A}$ mátrixának rangja megegyezik az $(\mathbf{A}, \underline{b})$ kibővített mátrix rangjával.*

Bizonyítás. Az $\mathbf{A}\underline{x} = \underline{b}$, azaz az

$$x_1\underline{a}_1 + \cdots + x_n\underline{a}_n = \underline{b}$$

lineáris egyenletrendszer akkor és csak akkor oldható meg, ha a $\underline{b}$ vektor benne van az $\underline{a}_1, \ldots, \underline{a}_n \in F^m$ vektorok által kifeszített $\langle \underline{a}_1, \ldots, \underline{a}_n \rangle$ altérben, azaz

$$\langle \underline{a}_1, \ldots, \underline{a}_n \rangle = \langle \underline{a}_1, \ldots, \underline{a}_n, \underline{b} \rangle,$$

ami azzal ekvivalens, hogy

$$dim(\langle \underline{a}_1, \ldots, \underline{a}_n \rangle) = dim(\langle \underline{a}_1, \ldots, \underline{a}_n, \underline{b} \rangle).$$

Ezen utóbbi feltétel a

$$rang\mathbf{A} = rang(\mathbf{A}, \underline{b}).$$

feltétellel ekvivalens a Tétel **??** miatt. □

Tétel 4.4 *Egy n ismeretlent tartalmazó $\mathbf{A}\underline{x} = \underline{b}$ lineáris egyenletrendszer akkor és csak akkor oldható meg egyértelműen, ha $rang(\mathbf{A}) = rang(\mathbf{A}, \underline{b}) = n$.*

Igy egy lineáris egyenletrendszernek akkor és csak akkor van több megoldása, ha az egyenletrendszer mátrixának rangja megegyezik a kibővített mátrixának rangjával, és ez a közös rang kisebb az ismeretlenek számánál.

Bizonyítás. Legyen $\mathbf{A}\underline{x} = \underline{b}$ olyan n ismeretlent tartalmazó lineáris egyenletrendszer, melynek van egy és csak egy megoldása az F^n vektortérben. A lineáris egyenletrendszer a következő alakban is tekinthető:

$$x_1\underline{a}_1 + \cdots + x_n\underline{a}_n = \underline{b}.$$

Mivel az egyenletrendszer megoldható, ezért

$$\langle \underline{a}_1, \ldots, \underline{a}_n \rangle = \langle \underline{a}_1, \ldots, \underline{a}_n, \underline{b} \rangle.$$

A Tétel bizonyításának jelöléseit használva, $I = H + \underline{c}$, ahol $\underline{c}$ az $\mathbf{A}\underline{x} = \underline{b}$ egyenletrendszer egyetlen megoldását jelöli. Mivel $I = \{\underline{c}\}$, ezért $H = \{\underline{0}\}$, azaz $x_1\underline{a}_1 + \cdots + x_n\underline{a}_n = \underline{0}$ akkor és csak akkor teljesül valamely $x_1, \ldots, x_n \in F$ skalárokkal, ha $x_1 = \ldots = x_n = 0$, ami definíció szerint annyit jelent, hogy az $\underline{a}_1, \ldots, \underline{a}_n$ vektorok lineárisan függetlenek. Ezért

$$n = dim\langle \underline{a}_1, \ldots, \underline{a}_n \rangle = dim\langle \underline{a}_1, \ldots, \underline{a}_n, \underline{b} \rangle,$$

azaz

$$rang\mathbf{A} = rang(\mathbf{A}, \underline{b}) = n.$$

Fordítva, tegyük fel, hogy $rang\mathbf{A} = rang(\mathbf{A}, \underline{b}) = n$, azaz

$$dim\langle \underline{a}_1, \ldots, \underline{a}_n \rangle = n$$

(ami annyit jelent, hogy az $\underline{a}_1, \ldots, \underline{a}_n$ vektorok lineárisan függetlenek, és így az $\langle \underline{a}_1, \ldots, \underline{a}_n \rangle$ altér egy bázisát alkotják) és

$$dim\langle \underline{a}_1, \ldots, \underline{a}_n, \underline{b} \rangle = n.$$

Mivel $\langle \underline{a}_1, \ldots, \underline{a}_n \rangle$ altere az $\langle \underline{a}_1, \ldots, \underline{a}_n, \underline{b} \rangle$ altérnek, ezért

$$\langle \underline{a}_1, \ldots, \underline{a}_n \rangle = \langle \underline{a}_1, \ldots, \underline{a}_n, \underline{b} \rangle;$$

amely vektortérnek egy bázisa az $\underline{a}_1, \ldots, \underline{a}_n$ vektorrendszer. Ezért a $\underline{b}$ vektor egy és csak egyféleképpen állítható elő az $\underline{a}_1, \ldots, \underline{a}_n$ vektorok lineáris kombinációjaként. Ez utóbbi éppen azzal ekvivalens, hogy az

$$x_1\underline{a}_1 + \cdots + x_n\underline{a}_n = \underline{b}$$

lineáris egyenletrendszer egyértelműen oldható meg. □

Megjegyzés 4.2 *Mint arról már korábban volt szó, egy homogén lineáris egyenletrendszernek mindig van megoldása (a triviális megoldás). Ezért a homogén lineáris egyenletrendszereknél igazából az a probléma, hogy mikor van a triviálistól különböző megoldása. A Tétel 4.4 alapján egyszerűen adódik a következő eredmény.*

Tétel 4.5 *Egy homogén lineáris egyenletrendszernek akkor és csak akkor van a triviálistól különböző megoldása, ha az egyenletrendszer mátrixának rangja kisebb az ismeretlenek számánál.*

Az utolsó két tétel olyan egyenletrendszerekre vonatkozik, amelyben az egyenletek száma megegyezik az ismeretlenek számával.

Tétel 4.6 *Egy olyan lineáris egyenletrendszer, amelyben az egyenletek száma megegyezik az ismeretlenek számával akkor és csak akkor oldható meg egyértelműen, ha az egyenletrendszer mátrixa reguláris.*

Bizonyítás. A Tétel 4.4 miatt az egyenletrendszer akkor és csak akkor oldható meg egyértelműen, ha az egyenletrendszer mátrixának rangja és kiegészített mátrixának rangja egyenlő, és ez a közös rang megegyezik az ismeretlenek számával. Ez a mi esetünkben akkor és csak akkor teljesül, ha az egyenletrendszer mátrixa reguláris. □

Tétel 4.7 *Egy olyan homogén lineáris egyenletrendszernek, amelyben az egyenletek száma megegyezik az ismeretlenek számával akkor és csak akkor van a triviálistól különböző megoldása, ha az egyenletredszer mátrixa szinguláris.*

Bizonyítás. A Tétel 4.6 felhasználásával az állítás nyílvánvaló. □

4.2. A Gauss-módszer

Legyen

$$\begin{array}{cccccccc} a_{11}x_1 & +a_{12}x_2 & +a_{13}x_3 & +\dots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +a_{23}x_3 & +\dots & +a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +a_{m3}x_3 & +\dots & +a_{mn}x_n & = & b_m \end{array}$$

egy F test feletti egyenletrendszer, amelyben $m \geq 2$. Tegyük fel, hogy az $x_1, \dots, x_n$ ismeretlenek mindegyike legalább egy helyen nem nulla együtthatóval szerepel az egyenletrendszerben.

A Gauss-módszert azzal kezdjük (ha erre szükség van), hogy az egyenletek, illetve az ismeretlenek egymás közötti (és persze az indexek megfelelő) cseréjével elérjük, hogy $a_{11} \neq 0$ legyen.

A Gauss-módszer első lépéseként az i-dik ($i \neq 1$) egyenlethez hozzáadjuk az első egyenlet $-\frac{a_{i1}}{a_{11}}$-szeresét. Ekkor egy

$$\begin{array}{cccccccc} a_{11}^{(1)}x_1 & +a_{12}^{(1)}x_2 & +a_{13}^{(1)}x_3 & +\dots & +a_{1n}^{(1)}x_n & = & b_1^{(1)} \\ 0 & +a_{22}^{(1)}x_2 & +a_{23}^{(1)}x_3 & +\dots & +a_{2n}^{(1)}x_n & = & b_2^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & +a_{m2}^{(1)}x_2 & +a_{m3}^{(1)}x_3 & +\dots & +a_{mn}^{(1)}x_n & = & b_m^{(1)} \end{array}$$

alakú lineáris egyenletrendszert kapunk, amelyben az első egyenlet minden együtthatója és a konstans rendre megegyezik az eredeti egyenletrendszer első egyenletében szereplő megfelelő együtthatóval, illetve a konstanssal, és ezért $a_{11}^{(1)} = a_{11} \neq 0$ is teljesül.

A Gauss módszer következő lépésére csak abban az esetben kerül sor, ha ebben az egyenletrendszerben az első egyenlet és az első ismeretlen kivételével az egyenleteket, illetve az ismeretleneket lehet egymás között úgy cserélni, hogy az indexek megfelelő cseréje után $a_{22}^{(1)} \neq 0$. Ha ez nem lehetséges (mert például az első egyenlet kivételével minden egyenlet bal oldalán nulla áll (ilyen helyzet adódik például akkor, amikor $n = 1$)), akkor a Gauss-módszert befejeztük. Ellenkező esetben végrehajtjuk a Gauss-módszer második lépését.

A Gauss-módszer második lépéseként az i-dik ($i \neq 2$) egyenlethez hozzáadjuk a második egyenlet $-\frac{a_{i2}}{a_{22}}$-szeresét. Ekkor egy

$$\begin{array}{cccccccc}
a_{11}^{(2)}x_1 & + \;\; 0 & +a_{13}^{(2)}x_3 & +\dots & +a_{1n}^{(2)}x_n & = & b_1^{(2)} \\
0 & +a_{22}^{(2)}x_2 & +a_{23}^{(2)}x_3 & +\dots & +a_{2n}^{(2)}x_n & = & b_2^{(2)} \\
0 & + \;\; 0 & +a_{33}^{(2)}x_3 & +\dots & +a_{3n}^{(2)}x_n & = & b_3^{(2)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\
0 & + \;\; 0 & +a_{m3}^{(2)}x_3 & +\dots & +a_{mn}^{(2)}x_n & = & b_m^{(2)}
\end{array}$$

alakú egyenletrendszert kapunk, amelyben a második egyenlet minden egyes együtthatója és a konstans megegyezik az előző egyenletrendszer második egyenletében szereplő megfelelő együtthatóval, illetve a konstanssal, és ezért $a_{22}^{(2)} \neq 0$. Az első egyenletben szereplő x_1 ismeretlen együtthatója sem változik, ezért $a_{11}^{(2)} \neq 0$ is teljesül.

A Gauss módszer következő lépésére csak abban az esetben kerül sor, ha $m \geq 3$ és ebben az egyenletrendszerben az első két egyenlet és az első két ismeretlen kivételével az egyenleteket, illetve az ismeretleneket lehet egymás között úgy cserélni, hogy az indexek megfelelő cseréje után $a_{33}^{(2)} \neq 0$. Ha ez nem lehetséges (mert például $m = 2$ vagy mert az első két egyenlet kivételével minden egyenlet bal oldalán nulla áll (ez a helyzet pl. ha $n = 2$)), akkor a Gauss-módszert befejeztük. Ellenkező esetben végrehajtjuk a Gauss-módszer következő lépését.

A gondolatmenetet folytatva, a Gauss-módszer utolsó (r-edik) lépésének végrehajtása után ($1 \leq r \leq \min\{n, m\}$) egy

$$\begin{array}{ccccccc}
a_{11}^{(r)}x_1 & & & +a_{1,r+1}^{(r)}x_{r+1} & +\dots & +a_{1n}^{(r)}x_n & = & b_1^{(r)} \\
 & \ddots & & \vdots & \ddots & \vdots & & \vdots \\
 & & a_{rr}^{(r)}x_r & +a_{r,r+1}^{(r)}x_{r+1} & +\dots & +a_{rn}^{(r)}x_n & = & b_r^{(r)}
\end{array}$$

alakú, vagy egy

$$\begin{array}{lllllll}
a_{11}^{(r)}x_1 & & +a_{1,r+1}^{(r)}x_{r+1} & +\ldots & +a_{1n}^{(r)}x_n & = & b_1^{(r)} \\
& \ddots & \vdots & \ddots & \vdots & & \vdots \\
& a_{rr}^{(r)}x_r & +a_{r,r+1}^{(r)}x_{r+1} & +\ldots & +a_{rn}^{(r)}x_n & = & b_r^{(r)} \\
& & & & 0 & = & b_{r+1}^{(r)} \\
& & & & \vdots & & \vdots \\
& & & & 0 & = & b_m^{(r)}
\end{array}$$

alakú egyenletrendszert kapunk, annak megfelelően, hogy $r = m$ vagy $r < m$ (mindkét esetben $a_{ii}^{(r)} \neq 0$ minden $i = 1, \ldots, r$ indexre).

Az világos, hogy az így kapott egyenletrendszerek ekvivalensek az eredeti egyenletrendszerrel, azaz (F-beni) megoldáshalmazuk azonos. A Gauss módszer alkalmazása során az egyenletrendszer mátrixán, illetve kibővített mátrixán elemi átalakításokat hajtunk végre, így a Gauss-módrszer alkalmazása után kapott egyenletrendszer mátrixának, illetve kibővített mátrixának rangja megegyezik az eredeti egyenletrendszer mátrixának, illetve kibővített mátrixának rangjával.

Vizsgáljuk az első esetet (azaz, amikor $r = m$).

Ha $n = r$ (ekkor $rangA = rang(A, b) = n$), akkor az egyenletrendszer egyértelműen oldható meg, mégpedig

$$x_k = \frac{b_k^{(r)}}{a_{kk}^{(r)}} \quad (k = 1, \ldots, r).$$

Ha $r < n$ (ekkor $rangA = rang(A, b) < n$), akkor az egyenletrendszenek több megoldása is van. Az $x_{r+1}^{(r)}, \ldots, x_n^{(n)}$ ismeretlenek helyébe az F test elemeiből képezett tetszőleges $[c_{r+1}, \ldots, c_n]$ rendezett elem $n - r$-esét behelyettesítve, az x_k ($k = 1, \ldots, r$) ismeretlen egyértelműen kifejezhető az

$$x_k = \frac{d_k}{a_{kk}^{(r)}}$$

képlettel, ahol

$$d_k = b_k^{(r)} - (a_{k,r+1}^{(r)}c_{r+1} + \ldots + a_{kn}^{(r)}c_n).$$

Vizsgáljuk a második esetet (azaz, amikor $r < m$).

Ha a $b_{r+1}^{(r)}, \ldots, b_m^{(r)}$ konstansok mindegyike 0 (ekkor $rangA = rang(A, b)$), akkor az egyenletrendszernek van megoldása, mert ekkor az egyenletrendszerből az utolsó $m - r$ darab $0 = 0$ alakú egyenlet elhagyható, s így az első esetnek megfelelő egyenletrendszerhez jutunk (ezt az esetet már elemeztük; az egyenletrendszer megoldható volt).

Ha a $b_{r+1}^{(r)}, \ldots, b_m^{(r)}$ konstansok közül legalább az egyik (mondjuk a j-dik) nem nulla (ekkor $rangA \neq rang(A, b)$), akkor az egyenletrendszer nem oldható meg, mert az tartalmazza az ellentmondást jelentő $0 = b_j^{(r)}$ egyenlőséget.

4.3. A Cramer-szabály

Ebben a fejezetben egy speciális egyenletrendszerre adunk megoldást.

Tétel 4.8 *(A Cramer-szabály) Ha az n egyenletből álló, n ismeretlent tartalmazó*

$$\begin{array}{llllll} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\ldots & +a_{nn}x_n & = & b_n \end{array}$$

lineáris egyenletrendszer

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

mátrixa reguláris, akkor az egyenletrendszer megoldható, mégpedig egyértelműen, és a megoldást a következő képlet szolgáltatja:

$$x_k = \frac{detA_k}{detA} \ (k = 1, \ldots, n),$$

ahol

$$detA_k = \begin{vmatrix} a_{11} & \ldots & a_{1,k-1} & b_1 & a_{1,k+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,k-1} & b_2 & a_{2,k+1} & \ldots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{n,k-1} & b_n & a_{n,k+1} & \ldots & a_{nn} \end{vmatrix}$$

(az A_k mátrixot (az un. k-dik módosított mátrixot) az A mátrixból úgy származtatjuk, hogy annak k-dik oszlopa helyére beírjuk az egyenletrendszer konstansainak oszlopvektorát).

5. fejezet

Mátrixok sajátértékei, sajátvektorai

Definíció 5.1 *Egy F test valamely λ elemét az F test feletti $\mathbf{A}_{n\times n}$ mátrix sajátértékének nevezzük, ha megadható olyan $\underline{x} \in F^n$ vektor, amelyre teljesül az*

$$\mathbf{A}\underline{x} = \lambda\underline{x}$$

egyenlőség. A feltételnek eleget tevő $\underline{x}$ vektorokat az A mátrix sajátvektorainak (pontosabban, a λ sajátértékhez tartozó sajátvektoroknak) nevezzük. Egy mátrix sajátértékeinek halmazát a mátrix spektrumának nevezzük.

Megjegyzés 5.1 *Adott $\mathbf{A} \in F_{n\times n}$ mátrix esetén F^n minden vektora legfeljebb egy sajátértékhez tartozó sajátvektor lehet. Ugyanis, ha $\underline{x}$ egy $\mathbf{A} \in F_{n\times n}$ mátrix $\alpha, \beta \in F$ sajátértékekhez tartozó sajátvektor, akkor*

$$\alpha\underline{x} = \mathbf{A}\underline{x} = \beta\underline{x},$$

amiből $\underline{x} \neq \underline{0}$ miatt $\alpha = \beta$ következik.

Tétel 5.1 *Tetszőleges $\mathbf{A} \in F_{n\times n}$ mátrix tetszőleges sajátértékéhez tartozó sajátvektorok a $\underline{0}$ vektorral együtt az F^n vektortér egy alterét alkotják.*

Bizonyítás. Jelölje W az F^n vektortér azon részhalmazát, amely a $\underline{0}$ vektorból és az $\mathbf{A}$ mátrix valamely λ sajátértékéhez tartozó összes sajátvektoraiból áll. Tetszőleges $\underline{x}, \underline{y} \in W$ és tetszőleges $\xi, \eta \in F$ skalárok esetén

$$\mathbf{A}(\xi\underline{x} + \eta\underline{y}) = \xi(\mathbf{A}\underline{x}) + \eta(\mathbf{A}\underline{y}) = \xi(\lambda\underline{x}) + \eta(\lambda\underline{y}) = \lambda(\xi\underline{x} + \eta\underline{y}).$$

Tehát

$$\xi\underline{x} + \eta\underline{y} \in W.$$

A Tétel **??** miatt W az F^n vektortér egy altere. □

Definíció 5.2 *Egy* $\mathbf{A}$ *mátrix* λ *sajátértékeiből, illetve a nullvektorból álló alteret az* $\mathbf{A}$ *mátrix* λ *sajátértḱhez tartozó un. sajátaltérnek nevezzük. Ennek dimenzióját a* λ *sajátérték geometriai multiplicitásának nevezzük.*

Definíció 5.3 *Legyen* R *egy gyűrű és* A *egy* R *feletti négyzetes mátrix. A*

$$k_{\mathbf{A}}(x) = (-1)^n det(A - xE)$$

polinomot az A *mátrix karakterisztikus polinomjának, a*

$$det(A - xE) = 0$$

egyenletet pedig az A *mátrix karakterisztikus egyenletének nevezzük.*

Tétel 5.2 *Egy* F *test valamely* λ *eleme akkor és csak akkor sajátértéke egy* F *test feletti* A *mátrixnak, ha* λ *gyöke az* A *mátrix karakterisztikus egyenletének, azaz telesül rá a* $det(A - \lambda E) = 0$ *egyenlőség.*

Bizonyítás. $\lambda \in F$ akkor és csak akkor sajátértéke egy $\mathbf{A} \in F_{n\times n}$ mátrixnak, ha van olyan $\underline{0} \neq \underline{x} \in F^n$ vektor, hogy

$$\mathbf{A}\underline{x} = \lambda\underline{x} = \lambda\mathbf{E}\underline{x},$$

azaz

$$[\mathbf{A} - \lambda\mathbf{E}]\underline{x} = \underline{0}.$$

Ezen utóbbi kifejezés egy homogén lineáris egyenletrendszer mátrixszorzatos alakja. λ tehát akkor és csak akkor sajátértéke $\mathbf{A}$-nak, ha ennek az egyenletrendszernek van nem-triviális megoldása, ami a Tétel **??** miatt azzal ekvivalens, hogy

$$det(\mathbf{A} - \lambda\mathbf{E}) = 0.$$

□

Megjegyzés 5.2 *Egy* $\mathbf{A} \in \mathbf{F_{n\times n}}$ *mátrix sajátértékeinek, illetve sajátvektorainak meghatározást́ tehát azzal kezdjük, hogy meghatározzuk a mátrix karakterisztikus egyenletének* F *testbeli gyökeit. Ha vannak ilyenek (jelöljön egy* *sajátértéket* λ*), akkor a* λ*-hoz tartozó sajátvektorok az* $[\mathbf{A} - \lambda\mathbf{E}]\underline{x} = \underline{0}$ *homogén lineáris egyenletrendszer nem-triviális megoldásvektorai lesznek. A megoldásvektorok halmaza (a sajátvektorok halmaza hozzávéve a nullvektort) a Tétel* **??** *miatt az* F^n *vektortér egy altere, ennek dimenziója, azaz a* λ *sajátérték geometriai multiplicitása megegyezik a fenti homogén lineáris egyenletrendszer megoldása során szabadon választható ismeretlenek számával (lásd a Tétel* **??***-t).*

Megjegyzés 5.3 *Egy mátrixnak nincs mindig sajátértéke. Például az*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix} \in (\mathbb{Z}_5)_{2\times 2}$$

mátrix karakterisztikus egyenlete

$$x^2 + x + 2 = 0.$$

Könnyen ellenőrízhető, hogy ennek az egyenletnek nincs gyöke $\mathbb{Z}_5$*-ben, ezért az* $\mathbf{A}$ *mátrixnak nincs sajátértéke* $\mathbb{Z}_5$*-ben.*

Ha az $\mathbf{A}$ *mátrixot, mint* $\mathbb{R}$ *feletti mátrixot tekintjük, akkor karakterisztikus egyenlete*

$$x^2 - 4x - 3 = 0,$$

melynek megoldásai $x_{1,2} = 2 \pm \sqrt{7}$*.*

Definíció 5.4 *Egy* A *mátrix valamely* λ *sajátértékének algebrai multiplicitásán azt a pozitív egász számot értjük, amely megegyezik* λ*-nak, mint a* $det(\mathbf{A} - x\mathbf{E}) = 0$ *karakterisztikus egyenlet gyökének multiplicitásával.*

Megjegyzés 5.4 *Egy* $\mathbf{A}$ *mátrix valamely* λ *sajátértékének geometriai multiplicitása kisebb vagy egyenlő mint az algebrai multiplicitása (lásd).*

Definíció 5.5 *Legyen* $\mathbf{A}$ *egy komplex sz számokból álló négyzetes mátrix. Az* $\mathbf{A}$ *mátrixot önadjungál mátrixnak nevezzük, ha* $\overline{\mathbf{A}}^T = \mathbf{A}$*. Az* $\mathbf{A}$ *mátrixot ferdén önadjungált mátrixnak nevezzük, ha* $\overline{\mathbf{A}}^T = -\mathbf{A}$*.*

Tétel 5.3 *Önadjungált mátrix minden sajátértéke valós szám. Ferdén önadjungált mátrix minden sajátértéke képzetes szám.*

Bizonyítás. Legyen $\mathbf{A}$ egy n-edrendű komplex elemű mátrix. Legey $\lambda = a + ib \in \mathbb{C}$ az $\mathbf{A}$ egy sajátértéke. Akkor van olyan $\underline{0} \neq \underline{v} = \underline{x} + i\underline{y} \in \mathbb{C}^n$ mátrix, hogy

$$\mathbf{A}\underline{v} = \lambda\underline{v}.$$

Szorozzuk be ezt az egyenlőséget balról a $\underline{v}$ vektor konjugáltjának transzponáltjával. Ekkor az

$$\overline{(\underline{v})}^T \mathbf{A}\underline{v} = \lambda(|\underline{x}|^2 + |\underline{y}|^2)$$

egyenlőséget kapjuk. Vegyük mindkét oldal konjugáltjának transzponáltját. Ekkor az

$$\overline{(\underline{v})}^T \overline{\mathbf{A}}^T \underline{v} = \overline{\lambda}(|\underline{x}|^2 + |\underline{y}|^2),$$

azaz

$$\overline{(\underline{v})}^T \mathbf{A}\underline{v} = \overline{\lambda}(|\underline{x}|^2 + |\underline{y}|^2)$$

adódik, mivel $\mathbf{A}$ önadjungált mátrix, $|\underline{x}|^2 + |\underline{y}|^2$ pedig valós szám. A második és az utolsó egyenlőségből azt kapjuk, hogy

$$\lambda(|\underline{x}|^2 + |\underline{y}|^2) = \overline{\lambda}(|\underline{x}|^2 + |\underline{y}|^2),$$

azaz

$$\lambda = \overline{\lambda}.$$

Figyelembe véve λ algebrai alakját,

$$a + ib = a - ib,$$

amiből $b = 0$, aza $\lambda = a$ következik. Igy az $\mathbf{A}$ önadjungált mátrix λ sajátértéke valós szám.

A ferdén önadjungált eset bizonyítása az önadjungált settől csak annyiban különbözik, hogy az

$$\overline{(\underline{v})}^T \overline{\mathbf{A}}^T \underline{v} = \overline{\lambda}(|\underline{x}|^2 + |\underline{y}|^2)$$

egyenlőségből

$$\overline{(\underline{v})}^T (-\mathbf{A})\underline{v} = \overline{\lambda}(|\underline{x}|^2 + |\underline{y}|^2)$$

következik, s ezért $\lambda = -\overline{\lambda}$ miatt $b = 0$, azaz $\lambda = ib$ adódik; tehát ferdén önadjungált esetben a λ sajátérték képzetes szám. □

Definíció 5.6 *Egy valós számokból álló* $\mathbf{A}$ *négyzetes mátrixot szimmetrikus mátrixnak nevezünk, ha* $\mathbf{A} = \mathbf{A}^T$. *Az* $\mathbf{A}$ *mátrixot ferdén szimmetrikus mátrixnak nevezzük, ha* $\mathbf{A} = -\mathbf{A}^T$.

Mivel minden szimmetrikus mátrix önadjungált, és minden ferdén szimmetrikus mátrix ferdén önadjungált, ezért érvényes a Következő tétel.

Tétel 5.4 *Szimmetrikus mátrix minden sajátértéke valós szám. Ferdén szimmetrikus mátrix minden sajátértéke képzetes szám.*

A bázistranszformáció vizsgálatánál kiderült, hogy a mátrixok közötti hasonlóság fontos fogalom. A következőkben hasonló mátrixok sajátértékeinek kapcsolatával foglakozunk.

Tétel 5.5 *Szimmetrikus mátrix különböző sajátértékeihez tartozó sajátvektorok egymásra merőlegesek, azaz skaláris szorzatuk* 0*.*

Bizonyítás. Legyenek $\underline{a}$ és $\underline{b}$ egy $\mathbf{A}$ szimmetrikus mátrix valamely λ és μ sajátértékeihez tartozó sajátvektorok, azaz $\mathbf{A}\underline{a} = \lambda\underline{a}$ és $\mathbf{A}\underline{b} = \mu\underline{b}$. Ekkor $\underline{b}^T\mathbf{A}\underline{a} = \lambda\underline{b}^T\underline{a}$, amiből (mindkét oldal transzponáltját véve) $\underline{a}^T\mathbf{A}^T\underline{b} = \lambda\underline{b}^T\underline{a}$ adódik. Mivel az $\mathbf{A}$ mátrix szimmetrikus, ezért $\lambda\underline{b}^T\underline{a} = \underline{a}^T\mathbf{A}\underline{b} = \underline{a}^T\mu\underline{b}$, amiből $\underline{b}^T\underline{a} = \underline{a}^T\underline{b}$ miatt $(\lambda - \mu)\underline{a}^T\underline{b} = 0$ adódik. Ha $\lambda \neq \mu$, akkor $\underline{a}^T\underline{b} = 0$, azaz $\underline{a}$ és $\underline{b}$ egymásra merőleges vektorok. □

Tétel 5.6 *Hasonló mátrixok karakterisztikus polinomjai egymással megegyeznek.*

Bizonyítás. Legyenek $\mathbf{A}$ és $\mathbf{B}$ hasonló mátrixok. Akkor van olyan reguláris $\mathbf{C}$ mátrix, hogy

$$\mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{B}.$$

Akkor

$$det(\mathbf{B} - x\mathbf{E}) = det(\mathbf{C}^{-1}\mathbf{A}\mathbf{C} - x\mathbf{C}^{-1}\mathbf{E}\mathbf{C}) = det(\mathbf{C}^{-1}[\mathbf{A} - x\mathbf{E}]\mathbf{C}) =$$

$$= det(\mathbf{C}^{-1})det(\mathbf{A}-x\mathbf{E})det(\mathbf{C})det(\mathbf{C}^{-1})det(\mathbf{C})det(\mathbf{C}-\mathbf{A}-x\mathbf{E}) = det(\mathbf{A}-x\mathbf{E}).$$

Ebből már következik, hogy $k_{\mathbf{A}}(x) = k_{\mathbf{B}}(x)$. □

Megjegyzés 5.5 *A Tétel* **??** *állításának megfordítása nem igaz. Például a valós elemű*

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

mátrixok esetén mindkét mátrix karakterisztikus polinomja $x^2 - 2x + 1$*, a mátrixok még sem hasonlóak, hiszen az* $\mathbf{E}$ *egységmátrix csak önmagával hasonló.*

Mátrixok hasonlóságára egy későbbi fejezetben fogunk szükséges és elégséges feltétel adni (Theorem 7.10).

6. fejezet

Vektorterek közötti lineáris leképezések

6.1. A lineáris leképezés fogalma

Definíció 6.1 *Legyenek V_1 és V_2 ugyanazon F test feletti vektorterek. V_1-nek V_2-be való φ leképezését a V_1 vektortérnek a V_2-vektortérbe való lineáris leképezésének nevezzük, ha tetszőleges $a, b \in V_1$ és tetszőleges $\alpha \in F$ esetén teljesülnek az alábbi egyenlőségek:*

$$\varphi(a + b) = \varphi(a) + \varphi(b),$$

$$\varphi(\alpha a) = \alpha\varphi(a).$$

A V_1-ből V_2-be történő összes lineáris leképezés halmazát $Hom(V_1, V_2)$-vel fogjuk jelölni.

Tétel 6.1 *Tetszőleges F test feletti V_1 és V_2 vektorterek esetén $\varphi \in Hom(V_1, V_1)$ akkor és csak akkor, ha tetszőleges $a, b \in V_1$ vektorok és tetszőleges $\alpha, \beta \in F$ skalárok esetén*

$$\varphi(\alpha a + \beta b) = \alpha\varphi(a) + \beta\varphi(b)$$

teljesül.

Bizonyítás. □

Megjegyzés 6.1 *Tétel 6.1 alapján, ha $\{b_1, \ldots, b_n\}$ egy F test feletti V vektortér bázisa, akkor tetszőleges*

$$a = \alpha_1 b_1 + \ldots + \alpha_n b_n \in V$$

vektor és V-nek valamely vektortérbe való φ lineáris leképezése esetén

$$\varphi(a) = \alpha_1(\varphi(b_1)) + \ldots \alpha_n(\varphi(b_n)).$$

Ez és a következő tétel azt mutatja, hogy egy lineáris leképezés egyértelműen meg van határozva a vektortér egy bázisára való leszűkítése által.

Tétel 6.2 *Legyenek V_1 és V_2 ugyanazon F test feletti vektorterek. A V_1 vektortér tetszőleges $\{b_1, \ldots, b_n\}$ bázisához és a V_2 vektortér tetszőleges $c_1, \ldots, c_n$ vektorsorozatához megadható egy és csak egy olyan $\varphi \in Hom(V_1, V_2)$ lineáris leképezés, amelyre*

$$\varphi(b_i) = c_i, \quad i = 1, \ldots, n$$

teljesül.

Bizonyítás. Tetszőleges

$$a = \alpha_1 b_1 + \ldots + \alpha_n b_n \in V_1$$

vektorra definiáljuk $\varphi(a)$-t a következőképpen:

$$\varphi(a) = \alpha_1 c_1 + \ldots + \alpha_n c_n.$$

Világos, hogy φ teljesíti a tétel állításában megfogalmazott feltételeket. Az is világos a Megjegyzés 6.1 szerint, hogy a feltételeket teljesítő lineáris leképezés egyértelműen meg van határozva. □

6.2. Műveletek lineáris leképezések között

Tétel 6.3 *Legyenek V_1 és V_2 ugyanazon F test feletti vektorterek. Tetszőleges $\varphi_1, \varphi_2 \in Hom(V_1, V_2)$ esetén $\varphi_1 + \varphi_2 \in Hom(V_1, V_2)$, ahol a $\varphi_1 + \varphi_2$ leképezés a következőképpen van értelmezve:*

$$(\varphi_1 + \varphi_2) \ : \ a \ \mapsto \ \varphi_1(a) + \varphi_2(a), \quad a \in V_1.$$

Bizonyítás. Tetszőleges $a, b \in V_1$ és tetszőleges $\alpha, \beta \in F$ esetén

$$(\varphi_1 + \varphi_2)(\alpha a + \beta b) = \varphi_1(\alpha a + \beta b) + \varphi_2(\alpha a + \beta b) =$$

$$\alpha\varphi_1(a) + \beta\varphi_1(b) + \alpha\varphi_2(a) + \beta\varphi_2(b) =$$
$$\alpha(\varphi_1(a) + \varphi_2(a)) + \beta(\varphi_1(b) + \varphi_2(b)) =$$
$$\alpha(\varphi_1 + \varphi_2)(a) + \beta(\varphi_1 + \varphi_2)(b).$$

Tehát, a Tétel 6.1 miatt, $\varphi_1 + \varphi_2 \in Hom(V_1, V_2)$. □

Definíció 6.2 *A Tétel 6.3-ben definiált $\varphi_1 + \varphi_2$ lineáris leképezést a φ_1 és φ_2 lineáris leképezések összegének nevezzük.*

Tétel 6.4 *$Hom(V_1, V_2)$ kommutatív csoport a lineáris leképezések összeadására nézve.*

Bizonyítás.

Tétel 6.5 *Legyenek V_1 és V_2 ugyanazon F test feletti vektorterek. Tetszőleges $\varphi \in Hom(V_1, V_2)$ és tetszőleges $\alpha \in F$ esetén $\alpha\varphi \in Hom(V_1, V_2)$, ahol az $\alpha\varphi$ leképezés a következőképpen van értelmezve:*

$$(\alpha\varphi) \ : \ a \ \mapsto \ \alpha(\varphi(a)), \quad a \in V_1.$$

Bizonyítás.

Definíció 6.3 *A Tétel 6.5-ben szereplő $\alpha\varphi$ lineáris leképezést a φ lineáris leképezés α skalárral képezett szorzatának nevezzük.*

Tétel 6.6 *Tetszőleges F test feletti V_1 és V_2 vektorterek esetén $Hom(V_1, V_2)$ vektorteret alkot F felett (a lineáris leképezések összeadására, illetve a lineáris leképezések skalárral képezett szorzására nézve).*

Bizonyítás.

Tétel 6.7 *Legyenek V_1, V_2, V_3 ugyanazon F test feletti vektorterek. Tetszőleges $\varphi_1 \in Hom(V_1, V_2)$ és $\varphi_2 \in Hom(V_2, V_3)$ esetén $\varphi_2\varphi_1 \in Hom(V_1, V_3)$, ahol a $\varphi_2\varphi_1$ leképezés a következőképpen van értelmezve:*

$$(\varphi_2\varphi_1)(a) = \varphi_2(\varphi_1(a)), \quad a \in V_1.$$

Bizonyítás.

Definíció 6.4 *A Tétel 6.7-ben definiált $\varphi_2\varphi_1 \in Hom(V_1, V_3)$ lineáris leképezést a szóban forgó lineáris leképezések szorzatának nevezzük.*

Tétel 6.8 *Tetszőleges F test feletti V vektortér esetén $Hom(V, V)$ gyűrűt alkot a lineáris transzformációk összeadására és szorzására nézve.*

Bizonyítás.

6.3. Vektorterek izomorfizmusa

Definíció 6.5 *Egy $\varphi \in Hom(V_1, V_2)$ lineáris leképezést izomorfizmusnak nevezünk, ha φ szürjektív és injektív. Azt mondjuk, hogy egy V_1 vektortér izomorf egy V_2 vektortérrel, ha megadható V_1-nek V_2-re való izomorfizmusa (ekkor persze V_2 is izomorf V_1-gyel, s ezért azt is szoktuk mondani, hogy V_1 és V_2 egymással izomorfak).*

Tétel 6.9 *Legyenek V_1 és V_2 ugyanazon F test feletti vektorterek, és legyen φ V_1-nek V_2-be való izomorfizmusa. Akkor a V_1 tér tetszőleges $a_1, \ldots, a_m$ vektorai akkor és csak akkor lineárisan függetlenek, ha a V_2 tér $\varphi(a_1), \ldots, \varphi(a_m)$ vektorai lineárisan függetlenek.*

Bizonyítás. Valamely $a_1, \ldots, a_m \in V_1$ vektorokra és $\alpha_1, \ldots, \alpha_m \in F$ skalárokra

$$\alpha_1 a_1 + \ldots + \alpha_m a_m = 0$$

akkor és csak akkor teljesül, ha

$$\alpha_1\varphi(a_1) + \ldots + \alpha_m\varphi(a_m) = 0.$$

Ebből már egyszerűen következik a tétel állítása. □

Tétel 6.10 *Ugyanazon test feletti véges dimenziós vektorterek akkor és csak akkor izomorf egymással, ha azonos dimenziójúak.*

Bizonyítás. Legyenek V_1 és V_2 ugyanazon test feletti véges dimenziós vektorterek. Tegyük fel, hogy $dim(V_1) = dim(V_2)$. Legyen $\{b_1, \ldots, b_n\}$ a V_1 vektortér, $\{c_1, \ldots, c_n\}$ pedig a V_2 vektortér egy-egy bázisa. Akkor van V_1-nek V_2-be olyan φ lineáris leképezése (Tétel 6.2), amely a b_i bázisvektorhoz a c_i bázisvektort rendeli ($i = 1, \ldots, n$). Mivel a $\{c_1, \ldots, c_n\}$ vektorrendszer bázisa V_2-nek, ezért φ szürjektív.

Legyenek
$$a = \alpha_1 b_1 + \ldots + \alpha_n b_n$$
és
$$b = \beta_1 b_1 + \ldots + \beta_n b_n$$
tetszőleges V_1-beli vektorok. Ha $\varphi(a) = \varphi(b)$, akkor
$$\alpha_1 \varphi(b_1) + \ldots + \alpha_n \varphi(b_n) = \beta_1 \varphi(b_1) + \ldots \beta_n \varphi(b_n),$$
és így
$$\alpha_1 c_1 + \ldots + \alpha_n c_n = \beta_1 c_1 + \ldots \beta_n c_n,$$
amiből
$$\alpha_i = \beta_i, \quad i = 1, \ldots n$$
adódik, azaz $a = b$. Tehát φ injective. Következésképpen φ V_1-nek V_2-re való izomorfizmusa.

Fordítva, tegyük fel, hogy V_1 és V_2 egymással izomorfak. A Tétel 6.9 alapján igen egyszerűen adódik, hogy $dim(V_1) = dim(V_2)$, ha figyelembe vesszük, hogy a bázis nem más mint egy maximális lineárisan független vektorrendszer. □

Tétel 6.11 *Legyen V egy F test feletti n-dimenziós vektortér, és legyen $\mathcal{B} = \{b_1, \ldots, b_n\}$ a V egy bázisa. Akkor V-nek az F^n-re való azon φ leképezése, amely a V tetszőleges*
$$a = \alpha_1 b_1 + \ldots + \alpha_n b_n$$
vektorához az
$$[a]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \in F^n$$
elem n-est rendeli, a V vektortérnek az F^n vektorérre való izomorfizmusa.

Bizonyítás. A Tétel 6.10 alapján nyílvánvaló. □

Definíció 6.6 *A Tétel 6.11-ben definiált $\varphi : a \mapsto [a]_{\mathcal{B}}$ hozzárendelést a $\mathcal{B}$ bázis szerinti koordinatizációnak nevezzük.*

Tétel 6.12 *Legyen V egy F test feletti vektortér és $\mathcal{B}$ a V egy bázisa. Akkor tetszőleges $a, b \in V$ vektorok és tetszőleges $\alpha \in F$ skalár esetén teljesülnek az alábbi egyenlőségek:*

$$[a + b]_{\mathcal{B}} = [a]_{\mathcal{B}} + [b]_{\mathcal{B}},$$

$$[\alpha a]_{\mathcal{B}} = \alpha [a]_{\mathcal{B}}.$$

Tétel 6.13 *Egy F test feletti n-dimenziós V vektortér bármely $\mathcal{B}$ bázisa esetén a V tetszőleges $\{a_1, \ldots, a_m\}$ vektorrendszere akkor és csak akkor lineárisan független, ha az F^n vektortér $\{[a_1]_{\mathcal{B}}, \ldots, [a_m]_{\mathcal{B}}\}$ vektorrendszere lineárisan független.*

Bizonyítás. Jelölje φ az $a \mapsto [a]_{\mathcal{B}}$ leképezést (lásd Tétel 6.11). Ez V-nek F^n-re való izomorfizmusa. Igy a Tétel 6.9 felhasználásával igen egyszerűen adódik a jelen tétel állítása. □

6.4. Duális tér, duális bázis

...

6.5. Lineáris leképezések mátrixa

Tétel 6.14 *Legyenek V_1, illetve V_2 ugyanazon F test feletti n, illetve m dimenziós vektorterek. Akkor V_1-nek V_2-be való tetszőlege φ lineáris leképezéséhez és V_1, illetve V_2 egy-egy rğzített $\mathcal{B}_1$, illetve $\mathcal{B}_2$ bázisához van egy és csak egy olyan $m \times n$-típusú, F feletti $[\varphi]_{\mathcal{B}_2/\mathcal{B}_1}$ mátrix, hogy V_1 tetszőleges a vektora esetén*

$$[\varphi(a)]_{\mathcal{B}_2} = [\varphi]_{\mathcal{B}_2/\mathcal{B}_1} [a]_{\mathcal{B}_1}$$

teljesül.

Bizonyítás.

Tétel 6.15

6.6. Bázistranszformáció

Definíció 6.7 *Egy V vektortér önmagába való φ lineáris leképezéseit lineáris transzformációknak nevezzük. Ennek a V vektortér egy $\mathcal{B}$ bázisához tartozó mátrixát $[\varphi]_{\mathcal{B}}$ módon fogjuk jelölni.*

Tétel 6.16 *Egy F test feletti n-dimenziós V vektortér tetszőleges τ lineáris transzformációja és tetszőleges $\mathcal{B}$ bázisa esetén a következő feltételek egymással ekvivalensek.*

1. *A $\mathcal{B}$ bázis τ-szerinti képe is bázisa V-nek.*
2. *$[\tau]_{\mathcal{B}}$ reguláris mátrix.*

Bizonyítás. A Tétel 6.13 és a Tétel 3.14 alapján nyílvánvaló. □

Definíció 6.8 *Egy V vektortér valamely τ lineáris transzformációját bázistranszformációnak nevezzük, ha V-nek van olyan bázisa, melynek τ szerinti képe bźisa V-nek.*

Tétel 6.17 *Legyen V egy n-dimenziós vektortér egy F test felett, és legyen $\mathcal{B} = \{e_1, \ldots, e_n\}$, illetve $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ a V két bázisa. Legyen τ a V vektortér azon bázistranszformációja, amelyre*

$$e'_i = \varphi(e_i), \quad i = 1, \ldots, n$$

teljesül. Akkor tetszőleges $a \in V$ vektor esetén

$$[a]_{\mathcal{B}'} = [\tau]_{\mathcal{B}}^{-1}[a]_{\mathcal{B}}.$$

Bizonyítás. Legyen $a \in V$ tetszőleges vektor. Az a alakja a $\mathcal{B}'$ bázisban (valamely $\alpha'_1, \ldots, \alpha'_n \in F$ skalárokkal):

$$a = \alpha'_1 e'_1 + \ldots + \alpha'_n e'_n = \alpha'_1 \tau(e_1) + \ldots + \alpha'_n \tau(e_n).$$

Ezért (használva a Tétel 6.14 és Tétel 6.12 tételeket is)

$$[a]_{\mathcal{B}} = \alpha'_1 [\tau]_{\mathcal{B}}[e_1]_{\mathcal{B}} + \ldots + \alpha'_n [\tau]_{\mathcal{B}}[e_n]_{\mathcal{B}} =$$

$$= [\tau]_{\mathcal{B}}[a]_{\mathcal{B}'}.$$

Ebből már adódik az

$$[a]_{\mathcal{B}'} = [\tau]_{\mathcal{B}}^{-1}[a]_{\mathcal{B}}$$

egyenlőség, mivel $[\tau]_{\mathcal{B}}$ reguláris (Tétel 6.16), így van inverze (Tétel 3.11). □

Tétel 6.18 *Legyen V egy n-dimenziós vektortér egy F test felett, és legyen $\mathcal{B} = \{e_1, \ldots, e_n\}$, illetve $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ a V két bázisa. Legyen τ a V vektortér azon bázistranszformációja, amelyre*

$$e'_i = \varphi(e_i), \quad i = 1, \ldots, n$$

teljesül. Akkor a V vektortér tetszőleges φ lineáris transzformációja esetén

$$[\varphi]_{\mathcal{B}'} = [\tau]_{\mathcal{B}}^{-1}[\varphi]_{\mathcal{B}}[\tau]_{\mathcal{B}}.$$

Bizonyítás. Legyen φ a V vektortér egy lineáris transzformációja. A Tétel 6.14 szerint φ mátrixa a $\mathcal{B}'$ bázisban a következő alakú:

$$[\varphi]_{\mathcal{B}'} = [[\varphi(e'_1)]_{\mathcal{B}'}, \ldots, [\varphi(e'_n)]_{\mathcal{B}'}] =$$

$$[[\varphi\tau(e_1)]_{\mathcal{B}'}, \ldots, [\varphi\tau(e_n)]_{\mathcal{B}'}] =$$

$$[[\tau]_{\mathcal{B}}^{-1}[\varphi\tau(e_1)]_{\mathcal{B}}, \ldots, [\tau]_{\mathcal{B}}^{-1}[\varphi\tau(e_n)]_{\mathcal{B}}] =$$

$$[[\tau]_{\mathcal{B}}^{-1}[\varphi]_{\mathcal{B}}[\tau]_{\mathcal{B}}[e_1]_{\mathcal{B}}, \ldots, [\tau]_{\mathcal{B}}^{-1}[\varphi]_{\mathcal{B}}[\tau]_{\mathcal{B}}[e_n]_{\mathcal{B}} =$$

$$[\tau]_{\mathcal{B}}^{-1}[\varphi]_{\mathcal{B}}[\tau]_{\mathcal{B}}.$$

Közben felhasználtuk a következő két tételt is: Tétel 6.17 és Tétel 6.15. □

Tétel 6.19 *Egy F test feletti véges dimenziós V vektortér valamely τ lineáris transzformációja akkor és csak akkor bázistranszformáció, ha V bármely bázisának τ szerinti képe bázisa V-nek.*

Bizonyítás. A Tétel 6.16, Tétel 6.18 és Tétel 3.12 alapján nyílvánvaló. □

6.7. Lineáris transzformációk sajátértékei, sajátvektorai

Definíció 6.9 *Legyen V egy F egy test feletti vektortér, és φ a V egy lineáris transzformációja. Az F valamely λ elemét a φ egy sajátértékének nevezzük, ha megadható plyan $\underline{0} \neq \underline{x} \in V$ vektor, hogy*

$$\varphi(\underline{x}) = \lambda \underline{xc}.$$

A fenti egyenlőségben szereplő $\underline{x} \neq \underline{0}$ vektorokat a φ lineáris transzformáció λ sajátértékéhez tartozó sajátvektorának nevezzük.

Tétel 6.20 *Egy F test feletti V vektortér valamely φ lineáris transzformációjának sajátértékei megegyeznek φ tetszőleges V-beli $\mathcal{B}$ bázisához tartozó mátrixának sajátértékeivel.*

Bizonyítás. Legyen $\mathcal{B}$ a V vektortér tetszőleges bázisa. Az F valamely λ eleme akkor és csak akkor sajátértéke φ-nek, ha

$$\varphi(\underline{x}) = \lambda \underline{x}$$

teljesül valamely $\underline{0} \neq \underline{x} \in V$ vektorra. Ezen utóbbi egyenlőség azzal ekvivalens, hogy

$$[\varphi(\underline{x})]_{\mathcal{B}} = [\lambda \underline{x}]_{\mathcal{B}},$$

azaz

$$[\varphi]_{\mathcal{B}}[\underline{x}]_{\mathcal{B}} = \lambda[\underline{x}]_{\mathcal{B}},$$

ami annyit jelent, hogy λ sajátértéke a $[\varphi]_{\mathcal{B}}$ mátrixnak. □

Tétel 6.21 *Egy V vektortér tetszőleges φ lineáris transzformációjának V valamely $\mathcal{B}$ bázisához tartozó mátrixa akkor és csak akkor diagonális, ha $\mathcal{B}$ minden vektora a φ egy-egy sajátvektora. Ha ez a helyzet, akkor a diagonális mátrix főátlójában álló elemek a φ sajátértékei (mindegyik annyiszor szerepel a főátlóban, amennyi az algebrai multiplicitása).*

Bizonyítás. Legyen $[\varphi]_{\mathcal{B}}$ diagonális a V egy

$$\mathcal{B}\{\underline{e}_1, \ldots, \underline{e}_n\}$$

bázisában. Ha a főátlóban álló elemek rendre $\lambda_1, \ldots, \lambda_n$, azaz

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

akkor

$$[\varphi(\underline{e}_i)]_{\mathcal{B}} = [\varphi]_{\mathcal{B}}[\underline{e}_i]_{\mathcal{B}} = [\lambda_i \underline{e}_i]_{\mathcal{B}},$$

azaz

$$\varphi(\underline{e}_i) = \lambda_i \underline{e}_i$$

teljesül minden $i = 1, \ldots, n$ indexre. Tehát a $\mathcal{B}$ bázis elemei a φ lineáris transzformáció sajátvektorai.

Fordítva, tegyük fel, hogy egy F test feletti V vektortér valamely

$$\mathcal{B}\{\underline{e}_1, \ldots, \underline{e}_n\}$$

bázisának elemei a V egy φ lineáris transzformációjának sajátvektorai. Akkor minden $i = 1, \ldots, n$ indexre

$$\varphi(\underline{e}_i) = \lambda_i \underline{e}_i$$

teljesül valamely $\lambda_i \in F$ skalárral. Ebből már világos, hogy

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

azaz φ $\mathcal{B}$ bázisbeli mátrixa diagonális. Mivel ezen diagonális mátrix karakterisztikus egyenlete

$$\prod_{i=1}^{n}(\lambda - \lambda_i),$$

ezért egy-egy sajátérték annyiszor szerepel a főátlóban, amennyi az algebrai multiplicitása. $\square$

7. fejezet

Polinomiális mátrixok

Definíció 7.1 *Polinomiális mátrixon olyan négyzetes mátrixot értünk, amelynek elemei egy adott F gyűrű feletti egyhatározatlanú polinomok; a polinomiális mátrixok tehát egy $R[x]$ polinomgyűrű feletti mátrixok. Egy polinomiális mátrix fokszámán a benne szereplő polinomok fokszámának maximumát értjük; ha a polinomiális mátrix minden eleme az R nulleleme, akkor nem definiálunk hozzá fokszámot.*

Megjegyzés 7.1 *Egy F test feletti négyzetes mátrix karakterisztikus mátrixa elsőfokú polinomiális mátrix.*

Megjegyzés 7.2 *Ha $A(x)$ egy $R[x]$ polinomgyűrű feletti $m \times n$-típusú mátrix úgy, hogy a benne szereplő polinomok fokszámánk maximuma k, akkor $A(x)$ kifejezhető*

$$A(x) = A_k x^k + \cdots + A_1 x + A_0$$

formában, ahol $A_k, \cdots, A_1, A_0$ az R gyűrű feletti $m \times n$-típusú mátrixok ($A_k \neq 0$).

Fordítva, ha $A_k, \cdots, A_1, A_0$ egy R gyűrű feletti $m \times n$-típusú mátrixok és $A_k \neq 0$, akkor egy $A_k x^k + \cdots + A_1 x + A_0$ alakú kifejezés olyan polinomiális mátrixformára hozható, amelyben szereplő polinomok fokszámának maximuma egyenlő k-val.

Ebben a fejezetben csak test feletti, négyzetes polinomiális mátrixokkal foglalkozunk.

Definíció 7.2 *Legyenek $A(x)$ és $B(x)$ egy F test feletti $n \times n$-típusú polinomiális mátrixok. $A(x)$-nek $B(x)$-szel való bal oldali maradékos osztásán* *olyan R feletti (szintén $n \times n$-típusú) $Q(x)$ és $R(x)$ polinomiális mátrixok meghatátozását értjük, amelyekre fennáll az*

$$A(x) = B(x)Q(x) + R(x)$$

egyenlőség, továbbá $R(x) = 0$ vagy $R(x)$ fokszáma kisebb a $B(x)$ fokszámánál. A jobb oldali maradékos osztás fogalma a bal oldali maradékos osztás fogalmának duálisa, azaz a fenti egyenlőség helyett az

$$A(x) = Q(x)B(x) + R(x)$$

egyenlőség teljesülését követeljük meg.

Tétel 7.1 *Ha az F test feletti $n \times n$-típusú $B(x) = B_k x^k + \cdots + B_1 x + B_0$ polinomiális mátrixban szereplő B_k mátrix reguláris, akkor tetszőleges F feletti $n \times n$-típusú $A(x)$ polinomiális mátrixnak a $B(x)$ polinomiális mátrixszal való mindkét oldali maradékos osztása egyértelműen elvégezhető.*

Bizonyítás. Mivel egy F test feletti $n \times n$-típusú mátrixok gyűrűt alkotnak a mátrixok összeadására és szorzására nézve, valamint az F feletti $n \times n$-típusú polinomiális mátrixok olyan polinomok, melynek együtthatói F feletti $n \times n$-típusú mátrixok, ezért a tétel állítása a Tétel 1.29 következménye. □

A következő tétel a karakterisztikus mátrixszokkal való maradékos osztással foglalkozik.

Tétel 7.2 *Legyen*

$$A(x) = A_m x^m + \cdots + A_1 x + A_0$$

egy F test feletti $n \times n$-típusú polinomiális mátrix, B pedig egy F feletti $n \times n$-típusú (konstans) mátrix. Ha

$$A(x) = (xE - B)Q(x) + R(x),$$

akkor

$$R(x) = B^m A_m + B^{m-1} A_{m-1} + \cdots + B A_1 + A_0.$$

Ha

$$A(x) = Q(x)(xE - B) + R(x),$$

akkor

$$R(x) = A_m B^m + A_{m-1} B^{m-1} + \cdots + A_1 B + A_0.$$

Bizonyítás. Elegendő a tétel két állítása közül csak az elsőt bizonyítani, mivel a második bizonyítása teljesen analóg módon történik. A bizonyítás annak közvetlen ellenőrzéséből áll, hogy az $A(x) = (xE - B)Q(x) + R(x)$ egyenlőség érvényben marad, ha az $A(x)$ mátrixot annak tételbeli $A(x) = A_m x^m + \cdots + A_1 x + A_0$ előállításával helyettesítjük, $R(x)$ helyébe behelyettesítjük az $R(x) = B^m A_m + B^{m-1} A_{m-1} + \cdots + BA_1 + A_0$ egyenlőség jobb oldalát, $Q(x)$-nek pedig a

$$Q(x) = A_m x^{m-1} + (BA_m + A_{m-1})x^{m-2} + (B^2 A_m + BA_{m-1} + A_{m-2})x^{m-3} + \cdots +$$

$$+ (B^{m-1} A_m + B^{m-2} A_{m-1} + \cdots + A_1)$$

polinomiális mátrixot választjuk. Ezt az ellenőrzést az olvasóra bízzuk. □

Definíció 7.3 *Egy polinomiális mátrixon végrehajtott elemi átalakításon a következő átalakítások valamelyikét értjük.*

- *Két sor (vagy két oszlop) cseréje.*
- *Valamelyik sornak (vagy oszlopnak) egy nem-nulla konstans polinommal való szorzása.*
- *Egyik sorhoz (vagy oszlophoz) egy tőle különböző sor (vagy oszlop) polinomszorosának hozzáadása.*

Megjegyzés 7.3 *Könnyen belátható, hogy két sor (illetve két oszlop) cseréje a sorokra (oszlopokra) vonatkozó másik két elemi átalkítás kellő számú alkalmazásával elérhető, így a sorok (oszlopok) cseréjét nem szükséges a továbbiakban külön elemezni, elegendő csak a másik két átalakításra koncentrálni. Az i-dik és j-dik sor cseréje a következőképpen is elérhető: adjuk a j-dik sort az i-dikhez, majd az új i-dik sort vonjuk ki az új i-dikből (ekkor az i-dik sorban az eredeti mátrix i-dik és j-dik sorának összege, a j-dik sorban pedig az eredeti mátrix i-dik sorának ellentettje szerepel), majd ezt követően adjuk az előző lépásben megszerkesztett j-dik sort az i-dik sorhoz, végül szorozzuk meg az így kapott j-dik sort -1-gyel.*

Definíció 7.4 *Azokat a polinomiális mátrixokat, amelyek az egységmátrixokból magkaphatók úgy, hogy azokon egy elemi átalakítás végrehajtunk, elemi mátrixoknak nevezzük.*

A következő mátrixok elemi mátrixok.

$$E^i_\alpha = \begin{array}{c} \begin{array}{ccccccccc} & & & & i-dik & & & & \\ & & & & oszlop & & & & \end{array} \\ \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \ddots & 0 & \alpha & 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 1 \end{bmatrix} \end{array}, \quad (\alpha \neq 0)$$

$$E^{i,j} = \begin{array}{c} \begin{array}{ccccccccc} & & & i-dik & & j-dik & & & \\ & & & oszlop & & oszlop & & & \end{array} \\ \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \end{array},$$

$$E^{i<j}_{\varphi(x)} = \begin{array}{c} \begin{array}{ccccccccc} & & & i-dik & & j-dik & & & \\ & & & oszlop & & oszlop & & & \end{array} \\ \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \dots & \varphi(x) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \end{array},$$

$$E_{\varphi(x)}^{i>j} = \begin{array}{c} \begin{array}{ccccccccc} & & & j-dik & & i-dik & & & \\ & & & oszlop & & oszlop & & & \end{array} \\ \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \varphi(x) & \dots & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \end{array}.$$

A fenti mátrixokat a következőképpen kaptuk az egységmátrixból.

1. $\mathbf{E}_\alpha^i$: Az i-dik sor (oszlop) szorzása az $\alpha \neq 0$ skalárral.

2. $\mathbf{E}^{i,j}$: Az i-dik és j-dik sorok (oszlopok) cseréje.

3. $\mathbf{E}^{i<j}$: Az i-dik sorhoz hozzáadtuk a j-dik sor $\varphi(x)$-szeresét, vagy a j-dik oszlophoz hozzáadtuk az i-dik oszlop $\varphi(x)$-szeresét $i < j$.

4. $\mathbf{E}^{i>j}$: Az i-dik sorhoz hozzáadtuk a j-dik sor $\varphi(x)$-szeresét, vagy a j-dik oszlophoz hozzáadtuk az i-dik oszlop $\varphi(x)$-szeresét $i > j$.

Megjegyzés 7.4 *Legyen $\mathbf{A}(x)$ egy F test feletti $n \times n$-típusú polinomiális mátrix. Könnyen ellenőrízhető, hogy ha az $\mathbf{A}(x)$ mátrixot balról (jobbról) szorozzuk egy elemi mátrixszal, akkor a szorzatot úgy is megkaphatjuk, hogy az $\mathbf{A}(x)$ mátrix sorain (oszlopain) végrehajtjuk azt az elemi átalakítást, amelyet az egységmátrix sorain (oszlopain) végrehajtottunk ahhoz, hogy a szóban forgó elemi mátrixot kapjuk.*

Megjegyzés 7.5 *A Megjegyzés 7.4-ben szereplő mátrixok olyan reguláris mátrixok, melyek determinánsa nem függ az x változótól:*

$$Det(\mathbf{E}_\alpha^i) = \alpha \neq 0,$$

$$Det(\mathbf{E}^{i,j}) = -1,$$

$$Det(\mathbf{E}_{\varphi(x)}^{i<j}) = Det(\mathbf{E}_{\varphi(x)}^{i>j}) = 1.$$

Igy ezek a mátrixok invertálhatóak. Könnyen ellenőrízhető, hogy

$$(\mathbf{E}^i_\alpha)^{-1} = \mathbf{E}^i_{1/\alpha},$$

$$(\mathbf{E}^{i,j})^{-1} = \mathbf{E}^{i,j},$$

$$(E^{i<j}_{\varphi(x)})^{-1} = E^{i<j}_{-\varphi(x)},$$

$$(E^{i>j}_{\varphi(x)})^{-1} = E^{i>j}_{-\varphi(x)},$$

Definíció 7.5 *Legyenek $\mathbf{A}(x)$ és $\mathbf{B}(x)$ ugyanazon F test feletti $n \times n$-típusú polinomiális mátrixok. Akkor mondjuk, hogy $A(x)$ ekvivalens $B(x)$-vel (jelölése: $A(x) \sim B(x)$), ha $B(x)$ megkapható $A(x)$-ből véges sok elemi átalakítás egymás után való alkalmazásával.*

Tétel 7.3 *Adott F test feletti $n \times n$-típusú polinomiális mátrixok halmazán értelmezett $\sim$ reláció reflexív, szimmetrikus és tranzitív, azaz ekvivalencia-reláció.*

Bizonyítás. A Megjegyzés 7.4 és Megjegyzés 7.5 alapján nyílvánvaló. □

Definíció 7.6 *Legyen $A(x)$ egy F test feletti $n \times n$-típusú polinomiális mátrix. Jelölje $d_{A,k}(x)$ $(k = 1, \ldots, n)$ az $A(x)$-ből kiválasztható aldeterminánsoknak, mint F feletti polinomoknak a legnagyobb közös osztóját.*

Tétel 7.4 *Ha $A(x)$ és $B(x)$ ugyanazon F test feletti, $n \times n$-típusú, egymással ekvivalens polinomiális mátrixok, akkor tetszőleges $k = 1, \ldots, n$ indexre*

$$d_{A,k}(x) = d_{B,k}(x).$$

Bizonyítás. □

Definíció 7.7 *Kanonikus polinomiális mátrixon olyan F test feletti*

$$\begin{bmatrix} e_1(x) & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & e_2(x) & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & e_3(x) & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & e_{n-2}(x) & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & e_{n-1}(x) & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & e_n(x) \end{bmatrix}$$

diagonális polinomiális mátrixot értünk, amelyre teljesülnek a következő feltételek:

- *Minden $i = 1, \ldots, n-1$ indexre, az $e_i(x)$ polinom osztója az $e_{i+1}(x)$ polinomnak.*

- *Ha valamely $i = 1, \ldots, n$ indexre $e_i(x) \neq 0$ teljesül, akkor $e_i(x)$ főpolinom, azaz főegyütthatója egyenlő 1-gyel (az F test egységelemével).*

Megjegyzés 7.6 *Ha egy kanonikus polinomiális mátrix főátlójában szerepel az F test 0-eleme, akkor az csak a főátló utolsó pozícióiban jelenhet meg. Ugyanígy, ha a főátlóban szerepel az F egységeleme, akkor az csak a főátló első pozícióiban jelenhet meg. Ilyen esetet szemléltet a következő mátrix.*

$$\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & e_3(x) & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & e_{n-2}(x) & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & e_{n-1}(x) & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & 0 \end{bmatrix}.$$

Tétel 7.5 *Minden F test feletti $n \times n$-típusú $A(x)$ polinomiális mátrixhoz van egy és csak egy olyan F feletti $n \times n$-típusú $A^{\sim}(x)$ kanonikus polinomiális mátrix, amelyre $A(x) \sim A^{\sim}(x)$ teljesül (azaz, $A(x)$ és $A^{\sim}(x)$ egymással ekvivalensek).*

Bizonyítás. Legyen $A(x)$ tetszőleges $n \times n$-típusú F test feletti polinomiális mátrix. Először azt mutatjuk meg, hogy a tételben szereplő $A^{\sim}(x)$ mátrix létezik. Ha $A(x)$ az azonosan 0 mátrix, akkor $A^{\sim}(x)$-ként választható az $A(x)$ mátrix.

Tegyük fel, hogy $A(x)$ nem a nullmátrix. Jelölje M az $A(x)$-szel ekvivalens F feletti $n \times n$-típusú polinomiális mátrixok halmazát. A bizonyítás ezen részében az n-re vonatkozó teljes indukciót alkalmazunk.

Ha $n = 1$, akkor $A(x)$ egyetlen F feletti nem-nulla

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \quad (a_n \neq 0)$$

polinomból áll. Ekkor a keresett mátrix az a mátrix, amely az

$$x^n + \frac{a_{n-1}}{a_n} x^{n-1} + \ldots + \frac{a_1}{a_n} x + \frac{a_0}{a_n}$$

polinomból áll.

Legyen $n > 1$ tetszőleges egész szám. Tegyük fel, hogy az állítást már bebizonyítottuk minden n-nél kisebb pozitív egész számra. Az világos, hogy M tartalmaz olyan $B(x) = [\beta_{ij}(x)]$ mátrixot, amelyben $e_1(x) = \beta_{11} \neq 0$ olyan főpolinom, hogy minden $M(x) = [\mu_{ij}(x)] \in M$ mátrixra $e_1(x)$ fokszáma kisebb vagy egyenlő mint $\mu_{11}(x)$ fokszáma. Megmutatjuk, hogy ebben a $B(x)$ mátrixban az $e_1(x)$polinom osztója az első sorban, illetve az első oszlopban szereplő összes többi polinomnak. A bizonyítást csak a sorra végezzük el. Az oszlopra vonatkozó bizonyítás hasonló. Mivel $e_1(x) \neq 0$, ezért $\beta_{1j}(x)$-nek $(j = 2, \ldots, n)$ $e_1(x)$-szel való maradékos osztása elvégezhető:

$$\beta_{1j}(x) = e_1(x)q_j(x) + r_j(x),$$

ahol $q_j(x), r_j(x) \in F[x]$, valamint $r_j(x) = 0$ vagy $r_j(x)$ fokszáma kisebb $e_1(x)$ fokszámánál. Az utóbbi eset nem állhat fenn, mert akkor M tartalmazna olyan mátrixot, melyben az első sor első eleme $r_j(x)$ lenne (ez a mátrix az $M(x)$ mátrixból úgy nyerhető, hogy annak j-dik oszlopából kivonjuk az első oszlop $q_j(x)$-szeresét, majd az így kapott mátrixban felcseréljük az első és a j-dik oszlopot). Mivel $r_j(x)$ fokszáma kisebb az $e_1(x)$ fokszámánál, ez ellentmondást jelentene. Tehát csak az $r_j(x) = 0$ eset állhat fenn, és így $\mu_{1j}(x) = e_1(x)q_j(x)$, azaz $e_1(x)$ osztója $\mu_{1j}(x)$-nek. Ha a j-dik $(j = 2, \ldots, n)$ oszlopból kivonjuk az első oszlop $q_j(x)$-szeresét, akkor olyan $B(x) \in M$ mátrixot kapunk, amelyben az első sor első eleme $\mu_{11}(x)$, a többi pedig 0. A sorokra vonatkozó hasonló átalakítások után pedig olyan $C(x) \in M$ mátrixot kapunk, amelyben az első sor első eleme $e_1(x)$, az első sor, illetve első oszlop többi eleme viszont 0, azaz

$$\begin{bmatrix} e_1(x) & 0 & \ldots & 0 \\ 0 & c_{22}(x) & \ldots & c_{2n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & c_{n2}(x) & \ldots & c_{nn}(x) \end{bmatrix}$$

alakú. Alkalmazzuk az indukciós feltételt az $(n-1) \times (n-1)$-típusú

$$C'(x) = \begin{bmatrix} c_{22}(x) & \ldots & c_{2n}(x) \\ \vdots & \ddots & \vdots \\ c_{n2}(x) & \ldots & c_{nn}(x) \end{bmatrix}$$

mátrixra. $C'(x)$ ekvivalens egy

$$\begin{bmatrix} e_2(x) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e_n(x) \end{bmatrix}$$

kanonikus mátrixszal. Megmutatjuk, hogy $e_1(x)$ osztója $e_2(x)$-nek. Végezzük el az $e_2(x)$ polinomnak az $e_1(x)$ polinommal való maradékos osztását. Akkor

$$e_2(x) = e_1(x)q(x) + r(x) \quad (q(x), r(x) \in F[x]).$$

Tegyük fel, hogy $r(x) \neq 0$. Az $A(x)$-szel ekvivalens

$$\begin{bmatrix} e_1(x) & 0 & \dots & 0 \\ 0 & e_2(x) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e_n(x) \end{bmatrix}$$

mátrix első oszlopának $q(x)$-szeresét adjuk hozzá a második oszlophoz, majd az így keletkezett mátrix első sorát vonjuk ki a második sorból. Ezzel olyan M-beli (azaz, $A(x)$-szel ekvivalens) mátrixot kapunk, amelyben a második sor második eleme egyenlő $r(x)$-szel. Az első két sor, majd az első két oszlop cseréjével olyan M-beli mátrixot kapunk, melyben az első sor első eleme $r(x)$. Ebből $r(x) = 0$ következik. Igy $e_2(x) = e_1(x)q(x)$. Tehát $e_1(x)$ osztója az $e_2(x)$ polinomnak. Ezért az

$$A^{\sim}(x) = \begin{bmatrix} e_1(x) & 0 & \dots & 0 \\ 0 & e_2(x) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e_n(x) \end{bmatrix}$$

mátrix kanonikus mátrix. Ez a mátrix ekvivalens az $A(x)$ polinomiális mátrixszal.

A bizonyítás hátralévő részében az egyértelműség igazolásával foglalkozunk. Legyen

$$A^{*}(x) = \begin{bmatrix} f_1(x) & 0 & \dots & 0 \\ 0 & f_2(x) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_n(x) \end{bmatrix}$$

is olyan kanonikus mátrix, amely (az $A^{\sim}(x)$ mátrixszal együtt) ekvivalens az $A(x)$ polinomiális mátrixszal.

A Tétel 7.4 miatt minden $k = 1, \ldots, n$ indexre

$$e_1(x) \cdots e_k(x) = d_{A^\sim,k}(x) = d_{A,k}(x) = d_{A^*,k}(x) = f_1(x) \cdots f_k(x).$$

Jelölje r az $A(x)$ (és így az $A^\sim(x)$ és az $A^*(x)$ mátrixok) rangját. Akkor

$$e_1(x) = d_{A,1}(x) = f_1(x),$$
$$e_i(x) = \frac{d_{A,i}(x)}{d_{A,i-1}(x)} = f_i(x) \quad (i = 2, \ldots, r)$$
$$e_j(x) = 0 = f_j(x) \quad (j = r+1, \ldots, n).$$

Ezzel az egyértelműséget is bebizonyítottuk. □

Definíció 7.8 *Egy $A(x)$ polinomiális mátrixszal ekvivalens (egyértelműen meghatározott) $A^\sim(x)$ kanonikus mátrixot az $A(x)$ mátrix kanonikus alakjának nevezzük.*

Definíció 7.9 *Egy $A(x)$ polinomiális mátrix kanonikus alakjában szereplő főátlóbeli $e_1(x), \ldots, e_n(x)$ polinomokat az $A(x)$ mátrix invariáns faktorainak nevezzük.*

Definíció 7.10 *Unimoduláris mátrixon olyan polinomiális mátrixot értünk, melynek kanonikus alakja az egységmátrix.*

Tétel 7.6 *Egy F test feletti polinomiális mátrix akkor és csak akkor unimoduláris, ha determinánsa egyenlő az F test egy nem-nulla elemeével (azaz, a determinánsa egy nem nulla konstans polinom).*

Bizonyítás. Egy $n \times n$-típusú $A(x)$ polinomiális mátrix (definíció szerint) unimoduláris, ha kanonikus alakja az $n \times n$-típusú egységmátrix. Ez akkor és csak akkor teljesül, ha $d_n(x) = 1$, ami azzal ekvivalens, hogy $A(x)$ determinánsa az F test valamely nem-nulla eleme. □

Tétel 7.7 *Unimoduláris mátrixok szorzata unimoduláris.*

Bizonyítás. Legyenek $A_1(x), \ldots, A_k(x)$ unimoduláris polinomiális mátrixok, azaz az $A_i(x)$ $(i = 1, \ldots, k)$ mátrixok determinánsa nem-nulla konstans polinom. A determinánsok szorzástétele miatt

$$det(A_1(x) \cdots A_k(x)) = det(A_1(x)) \cdots det(A_k(x)) = konstans \neq 0.$$

□

Tétel 7.8 *Egy polinomiális mátrix akkor és csak akkor unimoduláris, ha elemi mátrixok szorzata.*

Bizonyítás. Legyen $A(x)$ unimoduláris mátrix. Akkor megadhatók olyan $U_i(x)$ $(i = 1, \ldots m)$ és $V_j(x)$ $(j = 1, \ldots k)$ elemi mátrixok, hogy

$$A(x) = U_1(x) \cdots U_m(x) E V_1(x) \cdots V_k(x),$$

azaz $A(x)$ elemi mátrixok szorzata.

Fordítva, tegyük fel, hogy egy $A(x)$ polinomiális mátrix véges sok elemi mátrix szorzata. Mivel az elemi mátrixok unimodulárisak, ezért azok szorzata, és így az $A(x)$ mátrix is unimoduláris a Tétel 7.7 miatt. $\square$

Tétel 7.9 *Valamely $A(x)$ és $B(x)$ polinomiális mátrixok akkor és csak akkor ekvivalensek, ha megadhatók olyan $U(x)$ és $V(x)$ unimoduláris mtrixok, hogy $U(x)A(x)V(x) = B(x)$.*

Bizonyítás. Legyenek $A(x)$ és $B(x)$ egy F test feletti polinomiális mátrixok.

Tegyük fel, hogy $A(x)$ és $B(x)$ ekvivalensek. Akkor megadhatók olyan $U_i(x)$ $(i = 1, \ldots m)$ és $V_j(x)$ $(j = 1, \ldots k)$ elemi mátrixok, hogy

$$(U_1(x) \cdots U_m(x))A(x)(V_1(x) \cdots V_k(x)) = B(x).$$

A Tétel 7.7 miatt $U(x) = U_1(x) \cdots U_m(x)$ és $V(x) = V_1(x) \cdots V_k(x)$ unimoduláris mátrixok.

Fordítva, tegyük fel, hogy valamely F feletti $U(x)$ és $V(x)$ unimoduláris mátrixszal $U(x)A(x)V(x) = B(x)$ teljesül. A Tétel 7.7 miatt megadhatók olyan $U_i(x)$ $(i = 1, \ldots m)$ és $V_j(x)$ $(j = 1, \ldots k)$ elemi mátrixok, hogy

$$U(x) = U_1(x) \cdots U_m(x)$$

és

$$V(x) = V_1(x) \cdots V_k(x),$$

azaz

$$(U_1(x) \cdots U_m(x))A(x)(V_1(x) \cdots V_k(x)) = B(x).$$

Elvégezve az $A(x)$ mátrix sorain az $U_m(x), \ldots, U_1(x)$ elemi mátrixoknak, oszlopain pedig a $V_1(x), \ldots, V_k(x)$ mátrixoknak megfelelő elemei átalakításokat, eredményként a $B(x)$ mátrix adódik. Tehát $A(x)$ és $B(x)$ ekvivalens polinomiális mátrixok. $\square$

Tétel 7.10 *Egy F test feletti $n\times n$-típusú A és B mátrix akkor és csak akkor hasonló, ha karakterisztikus mátrixaik (mint F feletti polinomiális mátrixok) egymással ekvivalensek, azaz karakterisztikus mátrixaik kanonikus alakjai megegyeznek.*

Bizonyítás. Legyenek A és B egy F test elemeiből képezett $n \times n$-típusú mátrixok.

Tegyük fel, hogy A és B hasonlóak. Akkor (definíció szerint) megadható olyan reguláris (F feletti) C mátrix, hogy $C^{-1}AC = B$. Ezért

$$[B - \lambda E] = [C^{-1}AC - C^{-1}(\lambda E)C] = C^{-1}[A - \lambda E]C.$$

Mivel C reguláris mátrix, melynek elemei F-ből valók, azaz F feletti konstans polinomok, ezért C és C^{-1} unimoduláris polinomiális mátrixok. A Tétel 7.9 miatt A és B karakterisztikus mátrixai ekvivalensek.

Fordítva, tegyük fel, hogy A és B karakterisztikus mátrixai ekvivalensek. A Tétel 7.9 miatt megadhatók olyan $U(\lambda)$ és $V(\lambda)$ unimoduláris mátrixok, hogy

$$U(\lambda)[A - \lambda E]V(\lambda) = [B - \lambda E].$$

A Tétel **??** miatt megadhatók olyan F feletti $Q_1(\lambda), Q(\lambda), R_1(\lambda), R_2(\lambda)$ polinomiális mátrixok, hogy

$$U(\lambda) = [B - \lambda E]Q_1(\lambda) + R_1(\lambda)$$

és

$$V(\lambda) = Q_2(\lambda)[B - \lambda E] + R_2(\lambda).$$

Az világos, hogy ha $R_1(\lambda)$ és $R_2(\lambda)$ elemei konstans polinomok, azaz az F test elemei. Megmutatjuk, hogy $R_1[A - \lambda E]R_2 = [B - \lambda E]$. Ebből már fog következni a tétel állítása, mert az $R_1[A - \lambda E]R_2 = [B - \lambda E]$ egyenlőségből $R_1AR_2 - R_1R_2\lambda = B - E\lambda$ adódik, amelyből a mátrixpolinomok egyenlőségének definíciója miatt

$$R_1AR_2 = B$$

és

$$R_1R_2 = E$$

következik. Az utóbbi egyenlőség azt mutatja, hogy $R_1 = R_2^{-1}$, és így kapjuk az

$$R_2^{-1}AR_2 = B$$

egyenlőséget, azaz A és B hasonló mátrixok. Nincs tehát más hátra csask annak igazolása, hogy $R_1[A-\lambda E]R_2 = [B-\lambda E]$. A ... miatt $R_1[A-\lambda E]R_2 =$ ... $\square$

8. fejezet

Jordan-mátrix

Definíció 8.1 *Legyen F egy test és k egy pozitív egész szám. Az F valamely λ_0 eleméhez tartozó k-adrendű alsó Jordan-tömbön a következő $k \times k$-típusú mátrixot értjük:*

$$\begin{bmatrix} \lambda_0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & \lambda_0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \lambda_0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & \lambda_0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & \lambda_0 \end{bmatrix}.$$

Egy λ_0 elemhez tartozó alsó Jordan-tömb tehát egy olyan mátrix, melynek főátlójában mindenütt λ_0 áll és alattuk (a legalsó kivételével) minden elem 1; a mátrix összes többi eleme pedig egyenlő 0-val. A következő mátrixok különböző rendű Jordan-tömbök:

$$[\,3\,], \quad \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{bmatrix}.$$

Definíció 8.2 *n-edrendű Jordan-mátrixon olyan $n\times n$-típusú mátrixot értünk, melyben a főátló mentén Jordan-tömbök állnak, az összes többi elem pedig* 0.

A következő mátrix egy 6-odrendű Jordan-mátrix:

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix},$$

amelyben a főátló mentén a következő Jordan-tömbök állnak:

$$\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}, \quad [\,5\,], \quad \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{bmatrix}.$$

A következőkben meg fogjuk vizsgálni, hogy melyek azok a mátrixok, amelyek hasonlóak egy Jordan mátrixhoz. A Tétel 7.10 szerint ehhez világosan kell látnunk, hogy milyen alakú egy Jordan mátrix karakterisztikus mátrixának kanonikus alakja. Először a Jordan-tömböket vizsgáljuk.

Tétel 8.1 *Egy F test λ_0 eleméhez tartozó k-adrendű Jordan tömb karakterisztikus mátrixának kanonikus alakja*

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & (\lambda-\lambda_0)^k \end{bmatrix}.$$

Bizonyítás. Egy k-adrendű

$$\begin{bmatrix} \lambda_0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & \lambda_0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \lambda_0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & \lambda_0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & \lambda_0 \end{bmatrix}$$

Jordan-tömb

$$\begin{bmatrix} \lambda_0-\lambda & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & \lambda_0-\lambda & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \lambda_0-\lambda & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_0-\lambda & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & \lambda_0-\lambda & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & \lambda_0-\lambda \end{bmatrix}$$

karakterisztikus mátrixa esetén

$$d_1(x) = d_2(x) = \ldots = d_{k-1}(x) = 1$$

és

$$d_k(x) = (\lambda-\lambda_0)^k.$$

Ezért a karakterisztikus mátrix kanonikus alakjában

$$e_1(x) = e_2(x) = \ldots = e_{k-1} = 1$$

és

$$e_k(x) = (\lambda-\lambda_0)^k,$$

azaz a kanonikus alak

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & (\lambda-\lambda_0)^k \end{bmatrix}.$$

□

Megjegyezzük, hogy az előző tételben szereplő kanonikus alakból az eredeti Jordan-tömb felírható. Például, az

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & (\lambda-3)^4 \end{bmatrix}$$

mátrix a $\lambda_0 = 3$-hoz tartozó negyedrendű

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Jordan-tömb karakterisztikus mátrixának kanonikus alakja.

Legyünk óvatosak, mert például az

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & (\lambda-3)^5 \end{bmatrix}$$

polinomiális mátrix kanonikus mátrix, de nem egy Jordan-tömb karakterisztikus mátrixának kanonikus alakja, ugyanis a $\lambda - 3$ kitevője (ami a példánkban 5) nem egyezik meg a mátrix rendjével (azaz 4-gyel).

Mielőtt megvizsgáljuk a Jordan-mátrixok karakterisztikus mátrixának kanonikus alakját, bebizonyítunk egy lemmát, melyet a vizsgálat során fogunk használni.

Lemma 8.1 *Ha $\varphi_1(x), \ldots, \varphi_n(x)$ egy $\mathbb{F}$ test feletti páronként relatív prím polinomok, akkor a*

$$\begin{bmatrix} \varphi_1(x) & 0 & \ldots & 0 & 0 \\ 0 & \varphi_2(x) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \varphi_{n-1}(x) & 0 \\ 0 & 0 & \ldots & 0 & \varphi_n(x) \end{bmatrix}$$

polinomiális mátrix ekvivalens az

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & \varphi_1(x) \cdots \varphi_n(x) \end{bmatrix}$$

mátrixszal.

Bizonyítás. Mivel $\varphi_1(x)$ és $\varphi_2(x)$ az $\mathbb{F}$ test feletti relatív prím polinomok, ezért megadhatók olyan $\mathbb{F}$ feletti $u(x)$ és $v(x)$ polinomok, hogy

$$u(x)\varphi_1(x) + v(x)\varphi_2(x) = 1.$$

Adjuk a

$$\begin{bmatrix} \varphi_1(x) & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \varphi_2(x) & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \varphi_3(x) & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \varphi_4(x) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varphi_{n-1}(x) & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \varphi_n(x) \end{bmatrix}$$

mátrix első oszlopának $u(x)$-szeresét a második oszlophoz, majd az így keletkezett mátrix második sorának $v(x)$-szeresét az első sorhoz. Ekkor a

$$\begin{bmatrix} \varphi_1(x) & u(x)\varphi_1(x) + v(x)\varphi_2(x) & 0 & 0 & \dots & 0 & 0 \\ 0 & \varphi_2(x) & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \varphi_3(x) & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \varphi_4(x) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varphi_{n-1}(x) & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \varphi_n(x) \end{bmatrix},$$

azaz a

$$\begin{bmatrix} \varphi_1(x) & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \varphi_2(x) & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \varphi_3(x) & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \varphi_4(x) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varphi_{n-1}(x) & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \varphi_n(x) \end{bmatrix}$$

mátrixot kapjuk. Cseréljük fel az első két oszlopot, majd vonjuk ki a második oszlopból az eslő oszlop $\varphi_1(x)$-szeresét. Ezzel az

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ \varphi_2(x) & -\varphi_1(x)\varphi_2(x) & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \varphi_3(x) & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \varphi_4(x) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varphi_{n-1}(x) & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \varphi_n(x) \end{bmatrix}$$

mátrixot kapjuk. Ha ennek a mátrixnak a második sorából kivonjuk az első sor $\varphi_2(x)$-szeresét, majd az így keletkezett mátrix második sorát (-1)-gyel szorozzuk, akkor az

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \varphi_1(x)\varphi_2(x) & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \varphi_3(x) & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \varphi_4(x) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varphi_{n-1}(x) & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \varphi_n(x) \end{bmatrix}$$

mátrixszot kapjuk. Mivel $\varphi_3(x)$ relatív prím $\varphi_1(x)$-hez és $\varphi_2(x)$-hez, ezért $\varphi_3(x)$ relatív prím a $\varphi_1(x)\varphi_2(x)$ szorzathoz is (Tétel **??**). Az előző részben alkalmazott átalakításokat használva, megmutatható, hogy ezen utóbbi mátrix ekvivalens az

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \varphi_1(x)\varphi_2(x)\varphi_3(x) & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \varphi_4(x) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \varphi_{n-1}(x) & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & \varphi_n(x) \end{bmatrix}$$

mátrixszal. A gondolatmenetet folytatva adódik a lemma állítása. □

Tétel 8.2 *Legyen J olyan n-edrendű Jordan-mátrix, amely egy $\mathbb{F}$ test páronként különböző*

$$\lambda_1, \dots, \lambda_t$$

elemihez tartozó Jordan-tömbökből áll mégpedig úgy, hogy a λ_i $(i = 1, \ldots, t)$ *elemhez tartozó Jordan-tömbök száma* q_i*, s az egyes tömbök mérete*

$$k_{i1} \geq k_{i2} \geq \ldots \geq k_{iq_i}.$$

Legyen

$$q = \max\{q_1, \ldots, q_t\}.$$

Ha valamely i *indexre* $(i = 1, \ldots, t)$ $q_i < q$ *akkor legyen*

$$k_{iq_{i+1}} = \cdots = k_{iq} = 0.$$

Képezzük tetszőleges $j = 1, \ldots, q$ *indexekre az*

$$e_{n-j+1}(\lambda) = (\lambda - \lambda_1)^{k_{1j}} (\lambda - \lambda_2)^{k_{2j}} \cdots (\lambda - \lambda_t)^{k_{tj}}$$

polinomot. Akkor a J *mátrix karakterisztikus mátrixának kanonikus alakja*

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & e_{n-q+1}(\lambda) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & e_{n-1}(\lambda) & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & e_n(\lambda) \end{bmatrix}.$$

Bizonyítás. Használjuk a Tétel 8.1 eredményét: csak az egyes Jordan-tömbökhöz tartozó sorokon, illetve oszlopokon végezve elemi átalakításokat, a J mátrix karakterisztikus mátrixa ekvivalens azzal a polinomiális mátrixszal, amelyben az egyes tömbnek megfelelő helyeken

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & (\lambda - \lambda_i)^{k_{ij}} \end{bmatrix}$$

alakú mátrixok állnak $(i = 1, \ldots, t;\ \lambda_i \in \mathbb{F};\ j = 1, \ldots, q_i)$. Sorok és oszlopok cseréjét végezve könnyen adódik, hogy ez a mátrix ekvivalens azzal a diagonális mátrixszal, amelyben (felülről számítva) a főátló első $n-(q_1+\cdots+q_t)$

helyén mindenütt 1-es áll, a többi helyen pedig a $(\lambda - \lambda_i)^{k_{ij}}$ alakú hatványok állnak. ($i = 1, \ldots, t$; $\lambda_i \in \mathbb{F}$; $j = 1, \ldots, q_i$). Feltehetjük, hogy ezek a hatványok a következőképpen helyezkednek el. Az egyesek után először az $e_{n-q+1}(\lambda)$ polinom, majd az $e_{n-q+2}(\lambda)$ polinom, és így tovább, az $e_{n-1}(\lambda)$ polinom, majd végül az $e_n(\lambda)$ polinom tényezői következnek (egy polinom tényezői tetszőleges sorrendben követhetik egymást).

Mivel a $\lambda_1, \ldots, \lambda_t$ elemek páronként különbözőek, ezért az $e_{n-j+1}(\lambda)$ ($j = 1, \ldots, q$) polinomok főátlóban egymást követő tényezői páronként relatív prímek, ezért a Lemma 8.1 miatt az általunk vizsgát karakterisztikus mátrix ekvivalens azzal a mátrixszal, amelyben az $e_{n-j+1}(\lambda)$ polinomok tényezői által meghatározott főátlóbeli blokkok helyén az

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots 1 & 0 & \\ 0 & 0 & \ldots 0 & e_{n-j+1} & \end{bmatrix}$$

mátrixok állnak, az összes többi elem pedig 0. Sorok és oszlopok megfelelő cseréjével elérhetjük, hogy az eredeti karakterisztikus mátrixszal ekvivalens

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & e_{n-q+1}(\lambda) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & e_{n-1}(\lambda) & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & e_n(\lambda) \end{bmatrix}$$

mátrixot kapjuk. Ez kanonikus mátrix. Ezzel a tételt bebizonyítottuk. □

Tétel 8.3 *Két Jordan mátrix akkor és csak akkor hasonló, ha ugyanazokból a Jordan-tömbökből állnak, s esetleg csak a Jordan tömbök főátlóbeli elhelyezkedésének sorrendjében különböznek.*

Tétel 8.4 *Egy F test feletti $\mathbf{A}$ mátrix akkor és csak akkor hasonló egy F test feletti Jordan-mátrixszal, ha az $\mathbf{A}$ mátrix minimálpolinomja F feletti elsőfokú polinomok szorzatára bontható.*

Bizonyítás. Egy F test feletti $\mathbf{A}$ mátrix akkor és csak akkor hasonló egy F feletti Jordan mátrixszal, ha az $\mathbf{A}$ mátrix karakterisztikus mátrixának kanonikus alakja megegyezik egy Jordan mátrix karakterisztikus mátrixának kanonikus alakjával. Ez viszont azzal ekvivalens, hogy az $\mathbf{A}$ mátrix minimálpolinomja F feletti elsőfokú polinomok szorzatára bontható.

Tétel 8.5 *Egy F test feletti* $\mathbf{A}$ *mátrix akkor és csak akkor hasonló egy F feletti diagonális mátrixszal, ha az* $\mathbf{A}$ *mátrix minimálpolinomja F feletti elsőfokú polinomok szorzatára bontható, s a minimálpolinom minden gyöke egyszeres.*

Bizonyítás. A diagonális mátrixok speciális Jordan mátrixok; elsőrendű Jordan-tömbökből állnak. Igy állításunk az előző tétel miatt nyílvánvaló. $\square$

9. fejezet

Vektortérkonstrukciók

9.1. Faktortér

Legyen V egy F test feletti vektortér és W a V egy altere. Ekkor W a V kommutatív csoport egy (normális) részcsoportja. Tekintsünk egy tetszőleges W részcsoport szerinti $a + W$ mellékosztályt. Akkor tetszőleges $\alpha \in F$ skalárra

$$\alpha(a + W) = \alpha a + \alpha W \subseteq \alpha a + W.$$

Emiatt

$$\alpha(a + W) = \alpha a + W$$

egy jól definiált szorzás az F test elemei és a V/W faktorcsoport elemei között.

Tétel 9.1 *Legyen F tetszőleges test, V vektortér F felett és W a V egy altere. Akkor a V/W faktorcsoport vektorteret alkot F felett (az előzőekben definiált skalárral való szorzásra nézve). Ebben a vektortérben a $W = 0 + W$ mellékosztály a nullvektor.*

Bizonyítás.

Definíció 9.1 *Az előző tételben szereplő vektorteret a V vektortér W altere szerinti faktorterének nevezzük, és V/W módon jelöljük.*

Tétel 9.2 *Tetszőleges F test feletti tetszőleges véges dimenziós V vektortér tetszőleges W altere esetén*

$$dimV = dimW + dim(V/W).$$

Bizonyítás. Legyen $dimV = n$ és $dimW = m$.

Ha $n = m$, akkor $V = W$, s ebben az esetben $dim(V/W) = 0$. Ekkor tehát fennáll a bizonyítand'øegyenlőség.

Vizsgáljuk az $m < n$ esetet. Legyen $\{e_1, \ldots, e_m\}$ a W altér egy bázisa. Akkor ez kiegészíthető olyan $V - W$-beli $f_{m+1}, \ldots, f_n$ vektorokkal, hogy a $\{e_1, \ldots, e_m, f_{m+1}, \ldots, f_n\}$ vektorrendszer a V vektortér egy bázisa. Megmutatjuk, hogy a V/W faktortérben $\{f_{m+1}+W, \ldots, f_n+W\}$ egy bázis. Legyen $a+W$ tetszőleges V/W-beli vektorvektor ($a \in V$). Akkor megadhatók olyan

$$\alpha_1, \ldots, \alpha_m, \beta_{m+1}, \ldots, \beta_n \in F$$

skalárok, hogy

$$a = \alpha_1 e_1 + \ldots + \alpha_m e_m + \beta_{m+1} f_{m+1} + \ldots + \beta_n f_n.$$

Ebből

$$a + W = (\alpha_1 e_1 + \ldots + \alpha_m e_m + \beta_{m+1} f_{m+1} + \ldots + \beta_n f_n) + W =$$

$$= (\alpha_1 e_1 + W) \ldots + (\alpha_m e_m + W) + (\beta_{m+1} f_{m+1} + W) + \ldots + (\beta_n f_n + W) =$$

$$= \alpha_1 (e_1 + W) + \ldots + \alpha_m (e_m + W) + \beta_{m+1} (f_{m+1} + W) + \ldots + \beta_n (f_n + W) =$$

$$= \alpha_1 W + \ldots + \alpha_m W + \beta_{m+1} (f_{m+1} + W) + \ldots + \beta_n (f_n + W) =$$

$$= \beta_{m+1} (f_{m+1} + W) + \ldots + \beta_n (f_n + W)$$

adódik. Tehát az $\{f_{m+1} + W, \ldots, f_n + W\}$ vektorrendszer generálja a V/W faktorteret. Megmutatjuk, hogy ez a vektorrendszer lineárisan független. Tegyük fel, hogy

$$\gamma_{m+1}(f_{m+1} + W) + \ldots + \gamma_n (f_n + W) = 0 + W$$

valamely $\gamma_{m+1}, \ldots, \gamma_n \in F$ skalárokra. Ez azt jelenti, hogy

$$(\gamma_{m+1} f_{m+1} + \ldots + \gamma_n f_n) + W = 0 + W,$$

és így

$$(\gamma_{m+1} f_{m+1} + \ldots + \gamma_n f_n) = 0$$

teljesül a V vektortérben. Mivel az $f_{m+1}, \ldots, f_n$ vektorok a V vektortér egy bázisának elemei, ezért azok lineárisan függetlenek, és így

$$f_{m+1} = \ldots = f_n = 0.$$

Tehát a $\{f_{m+1} + W, \ldots, f_n + W\}$ vektorrendszer lineárisan független. Mivel a vizsgált vektorrendszer lineárisan független generátorrendszer, ezért bázis. Következésképpen $dim(V/W) = n - m$. Tehát

$$dimV = n = m + (n - m) = dimW + dim(V/W).$$

□

9.2. Vektorterek direkt összege

Tétel 9.3 *Legyenek V_1 és V_2 ugyanazon F test feletti vektorterek. Akkor a V_1 és V_2 additív Abel-csoportok $V_1 \oplus V_2$ külső direkt összege vektorteret alkot F felett, ahol a skalárral való szorzás a következőképpen van értelmezve: tetszőleges $\alpha \in F$ és tetszőleges $(\underline{a}, \underline{b}) \in V_1 \oplus V_2$ esetén $\alpha(\underline{a}, \underline{b}) = (\alpha\underline{a}, \alpha\underline{b})$.*

Bizonyítás. Mivel egy α skalárral való szorzás a koordináták α-val való szorzásával van értelmezve, ezért a vektortér definíciójában a skalárral való szorzásra vonatkozó négy feltétel nyílvánvalóan teljesül. □

Definíció 9.2 *A Tétel 9.3-ben szereplő $V_1 \oplus V_2$ vektorteret a V_1 és V_2 vektorterek külső direkt összegének nevezzük,*

Tétel 9.4 *Tetszőleges F test feletti véges dimenziós V_1 és V_2 vektorterek esetén $dim(V_1 \oplus V_2) = dimV_1 + dimV_2$.*

Bizonyítás. Legyen $dimV_1 = n$ és $dimV_2 = m$. Legyen $\{\underline{e}_i,\ i = 1, \ldots, n\}$ a V_1 vektortér, $\{\underline{f}_j,\ j = 1, \ldots, m\}$ pedig a V_2 egy-egy bázisa. Jelölje 0_1, illetve 0_2 a V_1, illetve a V_2 vektortér nullvektorát. Megmutatható, hogy az

$$\{(\underline{e}_1, 0_2); \ldots; (\underline{e}_n, 0_2); (0_1, \underline{f}_1); \ldots; (0_2; \underline{f}_m)\}$$

Tegyük fel, hogy

$$\alpha_1(\underline{e}_1, 0_2) + \cdots + \alpha_n(\underline{e}_n, 0_2) + \beta_1(0_1, \underline{f}_1) + \cdots + \beta_m(0_1, \underline{f}_m) = (\underline{0}_1, \underline{0}_2)$$

teljesül valamely F-beli α_i és β_j $(i = 1, \ldots, n;\ j = 1, \ldots, m)$ skalárokkal. Akkor

$$(\alpha_1\underline{e}_1 + \cdots + \alpha_n\underline{e}_n, \beta_1\underline{f}_1 + \cdots + \beta_m\underline{f}_m) = (\underline{0}_1, \underline{0}_2),$$

ami akkor és csak akkor teljesül, ha

$$\alpha_1\underline{e}_1 + \cdots + \alpha_n\underline{e}_n = \underline{0}_1$$

és

$$\beta_1 \underline{f}_1 + \cdots + \beta_m \underline{f}_m = \underline{0}_2,$$

amiből

$$\alpha_1 = \cdots = \alpha_n = \beta_1 = \cdots = \beta_m = 0$$

következik, mert az $\{\underline{e}_i,\ i = 1, \ldots, n\}$, illetve az $\{\underline{f}_j,\ j = 1, \ldots, m\}$ vektorrendszer lineárisan független F felett. Mivel ez a két vektorrendszer a V_1, illetve a V_2 vektortér egy-egy generátorrendszere, ezért tetszőleges $(\underline{a}, \underline{b}) \in V_1 \oplus V_2$ vektorhoz megadhatók olyan F-beli α_i és β_j ($i = 1, \ldots, n$; $j = 1, \ldots, m$) skalárok, hogy

$$\underline{a} = \alpha_1 \underline{e}_1 + \cdots + \alpha_n \underline{e}_n,$$

$$\underline{b} = \beta_1 \underline{f}_1 + \cdots + \beta_m \underline{f}_m,$$

azaz

$$(\underline{a}, \underline{b}) = (\alpha_1 \underline{e}_1 + \cdots + \alpha_n \underline{e}_n, \beta_1 \underline{f}_1 + \cdots + \beta_m \underline{f}_m),$$

amiből

$$(\underline{a}, \underline{b}) = \alpha_1(\underline{e}_1, 0_2) + \cdots + \alpha_n(\underline{e}_n, 0_2) + \beta_1(0_1, \underline{f}_1) + \cdots + \beta_m(0_1, \underline{f}_m)$$

következik. Tehát az

$$\{(\underline{e}_1, 0_2); \ldots; (\underline{e}_n, 0_2); (0_1, \underline{f}_1); \ldots; (0_2; \underline{f}_m)\}$$

vektorrendszer a $V_1 \oplus V_2$ vektortér generátorrendszere. Mivel egy független generátorrendszer bázis, ezért a tételt bebizonyítottuk. □

Tétel 9.5 *Ugyanazon F test feletti V_1 és V_2 vektorterek esetén*

$$A = \{(\underline{a}, \underline{0}_2)\ :\ \underline{a} \in V_1\},$$

illetve

$$B = \{(\underline{0}_1, \underline{b})\ :\ \underline{b} \in V_2\}$$

a $V_1 \oplus V_2$ külső direkt összeg V_1-gyel, illetve V_2-vel izomorf olyan alterei, amelyekre teljesülnek az alábbiak:

1. *Minden $\underline{v} \in V_1 \oplus V_2$ vektorhoz megadhatók olyan $\underline{a} \in A$ és $\underline{b} \in B$ vektorok, hogy $\underline{v} = \underline{a} + \underline{b}$.*

2. $A \cap B$ csak a nullvektort tartalmazza.

Bizonyítás. Mivel V_1 és V_2 vektorterek az F test felett és $\alpha \underline{0}_i = \underline{0}_i$ $(i = 1, 2)$ minden $\alpha \in F$, ezért nyílvánvaló, hogy A és B altarei a $V_1 \oplus V_2$ külső direkt szorzatnak. Az

$$\underline{a} \ \mapsto (\underline{a}, \underline{0}_2),$$

illetve

$$\underline{b} \ \mapsto (\underline{0}_1, \underline{b})$$

leképezések a V_1, illetve V_2 vektortérnek az A, illetve B altérre való izomorfizmusai.

Az világos, hogy $A \cap B = \{(\underline{0}_1, \underline{0}_2)\}$. Továbbá, tetszőleges $(\underline{a}, \underline{b}) \in V_1 \oplus V_2$ vektor esetén

$$(\underline{a}, \underline{b}) = (\underline{a}, \underline{0}_2) + (\underline{0}_1, \underline{b}).$$

Ezzel az állítást bebizonyítottuk. □

Definíció 9.3 *Legyen V egy F test feletti vektortér. Akkor mondjuk, hogy V az A és B alterek belső direkt összege, ha*

1. V minden $\underline{v}$ vektorához megadhatók olyan $\underline{a} \in A$ és $\underline{b} \in B$ vektorok, hogy $\underline{v} = \underline{a} + \underline{b}$, illetve

2. $A \cap B$ csak a nullvektort tartalmazza.

Tétel 9.6 *Egy F test feletti V vektortér akkor és csak akkor belső direkt összege a V valamely A és B altereinek, ha V minden $\underline{v}$ vektorához van egy és csak egy $\underline{a} \in A$ és $\underline{b} \in B$ vektorpár, hogy $\underline{v} = \underline{a} + \underline{b}$.*

Bizonyítás. Ha V az A és B altereinek belső direkt összege, akkor definíció szerint vannak olyan $\underline{a} \in A$ és $\underline{b} \in B$ vektorok, hogy

$$\underline{v} = \underline{a} + \underline{b}.$$

Tegyük fel, hogy

$$\underline{v} = \underline{a}' + \underline{b}'$$

is teljesül valamely $\underline{a}' \in A$ és $\underline{b}' \in B$ vektorokra. Akkor

$$\underline{a} - \underline{a}' = \underline{b}' - \underline{b}.$$

Mivel $\underline{a}-\underline{a}' \in A$ és $\underline{b}'-\underline{b} \in B$, és ez a két különbség egymással egyenlő, ezért mindketten az $A \cap B$ elemei. Mivel $A \cap B = \{\underline{0}\}$, ezért $\underline{a}-\underline{a}' = \underline{b}'-\underline{b} = \underline{0}$, azaz

$$\underline{a} = \underline{a}'$$

és

$$\underline{b} = \underline{b}'.$$

Fordítva, tegyük fel, hogy a V vektortér minden $\underline{v}$ vektorához van egy és csak egy $\underline{a} \in A$ és $\underline{b} \in B$ vektorpár, hogy $\underline{v} = \underline{a} + \underline{b}$. Ekkor persze V minden $\underline{v}$ vektora előáll egy A-beli és egy B-beli vektor összegeként. Igy Csak azt kell megmutatni, hogy $A \cap B = \{\underline{0}\}$. Ha $\underline{v} \in A \cap B$, akkor

$$\underline{v} = \underline{v} + \underline{0}$$

és

$$\underline{v} = \underline{0} + \underline{v}$$

a $\underline{v}$ vektornak két előĺlítása A-beli és B-beli vektorok összegeként, mivel a $\underline{0}$ és $\underline{v}$ vektorok benne vannak A-ban is és B-ben is. Ebből viszont

$$\underline{v} = \underline{0}$$

következik, azaz

$$A \cap B = \{\underline{0}\}.$$

Tehát a V vektortér az A és B alterek belső direkt összege.

Tétel 9.7 *Ugyanazon F test feletti V_1 és V_2 vektorterek esetén a $V_1 \oplus V_2$ külső direkt összeg előáll az $A = \{(\underline{a}, \underline{0}_2) \ : \ \underline{a} \in V_1\}$ és $B = \{(\underline{0}_1, \underline{b}) \ : \ \underline{b} \in V_2\}$ alterek belső direkt összegeként. Továbbá, ha egy F test feletti V vektortér valamely A és B altereinek belső direkt összege, akkor V izomorf az A és B alterek, mint F feletti vektorterek külső direkt összegével.*

Bizonyítás. A Definíció 9.3 és a Tétel 9.5 miatt a tétel első ĺlítása nyilvánvaló.

Fordítva, tegyük fel, hogy egy F test feletti V vektortér valamely A és B altereinek belső direkt összege. Képezzük ezen altereknek, mint F feletti vektortereknek a külső direkt összegét. Legyen φ az $A \oplus B$ külső direkt összegnek a V vektortérbe való következő leképezése. Rendelje φ az $A \oplus B$ vektortér tetszőleges $(\underline{a}, \underline{b})$ vektorához a V vektortér $\underline{a} + \underline{b}$ vektorát. Mivel V az A és B belső direkt összege, ezért V minden $\underline{v}$ vektorához megadhatók olyan $\underline{a} \in A$ és $\underline{b} \in B$ vektorok, hogy $\underline{v} = \underline{a}+\underline{b}$. Ezért a φ leképezés szürjektív.

A Tétel **??** miatt a φ leképezés injektív is. Az könnyen ellenőrízhető, hogy φ lineáris. Tehát φ az $A \oplus B$ külső direkt összegnek a V vektortérre való izomorfizmusa. □

Megjegyzés 9.1 *Az előző tétel miatt nincs lńyeges különbség a külső direkt összeg és a belső direkt összeg fogalmak között. Igy a belső direkt összeget is hasonló módon fogjuk jelölni. Tehét, ha a V vektortér az A és B alterek belső direkt összege, akkor azt is $V = A \oplus B$ módon jelöljük. A fentiek szerint a "belső", illetve "külső" jelzőknek nincs jelentősége, ezért azt a továbbiakbal el is hagyhatjuk, elegendő csak a direkt összeg kifejezést használni.*

9.3. Lineáris transzformációk direkt összege

Definíció 9.4 *Egy F test feletti V vektortér valamely A lineáris transzformációját nilpotensnek nevezzük, ha megadható olyan pozití egész k szám, hogy $A^k = 0$. Azt a legkisebb m pozitív egész számot, amelyre $A^m = 0$ teljesül, az A nilpotens lineáris transzformáció nilpotencia-fokának nevezzük.*

Tétel 9.8 *Ha A egy F test feletti véges dimenziójú V vektortér olyan nilpotens lineáris transzformációja, melynek nilpotencia-foka m, akkor megadható V-nek olyan $\underline{x} \neq \underline{0}$ vektora, hogy az*

$$\underline{x}, A(\underline{x}), \ldots, A^{m-1}(\underline{x})$$

vektorok lineárisan függetlenek F felett.

Bizonyítás. Az $m = 1$ eset triviális. Tegyük fel a továbbiakban, hogy $m > 1$. Mivel $A^{m-1} \neq 0$, ezért megadható olyan $\underline{0} \neq \underline{x} \in V$ vektor, hogy $A(\underline{x}) \neq \underline{0}$. Tegyük fel, hogy

$$\alpha_0\underline{x} + \alpha_1 A(\underline{x}) + \cdots + \alpha_{m-1}A^{m-1}(\underline{x}) = \underline{0}$$

teljesül valamely $\alpha_0, \ldots, \alpha_{m-1} \in F$ skalárokkal. Akkor

$$A^{m-1}(\alpha_0\underline{x} + \alpha_1 A(\underline{x}) + \cdots + \alpha_{m-1}A^{m-1}(\underline{x})) = \underline{0},$$

azaz

$$\alpha_0 A^{m-1}(\underline{x}) + \alpha_1 A^m(\underline{x}) + \cdots + \alpha_{m-1}A^{2m-2}(\underline{x}) = \underline{0},$$

amely lényegében

$$\alpha_0 A^{m-1}(\underline{x}) = \underline{0}$$

alakú, mivel

$$A^m(\underline{x}) = \cdots = A^{2m-2}(\underline{x}) = \underline{0}.$$

Mivel $A^{m-1}(\underline{x}) \neq \underline{x}$, ezért

$$\alpha_0 = 0,$$

és így

$$\alpha_1 A(\underline{x}) + \cdots + \alpha_{m-1} A^{m-1}(\underline{x}) = \underline{0}.$$

Alkalmazva mindkét oldalra A^{m-2}-t, az

$$\alpha_1 A^{m-1}(\underline{x}) = \underline{0}$$

egyenlőséget kapjuk, amelyből

$$\alpha_1 = 0$$

adódik. Folytatva ezt az eljárást, kapjuk, hogy

$$\alpha_0 = \alpha_1 = \cdots = \alpha_{m-1} = 0.$$

Tehát az

$$\underline{x}, A(\underline{x}), \ldots, A^{m-1}(\underline{x})$$

vektorok lineárisan függetlenek F felett. □

9.4. Vektorterek tenzori szorzata

10. fejezet

Valós, illetve komplex bilineáris funkcionálok

Definíció 10.1 *Legyen V egy valós vektortér (azaz vektortér a valós számok teste felett). A $V \oplus V$ direkt összegnek az $\mathbb{R}$-be való φ leképezését bilineáris funkcionálnak nevezzük, ha teljesülnek az alábbia*

1. $\varphi(\underline{a}_1 + \underline{a}_2, \underline{b}) = \varphi(\underline{a}_1, \underline{b}) + \varphi(\underline{a}_2, \underline{b})$,
2. $\varphi(\alpha\underline{a}, \underline{b}) = \alpha\varphi(\underline{a}, \underline{b})$,
3. $\varphi(\underline{a}, \underline{b}_1 + \underline{b}_2) = \varphi(\underline{a}, \underline{b}_1) + \varphi(\underline{a}, \underline{b}_2)$,
4. $\varphi(\underline{a}, \beta\underline{b}) = \beta\varphi(\underline{a}, \underline{b})$

tetszőleges $\underline{a}, \underline{a}_1, \underline{a}_2, \underline{b}, \underline{b}_1, \underline{b}_2 \in V$ és tetszőleges $\alpha, \beta \in F$ esetén.

Definíció 10.2 *Legyen V egy komplex vektortér (azaz vektortér a komplex számok teste felett). A $V \oplus V$ direkt összegnek $\mathbb{C}$-be való φ leképezését bilineáris funkcionálnak nevezzük, ha teljesülnek az alábbia:*

1. $\varphi(\underline{a}_1 + \underline{a}_2, \underline{b}) = \varphi(\underline{a}_1, \underline{b}) + \varphi(\underline{a}_2, \underline{b})$,
2. $\varphi(\alpha\underline{a}, \underline{b}) = \overline{\alpha}\varphi(\underline{a}, \underline{b})$,
3. $\varphi(\underline{a}, \underline{b}_1 + \underline{b}_2) = \varphi(\underline{a}, \underline{b}_1) + \varphi(\underline{a}, \underline{b}_2)$,
4. $\varphi(\underline{a}, \beta\underline{b}) = \beta\varphi(\underline{a}, \underline{b})$.

tetszőleges $\underline{a}, \underline{a}_1, \underline{a}_2, \underline{b}, \underline{b}_1, \underline{b}_2 \in V$ és tetszőleges $\alpha, \beta \in F$ esetén.

Megjegyzés 10.1 *Figyeljük meg, hogy a fenti két definícióban szereplő képletek között csak a másodikban van eltérés.*

Tétel 10.1 *Tetszőleges n-dimenziós valós V vektortér tetszőleges φ bilibeáris funkcionáljához és tetszőleges $\mathcal{B}$ bázisához megadható egy és csak egy olyan valós számokból álló n-edrendű $[\varphi]_\mathcal{B}$ mátrix, hogy minden V-beli $\underline{x}$ és $\underline{y}$ vektor esetén*

$$\varphi(\underline{x}, \underline{y}) = [\underline{x}]^T_\mathcal{B}[\varphi]_\mathcal{B}[\underline{y}]_\mathcal{B}$$

teljesül.

Bizonyítás. Legyenek a $\mathcal{B}$ bázis elemei $\underline{e}_1, \ldots, \underline{e}_n$. Jelölje $[\varphi]_\mathcal{B}$ azt az n-edrendű mátrixot, melynek az i-dik sorában álló j-dik elem egyenlő $\varphi(\underline{e}_i, \underline{e}_j)$-vel $(i, j = 1, \ldots, n)$. Akkor tetszőleges $\underline{x} = \xi_1\underline{e}_1 + \cdots + \xi_n\underline{e}_n$ és $\underline{y} = \eta_1\underline{e}_1 + \cdots + \eta_n\underline{e}_n$ vektorokra

$$\varphi(\underline{x}, \underline{y}) =$$

$$= \varphi(\xi_1\underline{e}_1 + \cdots + \xi_n\underline{e}_n, \eta_1\underline{e}_1 + \cdots + \eta_n\underline{e}_n) =$$

$$= \sum_{i,j=1}^{n} \xi_i\eta_j\varphi(\underline{e}_i, \underline{e}_j) = \sum_{i,j=1}^{n} \xi_i\eta_j a_{ij} = [\underline{x}]^T_\mathcal{B}[\varphi]_\mathcal{B}[\underline{y}]_\mathcal{B}.$$

Az egyértelműség bizonyításához tegyük fel, hogy valamely n-edrendű $\mathbf{A}$ és $\mathbf{B}$ mátrixokra és minden $\underline{x}, \underline{y} \in V$ vektorra teljesül az

$$[\underline{x}]^T_\mathcal{B}\mathbf{A}[\underline{y}]_\mathcal{B} = \varphi(\underline{x}, \underline{y}) = [\underline{x}]^T_\mathcal{B}\mathbf{B}[\underline{y}]_\mathcal{B}$$

egyenlőség. Akkor az $\underline{x} = \underline{e}_i$ és $\underline{y} = \underline{e}_j$ választás mellett

$$a_{ij} = [\underline{e}_i]^T_\mathcal{B}\mathbf{A}[\underline{e}_j]_\mathcal{B} = [\underline{e}_i]^T_\mathcal{B}\mathbf{B}[\underline{e}_j]_\mathcal{B} = b_{ij}$$

adódik. Tehát $\mathbf{A} = \mathbf{B}$. □

Tétel 10.2 *Tetszőleges n-dimenziós komplex V vektortér tetszőleges φ bilibeáris funkcionáljához és tetszőleges $\mathcal{B}$ bázisához megadható egy és csak egy olyan komplex számokból álló n-edrendű $[\varphi]_\mathcal{B}$ mátrix, hogy minden V-beli $\underline{x}$ és $\underline{y}$ vektor esetén*

$$\varphi(\underline{x}, \underline{y}) = \overline{[\underline{x}]}^T_\mathcal{B}[\varphi]_\mathcal{B}[\underline{y}]_\mathcal{B}$$

teljesül.

Bizonyítás. Az előző tétel bizonyításának jelöléseit használva,

$$\varphi(\underline{x}, \underline{y}) =$$
$$= \varphi(\xi_1\underline{e}_1 + \cdots + \xi_n\underline{e}_n, \eta_1\underline{e}_1 + \cdots + \eta_n\underline{e}_n) =$$
$$\sum_{i,j=1}^{n} \overline{\xi_i}\eta_j\varphi(\underline{e}_i, \underline{e}_j) = \sum_{i,j=1}^{n} \overline{\xi_i}\eta_j a_{ij} = \overline{[\underline{x}]}_{\mathcal{B}}^{T}[\varphi]_{\mathcal{B}}[\underline{y}]_{\mathcal{B}}.$$

Az egyértelműség bizonyítása hasonló a valós esethez. □

Definíció 10.3 *Egy valós V vektortéren értelmezett φ bilineáris funkcionált szimmetrikus bilineáris funkcionálnak nevezünk, ha a V vektortér tetszőleges $\underline{x}$ és $\underline{y}$ vektorai esetén*

$$\varphi(\underline{x}, \underline{y}) = \varphi(\underline{y}, \underline{x}).$$

Tétel 10.3 *Ha egy valós V vektortér valamely φ bilineáris funkcionálja szimmetrikus, akkor V tetszőleges $\mathcal{B}$ bázisa esetén a $[\varphi]_{\mathcal{B}}$ mátrix szimmetrikus. Fordítva, ha V valamely φ bilineáris funkcionĺjának valamely bźisban szimmetrikus a mátrixa, akkor φ szimmetrikus bilineáris funkcionĺja V-nek.*

Bizonyítás. Ha φ egy n-dimenziós valós V vektortér szimmetrikus bilineáris funkcionálja, akkor V bármely $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n\}$ bázisa esetén

$$\varphi(\underline{e}_i, \underline{e}_j) = \varphi(\underline{e}_j, \underline{e}_i),$$

ami annyit jelent, hogy a $[\varphi]_{\mathcal{B}}$ mátrix szimmetrikus (lásd a Tétel 10.1 bizonyítását is).

Fordítva, Legyen φ egy n-dimenziós valós V vektortér olyan bilineáris funkcionálja, melynek mátrixa a V egy $\mathcal{B} = \{\underline{]}_{\infty}, \ldots, \underline{]}_{\backslash}\}$ bázisában szimmetrikus. Akkor tetszőleges $\underline{x}, \underline{y} \in V$ vektorok esetén

$$\varphi(\underline{x}, \underline{y}) = (\varphi(\underline{x}, \underline{y}))^T = ([\underline{x}]_{\mathcal{B}}^T[\varphi]_{\mathcal{B}}[\underline{y}]_{\mathcal{B}})^T =$$
$$= [\underline{y}]_{\mathcal{B}}^T[\varphi]_{\mathcal{B}}^T[\underline{x}]_{\mathcal{B}} = [\underline{y}]_{\mathcal{B}}^T[\varphi]_{\mathcal{B}}[\underline{x}]_{\mathcal{B}} = \varphi(\underline{y}, \underline{x}).$$

Tehát φ szimmetrikus bilineáris funkcionál. □

Definíció 10.4 *Egy komplex V vektortéren értelmezett φ bilineáris funkcionált Hermite-féle bilineáris funkcionálnak nevezünk, ha a V vektortér tetszőleges $\underline{x}$ és $\underline{y}$ vektorai esetén*

$$\varphi(\underline{x}, \underline{y}) = \overline{\varphi(\underline{y}, \underline{x})}.$$

Tétel 10.4 *Ha egy komplex V vektortér valamely φ bilineáris funkcionálja Hermite-féle, akkor V tetszőleges $\mathcal{B}$ bázisa esetén a $[\varphi]_{\mathcal{B}}$ mátrix önadjungált (Definíció **??**). Fordítva, ha V valamely φ bilineáris funkcionĺjának valamely bźisban szimmetrikus a mátrixa, akkor φ szimmetrikus bilineáris funkcionĺja V-nek.*

Bizonyítás. Ha φ egy n-dimenziós komplex V vektortér szimmetrikus bilineáris funkcionálja, akkor V bármely $\mathcal{B} = \{\underline{e}_1, \dots, \underline{e}_n\}$ bázisa esetén

$$\varphi(\underline{e}_i, \underline{e}_j) = \overline{\varphi(\underline{e}_j, \underline{e}_i)},$$

ami annyit jelent, hogy a $[\varphi]_{\mathcal{B}}$ mátrix Hermit-féle (lásd a Tétel 10.2 bizonyítását is).

Fordítva, Legyen φ egy n-dimenziós komplex V vektortér olyan bilineáris funkcionálja, melynek mátrixa a V egy $\mathcal{B} = \{\underline{]}_{\infty}, \dots, \underline{]}_{\backslash}\}$ bázisában önadjungált. Akkor tetszőleges $\underline{x}, \underline{y} \in V$ vektorok esetén

$$\varphi(\underline{x}, \underline{y}) = \overline{\overline{(\varphi(\underline{x}, \underline{y}))^T}} = \overline{\overline{(\overline{[\underline{x}]}_{\mathcal{B}}^T [\varphi]_{\mathcal{B}} [\underline{y}]_{\mathcal{B}})}^T} =$$

$$= \overline{\overline{[\underline{y}]}_{\mathcal{B}}^T \overline{[\varphi]}_{\mathcal{B}}^T [\underline{x}]_{\mathcal{B}}} = \overline{\overline{[\underline{y}]}_{\mathcal{B}}^T [\varphi]_{\mathcal{B}} [\underline{x}]_{\mathcal{B}}} = \overline{\varphi(\underline{y}, \underline{x})}.$$

Tehát φ Hermite-féle bilineáris funkcionál. □

Tétel 10.5 *Legyen V egy n-dimenziós valós vektortér, és legyen $\mathcal{B} = \{e_1, \dots, e_n\}$, illetve $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ a V két bázisa. Legyen τ a V vektortér azon bázistranszformációja, amelyre*

$$e'_i = \varphi(e_i), \quad i = 1, \dots, n$$

teljesül. Akkor a V vektortér tetszőleges φ bilineáris funkcionálja esetén

$$[\varphi]_{\mathcal{B}'} = [\tau]_{\mathcal{B}}^T [\varphi]_{\mathcal{B}} [\tau]_{\mathcal{B}}.$$

Bizonyítás. A Tétel 6.17 felhasználásával

$$\varphi(\underline{x}, \underline{y}) = [\underline{x}]_{\mathcal{B}}^T [\varphi]_{\mathcal{B}} [\underline{y}]_{\mathcal{B}} = ([\tau]_{\mathcal{B}} [\underline{x}]_{\mathcal{B}'})^T [\varphi]_{\mathcal{B}} ([\tau]_{\mathcal{B}} [\underline{y}]_{\mathcal{B}'} =$$

$$= [\underline{x}]_{\mathcal{B}'}^T ([\tau]_{\mathcal{B}}^T [\varphi]_{\mathcal{B}} ([\tau]_{\mathcal{B}}) [\underline{y}]_{\mathcal{B}'},$$

amiből a bilineáris funkcionálok mátrixának egyértelműsége miatt

$$[\varphi]_{\mathcal{B}'} = [\tau]_{\mathcal{B}}^T [\varphi]_{\mathcal{B}} ([\tau]_{\mathcal{B}}$$

következik. □

Tétel 10.6 *Legyen V egy n-dimenziós komplex vektortér, és legyen $\mathcal{B} = \{e_1, \ldots, e_n\}$, illetve $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ a V két bázisa. Legyen τ a V vektortér azon bázistranszformációja, amelyre*

$$e'_i = \varphi(e_i), \quad i = 1, \ldots, n$$

teljesül. Akkor a V vektortér tetszőleges φ bilineáris funkcionálja esetén

$$[\varphi]_{\mathcal{B}'} = \overline{[\tau]}_{\mathcal{B}}^{T} [\varphi]_{\mathcal{B}} [\tau]_{\mathcal{B}}.$$

Bizonyítás. Komplex esetben az előző tétel bizonyítása a következőképpen módosul.

$$\varphi(\underline{x}, \underline{y}) = \overline{[\underline{x}]}_{\mathcal{B}}^{T} [\varphi]_{\mathcal{B}} [\underline{y}]_{\mathcal{B}} = \overline{([\tau]_{\mathcal{B}} [\underline{x}]_{\mathcal{B}'})}^{T} [\varphi]_{\mathcal{B}} ([\tau]_{\mathcal{B}} [\underline{y}]_{\mathcal{B}'}) =$$

$$= \overline{[\underline{x}]}_{\mathcal{B}'}^{T} (\overline{[\tau]}_{\mathcal{B}}^{T} [\varphi]_{\mathcal{B}} [\tau]_{\mathcal{B}}) [\underline{y}]_{\mathcal{B}'},$$

amiből a bilineáris funkcionálok mátrixának egyértelműsége miatt

$$[\varphi]_{\mathcal{B}'} = \overline{[\tau]}_{\mathcal{B}}^{T} [\varphi]_{\mathcal{B}} [\tau]_{\mathcal{B}}$$

következik. □

11. fejezet

Valós, illetve komplex kvadratikus alakok

11.1. A kvadratikus alak fogalma és mátrixa

Definíció 11.1 *Egy V valós vektortéren értelmezett (valós) kvadratikus alakon a V vektortérnek a valós számok testébe való olyan g funkcionálját értjük, amelyhez megadható a V vektortéren értelmezett olyan φ valós bilineáris funkcionál, hogy minden $\underline{x} \in V$ vektorra*

$$g(\underline{x}) = \varphi(\underline{x}, \underline{x})$$

teljesül.

Definíció 11.2 *Egy V komplex vektortéren értelmezett (komplex) kvadratikus alakon a V vektortérnek a komplex számok testébe való olyan g funkcionálját értjük, amelyhez megadható a V vektortéren értelmezett olyan φ komplex bilineáris funkcionál, hogy minden $\underline{x} \in V$ vektorra*

$$g(\underline{x}) = \varphi(\underline{x}, \underline{x})$$

teljesül.

Tétel 11.1 *Adott V komplex vektortér esetén kölcsöösen egyértelmű megfeleltetés van a V-n értelmezett kvadratikus alakok és V bilineáris funkcionáljai között. Ennél a megfeleltetésnél valós értékű kvadratikus alaknak Hermite-féle bilineáris funkcionál felel meg.*

Tétel 11.2 *Adott V valós vektortér esetén kölcsöösen egyértelmű megfeleltetés van a V-n értelmezett kvadratikus alakok és V szimmetrikus bilineáris funkcionáljai között.*

Definíció 11.3 *Egy valós értékű valós (illetve, komplex) kvadratikus alak valamely $\mathcal{B}$ bázishoz tartozó mátrixán a neki megfeleltetett szimmetrikus (illetve, Hermite-féle) bilineáris funkcionál $\mathcal{B}$-hez tartozó mátrixát értjük.*

Megjegyzés 11.1 *Az előző definíció szerint, valós értékű valós kvadratikus alak mátrixa minden bázisban szimmetrikus, valós értékű komplex kvadtarikus alak mátrixa pedig minden bázisban önadjungált.*

Megjegyzés 11.2 *Egy 2-dimenziós V valós vektortér esetén egy*

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

mátrixú bilineáris funkcionálból származtatott g kvadratikus alak

$$g(\underline{x}) = [x_1, x_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (\underline{x} = [x_1, x_2]^T)$$

képlettel írható fel. A szorzást elvégezve,

$$g(\underline{x}) = a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2.$$

Ha bevezetjük a $b_{12} = b_{21} = \frac{1}{2}(a_{12} + a_{21})$ jelölést, akkor

$$g(\underline{x}) = a_{11}x_1^2 + (b_{12} + b_{21})x_1x_2 + a_{22}x_2^2 = a_{11}x_1^2 + b_{12}x_1x_2 + b_{21}x_2x_1 + a_{22}x_2^2,$$

ami egyenlő az

$$[x_1, x_2] \begin{bmatrix} a_{11} & b_{12} \\ b_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

szorzattal. Tehát a g valós kvadratikus alakot kapjuk akkor is, ha a szimmetrikus

$$\begin{bmatrix} a_{11} & b_{12} \\ b_{21} & a_{22} \end{bmatrix}$$

mátrixú valós bilineáris funkcionálban a változókat azonosítjuk.

Ez az egyszerű példa jól szemlélteti, hogyan kaphatjuk meg egy

$$g(\underline{x}) \quad (\underline{x} = [x_1, \ldots x_n]^T)$$

valós kvadratikus alak mátrixát. Az i-dik sor i-dik eleme megegyezik az x_i^2-tag g-beli együtthatójával, az i-dik sor j-dik és a j-dik sor i-dik ($i \neq j$) eleme megegyezik az $x_i x_j$ tag g-beli együtthatójának felével. Például a

$$g(\underline{x}) = 3x_1^2 + 8x_1x_2 + 6x_2^2$$

valós kvadratikus alak mátrixa

$$\begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix}.$$

Megjegyzés 11.3 *Egy 2-dimenziós V komplex vektortér esetén egy*

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

mátrixú bilineáris funkcionálból származtatott g kvadratikus alak

$$g(\underline{x}) = [\overline{x}_1, \overline{x}_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (\underline{x} = [x_1, x_2]^T)$$

képlettel írható fel. A szorzást elvégezve,

$$g(\underline{x}) = a_{11}\overline{x}_1 x_1 + a_{12}\overline{x}_1 x_2 + a_{21}\overline{x}_2 x_1 + a_{22}\overline{x}_2 x_2.$$

Ez jól szemlélteti, hogyan kaphatjuk meg egy

$$g(\underline{x}) \quad (\underline{x} = [x_1, \dots x_n]^T)$$

komplex kvadratikus alak mátrixát: az i-dik sor j-dik eleme megegyezik az $\overline{x}_i x_j$ tag g-beli együtthatójával.

Például a

$$g(\underline{x}) = 2\overline{x}_1 x_1 + (2i)\overline{x}_1 x_2 - (2i)\overline{x}_2 x_1 + 6\overline{x}_2 x_2$$

komplex kvadratikus alak mátrixa

$$\begin{bmatrix} 2 & 2i \\ -2i & 6 \end{bmatrix}.$$

Mivel

$$\overline{\begin{bmatrix} 2 & 2i \\ -2i & 6 \end{bmatrix}^T} = \overline{\begin{bmatrix} 2 & -2i \\ 2i & 6 \end{bmatrix}} = \begin{bmatrix} 2 & 2i \\ -2i & 6 \end{bmatrix},$$

ezért a g mátrixa önadjungált. Tehát a példában szereplő g komplex kvadratikus alak egy Hermite-féle komplex bilineáris funkcionálból származott, s ezért minden értéke valós szám. Példaként tekintsük az

$$\underline{x}_0 = [3i, (1+i)]^T$$

komplex vektorhoz rendelt értéket:

$$g(\underline{x}_0) = 2(-3i)(3i) + (2i)(-3i)(1+i) - (2i)(1-i)(3i) + 6(1-i)(1+i) = 42.$$

11.2. A valós értékű kvadratikus alakok osztályozása

Definíció 11.4 *Egy valós vagy komplex vektortéren értelmezett valós értékű g kvadratikus alakról azt mondjuk, hogy*

1. *pozitív definit, ha minden $\underline{0} \neq \underline{x} \in V$ vektorra $g(\underline{x}) > 0$ teljesül;*
2. *pozitív szemidefinit, ha minden $\underline{x} \in V$ vektorra $g(\underline{x}) \geq 0$ teljesül, de van olyan $\underline{0} \neq \underline{x} \in V$ vektor, hogy $g(\underline{x}) = \underline{0}$;*
3. *negatív definit, ha minden $\underline{0} \neq \underline{x} \in V$ vektorra $g(\underline{x}) < 0$ teljesül;*
4. *negatív szemidefinit, ha minden $\underline{x} \in V$ vektorra $g(\underline{x}) \leq 0$ teljesül, de van olyan $\underline{0} \neq \underline{x} \in V$ vektor, hogy $g(\underline{x}) = \underline{0}$;*
5. *indefinit, ha vannak olyan, a nullvektortól különböző $\underline{x}, \underline{y} \in V$ vektorok, hogy $g(\underline{x}) > \underline{0}$ és $g(\underline{y}) < \underline{0}$.*

Definíció 11.5 *Egy szimmetrikus, illetve egy Hermite-féle bilineáris funkcionált pozitív definitnek nevezünk, ha a neki megfeleltetett valós értékű kvadratikus alak pozitív definit.*

12. fejezet

Valós, illetve komplex euklideszi tér

Definíció 12.1 *Egy valós vektortéren értelmezett pozitív definit szimmetrikus bilineáris funkcionált valós skaláris szorzásnak, egy komplex vektortéren értelmezett pozitív definit Hermite-féle bilineáris funkcionált komplex skaláris szorzásnak nevezünk.*

12.1. Az euklideszi tér fogalma

Definíció 12.2 *Egy valós (komplex) vektorteret valós (komplex) euklideszi térnek nevezünk, ha értelmezve van rajta egy valós (komplex) skaláris szorzás. Ha a vektorteret V-vel, a skaláris szorzást f-fel jelöljük, akkor az általuk definiált euklideszi teret (V, f)-fel fogjuk jelölni.*

Megjegyzés 12.1 *Ha $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n\}$ egy n-dimenziós V valós euklideszi tér egy bázisa, akkor az az $f : V \oplus V \mapsto \mathbb{R}$ kétváltozós funkcionál, amely V tetszőleges $\underline{x} = \xi_1\underline{e}_1 + \cdots + \xi_n\underline{e}_n$ és $\underline{y} = \eta_1\underline{e}_1 + \cdots + \eta_n\underline{e}_n$ vektoraihoz az*

$$f(\underline{x}, \underline{y}) = \sum_{i=1}^{n} \xi_i \eta_i$$

valós számot rendeli, valós skaláris szorzás.

Komplex esetben az

$$f(\underline{x}, \underline{y}) = \sum_{i=1}^{n} \overline{\xi}_i \eta_i$$

képlettel definiált kétváltozós funkcionál komplex skaláris szorzás.

Definíció 12.3 *Egy* (V, f) *valós (komplex) euklideszi tér* $\underline{x}$ *és* $\underline{y}$ *vektorairól azt mondjuk, hogy ortogonálisak, ha* $f(\underline{x}, \underline{y}) = 0$. *A* (V, f) *tér egy vektorrendszerét ortogonális vektorrendszernek nevezzük, ha a hozzá tartozó vektorok páronként ortogonálisak.*

Megjegyzés 12.2 *A nullvektor bármely vektorral ortogonális.*

Tétel 12.1 *Egy valós, illetve komplex euklideszi tér bármely olyan ortogonális vektorrendszere, amely nem tartalmazza a nullvektort, lineárisan független.*

Bizonyítás. Legyen $\{\underline{b}_1, \ldots, \underline{b}_k\}$ egy (V, f) euklideszi tér olyan ortogonális vektorrendszere, amely nem tartalmazza a nullvektort. Tegyük fel, hogy

$$\beta_1 \underline{b}_1 + \cdots + \beta_k \underline{b}_k = \underline{0}.$$

Akkor tetszőleges $i = 1, \ldots, k$ index esetén

$$0 = f(\underline{b}_i, \beta_1 \underline{b}_1 + \cdots + \beta_k \underline{b}_k) = \beta_i f(\underline{b}_i, \underline{b}_i).$$

Mivel f pozitív definit, ezért $f(\underline{b}_i, \underline{b}_i) > 0$, s ezért

$$\beta_i = 0.$$

Tehát a $\{\underline{b}_1, \ldots, \underline{b}_k\}$ vektorrendszer lineárisan független. □

Tétel 12.2 *Ha* $\{\underline{e}_1, \ldots \underline{e}_n\}$ *egy valós (komplex)* (V, f) *euklideszi tér ortogonális vektorrendszere, akkor* V *tetszőleges* $\underline{v}$ *vektora esetén a*

$$\underline{v}' = \underline{v} - \sum_{i=1}^{n} \alpha_i \underline{e}_i \quad \alpha_i = \frac{f(\underline{v}, \underline{e}_i)}{f(\underline{e}_i, \underline{e}_i)}$$

vektor ortogonális az $\underline{e}_j$ $(j = 1, \ldots n)$ *vektorok mindegyikére, így az azok által kifeszített altér minden vektorára.*

Bizonyítás. A következő bizonyítás mind a valós, mind a komplex esetre alkalmazható. Tetszőleges $j = 1, \ldots, n$ index esetén

$$f(\underline{v}', \underline{e}_j) = f(\underline{v}, \underline{e}_j) - \sum_{i=1}^{n} \alpha_j f(\underline{e}_i, \underline{e}_j) =$$

$$= f(\underline{v}, \underline{e}_j) - \frac{f(\underline{v}, \underline{e}_j)}{f(\underline{e}_j, \underline{e}_j)} f(\underline{e}_j, \underline{e}_j) = 0.$$

Tehát $\underline{v}'$ ortogonális az $\underline{e}_j$ $(j = 1, \ldots, n)$ vektorok mindegyikére. □

Tétel 12.3 *(Gram-Schmidt-féle ortogonalizáció) Egy (V, f) valós (komplex) euklideszi tér tetszőleges lineárisan független $\{\underline{v}_1, \ldots, \underline{v}_k\}$ vektorrendszeréhez van olyan $\{\underline{e}_1, \ldots, \underline{k}_k\}$ ortogonális vektorrendszere, hogy minden $i = 1, \ldots, k$ indexre*

$$\langle \underline{e}_1, \ldots, \underline{e}_i \rangle = \langle \underline{v}_1, \ldots, \underline{v}_i \rangle.$$

Bizonyítás. A bizonyítás mind a valós, mind a komplex esetre alkalmazható. Legyen

$$\{\underline{v}_1, \ldots, \underline{v}_k\}$$

egy (V, f) euklideszi tér lineárisan független vektorrendszere. Legyen $\underline{e}_1 = \underline{v}_1$. Legyen továbbá

$$\underline{e}_2 = \underline{v}_2 - \alpha_1 \underline{e}_1,$$

ahol

$$\alpha_1 = \frac{f(\underline{v}_2, \underline{e}_1)}{f(\underline{e}_1, \underline{e}_1)}.$$

A Tétel 12.2 miatt az $\{\underline{e}_1, \underline{e}_2\}$ vektorrendszer ortogonális. Továbbá,

$$\langle \underline{e}_1, \underline{e}_2 \rangle = \langle \underline{v}_1, \underline{v}_2 \rangle.$$

Folytatva az eljárást, ha már előállítottunk egy olyan

$$\{\underline{e}_1, \ldots, \underline{e}_j\}$$

$(j < k)$ ortogonális vektorrendszert, amelyre

$$\langle \underline{e}_1, \ldots, \underline{e}_j \rangle = \langle \underline{v}_1, \ldots, \underline{v}_j \rangle$$

teljesül, akkor képezzük az

$$\underline{e}_{j+1} = \underline{v}_{j+1} - \sum_{t=1}^{j} \alpha_t \underline{e}_t$$

vektort, ahol

$$\alpha_t = \frac{f(\underline{v}_{j+1}, \underline{e}_t)}{f(\underline{e}_t, \underline{e}_t)}.$$

A Tétel 12.2 miatt $\underline{e}_{j+1}$ ortogonális az $\underline{e}_1$ $(i = 1, \ldots, j)$ vektorok mindegyikére. Továbbá,

$$\langle \underline{e}_1, \ldots, \underline{e}_{j+1} \rangle = \langle \underline{v}_1, \ldots, \underline{v}_{j+1} \rangle$$

is teljesül. Az eljárás befejezéseként olyan

$$\{\underline{e}_1, \ldots, \underline{e}_k\}$$

ortogonális vektorrendszert kapunk, amelyre teljesül, hogy minden $i = 1, \ldots, k$ indexre

$$\langle \underline{e}_1, \ldots, \underline{e}_i \rangle = \langle \underline{v}_1, \ldots, \underline{v}_i \rangle.$$

□

Tétel 12.4 *Minden (V, f) euklideszi térnek van ortogonális bázisa. Továbbá (V, f) minden nem nullvektora benne van egy ortogonális bázisban.*

Bizonyítás. Legyen $\{\underline{b}_1, \ldots, \underline{b}_n\}$ egy n-dimenziós (V, f) euklideszi tér valamely bázisa. Akkor a Tétel 12.3 szerint van (V, f)-nek olyan $\{\underline{a}_1, \ldots, \underline{a}_n\}$ ortogonális vektorrendszere, amely V-nek bázisa (mivel ez a vektorrendszer V generátorrendszere a Tétel 12.3 szerint, és vektorai függetlenek a Tétel 12.1 miatt)

Legyen $\underline{x} \neq \underline{0}$ a V tetszőleges vektora. Akkor $\underline{x}$ benne van V egy $\{\underline{b}_1 = \underline{x}, \ldots, \underline{b}_n\}$ bázisban. Erre alkalmazva a Gram-Schmidt-féle ortogonalizációt, van V-nek olyan bázisa, amely tartalmazza az $\underline{x}$ vektort. □

Definíció 12.4 *Egy (V, f) euklideszi tér valamely $\underline{x}$ vektorát normálvektornak nevezzünk, ha $f(\underline{x}, \underline{x}) = 1$. Egy $\underline{x} \neq \underline{0}$ vektor normálásának nevezzük a*

$$\underline{y} = \frac{1}{\sqrt{f(\underline{x}, \underline{x})}} \underline{x}$$

vektorra való áttérést.

Megjegyzés 12.3 *Minden normálvektor nem nullvektor. Továbbá, egy nem nullvektor normálásával keletkezett vektor normálvektor. Az előző definíció jelöléseit használva, (mind valós, mind komplex euklideszi tér esetén)*

$$f(\underline{y}, \underline{y}) = (\frac{1}{\sqrt{f(\overline{x}, \overline{x})}})^2 f(\overline{x}, \overline{x}) = \frac{1}{f(\overline{x}, \overline{x})} f(\overline{x}, \overline{x}) = 1,$$

mivel $\frac{1}{\sqrt{f(\underline{x}, \underline{x})}}$ valós szám, s ezért megegyezik a konjugáltjával.

Definíció 12.5 *Egy (V, f) euklideszi tér $\{\underline{b}_1, \ldots, \underline{b}_k\}$ ortogonális vektorrendszerét ortonormált vektorrendszernek nevezzük, ha elemei normálvektorok.*

Tétel 12.5 *Minden (V, f) euklideszi térnek van ortonormált bázisa. Továbbá (V, f) minden normálvektora benne van egy ortonormált bázisban.*

Bizonyítás. A Tétel 12.4 szerint minden euklideszi térnek van ortogonális bázisa. Ezen bázis vektorainak normálásával keletkezett vektorok rendszere az euklideszi tér egy ortonormált bázisa.

Mivel egy normálvektor nem nullvektor, ezért az benne van a tér egy ortogonális bázisában, így a tér egy ortonormált bázisában is. □

Tétel 12.6 *Egy (V, f) valós euklideszi tér valamely $\{\underline{e}_1, \ldots, \underline{e}_n\}$ bázisa akkor és csak akkor ortonormált, ha V tetszőleges $\underline{x} = \xi_1\underline{e}_1 + \cdots + \xi_n\underline{e}_n$ és $\underline{y} = \eta_1\underline{e}_1 + \cdots + \eta_n\underline{e}_n$ vektorai esetén*

$$f(\underline{x}, \underline{y}) = \sum_{i=1}^{n} = \xi_i\eta_i.$$

Bizonyítás. Nyilvánvaló. □

Tétel 12.7 *Egy (V, f) komplex euklideszi tér valamely $\{\underline{e}_1, \ldots, \underline{e}_n\}$ bázisa akkor és csak akkor ortonormált, ha V tetszőleges $\underline{x} = \xi_1\underline{e}_1 + \cdots + \xi_n\underline{e}_n$ és $\underline{y} = \eta_1\underline{e}_1 + \cdots + \eta_n\underline{e}_n$ vektorai esetén*

$$f(\underline{x}, \underline{y}) = \sum_{i=1}^{n} = \overline{\xi}_i\eta_i.$$

Bizonyítás. Nyilvánvaló. □

Megjegyzés 12.4 *Az előző két tétel, valamint a Megjegyzés 12.1 szerint egy valós vagy komplex V vektortér bármely $\mathcal{B}$ bźisá esetén lehet értlemezni V-n olyan skaláris szorzást, amelyre nézve $\mathcal{B}$ ortonormált bázis.*

Tétel 12.8 *(Bessel-egyenlőtlenség) Ha $\{\underline{e}_1, \ldots \underline{e}_n\}$ egy valós (komplex) (V, f) euklideszi tér ortonormált vektorrendszere, akkor V tetszőleges $\underline{v}$ vektora esetén*

$$\sum_{i=1}^{n} |f(\underline{v}, \underline{e}_i|^2 \leq \| \underline{v} \|^2 .$$

Bizonyítás. A komplex esettel foglalkozunk, amelybő speciálisan adódik a valós eset is. Képezzük az

$$\underline{v}' = \underline{v} - \sum_{i=1}^{n} f(\underline{v}, \underline{e}_i)\underline{e}_i$$

vektort. Akkor

$$0 \leq f(\underline{v}', \underline{v}') = f(\underline{v}, \underline{v}) - \sum_{i=1}^{n} \overline{f(\underline{v}, \underline{e}_i)} f(\underline{v}, \underline{e}_i) - \sum_{i=1}^{n} f(\underline{v}, \underline{e}_i)\overline{f(\underline{v}, \underline{e}_i)} + \sum_{i=1}^{n} \overline{f(\underline{v}, \underline{e}_i)} f(\underline{v}, \underline{e}_i) =$$

$$f(\underline{v}, \underline{v}) - \sum_{i=1}^{n} \overline{f(\underline{v}, \underline{e}_i)} f(\underline{v}, \underline{e}_i) == f(\underline{v}, \underline{v}) - \sum_{i=1}^{n} |f(\underline{v}, \underline{e}_i)|^2,$$

amiből már következik a tételben felírt egyenlőtlenség. □

12.2. Ortogonális és unitér transzformációk

Definíció 12.6 *Egy valós (illetve, komplex) (V, f) euklideszi tér valamely φ lineáris transzformációját ortogonális transzformációnak nevezzük, ha tetszőleges $\underline{x}, \underline{y} \in V$ vektorokra*

$$f(\varphi(\underline{x}), \varphi(\underline{y})) = f(\underline{x}, \underline{y})$$

teljesül.

Tétel 12.9 *Legyen φ egy valós (illetve, komplex) euklideszi tér lineáris transzformációja. Ha φ ortogonális, akkor a tér bármely ortonormált bázisát ortonormált bázisba viszi át. Fordítva, ha φ a tér valamely ortonormált bázisát ortonormált bázisba viszi át, akkor φ ortogonális.*

Bizonyítás. Ha φ ortogonális, akkor a tér tetszőleges ortonormált $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n$ bázisa esetén tetszőleges $i, j \in \{1, \ldots, n\}$ indexekre

$$f(\varphi(\underline{e}_i), \varphi(\underline{e}_i)) = f(\underline{e}_i, \underline{e}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

teljesül. Tehát a $\mathcal{B}$ bázis φ szerinti képe ortonormált bázis.

Fordítva, legyen φ a tér olyan lineáris transzformációja, amely valamely $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n$ ortonormált bázisát ortonormált bázisba viszi át. Akkor a tér tetszőleges $\underline{x} = \sum_{i=1}^{n} \xi_i \underline{e}_i$ és $\underline{y} = \sum_{i=1}^{n} \eta_i \underline{e}_i$ vektorai esetén

$$f(\varphi(\underline{x}), \varphi(\underline{y})) = \sum_{i,j=1}^{n} \xi_i \eta_j f(\varphi(\underline{e}_i), \varphi(\underline{e}_j)) = \sum_{k=1}^{n} \xi_k \eta_k = \sum_{i,j=1}^{n} \xi_i \eta_j f(\underline{e}_i, \underline{e}_j) = f(\underline{x}, \underline{y}).$$

Igy φ ortogonális transzformáció. □

Definíció 12.7 *Egy valós elemű* $\mathbf{A}$ *mátrixot ortogonális mátrixnak nevezünk, ha*

$$\mathbf{A}^{-1} = \mathbf{A}^T.$$

Tétel 12.10 *Egy n-edrendű valós elemű* $\mathbf{A}$ *mátrix akkor és csak akkor ortogonális, ha tetszőleges* $i, j \in \{1, \ldots, n\}$ *indexekre*

$$\sum_{k=1} a_{ki} a_{kj} = \begin{cases} 1 & \textit{if } i = j \\ 0 & \textit{if } i \neq j. \end{cases}$$

Bizonyítás Legyen $\mathbf{A}$ egy n-edrendű valós elemű mátrix. Tegyük fel, hogy $\mathbf{A}$ ortogonális, azaz $\mathbf{A}^{-1} = \mathbf{A}^T$. Akkor $\mathbf{E} = \mathbf{A}^T\mathbf{A}$, amiből már adódik, hogy

$$\sum_{k=1} a_{ki} a_{kj} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j. \end{cases}$$

Fordítva, tegyük fel, hogy egy n-edrendű valós $\mathbf{A}$ mátrix elemeire tetszőleges $i, j \in \{1, \ldots, n\}$ indexpár eseténteljesül a

$$\sum_{k=1} a_{ki} a_{kj} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j \end{cases}$$

feltétel. Akkor

$$\mathbf{E} = \mathbf{A}^T\mathbf{A},$$

azaz

$$\mathbf{A}^{-1} = \mathbf{A}^T.$$

Tehát $\mathbf{A}$ ortogonális mátrix. □

Definíció 12.8 *Egy komplex elemű négyzetes* $\mathbf{A}$ *mátrixot unitér mátrixnak nevezünk, ha* $\mathbf{A}^{-1} = \overline{\mathbf{A}^T}$.

Tétel 12.11 *Egy n-edrendű komplex elemű* **A** *mátrix akkor és csak akkor unitér, ha tetszőleges $i, j \in \{1, \ldots, n\}$ indexekre*

$$\sum_{k=1} \overline{a}_{ki} a_{kj} = \begin{cases} 1 & \textit{if } i = j; \\ 0 & \textit{if } i \neq j. \end{cases}$$

Bizonyítás Legyen **A** egy n-edrendű komplex elemű mátrix. Tegyük fel, hogy **A** unitér, azaz $\mathbf{A}^{-1} = \overline{\mathbf{A}^T}$. Akkor $\mathbf{E} = \overline{\mathbf{A}^T}\mathbf{A}$, amiből már adódik, hogy

$$\sum_{k=1} \overline{a}_{ki} a_{kj} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j. \end{cases}$$

Fordítva, tegyük fel, hogy egy n-edrendű komplex **A** mátrix elemeire tetszőleges $i, j \in \{1, \ldots, n\}$ indexpár esetén teljesül a

$$\sum_{k=1} \overline{a}_{ki} a_{kj} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j \end{cases}$$

feltétel. Akkor

$$\mathbf{E} = \overline{\mathbf{A}^T}\mathbf{A},$$

azaz

$$\mathbf{A}^{-1} = \overline{\mathbf{A}^T}.$$

Tehát **A** unitér mátrix. □

Tétel 12.12 *Valós euklideszi tér tetszőleges ortogonális transzformációjának bármely ortonormált bázishoz tartozó mátrixa ortogonális. Fordítva, ha egy valós euklideszi tér valamely φ lineáris transzformációjának mátrixa a tér egy ortonormált bázisában ortogonális, akkor φ a tér ortogonális transzformációja.*

Bizonyítás. Legyen φ egy (V, f) valós euklideszi tér lineáris transzformációja. Tegyük fel, hogy φ ortogonális. Legyen $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n$ a V egy ortonormált bázisa. Legyen $\mathbf{A} = [\varphi]_{\mathcal{B}}$. Akkor tetszőleges $i, j \in \{1, dots, n\}$ indexekre

$$\sum k = 1^n a_{ki} a_{kj} = f(\varphi(\underline{e}_i), \varphi(\underline{e}_j)) = f(\underline{e}_i, \underline{e}_j),$$

amely egyenlő 1-gyel vagy 0-val, aszerint, hogy $i = j$ vagy $i \neq j$. A Tétel **??** miatt a $[\varphi]_{\mathcal{B}}$ mátrix ortogonális.

Fordítva, tegyük fel, hogy φ mátrixa a V valamely $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n$ ortonormált bázisában ortogonális. Legyen $[\varphi]_\mathcal{B} = \mathbf{A}$. Jelölje $\underline{a}_k$ az $\mathbf{A}$ mátrix k-dik oszlopvektorát. Mivel $\mathbf{A}$ ortogonális mátrix, ezért a Tétel **??** miatt

$$\sum_{k=1} a_{ki}a_{kj} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j \end{cases}$$

feltétel. Igy a V tér tetszőleges $\underline{x} = \sum_{i=1}^n \xi_i \underline{e}_i$ és $\underline{y} = \sum_{i=1}^n \eta_i \underline{e}_i$ vektorai esetén

$$f(\varphi(\underline{x}), \varphi(\underline{y})) = \sum_{i,j=1}^n \xi_i \eta_j f(\varphi(\underline{e}_i), \varphi(\underline{e}_j)) =$$

$$\sum_{i,j=1}^n \xi_i \eta_j f(\underline{a}_i, \underline{a}_j) = \sum_{i,j=1}^n \xi_i \eta_j (\sum k = 1^n a_{ki}a_{kj}) = \sum_{t=1}^n \xi_t \eta_t = f(\underline{x}, \underline{y}).$$

Tehát φ ortogonális transzformáció. □

Tétel 12.13 *Komplex euklideszi tér tetszőleges ortogonális transzformációjának bármely ortonormált bázishoz tartozó mátrixa unitér. Fordítva, ha egy komplex euklideszi tér valamely φ lineáris transzformációjának mátrixa a tér egy ortonormált bázisában unitér, akkor φ a tér ortogonális transzformációja.*

Bizonyítás. Legyen φ egy (V, f) komplex euklideszi tér lineáris transzformációja. Tegyük fel, hogy φ ortogonális. Legyen $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n$ a V egy ortonormált bázisa. Legyen $\mathbf{A} = [\varphi]_\mathcal{B}$. Akkor tetszőleges $i, j \in \{1, dots, n\}$ indexekre

$$\sum k = 1^n \overline{a}_{ki}a_{kj} = f(\varphi(\underline{e}_i), \varphi(\underline{e}_j)) = f(\underline{e}_i, \underline{e}_j),$$

amely egyenlő 1-gyel vagy 0-val, aszerint, hogy $i = j$ vagy $i \neq j$. A Tétel **??** miatt a $[\varphi]_\mathcal{B}$ mátrix unitér.

Fordítva, tegyük fel, hogy φ mátrixa a V valamely $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n$ ortonormált bázisában unitér. Legyen $[\varphi]_\mathcal{B} = \mathbf{A}$. Jelölje $\underline{a}_k$ az $\mathbf{A}$ mátrix k-dik oszlopvektorát. Mivel $\mathbf{A}$ unitér mátrix, ezért a Tétel **??** miatt

$$\sum_{k=1} \overline{a}_{ki}a_{kj} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j. \end{cases}$$

Igy a V tér tetszőleges $\underline{x} = \sum_{i=1}^n \xi_i \underline{e}_i$ és $\underline{y} = \sum_{i=1}^n \eta_i \underline{e}_i$ vektorai esetén

$$f(\varphi(\underline{x}), \varphi(\underline{y})) = \sum_{i,j=1}^n \overline{\xi}_i \eta_j f(\varphi(\underline{e}_i), \varphi(\underline{e}_j)) =$$

$$\sum_{i,j=1}^{n} \overline{\xi}_i \eta_j f(\underline{a}_i, \underline{a}_j) = \sum_{i,j=1}^{n} \overline{\xi}_i \eta_j (\sum k = 1^n \overline{a}_{ki} a_{kj}) = \sum_{t=1}^{n} \overline{\xi}_t \eta_t = f(\underline{x}, \underline{y}),$$

azaz φ ortogonális □

12.3. Szimmetrikus és önadjungált transzformációk

Definíció 12.9 *Egy valós (komplex) (V, f) euklideszi tér valamely φ lineáris transzformációját szimmetrikus (önadjungált) transzformációnak nevezzük, ha tetszőleges $\underline{x}, \underline{y} \in V$ vektorokra*

$$f(\varphi(\underline{x}), \underline{y}) = f(\underline{x}, \varphi(\underline{y}))$$

teljesül.

Tétel 12.14 *Valós euklideszi tér tetszőleges szimmetrikus transzformációjának bármely ortonormált bázishoz tartozó mátrixa szimmetrikus. Fordítva, ha egy valós euklideszi tér valamely φ lineáris transzformációjának mátrixa a tér egy ortonormált bázisában szimmetrikus, akkor φ a tér szimmetrikus transzformációja.*

Tétel 12.15 *Komplex euklideszi tér tetszőleges önadjungált transzformációjának bármely ortonormált bázishoz tartozó mátrixa önadjungált. Fordítva, ha egy valós euklideszi tér valamely φ lineáris transzformációjának mátrixa a tér egy ortonormált bázisában önadjungált, akkor φ a tér önadjungált transzformációja.*

Tétel 12.16 *Egy (V, f) valós (komplex) euklideszi tér tetszőleges szimmetrikus (önadjungált) lineáris traszformációjának különböző sajátértékeihez tartozó sajátvektorai egymásra ortogonálisak.*

Bizonyítás. Mivel a szimmetrikus, illetve komplex önadjungált lineáris transzformációk mátrixa ortonormált bázisban szimmetrikus, illetve önadjungált, ezért a Tétel **??** miatt a sajátértékei valós számok. Igy az alábbi bizonyítás mindkét esetre érvényes.

Legyen φ egy (V, f) euklideszi tér szimmetrikus, (komplex esetben önadjungált) lineáris transzformációja. Legyenek $\lambda_1 \neq \lambda_2$ a φ sajátértékei. Akkor megadhatók olyan V-beli $\underline{v}_1$ és $\underline{v}_2$ vektorok, hogy

$$\varphi(\underline{v}_1 = \lambda_1 \underline{v}_1 \quad \varphi(\underline{v}_1 = \lambda_1 \underline{v}_1.$$

Mivel
$$f(\varphi(\underline{v}_1), \underline{v}_2) = f(\lambda_1 \underline{v}_1, \underline{v}_2) = \lambda_1 f(\underline{v}_1, \underline{v}_2)$$
és
$$f(\underline{v}_1, \varphi(\underline{v}_2)) = f(\underline{v}_1, \lambda_2 \underline{v}_2) = \lambda_2 f(\underline{v}_1, \underline{v}_2),$$
ezért az
$$f(\varphi(\underline{v}_1), \underline{v}_2) = f(\underline{v}_1, \varphi(\underline{v}_2))$$
egyenlőségből
$$\lambda_1 f(\underline{v}_1, \underline{v}_2) = \lambda_2 f(\underline{v}_1, \underline{v}_2)$$
azaz, a $\lambda_1 \neq \lambda_2$ feltétel miatt,
$$f(\underline{v}_1, \underline{v}_2) = 0$$
adódik. Tehát $\underline{v}_1$ és $\underline{v}_2$ a (V, f) euklideszi tér ortogonális vektorai. □

Tétel 12.17 *Egy (V, f) valós euklideszi tér valamely φ lineáris transzformációja akkor és csak akkor szimmetrikus, ha V-nek létezik a φ sajátvektoraiból álló ortonormált bázisa.*

Bizonyítás. Tegyük fel, hogy (V, f)-nek van olyan $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n\}$ ortonormált bázisa, melynek vektorai egy φ lineáris transzformáció sajátvektorai. Akkor minden $i = 1, \ldots, n$ indexhez van olyan λ_i valós szám, hogy
$$\varphi(\underline{e}_i) = \lambda_i \underline{e}_i.$$
Igy
$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$
amely mátrix szimmetrikus. Igy, a Tétel 12.14 szerint, φ a (V, f) tér szimmetrikus lineáris transzformációja.

Fordítva, tegyük fel, hogy φ a (V, f) euklideszi tér szimmetrikus lineáris transzformációja. Megmutajuk, hogy ekkor (V, f)-nek van a φ sajátvektoraiból álló ortonormált bázisa. Ezt a (V, f) euklideszi tér dimenziójára vonatkozó teljes indukciós bizonyítás módszerével tesszük.

Ha $dimV = 1$, akkor V-nek minden nem nullvektora a φ-nek sajátvektora, így (V, f)-nek minden normálvektora φ sajátvektorából álló ortonormált bázis.

Tegyük fel, hogy $n \geq 1$, és az állítás igaz minden olyan (V, f) valós euklideszi tér tetszőleges szimmetrikus lineáris transzformációjára, amelyre $dimV \leq n$ teljesül. Legyen (V, f) $n+1$-dimenziós valós euklideszi tér. Legyen φ a (V, f) egy szimmetrikus lineáris transzformációja. A Tétel **??** és Tétel **??** miatt van olyan λ_1 valós szám, amely φ-nek sajátértéke. Jelölje $\underline{e}_1$ a φ egy normált sajátvektorát. A Tétel **??** miatt ez benne van a (V, f) tér egy ortonormált

$$\mathcal{B} = \{\underline{e}_1, \underline{e}'_2, \ldots, \underline{e}'_{n+1}\}$$

bázisában. Jelölje W a V vektortér

$$\langle \underline{e}'_2, \ldots, \underline{e}'_{n+1} \rangle$$

alterét. Az világos, hogy tetszőleges

$$\underline{x} = \xi_1 \underline{e}_1 + \xi_2 \underline{e}'_2 + \cdots + \xi_{n+1} \underline{e}'_{n+1} \in V$$

vektor akkor és csak akkor ortogonális az $\underline{e}_1$ vektorral, ha

$$0 = f(\underline{e}_1, \xi_1\underline{e}_1 + \xi_2\underline{e}'_2 + \cdots + \xi_{n+1}\underline{e}'_{n+1}) = \xi_1 f(\underline{e}_1, \underline{e}_1) + \xi_2 f(\underline{e}_1, \underline{e}'_2) + \cdots + \xi_{n+1} f(\underline{e}_1, \underline{e}'_{n+1}) = \xi_1,$$

azaz $\underline{x} \in W$. Igy, ha $\underline{x} \in W$ tetszőleges vektor, akkor

$$f(\underline{e}_1, \varphi(\underline{x})) = f(\varphi(\underline{e}_1), \underline{x}) = f(\lambda_1 \underline{e}_1, \underline{x}) = \lambda_1 f(\underline{e}_1, \underline{x}) = 0.$$

Tehát minden $\underline{x} \in W$ vektor esetén $\varphi(\underline{x}) \in W$ is teljesül. Tehát φ a W alteret önmagába képezi le. Továbbá az is igaz, hogy φ-nek W-re való leszűkítése a W egy szimmetrikus lineáris transzformációja. Mivel minden $\underline{x} \in W$ és minden $\xi \in \mathbb{R}$ esetén

$$f(\underline{e}_1, \xi\underline{x}) = \xi f(\underline{e}_1, \underline{x}) = 0,$$

ezért W nem csak altere a V vektortérnek, hanem euklideszi tér is az f skaláris szorzásra nézve. Mivel $dimW = n$, ezért az indukciós feltétel miatt W-nek van a φ sajátvektoraiból álló

$$\{\underline{e}_2, \ldots, \underline{e}_{n+1}\}$$

ortonormált bázisa. Ennek vektorai az $\underline{e}_1$ vektorral együtt a (V, f) euklideszi tér φ sajátvektoraiból álló ortonormált bázisát adják. □

Tétel 12.18 *Minden* **A** *szimmetrikus mátrixhoz van olyan* **Q** *ortogonális mátrix, hogy a* $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ *mátrix diagonális.*

Bizonyítás. Legyen $\mathbf{A}$ egy n-edrendű szimmetrikus mátrix. Legyen $\mathcal{B} = \{\underline{e}_1, \ldots, \underline{e}_n\}$ egy (V, f) valós euklideszi tér valamely ortonormált bázisa. Legyen φ a V azon lineáris transzformációja, melynek a $\mathcal{B}$ bázishoz tartozó mátrixa az $\mathbf{A}$ mátrix. Az előző tétel miatt, a (V, f) euklideszi térnek van olyan $\mathcal{B}'$ ortonormált bázisa, amelynek vektorai a φ sajátvektorai; ebben a bázisban a φ mátrixa egy $\mathbf{D}$ diagonális mátrix. A Tétel **??** szerint a $\mathcal{B}$ ortonormált bázisról a $\mathcal{B}'$ ortonormált bázisra való átmenet $\mathbf{Q}$ mátrixa ortogonális. Igy a Tétel **??** szerint

$$\mathbf{D} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}.$$

□

Tétel 12.19 *Egy (V, f) komplex euklideszi tér valamely φ lineáris transzformációja akkor és csak akkor önadjungált, ha φ minden sajátértéke valós szám, valamint V-nek létezik a φ sajátvektoraiból álló ortonormált bázisa.*

Bizonyítás. A Tétel 12.17 bizonyítása szinte szóról-szóra alkalmazható, ha figyelembe vesszük azt a tényt, hogy önadjungált lineáris transzformáció sajátértékei valós számok. □

Tétel 12.20 *Minden $\mathbf{A}$ önadjungált mátrixhoz van olyan $\mathbf{Q}$ unitér mátrix, hogy a $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ mátrix diagonális.*

Bizonyítás. Jelen állítás bizonyításhoz szinte szóról-szóra alkalmazhatjuk a Tétel 12.18 bizonyítását, csak szimmetrikus helyett önadjungált, ortogonális helyett pedig unitér mátrixokat kell tekinteni □

Tétel 12.21 *Bármilyen legyen is az az ortogonális (unitér) mátrix, melynek segítségével egy szimmetrikus (önadjungált) $\mathbf{A}$ mátrixot diagonális alakra transzformálunk, a kapott diagonális mátrix főátlójában az $\mathbf{A}$ mátrix sajátértékei állnak, mindegyik annyiszor, amennyi az illető sajátérték algebrai multiplicitása.*

Bizonyítás. Mivel az $\mathbf{A}$ mátrixból származtatott $\mathbf{D}$ diagonális mátrix hasonló $\mathbf{A}$-val, ezért sajátértékei megegyeznek az $\mathbf{A}$ sajátértékeivel úgy, hogy minden sajátérték algebrai multiplicitása $\mathbf{A}$-nál és a $\mathbf{D}$-nél ugyanaz. Ebből ás a Tétel 6.21-ből már adódik a tétel állítása. □

12.4. Valós értékű kvadratikus alakok

Tétel 12.22 *Legyenek $\mathcal{B}$ és $\mathcal{B}'$ egy valós (vagy komplex) (V, f) euklideszi tér két ortonormált bázisai. Akkor V tetszőleges φ bilineáris funkcionálja esetén a $[\varphi]_{\mathcal{B}}$ és $[\varphi]_{\mathcal{B}'}$ mátrixok karakterisztikus polinomjai megegyeznek.*

Bizonyítás. Legyen τ a $\mathcal{B}$ ortonármált bázist a $\mathcal{B}'$ ortonormált bázisra képező ortogonális transzformáció.

Vizsgáljuk először azt az esetet, amikor (V, f) valós euklideszi tér. Mivel τ ortogonális, ezért a Tétel **??** szerint $[\tau]_{\mathcal{B}}$ ortogonális mátrix, azaz $[\tau]_{\mathcal{B}}^{-1} = [\tau]_{\mathcal{B}}^{T}$. Ezért a Tétel **??** es Tétel **??** miatt

$$[\varphi]_{\mathcal{B}'} = [\tau]_{\mathcal{B}}^{-1}[\varphi]_{\mathcal{B}}[\tau]_{\mathcal{B}},$$

azaz $[\varphi]_{\mathcal{B}}$ és $[\varphi]_{\mathcal{B}'}$ hasonló mátrixok. A Tétel **??** miatt ezek karakterisztikus polinomjai megegyeznek.

Vizsgáljuk most azt az esetet, amikor (V, f) komplex euklideszi tér. Mivel τ ortogonális, ezért a Tétel **??** szerint $[\tau]_{\mathcal{B}}$ unitér mátrix, azaz $[\tau]_{\mathcal{B}}^{-1} = \overline{[\tau]^{T}}_{\mathcal{B}}$. Ezért a Tétel **??** es Tétel **??** miatt

$$[\varphi]_{\mathcal{B}'} = [\tau]_{\mathcal{B}}^{-1}[\varphi]_{\mathcal{B}}[\tau]_{\mathcal{B}},$$

azaz $[\varphi]_{\mathcal{B}}$ és $[\varphi]_{\mathcal{B}'}$ hasonló mátrixok. A Tétel **??** miatt ezek karakterisztikus polinomjai megegyeznek. $\square$

Tétel 12.23 *Egy V valós vagy komplex euklideszi téren értelmezett g kvadratikus alak akkor és csak akkor*

1. *pozitív definit, ha g tetszőleges V-beli ortonormált bázisához tartozó mátrixának minden sajátértéke pozitív;*

2. *pozitív szemidefinit, ha g tetszőleges V-beli ortonormált bázisához tartozó mátrixának minden sajátértéke nemnegatív, de a sajátértékek között van olyan, amelyik egyenlő a 0-val;*

3. *negatív definit, ha g tetszőleges V-beli ortonormált bázisához tartozó mátrixának minden sajátértéke negatív;*

4. *negatív szemidefinit, ha g tetszőleges V-beli ortonormált bázisához tartozó mtŕixának minden sajátértéke nempozitív, de a sajátértékek között van olyan, amelyik egyenlő a 0-val;*

5. *indefinit, ha g tetszőleges V-beli ortonormált bázisához tartozó mátrixának van pozitív és negatív sajátértéke.*

13. fejezet

Másodrendű görbék

Definíció 13.1 *Másodrendű görbén olyan síkgörbét értünk, amelynek egyenlete síkbeli* (x, y) *derékszögű koordináta-rendszerben olyan*

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

alakú egyenlet, amelyben szereplő a, b és c együtthatók közül legalább az egyik nem nulla.

Példák.

1. Ellipszis, melynek középponti egyenlete: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$.
2. Hiperbola, melynek középponti egyenlete: $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$.
3. Parabola, melynek középponti egyenlete: $y^2 - 2px = 0$.
4. Metsző egyenespár, melynek középponti egyenlete: $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$.
5. Párhuzamos egyenespár, melynek középponti egyenlete: $y^2 - b^2 = 0$.
6. Pont, melynek középponti egyenlete: $x^2 + y^2 = 0$.
7. Kettős egyenes, melynek középpont egyenlete: $y^2 = 0$

Megjegyzés 13.1 *Megmutatható, hogy minden másodrendű görbe a fenti hét típus valamelyike.*

Megjegyzés 13.2 *Könnyen ellenőrízhető, hogy egy másodrendű görbe*

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

egyenlete a következő mátrixszorzatos alakba írható:

$$[x, y, 1] \begin{bmatrix} a & \frac{b}{2} & \frac{d}{2} \\ \frac{b}{2} & c & \frac{e}{2} \\ \frac{d}{2} & \frac{e}{2} & f \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}.$$

Figyeljük meg, hogy a fenti szorzatban szereplő mátrix szimmetrikus!

Definíció 13.2 *Az $ax^2 + bxy + cy^2 + dx + ey + f = 0$ egyenletű másodrendű görbe mátrixán az*

$$\begin{bmatrix} a & \frac{b}{2} & \frac{d}{2} \\ \frac{b}{2} & c & \frac{e}{2} \\ \frac{d}{2} & \frac{e}{2} & f \end{bmatrix}$$

(szimmetrikus) mátrixot értjük. Ennek determinánsát a görbe determinánsának nevezzük.

Definíció 13.3 *Ha egy másodrendű görbe determinánsa nem nulla, akkor a görbét nem elfajuló (vagy közönséges) másodrendű görbének, ellenkező esetben elfajuló másodrendű görbének nevezzük.*

Megjegyzés 13.3 *Az ellipszis, a hiperbola és a parabola közönséges mśodrendű görbék, a metsző egyenespár, a párhuzamos egyenespár, a pont és a kettös egyenes elfajuló másodrendű görbék, ezen utóbbi négyből az első három mátrixának rangja 2, a negyedik (kettős egyenes) mátrixának rangja pedig 1.*

Megjegyezzük még, hogy a fenti hét típus közül csak a parabolának nincs szimmetriaközéppontja.

Tétel 13.1 *Egy S síkbeli $ax^2 + bxy + cy^2 + dx + ey + f = 0$, azaz*

$$[x, y, 1] \begin{bmatrix} a & \frac{b}{2} & \frac{d}{2} \\ \frac{b}{2} & c & \frac{e}{2} \\ \frac{d}{2} & \frac{e}{2} & f \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

egyenletű másodrendű görbének egy S-beli (u, v) pont akkor és csak akkor szimmetriaközéppontja, ha az (u, v) pár megoldása az

$$\begin{aligned} au + \tfrac{b}{2}v + \tfrac{d}{2} = &\quad 0 \\ \tfrac{b}{2}u + cv + \tfrac{e}{2} = &\quad 0 \end{aligned}$$

lineáris egyenletrendszernek.

Irodalomjegyzék

[1]

Tárgymutató